A LECTURE COURSE ON FUNCTIONAL ANALYSIS

JOHANNES EBERT

Contents

Date: January 29, 2021.

What this course is about

The minimal definition of Functional Analysis is that it deals with (usually infinite-dimensional) vector spaces over $\mathbb{R}$ or $\mathbb{C}$, which are in addition equipped with topology, and with continuous linear maps between them. A little more lively (quoted from the introduction of [3]):

"Functional Analysis might be described as a part of mathematics where analysis, topology, measure theory, linear algebra and algebra come together to create a rich and fascinating theory. The applications of this theory are then equally spread throughout mathmatics (and beyond)".

Methods of functional analysis are used in partial differential equations, differential geometry, algebraic topology, number theory, group theory, ...

The books which were used in preparation to this course

(1) Hirzebruch–Scharlau [6]. This classic introduction which covers a lot of material in a few pages is highly recommended.

(2) Rudin, Real and Complex Analysis [8]. This contains only a little bit of functional analysis proper, but a is very good introduction to some necessary background material from measure theory, Fourier analysis and basic complex analysis.
(3) Rudin, Functional analysis [9]. This well-known classic is in my opinion *not* a good introduction, because it is not an introduction at all.
(4) Conway [2] is highly recommended.
(5) Werner, Funktionalanalysis [11] is also highly recommended.
(6) Einsiedler, Ward [3].
(7) Tao [10].

Standard notations

(1) $\mathbb{K}$ is one of the fields $\mathbb{R}$ or $\mathbb{C}$. The real and imaginary part of $z \in \mathbb{C}$ are denoted $\Re(z), \Im(z)$, and the complex conjugate is $\overline{z}$. At some places, there are arguments from linear algebra which work for every field. Then we use the letter $\Bbbk$, to indicate that the argument is purely algebraic.
(2) For a $\Bbbk$-vector space V and $A, B \subset V$, $z \in \Bbbk$, we denote $A+B = \{a+b | a \in A, b \in B\}$ and $zA = \{za | a \in A\}$ and $A - B := A + (-B)$.
(3) $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
(4) $\underline{n} := \{1, \ldots, n\}$.
(5) $\mathcal{P}(X)$ is the set of all subsets of a set X (the *power set*).
(6) For a linear map $F : V \to W$ of vector space, we use the notation $\ker(F)$ and $\operatorname{im}(F)$ for kernel and image.
(7) For two $\Bbbk$-vector spaces V, W, we denote $\operatorname{Hom}(V, W)$ the vector space of all linear maps $V \to W$.

Prerequisites for this course

(1) Linear Algebra, as done in the introductory lectures. In particular, familiarity with the theory of euclidean and unitary vector spaces and normal endomorphisms of them will be of great help.
(2) Basic measure theory and Lebesgue integration, as covered in Analysis III, will be used throughout the course, both to have interesting examples to which the theory can be applied, as well as useful tools for the development of the theory. I collected the relevant facts with proofs in the appendix §C. This is intended as a reference.
(3) Basic point-set topology, as we have learnt it in the course "Analysis, Topologie und Geometrie". In the beginning, metric spaces suffice. We will make occasional use of theorems about continuous functions on compact Hausdorff spaces, such as Urysohn's lemma and the Stone–Weierstrass theorem. Later on, knowing the product topology (with infinitely many factors) and Tychonov's theorem will be crucial. We make occasional use of (ultra)filters and nets. The appendix B collects the material.
(4) There is no way around using some facts about holomorphic functions in one complex variable when it comes to the discussion of spectra of operators. The small amount of things which is needed in this course is collected in the appendix D. Unfortunately, this has been kicked out of the curriculum. Even though this is the case, I think that any serious student of mathematics must know the Cauchy integral formula.

(5) Some key theorems in this course depend crucially on the use of Zorn's lemma. This was introduced in the course "Analysis, Topologie und Geometrie", before we proved Tychonov's theorem. The appendix A contains a proof of this fundamental result, together with some typical applications.

Remarks

(1) Some subsections are marked with an asterisk. These contain material which is not so important, and in particular not examinable.

1. Normed spaces and Banach spaces

1.1. **Normed spaces.** Throughout these notes, $\mathbb{K}$ is one of the fields $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1. *Let V be a $\mathbb{K}$-vector space. A* seminorm *on V is a map*

$$\|\_\| : V \to [0,\infty),$$

such that

1. *for all $v \in V$ and $a \in \mathbb{K}$, we have $\|av\| = |a|\|v\|$,*
2. *for all $v,w \in V$, we have $\|v+w\| \leq \|v\| + \|w\|$.*

If in addition

$$\|v\| = 0 \Rightarrow v = 0$$

holds, $\|\_\|$ is a norm.

A normed vector space *is a $\mathbb{K}$-vector space V, together with a norm $\|\_\|$ on V.*

The absolute value on $\mathbb{K}$ is a norm. Other finite-dimensional examples are the norms

$$\|x\|_{\ell^1} := \sum j = 1^n |x_j|,$$
$$\|x\|_{\ell^\infty} := \max_{j=1,\ldots,n} |x_j|$$

and

$$\|x\|_{\ell^2} := \sqrt{\sum j = 1^n |x_j|^2}$$

on $\mathbb{K}^n$.

Definition 1.2. *Let $(V, \|\_\|)$ be a normed vector space. The distance function*

$$d : V \times V \to V,\ d(v,w) := \|v - w\|$$

makes V into a metric space, and the metric induces a topology on V.

We use the following notations for balls.

Notation 1.3. *Let $(V, \|\_\|)$ be a (semi)normed space. We write*

$$B_r(v) := \{w \in V \|v - w\| < r\}$$

for the open r-ball around v, and

$$\overline{B}_r(v) := \{w \in V \|v - w\| \leq r\}$$

for the closed ball. The closed ball $\overline{B}_r(0)$ appears often enough to be denoted by a special notation

$$D_r(V) := \{v \in V | \|v\| \leq r\} \subset V.$$

The metric topology on V is of course Hausdorff if $\|\_\|$ is a norm, and it is first countable. Recall that continuity of maps between metric spaces can equivalently expressed using open sets, the $\epsilon - \delta$-criterion and sequential continuity.

Lemma 1.4. *Let $(V, \|\_\|)$ be a normed vector space. Then*

1. *The norm $\|\_\| : V \to \mathbb{R}$ is continuous.*
2. *The addition $\alpha : V \times V \to V$, $\alpha(v,w) := v + w$, is continuous.*
3. *The scalar multiplication $\mu : \mathbb{K} \times V \to V$, $\mu(a,v) := av$, is continuous.*

The proof is trivial, once one uses the characterization of continuity by sequences. The lemma has an important consequence.

Lemma 1.5. *Let V and W be normed vector spaces, and let $\mathcal{L}(V, W)$ be the set of all linear continuous maps $F : V \to W$. Then $\mathcal{L}(V, W)$ is a vector space.*

Proof. This amounts to proving that when $F, G : V \to W$ are continuous and $a \in \mathbb{K}$, the two linear maps

$$aF, F + G$$

are continuous, which is a straightforward consequence of the previous lemma. $\square$

Continuous *linear* maps have a very important characterization.

Theorem 1.6. *Let V and W be normed vector spaces and let $F : V \to W$ be linear. The following are equivalent:*

1. *F is continuous,*
2. *F is continuous at 0,*
3. *there is $C \geq 0$ such that $\|Fv\| \leq C\|v\|$ for all $v \in W$.*
4. $\sup_{v \in V, \|v\| \leq 1} \|Fv\| < \infty$.

Definition 1.7. *Another name for a continuous linear map is* bounded operator. *If the target space is the ground field $\mathbb{K}$, we also say* functional.

Proof of Theorem 1.6. The implications $1 \Rightarrow 2$ and $3 \Rightarrow 4$ are trivial, and $4 \Rightarrow 1$ is very easy. $2 \Rightarrow 3$: there is $\delta > 0$, so that $\|v\| \leq \delta$ implies $\|Fv\| \leq 1$. Then for each v

$$\|Fv\| = \|\frac{\|v\|}{\delta} F(\frac{\delta}{\|v\|} v)\| \leq \frac{1}{\delta}\|v\|.$$

$\square$

Definition 1.8. *Let $F : V \to W$ be a continuous linear map between normed vector spaces. We define the* operator norm *of F as the quantity*

$$\|F\| := \sup_{v \in V, \|v\| \leq 1} \|Fv\|.$$

Note that for all $v \neq 0$, we have

$$\|Fv\| = \|F(\|v\| \frac{v}{\|v\|})\| = \|\|v\| F(\frac{v}{\|v\|})\| = \|v\| \|F(\frac{v}{\|v\|}\| \leq \|v\| \|F\|,$$

so that

$$\|Fv\| \leq \|F\| \|v\|$$

for all v.

Lemma 1.9. *Let V and W be normed vector spaces. Then the operator norm is a norm on $\mathcal{L}(V, W)$. If U is a third normed vector space, $F : V \to W$, $G : W \to U$ continuous linear maps, then*

$$\|GF\| \leq \|G\| \|F\|.$$

Proof. It is clear that $\|aF\| = |a| \|F\|$ when $F \in \mathcal{L}(V, W)$ and $a \in \mathbb{K}$. If $\|F\| = 0$, then $\|Fv\| = 0$ for all v with $\|v\| \leq 1$, but that implies $Fv = 0$ for all v. For the triangle inequality, let $F_0, F_1 \in \mathcal{L}(V, W)$ and $v \in V$ with $\|v\| \leq 1$. Then

$$\|F_0 v + F_1 v\| \leq \|F_0 v\| + \|F_1 v\| \leq \|F_0\| + \|F_1\|,$$

and pass to the supremum. Similarly

$$\|GFv\| \leq \|G\| \|Fv\| \leq \|G\| \|F\| \|v\|$$

implies $\|GF\| \leq \|G\| \|F\|$. $\square$

Definition 1.10. *Let V and W be normed space and let $F : V \to W$ be a bounded linear operator.*

1. *F is an* isometry *if it preserves norms, i.e. if*
$$\|Fv\| = \|v\|$$
for all $v \in V$.
2. *F is an* isomorphism *if F is bijective and if the inverse $F^{-1} : W \to V$ is also bounded.*
3. *F is an* isometric isomorphism *if F is both, an isometry and an isomorphism.*

The notions of "isomorphism" and "isometric isomorphism" of two normed spaces need to be distinguished. Both appear naturally. A special case is when a vector space V is equipped with two norms $\|\_\|_0$ and $\|\_\|_1$. A special case of Theorem 1.6 is the following:

Lemma 1.11. *Let $\|\_\|_0$ and $\|\_\|_1$ be two norms on V. Then the following are equivalent.*

1. *$\|\_\|_0$ and $\|\_\|_1$ induce the same topology on V.*
2. $\mathrm{id} : (V, \|\_\|_0) \to (V, \|\_\|_1)$ *and* $\mathrm{id} : (V, \|\_\|_1) \to (V, \|\_\|_0)$ *are continuous.*
3. *There are constants $0 < c \le C$ such that $c\|v\|_0 \le \|v\|_1 \le C\|v\|_0$ for all $v \in V$.* □

If the two norms satisfy these conditions, they are called equivalent*.*

The following crucial result is known from last term's course.

Theorem 1.12. *All norms on a finite-dimensional $\mathbb{K}$-vector space V are equivalent. Hence each linear map $F : V \to W$ from a finite-dimensional normed space to an arbitrary normed space is bounded.*

The second claim follows from the first, since $\|v\|_F := \|v\| + \|Fv\|$ is a norm on V, and hence there is C with $\|v\|_F \le C\|v\|$. The Hahn-Banach Theorem, which will be shown later in this course can be used to give a quick proof of Theorem 1.12. For sake of completeness:

Proof of Theorem 1.12. Without loss of generality, $V = \mathbb{R}^n$. Let $\|\_\|$ be a norm on $\mathbb{R}^n$. Then
$$\|v\| = \|\sum_j v_j e_j\| \le \sum_j |v_j| \|e_j\| \le \|v\|_{\ell^\infty} \sum_j \|e_j\|.$$
Hence there is $C > 0$ with
$$\|v\| \le C\|v\|_{\ell^\infty}. \tag{1.13}$$
The norm $\|\_\|_{\ell^\infty}$ induces the product topology on $\mathbb{R}^n$. The subset
$$S := \{v \in \mathbb{R}^n | \|v\|_{\ell^\infty} = 1\} \subset \mathbb{R}^n$$
is closed and bounded and hence compact. The function $\mathbb{R}^n \to \mathbb{R}$, $v \mapsto \|v\|$ is continuous by (1.13), and so it attains its minimum
$$0 < c := \min_{v \in S} \|v\|.$$
It follows that
$$c\|v\|_{\ell^\infty} \le \|v\| \le C\|v\|_{\ell^\infty}$$
for all $v \in S$, and hence by homogeneity for all $v \in \mathbb{R}^n$. □

Exercise 1.14. *Let $F : V \to W$ and $G : W \to V$ be bounded operators. Show:*

1. *If F is an isometry, it is injective.*
2. *If F is an isometry and bijective with inverse G, then G is an isometry.*
3. *If $GF = 1$ and $\|F\|, \|G\| \leq 1$, then F is an isometry. Proof: $\|v\| = \|GFv\| \leq \|G\|\|Fv\| \leq \|Fv\| \leq \|v\|$.*

1.2. **Completeness.**

Definition 1.15. *A normed vector space is a* Banach space *if V is complete with the metric induced by $\|\_\|$. In other words, each Cauchy sequence in V converges.*

Lemma 1.16. *Let V be a normed vector space. The following are equivalent:*

1. *V is complete.*
2. *Whenever $v_n \in V$ is a sequence such that $\sum_{n=1}^{\infty} \|v_m\| < \infty$, the sequence $\sum_{k=1}^{n} v_k$ of partial sums converges in V ("absolutely convergent series are convergent").*

Proof. $1 \Rightarrow 2$ is clear, by the argument form Analysis I which shows that absolutely convergent series in $\mathbb{R}$ are actually convergent.

$2 \Rightarrow 1$: let v_n be a Cauchy sequence in V. We can find a subsequence v_{n_k} such that for $k \geq l$, we have $\|v_{n_k} - v_{n_l}\| \leq \frac{1}{2^l}$. By hypothesis, the series

$$\sum_{m=1}^{\infty} (v_{n_{m+1}} - v_{n_m})$$

converges to some $w \in V$. But

$$\sum_{m=1}^{k-1} = v_{n_k} - v_{n_1}$$

and so we have $\lim_{k to \infty} v_{n_k} = w + v_{n_1}$. Since $\lim_{k\to\infty} \|v_k - v_{n_k}\| = 0$, this implies that $\lim_{k\to\infty} v_k = w + v_{n_1}$. □

Lemma 1.17. *Let $U \subset V$ be a linear subspace of a normed space. Then*

1. *if V is complete, and U is closed, then U is complete.*
2. *if U is complete, then U is closed.*

Proof. (1): a Cauchy sequence u_n in U converges to a limit $u \in V$, and as U is closed, $u \in U$, so u_n has a limit in U. (2): let u_n be a sequence in U which converges to a limit $u \in V$. As U is complete and u_n a sequence, it must converge to some limit $v \in U$, and necessarily $v = u$. □

Lemma 1.18. *Let $F : V \to W$ be a bounded linear operator and assume that there is $c > 0$ such that $c\|v\| \leq \|Fv\|$ for all $v \in V$ (this is often expressed by saying that F* is bounded away from zero*. Then if V is complete, the image* $\mathrm{im}(F) \subset W$ *is complete and hence closed (by Lemma 1.17).*

Proof. Let $v_n \in V$ be a sequence such that Fv_n is a Cauchy sequence. As $\|v_n - v_m\| \leq \frac{1}{c}\|F(v_n - v_m)\|$, v_n is Cauchy and converges to some $v \in V$. Because F is continuous, $Fv = \lim_n Fv_n \in \mathrm{im}(F)$, so that $\mathrm{im}(F)$ is complete. □

Corollary 1.19. *Let $\|\_\|_0$ and $\|\_\|_1$ be two equivalent norms on V. Then $(V, \|\_\|_0)$ is complete iff $(V; \|\_\|_1)$ is complete.*

Proof. Because equivalence of norms is an equivalence relation, it is enough to prove one implication. When $(V, \|\_\|_0)$ is complete, the identity $(V, \|\_\|_0) \to (V, \|\_\|_1)$ satisfies the hypotheses of Lemma 1.18, and so $(V, \|\_\|_1)$ is complete. $\square$

Corollary 1.20. *Each finite-dimensional normed space V is complete, and each finite-dimensional linear subspace $V \subset W$ of an arbitrary normed vector space is closed.*

Proof. V is isometrically isomorphic to $(\mathbb{R}^n, \|\_\|)$ for some norm on $\mathbb{R}^n$. But as $(\mathbb{R}^n, \|\_\|_{\ell^\infty})$ is complete by Analysis II, 1.19 shows that V is complete. The second part follows from the first one and Lemma 1.17. $\square$

A dramatic difference between finite-dimensional and infinite-dimensional normed spaces is that infinite-dimensional normed spaces can contain dense subspaces. These play an important role in many situations. The following Theorem describes the *standard method* of constructing bounded operators, and is very important.

Theorem 1.21 (Extension of bounded operators)**.** *Let V be a normed space, let W be a Banach space and let $U \subset V$ be a dense linear subspace. Let $F : U \to W$ be a continuous linear map. Then there is a unique extension of F to a continuous linear map $G : V \to W$, and $\|G\| = \|F\|$.*

Proof. Uniqueness: if $v \in V$, pick a sequence $u_n \in U$ with $\lim_n u_n = v$. By continuity, we must have $Gv = \lim_n Gu_n = \lim_n Fu_n$, in other words: F can have at most continuous extension.

For the existence, let $v \in V$ and pick a sequence $u_n \in U$ with $u_n \to v$. The sequence $Fu_n \in W$ is a Cauchy sequence, since

$$\|F(u_n - u_m)\| \leq \|F\|\|u_n - u_m\|,$$

and so

$$Gv := \lim_n Fu_n \in W$$

exists. We have to verify the following:

(1) $\|Gv\| \leq \|F\|\|v\|$. Because $\|Fu_n\| \leq \|F\|\|u_n\|$ for all n, we have

$$\|Gv\| = \|\lim_n Fu_n\| = \lim_n \|Fu_n\| \leq \|F\| \lim_n \|u_m\| = \|F\|\|v\|.$$

(2) G is well-defined, that is, the above definition of Gv does not depend on the choice of the sequence u_n which converges to v. But if $u_n \in U$ is another sequence with $w_n \to v$, then $\lim_n(u_n - u_n') = 0$, and since F is continuous, it follows that

$$\lim_n F(u_n - w_n) = 0,$$

and therefore

$$\lim_n Fu_n' = \lim_n Fu_n.$$

(3) G is linear and $G|_U = F$. This is easy.

$\square$

Theorem 1.22. *Let W be a normed space and let V be a Banach space. Then the vector space $\mathcal{L}(W, V)$ of bounded operators $W \to V$, equipped with the operator norm, is a Banach space. In particular, the* dual space *$W' := \mathcal{L}(W, \mathbb{K})$ of any normed vector space is complete.*

Proof. Let $(F_n)_n$ be a Cauchy sequence in $\mathcal{L}(V, W)$. Then there is C so that $\|F_n\| \le C$ for all n. For each $v \in V$, the sequence $F_n v \in W$ is Cauchy because $\|F_n v - F_m v\| \le \|F_n - F_m\|\|v\|$, and so has a limit

$$Fv := \lim_{n\to\infty} F_n v \in W.$$

It is easy to verify that the so defined map $F : V \to W$ is linear. For $v \in V$, we have

$$\|Fv\| = \lim_{n\to\infty} \|F_n v\| \le C\|v\|,$$

so that F is bounded. To check that $\|F - F_n\| \to 0$, let $\epsilon > 0$ and pick n_0 so that $\|F_n - F_m\| \le \epsilon$ for all $m, n \ge n_0$. For $v \in V$ and all such m, n, we have

$$\|(F_n - F_m)v\| \le \|F_n - F_m\|\|v\| \le \epsilon\|v\|.$$

Therefore

$$\|(F_n - F)v\| = \lim_{m\to\infty} \|(F_n - F_m)v\| \le \epsilon\|v\|,$$

as required. □

Definition 1.23. *Let V be a normed space. A* completion *of V is a Banach space W, together with an isometry $\iota : V \to W$ with dense image.*

A completion has a *universal property*, which makes it unique up to isometric isomorphism, once it exists.

Proposition 1.24. *Let V be a normed space.*

(1) *Let $\iota : V \to W$ be a completion of V, let U be a Banach space and let $F : V \to U$ be bounded. Then there is a unique bounded $G : W \to U$ with $G \circ \iota = F$, and $\|G\| = \|F\|$.*
(2) *Let $\iota_i : V \to W_i$, $i = 0, 1$, be two completions. Then there is a unique isometric isomorphism $F : W_0 \to W_1$ with $F \circ \iota_0 = \iota_1$.*

Proof. (1): Since ι is an isometry, the inverse map $\iota^{-1} : \iota(V) \to V$ is an isometry. The operator $F \circ \iota^{-1} : \iota(V) \to U$ is bounded, and so it has a unique extension to $G : W \to U$, by Theorem 1.21.

(2): part (1) gives two bounded operators $F_0 : W_0 \to W_1$ and $F_1 : W_1 \to W_0$ with $F_0 \circ \iota_0 = \iota_1$ and $F_1 \circ \iota_1 = \iota_0$. The operator $F_1 F_0 : W_0 \to W_0$ satisfies

$$F_1 \circ F_0 \circ \iota_0 = F_1 \circ \iota_1 = \iota_0,$$

and $\mathrm{id} : W_0 \to W_0$ has the same property. By the uniqueness statement of part (1), it follows that $F_1 \circ F_0 = \mathrm{id}$. Similarly $F_0 \circ F_1 = \mathrm{id}$.

Furthermore $\|F_i\| = \|\iota_i\| = 1$, and so Exercise 1.14 implies that F_0, F_1 are isometries. □

The proposition shows that a completion, if it exists, is unique. The following theorem guarantees that completions always exist.

Theorem 1.25. *Let V be a normed vector space. Then there exists a Banach space W and a linear isometry $\iota : V \to W$ such that $\iota(V)$ is dense in W.*

We will give two proofs later on. Many (but not all) important Banach spaces arise as a completion of a normed spaces, and many important bounded operators come from an application of Theorem 1.21.

Let us now explain another drastic difference between finite-dimensional and infinite-dimensional spaces: closed bounded subsets of $\mathbb{R}^n$ are compact (Heine-Borel theorem), and this plays a key role in analysis. This is *always* false in infinite dimensions.

Lemma 1.26 (Riesz Lemma)**.** *Let V be a normed vector space and let $W \subset V$ be a proper ($W \neq V$) closed linear subspace. Then for each $\delta > 0$, there is $v \in V$ with $\|v\| \leq 1$ and $\mathrm{dist}_W(v) \geq 1 - \delta$.*

Proof. Choose $u \in V \setminus W$. Then $r := \mathrm{dist}_W(u) > 0$ because W is closed. For each $\epsilon > 0$, there is $w \in W$ with $r \leq \|u - w\| \leq r + \epsilon$. Then $v := \frac{u-w}{\|u-w\|}$ has norm 1 and for each $x \in W$, we have

$$\|v-u\| = \frac{1}{\|u-w\|}(\|u-w-\|u-w\|x\|) = \frac{1}{\|u-w\|}(\|u-(w+\|u-w\|x)\|) \geq \frac{r}{r+\epsilon}.$$

Pick ϵ small enough so that $\frac{1}{r+\epsilon} \geq 1 - \delta$. □

Proposition 1.27. *Let V be a normed vector space and assume that the closed unit ball $D_1(V)$ is compact. Then V is finite-dimensional.*

Proof. We prove the following: if V is infinite-dimensional, there is a sequence $(v_n)_n$ in $D_1(V)$ with $\|v_n - v_m\| \geq \frac{1}{2}$ whenever $n \neq m$. This sequence cannot have a convergent subsequence, so that $D_1(V)$ is not compact.

To construct the sequence, pick $v_1 \in D_1(V)$ arbitrarily. If $v_1, \ldots, v_n$ are already constructed, let $V_n := \mathrm{span}\{v_j | j \leq n\} \subset V$. This is a closed subspace because V_n is finite-dimensional and a proper subspace because V is infinite-dimensional. By Lemma 1.26, there is $v_{n+1} \in D_1(V)$ with $\mathrm{dist}_{V_n}(v_{n+1}) \geq \frac{1}{2}$, in particular $\|v_{n+1} - v_j\| \geq \frac{1}{2}$ for all $j \leq n$. □

1.3. **Linear-algebraic constructions I: quotients.**

From linear algebra, one recalls the notion of a *quotient space*: if V is a $\mathbb{k}$-vector space and $U \subset V$ is a linear subspace, the quotient V/U is the set of all subsets of V of the form

$$v + U := \{v + u | u \in U\} \subset V.$$

The addition and scalar multiplication is defined by

$$(v+U) + (w+U) := (v+w) + U, \; a(v+U) := av + U,$$

and one checks that these are well-defined and give V/U the structure of a $\mathbb{k}$-vector space. The quotient map

$$q : V \to V/U, \; q(v) := v + U$$

is linear, surjective, and has $\ker(q) = U$. It is sometimes useful notation to write $[v] := v + W$, if the subspace W is understood. The most important feature of the quotient construction is its *universal property*: If $F : V \to W$ is a linear map to some other vector space, and $F|_U = 0$, then there is a unique linear $G : V/U \to W$ such that $G \circ q = F$. G is defined by the formula

$$G(v+U) := Fv.$$

Let us introduce (semi)norms into these constructions.

Definition 1.28. *Let V be a seminormed space and let $W \subset V$ be a linear subspace. Consider the quotient space V/W with the quotient map $q : V \to V/W$. We define a seminorm on V/W by*

$$\|x\| := \inf_{v \in V, q(v)=x} \|v\|.$$

Proof that this is a seminorm. The only point that requires thought is the triangle inequality. Let $x_0, x_1 \in V/W$ and let $\epsilon > 0$ be arbitrary. Choose $v_i \in V$ with $q(v_i) = x_i$ and $\|v_i\| \leq \|x_i\| + \epsilon$. As $q(v_0 + v_1) = x_0 + x_1$, we have

$$\|x_0 + x_1\| \leq \|v_0 + v_1\| \leq \|v_0\| + \|v_1\| \leq \|x_0\| + \|x_1\| + 2\epsilon.$$

This holds for all $\epsilon > 0$, and the triangle inequality follows. □

Occasionally, an alternative description for the quotient seminorm is useful. If $x = q(v) = v + W$, then any other $v' \in V$ with $q(v') = x$ can be written as $v' = v + w$ with $w \in W$. Hence we also have

$$\|q(v)\| = \inf_{w \in W} \|v + w\|. \tag{1.29}$$

By definition, we have

$$\|qv\| \leq \|v\|$$

for each $v \in V$, so that q is bounded with operator norm at most 1.

The seminorm on the quotient interacts nicely with the universal property:

Proposition 1.30. *Let V be a seminormed space and let $U \subset V$ be a linear subspace. Let W be a further seminormed space and let $F : V \to W$ be a bounded linear operator with $F|_U = 0$. Then the unique bounded linear map $G : V/U \to W$ with $G \circ q = F$ is bounded, and $\|G\| = \|F\|$.*

Proof. Let $x \in V/U$ and $\epsilon > 0$, and pick $v \in V$ with $q(v) = x$ and $\|v\| \leq \|x\| + \epsilon$. Then $G(x) = G(q(v)) = Fv$ and so

$$\|Gx\| \leq \|F\|\|v\| \leq \|F\|\|x\| + \|F\|\epsilon.$$

$\epsilon > 0$ was arbitrary, and so $\|Gx\| \leq \|F\|\|x\|$. Therefore G is bounded, and $\|G\| \leq \|F\|$. On the other hand

$$\|F\| = \|G \circ q\| \leq \|G\|\|q\| \leq \|G\|.$$

□

The first thing we want to do with this construction is to replace seminorms by norms.

Proposition 1.31. *Let V be a seminormed space. The* null space

$$N := \{v \in V | \|v\| = 0\} \subset V$$

is a linear subspace, and the quotient seminorm on V/W is a norm, and $\|q(v)\| = \|v\|$.

Proof. The triangle inequality shows that N is a linear subspace. Let $x \in V/N$ be a vector with $\|x\| = 0$ and write $x = q(v)$. We have to prove that $x = 0$, or that $v \in N$, or that $\|v\| = 0$.

For each $\epsilon > 0$, we can find, by (1.29), $w \in N$ with

$$\|v + w\| \leq \epsilon.$$

It follows that
$$\|v\| = \|(v+w)-w\| \leq \|v+w\| + \|w\| = \|v+w\| \leq \epsilon,$$
and hence $\|v\| = 0$. □

Here is a more substantial application of the quotient construction.

Proof of Theorem 1.25. (In the lecture, we give a short proof later on). Let $\mathrm{CF}(V)$ be the vector space of all Cauchy sequences in V. If $(v_n)_n \in \mathrm{CF}(V)$, then $(\|v_n\|)_n$ is a Cauchy sequence in $\mathbb{R}$. Define
$$\|(v_n)_n\|_{\mathrm{CF}} := \lim_{n\to\infty} \|v_n\|.$$
This is a seminorm, as it is easily checked.

The null space of $\|\_\|_{\mathrm{CF}}$ is the space $\mathrm{NF}(V) \subset \mathrm{CF}(V)$ of all null sequences. We define
$$W := \mathrm{CF}(V)/\mathrm{NF}(V)$$
with the induced quotient *norm* (see 1.31). Furthermore, we define $\iota : V \to W$ as the composition
$$V \xrightarrow{i} \mathrm{CF}(V) \xrightarrow{q} W$$
of the map which assigns to v the constant sequence with value v with the quotient map. It is clear that ι is a linear isometry.

It remains to prove that $\iota(V)$ is dense and that W is complete. To this end, we make the following observation. Let $x := (v_n)_n \in \mathrm{CF}(V)$, which defines a point $q(x) \in W$. Then
$$q(x) = \lim_n \iota(v_n). \tag{1.32}$$
This is because
$$\|\iota(v_n) - q(x)\|_W = \|i(v_n) - x\|_{\mathrm{CF}} = \lim_m \|v_n - v_m\|$$
implies
$$\lim_n \|\iota(v_n) - q(x)\|_W = \lim_n \lim_m \|v_n - v_m\| = 0$$
as (v_n) is Cauchy. Since each point of W can be written as $q(x)$ for some $x \in \mathrm{CF}(V)$, (1.32) immediately implies that $\iota(V) \subset W$ is dense.

To prove that W is complete, let x_n be a Cauchy sequence in W. By density of the image of ι, we can pick a sequence $y_n \in V$ such that $\|x_n - \iota(y_n)\| \leq \frac{1}{n}$. Since
$$\|\iota(y_n - y_m)\| \leq \|x_n - x_m\| + \frac{1}{n} + \frac{1}{m},$$
$\iota(y_n)$ is a Cauchy sequence and because ι is an isometry, $y := (y_n)_n$ is a Cauchy sequence in V. By 1.32,
$$\lim_n \|\iota(y_n) - q(y)\| = 0.$$
Therefore
$$\|x_n - q(y)\| \leq \|x_n - \iota(y_n)\| + \|\iota(y_n) - q(y)\| \leq \frac{1}{n} + \|\iota(y_n) - q(y)\|$$
tends to 0 as well. □

Lemma 1.33. *Let V be a* normed *space and let $W \subset V$ be a linear subspace. Then the quotient seminorm on V/W is a norm if and only if W is closed.*

Proof. Assume that the quotient seminorm is a norm, and let v_n be a sequence in W with $v = \lim_{n\to\infty} v_n \in V$. But

$$\|(v_n + W) - (v + W)\| \leq \|v_n - v\| \to 0$$

implies that $\|v + W\| = 0$, i.e. that $v \in W$.

Vice versa, let W be closed, and assume that $\|v + W\| = 0$. For each $n \in \mathbb{N}$, there is $w_n \in W$ with $\|v + w_n\| \leq \frac{1}{n}$. But then $-\lim_{n\to\infty} w_n = v$ proves that $v \in W$, i.e. $v + W = 0$. □

Lemma 1.34. *Let V be a Banach space and let $W \subset V$ be a closed linear subspace. Then the normed space V/W is complete.*

Proof. By Lemma 1.16, we have to show that if $v_n + W \in V/W$ is a sequence with $\sum_{n=1}^{\infty} \|v_n + W\| < \infty$, the series $\sum_{n=1}^{\infty} v_n + W$ converges in V/W.

Pick $w_n \in W$ with $\|v_n + w_n\| \leq \frac{1}{2^n}$. Then

$$\sum_{n=1}^{\infty} \|v_n + w_n\| \leq 1 + \sum_{n=1}^{\infty} \|v_n + W\| < \infty,$$

and because V is complete, the sequence

$$u_m := \sum_{n=1}^{m} v_n + w_n \in V$$

converges to some $u \in V$, again by Lemma 1.16. But $u_m + W = \sum_{n=1}^{m} (v_n + W) \in V/W$, and so the latter series converges in V/W, to $u + W$. □

1.4. **Examples I: spaces of continuous functions.** We fill the abstract notions with life and introduce the most important Banach spaces. We have essentially two important classes: spaces of continuous functions on topological spaces, and spaces of measurable functions on measure space.

Definition 1.35. *Let X be a topological space. We denote by $C_b(X;\mathbb{K})$ or just $C_b(X)$ the vector space of all bounded continuous functions $f : X \to \mathbb{K}$. The norm of f is defined by*

$$\|f\|_{C^0} := \sup_{x\in X} |f(x)|.$$

It is easy to see that this defines a norm.

Lemma 1.36. *$C_b(X)$ is complete.*

Proof. Let $(f_n)_n$ be a Cauchy sequence in $C_b(X)$. For each $x \in X$, the sequence $f_n(x))_n$ is Cauchy in $\mathbb{K}$ and hence convergent. Define

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

We show that f_n converges uniformly to f, so let $\epsilon > 0$ and pick n_0, such that for all $m, n \geq n_0$, we have $\|f_n - f_m\| \leq \epsilon$. For each $n \geq n_0$ and $x \in X$, we have

$$|f(x) - f_n(x)| = \lim_{m\to\infty} |f_m(x) - f_n(x)| \leq \limsup_{m\to\infty} \|f_m - f_n\|_{C^0} \leq \epsilon.$$

With the argument known from Analysis I, f is continuous, and we also get $\|f\|_{C^0} = \lim_n \|f_n\|_{C^0} < \infty$. □

The most important case is where X is compact and Hausdorff (the Hausdorff condition is not needed here, but compact and non-Hausdorff spaces do usually not amit many functions to $\mathbb{K}$). In that case, each continuous function $X \to \mathbb{K}$ is bounded and attains its maximum value. It is usual to write

$$C(X; \mathbb{K}) := C_b(X; \mathbb{K})$$

in that case.

An important variation occurs if X is locally compact and Hausdorff.

Definition 1.37. *Let X be locally compact and Hausdorff. We denote by*

$$C_0(X) \subset C_b(X)$$

the subspace of all continuous bounded functions which vanish at infinity. *That is, for each $\epsilon > 0$, the set $\{x | |f(x)| \geq \epsilon\} \subset X$ is compact.*

This is a closed subspace and hence a Banach space on its own right. Why is it closed? Let $f_n \to f$ be a sequence in $C_0(X)$ which converges to a limit in $C_b(X)$. For $\epsilon > 0$, choose n with $\|f - f_n\| \leq \epsilon$ and let $K \subset X$ be compact so that $|f_n(x)| \leq \epsilon$ if $x \notin K$. If $x \notin K$, we have $|f(x)| \leq 2\epsilon$. Altogether $\{x \in X | |f(x)| \geq 3\epsilon\}$ is compact.

The simplest topological spaces are discrete sets (all sets are open). In that case, we write

$$c_0(S) := C_0(S)$$

and

$$\ell^\infty(S) := C_b(S)$$

(we'll learn soon where this notation comes from). When $S = \mathbb{N}$, $c_0(\mathbb{N})$ is the space of all null sequences in $\mathbb{K}$, and $\ell^\infty(\mathbb{N})$ is the space of all bounded sequences.

When dealing with Banach spaces, it is useful to know dense subspaces, in order to apply Theorem 1.21. Here are some useful examples.

(1) The space $C_c(X)$ of compactly supported continuous functions on a locally compact Hausdorff space is a dense subspace of $C_0(X)$. When S is a discrete set, one usually writes

$$c_{00}(S) := C_c(S). \tag{1.38}$$

For $S = \mathbb{N}$, this is the space of all finite sequences (of arbitrary length). Let us prove the density. In the discrete case, let $f \in c_0(S)$ and let $\epsilon > 0$. Then $T := \{s \in S | |f(s)| \geq \epsilon\}$ is compact, hence finite. Let χ_T be the characteristic function of T, in other words

$$\chi_T(s) := \begin{cases} 1 & s \in T \\ 0 & s \notin T. \end{cases}$$

Then $\chi_T f \in c_{00}(S)$, and $\|f - \chi_T f\| \leq \epsilon$. For arbitrary locally compact Hausdorff spaces, recall that if $K \subset X$ is a compact subset and $K \subset U \subset X$, U open, there exists a continuous function $h : X \to [0, 1]$, which is 1 on K, 0 outside U and has compact support. This is a consequence of Urysohn's lemma, see Proposition B.79. Now if $f \in C_0(X)$ and $\epsilon > 0$, let $K \subset X$ be compact so that $|f(x)| \leq \epsilon$ when $x \notin K$. Let h be as above. Then $hf \in C_c(X)$ and $\|f - hf\| \leq \epsilon$.

(2) The *Stone–Weierstrass theorem* (Theorem B.72) plays a crucial role in the study of $C(X)$ when X is compact Hausdorff. It says that if $A \subset C(X)$ is a linear subspace, which contains 1, separates points, is closed under taking products and conjugates, is dense in $C(X)$.
(3) Important examples are the space of polynomials $P \subset C([a, b])$, and the space of trigonometric polynomials $\mathcal{T} \subset C(S^1)$.
(4) The space $C_c^\infty(\mathbb{R}^n) \subset C_0(\mathbb{R}^n)$ of compactly supported, smooth functions is dense.

1.5. **Examples II: spaces of measurable functions.** This is what we want to say about spaces of continuous functions for the moment. The second source of important Banach spaces comes from measure theory. For more details and proofs, we refer to §C.

Let $(X, \mathcal{B}, \mu)$ be a measure space. That is, X is a set, $\mathcal{B}$ is a σ-*algebra* on X and $\mu : \mathcal{B} \to [0, \infty]$ is a measure. A function $f : X \to \mathbb{K}$ is measurable if for each open $U \subset \mathbb{K}$, $f^{-1}(U) \in \mathcal{B}$. For $p \in [1, \infty)$, we define the L^p-*norm* of f by

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

and the L^∞-norm by

$$\|f\|_{L^\infty} := \sup_{x \in X} |f(x)| \in [0, \infty].$$

The most important cases are $p = 1, 2, \infty$. The following two inequalities are important. Let $p, q \in [1, \infty]$ be two numbers with $\frac{1}{p} + \frac{1}{q} = 1$, with the convention that $\frac{1}{\infty} = 0$. These are called *conjugate exponents*. The *Hölder inequality* (Theorem C.37) states that when $f, g : X \to \mathbb{K}$ are measurable, then

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q},$$

and the *Minkowski inequality* (Theorem C.40) states that

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

For $p \in [1, \infty)$, we define

$$\mathscr{L}^p(X, \mathcal{B}, \mu) := \{f : x \to \mathbb{K} | \|f\|_{L^p} < \infty\}.$$

The Minkowski inequality proves that $\mathscr{L}^p(X)$ is a vector space and that $\|\_\|_{L^p}$ is a *seminorm*. The kernel of the seminorm is

$$\mathscr{N}(X, \mathcal{B}, \mu) =: \{f : X \to \mathbb{K} | f \text{ measurable }, \mu(\{x | f(x) \neq 0\}) = 0\},$$

the space of all functions which vanish μ-almost everywhere. The quotient

$$L^p(X, \mathcal{B}, \mu) := \mathscr{L}^p(X, \mathcal{B}, \mu)/\mathscr{N}(X, \mathcal{B}, \mu)$$

is a normed vector space by Proposition 1.31.

For $p = \infty$, we define

$$\mathscr{L}^\infty(X, \mathcal{B}, \mu)$$

as the space of all bounded measurable functions. $\|\_\|_{L^\infty}$ is a *norm* on $\mathscr{L}^\infty$, and not merely a seminorm. Nevertheless, we define

$$L^\infty(X, \mathcal{B}, \mu) := \mathscr{L}^\infty(X, \mathcal{B}, \mu)/\mathscr{N}(X, \mathcal{B}, \mu)$$

with the induced norm. Because a union of countably many null sets is a null set, $\mathscr{N}(X, \mathcal{B}, \mu) \subset \mathscr{L}^\infty(X, \mathcal{B}, \mu)$ is a closed linear subspace, and it follows that $L^\infty(X, \mathcal{B}, \mu)$ is a normed space.

Usually, we write $L^p(X,\mu) := L^p(X,\mathcal{B},\mu)$ or even $L^p(X)$ or $L^p(\mu)$, depending on the context. Sometimes, we need to specify the field $\mathbb{K}$, in which case we use notations such as $L^p(X,\mathbb{K})$ or $L^p(X,\mu,\mathbb{K})$ or the like.

Remark 1.39. *We note that the elements of $L^p(X,\mu)$ are not functions on X, but equivalence classes of such functions, with $f \equiv g$ if and only if they agree μ-almost everywhere. We shall usually be sloppy about this distinction and refer to elements of $L^p(X,\mu)$ as functions. When doing so, one must of course keep in mind that not all concepts one is used to when dealing with functions make sense for equivalence classes. For example, the expression "$f(x)$" does not at all make sense when $f \in L^p(X,\mu)$.*

Let $f_n \in L^p(X,\mu)$ be a sequence and let $g: X \to \mathbb{K}$ be a measurable function. We say that f_n converges pointwise almost everywhere *to g, if for any choice of representatives $g_n \in \mathscr{L}^p(X,\mu)$ with $[g_n] = f_n$, there is a measurable set $S \subset X$ with $\mu(S) = 0$ such that $\lim_n g_n(x) = g(x)$ holds for all $x \in X \setminus S$. (this notion does not depend on a specific chice of representatives g_n: if h_n is another choice of representatives, the set $S_n := \{x | g_n(x) \neq h_n(x)\}$ has measure zero, and so does $\bigcup_{n=1}^\infty S_n \cup S$, and outside this set, $h_n(x) \to g(x)$).*

Let us recall the fundamental result of integration theory.

Theorem 1.40 (Theorems C.42 and C.43)**.** *Let $(X,\mathcal{B},\mu)$ be a measure space and $p \in [1,\infty]$. Let f_n be a sequence in $L^p(X,\mu)$.*

(1) *If f_n is a Cauchy sequence, there is a subsequence f_{n_m} such that f_{n_m} converges almost everywhere to a function $f \in L^p(X,\mu)$, and $\|f_n - f\|_{L^p} \to 0$. In particular, $L^p(X,\mu)$ is a Banach space.*
(2) *Let $p < \infty$. If $g: X \to [0,\infty]$ is a measurable function with $\int_X g(x)^p d\mu < \infty$, and if $|f_n| \leq g$ for all n, and if f_n converges pointwise to a function f, then $f \in L^p(X,\mu)$, $(f_n)_n$ is a Cauchy sequence, and $\|f_n - f\|_{L^p} \to 0$ (dominated convergence theorem).*

Important dense linear subspaces of $L^p(X,\mu)$ are the spaces of step functions. For $S \subset X$, the characteristic function χ_S of S is

$$\chi_S(x) := \begin{cases} 1 & x \in S \\ 0 & x \notin S. \end{cases}$$

Let $\mathrm{St}(X)$ be the space of all finite linear combinations of characteristic functions χ_S, $S \in \mathcal{B}$. Then $\mathrm{St}(X) \subset L^\infty(X,\mu)$ is dense (Proposition C.48).

Moreover, let $\mathrm{St}_f(X) \subset \mathrm{St}(X)$ be the subspace of all sums $\sum_{j=1}^n a_j \chi_{S_j}$, where S_j has *finite* measure. Then $\mathrm{St}_f(X) \subset L^p(X)$ is a dense linear subspace (Proposition C.48).

The definition of $L^p(X,\mu)$ and the proof of Theorem 1.40 uses integration, but only the integral of nonnegative functions. With the help of the dense linear subspace $\mathrm{St}_f(X)$, we can define the integral

$$\int_X : L^1(X,\mu) \to \mathbb{K}$$

as follows. For a step function $f = \sum_{j=1}^n a_j \chi_{S_j}$, we define

$$\int_X f(x) d\mu(x) := \sum_{j=1}^n a_j \mu(S_j).$$

The map $\int_X$ is linear. This is not completely obvious because a step function can be written as linear combination of characteristic functions in many different ways, but rather boring: one writes f as a characteristic function of pairwise disjoint sets.

Furthermore, if all S_j are disjoint, then

$$|\int_X f(x)d\mu(x)| \leq \sum_{j=1}^{n} |a_j|\mu(S_j) = \int_X |f(x)|d\mu(x) = \|f\|_{L^1}.$$

Therefore, by Theorem 1.21, $\int_X$ extends to a bounded functional $L^1(X,\mu) \to \mathbb{K}$, which we also denote by $\int_X$. There is usually *no* way to define $\int_X f(x)d\mu$ when $f \in L^p(X,\mu)$, $p > 1$.

Example 1.41. *The simplest measures are the* counting measures*. That is, let S be a set. The whole power set $\mathcal{P}(S)$ is a σ-algebra and we define*

$$\mu : \mathcal{P}(X) \to [0,\infty],\ \mu(T) := |T|,$$

the number of elements of T. It is customary to write

$$\ell^p(S) := L^p(S).$$

Note that every nonempty subset of S has positive measure, and so $\mathscr{N}(S,\mu) = \{0\}$. Therefore, the elements of $\ell^p(S)$ are really functions $f : S \to \mathbb{K}$, not equivalence classes.

The space $\ell^\infty(S)$ is the same as the space $C_b(S)$ of bounded continuous functions on S.

The space $\mathrm{St}_f(S) \subset \ell^p(S)$ is exactly the space of all functions $S \to \mathbb{K}$ with finite support, which is exactly the space $c_{00}(S)$ introduced in 1.38.

More interesting measures are hard to construct. The Lebesgue measure on $\mathbb{R}^n$ is known from Analysis III, see also C.5. Let us put this example in a greater context.

Definition 1.42. *Let X be a topological space. The* Borel-σ-algebra *is the smallest σ-algebra which contains all open subsets of X. A* Borel measure *on X is a measure which is defined on the Borel σ-algebra.*

Definition 1.43. *Let X be a locally compact Hausdorff space. A* Radon measure *is a Borel measure μ on X such that*

(1) *$\mu(K) < \infty$ for all compact $K \subset X$ (μ is locally finite),*
(2) *for each Borel set $S \subset X$, we have*

$$\mu(S) = \sup_{K \subset S\ compact} \mu(K) = \inf_{S \subset U\ open} \mu(U)$$

(μ is inner and outer regular).

Under favorable circumstances, regularity holds automatically: if X is second countable, then each locally finite Borel measure on X is a Radon measure, see Proposition C.59.

Proposition 1.44 (Proposition C.58)**.** *Let X be a locally compact Hausdorff space and let μ be a Radon measure on X. Then $C_c(X) \subset L^p(X,\mu)$ for each p, and if $p < \infty$, then this is a dense linear subspace.*

Examples 1.45. (1) *On every discrete space S, the counting measure is a Radon measure.*

(2) *The Lebesgue measure λ on $\mathbb{R}^n$ is a Radon measure. The Lebesgue measure has the important feature that it is translation invariant, and it is (up to multiplication by a constant factor) the only such measure on $\mathbb{R}^n$.*
(3) *For each $f \in L^1(\mathbb{R}^n, \lambda)$, the measure $f\lambda$ defined by*
$$f\lambda(S) := \int_S f d\lambda$$
is a Radon measure.

Now let X be a locally compact Hausdorff space and let μ be a Radon measure on X. Then since $C_c(X) \subset L^1(X, \mu)$, we have the linear functional
$$\int_X \_d\mu : C_c(X) \to \mathbb{K}$$
which sends f to $\int_X f d\mu$. This functional is *positive*, which means that
$$\int_X f d\mu \geq 0$$
whenever $f \geq 0$.

The space $C_c(X)$ has the supremum norm, but $\int_X \_d\mu$ is usually *not* bounded. If, however, μ is finite, e.h. $\mu(X) < \infty$, then by Hölder's inequality
$$|\int_X f d\mu| = \|f\|_{L^1} \leq \|1\|_{L^1} \|f\|_{L^\infty} = \mu(X) \|f\|_{C^0}.$$
Since $C_c(X) \subset C_0(X)$ is dense, the integral extends to a bounded linear functional on $C_0(X)$ in this case. We now state one of the main results of this course; the proof is beyond our capabilities yet.

Theorem 1.46 (Riesz-Markov-Kakutani representation theorem)**.** *Let X be a locally compact Hausdorff space. Let $F : C_c(X; \mathbb{K}) \to \mathbb{K}$ be a positive functional (not necessarily bounded). Then there exists a unique Radon measure μ on X such that*
$$\int_X f(x) d\mu(x) = F(f)$$
for all $f \in C_0(X, \mathbb{K})$.

Example 1.47. *The Riemann integral gives a positive functional $F : C_c(\mathbb{R}) \to \mathbb{K}$, $F(f) := \int_{-\infty}^{\infty} f(x) dx$. Theorem 1.46 provides a Radon measure on $\mathbb{R}$, which is nothing else than the Lebesgue measure.*

Example 1.48. *The functional $\mathrm{ev}_x : C_c(X) \to \mathbb{K}$, $f \mapsto f(x)$ is positive. The corresponding measure is the Dirac measure at x, which takes a set $S \subset X$ to 1 if $x \in S$ and to 0 otherwise.*

The space $C_c(X)$ has the supremum norm $\|f\|_{C^0}$; $C_c(X)$ is *not* complete with respect to that norm unless X is compact. The space $C_c(X)$ is a dense linear subspace of $C_0(X)$. A positive functional on $C_c(X)$ is usually not bounded, as the example of the Lebesgue measure shows.

Lemma 1.49. *Let X be a locally compact Hausdorff space and let μ be a Radon measure on X and let $F_\mu : C_c(X) \to \mathbb{K}$ be the integration functional. Then F_μ is bounded if and only if $\mu(X) < \infty$, and in that case*
$$\|F_\mu\| = \mu(X).$$

Proof. It suffices to prove that

$$\mu(X) = \sup_{K \subset X \text{compact}} \mu(K) \leq \sup_{0 \leq f \leq 1, f \in C_c(X)} \int_X f d\mu \leq \|F_\mu\| \leq \mu(X).$$

The first equality is immediate from the definition of a regular measure. If $K \subset X$ is compact, there is $f \in C_c(X)$ with $0 \leq f \leq 1$ and $f|_K = 1$. Then $\mu(K) \leq \int_X f d\mu$, from which the first inequality follows. If $0 \leq f \leq 1$, then

$$\int_T f d\mu = F_\mu(f) = |F_\mu(f)| \leq \|F_\mu\| \|f\| \leq \|F_\mu\|.$$

This proves the second inequality, and the third holds because if $f \in C_c(X)$ and $\|f\| \leq 1$, then $0 \leq |f| \leq 1$ and so

$$|F_\mu(f)| = |\int_X f d\mu| \leq \int_X |f| d\mu \leq \mu(X)$$

where we used the first equation in the last step. □

Example 1.50. *Consider $X = S^1$ and the map $p : \mathbb{R} \to S^1$, $p(t) := e^{2\pi i t}$. We define a functional $F : C(S^1) \to \mathbb{K}$ by*

$$F(f) := \int_0^1 f(e^{2\pi i t}) dt.$$

This is obviously positive. The Riesz-Markov-Kakutani theorem yields a Radon measure μ on S^1. Observe that $\mu(S^1) = F(1) = 1$. This measure has an important property: it is translation-invariant.

To formulate this, let $z \in S^1$ and define for $f \in C(S^1)$ the translated function

$$T_z f(y) := f(z^{-1} y).$$

Then $F(T_z f) = F(f)$ for all z and f. To see this, pick $s \in [0,1]$ with $e^{2\pi i s} = z$. Then

$$F(T_z f) = \int_0^1 f(e^{2\pi i(t-s)}) dt = \int_{-s}^{1-s} f(e^{2\pi i u}) du = \int_{-s}^0 f(e^{2\pi i t}) dt + \int_0^{1-s} f(e^{2\pi i t}) dt =$$

$$= \int_{1-s}^1 f(e^{2\pi i t}) dt + \int_0^{1-s} f(e^{2\pi i t}) dt = F(f).$$

The measure μ is the normalized Haar measure *on the circle.*

Let us give some context for this construction.

Definition 1.51. *A* locally compact group *G is a group which is equipped with a topology such that*

1. *the multiplication map $G \times G \to G$ is continuous,*
2. *the map $G \to G$, $g \mapsto g^{-1}$ is continuous,*
3. *G is locally compact and Hausdorff.*

For a function $f \in C_c(G)$ and $g \in G$, we define $T_g f(x) := f(g^{-1} x)$.

Definition 1.52. *Let G be a locally compact group. A* Haar measure *μ on G is a Radon measure which is not identically 0 and satisfies the following two equivalent conditions:*

1. *for each $g \in G$ and each Borel set $S \subset G$, we have $\mu(gS) = \mu(S)$,*
2. *for each $g \in G$ and each $f \in C_c(G)$, we have $\int_G f(x) d\mu(x) = \int_G T_g f(x) d\mu(x)$.*

Theorem 1.53 (Existence of Haar measure)**.** *Every locally compact group has a Haar measure which is unique up to multiplication by a positive constant.*

The proof of that result is beyond our scope (we develop enough theory to give a proof for abelian G).

1.6. **Duality relation between the sequence spaces.** Let S be a set. For $1 \leq p \leq q < \infty$, we have the inclusions

$$c_{00}(S) \subset \ell^1(S) \subset \ell^p(S) \subset \ell^q(S) \subset c_0(S) \subset \ell^\infty(S).$$

To see this, let $p \leq q$. Then, if $f \in c_{00}(S)$, we have

$$\|f\|_{\ell^q} = (\sum_s |f(s)|^q)^{1/q} = \|f\|_{\ell^p}(\sum_s (\frac{|f(s)|}{\|f\|_{\ell^p}})^q)^{1/q} \leq$$

(this inequality follows, as $0 \leq \frac{|f(s)|}{\|f\|_{\ell^p}} \leq 1$, and for $x \in [0,1]$, we have $x^q \leq x^p$)

$$\leq \|f\|_{\ell^p}(\sum_s (\frac{|f(s)|}{\|f\|_{\ell^p}})^p)^{1/q} = \frac{\|f\|_{\ell^p}}{\|f\|_{\ell^p}^{\frac{p}{q}}}(\sum_s |f(s)|^p)^{1/q} = \|f\|_{\ell^p}.$$

Therefore, the identity $(c_{00}(S), \|\_\|_{\ell^p}) \to (c_{00}(S), \|\_\|_{\ell^q})$ has operator norm ≤ 1 and so extends to a bounded operator $\ell^p(S) \to \ell^q(S)$. It is easy to see that this is injective.

Similarly, for $f \in c_{00}(S)$ and $s \in S$, we have

$$|f(s)| = (|f(s)|^p)^{1/p} \leq (\sum_s |f(s)|^p)^{1/p} = \|f\|_{\ell^p},$$

so that $\|f\|_{c_0} \leq \|f\|_{\ell^p}$. Finally, the inclusion $c_0(S) \subset \ell^\infty(S)$ is clear (and it is an isometry, unlike the other inclusions).

Exercise 1.54. *Let $p < q < \infty$ and let S be infinite. Prove that there is no $c > 0$ with $\|f\|_{\ell^p} \leq c\|f\|_{\ell^q}$ for all f, and that there is no $c > 0$ with $\|f\|_{\ell^p} \leq \|f\|_{c_0}$ for all f.*

The right way to think about the relation between these spaces is in terms of *duality*. We will now prove the following three results.

Theorem 1.55. *There is an isometric isomorphism*

$$\Phi : \ell^1(S) \to c_0(S)'$$

constructed explicitly in the course of the proof.

Theorem 1.56. *There is an isometric isomorphism*

$$\Phi : \ell^\infty(S) \to \ell^1(S)'$$

constructed explicitly in the course of the proof.

Theorem 1.57. *Let $1 < p, q < \infty$ be conjugate exponents ($\frac{1}{p} + \frac{1}{q} = 1$). Then there is an isometric isomorphism*

$$\Phi : \ell^q(S) \to \ell^p(S)'$$

constructed explicitly in the course of the proof.

Let us introduce the notation

$$\delta_s : S \to \mathbb{K},\ \delta_s(t) := \begin{cases} 1 & t = s \\ 0 & t \neq s. \end{cases}$$

Proof of Theorem 1.55. Let $f \in \ell^1(S)$ and $g \in \ell^\infty(S)$. By the Hölder inequality, we have

$$\sum_{s\in S} |f(s)g(s)| = \|fg\|_{\ell^1} \le \|f\|_{\ell^1}\|g\|_{\ell^\infty}.$$

It follows that the sum $\sum_{s\in S} f(s)g(s)$ is absolutely convergent. In particular, this holds when g is an element of the closed linear subspace $c_0(S) \subset \ell^\infty(S)$. We define Φ by the formula

$$\Phi(f)(g) := \sum_{s\in S} f(s)g(s) \in \mathbb{K}$$

Then $|\Phi(f)(g)| \le \|f\|_{\ell^1}\|g\|_{c_0}$, and hence $\Phi(f) \in c_0(S)'$ with $\|\Phi(f)\| \le \|f\|_{\ell^1}$. It follows that

$$\Phi : \ell^1(S) \to c_0(S)'$$

is a bounded operator of norm $\|\Phi\| \le 1$.

We claim that Φ is an isometric isomorphism. This involves two things: $\|\Phi(f)\| = \|f\|_{\ell^1}$, and that Φ is surjective.

Let us first show that $\|\Phi(f)\| = \|f\|_{\ell^1}$. We have already proven that $\|\Phi(f)\| \le \|\Phi\|\|f\| \le \|f\|_{\ell^1}$. For the reverse inequality, let $\epsilon > 0$. Pick a finite subset $T \subset S$, so that

$$\|f\|_{\ell^1} \ge \sum_{s\in T} |f(s)| \ge \|f\|_{\ell^1} - \epsilon.$$

Let $a_s \in S^1$ be a complex number so that $a_s f(s) \ge 0$. Define $g \in c_0(S)$ by the formula

$$g(s) := \begin{cases} a_s & s \in T \\ 0 & s \notin T. \end{cases}$$

Then $\|g\|_{c_0} \le 1$, and

$$|\Phi(f)(g)| = |\sum_{s\in T} f(s)a_s| = \sum_{s\in T} f(s)a_s = \sum_{s\in T} |f(s)| \ge \|f\|_{\ell^1} - \epsilon.$$

This proves that $\|\Phi(f)\| \ge \|f\|_{\ell^1} - \epsilon$, and since ϵ was arbitrary, that $\|\Phi(f)\| = \|f\|$.

For the surjectivity of Φ, let $F \in c_0(S)'$ be a bounded functional. We define a map $f : S \to \mathbb{K}$ by

$$f(s) := F(\delta_s),$$

and first claim that $f \in \ell^1(S)$. To prove this, let $T \subset S$ be an arbitrary finite subset. Pick $b_s \in S^1$ with $f(s)b_s \ge 0$. Then

$$\sum_{s\in T} |f(s)| = \sum_{s\in T} f(s)b_s = F(\sum_{s\in T} b_s\delta_s) = |F(\sum_{s\in T} b_s\delta_s)| \le \|F\|\|h\|_{c_0},$$

where $h \in c_{00}(S)$ is the function $h(s) := b_s$ if $s \in T$ and $h(s) = 0$ if $s \notin T$. Clearly $\|h\|_{c_0} \le 1$, so that

$$\sum_{s\in T} |f(s)| \le \|F\|$$

for each finite $T \subset S$, so that $\|f\|_{\ell^1} \le \|F\|$.

Now let $g \in c_{00}(S)$. Since

$$F(g) = \sum_{s\in S} g(s)F(\delta_s) = \sum_{s\in S} g(s)f(s)$$

(only finitely many terms in the sum are finite!), we see that

$$\Phi(f)(g) = F(g)$$

for all $g \in c_{00}(S)$, and hence by continuity that $\Phi(f) = F$. $\square$

Remark 1.58. *The formula for Φ in Theorem 1.55 really defines an isometry*

$$\Phi : \ell^1(S) \to \ell^\infty(S)'.$$

However, this map is far from being surjective. To prove that this is the case, more theory (and the axiom of choice) is needed.

We now turn to Theorems 1.56 and 1.57. These have analogues for measure spaces, which are isometric isomorphisms

$$L^q(X,\mu) \to L^p(X,\mu)'$$

and

$$L^\infty(X,\mu) \to L^1(X,\mu)'$$

(the latter only if μ is σ-finite). The first half of the proofs is the same for both, the discrete and the general case. To get started, we need a converse to the Hölder inequality (which is very useful in some other contexts, when L^p-norms need to be estimated).

Proposition 1.59 (Reverse Hölder inequality)**.** *Let (X,μ) be a measure space and let $f : X \to \mathbb{K}$ be measurable, and let p,q be conjugate exponents. When $p = \infty$, assume that X is locally finite. Then $f \in L^p(X)$ if and only if*

$$\sup_{\|g\|_{L^q} \leq 1} |\int_X fg d\mu| < \infty$$

and in that case

$$\|f\|_{L^p} = \sup_{\|g\|_{L^q} \leq 1} |\int_X fg d\mu| \in [0,\infty].$$

Proof. Let $C_f := \sup_{\|g\|_{L^q} \leq 1} |\int_X fg d\mu|$.

The Hölder inequality shows that $C_f \leq \|f\|_{L^p}$, and we only have to prove that $\|f\|_{L^p} \leq C_f$. We define $a : X \to \mathbb{K}$ by

$$a(x) = \begin{cases} \frac{|f(x)|}{f(x)} & f(x) \neq 0 \\ 1 & f(x) = 0. \end{cases}$$

Then $a \in L^\infty(X,\mu)$, $\|a\|_{L^\infty} = 1$ and $af = |f| \geq 0$.

For $p = 1$, let $g := a$ and observe that

$$C_f \geq |\int_X fad\mu| = \int_X fad\mu = \int_X |f| d\mu = \|f\|_{L^1}.$$

For $p = \infty$, let $0 < R < \|f\|_{L^\infty}$ and consider $S := \{x \in X | |f(x)| \geq R\}$. By definition of the L^∞-norm, $\mu(S) > 0$, and if μ is locally finite, we find $T \subset S$ with $0 < \mu(T) < \infty$. Put $g := \frac{1}{\mu(T)} a\chi_T$. Then

$$\|g\|_{L^1} = 1$$

and

$$C_f \geq |\int_X fd\mu| = \frac{1}{\mu(T)} \int_T |f| d\mu = \frac{1}{\mu(T)} \mu(T) R = R.$$

Therefore $R \leq C_f$. This holds for each $R < \|f\|_{L^\infty}$, and therefore $\|f\|_{L^p} \leq C_f$.

It remains the case $1 < p < \infty$. Assume first $\|f\|_{L^p} < \infty$. Set

$$h(x) := a(x)|f(x)|^{\frac{p}{q}}.$$

Then

$$\|h\|_{L^q}^q = \int_X |f|^p d\mu = \|f\|_{L^p}^p < \infty,$$

and for

$$g(x) := \frac{h(x)}{\|h\|_{L^q}} = \frac{h(x)}{\|f\|_{L^p}^{\frac{p}{q}}}$$

we have $\|g\|_{L^q} = 1$, $fg \geq 0$ and so

$$C_f \geq |\int_X fgd\mu| = \int_X fgd\mu = \frac{1}{\|f\|_{L^p}^{\frac{p}{q}}} \int_X |f|^{1+\frac{p}{q}} d\mu.$$

But $1 + \frac{p}{q} = p(\frac{1}{p} + \frac{1}{q}) = p$ and $\frac{p}{q} = p - 1$ and so

$$\int_X fg = \frac{1}{\|f\|_{L^p}^{\frac{p}{q}}} \int_X |f|^p d\mu = \|f\|_{L^p}^{1-p+p} = \|f\|_{L^p},$$

proving that $\|f\|_{L^p} \leq C_f$ provided that $\|f\|_{L^p} < \infty$.

It remains to prove that

$$\|f\|_{L^p} = \infty \Rightarrow C_f = \infty.$$

Define

$$f_n(x) := \begin{cases} f(x) & \frac{1}{n} \leq |f(x)| \leq n \\ 0 & \text{otherwise.} \end{cases}$$

Then $\|f_n\|_{L^p} < \infty$ and $0 \leq |f_1| \leq |f_2| \leq \ldots \to |f|$, and by the monotone convergence theorem,

$$\lim_{n\to\infty} \|f_n\|_{L^p}^p = \|f\|_{L^p}^p = \infty.$$

It hence suffices to prove that $\|f_n\|_{L^p} \leq C_f$ for each n. By what we already proved, there is $g \in L^q(X)$ with $\|g\|_{L^q}$ and $|\int_X f_n g d\mu| = \|f_n\|_{L^p}$. On the other hand

$$|\int_X f_n g d\mu| \leq \int_X |f_n g| d\mu = \int_X f_n a|g| d\mu \overset{!}{\leq} \int_X fa|g| d\mu \leq C_f$$

(the second inequality holds because $f_n a|g| = |f_n||g| \leq |f||g| = fa|g|$, and the third because $\|a|g|\|_{L^q} = \|g\|_{L^q}$). □

Corollary 1.60. *Let $1 \leq p, q \leq \infty$ be two conjugate exponents, and assume that μ is locally finite when $p = 1$. Then the formula*

$$\Phi : L^q(X, \mu) \to L^p(X, \mu)',\ f \mapsto (g \mapsto \int_X fgd\mu) \tag{1.61}$$

defines an isometry.

Proof. It is by now clear that Φ is well-defined, linear, and bounded with $\|\Phi(f)\| \leq \|f\|_{L^q}$. By the reverse Hölder inequality

$$\|\Phi(f)\| = \sup \|g\|_{L^p} \leq 1 |\int_X fgd\mu| = \|f\|_{L^q},$$

so that Φ is an isometry (if $p = 1$, we need that μ is locally finite). □

Proof of Theorem 1.56. By Corollary 1.60, the formula
$$\Phi(f)(g) := \sum_{s\in S} f(s)g(s)$$
defines an isometry
$$\Phi : \ell^\infty(S) \to \ell^1(S)'$$
(note that the counting measure on S is obviously locally finite). It remains to prove that Φ is surjective, so let $F \in \ell^1(S)'$ be a given linear bounded functional. Define a function $f : S \to \mathbb{K}$ by
$$f(s) := F(\delta_s).$$
Since
$$|F(\delta_s)| \leq \|F\|\|\delta_s\|_{\ell^1} = \|F\|$$
for all $s \in S$, f is an element of $\ell^\infty(S)$. For each $g \in c_{00}(S)$, we can write
$$g = \sum_{s\in S} g(s)\delta_s$$
(which is truely a finite sum), and so
$$F(g) = \sum_{s\in S} g(s)F(\delta_s) = \sum_{s\in S} g(s)f(s) = \Phi(f)(g).$$
Hence we have
$$\Phi(f)(g) = F(g)$$
for all g in the dense subspace $c_{00}(S) \subset \ell^1(S)$, and hence by continuity of $\Phi(f)$ and F also for all $g \in \ell^1(S)$. □

Proof of Theorem 1.57. By Corollary 1.60, the formula
$$\Phi(f)(g) := \sum_{s\in S} f(s)g(s)$$
defines an isometry
$$\Phi : \ell^q(S) \to \ell^p(S)'.$$
We claim that Φ is surjective. To that end, let $F \in \ell^p(S)'$ be a functional. Define $f : S \to \mathbb{K}$ by
$$f(s) := F(\delta_s).$$
We claim that $f \in \ell^q(S)$. If that is proven, it follows as in the proof of Theorem 1.56 that $\Phi(f)(g) = F(g)$ for all g in the dense subspace $c_{00}(S) \subset \ell^p(S)$, and hence by continuity for all $g \in \ell^p(S)$.

To prove that $\|f\|_{\ell^q} < \infty$, note that
$$\|f\|_{\ell^q} = \sup_{T\subset S \text{finite}} \|\chi_T f\|_{\ell^q}.$$
By the reverse Hölder inequality
$$\|\chi_T f\|_{\ell^q} = \sup_{\|g\|_{\ell^p}\leq 1} |\sum_{s\in T} f(s)g(s)|,$$
and since
$$|\sum_{s\in T} f(s)g(s)| = |\sum_{s\in T} F(\delta_s)g(s)| = |F(\sum_{s\in T}\delta_s g(s))| \leq \|F\|\|\sum_{s\in T}\delta_s g(s)\|_{\ell^p} =$$

$$= \|F\|(\sum_{s\in T}|g(s)|^p)^{\frac{1}{p}} \le \|F\|\|g\|_{\ell^p},$$

it holds that

$$\|\chi_T f\|_{\ell^q} \le \|F\|$$

and hence that

$$\|f\|_{\ell^q} \le \|F\|,$$

as claimed. □

1.7. **$*$-Absolute and unconditional convergence.** We often have to deal with infinite series whose terms are elements in a Banach space V. We begin with a recapitulation from Analysis I. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. One defines

$$\sum_{n=1}^{\infty} a(n) := \lim_m \sum_{n=1}^{m} a(n) \in V$$

if that limit exists. Sometimes, one has a better type of convergence: *absolute convergence*, which means that

$$\sum_{n=1}^{\infty} |a(n)| < \infty$$

(or, equivalently: there is C such that $\sum_{n=1}^{m} |a(n)| \le C$ for all m). We have the following basic result from Analysis I:

Theorem 1.62. *Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. Then*

(1) *if $\sum_{n\in\mathbb{N}} a(n)$ is absolutely convergent, then $\lim_{m\to\infty} \sum_{n=1}^{m} a(n)$ exists.*
(2) *More generally, if $\sum_{n\in\mathbb{N}} a(n)$ is absolutely convergent and $F_1 \subset F_2 \subset \ldots \subset \mathbb{N}$ is a sequence of finite subsets which is* exhausting *(in other words $\mathbb{N} = \bigoplus_{m=1}^{\infty} F_m$), then $\lim_{m\to\infty} \sum_{n\in F_m} a(n)$ exists, and is equal to $\lim_{m\to\infty} \sum_{n=1}^{m} a(n)$.*
(3) *Conversely, if the limit $\lim_{m\to\infty} \sum_{n\in F_m} a(n)$ exists for all exhausting sequences of finite subsets $F_1 \subset F_2 \subset \ldots$, then the limit is the same for each such sequence $(F_m)_m$ and $\sum_{n\in\mathbb{N}} a(n)$ converges absolutely.*

Condition (2) expresses the rearrangement property: if $\varphi : \mathbb{N} \to \mathbb{N}$ is bijective and $\sum_{n\in\mathbb{N}} a(n)$ is absolutely convergent, then $\sum_{n=1}^{\infty} a(\varphi(n))$ converges, and the limit does not depend on the choice of φ.

Proof. (1) is one of the main results of Analysis I. (2) is most conveniently shown using the dominated convergence theorem: consider the measure space $\mathbb{N}$; then absolute convergence of a means the same as $a \in \ell^1(S)$, and the sequence $\chi_{F_m} a$ converges pointwise to a and is dominated by a. (3): we show that if $\sum_{n\in\mathbb{N}} a(n)$ does not converge absolutely, there is an exhausting sequence of finite subsets (F_m) such that $\lim_{m\to\infty} \sum_{n\in F_m} a(n)$ does not converge. This will prove (3).

So suppose that $\sum_{n\in\mathbb{N}} a(n)$ does not converge absolutely. Without loss of generality $a(n) \ne 0$ for all n. Let $I_\pm = \{n \in \mathbb{N} | \pm a(n) > 0$. Then I_+ and I_- are disjoint subsets. At least one of the sums $\sum_{n\in I_\pm} \pm a(n)$ must be $+\infty$; assume that $\sum_{n\in I_\pm} a(n) = +\infty$. One can build a sequence of finite subsets $F_1 \subset F_2 \subset \ldots$ with $\bigcup_{m=1}^{\infty} F_m = \mathbb{N}$ such that $\sum_{n\in F_m} a(n) \ge m$. □

The proof of (3) very much depended on the order structure in $\mathbb{R}$, and one should expect this step to fail when $\mathbb{R}$ is replaced by a general Banach space.

Definition 1.63. *Let V be a Banach space, let S be a countable*[1] *set and let $a : S \to V$ be a map. We say that the series $\sum_{s\in S} a(s)$ is*

(1) absolutely convergent *if $\sum_{s\in S} \|v_s\| < \infty$.*
(2) unconditionally convergent *if for each $\epsilon > 0$, there is a finite $F \subset S$ such that for each finite $F \subset G \subset S$, $\| \sum_{s\in G\setminus F} v_s\| \leq \epsilon$.*
(3) unconditionally convergent *to $v \in V$ if for each $\epsilon > 0$, there is a finite $F \subset S$ such that for each finite $F \subset G \subset S$, $\|v - \sum_{s\in G} v_s\| \leq \epsilon$.*

Theorem 1.64. *Let V be a Banach space, let S be a countable set and let $a : S \to V$ be a map.*

(1) *If $\sum_{s\in S} a(s)$ converges absolutely, it converges unconditionally.*
(2) *If $\sum_{s\in S} a(s)$ converges unconditionally, it converges unconditionally to a unique $v \in V$.*

There is no parallel to the third part of Theorem 1.62 in general Banach space. For example, let $S = \mathbb{N}$, $V = \ell^2(\mathbb{N})$ and $a(n) = \frac{1}{n}\delta_n$. $\sum_{n\in\mathbb{N}} a(n)$ converges unconditionally to $a \in \ell^2(\mathbb{N})$, $a(n) = \frac{1}{n}$, but not absolutely.

Proof. (1) Absolute convergence of $\sum_{s\in S} a(s)$ can be expressed by saying that for each $\epsilon > 0$, there is a finite $F \subset S$ such that for each finite $F \subset G \subset S$, we have $\sum_{s\in G\setminus F} \|a(s)\| \leq \epsilon$. The unconditional convergence thus simply follows from the triangle inequality.

(2) Assume that $\sum_{s\in S} a(s)$ converges unconditionally. Choose finite $F_n \subset S$ so that if $F_n \subset G$ and G is finite, then $\| \sum_{s\in G-F_n} a(s)\| \leq \frac{1}{n}$. In particular $\|a(s)\| \leq \frac{1}{n}$ if $s \notin F_n$.

We can also assume that $F_1 \subset F_2 \subset F_3 \subset \ldots$. If $s \in S \setminus \bigcup_{n=1}^\infty F_n$, then $\|a(s)\| = 0$, and we may discard the set $S \setminus \bigcup_{n=1}^\infty F_n$ entirely, and may well assume that $\bigcup_{n=1}^\infty F_n = S$. The sequence

$$v_n := \sum_{s\in F_n} a(s)$$

is a Cauchy sequence, which converges to some $v \in V$. The series $\sum_{s\in S} a(s)$ converges unconditionally to v, more or less by definition. A routine argument shows that if $\sum_{s\in S} a(s)$ converges unconditionally to v and w, then $\|v - w\| \leq 2\epsilon$ for each $\epsilon > 0$, proving uniqueness of the unconditional limit. □

1.8. **Notes.** Most of the material in this chapter is standard material and can be found in every textbook on functional analysis. The exception is the part about absolute and unconditional convergence. I did work it out by myself.

[1] to make things a little simpler

2. Hilbert spaces

2.1. Inner products.

Definition 2.1. *Let V be a $\mathbb{K}$-vector space. An* inner product *on V is a map*

$$\langle \_, \_ \rangle : V \times V \to \mathbb{K}$$

such that

1. *$\langle \_, \_ \rangle$ is $\mathbb{R}$-bilinear*
2. *$\langle \_, \_ \rangle$ is $\mathbb{K}$-linear in the second variable and $\mathbb{K}$-antilinear in the first one,*
3. *$\langle \_, \_ \rangle$ is* symmetric*, i.e. $\langle v, w \rangle = \overline{\langle w, v \rangle}$ for all $v, w \in V$ and*
4. *$\langle \_, \_ \rangle$ is* positive*, i.e. $\langle v, v \rangle \geq 0$ for all $v \in V$, and*
5. *$\langle \_, \_ \rangle$ is* definite*, i.e. $\langle v, v \rangle = 0$ only holds for $v = 0$.*

We put

$$\|v\| := \sqrt{\langle v, v \rangle} \in [0, \infty).$$

Example 2.2. *Let (X, μ) be a measure space. On $L^2(X, \mu; \mathbb{K})$, we define*

$$\langle f, g \rangle := \int_X \overline{f(x)} g(x) d\mu(x).$$

This is an inner product, and the induced norm is just the L^2-norm. If we would have considered $\mathscr{L}^2(X; \mu)$ instead, we would only have gotten a semi-inner product. Then $\sqrt{\langle f, f \rangle} = \|f\|_{L^2}$.

If $X = S$ is a set with the counting measure, this becomes the inner product

$$\langle f, g \rangle = \sum_{s \in S} \overline{f(s)} g(s)$$

on $\ell^2(S)$. If $S = \underline{n}$, we get $\mathbb{K}^n$, with the standard inner product familiar from linear algebra II.

Let us first prove that $\|\_\|$ is indeed a norm on V. Because

$$\|zv\|^2 = \langle zv, zv \rangle = z\overline{z} \langle v, v \rangle = |z|^2 \|v\|^2$$

and

$$\|v\| = 0 \Rightarrow \langle v, v \rangle = 0 \Rightarrow v = 0,$$

only the triangle inequality needs to be proven.

Theorem 2.3. *Let $\langle \_, \_ \rangle$ be an inner product on V.*

1. *(Cauchy-Schwarz inequality) For all $v, w \in V$, we have*
$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$
2. *If $|\langle v, w \rangle| = \|v\| \|w\|$ and $\|v\| \neq 0$, then w is a multiple of v.*
3. *$\|v\| := \sqrt{\langle v, v \rangle}$ is a norm on V.*

In the proof (and several of the following proofs) we make use of the following elementary fact.

Lemma 2.4. *Let*

$$p(t) := at^2 + bt + c \in \mathbb{R}[t]$$

be a quadratic polynomial with real coefficients such that $p(t) \geq 0$ for all $t \in \mathbb{R}$. Then

$$b^2 \leq 4ac.$$

Proof. If $a = 0$, then $b = 0$ and the claim is trivial. If $a \neq 0$, the zeroes of $p(t)$ in $\mathbb{C}$ are
$$-\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c}{a}}.$$
If $p(t) \geq 0$ for all $t \in \mathbb{R}$, p cannot have two distinct real roots, which forces
$$\frac{b^2}{4a^2} - \frac{4ac}{4a^2} \leq 0.$$
□

Proof of Theorem 2.3. (1): For each $t \in \mathbb{R}$, we have
(2.5)
$$0 \leq \langle v+tw, v+tw\rangle = \langle v,v\rangle + t(\langle v,w\rangle + \langle w,v\rangle) + \langle w,w\rangle = \|v\|^2 + t^2\|w\|^2 + 2t\Re\langle v,w\rangle =: p(t).$$
Lemma 2.4 shows that
$$\Re\langle v,w\rangle^2 \leq \|v\|^2\|w\|^2.$$
Taking square roots yields
$$|\Re\langle v,w\rangle| \leq \|v\|\|w\|.$$
There is $z \in S^1$ such that $\langle v, zw\rangle = z\langle v,w\rangle > 0$. It follows that
$$|\langle v,w\rangle| = \langle v, zw\rangle = |\Re\langle v, zw\rangle| \leq \|v\|\|zw\| = \|v\||z|\|w\| = \|v\|\|w\|.$$
(2): As above, we find $z \in S^1$ such that $\langle v, zw\rangle = \|v\|\|zw\|$, so that we may assume $0 \leq \langle v,w\rangle = |\langle v,w\rangle| = \|v\|\|w\|$. We can also assume by scaling v that $\|v\| = 1$. Then
$$\|w - \langle v,w\rangle v\|^2 = \|w\|^2 - 2\langle v,w\rangle\langle v,w\rangle + \langle v,w\rangle\|v\|^2 = \|w\|^2 - \langle v,w\rangle^2 = \|w\|^2 - \|w\|^2\|v\|^2 = 0$$
or
$$w = \langle v,w\rangle v.$$
(3): For the triangle inequality, note that
$$\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2\Re\langle v,w\rangle \leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2$$
and take square roots. □

Remark 2.6. *Occasionally, it is useful to consider* semi-inner products, *which are maps $V \times V \to \mathbb{K}$, $(v,w) \mapsto \langle v,w\rangle$, which satisfy all axioms for inner products, except definiteness. In that case, the formula $\|v\| := \sqrt{\langle v,v\rangle}$ defines a seminorm on V, and the Cauchy-Schwarz inequality continues to hold.*

Definition 2.7. *A* Pre-Hilbert space *is a $\mathbb{K}$-vector space V, together with an inner product. A Pre-Hilbert space is a* Hilbert space *if V, with the norm induced by $\langle \_, \_ \rangle$ is complete.*

There are a couple of simple identities for inner products, which will be used very often. We begin with a definition.

Definition 2.8. *Let V be a Pre-Hilbert space and $v, w \in V$. We say that v and w are* orthogonal *and write $v \perp w$, provided that $\langle v,w\rangle = 0$. If $S \subset V$ is a subset, we write*
$$S^\perp := \{w \in V | w \perp v \forall v \in S\}.$$

Proposition 2.9. *Let V be a Hilbert space.*

(1) *For each $w \in V$, the functional $L_w : V \to \mathbb{K}$, $L_w(v) := \langle w,v\rangle$ is bounded and $\|L_w\| = \|w\|$.*

(2) *If $S \subset V$ is a subset, then $S^\perp$ is a closed linear subspace.*
(3) *If $W \subset V$ is a linear subspace, then $W \cap W^\perp = \{0\}$.*

Proof. (1): $|L_w(v)| \leq \|w\|\|v\|$ holds by Cauchy-Schwarz, and $|L_w(w)| = \|w\|^2$ by definition.

(2): We can write
$$S^\perp = \bigcap_{s \in S} \ker(L_s),$$
i.e. as the intersection of closed linear subspaces.

(3): if $v \in W \cap W^\perp$, we must have $\langle v, v\rangle = 0$, so $v = 0$. □

Proposition 2.10. *Let V be a Pre-Hilbert space and $v, w \in V$. Then*
(1) $\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\Re\langle v, w\rangle$.
(2) *(Pythagoras theorem) If $v \perp w$, then*
$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$
(3) *(Parallelogram identity)* $\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$.
(4) *(Polarization identity in the real case). If $\mathbb{K} = \mathbb{R}$, then*
$$\langle v, w\rangle = \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2).$$
(5) *(Polarization identity in the complex case). If $\mathbb{K} = \mathbb{C}$, then*
$$\langle v, w\rangle = \frac{1}{4}\sum_{k \in \mathbb{Z}/4} (-i)^k \|v + i^k w\|^2.$$

Proof. The identities (1)-(4) are obvious. For (5), note that
$$\|x + y\|^2 - \|x - y\|^2 = 4\Re\langle x, y\rangle$$
and that
$$\Re(-iz) = \Im(z)$$
for $x, y \in V$ and $z \in \mathbb{C}$. Therefore
$$\sum_{k \in \mathbb{Z}/4} (-i)^k \|v + i^k w\|^2 = (\|v + w\|^2 - \|v - w\|^2) + i(\|v - iw\|^2 - \|v + iw\|^2) =$$
$$= 4\Re\langle v, w\rangle + 4i(\Re\langle v, -iw\rangle) = 4\Re\langle v, w\rangle + 4i(\Re - i\langle v, w\rangle) = 4(\Re\langle v, w\rangle + i\Im\langle v, w\rangle) = 4\langle v, w\rangle.$$
□

Just as most Banach spaces arise as completions of normed spaces, most Hilbert spaces arise through a completion process.

Lemma 2.11. *The Banach space completion of a Pre-Hilbert space is a Hilbert space.*

Proof. Let $(W, \langle \_, \_\rangle_W)$ be a Pre-Hilbert space and let V be its Banach space completion. We have to prove that the inner product on W extends to an inner product on V which induces the norm. To that end, we note that
$$|\langle v, w\rangle - \langle v', w'\rangle| = |\langle v, w - w'\rangle + \langle v - v', w'\rangle| \leq \|v\|\|w - w'\| + \|v - v'\|\|w\|.$$
For $v, w \in V$, let $v_n \to v$ and $w_n \to w$ be sequences in W which converge to these vectors. The above estimate shows that the limit
$$\langle v, w\rangle := \lim_{n \to \infty} \langle v_n, w_n\rangle$$

exists, does not depend on the choice of the sequences v_n and w_n and yields an inner product on V. Furthermore, the continuity of the norm proves that

$$\|v\| = \lim_{n\to\infty} \|v_n\| = \lim_{n\to\infty} \sqrt{\langle v_n, v_n\rangle} = \sqrt{\langle v, v\rangle}$$

holds. □

2.2. **The projection theorem.** We recall that a subset $U \subset V$ of an $\mathbb{R}$-vector space is *convex* if for $x, y \in U$ and $t \in [0,1]$, the point $tx + (1-t)y$ also lies in U.

Theorem 2.12 (Projection theorem)**.** *Let V be a Hilbert space and let $K \subset V$ be a nonempty closed convex subset. Then there is a unique $v \in K$ of smallest norm. For all $w \in K$, we have $\Re\langle v - w, v\rangle \leq 0$.*

Proof. Let $d := \inf_{x\in K} \|x\|$. Let $t \geq 0$ and let $x, y \in K$ with $\|x\|^2, \|y\|^2 \leq d^2 + \frac{1}{4}t^2$. Since K is convex, $\frac{1}{2}(x+y) \in K$ and so

$$d^2 + \frac{1}{4}\|x-y\|^2 \leq \|\frac{1}{2}(x+y)\|^2 + \|\frac{1}{2}(x-y)\|^2 = \frac{1}{2}(\|x\|^2 + \|y\|^2) \leq d^2 + \frac{1}{4}t^2$$

by the parallelogram identity and therefore

$$\|x - y\| \leq t.$$

Uniqueness of v follows immediately.

To prove existence of v, let $x_n \in K$ with $\|x_n\| \to d$. Then for large m, $\|x_m\| \leq \|x_n\|$ and therefore

$$\limsup_{m\to\infty} \|x_n - x_m\| \leq \|x_n\| - d$$

which implies

$$\lim_{n\to\infty} \limsup_{m\to\infty} \|x_n - x_m\| = 0$$

as $\|x_n\| \to d$. Hence $(x_n)_n$ is a Cauchy sequence, which converges to some $v \in V$, and as K was assumed to be closed, we have $v \in K$, and furthermore

$$\|v\| = \lim_n \|x_n\| = d.$$

□

Lemma 2.13. *Let V be a Hilbert space and let $W \subset V$ be a closed linear subspace. Then*

(1) *each $x \in V$ can be written uniquely as a sum*
$$x = P(x) + Q(x)$$
with $P(x) \in W$ and $Q(x) \in W^\perp$,

(2) *for all $x, y \in V$, we have*
$$\|P(x)\|^2 + \|Q(x)\|^2 = \|x\|^2,$$

(3) *and the maps $P, Q : V \to V$ defined in this way are linear bounded operators of norm at most 1.*

Proof. (1) The uniqueness is easy: W and $W^\perp$ are linear subspaces, and $W \cap W^\perp = \{0\}$, so if $w, w' \in W$ and $v, v' \in W^\perp$ are given with

$$w + v = w' + v',$$

then $w = w'$ and $v = v'$. For the existence, note that $x + W$ is a nonempty closed convex subset of V. Let $y \in x+W$ be the element of smallest norm whose existence is guaranteed by 2.12. Then obviously

$$P(x) := x - y \in W,$$

and we put

$$Q(x) := x - P(x) = y.$$

To verify that $y \in W^\perp$, let $w \in W$ and $t \in \mathbb{R}$ be arbitrary. We have

$$\|y\|^2 \leq \|y + tw\|^2 = \|y\|^2 + 2t\Re\langle y, w\rangle + t^2\|w\|^2$$

or

$$t^2\|w\|^2 + 2t\Re\langle y, w\rangle \geq 0.$$

From Lemma 2.4, we conclude that

$$(\Re\langle y, w\rangle)^2 = 0$$

and hence

$$\Re\langle y, w\rangle = 0$$

for all $w \in W$. Pick $z \in S^1$ such that $z\langle y, w\rangle \in \mathbb{R}$ and we find that

$$z\langle y, w\rangle = \langle y, zw\rangle = \Re\langle y, zw\rangle = 0$$

and therefore $y \in W^\perp$, as claimed.

(2): Since $P(x) \in W$ and $Q(x) \in W^\perp$, we have $\langle P(x), Q(x)\rangle = 0$ and hence

$$\|x\|^2 = \|P(x)\|^2 + \|Q(x)\|^2$$

by the Pythagoras identity.

(3): The statement that each $v \in V$ can uniquely be written as a sum $P(x)+Q(x)$ with $P(x) \in W$ and $Q(x) \in W^\perp$ can be reformulated by saying that the linear map

$$W \oplus W^\perp \to V,\ (x, y) \mapsto x + y$$

is bijective, with inverse map $v \mapsto (P(v), Q(v))$. Because inverses of bijective linear maps are linear (a fact proven in Linear Algebra I), it follows that P and Q are linear. The equation $\|x\|^2 = \|P(x)\|^2+\|Q(x)\|^2$ immediately implies $\|Px\|, \|Qx\| \leq \|x\|$, so that both are bounded of norm ≤ 1. □

Theorem 2.14 (Projection theorem for linear subspaces)**.** *Let $W \subset V$ be a closed linear subspace of a Hilbert space, and let $P, Q \in \mathcal{L}(V, V)$ be the two maps constructed in Lemma 2.13. Then*

(1) *For all $x, y \in V$, we have*

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

(2) $W = \ker(Q) = \operatorname{im}(P)$, $W^\perp = \ker(P) = \operatorname{im}(Q)$.

(3) $P^2 = P$, $Q^2 = Q$, $QP = PQ = 0$, $P + Q = 1$.

Proof. (1): we compute

$$\begin{aligned}\langle P(x), y\rangle - \langle x, P(y)\rangle &= \langle P(x), P(y) + Q(y)\rangle - \langle P(x) + Q(x), P(y)\rangle = \\ &= \langle P(x), P(y)\rangle + \langle P(x), Q(y)\rangle - \langle P(x), P(y)\rangle - \langle Q(x), P(y)\rangle = \\ &= \langle P(x), P(y)\rangle - \langle P(x), P(y)\rangle = 0,\end{aligned}$$

using that W and $W^\perp$ are orthogonal. A similar computation shows the claim for Q.

(2): from $x = Px + Qx$, $Px \in W$, $Qx \in W^\perp$ and $W \cap W^\perp = \{0\}$, we get

$$x \in W \Rightarrow Qx = x - Px \in W \cap W^\perp \Rightarrow Qx = 0 \Rightarrow x = Px \Rightarrow x \in W.$$

This means

$$W \subset \ker(Q) \subset \mathrm{im}(P) \subset W$$

and therefore $W = \ker(Q) = \mathrm{im}(P)$. Similarly

$$x \in W^\perp \Rightarrow Px = x - Qx \in W \cap W^\perp \Rightarrow Px = 0 \Rightarrow x = Qx \Rightarrow x \in W^\perp$$

proves

$$W^\perp \subset \ker(P) \subset \mathrm{im}(Q) \subset W^\perp,$$

so $W^\perp = \ker(P) = \mathrm{im}(Q)$.

(3): $\mathrm{im}(P) = \ker(Q)$ implies $QP = 0$, and similarly $PQ = 0$ follows from $\mathrm{im}(Q) = \ker(P)$. It is obvious that $P + Q = 1$ and so

$$P = (P+Q)P = P^2 + QP = P^2$$

and

$$Q = (P+Q)Q = PQ + Q^2 = Q^2.$$

□

Definition 2.15. *The map $P = P_W$ from Lemma 2.13 is called the* orthogonal projection *onto W.*

The result has a converse.

Definition 2.16. *Let V be a Hilbert space. A* projection *in V is a bounded operator $P \in \mathcal{L}(V,V)$ such that*

$$P^2 = P,\ \forall v, w \in V : \langle Pv, w\rangle = \langle v, Pw\rangle.$$

Proposition 2.17. *Let V be a Hilbert space and let R be a projection. Then*

$$\mathrm{im}(R) = \ker(1-R)$$

is a closed subspace,

$$\mathrm{im}(R)^\perp = \ker(R),$$

and $R = P_{\mathrm{im}(R)}$ is the orthonormal projection onto $\mathrm{im}(R)$.

Similarly, $\mathrm{im}(1-R) = \ker(R)$, $\ker(R)^\perp$, *and* $1 - R = P_{\mathrm{im}(R)^\perp}$.

Proof. Since

$$(1-R)R = R - R^2 = 0,$$

it follows that $\mathrm{im}(R) \subset \ker(1-R)$. If $v \in \ker(1-R)$, then $v - Rv = (1-R)v = 0$, so $v = Rv \in \mathrm{im}(R)$. Since $(1-R)$ is continuous, $\ker(1-R)$ is a closed subspace.

If $w \perp \mathrm{im}(R)$, then $\langle v, Rw\rangle = \langle Rv, w\rangle = 0$ for all v, hence $Rw = 0$, so $\mathrm{im}(R)^\perp \subset \ker(R)$. If $Rv = 0$, then for each $w \in V$, we have $0 = \langle w, Rv\rangle = \langle Rw, v\rangle$, so $v \in \mathrm{im}(R)^\perp$.

To show that $R = P_{\mathrm{im}(R)}$, write $P = P_{\mathrm{im}(R)}$ and $Q = 1 - P_{\mathrm{im}(R)}$ and $W = \mathrm{im}(R)$. For $x \in W = \mathrm{im}(R)$, write $x = Rv$ and compute

$$Rx = R^2 v = Rv = x.$$

It follows that $RP = P$. If $x \in W^\perp$, then $\langle Rx, y\rangle = \langle x, Ry\rangle = 0$ for all $y \in V$ because $\mathrm{im}(R) = W$. Hence $Rx = 0$ when $x \in W^\perp$, and it follows that $RQ = 0$. Therefore

$$R = R(P+Q) = RP + RQ = P.$$

□

Proposition 2.18. *Let $W \subset V$ be a linear subspace of a Hilbert space. Then*

$$W^\perp = \overline{W}^\perp$$

and

$$(W^\perp)^\perp = \overline{W}.$$

Proof. For two subsets $S \subset T \subset V$, we have $T^\perp \subset S^\perp$. Applying this to $W \subset \overline{W}$ shows

$$\overline{W}^\perp \subset W^\perp.$$

If $v \in W^\perp$, then $\langle v, w\rangle = 0$ for all $w \in W$, hence by continuity also for all $w \in \overline{W}$, so $v \in \overline{W}^\perp$, which proves $W^\perp \subset \overline{W}^\perp$.

For the second part, we can assume without loss of generality that W is closed, by the first part. By Theorem 2.14, we have

$$W = \mathrm{im}(P_W),\ W^\perp = \mathrm{im}(1 - P_W)$$

and $1 - P_W = P_{W^\perp}$. It follows that

$$(W^\perp)^\perp = \mathrm{im}(1 - P_{W^\perp}) = \mathrm{im}(1 - (1 - P_W)) = \mathrm{im}(P_W) = W.$$

□

Corollary 2.19. *Let $W \subset V$ be a proper closed linear subspace of a Hilbert space (i.e. $W \neq V$). Then there is $v \in W^\perp$ with $\|v\| = 1$.*

Proof. We must have $W^\perp = \ker(P_W) \neq \{0\}$. □

In the finite-dimensional case, the statements of Theorem 2.14 and Proposition 2.18 are easily shown by methods of linear algebra. What we achieved so far was to prove that the geometric intuition derived from the finite-dimensional case is largely correct for Hilbert spaces. Completeness of both V and the linear subspace W was used in an essential way, in the proof of Theorem 2.12. We are no ready to harvest the fruits of these results and turn to linear functionals and linear operators.

2.3. **The Riesz representation theorem.**

Theorem 2.20 (Riesz representation theorem for functionals on Hilbert spaces)**.** *Let V be a Hilbert space and let $F : V \to \mathbb{K}$ be a continuous linear functional. Then there is a unique $w \in V$, such that*

$$F(v) = \langle w, v\rangle$$

for all $v \in V$, and we have

$$\|F\| = \|w\|.$$

Proof. The uniqueness statement is clear: if w, w' are two such vectors, then $\langle w - w', v\rangle = 0$ for all $v \in V$, in particular $\|w - w'\|^2 = \langle w - w', w - w'\rangle = 0$. The inequality $\|F\| \leq \|w\|$ follows from the Cauchy-Schwarz inequality, and as $|F(w)| = \langle w, w\rangle = \|w\|^2$, we also see $\|w\| \leq \|F\|$.

For the existence statement, we can assume without loss of generality that $F \neq 0$. As F is continuous, $\ker(F) \subset V$ is a closed linear subspace, and it has codimension 1. We choose $u \in V$ with $F(u) = 1$, and define

$$w_0 := (u - P_{\ker(F)} u) \in \ker(F)^\perp.$$

Because the codimension of $\ker(F)$ is 1, the complement $\ker(F)^\perp$ is 1-dimensional, and because $w_0 \neq 0$, $\ker(F)^\perp$ is spanned by w_0. Hence any $v \in V$ can be written in the form $v_0 + tw_0$, $v_0 \in \ker(F)$, $t \in \mathbb{K}$.

We claim that there is $a \in \mathbb{K}$ such that $F(v) = \langle aw_0, v\rangle$ for all $v \in V$. But this is now easily verified, writing $v = v_0 + tw_0$ as above:

$$F(v_0 + tw_0) = F(v_0) + tF(w_0) = tF(u - P_{\ker(F)}u) = tF(u) = t$$

and

$$\langle w_0, v_0 + tw_0\rangle = t\langle w_0, w_0\rangle$$

Hence $w := \frac{1}{\langle w_0, w_0\rangle} w_0$ solves the problem.

□

It is crucial for the validity of Theorem 2.20 that V is complete. Despite its simplicity, the result is very powerful. When applied to the Hilbert space $L^2(X,\mu)$, this result has an important consequence.

Corollary 2.21. *Let (X,μ) be a measure space. Then the map*

$$\Phi : L^2(X,\mu) \to L^2(X,\mu)',\ \Phi(f)(g) := \int_X fg d\mu$$

is an isometric isomorphism.

Proof. We can write

$$\Phi(f)(g) = \langle \overline{f}, g\rangle,$$

and the claim is immediate from Theorem 2.20. □

To demonstrate the power of these methods, we identify the dual space of $L^1(X,\mu)$. Recall from Corollary 1.60 that the map

$$\Phi : L^\infty(X,\mu) \to L^1(X,\mu)'$$

defined by

$$\Phi(f)(g) := \int_X fg d\mu$$

is an isometry if X is locally finite. We use Hilbert space methods to prove that Φ is in fact surjective.

Theorem 2.22. *Let (X,μ) be σ-finite. Then the map*

$$\Phi : L^\infty(X,\mu) \to L^1(X,\mu)'$$

is an isometric isomorphism.

Proof. A σ-finite measure is locally finite, and so Corollary 1.60 proves that Φ is an isometry. The only remaining issue is to prove that Φ is surjective.

We first consider the case where $\mu(X) < \infty$. The (quite ingenious) trick is to bring the Hilbert space $L^2(X,\mu)$ into play. So consider $F \in L^1(X,\mu)'$.

Since $\mu(X) < \infty$, the function 1 belongs to $L^2(X,\mu)$, and $\|1\|_{L^2} = \sqrt{\mu(X)}$. The Cauchy-Schwarz inequality shows that for $f \in L^2(X,\mu)$

$$\|f\|_{L^1} = \int_X |f| d\mu = \int_X 1|f| d\mu = \langle 1, |f|\rangle \leq \|1\|_{L^2} \||f|\|_{L^2} = \sqrt{\mu(X)}\|f\|_{L^2}.$$

This proves that $L^2(X,\mu) \subset L^1(X,\mu)$ and that the inclusion map $I : L^2(X,\mu) \to L^1(X,\mu)$ has norm $\|I\| \leq \sqrt{\mu(X)}$.

Therefore, $G = F \circ I : L^2(X, \mu) \to \mathbb{K}$ is a bounded linear functional. By the Riesz representation theorem or rather its Corollary 2.21, there is a unique $g \in L^2(X, \mu)$ with

$$F \circ I(f) = \int_X g f d\mu$$

for all $f \in L^2(X, \mu)$.

We claim that g is in $L^\infty(X, \mu) \subset L^2(X, \mu)$. For $c > 0$, let

$$S_c := \{x \in X | |g(x)| \geq c\} \subset X.$$

This is a measurable subset of finite measure. Let $a : X \to S^1$ be measurable such that $ag \geq 0$. Define

$$f(x) = \begin{cases} a(x) & x \in S_c \\ 0 & x \notin S_c. \end{cases}$$

Then f is a bounded measurable function and hence $f \in L^2(X, \mu)$. Observe that $\|f\|_{L^1} = \mu(S_c)$. We have constructed f so that

$$c\mu(S_c) \leq \int_X g f d\mu = F(f) = |F(f)| \leq \|F\| \|f\|_{L^1} = \mu(S_c)\|F\|.$$

If $c > \|F\|$, the above inequality is impossible unless $\mu(S_c) = 0$. It follows that

$$|g(x)| \leq \|F\|$$

for almost all $x \in X$. Hence $g \in L^\infty(X, \mu)$.

We have constructed an element $g \in L^\infty(X, \mu)$ such that $F(f) = \int_X f g d\mu$ for all $f \in L^2(X, \mu)$. Since $L^2(X, \mu) \subset L^1(X, \mu)$ is dense, the claim follows.

This settles the case of finite $\mu(X)$. The general case can be reduced to the finite case, using Lemma 2.23 below. Let $w \in L^1(X, \mu)$ be as in that Lemma. The measure $\nu = w\mu$ is finite, and the map

$$T : L^1(X, \nu) \to L^1(X, \mu),\ T(f) := wf$$

is an isometric isomorphism as

$$\int_X w f d\mu = \int_X f d\nu$$

(the inverse is $T^{-1}(g) = \frac{1}{w} g$). Let $F \in L^1(X, \mu)$ be a linear functional. By the finite case of the theorem, there is a unique $g \in L^\infty(X, \nu)$ with

$$FT(f) = \int_X g f d\nu$$

for all $f \in L^1(X, \nu)$. If $h \in L^1(X, \mu)$, then

$$F(h) = FT(T^{-1}h) = \int_X g T^{-1}(f) d\nu = \int_X g \frac{1}{w} f w d\mu = \int_X g f d\mu.$$

Since $0 < w(x) < \infty$ for all $x \in X$, $L^\infty(X, \nu) = L^\infty(X, \mu)$, and the Theorem is proven. $\square$

Lemma 2.23. *Let (X, μ) be a σ-finite measure space. Then there is a measurable function $0 < w \leq 1$ with $\|w\|_{L^1} < \infty$.*

Proof. Write $X = \bigcup_{n=1}^{\infty} X_n$ as the disjoint union of subsets of finite measure. The function w will be of the form

$$w = \sum_{n=1}^{\infty} a_n \chi_{X_n}$$

with certain $0 < a_n \leq 1$. Since

$$\|w\|_{L^1} = \sum_{n=1}^{\infty} a_n \mu(X_n),$$

we only have to pick a_n so that $a_n \mu(X_n) < \frac{1}{2^n}$, which is certainly possible. $\square$

Remark 2.24. *A similar method can be used to show that $L^q(X,\mu) \to L^p(X,\mu)'$ is surjective for σ-finite μ and $1 < p < 2$. The case $2 < p < \infty$ is more difficult because $L^2(X,\mu) \not\subset L^p(X,\mu)$ in that case. The proof given in* [8] *that $L^p(X,\mu)' \cong L^q(X,\mu)$ for $p > 2$ also uses the Riesz representation theorem, but through a longer detour (via the* Radon-Nikodym Theorem*), and we omit it.*

2.4. **The adjoint operator.**

Theorem 2.25. *Let V and W be Hilbert spaces, and let $F : V \to W$ be a bounded operator. Then there is a unique bounded operator $F^* : W \to V$ such that*

$$\langle w, Fv \rangle = \langle F^* w, v \rangle$$

for all $v \in V$, $w \in W$.

Proof. Uniqueness is clear, since if $F' : W \to V$ is another such operator, then $\langle F^* w - F' w, v \rangle = 0$ for all v and w, and so $F^* w = F' w$ for all w. For the existence, let $w \in W$ and consider the linear functional

$$L_w : V \to \mathbb{K},\ L_w(v) := \langle w, Fv \rangle.$$

This is bounded because $|L_w(v)| \leq \|w\| \|F\| \|v\|$. By the Riesz representation theorem, there is a unique $u \in V$ with

$$L_w(v) = \langle u, v \rangle$$

for all $v \in V$. Define a map $F^* : W \to V$ by

$$F^* w := u.$$

By construction, we have

$$\langle w, Fv \rangle = \langle F^* w, v \rangle$$

for all $v \in V$, $w \in W$. It follows that F^* is linear, since

$$\langle F^*(a_0 w_0 + a_1 w_1), v \rangle = \langle a_0 w_0 + a_1 w_1, Fv \rangle = \overline{a_0} \langle w_0, Fv \rangle + \overline{a_1} \langle w_1, Fv \rangle =$$
$$= \overline{a_0} \langle F^* w_0, v \rangle + \overline{a_1} \langle F^* w_1, v \rangle = \langle a_0 F^* w_0 + a_1 F^* w_1, v \rangle.$$

Finally

$$\|F^* w\| = \sup_{\|v\| \leq 1} \langle F^* w, v \rangle = \sup_{\|v\| \leq 1} \langle w, Fv \rangle \leq \|w\| \|F\| \|v\|,$$

so that F^* is bounded with $\|F^*\| \leq \|F\|$. $\square$

Examples 2.26.
(1) *Let $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$, both equipped with the standard inner product, and let $A \in \mathrm{Mat}_{m,n}(\mathbb{K}) = \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$. The adjoint of A is $\overline{A}^t$, the conjugate transpose matrix.*

(2) *If $W \subset V$ is a closed linear subspace, then $P_W^* = P_W$. Here, we view $P_W : V \to V$.*
(3) *If we view P_W as a map $V \to W$, its adjoint is the inclusion map $W \to V$.*
(4) *Let (X, μ) be a measure space and let $f \in L^\infty(X, \mu)$. This determines a multiplication operator $M_f : L^2(X, \mu) \to L^2(X, \mu)$, $M_f(g) := fg$, with adjoint $(M_f)^* = M_{\overline{f}}$.*

Theorem 2.27. *Let V, W, U be Hilbert spaces. Then for all $F, H \in \mathcal{L}(V, W)$, $G \in \mathcal{L}(W, U)$ and $a \in \mathbb{K}$, we have*
(1) $(F + H)^* = F^* + H^*$,
(2) $(aF)^* = \overline{a}F^*$,
(3) $(GF)^* = F^*G^*$,
(4) $(F^*)^* = F$,
(5) $\|F^*\| = \|F\|$,
(6) *(C^*-identity)* $\|F^*F\| = \|F\|^2$.
(7) $\ker(F^*) = \operatorname{im}(F)^\perp$, $\overline{\operatorname{im}(F^*)} = \ker(F)^\perp$.

Proof. The first four identities are straightforward to prove. For example, $(F^*)^*$ is the unique operator $V \to W$ such that

$$\langle (F^*)^* v, w\rangle = \langle v, F^* w\rangle$$

for all v, w. But F also has this property, as

$$\langle Fv, w\rangle = \langle v, F^* w\rangle.$$

It follows from the uniqueness statement of Theorem 2.25 that $(F^*)^* = F$.

At the very end of the proof of Theorem 2.25, we established that $\|F^*\| \leq \|F\|$. The reverse inequality follows from

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

For (6), first note that

$$\|F^*F\| \leq \|F^*\|\|F\| = \|F\|^2.$$

For the reverse inequality, let $\epsilon > 0$ and pick $v \in V$ with $\|v\| \leq 1$ and $\|Fv\|^2 \geq \|F\|^2 - \epsilon$. Then

$$\|F\|^2 \leq \|Fv\|^2 + \epsilon = \langle Fv, Fv\rangle + \epsilon = \langle F^*Fv, v\rangle + \epsilon \leq \|F^*F\|\|v\|^2 + \epsilon \leq \|F^*F\|^2 + \epsilon,$$

and the claim follows by letting $\epsilon \to 0$.

For the last point, let $u \in \ker(F^*)$ and $w = Fv \in \operatorname{im}(F)$. Then $\langle u, Fv\rangle = \langle F^*u, v\rangle = 0$ for all $v \in V$ and hence $u \in \operatorname{im}(F)^\perp$. The argument can be read backwards to show $\operatorname{im}(F)^\perp \subset \ker(F^*)$. Finally

$$\overline{\operatorname{im}(F^*)} = (\operatorname{im}(F^*)^\perp)^\perp = \ker(F)^\perp.$$

□

The innocent equation $\|F^*F\| = \|F\|^2$ is fundamental for the theory of algebras of operators on Hilbert spaces. We will only gradually see its impact.

Definition 2.28. *A bounded operator $T \in \mathcal{L}(V, V)$ is called* selfadjoint *if $T^* = T$.*

For example, a projection as defined in 2.16 is the same as a self-adjoint $P \in \mathcal{L}(V, V)$ with $P^2 = P$.

Lemma 2.29. *Let V and W be Hilbert spaces and let $T : V \to W$ be a bounded operator.*

(1) *Then T is an isometry if and only if $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in V$, and this if and only if $T^*T = 1$.*
(2) *In that case, TT^* is the orthogonal projection onto the closed linear subspace* $\mathrm{im}(T)$.
(3) *T is an isometric isomorphism if and only if $T^*T = 1$ and $TT^* = 1$.*

Definition 2.30. *A* unitary *operator on a Hilbert space V is an element $T \in \mathcal{L}(V, V)$ with $T^*T = TT^* = 1$; in other words, an isometric isomorphism $V \to V$.*

Proof. (1) If $T^*T = 1$, then

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle = \langle v, v\rangle = \|v\|^2$$

for all v and hence T is an isometry. If T is an isometry, we have (in the case $\mathbb{K} = \mathbb{C}$)

$$\langle Tv, Tw\rangle = \frac{1}{4}\sum_{k\in\mathbb{Z}/4}(-i)^k\|T(v + i^k w)\|^2 = \frac{1}{4}\sum_{k\in\mathbb{Z}/4}(-i)^k\|v + i^k w\|^2 = \langle v, w\rangle$$

for all v, w by the polarization identity. Similarly in the real case. If $\langle Tv, Tw\rangle = \langle v, w\rangle$ holds for all v, w, then

$$\langle v, w\rangle = \langle Tv, Tw\rangle = \langle T^*Tv, w\rangle$$

for all v, w and so $T^*T = 1$.

(2) If $T^*T = 1$, then $(TT^*)^2 = TT^*TT^* = TT^*$, and $(TT^*)^* = (T^*)^*T^* = TT^*$. Hence TT^* is a projection, and by Proposition 2.17, it is the projection onto the (closed) subspace $\mathrm{im}(TT^*)$. Finally $\mathrm{im}(TT^*) = \mathrm{im}(T)$, because

$$\mathrm{im}(TT^*) \subset \mathrm{im}(T) = \mathrm{im}(T(T^*T)) = \mathrm{im}((TT^*)T) \subset \mathrm{im}(TT^*).$$

(3) if T is an isometric isomorphism, then T is bijective, and since $T^*T = 1$, we must have[2] $T^{-1} = T^*$, which implies $TT^* = TT^{-1} = 1$. Vice versa, if T is an isometry with $TT^* = 1$, then 1 is the orthogonal projection onto $\mathrm{im}(T)$. Hence $\mathrm{im}(T) = W$, or T is surjective. ☐

2.5. **Orthogonal systems.**

Definition 2.31. *Let V be a Hilbert space. An* orthonormal system *in V is a map $S \to V$, $s \mapsto v_s$ from a set S such that*

$$\langle v_s, v_t\rangle = \begin{cases} 1 & s = t \\ 0 & s \neq t \end{cases}$$

for all $s, t \in S$.

Example 2.32. *Let S be a set and define $\delta_s \in \ell^2(S)$ by $\delta_s(t) = 0$ if $t \neq s$ and $\delta_s(s) = 1$. Then $S \to V$, $s \mapsto \delta_s$ is an orthonormal system.*

Example 2.33. *Let μ be the Haar measure on S^1, normalized by $\mu(S^1) = 1$. Then the functions z^n, $n \in \mathbb{Z}$, form an orthonormal system. This follows from*

$$\langle z^n, z^m\rangle = \int_{S^1}\overline{z^n}z^m d\mu = \int_{S^1} z^{m-n}d\mu = \int_0^1 e^{2\pi i(m-n)t}dt.$$

---

[2] This is a general fact: if $f : X \to Y$ is a bijective map of sets and $g : Y \to X$ is a map with $g \circ f = \mathrm{id}_X$, then $f \circ g = \mathrm{id}_Y$.

Construction 2.34. *Let V be a Hilbert space, and let $(v_s)_{s\in S}$ be an orthonormal system in V. Denote*

$$W_0 := \operatorname{span}\{v_s | s \in S\}$$

and

$$W := \overline{W_0}.$$

Then

$$T_0 : c_{00}(S) \to V,\ f \mapsto \sum_{s\in S} f(s)v_s \tag{2.35}$$

is a linear map (the sum is finite since f has finite support). Using that $(v_s)_s$ is orthonormal, compute

$$\|T_0 f\|^2 = \sum_{s,t\in S} \langle f(s)v_s, f(t)v_t\rangle = \sum_{s\in S} |f(s)|^2\langle v_s, v_s\rangle = \sum_{s\in S} |f(s)|^2 = \|f\|^2_{\ell^2(S)}.$$

It follows that T_0 is an isometry, and by continuity, it extends to a bounded operator

$$T : \ell^2(S) \to V$$

which is also an isometry. By definition $\operatorname{im}(T_0) = W_0$, and by continuity, it follows that

$$W_0 \subset T(\ell^2(S)) = T(\overline{c_{00}(S)}) \subset \overline{T_0(c_{00}(S))} = \overline{W_0} = W.$$

By Lemma 1.18, $\operatorname{im}(T)$ is closed, and it follows that $\operatorname{im}(T) = W$.

The bounded operator T has an adjoint

$$T^* : V \to \ell^2(S).$$

We know that $T^*T = 1$ and TT^* is the orthogonal projection onto $\operatorname{im}(T)$ (Lemma 2.29).

Let us compute a formula for T^*. For $x \in V$, T^*x is a function on S. But

$$(T^*x)(s) = \langle \delta_s, T^*x\rangle_{\ell^2} = \langle T\delta_s, x\rangle = \langle v_s, x\rangle.$$

Therefore T^* sends $x \in V$ to the function

$$(T^*x)(s) = \langle v_s, x\rangle. \tag{2.36}$$

Lemma 2.29 shows that TT^* is the orthogonal projection P_W onto W, and $1 - TT^*$ is the orthogonal projection $P_{W^\perp}$ onto $W^\perp$. For $x \in V$, we obtain

$$\|x\|^2 = \|P_W x\|^2 + \|P_{W^\perp} x\|^2 = \|TT^*x\|^2 + \|P_{W^\perp} x\|^2,$$

and because T is an isometry, this leads to

$$\|TT^*x\|^2 + \|P_{W^\perp}x\|^2 = \|T^*x\|^2 + \|P_{W^\perp}x\|^2 = \sum_{s\in S} |T^*x(s)|^2 + \|P_{W^\perp}x\|^2 = \sum_{s\in S} |\langle v_s, x\rangle|^2 + \|P_{W^\perp}x\|^2.$$

We have shown

$$\|x\|^2 = \sum_{s\in S} |\langle v_s, x\rangle|^2 + \|P_{W^\perp}x\|^2. \tag{2.37}$$

Remark 2.38. *We can write at least formally*

$$T^*x = \sum_{s\in S} \langle v_s, x\rangle \delta_s$$

and

$$TT^*x = \sum_{s\in S} \langle v_s, x\rangle v_s. \tag{2.39}$$

One might naively think that 2.39 converges absolutely, but this is false, as one can see by the following example.

Consider $V = \ell^2(\mathbb{N})$, and the orthonormal system $\{\delta_n | n \in \mathbb{N}\}$. In this example, the operator T is the identity, and so is T^* and hence TT^* as well. The function $x : \mathbb{N} \to \mathbb{K}$, $x(n) = \frac{1}{n}$, belongs to $\ell^2(\mathbb{N})$, because $\sum_{n=1}^\infty \frac{1}{n^2} < \infty$. Moreover $\langle x, \delta_n \rangle = \frac{1}{n}$. In this example, the formula (2.39) becomes

$$x = \sum_{n=1}^\infty \frac{1}{n}\delta_n.$$

The right hand side does not *converge absolutely because $\sum_{n=1}^\infty \frac{1}{n} = \infty$. Instead, it converges* unconditionally*. This notion has been explained in §1.7.*

Lemma 2.40. *The following conditions on an orthonormal system $(v_s)_{s\in S}$ with W_0, W, T as above in a Hilbert space V are equivalent.*

1. *For each $v \in V$, we have $\|v\|^2 = \sum_{s\in S} |\langle v_s, v\rangle|^2$.*
2. *$W^\perp = 0$.*
3. *If $\langle v_s, v\rangle = 0$ for all $s \in S$, then $v = 0$.*
4. *The subspace $W_0 := \mathrm{span}\{v_s | s \in S\} \subset V$ is dense.*
5. *The maps $T : \ell^2(S) \to V$ and $T^* : V \to \ell^2(S)$ are isometric isomorphisms.*

Such orthonormal systems are called complete*. In particular, if $(v_s)_{s\in S}$ is a complete orthonormal system, then V is isometrically isomorphic to $\ell^2(S)$.*

Proof. Let $W := \overline{W_0}$. $1 \Leftrightarrow 2$: by (2.37), (4) is equivalent to $P_{W^\perp} = 0$, which is the same as $W^\perp = 0$.

$2 \Leftrightarrow 3$: By linearity, (3) is equivalent to $W_0^\perp = 0$, and $W_0^\perp = W^\perp$ holds as $W = \overline{W_0}$.

$2 \Leftrightarrow 4$: since $P_{W^\perp} = 1 - TT^*$ (2) is equivalent to $TT^* = 1$, which is equivalent to T and T^* being isometric isomorphisms. □

Theorem 2.41. *Let V be a Hilbert space. Then V admits a complete orthonormal system. Hence V is isometrically isomorphic to $\ell^2(S)$, for some set S.*

Proof. This is an argument with Zorn's Lemma. Let $\mathcal{Z}$ be the set of all orthonormal subsets $S \subset V$. We give $\mathcal{Z}$ the partial ordering $S \leq T :\Leftrightarrow S \subset T$. Since $\emptyset \in \mathcal{Z}$, $\mathcal{Z}$ is nonempty. Let $\mathcal{C} \subset \mathcal{Z}$ be a chain. Then $\bigcup_{S\in\mathcal{C}} S$ is an orthonormal system, as one checks easily. Therefore, we can apply Zorn's Lemma and find a maximal orthonormal system $S \subset V$.

If $W := \overline{\mathrm{span}(S)} \neq V$, then $W^\perp \neq \{0\}$ by Corollary 2.19. Pick a unit vector $v \in W^\perp$. Then $S \cup \{v\}$ is an orthonormal system, contradicting the maximality of S. Therefore $\mathrm{span}(S) \subset V$ is dense, and S is complete. The last sentence follows immediately. □

Proposition 2.42. *Let V be a Hilbert space. The following are equivalent:*

1. *V is separable (i.e. it has a countable subset),*
2. *V has a countable complete orthonormal system.*

Proof. $2 \Rightarrow 1$ is clear: the set of all finitely supported functions $\mathbb{N} \to \mathbb{Q}$ is a countable dense subset of $\ell^2(\mathbb{N})$, and under hypothesis (2), V is isometrically isomorphic to $\ell^2(\mathbb{N})$.

$1 \Rightarrow 2$ is less easy. Let $\{v_n | n \in \mathbb{N}\} \subset V$ be dense, and put $W_n := \mathrm{span}\{v_1, \ldots, v_n\}$. Then W_n has dimension at most n, and $\bigcup_{n=1}^{\infty} W_n \subset V$ is a dense subspace of countable dimension. Let $V_1, V_2, \ldots$ be a subsequence of the sequence $W_1, W_2, \ldots$ so that $\dim(V_n) = n$ for all n. Put $L_1 := V_1$ and let $L_n = V_{n-1}^{\perp} \cap V_n$. The space L_n is 1-dimensional (there must be a unit vector in V_n orthogonal to V_{n-1}), and $L_n \perp L_m$ when $m \neq n$. Pick $w_n \in L_n$ of norm 1. Then $\{w_n | n \in \mathbb{N}\}$ is an orthonormal system, and as $(w_1, \ldots, w_n)$ is a basis of V_n, it follows that $\mathrm{span}(\{w_n | n \in \mathbb{N}\})$ is dense in V. □

2.6. **Example: Fourier series.** Let us look at an important example. Let μ be the Haar measure on S^1, normalized so that $\mu(S^1) = 1$. The functions $\chi_n(z) := z^n$ form an orthonormal system $(\chi_n)_{n \in \mathbb{Z}}$ in$L^2(S^1; \mathbb{C})$. For $f \in L^2(S^1)$, the scalar product $\langle \chi_n, f \rangle$ is

$$\hat{f}(n) := \langle \chi_n, f \rangle = \int_{S^1} \overline{\chi_n} f d\mu = \int_{S^1} \chi_{-n} f d\mu = \int_0^1 e^{-2\pi i n t} f(e^{2\pi i t}) dt,$$

which is also known as the *Fourier coefficient* of f. The series

$$\sum_{n \in \mathbb{N}} \hat{f}(n) \chi_n$$

is the *Fourier series* of f.

Theorem 2.43. *The orthonormal system $(\chi_n)_{n \in \mathbb{Z}}$ in $L^2(S^1)$ is complete.*

Proof. Since the Haar measure μ is regular, $C(S^1) \subset L^2(S^1)$ is dense. It follows from the Stone-Weierstrass theorem B.72 that

$$W_0 = \mathrm{span}\{\chi_n | n \in \mathbb{Z}\} \subset C(S^1)$$

is dense (in the C^0-norm). Because for $f \in C(S^1)$, we have $\|f\|_{L^2}^2 \leq \|f\|_{C^0}$, it follows that $W_0 \subset L^2(S^1)$ is dense. □

Remark 2.44. *The density of W_0 in $C(S^1)$ can be shown by more direct means, without recourse to the general Stone-Weierstrass theorem. To this end, let*

$$S_n f := \sum_{k=-n}^{n} \hat{f}(n) \chi_n$$

be the nth partial sum of its Fourier series. It is not true that the Fourier series of a continuous function f converges to f. It is a theorem by Fejer, however, that the sequence of arithmetic means

$$T_n f := \frac{1}{n+1} \sum_{k=0}^{n} S_n f$$

converges uniformly to f, for each continuous function f. It follows that $W_0 \subset C(S^1)$ is dense (and this can be used to prove *the general Stone–Weierstrass theorem.*

Corollary 2.45. *For $f \in L^2(S^1)$, the partial sums*

$$\sum_{k=-n}^{n} \hat{f}(k) z^k$$

of the Fourier series of f converge in the L^2-norm to f.

2.7. **Notes.** Everything in this chapter is standard material. I recommend [8] for its elegance. The details of the treatment of orthonormal systems are as in [10].

3. Baire's category theorem and its consequences

The main result of this section is the *open mapping theorem.* For the time being, let us only state a special case, which is easily digested.

Theorem 3.1. *Let V, W be Banach spaces, and let $F : V \to W$ be a bounded linear operator which is bijective. Then the inverse $F^{-1} : W \to V$ is bounded.*

Some remarks are in order. It is clear that F^{-1} is a linear map. The important part is the continuity of F^{-1}, and as we shall see, the completeness of both, V and W, is an essential hypothesis.

One can compare this to an analogous result in general topology. A continuous bijection $f : X \to Y$ of topological spaces is usually not a homeomorphism. On the other hand, if X and Y are *compact Hausdorff* spaces, then any continuous bijection $f : X \to Y$ is automatically a homeomorphism, which is a tremendously useful result.

3.1. **Baire's Theorem.**

The proof of Theorem 3.1 relies on a fundamental property of *complete* metric spaces, which is captured by the following result.

Theorem 3.2 (Baire category theorem)**.** *Let X be a complete metric space and let $U_n \subset X$, $n \in \mathbb{N}$, be open and dense. Then $\bigcap_{n=1}^{\infty} U_n \subset X$ is dense.*

Let us not dwell on the question why this result is called the Baire *category* theorem. It is essential that X is required to be complete: $\mathbb{Q}$ is countable, and for each $q \in \mathbb{Q}$, $\mathbb{Q} \setminus \{q\} \subset \mathbb{Q}$ is open and dense, but the countable intersection $\bigcap_{q\in\mathbb{Q}} \mathbb{Q} \setminus \{q\}$ is empty.

Proof. A subset $S \subset X$ is dense iff for each open and nonempty $U_0 \subset X$, the intersection $S \cap U_0$ is nonempty.

So let $U_0 \subset X$ be open and not empty. The goal is to prove that $U_0 \cap \bigcap_{n=1}^{\infty} U_n \neq \emptyset$.

Since U_1 is dense, $U_0 \cap U_1$ is not empty, and we pick $x_1 \in U_0 \cap U_1$. As $U_0 \cap U_1$ is open, there is $r_1 > 0$ such that $\bar{B}_{r_1}(x_1) \subset U_0 \cap U_1$.

Since U_2 is dense, $B_{r_1}(x_1) \cap U_2$ is not empty, and we may pick $x_2 \in B_{r_1}(x_1) \cap U_2$, and by openness a number $0 < r_2 \leq \frac{1}{2} r_1$ such that $\bar{B}_{r_2}(x_2) \subset B_{r_1}(x_1) \cap U_2$.

Continuing in this fashion, we obtain a sequence $x_n \in X$ and a sequence $r_n \to 0$ of positive numbers such that

$$\bar{B}_{r_n}(x_n) \subset \bar{B}_{r_{n-1}}(x_{n-1}) \cap U_n$$

for all $n \geq 0$. For $m \geq n$, we have $d(x_m, x_n) \leq r_n$, so that x_n is a Cauchy sequence. As X is complete, the sequence converges to some $x \in X$.

For each n and $m \geq n$, we have $x_m \in \bar{B}_{r_n}(x_n)$, and so $x = \lim_{m\to\infty} \in \bar{B}_{r_n}(x_n)$ for all $n \geq 1$. Therefore $x \in U_n$ for all $n \geq 0$, and hence $x \in U_0 \cap \bigcap_{n=1}^{\infty} U_n$, so in particular $U_0 \cap \bigcap_{n=1}^{\infty} U_n \neq \emptyset$. □

Corollary 3.3. *Let $X \neq \emptyset$ be a complete metric space, and let $A_n \subset X$, $n \geq 1$, be closed subsets such that $\bigcup_{n=1}^{\infty} A_n = X$. Then at least one A_n has nonempty interior.*

Proof. If the interior of each A_n is empty, the open sets $U_n := A_n^c$ are all dense and hence $\emptyset \neq \bigcap_{n=1}^{\infty} U_n = (\bigcup_{n=1}^{\infty} A_n)^c$, a contradiction. □

3.2. **The open mapping theorem.**

Definition 3.4. *Let X and Y be topological spaces. A map $f : X \to Y$ is* open *if $f(U) \subset Y$ is open whenever $U \subset X$ is open.*

It is clear that a continuous and open bijective map is a homeomorphism. Hence Theorem 3.1 is a special case of the following result.

Theorem 3.5 (Open mapping theorem)**.** *Let V, W be Banach spaces and let $F : V \to W$ be a bounded linear operator. The following are equivalent:*

(1) *F is surjective.*
(2) *There is $C \geq 0$ such that for each $w \in W$, there is $v \in V$ with $\|v\| \leq C\|w\|$ and $F(v) = w$.*
(3) *F is an open map, i.e. if $U \subset V$ is open, then so is $F(U)$.*

The implication $2 \Rightarrow 1$ is completely trivial, but we shall show $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$. The difficult and interesting implication is $1 \Rightarrow 2$.

Proof of the easy implications. $2 \Rightarrow 3$: the hypothesis (2) means that

$$\bar{B}_W(0, r) \subset F(\bar{B}_V(0, Cr))$$

for each $r > 0$. Let $U \subset V$ be open and $w \in F(U)$. Pick $v \in U$ with $F(v) = w$ and $\epsilon > 0$ such that $\bar{B}_V(v, \epsilon) \subset U$. It follows that

$$\bar{B}_W(w, \frac{\epsilon}{C}) = w + \bar{B}_W(0, \frac{\epsilon}{C}) \subset w + F(\bar{B}_V(0, \epsilon)) = F(\bar{B}_V(v, \epsilon)) \subset F(U).$$

So $F(U)$ is a neighborhood of w, and hence $F(U)$ is open.

$3 \Rightarrow 1$: Since $F(V) \subset W$ is an open neighborhood of 0, there is $r > 0$ such that $\bar{B}_W(0, r) \subset F(V)$. For $w \in W$, it follows that there is $v \in V$ with

$$\frac{rw}{\|w\|} = F(v)$$

and so

$$w = \frac{\|w\|}{r} \frac{rw}{\|w\|} = F(\frac{\|w\|}{r} v).$$

□

For the implication $1 \Rightarrow 2$, we have to use the completeness of both V and W. The first step uses the completeness of W through the Baire category theorem.

Lemma 3.6. *Let $F : V \to W$ be a surjective bounded linear operator from a normed space to a Banach space and $q \in (0, 1)$. Then there is $a > 0$ such that for each $w \in W$, there is $v \in V$ with*

$$\|v\| \leq a\|w\|$$

and

$$\|w - F(v)\| \leq q\|w\|.$$

Proof. For $n \in \mathbb{N}$, the set

$$A_n := \overline{F(\bar{B}_V(0, n))} \subset W$$

is closed and since F is surjective, we have

$$\bigcup_{n=1}^{\infty} A_n = W.$$

By the Baire category theorem, there is $n \in \mathbb{N}$ such that A_n contains a ball, say

$$\bar{B}_W(v, r) \subset A_n.$$

Then $\bar{B}_W(-v, r) \subset A_n$ and since A_n is convex (why?), we conclude that

$$\bar{B}_W(0, r) \subset A_n.$$

Let $a := \frac{n}{r}$. For $R > 0$ arbitrary, it follows that

$$\bar{B}_W(0, R) = \frac{R}{r}\bar{B}_W(0, r) \subset \overline{F(\bar{B}_V(0, aR))},$$

which implies that for $w \in W$, we have

$$w \in \overline{F(\bar{B}_V(0, a\|w\|))}.$$

This is what we had to prove. $\square$

Together with the next lemma, Lemma 3.6 implies the implication $1 \Rightarrow 2$ of the open mapping theorem. The following Lemma appeared in the course of the proof of the Tietze extension theorem.

Lemma 3.7. *Let $F : V \to W$ be a bounded linear operator from a Banach space to a normed space. Assume that there is $q \in (0, 1)$ and $a > 0$ such that for each $w \in W$, there is $v \in V$ with*

$$\|v\| \leq a\|w\|$$

and

$$\|F(v) - w\| \leq q\|w\|.$$

Then for each $w \in W$, there is $v \in V$ with $F(v) = w$ and

$$\|v\| \leq \frac{a}{1-q}\|w\|.$$

Proof. We construct elements $v_n \in V$, $n \in \mathbb{N}_0$, such that

$$\|v_n\| \leq aq^n\|w\|$$

and

$$\|w - F(v_0 + \ldots + v_n)\| \leq q^{n+1}\|w\|.$$

The existence of v_0 is required in the hypothesis of the Lemma. Suppose that $v_0, \ldots, v_{n-1}$ with the above two properties have been constructed. Let $w_n := w - F(v_0 + \ldots + v_{n-1})$; we have

$$\|w_n\| \leq q^n\|w\|.$$

By hypothesis, there is $v_n \in V$ with

$$\|v_n\| \leq a\|w_n\| = aq^n\|w\|$$

and

$$\|F(v_n) - w_n\| \leq q\|w_n\| \leq q^{n+1}\|w\|.$$

Then

$$\|w - F(v_0 + \ldots + v_n)\| = \|w - F(v_0 + \ldots + v_{n-1}) - F(v_n)\| = \|w_n - F(v_n)\| \leq q^{n+1}\|w\|.$$

The series $\sum_{n=0}^{\infty} v_n$ converges in V to, say v, and since F is continuous, we have $Fv = w$. (It is in the last sentence of the proof that we are using that F is continuous). $\square$

3.3. **Consequences of the open mapping theorem.** We want to discuss the most important consequences of the open mapping theorem by analogy with general topology. We already alluded to the first of those analogies:

Theorem 3.8. (1) *Let $f : X \to Y$ be a continuous bijective map between compact Hausdorff spaces. Then the inverse f^{-1} is continuous.*
(2) *Let $F : V \to W$ be a bijective continuous linear map between Banach spaces. Then the inverse F^{-1} is continuous.*

The first part is well-known from general topology, and the second one is just Theorem 3.1. The next analogy is the *closed graph theorem*.

If $f : X \to Y$ is a continuous map of topological spaces and Y is Hausdorff, then the graph $\mathrm{gra}(f) := \{(x, f(x)) \in X \times Y | x \in X\} \subset X \times Y$ is a closed subset.

Proof: If $(x, y) \notin \mathrm{gra}(f)$, pick disjoint open neighborhoods $y \in V_1 \subset Y$; $f(x) \in V_0 \subset Y$, and let $x \in U \subset X$ an open neighborhood such that $f(U) \subset V_0$. Then $U \times V_1 \subset X \times Y$ is a neighborhood disjoint from $\mathrm{gra}(f)$.

The converse only true under some restrictions, as the following example shows: $X = [0, 1]$, $Y = \mathbb{R}$, $f(x) = \frac{1}{x}$ for $x > 0$ and $f(0) = 0$ has closed graph, but is not continuous.

Theorem 3.9 (Closed graph theorem)**.** (1) *Let $f : X \to Y$ be a map of compact Hausdorff spaces. Then f is continuous if and only if the graph* $\mathrm{gra}(f) \subset X \times Y$ *is closed.*
(2) *Let $F : V \to W$ be a linear map of Banach spaces. Then F is continuous if and only if the graph* $\mathrm{gra}(F) \subset V \times W$ *is closed.*

Proof. For both statements, the "only if" part is proven above.

(1): Let $\mathrm{pr}_X : \mathrm{gra}(f) \to X$ and $\mathrm{pr}_Y : \mathrm{gra}(f) \to Y$ be the two projections. These are both continuous, pr_X is bijective, and the composition

$$X \overset{\mathrm{pr}_X^{-1}}{\to} \mathrm{gra}(f) \overset{\mathrm{pr}_Y}{\to} Y$$

is f. In order to prove the continuity, it suffices to show that the continuous bijection $\mathrm{pr}_X : \mathrm{gra}(f) \to X$ is a homeomorphism. But $\mathrm{gra}(f) \subset X \times Y$ is closed, hence compact and Hausdorff, and so by Theorem 3.8 a homeomorphism.

(2): replace symbols $X \rightsquigarrow V$, $Y \rightsquigarrow W$, $f \rightsquigarrow F$, add the adjective "linear" whereever it makes sense and replace "compact Hausdorff space" by "Banach space" in the above argument. □

3.4. **Limitations of the open mapping theorem.** We show that the assumption that both, V and W, are complete, is necessary for the open mapping theorem to hold. One of these is easy:

Example 3.10 (Completeness of the target is necessary for the open mapping theorem)**.** *On the vector space $C^1([0, 1])$ of all C^1-functions on $[0, 1]$, we have the two norms*

$$\|f\|_{C^0} := \max_x |f(x)|$$

and

$$\|f\|_{C^1} := \max_x |f(x)| + |f'(x)|.$$

Let $V = (C^1([0, 1]), \|\text{-}\|_{C^1})$ and $W = (C^1([0, 1]), \|\text{-}\|_{C^0})$. Then $\mathrm{id} : V \to W$ is bounded, and $\mathrm{id} : W \to V$ is not bounded, since $\|e^{inx}\|_{C^0} = 1$ and $\|e^{inx}\|_{C_1} = n+1$. Here V is complete, while W is not complete.

The completeness of the source is also necessary, but we have to dig a little deeper to see that. We first show:

Proposition 3.11. *Let V be an infinite-dimensional normed vector space. Then there exists a discontinuous linear map $F : V \to \mathbb{K}$.*

Proof. The vector space V has a basis $\mathcal{B}$ in the sense of linear algebra: each $v \in V$ is a unique finite linear combination of elements of $\mathcal{B}$. For finite-dimensional vector spaces, this is done in Linear Algebra I, for infinite-dimensional spaces, one needs Zorn's lemma (Theorem A.4).

Since V is infinite-dimensional, there is a countable subset $\{v_n | n \in \mathbb{N}\} \subset \mathcal{B}$. We define $F : V \to \mathbb{K}$ by prescribing it on $\mathcal{B}$. For $w \in \mathcal{B}$, we define

$$F(w) := \begin{cases} n\|v_n\| & w = v_n \\ 0 & w \in \mathcal{B} \setminus \{v_n | n \in \mathbb{N}\} \end{cases}$$

and extend F linearly to all of V. Since $|F(v_n)| = n\|v_n\|$, F is not continuous. □

Example 3.12 (Completeness of the source is necessary for the open mapping theorem)**.** *This is harder to come by. Let W be an infinite dimensional Banach space. By Proposition 3.11, there is a discontinuous linear $F : W \to \mathbb{K}$.*

Now consider $V := \mathrm{gra}(F) \subset W \oplus \mathbb{K}$. By the closed graph theorem, V is not closed and hence V is not complete. The projection $V \to W$ is bounded with norm at most 1 and it is bijective, but its inverse is not continuous, since otherwise F would be continuous.

3.5. **The Banach-Steinhaus Theorem.** There is another important result on Banach spaces which is proved by the Baire category theorem. I admit that it has less appeal than the open mapping theorem.

Theorem 3.13 (Principle of uniform boundedness)**.** *Let V be a Banach space and let $(W_i, f_i)_{i\in I}$ be a family of normed vector spaces together with bounded linear maps $f_i : V \to W_i$. Suppose that*

$$\sup_{i\in I} \|f_i(v)\| < \infty$$

for each $v \in V$. Then

$$\sup_{i\in I} \|f_i\| < \infty.$$

Proof. The set

$$A_n := \{v \in V | \forall i \in I : \|f_i(v)\| \leq n\} \subset V$$

is closed, and

$$\bigcup_{n=1}^{\infty} A_n = V$$

by our assumption. Corollary 3.3 implies that there exists n, $y \in V$ and $r > 0$ such that

$$\bar{B}_r(y) \subset A_n.$$

In other words:

$$z \in V, \|z - y\| \leq r \Rightarrow \forall i \in I : \|f_i(z)\| \leq n.$$

For $v \in V$ with $\|v\| \leq 1$ and $i \in I$, we therefore have

$$\|f_i(v)\| = \frac{1}{r}\|f_i(rv)\| = \frac{1}{r}(\|f_i(rv+y)-f_i(y)\|) \leq \frac{1}{r}\|f_i(rv+y)\|+\frac{1}{r}\|f_i(y)\| \leq \frac{n}{r}+\frac{1}{r}\sup_{i\in I}\|f_i(y)\|.$$

It follows that
$$\|f_i\| \leq \frac{n}{r} + \frac{1}{r} \sup_{i \in I} \|f_i(y)\|.$$
□

Theorem 3.14 (Banach-Steinhaus Theorem)**.** *Let V be a Banach space, let W be a normed space, and let $f_n : V \to W$ be a sequence of bounded operators. Assume that the limit*
$$f(v) := \lim_{n\to\infty} f_n(v) \in W$$
exists for all $v \in V$. Then f is linear, bounded, and
$$\|f\| \leq \liminf_n \|f_n\| \leq \sup_n \|f_n\| < \infty.$$

Proof. It is clear that f is linear and that $\liminf_n \|f_n\| \leq \sup_n \|f_n\|$, and it is immediate from the principle of uniform boundedness that $\sup_n \|f_n\| < \infty$. For each v with $\|v\| \leq 1$, we have
$$\|f(v)\| = \lim_n \|f_n(v)\| \leq \liminf_n \|f_n\|,$$
and so the claim follows. □

Remark 3.15. *It is very easy to give examples of discontinuous linear functionals $V \to \mathbb{K}$ when V is not complete (for example, equip $c_{00}(\mathbb{N})$ with the $\|\_\|_{c_0}$-norm. Then $\{\delta_n\}$ is a basis for $c_{00}(\mathbb{N})$, and the functional $F(\delta_n) := n$ is not continuous). It is much more difficult to construct discontinuous linear functionals on Banach spaces, and Theorem 3.14 partially explains why. It rules out the strategy to construct a discontinuous F, namely as* pointwise *limit of continuous functionals.*

3.6. **Notes.** Everything in this chapter is standard material which can be found in any textbook on functional analysis. There is a slightly alternative proof of the open mapping theorem which uses what is called Zabreiko's lemma; in my opinion this alternative is not very elegant.

The parallel between the proof of the open mapping theorem and the Tietze extension theorem was noted by Grabiner [5].

4. The Hahn-Banach theorem and some of its applications

4.1. **The theorem.** Let V be a normed space, which is not the zero vector space. Is the dual space V' nontrivial? Equivalently, does there exist a nonzero bounded linear functional $F : V \to \mathbb{K}$? This is not at all obvious to answer (if you think this should be obvious, think about the quotient $\ell^\infty(S)/c_0(S)$). The Hahn-Banach theorem gives a positive answer. In fact, "Hahn-Banach Theorem" is a name for a whole collection of results. All those results concern the existence of some linear functional $F : V \to \mathbb{K}$ on a $\mathbb{K}$-vector space V. The first version we shall learn is as follows:

Theorem 4.1 (Hahn-Banach Theorem for normed spaces)**.** *Let V be a normed $\mathbb{K}$-vector space, let $W \subset V$ be a linear subspace, and let $F : W \to \mathbb{K}$ be linear and bounded. Then there exists a bounded linear functional $G : V \to \mathbb{K}$ with $\|G\| = \|F\|$ and $G|_W = F$.*

Before we study the (nontrivial) proof, let us record some important consequences.

Corollary 4.2. *Let V be a normed space and let $v \in V$ be a vector. Then there is a bounded linear functional $F : V \to \mathbb{K}$ with $F(v) = \|v\|$ and $\|F\| \le 1$ (if $v \neq 0$, the latter can be improved to say $\|F\| = 1$).*

Proof. Let $W := \operatorname{span}\{v\}$. We define $G : W \to \mathbb{K}$ by $G(av) = a\|v\|$. Then G is bounded with norm ≤ 1 because $|G(av)| = |a|\|v\| = \|av\|$. By Theorem 4.1, we can extend G to $F : V \to \mathbb{K}$ with $\|F\| = \|G\|$. □

By *definition* of the operator norm, we have for $F \in V'$

$$\|F\| = \sup_{v \in V, \|v\| \le 1} |F(v)|.$$

The corollary shows that we also have

$$\|v\| = \sup_{F \in V', \|F\| \le 1} |F(v)| \tag{4.3}$$

(and the maximum is attained).

For each normed space V, we can form the *double dual* V'', the space of all bounded linear functionals on V'. There is a canonical map

$$\iota_V : V \to V'',\ \iota(v)(F) := F(v).$$

Corollary 4.4. *For each normed vector space V, the map $\iota_V : V \to V''$ is an isometry.*

Proof. Let $v \in V$. Then

$$\|\iota(v)\| = \sup_{F \in V', \|F\| \le 1} |\iota(v)(F)| = \sup_{F \in V', \|F\| \le 1} |F(v)| = \|v\|.$$

In the last equation, we used the Hahn-Banach theorem through (4.3). □

Since the double dual is automatically complete, we obtain a different proof of the existence of a completion.

Alternative proof of Theorem 1.25. Let $W \subset V''$ be the closure of the image of $\iota_V : V \to V''$. Since V'' is complete, so is W, and $\operatorname{im}(\iota_V) \subset W$ is dense. Use Corollary 4.4. □

Another application of the Hahn-Banach Theorem is a proof that all norms on finite-dimensional vector spaces are equivalent.

Lemma 4.5. *Let V be a finite-dimensional Banach space. Then each linear functional $V \to \mathbb{K}$ is continuous.*

Proof. First note that the topological dual space V' is a linear subspace of the algebraic dual $\mathrm{Hom}(V;\mathbb{R})$. Therefore

$$\dim(V') \le \dim(\mathrm{Hom}(V;\mathbb{R})) = \dim(V),$$

the latter equation follows from Linear Algebra I. On the other hand, Corollary 4.4 proves

$$\dim(V) \le \dim(V'') \le \dim(\mathrm{Hom}(V';\mathbb{R})) = \dim(V').$$

Hence $\dim(V') = \dim(\mathrm{Hom}_{\mathbb{R}}(V;\mathbb{R}))$, and therefore $V' = \mathrm{Hom}(V;\mathbb{R})$, so every element of $\mathrm{Hom}(V;\mathbb{R})$ is bounded, as claimed. $\square$

Proof of Theorem 1.12. Let V be a finite-dimensional normed space. Pick a basis $(v_1, \ldots, v_n)$ of V, which gives a linear isomorphism

$$F : \mathbb{R}^n \to V,\ F(e_i) := v_i$$

with inverse

$$G : V \to \mathbb{R}^n.$$

Equip $\mathbb{R}^n$ with the norm $\|\_\|_{\ell^\infty}$. The map F is continuous, since

$$\|F(\sum_i a_i e_i)\| = \|\sum_i a_i F e_i\| \le \sum_i |a_i|\|v_i\| \le (\sum_{j=1}^n \|v_j\|) \max_i |a_i| =: C \max_i |a_i| = C\|\sum_i a_i e_i\|_{\ell^\infty}.$$

The map G can be written as $G = (G_1, \ldots, G_n)$, where $G_i : V \to \mathbb{R}$ is bounded by Lemma 4.5. It follows that

$$\|G(v)\|_{\ell^\infty} = \max_j |G_j(v)| \le (\max_j \|G_j\|)\|v\|,$$

so that G is continuous. This proves that F and G are isomorphisms of normed spaces, *no matter what norm on V is chosen.* The claim follows. $\square$

Definition 4.6. *If V, W are normed spaces and $T : V \to W$ is a bounded operator, we define the* dual *operator by*

$$T' : W' \to V',\ T(F)(v) := F(Tv)$$

for $F \in W'$.

For example, let V, W be Hilbert spaces, and recall the antilinear isometries $\Lambda_V : V \to V'$ and $\Lambda_W : W \to W'$. Then

$$F^* = \Lambda_V^{-1} F' \Lambda_W$$

(in linear algebra, this formula is often taken as the *definition* of the adjoint).

We can also form the double dual $T'' : V'' \to W''$.

Proposition 4.7. *Let V, W, U be normed spaces and let $G : U \to V$, $F : V \to W$ be bounded. Then*

(1) $\|F'\| = \|F\|$,
(2) $(FG)' = G'F'$,

(3) *the diagram*

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \downarrow{\scriptstyle \iota_V} & & \downarrow{\scriptstyle \iota_W} \\ V'' & \xrightarrow{F''} & W'' \end{array}$$

commutes.

(4) *If F is an isomorphism, then so is F'. The converse holds when V, W are complete.*

Proof. For $L \in W'$, we have

$$\|F'L\| = \sup_{v\in V, \|v\|\le 1} |F'(L)(v)| = \sup_{v\in V, \|v\|\le 1} |L(Fv)| \le \|F\|\|L\|$$

and so $\|F'\| \le \|F\|$. For the reverse inequality, let $\epsilon > 0$ and choose $v \in V$ with $\|v\| \le 1$ and $\|Fv\| \ge \|F\| - \epsilon$. By Corollary, 4.2, there is $L \in W'$ with $\|L\| \le 1$ and $|L(Fv)| = \|Fv\|$. Because

$$\|F\| - \epsilon \le \|Fv\| = |L(Fv)| = |(F'L)(v)| \le \|F'L\|\|v\| \le \|F'\|\|L\|\|v\| \le \|F'\|,$$

we obtain $\|F\| - \epsilon \le \|F'\|$ for each $\epsilon > 0$ and hence $\|F\| \le \|F'\|$.

The equation $(FG)' = G'F'$ is trivial. The commutativity of the diagram is also easy: for $v \in V$ and $L \in W'$, we have

$$T''(\iota_V(v))(L) = (\iota_V(v))(T'L) = (T'L)(v) = L(Tv)$$

and

$$(\iota_W(Tv))(L) = L(Tv).$$

It is also clear that F' is an isomorphism if F is. If F' is an isomorphism, then so is F''. There is $c > 0$ such that $\|F''\varphi\| \ge c\|\varphi\|$ for each $\varphi \in V''$. Since ι_V and ι_W are isometries, we get for each $v \in V$ the estimate

$$\|Fv\| = \|\iota_W Fv\| = \|F''\iota_V v\| \ge c\|\iota_V v\| = c\|v\|.$$

Hence F is bounded away from 0 and is therefore injective with closed image. If $\mathrm{im}(F) \ne W$, we get, by Lemma 4.8 below a linear functional $L \in W'$ with $\|L\| = 1$ and $F'(L) = L \circ F = 0$, which is impossible since F' is injective. Therefore F is bijective and an isomorphism by the open mapping theorem (the use of the open mapping theorem was not essential). □ explain why

Lemma 4.8. *Let V be a normed space and let $W \subset V$ be a closed linear subspace, $W \ne V$. Then there is $L \in V'$ with $\|L\| = 1$ and $L|_W = 0$.*

Proof. Let $\pi : V \to V/W$ be the quotient map. The quotient V/W is a normed space, and $V/W \ne 0$. Hence by the Hahn-Banach theorem or rather Corollary 4.2, there is $F \in (V/W)'$ with $\|F\| = 1$. Then $L := F \circ \pi$ is the desired linear functional. □

4.2. **The proof of the Hahn-Banach Theorem.** We already said that Theorem 4.1 is just one of several results with the name “Hahn-Banach Theorem”. We prove a more general version, which implies Theorem 4.1, as well as the other Hahn-Banach Theorems.

Definition 4.9. *Let V be a $\mathbb{R}$-vector space. A function $p : V \to \mathbb{R}$ is* sublinear *if*
$$p(av) = ap(v)$$
for all $a > 0$ and $v \in V$ and
$$p(u + v) \leq p(u) + p(v)$$
for all $u, v \in V$.

For example, a seminorm is sublinear, and an $\mathbb{R}$-linear map is also sublinear.

Theorem 4.10 (Hahn-Banach Theorem)**.** *Let V be an $\mathbb{R}$-vector space and let $p : V \to \mathbb{R}$ be sublinear. Let $W \subset V$ be a linear subspace and let $f : W \to \mathbb{R}$ be linear such that $f(w) \leq p(w)$ for all $w \in W$. Then there exists a linear $g : V \to \mathbb{R}$ such that $g|_W = f$ and such that $g(v) \leq p(v)$ for all $v \in V$.*

Lemma 4.11. *Theorem 4.10 holds when* $\dim(V/W) = 1$.

Proof. Let $v \in V \setminus W$. Then each element of V can be written uniquely as $w + tv$, $w \in W$ and $t \in \mathbb{R}$. For $a \in \mathbb{R}$ fixed, the map
$$g(w + tv) = f(w) + ta$$
is an extension of f to a linear map on V. We claim that we can pick a such that $g(x) \leq p(x)$ for all $x \in V$. For $x, y \in W$, we estimate
$$f(x) + f(y) = f(x + y) \leq p(x + y) = p((x + v) + (y - v)) \leq p(x + v) + p(y - v).$$
It follows that
$$f(y) - p(y - v) \leq p(x + v) - f(x)$$
for all $x, y \in W$ and hence that
$$A := \sup_{y \in W} (f(y) - p(y - v)) \leq \inf_{x \in W} (p(x + v) - f(x)) =: B,$$
and we pick $a \in \mathbb{R}$ such that
$$A \leq a \leq B.$$
For $w \in W$ and $t > 0$, we have
$$g(w + tv) = f(w) + ta \leq f(w) + tB \leq f(w) + t(p(\frac{w}{t} + v) - f(\frac{w}{t})) =$$
$$= tp(\frac{w}{t} + v) = p(w + tv),$$
and
$$g(w - tv) = f(w) - ta \leq f(w) - tA \leq f(w) - t(f(\frac{w}{t}) - p(\frac{w}{t} - v)) =$$
$$= tp(\frac{w}{t} - v) = p(w - tv).$$
$\square$

Proof of Theorem 4.10. Let $\mathfrak{X}$ be the set of all pairs (U, g), where $U \subset V$ is a linear subspace containing W, and $g : U \to \mathbb{R}$ is linear with $g|_W = f$, and $g(x) \leq p(x)$ for all $x \in U$. We give $\mathfrak{X}$ the partial order
$$(U, g) \leq (U', g') :\Leftrightarrow U \subset U', g'|_U = g.$$
The set $\mathfrak{X}$ is nonempty, since $(W, f) \in \mathfrak{X}$. If $\mathfrak{C} \subset \mathfrak{X}$ is a chain, then $X := \cup_{(U,g) \in \mathfrak{C}} U$ is a linear subspace of V, the map $h = \cup_{(U,g) \in \mathfrak{C}} g : X \to \mathbb{R}$ is well-defined and linear, and $h(x) \leq p(x)$ for all $x \in X$. Hence the hypotheses of Zorn's lemma are satisfied, and so there exists a maximal element $(U, g) \in \mathfrak{X}$. If $U \neq V$, pick $v \in V \setminus U$ and

let $X := \operatorname{span}(U \cup \{v\})$. By Lemma 4.11, there is $h : X \to \mathbb{R}$ with $h|_U = g$ and $h \leq p$. This contradicts the maximality of (U, g), and so $U = V$, and the theorem is proven. $\square$

Proof of Theorem 4.1. There is not much left to be proven if $\mathbb{K} = \mathbb{R}$. The map $p(x) = \|f\|\|x\|$ is sublinear, and we have $f(x) \leq |f(x)| \leq p(x)$ for $x \in W$. By the Hahn-Banach Theorem, there is a linear map $g : V \to \mathbb{R}$ with $g(x) \leq \|f\|\|x\|$ for all $x \in V$. If $g(x) \geq 0$, we get

$$|g(x)| = g(x) \leq \|f\|\|x\|,$$

and if $g(x) \leq 0$, we get

$$|g(x)| = -g(x) = g(-x) \leq \|f\|\| - x\|,$$

so that $\|g\| \leq \|f\|$.

If $\mathbb{K} = \mathbb{C}$, we have to do a little more work and need a Lemma. $\square$

Lemma 4.12. *Let V be a complex normed space and let $V_{\mathbb{R}}$ denote the underlying real normed space. Then the map*

$$\psi : \mathcal{L}_{\mathbb{C}}(V; \mathbb{C}) \to \mathcal{L}_{\mathbb{R}}(V_{\mathbb{R}}; \mathbb{R});\ f \mapsto \Re(f)$$

is bijective and norm-preserving.

Proof. For $h \in \mathcal{L}_{\mathbb{R}}(V_{\mathbb{R}}; \mathbb{R})$, we put

$$g(x) := h(x) - ih(ix) \in \mathbb{C}.$$

The functional g is $\mathbb{C}$-linear because

$$g(ix) - ig(x) = h(ix) - ih(i^2x) - ih(x) - h(ix) = 0,$$

and clearly bounded. It is clear that $\Re(g(x)) = h(x)$, and if $h = \Re(f)$ for $f \in \mathcal{L}_{\mathbb{C}}(V; \mathbb{C})$, we have

$$g(x) = \Re(f(x)) - i\Re(f(ix)) = \Re(f(x)) + i\Re(-if(x)) = \Re(f(x)) + i\Im(f(x)) = f(x).$$

In other words,

$$\varphi : \mathcal{L}_{\mathbb{R}}(V_{\mathbb{R}}; \mathbb{R}) \to \mathcal{L}_{\mathbb{C}}(V; \mathbb{C});\ \varphi(h)(x) := h(x) - ih(ix)$$

is the inverse to ψ.

The map ψ is norm-preserving: it is clear that $|\Re(f(x))| \leq |f(x)| \leq \|f\|\|x\|$, so that $\|\Re(f)\| \leq \|f\|$. For the reverse inequality, let $\epsilon > 0$ and pick $x \in V$, $\|x\| \leq 1$ so that $|f(x)| \geq \|f\| - \epsilon$. There is $z \in S^1$ so that $f(zx) > 0$. Hence

$$\|Re(f)\| \geq \Re(f(zx)) = f(zx) = |f(x)| \geq \|f\| - \epsilon,$$

and as ϵ was arbitrary, we have

$$\|Re(f)\| \geq \|f\|.$$

$\square$

Proof of Theorem 4.1, complex case. Let $f : W \to \mathbb{C}$ be $\mathbb{C}$-linear and bounded. By the real version of the Theorem, there is a functional $h : V \to \mathbb{R}$ with $h|_W = \Re(f)$ and $\|h\| = \|\Re(f)\| = \|f\|$ (here we used Lemma 4.12 for the first time). By Lemma 4.12 again, there is a (unique) $g : V \to \mathbb{C}$ with $\Re(g) = h$, and $\|g\| = \|h\| = \|f\|$.

By construction, we have

$$\Re(g|_W) = \Re(g)|_W = h|_W = \Re(f)$$

and since f and g are both $\mathbb{C}$-linear, Lemma 4.12 again implies $g|_W = f$, as desired. $\square$

4.3. **Generalized limits.** Recall that

$$\Phi : \ell^1(S) \to \ell^\infty(S)', \ \Phi(f)(g) := \sum_{s \in S} f(s)g(s)$$

is an isometry. With the help of the Hahn-Banach theorem, we now show that Φ is not surjective.

Definition 4.13. *Let S be an infinite set. Let $c(S) \subset \ell^\infty(S)$ be the subspace of all $f \in \ell^\infty(S)$ such that there is $\lim f \in \mathbb{K}$ such that for all $\epsilon > 0$, the set of all s with $|f(s) - \lim f| \leq \epsilon$ is finite. It is easy to check that $c(S)$ is a closed linear subspace, that $\lim f$ is uniquely determined by f, that*

$$\lim : c(S) \to \mathbb{K}$$

is a functional with $\| \lim \| = 1$ and that $c_0(S) = \ker(\lim)$.

For example, $c(\mathbb{N})$ is the space of all convergent sequences, and $\lim f = \lim_{n \to \infty} f(n)$ is the ordinary limit.

Definition 4.14. *Let S be an infinite set. A* generalized limit *is a linear functional $L : \ell^\infty(S) \to \mathbb{K}$ such that $\|L\| = 1$ and $L|_{c(S)} = \lim$.*

The Hahn-Banach theorem proves that generalized limit functionals exist.

Lemma 4.15. *Let S be an infinite set and let $L : \ell^\infty(S) \to \mathbb{K}$ be a generalized limit. Then L does not lie in the image of $\Phi : \ell^1(S) \to \ell^\infty(S)'$. Hence Φ is not surjective.*

Proof. Assume that $g \in \ell^1(S)$ is such that $L = \Phi(g)$ is a generalized limit. Since L is a generalized limit, we have $c_0(S) = \ker(\lim) \subset \ker(L)$. Hence for each $f \in c_0(S)$, we must have

$$0 = L(f) = \Phi(g)(f) = \sum_{s \in S} g(s)f(s).$$

Inserting $f = \delta_s$ shows $g(s) = 0$, hence $g = 0$, hence $L = 0$, which is absurd, as $L(1) = 1$. $\square$

Let us now turn to the case $S = \mathbb{N}$. The ordinary limit

$$\lim : c(\mathbb{N}) \to \mathbb{K}$$

has two additional properties (besides $\lim(1) = 1$ and $\| \lim \| = 1$): It is an *algebra homomorphism*:

$$\lim(fg) = \lim(f)\lim(g), \tag{4.16}$$

and it is *translation-invariant*. For the latter, let $T : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ be the map

$$Tf(n) := f(n+1).$$

This is clearly a bounded operator which maps $c(\mathbb{N})$ to itself. Then

$$\lim(Tf) = \lim(f) \tag{4.17}$$

for each $f \in c(\mathbb{N})$ (this boils down to a statement about convergence which so utterly trivial that it is not even mentioned in introductory textbooks on Analysis). One can prove that lim is the uniquely linear map $c(\mathbb{N}) \to \mathbb{K}$ which satisfies (4.16), (4.17) and $\lim(1) = 1$. One might wish to have a generalized limit which is both, an algebra homomorphism and translation-invariant. This, however, is impossible:

Lemma 4.18. *There is no algebra homomorphism $L : \ell^\infty(\mathbb{N}) \to \mathbb{K}$ which is translation-invariant.*

Proof. Let $f = \chi_{2\mathbb{N}} \in \ell^\infty(\mathbb{N})$. Then $f + Tf = 1$ and $f \cdot Tf = 0$. If $L : \ell^\infty(\mathbb{N})$ were a translation-invariant algebra homomorphism, then

$$0 = L(f \cdot Tf) = L(f)L(Tf) = L(f)^2$$

enforces $L(f) = 0$, and

$$1 = L(1) = L(f + Tf) = L(f) + L(Tf) = 2L(f)$$

enforces $L(f) = \frac{1}{2}$, a contradiction. □

It turns out that there are generalized limits which are translation-invariant, and generalized limits which are algebra homomorphisms. Both concepts are useful enough to get names.

Definition 4.19. *A* Banach limit *is a generalized limit $L : \ell^\infty(\mathbb{N}) \to \mathbb{K}$ which is translation invariant ($LT = L$). An* ultralimit *is a generalized limit $L : \ell^\infty(S) \to \mathbb{K}$ which is an algebra homomorphism.*

For the construction of ultralimits, we need more theory, but Banach limits can be constructed now.

Proposition 4.20. *There exists a Banach limit $L : \ell^\infty(\mathbb{N}) \to \mathbb{K}$.*

Proof. Let $M : \ell^\infty(\mathbb{N}) \to \mathbb{K}$ be a generalized limit, as constructed by the Hahn-Banach theorem. The idea is to average M appropriately to make it translation invariant. We define $S : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ by

$$Sf(n) := \frac{1}{n}\sum_{k=1}^{n} f(k) = (\frac{1}{n}\sum_{k=1}^{n} T^{k-1}f)(1)$$

and put $Lf := MSf$. It is clear that $\|Sf\|_{\ell^\infty} \leq \|f\|_\infty$ and that $S(1) = 1$. Moreover $Sf \in c_0(\mathbb{N})$ if $f \in c_0(\mathbb{N})$. To see this, let $f \in c_0(\mathbb{N})$ and $\epsilon > 0$. There is n_0 such that $|f(n)| \leq \epsilon$ for $n \geq n_0$. This can be rephrased by the statement that $\|T^n f\|_{\ell^\infty} \leq \epsilon$ when $n \geq n_0$.

It follows that for $n \geq n_0$, we have

$$|Sf(n)| \leq \frac{1}{n}\sum_{k=1}^{m_0} |T^{k-1}f(1)| + \frac{1}{n}\sum_{k=m_0+1}^{n} |T^{k-1}f(1)| \leq \frac{1}{n}\sum_{k=1}^{m_0} |T^{k-1}f(1)| + \frac{n-m_0}{n}\epsilon \to 0,$$

so that $Sf \in c_0(\mathbb{N})$. It follows that $Lf = MSf = \lim(f)$ when $f \in c(\mathbb{N})$.

To prove that L is translation-invariant, let $f \in \ell^\infty(\mathbb{N})$ and observe that

$$STf(n) - Sf(n) = \frac{1}{n}\sum_{k=1}^{n} T^k f(1) - \frac{1}{n}\sum_{k=1}^{n} T^{k-1}f(1) =$$

$$= \frac{1}{n}\sum_{k=1}^{n} T^k f(1) - \frac{1}{n}\sum_{k=0}^{n-1} T^k f(1) = \frac{1}{n}(T^n f(1) - f(1)) = \frac{1}{n}(f(n+1) - f(1)).$$

Therefore

$$|STf(n) - Sf(n)| \leq \frac{2}{n}\|f\|_{\ell^\infty}.$$

It follows that $STf - Sf \in c_0(\mathbb{N})$ for all $f \in \ell^\infty(\mathbb{N})$. Therefore

$$M(STf - Sf) = 0$$

or

$$LTf - Lf = 0,$$

as claimed. □

4.4. Reflexivity.

Definition 4.21. *A Banach space V is* reflexive *if the natural isometry $\iota : V \to V''$ is surjective.*

Example 4.22. *Finite-dimensional Banach spaces are reflexive. This has been proven in the course of the proof of Lemma 4.5.*

Example 4.23. *Hilbert spaces are reflexive. We give the details of the proof in the real case. Let $\Phi : V \to V'$ be the isometric isomorphism $v \mapsto (w \mapsto \langle v, w\rangle)$ (here we use the Riesz representation theorem of course). The dual of Φ is an isomorphism $V'' \to V'$. The composition*

$$V \xrightarrow{\iota} V'' \xrightarrow{\Phi'} V'$$

is given by the formula

$$(\Phi'(\iota(v)))(w) = (\iota(v))(\Phi(w)) = (\Phi(w))(v) = \langle w, v\rangle = \langle v, w\rangle = \Phi(v)(w).$$

Therefore $\Phi' \circ \iota = \Phi$. As Φ and Φ' are bijective, so is ι.

In the complex case, one can apply the same argument, with a little more care because Φ is anti*linear in that case.*

Example 4.24. *Let (X, μ) be a measure space and let $1 < p, q < \infty$ be conjugate exponents. The formula*

$$\Phi_p(f)(g) := \int_X fg d\mu$$

defines an isometry

$$\Phi_p : L^q(X, \mu) \to L^p(X, \mu)',$$

as we saw in Corollary 1.60. Consider the composition

$$L^p(X, \mu) \xrightarrow{\iota_p} L^p(X, \mu)'' \xrightarrow{\Phi_p'} L^q(X, \mu)'.$$

Very much as in the previous example, we compute

$$(\Phi_p'(\iota_p(f)))(g) = (\iota_p(f))(\Phi_p(g)) = (\Phi_p(g))(f) = \int_X gf d\mu = \int_X fg d\mu = \Phi_q(f)(g),$$

so that

$$\Phi_p' \circ \iota_p = \Phi_q.$$

We have seen that Φ_p and Φ_q are isomorphisms when X is a set with the counting measure. Therefore $\ell^p(S)$ is reflexive for all $1 < p < \infty$. It is true that this holds for arbitrary measure spaces, and so $L^p(X, \mu)$ is reflexive when $1 < p < \infty$.

Most other Banach spaces which appear "in nature" are not reflexive.

Example 4.25. *$c_0(S)$ is not reflexive unlike S is finite. Recall the isometric isomorphisms*

$$\Phi : \ell^1(S) \to c_0(S)',\ f \mapsto (g \mapsto \sum_{s\in S} f(s)g(s))$$

and

$$\Psi : \ell^\infty(S) \to \ell^1(S)',\ f \mapsto (g \mapsto \sum_{s\in S} f(s)g(s)).$$

The two compositions

$$c_0(S) \xrightarrow{\iota} c_0(S)'' \xrightarrow{\Phi'} \ell^1(S)'$$

and

$$c_0(S) \subset \ell^\infty(S) \xrightarrow{\Psi} \ell^1(S)'$$

are equal: for $f \in c_0(S)$ and $g \in \ell^1(S)$, compute

$$(\Phi'(\iota(f)))(g) = (\iota(f))(\Phi(g)) = \Phi(g)(f) = \sum_{s\in S} g(s)f(s) = \Psi(f)(g).$$

It follows that $c_0(S)$ is reflexive if and only if $c_0(S) = \ell^\infty(S)$, which happens if and only if S is finite.

Neither $\ell^1(S)$ nor $\ell^\infty(S)$ is reflexive. This follows from the previous example and the following theorem.

Theorem 4.26. *Let V be a Banach space. The following are equivalent:*

(1) *V is reflexive.*
(2) *V' is reflexive.*

Proof. The map $\iota_V : V \to V''$ has a dual map $\iota_V' : V''' \to V'$. Look at the composition

$$V' \xrightarrow{\iota_{V'}} V''' \xrightarrow{\iota_V'} V'. \tag{4.27}$$

For $L \in V'$ and $v \in V$, we have by the definitions

$$\iota_V'(\iota_{V'}(L))(v) = (\iota_{V'}(L))(\iota_V v) = (\iota_V v)(L) = L(v).$$

Therefore, the composition (4.27) is the identity. Hence: ι_V is an isomorphism if and only if ι_V' is an isomorphism (use Proposition 4.7). Since the composition (4.27) is bijective, this is the case if and only if $\iota_{V'}$ is an isomorphism. □

4.5. **$*$-Complemented subspaces.** We begin with a general definition from algebra.

Definition 4.28. *A sequence*

$$V \xrightarrow{F} W \xrightarrow{G} U$$

of linear maps between $\Bbbk$-vector spaces is exact at W *if* $\ker(G) = \operatorname{im}(G)$. *A* short exact sequence *of vector spaces is a sequence*

$$0 \to V \xrightarrow{F} W \xrightarrow{G} U \to 0$$

which is exact at V, W and U. (In other words, $\ker(G) = \operatorname{im}(F)$, *$F$ is injective and G is surjective.*

For example, if $W \subset V$ is a linear subspace, then

$$0 \to W \to V \to V/W \to 0$$

is a short exact sequence.

Lemma 4.29. *Let*

$$0 \to W \overset{J}{\to} V \overset{Q}{\to} U \to 0$$

be a short exact sequence of Banach spaces. Then $\mathrm{im}(J)$ *is closed, and* $J : W \to \mathrm{im}(J)$*, as well as the the induced map* $V/\mathrm{im}(J) \to U$ *are isomorphism of Banach spaces.*

Proof. Since $\mathrm{im}(J) = \ker(Q)$, it is closed. The induced map $R : V/\mathrm{im}(J) \to U$ is bounded and bijective. By Lemma 1.34, $V/\mathrm{im}(J)$ is complete, and so the open mapping theorem proves that R is an isomorphism (of Banach spaces). Since $\mathrm{im}(J)$ is a Banach space, and $J : W \to \mathrm{im}(J)$ is bijective, the open mapping theorem proves that it is an isomorphism. $\square$

Lemma 4.30. *Let*

$$0 \to W \overset{J}{\to} V \overset{Q}{\to} U \to 0$$

be a short exact sequence of Banach spaces. The the sequence

$$0 \to U' \overset{Q'}{\to} V' \overset{J'}{\to} W' \to 0$$

is exact.

Proof. The following things are to be shown:

(1) Q' is injective (this is general nonsense).
(2) $J' \circ Q' = 0$, but this follows from $J'Q' = (QJ)' = 0$.
(3) $\ker(J') \subset \mathrm{im}(Q')$. If $F \in V'$ satisfies $J'(F) = 0$, then $F \circ J = 0$, and by an algebraic argument, there is a linear map $G : U \to \mathbb{K}$ with $G \circ Q = F$. G is bounded by the open mapping theorem: there is $C \geq 0$ such that for all $u \in U$, there is $v \in V$ with $\|v\| \leq C\|u\|$ and $Qv = u$. It follows that

$$|Gu| = |GQv| = |Fv| \leq \|F\|\|v\| \leq \|F\|C\|u\|,$$

as desired.
(4) J' is surjective: as $J : W \to \mathrm{im}(J)$ is an isomorphism of Banach spaces, each bounded linear functional $F : W \to \mathbb{K}$ is of the form $G \circ J$ for a bounded linear functional $G : \mathrm{im}(J) \to \mathbb{K}$, which can be extended to all of V by the Hahn-Banach theorem.

$\square$

Theorem 4.31. *Let V be a reflexive Banach space and let $W \subset V$ be a closed subspace. Then W and V/W are reflexive.*

Proof. We consider the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & W \arrow[r] \arrow[d, "\iota_W"] & V \arrow[r] \arrow[d, "\iota_V"] & V/W \arrow[r] \arrow[d, "\iota_{V/W}"] & 0 \\
0 \arrow[r] & W'' \arrow[r] & V'' \arrow[r] & (V/W)'' \arrow[r] & 0.
\end{tikzcd}$$

The rows are exact sequence. The hypothesis is that ι_V is bijective, and ι_W as well as $\iota_{V/W}$ are injective for general reasons. The *5-Lemma* from elementary homological algebra[3] proves that ι_W and $\iota_{V/W}$ must be surjective as well. □

Proposition 4.32. *Let*

$$0 \to W \xrightarrow{J} V \xrightarrow{Q} U \to 0$$

be a short exact sequence of Banach spaces and bounded operators. The following are equivalent:

(1) *There is a bounded operator $S : U \to V$ with $QS = 1$ (a "section").*
(2) *There is a bounded operator $R : V \to W$ with $RJ = 1$ (a "retraction").*

If that is the case, the maps

$$\Phi : W \oplus U \to V,\ (w, u) \mapsto Jw + Su$$

and

$$\Psi : V \to W \oplus U,\ v \mapsto (Rw, Qw)$$

are isomorphisms. Short exact sequences with this property are called split.

Proof. $1 \Rightarrow 2$: if $v \in V$, then $Q(v - SQv) = 0$, and so by exactness $v - SQv \in \ker(Q) = \mathrm{im}(J)$. Since J is injective, there is a *unique* $w \in W$ with $Jw = v - SQv$. Define $Rv := w$. To show that $RJ = 1$, let $x \in W$. Then $Jx - SQJx = Jx$, and so $RJx = x$. Note moreover that

$$JR = 1 - SQ.$$

It remains to prove that R is bounded. Since $\mathrm{im}(J) = \ker(Q)$ is closed, $\mathrm{im}(J)$ is complete, and $J : V \to \mathrm{im}(J)$ is a bijective bounded operator of Banach spaces, hence an isomorphism by the open mapping theorem, so that there is $c > 0$ with $\|Jw\| \geq c\|w\|$ for all $w \in W$. In the above construction of R, we therefore see that

$$\|Rv\| \leq \frac{1}{c}\|JRv\| \leq \frac{1}{c}(\|v\| + \|SQv\|) \leq \frac{1 + \|SQ\|}{c}\|v\|$$

and that R is bounded.

$2 \Rightarrow 1$: The operator $1 - JR : V \to V$ is zero on the image $\mathrm{im}(J)$ because $(1 - JR)J = J - J(RJ) = J - J = 0$. Hence there is a unique $T : V/\mathrm{im}(J) \to V$ with $Tq = 1 - JR$, where $q : V \to V/\mathrm{im}(J)$ is the quotient map, and T is bounded. Using the isomorphism $V/\mathrm{im}(J) \cong U$, we construct $S : U \to V$ with $SQ = 1 - JR$. But then

$$QSQ = Q - QJR = Q$$

and since Q is surjective, $QS = 1$ follows.

It is easy to verify that $\Phi \circ \Psi$ and $\Psi \circ \Phi$ are the identity maps. □

Corollary 4.33. *Let $W \subset V$ be a closed subspace of a Banach space. The following are equivalent:*

(1) *There is a bounded operator $P : V \to W$ such that $P|_W = \mathrm{id}$.*
(2) *There is a closed linear subspace $U \subset V$ such that $U \oplus W = V$.* □

Such linear subspaces are called *complemented*. We have seen that each closed subspace of a Hilbert space is complemented. Here is a non-example:

Theorem 4.34. *The subspace $c_0(\mathbb{N}) \subset \ell^\infty(\mathbb{N})$ is not complemented.*

[3] Check out wikipedia for the very easy proof.

Proof. Let $Q : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})/c_0(\mathbb{N})$ be the quotient map and let $S : \ell^\infty(\mathbb{N})/c_0(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ be a putative section to Q. The space $\ell^\infty(\mathbb{N})$ has the property that there is a countable set $F_n \in \ell^\infty(\mathbb{N})'$ which separates the points of $\ell^\infty(\mathbb{N})$. In other words, if $f \in \ell^\infty(\mathbb{N})$ is such that $F_n(f) = 0$ for all $n \in \mathbb{N}$, then $f = 0$. For example, one can take the evaluation functionals $F_n(f) := f(n)$.

If the section S exists, then the functionals $G_n := F_n \circ S : \ell^\infty(\mathbb{N})/c_0(\mathbb{N})$ separate the points. We shall show that no such countable collection $\{G_n\}$ of functionals on $\ell^\infty(\mathbb{N})/c_0(\mathbb{N})$ can exist, which is a contradiction to the existence of S.

The first step is to prove that there is an uncountable set I and a family A_i of subsets $A_i \subset \mathbb{N}$ such that $A_i \cap A_j$ is finite once $i \neq j$. This sounds utterly unplausible at first but is easy to show: for each $x \in \mathbb{R} \setminus \mathbb{Q}$, choose a sequence $y_{x,n}$ of rational numbers with $\lim_{n\to\infty} y_{x,n} = x$, and let $C_x \subset \mathbb{Q}$ be the set of all terms of $y_{x,n}$. If $x_0 \neq x_1$, then the sequences $y_{x_0,n}$ and $y_{x_1,n}$ can have only finitely many terms in common, so that $C_{x_0} \cap C_{x_1}$ is finite. Then $I := \mathbb{R} \setminus \mathbb{Q}$ is uncountable and after choosing a bijection $\varphi : \mathbb{Q} \cong \mathbb{N}$, this construction provides $A_i := \varphi(C_i)$.

The collection $\{\chi_{A_i} | i \in I\} \subset \ell^\infty(\mathbb{N})$ has the following property: $\|\chi_{A_i}\| = 1$, and if $J \subset I$ is a finite subset and $a_j \in \mathbb{K}$, $|a_j| = 1$, $j \in J$, then

$$\|Q(\sum_{j\in J} a_j \chi_{A_j})\| \leq 1 \tag{4.35}$$

improving the generic estimate $\|Q(\sum_{j\in J} a_j \chi_{A_j})\| \leq |J|$ a lot (why exactly does this hold?).

For $n \in \mathbb{N}$, let $I_n := \{i | G_n(Q(\chi_{A_i})) \neq 0\} \subset I$. Because the functionals G_n separate the points, $\bigcup_{n=1}^\infty I_n = I$, and hence at least one I_n is uncountable.

Let $I_{n,k} := \{i \in I_n || G_n(Q(\chi_{A_i}))| \geq \frac{1}{k}$. As $\bigcup_{k=1}^\infty I_{n,k} = I_n$, at least one $I_{n,k}$ is infinite (even uncountable).

Let $J \subset I_{n,k}$ be a finite subset, and for $j \in J$, choose $a_j \in \mathbb{K}$ with $|a_j| = 1$ and $a_j G_n(Q(\chi_{A_j})) \geq 0$. Then

$$G_n(\sum_{j\in J} a_j Q(\chi_{A_j}) = \sum_{j\in J} a_j G_n(Q(\chi_{A_j})) = \sum_{j\in J} |G_n(Q(\chi_{A_j}))| \geq |J| \frac{1}{k}$$

and

$$G_n(\sum_{j\in J} a_j Q(\chi_{A_j}) = |G_n(\sum_{j\in J} a_j Q(\chi_{A_j})| \leq \|G_n\| \|Q(\sum_{j\in J} a_j \chi_{A_j})\| \leq \|G_n\|.$$

Together, we get

$$|J| \leq k\|G_n\|$$

for *each* finite $J \subset I_{n,k}$. This is a contradiction to the ascertainment that $I_{n,k}$ is infinite. □

4.6. **Notes.** The Hahn-Banach theorem has essentially one proof, which can be found in every textbook. Likewise, the applications are fairly standard. The construction of Banach limits presented above is exactly the same as in [11], except that in loc.cit., the argument is phrased differently.

The proof that V is reflexive if and only if V' is reflexive is usually given using the Banach-Alaoglu theorem which we will cover later. I found the above argument by myself, but I am sure that it is written down somewhere.

The extremely elegant algebraic proof that subspaces and quotients of reflexive spaces are reflexive is taken from [12], but did not seem to have found its way into textbooks, where a much less memorable argument is being used. This is probably because the language of exact sequences and the 5-lemma is foreign to most anaylsts.

5. The Riesz-Markov-Kakutani representation theorem

The capstone of this first part of the course is the proof of Theorem 1.46. We will give the proof only in the special case when X is a compact metrizable space. Let us recall the statement.

Theorem 5.1 (Riesz-Markov-Kakutani theorem)**.** *Let X be a second countable locally compact Hausdorff space and let $L : C_c(X) \to \mathbb{K}$ be a positive linear functional (in other words, $L(f) \geq 0$ whenever $f \geq 0$). Then there exists a unique Radon measure μ on X such that*

$$L(f) = \int_X f d\mu$$

for all $f \in C_c(X)$.

Let us first explain first the significance of the condition that X is second countable.

(1) a second countable locally compact Hausdorff space is metrizable by Theorem B.71.
(2) X is σ-compact. In other words, X admits an exhaustion $X_1 \subset X_2 \subset X_3 \ldots \subset X$, where $X = \bigcup_{n=1} \infty X_n$, X_n is compact and $X_n \subset X_{n+1}^\circ$, by Lemma B.81.
(3) Proposition C.59 says that *any* locally finite Borel measure on X is a Radon measure. Therefore, regularity of μ is not an issue.

Proof of uniqueness. Let μ, ν be two Radon measures such that $\int_X f d\mu = \int_X f d\nu$ for all $f \in C_c(X)$. We have to prove that $\mu(S) = \nu(S)$ for all measurable $S \subset X$.

Now let $X_1 \subset X_2 \subset X_3 \ldots \subset X$ be an exaustion of X. Since $\mu(S) = \lim_{n\to\infty} \mu(S \cap X_n)$ and similarly for ν, we may assume that S is contained in a compact subset of X, and hence that $\mu(S)$ and $\nu(S)$ are both finite.

Let $\epsilon > 0$. Since μ is regular, there is a compact set K with $K \subset S$ and $\mu(S) \leq \mu(K) + \epsilon$, and since ν is regular, there is an open set U with $S \subset U$ and $\nu(U) \leq \nu(S) + \epsilon$. Let $f \in C(X)$ be a function with $\chi_K \leq f \leq \chi_U$. Then

$$\mu(S) - \epsilon \leq \mu(K) \leq \int_X f d\mu = \int_X f d\nu \leq \nu(U) \leq \nu(S) + \epsilon.$$

This is true for all $\epsilon > 0$, and so

$$\mu(S) \leq \nu(S)$$

for all S. Exchanging the roles of μ and ν also shows $\nu(S) \leq \mu(S)$, i.e. $\mu(S) = \nu(S)$. $\square$

We will prove the Theorem first for compact X, and generalize at the end to the locally compact case. From now on, until further notification, we assume that X is compact. Let us slowly work towards the core of the argument. The first step is a characterization of positive functionals.

Lemma 5.2. *Let X be a topological space and let $L : C_b(X) \to \mathbb{K}$ be a linear functional. The following are equivalent:*

(1) *L is positive.*
(2) *L is bounded and $L(1) = \|L\|$.*

(We stated the lemma in a more general version as it is useful in other contexts, for example when studying $\ell^\infty(S) = C_b(S)$).

Proof. We write down the proof in the case $\mathbb{K} = \mathbb{C}$, the real case is easier.

$1 \Rightarrow 2$: if f is real-valued, then $-R \le f \le R$ for some R. Hence by positivity we get $-RL(1) \le L(f) \le RL(1)$, so $L(f) \in \mathbb{R}$. For $f \in C_b(X)$ arbitrary, we write $f = \Re(f) + i\Im(f)$ and see that

$$L(\overline{f}) = L(\Re(f) - i\Im(f)) = L(\Re(f)) - iL(\Im(f)) = \overline{L(f)}.$$

Consider the map

$$B : C_b(X) \times C_b(X) \to \mathbb{K}$$

defined by

$$B(f,g) := L(\overline{f}g).$$

Then B is $\mathbb{C}$-sesquilinear, and

$$B(g,f) = L(\overline{g}f) = L(\overline{\overline{f}g}) = \overline{L(\overline{f},g)} = \overline{B(f,g)}.$$

Since L is positive, we also have

$$B(f,f) = L(|f|^2) \ge 0.$$

Altogether, B is a semi-inner product. If $f \ge 0$, we have, by the Cauchy-Schwarz inequality,

$$0 \le L(f) = B(f,1) \le B(f,f)^{1/2}B(1,1)^{1/2} = L(|f|^2)^{1/2}L(1)^{1/2} \le L(\|f\|^2 1)^{1/2}L(1)^{1/2} = \|f\|L(1).$$

Since $-|f| \le f \le |f|$, this proves

$$|L(f)| \le \|f\|L(1).$$

or $\|L\| \le L(1)$. Equality holds by inserting $f = 1$.

$2 \Rightarrow 1$: We first show that $L(f) \in \mathbb{R}$ when f is real valued. Without loss of generality, we assume $L(1) = 1$. For $s \in \mathbb{R}$, we have

$$|L(f+is)|^2 = |L(f)+is|^2 = \Re(L(f))^2 + (s + \Im(L(f)))^2 =$$

$$= \Re(L(f))^2 + \Im(L(f))^2 + s^2 + 2s\Im(L(f)) = |L(f)|^2 + s^2 + 2s\Im(L(f)).$$

Since $\|f + is\|^2 \le \|f\|^2 + s^2$ (picture!), we also see

$$|L(f+is)|^2 \le \|f+is\|^2 \le \|f\|^2 + s^2.$$

Together, we obtain

$$|L(f)|^2 + 2s\Im(L(f)) \le \|f\|^2$$

for all $s \in \mathbb{R}$, which is only possible if $\Im(L(f)) = 0$.

If $f : X \to \mathbb{R}$ is bounded and continuous, we get $L(f) \in \mathbb{R}$. The function

$$g := f - \frac{1}{2}\|f\|1$$

has norm $\|g\| = \frac{1}{2}\|f\|$ (why?). Therefore

$$L(f) = L(g + \frac{1}{2}\|f\|1) = L(g) + \frac{1}{2}\|f\|L(1) \ge -\|L\|\|g\| + \frac{1}{2}\|f\|L(1) = \frac{1}{2}\|f\|(\|L\| - L(1)) = 0.$$

□

The idea behind the proof of Theorem 5.1 is to reduce the statement to the case of a space X of particular type. The reduction step is as follows.

Lemma 5.3. *Let $\varphi : X \to Y$ be a surjective continuous map between compact metrizable spaces. If the conclusion of Theorem 5.1 holds for X, then it holds for Y.*

Proof. The map $\varphi^* : C(Y) \to C(X)$, $f \mapsto f \circ \varphi$ is an isometry because φ is surjective. Let $L : C(Y) \to \mathbb{K}$ be a positive functional. By Lemma 5.2, L is bounded and $\|L\| = L(1)$. By the Hahn-Banach theorem, there is a functional $F : C(X) \to \mathbb{K}$ with $\|F\| = \|L\| = L(1)$ and $F \circ \varphi^* = L$. Using Lemma 5.2 again, we get that F is positive. Therefore, by the hypothesis of the Lemma, there is a Radon measure μ on X such that

$$F(f) = \int_X f(x) d\mu(x)$$

for all $f \in C(X)$. For $g \in C(Y)$, we therefore have

$$L(g) = F(g \circ \varphi) = \int_X g(\varphi(x)) d\mu(x).$$

We define a Borel measure $\varphi_* \mu$ on Y by

$$\varphi_* \mu(S) = \mu(\varphi^{-1})(S).$$

By construction, we have

$$\int_Y \chi_S d(\varphi_* \mu) = \varphi_* \mu(S) = \mu(\varphi^{-1}(S)) = \int_X \chi_S(\varphi(x)) d\mu(s)$$

for each measurable subset. This implies the identity

$$\int_Y g(y) d(\varphi_* \mu)(y) = \int_X g(\varphi(x)) d\mu(x)$$

for each step function and hence for each integrable function, hence a fortiori for all continuous functions. Therefore

$$L(g) = \int_X g(\varphi(x)) d\mu(x) = \int_Y g(y) d(\varphi_* \mu)(y),$$

as desired. $\square$

Lemma 5.4. *Let X be a compact metric space. Then there is a closed subspace*

$$Y \subset \prod_{n=1}^{\infty} \{0,1\}$$

and a surjective continuous map $Y \to X$.

Remark 5.5. *The space $\prod_{n=1}^{\infty} \{0,1\}$ is homeomorphic to the Cantor set in $\mathbb{R}$. Even more is true: for each compact metric space X, there is a continuous surjective map $\prod_{n=1}^{\infty} \{0,1\} \to X$ (Hausdorff-Alexandroff theorem). For a proof of this fact and some applications, see* [1].

Proof. We first deal with the case $X = [0,1]$. For $m \in \mathbb{N}$, consider the function

$$g_m : \prod_{n=1}^{\infty} \{0,1\} \to [0,1]$$

given by

$$g_m((x_n)_n) := \sum_{n=1}^{m} \frac{x_n}{2^n}.$$

Since g_m can be written as the composition $\prod_{n=1}^{\infty}\{0,1\} \overset{p_m}{\to} \prod_{n=1}^{m}\{0,1\} \to [0,1]$ of the projection p_m onto the first m factors with some map from a discrete set to $[0,1]$, φ_m is continuous. As $|g_{m+1}((x_n)_n) - g_m((x_n)_n)| \leq \frac{1}{2^m}$, the sequence g_m converges uniformly to the function

$$g((x_n)_n) = \sum_{n=1}^{\infty} \frac{x_n}{2^n},$$

which is therefore continuous. Because any $t \in [0,1]$ can be written in the form $\sum_{n=1}^{\infty} \frac{x_n}{2^n}$ for $x_n = 0,1$ (Analysis I!), g is surjective, as claimed.

From that, we get a surjective continuous map

$$h : K := \prod_{n=1}^{\infty}\{0,1\} \cong \prod_{m=1}^{\infty}\prod_{n=1}^{\infty}\{0,1\} \to L := \prod_{m=1}^{\infty}[0,1].$$

Part of the proof of the Urysohn metrization theorem B.71 says that for each compact metric space X, there is an injective map $f : X \to \prod_{m=1}^{\infty}[0,1]$.

Now we let

$$Y := \{(x,k) \in X \times K | f(x) = h(k)\} \subset X \times K.$$

This is a closed subspace of $X \times K$, and $X \times K$ is a compact space by Tychonov's theorem and metrizable as K is metrizable, so that Y is a compact metric space. Moreover, the continuous map $Y \to K$, $(x,k) \mapsto k$, is injective, because f is injective. Hence Y is homeomorphic to a closed subspace of K.

Finally, the map $Y \to X$, $(x,k) \mapsto x$ is surjective because h is surjective, and clearly continuous. $\square$

What are the specific properties of closed subspaces of $\prod_{n=1}^{\infty}\{0,1\}$?

Lemma 5.6. *Let $Y \subset \prod_{n=1}^{\infty}\{0,1\}$ be a closed subspace. Let $\mathcal{CL} \subset \mathcal{P}(Y)$ be the set of all subsets which are simultaneously open and closed ("clopen"). Then*

1. *$\mathcal{CL}$ contains a countable basis for the topology of Y.*
2. *$\mathcal{CL}$ is a Boolean algebra (that means, complements, finite intersections and finite unions of clopen subsets are clopen, and $\emptyset$ and Y are clopen.*
3. *The σ-algebra $\langle\mathcal{CL}\rangle$ generated by $\mathcal{CL}$ agrees with the Borel-σ-algebra.*
4. *The set $A := \{\sum_{j=1}^{r} a_j\chi_{U_j} | U_j \in \mathcal{CL}, a_j \in \mathbb{K}\} \subset C(Y;\mathbb{K})$ of all linear combinations of characteristic functions of elements of $\mathcal{CL}$ is a dense linear subspace.*

Proof. (1): Let $q_m : \prod_{n=1}^{\infty}\{0,1\} \to \{0,1\}$ be the projection onto the mth factor. By the definition of the product topology, the sets of the form

$$q_{m_1}^{-1}(S_1) \cap \dots q_{m_r}^{-1}(S_r)$$

where $r \in \mathbb{N}$, $m_i \in \mathbb{N}$, $S_i \subset \{0,1\}$ open, form a basis for the topology of $\prod_{n=1}^{\infty}\{0,1\}$. All subsets of $\{0,1\}$ are open and closed, and hence $q_{m_i}^{-1}(S_i)$, as well as the intersection, are open and closed. Hence we found a basis $\mathcal{B}$ for the topology of $\prod_{n=1}^{\infty}\{0,1\}$ which consists of clopen subsets and is moreover countable. By the definition of the subspace topology, the intersections of the elements of $\mathcal{B}$ with Y are open and closed and form a basis for the topology.

(2): is trivial.

(3): From the proof of (1), we conclude that $\mathcal{CL}$ contains a *countable* basis of the topology. An open subset $U \subset Y$ can be written as a union of elements of $\mathcal{CL}$, and since $\mathcal{CL}$ contains a *countable* basis, countably many suffice. Hence $U \in \langle \mathcal{CL} \rangle$, and therefore, $\langle \mathcal{CL} \rangle$ is the Borel-σ-algebra on Y.

(4): Clearly χ_U is continuous when U is clopen. Using that $\chi_U \chi_V = \chi_{U \cap V}$ and $a\chi_U + b\chi_V = a\chi_{U \setminus V} + b\chi_{V \setminus U} + (a+b)\chi_{U \cap V}$, one shows that $A \subset C(Y;\mathbb{K})$ is a subalgebra. Of course $1 = \chi_Y \in A$, and $f \in A$ implies $\overline{f} \in A$. For two distinct points $x, y \in Y$, we find $U, V \in \mathcal{CL}$ with $U \cap V = \emptyset$ and $x \in U$, $y \in V$ (otherwise, $\mathcal{CL}$ would not form a basis for the topology which is Hausdorff). For the function $f = \chi_U$, we have $f(x) \neq f(y)$, and so A separates the points of Y. By the Stone-Weierstrass Theorem B.72, it follows that A is dense in $C(Y;\mathbb{K})$. □

Lemma 5.7. *Let $Y \subset \prod_{n=1}^{\infty} \{0,1\}$ be a closed subspace. The conclusion of Theorem 5.1 holds for Y.*

Proof. Let

$$L : C(Y) \to \mathbb{K}$$

be a positive linear functional. It is bounded with operator norm $\|L\| = L(1)$.

For $U \in \mathcal{CL}$, the characteristic function χ_U is continuous, and we set

$$\mu_0(U) := L(\chi_U) \in [0, \|L\|].$$

The set function

$$\mu_0 : \mathcal{CL} \to [0, \infty)$$

is finitely additive. We use the Caratheory extension theorem C.22 to extend μ_0 to a measure on the σ-algebra $\langle \mathcal{CL} \rangle$ generated by $\mathcal{CL}$. For that, we have to verify the following: if $U_n \in \mathcal{CL}$ are countably many disjoint open and closed subsets, such that $U = \bigcup_{n=1}^{\infty} U_n$ is again open and closed, then

$$\mu_0(U) = \sum_{n=1}^{\infty} \mu_0(U_n). \tag{5.8}$$

But U is compact because it is closed and Y is compact, and so it can be covered by finitely many of the sets U_n. As these are disjoint, almost all U_n's are empty, and so the right hand side of (5.8) is truely a finite sum and (5.8) holds as μ_0 is finitely additive. The Caratheodory extension theorem C.22 now proves that there is a unique measure μ on the σ-algebra $\mathcal{C} = \langle \mathcal{CL} \rangle$ which extends μ_0.

For each $U \in \mathcal{CL}$, we have

$$\int_X \chi_U d\mu = \mu(U) = \mu_0(U) = L(\chi_U),$$

and by linearity, it follows that the functionals $\int_X \_ d\mu$ and L agree on $A \subset C(Y;\mathbb{K})$. Since both functionals are continuous, it follows that $L(f) = \int_X f d\mu$ for each $f \in C(Y;\mathbb{K})$. □

At this point, we have proven Theorem 5.1 for compact X.

End of the proof of Theorem 5.1. Let X be second countable and locally compact. Let $L : C_c(X) \to \mathbb{K}$ be positive, and let us first assume that L is bounded. Since $C_c(X) \subset C_0(X)$ is dense, L extends to a continuous functional $C_0(X) \to \mathbb{K}$ which is also denoted by L and which is positive.

The 1-point compactfication X^+ of X is second countable and compact, and $C_0(X)$ can be identified with the space $\{f \in C(X^+) | f(\infty) = 0\}$. If $f \in C(X^+)$, then $f - f(\infty) \in C_0(X)$. For $a \in \mathbb{R}$, consider the bounded functional

$$F(f) := L(f - f(\infty)) + af(\infty).$$

Which conditions need to be put on a so that F is positive? If $g \in C_c(X)$ satisfies $g \geq -c$, $c \geq 0$, then there is $h \in C_c(X)$ with $0 \leq h \leq c$ and $g \geq -h$, so that

$$L(g) \geq -L(h) \geq -\|L\|c.$$

Let $f \in C(X^+)$ be positive. Then it follows that

$$F(f) = L(f - f(\infty)) + af(\infty) \geq (a - \|L\|)f(\infty).$$

Hence F is positive provided that $a \geq \|L\|$. By the compact case of the theorem, there is a Radon measure ν on X^+ such that

$$F(f) = \int_{X^+} f d\nu.$$

The formula $\mu(S) := \nu(S)$ defines a Radon measure on X, and since $F|_{C_0(X)} = L$, μ represents L, as claimed.

If L is not necessarily bounded, we pick a countable basis $\{U_n\}$ for the topology, with each $\overline{U_n}$ compact. Let $(h_n)_n$ be a subordinate partition of unity (Theorem B.82). Let $a_n > 0$ be chosen so that $C := \sum_{n=1}^{\infty} a_n L(h_n) < \infty$.

Define $F : C_c(X) \to \mathbb{K}$ by

$$F(f) := \sum_{n=1}^{\infty} a_n L(h_n f)$$

(this is a finite sum). Then

$$|F(f)| \leq \sum_{n=1}^{\infty} a_n \|f\| L(h_n) \leq C\|f\|,$$

so F is bounded. Let μ be the Radon measure which represents F, and let $h := \sum_{n=1}^{\infty} a_n h_n$, which is a positive continuous function.

We define a new Borel measure ν on X by

$$\nu(S) = \int_S \chi_S \frac{1}{h} d\mu$$

(σ-additivity is easily checked using the monotone convergence theorem). For $f \in C_c(X)$, we have

$$\int_X f d\nu = \int_X \frac{f}{h} d\mu = F(\frac{f}{h}) = \sum_{n=1}^{\infty} a_n L(\frac{h_n}{h} f) = L(\sum_{n=1}^{\infty} a_n \frac{h_n}{h} f) = L(f),$$

so that ν represents L. □

5.1. **Notes.** The standard proof of the Riesz-Markov-Kakutani theorem can be found in [8, Chapter 2]. The proof given above is unusual; I learnt it from the book [3]. A generalization of the idea works for arbitrary compact spaces X [4].

6. Locally convex spaces

From the viewpoint of applications in analysis, the theory of Banach spaces has several drawbacks. The first is that infinite-dimensional Banach spaces contain very few compact subsets. Given how important compactness is for proving existence theorems in analysis (it provides convergent subsequences of certain sequences, and the limit of such a subsequence is often the objects one is interested in), this is a serious problem, and we need compensation for it.

Secondly, the theory is designed to encode various notions of convergence, either of (continuous or measurable) functions on a topological or measure space, or of linear operators between spaces of such functions. There are several notions of convergence which cannot be encoded by a norm. Let us give some examples.

(1) Let X be a (say locally compact Hausdorff) topological space and let $C(X)$ be the space of all continuous functions $X \to \mathbb{K}$, bounded or not. A useful notion of convergence of such functions is *uniform convergence on compacta*: we want to say that $f_n \to f$ provided that for each compact $K \subset X$, the restrictions $f_n|_K$ converge uniformly to $f|_K$. For example, the series $\sum_{n=0}^\infty \frac{1}{n!} z^n$ converges uniformly on compacta to $\exp(z)$, but not uniformly.
(2) For a sequence $f_n \in C^\infty([0,1])$ of smooth functions, the appropriate notion of convergence $f_n \to f \in C^\infty([0,1])$ would be that all derivatives $f_n^{(k)}$ converge uniformly to $f^{(k)}$, for each $k \in \mathbb{N}_0$. Convergence of all $f^{(k)}$, up to $k = r$, could be encoded by the C^r-norm $\sum_{k=0}^r \|f^{(k)}\|_{C^0}$, but we cannot encode convergence of all derivatives by a norm.
(3) We have seen examples of sequences $F_n, F \in \mathcal{L}(V, W)$, where $F_n v \to F v$ for each $v \in V$, but $\|F_n - F\|$ does not converge to 0. For example, let V be a separable Hilbert spaces with orthonormal basis $\{v_n | n \in \mathbb{N}\}$. Let $V_n := \mathrm{span}\{v_k | k \leq n\}$ and let $P_n : V \to V$ be the orthogonal projection onto V_n. Then $P_n v \to v$ for all $v \in V$, but $\|1 - P_n\| = 1$. Or X could be a compact Hausdorff space and $x_n \to x$ be a convergent sequence in X. Let $F_n : C(X) \to \mathbb{K}$, $F_n(f) := x_n$. Then $F_n(f) \to F(f)$ for all $f \in C(X)$, but $\|F_n - F\|$ does not tend to 0, unless x_n is eventually constant. Such examples are simply too common to be viewed as pathological!

All these notions of convergence can be captured by a more general concept than a Banach space. We have to pass to *locally convex vector spaces.* We remark, however, that not all sensible notions of convergence can be described by a topology.

6.1. **Prelude: weak and weak∗-convergence.** Before we develop the general theory, we concentrate on the following two notions of convergence which are of central importance in the theory of Banach spaces (and in many applications).

Definition 6.1. *Let V be a Banach space. A sequence $v_n \in V$ is* weakly convergent *to $v \in V$ if for each $L \in V'$, we have*

$$\lim_{n\to\infty} L(v_n) = L(v) \in \mathbb{K}.$$

A sequence $L_n \in V'$ is weak-∗-convergent *to $L \in V'$ if for each $v \in V$, we have*

$$\lim_{n\to\infty} L_n(v) = L(v) \in \mathbb{K}.$$

Similarly, we say that $v_n \in V$ is a weak Cauchy sequence *if for each $L \in V'$, the sequence $L(v_n) \in \mathbb{K}$ is a Cauchy sequence, and that $L_n \in V'$ is a* weak-∗-Cauchy sequence *if for each $v \in V$, the sequence $L_n(v) \in \mathbb{K}$ is a Cauchy sequence.*

In contrast to these notions, we say that v_n is *norm convergent* to v if $\|v_n - v\| \to 0$ (and similarly for sequences in V'). Before we give examples, let us develop a little bit theory.

Lemma 6.2. *If $v_n \in V$ converges in norm to $v \in V$, then v_n is weakly convergent to the same limit. The same is true for sequences in V'.* □

The converse is only true in finite-dimensional spaces. This will be demonstrated by many examples below.

Lemma 6.3. *A weakly convergent sequence is a weak Cauchy sequence, and a weak-$*$-convergent sequence is a weak$*$-Cauchy sequence.* □

Proposition 6.4. *Let V be a Banach space and let L_n be a weak$*$-Cauchy sequence. Then L_n is bounded, and weakly$*$-convergent to a uniquely determined $L \in V'$.*

Proof. This is a special case of the Banach-Steinhaus theorem 3.14. □

The analogue of Proposition 6.4 is not true for weak convergence. What is true is the following.

Proposition 6.5. *Let V be a Banach space and let v_n be a weak Cauchy sequence in V. Then v_n is bounded, and if the weak limit exists, it is uniquely determined.*

Proof. Let $F_n : V' \to \mathbb{K}$ be the functional $F_n(L) := L(v_n)$. The hypothesis implies that $F_n(L) \in \mathbb{K}$ is a Cauchy sequence and hence bounded. Therefore, there is C_L such that $|F_n(L)| \leq C_L$ for all n. By the principle of uniform boundedness, there is $C \geq 0$ such that $\|F_n\| \leq C$ for all n. Note that $F_n = \iota_V(v_n)$, and so by the Hahn-Banach theorem, $\|F_n\| = \|v_n\|$. Therefore $\|v_n\|$ is bounded.

If v_n is weakly convergent to v and to $v' \in V$, then for each $L \in V'$, we have

$$L(v) = L(v'),$$

and the Hahn-Banach theorem implies that $v = v'$. □

Now we can discuss these two notions in the cases where we have determined the dual spaces of Banach spaces explicitly.

Example 6.6. *Let $V = \ell^p(\mathbb{N})$, $1 < p < \infty$, let q be the conjugate exponent and let $f_n \in \ell^p(\mathbb{N})$ be a sequence. The following are equivalent:*

1. *f_n is a weak Cauchy sequence.*
2. *The sequence $\|f_n\|$ is bounded, and $f_n(m)$ is a Cauchy sequence for each $m \in \mathbb{N}$.*

Each weak Cauchy sequence in $\ell^p(\mathbb{N})$ converges weakly to some limit in $\ell^p(\mathbb{N})$.

Proof. $1 \Rightarrow 2$: Let f_n be a weak Cauchy sequence. By Proposition 6.5, $\|f_n\| \leq C$ for some C and all n. For $m \in \mathbb{N}$, the functional $\delta_m : \ell^p(\mathbb{N}) \to \mathbb{K}$, $f \mapsto f(m)$ is bounded. Therefore, by the definition of weak Cauchy sequences, the limit $f(n) := \lim_{n\to\infty} f_n(m) \in \ell^p(\mathbb{N})$ exists. We claim that $f \in \ell^p(\mathbb{N})$ and that f_n is weakly convergent to f. For $k \in \mathbb{N}$, we have

$$\sum_{m=1}^{k} |f(m)|^p = \lim_n \sum_{m=1}^{k} |f_n(m)|^p \leq C^p.$$

Therefore $f \in \ell^p(\mathbb{N})$. To verify that f_n is weakly convergent to f, recall that each element of $\ell^p(\mathbb{N})'$ is of the form $f \mapsto \sum_m f(m)g(m)$ for some $g \in \ell^q(\mathbb{N})$. We have to prove that

$$\lim_n \sum_{m=1}^{\infty} g(m)(f_n(m) - f(m)) = 0$$

for each $g \in \ell^q(\mathbb{N})$. Let $\epsilon > 0$ be arbitrary and fix k such that $\sum_{m=k+1}^{\infty} |g(m)|^q \leq \epsilon$. Then

$$|\sum_{m=1}^{\infty} g(m)(f_n(m)-f(m))| = |\sum_{m=1}^{k} g(m)(f_n(m)-f(m))|+|\sum_{m=k+1}^{\infty} g(m)(f_n(m)-f(m))| \leq$$

$$\leq |\sum_{m=1}^{k} g(m)(f_n(m) - f(m))| + (\sum_{m=k+1}^{\infty} |g(m)|^q)^{\frac{1}{q}} + (\sum_{m=k+1}^{\infty} |f_n(m) - f(m)|^p)^{\frac{1}{p}}$$

by the Hölder inequality. The term $|\sum_{m=1}^{k} g(m)(f_n(m) - f(m))|$ converges to 0 because of pointwise convergence. The second one is bounded by

$$\epsilon^{\frac{1}{q}} \|f - f_n\| \leq 2C\epsilon^{\frac{1}{q}}.$$

$2 \Rightarrow 1$: let $\|f_n\| \leq C$ and assume that $f_n(m)$ is a Cauchy sequence for each m. Let $g \in \ell^q(\mathbb{N})$ and let $\epsilon > 0$. There is $h \in c_{00}(\mathbb{N})$ with $\|h - g\| \leq \epsilon$. Then

$$|\sum_m (f_n(m)-f_k(m))g(m)| \leq |\sum_m (f_n(m)-f_k(m))(g(m)-h(m))|+|\sum_m (f_n(m)-f_k(m))h(m)|.$$

The second term converges to 0 as h has finite support and by pointwise convergence. The first term is bounded by

$$|\sum_m (f_n(m) - f_k(m))(g(m) - h(m))| \leq \|f_n - f_k\|_{\ell^p} \|g - h\|_{\ell^q} \leq 2C\epsilon$$

and therefore $|\sum_m (f_n(m) - f_k(m))g(m)| \leq 3C\epsilon$ for sufficiently large n, k. □

Now $\ell^p(\mathbb{N}) \cong \ell^q(\mathbb{N})$, and we can also speak about weak$*$-convergence. It turns out that the two notions agree, and not accidentally this is linked to the reflexivity of $\ell^q(\mathbb{N})$.

Proposition 6.7. *Let V be a Banach space, let $W = V'$ and let $L_n \in W$ be a sequence. If L_n is weakly convergent, it is weak-$*$-convergent. The converse is true if V is reflexive.*

In particular, in Hilbert spaces, the two notions agree.

Proof. If L_n is weakly convergent, then $\varphi(L_n) \in \mathbb{K}$ converges for each $\varphi \in W' = V''$. In particular, $\iota_V(v)(L_n) = L_n(v)$ converges for each v, and L_n is weakly$*$-convergent.

If V is reflexive, then each $\varphi \in V''$ is of the form $\iota_V(v)$ for some $v \in V$, from which we conclude the converse. □

Example 6.8. *Recall that $c_0(\mathbb{N}) \cong \ell^1(\mathbb{N})$. A sequence $f_n \in c_0(\mathbb{N})$ is a weak Cauchy sequence if and only if $\|f_n\|$ is bounded and if $f_n(m)$ converges for each $m \in \mathbb{N}$.*

The limit function $f(m) := \lim_{n\to\infty} f_n(m)$ does not *need to be in $c_0(\mathbb{N})$. In particular, weak Cauchy sequences in $c_0(\mathbb{N})$ do not need to converge.*

Example 6.9. *The space $\ell^1(\mathbb{N})$ is the dual space of $c_0(\mathbb{N})$, so we can speak about weak∗-convergence in $\ell^1(\mathbb{N})$. A sequence $f_n \in \ell^1(\mathbb{N})$ is weakly∗-convergent if and only if it is bounded and pointwise convergent. This is done using the same method of Example 6.6.*

For weak convergence *in $\ell^1(\mathbb{N})$, we have the following surprising result by I. Schur: if $f_n \in \ell^1(\mathbb{N})$ is a weak Cauchy sequence, it converges* in norm *(!) to some $\ell^1(\mathbb{N})$.*

Example 6.10. *A sequence $f_n \in \ell^\infty(\mathbb{N}) \cong \ell^1(\mathbb{N})'$ is weakly∗-convergent if and only if it is bounded and pointwise convergent. I do not know a handy characterization of weak convergence.*

6.2. **Locally convex spaces.** Let us now define a framework in which the above sorts of convergence can be discussed.

Definition 6.11. *A* topological vector space *over $\mathbb{K}$ is a $\mathbb{K}$-vector space V, equipped with a topology, such that the addition*

$$\alpha_V : V \times V \to V,\ (u, v) \mapsto u + v$$

and scalar multiplication

$$\mu_V : \mathbb{K} \times V \to V,\ (z, v) \mapsto zv$$

are continuous.

For example, if V is a normed vector space, then V, together with the topology induced by the norm, is a topological vector space. Topological vector spaces are far to general to be useful as such. A useful class are the *locally convex spaces.*

Definition 6.12. *Let V be a $\mathbb{K}$-vector space. A subset $U \subset V$ is* balanced *if $v \in U$ and $z \in \mathbb{K}$, $|z| = 1$ implies $zv \in U$. A* locally convex space *is a topological vector space V if* 0 *has a neighborhood basis which consists of convex balanced subsets.*

For example each normed vector space V is a locally convex space. This is because the balls $B_r(0) \subset V$ are convex and balanced and form a form a neighborhood basis of 0.

Before we begin the development of the general theory, let us introduce the standard construction of locally convex spaces. The reader should recall the construction of the *induced topology* which is explained in §B.5.

Definition 6.13. *Let V be a $\mathbb{K}$-vector space, let $\mathcal{F} := (W_i, F_i)_{i\in I}$ be a family of normed vector spaces W_i, together with linear maps $F_i : V \to W_i$. We let $\mathcal{T}_\mathcal{F}$ be the topology on V which is induced by this family of maps.*

There are two aspects of the definition of the induced topology. Firstly, we have described its open sets explicitly, and secondly, it has a universal property.

The explicit definition is that a subset $U \subset V$ is open in $\mathcal{T}_\mathcal{F}$ if and only if for each $x \in U$, there are $i_1, \dots, i_n \in I$ and $\epsilon_1, \dots, \epsilon_n > 0$, such that

$$x + \bigcap_{j=1}^{n} F_{i_j}^{-1}(B_{\epsilon_j}(0)) \subset U.$$

Phrased differently, the sets of the form

$$\bigcap_{j=1}^{n} F_{i_j}^{-1}(B_{\epsilon_j}(0)),\ n \in \mathbb{N}, i_j \in I, \epsilon_j > 0$$

form a neighborhood basis of 0. It is easily verified that such sets are convex and balanced.

The universal property of the induced topology (Lemma B.32) immediately implies the first two items of the next result.

Lemma 6.14. *Let V and $\mathcal{F} = (W_i, F_i)_{i\in I}$ be as in 6.13 and let V carry the topology $\mathcal{T}_{\mathcal{F}}$ (and W_i the usual norm topology). Then*

1. *the maps $F_i : V \to W_i$ are continuous,*
2. *a map $g : X \to V$ from an arbitrary topological space is continuous if and only if the compositions $F_i \circ g : X \to W_i$ are all continuous,*
3. *a sequence $v_n \in V$ converges to v if and only if all sequences $F_i(v_n)$ converge to $F_i(v)$.*
4. *V, with the topology $\mathcal{T}_{\mathcal{F}}$, is a locally convex space.*
5. *$\mathcal{T}_{\mathcal{F}}$ is Hausdorff if and only if $\mathcal{F}$ separates the points of V, in other words, if and only if for each $0 \neq v \in V$, there is $i \in I$ such that $F_i(v) \neq 0$.*
6. *In that case, the injective map $\phi : V \to \prod_{i\in I} W_i$, $v \mapsto (F_i(v))_{i\in I}$, is a topological embedding (and hence each subset $Z \subset V$ is homeomorphic to the subset $\phi(Z) \subset \prod_{i\in I} W_i$.*

Proof. (3) follows from (2) by considering the space $X = \{0, 1, \frac{1}{2}, \ldots\} \subset \mathbb{R}$. Concerning (4), we already saw a neighborhood basis of 0 which consists of balanced convex sets. It remains to prove that the addition maps α_V and μ_V are continuous. We know that all the α_{W_i} and μ_{W_i} are continuous, since these are the addition and scalar multiplication maps of *normed spaces*. The linearity of F_i is encoded in the relations

$$F_i \circ \alpha_V = \alpha_{W_i} \circ (F_i \times F_i)$$

and

$$F_i \circ \mu_V = \mu_{W_i} \circ (\mathrm{id}_{\mathbb{K}} \times F_i).$$

But this already solves the problem: the right hand sides are continuous by (1), and (2) then proves that α_V and μ_V are continuous.

(5): The map

$$\phi : V \to \prod_{i\in I} W_i$$

is continuous. To see this, we have to show (by the universal property of the product topology) that the composition $p_i \circ \phi : V \to W_i$ is continuous, where p_i is the projection of the product onto the ith factor. But $p_i \circ \phi = F_i$ is continuous.

The map ϕ is injective if (and only if) the family $\mathcal{F}$ separates the points. The product $\prod_{i\in I} W_i$ is Hausdorff, and so if ϕ is injective, V must be Hausdorff.

(6): $\phi : V \to \phi(V) \subset \prod_{i\in I} W_i$ is continuous and bijective. To show that ϕ is a homeomorphism, it is enough to show that for an arbitrary topological space X, a map $g : X \to V$ is continuous if $\phi \circ g : X \to \prod_{i\in I} W_i$ is continuous. If $\phi \circ g$ is continuous, then so is the composition with the projection map $p_j : \prod_{i\in I} W_i \to W_j$, but $p_j \circ \phi = F_j$, so $F_j \circ g$ is continuous for each $j \in I$, hence g is continuous. □

Examples 6.15.

1. *Let X be a set and let $V = \mathbb{K}^X$, the set of all maps $X \to \mathbb{K}$. Let $I = X$, $W_x = \mathbb{K}$ and let $\mathrm{ev}_x : \mathbb{K}^X \to \mathbb{K}$ be the linear map $f \mapsto f(x)$, and write $\mathcal{F} = (\mathbb{K}, \mathrm{ev}_x)_{x\in X}$. The topology $\mathcal{T}_{\mathcal{F}}$ is called the* topology of pointwise convergence*, because a sequence $f_n \in \mathbb{K}^X$ converges in $\mathcal{T}_{\mathcal{F}}$ to f if and only if $f_n(x) \to f(x)$, for all x.*

(2) *Let X be a locally compact Hausdorff space. Let $C(X)$ be the vector space of all continuous functions $X \to \mathbb{K}$. Let $\mathcal{K}$ be the set of all compact subsets of X, and for $K \in \mathcal{K}$, consider $C(K)$ with the supremum norm and the linear map $r_K : C(X) \to C(K)$, $f \mapsto f|_K$. The family $\mathcal{F} := (C(K), r_K)_{K\in\mathcal{K}}$ induces the* topology of uniform convergence on compact subsets of X.
(3) *Let $V = C^\infty([0,1])$, and for $n \in \mathbb{N}_0$ consider the map $F_n : C^\infty([0,1]) \to C([0,1])$ (the target has the supremum norm), $F_n(f) := f^{(n)}$. The topology induced by the family $(C([0,1]), F_n)_{n\in\mathbb{N}_0}$ is the* C^∞-topology *and is Hausdorff.*
(4) *(very important example) Let V be a normed vector space. The family $(\mathbb{K}, F)_{F\in V'}$ induces the* weak topology wk *on V. More explicitly, we take the collection of maps $F : V \to \mathbb{K}$, where F ranges through the whole dual space of V. The weak topology is Hausdorff: if $0 \neq v \in V$, there is $F \in V'$ with $F(v) \neq 0$, by the Hahn-Banach theorem. A sequence $v_n \in V$ converges in the weak topology to v ("converges weakly") if and only if $F(v_n) \to F(v)$ for each $F \in V'$. A norm-convergent sequence converges weakly (since each F is continuous for the norm topology), but the opposite is not always true. For example, the sequence $\delta_n \in \ell^2(\mathbb{N})$ converges weakly to 0. The identity map* $\mathrm{id} : (V, \|\_\|) \to (V, \mathrm{wk})$ *is continuous, since each $F \in V'$ is continuous as a map $(V, \|\_\|) \to \mathbb{K}$, by definition. This means that the weak topology is coarser than the norm topology: the weak topology has fewer open and closed sets than the norm topology. It has more compact subsets than the norm topology, and more convergent sequences. There are more continuous maps into (V, wk) are fewer continuous maps out of (V, wk). We shall learn that the identity* $\mathrm{id} : (V, \mathrm{wk}) \to (V, \|\_\|)$ *is continuous if and only if V is finite-dimensional. In functional analysis, the usual wording is that the weak topology is* weaker *than the norm topology.*
(5) *(another very important example) Let V be a normed space. For $v \in V$, we let $\iota_v : V' \to \mathbb{K}$ be the map $F \mapsto F(v)$ (this is exactly $\iota_V(v) \in V''$, of course). The family $(\mathbb{K}, \iota_v)_{v\in V}$ induces the* weak-∗-topology wk^* *on V'. The weak-∗-topology on V is Hausdorff by definition: if $F \in V'$ is nonzero, there is $v \in V$ with $F(v) \neq 0$. The identity* $\mathrm{id} : (V', \|\_\|) \to (V', \mathrm{wk}^*)$ *is continuous (by construction again) and its inverse is continuous if and only if* $\dim(V) < \infty$.
(6) *The weak and weak-∗-topologies have many common features, and they can indeed both be obtained from the same construction. Let V, W be $\mathbb{K}$-vector spaces and let $\beta : V \times W \to \mathbb{K}$ be bilinear. Each $w \in W$ defines a linear map $\beta_w : V \to \mathbb{K}$, $v \mapsto \beta(v, w)$. The topology on V induced by the family $(\mathbb{K}, \beta_w)_{w\in W}$ is the* $\sigma(V, W)$-topology*. From*

$$V \times V' \to \mathbb{K},\ (v, F) \mapsto F(v),$$

one obtains the weak topology on V as the $\sigma(V, V')$-topology, and from

$$V' \times V \to \mathbb{K}\ (F, v) \mapsto F(v)$$

the weak-∗-topology on V' as the $\sigma(V', V)$-topology.

It should be said at the very beginning that the behaviour of the weak and the weak-∗-topology is very different in very important aspects.
(7) *For two normed spaces V, W, the* strong operator topology *on $\mathcal{L}(V, W)$ is the topology induced by the maps $E_v : \mathcal{L}(V, W) \to W$, $E_v(F) = F(v)$,*

$v \in V$. It is coarser than the norm topology. There is also a weak operator topology, which is induced from the maps ($(v, L) \in V \times V'$)

$$E_{v,L} : \mathcal{L}(V, W) \to \mathbb{K};\ F \mapsto L(Fv).$$

It is usually only considered when V, W are both Hilbert spaces, and not very important in the beginning.

From this list of examples, one sees that construction 6.13 gives rise to a plethora of examples, and promises to be a very useful tool to encode various notions of convergence.

6.3. **Locally convex spaces: some basic lemmas.** We now develop the general theory of locally convex spaces. This begins with some lemmas.

Lemma 6.16. *Let V and W be topological vector spaces and let $F : V \to W$ be linear. Then F is continuous if and only if F is continuous at 0.*

Proof. Let $v \in V$, and assume that F is continuous at 0. Let $T_v : V \to V$ be the translation map $x \mapsto x + v$. The translation map is a homeomorphism because the addition in V is continuous. It follows from the linearity of F that

$$F = T_{F(v)} \circ F \circ T_{-v}.$$

Because F is continuous at 0 and because T_{-v} and $T_{F(v)}$ are homeomorphisms, $T_{F(v)} \circ F \circ T_{-v}$ is continuous at $v = T_{-v}^{-1}(0)$, so F is continuous at v. □

Corollary 6.17. *Let V be a topological vector space, W a normed vector space and let $F : V \to W$ be linear. Then F is continuous if and only if $F^{-1}(B_1(0))$ is a neighborhood of 0.*

Proof. One direction is trivial. For the other, let $r > 0$; then $F^{-1}(B_r(0)) = rF^{-1}(B_1(0))$ is a neighborhood of 0 (because the map $S_r : V \to V$, $v \mapsto rv$ is a homeomorphism). The sets $B_r(0)$ form a neighborhood basis of $0 \in W$; hence this observation shows that F is continuous at 0, and by Lemma 6.16, F is continuous. □

Next, we need to understand *continuous seminorms* on topological vector spaces. We first introduce some notation. Let V be a vector space and let W be a normed vector space. To a linear map $F : V \to W$, we associate the seminorm

$$p_F(v) := \|Fv\|.$$

If p is a seminorm on V, we let

$$N_p := \{v \in V | p(v) = 0\} \subset V$$

which is a linear subspace by 1.31. Let $\pi_p : V \to V/N_p$ be the quotient and let $\|\_\|_p$ be the quotient norm on V/N_p, see Proposition 1.31. It was shown there that $\|\pi_p(v)\| = p(v)$.

Lemma 6.18. *Let V be a topological vector space and let W be a normed vector space with norm $\|\_\|_W$. Then*

1. *a linear map $F : V \to W$ is continuous if and only if the seminorm p_F defined by $p_F(v) := \|Fv\|_W$ is continuous,*
2. *a seminorm p is continuous if and only if the quotient map $\pi_p : V \to V/N_p$ is continuous,*

(3) *a seminorm p is continuous if and only if $p^{-1}([0,1))$ is a neighborhood of* 0.

Proof. (1) One direction is clear since $p_F = \|\_\|_W \circ F$ and since the norm on a normed vector space is continuous. If p_F is continuous, then $p_F^{-1}([0,1)) = F^{-1}(B_1(0))$ is a neighborhood of 0, and continuity of F follows from Corollary 6.17.

(2): Note that $p_{\pi_p} = p$. Hence the claim follows from (1).

(3): if p is continuous, then $p^{-1}([0,1))$ is a neighborhood. Vice versa, if $p^{-1}([0,1)) = \pi_p^{-1}(B_1(0))$ is a neighborhood of 0, Corollary 6.17 implies that π_p is continuous, and (2) implies that p is continuous. □

Next, we need to understand continuous linear maps maps from $(V, \mathcal{T}_\mathcal{F})$ to some normed vector space. This is *not* supported by the universal property of the construction, which concerns continuous maps *to* V.

Lemma 6.19. *Let V be a $\mathbb{K}$-vector space and let $\mathcal{F} = (W_i, F_i)_{i\in I}$ be a family of linear maps from V to normed spaces. Let X be a further normed vector space and let $G : V \to X$ be linear. Then $G : (V, \mathcal{T}_\mathcal{F}) \to X$ is continuous if and only if there are $i_1, \dots, i_n \in I$ and $c_j > 0$ such that*

$$\|G(v)\| \le \sum_{j=1}^{n} c_j \|F_{i_j}(v)\|$$

for all $v \in V$.

Proof. The inequality

$$\|G(v)\| \le \sum_{j=1}^{n} \|c_j F_{i_j}(v)\| \tag{6.20}$$

for all $v \in V$ implies

$$\bigcap_{j=1}^{n} F_{i_j}^{-1}(B_{\frac{1}{nc_j}}(0)) \subset G^{-1}(B_1(0)).$$

From that, we see that the inequality (6.20) with continuous F_{i_j} implies that $G^{-1}(B_1(0))$ is a neighborhood of 0 and hence by Corollary 6.17 the continuity of G.

Vice versa, if G is continuous, then $G^{-1}(B_1(0))$ is a neighborhood of 0. By the definition of the topology $\mathcal{T}_\mathcal{F}$, we find $i_1, \dots, i_n \in I$ and $\epsilon_j > 0$ such that

$$\bigcap_{j=1}^{n} F_{i_j}^{-1}(B_{\epsilon_j}(0)) \subset G^{-1}(B_1(0)).$$

In other words

$$\|F_{i_j} w\| < \epsilon_j,\ j = 1, \dots, n \Rightarrow \|Gw\| < 1.$$

Put $\delta := \min_j \epsilon_j > 0$. Then

$$\sum_{j=1}^{n} \|F_{i_j} w\| < \delta \Rightarrow \|Gw\| < 1.$$

For $v \in V$ with $\|Gv\| \neq 0$, put $w := \frac{1}{\|Gv\|} v$. Then $\|Gw\| = 1$, so that $\sum_{j=1}^n \|F_{i_j} w\| \geq \delta$. Hence

$$1 \leq \frac{1}{\delta} \sum_{j=1}^n \|F_{i_j} w\| = \frac{1}{\|Gv\|} \frac{1}{\delta} \sum_{j=1}^n \|F_{i_j} v\|$$

or

$$\|Gv\| \leq \frac{1}{\delta} \sum_{j=1}^n \|F_{i_j} v\|$$

for all $v \in V$, as desired. □

Furthermore, we need a general construction of sublinear maps on locally convex spaces.

Lemma 6.21. *Let V be a locally convex space and let $U \subset V$ be a convex open neighborhood of 0. Define the* Minkowski functional of U *by the formula*

$$q_U(v) := \inf_t \{t > 0 | \frac{1}{t} v \in U\} \in [0, \infty).$$

Then q_U is sublinear and $q_U^{-1}[0,1) = U$.

If U is balanced, q_U is a continuous seminorm. If U is not balanced, we find at least a continuous seminorm $p : V \to [0, \infty)$ with $q_U \leq p$.

One can prove without too much effort that q_U is continuous for all convex open U, but that is not needed for our purposes.

Proof. Since U is an open neighborhood of 0, there is for each $v \in V$ a $t > 0$ with $\frac{1}{t} v \in U$, so that the infimum is taken over a nonempty set of positive real numbers, hence well-defined.

It is clear that $q_U(0) = 0$. Let $s > 0$. Then $\frac{1}{t} v \in U \Leftrightarrow \frac{1}{st} sv \in U$. Therefore

$$q_U(sv) = s q_U(v).$$

Let $v, w \in V$ and let $s, t > 0$ be such that $\frac{v}{t} \in U$ and $\frac{1}{s} w \in U$. Since U is convex, we have

$$\frac{1}{s+t}(v+w) = \frac{t}{s+t} \frac{v}{t} + \frac{s}{s+t} \frac{w}{s} \in U.$$

This shows

$$q_U(v+w) \leq s + t$$

and by passing to the infimum, we obtain

$$q_U(v+w) \leq q_U(v) + q_U(w).$$

Hence q_U is sublinear.

If $v \in U$, there is $s < 1$ with $\frac{1}{s} v \in U$, because U is open and because $\mathbb{R} \to U$, $s \mapsto sv$ is continuous. That proves $q_U(v) < 1$ when $v \in U$, or

$$U \subset q_U^{-1}([0,1)).$$

For the reverse inclusion, assume $q_U(v) < 1$. Then there exists $t < 1$ such that $\frac{1}{t} v \in U$. Because U is convex and $0 \in U$, this proves that

$$v = (1-t)0 + t \frac{1}{t} v \in U.$$

Altogether

$$q_U^{-1}([0,1)) \subset U.$$

It is clear that q_U is a seminorm if U is balanced. In that case, the continuity of q_U follows from Lemma 6.18.

If U is not necessarily balanced, it contains a balanced convex neighborhood $0 \in O \subset U$. The Minkowsi functional q_O is a continuous seminorm, and from the construction of the Minkowski functionals, it is clear that $q_U \leq q_O$. $\square$

6.4. **Different constructions of locally convex spaces.** Most (and probably all) textbooks introduce the topology $\mathcal{T}_{\mathcal{F}}$ defined in 6.13 in a different way.

Definition 6.22. *Let V be a $\mathbb{K}$-vector space and let $\mathcal{P}$ be a family of seminorms on V. We define a topology $\mathcal{S}_{\mathcal{P}}$ on V as follows: a set $U \subset V$ is open, if for each $x \in V$, there are $p_1, \ldots, p_r \in \mathcal{P}$ and $\epsilon_1, \ldots, \epsilon_r > 0$ such that*

$$x + \bigcap_{j=1}^{r} p_j^{-1}([0, \epsilon_j)) \subset U.$$

Lemma 6.23. *Let V be a $\mathbb{K}$-vector space.*

(1) *Let $\mathcal{F} = (W_i, F_i)_{i \in I}$ be a family of normed spaces W_i, together with linear maps $F_i : V \to W_i$. For $i \in I$, let p_i be the seminorm defined by $p_i(v) := \|F_i(v)\|_{W_i}$, and let $\mathcal{P} := (p_i)_{i \in I}$. Then $\mathcal{S}_{\mathcal{P}} = \mathcal{T}_{\mathcal{F}}$.*

(2) *Let $\mathcal{P} = (p_i)_{i \in I}$ be a family of seminorms. Let $F_i : V \to V/N_{p_i}$ be the quotient map, and let $\mathcal{F} = (V/N_{p_i}, F_i)$. Then $\mathcal{T}_{\mathcal{F}} = \mathcal{S}_{\mathcal{P}}$.*

Proof. Exercise. $\square$

In most (but not all) examples, the construction with linear maps to normed spaces is in my opinion more natural. One advantage is that the viewpoint above makes the proof of the basic properties of the construction much more straightforward (consult a textbook and study the proof that $(V, \mathcal{S}_{\mathcal{P}})$ is a topological vector space if you do not believe me).

We will now prove that *each* locally convex space is isomorphic to one obtained by this construction. This lets some authors *define* locally convex spaces as the result of the construction. As a matter of taste, I dislike such definitions.

Proposition 6.24. *Let $(V, \mathcal{T})$ be a locally convex space. Then there is a family $\mathcal{F} = (W_i, F_i)_{i \in I}$ of normed spaces and linear maps $F_i : V \to W_i$ such that $\mathcal{T} = \mathcal{T}_{\mathcal{F}}$.*

Proof. Exercise $\square$

6.5. **The Hahn-Banach Theorem for locally convex spaces.** The reason why we proved the general version of the Hahn-Banach theorem for sublinear functions (Theorem 4.10) is that this version can be applied in locally convex spaces.

Lemma 6.25. *Let V be a locally convex space and let $0 \in U \subset V$ be open and convex, and $v \in V \setminus U$. Then there is a continuous linear map $F : V \to \mathbb{R}$ such that*

$$F(x) \leq F(v) = 1$$

for all $x \in U$.

Proof. Let $q_U : V \to [0, \infty)$ be the Minkowski functional of U (see 6.21), which is sublinear. Recall that $U = q_U^{-1}(-1, 1)$.

Let $W := \operatorname{span}\{v\} \subset V$. We define a linear map $G : W \to \mathbb{R}$ by $G(v) = 1$. Since $q_U(v) \geq 1$ and $q_U \geq 0$, we have

$$G(tv) = tG(v) = t \leq t q_U(v) = q_U(tv)$$

for $t \geq 0$ and

$$G(tv) < 0 \leq q_U(tv)$$

when $t < 0$. Hence $G \leq q_U$. By Theorem 4.10, there is a linear $F : V \to \mathbb{R}$ with $F|_W = G$ and $F \leq q_U$. By construction $F(v) = 1$. It remains to prove that F is continuous.

The formula $p_U(v) := q_U(v) + q_U(-v)$ defines a seminorm on V, and since $q_U \geq 0$, we have $F \leq p_U$. We claim that p_U is continuous.

By Lemma 6.21, there is a continuous seminorm p on V with $q_U \leq p$. It follows that $|F(v)| \leq p(v)$ for all v, and so

$$p^{-1}([0,1)) \subset F^{-1}(-1,1).$$

Therefore F is continuous by Lemma 6.17. □

Theorem 6.26. *Let V be a locally convex Hausdorff space and let $v \in V$, $v \neq 0$. Then there exists a continuous functional $F : V \to \mathbb{R}$ with $F(v) = 1$.*

Proof. Because V is Hausdorff and locally convex, there is a convex open neighborhood U of v that does not contain 0. Apply Lemma 6.25. □

Theorem 6.27. *Let $A, B \subset V$ be two disjoint nonempty convex subsets, and assume hat A is open. Then there is a continuous linear functional $F : V \to \mathbb{R}$ such that for all $x \in A$ and $y \in B$, we have*

$$F(x) < \sup_{x \in A} F(x) \leq F(y).$$

Proof. The set $B - A = \cup_{y \in B}(y - A)$ is open, convex (why?) and does not contain 0 (because $A \cap B = \emptyset$). Pick $v \in A$ and $w \in B$. Then

$$U := w - v - (B - A)$$

is a convex open neighborhood of 0, which does not contain $w - v$. By Lemma 6.25, there is a nonzero continuous linear $F : V \to \mathbb{R}$ such that

$$F(x) \leq F(w - v)$$

whenever $x \in U$. If $x \in A$ and $y \in B$ are arbitrary, then $w - v - y + x \in U$, so that

$$F(w - v - y + x) \leq F(w - v)$$

or

$$F(x) \leq F(y).$$

This is almost what we want, but what is missing is that $F(x) < \sup_{z \in A} F(z)$ when $x \in A$. This follows from Lemma 6.28 below, because we constructed F to be nonzero. □

Lemma 6.28. *Let V be a locally convex space and let $F : V \to \mathbb{R}$ be continuous and linear, and $F \neq 0$. Then F is open.*

Proof. Assume that F is not open; the goal is to show that $F \equiv 0$. If F is not an open map, there must be a convex open set U so that $F(U) \subset \mathbb{R}$ is not open. But $F(U)$ is a convex subset of $\mathbb{R}$, hence an interval. The only possibility for an interval $I \subset \mathbb{R}$ to be not open is if there is $a \in I$ such that $a \leq x$ for all $x \in I$ or $a \geq x$ for all $x \in I$. Assume that the former is the case, so that there is $x_0 \in U$ with

$$F(x_0) \leq F(x)$$

for all $x \in U$. It follows that

$$F(x) \geq 0$$

for all x in the neighborhood $-x_0 + U$ of 0. If $v \in V$ is arbitrary, then for some $t > 0$, we have $tv \in -x_0 + U$ and hence

$$F(v) = \frac{1}{t} F(tv) \geq 0.$$

for all $v \in V$. But then also

$$F(v) = -F(-v) \leq 0,$$

so that $F \equiv 0$. □

The most often used version of the Hahn-Banach theorem is as follows.

Theorem 6.29 (Hahn-Banach separation theorem)**.** *Let V be a locally convex space, and let $K, B \subset V$ be disjoint convex sets, where K is compact, and B is closed. Then there is a continuous linear $F : V \to \mathbb{R}$ such that*

$$\sup_{x \in K} F(x) < \inf_{y \in B} F(y).$$

Proof. Let $O := V \setminus B$, which is an open neighborhood of K. We first prove that there is a convex open neighborhood U of 0 such that $K + U \subset O$.

For each $x \in K$, pick a convex open neighborhood U_x of 0 such that $x + U_x \subset O$. Then

$$K \subset \bigcup_{x \in K} x + \frac{1}{2} U_x.$$

By compactness of K, there are $x_1, \ldots, x_r \in K$ with

$$K \subset \bigcup_{j=1}^{r} x_j + \frac{1}{2} U_{x_j}.$$

Put

$$U = \bigcap_{j=1}^{r} \frac{1}{2} U_{x_j},$$

which is a convex open neighborhood of 0. When $x \in K$ and $h \in U$, then there is j with $x \in x_j + \frac{1}{2} U_{x_j}$. As $h \in U \subset \frac{1}{2} U_{x_j}$, we have

$$x + h \in x_j + U_{x_j} \subset O.$$

Thus $K + U \subset O$.

The set $K + U = \bigcup_{x \in K} x + U$ is open and convex and disjoint from B. So by Theorem 6.27, there is a continuous linear F with

$$F(x) < \sup_{y \in K+U} F(y) \leq \inf_{z \in B} F(z)$$

for all $x \in K$. Because K is compact and F is continuous, F attains its maximum on K, and so

$$\max_{x \in K} F(x) < \sup_{y \in K+U} F(y) \leq \inf_{z \in B} F(z),$$

as desired. □

Theorem 6.30. *The only locally convex topology on $\mathbb{R}^n$ is the usual one.*

Proof. We show the following: if V is a finite-dimensional locally convex Hausdorff space, then all linear maps $\mathbb{R}^n \to V$ and $V \to \mathbb{R}^n$ are continuous, where $\mathbb{R}^n$ is equipped with its usual topology.

Any linear map $f : \mathbb{R}^n \to V$ to a locally convex space is continuous; this is because f is given by

$$f(x_1, \ldots, x_n) := \sum_{i=1}^n x_i v_i$$

for some $v_1, \ldots, v_n \in V$.

To prove that (if V is finite-dimensional and Hausdorff) any linear map $V \to \mathbb{R}^n$ is continuous, it suffices to prove that each linear form $V \to \mathbb{R}$ is continuous.

For that purpose, let $\mathrm{Hom}_{\mathbb{R}}(V; \mathbb{R})$ be the *algebraic* dual space and let $V' \subset \mathrm{Hom}_{\mathbb{R}}(V; \mathbb{R})$ be the space of all continuous linear functionals. Now we consider the linear map

$$\eta : V \to \mathrm{Hom}_{\mathbb{R}}(V'; \mathbb{R}),\ v \mapsto (L \mapsto L(v)).$$

The map η is injective: if $\eta(v) = 0$, then $L(v) = 0$ for every $L \in V'$, and this implies $v = 0$ by Theorem 6.26. Therefore

$$\dim(V) \leq \dim(\mathrm{Hom}_{\mathbb{R}}(V'; \mathbb{R})) = \dim(V') \leq \dim(\mathrm{Hom}_{\mathbb{R}}(V; \mathbb{R})).$$

Since $\dim(V) = \mathrm{Hom}_{\mathbb{R}}(V'; \mathbb{R})$, it follows that $\dim(V') = \dim(\mathrm{Hom}_{\mathbb{R}}(V; \mathbb{R}))$ and $V' = \mathrm{Hom}_{\mathbb{R}}(V; \mathbb{R})$. In other words, each linear form on V is continuous. $\square$

7. Duality and weak topologies

7.1. **Generalities.** We now put aside the general theory of locally convex spaces and focus on the example which is most central to what follows: the weak and the weak-$*$-topology on a Banach space and on its dual. These have been defined in 6.15, and we only recall that wk is induced from the family

$$(L : V \to \mathbb{K})_{L \in V'}$$

where L runs through V', and wk* is induced from the family

$$(\iota(v) : V' \to \mathbb{K})_{v \in V},$$

where $\iota(v) : V' \to \mathbb{K}$ is $\iota(v)(L) := L(v)$. The associated notion of convergence is the weak convergence and the weak$*$-convergence which we studied in some detail already.

Lemma 7.1. *Let V be a normed space. Then the weak-$*$-topology on V' and the weak topology on V are Hausdorff.*

Proof. We have to prove that the defining families of linear maps separate the points of V and V'.

The case of the weak-$*$-topology is almost trivial: if $0 \leq L \in V'$, there is $v \in V$ with $L(v) \neq 0$. The case of the weak topology uses the Corollary 4.2 of the Hahn-Banach theorem: if $0 \neq v \in V$, there is $L \in V'$ with $L(v) \neq 0$. □

Not the statement of the Lemma, but its proof, is a harbinger for what comes. The definition of the two topologies is very similar, but the behaviour of both is fundamentally different. As we will see, each of the topologies has a distinguished nice feature, and only for the class of reflexive Banach spaces, the nice features both hold. Let us inspect this feature of reflexive spaces.

Let V be a Banach space. Through the natural map $\iota : V \to V''$, we can identify V with the closed subspace $\iota(V) \subset V''$. Being a dual space, V'' has the wk*-topology (which should be denoted $\sigma(V'', V')$-topology to avoid confusion), and V carries the wk-topology (better $\sigma(V, V')$-topology).

Lemma 7.2. *Let V be a Banach space. Then the map*

$$\iota_V : (V, \sigma(V, V')) \to (V'', \sigma(V'', V'))$$

(which is injective by Hahn-Banach) is a homeomorphism onto its image.

Proof. To show that ι is continuous, we have to verify that for all $L \in V'$, the composition

$$\mathrm{ev}_L \circ \iota : (V, \sigma(V, V')) \to \mathbb{K}$$

is continuous for each $L \in V'$. But this sends v to $\iota(v)(L) := L(v)$ and is hence equal to L. The $\sigma(V, V')$-topology is designed so that each $L \in V'$ is continuous.

To show that ι is an embedding, we have to verify that for an arbitrary topological space X, a map $g : X \to (V, \sigma(V, V'))$ is continuous once $\iota \circ g : X \to (V'', \sigma(V'', V'))$ is continuous. But if $\iota \circ g$ is continuous, then for each $L \in V'$, $\mathrm{ev}_L \circ \iota \circ g = L \circ g$ is continuous, and by the design of the $\sigma(V, V')$-topology, this means that g is continuous. □

In particular, if V is reflexive, then $\iota : (V, \sigma(V, V')) \to (V, \sigma(V'', V'))$ is an isomorphism, and weak topologies share the nice features of both, wk and wk*.

Definition 7.3. *Let V and W be vector spaces and let $\beta : V \times W \to \mathbb{K}$ be bilinear. The* β-topology *(also called $\sigma(V,W)$-topology) on V is the locally convex topology induced by the seminorms*

$$p_w(v) := |\beta(v,w)|, w \in W.$$

Dually, the β-topology *(also called* $\sigma(W,V)$-topology*) on W is induced by the family of seminorms*

$$q_v(w) := |\beta(v,w)|, v \in V.$$

For a normed vector space V, we consider the bilinear map

$$\beta : V \times V' \to \mathbb{K}, \ (v,L) \mapsto L(v).$$

The $\sigma(V,V')$-topology on V is the weak topology wk, and the $\sigma(V',V)$-topology on V' is the weak-$*$-topology wk*.

The construction is symmetric in V and W: if $\beta : V \times W \to \mathbb{K}$ is bilinear, let $\beta^t : W \times V \to \mathbb{K}$ be defined by $\beta^t(w,v) = \beta(v,w)$. Then the β-topology on V agrees with the β^t-topology (and the same for W).

Hence for the development of the general theory, there is no need to consider the $\sigma(W,V)$-topology.

The identity map

$$\mathrm{id} : (V, \|\_\|) \to (V, \mathrm{wk})$$

is continuous. This follows easily from the universal property of the induced topology. This implies that each functional $V \to \mathbb{K}$ which is weakly continuous is also norm continuous. The opposite is also true: we have *defined* the weak topology so that each norm continuous $L : V \to \mathbb{K}$ is continuous in the weak topology.

Similarly

$$\mathrm{id} : (V', \|\_\|) \to (V', \mathrm{wk}^*)$$

is continuous, and hence each wk*-continuous $L : V' \to \mathbb{K}$ is norm continuous. The wk*-topology was designed so that all evaluation functionals $\iota_v(v) : V' \to \mathbb{K}$ are continuous. However, if V is not reflexive, there are more norm-continuous functionals on V' than the evaluation functionals. Are they also wk*-continuous? The answer is "no": the wk*-continuous functionals on V' are exactly the same as the evaluation functionals. This follows from the next result.

Theorem 7.4. *Let $\beta : V \times W \to \mathbb{K}$ be bilinear. Then each continuous linear functional $(V, \sigma(V,W)) \to \mathbb{K}$ is of the form $f_w := \beta(\_, w)$.*

Corollary 7.5. *Let V be a normed space. Then the continuous functionals*

$$(V, \mathrm{wk}) \to \mathbb{K}$$

are precisely the elements of V'. The continuous functionals

$$(V', \mathrm{wk}^*) \to \mathbb{K}$$

are precisely the evaluation functionals $L \mapsto L(v)$, $v \in V$.

Proof. It is clear that f_w is continuous. Vice versa, let $F : V \to \mathbb{K}$ be continuous. By Lemma 6.19, there are $w_1, \ldots, w_n \in W$ and $C \geq 0$, so that

$$|F(v)| \leq C \sum_{j=1}^{n} |f_{w_i}(v)|.$$

Therefore
$$K := \bigcap_{i=1}^{n} \ker(f_{w_i}) \subset \ker(F).$$
The linear subspace K has finite codimension, say m, and
$$m = \dim(V/K) \leq n$$
by elementary linear algebra. Let $\pi : V \to V/K$ be the quotient map. There are linear maps
$$h : V/K \to \mathbb{K},\ h \circ \pi = F$$
and
$$g_1, \dots, g_n : V/K \to \mathbb{K},\ g_i \circ \pi = f_{w_i}$$
by the universal property of the quotient of vector spaces.

The map
$$g = (g_1, \dots, g_n) : V/K \to \mathbb{K}^n$$
is by construction injective. Hence by elementary linear algebra again, there is a linear map $k : \mathbb{K}^n \to \mathbb{K}$ such that
$$k \circ g = h.$$
There is $(a_1, \dots, a_n) \in \mathbb{K}^n$ such that
$$k(x_1, \dots, x_n) = \sum_{i=1}^{n} a_i x_i$$
for some all $(x_1, \dots, x_n) \in \mathbb{K}^n$. Together, we get, for all $v \in V$, that
$$F(v) = h(\pi(v)) = k(g(\pi(v))) = k(g_1(\pi(v)), \dots, g_n(\pi(v))) = \sum_{i=1}^{n} a_i g_i(\pi(v)) = \sum_{i=1}^{n} a_i f_{w_i}(v) = f_{\sum_{i=1}^{n} a_i w_i}(v).$$
□

7.2. **The Banach-Alaoglu theorem.** Corollary 7.5 reflects the aspect in which wk and wk* are very similar. Now we turn to the differences, and first discuss the most distinguished feature of the wk*-topology, which is one of the central theorem of functional analysis.

Theorem 7.6 (Banach-Alaoglu theorem)**.** *Let V be a normed space. Then the closed unit ball $D_1(V')$ of the dual space is compact (and Hausdorff) when equipped with the weak-$*$-topology.*

Proof. We recall from Lemma 6.14 that the map
$$\phi : (V', \mathrm{wk}^*) \to \prod_{v \in V} \mathbb{K},\ \phi(L) := (L(v))_{v \in V}$$
is injective and identifies (V', wk^*) with the subspace $\phi(V')$ of the product. By definition, ϕ restricts to a map
$$\phi : D_1(V') \to \prod_{v \in V} D_{\|v\|}(\mathbb{K})$$
Since $D_{\|v\|}(\mathbb{K})$ is compact by the Heine-Borel theorem, the product is compact by Tychonov's theorem. It remains to be shown that the subspace $\phi(D_1(V')) \subset \prod_{v \in V} \mathbb{K}$ is closed.

But this is the intersection of the sets

$$(p_v + p_w - p_{v+w})^{-1}(0)$$

and

$$(ap_v - p_{av})^{-1}(0),$$

where v, w runs through V and a through $\mathbb{K}$, and $p_v : \prod_{v \in V} \mathbb{K} \to \mathbb{K}$ denotes the projection onto the vth factor. But the maps $p_v + p_w - p_{v+w}$ and $ap_v - p_{av}$ are continuous, and so their preimages are closed, and so is the intersection, which is $\Phi(D_1(V'))$. $\square$

The Banach-Alaoglu theorem is a powerful result, but many arguments in analysis are based on *sequences*. Therefore, it is desirable that the unit ball $D_1(V')$ is *sequentially compact*.

Let us recall the connection between these two notions of compactness (for Hausdorff spaces):

(1) A first-countable and compact space is sequentially compact.
(2) A second-countable and sequentially compact space is compact.
(3) For metrizable spaces, compactness and sequential compactness are equivalent.

This leaves the question: when is the unit ball $D_1(V')$ with the weak-$*$-topology first countable? There does not seem to be an easy general answer, but the following result gives a satisfactory answer.

Proposition 7.7. (1) *A compact Hausdorff space X is metrizable if and only if $C(X)$ is separable.*
(2) *The unit ball $D_1(V')$ with the weak-$*$-topology is metrizable if and only if V is separable.*

Hence $D_1(V')$ is sequentially compact when V is separable.

Proof. Step 1: If X is metrizable, then by Lemma 5.4 $C(X)$ is isometrically isomorphic to a subspace of $C(Y)$, where Y has a countable basis consisting of clopen sets. By Lemma 5.6, the linear combinations of their characteristic functions span a countably-dimensional dense subspace of $C(Y)$. Thus $C(Y)$ is separable, and so is the subspace $C(X)$.

Step 2: Let $D_1(V')$ be metrizable. Then it is a compact metric space, by the Banach-Alaoglu theorem. The space $C(D_1(V'))$ of functions (with the usual supremum norm) is then separable by the first step of the proof. If $v \in V$, then the function $R(v) : D_1(V') \to \mathbb{K}$ which sends L to $L(v)$, is continuous when $D_1(V')$ has the weak-$'$-topology. This defines a linear map

$$R : V \to C(D_1(V'))$$

which is an isometry because $\|R(v)\|_{C^0} = \sup_{L \in D_1(V')} |L(v)| = \|v\|$ (the last equation uses the Hahn-Banach theorem). Therefore, V is isometrically isomorphic to a subspace of the separable space $C(D_1(V'))$ and therefore itself separable.

Step 3: Assume that V is separable and let $\{v_n\} \subset V$ be a dense countable subset. The map

$$S : D_1(V') \to Z := \prod_{n=1}^{\infty} D_{\|v_n\|}(\mathbb{K})$$

which sends $L \in D_1(V')$ to the family $(L(v_n))_n$ is continuous by the definition of the weak-$*$-topology. It is furthermore injective: if $L_0 \neq L_1$, there is n so that

$L_0(v_n) \neq L_1(v_n)$ because $\{v_n\}$ is dense. Both $D_1(V')$ and Z are compact Hausdorff spaces, and so $D_1(V')$ is homeomorphic to the subspace $S(D_1(V')) \subset Z$. But Z, being a countable product of compact metric spaces is metrizable by Proposition B.35, and hence so is $S(D_1(V')$, and so $D_1(V')$ as well.

Step 4: Assume that $C(X)$ is separable. The map $g : X \to D_1(C(X)')$, $x \mapsto \mathrm{ev}_x$, is injective (Urysohn Lemma). It is continuous by the definition of the wk*-topology (when composed with $L \mapsto L(f)$, it becomes just f). Because X is compact, g is a homeomorphism onto the closed subspace $g(X) \subset D_1(C(X)')$. Since $C(X)$ is separable, $D_1(C(X)')$ is metrizable by step 3, and it is follows that X is metrizable. □

Remark 7.8. *The following fact is slightly confusing. If V is an infinite-dimensional Banach space, then neither (V, wk) nor (V', wk^*) are first countable, let alone metrizable.*

To see this, let $\beta : V \times W \to \mathbb{K}$ be a pairing and assume that $(V, \sigma(V, W))$ is first countable. We claim that W is at most countably-dimensional. Let $(U_n)_{n\in\mathbb{N}}$ be a neighborhood basis of $0 \in V$. For each n, there are $w_{n,1}, \dots, w_{n,r_n} \in W$ and $\epsilon > 0$ such that

$$\bigcap_{j=1}^{r_n} \beta^{-1}_{w_{n,j}}(B_\epsilon(0)) \subset U_n.$$

The collection $\mathcal{C} := \{w_{n,j} | n \in \mathbb{N}, 1 \leq j \leq r_n\}$ is countable. We claim that it spans W. To that end, let $w \in W$. Then $\beta_w : (V, \sigma(V, W)) \to \mathbb{K}$ is continuous, hence $\beta_w^{-1}(B_1(0))$ is open, hence there is n such that

$$\bigcap_{j=1}^{r_n} \beta^{-1}_{w_{n,j}}(B_\epsilon(0)) \subset \beta_w^{-1}(B_1(0)).$$

In particular

$$\bigcap_{j=1}^{r_n} \ker(\beta_{w_{n,j}}) \subset \beta_w^{-1}(B_1(0)),$$

but as the left hand side is a linear subspace, we get

$$\bigcap_{j=1}^{r_n} \ker(\beta_{w_{n,j}}) \subset \ker(\beta_w).$$

Exactly as in the proof of Theorem 7.4, this proves

$$w \in \mathrm{span}\{w_{n,1}, \dots, w_{n,r_n}\}.$$

Corollary 7.9. *Let V be a separable Banach space and let $L_n \in V'$ be a bounded sequence. Then there is a subsequence L_{n_k} and $L \in V'$ such that $L_{n_k}(v) \to L(v)$ for each $v \in V$.*

Here are some examples.

Example 7.10. *Let V be a separable Hilbert space and let $v_n \in V$ be a bounded sequence. Then there is a subsequence v_{n_k} and $v \in V$ such that $\langle w, v_{n_k}\rangle \to \langle w, v\rangle$ for all $w \in V$.*

Example 7.11. *Let X be a compact metric space and let* $\mathrm{Prob}(X)$ *be the set of all probability Radon measures on X. The map*

$$\psi : \mathrm{Prob}(X) \to C(X)',\ \mu \mapsto (f \mapsto \int_X f d\mu)$$

is injective, we identify $\mathrm{Prob}(X)$ *with* $\psi(\mathrm{Prob}(X))$ *and equip is with the* wk*-*topology. Clearly* $\mathrm{Prob}(X) \subset D_1(C(X)')$. *By the Riesz-Markov-Kakutani theorem,* $\mathrm{Prob}(X)$ *is the space of all $L \in C(X)'$ with $L(1) = 1$ and with $L(f) \geq 0$ for $f \geq 0$. Let $I_f : C(X)' \to \mathbb{K}$ be the map $I_f(L) = L(f)$. Hence*

$$\mathrm{Prob}(X) = I_1(1) \cap \bigcap_{f \geq 0} I_f^{-1}([0, \infty))$$

and since each I_f is continuous, it follows that $\mathrm{Prob}(X)$ *is closed in the* wk*-*topology. Being a subset of $D_1(C(X)')$, it is compact. Because X is metrizable, $C(X)$ is separable, and so* $\mathrm{Prob}(X)$ *is sequentially compact. This argument proves:*

Proposition 7.12. *Let X be a compact metric space and let μ_n be a sequence of probability Radon measures on X. Then there is a probability measure μ and a subsequence μ_{n_k} such that*

$$\lim_{k \to \infty} \int_X f d\mu_{n_k} = \int_X f d\mu.$$

Note that we had to use the Markov-Kakutani theorem to construct the measure μ.

7.3. **Mazur's Lemma.**

The identity map

$$(V, \|\_\|) \to (V, \mathrm{wk})$$

is clearly continuous. It follows that a weakly closed set $Z \subset V$ is also $\|\_\|$-closed, and that

$$\overline{Z}^{\|\_\|} \subset \overline{Z}^{\mathrm{wk}}$$

for each subset $Z \subset V$. The converse is not true:

Lemma 7.13. *Let V be an infinite-dimensional Banach space. Then the weak closure of the unit sphere $S(V) := \{v \in V | \|v\| = 1\}$ is $D_1(V)$.*

finish proof

Theorem 7.14. *Let V be a Banach space and let $K \subset V$ be convex. If K is norm closed, then it is weakly closed.*

Proof. The point is that the set of continuous functionals $V \to \mathbb{K}$ is the same for the weak and the norm topology.

If $v \in V$, $v \notin K$, we apply the Hahn-Banach separation theorem to the $\|\_\|$-closed set K and the $\|\_\|$-compact set $\{v\}$. The result is a $\|\_\|$-continuous functional $L_v : V \to \mathbb{R}$ with

$$a_v := \max_{x \in K} L_v(x) < L(v).$$

It follows that K is the intersection

$$\bigcap_{v \in V \setminus K} L_v^{-1}((-\infty, a_v]).$$

Since each L_v is continuous when V has the weak topology, it follows that K is weakly closed. $\square$

Theorem 7.14 has a nice consequence. First some notation. Let V be an $\mathbb{R}$-vector space and $Z \subset V$ a subset. The *convex hull* $\mathrm{co}(Z) \subset V$ is the smallest convex set containing Z (this is well-defined: it is the intersection of all convex subsets which contain Z). We denote

$$\Delta^n := \{(t_0, \dots, t_n) \in \mathbb{R}^{n+1} | t_j \geq 0, \sum_{j=0}^{n} t_j = 1\}.$$

If $v = (v_0, \dots, v_n) \in V^{n+1}$ is a tuple of $n+1$ points and $t \in \Delta^n$, the convex combination of those points is

$$t \cdot v := \sum_{j=0}^{n} t_j v_j.$$

One may alternatively describe $\mathrm{co}(Z)$ as the set

$$\mathrm{co}(Z) = \{t \cdot v | n \in \mathbb{N}_0, t \in \Delta^n, v \in V^{n+1}\}.$$

Elementary exercise: prove this.

Now let $v_n \in V$ be a weakly convergent sequence in V, with limit v. The terms of the sequence lie in the convex hull $Z := \mathrm{co}(\{v_n | n \in \mathbb{N}\})$, and the weak limit v lies in the weak closure

$$v \in \overline{Z}^{\mathrm{wk}}.$$

Because $\mathrm{id} : (V, \|\_\|) \to (V, \mathrm{wk})$ is continuous, we have

$$\overline{Z}^{\|\_\|} \subset \overline{Z}^{\mathrm{wk}}.$$

On the other hand, $\overline{Z}^{\mathrm{wk}}$ is norm closed by Theorem 7.14, and is therefore a norm-closed set containing Z. This shows (by the definition of the closure in a topological space) that

$$\overline{Z}^{\mathrm{wk}} \subset \overline{Z}^{\|\_\|}.$$

Therefore, the two closures agree. It follows that the weak limit v of the sequence v_n belongs to the *norm closure* of Z. Because

$$Z = \bigcup_{n=0}^{\infty} \mathrm{co}(v_0, \dots, v_n)$$

is an ascending union, we find a sequence $t_n \in \Delta^n$ such that we have *norm convergence*

$$\lim_{n \to \infty} t_n \cdot (v_0, \dots, v_n) = v.$$

This argument proves:

Theorem 7.15 (Mazur)**.** *Let $v_n \in V$ be a weakly convergent sequence in a Banach space, with weak limit v. Then there exists a sequence $w_n \in V$ of vectors such that*

$$\|w_n - v\| \to 0$$

and such that each w_n is a convex combination of the vectors $v_0, \dots, v_n$.

The remarkable point is that we upgraded weak convergence to norm convergence, not of the sequence, not of a subsequence, but of a sequence of convex combinations of the original sequence.

By itself, this is not a very powerful result. But when V is reflexive and separable, we can combine it with the Banach-Alaoglu theorem and obtain a corollary which is very useful for PDE theory. The hypotheses are satisfied for $L^p(X,\mu)$ where X is a second countable locally compact Hausdorff space, μ is a Radon measure and $1 < p < \infty$.

Theorem 7.16. *Let V be a reflexive Banach space with separable dual and let $v_n \in V$ be a bounded sequence. Then after passage to a subsequence, we find a sequence of convex combinations of the sequence which converges in norm.*

For Hilbert spaces, this result can be improved a little bit (with an easier proof). In that case, we get that after passage to a subsequence, the sequence $\frac{1}{n}\sum_{k=1}^n v_k$ is norm-convergent.

Proof. Without loss of generality $\|v_n\| \leq 1$. We identify $V = V''$. Under this identification, the weak topology on V and the wk*-topology coincide. By the Banach-Alaoglu theorem, the unit ball $D_1(V)$ is wk*- and hence wk-compact, and because V' is separable, also sequentially compact. Therefore v_n has a weakly convergent subsequence, and applying Theorem 7.15 finishes the proof. □

Theorem 7.17 (Goldstine's Theorem)**.** *Let V be a normed space, and let $\iota : V \to V''$ be the isometric embedding into its bidual. Then*

$$\iota(V) \subset (V'', \sigma(V'', V'))$$

and

$$\iota(D_1(V)) \subset (D_1(V''), \sigma(V'', V'))$$

are dense.

Proof. The closure $\overline{\iota(V)}$ is convex. If it is not all of V'', we can find $\varphi \in V'' \setminus \overline{\iota(V)}$. Apply the Hahn-Banach separation theorem 6.29 to the compact set $\{\varphi\}$ and the closed set $\overline{\iota(V)}$: it follows that there is a continuous functional $F : (V'', \sigma(V'', V')) \to \mathbb{K}$ such that

$$F(\varphi) < \inf_{v \in V} F(\iota(v)).$$

But F is an element of V' by Theorem 7.4, and so the above can be rewritten as

$$\varphi(F) < \inf_{v \in V} F(v) \in \mathbb{R}. \tag{7.18}$$

But V is a linear subspace, and because the linear map $F : V \to \mathbb{R}$ is bounded below, we must have $F = 0$. This is a contradiction, because (7.18) also shows $\varphi(F) < 0$.

The argument for the second claim is similar: $\overline{\iota(D_1(V))}$ is convex, and if it not all of $D_1(V'')$, there is $\varphi \in D_1(V'') \setminus \overline{\iota(D_1(V))}$ and $F \in V'$ such that

$$\varphi(F) < \inf_{v \in D_1(V)} Fv.$$

Because $\|\varphi\| \leq 1$, we also have

$$-\|F\| = \inf_{v \in D_1(V)} Fv < \varphi(F)$$

which is a contradiction. □

7.4. **Reflexivity.**

Theorem 7.19. *The following conditions for a Banach space V are equivalent:*

(1) *V is reflexive,*
(2) *The unit ball $D_1(V)$ is compact in the $\sigma(V, V')$-topology.*

Proof. If V is reflexive, the map $\iota : (V, \sigma(V, V')) \to (V'', \sigma(V'', V'))$ is a homeomorphism. Since the unit ball $D_1(V'')$ is $\sigma(V'', V')$-compact by the Banach-Alaoglu theorem, and therefore $\sigma(V, V')$-compact.

Vice versa, if the unit ball $D_1(V)$ is $\sigma(V, V')$-compact, its image under $\iota : (V, \sigma(V, V')) \to (V'', \sigma(V'', V'))$ is compact as well. On the other hand, by Goldstine's Theorem, $\iota(D_1(V)) \subset D_1(V'')$ is $\sigma(V'', V')$-dense. Since the $\sigma(V'', V')$-topology is Hausdorff, $\iota(D_1(V))$ is both closed and dense. This is only possible if $\iota(D_1(V)) = D_1(V'')$, but that implies easily that $\iota(V) = V''$, i.e. if V is reflexive. $\square$

8. Compact and Fredholm operators on Hilbert spaces

8.1. Compact operators.

Definition 8.1. *A bounded operator $F : V \to W$ between Banach spaces is* compact *if $\overline{F(D_1(V))} \subset W$ is compact. By $\mathcal{K}(V,W) \subset \mathcal{L}(V,W)$, we denote the set of all compact operators.*

Equivalently, each bounded sequence $v_n \in V$ has a subsequence v_{n_k} such that Fv_{n_k} is Cauchy.

Examples 8.2. *If the rank* $\operatorname{rank}(F) := \dim(\operatorname{im}(F))$ *is finite (for example because V or W are finite-dimensional), then F is compact. An isomorphism $F : V \to W$ is compact only if V and W are finite-dimensional.*

More generally:

Lemma 8.3. *Let $F : V \to W$ be a compact operator between Banach spaces and let $U \subset V$ be a closed linear subspace. Assume that there is $c > 0$ such that $\|Fu\| \geq c\|u\|$ for all $u \in U$. Then U is finite-dimensional.*

Proof. We show that $D_1(U)$ is compact and invoke Proposition 1.27. So let $u_n \in U$ with $\|u_n\| \leq 1$. We need to find a convergent subsequence. Because F is compact, we can assume that Fu_n is a Cauchy sequence, after passage to a subsequence. Then $\|u_n - u_m\| \leq \frac{1}{c}\|Fu_n - Fu_m\|$ proves that u_n is a Cauchy sequence. □

Theorem 8.4. *Let V, W, U be Banach spaces.*

1. *If $S \in \mathcal{L}(W,U)$ or $T \in \mathcal{L}(V,W)$ is compact, then ST is compact.*
2. *$\mathcal{K}(V,W) \subset \mathcal{L}(V,W)$ is a closed linear subspace.*

Proof. (1) is clear, since bounded operators map bounded sequences to bounded sequences, and Cauchy sequences to Cauchy sequences. (2): it is easy to see that $\mathcal{K}(V,W)$ is a linear subspace. It remains to prove that if $T_n \in \mathcal{K}(V,W)$ converges in norm to $T \in \mathcal{L}(V,W)$, then T is compact. Let v_n be a bounded sequence in V, and assume that $\|v_n\| \leq C$ for all n. There is a subsequence $v_{1,n}$ such that $K_1 v_{1,n}$ is Cauchy. Next, there is a subsequence $v_{2,n}$ of $v_{1,n}$ such that $K_2 v_{2,n}$ is Cauchy. Continuing in this way, we find subsequences $v_{k,n}$, $v_{k,n}$ a subsequence of $v_{k-1,n}$ such that $K_k v_{k,n}$ is Cauchy.

The diagonal sequence $n \mapsto v_{n,n}$ is a subsequence of the original sequence, and for each k, $K_k v_{n,n}$ is a Cauchy sequence. Then

$$\|Kv_{n,n} - Kv_{m,m}\| \leq \|Kv_{n,n} - K_k v_{n,n}\| + \|K_k(v_{n,n} - v_{m,m})\| + \|K_k v_{m,m} - Kv_{m,m}\| \leq$$
$$\leq \|K - K_k\|(\|v_{n,n}\| + \|v_{m,m}\|) + \|K_k(v_{n,n} - v_{m,m})\| \leq 2C\|K - K_k\| + \|K_k(v_{n,n} - v_{m,m})\|.$$

Therefore

$$\limsup_n \limsup_m \|Kv_{n,n} - Kv_{m,m}\| \leq 2C\|K - K_k\| + \limsup_n \limsup_m \|K_k(v_{n,n} - v_{m,m})\| = 2C\|K - K_k\|.$$

Since k was arbitrary and $\|K - K_k\| \to 0$, it follows that

$$\limsup_n \limsup_m \|Kv_{n,n} - Kv_{m,m}\| = 0,$$

i.e. that $Kv_{n,n}$ is a Cauchy sequence. □

Lemma 8.5. *Let V, W Banach spaces and let $T : V \to W$ be bounded. The following are equivalent:*

(1) *T is compact.*
(2) *The dual operator $T' \in \mathcal{L}(W', V')$ is compact.*

Proof. $1 \Rightarrow 2$: The proof uses the Arzela-Ascoli theorem B.63 which the reader should recall. If T is compact, then $K := \overline{T(B_1(V))} \subset W$ is a compact metric space and in particular separable. Let $L_n \in W'$ be a bounded sequence. Each L_n gives a continuous function $L_n|_K : K \to \mathbb{K}$. The set $\{L_n|_K\} \subset C(K, \mathbb{K})$ is equicontinuous and uniformly bounded. By the Arzela-Ascoli theorem, there is a subsequence L_{n_m} such that $L_{n_m}|_K$ is uniformly convergent. Unwinding the definitions, this means that $T'(L_{n_m}) = L_{n_m} \circ T \in V'$ is a Cauchy sequence.

$2 \Rightarrow 1$. Let T' be compact. By the already proven implication $1 \Rightarrow 2$, the bidual operator $T'' : V'' \to W''$ is compact. The diagram

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\iota_V"] & W \arrow[d, "\iota_W"] \\ V'' \arrow[r, "T''"] & W'' \end{tikzcd}$$

commutes. It follows that $T'' \circ \iota_V = \iota_W \circ T$ is compact. So when v_n is a bounded sequence in V, the sequence $\iota_W(Tv_n)$ has a subsequence which is Cauchy. By the Hahn-Banach theorem, ι_W is an isometric embedding, and so Tv_n has a subsequence which is Cauchy. □

Let us give some instructive examples of compact operators.

Example 8.6. *Let $a \in \ell^\infty(\mathbb{N})$. Then the formula $T_a(f) := af$ defines a bounded operator $T_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ with $\|T_a\| = \|a\|_{\ell^\infty}$. We claim that T_a is compact if and only if $a \in c_0(\mathbb{N})$. If $a \in c_{00}(\mathbb{N})$, then T_a has finite rank and is hence compact. For a general element $a \in c_0(\mathbb{N})$, there is a sequence $a_n \in c_{00}(\mathbb{N})$ which converges to a. It follows that $\|T_{a_n} - T_a\| \to 0$, so that T_a is compact.*

Vice versa, assume that T_a is compact. We have to prove that for $\epsilon > 0$, the set $I := \{n \in \mathbb{N} | |a(n)| \geq \epsilon\}$ is finite. The subspace $\ell^2(I) = \{f \in \ell^2(\mathbb{N}) | \operatorname{supp}(f) \subset I\} \subset \ell^2(\mathbb{N})$ is closed. For $f \in \ell^2(I)$, we have

$$\|T_a f\|^2 = \sum_{n \in I} |a(n)|^2 |f(n)|^2 \geq \epsilon^2 \sum_{n \in I} |f(n)|^2 = \epsilon^2 \|f\|^2.$$

It follows from Lemma 8.3 that $\ell^2(I)$ is finite-dimensional, or that I is finite.

8.2. **Compact operators on Hilbert spaces.** We now specialize to the case of Hilbert spaces.

Corollary 8.7. *A bounded operator $T : V \to W$ of Hilbert spaces is compact if and only if T^* is compact.*

Proof. It is enough to prove that T^* is compact if T is. Let $\mu_V : V \to V'$ be the map $v \mapsto (w \mapsto \langle v, w \rangle)$, which is a conjugate linear isometric isomorphism. Now

$$\mu_V \circ T^* = T' \circ \mu_W : W \to V'.$$

This is because for $w \in W$ and $v \in V$, we have

$$T'(\mu_W(w))(v) = (\mu_W(w))(Tv) = \langle w, Tv \rangle = \langle T^* w, v \rangle$$

and

$$\mu_V(T^*(w))(v) = \langle T^* w, v \rangle.$$

Because μ_V and μ_W are isomorphisms, and T' is compact if T is compact, we get that T^* is compact once T is compact. □

Proposition 8.8. *Let V and W be two Hilbert spaces and $T \in \mathcal{L}(V,W)$. Then T is compact if and only if there is a sequence T_n of operators with $\dim(\mathrm{im}(T_n)) < \infty$ and $\|T - T_n\| \to 0$.*

Proof. The "if" direction follows from Theorem 8.4 because operators of finite rank are compact (the fact that V and W are Hilbert spaces plays no role).

For the "only if" direction, we shall assume that W is separable (and the general case can be reduced to that). Pick an orthonormal basis $(w_n)_{n\in\mathbb{N}}$, and let $P_n : W \to W$ be the orthogonal projection onto $\mathrm{span}\{w_k | k \leq n\}$. Then P_n has finite rank, and hence so does P_nT. We claim that $P_nT \to T$ in norm. Because $\overline{T(D_1(V))}$ is a compact metric space and because the set $P_n|_{\overline{T(D_1(V))}}$ is uniformly bounded and equicontinuous, and converges pointwise to the identity, we get that $P_n \to \mathrm{id}$ uniformly on $\overline{T(D_1(V))}$ by the Arzela-Ascoli theorem. But this means exactly the same as norm convergence $P_nT \to T$. □

8.3. **The spectral theorem for self-adjoint compact operators on a Hilbert space.** We first recall the following result from Linear Algebra.

Theorem 8.9 (Spectral theorem for self-adjoint operators on finite-dimensional Hilbert spaces)**.** *Let V be a finite-dimensional complex Hilbert space and let $T \in \mathcal{L}(V)$ be selfadjoint operator. Then there exists an orthonormal basis $(v_1, \ldots, v_n)$ of V which consists of eigenvectors, i.e. $Tv_i = \lambda_i v_i$ for $i = 1, \ldots, n$, where $\lambda_i \in \mathbb{R}$.*

The goal of spectral theory of operators in Hilbert spaces is a generalization of this theorem to bounded operators on Hilbert spaces. The operators for which the situation is closest to the finite-dimensional case are the compact self-adjoint operators.

Let us introduce some notations first. Assume that I is a set, and V_i, $i \in I$ is a family of Hilbert spaces. On the (linear-algebraic) direct sum $\bigoplus_{i\in I} V_i$, we introduce the scalar product $\langle (v_i)_i, (w_i)_i \rangle := \sum_{i\in I} \langle v_i, w_i \rangle_{V_i}$. This is a scalar product, and the *Hilbert sum*

$$\bigoplus_{i\in I}^{(2)} V_i$$

of the spaces V_i is the Hilbert space completion of this scalar product.

As a matter of notation, we define

$$\mathrm{Eig}(T,\lambda) := \ker(T - \lambda),$$

the space of all eigenvectors of the linear map T to the eigenvalue λ. As in linear algebra, an *eigenvalue* of T is a $\lambda \in \mathbb{K}$ with $\mathrm{Eig}(T,\lambda) \neq \{0\}$.

Theorem 8.10 (Spectral theorem for selfadjoint compact operators)**.** *Let T be a compact self-adjoint operator on a Hilbert space V. Then*

1. *all eigenvalues of T are real,*
2. *the eigenspaces $\mathrm{Eig}(T,\lambda)$ and $\mathrm{Eig}(T,\mu)$ are orthogonal if $\mu \neq \lambda$,*
3. *for each $\epsilon > 0$, the direct sum $\bigoplus_{|\lambda|\geq\epsilon}^{(2)} \mathrm{Eig}(T,\lambda)$ is finite-dimensional,*
4. *V has a complete orthonormal system which consists of eigenvectors of T,*

This contains Theorem 8.9 as a special case. Usually, Theorem 8.9 is proven using determinants (characteristic polynomials). This tool is not available for infinite-dimensional Hilbert space, and we have to find an alternative proof. In fact, the following proof simplifies in finite-dimensional spaces and gives the simplest possible proof of Theorem 8.9.

Proof of Theorem 8.10 (1), (2) and (3). (1): if $Tv = \lambda v$, then

$$(\lambda - \overline{\lambda})\langle v, v\rangle = \langle v, \lambda v\rangle - \langle \lambda v, v\rangle = \langle v, Tv\rangle - \langle Tv, v\rangle = \langle v, Tv\rangle - \langle v, Tv\rangle = 0$$

proves $\lambda \in \mathbb{R}$ or $v = 0$.

(2): assume $Tv = \lambda v$ and $Tw = \mu w$, $\lambda, \mu \in \mathbb{R}$. Then

$$(\lambda - \mu)\langle v, w\rangle = \langle \lambda v, w\rangle - \langle v, \mu w\rangle = \langle Tv, w\rangle - \langle v, Tw\rangle = 0$$

proves $\lambda = \mu$ or $\langle v, w\rangle = 0$.

(3): a vector $v \in \bigoplus^{(2)}_{|\lambda|\geq\epsilon}$ can be written as $v = \sum_{|\lambda|\geq\epsilon} v_\lambda$, where $Tv_\lambda = \lambda v_\lambda$ (only countably many of these vectors can be nonzero). Then by the Pythagoras identity (all the vectors Tv_λ are orthogonal by (1))

$$\|Tv\|^2 = \sum_\lambda \|\lambda v_\lambda\|^2 = \sum_\lambda |\lambda|^2 \|v_\lambda\|^2 \geq \epsilon^2 \sum_\lambda \|v_\lambda\|^2 = \epsilon^2 \|v\|^2.$$

Lemma 8.3 shows that $\bigoplus^{(2)}_{|\lambda|\geq\epsilon}$ is finite-dimensional. □

Item (4) of Theorem 8.10 is a bit harder. The key step is the following lemma.

Lemma 8.11. *Let $T \in \mathcal{K}(V)$ be a compact self-adjoint operator on a Hilbert space. Then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. There is nothing to prove if $T = 0$, so we may assume

$$C := \|T\| > 0.$$

By the C^*-identity (Theorem 2.27), we have

$$C^2 = \|T\|^2 = \|T^*T\| = \|T^2\|.$$

Let

$$S(V) := \{v \in V | \|v\| = 1\} \subset V$$

be the unit sphere in V. By the definition of the operator norm, there is a sequence $v_n \in S(V)$ such that

$$\lim_n \|T^2 v_n\| = \|T^2\| = C^2.$$

Since T is compact, there is a subsequence v_{n_k} such that

$$u := \lim_k T v_{n_k} \in V$$

exists. Then

$$\|u\| \leq \|T\| = C$$

and

$$T^2 v_{n_k} \to Tu,$$

hence

$$\|Tu\| = C^2.$$

It follows that

$$C^2 = \|Tu\| \leq \|T\|\|u\| \leq C^2,$$

so that

$$\|u\| = C \tag{8.12}$$

and

$$C^4 = \|Tu\|^2 = \langle Tu, Tu\rangle = \langle TTu, u\rangle \le \|TTu\|\|u\| \le \|T\|\|Tu\|\|u\| = C^4. \tag{8.13}$$

It follows that

$$\langle T^2 u, u\rangle = \|T^2 u\|\|u\|. \tag{8.14}$$

Hence by the equality case of the Cauchy-Schwarz inequality (Theorem 2.3 (2)), we must have that T^2u is a multiple of u. In other words, $T^2u = \mu u$ for some $\mu \in \mathbb{C}$. But as

$$C^4 = \langle u, T^2 u\rangle = \langle u, u\rangle \mu = C^2 \mu,$$

we have $\mu = C^2$, so that

$$T^2 u = C^2 u,$$

and that C^2 is an eigenvalue of T^2. The operator T restricts to a linear map $T_0 : \mathrm{Eig}(T, C^2) \to \mathrm{Eig}(T, C^2)$, which satisfies

$$(T_0 - C)(T_0 + C) = T^2 - C^2 = 0.$$

Therefore at least one of $T_0 - C$ and $T_0 + C$ must have a nonzero kernel, and this implies that C or $-C$ is an eigenvalue of T_0, hence of T. $\square$

End of the proof of Theorem 8.10. Let $W \subset V$ be the closed subspace spanned by all eigenvectors of T. Then $T(W) \subset W$. If $v \in W^\perp$, then $Tv \in W^\perp$, because for each $w \in W$, we have $Tw \in W$ and hence

$$\langle Tv, w\rangle = \langle v, Tw\rangle = 0.$$

The operator $T|_{W^\perp}$ therefore maps $W^\perp$ into itself, and when viewed as an operator $W^\perp \to W^\perp$, $T|_{W^\perp}$ is self-adjoint. By Lemma 8.11, $W^\perp$ contains an eigenvector of T unless $W^\perp = \{0\}$. The first is impossible by construction. So $W^\perp = 0$, hence $W = V$. This argument shows that the eigenvectors span all of V. Picking a complete orthonormal system for each eigenspace provides a complete orthonormal system consisting of eigenvectors. $\square$

8.4. $*$-Fredholm operators.

Definition 8.15. *Let V, W be $\mathbb{k}$-vector spaces and let $F : V \to W$ be linear. F is* Fredholm *if the kernel* $\ker(F)$ *and the* cokernel $W/\mathrm{im}(F)$ *are both finite dimensional. The* index *of a Fredholm map is the integer*

$$\mathrm{ind}(F) = \dim(\ker(F)) - \dim(W/\mathrm{im}(F)) \in \mathbb{Z}.$$

If V and W are Banach spaces and F is continuous and Fredholm, then F is a Fredholm operator*.*

The first thing we prove has nothing to do with norms and continuity:

Proposition 8.16. *Let $U \xrightarrow{G} V \xrightarrow{F} W$ be two linear maps. If two of the three linear maps F, G and FG are Fredholm, then so is the third, and we have*

$$\mathrm{ind}(FG) = \mathrm{ind}(F) + \mathrm{ind}(G).$$

Proof. Recall that a sequence

$$U_0 \xrightarrow{f_1} U_1 \xrightarrow{f_2} U_2$$

of vector spaces is *exact* at U_1 if $\mathrm{im}(f_1) = \ker(f_2)$. If U_2 and U_0 are finite-dimensional, then U_1 is finite-dimensional. In the situation of the Proposition, there is a sequence

$$0 \to \ker(G) \xrightarrow{i} \ker(FG) \xrightarrow{G} \ker(F) \xrightarrow{\partial} V/\mathrm{im}(G) \xrightarrow{f} W/\mathrm{im}(FG) \xrightarrow{p} W/\mathrm{im}(F) \to 0$$

which is everywhere exact. The map i is the inclusion, ∂ sends $v \in \ker(F)$ to $v+\mathrm{im}(G)$, f sends $v+\mathrm{im}(G)$ to $Fv+\mathrm{im}(FG)$, and p sends $w+\mathrm{im}(FG)$ to $w+\mathrm{im}(F)$. It is left to the reader to prove that these maps are indeed well-defined and that the sequence is exact.

Together with the fact about finite-dimensionality of vector spaces in an exact sequence, this proves that if two of F, G and FG are exact, then so is the third.

To prove the formula for the index, it is enough to prove that when

$$0 \xrightarrow{f_0} U_0 \xrightarrow{f_1} U_1 \xrightarrow{f_2} U_2 \to \ldots \xrightarrow{f_n} U_n \xrightarrow{f_{n+1}} 0$$

is an exact sequence of finite-dimensional spaces, then

$$\sum_{j=0}^{n} (-1)^j \dim(U_j) = 0.$$

To this end, let $d_j := \dim(U_j)$, $r_j := \dim(\mathrm{im}(f_j))$ and $k_j := \dim(\ker(f_j))$. The rank-nullity theorem shows

$$d_j = r_{j+1} + k_{j+1},$$

and the exactness of the sequence implies

$$r_j = k_{j+1}.$$

Therefore

$$\sum_{j=0}^{n} (-1)^j d_j = \sum_{j=0}^{n} (-1)^j (r_{j+1} + k_{j+1}) = \sum_{j=0}^{n} (-1)^j (r_{j+1} + r_j) =$$

$$= \sum_{j=0}^{n} (-1)^j r_j + \sum_{j=1}^{n+1} (-1)^{j-1} r_j = r_0 + (-1)^n r_{n+1} = 0 + 0 = 0.$$

□

Proposition 8.17. *Let $F : V \to W$ be a Fredholm operator between Banach spaces. Then $\mathrm{im}(F) \subset W$ is closed.*

Proof. Choose a linear complement $Z \subset W$ of $\mathrm{im}(F)$; Z is finite-dimensional, equip Z with a norm and denote the inclusion by $I : Z \to W$. Then $V \oplus Z \xrightarrow{F+I} W$ is surjective and continuous, and hence an open map by the open mapping theorem. Since $V \times (Z \setminus 0) \subset V \oplus Z$ is open, $(F + I)(V \times (Z \setminus 0)) \subset W$ is open. Therefore, its complement $W \setminus (F + I)(V \times (Z \setminus 0)) = F(V)$ is closed. □

From now on, we turn our attention to Hilbert spaces (many of the following results remain true for Banach spaces). We denote by $\mathrm{Fred}(V, W) \subset \mathcal{L}(V, W)$ the set of all Fredholm operators (it is not a linear subspace).

Proposition 8.18. *The set* $\mathrm{Fred}(V,W) \subset \mathcal{L}(V,W)$ *is open, and the map*

$$\mathrm{ind} : \mathrm{Fred}(V,W) \to \mathbb{Z}$$

is locally constant.

Proof. Let $T : V \to W$ be Fredholm. Since $\mathrm{im}(T) \subset W$ is closed, there is the orthogonal projection map $P : W \to \mathrm{im}(T)$. Since $\ker(P) = \mathrm{im}(T)^\perp \cong W/\mathrm{im}(T)$ and P is surjective, P is Fredholm with index $\dim(W/\mathrm{im}(T))$

The inclusion $J : \ker(T)^\perp \to V$ is injective, and $V/\mathrm{im}(J) = V/\ker(T)^\perp \cong \ker(T)$, so J is Fredholm with index $-\dim(\ker(T)$.

The composition $\ker(T)^\perp \xrightarrow{J} V \xrightarrow{T} W \xrightarrow{P} \mathrm{im}(T)$ is invertible. Hence there is $\delta > 0$ such that when $\|S - T\| \leq \delta$, then PSJ is also invertible. Since P and PSJ are Fredholm, so is SJ, and since J is Fredholm, S is Fredholm. Therefore $\mathrm{Fred}(V,W)$ is open.

Moreover, for S as above

$$0 = \mathrm{ind}(PSJ) = \mathrm{ind}(P)+\mathrm{ind}(S)+\mathrm{ind}(J) = \dim(W/\mathrm{im}(T))+\mathrm{ind}(S)-\dim(\ker(T)) = \mathrm{ind}(S)-\mathrm{ind}(T).$$

□

Theorem 8.19 (Atkinson)**.** *For a bounded operator* $T : V \to W$ *between Hilbert spaces, the following are equivalent.*

(1) *T is Fredholm.*
(2) *There is $S \in \mathcal{L}(W,V)$ such that $1 - TS$ and $1 - ST$ are compact.*

An operator S as in the Theorem is a *parametrix* for T.

Proof of $1 \Rightarrow 2$. Assume that T is Fredholm. Then $\mathrm{im}(T)$ is closed by Proposition 8.17. Write $V = \ker(T)^\perp \oplus \ker(T)$ and $W = \mathrm{im}(T) \oplus \mathrm{im}(T)^\perp$. With respect to this decomposition, we can write

$$T = \begin{pmatrix} T_0 & 0 \\ 0 & 0 \end{pmatrix}$$

where $T_0 : \ker(T)^\perp \to \mathrm{im}(T)$ is the restriction of T and bijective. Let

$$S = \begin{pmatrix} T_0^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$1 - TS = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

where the nonzero entry is the identity on the finite-dimensional space $\mathrm{im}(T)^\perp$ and hence compact. Therefore $1 - TS$ is compact, and similarly, $1 - ST$ is compact. □

For the reverse implication, we isolate the main step as a separate lemma.

Lemma 8.20. *Let $T : V \to W$ be a bounded operator between Hilbert spaces, let $K : V \to U$ be a compact operator to a further Hilbert space and assume that*

$$\|v\| \leq C(\|Tv\| + \|Kv\|)$$

for some $C \geq 0$ and all $v \in V$. Then $\ker(T)$ is finite-dimensional and $\mathrm{im}(T)$ is closed.

Proof. Let $(v_n)_n$ be a sequence in $\ker(T)$ with $\|v_n\| \leq 1$. We claim that it has a convergent subsequence; this will show that $\dim(\ker(T)) < \infty$. After passing to a subsequence, we may assume that Kv_n is convergent. But then $\|v_n - v_m\| \leq C\|Kv_n - Kv_m\|$ proves that $(v_n)_n$ is Cauchy.

To prove that $\operatorname{im}(T)$ is closed, note that $\operatorname{im}(T) = \operatorname{im}(T|_{\ker(T)^\perp})$. In other words, we can assume without loss of generality that T is injective, and we do so from now on.

We show that there is $c > 0$ so that $\|Tv\| \geq c\|v\|$ for all $v \in V$; Lemma 1.18 implies that $\operatorname{im}(T)$ is closed.

If there is no such c, we find for each n a unit vector $v_n \in V$ with $\|Tv_n\| \leq \frac{1}{n}$. We therefore found a sequence of unit vectors v_n with $\|Tv_n\| \to 0$, and after passage to a subsequence we may assume that Kv_n is convergent. It follows from the estimate $\|v\| \leq C(\|Tv\| + \|Kv\|)$ that v_n is Cauchy and hence tends to some $v \in V$, with $\|v\| = 1$. But we also have $Tv = \lim_n Tv_n = 0$, which contradicts the injectivity of T. $\square$

Proof of $2 \Rightarrow 1$ *of Theorem 8.19.* If $1 - ST$ is compact, we get

$$\|v\| \leq \|STv\| + \|(1 - ST)v\| \leq \|S\|\|T\| + \|(1 - ST)v\|.$$

Lemma 8.20 implies that $\ker(T)$ is finite-dimensional and $\operatorname{im}(T)$ is closed. But $\operatorname{im}(T)^\perp = \ker(T^*)$, and since

$$1 - S^*T^* = (1 - TS)^*$$

is compact, we can apply the same argument and conclude that $\ker(T^*)$ is finite-dimensional. Therefore, $\operatorname{im}(T)^\perp$ is finite-dimensional, and because $\operatorname{im}(T)$ is closed, we get that $W/\operatorname{im}(T) \cong \operatorname{im}(T)^\perp$ is finite-dimensional as well. $\square$

Corollary 8.21. *Let* $T \in \operatorname{Fred}(V, W)$ *and let* $K \in \mathcal{K}(V, W)$. *Then* $T + K$ *is Fredholm, and* $\operatorname{ind}(T + K) = \operatorname{ind}(T)$.

Proof. Let $S \in \operatorname{Fred}(W, V)$ be a parametrix for T. Then

$$1 - S(T + K) = 1 - ST - SK$$

and

$$1 - (T + K)S = 1 - TS - KS$$

are compact, and so by Theorem 8.19, $T + K$ is compact. This argument shows that $T + tK$ is compact for each $t \in [0, 1]$, and because the index is locally constant, $\operatorname{ind}(T + K) = \int(T)$. $\square$

Corollary 8.22. *Let* $T \in \operatorname{Fred}(V, W)$. *Then the adjoint* $T^* \in \mathcal{L}(W, V)$ *is Fredholm, and* $\operatorname{ind}(T) = -\operatorname{ind}(T^*)$.

Proof. Let S be a parametrix for T. Then S^* is a parametrix for T^*, and so T^* is Fredholm. Observe that

$$\ker(T^*) = \operatorname{im}(T)^\perp \cong W/\operatorname{im}(T)$$

(the last is true because $\operatorname{im}(T) \subset W$ is closed). The same applies to T^* because this is also Fredholm, and hence

$$\ker(T) = \operatorname{im}(T^*)^\perp \cong V/\operatorname{im}(T^*).$$

It follows that

$$\operatorname{ind}(T) = \dim(\ker(T)) - \dim(W/\operatorname{im}(T)) = \dim(V/\operatorname{im}(T^*)) - \dim(\ker(T^*)) = -\operatorname{ind}(T^*).$$

□

8.5. ∗ **The Toeplitz index theorem.** We discuss an interesting (and fairly deep) connection between the index of a Fredholm operator and algebraic topology. This is the first instance of an *index theorem.*

Consider the Hilbert space $L^2(S^1;\mathbb{C})$. It contains the space

$$H(S^1) := \overline{\operatorname{span}\{z^n | n \geq 0\}}$$

of functions all whose Fourier coefficients of negative index are zero. Let $P : L^2(S^1) \to H(S^1) \subset L^2(S^1)$ be the orthogonal projection onto $H(S^1)$. Furthermore, let $f : S^1 \to \mathbb{C}$ be a continuous function. This gives rise to a multiplication operator $M_f(h) := fh$ on $L^2(S^1)$. Note that

$$\|M_f\| = \|f\|_{C^0},$$
$$M_{fg} = M_f M_g,$$
$$M_{f+g} = M_f + M_g,$$

and $M_1 = 1$, $M_{\overline{f}} = (M_f)^*$.

We define an operator

$$T_f : L^2(S^1) \to L^2(S^1),\ T_f = PM_fP + (1 - P),$$

which is called the *Toeplitz operator* associated with f.

Now recall from elementary algebraic topology the notion of the *winding number*: for a continuous function $f : S^1 \to \mathbb{C}^\times$, it associates the winding number $\deg(f) \in$ expand $\mathbb{Z}$.

Theorem 8.23 (Toeplitz index theorem)**.** *Let $f : S^1 \to \mathbb{C}^\times$ be continuous. Then T_f is a Fredholm operator, and $\operatorname{ind}(T_f) = -\deg(f)$.*

Lemma 8.24. *Let $f \in C(S^1)$. Then the commutator $[P, M_f]$ is compact.*

Proof. Let $A \subset C(S^1)$ be the subspace of those f for which $[P, M_f]$ is compact. We verify the following claims:

(1) A is a linear subspace (clear).
(2) $1 \in A$ (clear, $TM_1 - M_1T = 0$).
(3) $f, g \in A$, then $fg \in A$: for this, note that

$$[P, M_{fg}] = PM_{fg} - M_{fg}P = PM_fM_g - M_fM_gP = [P, M_f]M_g + M_f[P, M_g],$$

which is compact if $f, g \in A$.

(4) If $f \in A$, then $\overline{f} \in A$. This is because

$$[P, M_{\overline{f}}] = [P^*, (M_f)^*] = P^*(M_f)^* - (M_f)^*P^* = (M_fP - PM_f)^* = [M_f, P]^* = -[P, M_f]^*$$

is compact.

(5) $z \in A$. For that, we need to compute $[P, M_z]$ explicitly on the orthonormal basis $\{z^k\}$. If $k \geq 0$, we have

$$PM_z(z^k) - M_zP(z^k) = P(z^{k+1}) - zz^k = 0,$$

for $k = -1$, we have

$$PM_z(z^{-1}) - M_zP(z^{-1}) = Pz^0 - 0 = z^0,$$

and for $k \leq -2$, we have $[P, M_z]z^k = 0$. Therefore, $[P, M_z]$ is of rank 1 and therefore compact.

The Stone-Weierstrass theorem easily implies that $A = C(S^1)$, which is what we claimed. $\square$

Proof of the Toeplitz index theorem. Let $f, g \in C(S^1)$. Then (because $M_{fg} = M_f M_g$)

$$T_f T_g - T_{fg} = (PM_fP + (1-P))(PM_gP + (1-P)) - (PM_fM_gP + (1-P)) =$$

$$= PM_fPPM_gP + PM_fP(1-P) + (1-P)PM_gP + (1-P)^2 - PM_fM_gP - (1-P) =$$

(use that $P^2 = P$, $P(1-P) = (1-P)P = 0$ and $(1-P)^2 = 1-P$)

$$= PM_fPM_gP - PM_fM_gP = P[M_f, P]M_gP + PPM_fM_gP - PM_fM_gP =$$
$$= P[M_f, P]M_gP.$$

This is compact by the previous lemma.

If $f : S^1 \to \mathbb{C}^\times$ is continuous, we apply this argument to f and $g = f^{-1}$ and obtain that (note that $T_1 = \mathrm{id}$)

$$1 - T_f T_{f^{-1}},\ 1 - T_f T_{f^{-1}}$$

are compact. Hence $T_{f^{-1}}$ is a parametrix for T_f, whence T_f is Fredholm. If $f, g : S^1 \to \mathbb{C}^\times$ are continuous, the above argument also proves

$$\mathrm{ind}(T_{fg}) = \mathrm{ind}(T_f) + \mathrm{ind}(T_g).$$

As $\|T_f\| \leq \|f\|$, $f \mapsto T_f$ is continuous, and therefore $\mathrm{ind}(T_f)$ only depends on the homotopy class of f in $[S^1; \mathbb{C}^\times]$.

One of the key results of elementary algebraic topology is that two maps $S^1 \to \mathbb{C}^\times$ are homotopic if and only if they have the same winding number. Therefore, $\mathrm{ind}(T_f) = \mathrm{ind}(T_{z^k})$, where $k = \deg(f)$. But

$$\mathrm{ind}(T_{z^k}) = k\mathrm{ind}(T_z),$$

and all that remains is to compute $\mathrm{ind}(T_z)$. But

$$T_z(z^k) = \begin{cases} z^k & k \leq -1 \\ z^{k+1} & k \geq 0. \end{cases}$$

Therefore T_z is injective, and its image is the closed subspace $\overline{\mathrm{span}\{z^n | n \neq 0\}}$ which has codimension 1. Thus

$$\mathrm{ind}(T_z) = -1 = -\deg(z).$$

$\square$

9. The spectrum

9.1. Banach algebras.

Definition 9.1. *Let $\Bbbk$ be a field. A* $\Bbbk$-algebra *is a $\Bbbk$-vector space A, together with a $\Bbbk$-bilinear map*

$$A \times A \to A,\ (a, b) \mapsto ab,$$

such that this multiplication is associative, that is

$$(ab)c = a(bc).$$

We say that A is commutative *if $ab = ba$ for all $a, b \in A$. If there is a neutral element $e \in A$ (i.e. $ea = ae = a$ for all $a \in A$) and $e \neq 0$, we say that A is* unital.

If 0 is a neutral element, then $A = \{0\}$, so the condition $e \neq 0$ only excludes the trivial algebra $A = 0$. The neutral element e is uniquely determined, because if e, e' are two such neutral elements, then $e = ee' = e'$. It is customary to denote $1 := e$.

A *homomorphism* $\phi : A \to B$ of two algebras is a linear map such that $\phi(ab) = \phi(a)\phi(b)$ for all $a, b \in A$. If A and B are unital, a homomorphism $\phi : A \to B$ is *unital* if $\phi(1) = 1$. If A is unital, then $\phi : \Bbbk \to A$, $z \mapsto z1$ is an injective unital homomorphism. It is customary to write z instead of $z1$.

A *left-inverse* of $a \in A$ is a c with $ca = 1$, and a *right-inverse* is b with $ab = 1$. An element $a \in A$ may have a left (right) inverse but not a right (left) inverse, and many different left (right) inverses.

An element a of a unital $\Bbbk$-algebra is a *unit* if it has both, a left- and a right inverse. In that situation, left- and right inverses agree and are uniquely determined: if $c'a = ca = 1$ and $ab' = ab = 1$, then

$$c = cab = b = c'ab = c' = c'ab' = b'.$$

It is customary to denote the left-and-right-inverse of a by $a{-}1$.

The set $A^\times \subset A$ of units is a group under multiplication.

Definition 9.2. *A* Banach algebra *is a $\mathbb{K}$-algebra A, together with a norm $\|\_\|$ on A, such that*

$$\|ab\| \leq \|a\|\|b\|$$

for all a, b, and such that the normed vector space $(A, \|\_\|)$ is complete. A unital Banach algebra *is a Banach algebra, such that the naked $\mathbb{K}$-algebra A is unital and $\|1\| = 1$.*

Example 9.3. *If $X \neq \emptyset$ is a topological space, then $C_b(X)$ is a commutative unital $\mathbb{K}$-algebra by pointwise multiplication of functions. The usual norm makes is into a unital Banach algebra. An element $f \in C_b(X)$ is a unit if and only if $0 \notin \overline{f(X)}$.*

If X is locally compact Hausdorff, then $C_0(X)$ is a Banach algebra which is not unital (unless X is compact).

Example 9.4. *For each Banach space V, the space $\mathcal{L}(V) := \mathcal{L}(V, V)$ of bounded operators becomes a $\mathbb{K}$-algebra, with multiplication $(F, G) \mapsto FG$. The operator norm turns $\mathcal{L}(V, V)$ into a Banach algebra, which is unital except in the trivial case $V = 0$. $F \in \mathcal{L}(V)$ is a unit if and only if F is bijective, by the open mapping theorem.*

Example 9.5. *Let (X, μ) be a measure space. Then $L^\infty(X, \mu)$, with the L^∞-norm, is a Banach algebra, which is unital. An element $f \in L^\infty(X, \mu)$ is a unit if and only if there is $\delta > 0$ such that $\mu(f^{-1}(B_\delta(0))) = 0$.*

Example 9.6. *Let X be a measur*able *space (C.2). Then $\mathscr{L}^\infty(X)$, the space of all measurable bounded functions $f : X \to \mathbb{K}$, is a Banach algebra with the supremum norm $\|f\|_{L^\infty} := \sup_{x\in X} |f(x)|$. $f \in \mathscr{L}^\infty(X)$ is a unit if and only if $0 \notin \overline{f(X)}$.*

If X carries a measure μ, the quotient map $\mathscr{L}^\infty(X) \to L^\infty(X, \mu)$ which assigns to a measurable function its equivalence class, is a continuous surjective unital homomorphism. It is usually not injective. It is important to distinguish the Banach algebras $\mathscr{L}^\infty(X)$ and $L^\infty(X, \mu)$.

Some of these algebras have even more structure.

Definition 9.7. *A C^*-*algebra *A is a complex Banach algebra A, together with a $\mathbb{R}$-linear map $A \to A$, $a \mapsto a^*$, such that*

1. $(za)^* = \overline{z}a^*$
2. $(ab)^* = b^*a^*$,
3. $(a^*)^* = a$ *and*
4. $\|a\|^2 = \|a^*a\|$

for all $a, b \in A$ and $z \in \mathbb{C}$.

Examples 9.8. *For a Hilbert space V, $\mathcal{L}(V)$ with $T \mapsto T^*$ is a C^*-algebra. $C_b(X)$, $L^\infty(X, \mu)$ and $\mathscr{L}^\infty(X)$ are C^*-algebras with $f^*(x) := \overline{f(x)}$. All these C^*-algebras are unital, except in trivial cases.*

Since $\|a\|^2 = \|a^*a\| \le \|a^*\|\|a\|$, we have $\|a\| \le \|a^*\|$ for all a, and hence $\|a^*\| \le \|(a^*)^*\| = \|a\|$, so that $\|a^*\| = \|a\|$. Furthermore, if A is unital and $a \in A$, we have $1^*a = 1^*(a^*)^* = (a^*1)^* = (a^*)^* = a$ and $a1^* = (a^*)^*1^* = (1a^*)^* = a$, so that $1^* = 1$.

Definition 9.9. *Let A be a unital C^*-algebra and $a \in A$. We say that a is* self-adjoint *if $a^* = a$, that a is* unitary *if $a^*a = aa^* = 1$, and that a is* normal *if $a^*a = aa^*$. Selfadjoint and unitary elements are normal.*

Theorem 9.10. *Let A be a unital Banach algebra. Then the set $A^\times \subset A$ of invertible elements is open, and the inversion map $\iota : A^\times \to A^\times$ is continuous.*

The first step of the proof is a lemma that is very often used.

Lemma 9.11. *Let A be a unital Banach algebra and let $a \in A$ with $\|a\| < 1$. Then $1 + a$ is invertible in A, and*

$$\|(1 + a)^{-1}\| \le \frac{1}{1 - \|a\|}.$$

Proof. For $a \in A$, $\|a\| < 1$, the series

$$\sum_{n=0}^{\infty} (-1)^n a^n$$

is absolutely convergent in A. Let $b \in A$ be the limit. The same computation that computes the partial sum of a geometric series proves that

$$b(1 + a) = (1 + a)b = 1.$$

Therefore $1 + a \in A^\times$. We also get the estimate
$$\|(1+a)^{-1}\| \leq \sum_{n=0} \|a\|^n = \frac{1}{1-\|a\|}.$$
□

Proof of Theorem 9.10. Let $a \in A^\times$. We show that $B_{\frac{1}{\|a^{-1}\|}}(a) \subset A^\times$, and this implies that $A^\times$ is open. So let $h \in A$ with $\|h\| < \frac{1}{\|a^{-1}\|}$. Then $\|ha^{-1}\| < 1$, and $1 + ha^{-1} \in A^\times$ by Lemma 9.11. Since $A^\times$ is a group, we get
$$a + h = (1 + ha^{-1})a \in A^\times$$
and
$$(a + h)^{-1} = a^{-1}(1 + ha^{-1})^{-1}.$$
Therefore
$$(a + h)^{-1} - a^{-1} = a^{-1}((1 + ha^{-1})^{-1} - 1) = a^{-1} \sum_{k=1}^{\infty} (ha^{-1})^k$$
and
$$\|(a + h)^{-1} - a^{-1}\| \leq \|a^{-1}\| \sum_{k=1}^{\infty} \|ha^{-1}\|^k = \|a^{-1}\|\|ha^{-1}\| \frac{1}{1 - \|ha^{-1}\|}$$
The right hand side tends to 0 as $h \to 0$, proving that the inversion map is continuous at a. □

9.2. **The spectrum of an element in a Banach algebra.**

Definition 9.12. *Let A be a unital Banach algebra over $\mathbb{C}$ and $a \in A$. The* spectrum *of a is the set* $\operatorname{spec}(a) \subset \mathbb{C}$ *which consists of all $z \in \mathbb{C}$ such that $z - a$ is* not *invertible. The complement* $\operatorname{spec}(a)^c \subset \mathbb{C}$ *is the* resolvent set.

Lemma 9.13. *Let $a \in A$. Then* $\operatorname{spec}(a) \subset \mathbb{C}$ *is closed and contained in the closed disc $D_{\|a\|}(\mathbb{C})$ of radius $\|a\|$.*

Proof. The map $J : \mathbb{C} \to A$, $z \mapsto z - a$ is continuous, and hence $\operatorname{spec}(a)^c = J^{-1}(A^\times)$ is open, by Theorem 9.10. If $|z| > \|a\|$, then $z - a = z(1 - \frac{a}{z})$ is invertible as $\|\frac{a}{z}\| < 1$, by Lemma 9.11. □

Example 9.14. *If $a \in \operatorname{Mat}_{n,n}(\mathbb{C})$, then* $\operatorname{spec}(a)$ *is the set of all eigenvalues of a.*

Example 9.15. *If $f \in C(X)$, X compact,* $\operatorname{spec}(f) = f(X) \subset \mathbb{C}$.

It is not true that for $F \in \mathcal{L}(V)$, any point in $\operatorname{spec}(F)$ is an eigenvector of F (there will be an ample supply of examples later on).

Lemma 9.16. *Let A be a unital C^*-algebra and let $a \in A$.*
(1) *if a is self-adjoint, then* $\operatorname{spec}(a) \subset \mathbb{R}$.
(2) *if a is unitary, then* $\operatorname{spec}(a) \subset \mathbb{C}$.

Proof. (1): Let $\lambda \in \operatorname{spec}(a)$ and $t \in \mathbb{R}$. Then $\lambda + it \in \operatorname{spec}(a + it)$. Therefore $|\lambda + it| \leq \rho(a + it) \leq \|a + it\|$. Therefore
$$|\lambda + it|^2 \leq \|a + it\|^2 = \|(a + it)^*(a + it)\| = \|(a - it)(a + it)\| = \|a^2 + t^2\| \leq \|a^2\| + t^2.$$
On the other hand
$$|\lambda + it|^2 = |\lambda|^2 + t^2 + 2t\Im(\lambda).$$

Together, this shows that
$$|\lambda|^2 + 2t\Im(\lambda) \leq \|a^2\|$$
for all $t \in \mathbb{R}$. Such an inequality is impossible unless $\Im(\lambda) = 0$.

(2): Since $\|a\|^2 = \|a^*a\| = \|1\| = 1$, we have $\|a\| = 1$ and hence $\rho(a) \leq \|a\| = 1$, or $\operatorname{spec}(a) \subset D_1(\mathbb{C})$. On the other hand, let $|z| < 1$. Since $a \in A^\times$ and $a^{-1} = a^*$ has norm $\|a^{-1}\| = \|a^*\| = \|a\| = 1$, we have $\|za^{-1}\| < 1$, and so $(1 - za^{-1})$ is invertible, and hence so is
$$a - z = (1 - za^{-1})a,$$
in other words $z \notin \operatorname{spec}(a)$. □

9.3. **The spectrum is nonempty.** For the deeper investigations of the spectrum, we henceforth assume that $\mathbb{K} = \mathbb{C}$. The first thing we have to consider are the two *spectral mapping theorems.*

Theorem 9.17 (First spectral mapping theorem)**.** *Let $\Phi : A \to B$ be a homomorphism of unital Banach algebras and let $a \in A$. Then*
$$\operatorname{spec}_B(\Phi(a)) \subset \operatorname{spec}_A(a).$$

Proof. Let $z \in \operatorname{spec}_A(a)^c$, that is, $(z - a) \in A^\times$. Since Φ is multiplicative and $\Phi(1) = 1$, we get $\Phi(z - a) = z - \Phi(a) \in B^\times$, hence $z \in \operatorname{spec}_B(\Phi(a))^c$. Hence $\operatorname{spec}_A(a)^c \subset \operatorname{spec}_B(\Phi(a))^c$ and so $\operatorname{spec}_B(\Phi(a)) \subset \operatorname{spec}_A(a)$, as claimed. □

The second spectral mapping theorem is about inserting elements of a Banach algebra into polynomials. Let $p(x) = \sum_{k=0}^n c_k x^k \in \mathbb{C}[x]$ be a polynomial, and let $a \in A$. We define
$$p(a) := \sum_{k=0}^n c_k a^k \in A.$$
It is easy to check that
$$p \mapsto p(a)$$
is a homomorphism
$$\mathbb{C}[x] \to A$$
of unital algebras. Since we do not have a sensible norm on $\mathbb{C}[x]$, it does not make sense to say that it is continuous, bounded or something like that.

Theorem 9.18 (Second spectral mapping theorem)**.** *Let $p \in \mathbb{C}[x]$ and $a \in A$. Then*
$$\operatorname{spec}_A(p(a)) = p(\operatorname{spec}_A(a)) := \{p(z) | z \in \operatorname{spec}_A(a)\}.$$

Proof. Let $y \in \mathbb{C}$. Write
$$q(x) := p(x) - y \in \mathbb{C}[x],$$
and by the fundamental theorem of algebra, we can write
$$q(x) = \prod_{j=1}^n (x - z_j).$$
It follows that
$$p(a) - y = \prod_{j=1}^n (a - z_j).$$
The following statements are equivalent:

(1) $y \in p(\operatorname{spec}_A(a))$,

(2) there is $z \in \operatorname{spec}_A(a)$ with $p(z) = y$,
(3) at least one of the roots z_j of $q(x)$ lies in $\operatorname{spec}_A(a)$,
(4) at least one of the $(a - z_j)$ is not invertible in A,
(5) the product $\prod_{j=1}^{n}(a - z_j)$ is not invertible in A,
(6) $p(a) - y$ is not invertible in A,
(7) $y \in \operatorname{spec}(p(a))$.

The only step that needs to be justified is the equivalence $4 \Rightarrow 5$. What is relevant here is that the elements $(a - z_j)$ *commute* with each other; it follows that if the product is invertible, all of its factors are (exercise). □

We wish to prove that the spectrum of any element in a complex Banach algebra is nonempty. The actual result is a little more precise, and we need a definition.

Definition 9.19. *Let $a \in A$ be an element of a unital Banach algebra. The* spectral radius *of a is the number*

$$\rho(a) := \sup\{|z| | z \in \operatorname{spec}(a)\} \in [0, \infty).$$

The definition of the supremum assumes that $\operatorname{spec}(a) \neq \emptyset$, which we will show soon. We already know that the spectrum is contained in $D_{\|a\|}(\mathbb{C})$, from which we get

$$\rho(a) \leq \|a\|. \tag{9.20}$$

Our goal is to prove the following result.

Theorem 9.21. *Let A be a unital Banach algebra over $\mathbb{C}$ and let $a \in A$. Then*

$$\operatorname{spec}_A(a) \neq \emptyset$$

and the spectral radius is given by the formula

$$\rho(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}}.$$

First part of the proof. Let $\lambda \in \operatorname{spec}(a)$. By Theorem 9.18, we have

$$\lambda^n \in \operatorname{spec}(a^n),$$

and (9.20) shows that

$$|\lambda|^n = |\lambda^n| \leq \|a^n\|$$

and hence

$$\rho(a) \leq \|a^n\|^{\frac{1}{n}}$$

and hence

$$\rho(a) \leq \inf_n \|a^n\|^{\frac{1}{n}} \leq \limsup_n \|a^n\|^{\frac{1}{n}}.$$

□

It remains to be shown that $\operatorname{spec}(a) \neq \emptyset$ and that

$$\limsup_n \|a^n\|^{\frac{1}{n}} \leq \rho(a).$$

Let

$$U_a := \{z \in \mathbb{C} | 1 - za \in A^\times\} \subset \mathbb{C}.$$

This is an open subset. Note that

$$U_a = \{z \in \mathbb{C}^\times | \frac{1}{z} \in \operatorname{spec}(a)^c\} \cup \{0\}.$$

By (9.20), we have

$$B_{\frac{1}{\|a\|}}(0) \subset U_a.$$

Consider the *converse resolvent*[4] function

$$T_a : U_a \to A,\ T_a(z) := (1 - za)^{-1}.$$

For $|z| < \frac{1}{\|a\|}$, we have the (absolutely convergent) representation

$$T_a(z) = \sum_{n=0}^{\infty} z^n a^n. \tag{9.22}$$

For $z, z+h \in U_a$, we compute

$$\frac{1}{h}(T_a(z+h) - T_a(z)) = \frac{1}{h}((1-(z+h)a)^{-1} - (1-za)^{-1}) =$$

$$= \frac{1}{h}((1-(z+h)a)^{-1}(1-za)-1)(1-za)^{-1} = \frac{1}{h}(1-(z+h)a)^{-1}\Big((1-za)-(1-(z+h)a)\Big)(1-za)^{-1} =$$

$$= (1-(z+h)a)^{-1}a(1-za)^{-1}.$$

It follows that

$$\lim_{h\to 0}\frac{1}{h}(T_a(z+h) - T_a(z)) = (1-za)^{-1}a(1-za)^{-1}.$$

Now we call the basic results from the theory of holomorphic functions in one complex variable. Recall that a *holomorphic function* is a function $f : U \to \mathbb{C}$, defined on an open subset $U \subset \mathbb{C}$, such that for each $z \in U$, the limit

$$f'(z) := \lim_{h\to 0}\frac{1}{h}(f(z+h) - f(z))$$

exists. The above computation shows that for each $L \in A'$, the function

$$F_{L,a} : U_a \to \mathbb{C},\ F_{L,a}(z) := L(T_a(z))$$

is holomorphic.

Important examples are power series

$$f(z) = \sum_{n=0}^{\infty} c_n z^n.$$

There is a unique $R \in [0, \infty]$, the *radius of convergence*, so that the series converges absolutely on $B_R(0)$ and diverges outside $\bar{B}_R(0)$. One can show that f is holomorphic on $B_R(0)$. Furthermore, if $0 \le r < R$, the sequence

$$n \mapsto |c_n| r^n$$

is bounded (otherwise, we would not have absolute convergence). The fundamental theorem of elementary complex analysis says that all holomorphic functions can be written as power series. Here is the precise statement. We have written the proofs in §D.

---

[4]The *resolvent function* is $R_a : \mathrm{spec}(a)^c \to A$, $R_a := (z-a)^{-1}$. We find it more convenient to use T_a here.

Theorem 9.23. *Let $0 \in U \subset \mathbb{C}$ be open and let $f : U \to \mathbb{C}$ be holomorphic. Then there is a unique power series expansion*

$$f(z) = \sum_{n=0}^{\infty} c_n z^n$$

which converges absolutely on each disc $\bar{B}_r(0) \subset U$. The coefficients c_n are given by the integral formula

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(re^{it})}{r^n e^{int}} dt.$$

Corollary 9.24 (Liouville's theorem)**.** *Let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic and bounded. Then f is constant.*

Proof. Let $M := \sup_z |f(z)|$. By the formula for the coefficients, we get for each r that

$$|c_n| r^n \leq \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it}| dt \leq M.$$

This implies $c_n = 0$ for $n > 0$. The power series expansion (which converges on $\mathbb{C}$) shows that f is constant. $\square$

Second part of the proof of Theorem 9.21. We first prove that $\mathrm{spec}(a) \neq \emptyset$. Suppose, for sake of a contradiction, that $\mathrm{spec}(a) = \emptyset$. Then $a \in A^{\times}$, and $U_a = \mathbb{C}$. Then

$$T_a(z) = (1 - za)^{-1} = \frac{1}{z}(\frac{1}{z} - a)^{-1}$$

and

$$\lim_{z \to \infty} T_a(z) = 0.$$

Hence the function $z \mapsto \|T_a(z)\|$ is bounded. It follows that for each $L \in A'$, the function $F_{L,a} : \mathbb{C} \to \mathbb{C}$ is a bounded holomorphic function. By Liouville's Theorem, it must be constant, and hence equal to 0. Therefore

$$0 = F_{L,a}(0) = L(1)$$

for each $L \in A'$. By the Hahn-Banach theorem, this implies the absurdity $0 = 1$.

For the spectral radius formula, it remains to be shown that

$$\limsup_n \|a^n\|^{\frac{1}{n}} \leq \rho(a).$$

By the definition of U_a and the spectral radius, we have

$$\frac{1}{\rho(a)} = \sup\{r | D_r(\mathbb{C}) \subset \mathbb{C}\}.$$

Therefore, we must prove

$$D_r(\mathbb{C}) \subset U_a \Rightarrow r \limsup_n \|a^n\|^{\frac{1}{n}} \leq 1. \tag{9.25}$$

So let $r > 0$ such that $D_r(\mathbb{C}) \subset U_a$. For $L \in A'$, the power series expansion

$$F_{L,a}(z) = \sum_{n=0}^{\infty} z^n L(a^n)$$

converges around 0. On the other hand, $F_{L,a} : U_a \to \mathbb{C}$ is holomorphic, and by the uniqueness part of Theorem 9.23, it follows that the expansion converges absolutely on $D_r(\mathbb{C})$. Hence there is $M(r, L) \geq 0$ such that

$$r^n |L(a^n)| \leq M(r, L)$$

for all n. Now we consider the family of linear maps

$$G_n : A' \to \mathbb{C},\ G_n(L) := r^n L(a^n).$$

By the Hahn-Banach theorem, the operator norm of G_n is

$$\|G_n\| = r^n \|a^n\|,$$

and for each $L \in A'$, we have $|G_n(L)| \leq M(r, L)$. The principle of uniform boundedness implies that there is $C(r) \geq 0$ such that

$$r^n \|a^n\| = \|G_n\| \leq C(r)$$

for all n. It follows that

$$r \|a^n\|^{\frac{1}{n}} \leq C(r)^{\frac{1}{n}}$$

and hence

$$r \limsup_n \|a^n\|^{\frac{1}{n}} \leq \lim_n C(r)^{\frac{1}{n}}.$$

The latter limit equals 1, and hence we have proven (9.25) for each $r < \frac{1}{\rho(a)}$. The spectral radius formula follows. □

Theorem 9.21 has some interesting consequences for C^*-algebras.

Corollary 9.26. *Let A be a unital C^*-algebra and let $a \in A$ be normal. Then*

$$\rho(a) = \|a\|.$$

Proof. First note that by the C^*-identity

$$\|a^2\|^2 = \|(a^2)^* a^2\| = \|a^* a^* a a\| = \|(a^* a)(a^* a)\| = \|(a^* a)^* (a^* a)\| = \|a^* a\|^2 = \|a\|^4$$

or $\|a^2\| = \|a^2\|$. Because each power a^n is also normal, we obtain inductively that

$$\|a^{2^n}\| = \|a\|^{2^n}$$

for all $n \geq 0$. By the spectral radius formula

$$\rho(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|a^{2^n}\|^{\frac{1}{2^n}} = \|a\|.$$

□

Corollary 9.27. *Let $a \in A$ be an arbitrary element of a unital C^*-algebra. Then*

$$\|a\| = \sqrt{\rho(a^* a)}.$$

Note that the left hand side is entirely determined by the algebraic structure of A (the norm does not enter the definition).

Proof.

$$\|a\| = \sqrt{\|a^* a\|} = \sqrt{\rho(a)}.$$

□

Definition 9.28. *Let A and B be two C^*-algebras. A* $*$-homomorphism *is an algebra homomorphism $\Phi : A \to B$ such that $\Phi(a^*) = \Phi(a)^*$ for all $a \in A$.*

It follows automatically that $\Phi(a)$ is self-adjoint if a is self-adjoint.

Corollary 9.29. *Let A, B be unital Banach algebras and let $\Phi : A \to B$ be a unital $*$-homomorphism. Then Φ is continuous, and $\|\Phi\| = 1$, if one of the following hypotheses holds:*

1. *A, B are C^*-algebras and Φ is a $*$-homomorphism.*
2. *B is a commutative C^*-algebra.*

Proof. (1):

$$\|\Phi(a)\|^2 = \rho(\Phi(a)^*\Phi(a)) = \rho(\Phi(a^*a)) \leq \rho(a^*a) = \|a\|^2;$$

the inequality holds by Theorem 9.17.

(2): because B is commutative, $\Phi(a)$ is normal, and so

$$\|\Phi(a)\| = \rho(\Phi(a)) \leq \rho(a) \leq \|a\|$$

by Theorem 9.17 and the simple inequality $\rho(a) \leq \|a\|$. $\square$

10. The spectral theorem for normal bounded operators on Hilbert spaces

The *spectral theorem* is not a single theorem, but a collection of various related results about normal operators on Hilbert spaces. We take the finite-dimensional case as a model case. Recall from Linear Algebra:

Theorem 10.1. *Let V be a finite-dimensional complex Hilbert space and let $T \in \mathcal{L}(V)$ be normal. Then there is an orthonormal basis $(v_1, \ldots, v_n)$ of eigenvectors, $Tv_j = \lambda_j v_j$.*

In that case $\operatorname{spec}(T) = \{\lambda_1, \ldots, \lambda_n\}$. For each $v \in V$, we have

$$v = \sum_{k=1}^{n} \langle v_k, v \rangle v_k$$

and therefore

$$\|v\|^2 = \sum_{k=1}^{n} |\langle v_k, v \rangle|^2.$$

It follows that

$$\|Tv\|^2 = \sum_{k=1}^{n} |\langle \lambda_k v_k, v \rangle|^2 \le \max_k |\lambda_k|^2 \sum_{k=1}^{n} |\langle v_k, v \rangle|^2 = \max_k |\lambda_k|^2 \|v\|^2$$

and therefore

$$\|T\| = \max_k |\lambda_k|. \tag{10.2}$$

It is also easily seen that

$$T^* v_k = \overline{\lambda_k} v_k.$$

10.1. The continuous functional calculus.

From linear algebra, one remembers that the operation of inserting matrices into polynomials is very important for the understanding of the structure of matrices. For any unital $\mathbb{C}$-algebra and any $a \in A$, there is a unique algebra homomorphism

$$\Phi_a^p : \mathbb{C}[x] \to A$$

with

$$\Phi_a^p(1) = 1, \ \Phi_a^p(x) = a.$$

This is defined by the formula

$$\Phi_a^p\Big(\sum_{k=0}^{m} c_k x^k\Big) = \sum_{k=0}^{m} c_k a^k.$$

The homomorphism Φ_a^p is called the *polynomial functional calculus* for a, and it is common to denote

$$p(a) := \Phi_a^p(p).$$

The functional calculus tells us how to insert elements of A into polynomials.

One can insert elements of A into more general functions. For example, we can take $z \notin \operatorname{spec}(a)$ and define

$$\frac{1}{a-z} := (a-z)^{-1}.$$

More generally, one can define $f(a)$ for every rational function on $\mathbb{C}$ which has no poles on $\operatorname{spec}(a)$. Up to this point, no norm or topology on A is used.

If $a \in A$ is an element of a unital Banach algebra, we can define expressions such as

$$\exp(a) := \sum_{k=0}^{\infty} \frac{1}{k!} a^k.$$

The series on the right converges absolutely. Instead of the exponential series, one can take any other power series $f(x) := \sum_{k=0}^{\infty} c_k x^k$ with positive convergence radius R, and define $f(a)$ when $a \in A$ has norm $\|a\| < R$. These ideas can be pursued further and give the "holomorphic functional calculus". It defines $f(a)$ whenever f is a holomorphic function defined on an open neighborhood of $\mathrm{spec}(a)$, and a is an element of a Banach algebra. The construction of the holomorphic functional calculus requires deeper results from complex analysis.

Instead, we want to generalize the polynomial functional calculus for *normal* elements of a unital C^**-algebra*. The goal is the *continuous functional calculus*. To what this should accomplish, let us return to the situation of Theorem 10.1. Let $f : \mathrm{spec}(T) = \{\lambda_1, \ldots, \lambda_n\} \to \mathbb{C}$ be an arbitrary function, we define the linear map $\Phi_T^c(f) \in \mathcal{L}(V)$ by the formula

$$\Phi_T^c(f) v_j := f(\lambda_j) v_j.$$

It is easily proven that

$$\Phi_T^c : C(\mathrm{spec}(T)) \to \mathcal{L}(V)$$

is a unital $*$-homomorphism, and that $\Phi_T^c(x) = T$, and that Φ_T^c is the unique such homomorphism. It extends the polynomial functional calculus in the sense that $p \in \mathbb{C}[x]$ defines a map $p : \mathrm{spec}(T) \to \mathbb{C}$, and then

$$\Phi_T^c(p) = \Phi_T^p(p).$$

To verify this, one shows that

$$\Phi_T^c(p) v_j = p(\lambda_j) v_j = \Phi_T^p(p) v_j.$$

Theorem 10.3 (The spectral theorem: functional calculus version)**.** *Let A be a unital C^*-algebra and let $a \in A$ be normal. Then there is a unique algebra homomorphism*

$$\Phi_a^c : C(\mathrm{spec}(a)) \to A,$$

the continuous functional calculus*, such that*

$$\Phi_a^c(1) = 1,$$

$$\Phi_a^c(x) = a$$

and

$$\Phi_a^c(\overline{f}) = \Phi_a^c(f)^*.$$

Moreover,

$$\|\Phi_a^c(f)\| = \|f\|$$

for all $f \in C(\mathrm{spec}(a))$.

We usually write

$$f(a) := \Phi_a^c(f).$$

Remark 10.4. *It turns out that the proof is substantially simpler if a is assumed to be self-adjoint. However, we shall carry out all steps for normal a whenever possible, to clearly understand the additional difficulty.*

Proof of the uniqueness statement of Theorem 10.3. Let $\Phi_a, \Phi'_a : C(\mathrm{spec}(a)) \to A$ be two such homomorphisms. We consider

$$B := \{f \in C(\mathrm{spec}(a)) | \Phi_a(f) = \Phi'_a(f)\} \subset C(\mathrm{spec}(a)).$$

This is a subalgebra of $C(\mathrm{spec}(a))$. It contains 1 and the identity function x. Since Φ_a and Φ'_a are both $*$-homomorphisms, we have $f \in B \Rightarrow \overline{f}$. Since x is injective, the Stone-Weierstrass theorem implies that B is dense in $C(\mathrm{spec}(a))$. Finally, by Corollary 9.29, Φ_a and Φ'_s are continuous, so that B is closed. Therefore $B = C(\mathrm{spec}(a))$. □

Construction of the functional calculus for self-adjoint a. Let $p = \sum_{k=0}^n c_k x^k \in \mathbb{C}[x]$. Note that

$$p(a)^* = \sum_{k=0}^n \overline{c_k} (a^*)^k.$$

It follows that $p(a)$ is normal if a is normal. Therefore (Corollary 9.26), we have

$$\|p(a)\| = \rho(p(a)) = \sup_{y \in \mathrm{spec}(p(a))} |y| = \sup_{z \in \mathrm{spec}(a)} |p(z)|. \tag{10.5}$$

Now let

$$\mathcal{T} \subset C(\mathrm{spec}(a))$$

be the subalgebra of all polynomial functions. For $f \in \mathcal{T}$, pick a polynomial $p \in \mathbb{C}[x]$ such that $p(\lambda) = f(\lambda)$ for all $\lambda \in \mathrm{spec}(a)$. (When $\mathrm{spec}(a)$ is infinite, p is uniquely determined, but this is not the case if $\mathrm{spec}(a)$ is a finite set). We define

$$\Phi_a^c(f) := \Phi_a^p(p) = p(a).$$

The equation (10.5) shows that

$$\|\Phi_a^c(f)\| = \|f\|_{C(\mathrm{spec}(a))}.$$

Therefore, $\Phi_a^c(f)$ is well-defined (it does not depend on the choice of the polynomial p), and

$$\Phi_a^c : \mathcal{T} \to A$$

is an isometry.

Since the polynomial functional calculus is an algebra homomorphism, Φ_a^c is an algebra homomorphism, and it is clear that

$$\Phi_a^c(1) = 1,\ \Phi_a^c(x) = a.$$

Since Φ_a^c is an isometry, it extends by continuity to an isometry

$$\Phi_a^c : \overline{\mathcal{T}} \to A$$

from the closure to A. The closure is a subalgebra of $C(\mathrm{spec}(a))$, and the extended map Φ_a^c is an isometry and an algebra homomorphism by a straightforward limit argument.

Up to this point, the construction works for normal a. *Now assume that* $a^* = a$. Therefore

$$\mathrm{spec}(a) \subset \mathbb{R}$$

by Lemma 9.16. This has the following consequence: if $f = \sum_{k=0}^n c_k x^k \in \mathcal{T}$, then the conjugate $\overline{f} \in C(\mathrm{spec}(a))$ is given by the formula

$$\overline{f}(z) = \overline{\sum_{k=0}^n c_k z^k} = \sum_{k=0}^n \overline{c_k} z^k \tag{10.6}$$

for all $z \in \operatorname{spec}(a)$. Hence the conjugate of $f \in \mathcal{T}$ is again in $\mathcal{T}$.

Therefore the subalgebra $\mathcal{T} \subset C(\operatorname{spec}(a))$ is invariant under conjugation, and it is clear that it separates points and contains 1. Therefore, by the Stone–Weierstrass theorem, $\overline{\mathcal{T}} = C(\operatorname{spec}(a))$.

This finishes the construction of Φ_a^c and proves that it is an isometry. We have not yet checked that Φ_a^c is a $*$-homomorphism. But the identity (10.6), together with $a^* = a$, shows that

$$\Phi_a^c(\overline{f}) = \Phi_a^c(f)^*$$

whenever $f \in \mathcal{T}$, and this identity extends by continuity to all of $C(\operatorname{spec}(a))$. □

Remark 10.7. *When $a \in A$ is normal, but not self-adjoint, then $\mathcal{T} \subset C(\operatorname{spec}(a))$ is not preserved by conjugation. Hence the Stone–Weierstrass theorem cannot be applied (in fact, the closure $\overline{\mathcal{T}}$ is usually smaller than $C(\operatorname{spec}(a))$).*

The procedure must then be modified; one considers the subalgebra

$$\mathcal{Q} \subset C(\operatorname{spec}(a))$$

which consists of all polynomial functions of the form

$$\sum_{k,l=0}^{n} c_{kl} x^k \overline{x}^l.$$

This lies dense in $C(\operatorname{spec}(a))$ by the Stone–Weierstrass theorem, and the formula

$$\Phi_a^c\Big(\sum_{k,l=0}^{n} c_{kl} x^k \overline{x}^l\Big) := \sum_{k,l=0}^{n} c_{kl} a^k (a^*)^l$$

defines a $*$-homomorphism $\mathcal{Q} \to A$ (it is important that a is normal in order for Φ_a^c to be multiplicative). To extend Φ_a^c to all of $C(\operatorname{spec}(a))$, one needs to prove that Φ_a^c is bounded (the automatic continuity from Corollary 9.29 is of no use for that because $\mathcal{Q}$ is not complete). It is even an isometry, and the equation

$$\Big\| \sum_{k,l=0}^{n} c_{kl} a^k (a^*)^l \Big\| = \sup_{\lambda \in \operatorname{spec}(a)} \Big| \sum_{k,l=0}^{n} c_{kl} \lambda^k (\lambda^*)^l \Big|$$

holds, but is much more difficult to prove than the equation (10.5) *of the above proof.*

The naive estimate

$$\Big\| \sum_{k,l=0}^{n} c_{kl} a^k (a^*)^l \Big\| \leq \sum_{k,l=0}^{n} |c_{kl}| \|a\|^{k+l}$$

is not sufficient to prove that Φ_a^c is bounded.

10.2. **Multiplication operators.** Theorem 10.1 asserts that any normal operator on a finite-dimensional Hilbert space is equivalent to a certain *normal form*. Here, the normal form is a diagonal matrix. For normal operators on Hilbert spaces, the normal form is necessarily more complicated.

Let us first define the appropriate version of equivalence.

Definition 10.8. *Let (V_j, T_j), $j = 0, 1$, be two Hilbert spaces with $T_j \in \mathcal{L}(V_j)$. A* unitary equivalence $U : (V_0, T_0) \to (V_1, T_1)$ *is an isometric isomorphism $U : V_0 \to V_1$ such that $UT_0U^* = T_1$.*

Note that $\mathcal{L}(V_0) \to \mathcal{L}(V_1)$, $T \mapsto UTU^*$ is a $*$-isomorphism of C^*-algebras, and so it preserves adjoints, spectra and so on.

Next, we define the normal form.

Definition 10.9. *Let (X, μ) be a measure space and let $f \in L^\infty(X, \mu)$. We define the* multiplication operator *$M_f \in \mathcal{L}(L^2(X, \mu))$ by*

$$M_f(h) := fh.$$

Examples 10.10. (1) *Let T be as in Theorem 10.1. Let $X = \underline{n}$ with the counting measure, and let $f : \underline{n} \to \mathbb{C}$ be the function $f(j) := \lambda_j$. Let $U : \ell^2(\underline{n}) \to V$ be the isometric isomorphism determined by $U(\delta_j) := v_j$. Then $UTU^* = M_f$. Hence Theorem 10.1 might be restated by saying that each normal operator on a finite-dimensional Hilbert space is unitarily equivalent to an operator of the form M_f.*

(2) *Similarly, let $X = \mathbb{N}$ with the counting measure and let $a \in c_0(\mathbb{N})$. The spectral theorem for compact self-adjoint operators can be restated by saying that each compact selfadjoint operator on a separable Hilbert space is unitarily equivalent to a multiplication operator. The proof of the spectral theorem for compact selfadjoint operators can be adapted to deal with compact normal operators as well.*

Indeed, the general version reads as follows.

Theorem 10.11 (Spectral theorem, multiplication operator version)**.** *Let V be a Hilbert space and let $T \in \mathcal{L}(V)$ be normal. Then there is a locally finite measure space (X, μ), $f \in L^\infty(X, \mu)$ and a unitary equivalence*

$$U : (L^2(X, \mu), M_f) \cong (V, T).$$

If V is separable, we can take μ to be σ-finite.

One advantage of this formulation is that it gives a lot of insight of the behaviour of normal operators. To that end, let us first study the multiplication operators M_f in some detail.

Definition 10.12. *Let (X, μ) be a measure space and let $f : X \to \mathbb{C}$ be measurable. The* essential *range* $\mathrm{essrange}(f)$ *of f consists of all $z \in \mathbb{C}$ such that for each $\epsilon > 0$, the set $f^{-1}(B_\epsilon(z))$ has positive measure.*

It is clear that $\mathrm{essrange}(f) \subset \mathbb{C}$ is closed, and that $\|f\|L^\infty = \sup_{z \in \mathrm{essrange}(f)} |f(z)|$.

Lemma 10.13. *Let (X, μ) be a measure space.*

(1) *$L^\infty(X, \mu) \to \mathcal{L}(L^2(X, \mu))$, $f \mapsto M_f$, is a unital $*$-homomorphism, and $\|M_f\| \le \|f\|_{L^\infty}$.*
(2) *If μ is locally finite, then $\|M_f\| = \|f\|_{L^\infty}$ for all f.*
(3) *M_f is invertible if and only if there exists $\epsilon > 0$ such that $\mu(f^{-1}(B_\epsilon(0))) = 0$.*
(4) $\mathrm{spec}(M_f) = \mathrm{essrange}(f)$.

Proof. (1) is clear, except perhaps the identity $M_f^* = M_{\overline{f}}$, which follows from

$$\langle M_f h, g\rangle = \int_X \overline{fh} g d\mu = \int_X \overline{h\overline{f}} g d\mu = \langle h, M_{\overline{f}} g\rangle.$$

(2): the inequality $\|M_f\| \leq \|f\|_{L^\infty}$ is trivial. For the reverse one, let $\epsilon > 0$. The set $\{x \in X | |f(x)| \geq \|f\|_{L^\infty} - \epsilon\}$ has positive measure, and because (X, μ) is locally finite, this set contains a subset S with $0 < \mu(S) < \infty$. Then $\|\chi_S\|_{L^2}^2 = \mu(S)$ and

$$\|M_f \chi_S\|^2 = \int_S |f(x)|^2 \geq (\|f\|_{L^\infty} - \epsilon)^2 \mu(S)$$

together prove that $\|f\|_{L^\infty} - \epsilon \leq \|M_f\|$.

(3): The existence of such an ϵ shows that $|f(x)| \geq \epsilon$ almost everywhere. Hence $g(x) := \frac{1}{f(x)}$ is (almost everywhere) defined and (essentially) bounded, and $M_f M_g = M_g M_f = 1$. Vice versa, if M_f is invertible, the open mapping theorem shows that there is $c > 0$ such that $\|M_f h\|^2 \geq c^2 \|h\|^2$ for all $h \in L^2(X)$, or

$$\int_X |f|^2 |h|^2 d\mu \geq c^2 \int_X |h|^2.$$

Now let $h = \chi_{f^{-1}(B_\delta(0))}$. We get

$$\delta^2 \mu(f^{-1}(B_\delta(0))) \geq \int_X |f|^2 |h|^2 d\mu \geq c^2 \int_X |h|^2 = c^2 \mu(f^{-1}(B_\delta(0)))$$

When $\delta < c$, this is only possible if $\mu(f^{-1}(B_\delta(0))) = 0$.

(4): follows immediately from (3). □

Corollary 10.14. *M_f is self-adjoint if and only if* $\mathrm{essrange}(f) \subset \mathbb{R}$*, and this happens if and only if* $\mathrm{spec}(M_f) \subset \mathbb{R}$*. M_f is unitary if and only if* $\mathrm{essrange}(f) \subset S^1$*, and this happens if and only if* $\mathrm{spec}(M_f) \subset S^1$*.* □

10.3. **Proof of the multiplication operator version of the spectral theorem.** Consider a normal operator $T \in \mathcal{L}(V)$ on a Hilbert space. Theorem 10.3 provides the continuous functional calculus

$$\Phi_T^c : C(\mathrm{spec}(T)) \to \mathcal{L}(V)$$

of T, which for simplicity we denote by $f(T) := \Phi_T^c(f)$.

Lemma 10.15. *For each $v \in V$, there exists a unique Radon measure μ_v on* $\mathrm{spec}(T)$ *with $\mu_v(\mathrm{spec}(T)) = \|v\|^2$ and an isometry*

$$U_v : L^2(\mathrm{spec}(T), \mu_v) \to V$$

such that

$$U_v(1) = v$$

and such that

$$U_v \circ M_f = f(T)U$$

for all $f \in C(\mathrm{spec}(T))$.

We have $\mu_v(\mathrm{spec}(T)) = \|v\|^2$. The image of U_v is the closure of the linear subspace

$$\{f(T)v | f \in C(\mathrm{spec}(T))\} \subset V.$$

Proof. Let us first prove the uniqueness of μ_v, which help to prove existence. The conditions on U_v force

$$\int_{\mathrm{spec}(T)} f d\mu_v = \int_{\mathrm{spec}(T)} \overline{1} f d\mu_v = \langle 1, f \rangle_{L^2} = \langle U_v(1), U_v(M_f(1)) \rangle = \langle v, f(T)v \rangle$$

Hence the integration functional for μ_v is uniquely determined, and so μ_v itself is unique.

To show existence, we define a functional

$$L_v(f) := \langle v, f(T)v \rangle$$

on $C(\mathrm{spec}(T))$. Note that

$$L_v(1) = \langle v, v \rangle = \|v\|^2.$$

The functional L_v is positive: if $f \geq 0$, we can write $f = \overline{g}g$ for some $g \in C(\mathrm{spec}(T))$ and obtain

$$L_v(f) = \langle v, \overline{g}(T)g(T)v \rangle = \langle g(T)v, g(T)v \rangle \geq 0$$

because the continuous functional calculus is a $*$-homomorphism.

By the Riesz–Markov–Kakutani theorem, there is a unique Radon measure μ_v such that

$$\int_{\mathrm{spec}(T)} f d\mu_v = \langle v, f(T)v \rangle$$

for all $f \in C(\mathrm{spec}(T))$. Note that

$$\mu_v(\mathrm{spec}(T)) = \|v\|^2.$$

Let us now define

$$U_0 : C(\mathrm{spec}(T)) \to V$$

by

$$U_0(h) := h(T)v.$$

Clearly $U_0(1) = v$ and U_0 is linear. It is also clear that

$$U_0(fh) = f(T)U_0(h). \tag{10.16}$$

Calculate

$$\|U_0(h)\|^2 = \langle h(T)v, h(T)v \rangle = \langle v, h(T)^* h(T)v \rangle = \langle |h|^2(T)v, v \rangle = \int_X |h|^2 d\mu_v = \|h\|^2_{L^2(\mathrm{spec}(T),\mu_v)}.$$

Therefore U_0 is an isometry when $C(\mathrm{spec}(T))$ carries the L^2-norm of the measure μ_v. Therefore, it extends to an isometry

$$U_v : L^2(\mathrm{spec}(T), \mu_v) \to V.$$

The equation (10.16) shows that

$$U_v \circ M_f = f(T) \circ U_v$$

for all $f \in C(\mathrm{spec}(T))$. By construction, the image of U_0 is equal to

$$\{f(T)v | f \in C(\mathrm{spec}(T))\},$$

and so the image of U_v is the closure of that linear subspace. □

If the isometry U_v from Lemma 10.15 were surjective, the proof of Theorem 10.11 would already be complete: then $U_v : L^2(X, \mu_v) \to V_v$ is an isometric isomorphism with inverse U_v^*, and the lemma shows

$$U_v M_f U_v^* = f(T)$$

for all $f \in C(\mathrm{spec}(T))$. Applied to $f = x$, this gives Theorem 10.11.

Definition 10.17. *Let $T \in \mathcal{L}(V)$ be normal. A closed linear subspace $W \subset V$ is* T-invariant *if*

$$f(T)(W) \subset W$$

for all $f \in C(\mathrm{spec}(T))$.

The cyclic subspace *generated by $v \in V$ is*

$$V_v := \overline{\{f(T)v | f \in C(\mathrm{spec}(T))\}} \subset V.$$

The cyclic subspace V_v is T-invariant, and the image of the isometry U_v is precisely V_v. The proof of Theorem 10.11 is finished by breaking up V into cyclic pieces.

Consider two families $(V_i)_{i \in I}$ and $(W_i)_{i \in I}$ of Hilbert spaces. Let $T_i \in \mathcal{L}(V_i, W_i)$. Then

$$\bigoplus_{i \in I} T_i : \bigoplus_{i \in I} V_i \to \bigoplus_{i \in I} W_i$$

denotes the linear map

$$(v_i)_i \mapsto (T_i v_i)_i.$$

This is not necessarily bounded, but if $\sup_{i \in I} \|T_i\| < \infty$, then it is bounded, and one checks easily that

$$\|\bigoplus_{i \in I} T_i\| = \sup_{i \in I} \|T_i\|,$$

so that it extends to a bounded operator

$$\bigoplus_{i \in I} T_i : \bigoplus_{i \in I}^{(2)} V_i \to \bigoplus_{i \in I}^{(2)} W_i$$

of the Hilbert sums. It is easy to see that

$$\bigoplus_{i \in I} T_i^* = (\bigoplus_{i \in I} T_i)^*.$$

Example 10.18. *Let $(X_i, \mu_i)_{i \in I}$ be locally finite measure spaces, and let $f_i \in L^\infty(X_i, \mu_i)$. Let $X = \coprod_{i \in I} X_i$ be the disjoint union. The formula*

$$\mu(S) := \sum_{i \in I} \mu_i(S \cap X_i)$$

defines a measure on X (on the σ-algebra generated by all measurable subsets of some X_i). The union $f := \coprod_{i \in I} f : X \to \mathbb{C}$ is measurable and $\|f\|_{L^\infty} = \sup_{i \in I} \|f_i\|_{L^\infty}$.

Then $L^\infty(X, \mu) \cong \bigoplus_{i \in I} L^2(X_i, \mu_i)$, and M_f corresponds to $\bigoplus_{i \in I}^{(2)} M_{f_i}$.

Lemma 10.19. *Let V be a Hilbert space and let $T \in \mathcal{L}(V)$ be normal. Then there are Hilbert spaces V_i, normal $T_i \in \mathcal{L}(V_i)$ such that (V_i, T_i) is cyclic and such that*

$$(\bigoplus_i V_i, \bigoplus_i T_i)$$

is unitarily equivalent to (V, T).

Lemma 10.15, Lemma 10.19 and Example 10.18 together finish the proof of Theorem 10.11.

Before we give the proof of Lemma 10.19, we isolate the key observation.

Lemma 10.20. *Let $T \in \mathcal{L}(V)$ be normal and let $W \subset V$ be a closed, T-invariant linear subspace. Then the orthogonal complement $W^\perp$ is T-invariant.*

Proof. Let $v \in W^\perp$ and $f \in C(\mathrm{spec}(T))$. Then for each $w \in W$, we have

$$\langle f(T)v, w\rangle = \langle v, f(T)^* w\rangle = \langle v, \overline{f}(T)w\rangle = 0$$

because $\overline{f}(T)w \in W$. □

Proof of Lemma 10.19. The decomposition is obtained by an argument with Zorn's lemma, which is similar to the argument for the existence of complete orthonormal systems (Theorem 2.41).

Let $\mathcal{Z}$ be the set of all subsets $S \subset V$ with the following properties:

(1) For all $v \in S$, we have $\|v\| = 1$,
(2) if $v, w \in S$, $v \neq w$, the cyclic subspaces V_v and V_w are orthogonal.

The set $\mathcal{Z}$ is ordered by inclusion of subsets. It is not empty because $\emptyset \in \mathcal{Z}$. If $\mathcal{C} \subset \mathcal{Z}$ is a chain, it is easy to check that $\bigoplus_{S\in\mathcal{C}} S$ is an element of $\mathcal{Z}$. Hence each chain in $\mathcal{Z}$ has an upper bound, and Zorn's Lemma guarantees the existence of a maximal element $S \in \mathcal{Z}$.

We have to prove that $W := \bigoplus^{(2)}_{v\in S} V_v = V$, or that $W^\perp = 0$. For each $f \in C(\mathrm{spec}(T))$, we have $f(T)(W) \subset W$, and so by Lemma 10.19, $W^\perp$ is T-invariant. If $W^\perp \neq 0$, we could pick $w \in W^\perp$ with $\|w\| = 1$. Then the cyclic subspace V_w is contained in $W^\perp$, and so $S \cup \{w\} \in \mathcal{Z}$, contradicting the maximality of S. □

10.4. **The measurable functional calculus.** The measurable functional calculus is an extension of the continuous functional calculus to measurable functions. Unlike the continuous functional calculus, it can only be defined for normal operators on a Hilbert space, not for normal elements of a C^*-algebra.

One application of Theorem 10.11 is to extend the continuous functional calculus of a normal operator to the *measurable functional calculus.*

First some notation. For a measurable space $(X, \mathcal{B})$, we let $\mathscr{L}^\infty(X, \mathcal{B})$ be the space of all bounded measurable functions $f : X \to \mathbb{C}$ (here $\mathbb{C}$ has the Borel σ-algebra. This is a C^*-algebra, with the supremum norm, pointwise multiplication of functions and $f^*(x) := \overline{f(x)}$. $\mathscr{L}^\infty(X, \mathcal{B})$ is unital unless $\mu(X) = 0$. One has to distinguish between $\mathscr{L}^\infty(X, \mathcal{B})$ and $L^\infty(X, \mu)$: the latter is only defined when we specify a measure. The quotient map $\mathscr{L}^\infty(X) \to L^\infty(X, \mu)$ is surjective, but not injective in general.

For a compact Hausdorff space X, let $\mathcal{B}$ be the Borel-σ-algebra, so that $(X, \mathcal{B})$ is a measurable space. We denote $\mathscr{L}^\infty(X) := \mathscr{L}^\infty(X, \mathcal{B})$.

Theorem 10.21 (Measurable functional calculus)**.** *Let V be a Hilbert space and let $T \in \mathcal{L}(V)$ be normal. Then there exists a unique $*$-homomorphism*

$$\Phi^m_T : \mathscr{L}^\infty(\mathrm{spec}(T)) \to \mathcal{L}(V),$$

the measurable functional calculus *with the following properties:*

(1) *For $f \in C(\mathrm{spec}(T))$, $\Phi^m_T(f) = \Phi^c_T(f)$,*
(2) *if $f_n \in \mathscr{L}^\infty(\mathrm{spec}(T))$ is a bounded sequence and $f_n \to f \in \mathscr{L}^\infty(X)$ pointwise, then $\Phi^m_T(f_n)v \to \Phi^m_T(f)v$ for each $v \in V$.*

The construction that we shall give or Corollary 9.29 proves that $\|\Phi^m_T(f)\| \leq \|f\|_{L^\infty}$ for all f, but equality does not have to be true.

Proof. First we prove existence, and the first step is that we can change (V, T) up to unitary equivalence. Indeed, suppose that V_0, V_1 are two Hilbert spaces, $T_j \in \mathcal{L}(V_j)$ are normal, $U : (V_0, T_0) \to (V_1, T_1)$ is a unitary equivalence. Then

$$\mathrm{Ad}(U) : \mathcal{L}(V_0) \to \mathcal{L}(V_1),\ T \mapsto UTU^*$$

is a $*$-isomorphism of C^*-algebras. Hence it preserves multiplication, the spectrum, the norm and the adjoint operation. If $T_1 = \mathrm{Ad}(U)(T_0)$, then $\mathrm{spec}(T_1) = \mathrm{spec}(T_0)$. Suppose that

$$\Phi^m_{T_0} : \mathscr{L}^\infty(T_0) \to \mathcal{L}(V_0)$$

is a measurable functional calculus for T_0. Then

$$\Phi^m_{T_1} := \mathrm{Ad}(U) \circ \Phi^m_{T_2}$$

is a measurable functional calculus for T_1.

Hence by the multiplication operator version of the spectral theorem, we may assume that $(V, T) = (L^2(X, \mu), M_f)$. We can change f on a set of measure zero[5] such that $f(X) \subset \mathrm{essrange}(f)$. Because $f(X) \subset \mathrm{essrange}(f) = \mathrm{spec}(M_f)$, the composite $g \circ f$ makes sense when $g : \mathrm{spec}(M_f) \to \mathbb{C}$ is a function. If g is (Borel-Borel)-measurable, then $g \circ f$ is measurable. We define

$$\Phi^m_{M_f} : \mathscr{L}^\infty(\mathrm{essrange}(f)) \to \mathcal{L}(L^2(X, \mu))$$

by the formula

$$\Phi^m_{M_f}(g) := M_{g \circ f}.$$

It is straightforward to prove that $\Phi^m_{M_f}$ is a unital $*$-homomorphism (using the identities from Lemma 10.13), that $\|\Phi^m_{M_f}(g)\| \leq \|g\|_{L^\infty}$ and that $\Phi^m_{M_f}(x) = M_f$.

In particular, the restriction of $\Phi^m_{M_f}$ to $C(\mathrm{spec}(M_f))$ is a unital $*$-homomorphism which maps x to M_f. Therefore, by the uniqueness statement of Theorem 10.3, $\Phi^m_{M_f}|_{C(\mathrm{spec}(M_f))}$ agrees with the continuous functional calculus of M_f.

Now assume that $g_n \in \mathscr{L}^\infty(\mathrm{essrange}(f))$ is a sequence with $\|g_n\|_{L^\infty} \leq C$ and that $g_n(y) \to g(y)$ for all x. Let $h \in L^2(X, \mu)$. We need to prove that

$$\|\Phi^m_{M_f}(g_n - g)h\|^2_{L^2} \to 0. \tag{10.22}$$

But

$$\|\Phi^m_{M_f}(g_n - g)h\|^2_{L^2} = \int_X |g_n(f(x)) - g(f(x))|^2 |h(x)|^2 d\mu(x).$$

The integrand converges pointwise to 0, and

$$|g_n(f(x)) - g(f(x))|^2 |h(x)|^2 \leq 4C^2 |h(x)|^2.$$

Since $|h|^2$ is integrable, the dominated convergence theorem applies and proves (10.22).

This finishes the existence proof. For the uniqueness, suppose that $\Phi_0, \Phi_1 : \mathscr{L}^\infty(\mathrm{spec}(T)) \to \mathcal{L}(V)$ be two $*$-homomorphisms with properties (1) and (2) of the theorem. Put

$$B := \{f \in \mathscr{L}^\infty(\mathrm{spec}(T)) | \Phi_0(f) = \Phi_1(f)\} \subset \mathscr{L}^\infty(\mathrm{spec}(T)).$$

---

[5]Here is how this works: the preimage $f^{-1}(\mathrm{essrange}(f)^c) \subset X$ has measure zero, because the open set $\mathrm{essrange}(f)^c$ can be written as the countable union of open balls, and the preimage of each of those balls has measure zero. Therefore, we can achieve, by changing f on a set of measure 0, that $f(X) \subset \mathrm{essrange}(f)$. It follows that $\mathrm{essrange}(f) = \overline{f(X)}$.

We have to prove that $B = \mathscr{L}^\infty(\mathrm{spec}(T))$. Since Φ_0 and Φ_1 are $*$-homomorphisms, B is a subalgebra. By automatic continuity, Φ_0 and Φ_1 are continuous, and so B is closed. By property (1), $C(\mathrm{spec}(T)) \subset B$. Furthermore, if $f_n \in B$ is a bounded sequence which converges pointwise to $f \in \mathscr{L}^\infty(\mathrm{spec}(T))$, by property (2). Lemma 10.23 below implies $B = \mathscr{L}^\infty(\mathrm{spec}(T))$, which is what we had to prove. □

Lemma 10.23. *Let X be a compact metric space and let $B \subset \mathscr{L}^\infty(X)$ be a closed subalgebra such that*

(1) *$C(X) \subset B$,*
(2) *if $f_n \in B$ is a bounded sequence which converges pointwise to $f \in \mathscr{L}^\infty(X)$, then $f \in B$.*

Then $B = \mathscr{L}^\infty(X)$.

Proof. One might naively expect that each $f \in \mathscr{L}^\infty(X)$ is the pointwise limit of a sequence of continuous functions, which would make the proof almost trivial. However, that is not true (try to prove it to understand the problem), so the proof is a little more complicated.

For each bounded $f \in \mathscr{L}^\infty(X)$, there is a sequence of step functions which converge uniformly to f. Because B is norm-closed, it is enough to prove that each step function belongs to B, and by linearity, we must prove that for each Borel set $S \subset X$, $\chi_S \in B$.

So let $\mathcal{B}$ be the Borel-σ-algebra of X and define

$$\mathcal{C} := \{S \in \mathcal{B} | \chi_S \in B\} \subset \mathcal{B}.$$

We must show that $\mathcal{C} = \mathcal{B}$, and do this by proving that $\mathcal{C}$ is a σ-algebra and that each open subset $U \subset X$ belongs to $\mathcal{C}$.

Since $0, 1 \in B$, we have

$$\emptyset, X \in \mathcal{C}.$$

Because $\chi_{S^c} = 1 - \chi_S$, it follows that

$$S \in \mathcal{C} \Rightarrow S^c \in \mathcal{C}.$$

Since B is an algebra and $\chi_S \chi_{S'} = \chi_{S \cap S'}$, it follows that

$$S, S' \in \mathcal{C} \Rightarrow S \cap S' \in \mathcal{C}.$$

Because $S \cup S' = (S^c \cap S'^c)^c$, $\mathcal{C}$ is a Boolean algebra. If $S_1 \subset S_2 \subset \ldots \subset S = \bigcup_{n=1}^\infty$, then $\chi_{S_n} \to \chi_S$ pointwise, and so $S \in \mathcal{C}$ if all S_n belong to $\mathcal{C}$. If U_n, $n \in \mathbb{N}$ are elements of $\mathcal{C}$, then

$$S_n := \bigcup_{m=1}^n U_m \in \mathcal{C}$$

for all n, and $S_1 \subset S_2 \subset \ldots \bigcup_{n=1}^\infty S_n = \bigcup_{n=1}^\infty U_n$, we get that $\bigcup_{n=1}^\infty U_n \in \mathcal{C}$. Altogether, $\mathcal{C}$ is a σ-algebra.

For an open subset $U \subset X$, there is a sequence $f_n \in C(X)$ with $0 \leq f_n \leq 1$ which converges pointwise to χ_U. This is shown in the course of the proof of Proposition C.59. Hence $U \in \mathcal{C}$. □

10.5. **Spectral measures.** The measurable functional calculus is the most powerful version of the spectral theorem.

Theorem 10.24. *Let V be a Hilbert space and let $T \in \mathcal{L}(V)$ be normal. Let*

$$\Phi_T^m : \mathscr{L}^\infty(\operatorname{spec}(T)) \to \mathcal{L}(V)$$

be the measurable functional calculus provided by Theorem 10.21 and let $\mathcal{B}$ be te Borel-σ-algebra of $\operatorname{spec}(T)$. *We define*

$$E : \mathcal{B} \to \mathcal{L}(V)$$

by

$$E(A) := \Phi_T^m(\chi_A) \in \mathcal{L}(V).$$

Then (for Borel sets A, B, A_n)

1. *Each $E(A)$ is a projection.*
2. *$E(\emptyset) = 0$, $E(\operatorname{spec}(T)) = 1$.*
3. *$E(A)E(B) = E(B)E(A) = E(A \cap B)$.*
4. *if $A \cap B = \emptyset$, then $E(A \cup B) = E(A) + E(B)$, and* $\operatorname{im}(E(A)) \perp \operatorname{im}(E(B))$.
5. *If $A_1 \subset A_2 \subset \dots$ is an ascending sequence of Borel sets, then*
$$E(\bigcup_{n=1}^\infty A_n) = \lim_{n\to\infty} E(A_n)$$
in the strong operator topology.

The projection $E(A)$ is called the spectral projection *of A, and the map E is called the* spectral measure *of T.*

Proof. This is an almost trivial consequence of Theorem 10.21: (1) follows from $\chi_A = \chi_A^2 = \overline{\chi_A}$. (2) follows from $\chi_\emptyset = 0$ and $\chi_{\operatorname{spec}(T)} = 1$. (3) follows from $\chi_{A\cap B} = \chi_A \chi_B$. (4) follows from $\chi_{A\cup B} = \chi_A + \chi_B$ which holds for disjoint A, B. (5) follows from $\lim_n \chi_{A_n}(x) = \chi_{\bigcup_{n=1}^\infty A_n}(x)$. □

Examples 10.25.
1. *Let $\dim(V) < \infty$. For $A \subset \operatorname{spec}(T)$, $E(A)$ is the orthogonal projection onto $\bigoplus_{\lambda\in A} \operatorname{Eig}(T, \lambda)$.*
2. *If $f \in L^\infty(X, \mu)$ and $T = M_f$, $E(A)$ is multiplication by the characteristic function of the set $f^{-1}(A)$.*

Proposition 10.26. *Assume the situation of Theorem 10.24 and fix a vector $v \in V$. Define*

$$\nu_v : \mathcal{B} \to \mathbb{R},\ A \mapsto \langle E(A)v, v\rangle.$$

This is a (nonnegative) measure, and it agrees with the measure μ_v constructed in Lemma 10.15. For each $f \in \mathscr{L}^\infty(\operatorname{spec}(T))$, we have

$$\langle v, f(T)v\rangle = \int_{\operatorname{spec}(T)} f(x) d\nu_v(x). \tag{10.27}$$

Proof. Since $E(A) = E(A)E(A) = E(A)^*E(A)$, we have

$$\langle E(A)v, v\rangle = \langle E(A)v, E(A)v\rangle \geq 0,$$

so that ν_v is nonnegative. If $A \cap B = \emptyset$, then

$$\langle E(A \cup B)v, v\rangle = \langle E(A)v, v\rangle + \langle E(B)v, v\rangle$$

and if $A_1 \subset A_2 \subset \ldots$, then $E(A_n)v \to E(\bigcup_{n=1}^{\infty} A_n)v$ and so

$$\lim_n \nu_v(A_n) = \lim_n \langle E(A_n)v, v\rangle = \langle \lim_n E(A_n)v, v\rangle = \langle E(\bigcup_{n=1}^{\infty} A_n)v, v\rangle = \nu_v(\bigcup_{n=1}^{\infty} A_n).$$

Hence ν_v is a Borel measure with total mass

$$\langle E(\mathrm{spec}(T))v, v\rangle = \|v\|^2.$$

Since the step functions lie norm dense in $\mathscr{L}^\infty(\mathrm{spec}(T))$, it suffices to prove (10.27) for step functions. So let $f = \sum_{j=1}^n a_j \chi_{A_j}$ with disjoint A_j's, and compute

$$\langle v, f(T)v\rangle = \sum_{j=1}^n a_j \langle v, E(A_j)v\rangle = \sum_{j=1}^n \nu_v(A_j) = \int_{\mathrm{spec}(T)} f(x) d\nu_v(x).$$

For all continuous functions f, we have

$$\int_X f d\mu_v = \langle v, f(T)v\rangle = \int_{\mathrm{spec}(T)} f(x) d\nu_v(x),$$

and so the Borel measures ν_v and μ_v represent the same functional on $C(\mathrm{spec}(T))$. By the uniqueness part of the Riesz–Markov–Kakutani theorem, $\mu_v = \nu_v$, as claimed. $\square$

Proposition 10.28. *Assume the situation of Theorem 10.24 and let* $\lambda \in \mathrm{spec}(T)$. *Then* $E(\{\lambda\})$ *is the orthogonal projection onto* $\ker(T-\lambda)$. *Hence* λ *is an eigenvalue of* T *if and only if* $E(\{\lambda\}) \neq 0$.

Proof. The identity $(x-\lambda)\chi_{\{\lambda\}} = 0 \in \mathscr{L}^\infty(\mathrm{spec}(T))$ implies that

$$(T-\lambda)E(\{\lambda\}) = 0$$

and so that

$$\mathrm{im}(E(\{\lambda\})) \subset \ker(T-\lambda).$$

The reverse inclusion is less formal. Assume

$$u \in \ker(T-\lambda).$$

Suppose that $f \in C(\mathrm{spec}(T))$ vanishes in a neighborhood of λ. Then the function $g := \frac{f}{x-\lambda}$ is continuous, and since $g \cdot (x-\lambda) = f$, we get

$$f(T)u = g(T)(T-\lambda)u = 0.$$

Next, there is a sequence of continuous $f_n : \mathrm{spec}(T) \to [0,1]$ such that $f_n \equiv 0$ on a neighborhood of λ and such that $\lim_n f_n(x) = \chi_{\mathrm{spec}(T)\setminus\lambda}(x)$ for all x. Then $\chi_{\{\lambda\}} + f_n \to 1$ pointwise (and dominated), so that

$$u = \lim_n (\chi_{\{\lambda\}}(T)u + f_n(T)u) = E(\{\lambda\})u + 0.$$

Therefore $\ker(T-\lambda) \subset \mathrm{im}E(\{\lambda\})$. $\square$

11. Spectral theory via Banach algebras

In the previous chapter, we developped the different version of the spectral theorem out of the continuous functional calculus for normal operators (Theorem 10.3). The proof, however, was only given for self-adjoint operators. In this chapter, we give the proof for normal elements of a C^*-algebra, which is much harder, and requires a fresh look at commutative Banach algebras.

11.1. The spectrum of a commutative Banach algebra.

Lemma 11.1. *Let A be a unital complex Banach algebra and let $\varphi : A \to \mathbb{C}$ be an algebra homomorphism with $\varphi(1) = 1$. Then $\varphi(a) \in \operatorname{spec}_A(a)$ for all $a \in A$ and φ is bounded with operator norm $\|\varphi\| = 1$.*

Proof. Let $a \in A$. Then $\{\varphi(a)\} = \operatorname{spec}_{\mathbb{C}}(\varphi(a))$, and by Theorem 9.17, it follows that $\varphi(a) \in \operatorname{spec}_A(a)$. Moreover $|\varphi(a)| \leq \rho(a) \leq \|a\|$. □

Definition 11.2. *Let A be a commutative unital complex Banach algebra. The* spectrum $\operatorname{Spec}(A)$ *of A is the set of all unital algebra homomorphisms $\varphi : A \to \mathbb{C}$. By Lemma 11.1,* $\operatorname{Spec}(A) \subset D_1(A')$*, and we equip it with the* wk**-topology.*

Lemma 11.3. *The space* $\operatorname{Spec}(A)$ *is a compact Hausdorff space.*

Proof. By the Banach-Alaoglu theorem, $D_1(A')$ is compact and Hausdorff when equipped with the wk*-topology. Hence it suffices to prove that $\operatorname{Spec}(A) \subset A'$ is wk*-closed. Denote for $a \in A$

$$p_a : A' \to \mathbb{C},\ p_a(L) := L(a).$$

This map is continuous by the definition of the wk*-topology. Finally

$$\operatorname{Spec}(A) = p_1^{-1}(1) \cap \bigcap_{a,b \in A} (p_a p_b - p_{ab})^{-1}(0),$$

and this is clearly a closed subspace. □

Lemma 11.4. *Let X be a compact Hausdorff space. Then the map*

$$\eta : X \to \operatorname{Spec}(C(X)),\ \eta(x)(f) := f(x)$$

is a homeomorphism.

Proof. Urysohn's Lemma shows that η is injective: if $x_0 \neq x_1 \in X$, there is $f \in C(X)$ with $f(x_0) = 0$ and $f(x_1) = 1$. Then

$$\eta(x_0)(f) = 0 \neq 1 = \eta(x_1)(f)$$

shows that $\eta(x_0) \neq \eta(x_1)$.

The map η is continuous by the definition of wk*: it suffices to show that

$$p_f \circ \eta : X \to \mathbb{C}$$

is continuous for each $f \in C(X)$, where $p_f(\varphi) = \varphi(f)$. But

$$p_f(\eta(x)) = \eta(x)(f) = f(x)$$

is by definition continuous.

Finally, we have to prove that η is surjective. To this end, we argue by contradiction and assume that $\varphi \in \operatorname{Spec}(C(X))$ is not of the form η_x for any $x \in X$.

Then for each $x \in X$, we have $\varphi \neq \eta_x$, and hence there is a function $f_x \in C(X)$ with
$$\varphi(f_x) \neq \eta_x(f_x) = f_x(x).$$
Let
$$g_x := \frac{\varphi(f_x) - f_x}{\varphi(f_x) - f_x(x)} \in C(X)$$
(the denominator is a number). Then
$$g_x(x) = 1$$
and
$$\varphi(g_x) = 0.$$
Let $U_x := \{y \in X | g_x(y) \neq 0\} \subset X$; this is an open subset. Because X is compact, we find $x_1, \ldots, x_r \in X$ such that $X = \bigcup_{j=1}^r U_{x_j}$. Let
$$h := \sum_{j=1}^r \overline{g_{x_j}} g_{x_j} > 0 \in C(X).$$
It follows that
$$1 = \sum_{j=1}^r \frac{\overline{g_{x_j}}}{h} g_{x_j}.$$
Hence
$$1 = \varphi(1) = \sum_{j=1}^r \varphi(\frac{\overline{g_{x_j}}}{h})\varphi(g_{x_j}) = 0,$$
a contradiction. □

Lemma 11.5 (Gelfand transform)**.** *Let A be a commutative unital complex Banach algebra. The formula*
$$\hat{a}(\varphi) := \varphi(a)$$
defines a unital algebra homomorphism
$$\Gamma : A \to C(\mathrm{Spec}(A)),\ \Gamma(a)(\varphi) := \hat{a}(\varphi),$$
the Gelfand transform*. We have*
$$\hat{a}(\mathrm{Spec}(A)) \subset \mathrm{spec}_A(a)$$
and in particular
$$\|\hat{a}\| \leq \rho(a).$$

Proof. We first show that $\Gamma(a)$ is a continuous function when $a \in A$. But $\Gamma(a)$ is nothing else than the restriction of the evaluation map $p_a : A' \to \mathbb{C}$, $L \mapsto L(a)$, to the subspace $\mathrm{Spec}(A) \subset A'$, and hence clearly continuous, by the construction of the wk*-topology.

Since each $\varphi \in \mathrm{Spec}(A)$ is a unital algebra homomorphism, we have
$$\Gamma(ab)(\varphi) = \varphi(ab) = \varphi(a)\varphi(b) = \Gamma(a)(\varphi)\Gamma(b)(\varphi)$$
and
$$\Gamma(1)(\varphi) = \varphi(1) = 1.$$
The proof that $\Gamma(a + b) = \Gamma(a) + \Gamma(b)$ is the same. Hence Γ is a unital algebra homomorphism.

Lemma 11.1 shows that $\hat{a}(\varphi) = \varphi(a) \in \mathrm{spec}_A(a)$, and so $\hat{a}(\mathrm{Spec}(A)) \subset \mathrm{spec}_A(a)$. The estimate $\|\hat{a}\| \leq \rho(a)$ follows immediately. □

Example 11.6. *Let $A = C(X)$. The composition of Γ with the isomorphism $\eta^* : C(\mathrm{Spec}(C(X))) \to C(X)$ induced by the homeomorphism η from Lemma 11.4 is the map*
$$\eta^* \circ \Gamma : C(X) \to C(\mathrm{Spec}(C(X))) \to C(X)$$
given by
$$\eta^*(\Gamma(f))(x) = \Gamma(f)(\eta(x)) = \eta(x)(f) = f(x),$$
hence the identity. Therefore Γ is an isomorphism in that case.

11.2. **The Gelfand-Naimark theorem.**

Theorem 11.7 (Gelfand-Naimark Theorem)**.** *Let A be a commutative unital complex Banach algebra, $a \in A$ and $\lambda \in \mathrm{spec}_A(a)$. Then there exists $\varphi \in \mathrm{Spec}(A)$ such that $\varphi(a) = \lambda$.*

Corollary 11.8. *Let $a \in A$ be an element of a commutative unital Banach algebra. Then*
$$\hat{a}(\mathrm{Spec}(A)) = \mathrm{spec}_A(a)$$
and
$$\|\hat{a}\| = \rho(a). \quad \square$$

Corollary 11.9 (Gelfand-Naimark Theorem for C^*-algebras)**.** *Let A be a commutative unital C^*-algebra. Then the Gelfand transform $\Gamma : A \to C(\mathrm{Spec}(A))$ is an isometric $*$-isomorphism.*

Proof. If $a \in A$ is selfadjoint, then $\mathrm{spec}_A(a) \subset \mathbb{R}$ by Lemma 9.16. It follows from Lemma 11.1 that $\hat{a} \in C(\mathrm{Spec}(A))$ is real-valued.

A general element of A can be written in the form
$$a = a_1 + ia_2,$$
where
$$a_1 := \frac{1}{2}(a + a^*),\ a_2 := \frac{1}{2i}(a - a^*)$$
are self-adjoint. Then $a^* = a_1 - ia_2$, and it follows that
$$\hat{a^*}(\varphi) = \varphi(a_1) - i\varphi(a_2) = \overline{\varphi(a_1) + i\varphi(a_2)} = \overline{\hat{a}(\varphi)}. \tag{11.10}$$
Therefore Γ is a $*$-homomorphism.

Because A is commutative, every element of A is normal. From Corollary 11.8 and Corollary 9.26, we get
$$\|\hat{a}\| = \rho(a) = \|a\|,$$
so that Γ is an isometry.

The subalgebra $\Gamma(A) \subset C(\mathrm{Spec}(A))$ is closed and contains 1. It separates the points by definition: if $\varphi_0 \neq \varphi_1 \in \mathrm{Spec}(A)$, then there is $a \in A$ with $\hat{a}(\varphi_0) = \varphi_0(a) \neq \varphi_1(a) = \hat{a}(\varphi_1)$. Finally, because
$$\Gamma(a^*) = \Gamma(a)^* = \overline{\Gamma(a)},$$
$\Gamma(A)$ is invariant under conjugation. By the Stone-Weierstrass theorem, $\Gamma(A) = C(\mathrm{Spec}(A))$. $\square$

Equation (11.10) also shows the following result.

Lemma 11.11. *Let A be a unital C^*-algebra. Then each unital algebra homomorphism $\varphi : A \to \mathbb{C}$ is in fact a $*$-homomorphism.* $\square$

The first step in the proof of Theorem 11.7 is of independent interest.

Theorem 11.12 (Gelfand-Mazur)**.** *Let A be a commutative unital complex Banach algebra which is a field. Then $A \cong \mathbb{C}$.*

Proof. For $a \in A$, there exists $z \in \operatorname{spec}(a)$. Then $a - z1 \in A$ is not invertible, hence $a - z1 = 0$, hence $a = z1$. Therefore $\mathbb{C} \to A$, $z \mapsto z1$ is an isomorphism. □

Proof of Theorem 11.7. It is enough to prove that if $a \in A$ is not invertible, then there is $\varphi \in \operatorname{Spec}(A)$ with $\varphi(a) = 0$. The set

$$(a) := \{ba | b \in A\} \subset A$$

is an ideal in A (Definition A.22), and $(a) \neq 1$, since otherwise $1 = ba$ and a must be invertible since A is commutative. A standard application of Zorn's Lemma (Theorem A.23) proves that there is a maximal ideal $I \subset A$ with $(a) \subset I$ (this means that $I \neq A$ and that when $I \subset J \subset A$ is a larger ideal with $J \neq A$, then $J = I$.

The closure $\overline{I}$ of I is again an ideal. Since $A^\times \cap I = \emptyset$ and $A^\times \subset A$ is open, it must be true that $A^\times \cap \overline{I} = \emptyset$. Hence $\overline{I} \neq A$, and by maximality of I, it follows that $I = \overline{I}$, or that I is closed.

The quotient space A/I is a Banach space. The formula

$$(a + I)(b + I) := ab + I$$

gives a well-defined structure of a unital commutative algebra on A/I, and the quotient map $\pi : A \to A/I$ is an algebra homomorphism.

For $x, y \in A/I$, we claim that

$$\|xy\| \leq \|x\|\|y\|.$$

To see this, let $\epsilon > 0$ and pick $a, b \in A$ with $\pi(a) = x$ and $\pi(b) = y$ and $\|a\| \leq \|x\| + \epsilon$, $\|b\| \leq \|y\| + \epsilon$. Then

$$\|xy\| = \|\pi(ab)\| \leq \|ab\| \leq \|a\|\|b\| \leq (\|x\| + \epsilon)(\|y\| + \epsilon);$$

as this is true for all $\epsilon > 0$, we get $\|xy\| \leq \|x\|\|y\|$.

It is clear that

$$\|1\| = \|\pi(1)\| \leq \|1\| = 1;$$

and

$$\|1\| = \|11\| \leq \|1\|^2$$

implies $1 \leq \|1\|$. Together $\|1\| = 1$.

This proves that the quotient A/I of a unital commutative Banach algebra by a proper ideal is again a unital commutative Banach algebra.

If I is maximal, then A/I is a field: if $x \in A/I$ is not a unit, then the ideal $(x) \subset A/I$ is a proper ideal, and the preimage $\pi^{-1}((x)) \subset A$ is a proper ideal which contains I. Since I is maximal, $I = \pi^{-1}((x))$, and this implies $x = 0$.

Using Theorem 11.12, we find $A/I \cong \mathbb{C}$. The composition $A \xrightarrow{\pi} A/I \cong \mathbb{C}$ is the desired $\varphi \in \operatorname{Spec}(A)$. □

11.3. **The continuous functional calculus.**

Definition 11.13. *Let A be a unital C^*-algebra and let $a \in A$ be a normal element. We let $C^*(a) \subset A$ be the smallest closed $*$-subalgebra such that $1 \in C^*(a)$ and $a, a^* \in C^*(A)$.*

More constructively, $C^*(a)$ is the closure of the span of all elements $a^k(a^*)^l$, $k, l \in \mathbb{N}$. Note that $C^*(a)$ is commutative. By the Gelfand-Naimark theorem (or rather Corollary 11.9), the Gelfand transformation is a $*$-isomorphism

$$\Gamma : C^*(a) \cong C(\mathrm{Spec}(C^*(a))).$$

Lemma 11.14. *The function $\Gamma(a) = \hat{a} \in C(\mathrm{Spec}(C^*(a)))$ gives a homeomorphism*

$$\hat{a} : \mathrm{Spec}(C^*(a)) \to \mathrm{spec}_{C^*(a)}(a) \subset \mathbb{C}.$$

Proof. From Lemma 11.1, it follows that $\hat{a}(\mathrm{Spec}(C^*(a))) \subset \mathrm{spec}_{C^*(a)}(a)$, and from Theorem 11.7, it follows that $\hat{a}$ maps onto $\mathrm{spec}_{C^*(a)}(a)$.

Since both spaces are compact, it is enough to verify that $\hat{a}$ is injective. So assume

$$\hat{a}(\varphi_0) = \hat{a}(\varphi_1)$$

for $\varphi_0, \varphi_1 \in \mathrm{Spec}(C^*(a))$. This equation means that

$$\varphi_0(a) = \varphi_1(a).$$

The set

$$B := \{b \in C^*(a) | \varphi_0(b) = \varphi_1(b)\} \subset C^*(a)$$

is a closed subalgebra and contains 1 and a. Moreover, Lemma 11.11 shows that B is a $*$-subalgebra, and the definition of $C^*(a)$ proves that $B = C^*(a)$, in other words that $\varphi_0 = \varphi_1$. $\square$

By Theorem 9.17, we have

$$\mathrm{spec}_A(a) \subset \mathrm{spec}_{C^*(a)}(a) \subset \mathbb{C}.$$

The converse is also true:

Lemma 11.15. *Let A be a unital C^*-algebra, let $B \subset A$ be a closed unital $*$-subalgebra. Then for each $a \in B$, we have*

$$\mathrm{spec}_A(a) = \mathrm{spec}_B(a).$$

Proof. The inclusion $\subset$ follows from Theorem 9.17. For the reverse inclusion, we need to prove that if $a \in B$ is invertible in A, then its inverse lies in B (suppose this is shown and $\lambda \in \mathrm{spec}_A(a)^c$. Then $a - \lambda$ is invertible in A, hence invertible in B, hence $\lambda \in \mathrm{spec}_B(a)^c$.)

Assume that $a \in B$ is self-adjoint and invertible in A. Since $\mathrm{spec}_B(a) \subset \mathbb{R}$, $a + it \in B^\times$ whenever $t \in \mathbb{R} \setminus \{0\}$. So $(a + it)^{-1} \in B$. Since the inversion map is continuous, we have

$$a^{-1} = \lim_{t \to 0} (a + it)^{-1} \in B$$

since B is closed.

Now let $a \in B$ be a general element which is invertible in A. Then a^*a and aa^* are selfadjoint and invertible in A, and by what we have shown

$$(a^*a)^{-1}, (aa^*)^{-1} \in B.$$

Then
$$((a^*a)^{-1}a^*)a = 1$$
and
$$a(a^*(aa^*)^{-1}) = 1$$
prove that a is invertible in B. □

Theorem 11.16. *Let $a \in A$ be a normal element of a unital C^*-algebra. Then there is a unique unital $*$-homomorphism*
$$\Phi_a^c : C(\mathrm{spec}_A(a)) \to A$$
such that $\Phi_a^c(x) = a$. It is an isometry.

Proof. We have established uniqueness already in the proof of Theorem 10.3. We define Φ_a^c as the composition
$$C(\mathrm{spec}_A(a)) = C(\mathrm{spec}_{C^*(a)}(a)) \overset{-\circ\hat{a}}{\to} C(\mathrm{Spec}(C^*(a))) \overset{\Gamma^{-1}}{\to} C^*(a) \subset A.$$
The first equality comes from Lemma 11.15. The second is an isomorphism by Lemma 11.14. The third is the inverse of the Gelfand transform.

This defines a unital $*$-homomorphism which is also an isometry. To show that it sends the identity function x to a, we need to prove
$$\Gamma^{-1}(x \circ \hat{a}) = a \in A.$$
But this is a tautology: x is the identity and $\hat{a} = \Gamma(a)$ are just two notations for the same thing, so clearly
$$\Gamma^{-1}(x \circ \hat{a}) = \Gamma^{-1}(\hat{a}) = \Gamma^{-1}(\Gamma(a)) = a.$$
□

Appendix A. The Axiom of Choice and Zorn's Lemma

A.1. **The Axiom of Choice.**

Axiom of Choice. *If X_i, $i \in I$, are nonempty sets, and $I \neq \emptyset$, then the product*
$$\prod_{i \in I} X_i = \{(x_i)_{i \in I} | x_i \in X_i\}$$
is nonempty.

Lemma A.1. *The following three axioms are equivalent:*

1. *The axiom of choice.*
2. *Every surjective map $f : X \to Y$ has a right inverse, i.e. there is a map $g : Y \to X$ with $f \circ g = \mathrm{id}_Y$.*
3. *For every set X, there is a* choice function $\mathrm{ch} : \mathcal{P}(X) \setminus \{\emptyset\} \to X$, *i.e. a function such that* $\mathrm{ch}(S) \in S$ *for each* $S \in \mathcal{P}(X) \setminus \{\emptyset\}$.

Proof. $1 \Rightarrow 3$: Without loss of generality, $X \neq \emptyset$. The product
$$\prod_{S \in \mathcal{P}(X) \setminus \emptyset} S$$
is not empty, and we pick an element
$$(x_S)_{S \in \mathcal{P}(X) \setminus \emptyset}.$$
Then $\mathrm{ch}(S) := x_S$ is a choice function.

$3 \Rightarrow 2$: Since f is surjective, we have $f^{-1}(y) \neq \emptyset$ for each y. Let ch be a choice function for X. Then put

$$g(y) := \mathrm{ch}(f^{-1}(y)) \in X.$$

$2 \Rightarrow 1$: The map

$$f : \{(i,x) | i \in I, x \in X_i\} \to I;\ f(i,x) := i$$

is surjective, and let g be a right inverse of f. Then

$$(g(i))_{i \in I} \in \prod_{i \in I} X_i.$$

□

A.2. Zorn's lemma.

Definition A.2. *A* partial order *on a set X is a binary relation $\leq$ on X such that*

(1) *$x \leq x$ for all $x \in X$,*
(2) *if $x \leq y$, $y \leq z$, then $x \leq z$,*
(3) *if $x \leq y$ and $y \leq x$, then $x = y$.*

A total order *is a partial order, such that for all $x, y \in X$, one of the relations $x \leq y$ or $y \leq x$ holds.*

A partially ordered set *or* poset *is a set X, together with a partial order on X. A subset $Y \subset X$ is a* chain *if the induced partial order is a total order.*

We write $x < y$ if $x \leq y$, but $x \neq y$, $x \geq y$ if $y \leq x$, and $x > y$ if $y < x$.

Definition A.3. *Let X be a partially ordered set and $Y \subset X$.*

(1) *An* upper bound *for Y is an $x \in X$, such that $y \leq x$ for all $y \in Y$.*
(2) *An* lower bound *for Y is an $x \in X$, such that $y \geq x$ for all $y \in Y$.*
(3) *A* strict upper bound *for Y is an upper bound x such that $y < x$ for all $y \in Y$.*
(4) *A* strict lower bound *for Y is a lower bound x such that $y > x$ for all $y \in Y$.*
(5) *A* maximal element *of X is an $x \in X$, such that $y \in X$, $x \leq y$ implies $y = x$.*
(6) *A* minimal element *of X is an $x \in X$, such that $y \in X$, $x \geq y$ implies $y = x$.*
(7) *A* greatest element *of X is an upper bound for X.*
(8) *A* least element *of X is a lower bound for X.*

We remark that greatest elements are maximal, and that a poset X has at most one greatest element. Maximal elements do not need to be greatest elements, and a set might have many maximal elements. An analogous remark holds for least elements and minimal elements.

Theorem A.4 (Zorn's Lemma)**.** *Let X be a nonempty partially ordered set such that each chain in X has an upper bound. Then X has a maximal element.*

Zorn's lemma is logically equivalent to the Axiom of Choice. One direction is fairly easy, and we prove the other one further below.

Proof that Zorn's Lemma implies the Axiom of Choice. Let $f : X \to Y$ be surjective. We let $\mathfrak{Z}$ be the set of all pairs (Z, h), with $Z \subset Y$ and $h : Z \to X$ a map with $f(h(y)) = y$ for all $y \in Z$. We have to prove that there is an element $(Y, g) \in \mathfrak{Z}$.

The set $\mathfrak{Z}$ is partially ordered by inclusion:

$$(Z,h) \leq (Z',h') \Leftrightarrow Z \subset Z',\ h'|_Z = h.$$

Since $(\emptyset,\emptyset) \in \mathfrak{Z}$ (the empty set with the empty map to X), $\mathfrak{Z} \neq \emptyset$. Let $\mathfrak{C}$ be a chain in $\mathfrak{Z}$. Put

$$W := \bigcup_{(Z,h)\in\mathfrak{C}} Z \subset Y$$

For $y \in W$, pick $(Z,h) \in \mathfrak{C}$ such that $y \in Z$. If $(Z',h') \in \mathfrak{C}$ is another such element, then either $(Z,h) \leq (Z',h')$ or $(Z',h') \leq (Z,h)$ because $\mathfrak{C}$ is a chain. In the first case

$$h(y) = (h'|_Z)(y) = h'(y)$$

and similarly, in the second case $h(y) = h'(y)$. Therefore, the map

$$k : W \to X;\ k(y) = h(y)\ \ y \in Z, (Z,h) \in \mathfrak{C}$$

is well-defined and $f(k(y)) = y$ for all $y \in W$. By construction $(Z,h) \leq (W,k)$ for all $(Z,h) \in \mathfrak{C}$, so that $\mathfrak{C}$ has an upper bound.

By Zorn's Lemma, $\mathfrak{Z}$ has a maximal element (Z,h). If $Z \neq Y$, pick $y \in Y \setminus Z$ and $x \in X$ with $f(x) = y$. Then (Z',h'), defined by

$$Z' := Z \cup \{y\},;h'|_Z := h,\ h'(y) := x$$

is an element of $\mathfrak{Z}$ with $(Z',h') > (Z,h)$, contradicting the maximality of (Z,h). □

A.3. **Proof of Zorn's lemma from the Axiom of Choice.** The proof of Zorn's lemma that we give now follows [7]. We need some preliminaries.

Definition A.5. *A* well-ordering *on a set X is a total order $\leq$ on X, such that each nonempty subset $Y \subset X$ has a least element.*

The (unique) least element of Y will be denoted

$$\min(Y).$$

Note that a subset of a well-ordered set is again well-ordered. If X is partially ordered and $x \in X$, we write

$$P(X,x) := \{y \in X | y < x\}.$$

Definition A.6. *Let X be well-ordered. An* initial segment *of X is a subset $Y \subset X$ with the property that*

$$y \in Y, z \in X, z < y \Rightarrow z \in Y.$$

An initial segment Y is proper *if $Y \neq X$.*

Lemma A.7. *Let X be a well-ordered set.*

(1) *The proper initial segments $Y \subset X$ are exactly the subsets $P(X,x)$, $x \in X$.*
(2) *The union of initial segments is an initial segment.*

Proof. (1): It is clear that $P(X,x)$ is a proper initial segment. If $Y \neq X$ is an initial segment, let

$$x := \min(X \setminus Y).$$

If $y < x$, then $y \in Y$ by the definition of x, so that $P(X,x) \subset Y$. If $y \in Y$, then $x \leq y$ cannot hold by assumption, so that $x > y$. Hence $Y \subset P(X,x)$.

(2): If $Y = \bigcup_{i\in I} Y_i$ is a union of initial segments, $y \in Y$, $x \in X$, $x < y$, pick $i \in I$ with $y \in Y_i$. Then $x \in Y_i \subset Y$. □

The proof of Zorn's Lemma is by contradiction, and it is structured into two lemmas.

Lemma A.8. *Assume that $X \neq \emptyset$ is a partially ordered set such that each of its chains has a upper bound, but X has no maximal element. Let $\mathfrak{C}\mathfrak{H}(X)$ be the set of all chains in X. Then there is a map $g : \mathfrak{C}\mathfrak{H}(X) \to X$, such that $g(C)$ is a strict upper bound for C, for each $C \in \mathfrak{C}\mathfrak{H}(X)$.*

Proof. Let $\mathrm{Upp}(C) \subset X$ be the set of all strict upper bounds for the chain C. The hypothesis implies that

$$\mathrm{Upp}(C) \neq \emptyset$$

for each chain C: if y is an upper bound for C, there is $x \in X$ with $y < x$ since X has no maximal element, and $x \in \mathrm{Upp}(C)$. Now let ch be a choice function for X, using the axiom of choice. We define

$$g : \mathfrak{C}\mathfrak{H}(X) \to X;\ g(C) := \mathrm{ch}(\mathrm{Upp}(C)).$$

By construction, $g(C)$ is a strict upper bound for each C. □

Assumption A.9. *For the rest of the proof, let X be a partially ordered set in which each chain has an upper bound, but which has no maximal element, and let $g : \mathfrak{C}\mathfrak{H}(X) \to X$ be a map as constructed in Lemma A.8.*

Definition A.10. *A subset $A \subset X$ is* distinguished*, if*

(1) *A is well-ordered,*
(2) *for all $x \in A$, we have $x = g(P(A, x))$.*

Lemma A.11. *Let X, g be as in A.9.*

(1) *If $A, B \subset X$ are distinguished, either A is an initial segment of B or B is an initial segment of A.*
(2) *A union of distinguished subsets is distinguished.*
(3) *If A is distinguished, then $\overline{A} = A \cup \{g(A)\}$ is distinguished.*

Proof. (1): Let C be the union of all common initial segments of A and B, which is a common initial segment of A and B, by Lemma A.7 (2). We have to show that $C = A$ or $C = B$. For the sake of contradiction, assume that C is proper in both, A and B. Then there are $a \in A$ and $b \in B$ such that

$$C = P(A, a) = P(B, b).$$

Since A and B are distinguished, we have

$$a = g(P(A, a)) = g(P(B, b)) = b$$

and put

$$\overline{C} := C \cup \{a\} \subset A \cap B.$$

Let

$$c \in \overline{C},\ x \in A \text{ and } x < c.$$

Then $x \in \overline{C}$: if $c = a$, $x \in P(A, a) = C \subset \overline{C}$, and if $c \in C$, then $x < c < a$ implies $x \in P(A, a)$ as well. Therefore, $\overline{C}$ is an initial segment of A. For symmetry reasons, $\overline{C}$ is an initial segment of B. This is a contradiction as C is a proper subset of $\overline{C}$ and C contains all common initial segments.

(2): Let $A_i \subset X$, $i \in I$ be distinguished subsets and let

$$U := \bigcup_{i \in I} A_i \subset X.$$

We claim that U is distinguished.

Firstly, we show that U is totally ordered. Let $x, y \in U$, $x \in A_i$, $y \in A_j$. By (1), A_i is an initial segment of A_j or vice versa; without loss of generality $A_i \subset A_j$, and $x, y \in A_j$. Since A_j is totally ordered, $x \leq y$ or $y \leq x$.

To prove that U is well-ordered, let $Z \subset U$ be nonempty. There is $i \in I$ such that $Z \cap A_i \neq \emptyset$, and we put

$$y := \min(Z \cap A_i).$$

Claim: y is a least element of Z. To verify this, let $z \in Z$ and choose $j \in I$ with $z \in A_j$. If $A_j \subset A_i$, then $z \in Z \cap A_i$ and hence $z \geq y$. If $A_j \not\subset A_i$, A_i is an initial segment of A_j by (1). If $z \in A_i$, we have $z \geq y$ by definition. If $z \notin A_i$, we have $x < z$ whenever $x \in A_i$, in particular $y < z$. Therefore y is a least element of Z, and hence U is well-ordered.

For the last property of a distinguished set, we first show that when $A_i \subset U$ and $x \in A$, then

$$P(A_i, x) = P(U, x). \tag{A.12}$$

Assume that this is proven. Then for $x \in U$, pick $i \in I$ with $x \in A_i$ and conclude

$$x = g(P(A_i, x)) = g(P(U, x)).$$

To prove (A.12), observe that $P(A_i, x) \subset P(U, x)$ holds trivially. If $y \in P(U, x)$, pick $j \in I$ with $y \in A_j$. If $A_j \subset A_i$, $y \in A_i$ and $y < x$, so $y \in P(A_i, x)$. If $A_j \not\subset A_i$, $A_i \subset A_j$ is an initial segment. If $y \in A_i$, then again $y \in P(A_i, x)$. If $y \notin A_i$, we must have $y > z$ for all $z \in A_i$, in particular $y > x$, which is absurd. This finishes the proof of (A.12) and hence of part (2) of the lemma.

(3): it is clear that $\overline{A}$ is well-ordered, and it is also clear that $P(A, x) = P(\overline{A}, x)$ for all $x \in A$, so that

$$g(P(\overline{A}, x)) = x$$

when $x \in A$. We have arranged things so that $P(\overline{A}, g(A)) = A$ and so

$$g(\overline{A}, g(A)) = g(A)$$

as well. □

Proof of Zorn's lemma. We argue by contradiction, and let X and g be as in A.9. Let $U \subset X$ be the union of all distinguished subsets of X, which is distinguished by Lemma A.11 (2), and let $\overline{U} = U \cup \{g(U)\}$, which is also distinguished by Lemma A.11 (3). Since $g(U) \notin U$, U is a proper subset of $\overline{U}$, which is a contradiction. □

A.4. **Application I: Bases in vector spaces.** We give now three typical applications of Zorn's lemma.

Theorem A.13 (Bases in vector spaces)**.** *Let $\Bbbk$ be a field and let V be a $\Bbbk$-vector space, $Z \subset Y \subset V$ be two subsets such that Z is linearly independent and Y generates V. Then there is a basis X with $L \subset X \subset Y$.*

Proof. Let $\mathfrak{X}$ be the set of all linearly independent subsets L with $Z \subset L \subset Y$. We order $\mathfrak{X}$ by inclusion. Since $Z \in \mathfrak{X}$, $\mathfrak{X} \neq \emptyset$. Let $\mathfrak{C}$ be a chain in $\mathfrak{X}$. Then $B := \bigcup_{L \in \mathfrak{C}}$ contains Z and is contained in Y, and it is linearly independent: if $b_1, \ldots, b_n \in B$, $a_1, \ldots, a_n \in K$ are such that

$$\sum_{i=1}^{n} a_i b_i = 0,$$

pick $L_1, \ldots, L_n \in \mathfrak{C}$ with $b_i \in L_i$. Since $\mathfrak{C}$ is a chain, we can assume that $L_1 \subset \ldots \subset L_1$. So $b_1, \ldots, b_n \in L_n$ and since L_n is linearly independent, it follows that $a_1 = \ldots = a_n = 0$. Therefore B is linearly independent. Hence the hypotheses of Zorn's lemma hold, and there is a maximal element $C \in \mathfrak{X}$. If C would not generate V, we find $v \in Y \setminus \operatorname{span}(C)$, and $C \cup \{v\}$ is linearly independent, contradicting the maximality of C.

Hence C is a linearly independent generating set, in other words a basis of V. □

A.5. **Application II: The ultrafilter lemma.** Ultrafilters are a crucial tool for the proof of Tychonov's Theorem. Let us begin with some definitions.

Definition A.14. *Let X be a set. A* filter *on X is a nonempty subset $\mathcal{F} \subset \mathcal{P}(X)$ such that*

(1) $\emptyset \notin \mathcal{F}$,
(2) $F_1, \ldots, F_n \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{n} F_i \in \mathcal{F}$,
(3) $F \in \mathcal{F}$, $F \subset G \subset X \Rightarrow G \in \mathcal{F}$.

The second property is also called *finite intersection property*: a subset $\mathcal{A} \subset \mathcal{P}(X)$ has the finite intersection property if $F_1, \ldots, F_n \in \mathcal{A}$, then $\bigcap_{i=1}^{n} F_i \neq \emptyset$.

Example A.15. *If $x \in X$, then*

$$\mathcal{F}_x := \{S \subset X | x \in S\}$$

is a filter.

Example A.16. *If X is infinite, then*

$$\mathcal{F} := \{S \subset X || S^c| < \infty\}$$

is a filter.

Example A.17. *If $\mathcal{A} \subset \mathcal{P}(X)$ has the finite intersection property, then*

$$\langle \mathcal{A} \rangle := \{S \subset X | \exists F_1, \ldots, F_n \in \mathcal{A} : \bigcap_{i=1}^{m} F_i \subset S$$

is a filter, the filter generated by *$\mathcal{A}$.*

Example A.18. *If X is a topological space and $x \in X$ a point, the set $\mathcal{U}(x)$ of all neighborhoods of x is a filter, the* neighborhood filter *of x.*

Lemma A.19. *For a filter $\mathcal{F}$ on X, the following conditions are equivalent:*

(1) *$\mathcal{F}$ is a maximal filter, i.e. if $\mathcal{G}$ is a filter with $\mathcal{F} \subset \mathcal{G}$, then $\mathcal{G} = \mathcal{F}$.*
(2) *If $S \subset X$, then $S \in \mathcal{F}$ or $S^c \in \mathcal{F}$.*
(3) *If $S \subset X$ is a subset such that $S \cap F \neq \emptyset$ for all $F \in \mathcal{F}$, then $S \in \mathcal{F}$.*

Definition A.20. *An* ultrafilter *on X is a filter that satisfies the conditions of Lemma A.19.*

Proof of Lemma A.19. $1 \Rightarrow 3$: The hypothesis says that $\mathcal{F} \cup \{S\}$ has the finite intersection property. Then $\langle \mathcal{F} \cup \{S\} \rangle$ is a filter which contains $\mathcal{F}$. So by maximality of $\mathcal{F}$,

$$\langle \mathcal{F} \cup \{S\} \rangle = \mathcal{F} \Rightarrow S \in \mathcal{F}.$$

$3 \Rightarrow 2$: One of the sets $\mathcal{F} \cup \{S\}$ and $\mathcal{F} \cup \{S^c\}$ have the finite intersection property. If not, there are

$$F_1, \ldots, F_n, G_1, \ldots, G_m \in \mathcal{F}$$

such that

$$S \cap F_1 \cap \ldots \cap F_n = \emptyset$$

and

$$S^c \cap G_1 \cap \ldots \cap G_m = \emptyset.$$

It follows that

$$F_1 \cap \ldots \cap F_n \cap G_1 \cap \ldots \cap G_m = \emptyset,$$

violating the filter axioms.

If $\mathcal{F} \cup \{S\}$ has the finite intersection property, then $S \in \mathcal{F}$, and if $\mathcal{F} \cup \{S^c\}$ has the finite intersection property, then $S^c \in \mathcal{F}$.

$2 \Rightarrow 1$: Let $\mathcal{G}$ be a filter with $\mathcal{F} \subset \mathcal{G}$ and $S \in \mathcal{G}$. It cannot be that $S^c \in \mathcal{F}$, since $S^c \cap S = \emptyset$. Hence $S \in \mathcal{F}$, so $\mathcal{F} = \mathcal{G}$ and $\mathcal{F}$ is maximal. $\square$

Lemma A.21 (Ultrafilter-Lemma)**.** *Let $\mathcal{A} \subset \mathcal{P}(X)$ have the finite intersection property. Then there is an ultrafilter $\mathcal{F}$ containing $\mathcal{A}$.*

Proof. Let $\mathbb{F} \subset \mathcal{P}(\mathcal{P}(X))$ be the set of all filters $\mathcal{F}$ on X which contain $\mathcal{A}$. We order $\mathbb{F}$ by inclusion. We use Zorn's Lemma to show that $\mathbb{F}$ contains a maximal element. Since $\langle \mathcal{A} \rangle \in \mathbb{F}$, $\mathbb{F} \neq \emptyset$. Let $\mathbb{G}$ be a chain in $\mathbb{F}$. Then

$$\mathcal{H} = \bigcup_{\mathcal{G} \in \mathbb{G}} \mathcal{G}$$

is a filter, as one easily checks. For example, if $F_1, \ldots, F_n \in \mathcal{H}$, there are $\mathcal{G}_1, \ldots, \mathcal{G}_n \in \mathbb{G}$ with $F_j \in \mathcal{G}_j$. Since $\mathbb{G}$ is a chain, we can assume that $\mathcal{G}_n \subset \ldots \subset \mathcal{G}_1$, so that $F_j \in \mathcal{G}_1$. It follows that $\bigcap_{j=1}^n F_j \in \mathcal{G}_1 \subset \mathcal{H}$.

Therefore, the chain $\mathbb{G}$ has an upper bound. Zorn's Lemma guarantees the existence of a maximal element $\mathcal{F} \in \mathbb{F}$, which is our desired ultrafilter. $\square$

A.6. **Application III: Maximal ideals in algebras.**

Definition A.22. *Let $\Bbbk$ be a field. A $\Bbbk$-algebra is a $\Bbbk$-vector space A, together with a bilinear map*

$$A \times A \to A, \ (a, b) \mapsto ab,$$

such that

$$(ab)c = a(bc)$$

for all $a, b, c \in A$. An ideal *in A is a linear subspace $I \subset A$ such that $ax, xa \in I$ holds for all $a \in A$, $x \in I$.*

We say that A is *commutative* if $ab = ba$ holds for all $a, b \in A$, and *unital* if there is $1 \in A$, $1 \neq 0$, such that $a1 = 1a = a$ for all $a \in A$.

Theorem A.23 (Existence of maximal ideals)**.** *Let A be a commutative unital $\Bbbk$-algebra and let $I \subset A$ be a proper ideal, that is $I \neq A$. Then there exists a* maximal ideal *J with $I \subset J \subset A$, that is, J is an ideal, $J \neq A$, and for every other ideal $J \subset K \subset A$, we have $K = J$ or $K = A$.*

Proof. Let $\mathfrak{X}$ be the set of all proper ideals J with $I \subset J$, and order $\mathfrak{X}$ by inclusion. This satisfies the hypotheses of Zorn's lemma, as one verifies easily. Take a maximal element J of $\mathfrak{X}$. $\square$

Appendix B. General topology

B.1. The definition.

Notation B.1. *For a set X, we denote by $\mathcal{P}(X)$ its power set, i.e. the set of all subsets of X.*

Definition B.2. *Let X be set. A* topology *on X is a subset $\mathcal{T} \subset \mathcal{P}(X)$, such that*
1. $\emptyset, X \in \mathcal{T}$,
2. *whenever $U_i \in \mathcal{T}$, $i \in I$, then $\bigcup_{i\in I} U_i \in \mathcal{T}$,*
3. *whenever $U_1, \ldots, U_n \in \mathcal{T}$, then $\bigcap_{i=1}^n U_i \in \mathcal{T}$.*

A topological space *is a pair $(X, \mathcal{T})$, consisting of a set X and a topology $\mathcal{T}$ on X. The elements of $\mathcal{T}$ are called* open *subsets of X.*

Usually, we denote a topological space $(X, \mathcal{T})$ just by the symbol X.

Definition B.3. *Let X be a topological space. A subset $A \subset X$ is* closed *if the complement $A^c = X \setminus A$ is open.*

We note:
1. $\emptyset$ and X are closed,
2. whenever A_i, $i \in I$ are closed subsets, then $\bigcap_{i\in I} U_i$ is closed,
3. whenever $A_1, \ldots, A_n$ are closed, then $\bigcup_{i=1}^n A_i$ is closed.

Definition B.4. *Let X be a topological space and let $x \in X$. A subset $N \subset X$ is a* neighborhood *of x, if $x \in N$ and if there is an open subset U such that $x \in U \subset N$. By $\mathcal{U}(x)$, we denote the set of all neighborhoods of x.*

We observe:

Lemma B.5. *A subset $V \subset X$ is open if and only if V is a neighborhood of each of its points.* □

Definition B.6. *Let X be a topological space and, $Z \subset X$ be a subset and let $x \in X$.*
1. *We say that x is an* interior point *of Z if $Z \in \mathcal{U}(x)$, and we denote by Z° the set of all interior points of Z.*
2. *We say that x is a* limit point *of Z if $U \cap Z \neq \emptyset$ for each $U \in \mathcal{U}(x)$, and we denote by $\overline{Z}$ the set of all limit points of Z, which is also called the* closure of Z in X.

It is clear that
$$Z^\circ \subset Z \subset \overline{Z}.$$
The difference is denoted
$$\partial Z := \overline{Z} \setminus Z^\circ$$
and its points are called *boundary points of Z*.

Lemma B.7. (1) *Z° is open and it is the largest open subset of Z.*
2. $\overline{Z}^c = (Z^c)^\circ$,
3. $\overline{Z^c} = (Z^\circ)^c$,
4. *$\overline{Z}$ is closed, and it is the smallest closed set containing Z.*
5. *Z is open iff $Z = Z^\circ$.*
6. *Z is closed iff $Z = \overline{Z}$.*

Proof. (1) For $x \in Z^\circ$, there is an open U with $x \in U \subset Z$. If $y \in U$, then Z is a neighborhood of y, so $y \in Z^\circ$. This shows that Z° is a neighborhood of x, and therefore Z° is open. If $U \subset Z$ is open and $y \in U$, Z is a neighborhood of y and so $y \in Y^\circ$, so that altogether $U \subset Z^\circ$.

(2): $x \in \overline{Z}^c$ is equivalent to the existence of a neighborhood U of x with $U \cap Z = \emptyset$, i.e. $U \subset Z^c$; and this is equivalent to $x \in (Z^c)^\circ$.

(3) follows from (2) by taking complements:

$$\overline{Z^c} = ((\overline{Z^c})^c)^c \overset{(2)}{=} (((Z^c)^c)^\circ)^c = (Z^\circ)^c.$$

(4): follows from (1) and (2) by taking complements. (5) and (6) are also easy consequences. $\square$

Definition B.8. *Let X and Y be topological spaces and $x \in X$. A map $f : X \to Y$ is* continuous at x *if for each neighborhood U of $f(x)$, the preimage $f^{-1}(U)$ is a neighborhood of x. f is* continuous *if it is continuous at each $x \in X$.*

Lemma B.9. *$f : X \to Y$ is continuous if and only if $f^{-1}(U) \subset X$ is open for each open $U \subset Y$.*

Proof. Let f be continuous, $U \subset Y$ open and $x \in f^{-1}(U)$. Then U is a neighborhood of $f(x)$, hence by continuity $f^{-1}(U)$ is a neighborhood of x, and since this holds for all $x \in f^{-1}(U)$, $f^{-1}(U)$ is open. Vice versa, assume that preimages of open subsets are open and let $x \in X$. Let V be a neighborhood of $f(x)$ and pick an open $x \in U \subset V$. Then $x \in f^{-1}(U)$ is open and so $f^{-1}(V)$ is a neighborhood of x. $\square$

Lemma B.10. *The composition $f \circ g$ of two continuous maps is continuous.* $\square$

Definition B.11. *A* homeomorphism *$f : X \to Y$ between topological spaces is a bijective continuous map such that the inverse map is continuous as well.*

Lemma B.12. *Let X and Y be topological spaces. A map $f : X \to Y$ is continuous if and only if for each $Z \subset X$, we have*

$$f(\overline{Z}) \subset \overline{f(Z)}.$$

Proof. Let $f : X \to Y$ be continuous and let $Z \subset X$. To check that $f(\overline{Z}) \subset \overline{f(Z)}$, we have to check that for each $y \in \overline{Z}$ and each neighborhood $U \subset Y$ of $f(y)$, we have $U \cap f(Z) \neq \emptyset$. But $f^{-1}(U)$ is a neighborhood of y, and therefore $f^{-1}(U) \cap Z \neq \emptyset$, so that there is $z \in f^{-1}(U) \cap Z$. It follows that

$$f(z) \in U \cap f(Z),$$

and in particular $U \cap f(Z)$ is not empty.

For the reverse implication, assume that f is not continuous at $x \in X$ and pick a neighborhood $U \subset Y$ of $f(x)$ such that $f^{-1}(U)$ is not a neighborhood of x. It follows that $x \in \overline{f^{-1}(U)^c} = \overline{f^{-1}(U^c)}$, and that

$$f(x) \in f(\overline{f^{-1}(U^c)}) \subset \overline{f(f^{-1}(U^c))} \subset \overline{U^c}.$$

But that means that U is not a neighborhood of $f(x)$, a contradiction. $\square$

Definition B.13. *Let X be a topological space and let $A \subset X$ be a subset. The* subspace topology *on A is defined as follows. A subset $U \subset A$ is open if and only if it is of the form $U \cap A$ with $U \subset X$ open.*

Definition B.14. *An injective continuous map $f : X \to Y$ is called an* embedding *or a* homeomorphism onto its image *if the map $f : X \to f(X) \subset Y$ is a homeomorphism, where $f(X)$ carries the subspace topology.*

B.2. **Metric spaces.**

Definition B.15. *Let X be a set. A* pseudometric *d on X is a map*

$$d : X \times X \to [0, \infty)$$

such that

$$d(x, x) = 0,$$

$$d(x, y) = d(y, x)$$

(symmetry) and

$$d(x, y) \leq d(x, z) + d(z, y)$$

(triangle inequality) hold for all $x, y, z \in X$. A pseudometric space *(X, d) is a pair, consisting of a set X and a pseudometric d on X.*

A pseudometric d is a metric *if in addition*

$$d(x, y) = 0 \Rightarrow x = y.$$

A metric space *is a pair (X, d), consisting of a set X and a metric on X.*

On the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, a metric is given by the absolute value

$$d(x, y) := |x - y|.$$

More generally, if V is a $\mathbb{K}$-vector space and $\|\_\|$ a seminorm, then

$$d(v, w) := \|v - w\|$$

is a pseudometric, which is a metric if and only if $\|\_\|$ is a norm.

We are usually interested in metric spaces, but occasionally it is handy to have the more general definition at hand. For a pseudometric space (X, d), $x \in X$ and $r > 0$, we denote by

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

the *open ball* of radius r around x, and by

$$\overline{B}_r(x) := \{y \in X | d(x, y) \leq r\}$$

the *closed ball* of radius r around x.

Definition B.16. *Let (X, d) be a pseudometric space and let $U \subset X$. We say that U is* open *if for each $x \in U$, there is $r > 0$ such that $B_r(x) \subset U$. The collection of all open subsets of X is a topology, the* metric topology *induced by d.*

Definition B.17. *A topological space X is* metrizable *if there is a metric d on X which induces the given topology on X.*

It is easily verified that this is indeed a topology. Let us list some easy properties of the metric topology.

Lemma B.18. *Let X be a pseudometric space, $x \in X$ and $r > 0$. Then $B_r(x)$ is open and $\overline{B}_r(x)$ is closed.*

Proof. Let $y \in B_r(x)$. Let $\delta > 0$ be such that $\delta + d(x,y) \leq r$. The triangle inequality shows that $B_\delta(y) \subset B_r(x)$, in other words, that $B_r(x)$ is open. Similarly, if $y \in \overline{B}_r(x)^c$, then $d(x,y) > r$, and there is $\epsilon > 0$ such that $r + \epsilon \leq d(x,y)$. The triangle inequality proves that $B_\epsilon(y) \cap \overline{B}_r(x) = \emptyset$, so that $\overline{B}_r(x)^c$ is a neighborhood of each of its points, therefore open. $\square$

Warning: it is not always true that $\overline{B}_r(x)$ is the closure of $B_r(x)$.

Definition B.19. *A topological space X is called* Hausdorff space *or* T_2-space *if for $x, y \in X$ with $x \neq y$, there are neighborhoods $U \in \mathcal{U}(x)$ and $V \in \mathcal{U}(y)$ with $U \cap V = \emptyset$.*

A metric space is Hausdorff: if $x \neq y$, then $B_{\frac{1}{2}d(x,y)}(x)$ and $B_{\frac{1}{2}d(x,y)}(y)$ are disjoint open neighborhoods of x and y.

For metric spaces, one can express all topological concepts using convergent sequences.

Definition B.20. *A sequence $\mathbb{N} \to X$, $n \mapsto x_n$ in a topological space converges to $x \in X$, in symbol $\lim_{n\to\infty} x_n = x$, if for each neighborhood $U \in \mathcal{U}(x)$, there is $n_0 \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq n_0$.*

In a metric space, this can be expressed more conveniently: $\lim_{n\to\infty} x_n = x$ if and only if for each $\epsilon > 0$, there is $n_0 \in \mathbb{N}$ such that $d(x, x_n) \leq \epsilon$ for all $n \geq n_0$.

Lemma B.21. *Let X be a metric space and $Z \subset X$.*

1. *A point $x \in X$ lies in $\overline{Z}$ if and only if there is a sequence $x_n \in Z$ with $x = \lim_n x_n$.*
2. *The set $Z \subset X$ is closed if and only if for each sequence $x_n \in Z$ which converges to $x \in X$, we have $x \in Z$.*

Proof. 1: if $x \in \overline{Z}$, then $B_{\frac{1}{n}}(x) \cap Z \neq \emptyset$, and we can therefore find points $x_n \in B_{\frac{1}{n}}(x) \cap Z$. The sequence x_n converges to x. Vice versa, if $x_n \in Z$ and $\lim_{n\to\infty} x_n = x$, then for each neighborhood U of x, we have $x_n \in U$ for sufficiently large n, in particular $U \cap Z \neq \emptyset$.

2: Assume that Z is closed, $x_n \in Z$ and $\lim_n x_n = x \in X$. Then by part (1), $x \in \overline{Z} = Z$. Vice versa, assume that limits of sequences in Z lie in Z and let $x \in \overline{Z}$. Then by (1), there is a sequence $x_n \in Z$ which converges to x, proving that $x \in Z$, so $Z = \overline{Z}$. $\square$

Theorem B.22. *Let X and Y be metric spaces and let $f: X \to Y$ be a map. The following are equivalent:*

1. *f is continuous.*
2. *For each $x \in X$ and each $\epsilon > 0$, there is $\delta > 0$ so that $f(B_\delta(x)) \subset B_\epsilon(f(x))$.*
3. *f is sequentially continuous. In other words, if x_n is a convergent sequence in X, then $f(\lim_n x_n) = \lim_n f(x_n)$.*

Proof. $1 \Rightarrow 2$: $B_\epsilon(f(x))$ is open, and since f is continuous, $f^{-1}(B_\epsilon(f(x)))$ is open. Therefore, there is $\delta > 0$ with $B_\delta(x) \subset f^{-1}(B_\epsilon(f(x)))$, which means $f(B_\delta(x)) \subset B_\epsilon(f(x))$.

$2 \Rightarrow 3$: is straightforward.

$3 \Rightarrow 1$: we use Lemma B.12. Let $Z \subset X$; we have to prove that

$$f(\overline{Z}) \subset \overline{f(Z)}.$$

For $x \in \overline{Z}$, pick a sequence $x_n \in Z$ with $\lim_n x_n = x$. Then $f(x) = \lim_n f(x_n)$, hence $f(x) \in \overline{f(Z)}$. □

Definition B.23. *A metric space X is* complete *if each Cauchy sequence in X converges. More precisely: if (x_n) is a Cauchy sequence in X (that is, for each $\epsilon > 0$, there is n_0 such that for all $m, n \geq n_0$, we have $d(x_n, x_m) \leq \epsilon$), then there exists $x \in X$ with $x = \lim_n x_n$.*

It is of course known that $\mathbb{R}$, $\mathbb{C}$ are complete.

B.3. **Bases and subbases of a topology.**

Definition B.24 (Neighborhood basis)**.** *Let X be a topological space and $x \in X$. A* neighborhood basis *of x is a set $\mathcal{U}$ of neighborhoods of x such that for each neighborhood $U \in \mathcal{U}(x)$, there is $V \in \mathcal{U}$ with $V \subset U$.*

Definition B.25 (Basis and subbasis)**.** *Let $(X, \mathcal{T})$ be a topological space. A* basis *for the topology of X is a subset $\mathcal{B} \subset \mathcal{T}$ such that each element U of $\mathcal{T}$ can be written as the union of elements of $\mathcal{B}$.*

A subbasis *for $\mathcal{T}$ is a subset $\mathcal{A} \subset \mathcal{T}$ containing X such that the set*

$$\{O \subset X | \exists U_1, \dots, U_n \in \mathcal{A} : O = U_1 \cap \dots \cap U_n\}$$

is a basis for $\mathcal{T}$.

Definition B.26. *A subset $Z \subset X$ of a topological space is called* dense *if $\overline{Z} = X$.*

Definition B.27 (Countability axioms)**.** *Let X be a topological space.*

(1) *X is* first countable *if each $x \in X$ has a countable neighborhood basis.*
(2) *X is* second countable *if the topology of X has a countable basis.*
(3) *X is* separable *if X has a countable dense subset.*

Examples B.28. (1) *A metric space X is first countable: the set $\{B_{\frac{1}{n}}(x) | n \in \mathbb{N}\}$ is a neighborhood basis for x.*
(2) *A second countable space is first countable: if $\mathcal{B}$ is a countable basis, then $\{U \in \mathcal{B} | x \in U\}$ is a neighborhood basis for x.*
(3) *A second countable space is separable: let $\mathcal{B}$ be a countable basis. Without loss of generality, each $U \in \mathcal{B}$ is nonempty. Pick $x_U \in U$, for each $U \in \mathcal{B}$. Then the set $\{x_U | U \in \mathcal{B}\} \subset X$ is dense.*
(4) *A separable metric space X is second countable: assume that $Z = \{z_n | n \in \mathbb{N}\} \subset X$ is a countable dense subset. We claim that $\mathcal{U} := \{B_{\frac{1}{m}}(z_n) | m, n \in \mathbb{N}\}$ is a basis for the topology of X. It suffices to show that if $U \subset X$ is open and $x \in U$, then there are m, n with $x \in B_{\frac{1}{m}}(z_n)$.*

    There is m such that $B_{\frac{2}{m}}(x) \subset U$, and there is n such that $d(x, z_n) < \frac{1}{2m}$. Then $x \in B_{\frac{1}{m}}(z_n) \subset U$.
(5) *$\mathbb{R}^n$ is separable: $\mathbb{Q}^n \subset \mathbb{R}^n$ is dense.*

B.4. **The topology generated by a set of subsets.**

Definition B.29. *Let X be a set and let $\mathcal{T}_0$, $\mathcal{T}_1$ be two topologies on X. We say that $\mathcal{T}_1$ is* coarser than *$\mathcal{T}_0$ - or that $\mathcal{T}_0$ is* finer than *$\mathcal{T}_1$ if $\mathcal{T}_1 \subset \mathcal{T}_0$, or equivalently, if $\mathrm{id} : (X, \mathcal{T}_0) \to (X, \mathcal{T}_1)$ is continuous.*

On each set X, there is a coarsest topology, which consists only of the two sets $\emptyset$ and X. This is the *trivial topology*. The finest topology on X is the *discrete topology* which consists of all subsets of X. These two topologies are rather uninteresting. More interesting is the coarsest topology which contains a certain supply of subsets of X.

Lemma B.30. *Let X be a set and let $\mathcal{A} \subset \mathcal{P}(X)$.*

1. *There is a uniquely determined* coarsest topology *$\mathcal{T}$ with $\mathcal{A} \subset \mathcal{T}$.*
2. *$\mathcal{T}$ consists of X, and all sets U with the following property: for each $x \in U$, there are $U_1, \ldots, U_n \in \mathcal{A}$ such that*
$$x \in \bigcap_{j=1}^{n} U_j \subset U.$$
3. *Let Z be a topological space and let $f : Z \to X$ be a map. Then $f : Z \to (X, \mathcal{T})$ is continuous if and only if $f^{-1}(U) \subset Z$ is open for each $U \in \mathcal{A}$.*
4. *$\mathcal{A}$ is a subbasis for $\mathcal{T}$ if $\bigcup_{U \in \mathcal{A}} U = X$.*
5. *$\mathcal{A}$ is a basis for $\mathcal{T}$ if $\bigcup_{U \in \mathcal{A}} U = X$ and if for $U, V \in \mathcal{A}$, and each $x \in U \cap V$, there is $W \in \mathcal{A}$ with $x \in W \subset U \cap V$.*

Proof. 1: Let $\mathbb{T}$ be the set of all topologies on X (this is a subset of $\mathcal{P}(\mathcal{P}(X))$). If $\mathcal{T}_i$, $i \in I$, are topologies on X, then so is the intersection
$$\bigcap_{i \in I} \mathcal{T}_i \subset \mathcal{P}(X).$$
We *must* define
$$\mathcal{T} := \bigcap_{\mathcal{S} \in \mathbb{T}, \mathcal{A} \subset \mathcal{S}} \mathcal{S}.$$
This is the desired topology.

2: let $\mathcal{S}$ be the set of all sets U with the stated property, together with X. It is easy to check that $\mathcal{S}$ is a topology and that $\mathcal{A} \subset \mathcal{S}$. Therefore $\mathcal{T} \subset \mathcal{S}$. On the other hand, if $\mathcal{U}$ is a topology containing $\mathcal{A}$ and $U \in \mathcal{S}$, then it follows from Lemma B.5 that $U \in \mathcal{U}$. Therefore $\mathcal{S} \subset \mathcal{U}$, which in particular implies $\mathcal{S} \subset \mathcal{T}$.

3: since $\mathcal{A} \subset \mathcal{T}$, any continuous map $f : Z \to (X, \mathcal{T})$ has the property that $f^{-1}(U) \subset Z$ is open whenever $U \in \mathcal{A}$. On the other hand, if $f^{-1}(U)$ is open for each $U \in \mathcal{A}$, then $f^{-1}(U_1 \cap \ldots \cap U_n) = f^{-1}(U_1) \cap \ldots \cap f^{-1}(U_n)$ is open when $U_1, \ldots, U_n \in \mathcal{A}$. Any set in $\mathcal{T}$ can be written as the union of such finite intersections, and so its preimage is open in Z, so f is continuous.

4, 5: This is straightforward from the definition of a (sub)basis and the explicit description of $\mathcal{T}$ given in (2). □

B.5. **The induced topology and products.**

Definition B.31. *Let X be a set, let Y_i, $i \in I$, be topological spaces and let $\mathcal{F} := \{f_i : X \to Y_i | i \in I\}$ be a family of maps. The topology on X induced by $\mathcal{F}$ is the coarsest topology on X which contains all the sets $f_j^{-1}(V_j)$, where $V_j \subset Y_j$ is open (see Lemma B.30).*

Using Lemma B.30, we see that $U \subset X$ is open in the topology induced by $\mathcal{F}$ if and only if for each $x \in U$, there is a finite subset $J \subset I$ and open $V_j \subset Y$, $j \in J$, such that
$$x \in \bigcap_{j \in J} f_j^{-1}(V_j) \subset U.$$

Lemma B.32. *The induced topology has the following properties.*

(1) *The maps $f_i : X \to Y_i$ are continuous.*
(2) *The induced topology is the coarsest topology with this property.*
(3) *If Z is a further topological space and $g : Z \to X$ is a map, then g is continuous if and only if $f_i \circ g$ is continuous for each $i \in I$.*

Proof. 1: By the concrete description of the induced topology, it is clear that all maps f_i are continuous.

3: By Lemma B.30, g is continuous if and only if $g^{-1}(f_i^{-1}(V))$ is open for each $i \in I$ and each open subset $V \subset Y_i$. Since $g^{-1}(f_i^{-1}(V)) = (f_i \circ g)^{-1}(V)$, this condition is equivalent to the continuity of $f_i \circ g$.

2: denote the induced topology by $\mathcal{T}$, and let $\mathcal{S}$ be another topology on X so that each $f_i : (X;\mathcal{S}) \to Y_i$ is continuous. By (3), the map $\mathrm{id} : (X,\mathcal{S}) \to (X,\mathcal{T})$ is continuous, but that means that $\mathcal{T}$ is coarser than $\mathcal{S}$. $\square$

The most interesting special case of this construction is the *product topology.* Let X_i, $i \in I$, be topological spaces. We denote by $p_j : \prod_{i\in I} X_i \to X_j$ be the projection map from the cartesian product. We can view $\prod_{i\in I} X_i$ as the set of all families $(x_i)_{i\in I}$ with $x_i \in X_i$, and the projection map is given by $p_j((x_i)_{i\in I}) := x_j$.

Definition B.33. *The* product topology *on $\prod_{i\in I} X_i$ is the topology induced by the family $\{p_j : \prod_{i\in I} X_i \to X_j\}$ of all projection maps.*

One can easily prove that in the case of a finite product $X_1 \times \ldots \times X_n$, the product topology has a basis which consists of all the sets $U_1 \times \ldots \times U_n$, with $U_j \subset X_j$ open. The general case is a little less intuitive: a basis is given by all the sets of the form $p_{i_1}^{-1}(U_{i_1}) \cap \ldots \cap p_{i_n}^{-1}(U_{i_n})$.

B.6. **The product topology for metric spaces.**

Lemma B.34. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then the metric*

$$d_{X\times Y}((x_0, y_0),(x_1, y_1)) := \max\{d_X(x_0, x_1), d_Y(y_0, y_1)\}$$

induces the product topology on $X \times Y$.

This is an exercise. More interesting is the case of countable products.

Proposition B.35. *Let $(X_n, d_n)_{n\in\mathbb{N}}$ be metric spaces. On the product $\prod_{n=1}^{\infty} X_n$, the formula*

$$d_{\prod}((x_n)_n,(y_n)_n) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}$$

defines a metric, and $d_{\prod}$ induces the product topology.

Proof. The only property of a metric which is nontrivial is the triangle inequality. To prove this, let $0 \le a, b, c$ be real numbers with $a \le b + c$. Then

$$\frac{a}{1+a} \le \frac{b}{1+b} + \frac{c}{1+c}. \tag{B.36}$$

This is because $f(t) := \frac{1}{1+t}$ is increasing, so that $\frac{a}{1+a} \le \frac{b+c}{1+b+c}$, and the easily verified inequality $\frac{b+c}{1+b+c} \le \frac{b}{1+b} + \frac{c}{1+c}$. The inequality (B.36) immediately implies the triangle inequality for $d_{\prod}$.

The projection map

$$p_m : (\prod_{n=1}^{\infty} X_n, d_{\prod}) \to X_m$$

is continuous (since the function f is continuous at 0). Therefore, by Lemma B.32, the identity

$$(\prod_{n=1}^{\infty} X_n, d_{\prod}) \to (\prod_{n=1}^{\infty} X_n, \mathcal{T})$$

is continuous, where $\mathcal{T}$ denotes the product topology. Hence the topology induced by $d_{\prod}$ is finer than the product topology. We also need to show that it is coarser than the product topology, in other words: each subset which is open in the metric topology is open in the product topology. If we can show that each ball $B_r((x_n)_n)$ contains a neighborhood of $(x_n)_n$ in the product topology, we are done with that.

Choose n_0 with $\sum_{n=n_0+1}^{\infty} \frac{1}{2^n} \leq \frac{r}{2}$. Then there are $\delta_1, \ldots, \delta_{n_0} > 0$, so that

$$B_{\delta_1}(x_1) \times \ldots \times B_{\delta_{n_0}}(x_{n_0}) \times \prod_{n=n_0+1} X_n \subset B_r((x_n)_n).$$

This proves all claims. □

B.7. **Net convergence.** Lemma B.21 and Theorem B.22 show that the topology of metric spaces is adequately reflected by convergence of sequences, which is very convenient for all purposes of analysis. In general topological spaces, this does not need to hold, and a more general notion of convergence is necessary. There are two such concepts: nets and filters. We develop the language of nets; for convergence of filters, we refer to the literature.

Definition B.37. *A* directed set *is a set I, together with a binary relation $\leq$ on I such that*

1. $x \leq x$,
2. $x \leq y$, $y \leq z \Rightarrow x \leq z$,
3. *for all $x, y \in I$, there is $z \in I$ such that $x \leq z$ and $y \leq z$.*

A net *in a set X consists of a directed set I, together with a map $N : I \to X$.*

If $N : I \to X$ is a net and $Z \subset X$, we say that N is eventually in Z *if there is $i \in I$ such that $N(j) \in Z$ for all $j \geq i$. We say that N is* frequently in Z *if for each $i \in I$, there is $j \geq i$ with $N(j) \in Z$.*

Definition B.38. *Let X be a topological space and let $N : I \to X$ be a net. We say that N converges to $x \in X$, in symbols $\lim_i N(i) = x$ if for each neighborhood U of x, N is eventually in U.*

Remark B.39. *If X is Hausdorff, then a net $N : I \to X$ has at most one limit.*

Examples B.40.
1. *$\mathbb{N}$ with the usual order relation is a directed set. A net $N : \mathbb{N} \to X$ converges to x if and only if $\lim_{n\to\infty} N(n) = x$.*
2. *The set of neighborhoods $\mathcal{U}(x)$ of $x \in X$ is a directed set with the relation $U \leq V :\Leftrightarrow V \subset U$. Picking a point $x_U \in U$ for each $U \in \mathcal{U}(x)$ gives a net $N : \mathcal{U}(x) \to X$, $N(U) := x_U$, and this net converges to x.*
3. *The Riemann integral can be formulated using nets: the set $\mathcal{Z}$ of all partitions of the interval $[a, b]$ (that is, just finite subsets of (a, b)) is a directed set, with the relation $Z_0 \leq Z_1 \Leftrightarrow Z_0 \subset Z_1$. For a bounded function $f : [a, b] \to \mathbb{R}$, we define the upper sum $U(f, Z)$ and the lower sum $L(f, Z)$.*

Then $N_0, N_1 : \mathcal{Z} \to \mathbb{R}$, $N_0(Z) = L(Z, f)$ and $N_1(Z) = U(T, f)$ are two nets, and f is Riemann integrable if and only if both nets converge to the same limit.

Lemma B.21 has the following generalization in terms of nets, valid for any space.

Lemma B.41. *Let $Z \subset X$ be a subset of a topological space and $x \in X$. Then $x \in \overline{Z}$ if and only if there is a net $N : I \to Z \subset X$ with $\lim_i N = x$.*

Proof. If $N : I \to Z \subset X$ is a net with $x = \lim_i N(i)$, then each neighborhood U of x contains some $N(i)$, and therefore $U \cap Z \neq \emptyset$, so that $x \in \overline{Z}$.

On the other hand, if $x \in \overline{Z}$, then each neighborhood $U \in \mathcal{U}(x)$ contains a point $x_U \in U \cap Z$. We define a map[6]

$$N : \mathcal{U}(x) \to Z,\ U \mapsto x_U.$$

On $\mathcal{U}(x)$, we introduce the relation $U \leq V :\Rightarrow V \subset U$, so that N is a net. Clearly $\lim_U x_U = x$. □

Theorem B.42. *A map $f : X \to Y$ of topological spaces is continuous if and only if for each net $N : I \to X$ which converges to some $x \in X$, the net $f \circ N : I \to Y$ converges to $f(x)$.*

Proof. Let $f : X \to Y$ be continuous and let $N : I \to X$ be a net with limit x. Let $U \subset Y$ be a neighborhood of $f(x)$. Then $f^{-1}(U)$ is a neighborhood of x, and so N is eventually in $f^{-1}(U)$. But this means that $f \circ N$ is eventually in U, and hence $\lim_i f(N(i)) = f(x) = f(\lim_i N(i))$.

Vice versa, let $Z \subset X$ be a subset. We have to prove that $f(\overline{Z}) \subset \overline{f(Z)}$, by Lemma B.12. For $x \in \overline{Z}$, pick a net $N : I \to Z$ with $\lim_i N(i) = x$, using Lemma B.41. It follows that

$$f(x) = f(\lim_i N(i)) = \lim_i f(N(i)) \in \overline{f(Z)},$$

using Lemma B.41 again. □

There is a notion of a *subnet* of a net, whose precise formulation is quite abstract.

Definition B.43. *Let I, J be directed sets and let $\varphi : I \to J$ be a map. We say that φ is* cofinal *if for all $j \in J$, there is $i_0 \in I$, such that for all $i \geq i_0$, we have $\varphi(i) \geq j$.*

A subnet *of a net $N \to X$ is the compostion $N \circ \varphi : J \to X$ with a cofinal map $\varphi : J \to I$.*

B.8. **Compactness.**

Definition B.44. *An* open cover *of $(X, \mathcal{T})$ is a subset $\mathcal{U} \subset \mathcal{T}$ such that $\bigcup_{U \in \mathcal{U}} U = X$. A* subcover *of $\mathcal{U}$ is a subset $\mathcal{V} \subset \mathcal{U}$ which is also a cover.*

A topological space X is compact *if each open cover of X has a finite subcover.*

There is a dual formulation of the definition using closed subsets.

Definition B.45. *A subset $\mathcal{A} \subset \mathcal{P}(X)$ has the* finite intersection property *if the intersection $A_1 \cap \ldots \cap A_n$ of finitely many elements of $\mathcal{A}$ is nonempty.*

Theorem B.46. *Let X be a topological space. The following are equivalent:*

[6] The way we wrote this requires the axiom of choice, but that can be circumvented.

(1) *X is compact,*
(2) *Let $\mathcal{A} \subset \mathcal{P}(X)$ be a family of closed subsets with the finite intersection property. Then*
$$\bigcap_{A \in \mathcal{A}} A \neq \emptyset$$

Proof. For a family $\mathcal{A} \subset \mathcal{P}(X)$ of closed subsets, we let
$$\mathcal{A}^\sharp := \{A^c | A \in \mathcal{A}\}$$
which is a family of open subsets. For a family $\mathcal{O}$ of open subsets, we let
$$\mathcal{O}^\flat := \{U^c | U \in \mathcal{O}\},$$
a family of closed subsets.

$1 \Rightarrow 2$: Let X be compact and let $\mathcal{A}$ be a family of closed subsets with the finite intersection property. Then no finite subset of $\mathcal{A}^\sharp$ covers X. As X is compact, and all elements of $\mathcal{A}^\sharp$ are open, this implies that $\mathcal{A}^\sharp$ is not a cover of X. In other words
$$\emptyset \neq (\bigcup_{A \in \mathcal{A}} A^c)^c = \bigcap_{A \in \mathcal{A}} (A^c)^c = \bigcap_{A \in \mathcal{A}} A.$$

$2 \Rightarrow 1$: Let $\mathcal{O}$ be an open cover of X. Then
$$\bigcap_{U \in \mathcal{O}} U^c = \emptyset,$$
hence $\mathcal{O}^\flat$ does not have the finite intersection property. By the hypothesis on X, there are $U_1, \ldots, U_n \in \mathcal{O}$ with $\bigcap_{i=1}^n U_i^c = \emptyset$. Then $(U_i)_{i=1}^n$ is a finite subcover of $\mathcal{O}$. $\square$

Proposition B.47. *Let X be compact and let $f : X \to \mathbb{R}$ be continuous. Then f attains maximum and minimum.*

Among the many useful properties of compact spaces, we single out one, which is very often used.

Proof. For $r < R := \sup_{x \in X} f(x)$, the set $A_r := \{x | f(x) \geq r\}$ is closed, and nonempty. The system
$$\mathcal{A} = \{A_r | r < R\}$$
of closed subsets has the finite intersection property. Therefore, there is $x \in \bigcap_{r<R} A_r$, and f attains its maximum at x. Similarly, one proves the existence of a minimum. $\square$

One of the most important theorems about compactness is *Tychonov's Theorem* which asserts that the product $\prod_{i \in I} X_i$ of compact spaces is compact. We prove this later on; the case of finite products is easier. We start with a helpful remark.

Remark B.48. *Let X be a topological space and let $\mathcal{B}$ be a basis for the topology. Then X is compact if and only if each open cover $\mathcal{U} \subset \mathcal{B}$ has a finite subcover. In other words, it is enough to prove that covers by sets of a basis have finite subcovers. This is because $\mathcal{A} = \{V \in \mathcal{B} | \exists U \in \mathcal{U} : V \subset U\}$ is an open cover, and if $\mathcal{A}$ has a finite subcover, then so does $\mathcal{U}$.*

Theorem B.49 (Baby Tychonov)**.** *Let X and Y be compact. Then $X \times Y$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $X \times Y$. Since the sets of the form $U \times V$, $U \subset X$, $V \subset Y$ open, form a basis of the topology of $X \times Y$, we can assume that $\mathcal{U} = \{U_i \times V_i | i \in I\}$, by Remark B.48.

For each $x \in X$, the sets $(U_i \times V_i) \cap \{x\} \times Y$ form a cover of the compact space $\{x\} \times Y \cong Y$, and so there is a finite subset $I_x \subset I$ such that $\{x\} \times Y \subset \bigcap_{i \in I_x} U_i \times V_i$. The set $W_x := \bigcap_{i \in I_x} U_i \subset X$ is an open neighborhood of x, and as X is compact, we find $x_1, \ldots, x_n \in X$ with $W_{x_1} \cup \ldots \cup W_{x_n} = X$. Then

$$\{U_i \times V_i | i \in I_{x_1} \cup \ldots \cup I_{x_n}\} \subset \mathcal{U}$$

is a finite subcover of $\mathcal{U}$. □

B.9. **Compactness and Hausdorff property.** Spaces which are both compact and Hausdorff are extremely well-behaved.

Lemma B.50. *Let $f : X \to Y$ be a map of topological spaces.*

(1) *If f is injective and Y is Hausdorff, then X is Hausdorff.*
(2) *If f is surjective and X is compact, then Y is compact.*

Proof. (1): Let $x_0 \neq x_1 \in X$. Then $f(x_0) \neq f(x_1)$, and there are open disjoint neighborhoods U_0 of x_0 and U_1 of x_1. Then $f-1(U_i)$ is an open neighborhood of x_i and $f^{-1}(U_0) \cap f^{-1}(U_1) = \emptyset$.

(2): Let $\mathcal{U}$ be an open cover of Y. Then $f^{-1}(\mathcal{U}) = \{f^{-1}(U) | U \in \mathcal{U}\}$ is an open cover of X, which has a finite subcover $\{f^{-1}(U_1), \ldots, f^{-1}(U_n)\}$. Since f is surjective, we have $U_i = f(f^{-1}(U_i))$ and

$$Y = f(X) = f(f^{-1}(U_1)) \cup \ldots f(f^{-1}(U_n)) = U_1 \cup \ldots \cup U_n,$$

so that $\mathcal{U}$ has a finite subcover. □

Lemma B.51. *Let X be a topological space and $Z \subset X$ a subspace.*

(1) *If X is compact and Z is closed, then Z is compact.*
(2) *If X is Hausdorff and Z is compact, then Z is closed.*

Proof. (1): Let $\mathcal{U}$ be an open cover of Z. For each $U \in \mathcal{U}$, there is an open $V_U \subset X$ with $V_U \cap Z = U$. Then

$$\mathcal{V} := \{V_U | U \in \mathcal{U}\} \cup \{X \setminus Z\}$$

is an open cover which has a finite subcover $\{V_{U_1}, \ldots, V_{U_n}, X \setminus Z\}$. Then $\{U_1, \ldots, U_n\}$ is a finite subcover of $\mathcal{U}$.

(2): Let $x \in Z^c$ be a point. For each $y \in Z$, there are open neighborhoods $y \in U_y$, $x \in V_y$ which are disjoint. Since Z is compact, there are $y_1, \ldots, y_n \in Z$ such that

$$Z \subset U_{y_1} \cup \ldots \cup U_{y_n} =: U.$$

Then $x \in V_{y_1} \cap \ldots \cap V_{y_n} =: V$, V is open and $V \cap U = \emptyset$. It follows that $V \subset Z^c$, and since x was arbitrary, that Z^c is open, hence Z is closed. □

Theorem B.52. *Let $f : X \to Y$ be bijective, X compact and Y Hausdorff. Then f is a homeomorphism.*

Proof. By Lemma B.50, X is Hausdorff and Y is compact. Let $g : Y \to X$ be the inverse map to f. In order to prove that g is continuous, it suffices to show that $f(A) \subset Y$ is closed for each closed $A \subset X$.

By Lemma B.51, A is compact, hence by Lemma B.50, $f(A)$ is compact, hence $f(A) \subset Y$ is closed by Lemma B.51. □

B.10. **Compactness and sequential compactness.**

Definition B.53. *A space X is* sequentially compact *if each sequence in X has a convergent subsequence.*

Theorem B.54. (1) *Let X be first countable and compact. Then X is sequentially compact.*
(2) *Let X be second countable and sequentially compact. Then X is compact.*

We will soon prove that for *metric spaces*, compactness and sequential compactness are equivalent.

Proof. (1): Let (x_n) be a sequence in X and let

$$A_m := \overline{\{x_n | n \geq m\}} \subset X.$$

Then $A_m \neq \emptyset$ is closed, and we have

$$A_1 \supset A_2 \supset \ldots$$

Since X is compact, $A := \bigcap_{m=1}^{\infty} A_m \neq \emptyset$, and we pick $x \in A$. Let $(V_n)_n$ be a neighborhood basis of x. The sets $U_n := V_1 \cap \ldots \cap V_n$ form a neighborhood basis of x with the extra property that $U_1 \supset U_2 \supset \ldots$. Then $U_n \cap A_m \neq \emptyset$ for all m and n, and so there is $k \geq m$ with

$$x_k \in U_n.$$

We can inductively pick a sequence x_{n_k} with $x_{n_k} \in U_k$, and since $(U_n)_n$ is a neighborhood *basis*, $\lim_{k\to\infty} x_{n_k} = x$.

(2): Let $\mathcal{B}$ be a countable basis for the topology of X and let $\mathcal{U}$ be an open cover of X. The set

$$\mathcal{B}' := \{O \in \mathcal{B} | \exists U \in \mathcal{U} : O \subset U\} \subset \mathcal{B}$$

is an countable open cover of X, and it is enough to cover X by finitely elements from $\mathcal{B}'$. This argument shows that it is enough to show that each *countable* open cover of X has a finite subcover.

So let $\mathcal{U} = \{U_n | n \in \mathbb{N}\}$ be an open cover of X. Assume that $\mathcal{U}$ has no finite subcover. This means that $V_n := \bigcup_{k=1}^{n} U_k \neq X$. Hence there is a sequence $x_n \in X$, such that $x_n \notin V_n$. There is a subsequence x_{n_m} with $\lim_m x_{n_m} = x \in X$. Then $x \in U_N$ for some N, and so $x_{n_m} \in U_N$ for $m \geq m_0$. This is a contradiction when $n_m > N$. □

Corollary B.55 (Heine-Borel Theorem)**.** *A subset $X \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded (here the topology on $\mathbb{R}^n$ is the product topology, which is the same as the one induced by the norm $\|x\|_{\ell^\infty} := \max\{|x_1|, \ldots, |x_n|\}$.*

Proof. If $X \subset \mathbb{R}^n$ is compact, it must be closed because $\mathbb{R}^n$ is Hausdorff, by Lemma B.51. The function $f = \|\_\|_{\ell^\infty} : \mathbb{R}^n \to \mathbb{R}$ is continuous, and by Proposition B.47, f attains its maximum on X and is in particular bounded.

If $X \subset \mathbb{R}^n$ is bounded and closed, it is a closed subset of $[-R, R]^n$ for some $R \geq 0$. Using Lemma B.51, it suffices to prove that $[-R, R]^n$ is compact, and by Theorem B.49, it suffices to do so in the case $n = 1$. The interval $[-R, R]$ is sequentially compact by the Bolzano-Weierstrass Theorem, and therefore compact by Theorem B.54, because $\mathbb{R}$ is second countable because it is a separable metric space. □

B.11. **Tychonov's Theorem.**

Theorem B.56 (Tychonov)**.** *Let X_i, $i \in I$, be compact. Then the product*

$$X = \prod_{i \in I} X_i$$

is compact.

Proof. We denote the projections by $p_i : X \to X_i$. Let $\mathcal{A} \subset \mathcal{P}(X)$ be a family of closed subsets with the finite intersection property. By the Ultrafilter Lemma A.21, there is an ultrafilter $\mathcal{F}$ on X which contains $\mathcal{A}$. For each $i \in I$, we let

$$\mathcal{G}_i := \{\overline{p_i(S)} | S \in \mathcal{F} \subset \mathcal{P}(X_i).$$

This is a family of closed subsets of X_i, and it has the finite intersection property, because

$$\overline{p_i(S_1)} \cap \ldots \cap \overline{p_i(S_n)} \supset p_i(S_1) \cap \ldots \cap p_i(S_n) \supset p_i(S_1 \cap \ldots \cap S_n) \neq \emptyset.$$

Since X_i is compact, the set

$$Z_i := \bigcap_{S \in \mathcal{F}} \overline{p_i(S)}$$

is not empty. Using the Axiom of Choice, the product

$$Z = \prod_{i \in I} Z_i \subset X$$

is not empty, and we pick an element

$$x = (x_i)_{i \in I} \in Z$$

(note that $x_i \in Z_i$). We claim that

$$x \in \overline{S}$$

for each $S \in \mathcal{F}$. From this, it will follow that

$$x \in \bigcap_{S \in \mathcal{F}} \overline{S} \subset \bigcap_{A \in \mathcal{A}} \overline{A} = \bigcap_{A \in \mathcal{A}} A,$$

and in particular that the intersection of $\mathcal{A}$ is not empty, which shows that X is compact.

Let $U \subset X$ be a neighborhood of x. By the definition of the product topology, there is a finite subset $J \subset I$ and open neighborhoods U_j of x_j, $j \in J$, such that

$$x \in \bigcap_{j \in J} p_j^{-1}(U_j) \subset U.$$

Since $x_j \in \overline{p_j(S)}$, it follows that $U_j \cap p_j(S) \neq \emptyset$ for each $S \in \mathcal{F}$, and this implies

$$p_j^{-1}(U_j) \cap S \neq \emptyset$$

for all $S \in \mathcal{F}$. Since $\mathcal{F}$ is an ultrafilter, it follows Lemma A.19 that

$$p_j^{-1}(U_j) \in \mathcal{F}$$

and hence that

$$U \in \mathcal{F},$$

so that

$$U \cap S \neq \emptyset$$

for each $S \in \mathcal{F}$. As U was an arbitrary neighborhood of x, we conclude that

$$x \in \overline{S}$$

for each $S \in \mathcal{F}$, as claimed. □

Corollary B.57. *Let X_n, $n \in \mathbb{N}$, be compact metric spaces. Then $\prod_{n=1}^\infty X_n$ is sequentially compact.*

Proof. By Tychonov's theorem, $\prod_{n=1}^\infty X_n$ is compact. Proposition B.35 shows that $\prod_{n=1}^\infty X_n$ is a metric space, hence first countable, and hence sequentially compact by Theorem B.54. □

B.12. **Metric spaces: completeness and compactness.**

Definition B.58. *A metric space X is* totally bounded *if for each $\epsilon > 0$, there are finitely many points $x_1, \ldots, x_r \in X$ such that $\bigcup_{j=1}^r B_\epsilon(x_j) = X$.*

Theorem B.59. *Let X be a metric space. The following are equivalent:*

(1) *X is compact.*
(2) *X is sequentially compact.*
(3) *X is complete and totally bounded.*

Proof. $1 \Rightarrow 2$: since metric spaces are first countable, this follows from Theorem B.54.

$2 \Rightarrow 3$: Let $(x_n)_n$ be a Cauchy sequence in X and let $(x_{n_m})_m$ be a convergent sequence with $\lim_{m\to\infty} x_{n_m} = x$. The sequence $d(x_m, x_{n_m})$ in $\mathbb{R}$ converges to 0, and so

$$\lim_{m\to\infty} (d(x, x_m) \leq d(x, x_{n_m}) + d(x_{n_m}, x_m)) = 0.$$

Hence X is complete. If X is not totally bounded, there is $\delta > 0$, so that for arbitrarily chosen $x_1, \ldots, x_n \in X$, we have

$$X \setminus \bigcup_{j=1}^n B_\delta(x_j) \neq \emptyset.$$

Hence there is a sequence x_n with $d(x_n, x_m) \geq \delta$ whenever $m \neq n$. Each subsequence has the same property and hence cannot converge, so that X is not sequentially compact.

$3 \Rightarrow 2$: let $(x_n)_n$ be a sequence in X. We claim that there is a Cauchy sequence $(y_m)_m$ in X and infinite subsets $\mathbb{N} \supset J_1 \supset J_2 \supset \ldots$, such that $d(x_n, y_m) \leq \frac{1}{2^m}$ whenever $m \in J_n$.

Since X is totally bounded, we can cover X by finitely many balls of radius $\frac{1}{2}$. One of those balls, say $B_{\frac{1}{2}}(y_1)$ contains infinitely many terms of $(x_n)_n$, and we let J_1 be the set of those indices. The ball $B_{\frac{1}{2}}(y_1)$ is totally bounded. We can repeat the argument and find $y_2 \in B_{\frac{1}{2}}(y_1)$, such that $B_{\frac{1}{4}}(y_2)$ contains infinitely many of the x_m's with $m \in J_1$, and we let $J_2 \subset J_1$ be the set of those indices. We pick a subsequence x_{n_m} with $n_m \in J_m$.

The sequence (y_m) is Cauchy and hence has a limit $y := \lim_m y_m$. Since $d(y_m, x_{n_m}) \leq \frac{1}{2^m}$, we conclude that $\lim_m x_{n_m} = y$. Therefore (x_n) has a convergent subsequence.

$3 \Rightarrow 1$: For each n, cover X by finitely many balls $B_{\frac{1}{2^n}}(x_{n,1}), \ldots, B_{\frac{1}{2^n}}(x_{n,r_n})$. Then the set $\{x_{n,j} | n \in \mathbb{N}, 1 \leq j \leq r_n\} \subset X$ is dense. Therefore X is separable, and hence second countable. We have already established that X is sequentially compact, and Theorem B.54 finishes the proof. $\square$

In the proof of the implication $3 \Rightarrow 1$, we have seen:

Corollary B.60. *A compact metric space is second countable.*

Another important property of compact metric spaces is the

Lemma B.61 (Lebesgue-Lemma)**.** *Let X be a sequentially compact metric space and let $\mathcal{U}$ be an open cover of X. Then there is $\delta > 0$ such that for each $x \in X$, there is $U \in \mathcal{U}$ with $B_\delta(x) \subset U$. Such a δ is called a* Lebesgue number *of $\mathcal{U}$.*

Proof. Suppose that the Lemma fails for some sequentially compact metric space and some open cover $\mathcal{U}$ of X. Then for each $\delta > 0$, there is $y_\delta \in X$ such that $B_\delta(y_\delta)$ is not contained in any of the open sets $U \in \mathcal{U}$.

Some subsequence $(x_n)_n$ of the sequence $n \mapsto y_{\frac{1}{n}}$ converges to some $x \in X$. Hence there is a sequence a_n of positive numbers with $a_n \to 0$, a sequence $x_n \to x$, so that no $B_{a_n}(x_n)$ is contained in one of the elements of $\mathcal{U}$. But there is $U_0 \in \mathcal{U}$ and $\epsilon > 0$, so that $B_\epsilon(x) \subset U_0$. For sufficiently large n, we have $B_{a_n}(x_n) \subset B_\epsilon(x) \subset U_0$. This is a contradiction. $\square$

Now we turn to the Arzela-Ascoli theorem, which is often used.

For two topological spaces X and Y, we denote by $C(X, Y)$ the set of all continuous maps $X \to Y$.

Definition B.62. *Let X be a topological space and let Y be a metric space. A subset $\mathcal{F} \subset C(X, Y)$ is* equicontinuous*, if for each $\epsilon > 0$ und each $x \in X$, there is an open neighborhood $U \subset X$ of x, so that for all $f \in \mathcal{F}$ and $x' \in U$, we have*

$$d(f(x), f(x')) \leq \epsilon.$$

If $\mathcal{F} = \{f\}$ is a single map, the condition of equicontinuity amounts to the continuity of f.

Theorem B.63 (Arzela -Ascoli)**.** *Let X be a separable topological space and let Y be a complete metric space. Let $\mathcal{F} \subset C(X, Y)$ be equicontinuous, and assume that for each $x \in X$, the set*

$$Z_x := \overline{\{f(x) | f \in \mathcal{F}\}} \subset Y$$

is compact. Then

1. *every sequence $f_n \in \mathcal{F}$ has a subsequence which converges pointwise.*
2. *every sequence $f_n \in \mathcal{F}$ which converges pointwise converges uniformly on each compact subset of X, and the limit function is continuous.*

Proof. (1) Let $\{x_m | m \in \mathbb{N}\} \subset X$ be dense. The metric space $Z := \prod_{m=1}^\infty Z_{x_m}$ is sequentially compact, by Corollary B.57. Let $g_n \in Z$ be the element $(f_n(x_m))_m$. Since Z is sequentially compact, a subsequence g_{n_k} of g_n is convergent in Z. But that means nothing else than that $\lim_n f_{n_k}(x_m) \in Y$ exists for each m. For $x \in X$, pick an open neighborhood $U_x \subset X$ as in the definition of equicontinuity and pick $x_m \in U_x$. Then for each $y \in U_x$, we have

(B.64)
$$d(f_{n_k}(y), f_{n_l}(y)) \leq d(f_{n_k}(y), f_{n_k}(x_m)) + d(f_{n_k}(x_m), f_{n_l}(x_m)) + d(f_{n_l}(x_m), f_{n_l}(y)) \leq$$
$$\leq \epsilon + d(f_{n_k}(x_m), f_{n_l}(x_m)) + \epsilon.$$

As $\lim_k f_{n_k}(x_m) \in Y$ exists, this shows that $f_{n_k}(y)$ is a Cauchy sequence in Y and hence converges.

(2) the estimate (B.64) proves more: if $f_n(x)$ converges for all x, then each $x \in X$ has a neighborhood U_x over which the convergence is uniform. This proves the continuity of the limit. Now let $K \subset X$ be compact. We have to prove that $f_n|_K$ converges uniformly. So let $\epsilon > 0$, and for each $x \in K$, choose a neighborhood $U_x \subset X$ of X over which the convergence is uniform. There are finitely many $x_1, \ldots, x_r \in K$ so that $K \subset U = U_{x_1} \cup \ldots \cup U_{x_r}$, and the sequence is uniformly convergent on the union U. □

B.13. **Urysohn's Lemma.**

Theorem B.65 (Urysohn's Lemma)**.** *Let X be a topological space. The following conditions on X are equivalent.*

1. *If $A, B \subset X$ are disjoint closed sets, there are disjoint open sets U, V with $A \subset U$ and $B \subset V$.*
2. *If $A \subset W \subset X$, A closed, W open, there is an open subset U with $A \subset U \subset \overline{U} \subset W$.*
3. *If A, B are disjoint closed sets, there is a continuous function $f : X \to [0,1]$ with $f|_A = 0$ and $f|_B = 1$.*

Proof. $3 \Rightarrow 1$: put $U := f^{-1}([0, \frac{1}{2}))$ and $V := f^{-1}((\frac{1}{2}, 1])$.

$1 \Rightarrow 2$: let $B := W^c$ and pick $A \subset U$, $B \subset V$, $U \cap V = \emptyset$, U, V open. As V is open, we even have $\overline{U} \cap V = \emptyset$, which implies $\overline{U} \subset W$.

$2 \Rightarrow 3$: this is the interesting direction. We first claim that there are open subsets U_a, $a \in \mathbb{Q} \cap [0,1]$, with $A \subset U_0$, $U_1 = B^c$, and

$$\overline{U_s} \subset U_r$$

whenever $s < r$. To this end, choose an enumeration $q_0 = 1, q_1 = 0, q_2, q_3, \ldots$ of $\mathbb{Q} \cap [0,1]$. Put $U_1 = B^c$, and choose U_0 so that $A \subset U_0 \subset \overline{U_0} \subset U_1$.

Suppose the sets $U_0, U_1, U_{q_1}, \ldots, U_{q_{k-1}}$ have already been constructed. There are $i, j < k$ with $q_i < q_k < q_j$, and none of the numbers q_l with $l < k$ lies in (q_i, q_j). Now construct U_{q_k} so that

$$\overline{U_{q_i}} \subset U_{q_k} \subset \overline{U_{q_k}} \subset U_{q_l}.$$

Now we define $f : X \to [0,1]$ by

$$f(x) := \begin{cases} \inf\{r | x \in U_r\} & x \in U_1 \\ 1 & x \in U_1^c. \end{cases}$$

It is clear that $f|_B = 1$ and $f|_A = 0$, and that $0 \leq f \leq 1$. It remains to prove that f is continuous. It suffices to check that for each $a \in \mathbb{R}$, the sets $f^{-1}(-\infty, a)$ and $f^{-1}(a, \infty)$ are open. The following cases are easy:

1. If $a > 1$, then $f^{-1}(-\infty, a) = X$ and $f^{-1}(a, \infty) = \emptyset$ are open.
2. If $a < 0$, then $f^{-1}(-\infty, a) = \emptyset$ and $f^{-1}(a, \infty) = X$ are open.
3. $f^{-1}(-\infty, 0) = \emptyset = f^{-1}(1, \infty)$ is open.

For $r \in \mathbb{Q} \cap [0,1]$, we have

1. $f(x) < r \Rightarrow x \in U_r$ (clear),
2. $x \in U_r^c \Rightarrow f(x) \geq r$ (negation of (1)),
3. $x \in U_r \Rightarrow f(x) \leq r$ (clearr),
4. $f(x) > r \Rightarrow x \in U_r^c$ (negation of (3)).

For $0 < a \leq 1$, we therefore have

$$f^{-1}(-\infty, a) \subset \bigcup_{r<a, r \in \mathbb{Q}\cap[0,1]} U_r \subset \bigcup_{r<a, r \in \mathbb{Q}\cap[0,1]} f^{-1}((-\infty, r]) = f^{-1}(-\infty, a).$$

Both inclusions must be equalities of sets. Hence $f^{-1}((-\infty, a))$ is open.

For $0 \leq a < 1$, we have

$$f^{-1}((a, \infty)) \subset \bigcup_{s>a, s \in \mathbb{Q}\cap[0,1]} U_s^c.$$

If $a < r < s$, then $\overline{U_r} \subset U_s$ or $U_s^c \subset \overline{U_r}^c$, and hence

$$f^{-1}((a, \infty)) \subset \bigcup_{r>a, r \in \mathbb{Q}\cap[0,1]} \overline{U_r}^c \subset \bigcup_{r>a, r \in \mathbb{Q}\cap[0,1]} U_r^c \subset \bigcup_{r>a, r \in \mathbb{Q}\cap[0,1]} f^{-1}([r, \infty)) = f^{-1}((a, \infty)).$$

Again, both inclusions are equalities, and this shows

$$f^{-1}((a, \infty)) = \bigcup_{r>a, r \in \mathbb{Q}\cap[0,1]} \overline{U_r}^c,$$

and the latter set is open. $\square$

Definition B.66. *A space X that satisfies the conditions of Theorem B.65 is called* normal. *A normal Hausdorff space is a T_4*-space.

Lemma B.67. *Metric spaces are normal (and T_4).*

Proof. For a subset $A \subset X$, we define $d_A : X \to \mathbb{R}$ by

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

One verifies that

$$|d_A(x) - d_A(y)| \leq d(x, y),$$

so that d_A is continuous. If A is closed, then $d_A^{-1}(0) = A$.

If A, B are disjoint, then put

$$f(x) := \frac{d_A(x)}{d_A(x) + d_B(x)}.$$

If A and B are closed, then $d_A + d_B > 0$, so that the quotient is defined and continuous. We have clearly $0 \leq f \leq 1$, $f|_A = 0$ and $f|_B = 1$. $\square$

Lemma B.68. *Compact Hausdorff spaces are normal and hence T_4.*

Proof. Let X be compact Hausdorff. We first prove that when $A \subset X$ is closed and $y \in X \setminus A$, there are disjoint open sets U, V with $A \subset U$ and $y \in V$.

Reason: for each $x \in A$, there are disjoint neighborhoods $x \in U_x$ and $y \in V_x$. Because A is compact, we find $x_1, \ldots, x_r \in A$ with $A \subset U := U_{x_1} \cup \ldots \cup U_{x_r}$. Then $V := V_{x_1} \cap \ldots \cap V_{x_r}$ is an open neighborhood of y, and $U \cap V = \emptyset$.

Now let A, B be disjoint closed subsets of X. For each $y \in B$, there are disjoint open sets U_y, V_y with $A \subset U_y$ and $x \in V_y$. Since B is compact, we find $y_1, \ldots, y_s \in B$ with $B \subset V := V_{y_1} \cup \ldots \cup V_{y_s}$. Then $A \subset U_{y_1} \cap \ldots U_{y_s} =: U$, U is open, and $U \cap V = \emptyset$. $\square$

The first part of the proof of Lemma B.68 proves a more general statement:

Lemma B.69. *Let X be a Hausdorff space, $A \subset X$ compact, $x \in X \setminus A$. Then there are disjoint open sets U, V with $x \in V$ and $A \subset U$.* $\square$

B.14. **Consequences of Urysohn's Lemma: Tietze extension and a metrization theorem.**

Theorem B.70 (Tietze extension theorem)**.** *Let X be a normal space, let $Y \subset X$ be a closed subspace and let $f : Y \to \mathbb{R}$ be a continuous function. Then there is a continuous function $g : X \to \mathbb{R}$ with $g|_Y = f$.*

Proof. We give the proof only in the case where f is bounded; the general case can be reduced to this case. We formulate the result in functional-analytic language: we have to prove that the map

$$T : C_b(X;\mathbb{R}) \to C_b(Y;\mathbb{R});\ T(g) := g_f$$

which restricts a bounded continuous function to Y, is surjective. Clearly, $C_b(X)$ and $C_b(Y)$ are Banach spaces (when equipped with the supremum norm) and T is a bounded operator of norm at most 1 (equal to 1 unless $Y = \emptyset$).

Let $f \in C_b(Y)$, $R := \|f\|$. Put $A := \{y \in Y | f(y) \geq \frac{R}{3}\}$ and $B := \{y \in Y | f(y) \leq -\frac{R}{3}\}$. These are disjoint subsets of Y and closed in the subspace topology. Since Y is closed in X, A and B are also closed in X.

By Ursohn's lemma, there is a continuous function $h : X \to [-\frac{R}{3}, \frac{R}{3}]$, $h|_A = \frac{R}{3}$ and $h|_B = -\frac{R}{3}$. Then

$$\|h\| \leq \frac{1}{3}\|f\|$$

and

$$\|Th - f\| \leq \frac{2}{3}\|f\|.$$

A Lemma which appears in the proof of the open mapping theorem[7], proves from this that for each $f \in C_b(Y;\mathbb{R})$, there is $g \in C_b(X;\mathbb{R})$ with $Tg = f$ and

$$\|g\| \leq \frac{1}{3}\frac{1}{1-\frac{2}{3}}\|f\| = \|f\|,$$

which is what we wanted to prove. □

Theorem B.71 (Special case of Urysohn's metrization theorem)**.** *Let X be a compact Hausdorff space. Then the following are equivalent:*

1. *X is second countable.*
2. *There are countably many continuous functions $f_k : X \to [0,1]$, $k \in \mathbb{N}$ which* separate the points of X, *in other word for $x \neq y \in X$, there is k with $f_k(x) \neq f_k(y)$.*
3. *X is metrizable.*

Proof. $2 \Rightarrow 1$: this is the content of Corollary B.60.

$1 \Rightarrow 2$: Let $\mathcal{B} = \{U_n | n \in \mathbb{N}\}$ be a basis for the topology of X. Let $I := \{(m,n) \in \mathbb{N} \times \mathbb{N} | \overline{U_n} \subset U_m\}$, which is a countable set. For $(m,n) \in I$, pick a function $f_{m,n} : X \to [0,1]$ with $f|_{\overline{U_n}} = 1$ and $f|_{U_m^c} = 0$, which exists by Urysohn's lemma and Lemma B.68. If $x \neq y \in X$, there is $m \in \mathbb{N}$ with $x \in U_m$, $y \notin U_m$. Since $\{x\}$ is closed as X is Hausdorff, there is an open V with $x \in V \subset \overline{V} \subset U_m$, and so there is $n \in \mathbb{N}$ with $x \in U_n \subset V$. Then $(m,n) \in I$ and $f_{m,n}(x) = 1$, $f_{m,n}(y) = 0$.

[7]This Lemma will be included in the lecture notes.

$2 \Rightarrow 3$: Let $(f_k)_k$ be a family of functions as in (2). Consider the function

$$f : X \to Z := \prod_{k=1}^{\infty} [0,1];\ f(x) := (f_k(x))_k.$$

This is continuous and injective since (f_k) separates the points of X. As Z is Hausdorff, so is the image $f(X) \subset Z$. By Theorem B.52, $f : X \to f(X)$ is a homeomorphism. But Z is metrizable by B.35, and hence so is $f(X) \cong X$. □

B.15. **The Stone–Weierstrass Theorem.**

Theorem B.72 (Stone–Weierstrass Theorem)**.** *Let X be a compact Hausdorff space. Let $A \subset C(X;\mathbb{K})$ be a subalgebra with the following properties:*

(1) $1 \in A$,
(2) *A separates the points of X (i.e. for $x \neq y \in X$, there is $f \in A$ with $f(x) \neq f(y)$),*
(3) *if $f \in A$, then the complex conjugate $\overline{f} \in A$.*

Then A is dense.

Corollary B.73. *For each continuous function $f : [a,b] \to \mathbb{K}$, there is a sequence p_n of polynomials which converges uniformly on $[a,b]$ to f.*

Proof. Let $A \subset C([a,b];\mathbb{K})$ be the subalgebra of polynomials. This satisfies the hypotheses of Theorem B.72, so it lies dense in $C([a,b];\mathbb{K})$, as claimed. □

Corollary B.74. *For each continuous function $f : S^1 \to \mathbb{C}$, there is a sequence p_n of trigonometric polynomials, in other words, functions of the form $\sum_{k=-n}^{n} a_n z^n$, $a_n \in \mathbb{C}$, which converges uniformly to f.*

Proof. Apply Theorem B.72 to the algebra $\mathcal{T} \subset C(S^1;\mathbb{C})$ of all trigonometric polynomials. □

For the proof of Theorem B.72, we first have to verify one special case of Corollary B.73.

Lemma B.75. (1) *There is a sequence of polynomials g_n which converges uniformly on each compact subinterval of $(0,2)$ to the function $\sqrt{x}$.*
(2) *There is a sequence of polynomials f_n which converges uniformly on $[0,1]$ to $\sqrt{x}$.*
(3) *For each $S > 0$, there is a sequence of polynomials p_n which converges uniformly on $[0,S]$ to $\sqrt{x}$.*
(4) *For each $R > 0$, there is a sequence of polynomials q_n which converges uniformly on $[-R,R]$ to $|x|$.*

Proof. 1: it is known from calculus that the function $\sqrt{1+x}$ can be expanded into the power series

$$\sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} x^k,$$

where the binomial coefficients for $s \in \mathbb{R}$ are defined as

$$\binom{s}{k} := \frac{s(s-1)\cdots(s-k+1)}{k!};\ \binom{s}{0} := 1.$$

The power series has radius of convergence 1, and so the sequence of partial sums

$$g_n(x) := \sum_{k=0}^{n} \binom{\frac{1}{2}}{k} x^k$$

converges uniformly on each compact subinterval of $(-1,1)$.

$1 \Rightarrow 2$: The sequence $\sqrt{\frac{1}{m}+x}$ converges uniformly on $[0,\infty)$ to $\sqrt{x}$, and by (1), each $\sqrt{\frac{1}{m}+x}$, $m \geq 2$, can be approximated uniformly by polynomials on $[0,1]$.

$2 \Rightarrow 3$: put $p_n(x) := \sqrt{S} f_n(\frac{x}{S})$.

$3 \Rightarrow 4$: put $q_n(x) := p_n(x^2)$, for suitable S. □

Proof of Theorem B.72. First, we reduce to the case $\mathbb{K} = \mathbb{R}$. Suppose that $A \subset C(X;\mathbb{C})$ is a subalgebra which satisfies the hypotheses of the Theorem, and that the theorem is proven for $\mathbb{K} = \mathbb{R}$. Put $B := A \cap C(X;\mathbb{R})$. Then B is a subalgebra, $1 \in B$, and B separates points. To see the latter claim, let $f \in A$ with $f(x) \neq f(y)$. Then either $\Im(f(x)) \neq \Im(f(y))$ or $\Re(f(x)) \neq \Re(f(y))$, but $\Im(f) = \frac{1}{2i}(f - \overline{f}) \in A \cap C(X;\mathbb{R})$ and $\Re(f) = \frac{1}{2}(f + \overline{f}) \in A \cap C(X;\mathbb{R})$. By the real version of the theorem, $B \subset C(X;\mathbb{R})$ is dense. If $f \in C(X;\mathbb{C})$, pick $g, h \in B$ with $\|g - \Re(f)\|, \|h - \Im(f)\| \leq \epsilon$. Then $g + ih \in B$ and $\|f - (g + ih)\| \leq 2\epsilon$.

Hence it is enough to prove the real case. We shall prove that the closure $\overline{A} \subset C(X;\mathbb{R})$ is dense. By Lemma B.75, we see that

$$f \in A \Rightarrow |f| \in \overline{A}$$

(since $q_n(f) \in A$ converges uniformly to $|f|$). Because

$$\max(f,g) = \frac{1}{2}(g + f + |g - f|)$$

and

$$\min(f,g) = \frac{1}{2}(g + f - |g - f|),$$

we get

$$f, g \in A \Rightarrow \max(f,g), \min(f,g) \in \overline{A}.$$

The properties of A that have been used in the proof also hold for $\overline{A}$, and therefore we conclude

$$f, g \in \overline{A} \Rightarrow \max(f,g), \min(f,g) \in \overline{A}. \tag{B.76}$$

Now consider $h \in C(X;\mathbb{R})$ and let $\epsilon > 0$ be arbitrary. First fix $x \in X$. For $y \in X$, there is a function $h_{x,y} \in A$ with $h_{x,y} = f(x)$ and $h_{x,y}(y) = f(y)$ (this follows since $1 \in A$ and since A separates the points of X). The set

$$U_y := \{z \in X | h_{x,y}(z) < f(z) + \epsilon$$

is an open neighborhood of y. Since X is compact, we find finitely many points $y_1, \ldots, y_r$, such that $U_{y_1} \cup \ldots \cup U_{y_r} = X$. The function

$$h_x := \min(h_{x,y_1}, \ldots, h_{x,y_r})$$

is an element of $\overline{A}$ by (B.76).

Let $z \in X$ be arbitrary, say $z \in U_{y_j}$. Then

$$h_x(z) \leq h_{x,y_j}(z) < f(z) + \epsilon,$$

and moreover

$$h_x(x) = f(x).$$

Let

$$V_x := \{z \in X | f(z) - \epsilon < h_x(z)\},$$

which is an open neighborhood of x. Using the compactness of X again, we find $z_1, \ldots, z_s \in X$ with $X = V_{z_1} \cup \ldots \cup V_{z_s}$. Now put

$$h := \max(h_{z_1}, \ldots, h_{z_s}).$$

By (B.76), $h \in \overline{A}$. Because $h_{z_i} < f + \epsilon$, we have $h < f + \epsilon$. Moreover, for $z \in V_{z_j}$,

$$h(z) \geq h_{z_j}(z) > f(z) - \epsilon.$$

Therefore $h > f - \epsilon$, and altogether $\|f - h\| < \epsilon$. Therefore $\overline{A} \subset C(X; \mathbb{R})$ and a fortiori A is dense. □

B.16. **Locally compact spaces.**

Definition B.77. *A space X is* locally compact *if each point $x \in X$ of X has a compact neighborhood.*

For example $\mathbb{R}^n$ and each open or closed subset of $\mathbb{R}^n$ is locally compact. Each compact space is locally compact, for trivial reasons. Without a Hausdorff hypothesis, locally compact spaces can be equite pathological (for example, open subsets of locally compact spaces do not need to be locally compact).

Lemma B.78. *Let X be a locally compact Hausdorff space and let $x \in U \subset X$ be an open neighborhood of x. Then there is an open set O with $x \in O \subset \overline{O} \subset U$ and $\overline{O}$ compact.*

Hence open or closed subspaces of locally compact Hausdorff spaces are locally compact.

Proof. Let C be a compact neighborhood of x. By Lemma B.51, $C \subset X$ is closed. Hence $C \cap U^c \subset X$ is closed, and as a subset of the compact set C compact, again by Lemma B.51.

Since $x \notin C \cap U^c$, Lemma B.69 shows that there are disjoint open sets V, W with $x \in W$ and $C \cap U^c \subset V$.

Using Lemma B.51 again, we see that $\overline{W} \cap C$ is a compact neighborhood of x. But

$$(\overline{W} \cap C) \cap U^c = \overline{W} \cap (C \cap U^c) = \emptyset$$

and therefore $\overline{W} \cap C \subset U$. So we have found a compact neighborhood D of x which is contained in U. By definition of the term "neighborhood"', D contains an open neighborhood $x \in O$ of x. Since D is closed (again Lemma B.51), $\overline{O} \subset \overline{D} = D \subset U$. □

A locally compact Hausdorff space is not necessarily normal, but a weaker version of Urysohn's Lemma is true for locally compact spaces.

Proposition B.79. *Let X be locally compact, $A \subset X$ compact, $B \subset X$ closed, $A \cap B = \emptyset$. Then there is a continuous function $f : X \to [0, 1]$ with $f|_A = 1$ and $f|_B = 0$, and the* support $\mathrm{supp}(f) := \overline{\{x \in X | f(x) \neq 0\}}$ *is compact.*

Proof. Put $U := A^c$, so that $B \subset U$. For $x \in B$, there is an open set V_x with $x \in V_x \subset \overline{V_x} \subset U$, by Lemma B.78. Since B is compact, we can cover B by finitely many such sets $V_{x_1}, \ldots, V_{x_r}$. The closure K of the open set $V = V_{x_1} \cup \ldots \cup V_{x_r}$ is compact, being a union of finitely many compact sets. The closed subsets ∂V and A are disjoint subsets of K.

The space K is compact Hausdorff and hence normal. Using Urysohn's Lemma, there is $g : K \to [0,1]$ with $g|_A = 1$ and $g|_{\partial V} = 0$.

Note that $\overline{K^c} = V^c$. Define $f : X \to [0,1]$ by

$$f(x) = \begin{cases} g(x) & x \in K \\ 0 & x \in V^c, \end{cases}$$

which is continuous. □

Definition B.80. *A locally compact space X is* σ-compact *if there are compact subspaces $K_n \subset X$, $n \in \mathbb{N}$ with $K_1 \subset K_2^\circ \subset K_2 \subset K_3^\circ \subset K_3 \subset \dots$ and $\bigcup_{n=1}^\infty K_n = X$. Such a sequence of compact subsets is called a* compact exhaustion *of X.*

Lemma B.81. *A second countable locally compact Hausdorff space is σ-compact.*

Proof. Let $\mathcal{B}$ be a countable basis for the topology of X, and let $\mathcal{B}' \subset \mathcal{B}$ be the subset of those U with compact closure. Using Lemma B.78, $\mathcal{B}' = \{U_1, \dots, \}$ is a basis for the topology.

Now put $K_1 := \overline{U_1}$. There is n_2 with $K_1 \subset \bigcup_{k=1}^{n_2} U_k$, and put $K_2 := \overline{\bigcup_{k=1}^{n_2+1} U_k}$. There is n_3 with $K_2 \subset \bigcup_{k=1}^{n_3} U_k$ and put $K_3 := \overline{\bigcup_{k=1}^{n_3+1} U_k}$. Continue in this fashion. □

Theorem B.82 (Partitions of unity)**.** *Let X be a σ-compact locally compact Hausdorff space and let $\mathcal{U}$ be an open cover of X. Then there exists a partition of unity subordinate to the cover $\mathcal{U}$, that is, a family $(f_j)_{j \in J}$ of continuous functions $f_j : X \to [0,1]$, such that*

1. *The family f_j is* locally finite*, which means that each $x \in X$ has a neighborhood which intersects only finitely many of the* supports $\mathrm{supp}(f_j) := \overline{\{y \in X | f_j(y) \neq 0\}}$.
2. *For each $j \in J$, there is $U \in \mathcal{U}$ with* $\mathrm{supp}(f_j) \subset U$.
3. $\sum_{j \in J} f_j(x) = 1$ *for all $x \in X$ (this is a finite sum).*

Proof. The set $L_n := K_n \setminus K_{n-1}^\circ$ is compact. For each $x \in L_n$, pick a set $U_x \in \mathcal{U}$, a compact neighhorhood C_x of x and an open set V_x with $C_x \subset V_x \subset U_x \cap (K_{n+1} \setminus K_{n-2}^\circ)$ (hence $\overline{V_x}$ is compact), and a function $h_x : X \to [0,1]$ with $\mathrm{supp}(h_x) \subset V_x$, $h_x|_{C_x} = 1$, according to Proposition B.79. Finitely many of such C_x cover L_n.

Hence we have found a finite set I_n, and for each $i \in I_n$ a function $h_i : X \to [0,1]$ such that the support of h_i is contained in one of the sets of $\mathcal{U}$, and in the set $K_{n+1} \setminus K_{n-2}^\circ$, and the sum $\sum_{i \in I_n} h_i \geq 0$ is positive everywhere on L_n.

Putting all those finite sets I_n, $n \in \mathbb{N}$ and these functions together, we found a countable set I, and functions $h_i : X \to [0,1]$ each of which has support in one of the sets of $\mathcal{U}$. Moreover, only finitely many of the supports of h_i intersect K_n, and the (finite) sum $\sum_{i \in I} h_i(x) > 0$ for all x. Finally, we put

$$f_i(x) := \frac{h_i(x)}{\sum_{i \in I} h_i(x)}.$$

□

B.17. **The 1-point compactification.**

Definition B.83. *Let X be a topological space. We define $X^+ := X \cup \{\infty\}$, where ∞ is an element not contained in X, with the following topology. A set $U \subset X^+$ is open if either*

(1) *$U \subset X \subset X^+$ and $U \subset X$ is open (type 1) or*
(2) *$\infty \in U$, and $X^+ \setminus U \subset X$ is a closed compact set (type 2).*

Lemma B.84. *The collection of open sets in X^+ described above is a topology, and X^+ is compact.*

Proof. In order to prove the first claim, we need to verify:

(1) X^+ and $\emptyset$ are open (clear).
(2) Arbitrary unions and finite intersections of type 1 open sets are type 1 open sets (clear).
(3) Finite intersections of two type 2 open sets are type 1 open sets (this is because finite unions of compact closed subsets of X are closed and compact).
(4) Arbitrary unions of type 2 open sets are type 2 open sets. This is because an intersection of arbitrarily many closed compact sets is closed (clear) and compact (by Lemma B.51).
(5) Let U_1 be open of type 1 and U_2 open of type 2. Then $U_1 \cap U_2$ is open of type 1 (clear).
(6) Let U_1 be open of type 1 and U_2 open of type 2. Then $U_1 \cup U_2$ is open of type 2, because $X^+ \setminus (U_1 \cup U_2) = (X^+ \setminus U_1) \cap (X^+ \setminus U_2)$ is the intersection of a closed compact set with a closed set and hence also compact by Lemma B.51.

Let $\mathcal{U}$ be an open cover of X^+. The point ∞ is contained in one of the sets of $\mathcal{U}$, say $x \in U$. Then $X^+ \setminus U$ is compact, and $\mathcal{U} \setminus \{U\}$ is an open cover of $X^+ \setminus U$, hence has a finite subcover. $\square$

We denote by $\iota : X \to X^+$ the inclusion map. It is continuous, and its image is open and dense.

Proposition B.85. (1) *ι is a homeomorphism.*
(2) *X^+ is Hausdorff if and only if X is locally compact and Hausdorff.*
(3) *X^+ is Hausdorff and second countable if and only if X is second countable locally compact and Hausdorff.*

Proof. 1: it is clear that ι is continuous and $\iota : X \to \iota(X)$ is bijective. If $U \subset X$ is open, then $\iota(U) \subset X^+$ is a (type 1) open subset, and hence also open in the subspace topology of $\iota(X)$.

2: if X^+ is Hausdorff, then $X \cong \iota(X)$ is clearly Hausdorff, and locally compact by Lemma B.78. If X is Hausdorff, then any two points $x, y \in x \subset X^+$ can be separated by disjoint open sets of type 1. If X is also locally compact and $x \in X$, we can find $x \in U \subset K$, where U is open and K compact (and hence closed). Then $X^+ \setminus K$ is a type 2 open set which contains ∞, but is disjoint from the open neighborhood U of x.

3: Subspaces of second countable spaces are always second countable. This proves “only if”. For the other implication, note that a neighborhood basis of X, together with the sets of the form $X^+ \setminus K_n$ for a compact exhaustion of X, makes up a basis for the topology of X^+. $\square$

Corollary B.86. *A second countable locally compact Hausdorff space is metrizable.*

Appendix C. Measure and integration

C.1. σ-algebras and measurable spaces.

Notation C.1. *If X is a set, we denote by $\mathcal{P}(X)$ the power set of X, i.e. the set of all subsets of X. Furthermore, for a subset $S \in \mathcal{P}(X)$, we denote by S^c the complement of S, $S^c := X \setminus S$.*

Definition C.2. *Let X be a set. A* σ-algebra *on X is a subset $\mathcal{B} \subset \mathcal{P}(X)$, such that the following hold:*

1. *$\emptyset \in \mathcal{B}$,*
2. *$\mathcal{B}$ is closed under forming complements, i.e. $S \in \mathcal{B} \Rightarrow S^c \in \mathcal{B}$,*
3. *$\mathcal{B}$ is closed under forming countable unions, i.e. $S_n \in \mathcal{B}$, $n \in \mathbb{N}$, then $\bigcup_{n=1}^\infty S_n \in \mathcal{B}$.*

A measurable space *is a pair $(X, \mathcal{B})$, consisting of a set X together with a σ-algebra $\mathcal{B}$ on X. We call the elements of $\mathcal{B}$ the* measurable *subsets of X.*

Lemma C.3. *Let $\mathcal{B}$ be a σ-algebra, and $S, T \in \mathcal{B}$, $S_n \in \mathcal{B}$, $n \in \mathbb{N}$. Then $S \cap T, S - T \in \mathcal{B}$, $\bigcap_{n=1}^\infty S_n \in \mathcal{B}$.*

Proof. Use
$$(\bigcap_n S_n)^c = \bigcup_n S_n^c$$
and
$$S - T = S \cap T^c.$$
□

Lemma C.4. *Let $\mathcal{A} \subset \mathcal{P}(X)$ be a collection of subsets of a set X. Then there is a unique smallest σ-algebra $\langle \mathcal{A} \rangle$ which contains $\mathcal{A}$.*

Proof. Let $\mathbb{S}$ be the set of all σ-algebras on X which contain $\mathcal{A}$. The set $\mathbb{S}$ is nonempty, since $\mathcal{P}(X)$ is an element of $\mathbb{S}$. Then put
$$\langle \mathcal{A} \rangle := \bigcap_{\mathcal{B} \in \mathbb{S}} \mathcal{B}.$$
□

Definition C.5. *Let X be a topological space, with topology $\mathcal{T} \subset \mathcal{P}(X)$. The σ-algebra $\langle \mathcal{T} \rangle$ on X is called the* Borel-σ-algebra, *and its elements are the* Borel sets.

If X is second-countable and $\mathcal{U}$ a countable basis for its topology, then $\langle \mathcal{T} \rangle = \langle \mathcal{U} \rangle$.

The Dynkin lemma.

Definition C.6. *Let X be a set. A* π-system *on X is a subset $\mathcal{F} \subset \mathcal{P}(X)$, such that*
$$S, T \in \mathcal{F} \Rightarrow S \cap T \in \mathcal{F}.$$
A Dynkin system *on X is a subset $\mathcal{D} \subset \mathcal{P}(X)$ such that*

1. *$X \in \mathcal{D}$,*
2. *$A, B \in \mathcal{D}, A \subset B \Rightarrow B \setminus A \in \mathcal{D}$,*
3. *$A_n \in \mathcal{D}$, $A_1 \subset A_2 \subset \ldots \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{D}$.*

For each subset $\mathcal{A} \subset \mathcal{P}(X)$, there is a unique smallest Dynkin system $\lambda(\mathcal{A})$ which contains $\mathcal{A}$. The proof of its existence is as the proof of Lemma C.4.

Lemma C.7. *A subset $\mathcal{B} \subset \mathcal{P}(X)$ which is both, a π-system and a Dynkin system, is a σ-algebra.*

Proof. If $S \in \mathcal{B}$, then $S^c = X - S \in \mathcal{B}$, by the first two axioms of a Dynkin system. Now let S_n, $n \in \mathbb{N}$, be an arbitrary countable collection of elements of $\mathcal{B}$. We have to show that $\bigcup_{n=1}^{\infty} S_n \in \mathcal{B}$. It is, by the third axiom of a Dynkin system, enough to prove that $\bigcup_{n=1}^{m} S_n \in \mathcal{B}$, for each m. This is shown by induction on m, and the inductive step amounts to proving that $S, T \in \mathcal{B} \Rightarrow S \cup T \in \mathcal{B}$. But $S^c, T^c \in \mathcal{B}$, and because $\mathcal{B}$ is a π-system, we have

$$(S \cup T)^c = S^c \cap T^c \in \mathcal{B},$$

which implies $S \cup T \in \mathcal{B}$. □

Lemma C.8 (Dynkin lemma)**.** *Let $\mathcal{A}$ be a π-system on X and let $\mathcal{D}$ be a Dynkin system which contains $\mathcal{A}$. Then the σ-algebra $\mathcal{B}(\mathcal{A})$ generated by $\mathcal{A}$ is contained in $\mathcal{D}$.*

Proof. It is enough to consider the case where $\mathcal{D} = \lambda(\mathcal{A})$ is the smallest Dynkin system containing $\mathcal{A}$. Let

$$\mathcal{D}_1 := \{T \subset X | T \cap S \in \mathcal{D} \; \forall S \in \mathcal{A}\}.$$

As $\mathcal{A}$ is a π-system, we have $\mathcal{A} \subset \mathcal{D}_1$, and we claim that $\mathcal{D}_1$ is a Dynkin system. As $\mathcal{A} \subset \mathcal{D}$, it follows that $X \in \mathcal{D}_1$. If $A \subset B$ and $A, B \in \mathcal{D}_1$ and $S \in \mathcal{A}$, then

$$(B - A) \cap S = B \cap S - A \cap S.$$

As $A \cap S$ and $B \cap S$ are in $\mathcal{D}$ and as $\mathcal{D}$ is a Dynkin system, we get $(B - A) \cap S \in \mathcal{D}$ and hence $B - A \in \mathcal{D}_1$. If $A_1 \subset A_2 \subset \ldots$ is an increasing sequence of elements of $\mathcal{D}_1$, and $S \in \mathcal{A}$, then

$$(\bigcup_{n=1}^{\infty} A_n) \cap S = \bigcup_{n=1}^{\infty} A_n \cap S.$$

Since $A_n \cap S \in \mathcal{D}$ and since $\mathcal{D}$ is a Dynkin system, it follows that $(\bigcup_{n=1}^{\infty} A_n) \cap S \in \mathcal{D}$ and hence $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}_1$. This finishes the proof that $\mathcal{D}_1$ is a Dynkin system containing $\mathcal{A}$. As $\mathcal{D} = \lambda(\mathcal{A})$, this implies $\mathcal{D} \subset \mathcal{D}_1$. In particular:

$$T \in \mathcal{D}, S \in \mathcal{A} \Rightarrow T \cap S \in \mathcal{D}. \tag{C.9}$$

Now put

$$\mathcal{D}_2 := \{T \subset X | T \cap S \in \mathcal{D} \forall S \in \mathcal{D}\}.$$

By (C.9), we find that $\mathcal{A} \subset \mathcal{D}_2$. We claim that $\mathcal{D}_2$ is a Dynkin system. It is clear that $X \in \mathcal{D}_2$, and the other two axioms are verified by the same argument as above, replacing the statement "$S \in \mathcal{A}$" by "$S \in \mathcal{D}$". As $\mathcal{D}_2$ is a Dynkin system which contains $\mathcal{A}$, it again follows that $\mathcal{D} \subset \mathcal{D}_2$. This means that

$$S, T \in \mathcal{D} \Rightarrow S \cap T \in \mathcal{D},$$

or in other words that $\mathcal{D}$ is a π-system. It follows that $\mathcal{D}$ is a σ-algebra, by Lemma C.7. Since $\mathcal{A} \subset \mathcal{D}$, the σ-algebra $\sigma(\mathcal{A})$ must be contained in $\mathcal{D}$. □

C.2. **Measurable maps.**

Definition C.10. *Let $(X, \mathcal{B})$ and $(Y, \mathcal{C})$ be measurable spaces. A map $f : X \to Y$ is* measurable *if $f^{-1}(S) \in \mathcal{B}$ for each $S \in \mathcal{C}$.*

If $\mathcal{C}$ is the σ-algebra generated by $\mathcal{A} \subset \mathcal{P}(X)$, it is enough to prove that $f^{-1}(S) \in \mathcal{B}$ for each $S \in \mathcal{A}$. It is clear that compositions of measurable maps are measurable.

When $f : X \to Y$ is a continuous map between topological spaces and $\mathcal{C}_X, \mathcal{C}_Y$ are the respective Borel-σ-algebras, then $f : (X, \mathcal{C}_X) \to (Y, \mathcal{C}_Y)$ is clearly measurable.

Lemma C.11. *Let $(X, \mathcal{B})$ be a measurable space, let Y be a topological space and let $f_n : X \to Y$ be a sequence of measurable maps. Suppose that $f(x) = \lim_n f_n(x)$ exists for all $x \in X$. Then the limit map f is measurable.*

Proof. If $U \subset Y$ is open, then $x \in f^{-1}(U)$ if and only if $x \in f_n^{-1}(U)$ for all sufficiently large n. Therefore

$$f^{-1}(U) = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} f_n^{-1}(U)$$

and since each f_n is measurable, it follows that $f^{-1}(U) \in \mathcal{B}$. □

Lemma C.12. *Let X be a measurable space and let $f_n : X \to [-\infty, \infty]$ be measurable. Then $\sup_n f_n$, $\inf_n f_n$, $\liminf_n f_n$ and $\limsup_n f_n$ are measurable.*

Proof. Since $\max : \mathbb{R}^n \to \mathbb{R}$ is continuous, the functions $g_n := \max(f_1, \ldots, f_n)$ are measurable, and $\sup_n f_n = \lim_n g_n$, and so $\sup_n f_n$ is measurable. Similarly, $\inf_n f_n$ is measurable. Furthermore

$$\liminf_n f_n = \sup_n \inf_{k \geq n} f_k$$

and

$$\limsup_n f_n = \inf_n \sup_{k \geq n} f_k$$

are measurable. □

C.3. **Measures.**

Definition C.13. *Let X be a set and let $\mathcal{B}$ be a σ-algebra on X. A* measure *on $\mathcal{B}$ is a function $\mu : \mathcal{B} \to [0, \infty]$ such that*

(1) *$\mu(\emptyset) = 0$,*
(2) *if $S_n \in \mathcal{B}$, $n \in \mathbb{N}$, are pairwise disjoint, then $\mu(\bigcup_{n=1}^{\infty} S_n) = \sum_{n=1}^{\infty} \mu(S_n)$.*

A measure space *is a tuple $(X, \mathcal{B}, \mu)$, consisting of a set X, a σ-algebra and a measure μ on $\mathcal{B}$. The elements of $\mathcal{B}$ are called the* measurable subsets *of X.*

Example C.14. *Let X be any set. The* counting measure *on X is $\mu : \mathcal{P}(X) \to [0, \infty]$, $\mu(S) := |S|$. Slightly more generally, for a function $a : X \to [0, \infty]$, the weighted counting measure $\mu(S) := \sum_{s \in S} a(s)$ is a measure.*

The construction of other measures is a nontrivial task. In the next subsection, we learn the key device for the construction of measures.

Lemma C.15. (1) *If $S \subset T$, then $\mu(S) \leq \mu(T)$.*
(2) *If $S_1 \subset S_2 \subset \ldots$ is an ascending sequence of elements of $\mathcal{B}$ and $S := \bigcup_{n=1}^{\infty} S_n$, then $\mu(S) = \lim_n \mu(S_n)$.*
(3) *For an arbitrary countable collection $S_n \in \mathcal{B}$, we have $\mu(\bigcup_{n=1}^{\infty}) \leq \sum_{n=1}^{\infty} \mu(S_n)$.*

(4) *If $S_1 \supset S_2 \supset \dots$ is a descending sequence of elements of $\mathcal{B}$, and $\mu(S_1) < \infty$, then*

$$\mu(\bigcap_{n=1}^{\infty} S_n) = \lim_n \mu(S_n).$$

Proof. (1): Since $T = S \dot\cup (T - S)$, we have $\mu(T) = \mu(S) + \mu(T - S)$.

(2): by (1), we have $\mu(S_1) \leq \mu(S_2) \leq \dots \leq \mu(S)$, so that $\lim_n \mu(S_n) \leq \mu(S)$. The reverse inequality is nontrivial only if $\lim_n \mu(S_n) < \infty$. Write $T_n := S_n - S_{n-1}$, $T_1 = S_1$, so that S is the *disjoint* union $\bigcup_{n=1}^{\infty} T_n$, and

$$\mu(S) = \sum_{n=1}^{\infty} \mu(T_n).$$

All the sets T_n have finite measure, and $\mu(T_n) = \mu(S_n) - \mu(S_{n-1})$, so that $\sum_{n=1}^{m} \mu(T_n) = \mu(S_m)$ (telescope sum).

(3): Let $T_0 = \emptyset$ and $T_n := \bigcup_{m=1}^{n} S_m$; which is an ascending sequence with union $\bigcup_{n=1}^{\infty} S_n$. It follows from (2) that

$$\mu(\bigcup_{n=1}^{\infty} S_n) = \lim_{n\to\infty} \mu(T_n)$$

and therefore, it is enough to prove that $\mu(T_n) \leq \sum_{m=1}^{n} \mu(S_m)$. But $T_m - T_{m-1} \subset S_m$, and so

$$\mu(T_n) = \sum_{m=1}^{n} \mu(T_m - T_{m-1}) \leq \sum_{m=1}^{n} \mu(S_m).$$

(4): Since

$$S_1 = \bigcup_{n=1}^{\infty} (S_1 - S_n) \dot\cup \bigcap_{n=1}^{\infty} S_n,$$

it follows from (2) that

$$\mu(S_1) = \lim_{n=1}^{\infty} \mu(S_1 - S_n) + \mu(\bigcap_{n=1}^{\infty} S_n)$$

or that (here we are using that $\mu(S_1) < \infty$)

$$\mu(\bigcap_{n=1}^{\infty} S_n) = \mu(S_1) - \lim_n \mu(S_1 - S_n) = \lim_n \mu(S_1) - \mu(S_1 - S_n) = \lim_n \mu(S_n).$$

$\square$

Definition C.16. *Let $(X, \mathcal{B}, \mu)$ be a measure space. A* null set *is a set $S \subset X$ such that there is $T \in \mathcal{B}$ with $\mu(S) = 0$ and $S \subset T$. We say that $(X, \mathcal{B}, \mu)$ is* complete *if every null set is an element of $\mathcal{B}$.*

Lemma C.17. *Let $(X, \mathcal{B}, \mu)$ be a measure space. Let $\mathcal{C} \subset \mathcal{P}(X)$ be the set of all subsets T, such that there is $S \in \mathcal{B}$, so that the symmetric difference $S \Delta T := (S - T) \cup (T - S)$ is a null set. Then $\mathcal{C}$ is a σ-algebra, μ admits a unique extension to a measure μ' on $\mathcal{C}$ and $(X, \mathcal{C}, \mu')$ is complete.*

Proof. The symmetric difference is $T\Delta S = (T\cap S^c)\cup(T^c\cap S)$ and so
$$T^c\Delta S^c = T\Delta S.$$
Hence the complement of a set in $\mathcal{C}$ also belongs to $\mathcal{C}$. Assume that $S_n\in\mathcal{C}$, $T_n\in\mathcal{B}$ with $S_n\Delta T_n$ a null set. Then
$$(\bigcup_{n=1}^{\infty} T_n)\Delta(\bigcup_{n=1}^{\infty} S_n)\subset \bigcup_{n=1}^{\infty} T_n\Delta S_n,$$
and the union of countable many null sets is a null set.

Let μ' be a measure on $\mathcal{C}$ extending μ, $S\in\mathcal{C}$, and $T\in\mathcal{B}$ with $S\Delta T$ a null set, we must have
$$\mu'(S\Delta T)=0,$$
and because
$$(S\cap T)\cup(S-T)=S$$
$$(S\cap T)\cup(T-S)=T,$$
we must have $\mu'(S)=\mu(T)$. This proves uniqueness of μ'. To prove existence, first observe that if $T_0,T_1\in\mathcal{B}$ are two sets such that $T_i\Delta S$ is a null set, it follows that $T_0\Delta T_1$ is a null set, hence $\mu(T_1)=\mu(T_0)$. Hence defining $\mu'(S):=\mu(T)$ is unambiguous, and the σ-additivity is left as an exercise. $\square$

Definition C.18. *Let $(X,\mathcal{B},\mu)$ be a measure space. We say:*

(1) *μ is* finite *if $\mu(X)<\infty$,*
(2) *μ is* locally finite *if each measurable subset $S\subset X$ with $\mu(S)>0$ contains a measurable subset $T\subset S$ with $0<\mu(T)<\infty$,*
(3) *μ is* σ-finite *if there is a decomposition $X=\bigcup_{n=1}^{\infty}X_n$ of X into countably many disjoint measurable subsets such that $\mu(X_n)<\infty$ for all n.*

Remark C.19. *A σ-finite measure is locally finite: if $S\subset X$ has positive measure and $X=\bigcup_{n=1}^{\infty}X_n$ is a decomposition of X into disjoint subsets of finite measure, then $\mu(S)=\sum_{n=1}^{\infty}\mu(S\cap X_n)$. At least one of $\mu(X_n\cap S)$ is positive, and is necessarily finite.*

C.4. **The Caratheodory extension theorem.**

Definition C.20. *A (concrete)* Boolean algebra *on a set X is a subset $\mathcal{A}\subset\mathcal{P}(X)$, so that*

(1) *$\emptyset\in\mathcal{A}$,*
(2) *$S,T\in\mathcal{A}\Rightarrow S^c\in\mathcal{A}, S\cup T\in\mathcal{A}$.*

It follows that $X=\emptyset^c\in\mathcal{A}$ and with $S,T\in\mathcal{A}$ that

(1) $S\cap T=(S^c\cup T^c)^c\in\mathcal{A}$,
(2) $S\setminus T:=S\cap T^c\in\mathcal{A}$,
(3) the *symmetric difference* $S\triangle T:=(S\cup T)\setminus(S\cap T)\in\mathcal{A}$.

Calculations with these set operations are most easily performed when one identifies $\mathcal{P}(X)$ with the set $\mathbb{F}_2^X$ of all functions $X\to\mathbb{F}_2$ (identify S with the characteristic function χ_S). Then
$$\chi_\emptyset=0,\chi_{S\cap T}=\chi_S\chi_T,\chi_{S\triangle T}=\chi_S+\chi_T,\chi_{S\cup T}=\chi_S\chi_T+\chi_S+\chi_T.$$

Definition C.21. *Let $\mathcal{A}$ be a Boolean algebra on X. A* premeasure *on $\mathcal{A}$ is a function $\mu:\mathcal{A}\to[0,\infty]$ such that*

(1) $\mu(\emptyset) = 0$,
(2) $\mu(S \cup T) = \mu(S) + \mu(T)$ *for all* $S, T \in \mathcal{A}$, *that is,* μ *is* finitely additive,
(3) *If* $S_1, S_2, \dots$ *are countably many elements of* $\mathcal{A}$, *so that their union* $S = \cup_{n=1}^{\infty} S_n$ *also lies in* $\mathcal{A}$, *then* $\mu(\cup_{n=1}^{\infty} S_n) \leq \sum_{n=1}^{\infty} \mu(S_n)$.
(4) *A premeasure is* finite *if* $\mu(X) < \infty$, *and it is* σ-finite *if* X *can be written as the union of countably many elements* $X_1, X_2, \dots \in \mathcal{A}$ *with* $\mu(X_n) < \infty$ *for all* n.

The axioms imply that
$$\mu(S) \leq \mu(T)$$
whenever $S \subset T$. It also follows from the axioms that if $S_1, S_2, \dots \in \mathcal{A}$ are pairwise disjoint subsets whose union lies in $\mathcal{A}$, then
$$\mu(\cup_{n=1}^{\infty} S_n) = \sum_{n=1}^{\infty} \mu(S_n).$$
To see this, we only have to verify that $\sum_{n=1}^{\infty} \mu(S_n) \leq \mu(\cup_{n=1}^{\infty} S_n)$, but this is true since for each m
$$\sum_{n=1}^{m} \mu(S_n) = \mu(\cup_{n=1}^{m} S_n) \leq \mu(\cup_{n=1}^{\infty} S_n)$$
by finite additivity.

Theorem C.22 (Caratheodory extension theorem)**.** *Let* $\mu_0 : \mathcal{A} \to [0, \infty]$ *be a* σ-*finite premeasure defined on a Boolean algebra on a set* X. *Then there is a unique measure* μ *on the* σ-*algebra* $\langle \mathcal{A} \rangle$ *generated by* $\mathcal{A}$ *such that* $\mu|_{\mathcal{A}} = \mu_0$.

The standard proof of this result is notorious for being completely unintuitive. We present an alternative proof that we learnt from [10, §2.1].

Outline of the proof in the case $\mu(X) < \infty$. We define the *outer measure* $\mu^* : \mathcal{P}(X) \to [0, \infty)$ by
$$\mu^*(S) := \inf_{S \subset \cup_{n=1}^{\infty} A_n, A_n \in \mathcal{A}} \sum_{n=1}^{\infty} \mu_0(A_n).$$
Furthermore, we define $\mathcal{B} \subset \mathcal{P}(X)$ as the set of all $S \in \mathcal{P}(X)$ such that for each $\epsilon > 0$, there is $T \in \mathcal{A}$ with
$$\mu^*(S \triangle T) \leq \epsilon,$$
and for $S \in \mathcal{B}$, we define
$$\mu(S) := \mu^*(S).$$
We have to prove that $\mathcal{B}$ is a σ-algebra which contains $\mathcal{A}$, that μ is a measure and that $\mu|_{\mathcal{A}} = \mu_0$, and that μ is uniquely determined by μ_0. □

The details of the proof are given in a sequence of lemmas.

Lemma C.23. *If* $(X, \mathcal{A})$ *and* μ_0 *are as in the theorem, and* $\mu_0(X) < \infty$, *and if* μ, μ' *are two extensions of* μ_0 *to a measure on* $\langle \mathcal{A} \rangle$, *then* $\mu = \mu'$.

Proof. Let $\mathcal{D} \subset \langle \mathcal{A} \rangle$ be the set of all S such that $\mu(S) = \mu'(S)$. Then $\mathcal{A} \subset \mathcal{D}$. It is easily verified that $\mathcal{D}$ is a Dynkin system, and that $\mathcal{A}$ is a π-system. Hence by the Dynkin lemma C.8, $\mathcal{D} = \langle \mathcal{A} \rangle$, as claimed. □

Lemma C.24 (Properties of the outer measure)**.** *We have*

(1) $\mu^*(A) = \mu_0(A)$ *for all* $A \in \mathcal{A}$,

(2) *$\mu^*(S) \leq \mu^*(T)$ if $S \subset T$,*
(3) *$\mu^*(\cup_{n=1}^\infty S_n) \leq \sum_{n=1}^\infty \mu^*(S_n)$.*

Proof. (1): Since $A \subset A \cup \emptyset \cup \emptyset \cup \ldots$, it is clear that $\mu^*(A) \leq \mu_0(A)$. For the reverse inequality, assume $A \subset \cup_{n=1}^\infty A_n$. Define $B_1 := A \cap A_1$ and $B_n := (A \cap A_n) \setminus \cup_{k=1}^{n-1}(A \cap A_k) \subset A_n$. Then $B_n \in \mathcal{A}$, and A is the disjoint union $\cup_{n=1}^\infty B_n$. It follows that

$$\mu_0(A) = \sum_{n=1}^\infty \mu_0(B_n) \leq \sum_{n=1}^\infty \mu_0(A_n),$$

and passage to the infimum proves the claim.

(2) is clear.

(3): Let $\epsilon > 0$ and let $S_n \subset \cup_{m=1}^\infty A_{n,m}$, $A_{n,m} \in \mathcal{A}$ with $\sum_{m=1}^\infty \mu_0(A_{n,m}) \leq \mu^*(S_n) + \frac{\epsilon}{2^n}$. Then $\cup_{n=1}^\infty S_n \subset \cup_{m,n=1}^\infty A_{n,m}$ and hence

$$\mu^*(\cup_{n=1}^\infty S_n) \leq \sum_{n,m=1}^\infty \mu_0(A_{n,m}) \leq \sum_{n=1}^\infty \mu^*(S_n) + \frac{\epsilon}{2^n} = (\sum_{n=1}^\infty \mu^*(S_n)) + \epsilon.$$

□

Lemma C.25. *For $R, S, T \subset X$, we have*
(1) *$\mu^*(R \triangle T) \leq \mu^*(R \triangle S) + \mu^*(S \triangle T)$.*
(2) *$|\mu^*(S) - \mu^*(T)| \leq \mu^*(S \triangle T)$.*

Proof. (1): Since $R \triangle T = (R \triangle S) \triangle (S \triangle T) \subset (R \triangle S) \cup (S \triangle T)$, we have

$$\mu^*(R \triangle T) \leq \mu^*((R \triangle S) \cup (S \triangle T)) \leq (R \triangle S) + \mu^*(S \triangle T).$$

(2): We have obviously $S = (S \cap T) \cup (S \setminus T) \subset (S \cap T) \cup (S \triangle T)$ and hence

$$\mu^*(S) \leq \mu^*(S \cap T) + \mu^*(S \triangle T) \leq \mu^*(T) + \mu^*(S \triangle T)$$

and by symmetry also

$$\mu^*(T) \leq \mu^*(S) + \mu^*(S \triangle T),$$

and the claim follows. □

Lemma C.26. (1) *$\mathcal{A} \subset \mathcal{B}$,*
(2) *$S \in \mathcal{B}$, then $S^c \in \mathcal{B}$,*
(3) *$S, T \in \mathbb{C}$, then $S \cup T \in \mathcal{B}$,*
(4) *If $S_1, S_2, \ldots \in \mathcal{B}$, then $\cup_{n=1}^\infty S_n \in \mathcal{B}$.*

In short, $\mathcal{B}$ is a σ-algebra which contains $\mathcal{A}$ and hence $\langle \mathcal{A} \rangle$.

Proof. (1): follows from the fact that $\mu^*(A) = \mu_0(A)$ when $A \in \mathcal{A}$. (2): is clear because $S^c \triangle A^c = S \triangle A$. (3): let $\epsilon > 0$ and pick $A, B \in \mathcal{A}$ with $\mu^*(S \triangle A), \mu^*(T \triangle B) \leq \epsilon$. It is easily seen that

$$(S \cap T) \triangle (A \cap B) \subset (S \triangle A) \cup (T \triangle B).$$

Therefore

$$\mu^*((S \cap T) \triangle (A \cap B)) \leq \mu^*(S \triangle A) + \mu^*(T \triangle B) \leq 2\epsilon.$$

(4) we first prove that $A_n \in \mathcal{A}$ implies $\cup_{n=1}^\infty A_n \in \mathcal{B}$. To that end, we can assume without loss of generality that the sets A_n are disjoint. Because

$$\sum_{n=1}^m \mu^*(A_n) = \mu^*(\cup_{n=1}^m A_n) \leq \mu^*(\cup_{n=1}^\infty A_n) \leq \mu_0(X) < \infty,$$

the series

$$\sum_{n=1}^{\infty} \mu^*(A_n)$$

is (absolutely) convergent. For $\epsilon > 0$, choose m so that

$$\sum_{n=m+1}^{\infty} \mu^*(A_n) \leq \epsilon.$$

Then $\cup_{n=1}^{m} A_n \in \mathcal{A}$, and

$$\mu^*((\cup_{n=1}^{\infty} A_n) \triangle (\cup_{n=1}^{m} A_n)) = \mu^*(\cup_{n=m+1}^{\infty} A_n) \leq \sum_{n=m+1}^{\infty} \mu^*(A_n) \leq \epsilon.$$

Therefore, $\cup_{n=1}^{\infty} A_n \in \mathcal{B}$. Now let $S_1, S_2, \ldots \in \mathcal{B}$. To show that $\cup_{n=1}^{\infty} S_n \in \mathcal{B}$, we can assume that all the sets S_n are disjoint, and pick $A_n \in \mathcal{A}$ with $\mu(A_n \triangle S_n) \leq \frac{\epsilon}{2^n}$. Because $(\cup_{n=1}^{\infty} A_n) \triangle (\cup_{n=1}^{\infty} S_n) \subset \cup_{n=1}^{\infty} (A_n \triangle S_n)$, we obtain

$$\mu^*((\cup_{n=1}^{\infty} A_n) \triangle (\cup_{n=1}^{\infty} S_n)) \leq \sum_{n=1}^{\infty} \mu^*(A_n \triangle S_n) \leq \epsilon,$$

which as $\cup_{n=1}^{\infty} A_n \in \mathcal{B}$ proves $\cup_{n=1}^{\infty} S_n \in \mathcal{B}$. □

Lemma C.27. *$\mu^* : \mathcal{B} \to [0, \infty)$ is a measure.*

Proof. We first prove finite additivity and assume that $S, T \in \mathcal{B}$ are disjoint. We already know that $\mu(S \cup T) \leq \mu^*(S) + \mu^*(T)$. For the reverse inequality, pick $A, B \in \mathcal{A}$ with $\mu^*(S \triangle A), \mu^*(T \triangle B) \leq \epsilon$. Then

$$\mu^*(S) + \mu^*(T) \leq \mu^*(A) + \mu^*(S \triangle A) + \mu^*(B) + \mu^*(T \triangle B) \leq \mu^*(A) + \mu^*(B) + 2\epsilon.$$

Note that

$$\mu^*(A) + \mu^*(B) = \mu^*(A \setminus B) + \mu^*(A \cap B) + \mu^*(B) = \mu^*(A \cup B) + \mu^*(A \cap B).$$

Because $S \cap T = \emptyset$, we have

$$A \cap B \subset (A \triangle S) \cup (B \triangle T)$$

and so $\mu^*(A \cap B) \leq 2\epsilon$, and therefore

$$\mu^*(A) + \mu^*(B) \leq \mu^*(A \cup B) + 2\epsilon.$$

On the other hand

$$\mu^*((A \cup B) \triangle (S \cup T)) \leq \mu^*(A \triangle S) + \mu^*(B \triangle T) \leq 2\epsilon$$

and therefore

$$\mu^*(A \cup B) \leq \mu^*(S \cup T) + 2\epsilon.$$

Putting everything together yields

$$\mu^*(S) + \mu^*(T) \leq \mu^*(A) + \mu^*(B) + 2\epsilon \leq \mu^*(A \cup B) + 4\epsilon \leq \mu^*(S \cup T) + 6\epsilon.$$

At this point, we have verified that $\mu^* : \mathcal{B} \to [0, \infty)$ is finitely additive. For the σ-additivity, let $S_1, S_2, \ldots \in \mathcal{B}$ be disjoint. For each m, we have

$$\sum_{n=1}^{m} \mu^*(S_n) = \mu^*(\cup_{n=1}^{m} S_n) \leq \mu^*(\cup_{n=1}^{\infty} S_n)$$

and passage $m \to \infty$ proves

$$\sum_{n=1}^{\infty} \mu^*(S_n) \le \mu^*(\cup_{n=1}^{\infty} S_n).$$

The reverse inequality is already proven. □

At this point, the proof in the case $\mu_0 < \infty$ is complete.

Proof in the σ-finite case. Choose subsets $X_1 \subset X_2 \subset \ldots \subset X$ with $\mu_0(X_n) < \infty$ and $\cup_{n=1}^{\infty} X_n = X$. For each n, let

$$\lambda_n : \mathcal{A} \to [0, \infty), \lambda_n(A) := \mu_0(A \cap X_n).$$

Then λ_n is a finite premeasure, and there are unique measures $\mu_n : \langle \mathcal{A} \rangle \to [0, \infty)$ extending λ_n. For $S \in \mathcal{B}$, we have $\mu_n(S) \le \mu_{n+1}(S)$ and so can define

$$\mu(S) := \lim_{n \to \infty} \mu_n(S) \in [0, \infty].$$

From the monotone convergence theorem C.32 for the counting measure[8], we see that μ is a measure, and one checks that $\mu|_{\mathcal{A}} = \mu_0$.

By construction, it is clear that $\mu_n(S) = \mu(S \cap X_n)$ (since the right hand side defines a measure on $\langle \mathcal{A} \rangle$ which extends λ_n), and so we *must* have defined $\mu(S)$ as above. This shows uniqueness. □

C.5. **The construction of the Lebesgue measure.** As a simple example for how the Caratheodory extension theorem works, let us present the construction of the Lebesgue measure, firstly on $[0, 1]$. Let $\mathcal{A} \subset \mathcal{P}([0, 1])$ be the set of all sets which can be written as the finite union of (open, half-open, or closed) intervals. We note that each element of $\mathcal{A}$ can be written as the finite union of *disjoint* intervals, and that $\mathcal{A}$ is a Boolean algebra on $\mathbb{R}$. Let $a \le b \in \mathbb{R}$ and let I be one of the sets $[a, b], (a, b), (a, b], [a, b)$. Then one easily checks that

$$\lim_{n \to \infty} \frac{1}{n} |I \cap \frac{1}{n}\mathbb{Z}| = b - a.$$

Now if $A \in \mathcal{A}$, we define the *elementary volume* by

$$\mu_0(A) := \lim_{n \to \infty} \frac{1}{n} |A \cap \frac{1}{n}\mathbb{Z}|.$$

Since each $A \in \mathcal{A}$ is a finite disjoint union of intervals, the limit exists and is equal to the sum of the lengths of the individual intervals. It is clear that μ_0 is finitely additive. Now let $A_1, A_2, \ldots$ are countably many elements of $\mathcal{A}$, so that their union $A = \cup_{n=1}^{\infty} A_n$ also lies in $\mathcal{A}$. We have to prove that $\mu(\cup_{n=1}^{\infty} A_n) \le \sum_{n=1}^{\infty} \mu(A_n)$. We can assume without of generality that A is an interval, and that each A_n is an interval. Let $\epsilon > 0$. There is a compact interval $B \subset A$ with $\mu_0(A) \le \mu_0(B) + \epsilon$, and there are open intervals $A_n \subset O_n$ with $\mu_0(O_n) \le \mu_0(A_n) + \frac{\epsilon}{2^n}$.

The sets O_n, $n \in \mathbb{N}$, form an open cover of B, and by compactness $B \subset \cup_{n=1}^{m} O_n$. It follows that

$$\mu_0(A) \le \mu_0(B) + \epsilon \le \sum_{n=1}^{m} \mu_0(O_n) + \epsilon \le \epsilon + \sum_{n=1}^{m} \mu_0(A_n) + \sum_{n=1}^{m} \frac{\epsilon}{2^n} \le 2\epsilon + \sum_{n=1}^{\infty} \mu_0(A_n).$$

Hence the Caratheodory extension theorem yields a Borel measure μ on $[0, 1]$ which extends μ_0. This μ is the Lebesgue measure. The Lebesgue measure on $\mathbb{R}$ can be

[8]Check for yourself that this forward reference does not produce a circular argument!

obtained by either taking the limit of the restrictions to $[-n,n]$, or by a slight modification of the above construction.

C.6. **The integral of nonnegative functions.** For the construction of the integral, we first study the integration of functions $X \to [0,\infty]$ with respect to a measure μ on X. For the rest of this section, we fix a measure space $(X,\mathcal{B},\mu)$.

Definition C.28. *A* nonnegative step function *on X is a measurable function $f : X \to [0,\infty)$ which only assumes finitely many values. We denote by* $\mathrm{St}_+(X)$ *the set of all nonnegative step functions on X.*

Alternatively, we can write $f = \sum_{j=1}^n a_n \chi_{S_n}$ as a finite linear combination of characteristic functions, but this representation is not unique.

Definition C.29. *Let f be a nonnegative step function and let $0 < a_1 < \ldots < a_r < \infty$ be the finitely many values it takes. We define*

$$\int_X f(x)d\mu(x) := \sum_{j=1}^r a_j \mu(f^{-1}(a_j)) \in [0,\infty].$$

Lemma C.30. *Let $f, g \in \mathrm{St}_+(X)$*

(1) *if $a \geq 0$, then $\int_X af d\mu = a\int_X f d\mu$.*
(2) *$\int_X (f+g)d\mu = \int_X f d\mu + \int_X g d\mu$.*
(3) *If $f \leq g$, then $\int_X f d\mu \leq \int_X g d\mu$.*
(4) *If $f \in \mathrm{St}_+(X)$, the function $\nu : \mathcal{B} \to [0,\infty]$, $\nu(S) := \int_X \chi_S f d\mu$ is a measure.*

(1) and (2) of the Lemma imply that

$$\int_X \sum_{j=1}^n a_j \chi_{S_j} d\mu = \sum_{j=1}^n a_j \mu(S_j).$$

It seems easier to *define* the integral by that formula, but this leaves the issue of showing that the result does not depend on the way in which f is written as a linear combination of characteristic functions.

Proof. The first claim is trivial. For the second one, note that we can write

$$\int_X f d\mu = \sum_{a\in[0,\infty)} a\mu(f^{-1}(a))$$

(this is a finite sum of course). Furthermore,

$$(f+g)^{-1}(c) = \bigcup_{a\in[0,c]} f^{-1}(a) \cap g^{-1}(c-a),$$

and this is a finite disjoint union. It follows that

$$\int_X (f+g)d\mu = \sum_{c\in[0,\infty)} c\mu((f+g)^{-1}(c)) = \sum_{c\in[0,\infty)} \sum_{a\in[0,c]} c\mu(f^{-1}(a)\cap g^{-1}(c-a)) =$$

$$\sum_{a\in[0,\infty)} \sum_{b\in[0,\infty)} (a+b)\mu(f^{-1}(a)\cap g^{-1}(b)) =$$

$$\sum_{a\in[0,\infty)} \sum_{b\in[0,\infty)} a\mu(f^{-1}(a)\cap g^{-1}(b)) + \sum_{a\in[0,\infty)} \sum_{b\in[0,\infty)} b\mu(f^{-1}(a)\cap g^{-1}(b)) =$$

$$\sum_{a\in[0,\infty)} a\mu(\bigcup_{b\in[0,\infty)} f^{-1}(a)\cap g^{-1}(b)) + \sum_{b\in[0,\infty)} b\mu(\bigcup_{a\in[0,\infty)} f^{-1}(a)\cap g^{-1}(b)) =$$

$$\sum_{a\in[0,\infty)} a\mu(f^{-1}(a)) + \sum_{b\in[0,\infty)} b\mu(g^{-1}(b)) = \int_X f d\mu + \int_Y g d\mu.$$

(3): the difference $g - f$ is an element of $\mathrm{St}_+(X)$. (4): Since a finite linear combination of measures with nonnegative coefficients is a measure, it is enough to consider the case $f = \chi_T$, $T \in \mathcal{B}$. It is straightforward to prove that ν is a measure in that case. □

Definition C.31. *Let $f : X \to [0,\infty]$ be measurable. We define*

$$\int_X f(x)d\mu(x) := \sup_{g\in \mathrm{St}_+(X), g\leq f} \int_X g(x)d\mu(x).$$

Two of the properties proved in Lemma C.30 carry over to the integral of non-negative functions without any difficulty. These are

(1) if $a \geq 0$, then $\int_X af d\mu = a\int_X f d\mu$.
(2) If $f \leq g$, then $\int_X f d\mu \leq \int_X g d\mu$.

The other properties also hold, but this requires more work.

Theorem C.32 (Monotone convergence theorem)**.** *Let $f_1 \leq f_2 \leq \ldots$ be an increasing sequence of nonnegative measurable functions on X. Then*

$$\lim_{n\to\infty} \int_X f_n(x)d\mu(x) = \int_X \lim_{n\to\infty} f_n(x)d\mu(x)$$

This is the key step for the development of the theory. Up to that point, we have not used the σ-additivity of μ, and it is being used in the proof of the monotone convergence theorem exactly once. The reader should look out to identify the step where it is used.

Proof. Put $f(x) := \lim_{n\to\infty} f_n(x)$, which is measurable. Since $\int_X f_n(x)d\mu(x) \leq \int_X f(x)d\mu(x)$, we obtain

$$\lim_{n\to\infty} \int_X f_n(x)d\mu(x) \leq \int_X f(x)d\mu(x).$$

For the reverse inequality, let $g \in \mathrm{St}_+(X)$ with $g \leq f$ and let $\epsilon > 0$. Put

$$S_n := \{x \in X | f_n(x) \geq (1-\epsilon)g(x)\} \subset X.$$

The set S_n is measurable, we have $S_1 \subset S_2 \subset \ldots$, and $\bigcup_{n=1}^\infty S_n = X$ (for the latter, note that $g^{-1}(0) \subset S_1$ and if $g(x) > 0$, then we have $f_n(x) \geq (1-\epsilon)g(x)$ for sufficiently large n). Since $S \mapsto \int_S (1-\epsilon)g(x)d\mu(x)$ is a measure, we know that

$$(1-\epsilon)\int_X g(x)d\mu(x) = \int_X (1-\epsilon)g(x)d\mu(x) = \lim_{n\to\infty} \int_{S_n} (1-\epsilon)g(x)d\mu(x).$$

Because $\int_{S_n} (1-\epsilon)g(x)d\mu(x) \leq \int_{S_n} f_n(x) \leq \int_X f_n(x)d\mu(x)$, we obtain

$$\lim_{n\to\infty} \int_{S_n} (1-\epsilon)g(x)d\mu(x) \leq \lim_{n\to\infty} \int_X f_n(x)d\mu(x).$$

Altogether

$$(1-\epsilon)\int_X g(x)d\mu(x) \leq \lim_{n\to\infty} \int_X f_n(x)d\mu(x).$$

This estimate holds for all $\epsilon > 0$, and so

$$\int_X g(x)d\mu(x) \leq \lim_{n\to\infty} \int_X f_n(x)d\mu(x).$$

This is true for all step functions $g \leq f$, and so

$$\int_X f(x)d\mu(x) \leq \lim_{n\to\infty} \int_X f_n(x)d\mu(x),$$

as desired. □

Lemma C.33. *Let $f : X \to [0,\infty]$ be measurable. Then there is a sequence $f_1 \leq f_2 \leq f_3 \dots$ of nonnegative step functions which converges pointwise to f.*

Proof. For $a > 0$, let $A_a := \{x \in X | f(x) \geq a\}$, which is measurable. The functions

$$f_n = \sum_{k=1}^{4^n} \frac{1}{2^n} \chi_{A_{\frac{k}{2^n}}}$$

do the job. □

Lemma C.34. *Let $f, g : X \to [0,\infty]$ be measurable.*

(1) *if $a \geq 0$, then $\int_X af d\mu = a \int_X f d\mu$.*
(2) *$\int_X f + g d\mu = \int_X f d\mu + \int_X g d\mu$.*
(3) *If $f \leq g$, then $\int_X f d\mu \leq \int_X g d\mu$.*
(4) *The function $\nu : \mathcal{B} \to [0,\infty]$, $\nu(S) := \int_X \chi_S f d\mu$ is a measure.*

Proof. Properties (1) and (3) have been used in the proof of the monotone convergence theorem and follow easily from the definitions. Properties (2) and (4) follow from Lemma C.30, Lemma C.33 and the monotone convergence theorem. □

Theorem C.35 (Fatou's Lemma)**.** *For sequences of measurable nonnegative functions, we have*

$$\int_X \liminf_n f_n(x)d\mu(x) \leq \liminf_n \int_X f_n(x)d\mu(x).$$

Proof. Put

$$g_n(x) := \inf_{k\geq n} f_k(x).$$

This function is measurable, and $g_1 \leq g_2 \leq \dots$, and $\lim_n g_n = \liminf_n f_n$. By the monotone convergence theorem, we see that

$$\int_X \liminf_n f_n(x)d\mu(x) = \lim_n \int_X g_n(x)d\mu(x).$$

For each $k \geq n$, we have $g_n \leq f_k$, and therefore

$$\int_X g_n(x)d\mu(x) \leq \inf_{k\geq n} \int_X f_k(x)d\mu(x),$$

which implies

$$\lim_n \int_X g_n(x)d\mu(x) \leq \liminf_n \int_X f_n(x)d\mu(x).$$

□

The following simple facts are sometimes useful.

Lemma C.36. *Let X be a measure space and $f : X \to [0,\infty]$.*

(1) *If* $\int_X f d\mu < \infty$, $a > 0$ *and* $A_a := \{x \in X | f(x) \geq a\}$, *then* $\mu(A_a) < \infty$.
(2) *If* $\int_X f d\mu = 0$, *then* $f(x) = 0$ *for almost all* x.

Proof. (1) this is simply the estimate

$$0 \leq a\mu(A_a) = \int_X a\chi_{A_a} d\mu \leq \int_X f d\mu.$$

(2) The above inequality proves that $\mu(A_a) = 0$ for all $a > 0$, and hence that

$$\{x | f(x) \neq 0\} = \bigcup_{n=1}^{\infty} A_{\frac{1}{n}}$$

has measure zero. □

C.7. **The Hölder and Minkowski inequality.** So far, we only have the notion of the integral of a nonnegative measurable function $X \to [0, \infty]$. Of course, we want to integrate real-valued, or complex-valued functions. The functions we will be able to integrate are the functions in $L^1(X, \mu)$. There are other spaces $L^p(X, \mu)$ for each $p \in [1, \infty]$, which are also important. As many arguments are parallel in the case $p = 1$ and $p > 1$, we start immediately with the case of an arbitrary p.

We consider a measure space (X, μ). For a measurable function $f : X \to [0, \infty]$ and $p \in [1, \infty)$, we define

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

Furthermore, we put

$$\|f\|_{L^\infty} := \inf_{\mu(S)=0} \sup_{x \in X - S} f(x) \in [0, \infty].$$

Theorem C.37 (Hölder inequality)**.** *Let* $f, g : X \to [0, \infty]$ *be measurable and let* $p, q \in [1, \infty]$ *satisfy* $\frac{1}{p} + \frac{1}{q} = 1$ *(with the convention that* $\frac{1}{\infty} = 0$*). Then*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Proof. In the following cases, the inequality is trivial: $p = 1$, $q = 1$, $\|f\|_{L^p} = \infty$, $\|f\|_{L^p} = 0$, $\|g\|_{L^q} = \infty$ and $\|g\|_{L^q} = 0$. Hence we may assume $p, q \in (1, \infty)$, and $0 < \|f\|_{L^p}, \|g\|_{L^q} < \infty$.

For $x, y \geq 0$, the Young inequality

$$xy \leq \frac{1}{p} x^p + \frac{1}{q} y^q$$

holds. It is trivial if $x = 0$ or $y = 0$. For $x, y > 0$, we use convexity of the exponential function

$$e^{\frac{1}{p}a + \frac{1}{q}b} \leq \frac{1}{p} e^a + \frac{1}{q} e^b$$

and insert $a = p \log(x)$, $b = q \log(y)$ to obtain

$$xy = e^{\log(x) + \log(y)} = e^{\frac{1}{p} p \log(x) + \frac{1}{q} q \log(y)} \leq \frac{1}{p} e^{p \log(x)} + \frac{1}{q} e^{q \log(y)} = \frac{1}{p} x^p + \frac{1}{q} y^q.$$

Integrating gives

$$\|fg\|_{L^1} = \int_X f g d\mu \leq \frac{1}{p} \int_X f^p d\mu + \frac{1}{q} \int_X g^q d\mu = \frac{1}{p} \|f\|_{L^p}^p + \frac{1}{q} \|g\|_{L^q}^q.$$

For $t > 0$ arbitrary, we obtain

$$\|fg\|_{L^1} = \|(tf)(\frac{g}{t})\|_{L^1} \leq \frac{1}{p}t^p\|f\|^p_{L^p} + \frac{1}{q}\frac{1}{t^q}\|g\|^q_{L^q} =: F(t), \tag{C.38}$$

and therefore

$$\|fg\|_{L^1} \leq \inf_{t\in(0,\infty)} F(t). \tag{C.39}$$

Since $\lim_{t\to 0} F(t) = \lim_{t\to\infty} F(t) = +\infty$, the infimum of F is attained at a t_0 with $F'(t_0) = 0$. But $F'(t_0) = 0$ holds if and only if

$$t_0 = (\frac{\|g\|^q_{L^q}}{\|f\|^p_{L^p}})^{\frac{1}{p+q}},$$

by an easy computation. Hence, inserting this value of t_0 into (C.38) yields

$$\|fg\|_{L^1} \leq \frac{1}{p}(\frac{\|g\|^q_{L^q}}{\|f\|^p_{L^p}})^{\frac{p}{p+q}}\|f\|^p_{L^p} + \frac{1}{q}(\frac{\|g\|^q_{L^q}}{\|f\|^p_{L^p}})^{\frac{-q}{p+q}}\|g\|^q_{L^q} =$$

$$= \frac{1}{p}\|g\|^{\frac{pq}{p+q}}_{L^q}\|f\|^{p-\frac{p^2}{p+q}}_{L^p} + \frac{1}{q}\|g\|^{q-\frac{-q^2}{p+q}}_{L^q}\|f\|^{\frac{pq}{p+q}}_{L^p}.$$

But $\frac{p+q}{pq} = \frac{1}{q} + \frac{1}{p} = 1$, $p - \frac{p^2}{p+q} = \frac{p^2+pq-p^2}{p+q} = \frac{pq}{p+q} = 1$ and $q - \frac{q^2}{p+q} = 1$, and so the last expression is

$$= \frac{1}{p}\|g\|_{L^q}\|f\|_{L^p} + \frac{1}{q}\|g\|_{L^q}\|f\|_{L^p} = \|f\|_{L^p}\|g\|_{L^q}.$$

□

Theorem C.40 (Minkowski inequality)**.** *If $p \in [1,\infty]$ and $f, g : X \to [0,\infty]$ are measurable, we have*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

Proof. The cases $p = 1, \infty$ are trivial, so we assume $p \in (1,\infty)$. Furthermore, the cases where $\|f + g\|_{L^p} = 0$, $\|f\|_{L^p} = \infty$ and $\|g\|_{L^p} = \infty$ are trivial, and we can exclude them. The crude estimate

$$\|f + g\|^p_{L^p} = \int_X |f + g|^p d\mu \leq 2^p \int_X \max(|f|^p, |g|^p)d\mu \leq$$

$$\leq 2^p \int_X (|f|^p + |g|^p)d\mu = 2^p(\|f\|_{L^p} + \|g\|_{L^p})$$

shows that $\|f\|_{L^p}, \|g\|_{L^p} < \infty$ implies $\|f + g\|_{L^p} < \infty$. Compute

$$\|f+g\|^p_{L^p} = \int_X (f+g)^{p-1}fd\mu + \int_X (f+g)^{p-1}gd\mu = \|(f+g)^{p-1}f\|_{L^1} + \|(f+g)^{p-1}g\|_{L^1}.$$

Let $q = (1 - \frac{1}{p})^{-1}$. The Hölder inequality implies

$$\|f+g\|^p_{L^p} \leq \|(f+g)^{p-1}\|_{L^q}\|f\|_{L^p} + \|(f+g)^{p-1}\|_{L^q}\|g\|_{L^p} = \|(f+g)^{p-1}\|_{L^q}(\|f\|_{L^p} + \|g\|_{L^p}).$$

But $\frac{p}{q} = p(1 - \frac{1}{p}) = p - 1$ and $(p - 1)q = p$, so that

$$\|(f + g)^{p-1}\|_{L^q} = (\int_X (f + g)^{(p-1)q})^{\frac{1}{q}} = ((\int_X (f + g)^p)^{\frac{1}{p}})^{\frac{p}{q}} = \|f + g\|^{p-1}_{L^p}.$$

Now divide by $\|f + g\|^{p-1}_{L^p} \in (0,\infty)$. □

C.8. L^p-spaces.

Definition C.41. *Let $p \in [1, \infty]$. We define $\mathscr{L}^p(X, \mathbb{K})$ as the space of all functions $f : X \to \mathbb{K}$ such that*

(1) *f is measurable,*
(2) *$\|f\|_{L^p} = \||f|\|_{L^p} < \infty$.*

It follows from Theorem C.40 that $\mathscr{L}^p(X, \mathbb{K})$ is a vector space and $\|\_\|_{L^p}$ is a seminorm. The space $\mathscr{N}(X, \mathbb{K}) := \{f \in \mathscr{L}^p(X, \mathbb{K}) | \|f\|_{L^p} = 0\}$ is exactly the subspace of all null functions, i.e. of all functions which vanish almost everywhere.

Theorem C.42. *Let $f_n \in \mathscr{L}^p(X, \mathbb{K})$ be an L^p-Cauchy sequence. Then*

(1) *There is a subsequence which converges pointwise almost everywhere to a function $f : X \to \mathbb{K}$.*
(2) *The limit function belongs to $\mathscr{L}^p(X, \mathbb{K})$.*
(3) $\lim_n \|f_n\|_{L^p} = \|f\|_{L^p}$, $\|f - f_n\|_{L^p} = 0$.

Proof. We give the detailed proof in the case $p < \infty$. The case $p = \infty$ is easier.

(1) After passage to a subsequence, we can assume that

$$\|f_{n+1} - f_n\|_{L^p} \leq \frac{1}{4^n}$$

for all n. Let

$$X_n := \{x \in X | |f_{n+1}(x) - f_n(x)| \geq \frac{1}{2^n}\} \subset X,$$

which is a measurable subset. We estimate

$$(\frac{1}{4^n})^p \geq \|f_{n+1} - f_n\|_{L^p}^p \geq \int_{X_n} |f_{n+1}(x) - f_n(x)|^p d\mu(x) \geq \mu(X_n)(\frac{1}{2^n})^p$$

and therefore, we get

$$\mu(X_n) \leq (\frac{1}{2^n})^p.$$

Let

$$Y_n := \bigcup_{m \geq n} X_m$$

and

$$Y := \bigcap_{n=1}^{\infty} Y_n$$

(note that $Y_1 \supset Y_2 \supset \ldots$). Then

$$\mu(Y_n) \leq (\frac{1}{2^n})^p \frac{1}{1 - (\frac{1}{2})^p},$$

and hence

$$\mu(Y) = 0.$$

If $x \in Y^c$, then $x \in Y_n^c = \bigcap_{m=n}^{\infty} X_m^c$ for some n, and so for $m \geq n$, we have

$$|f_{m+1}(x) - f_m(x)| \leq \frac{1}{2^m}.$$

Since $\mathbb{K}$ is complete, this implies that $f_m(x)$ converges. Hence $f_m|_{Y^c}$ is pointwise convergent.

(2): Being the pointwise limit of measurable functions, f is measurable. By Fatou's lemma

$$\int_X |f(x)|^p d\mu(x) \leq \liminf_n \int_X |f_n(x)|^p d\mu(x) = \liminf_n \|f_n\|_{L^p}^p < \infty.$$

(3): Again by Fatou's lemma

$$\|f_n - f\|_{L^p}^p = \int_X \lim_m |f_n(x) - f_m(x)|^p d\mu(x) \leq \liminf_m \|f_n(x) - f_m(x)\|_{L^p}^p \leq \epsilon$$

for sufficiently large n. On the other hand

$$\|f_n\|_{L^p} \leq \|f_n - f\|_{L^p} + \|f\|_{L^p}$$

and therefore

$$\limsup_n \|f_n\|_{L^p} \leq \|f\|_{L^p},$$

while

$$\|f\|_{L^p} \leq \|f - f_n\|_{L^p} + \|f_n\|$$

and so

$$\|f\|_{L^p} \leq \liminf_n \|f - f_n\|_{L^p} + \liminf_n \|f_n\| = \liminf_n \|f_n\|.$$

□

Theorem C.43 (Dominated convergence theorem)**.** *Let $p \in [1, \infty)$. Let $f_n \in \mathscr{L}^p(X, V)$ be a sequence which converges pointwise almost everywhere to a function $f : X \to \mathbb{K}$. Assume that there is a measurable function $g : X \to [0, \infty]$ with $\int_X g(x)^p d\mu < \infty$, and assume that $|f_n(x)| \leq g(x)$ for all n and almost all x. Then $f \in \mathscr{L}^p(X, \mathbb{K})$ and $\|f_n - f\|_{L^p} \to 0$.*

Proof. Without loss of generality, we can assume that $|f_n(x)| \leq g(x)$ and $\lim_n |f_n(x) - f(x)| = 0$ for all x. We have $0 \leq |f_n(x) - f_m(x)| \leq 2g(x)$ and

$$\lim_n \limsup_m |f_n(x) - f_m(x)| = 0.$$

Therefore, by a repeated application of Fatou's lemma,

$$\begin{aligned}
\int_X 2^p g(x)^p d\mu &= \int_X 2^p g(x)^p - \lim_n \limsup_m |f_n(x) - f_m(x)|^p d\mu = \\
&= \int_X \lim_n (2^p g(x)^p - \limsup_m |f_n(x) - f_m(x)|^p) d\mu \leq \\
&\leq \liminf_n \int_X (2^p g(x)^p - \limsup_m |f_n(x) - f_m(x)|^p) d\mu = \\
&= \liminf_n \int_X \liminf_m (2^p g(x)^p - |f_n(x) - f_m(x)|^p) d\mu \leq \\
&\leq \liminf_n \liminf_m \int_X (2^p g(x)^p - |f_n(x) - f_m(x)|^p) d\mu = \\
&= \int_X 2^p g(x)^p d\mu - \limsup_n \limsup_n \int_X |f_n(x) - f_m(x)|^p d\mu.
\end{aligned}$$

We conclude that

$$\limsup_n \limsup_n \int_X |f_n(x) - f_m(x)|^p d\mu = 0.$$

Therefore f_n is an L^p-Cauchy sequence, which implies all claims by Theorem C.42.

□

Lemma C.44. *Let $\mathscr{N}(X,\mathbb{K})$ be the space of null functions $X \to \mathbb{K}$, i.e. functions which are nonzero only on a set of measure* 0. *Let $p \in [1,\infty)$. Then*

$$\{f \in \mathscr{L}^p(X,\mathbb{K}) | \|f\|_{L^p} = 0\} = \mathscr{N}(X,\mathbb{K}).$$

Definition C.45. *We let $L^p(X,\mathbb{K}) := \mathscr{L}^p(X,\mathbb{K})/\mathscr{N}(X,\mathbb{K})$ with the induced quotient norm.*

The space $L^p(X,\mathbb{K})$ is a normed space, since we divided out all elements of norm 0. Moreover

Proposition C.46. *The normed space $L^p(X,\mathbb{K})$ is complete.*

Proof. For sake of notational clarity, we denote by the letters f_n and f measurable functions in $\mathscr{L}^p(X)$, and by $[v_n],[v] \in L^p(X)$ their *equivalence classes*. Typical elements of $L^p(X)$ will be denoted v_n, v.

Let now v_n be a Cauchy sequence in $L^p(X)$. We have to prove that v_n converges, and for that, it is enough to prove that some subsequence is convergent. Pick representatives $f_n \in \mathscr{L}^p(X)$, i.e. $[f_n] = v_n$. Then $f_n \in \mathscr{L}^p(X)$ is an L^p-Cauchy sequence, and by Theorem C.42, we can assume that f_n converges in the L^p-norm to some $f \in \mathscr{L}^p(X)$. Then v_n converges in $L^p(X)$ to $v := [f] \in L^p(X)$. □

Definition C.47. *Let $\mathrm{St}(X,\mathbb{K})$ be the set of all step functions $X \to \mathbb{K}$, i.e. all measurable functions which can be written in the form $f = \sum_{j=1}^r a_j \chi_{S_j}$, where $S_j \subset X$ is a measurable subset and $a_j \in \mathbb{K}$. We denote by $\mathrm{St}_f(X,\mathbb{K}) \subset \mathrm{St}(X,\mathbb{K})$ the subset of all step functions f with $\mu(\{x | f(x) \neq 0\}) < \infty$.*

Proposition C.48. *If $p \in [1,\infty)$, then $\mathrm{St}_f(X) \subset \mathscr{L}^p(X)$ is a dense subspace. If $p = \infty$, then $\mathrm{St}(X) \subset \mathscr{L}^\infty(X)$ is dense.*

Proof. Let us first assume that $\mu(X) < \infty$. Let $f \in \mathscr{L}^p(X)$ and $\epsilon > 0$. Put $\delta := \frac{\epsilon}{\mu(X)^{1/p}}$. Let $Z \subset \mathbb{K}$ be a countable dense subset. Pick $z_i \in Z$, $i \in \mathbb{N}$, such that $f(X) \subset \bigcup_{i=1}^\infty B_\delta(z_i)$. We can write X as a disjoint union of measurable sets X_i, so that $|f(x) - z_i| \leq \delta$ for $x \in X_i$. The measurable function

$$g = \sum_{i=1}^\infty z_i \chi_{X_i}$$

satisfies $|g(x) - f(x)| \leq \delta$ for all x, and hence

$$\|g - f\|_{L^p}^p \leq \delta \mu(X)^{1/p} = \epsilon.$$

Furthermore

$$g_n = \sum_{i=1}^n z_i \chi_{X_i}$$

is a step function, and $|g_n| \to |g|$ monotonously increasing. Hence the monotone convergence theorem implies that

$$\lim_n \|g - g_n\|_{L^p} = 0.$$

Putting everything together, we find a step function $h := g_n$ such that

$$\|h - f\|_{L^p} \leq \epsilon.$$

This finishes the proof when $\mu(X) < \infty$.

In the general case, we consider for $m \in \mathbb{N}$ the set $S_m := \{x||f(x)|^p \geq \frac{1}{m}\}$. Then

$$\mu(S_m)\frac{1}{m} \leq \int_X |f|^p d\mu,$$

and so S_m has finite measure. Since $\lim_m \int_{S_m} |f|^p d\mu = \int_X |f|^p d\mu$, there is m so that

$$\int_X |f|^p d\mu \geq \int_{S_m} |f|^p d\mu \geq \int_X |f|^p d\mu - \epsilon.$$

It follows that there is $T \subset X$ of finite measure such that

$$\|f - \chi_T f\|_{L^p}^p \leq \epsilon/3.$$

It suffices to approximate $\chi_T f$ by elements of $\mathrm{St}_f(X)$ with arbitrary precision, and this has been done in the first step of the proof. □

C.9. **Construction of the integral.** After the work from the previous section, we are ready to define the integral of a function $f \in \mathscr{L}^1(X, \mathbb{K})$.

Definition C.49. *Let $f \in \mathrm{St}_f(X, \mathbb{K})$. We define*

$$\int_X f d\mu = \sum_{v \in V\mathbb{K}} v\mu(f^{-1}(v)) \in \mathbb{K}$$

(which is a finite sum).

Lemma C.50. *The map $\mathrm{St}_f(X, \mathbb{K}) \to \mathbb{K}$ given by $f \mapsto \int_X f d\mu$ is linear, and*

$$|\int_X f d\mu| \leq \|f\|_{L^1}.$$

Proof. It is clear that $\int_X af d\mu = a \int_X f d\mu$ when $f \in \mathrm{St}_f(X, \mathbb{K})$ and $a \in \mathbb{K}$. For the additivity, we note that

$$(f+g)^{-1}(v) = \bigcup_{u \in \mathbb{K}} f^{-1}(u) \cap g^{-1}(v-u)$$

(finite disjoint union) and compute

$$\int_X (f+g) d\mu = \sum_{v \in \mathbb{K}} v\mu((f+g)^{-1}(v)) = \sum_{v \in \mathbb{K}} \sum_{u \in \mathbb{K}} (u + (v-u))\mu(f^{-1}(u) \cap g^{-1}(v-u)) =$$

$$= \sum_{w \in \mathbb{K}} \sum_{u \in \mathbb{K}} (u+w)\mu(f^{-1}(u) \cap g^{-1}(w)) =$$

$$= \sum_{w \in \mathbb{K}} \sum_{u \in \mathbb{K}} u\mu(f^{-1}(u) \cap g^{-1}(w)) + \sum_{w \in \mathbb{K}} \sum_{u \in V} w\mu(f^{-1}(u) \cap g^{-1}(w)) =$$

$$= \sum_{u \in \mathbb{K}} u\mu(f^{-1}(u)) + \sum_{w \in \mathbb{K}} w\mu(g^{-1}(w)) = \int_X f d\mu + \int_X g d\mu.$$

From the linearity, we conclude

$$\int_X \sum_{j=1}^r v_j \chi_{S_j} d\mu = \sum_{j=1}^r v_j \mu(S_j),$$

and this easily implies the estimate. □

It follows that $\int_X \_d\mu$ extends to a linear map $\mathscr{L}^1(X, \mathbb{K}) \to \mathbb{K}$ and descends to a linear map $L^1(X, \mathbb{K}) \to \mathbb{K}$. We need to perform a last consistency check:

Lemma C.51. *Assume that $f : X \to [0, \infty)$ belongs to $\mathscr{L}^1(X, \mathbb{R})$. Then the integral of f as a nonnegative function and as an element of $\mathscr{L}^1(X, \mathbb{R})$ agree.*

Proof. For the purpose of this proof, let $\int'_X f d\mu$ be the integral of f as a nonnegative function. It is clear that

$$\int'_X g d\mu = \int_X g d\mu$$

when $g \in \mathrm{St}_+(X) \cap \mathrm{St}_f(X, \mathbb{R})$. By Lemma C.33, we find a sequence f_n of non-negative step functions which converges monotonously to f. By the monotone convergence theorem, we have

$$\int'_X f d\mu = \lim_n \int'_X f_n d\mu = \lim_n \int_X f_n d\mu.$$

It follows that f_n is an L^1-Cauchy sequence, and therefore $\lim_n \int_X f_n d\mu = \int_X f d\mu$. $\square$

Proposition C.52. *Let X be σ-finite. Let $f \in L^1(X, \mu, \mathbb{K})$ and let $K \subset \mathbb{K}$ be a closed convex subset. The following are equivalent:*

(1) *For each measurable S with $0 < \mu(S) < \infty$, we have*

$$\frac{1}{\mu(S)} \int_S f d\mu \in K.$$

(2) *For almost all $x \in X$, $f(x) \in K$.*

Proof. $2 \Rightarrow 1$: We first consider the case of a step function f. We write

$$f = \sum_{j=1}^{n} a_j \chi_{T_j}$$

with *disjoint* sets T_j such that $\bigcup_{j=1}^n T_j = X$ and $\mu(T_j) > 0$ (we possibly have to include $f^{-1}(0)$, which might have infinite measure). Then $a_1, \ldots, a_n \in K$, and

$$\frac{1}{\mu(S)} \int_S f d\mu = \frac{1}{\mu(S)} \sum_{j=1}^{n} \mu(S \cap T_j) a_j = \sum_{j=1}^{n} \frac{\mu(S \cap T_j)}{\mu(S)} a_j$$

shows that the left hand side is a convex combination of the points $a_1, \ldots, a_n$ and hence belongs to K. The case of a general f is done by picking a sequence $f_n \to f$ of step functions which converge in L^1-norm to f.

$1 \Rightarrow 2$: The complement $\mathbb{K} \setminus K$ can then be written as a countable union of closed convex sets. If $f(x) \in K$ does not hold for almost all x, we find a closed convex set $L \subset \mathbb{K}$ with $L \cap K = \emptyset$ such that $S := f^{-1}(L)$ has positive finite measure. Then by the already proven implication, we get

$$\frac{1}{\mu(S)} \int_S f d\mu \in L,$$

and hence not in K. $\square$

C.10. **Product measures and the Fubini theorem.** Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two complete measure spaces. A set of the form $S \times T$, $S \in \mathcal{A}$ and $T \in \mathcal{B}$, is called a *measurable box*. We define $\mathcal{A} \otimes \mathcal{B}$ as the σ-algebra generated by all measurable boxes. Note that

$$(S \times T) \cap (S' \times T') = (S \cap S') \times (T \cap T'),$$

so that the measurable boxes form a π-system. For a subset $S \subset X \times Y$, $x \in X$ and $y \in Y$, we define

$$S_x := \{y \in Y | (x, y) \in S\} \subset Y$$

and

$$S^y := \{x \in X | (x, y) \in S\} \subset X.$$

Lemma C.53. *Assume that X and Y are both σ-finite. Then for each $S \in \mathcal{A} \otimes \mathcal{B}$, the following statements hold:*

(1) *For almost all $x \in X$, the set S_x belongs to $\mathcal{B}$.*
(2) *For almost all $Y \in y$, the set S^y belongs to $\mathcal{A}$.*
(3) *The (almost everywhere defined) function $f_S : X \to [0, \infty]$, $f_S(x) := \nu(S_x)$ is measurable.*
(4) *The (almost everywhere defined) function $g_S : Y \to [0, \infty]$, $g_S(y) := \mu(S^y)$ is measurable.*
(5) *We have $\int_Y g_S(y) d\nu(y) = \int_X f_S(x) d\mu(x)$.*
(6) *The formula $(\mu \otimes \nu)(S) := \int_X f_S(x) d\mu(x)$ is a measure on the σ-algebra $\mathcal{A} \otimes \mathcal{B}$.*

Proof. Let $X_1 \subset X_2 \subset \ldots X$ and $Y_1 \subset Y_2 \subset \ldots Y$ be exhaustions of X and Y by subsets of finite measure. We let $\mathcal{C} \subset \mathcal{P}(X \times Y)$ be the set of all $S \subset X \times Y$ such that S and all the sets $S \cap (X_n \times Y_n)$ satisfy the statements (1)–(5). Clearly, $\mathcal{C}$ contains the set $\mathcal{M}$ of all measurable boxes. We will prove that $\mathcal{C}$ is a Dynkin system. As $\mathcal{M}$ is a π-system, the Dynkin lemma will then imply that $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{C}$, and this finishes the proof of the claims (1)–(5). It is clear that $X \times Y \in \mathcal{C}$. Assume that $S_n \in \mathcal{C}$, $n \in \mathbb{N}$. Then $S = \bigcup_{n=1}^{\infty} S_n \in \mathcal{C}$: claims (1) and (2) follows from the fact that $\mathcal{A}$ and $\mathcal{B}$ are σ-algebras, claims (3) and (4) from the fact that pointwise limits of measurable functions are measurable, and claim (5) from the monotone convergence theorem.

Now let $S, T \in \mathcal{C}$ and $S \subset T$. The task is to prove that $T - S \in \mathcal{C}$. Let $S_n := S \cap (X_n \times Y_n)$ and define T_n analogously. It is clear that $(T - S)_x = T_x - S_x$ and $(T - S)^y = T^y - S^y$ are measurable for almost all x or y, so that the functions f_{T-S} and g_{T-S} are almost everywhere defined.

Since $f_{S_n} + f_{T_n - S_n} = f_{T_n}$ and $g_{S_n} + g_{T_n - S_n} = g_{T_n}$, and as f_{S_n} and g_{S_n} are finite, it follows that $f_{T_n - S_n}$ and $g_{T_n - S_n}$ are measurable. The same is then true for $g_{T-S} = \lim_n g_{T_n - S_n}$ and f_{T-S}. Furthermore, the computation

$$\int_Y g_{S_n}(y) d\nu(y) + \int_Y g_{T_n - S_n}(y) d\nu(y) = \int_Y g_{T_n}(y) d\nu(y) =$$
$$= \int_X f_{T_n}(y) d\mu(y) = \int_X f_{S_n}(y) d\mu(y) + \int_X f_{T_n - S_n}(y) d\mu(y)$$

finishes the verification that $T - S \in \mathcal{C}$. This also finishes the proof that all sets in $\mathcal{A} \otimes \mathcal{B}$ satisfy (1)–(5).

It is an easy application of the monotone convergence theorem that $\mu \otimes \nu$ is a measure. $\square$

For a function $f : X \times Y \to \mathbb{K}$, we denote by $f_x : Y \to \mathbb{K}$ the function $f_x(y) := f(x,y)$, and similarly $f^y(x) := f(x,y)$.

Theorem C.54 (Tonelli and Fubini theorem)**.** *Let (X,μ) and (Y,ν) be σ-finite measure spaces.*

(1) *If $f : X \times Y \to [0,\infty]$ is measurable, then*

$$\int_{X\times Y} f d\mu \otimes \nu = \int_X (\int_Y f_x(y) d\nu(y)) d\mu(x) = \int_Y (\int_X f^x(y) d\mu(x)) d\nu(y).$$

(2) *Assume that $f \in \mathscr{L}^1(X \times Y)$. Then for almost all $x \in X$, the function $f_x : Y \to \mathbb{K}$, $f_x(y) := f(x,y)$, belongs to $\mathscr{L}^1(Y,\mathbb{K})$, the (almost everywhere defined) function $g : X \to \mathbb{K}$, $g(x) := \int_Y f_x(y)d\nu(y)$ is in $\mathscr{L}^1(X,\mathbb{K})$, and*

$$\int_{X\times Y} f(x,y) d\mu \otimes \nu(x,y) = \int_X g(x) d\mu(x).$$

(3) *If $f : X \times Y \to \mathbb{K}$ is measurable and $\int_X(\int_Y |f(x,y)|d\nu(y))d\mu(x) < \infty$, then $f \in \mathscr{L}^1(X \times Y)$.*

Proof. (1) By Lemma C.33, there is an increasing sequence of step functions which converge pointwise to f. For a step function, statement (1) is an easy consequence of Lemma C.53, and the general case follows from the monotone convergence theorem.

(2) Let $\mathcal{F} \subset \mathscr{L}^1(X\times Y)$ be the subspace of function which satisfy the conclusion of the theorem. It is clear that $\mathcal{F}$ is a linear subspace.

Each characteristic function $f = \chi_S$ with $\mu \otimes \nu(S) < \infty$ belongs to $\mathcal{F}$. This follows immediately from Lemma C.53. Therefore by linearity $\mathrm{St}_f(X \times Y) \subset \mathcal{F}$.

Therefore, it remains to be proven that L^1-limits of functions in $\mathcal{F}$ belong to $\mathcal{F}$, so let f_n be an L^1-Cauchy sequence of step functions which converges almost everywhere to f. Then for almost all x, $(f_n)_x$ converges almost everywhere to f_x. By part (1)

$$\|f_n - f\|_{L^1} = \int_X \|(f_n)_x - f_x\|_{L^1} d\mu(x),$$

and so the functions $h_n : X \to [0,\infty]$, $h_n(x) := \|(f_n)_x - f_x\|_{L^1}$ converge to 0 in the L^1-norm. By Theorem C.42, it follows that $\|(f_n)_x \to f_x\|_{L^1} \to 0$ for almost all x (after passing to a subsequence, of course). It follows that $f_x \in \mathscr{L}^1(Y)$ for almost all x. It also follows that $g_n(x) \to g(x)$ for almost all x. Moreover

$$\|g_n - g_m\|_{L^1(X)} = \|f_n - f_m\|_{L^1(X\times Y)},$$

and so g_n is an L^1-Cauchy sequence. Therefore $g \in \mathscr{L}^1(X)$. Finally

$$\int_X g(x)d\mu(x) = \lim_n \int_X g_n(x)d\mu(x) = \lim_n \int_{X\times Y} f_n(x,y)d\mu\otimes\nu(x,y) = \int_{X\times Y} f(x,y)d\mu\otimes\nu(x,y).$$

(3) follows from (1), (2). $\square$

Lemma C.55. *Let (X,μ) be σ-finite and let $f : X \to [0,\infty]$ be measurable. For $\epsilon > 0$, let*

$$S_\epsilon := \{x \in X | f(x) \geq \epsilon\}.$$

Then

$$\int_X f d\mu = \int_0^\infty \mu(S_\epsilon) d\epsilon.$$

Proof. Let
$$Y := \{(x,t) \in X \times \mathbb{R} | 0 \leq t \leq f(x)\} \subset X \times \mathbb{R}.$$
This is measurable. Let λ be the Lebesgue measure on $\mathbb{R}$. By Fubini, we have
$$(\mu \otimes \lambda)(Y) = \int_X \lambda(Y_x) d\mu(x) = \int_X f(x) d\mu(x),$$
as well as
$$(\mu \otimes \lambda)(Y) = \int_0^\infty \mu(Y^t) d\lambda(t) = \int_0^\infty \mu(S_t) d\lambda(t).$$
□

C.11. **Borel measures.**

Definition C.56. *Let X be a topological space. The* Borel-σ-algebra *is the smallest σ-algebra which contains all open subsets of X. A* Borel measure *on X is a measure which is defined on the Borel σ-algebra.*

Definition C.57. *Let X be a locally compact Hausdorff space. A Borel measure μ on X is a* Radon measure *if*

(1) *$\mu(K) < \infty$ for all compact $K \subset X$ ("μ is locally finite"),*
(2) *for each Borel set $S \subset X$, we have*
$$\mu(S) = \sup_{K \subset S \text{ compact}} \mu(K)$$
("μ is inner regular") and
(3) *for each Borel set $S \subset X$, we have*
$$\mu(S) = \inf_{S \subset U \text{ open}} \mu(U).$$
("μ is outer regular").

Proposition C.58. *Let X be a locally compact Hausdorff space and let μ be a Radon measure on X. Then $C_c(X) \subset L^p(X,\mu)$ for each p, and if $p < \infty$, then this is a dense linear subspace.*

Proof. Let $\mathcal{B}$ be the Borel-σ-algebra and let $\mathcal{T}$ be the topology on X. Since $\mathcal{T} \subset \mathcal{B}$, each continuous function $X \to \mathbb{K}$ is measurable. If $f \in C_c(X)$, then $\mu(\mathrm{supp}(f)) < \infty$, and it follows that $|f| \leq \|f\|_{C^0} \mu(\mathrm{supp}(f)) < \infty$, so that $f \in L^p(X,\mu)$.

For the density, it suffices to prove that for each measurable subset $S \subset X$ with $\mu(S) < \infty$ and for each $\epsilon > 0$, there is $f \in C_c(X)$ with $\|f - \chi_S\|_{L^p} \leq \epsilon$.

Since μ is regular, there is a compact $K \subset S$ and an open $S \subset U$ with $\mu(U - K) \leq \epsilon$. A consequence of Urysohn's Lemma (Proposition B.79) yields a continuous function $f : X \to [0,1]$ with $\chi_K \leq f \leq \mu_U$. Then
$$\|f - \chi_S\|_{L^p} \leq \|f - \chi_K\|_{L^p} + \|\chi_S - \chi_K\|_{L^p}.$$
But $0 \leq f - \chi_K \leq \chi_U - \chi_K$, and $0 \leq \chi_S - \chi_K \leq \chi_U - \chi_K$ and so
$$\|f - \chi_S\|_{L^p} \leq 2\|\chi_U - \chi_K\|_{L^p} = 2(\mu(U \setminus K))^{1/p} \leq 2\epsilon^{1/p}.$$
□

Proposition C.59. *Let X be a* second countable *locally compact space and let μ be a locally finite Borel measure. Then μ is a Radon measure.*

Proof. Let us first assume that X is compact. Then μ is finite. Let $\mathcal{D} \subset \mathcal{B}$ be the set of all S such that $\mu(S) = \sup_{K \subset S \text{ compact}} \mu(K) = \inf_{S \subset U \text{ open}} \mu(U)$.

(1) Each open $U \subset X$ belongs to $\mathcal{D}$. Outer regularity is not an issue. By Corollary B.86, X is metrizable. Let $K_n := \{x \in U | \operatorname{dist}_{U^c}(x) \leq \frac{1}{n}\}$. Then K_n is compact, $K_n \subset K_{n+1}^\circ \subset K_{n+1} \subset \ldots$ and $\bigcup_n K_n = U$. Let $f_n : X \to [0,1]$ be continuous with $f_n|_{K_n} = 1$ $\operatorname{supp}(f_n) \subset K_{n+1}$. Then $f_n \to \chi_U$ pointwise almost everywhere, and the dominated convergence theorem implies
$$\mu(U) = \lim_{n\to\infty} \int_X f_n d\mu.$$
But $\int_X f_n d\mu \leq \mu(K_{n+1})$, whence $\lim_{n\to\infty} \mu(K_{n+1}) = \mu(U)$.

(2) Intersections of elements of $\mathcal{D}$ are in $\mathcal{D}$. To see this, note that $S \in \mathcal{D}$ if and only if there are compact K_n and open U_n with $K_n \subset S \subset U_n$, so that χ_{K_n} and χ_{U_n} converge almost everywhere to χ_S. If T is another set in $\mathcal{D}$, consider $L_n \subset T \subset V_n$ and observe that $\chi_{K_n \cap L_n} = \chi_{K_n}\chi_{L_n}$ converges almost everywhere to $\chi_{S\cap T}$, and so does $\chi_{U_n \cap V_n}$.

(3) Complements of elements in $\mathcal{D}$ are in $\mathcal{D}$. This is clear, and uses that we assumed X to be compact.

(4) Ascending unions of elements in $\mathcal{D}$ are in $\mathcal{D}$. Let $S_1 \subset S_2 \ldots$ be a sequence in $\mathcal{D}$ with $S = \bigcup_n S_n$. Pick $K_n \subset S_n \subset U_n$ with $\mu(U_n \setminus K_n) \leq \frac{\epsilon}{n}$. Then $K_n \subset S$, and $\mu(S \setminus K_n)$ becomes arbitrarily small, and $S \subset \bigcup_n U_n$. Then
$$\mu(\bigcup_{n=1}^m U_n \setminus \bigcup_{n=1}^m S_n) \leq \mu(\bigcup_n (S_n \setminus K_n)) \leq \sum_n \mu(S_n \setminus K_n) \leq \epsilon.$$

(2), (3) and (4) prove that $\mathcal{D}$ is a Dynkin system and a π-system, hence a σ-algebra by Lemma C.7, and since $\mathcal{D}$ contains all open sets, $\mathcal{D} = \mathcal{B}$. This finishes the proof in the compact case.

In the general case, pick an exhaustion $X_1 \subset X_1^\circ \subset X_2 \subset \ldots X$ of X by compact sets. The measure μ induces a measure μ_n on X_n by $\mu_n(S) := \mu(S)$. The measure μ_n is inner and outer regular by the first part of the proof.

For a measurable S, let $S_n := S \cap X_n$. By the regularity of μ_{n+1}, there is a closed and hence compact $K_n \subset S_n$ and a subset $V_n \subset X_{n+1}$ which is open in the subspace topology, such that $\mu_{n+1}(V_n \setminus K_n) \leq \frac{\epsilon}{2^n}$. The set $U_n := V_n \cap X_{n+1}^\circ$ is open in X, $K_n \subset S_n \subset U_n$ and $\mu(U_n \setminus K_n) \leq \frac{\epsilon}{2^n}$.

For inner regularity, there are two cases to distinguish: if $\mu(S) < \infty$, then $\mu(S) = \lim_{n\to\infty} \mu(S_n)$, and $\mu(K_n) \to \mu(S)$. If $\mu(S) = \infty$, then $\mu(S_n)$ and hence $\mu(K_n)$ becomes arbitrarily large.

Outer regularity is only an issue if $\mu(S) < \infty$. But we have $S \subset \bigcup_n U_n$, and the latter is open, and the same argument as in the proof of item (4) above proves that. $\square$

Appendix D. Holomorphic functions

We develop the basic theory of holomorphic functions on open subsets of the complex plane. We present the absolute minimum which is necessary for the development of spectral theory. As our exposition is geared towards the applications in functional analysis, we consider holomorphic functions with values in complex Banach spaces throughout (in the usual treatments of the theory, only functions with values in $\mathbb{C}$ are studied). This does not make any of the proofs much harder.

D.1. **Remarks about differential calculus.** We begin by recalling the definition of differentiability in several variables. We start with functions from (open subsets of) Banach spaces to Banach spaces. We consider real Banach spaces V, W whose norms we denote by the symbol $|v|$, to avoid heavy notation. On the vector space $\mathcal{L}(V; W)$ of continuous linear maps, we consider the operator norm which we denote as usual by $\|F\|$. We also agree that we write Fv for the value of $F \in \mathcal{L}(V, W)$ at $v \in V$.

Definition D.1. *Let V, W be Banach spaces, let $U \subset V$ be open and let $f : U \to W$ be a function. For $x \in U$, we let*

$$-x + U = \{-x + y | y \in U\} = \{h \in V | x + h \in U\}$$

and note that $0 \in -x + U$. We say that f is differentiable *at $x \in U$ if there is a function*

$$-x + U \to \mathcal{L}(V; W),\ h \mapsto F(x, h),$$

which is continuous at x and such that

$$f(x + h) = f(x) + F(x, h)h$$

for all $h \in -x + U$.

The differential *of f at x is the linear map*

$$Df(x) := F(x, 0) \in \mathcal{L}(V; W).$$

For example, linear maps are differentiable, and $DF(x) = F$. Linear combinations of differentiable maps are differentiable, and the formulas

$$D(f + g)(x) = Df(x) + Dg(x),\ D(af)(x) = aDf(x)$$

hold.

Example D.2. *Let $U \subset \mathbb{R}$ be open. A function $f : U \to V$ is differentiable at $x \in U$ if and only if the limit*

$$f'(x) := \lim_{h\to 0} \frac{1}{h}(f(x + h) - f(x)) \in V$$

exists, and we have

$$f'(x) = Df(x)1 \in V.$$

Reason: if the limit $f'(x)$ exists, we define $F(x, h) \in \mathcal{L}(\mathbb{R}; V) \cong V$ by

$$F(x, h) = \begin{cases} \frac{1}{h}(f(x + h) - f(x)) & h \neq 0 \\ f'(x) & h = 0. \end{cases}$$

Then $F(x, h)$ is continuous at $h = 0$, and $f(x + h) = f(x) + F(x, h)h$. Vice versa, if f is differentiable, then

$$\frac{1}{h}(f(x+h) - f(x)) - Df(x)1 = \frac{1}{h}(F(x, h)h) - F(x, 0)1 = (F(x, h) - F(x, 0))1 \to 0,$$

so that the limit $f'(x)$ exists and is equal to $Df(x)1$.

Example D.3. *If A is a real (or complex) unital Banach algebra and $A^\times \subset A$ is its (open) set of invertible elements, the function*

$$\iota : A^\times \to A,\ \iota(x) = x^{-1}$$

is differentiable. To see this, write

$$(x+h)^{-1} - x^{-1} = -(x+h)^{-1} h x^{-1}.$$

We define $F(x,h) \in \mathcal{L}(A,A)$ by

$$F(x,h)v := -(x+h)^{-1} h x^{-1}.$$

The function $h \mapsto F(x,h)$ is continuous. We conclude that ι is differentiable and

$$D\iota(x)v = -x^{-1} v x^{-1}.$$

Important is the *chain rule*.

Theorem D.4. *Let V_0, V_1, V_2 be Banach spaces, $U_i \subset V_i$ be open and let $f : U_0 \to U_1 \subset V_1$ be differentiable at $x \in U_0$, and $g : U_1 \to V_2$ be differentiable at $f(x)$. Then $g \circ f$ is differentiable at x, and*

$$D(g \circ f)(x) = Dg(f(x))Df(x).$$

Proof. Let $F(x, \_)$ and $G(f(x), \_)$ be functions certifying the differentiability of f and g at x and $f(x)$. Then

$$g(f(x+h)) = g(f(x) + F(x,h)h) = g(f(x)) + G(f(x), F(x,h)h)F(x,h)h.$$

Define $H(x,h) \in \mathcal{L}(V_0; V_1)$ by

$$H(x,h)v := G(f(x), F(x,h)h)F(x,h)v.$$

This is continuous at $h = 0$ since G and F are continuous. Therefore $g \circ f$ is differentiable at x, and

$$D(g \circ f)(x) = H(x,0) = G(f(x),0)F(x,0) = Dg(f(x))Df(x).$$

□

Let $U \subset V$ be open and let $f : U \to W$ be differentiable at x. Let $v \in V$. The curve $c : \mathbb{R} \to V$, $c(t) = x + tv$ is differentiable. By the chain rule, $f \circ c$ is differentiable at 0, and

$$D(f \circ c)(0) = Df(x)Dc(0).$$

But $Dc(0)1 = v$, and $D(f \circ c)(0)1 = \lim_{h \to 0} \frac{1}{h}(f(x+hv) - f(x))$. Therefore, we see that

$$Df(x)v = \lim_{h \to 0} \frac{1}{h}(f(x+hv) - f(x)).$$

D.2. **Holomorphy.**

Definition D.5. *Let $U \subset \mathbb{C}$ be an open subset, let V be a complex Banach space, and let $f : U \to V$ be a function. We say that f is* complex differentiable *at $z \in U$ if the limit*

$$f'(z) := \lim_{h \to 0} \frac{1}{h}(f(z+h) - f(z)) \in V$$

exists. We say that f is holomorphic *if f is complex differentiable at each point $z \in U$, and if the function $U \to V$, $z \mapsto f'(z)$ is continuous*[9]*.*

To relate this notion to the notion of differentiablity from D.1, we do some linear algebra first.

Definition D.6. *Let V, W be two complex Banach spaces. We let $\mathcal{L}_{\mathbb{C}}(V;W)$ be the space of bounded $\mathbb{C}$-linear maps $V \to W$, with the usual operator norm. This is a (real) subspace of the real Banach space $\mathcal{L}_{\mathbb{R}}(V,W)$ of all $\mathbb{R}$-linear bounded linear maps. Moreover, let $\mathcal{L}^-_{\mathbb{C}}(V;W) subset \mathcal{L}_{\mathbb{R}}(V;W)$ be the space of all $\mathbb{C}$-antilinear maps $F : V \to W$, which are the bounded $\mathbb{R}$-linear maps such that $F(zv) = \overline{z}F(v)$, for all $z \in \mathbb{C}$ and $v \in V$.*

Lemma D.7.

$$\mathcal{L}_{\mathbb{R}}(V;W) \cong \mathcal{L}_{\mathbb{C}}(V;W) \oplus \mathcal{L}^-_{\mathbb{C}}(V;W).$$

Proof. We have $\mathcal{L}_{\mathbb{C}}(V;W) \cap \mathcal{L}^-_{\mathbb{C}}(V;W) = 0$, because a linear map F which is both, $\mathbb{C}$-linear and $\mathbb{C}$-antilinear satisfies

$$iF(v) = F(iv) = -iF(v)$$

for all $v \in V$ and hence $F = 0$.

If $F : V \to W$ is $\mathbb{R}$-linear, then

$$F_l(v) := \frac{1}{2i}(iF(v) + F(iv))$$

is $\mathbb{C}$-linear because

$$F_l(iv) - iF_l(v) = \frac{1}{2i}(iF(iv) + F(-v) - i^2F(v) - iF(iv)) = 0$$

and

$$F_a(v) := F(v) - F_l(v) = \frac{1}{2i}(iF(v) - F(iv))$$

is $\mathbb{C}$-antilinear. □

Proposition D.8. *Let W be a complex Banach space, $z \in U \subset \mathbb{C}$ open and $f : U \to W$. The following are equivalent:*

(1) *f is complex differentiable at z.*
(2) *f is differentiable at z when viewed as a map between real Banach space, and the differential $Df(z) \in \mathcal{L}_{\mathbb{R}}(V;W)$ is $\mathbb{C}$-linear, i.e. an element of $\mathcal{L}_{\mathbb{C}}(V;W) \subset \mathcal{L}_{\mathbb{R}}(V;W)$.*

In that case, we have

$$Df(z)v = f'(z)v \in W.$$

for all $v \in \mathbb{C}$.

[9]One can drop the hypothesis that f' is continuous, it *follows* from the existence of $f'(z)$ for ech z. However, the stronger hypothesis simplifies the proofs in my opinion, and usually it does not take much work to verify this assumptions in practice.

Proof. $1 \Rightarrow 2$. We define, for $h \in -z+U$, an element $A(h) \in W$ by

$$A(h) := \begin{cases} \frac{1}{h}(f(z+h)-f(z)) & h \neq 0 \\ f'(z) & h = 0. \end{cases}$$

Then $A : -z+U \to W$ is continuous function, and

$$f(z+h) = f(z) + A(h)h$$

for $h \neq 0$ and by continuity also if $h = 0$. The map $\mathbb{C} \to W$, $v \mapsto A(h)v$ is complex linear. It follows that f is differentiable at z, $Df(z)$ is the $\mathbb{C}$-linear map $v \mapsto f'(z)v$.

$2 \Rightarrow 1$: let $A : -z+U \to \mathcal{L}_{\mathbb{R}}(\mathbb{C}; W)$ be a map certifying that f is differentiable at z. In other words,

$$f(z+h) = f(z) + A(h)h$$

and A is continuous at 0. The hypothesis says that $A(0) = Df(z)$ is $\mathbb{C}$-linear. We can write $A(h) = B(h) + C(h)$, where $B(h)$ is $\mathbb{C}$-linear and $C(h)$ is $\mathbb{C}$-antilinear, using Lemma D.7. Both, B and C are continuous at 0, and $C(0) = 0$. Therefore

$$\frac{1}{h}(f(z+h)-f(z)) = \frac{1}{h}(B(h)h + C(h)h).$$

Since $B(h)$ is $\mathbb{C}$-linear, we have

$$\frac{1}{h}B(h)h = B(h)1$$

and

$$|\frac{1}{h}C(h)h| \leq \|C(h)\|.$$

Altogether,

$$\lim_{h\to 0} \frac{1}{h}(f(z+h)-f(z)) - B(h)1 = 0$$

as $C(0) = 0$. It follows that $f'(z)$ exists and is equal to $B(0)1 = A(0)1 = Df(z)1$. $\square$

Le us give some typical examples.

Example D.9. *A rational functions $f(z) = \frac{p(z)}{q(z)}$, where $p, q \in \mathbb{C}[z]$ are two polynomials, is a holomorphic function $U \to \mathbb{C}$, where $U = \{z \in \mathbb{C} | q(z) \neq 0\}$.*

The functions $f(z) := \overline{z}$, $f(z) = |z|^2$ are not *holomorphic.*

Example D.10. *Let A be a unital complex Banach algebra. Then the function*

$$\iota : A^{\times} \to A, \; a \mapsto a^{-1}$$

whose real differentiability was shown in D.3, is in fact holomorphic.

D.3. **Power series.** Let $a_k \in V$ and consider the series

$$\sum_{k=0}^{\infty} a_k z^k. \tag{D.11}$$

Lemma D.12. (1) *If $|a_k r^k| \leq C$ for all k, and $|z| < r$, the series* (D.11) *is absolutely convergent.*

(2) *There is a unique $R \in [0, \infty]$, such that* (D.11) *converges absolutely for all $|z| < R$ and diverges for all $|z| > R$. R is called the* radius of convergence.

(3) *R is given by the formula*

$$R = \frac{1}{\limsup_n |a_n|^{1/n}}. \tag{D.13}$$

Proof. 1 follows from the estimate $|a_k z^k| \leq C(\frac{|z|}{r})^k$.

2: We let

$$R := \sup\{r \geq 0 | \exists C \forall k : |a_k r^k| \leq C\} \in [0, \infty].$$

If $|z| > R$, then $|a_k z^k|$ is unbounded and hence (D.11) diverges. If $|z| < R$, there is $|z| < r < R$ and C such that $|a_k r^k| \leq C$ for all k. Then $\sum_{k=0}^{\infty} a_k z^k$ is absolutely convergent.

3: If

$$|z| > \frac{1}{\limsup_n |a_n|^{1/n}} = \liminf_n \frac{1}{|a_n|^{1/n}},$$

there are infinitely many n such that

$$|z| \geq \frac{1}{|a_n|^{1/n}}$$

and hence

$$|a_n z^n| \geq 1.$$

Therefore (D.11) diverges for such z.

If

$$|z| < \liminf_n \frac{1}{|a_n|^{1/n}},$$

pick $s > |z|$ such that

$$s < \liminf_n \frac{1}{|a_n|^{1/n}},$$

which means that

$$s < \frac{1}{|a_n|^{1/n}}$$

for almost all n, and hence

$$s^n |a_n| < 1$$

for almost all n. Therefore

$$|a_n z^n| = s^n |a_n| (\frac{z}{s})^n$$

for almost all n, and (D.11) converges for such z. □

The formula (D.13) for the radius is often not very convenient, and can sometimes be replaced by a simpler formula.

Lemma D.14. *Let $(a_n)_n$ be a sequence in $\mathbb{C}$, such that $a_n \neq 0$ for all n. Assume that*

$$\lim_n \frac{|a_{n+1}|}{|a_n|} \in [0, \infty)$$

exists. Then

$$\lim_n \frac{|a_{n+1}|}{|a_n|} = \limsup_n |a_n|^{1/n}.$$

Proof. Write $b_n := |a_n|$,

$$P := \limsup_n |a_n|^{1/n},$$

and

$$Q := \lim_n \frac{|a_{n+1}|}{|a_n|}.$$

Let $\epsilon > 0$. Let m such that $|\frac{b_{n+1}}{b_n} - Q| \leq \epsilon$ for all $n \geq m$. For such n, we have

$$b_n = b_m \frac{b_{m+1}}{b_m} \cdots \frac{b_n}{b_{n-1}}$$

and hence

$$(b_n)^{1/n} \leq (b_m)^{1/n}(Q+\epsilon)^{\frac{n-m}{n}}$$

as well as

$$(b_n)^{1/n} \geq (b_m)^{1/n}(Q-\epsilon)^{\frac{n-m}{n}}$$

For $q > 0$ and $C > 0$, we have

$$\lim_n C^{1/n} q^{\frac{n-m}{n}} = \lim_n C^{1/n} q(\frac{1}{q^m})^{\frac{1}{n}} = q.$$

Therefore, if $Q = 0$, we have

$$\limsup_n b_n^{1/n} \leq \epsilon$$

for each $\epsilon > 0$ and hence $P = 0$. For $Q > 0$, we get that

$$Q - \epsilon \leq P \leq Q + \epsilon$$

for all $\epsilon > 0$ and hence $P = Q$ as well. □

Now let $\sum_{n=0}^{\infty} a_n z^n$ have have convergence radius $R > 0$ and consider the function

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

from $B_R(0) \subset \mathbb{C}$ to V. Our goal is to prove that f is holomorphic, and the derivative is the limit of the formally differentiated series

$$\sum_{n=1}^{\infty} a_n n z^{n-1}. \tag{D.15}$$

Lemma D.16. *The series* (D.15) *has the same radius of convergence as the original series* (D.11). *The function*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

is holomorphic in the open disc of convergence, and the formula

$$f'(z) = \sum_{n=1}^{\infty} a_n n z^{n-1}$$

holds.

Proof. 1: if (D.15) converges for $z \neq 0$ and $n \geq 1$, then

$$|a_n z^n| \leq |z||n a_n z^{n-1}|,$$

so that (D.11) is also convergent for z. Vice versa, let R be the radius of convergence for (D.11) and $|z| < R$. Pick $q \in (0,1)$ such that $|z| < qR < R$. There is $q < s < 1$ and C such that

$$|a_n s^n R^n| \leq C$$

for all n. Then

$$|(n+1)a_{n+1} q^n R^n| \leq (n+1)(\frac{q}{s})^n C \to 0$$

is in particular bounded, so that (D.15) converges for z.

2: We compute

$$\frac{1}{h}((z+h)^n - z^n) - nz^{n-1}h = \frac{1}{h}\sum_{k=0}^{n-2}\binom{n}{k} z^k h^{n-k} = h\sum_{k=0}^{n-2}\binom{n}{k} z^k h^{n-k-2} =$$

$$= h\sum_{k=0}^{n-2}\binom{n-2}{k} z^k h^{n-k-2}\frac{\binom{n}{k}}{\binom{n-2}{k}}.$$

But

$$\frac{\binom{n}{k}}{\binom{n-2}{k}} = \frac{n!k!(n-k-2)!}{k!(n-k)!(n-2)!} = \frac{n(n-1)}{(n-k)(n-k-1)} \leq \frac{1}{2}n(n-1)$$

if $0 \leq k \leq n$. Therefore

$$|\frac{1}{h}((z+h)^n - z^n) - nz^{n-1}h| \leq \frac{n(n-1)}{2}|h|\sum_{k=0}^{n-2}\binom{n-2}{k}|z|^k|h|^{n-k-2} = \frac{n(n-1)}{2}|h|(|z|+|h|)^{n-2}.$$

Let

$$g(z) = \sum_{n=1}^{\infty} a_n n z^{n-1}.$$

Then

$$|\frac{1}{h}(f(z+h) - f(z)) - g(z)| = |\sum_{n=1}^{\infty} a_n(\frac{(z+h)^n - z^n}{h} - nz^{n-1})| \leq$$

$$\leq |h|\sum_{n=1}^{\infty}|a_n|\frac{n(n-1)}{2}(|z|+|h|)^{n-2}.$$

If $|z| < R$ and $|h| \leq \frac{1}{2}(R - |z|)$, we have $|z| + |h| \leq \frac{1}{2}(R + |z|)$, and so $\sum_{n=1}^{\infty}|a_n|\frac{n(n-1)}{2}(|z| + |h|)^{n-2} < \infty$. It follows that

$$\lim_{h\to 0}\frac{1}{h}(f(z+h) - f(z)) - g(z) = 0 \qquad \square.$$

The most prominent example of a holomorphic function defined by a power series is the exponential function

$$e^z := \exp(z) := \sum_{n=0}\frac{1}{n!}z^n \in \mathbb{C}$$

which converges absolutely for each $z \in \mathbb{C}$.

D.4. **The local Cauchy theorem.**

D.5. **Curve integrals.** The fundamental theorem of calculus implies that each continuous function $f : I \to \mathbb{R}$ defined on an interval in $\mathbb{R}$ has a primitive function F, i.e. a differentiable function with $F' = f$. For continuous functions on open subsets of $\mathbb{C}$, this is not true, and the failure is quite interesting. We will from now on only consider holomorphic functions with values in $\mathbb{C}$, but the extension to Banach space valued functions is not very difficult.

As a first motivating step, let $U \subset \mathbb{C}$ be open and let $f : U \to \mathbb{C}$ be continuous. *Suppose* that f has a primitive function $F : U \to \mathbb{C}$, which by definition means a holomorphic function F with $F' = f$. Let $\gamma : [0,1] \to U$ be a smooth map. The function $[0,1] \to \mathbb{C}$, $t \mapsto F(\gamma(t))$ is differentiable, and by the fundamental theorem of calculus, we get

$$F(\gamma(1)) - F(\gamma(0)) = \int_0^1 \frac{d}{dt} F(\gamma(t))dt = \int_0^1 f(\gamma(t))\gamma'(t)dt. \tag{D.17}$$

In the second equality, we used the chain rule: if f is just differentiable, we have for $v \in \mathbb{C}$ that

$$D(F \circ \gamma)(t)v = DF(\gamma(t))D\gamma(t)v,$$

but $D\gamma(t)v = \dot{\gamma}(t)v$ and $DF(\gamma(t))D\gamma(t)v = F'(\gamma(t))\dot{\gamma}(t)v$. We take equation (D.17) as motivation for the following definition.

Definition D.18. *Let $U \subset \mathbb{C}$ be open, let $f : U \to \mathbb{C}$ be a continuous function and let $\gamma : [0,1] \to U$ be a smooth path. We define*

$$\int_\gamma f(z)dz := \int_0^1 f(\gamma(t))\gamma'(t)dt.$$

It follows that for a holomorphic function f and any path $\gamma : [0,1] \to U$, we have

$$f(\gamma(1)) - f(\gamma(0)) = \int_\gamma f'(z)dz.$$

Theorem D.19. *Let $U \subset \mathbb{C}$ be open, $f : U \to \mathbb{C}$ holomorphic, and let $\Gamma : (-\epsilon, \epsilon) \times [0,1] \to U$ be a smooth map. Define smooth curves*

$$\gamma_s(t) := \Gamma(s,t).$$

Then the function $F : (-\epsilon, \epsilon) \to \mathbb{C}$ defined by

$$F(s) := \int_{\gamma_s} f(z)dz$$

is differentiable and satisfies

$$F'(s) = f(\gamma_s(1))\frac{\partial}{\partial s}\gamma_s(1) - f(\gamma_s(0))\frac{\partial}{\partial s}\gamma_s(0).$$

Proof. Compute

$$F(s) = \int_{\gamma_s} f(z)dz = \int_0^1 f(\Gamma(s,t))\frac{\partial}{\partial t}\Gamma(s,t)dt.$$

This is differentiable by the theorem on differentation under the integral sign, and the derivative is

$$F'(s) = \int_0^1 \frac{\partial}{\partial s}\Big(f(\Gamma(s,t))\frac{\partial}{\partial t}\Gamma(s,t)\Big)dt =$$

$$= \int_0^1 (\frac{\partial}{\partial s} f(\Gamma(s,t))) \frac{\partial}{\partial t}\Gamma(s,t)dt + \int_0^1 f(\Gamma(s,t))(\frac{\partial}{\partial s}\frac{\partial}{\partial t}\Gamma(s,t))dt. \tag{D.20}$$

The first integral is equal to

$$\int_0^1 f'(\Gamma(s,t))\frac{\partial}{\partial s}\Gamma(s,t)\frac{\partial}{\partial t}\Gamma(s,t)dt =$$

$$= \int_0^1 \frac{\partial}{\partial t} f'(\Gamma(s,t)))\frac{\partial}{\partial s}\Gamma(s,t)dt = \text{ (partial integration)}$$

$$= -\int_0^1 f'(\Gamma(s,t)))\frac{\partial}{\partial t}\frac{\partial}{\partial s}\Gamma(s,t)dt + [f(\Gamma(s,t)\frac{\partial}{\partial s}\Gamma(s,t)]_{t=0}^1.$$

By Schwarz' theorem on the symmetry of higher partial derivatives, the first summand cancels against the second summand in (D.20), and the theorem is proven. □

To make use out of this computation, we introduce a couple of notions.

Definition D.21. *Let $U \subset \mathbb{C}$ be open. Let $\gamma_0, \gamma_1 : [0,1] \to U$ be smooth.*

(1) *A smooth* homotopy $\Gamma : \gamma_0 \rightsquigarrow \gamma_1$ *is a smooth map*

$$\Gamma : [0,1] \times [0,1] \to U$$

such that $\Gamma(i,t) = \gamma_i(t)$ for all t and $i = 0, 1$.

(2) *If $\gamma_0(0) = \gamma_1(0)$, $\gamma_0(1) = \gamma_1(1)$, a homotopy Γ will be* relative *endpoints, provided that*

$$\Gamma(s,i) = \gamma_0(i)$$

for $i = 0, 1$ and all s.

(3) *If γ_0 and γ_1 are* closed *curves, i.e. $\gamma_i(0) = \gamma_i(1)$, Γ is a* homotopy of closed curves*, provided that*

$$\Gamma(s,0) = \Gamma(s,1)$$

for all s.

It is easy to prove that these various notions of homotopy are equivalence relations on the sets of curves.

D.6. **The main theorem of the local theory.**

Definition D.22. *A subset $U \subset \mathbb{C}$ is* star-shaped *with respect to $z_0 \in U$ if for each $z \in U$ and each $t \in [0,1]$, $z_0 + t(z - z_0) \in U$.*

For example $\mathbb{C}$, as well as balls $B_r(z)$ are star-shaped.

Corollary D.23 (Cauchy integral theorem)**.** *Let $U \subset \mathbb{C}$ be open and star-shaped with respect to z_0 and let $f : U \to \mathbb{C}$ be holomorphic. Then for each closed curve γ in U, we have*

$$\int_\gamma f(z)dz = 0.$$

Proof. Let $\Gamma : [0,1] \times [0,1] \to U$, $\Gamma(s,t) := z_0 + s(\gamma(t) - z_0)$. Then $\gamma_s(t) := \Gamma(s,t)$ is a smooth closed curve for each s. By Theorem D.19, the function

$$G : [0,1] \to \mathbb{C},\ G(s) := \int_{\gamma_s} f(z)dz$$

is constant, but $G(1) = \int_\gamma f(z)dz$ and

$$G(0) = \int_{\gamma_0} f(z)dz = \int_0^1 f(z_0)\dot{\gamma_s}(t)dt = 0.$$

$\square$

For $w \in \mathbb{C}$ and $r > 0$, we let $C_r(w)$ be the curve

$$t \mapsto w + re^{2\pi i t}.$$

Then

$$\int_{C_r(w)} \frac{1}{z-w} dz = \int_0^1 \frac{1}{re^{2\pi i t}} 2\pi i r e^{2\pi i t} dt = 2\pi i \int_0^1 1 dt = 2\pi i.$$

Theorem D.24 (Cauchy integral formula)**.** *Let $U \subset \mathbb{C}$ be open and let $f : U \to \mathbb{C}$ be holomorphic, $z \in U$ and $\overline{B}_r(z_0) \subset U$. Then for all $z \in B_r(z_0)$, we have*

$$f(z) = \frac{1}{2\pi i} \int_{C_r(z)} \frac{f(w)}{w-z} dw.$$

Proof. The function $g(w) := \frac{f(w)}{w-z}$ is holomorphic on $U \setminus z$, and the closed curves $C_r(z_0)$ and $C_\delta(z)$ are homotopic through closed curves in $U \setminus z$, for small enough $\delta > 0$. From Theorem D.19, we see that

$$\frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{w-z} dw = \frac{1}{2\pi i} \int_{C_\delta(z)} \frac{f(w)}{w-z} dw =$$

$$= \frac{1}{2\pi i} \int_0^1 \frac{f(z + \delta e^{2\pi i t})}{\delta e^{2\pi i t}} 2\pi i \delta e^{2\pi i t} dt =$$

$$= \int_0^1 f(z + \delta e^{2\pi i t}) dt.$$

But

$$\lim_{\delta \to 0} \int_0^1 f(z + \delta e^{2\pi i t}) dt = f(z).$$

$\square$

Now assume that $\overline{B}_r(z_0) \subset U$. Then whenever $z \in B_r(z_0)$, we have by Theorem D.24

$$f(z) = \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{w-z} dw = \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{(w-z_0)-(z-z_0)} dw =$$

$$= \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{w-z_0} \frac{1}{1 - \frac{z-z_0}{w-z_0}} dw =$$

$$= \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{w-z_0} \sum_{n=0}^\infty \frac{(z-z_0)^n}{(w-z_0)^n} dw =$$

$$= \sum_{n=0}^\infty \left(\frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{(w-z_0)^{n+1}} dw\right)(z-z_0)^n.$$

The exchange of integration and summation is justified because the geometric series converges uniformly on $\operatorname{im}(C_r(z_0))$. We arrive at the following fundamental result.

Theorem D.25. *Let $U \subset \mathbb{C}$ be open, $z_0 \in U$ and $\overline{B}_r(z_0) \subset U$. Let $f : U \to \mathbb{C}$ be holomorphic. Define*

$$a_n := \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{(w - z_0)^{n+1}} dw \in \mathbb{C}.$$

Then

(1) *The power series*

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

converges absolutely and locally uniformly on $B_r(z_0)$ to f.

(2) *f is smooth, and all derivatives of f are holomorphic.*

(3) *The estimate*

$$|a_n| \leq \|f|_{\partial B_r(z_0)}\| \frac{2\pi}{r^n}$$

holds.

Proof. We have shown (1). Part (2) follows, because sums of power series are holomorphic in the interior of the disc of convergence. The estimate follows from

$$a_n = \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(w)}{(w - z_0)^{n+1}} dw = 2\pi i \int_0^1 \frac{f(z_0 + re^{2\pi i t})}{r^n e^{2\pi i n t}} dt$$

and the standard estimate for integrals over compact intervals. □

Corollary D.26 (Liouville's Theorem)**.** *A bounded holomorphic function $f : \mathbb{C} \to \mathbb{C}$ is constant.*

Proof. If $|f(z)| \leq C$ for all $z \in \mathbb{C}$, we have

$$|a_n| \leq \frac{2\pi C}{r^n}$$

for all $r > 0$, and hence $a_n = 0$ whenever $n > 0$. It follows that f is constant. □

References

[1] Yoav Benyamini. Applications of the universal surjectivity of the Cantor set. *Amer. Math. Monthly*, 105(9):832–839, 1998.

[2] John B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[3] Manfred Einsiedler and Thomas Ward. *Functional analysis, spectral theory, and applications*, volume 276 of *Graduate Texts in Mathematics*. Springer, Cham, 2017.

[4] D. J. H. Garling. A "short" proof of the Riesz representation theorem. *Proc. Cambridge Philos. Soc.*, 73:459–460, 1973.

[5] Sandy Grabiner. The Tietze extension theorem and the open mapping theorem. *Amer. Math. Monthly*, 93(3):190–191, 1986.

[6] Friedrich Hirzebruch and Winfried Scharlau. *Einführung in die Funktionalanalysis*, volume 296 of *B.I.-Hochschultaschenbücher [B.I. University Paperbacks]*. Bibliographisches Institut, Mannheim, 1991. Reprint of the 1971 original.

[7] Jonathan Lewin. A simple proof of Zorn's lemma. *Amer. Math. Monthly*, 98(4):353–354, 1991.

[8] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[9] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[10] Terence Tao. *An epsilon of room, I: real analysis*, volume 117 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Pages from year three of a mathematical blog.

[11] Dirk Werner. *Funktionalanalysis*. Springer-Verlag, Berlin, extended edition, 2000.
[12] Kung-wei Yang. A note on reflexive Banach spaces. *Proc. Amer. Math. Soc.*, 18:859–861, 1967.