Lecture Notes

Advanced Functional Analysis

Günther Hörmann
Fakultät für Mathematik
Universität Wien
guenther.hoermann@univie.ac.at

Sommersemester 2023

(Corrected and updated version as of September 13, 2024)

Preface

These lecture notes draw a substantial part of their material and style as well as its logical build-up from Chapters VII and VIII of Dirk Werner's excellent textbook [Wer18] (in German). Regarding the prerequisites according to an introductory course on functional analysis at our faculty we may refer to Chapters I–VI in [Wer18] or [Hoe21], both in German, or for English texts to the corresponding chapters in [Con10, Con16, Tes14]. (The latter sources contain also many further aspects and material.)

In course of the semester we might occasionally provide hints to supplementary concepts, examples, or further applications not covered in these notes. In fact, these notes do certainly not replace a book on the subject and are particularly sparse with intermediate and explanatory or motivating texts in between mathematical statements. However, Sections 1-7 as presented in the notes define the compulsory material for the exam. (Thus excluding Section 0 and the appendix.)

Many thanks go to Michael Kunzinger for several suggestions and corrections regarding a previous version of these notes and to Nobuya Kakehashi for several subtle mathematical comments leading to further improvements later on.

Günther Hörmann

Contents

Part I.

Bounded and unbounded self-adjoint operators

0. Before we begin ...

Recalling a few functional analytic basics

Notation and conventions: $\mathbb{N} = \{1, 2, \ldots\}$, H denotes a *complex* Hilbert space with scalar product $\langle ., . \rangle$ (conjugate-linear in its second argument), $L(H)$ is the space of bounded linear operators on H; sequences $(f_n)_{n \in \mathbb{N}}$ are often written simply in the form (f_n).

If $S \in L(H)$, then its adjoint $S^* \in L(H)$ can be characterized by the condition $\langle Sx, y \rangle = \langle x, S^* y \rangle$ for all $x, y \in H$; we have $S^{**} := (S^*)^* = S$. The kernel $\ker(S) := \{x \in H \mid Sx = 0\}$ is always a closed subspace of H, while the range (or image) $\operatorname{ran}(S) := \{Sx \mid x \in H\}$ is a subspace of H that is not necessarily closed. A basic relation involving these is $\operatorname{ran}(S)^\perp = \ker(S^*)$, hence also $\ker(S) = \operatorname{ran}(S^*)^\perp$ and $\overline{\operatorname{ran}(S)} = \operatorname{ran}(S)^{\perp\perp} = \ker(S^*)^\perp$.

Recall that $T \in L(H)$ is *self-adjoint*, if $T = T^*$, i.e., $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in H$; and that T is *normal*, if $TT^* = T^*T$. Clearly, every self-adjoint operator is normal. We call $T \in L(H)$ *positive*, if $\langle Tx, x \rangle \geq 0$ holds for all $x \in H$. On *complex* Hilbert spaces, positive operators in this sense are automatically self-adjoint ([Wer18, Satz V.5.6]). Finally, an operator $P \in L(H)$ is an *orthogonal projection*—onto the closed subspace $\operatorname{ran}(P)$ and we have $H = \ker(P) \oplus \operatorname{ran}(P)$ as an orthogonal direct sum—, if and only if P is self-adjoint and $P^2 = P$. In the sequel we will simply speak of *projections* to mean orthogonal projections.

0.1. Proposition: If $T \in L(H)$ is self-adjoint, then

(i) $\langle Tx, x \rangle \in \mathbb{R}$ for every $x \in H$,

(ii) $\|T\| = \sup\limits_{\|x\| \leq 1} |\langle Tx, x \rangle|$.

Remark: Property (i) is, in fact, also sufficient for self-adjointness of a bounded operator on *complex* Hilbert spaces (cf. [Wer18, Satz V.5.6]).

Proof: (i): $\langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle}$.

(ii): For any $x \in H$ with $\|x\| \leq 1$, we have $|\langle Tx, x \rangle| \leq \|Tx\| \cdot \|x\| \leq \|T\| \cdot \|x\|^2 \leq \|T\|$, hence it remains to show that $\|T\| \leq \sup_{\|x\| \leq 1} |\langle Tx, x \rangle| =: M$.

By elementary maneuvering,

$$\begin{aligned}\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle &= 2\langle Tx, y\rangle + 2\langle Ty, x\rangle = 2\langle Tx, y\rangle + 2\overline{\langle x, Ty\rangle} \\ &= 2(\langle Tx, y\rangle + \overline{\langle Tx, y\rangle}) = 4\,\mathrm{Re}\langle Tx, y\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}4\,\mathrm{Re}\langle Tx, y\rangle \leq |\langle T(x+y), x+y\rangle| + |\langle T(x-y), x-y\rangle| &\leq M\|x+y\|^2 + M\|x-y\|^2 \\ &= M(2\|x\|^2 + 2\|y\|^2) \quad \text{(by the parallelogram law)}.\end{aligned}$$

If $\|x\|, \|y\| \leq 1$, then we obtain $4\,\mathrm{Re}\langle Tx, y\rangle \leq M \cdot 4$, hence $\mathrm{Re}\langle Tx, y\rangle \leq M$. Replacing x by λx we deduce

$$\mathrm{Re}(\lambda\langle Tx, y\rangle) \leq M \qquad \forall \lambda \in \mathbb{C}, |\lambda| = 1, \forall x, y \in H, \|x\|, \|y\| \leq 1,$$

thus, $|\langle Tx, y\rangle| \leq M$, if $\|x\|, \|y\| \leq 1$. We conclude that $\|Tx\| \leq M$, if $\|x\| \leq 1$, which in turn yields $\|T\| \leq M$. □

0.2. Basic closed range condition: Recall that any $T \in L(H)$ satisfying for some $c > 0$ the following lower bound condition

$$\forall x \in H: \quad \|Tx\| \geq c\|x\|$$

has a closed range $\mathrm{ran}(T) = \{Tx \mid x \in H\}$ in H and the inverse $T^{-1}\colon \mathrm{ran}(T) \to H$ is continuous with $\|T^{-1}\| \leq 1/c$. In fact, a convergent sequence (Tx_n) in $\mathrm{ran}(T)$ is a Cauchy sequence in H and the estimate $\|x_n - x_m\| \leq \|Tx_n - Tx_m\|/c$ shows that (x_n) is a Cauchy sequence in H, hence convergent to some $x \in H$, thus (Tx_n) converges to $Tx \in \mathrm{ran}(T)$, thus $\mathrm{ran}(T)$ is closed. Clearly, T is injective and with $y = Tx$ we have $\|T^{-1}y\| = \|x\| \leq \|y\|/c$.

0.3. Theorem (Lax-Milgram): Let $B\colon H \times H \to \mathbb{C}$ be a sesquilinear form.

(a) The following are equivalent:

(i) B is continuous,

(ii) B is separately continuous,

(iii) $\exists M \geq 0\ \forall x, y \in H$: $|B(x, y)| \leq M\,\|x\|\,\|y\|$.

(b) If B is continuous and M is as in (a), part (iii), then there exists a unique $T \in L(H)$ such that $B(x, y) = \langle Tx, y\rangle$ for all $x, y \in H$ and we have $\|T\| \leq M$.
If, in addition, B is bounded below in the sense that $|B(x, x)| \geq c\,\|x\|^2$ holds with some $c > 0$ for all $x \in H$, then T is invertible and $\|T^{-1}\| \leq 1/c$.

Proof: (a): (i) $\Rightarrow$ (ii) is clear.

(ii) $\Rightarrow$ (iii): For $y \in H$ define the continuous linear functional l_y on H by $l_y(x) := B(x, y)$. If $x \in H$, then $y \mapsto B(x, y)$ is conjugate-linear and continuous, hence there is some $C_x \geq 0$

such that $|l_y(x)| = |B(x,y)| \leq C_x \|y\|$. Thus, the family $A := \{l_y \in H' \mid y \in H, \|y\| \leq 1\}$ of linear functionals is pointwise bounded on H, since $\sup_{\|y\|\leq 1} |l_y(x)| \leq C_x$. By the Banach-Steinhaus theorem (or uniform boundedness principle), A is bounded in H', i.e., there is an $M \geq 0$ such that $\|l_y\| \leq M$ for all $y \in H$ with $\|y\| \leq 1$. In other words,

$$\forall x, y \in H, \|x\| \leq 1, \|y\| \leq 1 : \quad |B(x,y)| \leq M.$$

Since $B(x,y) = 0$ if $x = 0$ or $y = 0$, and $|B(\frac{x}{\|x\|}, \frac{y}{\|y\|})| \leq M$, if $x \neq 0$ and $y \neq 0$, by the above inequality, we obtain $|B(x,y)| \leq M\|x\|\,\|y\|$.

(iii) $\Rightarrow$ (i): Let $(x_0, y_0) \in H$ and $\varepsilon > 0$ be given. Choose $\delta \in]0,1]$ such that $\delta < \varepsilon/(M(1 + \|x_0\| + \|y_0\|))$, and put $U_\delta := \{(x,y) \in H \times H \mid \|x - x_0\| + \|y - y_0\| < \delta\}$. Then U_δ is a neighborhood of (x_0, y_0) in $H \times H$ (with respect to the product topology) and we have for $(x,y) \in U_\delta$:

$$\begin{aligned} |B(x,y) - B(x_0,y_0)| &= |(B(x,y) - B(x_0,y)) + (B(x_0,y) - B(x_0,y_0))| \\ &\leq |B(x - x_0, y)| + |B(x_0, y - y_0)| \leq M\,\|x - x_0\|\,\|y\| + M\,\|x_0\|\,\|y - y_0\| \\ &\leq M\,\delta\,(\|y_0\| + \delta) + M\,\|x_0\|\,\delta = \delta\,M\,(\|x_0\| + \|y_0\| + \delta) \leq \delta\,M\,(\|x_0\| + \|y_0\| + 1) < \varepsilon. \end{aligned}$$

(b): If $x \in H$ then $h_x(y) := B(x,y)$ defines a conjugate-linear continuous functional on H and the conjugate-linear variant of the Riesz-Fréchet theorem provides us with a unique vector $T(x) \in H$ such that $B(x,y) = h_x(y) = \langle T(x), y\rangle$ for every $y \in H$. We obtain a map $T\colon H \to H$, $x \mapsto T(x)$, which is easily seen to be linear, since $\langle T(\lambda_1 x_1 + \lambda_2 x_2), y\rangle = B(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 B(x_1, y) + \lambda_2 B(x_2, y) = \langle \lambda_1 T(x_1) + \lambda_2 T(x_2), y\rangle$ for every y implies $T(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 T(x_1) + \lambda_2 T(x_2)$.

We have $|\langle Tx, y\rangle| = |B(x,y)| \leq M\,\|x\|\,\|y\|$, hence $\sup_{\|y\|\leq 1} |\langle Tx, y\rangle| \leq M\|x\|$, i.e., $\|Tx\| \leq M\|x\|$, and therefore $\|T\| \leq M$.

Uniqueness of T: If $\langle Tx, y\rangle = B(x,y) = \langle Sx, y\rangle$ for all $x, y \in H$, then $Tx - Sx \perp H$ for all $x \in H$, hence $T = S$.

Finally, $c\,\|x\|^2 \leq |B(x,x)| = |\langle Tx, x\rangle| \leq \|Tx\|\,\|x\|$ implies $c\|x\| \leq \|Tx\|$, which shows (cf. 0.2) that $\mathrm{ran}(T)$ is closed and that T^{-1} is continuous as linear map $\mathrm{ran}(T) \to H$ with $\|T^{-1}\| \leq 1/c$. We claim that $\mathrm{ran}(T) = H$, for otherwise there is some $z \neq 0$ with $z \perp \mathrm{ran}(T)$, i.e., $0 = |\langle Tz, z\rangle| = |B(z,z)| \geq c\,\|z\|^2$, a contradiction. □

0.4. The spectrum of a bounded operator $\mathbf{T} \in \mathbf{L}(\mathbf{H})$: For proofs not given here and the more general Banach space context see [Wer18, Abschnitt VI.1].

The *resolvent set*

$$\rho(T) := \{\lambda \in \mathbb{C} \mid \exists (\lambda - T)^{-1} \text{ in } L(H)\}$$

is an open subset of $\mathbb{C}$ and the *resolvent map* $R\colon \rho(T) \to L(H)$, $R_\lambda := R(\lambda) := (\lambda - T)^{-1}$ is analytic, i.e., is locally given by a power series with coefficients from $L(H)$.

The *spectrum* $\sigma(T) := \mathbb{C} \setminus \rho(T)$ is a compact non-empty subset of $\mathbb{C}$. We have $\sigma(T^*) = \{\overline{\lambda} \mid \lambda \in \sigma(T)\}$, since $((\lambda - T)^{-1})^* = ((\lambda - T)^*)^{-1} = (\overline{\lambda} - T^*)^{-1}$.

If $|\lambda| > \|T\|$ we can apply the Neumann series to find $(\lambda - T)^{-1} = (\mathrm{I} - T/\lambda)^{-1}/\lambda$, thus $\sigma(T) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|T\|\}$.

We obtain a slightly sharper "encirclement" of the spectrum by the *spectral radius* $r(T) := \inf_{n\in\mathbb{N}} \|T^n\|^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n}$, which clearly satisfies $r(T) \leq \|T\|$ (since $\|T^n\|^{1/n} \leq (\|T\|^n)^{1/n} = \|T\|$), namely

$$r(T) = \max\{|\lambda| \mid \lambda \in \sigma(T)\}.$$

If T is normal, then $r(T) = \|T\|$. In particular, we can then always find a spectral value $\lambda \in \sigma(T)$ with $|\lambda| = \|T\|$.

A spectral value $\lambda \in \mathbb{C}$ is defined by the failure of $(\lambda - T)^{-1}$ to exist in $L(H)$. Note that, due to the open mapping principle, the bijectivity of $(\lambda - T)\colon H \to H$ already implies continuity of its inverse. Hence the "cases of failure" can be separated into the following three classes:

1. $\lambda - T$ fails to be injective (it has a nontrivial kernel and thus λ is an eigenvalue),
2. $\lambda - T$ is injective, not surjective, and $\operatorname{ran}(\lambda - T)$ is dense in H,
3. $\lambda - T$ is injective and $\operatorname{ran}(\lambda - T)$ is not dense in H (hence T is also not surjective).

Accordingly, we have the following decomposition of the spectrum (as a disjoint union)

$$\sigma(T) = \{\lambda \in \mathbb{C} \mid \nexists(\lambda - T)^{-1} \text{ in } L(H)\} = \underbrace{\{\lambda \in \mathbb{C} \mid \ker(\lambda - T) \neq \{0\}\}}_{\sigma_p(T)}$$
$$\cup \underbrace{\{\lambda \in \mathbb{C} \mid \ker(\lambda - T) = \{0\}, \operatorname{ran}(\lambda - T) \neq H, \operatorname{ran}(\lambda - T) \text{ dense}\}}_{\sigma_c(T)}$$
$$\cup \underbrace{\{\lambda \in \mathbb{C} \mid \ker(\lambda - T) = \{0\}, \operatorname{ran}(\lambda - T) \text{ not dense in } H\}}_{\sigma_r(T)}.$$

The *point spectrum* $\sigma_p(T)$ consists of the *eigenvalues* of T, $\sigma_c(T)$ is called the *continuous spectrum*, and $\sigma_r(T)$ is the *residual spectrum*[1]. However, we show that the latter does not play any role for normal (or self-adjoint) operators.

0.5. Lemma: If T is normal, then $\sigma_r(T) = \emptyset$.

Proof: The operator $\lambda - T$ is normal as well, hence $\ker(\lambda - T) = \ker(\lambda - T)^*$. (In fact, $\|(\lambda - T)x\| = \|(\lambda - T)^* x\|$ holds, since $0 = \langle((\lambda - T)^*(\lambda - T) - (\lambda - T)(\lambda - T)^*)x, x\rangle = \langle(\lambda - T)^*(\lambda - T)x, x\rangle - \langle(\lambda - T)(\lambda - T)^* x, x\rangle = \|(\lambda - T)x\|^2 - \|(\lambda - T)^* x\|^2$.) Thus, $\lambda \in \sigma_r(T)$ would imply $\{0\} \neq \operatorname{ran}(\lambda - T)^\perp = \ker(\lambda - T)^* = \ker(\lambda - T)$, i.e., $\lambda \in \sigma_p(T)$, a contradiction. □

0.6. Proposition (on approximate eigenvalues of normal operators): Let $T \in L(H)$ be normal and $\lambda \in \mathbb{C}$, then the following are equivalent:

[1] A simple example for the existence of a residual spectrum is $\lambda = 0$ for the right-shift $R(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$ on $l^2(\mathbb{N})$, since it has $(1, 0, 0, \ldots)$ orthogonal to its range.

(i) $\lambda \in \sigma(T)$,

(ii) there is a sequence (x_n) in H with $\|x_n\| = 1$ and $\lim_{n\to\infty} (\lambda x_n - Tx_n) = 0$.

Proof: (i) $\Rightarrow$ (ii): By the above lemma we have $\sigma(T) = \sigma_p(T) \cup \sigma_c(T)$. If λ is an eigenvalue with normalized eigenvector x, then we may simply put $x_n := x$ and obtain $\lambda x_n - Tx_n = 0$.

It remains to consider $\lambda \in \sigma_c(T)$. Then $\lambda - T$ is injective and $\mathrm{ran}(\lambda - T)$ is dense in H, but $\mathrm{ran}(\lambda - T) \neq H$. There can be no constant $c > 0$ such that $\|(\lambda - T)x\| \geq c\,\|x\|$ holds for all $x \in H$, since this would imply continuity of $(\lambda - T)^{-1}$ as a linear map $\mathrm{ran}(\lambda - T) \to H$ with a continuous extension to H due to the density of $\mathrm{ran}(\lambda - T)$ (and uniform continuity of continuous linear maps). Therefore, for every $n \in \mathbb{N}$ we can find $y_n \in H$ such that $\|(\lambda - T)y_n\| < \|y_n\|/n$. Putting $x_n := y_n/\|y_n\|$ we obtain $\|\lambda x_n - Tx_n\| < 1/n$.

(ii) $\Rightarrow$ (i): If $\lambda \notin \sigma(T)$, then $(\lambda - T)^{-1} \colon H \to H$ is bounded and hence there is some $M > 0$ such that $\|(\lambda - T)^{-1}z\| \leq M\,\|z\|$ holds for all $z \in H$. Equivalently, upon putting $z = (\lambda - T)x$, we have $\|x\| \leq M\,\|(\lambda - T)x\|$ for all $x \in H$, which contradicts (ii), since this means $\|\lambda x - Tx\| \geq 1/M > 0$ for every $x \in H$ with $\|x\| = 1$. □

And a bit of measure theory

Here we collect a few measure theoretic notions and results (available also, e.g, in Appendix A and Chapters I, II, VII of [Wer18]); more on the measure theoretic background can be found in the excellent textbooks [Bau90, Els11] (in German) or [Bau01, Coh80, Con16, Fol99, Rud86] (in English). Prerequisites from basic topology courses as in [Hoe20] (contained in English in the book [Wil70]) will be used throughout the course without special notice. We denote by $\mathcal{P}(\Omega)$ the power set of the set Ω.

(A) Measures

0.7. Definition: Let Ω be a set. A *sigma algebra* on Ω is a subset $\Sigma \subseteq \mathcal{P}(\Omega)$ satisfying

(i) $\emptyset \in \Sigma$,

(ii) $A \in \Sigma \quad \Rightarrow \quad \Omega \setminus A \in \Sigma$,

(iii) $A_j \in \Sigma \ (j \in \mathbb{N}) \quad \Rightarrow \quad \bigcup_{j\in\mathbb{N}} A_j \in \Sigma$.

0.8. Definition: Let Ω be a topological space. The *Borel sigma algebra* on Ω is the smallest sigma algebra $\mathcal{B}(\Omega)$ containing the topology, i.e., the system of open subsets of Ω. The elements in $\mathcal{B}(\Omega)$ are called *Borel sets*.

Clearly, $\mathcal{B}(\Omega)$ contains all closed subsets, and in case of a Hausdorff space also all compact subsets. In $\mathbb{R}$ every interval is a Borel set, in $\mathbb{C} \cong \mathbb{R}^2$ any product of two intervals belongs to $\mathcal{B}(\mathbb{C})$. For more on these issues see [Els11, Kapitel I, §4].

0.9. Definition: Let Σ be a sigma algebra on the set Ω. A measure is a map $\mu \colon \Sigma \to [0, \infty]$ satisfying

(i) $\mu(\emptyset) = 0$,

(ii) *σ-additivity*: If $A_1, A_2, \ldots$ is a sequence of pairwise disjoint sets in Σ, then

$$\mu(\bigcup_{j \in \mathbb{N}} A_j) = \sum_{j \in \mathbb{N}} \mu(A_j).$$

If $\mu(\Omega) < \infty$, μ is a *finite measure*, and, if $\mu(\Omega) = 1$, it is said to be a *probability measure*. The triple (Ω, Σ, μ), or often simply the pair (Ω, μ), is called a measure space (or probability space, if $\mu(\Omega) = 1$). If Ω is a topological space and $\Sigma \supseteq \mathcal{B}(\Omega)$, then a measure μ defined on Σ (or rather its restriction $\mu|_{\mathcal{B}(\Omega)}$) is called a *Borel measure* on Ω.

The two simplest (nontrivial) examples of measures on an arbitrary set Ω with sigma algebra $\mathcal{P}(\Omega)$ are the *Dirac measure* δ_p concentrated at $p \in \Omega$ with $\delta_p(A) = 1$, if $p \in A$, $\delta_p(A) = 0$ otherwise, and the *counting measure* μ with $\mu(A) = \infty$, if A is an infinite subset of Ω, and $\mu(A)$ equal to the number of elements of A, if A is finite.

0.10. Theorem: For every $d \in \mathbb{N}$ there is a unique measure (defined at least) on the Borel sigma algebra $\mathcal{B}(\mathbb{R}^d)$ which is translation invariant and assigns the value $(b_1 - a_1) \cdots (b_d - a_d)$ to the product of closed bounded intervals $[a_1, b_1] \times \cdots \times [a_d, b_d]$. This measure is called the *d-dimensional Lebesgue measure*.

(B) Construction of the integral

0.11. Definition: Let Σ be a sigma algebra on Ω. A function $f \colon \Omega \to \mathbb{R}$ (or $\Omega \to [0, \infty]$) is *measurable* (or Σ-measurable), if $f^{-1}([a, b[) \in \Sigma$ for any $a \leq b$. A complex function $f \colon \Omega \to \mathbb{C}$ is *measurable*, if $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable functions $\Omega \to \mathbb{R}$. In case of a topological space with $\Sigma = \mathcal{B}(\Omega)$, a measurable function is called *Borel measurable*.

Continuous functions on a topological space are easily seen to be Borel measurable.

For any $A \subseteq \Omega$ we have the *characteristic function* (or *indicator function*) of A, defined by $\chi_A(q) = 1$, if $q \in A$, and $\chi_A(q) = 0$, if $q \notin A$. Clearly, χ_A is measurable if and only if $A \in \Sigma$.

Integral of step functions: A *simple function* or *step function* (or an *elementary function*) on Ω is a function of the form

$$f = \sum_{j=1}^{m} c_j \, \chi_{A_j},$$

where $c_j \in \mathbb{C}$ and $A_j \in \Sigma$ ($j = 1, \ldots, m$), and $A_1, A_2, \ldots A_m$ are pairwise disjoint. The *integral* of a non-negative step function f with respect to a measure μ on Σ is defined by

$$\int f \, d\mu = \sum_{j=1}^{m} c_j \, \mu(A_j).$$

Integral of non-negative measurable functions: It can be shown that pointwise limits of sequences of measurable functions are measurable and that any measurable function $f\colon \Omega \to [0, \infty]$ is the pointwise increasing limit of a sequence (φ_n) of non-negative step functions $0 \leq \varphi_1(q) \leq \varphi_2(q) \leq \cdots$ (cf. [Els11, Kapitel III, Satz 4.13]). The real sequence of integrals $(\int \varphi_n \, d\mu)_{n \in \mathbb{N}}$ is thus increasing and one puts

$$\int f \, d\mu := \lim_{n \to \infty} \int \varphi_n \, d\mu \quad \in [0, \infty].$$

In case $\int f \, d\mu < \infty$ the function $f\colon \Omega \to [0, \infty]$ is called *integrable* (or μ-integrable).

Integral of real- or complex-valued measurable functions: Let $f\colon \Omega \to \mathbb{R}$ be measurable, then $f_+ := \max(f, 0)$ and $f_- := \max(-f, 0)$ are measurable functions $\Omega \to [0, \infty[$ and $f = f_+ - f_-$. If both f_+ and f_- are integrable, then f is called *integrable* and we put

$$\int f \, d\mu := \int f_+ \, d\mu - \int f_- \, d\mu.$$

A complex-valued measurable function $f\colon \Omega \to \mathbb{C}$ is called *integrable*, if $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable, and we then put $\int f \, d\mu := \int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu$.

A property Φ depending on the points in Ω is said to hold *μ-almost everywhere* (μ-a.e.), if there is a set $N \in \Sigma$ with $\mu(N) = 0$ (N is a *null set*) such that $\Phi(q)$ holds whenever $q \notin N$.

0.12. Theorem (on dominated convergence): Let f and $f_1, f_2, \ldots$ be measurable functions on Ω and suppose that f is the pointwise limit of f_n μ-almost everywhere. If there is an integrable function g on Ω such that $|f_n| \leq g$ holds for all $n \in \mathbb{N}$ and μ-almost everywhere, then f as well as every f_n is integrable and we have

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu.$$

(In fact, the slightly stronger statement $\lim\limits_{n \to \infty} \int |f_n - f| \, d\mu = 0$ is true.)

Brief review of the construction of $\mathbf{L^p}$-spaces: Let (Ω, Σ, μ) be a measure space. If $1 \leq p < \infty$, we define $L^p(\Omega, \mu)$ as vector space quotient of $\mathcal{L}^p := \{f : \Omega \to \mathbb{C} \text{ measurable} \mid \int |f|^p \, d\mu < \infty\}$ modulo $\mathcal{N} := \{f \text{ measurable} \mid f = 0 \ \mu\text{-a.e.}\}$ and equip it with the norm $\|\text{class of } f\|_p := (\int |f|^p \, d\mu)^{1/p}$. We obtain a Banach space, which in case of a compact subset $\Omega \subseteq \mathbb{R}^d$ and μ the (restriction of) d-dimensional Lebesgue measure is the completion of the space of continuous functions on Ω with respect to $\|.\|_p$. (More generally, the compactly supported continuous functions on a locally compact Hausdorff space Ω are dense in $L^p(\Omega, \mu)$, if μ is a *regular* Borel measure.)
In case $p = \infty$ we define $\mathcal{L}^\infty(\Omega)$ to be the set of all μ-measurable functions $f : \Omega \to \mathbb{C}$ that are bounded μ-a.e., i.e., there exists a set $N \in \Sigma$ with $\mu(N) = 0$ such that the restriction $f \mid_{\Omega \setminus N}$ is bounded. On the quotient vector space $L^\infty(\Omega, \mu) := \mathcal{L}^\infty / \mathcal{N}$ we have the norm $\|\text{class of } f\|_\infty := \inf\{\sup_{x \in \Omega \setminus N} |f(x)| \mid N \in \Sigma, \mu(N) = 0\}$.
Note that $\Omega = \mathbb{N}$ and μ the counting measure gives $l^p(\mathbb{N})$ as special cases.

(C) The Banach space of bounded Borel functions

Let $\Omega \subseteq \mathbb{C}$ and denote by $B_b(\Omega)$ the vector space of bounded Borel measurable functions $f : \Omega \to \mathbb{C}$. We obtain $B_b(\Omega)$ as a closed subspace of the Banach space of all bounded functions on Ω equipped with the supremum norm $\|f\|_\infty := \sup_{x \in \Omega} |f(x)|$, since measurability is preserved even under pointwise limits. The proof of the monotone pointwise approximation of non-negative measurable functions shows, in fact, that a bounded measurable function can be uniformly approximated by step functions (cf. [Els11, Kapitel III, Korollar 4.14(a)]), i.e., the step functions are dense in the Banach space $(B_b(\Omega), \|.\|_\infty)$.

We state a technical lemma (cf. [Wer18, Lemma VII.1.5]) that will be useful in constructing a measurable functional calculus for self-adjoint operators on Hilbert spaces. (A proof is given in the appendix.)

0.13. Lemma: Let $\Omega \subset \mathbb{C}$ be compact and $(B_b(\Omega), \|.\|_\infty)$ be the Banach space of bounded Borel measurable functions $\Omega \to \mathbb{C}$. Suppose $U \subseteq B_b(\Omega)$ has the following properties:

(a) $C(\Omega) \subseteq U$,

(b) $f_n \in U$ $(n \in \mathbb{N})$, $\sup_{n \in \mathbb{N}} \|f_n\|_\infty < \infty$, and $f(t) := \lim_{n \to \infty} f_n(t)$ exists for every $t \in \Omega$
$\implies \quad f \in U$.

Then $U = B_b(\Omega)$.

(D) Signed and complex measures

0.14. Definition: Let Σ be a σ-algebra on the set Ω. A (finite) *signed measure* is a σ-additive map $\mu : \Sigma \to \mathbb{R}$. A *complex measure* is a σ-additive map $\mu : \Sigma \to \mathbb{C}$.

In both cases, $\mu(\emptyset) = 0$ follows from σ-additivity, since we excluded the possibility of infinite values in the above definition: $\mu(\emptyset) = \mu(\emptyset \cup \emptyset) = \mu(\emptyset) + \mu(\emptyset)$. A map $\mu \colon \Sigma \to \mathbb{C}$ is a complex measure if and only if $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$ are finite signed measures.

Let μ be a signed or complex measure. We define the *variation* $|\mu| \colon \Sigma \to [0, \infty[$ of μ by

$$|\mu|(A) := \sup\{\sum_{k=1}^{n} |\mu(E_k)| \mid E_1, \ldots, E_n \in \Sigma \text{ pairwise disjoint}, A = E_1 \cup \ldots \cup E_n\}.$$

The variation $|\mu|$ is a finite (positive) measure and the so-called *Jordan decomposition* holds for a signed measure: There are finite (positive) measures μ^+ and μ^- on Σ (concentrated on measurable subsets with $|\mu|$-null overlap only) such that $\mu = \mu^+ - \mu^-$ and $|\mu| = \mu^+ + \mu^-$.

Integrability of measurable functions f with respect to a (finite) signed measure μ is defined in terms of $|\mu|$-integrability and the integral is given by $\int f \, d\mu := \int f \, d\mu^+ - \int f \, d\mu^-$. For a complex measure $\mu = \mu_1 + i\mu_2$ with real part μ_1 and imaginary part μ_2 the notion of μ-integrability is also defined in terms of $|\mu|$-integrability and the integral is then given by $\int f \, d\mu := \int f \, d\mu_1 + i \int f \, d\mu_2$.

(E) Regular measures

0.15. Definition: Let Ω be a Hausdorff space. A Borel measure μ on Ω is *regular*, if

(i) $\mu(C) < \infty$ for every compact subset $C \subseteq \Omega$,

(ii) for every $A \in \mathcal{B}(\Omega)$,

$$\mu(A) = \sup\{\mu(C) \mid C \subseteq A,\ C \text{ compact}\} = \inf\{\mu(O) \mid A \subseteq O,\ O \text{ open}\}.$$

A signed or complex Borel measure is called *regular*, if the variation $|\mu|$ is regular. We denote by $M(\Omega)$ the vector space of all signed or complex regular Borel measures on Ω.

0.16. Lemma: If Ω is a compact metric space or a complete separable metric space or an open subset of $\mathbb{R}^d$, then every finite Borel measure on Ω is regular. The Lebesgue measure is regular. (This result is included in Ulam's theorem [Els11, Kapitel VIII, Satz 1.16].)

(F) Riesz representation theorem

Let Σ be a sigma algebra on the set Ω. The vector space of signed (or complex) measures on Σ becomes a Banach space when equipped with the variation norm $\|\mu\| := |\mu|(\Omega)$ (cf. [Wer18, Abschnitt I.1, Beispiel (j), Seiten 22-24]). If Ω is a Hausdorff space and $\Sigma = \mathcal{B}(\Omega)$, then the set of signed (or complex) *regular* Borel measures $M(\Omega)$ is a closed subspace of the former ([Els11, Kapitel VIII, Folgerung 2.22.a)]), hence $(M(\Omega), \|.\|)$ is a Banach space.

0.17. Theorem (Riesz representation theorem): Let Ω be a compact metric space. Then the normed dual $C(\Omega)'$ of $(C(\Omega), \|.\|_\infty)$ is isometrically isomorphic to $(M(\Omega), \|.\|)$ via the map $R\colon M(\Omega) \to C(\Omega)'$, given by

$$(R\mu)(f) := \int_\Omega f \, d\mu.$$

For a proof we refer to [Wer18, Theorem II.2.5] (or [Els11, Kapitel VIII, §2] and [Bau01] for more general variants of the theorem).

(G) Absolutely continuous functions

We introduce a notion that is stronger than plain continuity, weaker than Lipschitz continuity, and provides the perfect setting for a general fundamental theorem of calculus.

0.18. Definition: A function $f\colon [a,b] \to \mathbb{C}$ is *absolutely continuous*, if it satisfies the following: $\forall \varepsilon > 0 \; \exists \delta > 0$ such that for any $n \in \mathbb{N}$ and sequences $[a_1, b_1], \ldots, [a_n, b_n]$ of subintervals with $a \le a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n \le b$ of $[a,b]$ we have

$$\sum_{k=1}^{n} (b_k - a_k) < \delta \quad \Longrightarrow \quad \sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon.$$

The Cantor function and $f\colon [0,1] \to \mathbb{R}$, $f(0) := 0$, $f(x) := x\sin(1/x)$ $(x > 0)$, are prominent examples of continuous functions which are not absolutely continuous. The function $g\colon [0,1] \to \mathbb{R}$, $g(x) := \sqrt{x}$, is absolutely continuous, but not Lipschitz continuous.

0.19. Theorem (Fundamental theorem of calculus): A function $f\colon [a,b] \to \mathbb{C}$ is absolutely continuous if and only if it is differentiable almost everywhere and f' defines a Lebesgue integrable function on $[a,b]$. In this case we have for every $t \in [a,b]$

$$f(t) = f(a) + \int_a^t f' \, d\mu,$$

where μ denotes the one-dimensional Lebesgue measure and $\int_a^t f' \, d\mu$ means $\int \chi_{[a,t]} f' \, d\mu$.

We may even extend the formula of integration by parts to the case of absolutely continuous functions (see [Els11, Kapitel VII, 4.16] for (a) and [Bog07, 5.8.43] for (b)).

0.20. Proposition (Integration by parts): (a) If f and g are absolutely continuous functions $[a,b] \to \mathbb{C}$, then

$$\int_a^b f' g \, d\mu = f(b)g(b) - f(a)g(a) - \int_a^b f g' \, d\mu.$$

(b) Suppose f and g are functions on $\mathbb{R}$ that are absolutely continuous on bounded intervals and such that $f, f'g, fg' \in L^1(\mathbb{R})$, then

$$\int_{-\infty}^{\infty} f'g\, d\mu = -\int_{-\infty}^{\infty} fg'\, d\mu.$$

(H) Image measures

If Σ_j is a σ-algebra on Ω_j ($j = 1, 2$) and the map $F\colon \Omega_1 \to \Omega_2$ has the property that $F^{-1}(A_2) \in \Sigma_1$ whenever $A_2 \in \Sigma_2$ (this is the notion of Σ_1-Σ_2-measurability), then a measure μ on Σ_1 can be "transported" to a measure ν on Σ_2 by setting $\nu(A_2) := \mu(F^{-1}(A_2))$. We call ν the *image measure* of μ with respect to F and write $\nu =: F(\mu)$.

0.21. Theorem: A Σ_2-measurable function $f\colon \Omega_2 \to \mathbb{C}$ is $F(\mu)$-integrable if and only if the function $f \circ F\colon \Omega_1 \to \mathbb{C}$ is μ-integrable. In this case we have the transformation formula

$$\int_{\Omega_2} f\, dF(\mu) = \int_{\Omega_1} (f \circ F)\, d\mu.$$

(We refer to [Els11, Kapitel V, §3, Unterabschnitt 1] for a proof and to [Els11, Kapitel V, §4] for further variants of transformation formulae.)

1. The spectral theorem for bounded self-adjoint operators

1.1. The finite-dimensional case: If $H = \mathbb{C}^n$, equipped with the standard inner product, and T is a self-adjoint (or normal) operator on $\mathbb{C}^n$, then we know from linear algebra that there is an orthonormal basis B of $\mathbb{C}^n$ consisting of eigenvectors of T, i.e., the matrix of T with respect to the basis B is diagonal and the eigenvalues of T, with their multiplicities, occur as the entries along the diagonal. Identifying vectors v in $\mathbb{C}^n$ with functions $v\colon \{1, 2, \ldots, n\} \to \mathbb{C}$, $v_l := v(l)$ being the l-th component of v, a diagonal matrix D with diagonal entries $d(1), \ldots, d(n) \in \mathbb{C}$ acts on v as a multiplication operator, because $(D \cdot v)(l) = d(l)v(l)$ for $l = 1, \ldots, n$; thus, we may state that the unitary transformation $U\colon \mathbb{C}^n \to \mathbb{C}^n$ associated with the orthonormal basis B (used as column vectors of U) maps T to a multiplication operator D on $l^2(\{1, \ldots, n\}) \cong \mathbb{C}^n$ via

$$D = U^{-1}TU. \tag{1.1}$$

In this finite-dimensional case, the spectrum $\sigma(T) = \{\lambda_1, \ldots, \lambda_m\}$ is the set of pairwise distinct eigenvalues. Let E_j denote the orthogonal projection onto the eigenspace of λ_j ($j = 1, \ldots, m$), then we may write the diagonal representation of T in the abstract form

$$T = \sum_{j=1}^{m} \lambda_j E_j. \tag{1.2}$$

For any polynomial function p on $\mathbb{C}$ we easily derive the following formula (using $E_j E_k = 0$, if $j \neq k$, and ${E_j}^l = E_j$ for every $l \in \mathbb{N}$): $\quad p(T) = \sum_{j=1}^{m} p(\lambda_j) E_j$.

Nothing prevents us from using the scheme of this formula to define $f(T)$ for an arbitrary function $f\colon \sigma(T) \to \mathbb{C}$, namely

$$f(T) := \sum_{j=1}^{m} f(\lambda_j) E_j.$$

As an example, it is straight-forward to show that the above definition with $f = \exp$ agrees with the usual power series definition of $e^T := \sum_{k=0}^{\infty} T^k/k!$, since $T^k = \sum_{j=1}^{m} \lambda_j^k E_j$.

1.2. The case of compact operators: If T is a compact self-adjoint (or normal) operator on the complex Hilbert space H, then the analogue of equation (1.2) holds with an infinite sum, convergent in the operator norm ([Wer18, Korollar VI.3.3]) and a multiplication operator analogue of the diagonal matrix representation (1.1) can be constructed via a unitary map $U\colon l^2(S) \to H$, where S is a set of cardinality equal to the Hilbert dimension of H (i.e., the cardinality of a, hence any, complete orthonormal system in H), such that $(U^{-1}TUx)(s) = h(s)\,x(s)$ $(s \in S)$ holds with some $h \in l^\infty(S)$ and for every $x \in l^2(S)$. (We give a sketch of the construction in the appendix.)

In the current section we discuss the unitary equivalence of a bounded self-adjoint operator T on a complex Hilbert space H with a multiplication operator on an appropriate space $L^2(\Omega, \mu)$. We will also discuss how to define $f(T)$ for a bounded Borel measurable function f on $\sigma(T)$ based on an integral representation generalizing (1.2).

1.3. Lemma: Let the *numerical range* of $T \in L(H)$ be $W(T) := \{\langle Tx, x\rangle \mid \|x\| = 1\}$. Then $W(T)$ is a bounded subset of $\mathbb{C}$ satisfying $\sigma(T) \subseteq \overline{W(T)}$.

Proof: Boundedness of $W(T)$ follows from $|\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2$. Let $\lambda \in \mathbb{C} \setminus \overline{W(T)}$ and d denote the distance from λ to the compact set $\overline{W(T)}$, hence $d > 0$. If $x \in H$ with $\|x\| = 1$, then

$$0 < d\|x\| = d \leq |\lambda - \langle Tx, x\rangle| = |\langle(\lambda - T)x, x\rangle| \leq \|(\lambda - T)x\|\|x\| = \|(\lambda - T)x\|,$$

which proves injectivity of $\lambda - T$ and that the inverse $(\lambda - T)^{-1}\colon \operatorname{ran}(\lambda - T) \to H$ is continuous with norm bounded by $1/d$; the estimate $d\|x\| \leq \|(\lambda - T)x\|$ holds for all $x \in H$ upon rescaling, which proves closedness of $\operatorname{ran}(\lambda - T)$ (cf. 0.2). We show that $\operatorname{ran}(\lambda - T)$ is also dense, thus $\operatorname{ran}(\lambda - T) = H$ and $\lambda \in \rho(T)$ and the lemma will be proved. If $\operatorname{ran}(\lambda - T)$ were not dense in H, then we had some $x_0 \in \operatorname{ran}(\lambda - T)^\perp$ with $\|x_0\| = 1$ and thus

$$0 = \langle(\lambda - T)x_0, x_0\rangle = \lambda - \langle Tx_0, x_0\rangle,$$

contradicting the fact $\lambda \notin W(T)$. □

Recalling Proposition 0.1 we immediately obtain for self-adjoint T the relations $W(T) \subset \mathbb{R}$, $\|T\| = \max\{|\lambda| \mid \lambda \in \overline{W(T)}\}$, and in combination with the lemma the following statement.

1.4. Corollary: If $T \in L(H)$ is self-adjoint, then $\sigma(T) \subset \mathbb{R}$, more precisely,

$$\sigma(T) \subseteq [m(T), M(T)],$$

where $m(T) := \inf\{\langle Tx, x\rangle \mid \|x\| = 1\}$ and $M(T) := \sup\{\langle Tx, x\rangle \mid \|x\| = 1\}$. In particular, $\sigma(T) \subset [0, \infty[$ for a positive operator T).

Let $T \in L(H)$ be self-adjoint. A first task is the definition of $f(T) \in L(H)$ for any function $f \in C(\sigma(T))$. This is very easy, if f is a polynomial function $f(t) = \sum_{k=0}^n a_k t^k$, since then the only reasonable choice is $f(T) := \sum_{k=0}^n a_k T^k$, with the convention $T^0 := I$.

For general $f \in C(\sigma(T))$, one may continuously extend[1] f to a function $\tilde{f}$ on the real interval $[m(T), M(T)] \supset \sigma(T)$ and then use approximation by polynomials according to the Weierstraß theorem (cf., e.g., [Wer18, Satz I.2.11]). This leads to the so-called continuous functional calculus, which is summarized in the following theorem. Since this construction is often contained already in standard introductory courses on functional analysis, we repeat the technical details of its proof only in the appendix.

1.5. Theorem (Continuous functional calculus): If $T \in L(H)$ is self-adjoint, then there is a unique map $\Phi\colon C(\sigma(T)) \to L(H)$ with the following properties:

(a) $\Phi(\mathrm{id}) = T$ and $\Phi(1) = I$,

(b) Φ is an involutive algebra homomorphism, i.e.,

- Φ is $\mathbb{C}$-linear,
- $\forall f, g \in C(\sigma(T))$: $\Phi(f \cdot g) = \Phi(f) \cdot \Phi(g)$,
- $\Phi(\bar{f}) = \Phi(f)^*$,

(c) Φ is continuous with respect to the norm $\|.\|_\infty$ on $C(\sigma(T))$, in fact, isometric, i.e., $\|\Phi(f)\| = \|f\|_\infty$.

We write $f(T)$ instead of $\Phi(f)$ and call $f \mapsto f(T)$ the *continuous functional calculus* of T.

The following list collects a few more properties of the continuous functional calculus.

1.6. Theorem: Let $T \in L(H)$ be self-adjoint and $f \in C(\sigma(T))$, then the following hold:

(i) $\|f(T)\| = \|f\|_\infty$,

(ii) $f \geq 0$ implies that $f(T)$ is positive,

(iii) $x \in H$ and $Tx = \lambda x$ implies $f(T)x = f(\lambda)x$,

(iv) $f(T)$ is normal; $f(T)$ is self-adjoint, if and only if f is real-valued,

(v) the *spectral mapping theorem:* $\sigma(f(T)) = f(\sigma(T))$.

Moreover, $C^*(T) := \{f(T) \in L(H) \mid f \in C(\sigma(T))\}$ is a closed involutive commutative subalgebra of $L(H)$.

Proof: (i): This is only a restatement of part (c) in the above theorem.

(ii): We may write $f = g^2$ with $g \in C(\sigma(T))$ and $g \geq 0$. We get $f(T) = g(T)g(T)$ and $g(T)^* = \overline{g}(T) = g(T)$, hence for any $x \in H$,

$$\langle f(T)x, x\rangle = \langle g(T)x, g(T)^*x\rangle = \langle g(T)x, \overline{g}(T)x\rangle = \langle g(T)x, g(T)x\rangle = \|g(T)x\|^2 \geq 0.$$

[1]With the Euclidean topology, $\mathbb{R}$ and any closed subset is a normal topological space.

(iii): This is elementary, if f is a polynomial, and extends to general continuous f by uniform polynomial approximation $p_n \to f$ (thanks to the Stone-Weierstraß theorem applied to the compact subset $\sigma(T) \subseteq \mathbb{R}$; cf., e.g., [Wer18, Satz VIII.4.7]), since this implies $p_n(T) \to f(T)$ with convergence in the operator norm.

(iv): $f(T)^* f(T) = (\bar{f} f)(T) = (f \bar{f})(T) = f(T) f(T)^*$; $\bar{f} = f$ implies $f(T)^* = \bar{f}(T) = f(T)$, and for the converse note that $f(T) = f(T)^* = \bar{f}(T)$ in view of (i) yields $\|f - \bar{f}\|_\infty = \|f(T) - \bar{f}(T)\| = 0$, hence $f = \bar{f}$.

(v): For the case of a polynomial function, this relation is elementary (and already established, e.g., in course of the proof of Theorem 1.5; cf. Claim (i) in the proof given in the appendix).

Let $\mu \in \mathbb{C} \setminus f(\sigma(T))$. Then $g := \frac{1}{f-\mu} \in C(\sigma(T))$ and $g \cdot (f - \mu) = (f - \mu) \cdot g = 1$, hence

$$g(T)(f(T) - \mu) = (f(T) - \mu)g(T) = I,$$

which shows that $\mu \in \rho(f(T))$; i.e., $\sigma(f(T)) \subseteq f(\sigma(T))$.

Let $\mu \in f(\sigma(T))$. There is some $\lambda \in \sigma(T)$ such that $\mu = f(\lambda)$ and we have to show that $f(\lambda) \in \sigma(f(T))$. For every $n \in \mathbb{N}$ choose a polynomial p_n with $\|f - p_n\|_\infty \leq 1/n$ (again thanks to the Stone-Weierstraß theorem). It follows that $\|f(T) - p_n(T)\| \leq 1/n$ as well as $|p_n(\lambda) - f(\lambda)| \leq 1/n$. Since p_n is a polynomial, we have $p_n(\lambda) \in \sigma(p_n(T))$. By Proposition 0.6 there is some approximate eigenvector $x_n \in H$ with $\|x_n\| = 1$ and $\|(p_n(T) - p_n(\lambda))x_n\| \leq 1/n$. In summary, we obtain for every $n \in \mathbb{N}$,

$$\begin{aligned}
\|(f(T) - \mu)x_n\| &= \|(f(T) - p_n(T) + p_n(T) - p_n(\lambda) + p_n(\lambda) - f(\lambda))x_n\| \\
&\leq \underbrace{\|f(T) - p_n(T)\|}_{\leq 1/n} \|x_n\| + \underbrace{\|(p_n(T) - p_n(\lambda))x_n\|}_{\leq 1/n} + \underbrace{|p_n(\lambda) - f(\lambda)|}_{\leq 1/n} \|x_n\| \leq \frac{3}{n},
\end{aligned}$$

which means, again by Proposition 0.6, that $\mu \in \sigma(f(T))$.

Since the map $f \mapsto f(T) = \Phi(f)$ is multiplicative and involutive, $C^*(T) = \Phi(C(\sigma(T)))$ is an involutive subalgebra of $L(H)$ and commutativity follows from the same property of the pointwise multiplication in $C(\sigma(T))$. The closedness of $C^*(T)$ follows from the completeness of $C(\sigma(T))$ together with the fact that $f \mapsto f(T) = \Phi(f)$ is an isometry: If $(f_n(T))_{n\in\mathbb{N}}$ is a sequence in $C^*(T)$ that converges in $L(H)$, then it is a Cauchy sequence with respect to the operator norm; isometry implies that $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $C(\sigma(T))$; hence there is some $f \in C(\sigma(T))$ such that $f_n \to f$ uniformly and thus $f_n(T) \to f(T) \in C^*(T)$. □

1.7. Remark: The notation for the closed involutive abelian algebra $C^*(T)$ shall indicate the fact that it is exactly the C^*-subalgebra of $L(H)$ generated by the normal element T and I. On the background of this "Banach algebra point of view", the above continuous functional calculus appears as a special case of isometric embeddings (Gelfand transforms) or isomorphisms of abstract commutative Banach or C^*-algebras (cf. [Con10, Chapter VIII] or [Wer18, Korollar IX.3.8]).

In the finite-dimensional case with spectral representation (1.2) the spectrum $\sigma(T) = \{\lambda_1, \ldots, \lambda_m\}$ is a discrete topological space and the corresponding projections $E_1, \ldots, E_m$ onto the eigenspaces can be obtained via functional calculus: Let $f\colon \sigma(T) \to \mathbb{C}$ be defined by $f(\lambda_j) = 0$, if $j \neq k$, and

$f(\lambda_k) = 1$, then $E_k = \sum_{j=1}^m f(\lambda_j)E_j = f(T)$. Note that $f = \chi_{\{\lambda_k\}}$, the characteristic function of the singleton $\{\lambda_k\}$. In the general case, $\sigma(T)$ will not be discrete and hence a characteristic function χ_A, $A \subset \sigma(T)$, will not be continuous. If we are able to extend the functional calculus appropriately, then $\chi_A(T)$ is a perfect way to produce an orthogonal projection, since $\chi_A^2 = \chi_A$ and χ_A is real-valued. From characteristic functions we get to step functions by linear combinations and then pointwise limits would bring us into the realm of measurable functions.

To investigate an extension of the continuous functional calculus for a self-adjoint operator $T \in L(H)$, let $x, y \in H$ and write for any $f \in C(\sigma(T))$,

$$l_{x,y}(f) := \langle f(T)x, y\rangle.$$

The map $l_{x,y} \colon C(\sigma(T)) \to \mathbb{C}$ is linear and

$$|l_{x,y}(f)| \leq \|f(T)\| \|x\| \|y\| = \|f\|_\infty \|x\| \|y\|,$$

thus, $l_{x,y}$ is an element in the dual space of $(C(\sigma(T)), \|.\|_\infty)$ with $\|l_{x,y}\| \leq \|x\| \|y\|$. By the Riesz representation theorem (see Theorem 0.17) there is a complex Borel measure $\mu_{x,y}$ on $\sigma(T)$ with $\|\mu_{x,y}\| = \|l_{x,y}\|$, such that

$$\langle f(T)x, y\rangle = \int_{\sigma(T)} f \, d\mu_{x,y} \qquad \forall f \in C(\sigma(T)). \tag{1.3}$$

Now we observe that the right-hand side of (1.3) makes sense also for $f \in B_b(\sigma(T))$ and, considering its dependence on (x, y), defines a continuous sesquilinear form $b_f \colon H \times H \to \mathbb{C}$, since $(x, y) \mapsto \mu_{x,y}$ clearly is sesquilinear $H \times H \to M(\sigma(T))$ as seen from the defining Equation (1.3) and

$$|b_f(x, y)| = \Big|\int_{\sigma(T)} f \, d\mu_{x,y}\Big| \leq \|f\|_\infty \|\mu_{x,y}\| = \|f\|_\infty \|l_{x,y}\| \leq \|f\|_\infty \|x\| \|y\|. \tag{1.4}$$

By the Lax-Milgram theorem (Theorem 0.3(b)), b_f defines an operator $f(T) \in L(H)$ such that

$$\langle f(T)x, y\rangle = b_f(x, y) \qquad \forall x, y \in H. \tag{1.5}$$

Thereby the main ingredient for the *measurable functional calculus* has been constructed, namely a map $B_b(\sigma(T)) \to L(H)$ given by $f \mapsto f(T)$.

1.8. Theorem (Measurable functional calculus): Let $T \in L(H)$ be self-adjoint, then there is a unique map $\widehat{\Phi} \colon B_b(\sigma(T)) \to L(H)$ with the following properties (we will typically write $f(T)$ to mean $\widehat{\Phi}(f)$):

(a) $\widehat{\Phi}$ is an involutive algebra homomorphism that extends the continuous functional calculus Φ, i.e., $\widehat{\Phi}|_{C(\sigma(T))} = \Phi$, and is $\|.\|_\infty$-continuous, more precisely, $\|f(T)\| \leq \|f\|_\infty$,

(b) if $f_n \in B_b(\sigma(T))$ ($n \in \mathbb{N}$) is such that $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ and $f_n(t) \to f(t)$ pointwise for $t \in \sigma(T)$, then $f_n(T)\, x \to f(T)\, x$ for all $x \in H$.

Proof: Uniqueness: Suppose $\widehat{\Psi}$ is another measurable functional calculus satisfying (a) and (b). Define $U := \{f \in B_b(\sigma(T)) \mid \widehat{\Psi}(f) = \widehat{\Phi}(f)\}$. Then $C(\sigma(T)) \subseteq U$ and U satisfies also condition (b) in Lemma 0.13, hence $U = B_b(\sigma(T))$, which proves $\widehat{\Psi} = \widehat{\Phi}$.

Existence: If $f \in B_b(\sigma(T))$, then we define $\widehat{\Phi}(f) := f(T) \in L(H)$ as above via the sesquilinear form b_f, based on the complex Borel measures $\mu_{x,y} \in M(\sigma(T))$ for all $x, y \in H$, by the property in (1.5). Linearity of $\widehat{\Phi}\colon B_b(\sigma(T)) \to L(H)$ is clear from (1.3) and we have $|\langle f(T)x, y\rangle| = |b_f(x,y)| \leq \|f\|_\infty \|x\| \|y\|$ by (1.4), hence $\|f(T)\| \leq \|f\|_\infty$, which shows continuity of $\widehat{\Phi}$.

We easily obtain an intermediate, weaker, variant of (b): If $x, y \in H$ are arbitrary and f_n $(n \in \mathbb{N})$, f are as in the premise of (b), then by the theorem on dominated convergence

$$(1.6) \qquad \langle f_n(T)x, y\rangle = \int f_n \, d\mu_{x,y} \to \int f \, d\mu_{x,y} = \langle f(T)x, y\rangle \quad (n \to \infty),$$

i.e., $f_n(T)x \to f(T)x$ weakly in H for every x.

We will first show that $\widehat{\Phi}$ is multiplicative and involutive before getting back to the improvement of (1.6) establishing (b).

We already know that $\widehat{\Phi}(fg) = \widehat{\Phi}(f)\,\widehat{\Phi}(g)$, if both f and g are continuous.
Let g be continuous and put $U := \{f \in B_b(\sigma(T)) \mid \widehat{\Phi}(fg) = \widehat{\Phi}(f)\,\widehat{\Phi}(g)\}$, then we clearly have $C(\sigma(T)) \subseteq U$. Suppose that the sequence $f_n \in U$ $(n \in \mathbb{N})$ is uniformly bounded and pointwise convergent to $f \in B_b(\sigma(T))$. Then by (1.6),

$$\langle \widehat{\Phi}(f)(\widehat{\Phi}(g)x), y\rangle = \lim_{n\to\infty} \langle \widehat{\Phi}(f_n)(\widehat{\Phi}(g)x), y\rangle = \lim_{n\to\infty} \langle \widehat{\Phi}(f_n g)x, y\rangle = \langle \widehat{\Phi}(fg)x, y\rangle \quad \forall x, y \in H,$$

which implies $\widehat{\Phi}(f)\widehat{\Phi}(g) = \widehat{\Phi}(fg)$, hence $f \in U$ and Lemma 0.13 yields $U = B_b(\sigma(T))$.
Now let f be measurable and bounded and put $V := \{g \in B_b(\sigma(T)) \mid \widehat{\Phi}(fg) = \widehat{\Phi}(f)\,\widehat{\Phi}(g)\}$. We have shown above that $C(\sigma(T)) \subseteq V$, and, employing Lemma 0.13 in the same way again (strictly speaking, upon rewriting $\langle \widehat{\Phi}(f)(\widehat{\Phi}(g)x), y\rangle = \langle \widehat{\Phi}(g)x, \widehat{\Phi}(f)^* y\rangle$ and then using $g_n \to g$), we find $V = B_b(\sigma(T))$, thus, multiplicativity of $\widehat{\Phi}$.

A similar technique is used to show $\widehat{\Phi}(f)^* = \widehat{\Phi}(\bar{f})$, which clearly holds for continuous functions: Put $U := \{f \in B_b(\sigma(T)) \mid \widehat{\Phi}(f)^* = \widehat{\Phi}(\bar{f})\}$ and note that $C(\sigma(T)) \subseteq U$. If (f_n) is a uniformly bounded sequence in U and converges pointwise to f, then

$$\langle \widehat{\Phi}(f)^* x, y\rangle = \langle x, \widehat{\Phi}(f)y\rangle = \overline{\langle \widehat{\Phi}(f)y, x\rangle} = \lim_{n\to\infty} \overline{\langle \widehat{\Phi}(f_n)y, x\rangle} = \lim_{n\to\infty} \langle \widehat{\Phi}(\overline{f_n})x, y\rangle = \langle \widehat{\Phi}(\bar{f})x, y\rangle$$

holds for arbitrary $x, y \in H$ and hence $f \in U$; thus, $U = B_b(\sigma(T))$ again by Lemma 0.13, i.e., $\widehat{\Phi}$ is involutive.

Finally, we prove (b): Let f_n $(n \in \mathbb{N})$, f be as in the premise of (b) and $x \in H$. By (1.6) we have $f_n(T)x \to f(T)x$ weakly in H. By the properties of $\widehat{\Phi}$ we obtain in addition

$$\begin{aligned} \|f_n(T)x\|^2 &= \langle f_n(T)x, f_n(T)x\rangle = \langle f_n(T)^* f_n(T)x, x\rangle = \langle (\overline{f_n} f_n)(T)x, x\rangle \\ &\to \langle (\bar{f} f)(T)x, x\rangle = \langle f(T)^* f(T)x, x\rangle = \|f(T)x\|^2 \quad (n \to \infty). \end{aligned}$$

We may therefore conclude that $f_n(T)x \to f(T)x$ in H, since $\|f_n(T)x - f(T)x\|^2 = \|f_n(T)x\|^2 - 2\operatorname{Re}\langle f_n(T)x, f(T)x\rangle + \|f(T)x\|^2 \to \|f(T)x\|^2 - 2\operatorname{Re}\langle f(T)x, f(T)x\rangle + \|f(T)x\|^2 = 0$. □

1.9. Remark: While Φ is injective, since $\|\Phi(f)\| = \|f\|_\infty$, its extension $\widehat{\Phi}$ is in general not. For example, we will prove later in this section that for any $\lambda \in \sigma(T)$, $\widehat{\Phi}(\chi_{\{\lambda\}}) \neq 0$ if and only if λ is an eigenvalue.

As first application of the measurable functional calculus we investigate the orthogonal projections obtained from characteristic functions.

1.10. Lemma: Let $T \in L(H)$ be self-adjoint, then the following hold:

(i) for every Borel set $A \subseteq \sigma(T)$, $\chi_A(T)$ is an orthogonal projection,

(ii) $\chi_\emptyset(T) = 0$ and $\chi_{\sigma(T)}(T) = I$,

(iii) if $A_1, A_2, \dots$ are pairwise disjoint Borel subsets of $\sigma(T)$ and $x \in H$, then we have with $A := \bigcup_{j=1}^\infty A_j$

$$\sum_{j=1}^{\infty} \chi_{A_j}(T)x = \chi_A(T)x,$$

(iv) $\chi_A(T)\chi_B(T) = \chi_{A\cap B}(T)$ for Borel sets $A, B \subseteq \sigma(T)$.

Proof: (i) follows from $\chi_A^2 = \chi_A$ and $\overline{\chi_A} = \chi_A$.

(ii): $\chi_\emptyset = 0 \in C(\sigma(T))$ and $\Phi(0) = 0$ (by linearity); $\chi_{\sigma(T)} = 1$ in $C(\sigma(T))$ and $\Phi(1) = I$.

(iii): Put $f_n = \sum_{j=1}^n \chi_{A_j}$, $f := \chi_A$ and apply Theorem 1.8(b).

(iv) follows from $\chi_A\chi_B = \chi_{A\cap B}$. □

1.11. Remark: In general, there is no operator norm convergent analogue of (iii), since the projection $\chi_{A_j}(T)$ has operator norm 1, unless it is the zero operator (corresponding to the case $\mu_{x,y}(A_j) = 0$ for all $x, y \in H$).

The previous lemma opens the door to so-called *projection-valued measures* on the Borel sigma algebra $\mathcal{B}(\mathbb{R})$ by considering the map $E\colon \mathcal{B}(\mathbb{R}) \to L(H)$, $A \mapsto \chi_{A\cap\sigma(T)}(T)$, which is an example of the following concept.

1.12. Definition: A map $E\colon \mathcal{B}(\mathbb{R}) \to L(H)$, $A \mapsto E_A$ is called a *spectral measure*, if every E_A is an orthogonal projection and

(i) $E_\emptyset = 0$, $E_\mathbb{R} = I$,

(ii) if $A_1, A_2, \dots$ are pairwise disjoint Borel sets in $\mathbb{R}$, $A := \bigcup_{j=1}^\infty A_j$, and $x \in H$, then

$$\sum_{j=1}^{\infty} E_{A_j}x = E_A x.$$

A spectral measure E has *compact support*, if there is a compact subset $K \subset \mathbb{R}$ such that $E_K = I$.

It is an exercise to deduce $E_A E_B = E_{A\cap B} = E_B E_A$ and monotonicity, i.e., $E_A \leq E_B$ (in the sense that $E_B - E_A$ is a positive operator), if $A \subseteq B$.

1.13. Integration with respect to a spectral measure: Let E be a spectral measure. We first define the integral of a step function $h = \sum_{j=1}^m \alpha_j \, \chi_{A_j}$ on $\mathbb{R}$ simply by

$$\int h \, dE := \sum_{j=1}^m \alpha_j E_{A_j}$$

and note that this is independent of the representation of h: If $\sum_{j=1}^m \alpha_j \, \chi_{A_j} = \sum_{l=1}^n \beta_l \, \chi_{B_l}$, we may suppose that both $A_1, \ldots, A_m$ and $B_1, \ldots, B_n$ are partitions of $\mathbb{R}$ (by possibly adding terms with coefficient 0 in the sum representations). Then $E_{A_j} = \sum_{l=1}^n E_{A_j \cap B_l}$, $E_{B_l} = \sum_{j=1}^m E_{A_j \cap B_l}$ and $\alpha_j = \beta_l$, if $A_j \cap B_l \neq \emptyset$, therefore

$$\sum_{j=1}^m \alpha_j E_{A_j} = \sum_{\substack{1\leq j\leq m\\ 1\leq l\leq n}} \alpha_j E_{A_j\cap B_l} = \sum_{l=1}^n \beta_l E_{B_l}.$$

We claim that $\| \int h \, dE \| \leq \|h\|_\infty$. In fact, if $\|x\| \leq 1$ and $h = \sum_{j=1}^m \alpha_j \chi_{A_j}$, then (since the $E_{A_j} x$ are pairwise orthogonal)

$$\left\| \left(\int h \, dE \right) x \right\|^2 = \left\| \sum_{j=1}^m \alpha_j E_{A_j} x \right\|^2 = \sum_{j=1}^m |\alpha_j|^2 \|E_{A_j} x\|^2 \leq \left(\max_{j=1,\ldots,m} |\alpha_j|^2 \right) \sum_{j=1}^m \|E_{A_j} x\|^2$$
$$= \|h\|_\infty^2 \left\| \sum_{j=1}^m E_{A_j} x \right\|^2 = \|h\|_\infty^2 \left\| E_{\bigcup_{j=1}^m A_j} x \right\|^2 \leq \|h\|_\infty^2 \|x\|^2.$$

Let $f \in B_b(\mathbb{R})$, then by density of the step functions, there is a uniformly convergent sequence of step functions $h_n \to f$. We have

$$\| \int h_n \, dE - \int h_m \, dE \| = \| \int (h_n - h_m) \, dE \| \leq \|h_n - h_m\|_\infty,$$

hence $(\int h_n \, dE)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L(H)$ and possesses a limit; furthermore, this limit is independent of the choice of approximating sequence (h_n), since any mix of such sequences yields a Cauchy sequence of integrals, which cannot have a different accumulation point. Therefore, we may define

$$\int f \, dE := \lim_{n\to\infty} \int h_n \, dE.$$

Notations such as $\int f \, dE$, $\int f(\lambda) \, dE_\lambda$, and $\int_{\mathbb{R}} f \, dE$ will be used in the sequel. The map $f \mapsto \int f \, dE$ clearly satisfies $\int \overline{f} \, dE = (\int f \, dE)^*$, is linear $B_b(\mathbb{R}) \to L(H)$, and continuous, since

$$\| \int f \, dE \| \leq \|f\|_\infty.$$

Furthermore, a real-valued function $f \in B_b(\mathbb{R})$ gives a self-adjoint operator $\int f\,dE \in L(H)$.

If E has compact support, say $E_K = I$ for a compact set $K \subset \mathbb{R}$, and $f \in B_b(K)$, then we define $\int_K f\,dE := \int \chi_K f\,dE$ (formally, upon extending f to $\mathbb{R}$, e.g., by the value 0 outside K); we will often abuse notation by still writing $\int f\,dE$ in this case. The choice of K with $E_K = I$ is not essential, since $A \in \mathcal{B}(\mathbb{R})$ with $A \cap K = \emptyset$ implies $E_A = E_A I = E_A E_K = E_{A\cap K} = 0$ and hence $\int \chi_A f\,dE = 0$.

The measurable functional calculus for a self-adjoint operator $T \in L(H)$ provided us with the means to define a compactly supported spectral measure E associated with T by $E_A := \chi_{A\cap\sigma(T)}(T)$ $(A \in \mathcal{B}(\mathbb{R}))$.

Conversely, if a compactly supported spectral measure E is given, say $E_K = I$ with $K \subset \mathbb{R}$ compact, then any $f \in C(K)$ is bounded and Borel measurable. Hence $T := \int \lambda\,dE_\lambda \in L(H)$ is defined and self-adjoint, since id_K is real-valued. We will show that the spectral measure of T is exactly E and that the integrals with respect to E give the (uniquely determined) measurable functional calculus of T.

1.14. Proposition: Let E be a compactly supported spectral measure and $T = \int \lambda\,dE_\lambda$. Then $E_{\sigma(T)} = I$, T is self-adjoint, and its measurable functional calculus is given by

$$\Psi\colon B_b(\sigma(T)) \to L(H), \quad f \mapsto \int_{\sigma(T)} f\,dE.$$

In particular, we may conclude that

$$\forall A \in \mathcal{B}(\sigma(T)): \quad \chi_A(T) = \Psi(\chi_A) = \int \chi_A\,dE = E_A.$$

Proof: We first prove condition (a) of Theorem 1.8: We already know that Ψ is involutive, linear, and continuous. Multiplicativity is clear in case of step functions, since $\Psi(\chi_A)\Psi(\chi_B) = E_A E_B = E_{A\cap B} = \Psi(\chi_{A\cap B}) = \Psi(\chi_A\chi_B)$, and follows in general by a routine argument employing uniform approximation by step functions. We clearly have $\Psi(\mathrm{id}) = T$ and, by multiplicativity, $\Psi(\mathrm{id}^n) = T^n$ for $n \in \mathbb{N}$. It remains to show $\Psi(1) = I$, then we have that Ψ coincides with the continuous functional calculus on polynomials, hence on all of $C(\sigma(T))$.

Note that $\Psi(1) = \int_{\sigma(T)} 1\,dE = E_{\sigma(T)}$, thus we need to show that $E_{\sigma(T)} = I$. In fact, we will show $E_{\rho(T)\cap\mathbb{R}} = 0$, which implies $E_{\sigma(T)} = E_{\sigma(T)} + E_{\rho(T)\cap\mathbb{R}} = E_{(\sigma(T)\cup\rho(T))\cap\mathbb{R}} = E_{\mathbb{R}} = I$.

Here, and also in the sequel, we will simplify notation by writing E_A instead of $E_{A\cap\mathbb{R}}$, if $A \subseteq \mathbb{C}$ is a Borel set.

Choose $a < b$ such that $E_{]a,b]} = I$, i.e., $E_K = I$ for some compact subset K of $]a,b]$. Let $\mu \in \rho(T)$. If $\mu \notin]a,b]$, then choosing an open neighborhood U of μ with $U \cap K = \emptyset$ yields $E_U = 0$.

It remains to consider the case $\mu \in]a,b]$. Recall that the set of invertible operators is open in $L(H)$ and that inversion is a continuous map on that set. Therefore, given $C := \|(\mu - T)^{-1}\| + 1$, there is $\delta > 0$ such that

$$\forall S \in L(H): \quad \|S - (\mu - T)\| \le \delta \quad \Rightarrow \quad S \text{ is invertible and } \|S^{-1}\| \le C.$$

We may suppose that $\delta = (b-a)/N$ for some $N \in \mathbb{N}$ and that $\delta < 1/C$.

Put $a_k := a + k\delta$ $(k = 0, \dots, N)$, $A_k :=]a_{k-1}, a_k]$ $(k = 1, \dots, N)$, and consider the step function $h := \sum_{k=1}^N a_k \chi_{A_k}$. Then $|\lambda - h(\lambda)| \leq \delta$ for all $\lambda \in]a, b]$. Note that $\sum_{k=1}^N E_{A_k} = E_{]a,b]} = I$ and $E_{A_k} E_{A_l} = E_{A_k \cap A_l} = \delta_{kl} E_{A_k}$. Writing $\mu I = \sum_{k=1}^N \mu E_{A_k}$ gives

$$\|\sum_{k=1}^N (\mu - a_k) E_{A_k} - (\mu - T)\| = \|T - \sum_{k=1}^N a_k E_{A_k}\| = \|T - \int h\, dE\| \leq \sup_{a < \lambda \leq b} |\lambda - h(\lambda)| \leq \delta,$$

hence $B := \sum_{k=1}^N (\mu - a_k) E_{A_k}$ is invertible and $\|B^{-1}\| \leq C$. Denoting by Σ' the sum only over those $k \in \{1, \dots, N\}$ such that $E_{A_k} \neq 0$ we may write $B = \Sigma'(\mu - a_k) E_{A_k}$ and $B^{-1} = \Sigma'(\mu - a_k)^{-1} E_{A_k}$. Therefore,

$$C \geq \|B^{-1}\| = \max\{\frac{1}{|\mu - a_k|} \mid 1 \leq k \leq N \text{ and } E_{A_k} \neq 0\}.$$

Since $\mu \in]a, b] = A_1 \cup \ldots \cup A_N$ there is some k such that $\mu \in A_k =]a_{k-1}, a_k]$. Therefore, $|\mu - a_k| \leq |a_{k-1} - a_k| = \delta < 1/C$, which implies $E_{A_k} = 0$. If $\mu < a_k$, then $U :=]a_{k-1}, a_k[$ is an open neighborhood of μ such that $E_U = 0$. If $\mu = a_k$, then $|\mu - a_{k+1}| = \delta < 1/C$ and also $E_{A_{k+1}} = 0$, so that $U :=]a_{k-1}, a_{k+1}[$ is an open neighborhood of μ with $E_U = 0$.

To summarize, we found for any $\mu \in \rho(T)$ an open neighborhood U of μ such that $E_U = 0$.

If $M \subset \rho(T)$ is compact, we can cover M by neighborhoods $U(\mu)$ $(\mu \in M)$ constructed as above with $E_{U(\mu)} = 0$. By compactness, finitely many $U(\mu_1), \dots, U(\mu_m)$ suffice, hence $0 \leq E_M \leq E_{\bigcup_{j=1}^m U(\mu_j)} \leq$ [2] $\sum_{j=1}^m E_{U(\mu_j)} = 0$, hence $E_M = 0$. For arbitrary $x \in H$, the finite positive Borel measure $A \mapsto \langle E_A x, x\rangle$, $\mathcal{B}(\mathbb{R}) \to [0, \infty[$ is regular due to Lemma 0.16; it vanishes on compact subsets of $\rho(T)$ and we therefore have $\langle E_{\rho(T)} x, x\rangle = 0$; since x was arbitrary and $E_{\rho(T)}$ is self-adjont, we conclude from Proposition 0.1(ii) that $\|E_{\rho(T)}\| = \sup_{\|x\| \leq 1} |\langle E_{\rho(T)} x, x\rangle| = 0$, hence $E_{\rho(T)} = 0$.

We finally show that Ψ satisfies also condition (b) of Theorem 1.8: It suffices to show the weaker variant (1.6), since the final argument in the proof of Theorem 1.8 could be repeated here with $\Psi(f_n)$ in place of $f_n(T)$.

Let $x, y \in H$ and denote by $\nu_{x,y}$ the complex measure on $\mathcal{B}(\sigma(T))$ given by $A \mapsto \langle E_A x, y\rangle$. Then for any step function $h = \sum_{j=1}^m \alpha_j \chi_{A_j}$ we have

$$\langle \Psi(h) x, y\rangle = \langle (\int h\, dE) x, y\rangle = \langle \sum_{j=1}^m \alpha_j E_{A_j} x, y\rangle = \sum_{j=1}^m \alpha_j \underbrace{\langle E_{A_j} x, y\rangle}_{\nu_{x,y}(A_j)} = \int h\, d\nu_{x,y},$$

and for $g \in B_b(\sigma(T))$ we apply uniform approximation by step functions and pass to the limit to obtain

$$\langle \Psi(g) x, y\rangle = \int g\, d\nu_{x,y}.$$

[2] $E_{A \cup B} = E_{A \setminus B} + E_B \leq E_A + E_B$ and similarly for finitely many terms

If (f_n) and f are as in the premise of condition (b), then (1.6) follows from the theorem on dominated convergence, since

$$\langle \Psi(f_n)x, y\rangle = \int f_n \, d\nu_{x,y} \to \int f \, d\nu_{x,y} = \langle \Psi(f)x, y\rangle \qquad (n \to \infty).$$

□

Let $T \in L(H)$ be self-adjoint with spectral measure E. Put $E_A := \chi_{A\cap\sigma(T)}(T)$ $(A \in \mathcal{B}(\mathbb{R}))$, so that $E_{\sigma(T)} = \chi_{\sigma(T)}(T) = I$, and define $S := \int_{\sigma(T)} \lambda \, dE_\lambda$.

We claim that $S = T$.

Let $\varepsilon > 0$ and h be a step function on $\sigma(T)$ such that $\|\operatorname{id} - h\|_\infty \le \varepsilon/2$; suppose $h = \sum_{j=1}^m \alpha_j \chi_{A_j}$ (with $A_j \in \mathcal{B}(\sigma(T))$ pairwise disjoint). Denote by $f(T)$ the functional calculus of T and by $\Psi(f)$ that of S (according to the above proposition), then

$$\|T - S\| \le \|T - h(T)\| + \|h(T) - \Psi(h)\| + \|\Psi(h) - S\| =$$
$$\|(\operatorname{id} - h)(T)\| + \|\sum_{j=1}^m \alpha_j \chi_{A_j}(T) - \int\limits_{\sigma(T)} h(\lambda) \, dE_\lambda\| + \|\int\limits_{\sigma(T)} h(\lambda) \, dE_\lambda - \int\limits_{\sigma(T)} \lambda \, dE_\lambda\| =$$
$$\|(\operatorname{id} - h)(T)\| + \|\sum_{j=1}^m \alpha_j \chi_{A_j}(T) - \sum_{j=1}^m \alpha_j E_{A_j}\| + \|\int\limits_{\sigma(T)} (h - \operatorname{id}) \, dE\| \le$$
$$\|\operatorname{id} - h\|_\infty + \|\sum_{j=1}^m \alpha_j \underbrace{(\chi_{A_j}(T) - E_{A_j})}_{=0}\| + \|h - \operatorname{id}\|_\infty \le \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{2} = \varepsilon.$$

Thus, we have proved the following statement.

1.15. Theorem (Spectral theorem for self-adjoint bounded operators): Let $T \in L(H)$ be self-adjoint, then there exists a unique spectral measure E with compact support on $\mathbb{R}$ such that

$$T = \int\limits_{\sigma(T)} \lambda \, dE_\lambda.$$

The map $f \mapsto \int_{\sigma(T)} f \, dE$, $B_b(\sigma(T)) \to L(H)$, defines the measurable functional calculus of T, and $f(T)$ is determined by the measures $\mu_{x,y} \in M(\sigma(T))$ $(x, y \in H)$, $\mu_{x,y}(A) = \langle E_A x, y\rangle$, via

$$\langle f(T)x, y\rangle = \int\limits_{\sigma(T)} f \, d\mu_{x,y} =: \int\limits_{\sigma(T)} f(\lambda) \, d\langle E_\lambda x, y\rangle \qquad \forall x, y \in H.$$

1.16. The finite-dimensional case revisited: Let $H = \mathbb{C}^n$, T self-adjoint on H with $\sigma(T) = \{\lambda_1, \ldots \lambda_m\}$, and $E_1, \ldots, E_m$ denote the projections onto the eigenspaces. Then $\mathcal{B}(\sigma(T)) = \mathcal{P}(\sigma(T))$ and putting for any $A \subseteq \sigma(T)$

$$E_A := \sum_{\{j \in \{1,\ldots,m\} | \lambda_j \in A\}} E_j$$

defines the spectral measure of T, since $E_{\sigma(T)} = \sum_{j=1}^m E_j = I$ and

$$T = \sum_{j=1}^{m} \lambda_j E_j = \sum_{j=1}^{m} \lambda_j E_{\{\lambda_j\}} = \int_{\sigma(T)} \lambda \, dE_\lambda.$$

1.17. Example: Let $H = L^2([0,1])$ and $T\colon L^2([0,1]) \to L^2([0,1])$ be the multiplication operator $(Tx)(t) = tx(t)$ $(t \in [0,1],\ x \in L^2([0,1]))$. Clearly, T is self-adjoint and $T - \lambda$ possesses a continuous inverse, namely, $((T-\lambda)^{-1}x)(t) = x(t)/(t-\lambda)$, if and only if $\lambda \in \mathbb{C} \setminus [0,1]$, hence $\sigma(T) = [0,1]$. It is an exercise to show that T has no eigenvalues, thus, by Lemma 0.5 we have $\sigma(T) = \sigma_c(T) = [0,1]$. Let $x, y \in L^2([0,1])$, then

$$\langle Tx, y\rangle = \int_0^1 \lambda x(\lambda)\overline{y(\lambda)}\, d\lambda = \int_{\sigma(T)} \lambda\, x(\lambda)\overline{y(\lambda)}\, d\lambda,$$

which suggests that $\mu_{x,y}$ is Lebesgue measure on $[0,1]$ with density function $x\overline{y}$, i.e, that the spectral measure E is defined by $(E_A x)(t) := \chi_{A\cap[0,1]}(t)x(t)$ $(A \in \mathcal{B}(\mathbb{R}))$. Indeed[3], $\langle E_A x, y\rangle = \int_0^1 \chi_{A\cap[0,1]}(\lambda)x(\lambda)\overline{y(\lambda)}\, d\lambda = \int_{A\cap[0,1]} x(\lambda)\overline{y(\lambda)}\, d\lambda$ and

$$\int_{\sigma(T)} \lambda \, d\langle E_\lambda x, y\rangle = \int_{\sigma(T)} \mathrm{id}\ d\mu_{x,y} = \langle Tx, y\rangle = \int_0^1 \lambda x(\lambda)\overline{y(\lambda)}\, d\lambda.$$

1.18. Remark: Let $T \in L(H)$ be self-adjoint with spectral measure E. An operator $S \in L(H)$ commutes with T, if and only if S commutes with every E_A $(A \in \mathcal{B}(\mathbb{R}))$. The statement can be shown by first noting that $ST = TS$ is equivalent to S commuting with all powers T^n, in other words, that $\langle ST^n x, y\rangle = \langle T^n Sx, y\rangle$ holds for all $x, y \in H$. Using

$$\int \lambda^n \, d\langle E_\lambda Sx, y\rangle = \langle T^n Sx, y\rangle = \langle ST^n x, y\rangle = \langle T^n x, S^* y\rangle$$
$$= \int \lambda^n \, d\langle E_\lambda x, S^* y\rangle = \int \lambda^n \, d\langle SE_\lambda x, y\rangle$$

and density of polynomials, this in turn can be seen to mean that the complex measures $\nu_1(A) := \langle SE_A x, y\rangle$ and $\nu_2(A) := \langle E_A Sx, y\rangle$ agree when considered as functionals on $C(\sigma(T))$. Varying x and y we obtain $SE_A = E_A S$.

1.19. Proposition: If $S \in L(H)$ is self-adjoint, $g\colon \sigma(S) \to \mathbb{R}$ and $f\colon \mathbb{R} \to \mathbb{C}$ are bounded and Borel measurable, then

$$(f \circ g)(S) = f(g(S)).$$

Proof: The composition $f \circ g\colon \sigma(S) \to \mathbb{C}$ is Borel measurable and bounded, therefore $(f \circ g)(S)$ is defined by the measurable functional calculus. Let F denote the spectral measure of S. Since g is real-valued, $g(S)$ is self-adjoint and possesses as a unique spectral measure E. It suffices to prove

[3]Strictly speaking,we also need to check that E is a spectral measure, but this is easy.

the statement of the proposition in case $f = \chi_A$ ($A \in \mathcal{B}(\mathbb{R})$), since it can then be extended to step functions and measurable functions by the usual techniques and properties of the functional calculus.

We have $\chi_A \circ g = \chi_{g^{-1}(A)}$, thus it suffices to show $F_{g^{-1}(A)} = E_A$ for every $A \in \mathcal{B}(\mathbb{R})$. This in turn is equivalent to the equality of the measures $\mu_{x,y}(A) := \langle E_A x, y\rangle$ and $\nu_{x,y}(A) := \langle F_{g^{-1}(A)} x, y\rangle$ for arbitrary $x, y \in H$. Note that $\nu_{x,y}$ is the image measure of the measure $\rho_{x,y}(B) := \langle F_B x, y\rangle$ ($B \in \mathcal{B}(\mathbb{R})$) associated with the spectral measure F of S, since $\nu_{x,y}(A) = \langle F_{g^{-1}(A)} x, y\rangle = \rho_{x,y}(g^{-1}(A))$, i.e., $\nu_{x,y} = g(\rho_{x,y})$. We obtain

$$\int \lambda^n \, d\nu_{x,y} = \int g(\lambda)^n \, d\rho_{x,y} = \langle g(S)^n x, y\rangle = \int \lambda^n \, d\langle E_\lambda x, y\rangle = \int \lambda^n \, d\mu_{x,y},$$

therefore $\nu_{x,y}$ and $\mu_{x,y}$ are equal as (continuous) linear functionals on polynomial functions on $\sigma(S)$. By the theorem of Weierstraß, the two measures are equal on $C(\sigma(S))$. Since x, y were arbitrary, we have shown $F_{g^{-1}(A)} = E_A$. □

1.20. Example (roots of positive operators): Let $T \in L(H)$ be positive and $n \in \mathbb{N}$, then there is a unique positive $S \in L(H)$ such that $S^n = T$. Existence is clear, if we note that $\sigma(T) \subseteq [0, \infty[$ and put $S := T^{1/n}$. To show uniqueness, suppose S is positive, hence $\sigma(S) \subseteq [0, \infty[$, and satisfies $S^n = T$. Let $g\colon \sigma(S) \to \mathbb{R}$ be defined by $g(s) := s^n$ and $f\colon \mathbb{R} \to \mathbb{R}$, $f(t) := (t\chi_{\sigma(T)}(t))^{1/n}$. Then $f \circ g = \mathrm{id}_{\sigma(S)}$, since $\sigma(T) = \sigma(S^n) = g(\sigma(S))$, and we obtain from Proposition 1.19

$$S = (f \circ g)(S) = f(g(S)) = f(S^n) = f(T) = T^{1/n}.$$

1.21. Remark (polar decomposition): For arbitrary $T \in L(H)$, the operator T^*T is positive and we may define the positive operator $|T| := (T^*T)^{1/2}$. Note that $\||T|x\|^2 = \langle x, |T|^2 x\rangle = \langle x, T^*Tx\rangle = \|Tx\|^2$, which enables us to define $U\colon \mathrm{ran}(|T|) \to \mathrm{ran}(T)$ by $U(|T|x) := Tx$ and extend it as an isometry to $\overline{\mathrm{ran}(|T|)} \to \overline{\mathrm{ran}(T)}$. Putting $U = 0$ on $\overline{\mathrm{ran}(|T|)}^\perp = \mathrm{ran}(|T|)^\perp = \ker(|T|) = \ker(T)$, we obtain an operator $U \in L(H)$ with $\ker(U) = \ker(T)$, which acts as an isometry $\ker(T)^\perp \to \overline{\mathrm{ran}(T)}$, a so-called *partial isometry*, with the property $U|T| = T$. This relation is called *polar decomposition* of T, since it resembles the similar formula $z = e^{i\alpha}|z|$ for complex numbers. (It can be shown that $U|T| = T = (TT^*)^{1/2}\, U$, see, e.g., [KR, Theorem 6.1.2].)

1.22. Description of the spectrum in terms of the spectral measure: Let $T \in L(H)$ be self-adjoint with spectral measure E and λ be a complex number.

(a) $\lambda \notin \sigma(T)$ if and only if there is an open neighborhood $U \subseteq \mathbb{C}$ of λ such that $E_U = 0$.

Proof: By construction, $E_{\sigma(T)} = I$ (Proposition 1.14), thus $E_{\rho(T)} = 0$, and the resolvent set $\rho(T)$ is open, which proves the 'only if' part. To show the reverse implication, suppose U is an open neighborhood of λ such that $E_U = 0$. Define the bounded measurable function $f\colon \sigma(T) \to \mathbb{C}$ by $f(t) := 1/(\lambda - t)$, if $t \in \sigma(T) \setminus U$, and $f(t) := 0$, if $t \in U \cap \sigma(T)$. The function $g\colon \sigma(T) \to \mathbb{C}$, $g(t) := \lambda - t$ is also bounded and measurable and $f \cdot g = \chi_{\sigma(T)\setminus U}$, hence

$$f(T)(\lambda - T) = f(T)g(T) = (fg)(T) = \chi_{\sigma(T)\setminus U}(T) = E_{\sigma(T)\setminus U} = E_{\sigma(T)} - E_U = I.$$

The same holds for $(\lambda - T)f(T) = g(T)f(T) = (gf)(T) = (fg)(T)$, thus showing that $f(T) \in L(H)$ is an inverse of $\lambda - T$, i.e., $\lambda \in \rho(T)$. □

(b) $\lambda \in \sigma_p(T)$ (i.e., λ is an eigenvalue of T) if and only if $E_{\{\lambda\}} \neq 0$.
In this case, $E_{\{\lambda\}}$ projects onto the eigenspace of λ.

Proof: We show that $\operatorname{ran}(E_{\{\lambda\}}) = \ker(\lambda - T)$, then the result follows immediately.
• $\operatorname{ran}(E_{\{\lambda\}}) \subseteq \ker(\lambda - T)$: $x \in \operatorname{ran}(E_{\{\lambda\}})$ means $E_{\{\lambda\}}x = x$; let $y \in H$ arbitrary and note that $(\lambda - t)\chi_{\{\lambda\}}(t) = 0$ for all t to obtain

$$\langle(\lambda - T)x, y\rangle = \langle(\lambda - T)E_{\{\lambda\}}x, y\rangle = \int (\lambda - t)\chi_{\{\lambda\}}(t)\, d\langle E_t x, y\rangle = 0;$$

thus, $(\lambda - T)x = 0$.
• $\ker(\lambda - T) \subseteq \operatorname{ran}(E_{\{\lambda\}})$: If $Tx = \lambda x$, then $f(T)x = f(\lambda)x$ for every $f \in C(\sigma(T))$ by Theorem 1.6(iii); by the properties of the measurable functional calculus(Theorem 1.8) and an application of Lemma 0.13, we deduce that the relation also holds with $f \in B_b(\sigma(T))$; putting $f = \chi_{\{\lambda\}}$ yields $E_{\{\lambda\}}x = \chi_{\{\lambda\}}(\lambda)x = x$, thus, $x \in \operatorname{ran}(E_{\{\lambda\}})$. □

(c) If λ is an isolated point in $\sigma(T)$, then $\lambda \in \sigma_p(T)$.

Proof: Choose an open set U with $U \cap \sigma(T) = \{\lambda\}$; we have $E_U = E_{U\setminus\{\lambda\}} + E_{\{\lambda\}} = E_{\{\lambda\}}$, since $U \setminus \{\lambda\} \subseteq \rho(T)$ and $E_{\rho(T)} = 0$; U is an open neighborhood of λ and $\lambda \in \sigma(T)$, hence $E_U \neq 0$ by (a); thus, $E_{\{\lambda\}} \neq 0$ and therefore $\lambda \in \sigma_p(T)$ by (b). □

(d) $\sigma(T)$ is the smallest compact subset with $E_{\sigma(T)} = I$.

Proof: Let $K \subseteq \mathbb{R}$ be compact with $E_K = I$. Then $E_{\mathbb{R}\setminus K} = 0$ and, by (a), $\mu \in \mathbb{R} \setminus K$ implies $\mu \notin \sigma(T)$, i.e., $\sigma(T) \subseteq K$. □

We will now discuss the aspect of diagonalization in terms of unitary equivalence with a multiplication operator.

1.23. Multiplication operator version of the spectral theorem: A vector $x \in H$ is said to be a *cyclic vector* for an operator $S \in L(H)$, if the linear span of the set $\{x, Sx, S^2x, \ldots\}$ is dense in H. (Note that then H is necessarily separable, hence possesses a complete orthonormal system of at most countable cardinality.) The multiplication operator T in Example 1.17 possesses the constant function $1 \in L^2([0,1])$ as a cyclic vector, because $(T^n x)(t) = t^n x(t)$ and therefore $\operatorname{span}\{1, T1, T^2 1, \ldots\}$ contains the L^2-dense subspace of all polynomial functions on $[0,1]$. In contrast, the identity operator $I \in L(H)$ never possesses a cyclic vector, unless the dimension of H is 0 or 1.

Theorem A: Let $T \in L(H)$ be self-adjoint with spectral measure E and suppose $x \in H$ is a cyclic vector for T. If $\mu := \mu_{x,x}$ (recall $\mu_{x,x}(A) = \langle E_A x, x\rangle$), then there exists a unitary operator $U\colon H \to L^2(\sigma(T), \mu)$ such that we have, for every $\varphi \in L^2(\sigma(T), \mu)$,

$$(UTU^{-1}\varphi)(t) = t\,\varphi(t) \qquad \mu\text{-almost everywhere.}$$

Proof: Start out with $\varphi \in C(\sigma(T)) \subseteq L^2(\sigma(T), \mu)$, then we have

$$\|\varphi\|_2^2 = \int_{\sigma(T)} |\varphi(t)|^2\, d\mu(t) = \int_{\sigma(T)} \overline{\varphi(t)}\,\varphi(t)\, d\langle E_t x, x\rangle = \langle \varphi(T)^*\varphi(T)x, x\rangle = \|\varphi(T)x\|^2,$$

i.e., the linear map $V_0\colon C(\sigma(T)) \to H$, $\varphi \mapsto \varphi(T)x$ is isometric with respect to $\|.\|_2$. The unique continuous extension $V\colon L^2(\sigma(T),\mu) \to H$ is linear and also isometric. Moreover, $\operatorname{ran}(V) \supseteq \operatorname{ran}(V_0) \supseteq \operatorname{span}\{x, Tx, T^2x, \ldots\}$ is dense in H, hence V is also surjective (since $\|V\psi\| = \|\psi\|_2$ shows that V has closed range; cf. 0.2); therefore V is a unitary map. It satisfies the following relation for every φ in the dense subspace $C(\sigma(T)) \subseteq L^2(\sigma(T),\mu)$:

$$T(V\varphi) = T(\varphi(T)x) = (\mathrm{id}(T)\varphi(T))x = (\mathrm{id}\cdot\varphi)(T)x = V(\mathrm{id}\cdot\varphi),$$

in other words, $(V^{-1}TV\varphi)(t) = t\,\varphi(t)$ holds for all $\varphi \in L^2(\sigma(T),\mu)$ for μ-almost all $t \in \sigma(T)$. Thus, the theorem is proved upon setting $U := V^{-1}$. □

In the sense of Theorem A, the multiplication operator in Example 1.17 is the prototype of any self-adjoint operator with (purely continuous) spectrum $[0,1]$ and a cyclic vector x such that $\mu_{x,x}$ is (equivalent to) the Lebesgue measure.

The extension of the above theorem to the general situation of a self-adjoint operator T on a Hilbert space H faces merely technical difficulties. The major step in overcoming these is to show that H can be decomposed into T-invariant subspaces H_j ($j \in J$, J some index set) with cyclic vectors in each H_j. Then Theorem A applies to each restriction T_j of T to these components and we come up with an orthogonal direct sum of multiplication operators on L^2-spaces with respect to different measure spaces $(\sigma(T_j), \mathcal{B}(\sigma(T_j)), \mu_j)$. This still yields a multiplication operator—although not given as multiplication by a "global identity function"—on the direct sum, which can be implemented as a single space $L^2(\Omega,\mu)$ by artificially producing a disjoint union Ω of the sets $\sigma(T_j)$ and consider the "sum" μ of the measures μ_j on it. We do not go into the details of the constructions here, but refer to [Wer18, Satz VII.1.21 and Lemma VII.1.22] for a proof in the separable case and to [Con10, Chapter IX, Theorem 4.6] or [Kab14, Abschnitt 15.3] for a proof of the general situation.

Theorem B: Let $T \in L(H)$ be self-adjoint, then there exists a measure space (Ω, Σ, μ), a bounded measurable function $f\colon \Omega \to \mathbb{R}$, and a unitary operator $U\colon H \to L^2(\Omega,\mu)$ such that we have, for every $\varphi \in L^2(\Omega,\mu)$,

$$UTU^{-1}\varphi = f\varphi \qquad \mu\text{-almost everywhere.}$$

1.24. Example (convolution operators): [In this example we call on a few more basic results from measure theory and Fourier analysis that are typically covered by the companion master level course on Real Analysis; we may also refer to [Con16, Fol99].] Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $x \in L^2(\mathbb{R})$, then for every $s \in \mathbb{R}$, the function $t \mapsto f(s-t)$ belongs to L^2 as well and

$$(f * x)(s) := \int_{\mathbb{R}} f(s-t)x(t)\,dt$$

defines a measurable function $f * x\colon \mathbb{R} \to \mathbb{C}$, the so-called *convolution* of f and x, which obeys the estimate (a special case of Young's inequality)

$$\|f * x\|_2 \le \|f\|_1 \|x\|_2.$$

The map $x \mapsto f * x$ defines a linear continuous operator $C_f \colon L^2(\mathbb{R}) \to L^2(\mathbb{R})$, which is self-adjoint, if $f(r) = \overline{f(-r)}$ holds for all $r \in \mathbb{R}$, since

$$\langle C_f x, y\rangle = \int\int f(s-t)x(t)\,dt\,\overline{y(s)}\,ds = \int x(t)\overline{\int \overline{f(s-t)}\,y(s)\,ds}\,dt,$$

$$\text{while } \langle x, C_f y\rangle = \int x(t)\overline{\int f(t-s)\,y(s)\,ds}\,dt.$$

The Fourier transform, originally defined for functions $g \in L^1(\mathbb{R})$ via a formula like $\widehat{g}(\tau) = \int \exp(-it\tau)g(t)\,dt/\sqrt{2\pi}$ (factors involving 2π are a matter of convention), extends to a unitary operator $\mathcal{F}$ on $L^2(\mathbb{R})$ (Plancherel's theorem) and is famous for intertwining differentiation with multiplication by id (on functions with appropriate regularity or decay properties) as well as transforming convolution into multiplication, i.e.,

$$\mathcal{F}(f * x) = \sqrt{2\pi}\,\widehat{f}\cdot\widehat{x} \qquad (f \in L^1 \cap L^2, x \in L^2).$$

By the Riemann-Lebesgue lemma, $\widehat{f}$ is continuous and vanishes at infinity.

Thus, C_f is unitarily equivalent to the operator M_h of multiplication by $h = \sqrt{2\pi}\,\widehat{f}$, since $\mathcal{F}C_f x = \mathcal{F}(f * x) = h \cdot \widehat{x} = M_h\widehat{x}$, i.e.,

$$\mathcal{F}C_f\mathcal{F}^{-1} = M_h.$$

Unitarily equivalent operators have the same spectrum, since $\lambda - T$ is invertible if and only if $\lambda - UTU^{-1}$ is. Hence we may determine $\sigma(C_f)$ via $\sigma(M_h)$, which in case of a continuous function h vanishing at infinity can be shown to be $\overline{h(\mathbb{R})} = h(\mathbb{R}) \cup \{0\}$. Furthermore, a number λ can be seen to be an eigenvalue of M_h, hence also of C_f, if and only if the set $\{t \in \mathbb{R} \mid h(t) = \lambda\} = h^{-1}(\{\lambda\})$ has positive Lebesgue measure.

1.25. Towards a spectral theorem for normal operators: If $T \in L(H)$ is a normal operator, then T and T^* commute and

$$S_1 := \frac{1}{2}(T + T^*), \quad S_2 := \frac{1}{2i}(T - T^*)$$

are commuting self-adjoint operators such that $T = S_1 + iS_2$. Let E, F denote the spectral measure of S_1, S_2, respectively. For arbitrary $A, B \in \mathcal{B}(\mathbb{R})$, the projections E_A and F_B commute due to Remark 1.18. The operator $G_{A\times B} := E_A F_B$ is an orthogonal projection, since $G^*_{A\times B} = F^*_B E^*_A = F_B E_A = E_A F_B = G_{A\times B}$ and $G^2_{A\times B} = (E_A F_B)(E_A F_B) = E^2_A F^2_B = E_A F_B = G_{A\times B}$. For arbitrary $x, y \in H$, it can be shown that the map $A \times B \mapsto \langle G_{A\times B}x, y\rangle$ can be extended to a complex measure on $\mathcal{B}(\mathbb{R}^2)$, hence also on $\mathcal{B}(\mathbb{C})$, and that the following spectral theorem can be obtained for $T = S_1 + iS_2$ along the lines of the constructions for self-adjoint operators carried out above. We skip the technical details and may refer to, e.g., [Con10, Chapters VIII and IX] or [Kab14, Abschnitte 15.3-5] for the statements following below. We observe that the spectral representation at least becomes plausible by this sketch of a calculation:

$$T = S_1 + iS_2 = S_1 \cdot I + iI \cdot S_2 = \int \lambda\,dE_\lambda \cdot \int 1\,dF_\mu + i\int 1\,dE_\lambda \cdot \int \mu\,dF_\mu$$

$$= \iint (\lambda + i\mu)\,dE_\lambda\,dF_\mu = \int (\lambda + i\mu)\,dG_{(\lambda,\mu)} = \int z\,dG_z.$$

Theorem: Let $T \in L(H)$ be normal, then there exists a unique spectral measure $G\colon \mathcal{B}(\mathbb{C}) \to L(H)$ with compact support such that

$$T = \int\limits_{\sigma(T)} \lambda \, dG_\lambda.$$

The map $f \mapsto \int f(\lambda)\, dG_\lambda$, $B_b(\sigma(T)) \to L(H)$ defines the measurable functional calculus for T. Moreover, T is unitarily equivalent to an operator on some $L^2(\Omega, \mu)$ of multiplication by a complex bounded measurable function.

Spectral mapping theorem: If $f \in C(\sigma(T))$, then $\sigma(f(T)) = f(\sigma(T)) = \{f(\lambda) \mid \lambda \in \sigma(T)\}$.

1.26. Example (unitary operators): Let $U \in L(H)$ be unitary, then U is normal. Recall that

$$\sigma(U) \subseteq S^1 = \{z \in \mathbb{C} \mid |z| = 1\},$$

which can be argued as follows: Since U is bijective and $\|U\| = 1$, we have $0 \in \rho(U)$ and $\{\lambda \in \mathbb{C} \mid |\lambda| > 1\} \subseteq \rho(U)$; if $0 < |\lambda| < 1$, then λU as well as $(\frac{1}{\lambda} - U^*)$ are bijective, hence $\lambda - U = (-\lambda U)(\frac{1}{\lambda} - U^*)$ is continuously invertible showing that $\lambda \in \rho(U)$.

Consider the function $\arg\colon S^1 \to \mathbb{R}$ defined uniquely by the requirements $e^{i\arg(z)} = z$ and $\arg(S^1) \subseteq]-\pi, \pi]$. The function arg is bounded by construction and continuous except for a jump at $z = -1$, hence Borel measurable. The measurable functional calculus for U allows us to define $A := \arg(U) \in L(H)$, which is self-adjoint, because arg is real-valued. The relation $(\exp \circ (i \arg))(z) = \exp(i \arg(z)) = z$ now implies

$$\exp(iA) = \exp(i \arg(U)) = U.$$

Note that $\|A\| = \|\arg(U)\| \leq \|\arg\|_\infty = \pi$ implies $\sigma(A) \subseteq [-\pi, \pi]$.

1.27. Remark (Gelfand theory for commutative unital C*-algebras): As with the continuous functional calculus for self-adjoint bounded operators described in Theorem 1.6, we may consider the commutative closed subalgebra $C^*(T) := \{f(T) \mid f \in C(\sigma(T))\}$ of $L(H)$ now for an arbitrary *normal operator* $T \in L(H)$. Both $L(H)$ and $C^*(T)$ are examples of a unital C^*-algebra, i.e., a Banach algebra $\mathcal{A}$ with a unit element $I \in \mathcal{A}$ and a conjugate linear involution map $\mathcal{A} \to \mathcal{A}$, $A \mapsto A^*$ satisfying $(A^*)^* = A$ and $(AB)^* = B^*A^*$ such that $\|A^*A\| = \|A\|^2$. Since $\|AB\| \leq \|A\|\|B\|$ holds in any normed algebra, one easily deduces that involution is an isometry, i.e., $\|A^*\| = \|A\|$ (see [Con10, Chapter VIII, Proposition 1.7] or [Wer18, Lemma IX.3.3]).

While $L(H)$ is not commutative[4], for any normal operator T, the C^*-subalgebra $C^*(T) \subseteq L(H)$ is commutative and is, in fact, the smallest C^*-algebra of $L(H)$ containing I, T, and T^*. Further examples of *commutative* unital C^*-algebras are provided by $C(X) = \{f\colon X \to \mathbb{C} \mid f \text{ continuous}\}$ for any compact Hausdorff space X, with involution $f^*(x) := \overline{f(x)}$ $(x \in X)$ and the supremum norm $\|.\|_\infty$. According to the so-called Gelfand-Naimark theorem, every unital commutative C^*-algebra is of this latter form (cf. [Con10, Chapter VIII, Theorem 2.1] or [Wer18, Theorem IX.3.4]).

[4] unless H is one-dimensional

The basic construction of the *Gelfand isomorphism* establishing the Gelfand-Naimark theorem proceeds as follows: Let $\mathcal{A}$ be a unital *commutative* C^*-algebra and denote by $X(\mathcal{A})$ the set of all nonzero *characters* on $\mathcal{A}$, i.e., multiplicative linear functionals $\mathcal{A} \to \mathbb{C}$. It can be shown that $X(\mathcal{A}) \subseteq \mathcal{A}'$ and that every character is also continuous with respect to the weak* topology on $\mathcal{A}$ (cf. Example 4.12.1)). Moreover, the character space $X(\mathcal{A})$ is easily seen to be a weak* closed subset of the unit ball in $\mathcal{A}'$, hence is weak* compact by the Alaoglu-Bourbaki theorem (see Section 6). Recall the canonical isometry $\iota\colon \mathcal{A} \to \mathcal{A}''$, given by $\iota_A(\mu) := \mu(A)$ for all $A \in \mathcal{A}$ and $\mu \in \mathcal{A}'$. Using the fact $X(\mathcal{A}) \subseteq \mathcal{A}'$, the Gelfand transformation $\mathcal{A} \to C(X(\mathcal{A}))$ is given by $A \mapsto \widehat{A} := \iota_A|_{X(\mathcal{A})}$ and can be shown to be an isometric *-isomorphism

$$\mathcal{A} \cong C(X(\mathcal{A})).$$

In case $\mathcal{A} = C^*(T)$ one can show that $X(C^*(T))$ is homeomorphic to $\sigma(T)$ (see [Wer18, Satz IX.3.6] or [Con10, Chapter VIII, Proposition 2.3]) and therefore the Gelfand transformation implies

$$C^*(T) \cong C(\sigma(T)).$$

This provides yet another view on the continuous functional calculus for the normal operator T.

2. Unbounded operators

Differential operators, in applications often supplied with boundary conditions, are prominent examples of linear maps between function spaces, but they can typically not be implemented as bounded operators with respect to L^2-norms. For example, let $V := \{f \in C^1([0,1]) \mid f(0) = 0 = f(1)\}$ and $T_0 \colon V \to C([0,1])$ be given by $f \mapsto if'$. Then T_0 is bounded, if we equip V with the norm $\|f\|_{\infty,1} = \|f\|_\infty + \|f'\|_\infty$ and $C([0,1])$ with $\|.\|_\infty$ or $\|.\|_2$. But the map $f \mapsto if'$ is unbounded (hence discontinuous) considered as linear map $T \colon V \to L^2([0,1])$ with the L^2-norm put on both spaces[1]. But the L^2-setting still has the advantage of reflecting the following symmetry property of T, being defined on the (dense) subspace $V \subseteq L^2([0,1])$:

$$\forall x, y \in V \colon \quad \langle Tx, y\rangle = i \int_0^1 x'(t)\, \overline{y(t)}\, dt = ix(t)\, \overline{y(t)}\Big|_{t=0}^{t=1} - i \int_0^1 x(t)\, \overline{y'(t)}\, dt = \langle x, Ty\rangle.$$

2.1. Definition: (a) Consider a linear map $T \colon \mathrm{dom}(T) \to H$, where $\mathrm{dom}(T) \subseteq H$ is a subspace. Then T is called an *operator on H with domain* $\mathrm{dom}(T)$. The operator T is said to be *densely defined*, if $\mathrm{dom}(T)$ is dense in H.

(b) An operator $S \colon \mathrm{dom}(S) \to H$ is said to be an *extension* of T, in notation $T \subseteq S$, if $\mathrm{dom}(T) \subseteq \mathrm{dom}(S)$ and $Sx = Tx$ holds for all $x \in \mathrm{dom}(T)$.

(c) Two operators T and S are equal, we write $T = S$, if $T \subseteq S$ and $S \subseteq T$, i.e., $\mathrm{dom}(T) = \mathrm{dom}(S)$ and $Tx = Sx$ for all $x \in \mathrm{dom}(T)$.

(d) An operator $T \colon \mathrm{dom}(T) \to H$ is *symmetric*, if

$$\langle Tx, y\rangle = \langle x, Ty\rangle \quad \forall x, y \in \mathrm{dom}(T).$$

In the above sense, any element $S \in L(H)$ is an operator on H with $\mathrm{dom}(S) = H$. Clearly, if T is a densely defined operator on H which happens to be *continuous* (with respect to the Hilbert space norm restricted to $\mathrm{dom}(T)$), then T has a unique extension $T \subseteq S \in L(H)$. Finally, by the Hellinger-Toeplitz theorem ([Wer18, Satz V.5.5] or [Tes14, §4.1, Theorem 4.9]), a symmetric operator with $\mathrm{dom}(T) = H$ is bounded and self-adjoint, i.e., $T \in L(H)$ with $T^* = T$.

The operator T on $L^2([0,1])$ mentioned above, with domain $V = \{f \in C^1([0,1]) \mid f(0) = 0 = f(1)\}$ and given by $Tf = if'$, is symmetric. If $Sf = if'$ with $\mathrm{dom}(S) := \{f \in C^1([0,1]) \mid f(0) = f(1)\}$, then S is a symmetric extension of T.

[1] To have an explicit example at hands: We find that $f_n(t) := \sin(n\pi t)$ satisfies $f_n \in V$ and $\|f_n\|_2 = 1/\sqrt{2}$, whereas $\|Tf_n\|_2 = n\pi/\sqrt{2} \to \infty$ as $n \to \infty$.

Recall that the adjoint T^* of a bounded operator T is uniquely defined by the requirement $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in H$ and can be obtained by assigning to $y \in H$ the unique vector $z \in H$ such that the linear functional $x \mapsto \langle Tx, y\rangle$ is represented by $x \mapsto \langle x, z\rangle$. This construction still works in the more general context, if $x \mapsto \langle Tx, y\rangle$ is defined on a *dense* subspace, namely $\mathrm{dom}(T)$, and is a *continuous* linear functional.

2.2. Definition: Let T be a densely defined operator on H. On the subspace

$$\mathrm{dom}(T^*) := \{y \in H \mid x \mapsto \langle Tx, y\rangle \text{ is continuous } \mathrm{dom}(T) \to \mathbb{C}\}$$

we define the *adjoint* T^* of T as follows: If $y \in \mathrm{dom}(T^*)$, then $x \mapsto \langle Tx, y\rangle$ has a unique continuous extension to a linear functional on H; by the Riesz-Fréchet theorem this functional can be uniquely represented in the form $x \mapsto \langle x, z\rangle$ with $z \in H$; we put $T^*y := z$.
By construction,

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad \forall x \in \mathrm{dom}(T), \forall y \in \mathrm{dom}(T^*).$$

The operator T is called *self-adjoint*, if $T^* = T$.

As noted above, the Hellinger-Toeplitz theorem guarantees consistency of this new notion of self-adjointness with the corresponding one for bounded operators. Obviously, self-adjoint operators are symmetric. Conversely, symmetric operators need not be self-adjoint (in general, they need not even be densely defined): In the above example, $\mathrm{dom}(T) \subsetneq \mathrm{dom}(S) \subseteq \mathrm{dom}(T^*)$, since $f \mapsto \langle if', g\rangle = \langle f, ig'\rangle$ is L^2-continuous on $\mathrm{dom}(T)$ for every $g \in \mathrm{dom}(S)$; therefore, T is symmetric, densely defined, but not self-adjoint.

A densely defined symmetric operator satisfies $\mathrm{dom}(T) \subseteq \mathrm{dom}(T^*)$, since $x \mapsto \langle Tx, y\rangle = \langle x, Ty\rangle$ is continuous $\mathrm{dom}(T) \to \mathbb{C}$ for every $y \in \mathrm{dom}(T)$, thus $T \subseteq T^*$ and T^* is densely defined as well; moreover, $T^{**} := (T^*)^*$ can be defined in this case. In general (without symmetry), a densely defined operator T may have non-dense $\mathrm{dom}(T^*)$, even examples with $\mathrm{dom}(T^*) = \{0\}$ can be found (see, e.g, [Wer18, page 371]).

2.3. Definition: Let T be an operator on H with *graph*

$$\mathrm{gr}(T) := \{(x, Tx) \in H \times H \mid x \in \mathrm{dom}(T)\}$$

and equip $H \times H$ with the inner product $\langle (x, y), (u, v)\rangle_2 := \langle x, u\rangle + \langle y, v\rangle$, defining the norm $\|(x, y)\|_2 = \sqrt{\|x\|^2 + \|y\|^2}$ (which induces the product topology on $H \times H$), thus obtaining a Hilbert space structure on $H \times H$. The operator T is *closed* (or *has closed graph*), if $\mathrm{gr}(T)$ is closed in $H \times H$. Equivalently, for any sequence (x_n) in $\mathrm{dom}(T)$ such that $x_n \to x$ and $Tx_n \to y$ in H, we have $x \in \mathrm{dom}(T)$ and $y = Tx$.

Recall that the closed graph theorem states that a closed operator T with $\mathrm{dom}(T) = H$ is continuous. In case of unbounded operators the closedness serves as a partial replacement of continuity properties. To give an example of an operator that is not closed we turn again to $Tx = ix'$ with $\mathrm{dom}(T) = \{x \in C^1([0,1]) \mid x(0) = 0 = x(1)\} \subseteq L^2([0,1])$: Consider $x_n(t) = (\frac{1}{4} + \frac{1}{n})^{1/2} - ((t - \frac{1}{2})^2 + \frac{1}{n})^{1/2}$, $x(t) = \frac{1}{2} - |t - \frac{1}{2}|$, and $y(t) = i\,\mathrm{sgn}(\frac{1}{2} - t)$; then $x_n \to x$ uniformly, hence in $L^2([0,1])$, and $Tx_n \to y$ in $L^2([0,1])$ (dominated convergence), but $x \notin \mathrm{dom}(T)$. (A pity, since $Tx = y$ almost everywhere; the chosen domain is too narrow).

2.4. Proposition: Let $T\colon \operatorname{dom}(T) \to H$ be densely defined, then:

(i) T^* is closed,

(ii) if T^* is densely defined, then $T \subseteq T^{**}$,

(iii) if T^* is densely defined and S is a closed extension of T, then $T^{**} \subseteq S$.
This means that T^{**} is the "smallest" closed extension of T, the so-called *closure* of T.

Proof: (i): Suppose $y_n \in \operatorname{dom}(T^*)$, $y_n \to y$, and $T^* y_n \to z$. Then we have for every $x \in \operatorname{dom}(T)$,

$$\langle Tx, y\rangle = \lim \langle Tx, y_n\rangle = \lim \langle x, T^* y_n\rangle = \langle x, z\rangle.$$

Therefore $x \mapsto \langle Tx, y\rangle = \langle x, z\rangle$ is continuous $\operatorname{dom}(T) \to \mathbb{C}$, hence $y \in \operatorname{dom}(T^*)$ and $z = T^* y$.

(ii): If $x \in \operatorname{dom}(T)$ and $y \in \operatorname{dom}(T^*)$, then $\langle Tx, y\rangle = \langle x, T^* y\rangle$. Hence $y \mapsto \langle T^* y, x\rangle = \langle y, Tx\rangle$ is continuous $\operatorname{dom}(T^*) \to \mathbb{C}$, hence $x \in \operatorname{dom}(T^{**})$ and $\langle Tx, y\rangle = \langle x, T^* y\rangle = \langle T^{**} x, y\rangle$ for every $y \in \operatorname{dom}(T^*)$. Since $\operatorname{dom}(T^*)$ is dense, we obtain $Tx = T^{**} x$ for every $x \in \operatorname{dom}(T)$, in summary, $T \subseteq T^{**}$.

(iii): The statement follows once we showed $\overline{\operatorname{gr}(T)} = \operatorname{gr}(T^{**})$. Here, $\overline{\operatorname{gr}(T)} \subseteq \operatorname{gr}(T^{**})$ holds, since $T^{**} = (T^*)^*$ has closed graph by (i) and $T \subseteq T^{**}$ by (ii). It remains to prove $\overline{\operatorname{gr}(T)} \supseteq \operatorname{gr}(T^{**})$, which follows, if we show $\operatorname{gr}(T)^\perp \subseteq \operatorname{gr}(T^{**})^\perp$.

Let $(u, v) \in \operatorname{gr}(T)^\perp$. For every $(x, Tx) \in \operatorname{gr}(T)$, we have $0 = \langle (x, Tx), (u, v)\rangle_2 = \langle x, u\rangle + \langle Tx, v\rangle$, hence $\langle Tx, v\rangle = -\langle x, u\rangle$. This shows that $v \in \operatorname{dom}(T^*)$ and $T^* v = -u$.

Let $(z, T^{**} z) \in \operatorname{gr}(T^{**})$ arbitrary, then

$$\langle (z, T^{**} z), (u, v)\rangle_2 = \langle z, u\rangle + \langle T^{**} z, v\rangle = \langle z, u\rangle + \langle z, T^* v\rangle = \langle z, \underbrace{u + T^* v}_{u-u}\rangle = 0,$$

thus, $(u, v) \in \operatorname{gr}(T^{**})^\perp$. □

2.5. Corollary: Let $T\colon \operatorname{dom}(T) \to H$ be densely defined, then:

(i) T is symmetric, if and only if $T \subseteq T^*$.
In this case, $T \subseteq T^{**} \subseteq T^* = T^{***}$ and also T^{**} is symmetric.

(ii) T is closed and symmetric, if and only if $T = T^{**} \subseteq T^*$.

(iii) T is self-adjoint, if and only if $T = T^{**} = T^*$.

Proof: (i): That symmetry of T implies $T \subseteq T^*$ has been noted already in the paragraph preceding Definition 2.3; the converse is clear. By Proposition 2.4, T^* is closed and T^{**} is the closure of T, therefore $T \subseteq T^*$ implies $T \subseteq T^{**} \subseteq T^*$. Since T^* is densely defined as well, the closure of T^* is T^{***}; since T^* is closed, we have $T^* = T^{***}$. In particular, we obtained $T^{**} \subseteq T^{***}$, hence T^{**} is symmetric.

(ii): If $T = T^{**} \subseteq T^*$, then T is symmetric by (i) and closed by Proposition 2.4. [$T = (T^*)^*$]
If T is closed and symmetric, then $T = T^{**}$ by Proposition 2.4 and $T \subseteq T^*$ by (i).

(iii): Self-adjointness is defined by the equality $T = T^*$, which implies that T is closed by Proposition 2.4, hence $T = T^{**}$. □

There is an interesting intermediate case between self-adjointness and symmetry for densely defined operators, namely self-adjointness of the adjoint.

2.6. Definition: A densely defined symmetric operator T on H is *essentially self-adjoint*, if its closure T^{**} is self-adjoint, or, equivalently, $T \subseteq T^{**} = T^*$ $_{(= T^{***} \text{ by the above corollary})}$.

In general, self-adjoint extensions of symmetric operators need not exist, but if T is essentially self-adjoint it possesses its closure $T^{**} = T^*$ as the unique self-adjoint extension. (This is easily proved upon observing that $T \subseteq S$ implies $S^* \subseteq T^*$.)

2.7. An example revisited: We study again the example $Tx = ix'$ with $\mathrm{dom}(T) = \{x \in C^1([0,1]) \mid x(0) = 0 = x(1)\} \subseteq L^2([0,1])$. We have seen above that T is symmetric, but not self-adjoint, and not even closed.

We claim that $\mathrm{dom}(T^*) = \{x\colon [0,1] \to \mathbb{C} \mid x \text{ is absolutely continuous and } x' \in L^2([0,1])\}$ and $T^*x = ix'$.

Let $x \in \mathrm{dom}(T^*)$ and $y := T^*x \in L^2([0,1])$. We put $F(t) := \int_0^t y(s)\,ds$ and note that, by the fundamental theorem of calculus, F is absolutely continuous[2] and satisfies $F' = y$ almost everywhere. We may use the formula of integration by parts (Proposition 0.20) and obtain for any $z \in \mathrm{dom}(T)$ (recall that $z(0) = 0 = z(1)$)

$$\langle Tz, x\rangle = \langle z, T^*x\rangle = \langle z, y\rangle = \langle z, F'\rangle = \int_0^1 z(t)\overline{F'(t)}\,dt = z(t)\overline{F(t)}\Big|_{t=0}^{t=1} - \int_0^1 z'(t)\overline{F(t)}\,dt$$
$$= -\langle z', F\rangle = \langle iz', -iF\rangle = \langle Tz, -iF\rangle,$$

i.e., $\langle Tz, x + iF\rangle = 0$, and therefore

$$(*) \qquad x + iF \in \mathrm{ran}(T)^\perp.$$

Observe that $Tz \in C([0,1])$ with $\int_0^1 Tz = i\int_0^1 z' = i(z(1) - z(0)) = 0$ for every $z \in \mathrm{dom}(T)$; on the other hand, any continuous function w with $\int_0^1 w = 0$ is in $\mathrm{ran}(T)$, since $w = iW'$ with $W \in \mathrm{dom}(T)$, if $W(t) := -i\int_0^t w(s)\,ds$ (which yields $W(0) = W(1) = 0$). Thus,

$$\mathrm{ran}(T)^\perp = \{w \in C([0,1]) \mid \int_0^1 w = 0\}^\perp = \{w \in C([0,1]) \mid \langle w, 1\rangle = 0\}^\perp$$
$$= \left(\overline{\{w \in C([0,1]) \mid \langle w, 1\rangle = 0\}}\right)^\perp = \{w \in L^2([0,1]) \mid \langle w, 1\rangle = 0\}^\perp = (\{1\}^\perp)^\perp = \mathrm{span}\{1\}.$$

We conclude from $(*)$ that $x + iF \in \mathrm{span}\{1\}$, hence $x = -iF + \alpha 1$ with some $\alpha \in \mathbb{C}$, which shows that x is absolutely continuous and $ix' = F' + 0 = y = T^*x \in L^2([0,1])$.

If $x\colon [0,1] \to \mathbb{C}$ is absolutely continuous with $x' \in L^2([0,1])$, then we obtain for every $z \in \mathrm{dom}(T)$ (again employing integration by parts)

$$\langle Tz, x\rangle = \langle iz', x\rangle = \langle z, ix'\rangle,$$

[2] $y \in L^1([0,1])$, since $\int_0^1 |y|\,ds = \int_0^1 1 \cdot |y|\,ds \le \|1\|_2 \cdot \|y\|_2 = \|y\|_2$

which proves that $z \mapsto \langle Tz, x\rangle$ is continuous, thus $x \in \operatorname{dom}(T^*)$.

As an exercise, one can show along similar lines that $\operatorname{dom}(T^{**}) = \{x \in \operatorname{dom}(T^*) \mid x(0) = 0 = x(1)\} \subsetneq \operatorname{dom}(T^*)$ and $T^{**}x = ix'$, i.e, $T^{**} \subsetneq T^*$. We conclude that T is not essentially self-adjoint (and T^* is not symmetric).

2.8. Example (unbounded multiplication operators): Let (Ω, Σ, μ) be a measure space, $f\colon \Omega \to \mathbb{R}$ measurable, $\operatorname{dom}(T) := \{x \in L^2(\Omega,\mu) \mid f \cdot x \in L^2(\Omega,\mu)\}$, and $Tx := f \cdot x$. Since f is real-valued, T is symmetric.

We show that $\operatorname{dom}(T)$ is dense: For $n \in \mathbb{N}$ consider $\Omega_n := \{\omega \in \Omega \mid |f(\omega)| \le n\} \in \Sigma$; then $\bigcup_{n\in\mathbb{N}} \Omega_n = \Omega$ and $V_n := \{x \in L^2(\Omega,\mu) \mid \forall \omega \in \Omega \setminus \Omega_n\colon x(\omega) = 0\} \subseteq \operatorname{dom}(T)$ $(n \in \mathbb{N})$; if $x \in L^2(\Omega,\mu)$, then $x_n := x\,\chi_{\Omega_n} \in V_n \subseteq \operatorname{dom}(T)$ and $x_n \to x$ in $L^2(\Omega,\mu)$ (by dominated convergence).

So far, we have seen that T is a densely defined symmetric operator, hence $T \subseteq T^*$. We claim that T is self-adjoint. It suffices to show that $\operatorname{dom}(T^*) \subseteq \operatorname{dom}(T)$, since symmetry then implies $T^* \subseteq T$, which yields $T^* = T$.

Let $x \in \operatorname{dom}(T^*)$. We note that $\chi_{\Omega_n} f \in L^\infty(\Omega,\mu)$ and that $\chi_{\Omega_n} z \in \operatorname{dom}(T)$ for every $z \in \operatorname{dom}(T)$, hence we have for every $n \in \mathbb{N}$,

$$\langle z, \chi_{\Omega_n} T^*x\rangle = \langle \chi_{\Omega_n} z, T^*x\rangle = \langle T(\chi_{\Omega_n} z), x\rangle = \langle f\chi_{\Omega_n} z, x\rangle = \langle z, \chi_{\Omega_n} f x\rangle.$$

By density of $\operatorname{dom}(T)$, we deduce that $\chi_{\Omega_n} T^*x = \chi_{\Omega_n} f x$ holds in $L^2(\Omega,\mu)$ for every $n \in \mathbb{N}$. Since $\chi_{\Omega_n} \to 1$ pointwise, we obtain $T^*x = fx$ as measurable functions almost everywhere on Ω, hence $fx \in L^2(\Omega,\mu)$ (because T^*x belongs to $L^2(\Omega,\mu)$), which proves $x \in \operatorname{dom}(T)$.

Operators of the form $\lambda - T$, with $\lambda \in \mathbb{C}$, can always be defined on $\operatorname{dom}(\lambda - T) := \operatorname{dom}(T)$. We will make use of this convention from now on.

2.9. Lemma: Let T be a densely defined operator on H:

(i) $\ker(T^* \mp i) = \operatorname{ran}(T \pm i)^\perp$, in particular,

$$\ker(T^* \mp i) = \{0\} \quad \Leftrightarrow \quad \operatorname{ran}(T \pm i) \text{ dense in } H.$$

(Moreover, being orthogonal complements, the subspaces $\ker(T^* \mp i)$ are closed.)

(ii) If T is closed and symmetric, then $\operatorname{ran}(T \pm i)$ is closed in H.

Proof: (i): Clearly, $(T \pm i)^* = T^* \mp i$, thus we have

$$\begin{aligned} y \in \operatorname{ran}(T \pm i)^\perp \quad &\Leftrightarrow \quad \forall z \in \operatorname{dom}(T)\colon \langle (T \pm i)z, y\rangle = 0 \quad \Leftrightarrow \\ &y \in \operatorname{dom}(T^*) \text{ and } \forall z \in \operatorname{dom}(T)\colon \langle z, (T^* \mp i)y\rangle = 0 \quad \Leftrightarrow \quad y \in \ker(T^* \mp i). \end{aligned}$$

(ii): By symmetry of T, $\langle Tx, x\rangle \in \mathbb{R}$ for every $x \in \operatorname{dom}(T)$, thus we have

$$\|(T \pm i)x\|^2 = \|Tx\|^2 + \|x\|^2 \pm 2\operatorname{Re} \underbrace{\langle Tx, ix\rangle}_{-i\langle Tx,x\rangle \,\in\, i\mathbb{R}} = \|Tx\|^2 + \|x\|^2 \ge \|x\|^2. \tag{2.1}$$

Therefore $(T \pm i)^{-1}\colon \operatorname{ran}(T \pm i) \to \operatorname{dom}(T)$ exists and is continuous.

To show closedness of $\operatorname{ran}(T \pm i)$, suppose (x_n) is a sequence in $\operatorname{dom}(T)$ such that the sequence $((T \pm i)x_n)$ in $\operatorname{ran}(T \pm i)$ converges to $y \in H$. By continuity of $(T \pm i)^{-1}$, the Cauchy sequence $((T \pm i)x_n)$ in $\operatorname{ran}(T \pm i)$ is then mapped to the Cauchy sequence (x_n) in $\operatorname{dom}(T)$. Hence $x := \lim x_n$ exists in H and $Tx_n \to y \mp ix$. Since T is a closed operator, $x \in \operatorname{dom}(T)$ and $y \mp ix = Tx$, i.e., $y = (T \pm i)x \in \operatorname{ran}(T \pm i)$. □

2.10. Theorem: Let T be a symmetric and densely defined operator on H, then the following are equivalent:

(i) T is self-adjoint,

(ii) T is closed and $\ker(T^* \pm i) = \{0\}$,

(iii) $\operatorname{ran}(T \pm i) = H$.

Proof: (i) ⇒ (ii): $T = T^*$ is closed by Proposition 2.4. By symmetry of T^* and Equation (2.1) in the proof of Lemma 2.9 (applied to T^* in place of T), we find that $(T^* \pm i)x = 0$ implies $x = 0$.

(ii) ⇒ (iii): By the above lemma, $\operatorname{ran}(T \pm i)^\perp = \ker(T^* \mp i) = \{0\}$ and $\operatorname{ran}(T \pm i)$ is closed, since T is closed. Thus, we obtain $\operatorname{ran}(T \pm i) = H$.

(iii) ⇒ (i): The densely defined symmetric operator T satisfies $T \subseteq T^*$, hence it suffices to show $\operatorname{dom}(T^*) \subseteq \operatorname{dom}(T)$. Let $y \in \operatorname{dom}(T^*)$. Since $H = \operatorname{ran}(T - i)$ we can find some $x \in \operatorname{dom}(T)$ such that $(T^* - i)y = (T - i)x$. Due to $T \subseteq T^*$ we may thus write $(T^* - i)y = (T^* - i)x$. The previous lemma gives $\ker(T^* - i) = \operatorname{ran}(T + i)^\perp = H^\perp = \{0\}$, which implies that $T^* - i$ is injective, hence $y = x \in \operatorname{dom}(T)$. □

The theorem provides us with a very short argument that the operator T of multiplication by the real-valued measurable function f from Example 2.8 is self-adjoint: The functions $1/(f \pm i)$ and $f/(f \pm i)$ are measurable and bounded, hence any $h \in L^2(\Omega, \mu)$ is in the range of $(T \pm i)$, since $(f \pm i) \cdot \frac{h}{f \pm i} = h$ is clear and $f \cdot \frac{h}{f \pm i} = \frac{f}{f \pm i} \cdot h \in L^2(\Omega, \mu)$ shows that $\frac{h}{f \pm i} \in \operatorname{dom}(T)$.

2.11. Corollary: Let T be a symmetric and densely defined operator on H, then the following are equivalent:

(i) T is essentially self-adjoint,

(ii) $\ker(T^* \pm i) = \{0\}$,

(iii) $\operatorname{ran}(T \pm i)$ is dense in H.

Proof: (ii) ⇔ (iii): Follows directly from Lemma 2.9(i).

(i) ⇔ (ii): Recall that we have $T \subseteq T^{**} \subseteq T^* = T^{***}$ by Corollary 2.5(i). The above theorem, applied to T^{**} (which is closed), states that T^{**} is self-adjoint, if and only if $\ker(T^* \pm i) = \ker(T^{***} \pm i) = \{0\}$. □

For a densely defined symmetric operator T, the Hilbert dimensions of $\ker(T^* \pm i)$ are called *deficiency indices* of T. The corollary says that T is essentially self-adjoint, if and only if both deficiency indices are 0. More generally, we will show in the following theorem, that self-adjoint extensions exist, if and only if the deficiency indices are equal. Recall that $\ker(T^* \pm i) = \operatorname{ran}(T \mp i)^\perp$ by Lemma 2.9(i).

2.12. Theorem: A symmetric and densely defined operator T on H possesses self-adjoint extensions, if and only if its deficiency indices are equal, which means that $\dim\ker(T^* + i) = \dim\ker(T^* - i)$, where dim refers to the cardinality of a complete orthonormal system.

Remark on the following proof: The basic idea will be to transfer some of the properties of the bijective map $t \mapsto \frac{t+i}{t-i}$ between $\mathbb{R}$ and $\{z \in \mathbb{C} \mid |z| = 1, z \neq 1\}$, with inverse $u \mapsto i\frac{u+1}{u-1}$, to the level of operators via the so-called *Cayley-transform* $U := (T+i)(T-i)^{-1}$ of T.

Proof: Suppose that S is a self-adjoint extension of T. Recall from (2.1) that $T - i$ is injective, hence $U := (T + i)(T - i)^{-1}$ can be defined on $\mathrm{dom}(U) := \mathrm{ran}(T - i)$. By Theorem 2.10, $\mathrm{ran}(S - i) = H$ and we may define $V := (S + i)(S - i)^{-1}$ on $\mathrm{dom}(V) := H$. Equation (2.1) and Theorem 2.10 applied to S show that V is isometric and surjective, thus unitary. Since $T \subseteq S$, V maps $\mathrm{ran}(T - i)$ onto $\mathrm{ran}(T + i)$; moreover, as unitary operator, V also maps the respective orthogonal complements onto one another. Therefore, Lemma 2.9(i) implies that V maps $\ker(T^* + i) = \mathrm{ran}(T - i)^\perp$ unitarily onto $\ker(T^* - i) = \mathrm{ran}(T + i)^\perp$, hence the deficiency indices have to be equal.

In proving the reverse implication, start by supposing $\dim\ker(T^* + i) = \dim\ker(T^* - i)$. We learn from (2.1) that

$$\|(T + i)x\| = \|(T - i)x\| \quad \forall x \in \mathrm{dom}(T),$$

therefore, $U\colon \mathrm{ran}(T - i) \to \mathrm{ran}(T + i)$, $(T - i)x \mapsto (T + i)x$ is a well-defined isometric (hence injective) and surjective operator, which can be uniquely extended to an isometry between the corresponding closures of the subspaces. Appealing to Lemma 2.9(i), we obtain $\dim\mathrm{ran}(T - i)^\perp = \dim\ker(T^* + i) = \dim\ker(T^* - i) = \dim\mathrm{ran}(T + i)^\perp$, which allows us to extend U to a unitary operator $V\colon H \to H$ by simply mapping a complete orthonormal system of $\mathrm{ran}(T - i)^\perp$ into such for $\mathrm{ran}(T + i)^\perp$. We proceed in four steps:

1. $V - I$ is injective: If $y \in H$ with $(V - I)y = 0$, then also $(V^* - I)y = 0$, since $V - I$ is normal[3]. We therefore have for every $x \in \mathrm{dom}(T)$,

$$\begin{aligned} 2i\langle x, y\rangle = \langle (T + i)x - (T - i)x, y\rangle &= \langle V(T - i)x - (T - i)x, y\rangle \\ &= \langle (V - I)(T - i)x, y\rangle = \langle (T - i)x, (V^* - I)y\rangle = 0, \end{aligned}$$

i.e., $y \in \mathrm{dom}(T)^\perp = \{0\}$.

2. Construction of an extension S of T: Let $\mathrm{dom}(S) := \mathrm{ran}(V - I)$ and use the injectivity of $V - I$ to see that $(V - I)z \mapsto i(Vz + z)$ gives a well-defined linear map $S\colon \mathrm{ran}(V - I) \to H$. If $x \in \mathrm{dom}(T)$, then $(V - I)(T - i)x = V(T - i)x - (T - i)x = (T + i)x - (T - i)x = 2ix$, which shows that $x \in \mathrm{ran}(V - I) = \mathrm{dom}(S)$ and, moreover, that

$$\begin{aligned} Sx = S\left(\frac{1}{2i}(V - I)(T - i)x\right) = \frac{1}{2i}i(V + I)(T - i)x &= \frac{1}{2}\big(V(T - i)x + (T - i)x\big) \\ &= \frac{1}{2}\big((T + i)x + (T - i)x\big) = \frac{1}{2}(2Tx) = Tx. \end{aligned}$$

[3]Again by [Wer18, Lemma V.5.10], $\|Rx\| = \|R^*x\|$, if $R \in L(H)$ is normal. (Recall once more the quick proof: $0 = \langle (R^*R - RR^*)x, x\rangle = \langle R^*Rx, x\rangle - \langle RR^*x, x\rangle = \|Rx\|^2 - \|R^*x\|^2$.)

3. S is symmetric: If $x \in \operatorname{dom}(S) = \operatorname{ran}(V - I)$, say $x = (V - I)y$ for some $y \in H$, then

$$\begin{aligned}\langle Sx, x\rangle = \langle i(V+I)y, (V-I)y\rangle &= i\left(\langle Vy, Vy\rangle + \langle y, Vy\rangle - \langle Vy, y\rangle - \langle y, y\rangle\right)\\ &= i\left(\langle y, Vy\rangle - \langle Vy, y\rangle\right) = i2i\,\operatorname{Im}\langle y, Vy\rangle = -2\operatorname{Im}\langle y, Vy\rangle\end{aligned}$$

is a real number, in particular $\langle Sx, x\rangle = \overline{\langle Sx, x\rangle} = \langle x, Sx\rangle$. As an exercise, one can show that this implies symmetry of S (e.g., by comparing the expanded expressions for $\langle S(x_1 + x_2), x_1 + x_2\rangle = \langle x_1 + x_2, S(x_1 + x_2)\rangle$ with arbitrary $x_1, x_2 \in \operatorname{dom}(S)$).

4. S is self-adjoint: We show that $\operatorname{ran}(S \pm i) = H$, then the proof is complete by Theorem 2.10. For every $x \in H$, we have $(S - i)(V - I)x = i(V + I)x - i(V - I)x = 2ix$ and $(S + i)(V - I)x = i(V + I)x + i(V - I)x = 2iVx$, thus,

$$x = (S - i)\left(\frac{1}{2i}(V - I)x\right) = (S + i)\left(\frac{1}{2i}(V - I)V^*x\right),$$

hence $x \in \operatorname{ran}(S - i)$ and $x \in \operatorname{ran}(S + i)$. □

2.13. Examples: 1) Let $\mathbb{R}^+ :=]0, \infty[$ and consider $Tx = ix'$ on $\operatorname{dom}(T) = \mathscr{D}(\mathbb{R}^+) := \{h \in C^\infty(\mathbb{R}^+) \mid \operatorname{supp}(h) \text{ compact}\} \subseteq L^2(\mathbb{R}^+)$. It is shown, e.g., in [Wer18, Lemma V.1.10], hat $\mathscr{D}(\mathbb{R}^+)$ is dense. By techniques very similar to those applied in Example 2.7 above, one can show that

$$\begin{aligned}\operatorname{dom}(T^*) = \{x \in L^2(\mathbb{R}^+) \mid \text{the restriction } x|_I \text{ is absolutely continuous}\\ \text{for every compact interval } I \subset \mathbb{R}^+ \text{ and } x' \in L^2(\mathbb{R}^+)\}\end{aligned}$$

and $T^*x = ix'$. Therefore, T is symmetric and densely defined.

Determining $\ker(T^* \pm i)$ means solving the ordinary differential equation $y' \pm y = 0$ for $y \in \operatorname{dom}(T^*)$. The solutions are automatically in $C^1(\mathbb{R}^+)$ and have to be of the form $\alpha e^{\mp t}$ with $\alpha \in \mathbb{C}$. The requirement $y \in L^2(\mathbb{R}^+)$ reduces the solution spaces to

$$\ker(T^* + i) = \operatorname{span}\{e^{-t}\}, \quad \ker(T^* - i) = \{0\},$$

and shows that the deficiency indices are different. By Corollary 2.11, T is not essentially self-adjoint and Theorem 2.12 shows that there is no self-adjoint extension of T.

2) In Example 2.8 put $\Omega = \mathbb{R}^n$, $\Sigma = \mathcal{B}(\mathbb{R}^n)$, μ the Lebesgue measure, and $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(\xi) = |\xi|^2$. We immediately obtain that the operator M of multiplication by f with domain $\operatorname{dom}(M) = \{\psi \in L^2(\mathbb{R}^n) \mid \xi \mapsto |\xi|^2\,\psi(\xi) \text{ belongs to } L^2(\mathbb{R}^n)\}$ is self-adjoint.

Consider $M_0\,\varphi := f \cdot \varphi$ with dense domain

$$\begin{aligned}(2.2)\quad \mathscr{S}(\mathbb{R}^n) := \{\varphi \in C^\infty(\mathbb{R}^n) \mid \forall \alpha, \beta \in \mathbb{N}^n\colon \xi \mapsto \xi^\alpha \partial^\beta \varphi(\xi) \text{ is bounded}\}\\ \supseteq \mathscr{D}(\mathbb{R}^n) := \{h \in C^\infty(\mathbb{R}^n) \mid \operatorname{supp}(h) \text{ compact}\},\end{aligned}$$

then M_0 is symmetric, densely defined, and $M_0 \subseteq M$. Obviously, $\operatorname{ran}(M_0 \pm i) \supseteq \mathscr{D}(\mathbb{R}^n)$ shows that these ranges are dense (since $\mathscr{D}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, again by [Wer18, Lemma V.1.10]), hence Corollary 2.11 implies that M_0 is essentially self-adjoint. Its unique self-adjoint extension and closure is given by M, i.e., $M = M_0^{**}$ (refer to Proposition 2.4 and directly show $\operatorname{gr}(M) \subseteq \overline{\operatorname{gr}(M_0)}$.)

Well-known properties of the Fourier transform imply

$$(\mathcal{F}\Delta h)(\xi) = \sum_{j=1}^{n} \mathcal{F}(\partial_j^2 h)(\xi) = \sum_{j=1}^{n} (i\xi_j)^2 \widehat{h}(\xi) = -\sum_{j=1}^{n} \xi_j^2\, \widehat{h}(\xi) = -|\xi|^2\, \widehat{h}(\xi) = -(M\widehat{h})(\xi),$$

thus M is unitarily equivalent to the operator $-\Delta$ on the domain $H^2(\mathbb{R}^n) := \mathcal{F}^{-1}\mathrm{dom}(M)$, which can be identified with the *Sobolev space* of order 2 (i.e., the L^2-functions with weak derivatives up to order 2 belonging to L^2). We deduce that $-\Delta$ with domain $H^2(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n)$ is self-adjoint.

2.14. Revisiting the example revisited in Example 2.7: We study the self-adjoint extensions of $Tx = ix'$ with $\mathrm{dom}(T) = \{x \in C^1([0,1]) \mid x(0) = 0 = x(1)\} \subseteq L^2([0,1])$.

The *deficiency spaces* $\ker(T^* \pm i)$ are easily determined by solving $y' \pm y = 0$ for $y \in \mathrm{dom}(T^*)$. Both basic solutions $t \mapsto e^{\mp t} =: e_{\mp}(t)$ belong to $\mathrm{dom}(T^*)$, which yields

$$\mathrm{ran}(T-i)^{\perp} = \ker(T^* + i) = \mathrm{span}\{e_-\} \quad \text{and} \quad \mathrm{ran}(T+i)^{\perp} = \ker(T^* - i) = \mathrm{span}\{e_+\}.$$

The operator U constructed in the proof of Theorem 2.12 maps $\mathrm{ran}(T-i) \to \mathrm{ran}(T+i)$, $i(x'-x) \mapsto i(x'+x)$ for any $x \in \mathrm{dom}(T)$. It is extended to a unitary operator V on $L^2([0,1])$ by assigning a fixed normalized vector in $\mathrm{ran}(T-i)^{\perp}$, say $\widetilde{e_-} := ec_0 e_-$ with $c_0 := \sqrt{2/(e^2-1)}$, to a normalized vector in $\mathrm{ran}(T+i)^{\perp}$, say $V(\widetilde{e_-}) := e^{i\gamma} \cdot \widetilde{e_+}$ for some $\gamma \in \mathbb{R}$, where $\widetilde{e_+} := c_0 e_+$. Every $\gamma \in [0, 2\pi[$ gives a different extension V and these are all possible unitary extensions of U. The corresponding self-adjoint extension $S \supset T$ is then given on $\mathrm{dom}(S) = \mathrm{ran}(V - I)$ by $(V-I)z \mapsto i(V+I)z$.

Since T^{**} is the closure of T, the self-adjoint extensions of T and T^{**} agree. The latter can be shown to be described by the family of operators S_α ($\alpha \in \mathbb{C}$, $|\alpha| = 1$) with $S_\alpha x = ix'$ and $\mathrm{dom}(S_\alpha) = \{x \colon [0,1] \to \mathbb{C} \mid x \text{ is absolutely continuous and } x' \in L^2([0,1]), x(0) = \alpha x(1)\}$ (see, e.g., [Sch00, 5.3, Beispiel 3 and 5.1, Beispiel 4, Beispiel 5] or [RS75, Section X.1, Example 1], or also [Con10, Chapter X, Examples 2.21 and 1.11]).

2.15. Extension of the notion of a spectrum: In the most general case, the definition of the spectrum of an operator which is unbounded or/and not defined on the whole Hilbert space is not completely uniform in the literature. Thus, let us illustrate two technical issues before giving a definition.

Artefacts with non-dense domains: Let $H = l^2(\mathbb{N})$ and consider the left-shift, given by $(Lx)_n = x_{n+1}$ for $x = (x_n) \in l^2(\mathbb{N})$. The bounded operator $L \in L(H)$ clearly has $\ker(L) = \mathrm{span}\{e_1\}$, in particular $0 \in \sigma(L)$. If $\mathrm{dom}(L_0) := \{e_1\}^{\perp}$ and L_0 is L restricted to $\mathrm{dom}(L_0)$, then we have the bounded inverse $R_0 \colon l^2(\mathbb{N}) \to \mathrm{dom}(L_0)$ given by the right-shift. Hence we could not consider 0 to be a spectral value of L_0. However, if $\mathrm{dom}(L_1)$ is any dense subspace of $l^2(\mathbb{N})$ and L_1 is the corresponding restriction of L, then any sequence from $\mathrm{dom}(L_1)$ approximating e_1 can be used to show that L_1 is not continuously invertible, hence 0 could be considered a spectral value of L_1. This indicates why we will define the spectrum only for densely defined operators.

Non-closed operators have empty resolvent sets: Let $T \colon \mathrm{dom}(T) \to H$ be a linear operator (not necessarily densely defined) and suppose that $\lambda \in \mathbb{C}$ is such that $\lambda - T$ is bijective $\mathrm{dom}(T) \to H$ with bounded inverse. We will show that T has to be closed.

We first show that $\lambda - T$ is closed: Let (x_n) be a sequence in $\mathrm{dom}(T) = \mathrm{dom}(\lambda - T)$ such that

$x_n \to x$ in H and $(\lambda - T)x_n \to y$ in H. Then $z_n := (\lambda - T)x_n$ belongs to $(\lambda - T)(\mathrm{dom}(T))$ and $z_n \to y$, hence continuity of $(\lambda - T)^{-1}$ implies

$$x_n = (\lambda - T)^{-1} z_n \to (\lambda - T)^{-1} y \quad \text{in addition to} \quad x_n \to x,$$

therefore, $x = (\lambda - T)^{-1} y \in \mathrm{dom}(T)$ and $(\lambda - T)x = y$.

Now we find that also T is closed: Let (x_n) be a sequence in $\mathrm{dom}(T)$ such that $x_n \to x$ in H and $Tx_n \to z$ in H. Then we have

$$(\lambda - T)x_n = \lambda x_n - Tx_n \to \lambda x - z.$$

Since $\lambda - T$ is closed, this implies $x \in \mathrm{dom}(\lambda - T) = \mathrm{dom}(T)$ and $\lambda x - z = (\lambda - T)x$, i.e., $Tx = z$.

This fact is the reason why some authors prefer to define the spectrum only in case of closed operators, since non-closed operators automatically have spectrum equal to $\mathbb{C}$. However, there are also closed unbounded densely defined operators with spectrum $\mathbb{C}$, i.e., empty resolvent set. (An explicit example is given by the multiplication operator $(Tf)(z) = zf(z)$ on $\mathrm{dom}(T) := \{f \in L^2(\mathbb{C}) \mid z \mapsto zf(z) \in L^2(\mathbb{C})\} \subset L^2(\mathbb{C})$; details of arguments, e.g., that T is closed etc., can be easily supplied with the help of [Wei00, Abschnitt 6.1].)

Definition: Let $T\colon \mathrm{dom}(T) \to H$ be a densely defined linear operator, then

(i) $\rho(T) := \{\lambda \in \mathbb{C} \mid \lambda - T \text{ is bijective } \mathrm{dom}(T) \to H \text{ with bounded inverse}\}$ is called the *resolvent set* of T,

(ii) $R\colon \rho(T) \to L(H)$, $\lambda \mapsto R_\lambda := (\lambda - T)^{-1}$, is the *resolvent map* of T,

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

Remark: (i) For any closed subset $S \subseteq \mathbb{C}$ there is some closed densely defined operator T on a separable Hilbert space H such that $\sigma(T) = S$ (see [Wei00, Beispiel 5.10]).

(ii) If T is closed, then it suffices to check whether $\lambda - T$ is bijective to conclude $\lambda \in \rho(T)$ (this follows from [Wer18, Satz IV.4.4]).

(iii) By the observation made above, $\sigma(T) = \mathbb{C}$ for every non-closed densely defined operator T. In particular, $\sigma(T)$ need not be compact for unbounded operators. Neither is it guaranteed that the spectrum of an unbounded operator is non-empty. (We will give an example below.)

Proposition: Let $T\colon \mathrm{dom}(T) \to H$ be a densely defined linear operator, then

(i) $\rho(T)$ is open,

(ii) the resolvent map $\lambda \mapsto R_\lambda$ is analytic $\rho(T) \to L(H)$ and we have

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu,$$

(Note that, in particular, R_λ and R_μ commute.)

(iii) $\sigma(T)$ is closed.

Proof: Clearly, (iii) follows from (i) due to $\sigma(T) = \mathbb{C} \setminus \rho(T)$. The proofs of (i) and (ii) are very similar to the case of bounded T, since, by definition, the resolvent $R_\lambda = (\lambda - T)^{-1}$ belongs to $L(H)$ for every $\lambda \in \rho(T)$ and maps H into $\operatorname{dom}(T)$: In detail, we obtain the resolvent equation in (ii) from

$$R_\lambda = R_\lambda(\mu - T)R_\mu = R_\lambda(\mu - \lambda + \lambda - T)R_\mu = (\mu - \lambda)R_\lambda R_\mu + R_\mu.$$

As for (i), suppose $\lambda \in \rho(T)$ and $\mu \in \mathbb{C}$ satisfies $|\lambda - \mu| < 1/\|R_\lambda\|$. Then $\|(\lambda-\mu)R_\lambda\| < 1$ and hence $1 - (\lambda - \mu)R_\lambda$ is invertible with inverse given by the Neumann series, thus as a power series in the variable μ with coefficients in $L(H)$. The resolvent equation would suggest $R_\lambda = R_\mu(1-(\lambda-\mu)R_\lambda)$, and indeed $S := R_\lambda(1 - (\lambda - \mu)R_\lambda)^{-1} \in L(H)$ is a power series in the variable μ such that $(\mu - T)Sx = S(\mu - T)x = x$ holds for all x in the dense subspace $\operatorname{dom}(T) \subseteq H$ (note that R_λ commutes with $(1 - (\lambda - \mu)R_\lambda)^{-1}$ use $\mu - T = \mu - \lambda + \lambda - T$). Therefore $\mu \in \rho(T)$, thus $\rho(T)$ is open. Obviously, we have also proven the claim of analyticity in (ii) along the way. □

Examples: 1) On $H = L^2([0, 1])$ consider $Tx = ix'$ with

$$\operatorname{dom}(T) = \{x \in C([0, 1]) \mid x \text{ is absolutely continuous}, x' \in L^2([0, 1]), x(0) = 0\}.$$

The domain is dense and T is a closed operator (without proof, but recommended as an exercise). We claim that $\sigma(T) = \emptyset$.

Let $\lambda \in \mathbb{C}$ and $x \in L^2([0, 1])$ be arbitrary and consider the equation $(\lambda - T)y = x$, which translates into the ordinary differential equation $y' + i\lambda y = ix$. The solutions to the homogeneous equation are complex multiples of the function $t \mapsto \exp(-i\lambda t)$ and the usual "variation of constant" and adaptation to the "initial condition" $y(0) = 0$ yields the explicit solution formula

$$y(t) = i \int_0^t e^{i\lambda(s-t)} x(s)\, ds =: (S_\lambda x)(t) \quad (t \in [0, 1]).$$

We see that indeed this defines a solution $y \in \operatorname{dom}(T)$ and thus $S_\lambda \colon L^2([0, 1]) \to \operatorname{dom}(T)$ is the inverse of $\lambda - T$. Closedness of T (hence of $\lambda - T$) implies boundedness of S_λ (again by [Wer18, Satz IV.4.4]), but alternatively it is not difficult to see directly that $\|S_\lambda x\|_2 \le e^{|\operatorname{Im}(\lambda)|}\|x\|_2$. Thus, S_λ is the resolvent R_λ and $\lambda \in \rho(T)$.

2) Let (Ω, Σ, μ) be a σ-finite[4] measure space and $f \colon \Omega \to \mathbb{R}$ be measurable. Consider the multiplication operator $Tx = fx$ on $L^2(\Omega, \mu)$ with domain $\operatorname{dom}(T) = \{x \in L^2(\Omega, \mu) \mid fx \in L^2(\Omega, \mu)\}$. By Example 2.8, T is self-adjoint.

We claim that $\sigma(T) = \{\lambda \in \mathbb{R} \mid \forall \varepsilon > 0 \colon \mu\big(f^{-1}([\lambda - \varepsilon, \lambda + \varepsilon])\big) > 0\} =:$ essential range of f.

If $\lambda \in \mathbb{C}$ is not in the essential range of f, then $1/(\lambda - f) \in L^\infty(\Omega, \mu)$ and the corresponding multiplication operator is bounded; moreover, it is an inverse of $\lambda - T$, since $\frac{fy}{\lambda - f} = -y + \frac{\lambda}{\lambda - f} y \in L^2(\Omega, \mu) + L^\infty(\Omega, \mu) \cdot L^2(\Omega, \mu) \subseteq L^2(\Omega, \mu)$ for every $y \in L^2(\Omega, \mu)$. Therefore, $\lambda \in \rho(T)$.

If $\lambda \in \rho(T)$, then we may argue pointwise μ-almost everywhere to show that the resolvent $R_\lambda = (\lambda - T)^{-1} \colon L^2(\Omega, \mu) \to \operatorname{dom}(T)$ has to be given as operator of multiplication by the function

[4] i.e., $\exists \Omega_n \in \Sigma$ $(n \in \mathbb{N})$: $\mu(\Omega_n) < \infty$ and $\bigcup_{n \in \mathbb{N}} \Omega_n = \Omega$.

$h := 1/(\lambda - f)$. Note that for any $E \in \Sigma$ of finite measure we have

$$\int_E |h(\omega)|^2 \, d\mu(\omega) = \int |h(\omega)\chi_E(\omega)|^2 \, d\mu(\omega) = \|R_\lambda \chi_E\|_2^2 \leq \|R_\lambda\|^2 \|\chi_E\|_2^2 = \|R_\lambda\|^2 \mu(E).$$

We show that $\mu\big(\{\omega \in \Omega \mid |h(\omega)| > \|R_\lambda\|\}\big) = 0$: For $n \in \mathbb{N}$ let $A_n := \{\omega \in \Omega \mid |h(\omega)| \geq \frac{n+1}{n}\|R_\lambda\|\}$; if E is a measurable set of finite measure and with $E \subseteq A_n$, then we deduce from the above that

$$\|R_\lambda\|^2 \mu(E) \geq \int_E |h(\omega)|^2 \, d\mu(\omega) \geq \frac{(n+1)^2}{n^2} \|R_\lambda\|^2 \int_E 1 \, d\mu = \frac{(n+1)^2}{n^2} \|R_\lambda\|^2 \mu(E),$$

hence $\mu(E) = 0$ (since certainly $R_\lambda \neq 0$); therefore, $\mu(A_n) = 0$ for every $n \in \mathbb{N}$, and $\bigcup_{n \in \mathbb{N}} A_n = \{\omega \in \Omega \mid |h(\omega)| > \|R_\lambda\|\}$ has measure 0.

We conclude that $\frac{1}{|\lambda - f(\omega)|} = |h(\omega)| \leq \|R_\lambda\|$ holds for μ-almost all $\omega \in \Omega$, which implies that λ cannot belong to in the essential range of f.

Theorem: Let T be a symmetric densely defined operator on H, then

$$T \text{ is self-adjoint} \quad \Longleftrightarrow \quad \sigma(T) \subseteq \mathbb{R}.$$

Proof: $\Leftarrow$: We have $\mp i \in \rho(T)$, which implies $\operatorname{ran}(T \pm i) = H$. By Theorem 2.10, T is self-adjoint.

$\Rightarrow$: Let $z = \lambda + i\mu$ with real $\mu \neq 0$ and consider $S := (T - \lambda)/\mu$ on $\operatorname{dom}(S) := \operatorname{dom}(T)$. The operator S is self-adjoint and by Equation (2.1) (on page 37), applied to S, we obtain for every $x \in \operatorname{dom}(T)$:

$$\|(z - T)x\|^2 = \|(\lambda + i\mu)x - (\lambda + \mu S)x\|^2 = \|i\mu x - \mu Sx\|^2 = \mu^2 \|(i - S)x\|^2 \geq \mu^2 \|x\|^2.$$

From this we learn that the inverse $(z - T)^{-1}$ exists as linear map $\operatorname{ran}(z - T) \to \operatorname{dom}(T)$ and is bounded. Since S is self-adjoint, $\operatorname{ran}(z - T) = \operatorname{ran}(i - S) = H$ by Theorem 2.10 and therefore $z \in \rho(T)$. □

Thus, symmetric operators need not have real spectrum. Recall that non-closed operators have spectrum equal to $\mathbb{C}$, hence, in particular, any non-closed densely defined symmetric operator T has $\sigma(T) = \mathbb{C} \not\subseteq \mathbb{R}$ (for example, $Tx = ix'$ on $L^2([0,1])$) with $\operatorname{dom}(T) = \{x \in C^1([0,1]) \mid x(0) = 0 = x(1)\}$, which has been studied repeatedly above).

3. The spectral theorem for unbounded self-adjoint operators

3.1. Theorem (Multiplication operator version of the spectral theorem): Let the operator $T\colon \operatorname{dom}(T) \to H$ be self-adjoint, then there exists a measure space (Ω, Σ, μ), a measurable function $f\colon \Omega \to \mathbb{R}$, and a unitary operator $U\colon H \to L^2(\Omega, \mu)$ such that

(a) $\forall x \in H\colon \quad x \in \operatorname{dom}(T) \quad \text{if and only if} \quad f \cdot Ux \in L^2(\Omega, \mu)$,

(b) T is unitarily equivalent via U to the multiplication operator $M_f\,\varphi := f\varphi$ on $L^2(\Omega, \mu)$ with $\operatorname{dom}(M_f) := \{\varphi \in L^2(\Omega, \mu) \mid f\varphi \in L^2(\Omega, \mu)\}$, i.e., for every $\varphi \in \operatorname{dom}(M_f)$,

$$UTU^{-1}\varphi = M_f\,\varphi = f\varphi \quad \mu\text{-almost everywhere.}$$

Proof: By Theorem 2.15, $\sigma(T) \subseteq \mathbb{R}$ and therefore $R := (T+i)^{-1}$ and $(T-i)^{-1}$ exist as bounded operators $H \to \operatorname{dom}(T)$. The plan of the proof is to show that R is normal, apply the spectral theorem for bounded normal operators to R, and "map" everything back to T via the formal relation $T = R^{-1} - i$.

Let $z_1, z_2 \in H$ arbitrary. There exist $x, y \in \operatorname{dom}(T)$ such that $z_1 = (T+i)x$ and $z_2 = (T-i)y$ and we obtain

$$\langle Rz_1, z_2\rangle = \langle x, (T-i)y\rangle = \langle (T+i)x, y\rangle = \langle z_1, (T-i)^{-1}z_2\rangle,$$

i.e., $R^* = (T-i)^{-1}$ and the resolvent equation, Proposition 2.15(ii), shows that

$$(*) \qquad RR^* = R^*R = \frac{1}{2i}(R^* - R).$$

Thus, R is normal and by Theorem 1.25 there is a measure space (Ω, Σ, μ), a unitary map $U\colon H \to L^2(\Omega, \mu)$, and a bounded measurable function $g\colon \Omega \to \mathbb{C}$ such that $URU^{-1} = M_g$, where M_g denotes the operator of multiplication by g on $L^2(\Omega, \mu)$.

Being an inverse map, R is injective, hence M_g is and $N := \{\omega \in \Omega \mid g(\omega) = 0\}$ has to be a μ-null set. Therefore, $f(\omega) := \dfrac{1}{g(\omega)} - i$ is defined for μ-almost every $\omega \in \Omega$. Moreover, from $(*)$ we obtain

$$|g|^2 = \frac{1}{2i}(\overline{g} - g) = \frac{1}{2i}(-2i\operatorname{Im}(g)) = -\operatorname{Im}(g)$$

and may deduce that

$$\forall \omega \in \Omega \setminus N\colon \quad f(\omega) = \frac{1}{g(\omega)} - i = \frac{\overline{g(\omega)}}{g(\omega)\overline{g(\omega)}} - i = \frac{\operatorname{Re}(g(\omega)) + i|g(\omega)|^2}{|g(\omega)|^2} - i = \frac{\operatorname{Re}(g(\omega))}{|g(\omega)|^2} \in \mathbb{R},$$

in particular, f can be considered as a *real-valued* measurable function $f\colon \Omega \to \mathbb{R}$ such that $fg = 1 - ig$ holds μ-almost everywhere, which also shows that fg is essentially bounded.

We prove (a): Let $x \in \mathrm{dom}(T)$, then there is $y \in H$ such that $x = Ry$ and we obtain

$$f \cdot Ux = f \cdot URy = f \cdot (M_g Uy) = fg \cdot Uy \in L^2(\Omega, \mu),$$

because fg is essentially bounded and $Uy \in L^2(\Omega, \mu)$.
To show the reverse implication, suppose $f \cdot Ux \in L^2(\Omega, \mu)$. Then $(f+i)Ux \in L^2(\Omega, \mu)$ as well and we may choose $y \in H$ such that $(f+i)Ux = Uy$, hence $g \cdot Uy = g(f+i)Ux = (gf + ig)Ux = Ux$. In other words, $x = (U^{-1}M_g U)y = Ry = (T+i)^{-1}y \in \mathrm{dom}(T)$.

We prove (b): By (a) we have $\mathrm{dom}(M_f) = U(\mathrm{dom}(T))$. Let $\varphi \in \mathrm{dom}(M_f)$ and $x \in \mathrm{dom}(T)$ such that $\varphi = Ux$. Choose $y \in H$ with $x = Ry = (T+i)^{-1}y$, then $y = Tx + ix$ and $g \cdot Uy = Ux$ as above. We obtain μ-almost everywhere

$$UTU^{-1}\varphi = UTx = U(y - ix) = Uy - iUx = \frac{1}{g}Ux - iUx = (\frac{1}{g} - i)Ux = f \cdot Ux = f \cdot \varphi.$$

□

3.2. Example: [As earlier in Example 1.24 we will make use of basic facts about the Fourier transform and refer again to the course on Real Analysis or [Con16, Fol99].]
We consider the operator $Tx = -ix'$ with $\mathrm{dom}(T) = \mathscr{S}(\mathbb{R}) \subseteq L^2(\mathbb{R})$.
Let $\mathcal{F}$ denote the Fourier transform as unitary operator $L^2(\mathbb{R}) \to L^2(\mathbb{R})$, then $\mathcal{F}(\mathscr{S}(\mathbb{R})) = \mathscr{S}(\mathbb{R})$ and, similarly to Example 2.13, 2), we obtain $(\mathcal{F}\, T\, \mathcal{F}^{-1}\varphi)(t) = t\,\varphi(t)$ (from the so-called "exchange" of differentiation with multiplication by the Fourier transform), i.e., the unitary equivalence of T with the operator M, with $\mathrm{dom}(M) = \mathscr{S}(\mathbb{R})$ of multiplication by the real variable. Clearly, T and M are symmetric and densely defined; moreover, both subspaces $\mathrm{ran}(T \pm i) = \mathcal{F}^{-1}(\mathrm{ran}(M \pm i))$ and $\mathrm{ran}(M \pm i)$ are dense in $L^2(\mathbb{R})$, since the latter obviously contains $\mathscr{D}(\mathbb{R})$. Thus, T and M are essentially self-adjoint by Corollary 2.11 and their unique self-adjoint extensions are given by (their closures) T^* and M^*.
We claim that $\mathrm{dom}(M^*) = \{\psi \in L^2(\mathbb{R}) \mid t \mapsto t\,\psi(t) \in L^2(\mathbb{R})\}$: It is clear that any $\psi \in L^2(\mathbb{R})$ such that $t \mapsto t\,\psi(t) \in L^2(\mathbb{R})$ belongs to $\mathrm{dom}(M^*)$, since we may then write $\langle M\varphi, \psi\rangle = \int \varphi(t)\,\overline{t\psi(t)}\,dt$ for every $\varphi \in \mathrm{dom}(M) = \mathscr{S}(\mathbb{R})$ and obtain $|\langle M\varphi, \psi\rangle| \leq (\int t^2 |\psi(t)|^2\,dt)^{1/2} \|\varphi\|_2$. To show the reverse inclusion, suppose $\psi \in L^2(\mathbb{R})$ is such that $l_\psi\colon \varphi \to \langle M\varphi, \psi\rangle$ is an L^2-norm continuous linear functional $\mathscr{S}(\mathbb{R}) \to \mathbb{C}$. By density of $\mathscr{S}(\mathbb{R}) \subseteq L^2(\mathbb{R})$, it means that there is some $\eta \in L^2(\mathbb{R})$ such that

$$(*) \qquad \forall \varphi \in \mathscr{S}(\mathbb{R})\colon \int \varphi(t)\, t\, \overline{\psi(t)}\, dt = \int \varphi(t)\, \overline{\eta(t)}\, dt.$$

(Note that we have $l_\psi(\varphi) = \langle \varphi, \eta\rangle$ for every $\varphi \in L^2(\mathbb{R})$, but at this moment we have the formula $l_\psi(\varphi) = \int t\varphi(t)\overline{\psi(t)}\,dt$ only if $\varphi \in \mathscr{S}(\mathbb{R})$.) Since both $t \mapsto t\psi(t)$ and η are integrable on every compact subset of $\mathbb{R}$, we may use the following fact, which is shown later in course of an injectivity argument in Example 1) of 7.5: If f_1 and f_2 are locally integrable functions on $\mathbb{R}$ such that $\int \varphi f_1 = \int \varphi f_2$ holds for every $\varphi \in \mathscr{S}(\mathbb{R})$, then $f_1 = f_2$ almost everywhere. Thus, $(*)$ implies that $t\,\psi(t) = \eta(t)$ for almost all t, hence the function $t \mapsto t\,\psi(t)$ belongs to $L^2(\mathbb{R})$.

We also obtain $(M^*\psi)(t) = t\,\psi(t)$ and recognize M^* as a special case of the self-adjoint multiplication operator in Example 2.8. Therefore, $T^*x = -ix'$ on the domain $H^1(\mathbb{R}) := \mathcal{F}^{-1}(\mathrm{dom}(M^*))$, which can be shown to consist of all L^2-functions that are absolutely continuous on compact intervals and have derivative in $L^2(\mathbb{R})$—this is the (L^2-based) *Sobolev space* of order 1 on $\mathbb{R}$.

3.3. Spectral representation of an unbounded self-adjoint operator: Let T be self-adjoint and M_f as in Theorem 3.1. If $h\colon \mathbb{R} \to \mathbb{C}$ is a bounded measurable function, then $h \circ f\colon \Omega \to \mathbb{C}$ is bounded measurable and the multiplication operator $h(M_f) := M_{h\circ f}$ of multiplication by $h \circ f$ is bounded on $L^2(\Omega, \mu)$. The map $h \mapsto h(M_f)$ is continuous and multiplicative $B_b(\mathbb{R}) \to L(L^2(\Omega, \mu))$, since clearly $\|M_{h\circ f}\| \leq \|h\|_\infty$ and $(h_1 \cdot h_2) \circ f = (h_1 \circ f) \cdot (h_2 \circ f)$.

For any Borel subset $A \subseteq \mathbb{R}$ we put

$$F_A := \chi_A(M_f) = M_{\chi_A \circ f} = M_{\chi_{f^{-1}(A)}}.$$

It is not difficult to check that $F\colon A \mapsto F_A$ is a spectral measure and that

$$h(M_f) = \int_{\mathbb{R}} h(\lambda)\,dF_\lambda \quad \forall h \in B_b(\mathbb{R}),$$

since the latter obviously holds for step functions and the multiplication operator action reduces the question about the relevant limits to the approximation of bounded measurable functions by step functions. In general, F will not have compact support.

Let $E_A := U^{-1}F_AU$ $(A \in \mathcal{B}(\mathbb{R}))$, then $E\colon A \mapsto E_A$ is a spectral measure $\mathcal{B}(\mathbb{R}) \to L(H)$ and for any bounded measurable function $h\colon \mathbb{R} \to \mathbb{C}$ we have

$$\int_{\mathbb{R}} h(\lambda)\,dE_\lambda = U^{-1}h(M_f)U,$$

again by extending the obvious relation for characteristic functions to step functions and further to bounded measurable functions via approximation. We put $h(T) := U^{-1}h(M_f)U$ and obtain the usual properties of a functional calculus for the map $h \mapsto h(T)$.

We will extend this functional calculus to allow for a measurable unbounded *real-valued function* $h\colon \mathbb{R} \to \mathbb{R}$, thereby obtaining a self-adjoint unbounded operator $h(T)$ on the dense domain $D_h := \{x \in H \mid \int |h(\lambda)|^2\,d\langle E_\lambda x, x\rangle < \infty\}$ (the density is not obvious and will be proved below).

Let $x, y \in H$ and $\varphi, \psi \in L^2(\Omega, \mu)$ such that $Ux = \varphi$, $Uy = \psi$, then for any $A \in \mathcal{B}(\mathbb{R})$,

$$\langle E_A x, y\rangle = \langle U^{-1}F_AUx, U^{-1}Uy\rangle = \langle F_A\varphi, \psi\rangle = \int \chi_{f^{-1}(A)}\,\varphi\,\overline{\psi}\,d\mu = \int_{f^{-1}(A)} \varphi\,\overline{\psi}\,d\mu.$$

Putting $\nu(B) := \int_B \varphi\overline{\psi}\,d\mu$ we may interpret the last expression above as image measure $f(\nu)(A) = \nu(f^{-1}(A))$, hence $\langle E_A x, y\rangle = f(\nu)(A)$. The transformation formula for the integral of a $d\langle E_\lambda x, y\rangle$-integrable function $g\colon \mathbb{R} \to \mathbb{R}$ with respect to image measures then gives

$$\int_{\mathbb{R}} g(\lambda)\,d\langle E_\lambda x, y\rangle = \int_{\Omega} (g \circ f)\,d\nu = \int_{\Omega} (g \circ f) \cdot \varphi \cdot \overline{\psi}\,d\mu. \tag{3.1}$$

Applying the above equation to $|h|^2$ we find that $x \in D_h$, if and only if $\varphi = Ux$ satisfies $\int_\Omega |h \circ f|^2 \, |\varphi|^2 \, d\mu < \infty$. Note that Example 2.8, now with $h \circ f$ in place of f there, shows that the set of such φ, namely, $\mathrm{dom}(M_{h\circ f})$, is dense in $L^2(\Omega, \mu)$. Therefore, the corresponding set of vectors $x = U^{-1}\varphi$, which is D_h, is dense in H.

For every $x \in D_h$ and $y \in H$, the integral $\int_\mathbb{R} h(\lambda)\, d\langle E_\lambda x, y\rangle$ exists: Using $\varphi = Ux$, $\psi = Uy$ and the analogue of (3.1) for the variations of the complex measures, we have

$$\int_{\mathbb{R}} |h(\lambda)|\, d|\langle E_\lambda x, y\rangle| = \int_{\Omega} |h \circ f| \cdot |\varphi| \cdot |\overline{\psi}| \, d\mu \leq \left(\int_{\Omega} |h \circ f|^2 \cdot |\varphi|^2 \, d\mu \right)^{1/2} \left(\int_{\Omega} |\psi|^2 \, d\mu \right)^{1/2}$$

$$= \left(\int_{\mathbb{R}} |h(\lambda)|^2 \, d\langle E_\lambda x, x\rangle \right)^{1/2} \underbrace{\|Uy\|_2}_{=\|y\|} < \infty.$$

Hence, for every $x \in D_h$ there is a unique vector $z \in H$ such that $\int_\mathbb{R} h(\lambda)\, d\langle E_\lambda x, y\rangle = \langle z, y\rangle$ for all $y \in H$; we put $h(T)x := z$. Therefore, we obtain

$$\langle h(T)x, y\rangle = \int_{\mathbb{R}} h(\lambda)\, d\langle E_\lambda x, y\rangle \quad \forall x \in D_h, \forall y \in H. \tag{3.2}$$

This defines an operator $h(T)$ on H with dense domain D_h. The formal notation $h(T) = \int h(\lambda)\, dE_\lambda$ is commonly used, although this integral is in general not convergent with respect to the operator norm; the precise statement is Equation (3.2).

The special case $h(t) = t$ gives $x \in D_h$, if and only if $\varphi = Ux$ satisfies $f\varphi \in L^2(\Omega, \mu)$, i.e., $Ux \in \mathrm{dom}(M_f)$, equivalently $x \in \mathrm{dom}(T)$ by Theorem 3.1; moreover, we have

$$\int_{\mathbb{R}} \lambda \, d\langle E_\lambda x, y\rangle = \int_{\Omega} f\varphi\,\overline{\psi}\, d\mu = \langle M_f \varphi, \psi\rangle = \langle M_f Ux, Uy\rangle = \langle U^{-1} M_f U x, y\rangle = \langle Tx, y\rangle.$$

Similarly, one may show that $\langle h(T)x, y\rangle = \langle U^{-1} M_{h\circ f} U x, y\rangle$ holds for all $x \in D_h, y \in H$, which implies also the self-adjointness of $h(T)$ due to the established unitary equivalence with the self-adjoint multiplication operator $M_{h\circ f}$ with domain $\mathrm{dom}(M_{h\circ f}) = \{\varphi \in L^2(\Omega, \mu) \mid \int_\Omega |h \circ f|^2 \, |\varphi|^2 \, d\mu < \infty\}$.

To summarize, we have proved the following spectral representation.

Theorem: Let $T\colon \mathrm{dom}(T) \to H$ be self-adjoint, then there exists a unique spectral measure E such that

$$\langle Tx, y\rangle = \int_{\mathbb{R}} \lambda \, d\langle E_\lambda x, y\rangle \quad \forall x \in \mathrm{dom}(T), y \in H.$$

If $h\colon \mathbb{R} \to \mathbb{R}$ is measurable and $D_h := \{x \in H \mid \int |h(\lambda)|^2 \, d\langle E_\lambda x, x\rangle < \infty\}$, then

$$\langle h(T)x, y\rangle = \int_{\mathbb{R}} h(\lambda)\, d\langle E_\lambda x, y\rangle \quad \forall x \in D_h, y \in H,$$

defines a self-adjoint operator $h(T)$ on H with domain D_h.

3.4. A glimpse at the formalism of quantum mechanics: [We discuss here the formalism based directly on Hilbert spaces containing the physical states as in the so-called Schrödinger picture described in [Kab14, Hal13, Tes14b, Wei00]. The more abstract algebraic approach is based on C^*-algebras of observables and linear functionals on these as states (see, e.g., [Ara99, Thi03]).]

The basic idea is that (the state of) a single particle in $\mathbb{R}^d$ is described by a so-called complex wave function $\mathbb{R}^d \times \mathbb{R} \to \mathbb{C}$, $(x,t) \mapsto \psi(x,t)$, where $\rho_t(x) := |\psi(x,t)|^2$ is interpreted as the probability density associated with the (state of the) particle at time t. Thus, the number

$$\int_K |\psi(x,t)|^2 \, dx = \int_{\mathbb{R}^d} \chi_K(x) |\psi(x,t)|^2 \, dx$$

corresponds to the probability of finding the particle (in state ψ) at time t inside the measurable region $K \subseteq \mathbb{R}^d$ and, in particular, $\int_{\mathbb{R}^d} |\psi(x,t)|^2 \, dx = 1$, since the particle has to be somewhere.

To put it differently, $\psi(.,t) \in L^2(\mathbb{R}^d)$ for every $t \in \mathbb{R}$, $\|\psi(.,t)\|_{L^2} = 1$, and the value of the L^2-inner product $\langle \chi_K \, \psi(.,t), \psi(.,t) \rangle$ gives the expectation of the random variable $x \mapsto \chi_K(x)$ at time t (for a particle in state ψ). The bounded linear operator $\psi \mapsto \chi_K \, \psi$ is but one example of an observable, the most prominent other examples are the unbounded operators of spatial (coordinate) position $X_j \colon \psi \mapsto x_j \psi$ and momentum (coordinates) $P_j \colon \psi \mapsto \frac{\hbar}{i} \partial_{x_j} \psi$ ($\psi \in \mathscr{S}(\mathbb{R}^d)$).

In more abstract terms, the *quantum mechanical (vector) states* are represented by elements ψ of a complex (usually separable) Hilbert space H that are normalized such that $\|\psi\| = 1$. The *observables* are represented by densely defined linear operators on H. (Densely defined, since the observable should be "observable in most of the possible states".) The *expectation* (or mean value of measurement) of an observable A in the state $\psi \in \mathrm{dom}(A)$ is given by

$$E_\psi(A) := \langle A\psi, \psi \rangle.$$

In order to guarantee that all expectation values are real numbers, we suppose that A is symmetric (then $\overline{\langle A\psi, \psi \rangle} = \langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle$). But there are clear indications to require even more: The set of all possible *measurements* of an observable A, represented by the spectrum $\sigma(A)$, should be real, thus, calling on Theorem 2.15, we further suppose that the observables are self-adjoint.

Note that in case ψ is an eigenvector for A corresponding to the eigenvalue λ_0 we obtain

$$E_\psi(A) = \lambda_0 \langle \psi, \psi \rangle = \lambda_0,$$

since $A\psi = \lambda_0 \psi$. In general, the *mean deviation* $\Delta_\psi(A) := \|A\psi - E_\psi(A)\psi\|$ of A in the state ψ will not vanish. We have just seen that precisely the eigenvalues occur as sharp measurements (i.e., with zero deviation).

Recall from linear algebra that two self-adjoint operators A and B on a finite-dimensional Hilbert space are simultaneously diagonalizable if and only if they commute, i.e., $[A, B] := AB - BA = 0$. Considering now the infinite-dimensional case and interpreting the eigenvalues as the possible sharp measurements of observables, one can generalize this and ask to what extent simultaneous measurements $E_\psi(A)$, $E_\psi(B)$ can be achieved at least with small mean deviations. There is a general lower bound, which in the example $A = X_j$, $B = P_j$ of the position and momentum operators (in the same coordinate) implies the famous *Heisenberg uncertainty principle* from the fact $[X_j, P_j] = i\hbar I$: If $\psi \in \{\varphi \in \mathrm{dom}(A) \cap \mathrm{dom}(B) \mid B\varphi \in \mathrm{dom}(A)$ and $A\varphi \in \mathrm{dom}(B)\}$, then

$$\Delta_\psi(A)\,\Delta_\psi(B) \geq \frac{1}{2}\,|\langle [A,B]\psi, \psi\rangle|.$$

Proof: Let $A' := A - E_\psi(A)$ and $B' := B - E_\psi(B)$. Since I commutes with every operator, we have[1]

$$\begin{aligned}\langle [A,B]\psi,\psi\rangle &= \langle (AB-BA)\psi,\psi\rangle = \langle (A'B'-B'A')\psi,\psi\rangle = \langle A'B'\psi,\psi\rangle - \langle \psi,(B'A')^*\psi\rangle \\ &= \langle A'B'\psi,\psi\rangle - \langle \psi, A'B'\psi\rangle = \langle A'B'\psi,\psi\rangle - \overline{\langle A'B'\psi,\psi\rangle} = 2i\,\mathrm{Im}\langle A'B'\psi,\psi\rangle.\end{aligned}$$

Therefore, we deduce

$$|\langle [A,B]\psi,\psi\rangle| \leq 2\,|\langle A'B'\psi,\psi\rangle| = 2\,|\langle B'\psi, A'\psi\rangle| \leq 2\,\|B'\psi\|\,\|A'\psi\| = 2\,\Delta_\psi(A)\,\Delta_\psi(B).$$

□

The dynamical aspect of the quantum mechanical particle in $\mathbb{R}^d$ as introduced above is in the t-evolution of the wave function $\psi(.,t) \in L^2(\mathbb{R}^d)$. The model assumption is that $\psi(.,t) = U(t)\psi(.,0)$, where $U(t)$ is linear (superposition principle) and maps states into states, thus $\|U(t)\psi(.,0)\|_2 = \|\psi(.,0)\|_2$. Hence it is plausible to suppose that $U(t)$ is a unitary operator for every t. Moreover, it is natural to suppose that the evolution of duration t starting at time s gives the same result as the evolution of duration $t+s$ starting at time 0, i.e., to require $U(t)U(s)\psi(.,0) = U(t)\psi(.,s) = \psi(.,t+s) = U(t+s)\psi(.,0)$, or, $U(t)U(s) = U(t+s)$ on the level of operators. Thus, $t \mapsto U(t)$ is a group homomorphism from $(\mathbb{R},+)$ to the group of unitary operators on $L^2(\mathbb{R}^d)$. Finally, we will expect a certain continuity property of the evolution, at least along each trajectory of a given initial state, i.e., continuity of the map $t \mapsto U(t)\psi(.,0)$, $\mathbb{R} \to L^2(\mathbb{R}^d)$.

In abstract terms, we implement the dynamical aspect of a quantum mechanical model on the Hilbert space H of states by a so-called *strongly continuous (s-continuous) unitary group* of operators $U(t)$ ($t \in \mathbb{R}$), which means that $U(t) \in L(H)$ is unitary for every t, $U(t)U(s) = U(t+s)$ holds for all $s,t \in \mathbb{R}$, and for every $\psi \in H$ and $t_0 \in \mathbb{R}$ we have $\lim_{t\to t_0} U(t)\psi = U(t_0)\psi$. (We see that the term 'strong continuity' here simply refers to the pointwise continuity of the map $t \mapsto U(t)$, since the convergence $U(t) \to U(t_0)$ is tested pointwise on H; the topology on $L(H)$ describing this pointwise convergence is usually called the *strong operator topology*, hence the notion of s-continuity.)

We can obtain a strongly continuous unitary group on H from any self-adjoint operator S via functional calculus. We may simply put $U(t) := \exp(itS)$ and check, basically as an exercise, that all the required properties hold (cf. [Con10, Chapter X, Theorem 5.1] or [Kab14, Satz 16.16] or [Wei00, Satz 7.4]). Moreover, it is not difficult to see that $\mathrm{dom}(S)$ is invariant under $U(t)$ and

$$\forall \psi \in \mathrm{dom}(S): \quad \frac{d}{dt}\big(U(t)\psi\big) := \lim_{h\to 0} \frac{U(t+h)\psi - U(t)\psi}{h} = iS\,U(t)\psi,$$

[1] boldly using the relation $(ST)^* = T^*S^*$ for unbounded operators here without proof

i.e., $\varphi(t) := U(t)\psi$ satisfies the initial value problem

$$\frac{d}{dt}\varphi(t) = iS\varphi(t), \quad \varphi(0) = \psi. \tag{3.3}$$

By the famous theorem of Stone (see [Con10, Chapter X, Theorem 5.6] or [Kab14, Satz 16.18] or [Wei00, Satz 7.3]), every s-continuous unitary group $U(t)$ ($t \in \mathbb{R}$) is given in the above form $U(t) = \exp(itS)$, where S is a unique self-adjoint operator. In fact, S may be constructed on $\mathrm{dom}(S) := \{\psi \in H \mid \exists \lim_{h\to 0}(U(h)\psi - \psi)/h\}$ by the assignment $S\psi := -i \cdot \lim_{h\to 0}(U(h)\psi - \psi)/h$.

Therefore, the dynamics of the quantum mechanical system is determined by this self-adjoint operator S. In terms of physics, S corresponds to the energy of the system and is called the *Hamiltonian* and the differential equation in (3.3) is called the *Schrödinger equation*. Observe that in case the initial value ψ happens to be an eigenvector for S with eigenvalue $\lambda \in \mathbb{R}$, i.e., $S\psi = \lambda\psi$, then the solution to (3.3) is given by $\varphi(t) = e^{i\lambda t}\psi$, which corresponds to a *bound state* in physics, since $\varphi(t) \in \mathrm{span}\{\psi\}$ for all times $t \in \mathbb{R}$; in particular, the expectation values of observables remain constant: $E_{\varphi(t)}(A) = \langle A\varphi(t), \varphi(t)\rangle = e^{i\lambda t}e^{-i\lambda t}\langle A\psi, \psi\rangle = E_{\psi}(A)$.

Part II.

Locally convex vector spaces

4. Vector space topologies

In this second part of the lecture notes, X will denote a vector space with scalar field $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

4.1. Definition: A *seminorm* on X is a map $p\colon X \to [0, \infty[$ such that $\forall \lambda \in \mathbb{K}, x, y \in X$:

(a) $p(\lambda x) = |\lambda|\, p(x)$,

(b) $p(x + y) \leq p(x) + p(y)$.

Note that $p(x) \geq 0$ and that $p(0) = 0$ (by (a)), but $p(x) = 0$ does not imply $x = 0$.

4.2. Example: Let $X = \mathbb{C}^{[0,1]} = \{x\colon [0, 1] \to \mathbb{C}\}$ and for every $t \in [0, 1]$ let $p_t\colon X \to [0, \infty[$ be defined by $p_t(x) := |x(t)|$. Then p_t is a seminorm on X and we see that $p_t(x) = 0$ for any function x with $x(t) = 0$. The set of seminorms $P := \{p_t \mid t \in [0, 1]\}$ can be used to describe the pointwise convergence of functions in the sense that $x_n \to x$ pointwise on $[0, 1]$, if and only if

$$\forall t \in [0, 1]: \quad p_t(x_n - x) \to 0 \quad (n \to \infty),$$

written in more abstract terms: $\forall p \in P$ we have $p(x_n - x) \to 0$ as $n \to \infty$.
The statement for the pointwise convergence of a net $(x_j)_{j\in J}$ to x in X is completely analogous: It is equivalent to the fact that, for every $p \in P$, the net $(p(x_j - x))_{j\in J}$ converges to 0 in $\mathbb{C}$, i.e., for every $\varepsilon > 0$ there is some $j_0 \in J$ such that $p(x_j - x) < \varepsilon$, if $j \geq j_0$.
(Recall that $(J, \leq)$ is a *directed set*, if $\leq$ is a reflexive and transitive relation on J such that for any $j_1, j_2 \in J$ we can find $j_3 \in J$ with $j_1 \leq j_3$ and $j_2 \leq j_3$.)

We will show how to define vector space topologies with good properties from seminorms. The very basic notion is as follows.

4.3. Definition: Let X be a vector space over the field $\mathbb{K}$ and τ be a topology on X. Then (X, τ) is a *topological vector space*, if addition $X \times X \to X$, $(x, y) \mapsto x + y$, and scalar multiplication $\mathbb{K} \times X \to X$, $(\lambda, x) \mapsto \lambda x$, are continuous maps with respect to the product topologies on $X \times X$ and $\mathbb{K} \times X$.

The requirements for a topology τ on X to turn (X, τ) into a topological vector space are not trivial, as the example of the discrete topology on $X \neq \{0\}$ illustrates: The scalar multiplication then cannot be continuous $\mathbb{K} \times X \to X$, since with any $x \neq 0$ and $\lambda_n = 1/n$ we have $(\lambda_n, x) \to (0, x)$ and $0 \cdot x = 0$, but $0 \neq \lambda_n \cdot x$ does not converge to 0 in X with respect to the discrete topology.

As a preparation for the study of the so-called *locally convex vector space topologies* we need the following notions describing properties of subsets of $A \subseteq X$. Recall that A is *convex*, if for every $x, y \in A$ and for every $\lambda \in \mathbb{R}$, $0 \leq \lambda \leq 1$ we have $\lambda x + (1 - \lambda)y \in A$.

4.4. Definition: Let $A \subseteq X$, then A is called

(a) *balanced*, if for every $x \in A$ and every $\lambda \in \mathbb{K}$, $|\lambda| \leq 1$, we have $\lambda x \in A$,

(b) *absolutely convex*, if A is convex and balanced,

(c) *absorbing*, if for every $x \in X$ there is a number $c_x > 0$ such that for all $\lambda \in \mathbb{K}$, $0 \leq \lambda \leq c_x$, we have $\lambda x \in A$. (Since $x \in \frac{1}{\lambda} A$, if $0 < \lambda \leq c_x$, we may say: "Any point in X is swallowed by appropriate dilations of the subset A.") [Slightly different definition in [MV92, Tre67].]

4.5. Lemma: A subset $A \subseteq X$ is absolutely convex, if and only if it satisfies

$$x, y \in A \text{ and } \lambda, \mu \in \mathbb{K} \text{ with } |\lambda| + |\mu| \leq 1 \quad \Longrightarrow \quad \lambda x + \mu y \in A.$$

Proof: Suppose that A is absolutely convex and let x, y, λ, μ be as specified above. If $\lambda = 0$ or $\mu = 0$, the implication holds, since A is balanced. Let $\lambda \neq 0$ and $\mu \neq 0$, then both $\frac{\lambda}{|\lambda|} x$ and $\frac{\mu}{|\mu|} y$ belong to A, since A is balanced. We obtain

$$\lambda x + \mu y = \underbrace{(|\lambda| + |\mu|)}_{\leq 1} \cdot \underbrace{\left(\frac{|\lambda|}{|\lambda| + |\mu|} \cdot \frac{\lambda}{|\lambda|} x + \frac{|\mu|}{|\lambda| + |\mu|} \cdot \frac{\mu}{|\mu|} y \right)}_{\text{convex combination of } \frac{\lambda}{|\lambda|} x \text{ and } \frac{\mu}{|\mu|} y} \in A,$$

since A is convex and balanced.

Suppose the implication in the statement of the lemma holds. Clearly, convex combinations are a special case of its premise, thus A is convex. To see that A is also balanced, let $x \in A$ and $\nu \in \mathbb{K}$ with $|\nu| \leq 1$. Then $\nu x = \nu x + 0x \in A$. □

4.6. Neighborhood systems from seminorms: Let P be a set of seminorms on X. For any finite subset $F \subseteq P$ and any $\varepsilon > 0$ we define

$$U_{F,\varepsilon} := \{x \in X \mid \forall p \in F \colon p(x) \leq \varepsilon\}.$$

We collect all of these subsets in the family $\mathfrak{U} := \{U_{F,\varepsilon} \mid \varepsilon > 0, F \subseteq P \text{ finite}\}$ and list its basic properties:

(1) $\forall U \in \mathfrak{U}$ we have $0 \in U$.

This is clear, since $p(0) = 0$.

(2) $\forall U_1, U_2 \in \mathfrak{U}$ there exists $U \in \mathfrak{U}$ such that $U \subseteq U_1 \cap U_2$.

Let $U_j = U_{F_j, \varepsilon_j}$ $(j = 1, 2)$ and put $F = F_1 \cup F_2$, $\varepsilon := \min(\varepsilon_1, \varepsilon_2)$, then $U := U_{F,\varepsilon}$ fulfills the requirement, since $U_{F,\varepsilon} \subseteq U_{F_1,\varepsilon_1} \cap U_{F_2,\varepsilon_2}$.

(3) $\forall U \in \mathfrak{U}\ \exists V \in \mathfrak{U}$ such that $V + V \subseteq U$.

If $U = U_{F,\varepsilon}$ we may put $V := U_{F,\varepsilon/2}$, since $U_{F,\varepsilon/2} + U_{F,\varepsilon/2} \subseteq U_{F,\varepsilon}$.

(4) Every $U \in \mathfrak{U}$ is absorbing.

Let $U = U_{F,\varepsilon}$ and $x_0 \in X$. Put $\alpha_{x_0} := \max_{p \in F} p(x_0)$. If $\alpha_{x_0} = 0$, then $x_0 \in U_{F,\varepsilon}$. If $\alpha_{x_0} > 0$, then we put $c_{x_0} := \varepsilon / \alpha_{x_0}$ and obtain for every $p \in F$ and every $\lambda \in \mathbb{K}$ with $0 \leq \lambda \leq c_{x_0}$ that $p(\lambda x_0) = \lambda\, p(x_0) \leq c_{x_0} \alpha_{x_0} \leq \varepsilon$, i.e, $\lambda x_0 \in U_{F,\varepsilon}$.

(5) $\forall U \in \mathfrak{U}\ \forall \lambda \in \mathbb{K}$, $\lambda > 0$, there is some $V \in \mathfrak{U}$ with $\lambda V \subseteq U$.

Let $U = U_{F,\varepsilon}$, $\lambda > 0$, and note that $\lambda \cdot U_{F,\varepsilon/\lambda} = U_{F,\varepsilon}$, thus $V := U_{F,\varepsilon/\lambda}$ does the job.

(6) Every $U \in \mathfrak{U}$ is balanced.

This is clear here from the seminorm properties.

(7) Every $U \in \mathfrak{U}$ is absolutely convex.

This is also clear here from the seminorm properties.

4.7. Proposition: If $\mathfrak{U}$ is a non-empty system of subsets of X satisfying (1)-(6) as in 4.6, we obtain a topology τ on X by specifying for any $O \subseteq X$ that

$$O \in \tau \quad \Longleftrightarrow \quad \forall x \in O\ \exists U \in \mathfrak{U}\colon x + U \subseteq O.$$

Then $\mathfrak{U}$ is a basis of τ-neighborhoods at 0 and (X, τ) is a topological vector space.

(Moreover, then clearly $\{x + U \mid U \in \mathfrak{U}\}$ is a basis of neighborhoods at x, for every $x \in X$.)

Proof: We show that τ is a topology on X: Clearly, $X \in \tau$ and $\emptyset \in \tau$ (the latter as a "formality"). If $O_1, O_2 \in \tau$ and $x \in O_1 \cap O_2$, then choose $U_1, U_2 \in \mathfrak{U}$ such that $x + U_j \subseteq O_j$ $(j = 1, 2)$. By (2), we find $U \in \mathfrak{U}$ with $x + U \subseteq x + U_1 \cap U_2 \subseteq O_1 \cap O_2$. Thus $O_1 \cap O_2 \in \tau$. Finally, if M is some set, $O_q \in \tau$ for every $q \in M$, and $x \in \bigcup_{q \in M} O_q$, then $x \in O_{q_0}$ for some $q_0 \in M$. There exists some $U \in \mathfrak{U}$ such that $x + U \subseteq O_{q_0} \subseteq \bigcup_{q \in M} O_q$. Thus, $\bigcup_{q \in M} O_q$ is open.

Every $U \in \mathfrak{U}$ is a τ-neighborhood of 0: Let $O := \{x \in U \mid \exists W \in \mathfrak{U}\colon x + W \subseteq U\}$. By construction, $0 \in O \subseteq U$ and it remains to show that O is open. Let $x \in O$ and $W \in \mathfrak{U}$ with $x + W \subseteq U$. By (3), there is some $V \in \mathfrak{U}$ such that $V + V \subseteq W$. We show that $x + V \subseteq O$, which implies then that O is open. If $y \in x + V$, then $y + V \subseteq x + V + V \subseteq x + W \subseteq U$, hence $y \in O$.

$\mathfrak{U}$ is a basis of neighborhoods at 0: If $B \subseteq X$ is a τ-neighborhood of 0, then there is some $O \in \tau$ with $0 \in O \subseteq B$. By construction of τ, there is some $U \in \mathfrak{U}$ such that $U = 0 + U \subseteq O \subseteq B$.

Addition $a\colon X \times X \to X$, $a(x, y) := x + y$, is continuous: Given $O \in \tau$, we prove that $a^{-1}(O)$ is open in $X \times X$. Let $(x, y) \in a^{-1}(O)$, i.e, $x + y \in O$. By construction of τ, there is some $U \in \mathfrak{U}$ with $(x+y)+U \subseteq O$. Choose $V \in \mathfrak{U}$ according to (3) with $V + V \subseteq U$. Then $(x+V) \times (y+V)$ is a neighborhood of (x, y) and $(x+V) \times (y+V) \subseteq a^{-1}(O)$, since $(x+V)+(y+V) = x+y+V+V \subseteq x + y + U \subseteq O$.

Scalar multiplication $m\colon \mathbb{K} \times X \to X$, $m(\lambda, x) := \lambda x$, is continuous: Given $O \in \tau$, we prove that $m^{-1}(O)$ is open in $\mathbb{K} \times X$. Let $(\lambda, x) \in m^{-1}(O)$, i.e, $\lambda x \in O$. By construction of τ, there is some $U \in \mathfrak{U}$ with $\lambda x + U \subseteq O$. Our aim is to find some $\varepsilon > 0$ and $W \in \mathfrak{U}$ such that $m(D_\varepsilon(\lambda) \times (x + W)) \subseteq O$, where $D_\varepsilon(\lambda) = \{\mu \in \mathbb{K} \mid |\mu - \lambda| < \varepsilon\}$.

Choose $V \in \mathfrak{U}$ with $V + V \subseteq U$ (by property (3)) and then $1 \geq \varepsilon > 0$ such that $\varepsilon x \in V$ (by property (4)). Since V is balanced, by (6), we have $(\mu - \lambda)x \in V$ for every $\mu \in D_\varepsilon(\lambda)$. Finally, we employ (5) to choose $W \in \mathfrak{U}$ such that $(|\lambda| + \varepsilon)W \subseteq V$; by (6), we further have $\mu W \subseteq V$, if $|\mu| \leq |\lambda| + \varepsilon$. We obtain, for every $\mu \in D_\varepsilon(\lambda)$ and for every $w \in W$,

$$\mu(x + w) = \lambda x + (\mu - \lambda)x + \mu w \in \lambda x + V + V \subseteq \lambda x + U \subseteq O.$$

□

4.8. Remark: It is not difficult to see that every topological vector space possesses a basis of neighborhoods at 0 satisfying properties (1)-(6). Clearly, (1) and (2) hold for any neighborhood basis at 0. Properties (3)-(5) hold for the system of all neighborhoods at 0: Property (3) follows from continuity of addition at $(0,0)$. Property (4) follows from continuity of $\lambda \mapsto \lambda x$ at $0 \in \mathbb{K}$ for any $x \in X$, and (5) from continuity of the map $x \mapsto \lambda x$ at $0 \in X$ for any λ. Finally, every neighborhood U of 0 contains a balanced 0-neighborhood W: By continuity of scalar multiplication at $(0,0) \in \mathbb{K} \times X$, we can find $\varepsilon > 0$ and a neighborhood V of 0 such that $\lambda V \subseteq U$, if $|\lambda| \leq \varepsilon$; setting $W := \{\lambda v \mid |\lambda| \leq \varepsilon, v \in V\} = \bigcup_{|\lambda|\leq\varepsilon} \lambda V$ we clearly obtain a balanced subset, which is also a 0-neighborhood, since each λV with $\lambda \neq 0$ is one (as the image of V under the homeomorphism $x \mapsto \lambda x$). In conclusion, the set of all balanced neighborhoods at 0 satisfies (1)-(6).

We learn from 4.6 and 4.7 that a family of seminorms generates a topology for a topological vector space. In this topology, the basic neighborhoods $U_{F,\varepsilon}$ have the additional property (7) of being absolutely convex, which has not been used in Proposition 4.7. As it turns out, this additional property is crucial, for example, to guarantee a rich theory of linear functionals. Moreover, most of the topological vector spaces arising in applications are built from seminorms and it can be shown that topologically this is equivalent to the existence of a basis of neighborhoods at 0 consisting of (absolutely) convex subsets (see, e.g., [MV92, Lemmata 22.2 und 22.4] or [Tre67, Proposition 7.6], and [Wer18, Satz VIII.1.5] for a sketch of the proof). The key ingredient in constructing seminorms from absolutely convex neighborhoods of 0 is the *Minkowski functional*, $p_U(x) := \inf\{\lambda > 0 \mid x \in \lambda U\}$ (see also a corresponding lemma in 5.10).

4.9. Definition: A topological vector space is said to be *locally convex*, if it possesses a basis of neighborhoods at 0 consisting of convex sets.

From the discussion and the references given above, we have the following statement.

4.10. Proposition: A topological vector space (X, τ) is locally convex, if and only if τ is defined by a family of seminorms as in 4.6 and 4.7.

In particular, every normed vector space is a locally convex space, which happens to be metrizable as well and thus also a Hausdorff space. In general, locally convex or topological vector spaces need not be Hausdorff as the example of the chaotic topology (corresponding to the single seminorm $x \mapsto 0$) illustrates. However, there are the following useful criteria. (We note that the equivalence (i) $\Leftrightarrow$ (iii) holds in general topological vector spaces).

4.11. Lemma: Let P be a set of seminorms on X generating the locally convex vector space topology τ. Then the following are equivalent:

(i) (X, τ) is a Hausdorff space,

(ii) for every $x \in X$, $x \neq 0$, there is a $p \in P$ such that $p(x) \neq 0$,

(iii) there is a neighborhood basis $\mathfrak{U}$ of 0 such that $\bigcap_{U \in \mathfrak{U}} U = \{0\}$.

Proof: (i) $\Rightarrow$ (ii): Let $x \in X$, $x \neq 0$. The neighborhoods at x are of the form $x + U$, where U is a neighborhood of 0. By the Hausdorff property, we can separate x and 0, hence we find two neighborhoods U and V of 0 such that $(x + U) \cap V = \emptyset$. We may suppose that $V = U_{F,\varepsilon} = \{y \in X \mid p(y) \leq \varepsilon, p \in F\}$ with a finite subset $F \subseteq P$ and $\varepsilon > 0$. Since $x \notin V$, there is some $p \in F$ such that $p(x) > \varepsilon > 0$.

(ii) $\Rightarrow$ (iii): We only have to note that $p(x) = 0$ for every $p \in P$, if and only if $x \in U_{F,\varepsilon}$ for every finite subset $F \subseteq P$ and every $\varepsilon > 0$.

(iii) $\Rightarrow$ (i): Let $x, y \in X$, $x \neq y$, hence $x - y \neq 0$ and there has to be some $U \in \mathcal{U}$ such that $x - y \notin U$. The map $(x, y) \mapsto x - y$ is continuous (being the composition of addition with scalar multiplication by -1 in the second component), thus, there are neighborhoods V and W of 0 such that $W - V \subseteq U$. Therefore $x - y \notin W - V$, which means $(x + V) \cap (y + W) = \emptyset$, i.e., we have found disjoint neighborhoods of x and y. □

4.12. Examples: 1) *Topology of pointwise convergence*: Let T be a set and X be a subspace of the $\mathbb{K}$-vector space of functions $T \to \mathbb{K}$, e.g., $T = [0, 1]$ and $X = C([0, 1])$. For every $t \in T$, consider the seminorm $p_t(x) = |x(t)|$ $(x \in X)$. The family $P = \{p_t \mid t \in T\}$ defines a locally convex Hausdorff topology on X.

2) *Topology of uniform convergence on compact sets*: Let T be a topological space and X be a subspace of $C(T, \mathbb{K}) = C(T)$. For every compact subset $K \subseteq T$, we put $p_K(x) := \sup_{t \in K} |x(t)|$ $(x \in X)$. The family of seminorms $P = \{p_K \mid K \subseteq T \text{ compact}\}$ generates a locally convex Hausdorff topology on X.

3) Let $\Omega \subseteq \mathbb{R}^n$ be open and $\mathscr{E}(\Omega)$ be the vector space $C^\infty(\Omega)$ equipped with the following seminorms: For $m \in \mathbb{N}_0$ and $K \subset \Omega$ compact, define

$$p_{K,m}(x) := \max_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq m} \sup_{t \in K} |\partial^\alpha x(t)| \qquad (x \in \mathscr{E}(\Omega)).$$

The family of seminorms $P = \{p_{K,m} \mid m \in \mathbb{N}_0, K \subset \Omega \text{ compact}\}$ generates a locally convex Hausdorff topology on $\mathscr{E}(\Omega)$.

4) The so-called *Schwartz space* $\mathscr{S}(\mathbb{R}^n)$ is the set of functions $\varphi \in C^\infty(\mathbb{R}^n)$ such that every derivative $\partial^\beta \varphi$, $\beta \in \mathbb{N}_0^n$, is *rapidly decreasing*, i.e., for every $\alpha \in \mathbb{N}_0^n$, the function $x \mapsto x^\alpha \partial^\beta \varphi(x)$ is bounded on $\mathbb{R}^n$. We define seminorms on $\mathscr{S}(\mathbb{R}^n)$ for any $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$ by

$$p_{\alpha,m}(\varphi) := \sup_{x \in \mathbb{R}^n} (1 + |x|^m) \, |\partial^\alpha \varphi(x)| \qquad (\varphi \in \mathscr{S}(\mathbb{R}^n)),$$

or occasionally also $p_{l,m}(\varphi) := \max_{|\alpha| \leq l} p_{\alpha,m}(\varphi)$ $(l, m \in \mathbb{N}_0)$. The corresponding families of seminorms define a locally convex Hausdorff topology on $\mathscr{S}(\mathbb{R}^n)$.

5) Let $\Omega \subseteq \mathbb{R}^n$ be open, $K \subset \Omega$ compact, and $\mathscr{D}_K(\Omega)$ denote the set of all functions $\varphi \in C^\infty(\Omega)$ with $\operatorname{supp}(\varphi) \subseteq K$.
(Recall that $\operatorname{supp}(\varphi)$ is the closure, in the trace topology of Ω, of the set $\{x \in \Omega \mid \varphi(x) \neq 0\}$. Equivalently, it is the complement, within Ω, of the largest open set, where φ vanishes. I suppose that the existence of nonzero functions $\varphi \in C^\infty(\Omega)$ with compact support has been shown in your second or third semester Analysis course; if not, a reference is [Wer18, Beispiel zwischen Definition V.1.9 und Lemma V.1.10]. Clearly, if $K^\circ = \emptyset$, then $\mathscr{D}_K(\Omega) = \{0\}$.)
We equip $\mathscr{D}_K(\Omega)$ with the structure of a locally convex Hausdorff space generated by the family of seminorms

$$p_m(\varphi) := \max_{|\alpha| \leq m} \sup_{x \in \Omega} |\partial^\alpha \varphi(x)| \qquad (m \in \mathbb{N}_0, \varphi \in \mathscr{D}_K(\Omega)).$$

6) Let $\Omega \subseteq \mathbb{R}^n$ be open. The space of *test functions* $\mathscr{D}(\Omega)$ is the set of all functions $\varphi \in C^\infty(\Omega)$ with compact support, i.e., $\mathscr{D}(\Omega) = \bigcup_{\substack{K \subset \Omega \\ K \text{ compact}}} \mathscr{D}_K(\Omega)$.

We could consider seminorms p_m ($m \in \mathbb{N}_0$) as above also on $\mathscr{D}(\Omega)$, the only difference is that the functions now do not have their support inside some fixed compact set K. But it turns out that the corresponding topology is not appropriate to reflect the structure of $\mathscr{D}(\Omega)$ as a union of the spaces $\mathscr{D}_K(\Omega)$ with their topology τ_K defined in the previous example. Such a structural feature is desirable to obtain a function space with good localization properties and convenient approximation procedures being concentrated on bounded regions. Therefore, we consider the set P of seminorms p on $\mathscr{D}(\Omega)$ such that the restriction $p|_{\mathscr{D}_K(\Omega)}$ is continuous with respect to τ_K for every compact set $K \subset \Omega$. The locally convex topology on $\mathscr{D}(\Omega)$ generated from P is a special case of the so-called inductive limit topologies from the general theory of topological vector spaces. It is also Hausdorff, which we will show later in the section on distribution theory.

7) If $(X, \|.\|)$ is a normed space, then the locally convex topology generated by $P = \{\|.\|\}$ is the norm topology on X.

8) Let X be a normed space with dual space X' (i.e., the space of all continuous linear functionals $X \to \mathbb{K}$). Then the family of seminorms $P = \{p_{x'} \mid x' \in X'\}$, where

$$p_{x'}(x) := |x'(x)| \quad (x \in X),$$

generate the *weak topology* $\sigma(X, X')$ on X, which is locally convex and Hausdorff (due to the theorem of Hahn-Banach).

9) Let X' be the dual space of the normed space X and consider for every $x \in X$ the following seminorm p_x on X':

$$p_x(x') := |x'(x)| \quad (x' \in X').$$

The set of seminorms $P = \{p_x \mid x \in X\}$ generates the *weak* topology* $\sigma(X', X)$ on X'. It coincides with the topology of pointwise convergence for the subspace of continuous linear functions $X \to \mathbb{K}$. In general, $\sigma(X', X)$ is different from the weak topology $\sigma(X', X'')$.

10) Let X and Y be normed spaces. Apart from the (operator) norm topology on $L(X, Y)$, the following two locally convex topologies are also of interest in operator theory: The *strong operator topology* is generated by the seminorms

$$p_x(T) := \|Tx\| \quad (T \in L(X, Y), x \in X),$$

and the *weak operator topology*, which is defined by the seminorms

$$p_{x,y'}(T) := |y'(Tx)| \quad (T \in L(X, Y), x \in X, y' \in Y').$$

Both are Hausdorff topologies, which in case of the weak operator topology follows by appealing to the Hahn-Banach theorem. Recall that we have used the notions of (sequential) convergence with respect to these topologies already implicitly while discussing the measurable functional calculus of self-adjoint operators on Hilbert spaces (with $X = Y = H$ and $y'(x)$ corresponding to an inner product).

11) The notion of weak topology in probability theory: Let Ω be a metric space and $M(\Omega)$ be the space of regular signed or complex Borel measures on Ω. For every f in the space of bounded continuous functions $C_b(\Omega)$ we define the seminorm

$$p_f(\mu) := |\int_\Omega f \, d\mu|$$

on $M(\Omega)$. The family of seminorms $P = \{p_f \mid f \in C_b(\Omega)\}$ generates a locally convex topology on $M(\Omega)$, which in case of compact Ω corresponds to the weak* topology due to the Riesz representation theorem (since we then have $C_b(\Omega) = C(\Omega)$ and $C(\Omega)' \cong M(\Omega)$). This “probabilistic weak topology” always has the Hausdorff property, the proof is based on the regularity of the measures (cf. [Els11, Kapitel VIII, §1, Satz 4.6]).

5. Continuous linear maps and functionals

Some of the locally convex spaces of functions have been introduced essentially in order to provide a setting for linear functionals on these spaces, which are considered as *generalized functions* or *distributions*. It is the property of local convexity that guarantees the existence of continuous linear functionals via the Hahn-Banach theorem. Moreover, we have seen in the definition of the test function space $\mathscr{D}(\Omega)$ that seminorms with continuous restrictions to any $\mathscr{D}_K(\Omega)$ are used to define the locally convex topology on $\mathscr{D}(\Omega)$.

5.1. Lemma: Let (X, τ) be a locally convex vector space and P be a family of seminorms generating the topology.

(a) If $q\colon X \to [0, \infty[$ is a seminorm, then the following are equivalent:

  (i) q is continuous,

  (ii) q is continuous at 0,

  (iii) the set $U_{q,1} := \{x \in X \mid q(x) \leq 1\}$ is a neighborhood of 0.

(b) Every $p \in P$ is continuous.

(c) A seminorm q on X is continuous, if and only if there is a finite subset $F \subseteq P$ and a real number $M > 0$ such that

$$\forall x \in X\colon \quad q(x) \leq M \max_{p \in F} p(x).$$

Proof: (a): The implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) are immediate. Suppose that (iii) holds. We show that then (i) follows, which completes the proof of (a): Let $x \in X$ and $\varepsilon > 0$; we set $U := \varepsilon \cdot U_{q,1} = \{y \in X \mid q(y) \leq \varepsilon\}$ and note that $x + U$ is a neighborhood of x; if $y \in U$, then[1]

$$|q(x + y) - q(x)| \leq q((x + y) - x) = q(y) \leq \varepsilon,$$

thus, $q(x + U) \subseteq [q(x) - \varepsilon, q(x) + \varepsilon]$, proving continuity of q at x.

(b): This follows from (a),(iii), since by definition of τ, the sets $\{x \in X \mid p(x) \leq 1\}$ with $p \in P$ are neighborhoods of 0.

[1] Here, we simply use the “reverse triangle inequality”: $q(w) = q((w - z) + z) \leq q(w - z) + q(z)$ and $q(z) = q((z - w) + w) \leq q(z - w) + q(w) = q(w - z) + q(w)$ implies $|q(w) - q(z)| \leq q(w - z)$.

(c): By definition of τ and (a), (iii), the continuity of q is equivalent to the existence of a finite set $F \subseteq P$ and $\varepsilon > 0$ such that $U_{F,\varepsilon} \subseteq U_{q,1}$. Recalling $U_{F,\varepsilon} = \{x \in X \mid \forall p \in F\colon p(x) \leq \varepsilon\}$, the inclusion relation $U_{F,\varepsilon} \subseteq U_{q,1}$ is equivalent to

$$(*) \qquad \forall x \in X\colon \quad \frac{1}{\varepsilon} \max_{p \in F} p(x) \leq 1 \quad \Rightarrow \quad q(x) \leq 1.$$

If the inequality stated in (c) holds, then $(*)$ follows upon putting $\varepsilon := 1/M$. Conversely, suppose $(*)$ holds and let $x \in X$. If we had $q(x) > \max_{p \in F} p(x)/\varepsilon$, then we could choose $\mu > 0$ with $q(x) > \mu > \max_{p \in F} p(x)/\varepsilon$ and obtain a contradiction to $(*)$ for $y := x/\mu$, since $q(y) = q(x)/\mu > 1$ and $\max_{p \in F} p(y)/\varepsilon = \max_{p \in F} p(x)/(\varepsilon\mu) < 1$. Thus, the inequality stated in (c) holds with $M := 1/\varepsilon$. □

5.2. Corollary: Let (X, τ) be a locally convex vector space and P be a family a seminorms generating the topology. If $Q \supseteq P$ is a family of τ-continuous (!) seminorms on X, then Q also generates the topology τ.

Proof: Clearly, the topology generated by $Q \supseteq P$ is finer than τ, but by the above lemma and the continuity of each $q \in Q$ it is also coarser than τ. □

We can now provide simple criteria for the continuity of linear maps between locally convex vector spaces.

5.3. Theorem: Let (X, τ_P) and (Y, τ_Q) be two locally convex vector spaces, where τ_P is generated by the family P of seminorms on X and τ_Q by the family of seminorms Q on Y. Let $T\colon X \to Y$ be a linear map, then the following are equivalent:

(i) T is continuous,

(ii) T is continuous at 0,

(iii) if q is a continuous seminorm on Y, then $q \circ T$ is a continuous seminorm on X,

(iv) for every $q \in Q$ there is a finite subset $F \subseteq P$ and a real number $M > 0$ such that

$$\forall x \in X\colon \quad q(Tx) \leq M \max_{p \in F} p(x).$$

Proof: (i) $\Rightarrow$ (ii) is clear and (ii) $\Rightarrow$ (i) is easy: Let $x \in X$; a neighborhood of Tx is of the form $Tx + V$, where V is a τ_Q-neighborhood of 0 in Y; choose a τ_P-neighborhood U of 0 in X such that $T(U) \subseteq V$, then clearly $T(x + U) = Tx + T(U) \subseteq Tx + V$.

(ii) $\Rightarrow$ (iii): The composition $q \circ T$ is continuous at 0 and clearly defines seminorm on X. Hence, by (a) in the lemma above, $q \circ T$ is a continuous seminorm.

(iii) $\Rightarrow$ (iv): This follows from (b), applied to $q \in Q$, and (c), applied to $q \circ T$, in the lemma shown previously.

(iv) $\Rightarrow$ (ii): Let V be a τ_Q-neighborhood of 0 in Y. We may suppose that V is given by finitely many seminorms $q_1, \ldots, q_n \in Q$ and $\varepsilon > 0$ in the form $V = \{y \in Y \mid q_j(y) \leq \varepsilon \ (j = 1, \ldots, n)\}$. For every $j = 1, \ldots, n$, choose $F_j \subseteq P$ and $M_j > 0$ according to (iv) with q_j in place of q. We put $F := \bigcup_{j=1}^n F_j$, $M := \max_{j=1}^n M_j$, and consider the τ_P-neighborhood $U := U_{F,\varepsilon/M}$ of 0 in X. If $x \in U$ and $j \in \{1, \dots, n\}$, then

$$q_j(Tx) \le M_j \max_{p \in F_j} p(x) \le M \max_{p \in F} p(x) \le M \frac{\varepsilon}{M} = \varepsilon,$$

hence $Tx \in V$. Therefore, we have shown that $T(U) \subseteq V$, i.e., T is continuous at 0. □

The most important special case $Y = \mathbb{K}$ (with the Euclidean topology, generated from the single norm $|.|$) deserves an explicit statement.

5.4. Corollary: Let X be a locally convex vector space and P be a family a seminorms generating the topology. A linear functional l on X is continuous, if and only if there are finitely many seminorms $p_1, \dots, p_m \in P$ and a constant $M > 0$ such that

$$\forall x \in X\colon \quad |l(x)| \le M \max_{j=1,\dots,m} p_j(x).$$

5.5. Definition: (i) Let (X, τ) be a locally convex vector space. Then its *dual space* X', or $(X_\tau)'$, is the vector space of continuous linear functionals on X.

(ii) Let (X, τ_1) and (Y, τ_2) be locally convex vector spaces, then $L(X, Y)$, or $L(X_{\tau_1}, Y_{\tau_2})$, denotes the vector space of continuous linear maps $X \to Y$.

Note that we have $X' = L(X, \mathbb{K})$ and the fact that $L(X, Y)$ is a vector space follows from Theorem 5.3.

Recall that a topology τ_1 on a set X is finer than τ_2, if $\tau_2 \subseteq \tau_1$, or, equivalently, the identity map on X is continuous as a map $\mathrm{id}\colon (X, \tau_1) \to (X, \tau_2)$. If τ_1 and τ_2 are locally convex topologies on the vector space X, then the latter is equivalent to $\mathrm{id} \in L(X_{\tau_1}, X_{\tau_2})$. The topologies are equal, if both $\mathrm{id} \in L(X_{\tau_1}, X_{\tau_2})$ and $\mathrm{id} \in L(X_{\tau_2}, X_{\tau_1})$ hold.

5.6. Examples: 1) Let $(X, \|.\|)$ be a normed space and τ denote the topology induced by the norm, then τ is finer than the weak topology $\sigma(X, X')$, since for any $x' \in X'$ we have

$$\forall x \in X\colon \quad p_{x'}(x) = |x'(x)| \le \|x'\| \|x\|,$$

which shows that $\mathrm{id} \in L(X_\tau, X_{\sigma(X,X')})$.

2) Let $X = C(\mathbb{R}^n)$ and denote by τ_1 the topology of uniform convergence on compact sets, by τ_2 the topology of pointwise convergence. Then τ_1 is finer than τ_2, since for any $t \in \mathbb{R}^n$ we may take $K = \{t\}$ as compact set and obtain

$$\forall x \in C(\mathbb{R}^n)\colon \quad p_t(x) = |x(t)| = \sup_{t \in K} |x(t)| = p_K(x).$$

We briefly study convergence of nets in locally convex vector spaces. Recall that in general, sequences are not sufficiently powerful to characterize continuity of maps or closures of subsets (cf. Example 5.8 below) in topological spaces that are non-metrizable. (More precisely, the “barrier” are the first countable, or AA1, spaces).

5.7. Proposition: Let the locally convex topology τ on X be generated by the family of seminorms P. Then a net $(x_j)_{j\in I}$ converges to x in X with respect to τ, if and only if $\lim p(x_j - x) = 0$ for every $p \in P$.

Proof: The net (x_j) converges to x, if and only if the net $(x_j - x)$ converges to 0. Thus, we may suppose that $x = 0$.

If (x_j) converges to 0, then $\lim p(x_j) = 0$ for every $p \in P$ due to continuity of p (Lemma 5.1(b)).

Now suppose that $\lim p(x_j) = 0$ for every $p \in P$ and let $U = U_{F,\varepsilon}$ be a typical neighborhood of 0 with $F \subseteq P$ finite and $\varepsilon > 0$. In particular, $\lim p(x_j) = 0$ for every $p \in F$. Choose $j_p \in I$ such that $p(x_j) \leq \varepsilon$ for every $j \geq j_p$. Since $(I, \leq)$ is a directed set (and F is finite), there is some $j' \in I$ such that $j' \geq j_p$ for every $p \in F$. We obtain $p(x_j) \leq \varepsilon$ for every $j \geq j'$, hence $x_j \in U$ if $j \geq j'$. Thus, we have shown $x_j \to 0$ in (X, τ). □

5.8. Example: Let $X = l^2$, equipped with its weak topology $\sigma := \sigma(l^2, l^2)$, and consider the subset

$$A := \{e_m + m\, e_n \mid 1 \leq m < n\} \subseteq l^2,$$

where $e_k \in l^2$ denotes the vector with k-th component 1 and all other components 0.

Claim 1: $0 \in \overline{A}$ (weak closure).

Let U be a typical σ-neighborhood of 0, i.e., $U = \{x \in l^2 \mid |\langle x, y_j\rangle| \leq \varepsilon \ (j = 1, \ldots, r)\}$ with $\varepsilon > 0$ and $y_1, \ldots, y_r \in l^2$. Choose m such that $|y_j(m)| \leq \varepsilon/2$ for $j = 1, \ldots, r$. Then choose $n > m$ such that $|y_j(n)| \leq \varepsilon/(2m)$ for $j = 1, \ldots, r$. We obtain

$$|\langle e_m + m\, e_n, y_j\rangle| \leq |y_j(m)| + m\, |y_j(n)| \leq \frac{\varepsilon}{2} + m\, \frac{\varepsilon}{2m} = \varepsilon \quad (j = 1, \ldots, r),$$

i.e., $e_m + m\, e_n \in A \cap U$. We have therefore shown that $V \cap A \neq \emptyset$ for every neighborhood V of 0.

Claim 2: There is no sequence in A that converges weakly to 0.

Suppose $(e_{m_k} + m_k\, e_{n_k})_{k\in\mathbb{N}}$ is a sequence in A that converges weakly to 0. By Proposition 5.7, we conclude that

$$\forall y \in l^2: \qquad \lim_{k\to\infty} |y(m_k) + m_k\, y(n_k)| = \lim_{k\to\infty} |\langle e_{m_k} + m_k\, e_{n_k}, y\rangle| = 0.$$

If $(m_k)_{k\in\mathbb{N}}$ is bounded, there is some $q \in \mathbb{N}$ such that $m_k = q$ infinitely often. Choosing $y = e_q$ and recalling $n_k > m_k$, then produces the contradiction that $y(m_k) + m_k\, y(n_k) = 1$ infinitely often and converges to 0 as $k \to \infty$.
If $(m_k)_{k\in\mathbb{N}}$ is unbounded, we may suppose that both (m_k) and (n_k) are strictly increasing. Let $y \in l^2$ be given by $y(j) := 1/k$, if $j = n_k$, and $y(j) := 0$ otherwise. Then $y(m_k) + m_k\, y(n_k) \geq m_k/k \geq 1$ for every $k \in \mathbb{N}$, which also contradicts $y(m_k) + m_k\, y(n_k) \to 0$.

Remark: We can obtain a net in A converging weakly to 0 as follows. The set of all typical neighborhoods $U_{F,\varepsilon}$ of 0 becomes a directed set, if we introduce the relation $U_{F_1,\varepsilon_1} \leq U_{F_2,\varepsilon_2}$ meaning $U_{F_1,\varepsilon_1} \supseteq U_{F_2,\varepsilon_2}$. To the index set element $U_{F,\varepsilon}$ we assign the element $a_{U_{F,\varepsilon}} := e_m + m\, e_n$ constructed as in the proof of Claim 1 for $F = \{y_1, \ldots, y_r\}$ and $\varepsilon > 0$.

5.9. Examples: 1) The dual space of the Schwartz space $\mathscr{S}(\mathbb{R}^n)$, Example 4.12.4), is denoted by $\mathscr{S}'(\mathbb{R}^n)$ and called the space of *temperate (or tempered) distributions.* The locally convex topology on $\mathscr{S}(\mathbb{R}^n)$ is generated by the seminorms $p_{\alpha,m}(\varphi) := \sup_{x\in\mathbb{R}^n}(1+|x|^m)\,|\partial^\alpha\varphi(x)|$ ($m \in \mathbb{N}_0$, $\alpha \in \mathbb{N}_0^n$). We show that every $f \in L^p(\mathbb{R}^n)$ $(1 \le p \le \infty)$ provides an element in $\mathscr{S}'(\mathbb{R}^n)$, in the sense that the map $T_f\colon \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$,

$$T_f(\varphi) := \int\limits_{\mathbb{R}^n} f(x)\,\varphi(x)\,dx \quad (\varphi \in \mathscr{S}(\mathbb{R}^n)),$$

defines a continuous linear functional. Recall Hölder's inequality, which states that for any $g \in L^q(\mathbb{R}^n)$, $\frac{1}{p}+\frac{1}{q}=1$ with the convention $\frac{1}{\infty}=0$, we have $\|fg\|_{L^1} \le \|f\|_{L^p}\|g\|_{L^q}$. Since $\mathscr{S}(\mathbb{R}^n) \subseteq L^q(\mathbb{R}^n)$ $(1 \le q \le \infty)$, we obtain for any $\varphi \in \mathscr{S}(\mathbb{R}^n)$,

$$|T_f(\varphi)| \le \int\limits_{\mathbb{R}^n} |f(x)\,\varphi(x)|\,dx = \|f\varphi\|_{L^1} \le \|f\|_{L^p}\|\varphi\|_{L^q}.$$

If $p = 1$, then $q = \infty$ and we have $2\|\varphi\|_{L^\infty} = p_{0,0}(\varphi)$, thus, $|T_f(\varphi)| \le \|f\|_{L^1}\,p_{0,0}(\varphi)$ proves the continuity of T_f.
If $p > 1$, then $1 \le q < \infty$ and

$$\begin{aligned}\|\varphi\|_{L^q} = \left(\int\limits_{\mathbb{R}^n} |\varphi(x)|^q\,dx\right)^{1/q} &= \left(\int\limits_{\mathbb{R}^n} \frac{1}{(1+|x|^{n+1})^q}\underbrace{\big|(1+|x|^{n+1})\varphi(x)\big|^q}_{\le p_{0,n+1}(\varphi)^q}\,dx\right)^{1/q}\\ &\le \left\|\frac{1}{1+|.|^{n+1}}\right\|_{L^q}\cdot p_{0,n+1}(\varphi) =: M_q\cdot p_{0,n+1}(\varphi),\end{aligned}$$

which implies $|T_f(\varphi)| \le \|f\|_{L^p} M_q\,p_{0,n+1}(\varphi)$ and proves the continuity of T_f.

2) Let X be a normed space, τ denote the topology induced by the norm on X, and $\sigma := \sigma(X, X')$. As noted above, τ is finer than σ. Thus, if $\mu\colon X \to \mathbb{K}$ is a weakly continuous linear functional, then μ is also norm continuous, i.e., we have $(X_\sigma)' \subseteq (X_\tau)' = X'$.
If $x' \in X'$, then $|x'(x)| = p_{x'}(x)$ for every $x \in X$, therefore x' is also weakly continuous by Corollary 5.4; hence, also $X' = (X_\tau)' \subseteq (X_\sigma)'$ holds. To summarize, we have shown that

$$(X_{\sigma(X,X')})' = X'.$$

3) Let X and Y be normed spaces and consider the strong operator topology on $L(X,Y)$, described in Example 4.12.10) and generated by the seminorms $p_x(T) := \|Tx\|$ ($T \in L(X,Y)$, $x \in X$). We denote this locally convex space by $L_{\mathrm{st}}(X,Y)$ and claim that a linear functional Φ on $L(X,Y)$ belongs to $L_{\mathrm{st}}(X,Y)'$, if and only if Φ is of the following form: There are $n \in \mathbb{N}$, $x_1,\ldots,x_n \in X$, and $\mu_1,\ldots,\mu_n \in Y'$ such that

$$\Phi(T) = \sum_{j=1}^{n} \mu_j(Tx_j) \quad (T \in L(X,Y)).$$

Clearly, any Φ of the above form is continuous with respect to the strong operator topology by Corollary 5.4, since $|\Phi(T)| \le n\max_{j=1}^n \|\mu_j\|\|Tx_j\| \le (n\max_{j=1}^n \|\mu_j\|)\cdot\max_{l=1}^n p_{x_l}(T)$.

Suppose $\Phi \in L_{\rm st}(X,Y)'$, then Corollary 5.4 implies that there are $n \in \mathbb{N}$, $M > 0$, and $x_1, \ldots, x_n \in X$ such that

$$(*) \qquad |\Phi(T)| \leq M \max_{1\leq j\leq n} p_{x_j}(T) = M \max_{1\leq j\leq n} \|Tx_j\|.$$

Let $l^\infty_n(Y)$ denote the $\mathbb{K}$-vector space Y^n equipped with the norm $\|(y_1,\ldots,y_n)\|_\infty := \max_{j=1}^n \|y_j\|$. We claim that any norm continuous linear functional $\mu \in l^\infty_n(Y)'$ is given in the form

$$(**) \qquad \mu((y_1,\ldots,y_n)) = \sum_{j=1}^n \mu_j(y_j)$$

with certain $\mu_1,\ldots,\mu_n \in Y'$: In fact, define $\mu_j(z) := \mu((0,\ldots,z,0,\ldots))$ (here, z is in the jth component; $j = 1,\ldots,n$), then $\mu_j \colon Y \to \mathbb{K}$ is linear and continuous, since $|\mu_j(z)| \leq \|\mu\| \|(0,\ldots,z,\ldots,0)\|_\infty = \|\mu\| \|z\|$; finally, $(y_1,\ldots,y_n) = (y_1,0,\ldots,0) + \ldots + (0,\ldots,0,y_n)$ yields $\mu((y_1,\ldots,y_n)) = \sum_{j=1}^n \mu_j(y_j)$.

Consider the subspace $V := \{(Tx_1,\ldots,Tx_n) \in Y^n \mid T \in L(X,Y)\}$ of $l^\infty_n(Y)$. Note that by $(*)$ we have $\Phi(T) = 0$, if $(Tx_1,\ldots,Tx_n) = (0,\ldots,0)$, which shows that the linear functional $\nu \colon V \to \mathbb{K}$, $\nu((Tx_1,\ldots,Tx_n)) := \Phi(T)$ is well-defined. Moreover, $(*)$ also implies that $|\nu((Tx_1,\ldots,Tx_n))| \leq M\|(Tx_1,\ldots,Tx_n)\|_\infty$, i.e., ν is a norm continuous linear functional on the subspace $V \subseteq l^\infty_n(Y)$. By the Hahn-Banach theorem (for normed spaces), there is an extension $\mu \in l^\infty_n(Y)'$ of ν. Let μ be given by $\mu_1,\ldots,\mu_n \in Y'$ according to $(**)$. Then we obtain, putting $(y_1,\ldots,y_n) = (Tx_1,\ldots,Tx_n)$ with $T \in L(X,Y)$,

$$\Phi(T) = \nu((Tx_1,\ldots,Tx_n)) = \mu((Tx_1,\ldots,Tx_n)) = \sum_{j=1}^n \mu_j(Tx_j).$$

5.10. Hahn-Banach theorem(s) for locally convex vector spaces: Here, I suppose that you have already seen at least a proof of the extension theorem for normed spaces. Most likely, it was based on a so-called linear algebraic version as lemma, involving the crucial "lemma of Zorn argument" and an upper bound of the linear functional in terms of a seminorm or a convex function or a sublinear function (e.g., as in [Con10, Chapter III, Section 6] or [Tes14, Theorem 4.11] or [Wer18, Sätze III.1.2, III.1.4 und Lemma III.1.3]; note that a seminorm is a convex function and also an example of a sublinear function). Therefore, we "recall" the following statement without repeating a proof.

Linear algebraic version of Hahn-Banach's extension theorem: Let V be a subspace of the $\mathbb{K}$-vector space X and $p \colon X \to \mathbb{R}$ be sublinear, i.e., $p(\lambda x) = \lambda p(x)$, if $\lambda \geq 0$, and $p(x+y) \leq p(x) + p(y)$. If $l \colon V \to \mathbb{K}$ is a linear functional satisfying $\operatorname{Re} l(x) \leq p(x)$ for every $x \in V$, then there is a linear functional $L \colon X \to \mathbb{K}$ such that $L(x) = l(x)$, if $x \in V$, and $\operatorname{Re} L(x) \leq p(x)$ for every $x \in X$.

Extension theorem: Let V be a subspace of the locally convex vector space X and $l \in V'$. Then there exists an extension $L \in X'$ of l.

Proof: Suppose that the locally convex topology on X is generated by the family of seminorms P. Then the subspace topology on V is generated by the set of restrictions $\{p|_V \mid p \in P\}$ as family of seminorms on V (since the basic neighborhoods of 0 in V are just obtained from those in X intersected with V). By continuity of l, there are $M > 0$ and $p_1, \dots p_m \in P$ such that $|l(x)| \leq M \max_{j=1}^m p_j(x)$ for every $x \in V$. Define $p\colon X \to [0, \infty[$ by $p(x) := M \max_{j=1}^m p_j(x)$ ($x \in X$), then p is a continuous seminorm on X and $\operatorname{Re} l(x) \leq |l(x)| \leq p(x)$ for every $x \in V$. By the linear algebraic extension theorem, there is an extension L of l to X satisfying $\operatorname{Re} L(x) \leq p(x)$ for every $x \in X$. Since p is a seminorm we have $p(\lambda x) = p(x)$ for every $\lambda \in \mathbb{K}$ with $|\lambda| = 1$, hence we may conclude that $|L(x)| \leq p(x)$ holds for every $x \in X$. This shows that $L \in X'$ (strictly speaking, upon calling on Lemma 5.1(c) and Corollary 5.4). □

We turn now to so-called *geometric versions* of the Hahn-Banach theorem, which state properties on separation of convex subsets by values of linear functionals.

Lemma on Minkowski functionals: Let W be a convex neighborhood of 0 in the locally convex vector space X. Then the *Minkowski functional* $p_W\colon X \to [0, \infty]$,

$$p_W(x) := \inf\{\alpha > 0 \mid x \in \alpha W\},$$

has finite values and defines a sublinear function on X. If W is an absolutely convex neighborhood of 0, then p_W is a continuous seminorm on X.

Proof: Since W is absorbing[2], we have $p_W(x) < \infty$ for every $x \in X$. From the definition of p_W, we immediately obtain $p_W(\lambda x) = \lambda\, p_W(x)$, if $\lambda \geq 0$.

We show subadditivity of p_W, i.e., that $p_W(x + y) \leq p_W(x) + p_W(y)$ holds for all $x, y \in X$: Let $x, y \in X$ and $\varepsilon > 0$ be arbitrary. By definition of p_W as an infimum, there are $\lambda, \mu > 0$ such that $\lambda < p_W(x) + \frac{\varepsilon}{2}$, $\mu < p_W(y) + \frac{\varepsilon}{2}$ and $\frac{x}{\lambda} \in W$, $\frac{y}{\mu} \in W$. By convexity of W,

$$\frac{x+y}{\lambda+\mu} = \frac{\lambda}{\lambda+\mu} \cdot \frac{x}{\lambda} + \frac{\mu}{\lambda+\mu} \cdot \frac{y}{\mu} \in W,$$

which implies $p_W(x + y) \leq \lambda + \mu < p_W(x) + p_W(y) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we obtain $p_W(x + y) \leq p_W(x) + p_W(y)$.

If, in addition, W is balanced, then $\lambda W = W$ holds whenever $|\lambda| = 1$, thus $p_W(\lambda x) = p_{\lambda W}(\lambda x) = p_W(x)$ in this case. Therefore, we obtain for any $\lambda \neq 0$,

$$p_W(\lambda x) = p_W(\frac{\lambda}{|\lambda|}|\lambda| x) = p_W(|\lambda| x) = |\lambda|\, p_W(x) \quad (x \in X),$$

and may conclude that p_W is a seminorm. The continuity of p_W on X now follows from Lemma 5.1(a)(iii), since $\{x \in X \mid p_W(x) \leq 1\} = W$ is a neighborhood of 0. □

Separation lemma: Let V be a nonempty convex open subset of the locally convex vector space X. If $0 \notin V$, then there exists $x' \in X'$ such that $\operatorname{Re} x'(x) < 0$ for every $x \in V$.

Proof: Let $x_0 \in V$ and $U := -x_0 + V$. Then U is convex, open, $0 \in U$, and $-x_0 \notin U$. There is an absolutely convex neighborhood W of 0 such that $W \subseteq U$. By the lemma on Minkowski functionals, we have that p_U is sublinear and p_W is a continuous seminorm.

[2]Recall: Any neighborhood of 0 is absorbing, since it contains an absorbing basic neighborhood defined by finitely many seminorms.

Let $Y := \mathbb{R}\text{-span}\{-x_0\}$ and define the $\mathbb{R}$-linear functional $l\colon Y \to \mathbb{R}$ by $l(t(-x_0)) := t\, p_U(-x_0)$. We claim that $l(y) \leq p_U(y)$ for every $y \in Y$: If $t > 0$, then $l(t(-x_0)) = tp_U(-x_0) = p_U(t(-x_0))$; if $t \leq 0$, then $l(t(-x_0)) = t\, p_U(-x_0) \leq 0 \leq p_U(t(-x_0))$. By real linear algebraic extension, there is an $\mathbb{R}$-linear extension $L\colon X \to \mathbb{R}$ such that $L(x) \leq p_U(x)$ for every $x \in X$. By $W \subseteq U$ we have also $L(x) \leq p_U(x) \leq p_W(x)$, i.e., $|L(x)| \leq p_W(x)$ for every $x \in X$, since p_W is a seminorm. Thus, L is a continuous $\mathbb{R}$-linear functional on X, since p_W is a continuous seminorm (again employing Lemma 5.1(c) and Corollary 5.4)).

Let $x \in V = x_0 + U$, say $x = x_0 + u$ with $u \in U$. We claim that $p_U(u) < 1$: In fact, if $p_U(u) \geq 1$, then $\frac{u}{t} \in X \backslash U$ for every $0 < t < 1$; since $X \backslash U$ is closed, $u = \lim_{t<1, t\to 1} \frac{u}{t} \in X \backslash U$, a contradiction. Furthermore, note that $p_U(-x_0) \geq 1$, since $-x_0 \notin U$. Therefore, we obtain

$$L(x) = L(u) + L(x_0) = L(u) + l(x_0) \leq p_U(u) + l(x_0) = p_U(u) - p_U(-x_0) < 0.$$

Finally, in the case of a complex vector space we obtain a $\mathbb{C}$-linear continuous functional $x' \in X'$ with $\operatorname{Re} x' = L$ by setting $x'(x) := L(x) - iL(ix)$ (direct computation shows $x'((a+ib)x) = \ldots = (a+ib)x'(x)$), which clearly satisfies $\operatorname{Re} x'(v) = L(v) < 0$ for every $v \in V$. □

Separation theorem I: Let V_1 and V_2 be convex subsets of the locally convex vector space X. If V_1 is open and $V_1 \cap V_2 = \emptyset$, then there exists $x' \in X'$ such that

$$\forall v_1 \in V_1, \forall v_2 \in V_2\colon \quad \operatorname{Re} x'(v_1) < \operatorname{Re} x'(v_2).$$

Proof: Let $V := V_1 - V_2$, then V is convex (a convex combination of differences is a difference of convex combinations) and open (write $V = \bigcup_{x \in V_2}(V_1 - x)$ as union of open subsets; translation by $-x$ is a homeomorphism!). Since $V_1 \cap V_2 = \emptyset$, we have $0 \notin V$. By the lemma we may find $x' \in X'$ such that $\operatorname{Re} x'(v_1) - \operatorname{Re} x'(v_2) = \operatorname{Re} x'(v_1 - v_2) < 0$ for all $v_1 \in V_1$ and $v_2 \in V_2$. □

Separation theorem II: Let V be a closed convex subset of the locally convex vector space X. If $x \in X \setminus V$, then there exist $x' \in X'$ and $\varepsilon > 0$ such that

$$\forall v \in V\colon \quad \operatorname{Re} x'(x) + \varepsilon \leq \operatorname{Re} x'(v).$$

If, in addition, V is absolutely convex, then there exist $y' \in X'$ and $\varepsilon > 0$ such that

$$\forall v \in V\colon \quad |y'(v)| + \varepsilon \leq \operatorname{Re} y'(x).$$

Proof: Choose an absolutely convex open neighborhood U of 0 such that $(x + U) \cap V = \emptyset$. By the separation theorem I there exists $x' \in X'$ such that

$$(*) \qquad \forall u \in U, \forall v \in V\colon \quad \operatorname{Re} x'(x) + \operatorname{Re} x'(u) < \operatorname{Re} x'(v).$$

Since U is absorbing, there is some $u_0 \in U$ with $x'(u_0) \neq 0$ (for otherwise, we had $x' = 0$, which contradicts the above inequality). Being also a balanced subset, U contains every multiple λu_0 with $|\lambda| \leq 1$, hence there is some $u_1 \in U$ with $\operatorname{Re} x'(u_1) > 0$. Therefore $\varepsilon := \sup\{\operatorname{Re} x'(u) \mid u \in U\}$ is positive. Picking an arbitrary $v_0 \in V$, the inequality in $(*)$ implies $\operatorname{Re} x'(u) \leq \operatorname{Re} x'(v_0) - \operatorname{Re} x'(x)$ for every $u \in U$, hence $\operatorname{Re} x'(u)$ is bounded above and $\varepsilon < \infty$. We have thus shown that $\operatorname{Re} x'(x) + \varepsilon \leq \operatorname{Re} x'(v)$ for every $v \in V$, which proves the first part of the statement.

Note that, running through the reasoning above with $y' := -x'$, we obtain $\operatorname{Re} y'(x) - \varepsilon \geq \operatorname{Re} y'(v)$ for every $v \in V$. If V is absolutely convex, then $V = \lambda V$ for every $\lambda \in \mathbb{K}$ with $|\lambda| = 1$, hence we may conclude that $\operatorname{Re} y'(x) - \varepsilon \geq |y'(v)|$ holds for every $v \in V$. □

Corollary: If X is a Hausdorff locally convex vector space, then X' separates points, i.e., for any $x, y \in X$ with $x \neq y$, there is $x' \in X'$ such that $x'(x) \neq x'(y)$.

Proof: Apply the separation theorem II to $V := \{y\}$, which is convex and also closed (by the Hausdorff property). □

Remark: The proofs of the Hahn-Banach theorems of extension and separation II do rely in a crucial way on basic properties of locally convex topological vector spaces, since they employ seminorms or nontrivial convex open neighborhoods of 0. In general topological vector spaces the extension theorem may fail and it may happen that convex open 0-neighborhoods are trivial. Both is illustrated by the following example: Consider $L^p([0,1])$, but with $0 < p < 1$. Then it is easy to see that $d(f,g) := \int_0^1 |f(t) - g(t)|^p \, dt$ defines a metric on $L^p([0,1])$, which provides the topology of a topological vector space. However, one can show the following properties (cf. [Wer18, Aufgabe VIII.8.3] or [Con10, Chapter IV, Example 3.16]):

- If U is a convex neighborhood of 0, then $U = L^p([0,1])$.
- There is no nonzero continuous linear functional on $L^p([0,1])$, i.e., $L^p([0,1])' = \{0\}$.

The separation theorem I does hold in topological vector spaces, though the one convex subset which is supposed to be open might be a trivial subset (cf. [Con10, Chapter IV, Theorem 3.7.; see also the paragraph preceding Example 3.16]).

5.11. An application of Hahn-Banach separation—the Krein-Milman theorem: An *extreme point* x_0 in a convex set K in a $\mathbb{K}$-vector space X is a point $x_0 \in K$ that cannot be part of a non-degenerate line segment inside K, i.e.,

$$\forall x_1, x_2 \in K, \forall \lambda \in \mathbb{R}, 0 < \lambda < 1\colon \quad x_0 = \lambda x_1 + (1-\lambda) x_2 \quad \Rightarrow \quad x_1 = x_2 = x_0.$$

The Krein-Milman theorem states that a nonempty *compact* convex subset of a locally convex space is the closed convex hull of its extreme points. Before stating and proving this theorem we clarify or introduce some of the relevant notation.

If B is a subset of a vector space let $\operatorname{co} B$ denote its *convex hull* (i.e., the intersection of all convex subsets containing B), and, in case of a topological vector space, let $\overline{\operatorname{co}}\, B$ denote the closure of the convex hull. The latter is a convex set: Let $x, y \in \overline{\operatorname{co}}\, B$ and $0 \leq \lambda \leq 1$; there are nets (x_j) and (y_j) in $\operatorname{co} B$ with $x = \lim x_j$ and $y = \lim y_j$; by continuity of the vector space operations, we obtain $\lambda x + (1-\lambda) y = \lim(\lambda x_j + (1-\lambda) y_j) \in \overline{\operatorname{co}}\, B$. Therefore, $\overline{\operatorname{co}}\, B$ is the smallest closed convex set containing B and is called the *closed convex hull of* B.

As a generalization of the notion of extreme points, a nonempty subset F of a convex set K is called a *face of* K, if F is convex and satisfies

$$\forall x_1, x_2 \in K, \forall \lambda \in \mathbb{R}, 0 < \lambda < 1\colon \quad \lambda x_1 + (1-\lambda) x_2 \in F \quad \Rightarrow \quad x_1, x_2 \in F.$$

Thus, x_0 is an extreme point of K, if and only if $F := \{x_0\}$ is a face of K. We denote by $\operatorname{ex} K$ the set of extreme points of K.

As an interesting example we may mention, without proof, the case $X = C(\Omega)' \cong M(\Omega)$, where Ω is a compact metric space (and $M(\Omega)$ denotes the space of signed or complex regular Borel measures, cf. the Riesz representation theorem 0.17). Let K denote the subset of all Borel probability measures on Ω. Then clearly $K \neq \emptyset$ and K is convex. Particular elements in K are the Dirac measures δ_ω concentrated at $\omega \in \Omega$ and it turns out that (cf. [Wer18, Beispiel (f) in VIII.4])

$$\operatorname{ex} K = \{\delta_\omega \mid \omega \in \Omega\}.$$

Lemma: Let K be a non-empty compact convex subset of a Hausdorff locally convex vector space X and $\rho \in X'$. Let $c := \max\{\operatorname{Re}\rho(x) \mid x \in K\}$ (making use of the continuity of $\operatorname{Re}\rho\colon X \to \mathbb{R}$ and of the compactness of K), then $F := \{x \in K \mid \operatorname{Re}\rho(x) = c\}$ is a compact face of K.

Proof: Clearly, F is nonempty, convex, and closed, hence also compact. If $x_1, x_2 \in K$, $0 < \lambda < 1$, and $\lambda x_1 + (1-\lambda)x_2 \in F$, then $\operatorname{Re}\rho(x_j) \leq c$ $(j = 1, 2)$ and $c = \operatorname{Re}\rho(\lambda x_1 + (1-\lambda)x_2) = \lambda \operatorname{Re}\rho(x_1) + (1-\lambda)\operatorname{Re}\rho(x_2)$, hence $\operatorname{Re}\rho(x_1) = \operatorname{Re}\rho(x_2) = c$, which implies $x_1, x_2 \in F$. □

Theorem (Krein-Milman): If K is a nonempty compact convex subset of a Hausdorff locally convex vector space X, then $\operatorname{ex} K \neq \emptyset$ and $K = \overline{\operatorname{co}}\,(\operatorname{ex} K)$.

Proof: Let $\mathfrak{F}$ denote the set of all compact faces of K, which is nonempty since $K \in \mathfrak{F}$ and partially ordered by the inclusion relation. If $\mathfrak{F}_0 \subseteq \mathfrak{F}$ is totally ordered, then we may conclude that $F_0 := \bigcap_{F \in \mathfrak{F}_0} F$ is nonempty (by the finite intersection property for compact sets) and certainly compact and a face of K. Thus, F_0 is a lower bound of $\mathfrak{F}_0$. We may therefore apply Zorn's lemma and deduce that there exists $F \in \mathfrak{F}$ that is minimal with respect to inclusion.

We will show that F consists of a single point $x \in K$, i.e., $F = \{x\}$ and hence x is an extreme point of K, which then implies that $\operatorname{ex} K \neq \emptyset$.

Recall that as a face of K, the set F is nonempty. We prove the above assertion by contradiction. Suppose there are two distinct elements x_1 and x_2 of F. By the corollary to the Hahn-Banach theorems there is some $\rho \in X'$ such that $\operatorname{Re}\rho(x_1) \neq \operatorname{Re}\rho(x_2)$.

Applying the above lemma to F in place of K we may conclude that there is some $c \in \mathbb{R}$ such that $G := \{x \in F \mid \operatorname{Re}\rho(x) = c\}$ is a compact face of F. It follows easily from the definition that G is also a face of K, hence $G \in \mathfrak{F}$. Since it cannot be that both x_1 and x_2 belong to G, we obtain that G is a compact face proper subset of F, contradicting the minimality of F.

It remains to show that $K = \overline{\operatorname{co}}\,(\operatorname{ex} K)$. The relation $K \supseteq \overline{\operatorname{co}}\,(\operatorname{ex} K)$ is obvious and we will prove $K \subseteq \overline{\operatorname{co}}\,(\operatorname{ex} K)$ by contradiction. Denote $E := \overline{\operatorname{co}}\,(\operatorname{ex} K)$ and suppose $x_0 \in K \setminus E$. By (an obvious variant of) Hahn-Banach separation II there exist $\rho \in X'$ and a real number a such that

$$(*) \qquad \forall y \in E\colon \quad \operatorname{Re}\rho(x_0) > a \geq \operatorname{Re}\rho(y).$$

Let $c_1 := \max\{\operatorname{Re}\rho(x) \mid x \in K\}$, then $c_1 > a$ and the set $F_1 := \{x \in K \mid \operatorname{Re}\rho(x) = c_1\}$ is a compact face of K by the previous lemma. Since F_1 is also nonempty, compact, and convex, the first part of the proof shows that F_1 has an extreme point x_1, which is then[3] also an extreme point of K. In particular, $x_1 \in E$, while $\operatorname{Re}\rho(x_1) = c_1 > a$, which contradicts $(*)$. □

[3]It is easy to see that an extreme point of a face of K is an extreme point of K as well.

Corollary: Let K be a non-empty compact convex subset of a Hausdorff locally convex vector space X. If $\rho \in X'$, then there is an extreme point x_0 of K such that $\operatorname{Re}\rho(x) \leq \operatorname{Re}\rho(x_0)$ holds for all $x \in K$.

Proof: We may put $c := \max\{\operatorname{Re}\rho(x) \mid x \in K\}$ and obtain from the above lemma that $F := \{x \in K \mid \operatorname{Re}\rho(x) = c\}$ is a compact face of K. By the Krein-Milman theorem the nonempty compact convex subset F has an extreme point x_0. It follows from the definition that an extreme point of a face of K is an extreme point of K, thus x_0 is an extreme point of K and, since $x_0 \in F$, we have $\operatorname{Re}\rho(x_0) = c \geq \operatorname{Re}\rho(x)$ for every $x \in K$. □

Remark: One can also show the following related result (cf. [Wer18, Theorem VIII.4.4(c)] or [Con10, Chapter V, Theorem 7.8] or [KR, Theorem 1.4.5]): If K is a nonempty compact convex subset of a Hausdorff locally convex vector space X and B is a closed subset of K such that $K = \overline{\operatorname{co}}\, B$, then B contains the extreme points of K, i.e., $B \supseteq \operatorname{ex} K$.

6. Weak and weak* topologies

6.1. Dual pairs: Let X, Y be $\mathbb{K}$-vector spaces and $(x, y) \mapsto \langle x, y \rangle$ be a bilinear map $X \times Y \to \mathbb{K}$. We call (X, Y) a *dual pair*, if

$$\forall x \in X, x \neq 0, \ \exists y \in Y \colon \langle x, y \rangle \neq 0 \quad \text{and} \quad \forall y \in Y, y \neq 0, \ \exists x \in X \colon \langle x, y \rangle \neq 0. \tag{6.1}$$

For any $y \in Y$, we have the induced linear functional $l_y(x) := \langle x, y \rangle$ on X and the map $y \mapsto l_y$ is linear from Y into the algebraic dual X^* of X. The requirement above says that the latter is an injective map, and the same holds for the assignment $x \mapsto \langle x, . \rangle$, $X \to Y^*$. In this sense, we have $Y \hookrightarrow X^*$ and $X \hookrightarrow Y^*$ for any dual pair (X, Y). Note that the ranges of these injective maps are point separating subspaces in the respective duals.

For any $y \in Y$ we have the seminorm $p_y(x) := |\langle x, y \rangle|$ on X and the set P of these seminorms generates a locally convex vector space topology on X, which we call the $\sigma(X,Y)$-*topology* (or *weak topology*) on X. The $\sigma(Y,X)$-*topology* on Y is defined analogously. It follows from the definition of a dual pair and Lemma 4.11 that the $\sigma(X,Y)$-topology and the $\sigma(Y,X)$-topology are always Hausdorff. By Proposition 5.7, a net $(x_j)_{j \in I}$ in X converges to 0 in the $\sigma(X,Y)$-topology, if and only if $\langle x_j, y \rangle \to 0$ for every $y \in Y$. Considering x_j $(j \in I)$ as elements in Y^*, this translates into pointwise convergence of the net (x_j).

Examples: 1) Let X be a locally convex Hausdorff vector space and $Y = X'$, then (X, X') is a dual pair with respect to the bilinear map $(x, x') \mapsto x'(x)$, since the second part in (6.1) follows from the definition of maps $X \to \mathbb{K}$ and the first part from the corollary of the Hahn-Banach theorem stated in 5.10. The corresponding topology $\sigma(X, X')$ is also called the *weak topology* on X and this is consistent with the topology described in Example 4.12.8) in case of a normed space.

2) In the situation of 1), we also obtain a dual pair (X', X) and $\sigma(X', X)$ is the topology of pointwise convergence, which we call weak* topology as in the case of normed spaces discussed in Example 4.12.9).

3) Coming back to Example 4.12.11), consider a metric space Ω, $X = C_b(\Omega)$ the space of bounded continuous functions, and $Y = M(\Omega)$ the space of regular signed or complex Borel measures on Ω. The bilinear map

$$(f, \mu) \mapsto \int_\Omega f \, d\mu \quad (f \in C_b(\Omega), \mu \in M(\Omega))$$

defines the structure of a dual pair (the argument is again based on the regularity of the measures, see [Els11, Kapitel VIII, §1, Satz 4.6]). The $\sigma(M(\Omega), C_b(\Omega))$-topology corresponds to the notion of weak topology of probability theory as indicated in Example 4.12.11).

4) On $X = \mathbb{R}^\mathbb{R}$ (the space of functions $\mathbb{R} \to \mathbb{R}$) consider for any $t \in \mathbb{R}$ the evaluation functional $\delta_t : \mathbb{R}^\mathbb{R} \to \mathbb{R}$, $f \mapsto f(t)$. Let $Y := \operatorname{span}\{\delta_t \mid t \in \mathbb{R}\}$ (a subspace of $(\mathbb{R}^\mathbb{R})^*$). We obtain a dual pair by defining the bilinear map $X \times Y \to \mathbb{R}$,

$$\langle f, \sum_{k=1}^m \lambda_k \delta_{t_k} \rangle := \sum_{k=1}^m \lambda_k f(t_k),$$

and $\sigma(X,Y)$ is the topology of pointwise convergence as in Example 4.12.1).

5) Let X and Y be normed spaces. We will describe the weak operator topology on $L(X,Y)$ from Example 4.12.10) in terms of a σ-topology of a dual pairing as follows: The (purely algebraic $\mathbb{K}$-)tensor product $X \otimes Y'$ can be described equivalently as the set of all linear maps $u : X' \to Y'$ of the form $u(x') = \sum_{j=1}^m x'(x_j) y'_j$ with $m \in \mathbb{N}$ and $x_j \in X$, $y'_j \in Y'$ ($j = 1, \dots, m$), i.e., $u = \sum_{j=1}^m x_j \otimes y'_j$. We claim that we obtain a dual pair $(L(X,Y), X \otimes Y')$ via the bilinear map $\langle .,. \rangle : L(X,Y) \times (X \otimes Y') \to \mathbb{K}$, given by

$$\left\langle T, \sum_{j=1}^m x_j \otimes y'_j \right\rangle := \sum_{j=1}^m y'_j(T x_j).$$

To show that for any $T \neq 0$ there is a $u \in X \otimes Y'$ with $\langle T, u \rangle \neq 0$, first choose $x_0 \in X$ such that $T x_0 \neq 0$, then there is some $y'_0 \in Y'$ with $y'_0(T x_0) \neq 0$ by the Hahn-Banach theorem; thus, $\langle T, x_0 \otimes y'_0 \rangle \neq 0$. Furthermore, for given $u = \sum_{j=1}^m x_j \otimes y'_j \neq 0$ we have to find $T \in L(X,Y)$ such that $\langle T, u \rangle \neq 0$; there is some $x'_0 \in X'$ such that $0 \neq u(x'_0) \in Y'$; hence, there is also some $y_0 \in Y$ such that $u(x'_0)(y_0) \neq 0$; we define $T \in L(X,Y)$ by $Tx := x'_0(x) y_0$ and arrive at

$$\langle T, u \rangle = \sum_{j=1}^m y'_j(T x_j) = \sum_{j=1}^m y'_j(x'_0(x_j) y_0) = \left(\sum_{j=1}^m x'_0(x_j) y'_j \right)(y_0) = u(x'_0)(y_0) \neq 0.$$

Any of the seminorms $p_u(T) = |\langle T, u \rangle|$, with $u = \sum x_j \otimes y'_j$, is obviously continuous with respect to the weak operator topology. Therefore, Corollary 5.2 shows that the $\sigma(L(X,Y), X \otimes Y')$-topology coincides with the weak operator topology.

6.2. The dual space of $(X, \sigma(X,Y))$: We need a preparatory result from linear algebra.

Lemma: Let $l, l_1, \dots, l_n : X \to \mathbb{K}$ be linear functionals. Then the following are equivalent:

(i) $l \in \operatorname{span}\{l_1, \dots, l_n\}$,

(ii) $\exists M > 0$ such that $|l(x)| \leq M \max_{j=1}^n |l_j(x)|$ holds for every $x \in X$,

(iii) $\bigcap_{j=1}^n \ker(l_j) \subseteq \ker(l)$.

Proof: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) is clear. It remains to show (iii) $\Rightarrow$ (i).
Let $V := \{(l_1(x), \dots, l_n(x)) \in \mathbb{K}^n \mid x \in X\}$, then (iii) guarantees that the linear map $\phi_0 : V \to \mathbb{K}$, $\phi_0(l_1(x), \dots, l_n(x)) = l(x)$ is well-defined. Let $\phi : \mathbb{K}^n \to \mathbb{K}$ be a linear extension of ϕ_0, then there are $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that $\phi(\xi_1, \dots, \xi_n) = \sum_{j=1}^n \alpha_j \xi_j$ and we obtain

$$\forall x \in X: \quad l(x) = \phi_0(l_1(x), \dots, l_n(x)) = \phi(l_1(x), \dots, l_n(x)) = \sum_{j=1}^n \alpha_j l_j(x).$$

□

Corollary: Let (X, Y) be a dual pair. A linear functional on X is $\sigma(X, Y)$-continuous, if and only if it is of the form $x \mapsto \langle x, y\rangle$ for some $y \in Y$. Therefore,

$$(X_{\sigma(X,Y)})' = Y.$$

Proof: By Corollary 5.4, a linear functional l on X is $\sigma(X, Y)$-continuous, if and only if it satisfies an estimate as in statement (ii) of the above lemma with $l_j(x) = \langle x, y_j\rangle$ and $y_j \in Y$. Therefore, the $\sigma(X, Y)$-continuity of l is also equivalent to being a linear combination $l = \sum \alpha_j l_i$, i.e., $l(x) = \langle x, y\rangle$ with $y = \sum \alpha_j y_j$. □

In particular, in case of a locally convex vector space (X, τ) with the dual pairs (X, X') or (X', X), we obtain the following statements as direct **consequences**:

(a) A linear functional on X is weakly continuous, if and only if it is τ-continuous. (Compare with Example 5.9.2), where this result has been noted for normed spaces.)

(b) A linear functional on X' is weak* continuous, if and only if it is of the form of an evaluation functional $x' \mapsto x'(x)$ with some $x \in X$, i.e., $(X'_{\sigma(X',X)})' = X$.

6.3. Proposition: Let (X, Y) be a dual pair. The weak topology $\sigma(X, Y)$ has the following property: A map f from any topological space (T, τ) into $(X, \sigma(X, Y))$ is continuous, if and only if the composition map $y \circ f : t \mapsto \langle f(t), y\rangle$ is continuous $T \to \mathbb{K}$ for every $y \in Y$.

We may remark that, in fact ([Sch71, Sections II.5 and IV.1]), $\sigma(X, Y)$ is the coarsest topology on X such that every $y \in Y$ defines a continuous function and thus is the initial or projective topology on X with respect to the linear functionals given by the elements $y \in Y$.

Proof: Clearly, if $f : (T, \tau) \to (X, \sigma(X, Y))$ is continuous, then $y \circ f$ is continuous for every $y \in Y$ due to the corollary in 6.2.

To prove the converse, suppose $y \circ f$ is continuous for every $y \in Y$. Let $t \in T$ and U be a $\sigma(X, Y)$-neighborhood of 0 in X. We have to show that there is a τ-neighborhood W of t in T such that $f(W) \subseteq f(t) + U$. We may assume that U is a basis neighborhood of the form $U = \{x \in X \mid |\langle x, y_j\rangle| \leq \varepsilon \ (j = 1, \ldots, n)\}$ with $y_1, \ldots, y_n \in Y$. Every function $y_j \circ f$ $(1 \leq j \leq n)$ is continuous, thus, there exist τ-neighborhoods W_j of t in T such that $|\langle f(s) - f(t), y_j\rangle| \leq \varepsilon$, if $s \in W_j$. Putting $W := W_1 \cap \ldots \cap W_n$ we obtain a neighborhood of t such that $f(W) - f(t) \subseteq U$. □

6.4. Polar subsets: Let (X, Y) be a dual pair, $A \subseteq X$, and $B \subseteq Y$. The *polar* of A is

$$A^\circ := \{y \in Y \mid \forall x \in A\colon \operatorname{Re}\langle x, y\rangle \leq 1\}$$

and the *polar* of B is

$$B^\circ := \{x \in X \mid \forall y \in B\colon \operatorname{Re}\langle x, y\rangle \leq 1\}.$$

Note that $A^{\circ\circ} := (A^\circ)^\circ$ is defined and a subset of X.

Remark: (i) The following can be shown as an exercise: If X is a normed space and B_X denotes the closed unit ball, then the duality (X, X') yields $(B_X)^\circ = B_{X'}$. Furthermore, if U is a subspace of X, then U° coincides with the *annihilator* $U^\perp = \{x' \in X' \mid x'|_U = 0\}$.

(ii) In the literature, one often finds the definition of the polar of a subset $A \subseteq X$ in the form $A^\circ = \{y \in Y \mid \forall x \in A\colon |\langle x, y\rangle| \leq 1\}$, which we would call here the *absolute polar* of A.

For subsets of X or Y of the dual pair (X, Y), closures shall always refer to the weak topology $\sigma(X, Y)$ or $\sigma(Y, X)$, unless stated otherwise. Let A and A_j ($j = 1, 2$ or $j \in J$) denote subsets of X, then we collect the following list of properties[1]:

(a) $A_1 \subseteq A_2 \;\Rightarrow\; A_2^\circ \subseteq A_1^\circ$,

(b) $0 \in A^\circ$ and $A \subseteq A^{\circ\circ}$,

(c) $A^\circ = \overline{\mathrm{co}}\,(A^\circ)$ and $A^\circ = (\overline{\mathrm{co}}\,A)^\circ$,

(d) if A is balanced, then $A^\circ = \{y \in Y \mid \forall x \in A\colon |\langle x, y\rangle| \leq 1\}$ (absolute polar),

(e) if $\lambda > 0$, then $(\lambda A)^\circ = \frac{1}{\lambda} A^\circ$,

(f) $(\bigcup_{j\in J} A_j)^\circ = \bigcap_{j\in J} A_j^\circ$,

(g) $(\bigcap_{j\in J} A_j)^\circ \supseteq \overline{\mathrm{co}} \bigcup_{j\in J} A_j^\circ$.

Proof: Properties (a), (b), and (d)-(f) are clear or very easily shown.

(c): $A^\circ = \bigcap_{x\in A}\{y \in Y \mid \mathrm{Re}\langle x, y\rangle \leq 1\}$ is an intersection of closed convex sets, hence it is itself convex and closed, and therefore $A^\circ = \overline{\mathrm{co}}\,(A^\circ)$. Since $A \subseteq \overline{\mathrm{co}}\,A$, we have $(\overline{\mathrm{co}}\,A)^\circ \subseteq A^\circ$ by (a), hence it remains to show that any $y \in A^\circ$ belongs to $(\overline{\mathrm{co}}\,A)^\circ$. This follows upon observing that $\mathrm{Re}\langle \lambda x_1 + (1-\lambda)x_2, y\rangle = \lambda\,\mathrm{Re}\langle x_1, y\rangle + (1-\lambda)\,\mathrm{Re}\langle x_2, y\rangle$ (for any real λ) and $\mathrm{Re}\langle \lim x_j, y\rangle = \lim \mathrm{Re}\langle x_j, y\rangle$.

(g): Put $A := \bigcap_{j\in J} A_j$, then $A \subseteq A_j$ for every $j \in J$, hence $A^\circ \supseteq A_j^\circ$ for every $j \in J$, and therefore $A^\circ \supseteq \bigcup_{j\in J} A_j^\circ$. By (c), A° is convex and closed, thus also $A^\circ \supseteq \overline{\mathrm{co}} \bigcup_{j\in J} A_j^\circ$. □

Bipolar theorem: $A^{\circ\circ} = \overline{\mathrm{co}}\,(A \cup \{0\})$.

Proof: By (b), $0 \in (A^\circ)^\circ = A^{\circ\circ}$ and $A \subseteq A^{\circ\circ}$, hence $A \cup \{0\} \subseteq A^{\circ\circ}$. By (c), $A^{\circ\circ}$ is convex and closed, hence $\overline{\mathrm{co}}\,(A \cup \{0\}) \subseteq A^{\circ\circ}$.

Let $V := \overline{\mathrm{co}}\,(A \cup \{0\})$ and suppose $x_0 \in A^{\circ\circ} \setminus V$. Since V is closed and convex, the Hahn-Banach separation theorem II and Corollary 6.2 allow us to find some $y_0 \in Y$ and $\varepsilon > 0$ such that $\mathrm{Re}\langle x_0, y_0\rangle + \varepsilon \leq \mathrm{Re}\langle v, y_0\rangle$ holds for every $v \in V$, or, as noted towards the end in the proof of Hahn-Banach separation II, with $y_1 := -y_0$ we obtain $\mathrm{Re}\langle x_0, y_1\rangle - \varepsilon \geq \mathrm{Re}\langle v, y_1\rangle$ for every $v \in V$. Hence there is a real number α such that

$$\forall v \in V\colon \qquad \mathrm{Re}\langle x_0, y_1\rangle > \alpha > \mathrm{Re}\langle v, y_1\rangle.$$

Since $0 \in V$, we have $\alpha > 0$, and putting $y_2 := y_1/\alpha$ we obtain for every $a \in A \subseteq V$,

$$\mathrm{Re}\langle x_0, y_2\rangle > 1 > \mathrm{Re}\langle a, y_2\rangle.$$

By the second inequality, $y_2 \in A^\circ$. But then the first inequality implies $x_0 \notin A^{\circ\circ}$, which gives a contradiction. Thus, we also have $A^{\circ\circ} \subseteq V$ and the theorem is proved. □

[1]The analogous properties hold for subsets of Y.

Corollary: Let $C \subseteq X$ be convex with $0 \in C$, then we have:

$$C \text{ is } \sigma(X,Y)\text{-closed} \quad \iff \quad \exists B \subseteq Y : C = B^\circ.$$

Proof: The implication '$\Leftarrow$' follows by (c), and '$\Rightarrow$' follows upon setting $B := C^\circ$ and appealing to the bipolar theorem. $\square$

6.5. Alaoglu-Bourbaki theorem: Let X be a Hausdorff locally convex vector space and U be a neighborhood of 0 in X. Then U° (in the sense of the dual pair (X, X')) is weak* compact.

Proof: Let τ_p denote the topology of pointwise convergence on the set $\mathbb{K}^X$ of all functions $X \to \mathbb{K}$. Recall that τ_p corresponds to the product topology on $\mathbb{K}^X = \prod_{x \in X} \mathbb{K}$ and therefore we have Tychonoff's theorem as a convenient criterion for compactness of product sets in $(\mathbb{K}^X, \tau_p)$. We have $X' \subseteq \mathbb{K}^X$ and $\tau_p|_{X'} = \sigma(X', X)$ (as noted in 6.1, Example 1) above).

We may suppose that U is absolutely convex, for we always have an absolutely convex neighborhood V of 0 with $V \subseteq U$, which gives that $U^\circ \subseteq V^\circ$ and compactness of U° follows from that of V°. By property (d) in 6.4, we may thus suppose that

$$U^\circ = \{x' \in X' \mid \forall x \in U\colon |x'(x)| \leq 1\}.$$

For every $x \in X$, there is some $\lambda_x > 0$ with $x \in \lambda_x \cdot U$ (since U is absorbing). If $x \in U$, we choose $\lambda_x \leq 1$ (e.g., $\lambda_x = 1$ is always possible). Since $x/\lambda_x \in U$ by construction, we have for all $x' \in U^\circ$ and $x \in X$,

$$(*) \qquad |x'(x)| = \lambda_x \, |x'(\frac{x}{\lambda_x})| \leq \lambda_x.$$

By Tychonoff's theorem, the set

$$K := \prod_{x \in X} \underbrace{\{\lambda \in \mathbb{K} \mid |\lambda| \leq \lambda_x\}}_{=: K_x} = \{f \in \mathbb{K}^X \mid \forall x \in X\colon f(x) \in K_x\}$$

is τ_p-compact and $(*)$ shows that $U^\circ \subseteq K$. It remains to show that U° is τ_p-closed in K.

If $f \in K$ is in the τ_p-closure of U°, then being a pointwise limit of a net of linear functionals from U°, f is a linear map. Due to our choice of $\lambda_x \leq 1$ for $x \in U$ and knowing that $f \in K$, we have $f(U) \subseteq \{\lambda \in \mathbb{K} \mid |\lambda| \leq 1\}$, which implies that f is continuous[2] at 0, hence $f \in X'$. Thus, we have shown that $\overline{U^\circ}^{\tau_p} \subseteq X'$ and therefore deduce $\overline{U^\circ}^{\tau_p} = \overline{U^\circ}^{\sigma(X',X)} = U^\circ$. $\square$

Applying the Alaoglu-Bourbaki theorem to the closed unit ball B_X in a normed space X and recalling from Remark (i) above that $U := B_X$ implies $U^\circ = (B_X)^\circ = B_{X'}$, we obtain the following special case (Alaoglu's or Banach-Alaoglu theorem).

6.6. Corollary: Let X be a normed space, then the closed unit ball $B_{X'}$ of X' is weak* compact.

6.7. Remark: (i) If X is a separable locally convex vector space, then the topology induced by $\sigma(X', X)$ on the polar of a 0-neighborhood is metrizable ([Sch71, Chapter IV, 1.7]) and compactness means sequential compactness. Thus, in case of a separable normed space X, bounded

[2]Since, e.g., for every $\varepsilon > 0$ we obtain $f(\varepsilon U) \subseteq \{\lambda \mid |\lambda| \leq \varepsilon\}$ and εU is a neighborhood of 0 in X.

sequences in the dual X' always possess weak* convergent subsequences. Without separability this is no longer true. An example of this failure is the following: Consider the evaluation functionals $\delta_n \in l^{\infty\prime}$ ($n \in \mathbb{N}$) given by $\delta_n(x) := x_n$ for every $x = (x_m)_{m\in\mathbb{N}} \in l^\infty$. We have $\delta_n \in B_{l^{\infty\prime}} = (B_{l^\infty})^\circ$. Suppose we had a weak* convergent subsequence $(\delta_{n_k})_{k\in\mathbb{N}}$ of $(\delta_n)_{n\in\mathbb{N}}$. For every $x \in l^\infty$ put $y(x) := \lim_{k\to\infty} \delta_{n_k}(x)$. Then $y \in B_{l^{\infty\prime}}$, since $B_{l^{\infty\prime}}$ is weak* closed (being a weak* compact subset). Let $z = (z_m)_{m\in\mathbb{N}} \in l^\infty$ be defined by

$$z_m := \begin{cases} 0, & \text{if } \forall k \in \mathbb{N} : m \neq n_k, \\ 1, & \text{if } \exists l \in \mathbb{N} : m = n_{2l}, \\ 0, & \text{if } \exists l \in \mathbb{N}_0 : m = n_{2l+1}. \end{cases}$$

We obtain $\delta_{n_{2l}}(z) = 1$ for $l \in \mathbb{N}$ and $\delta_{n_{2l+1}}(z) = 0$ for $l \in \mathbb{N}_0$ and therefore arrive at the contradiction $y(z) = \lim_{k\to\infty} \delta_{n_k}(z) = \lim(0, 1, 0, 1, 0, 1, \ldots)$.

(ii) Recall that a Banach space X is *reflexive*, if it is canonically isomorphic to X'' via the map $\iota\colon X \to X''$, $\iota(x)(x') := x'(x)$. Combining the bipolar and the Alaoglu-Bourbaki theorem, one can obtain characterizations of reflexivity for Banach spaces (cf. [Wer18, Satz VIII.3.18] or [Con10, Chapter V, Theorem 4.2]), e.g., in the following form: The Banach space X is reflexive, if and only if B_X is weakly compact.

6.8. Weak* compactness of convex sets and extreme points: Although we always have plenty of convex sets in any locally convex space, the compactness of these subsets is often true only in the weak or weak* topologies and the Alaoglu-Bourbaki theorem plays an important part in many applications of the Krein-Milman theorem 5.11. A particular case is again the closed unit ball in the dual of a normed space, where the Krein-Milman theorem applies for the weak* topology and can be used, e.g., to show that the Banach spaces c_0 (the space of real or complex sequences converging to 0) and $L^1([0, 1])$ cannot be the dual of a normed space, since their closed unit balls do not have any extreme points (cf. [Wer18, Korollar VIII.4.6]).

7. Basic theory of distributions

The theory of distributions aims at an extension of the notion of scalar functions on an open subset of $\Omega \subseteq \mathbb{R}^n$ while maintaining a calculus of differentiation. The basic idea is to consider linear functionals on function spaces and we can make the following observations:

(a) Functions can be considered as linear functionals (see also Example 5.9.1)): For example, suppose $f \in L^1_{\mathrm{loc}}(\Omega)$, then[1]

$$T_f(\varphi) := \int_{\Omega} f(x)\varphi(x)\,dx$$

defines a linear functional $T_f \colon C_c(\Omega) \to \mathbb{C}$. But now we are not tied up with functions anymore in implementing such functionals. In particular, we have finite measures acting as functionals (similarly as in the Riesz representation theorem). One of the most prominent functionals is given by the Dirac measure δ_{x_0} concentrated at the point $x_0 \in \Omega$, acting by

$$\delta_{x_0}(\varphi) = \int_{\Omega} \varphi(x)\,d\delta_{x_0} = \varphi(x_0).$$

(b) Differentiation of non-smooth functions can be implemented taking up the idea of integration by parts (which is the concept of a *weak derivative*): On $\Omega = \,]0,1[$, for example, if $f \in C^1(]0,1[)$ and $\varphi \in C^1_c(]0,1[)$, then

$$T_{f'}(\varphi) = \int_0^1 f'(x)\varphi(x)\,dx = -\int_0^1 f(x)\varphi'(x)\,dx = -T_f(\varphi').$$

However, if f is not differentiable and we have merely $f \in L^1_{\mathrm{loc}}(]0,1[)$, then we may still define the linear functional $T_f{}'(\varphi) := -T_f(\varphi')$ on $C^1_c(]0,1[)$ and consider $T_f{}'$ to be the derivative of T_f. We observe that we may obtain definitions of derivatives up to order $k \in \mathbb{N}$, if we define the functionals instead on the smaller spaces $C^k_c(\Omega)$, or, to be prepared for derivatives of arbitrary order, consider the space $C^\infty_c(\Omega) = \mathscr{D}(\Omega)$ of test functions known from Example 4.12.6).

7.1. The space of test functions: Recall from Examples 5) and 6) in 4.12, that we defined the locally convex topology τ on $\mathscr{D}(\Omega)$ via the family of seminorms P on $\mathscr{D}(\Omega)$ such that, for every $p \in P$, the restriction $p|_{\mathscr{D}_K(\Omega)}$ is continuous $(\mathscr{D}_K(\Omega), \tau_K) \to [0,\infty[$ for every compact subset

[1] Recall that $L^1_{\mathrm{loc}}(\Omega)$ is the set of measurable functions that are integrable on every compact subset of Ω, and $C_c(\Omega)$ denotes the space of continuous functions with compact support.

$K \subset \Omega$. Here, $\mathscr{D}_K(\Omega)$ is the set of functions $\varphi \in C_c^\infty(\Omega)$ with $\operatorname{supp}(\varphi) \subseteq K$ and the locally convex topology τ_K is generated by the family of seminorms

$$p_m(\varphi) := \max_{|\alpha| \leq m} \sup_{x \in \Omega} |\partial^\alpha \varphi(x)| \qquad (m \in \mathbb{N}_0, \varphi \in \mathscr{D}_K(\Omega)).$$

Note that $p_l \leq p_m$, if $l \leq m$, thus, any finite number of seminorms $p_{m_1}, \ldots, p_{m_r}$ is bounded by the single seminorm p_m, if $m \geq m_j$ $(j = 1, \ldots, r)$. Strictly speaking, we should indicate the dependence on K in the notation for the seminorms p_m, but we do not want to overload the notation and it is convenient to think of p_m as denoting the appropriate restriction to $\mathscr{D}_K(\Omega)$ of a similar seminorm given on $\mathscr{D}(\Omega)$.

By Lemma 5.1(c) and the monotonicity of p_m with respect to m observed above, a seminorm p on $\mathscr{D}(\Omega)$ belongs to P, if and only if for every compact subset $K \subset \Omega$ there are $c > 0$ and $m \in \mathbb{N}_0$ such that

$$\forall \varphi \in \mathscr{D}_K(\Omega)\colon \quad p(\varphi) \leq c\, p_m(\varphi). \tag{7.1}$$

Note that, in general, c and m will depend on K.

Lemma: (a) $\tau|_{\mathscr{D}_K(\Omega)} = \tau_K$.

(b) $\mathscr{D}_K(\Omega)$ is a τ-closed subspace of $\mathscr{D}(\Omega)$.

(c) $(\mathscr{D}(\Omega), \tau)$ is a Hausdorff space.

(d) Let Y be a locally convex space and $L\colon \mathscr{D}(\Omega) \to Y$ be linear. Then L is τ-continuous, if and only if the restriction $L|_{\mathscr{D}_K(\Omega)}$ is τ_K-continuous for every compact set $K \subset \Omega$.

Proof: (a): The subspace topology $\tau|_{\mathscr{D}_K(\Omega)}$ is generated by the family of seminorms $Q = \{p|_{\mathscr{D}_K(\Omega)} \mid p \in P\}$ and we have by construction $\{p_m \mid m \in \mathbb{N}_0\} \subseteq Q$ and that Q consists of τ_K-continuous seminorms, thus, Corollary 5.2 proves the claim.

(b): We have $\mathscr{D}_K(\Omega) = \bigcap\limits_{x \in \Omega \setminus K} \pi_x^{-1}(\{0\})$, where every $\pi_x\colon \mathscr{D}(\Omega) \to [0, \infty[$, $\pi_x(\varphi) := |\varphi(x)|$, is a seminorm belonging to P, since $\pi_x(\varphi) \leq p_0(\varphi)$ for every $\varphi \in \mathscr{D}_M(\Omega)$ and $M \subset \Omega$ compact. Therefore, every subset $\pi_x^{-1}(\{0\})$ is τ-closed.

(c): We appeal to Lemma 4.11 upon observing that for any $\mathscr{D}(\Omega) \ni \varphi \neq 0$ there is some $x \in \Omega$ with $\varphi(x) \neq 0$ and hence π_x as in (b) is a seminorm in P with $\pi_x(\varphi) \neq 0$.

(d): If L is τ-continuous, then by (a) the restriction is τ_K-continuous. To prove the converse we apply Theorem 5.3(iii): Let q be a continuous seminorm on Y, then by assumption $q \circ (L|_{\mathscr{D}_K(\Omega)}) = (q \circ L)\,|_{\mathscr{D}_K(\Omega)}$ is a continuous seminorm on $\mathscr{D}_K(\Omega)$, i.e., $q \circ L \in P$, in particular, $q \circ L$ is a τ-continuous seminorm on $\mathscr{D}(\Omega)$. □

In applications of distribution theory it is extremely convenient that linear functionals on $\mathscr{D}(\Omega)$ are automatically τ-continuous, if they are sequentially continuous. We will prove this fact in Theorem 7.3 below, where we will see that it is based essentially on property (d) above. This is also the reason why we are content here with describing convergence of sequences in $\mathscr{D}(\Omega)$.

Proposition: Let $(\varphi_l)_{l \in \mathbb{N}}$ be a sequence in $\mathscr{D}(\Omega)$, then the following are equivalent:

(i) $\varphi_l \to 0$ with respect to τ,

(ii) there exists a compact set $K \subset \Omega$ such that $\operatorname{supp}(\varphi_l) \subseteq K$ for every $l \in \mathbb{N}$ and $\varphi_l \to 0$ holds in $\mathscr{D}_K(\Omega)$,

(iii) there exists a compact set $K \subset \Omega$ such that $\operatorname{supp}(\varphi_l) \subseteq K$ for every $l \in \mathbb{N}$ and for every $\alpha \in \mathbb{N}_0^n$ the sequence $(\partial^\alpha \varphi_l)_{l\in\mathbb{N}}$ converges uniformly to 0.

Proof: Statement (iii) is just a reformulation of (ii) and the implication (ii) $\Rightarrow$ (i) is clear from the definition of τ (and (a) in the above lemma), hence it suffices to prove (i) $\Rightarrow$ (ii).

Suppose (i) holds but (ii) were wrong, i.e., there is no compact subset $K \subset \Omega$ with the properties stated. Then we can find an increasing sequence of compact sets $K_1 \subset \overset{\circ}{K}_2 \subset K_2 \subset \overset{\circ}{K}_3 \subset K_3 \subset \ldots \subset \Omega$ with $\bigcup_{m\in\mathbb{N}} \overset{\circ}{K}_m = \Omega$ and a subsequence $(\varphi_{l_m})_{m\in\mathbb{N}}$ such that $\varphi_{l_m} \in \mathscr{D}_{K_m}(\Omega)\setminus\mathscr{D}_{K_{m-1}}(\Omega)$ for every $m \in \mathbb{N}$ (note that (i) and $\operatorname{supp}(\varphi_l) \subseteq K$ for every $l \in \mathbb{N}$ would imply $\varphi_l \to 0$ in $\mathscr{D}_K(\Omega)$). Any compact subset of Ω is contained in some K_m, since the sets $\overset{\circ}{K}_m$ ($m \in \mathbb{N}$) provide an increasing open cover of Ω. Therefore, $\mathscr{D}(\Omega) = \bigcup_{m\in\mathbb{N}} \mathscr{D}_{K_m}(\Omega)$.

Choose $x_m \in \overset{\circ}{K}_m \setminus K_{m-1}$ with $\alpha_m := |\varphi_{l_m}(x_m)| > 0$. We have seen in the previous proof of the lemma that the seminorms $\pi_m \colon \varphi \mapsto |\varphi(x_m)|/\alpha_m$ are τ-continuous on $\mathscr{D}(\Omega)$. We note that $\pi_m |_{\mathscr{D}_{K_r}(\Omega)} = 0$, if $m > r$. Hence for every $\varphi \in \mathscr{D}(\Omega)$ there are at most finitely many $m \in \mathbb{N}$ such that $\pi_m(\varphi) \neq 0$ and we may define the seminorm π on $\mathscr{D}(\Omega)$ by $\pi(\varphi) = \sum_{m=1}^\infty \pi_m(\varphi)$. If $M \subset \Omega$ is compact, there is some $N \in \mathbb{N}$ such that $M \subseteq K_N$, which implies $\pi |_{\mathscr{D}_M(\Omega)} = \sum_{m=1}^N \pi_m |_{\mathscr{D}_M(\Omega)}$ and shows τ_M-continuity of the restriction. Thus, $\pi \in P$, in particular, π is τ-continuous.

Since $\varphi_l \to 0$ by assumption, we have $\pi(\varphi_{l_m}) \to \pi(0) = 0$ ($m \to \infty$), but at the same time $\pi(\varphi_{l_m}) \geq \pi_m(\varphi_{l_m}) = 1$, a contradiction. □

7.2. Definition: The dual space of $(\mathscr{D}(\Omega), \tau)$ is denoted by $\mathscr{D}'(\Omega)$ and called the space of *distributions* on Ω.

7.3. Theorem: Let $T \colon \mathscr{D}(\Omega) \to \mathbb{C}$ be linear, then the following are equivalent:

(i) $T \in \mathscr{D}'(\Omega)$, i.e., T is continuous,

(ii) for every compact set $K \subset \Omega$, the restriction $T|_{\mathscr{D}_K(\Omega)}$ is continuous on $\mathscr{D}_K(\Omega)$,

(iii) for every compact set $K \subset \Omega$, there are $m \in \mathbb{N}_0$ and $c > 0$ such that

$$\forall \varphi \in \mathscr{D}_K(\Omega)\colon \quad |T(\varphi)| \leq c\, p_m(\varphi) = c \max_{|\alpha|\leq m} \sup_{x\in\Omega} |\partial^\alpha \varphi(x)|,$$

(iv) sequential continuity of T at 0, i.e., $\varphi_l \to 0$ in $\mathscr{D}(\Omega)$ implies $T\varphi_l \to 0$ in $\mathbb{C}$.

Proof: (i) $\Leftrightarrow$ (ii) is a special case of (d) in Lemma 7.1.

(ii) $\Leftrightarrow$ (iii) holds, because the seminorms p_m generate the topology τ_K on $\mathscr{D}_K(\Omega)$.

(i) $\Rightarrow$ (iv) is the general topological fact that continuity implies sequential continuity.

(iv) $\Rightarrow$ (ii): We immediately obtain the sequential continuity of $T|_{\mathscr{D}_K(\Omega)}$ on $\mathscr{D}_K(\Omega)$. The locally convex topology τ_K on $\mathscr{D}_K(\Omega)$ is generated by the countable set of seminorms $\{p_m \mid m \in \mathbb{N}_0\}$. It is a routine exercise to show that, generally in a locally convex vector space in these circumstances, we obtain a metric inducing the same topology by setting (see, e.g., [Kab14, Satz 1.1])

$$d(\varphi, \psi) := \sum_{m=0}^{\infty} 2^{-m} \frac{p_m(\varphi - \psi)}{1 + p_m(\varphi - \psi)} \qquad (\varphi, \psi \in \mathscr{D}_K(\Omega)).$$

(For the proof of the triangle inequality, it is advisable to make use of the fact that $t \mapsto t/(1+t)$ is an increasing function $[0, \infty[\to \mathbb{R}$.

Thus, we have seen that $(\mathscr{D}_K(\Omega), \tau_K)$ is a metrizable space, therefore sequential continuity of $T|_{\mathscr{D}_K(\Omega)}$ implies continuity. □

7.4. Remark: The constant c as well as the nonnegative integer m in condition (iii) of the above theorem depend, in general, on the compact set K. In case a uniform m can be found for all K, then the minimum of these numbers m is called the *order* of the distribution.

7.5. Examples: 1) The linear functional T_f, defined for any $f \in L^1_{\text{loc}}(\Omega)$ in observation (a) of the introduction to the current section, is indeed a distribution. For every compact $K \subset \Omega$ and $\varphi \in \mathscr{D}_K(\Omega)$, we have

$$|T_f(\varphi)| \le \int\limits_{\Omega} |f(x)||\varphi(x)|\, dx \le \int\limits_{K} |f(x)|\, dx \cdot \sup_{x \in \Omega} |\varphi(x)| = \|f|_K\|_{L^1} \cdot p_0(\varphi),$$

hence $T_f \in \mathscr{D}'(\Omega)$ (and of order 0).

Moreover, the map $f \mapsto T_f$ is linear and injective and the distributions in the image of this embedding $L^1_{\text{loc}}(\Omega) \hookrightarrow \mathscr{D}'(\Omega)$ are called *regular distributions.*
(Proof of the injectivity: If $T_f = 0$, then $\int_\Omega f\varphi = 0$ for every $\varphi \in \mathscr{D}(\Omega)$. Let $K \subset \Omega$ be a compact set, then we can apply standard approximation procedures[2] to manufacture a uniformly bounded sequence $(\varphi_l)_{l \in \mathbb{N}}$ in $\mathscr{D}(\Omega)$ with $\varphi_l \to \chi_K$ pointwise, thus, dominated convergence implies $\int_K f = \int \chi_K f = \lim \int \varphi_l f = 0$. Since K was arbitrary and the compact sets generate the Borel sigma algebra $\mathcal{B}(\Omega)$ (on the σ-compact Hausdorff space Ω, [Els11, Kapitel 1, Folgerung 4.2]), we obtain[3] $\int_{A \cap K} f = 0$ for every $A \in \mathcal{B}(\Omega)$ and every compact set $K \subset \Omega$. Therefore, $f|_K = 0$ almost everywhere for every compact subset K (by [Els11, Kapitel IV, Satz 4.4]), which implies $f = 0$ almost everywhere, since Ω is σ-compact.)

2) Let μ be a complex Borel measure on Ω (recall that μ is a regular measure by Lemma 0.16). Then

$$\mu(\varphi) := \int\limits_{\Omega} \varphi\, d\mu$$

defines a distribution (of order 0), since $|\mu(\varphi)| \le |\mu|(K) \cdot p_0(\varphi)$ for every $\varphi \in \mathscr{D}_K(\Omega)$. An argument as in 1) shows that also here the map assigning the functional to each measure is

[2] The construction sketched in [Wer18, S. 461] is as follows: Consider the compact sets $K_l := \{x \in \Omega \mid d(x, K) \le 1/l\}$; we have $K = \bigcap_{l=1}^{\infty} K_l$; choose $\varphi_0 \in \mathscr{D}(\mathbb{R}^n)$ with $\operatorname{supp}(\varphi_0) \subseteq \{|x| \le 1\}$ and $\int \varphi_0 = 1$ and put $\varphi_l(x) := \int_{K_l} l^n \varphi_0(l(x-y))\, dy$ for every $x \in \Omega$; then φ_l is smooth (integration with parameters), has $\operatorname{supp}(\varphi_l) \subseteq K_l + \{|x| \le 1/l\}$, and satisfies $\varphi_l(x) = 1$, if $x \in K$.

[3] Note that we may not suppose that $\int_A f$ exists for every $A \in \mathcal{B}(\Omega)$, since f need not be integrable on non-compact sets.

injective. A prominent example, mentioned already in observation (a) of the introduction, is the *Dirac distribution* δ_{x_0} concentrated at the point $x_0 \in \Omega$, which is obtained from the Dirac measure at x_0 and acts by $\delta_{x_0}(\varphi) = \varphi(x_0)$.

We can easily show that δ_{x_0} is not a regular distribution: Suppose $\delta_{x_0} = T_f$ for some $f \in L^1_{\text{loc}}(\Omega)$, then $\Omega_0 := \Omega \setminus \{x_0\}$ is an open subset of $\mathbb{R}^n$ and $\mathscr{D}(\Omega_0) \hookrightarrow \mathscr{D}(\Omega)$, since any smooth function with compact support in Ω_0 can be extended by 0 outside to give an element in $\mathscr{D}(\Omega)$ (compact sets have finite distance to the boundary of any surrounding open set). By slight abuse of notation, we have $T_{f|_{\Omega_0}} = (T_f)|_{\mathscr{D}(\Omega_0)} = \delta_{x_0}|_{\mathscr{D}(\Omega_0)} = 0$ and injectivity of $L^1_{\text{loc}}(\Omega_0) \ni g \mapsto T_g \in \mathscr{D}'(\Omega_0)$ yields that $f|_{\Omega_0} = 0$ almost everywhere. Therefore, we obtain $f = 0$ almost everywhere, which would imply $\delta_{x_0} = 0$, a contradiction, since there are functions $\varphi \in \mathscr{D}(\Omega)$ with $\varphi(x_0) = 1$.

3) The linear functional $T(\varphi) := \varphi'(0)$ is a distribution (of order 1) on $\mathbb{R}$, i.e., $T \in \mathscr{D}'(\mathbb{R})$, since $|T(\varphi)| \leq p_1(\varphi)$. (A proof that T is not of order 0 can be obtained from testing with functions $\varphi_n(x) := \varphi_0(nx)$, where $\varphi_0 \in \mathscr{D}(\mathbb{R})$ satisfies $\varphi_0(0) = 1$, $\varphi_0'(0) = 1$.)

4) The linear functional $T(\varphi) := \sum_{n=0}^{\infty} \varphi^{(n)}(n)$ is defined on all of $\mathscr{D}(\mathbb{R})$ and continuous, since with $N \in \mathbb{N}$ sufficiently large such that $\operatorname{supp}(\varphi) \subseteq [-N, N]$, we obtain[4] $|T(\varphi)| \leq \sum_{n=0}^{N-1} |\varphi^{(n)}(n)| \leq N \cdot p_{N-1}(\varphi)$. This distribution is not of finite order.

In observation (b) of the introductory paragraphs to the current section we have indicated a way to extend (the linear operation of) differentiation from function spaces to distributions by mimicking the integration by parts formula and "letting the derivatives fall on the test function". This can be described most systematically in terms of the general concept of an adjoint to a linear map.

7.6. Definition: Let X and Y be locally convex spaces and $L\colon X \to Y$ be a continuous linear map. The *adjoint* of L is defined as the linear map $L'\colon Y' \to X'$, given by $L'(y') := y' \circ L$ for every $y' \in Y'$.

7.7. Remark: Recall that $(X'_{\sigma(X',X)})' = X$ by Corollary 6.2. For every $x \in X = (X'_{\sigma(X',X)})'$, the map $x \circ L'\colon Y' \to \mathbb{K}$ is given by $(x \circ L')(y') = L'(y')(x) = (y' \circ L)(x) = y'(Lx)$ and thus a $\sigma(Y', Y)$-continuous linear functional. Therefore, Proposition 6.3 implies that L' is continuous as linear map $(Y', \sigma(Y', Y)) \to (X', \sigma(X', X))$, i.e., adjoints of continuous linear maps are always weak* continuous.

7.8. Differentiation of distributions: If $\alpha \in \mathbb{N}_0^n$, then $\varphi \mapsto (-1)^{|\alpha|}\partial^\alpha \varphi$ defines a continuous linear map $\mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$, since $p_m(\partial^\alpha \varphi) \leq p_{m+|\alpha|}(\varphi)$ shows continuity $\mathscr{D}_K(\Omega) \to \mathscr{D}_K(\Omega)$ for every compact set $K \subset \Omega$ and Lemma 7.1(d) applies. We define $\partial^\alpha\colon \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$ to be the adjoint of this map, i.e.,

$$(\partial^\alpha T)(\varphi) := (-1)^{|\alpha|}\, T(\partial^\alpha \varphi) \qquad (\varphi \in \mathscr{D}(\Omega), T \in \mathscr{D}'(\Omega)).$$

By the above remark, ∂^α is weak* continuous, and the formula of integration by parts (in several variables with iterated integrals and vanishing boundary terms due to compact support of the test functions) shows that the new definition of ∂^α is consistent when applied to regular distributions

[4] Note that φ vanishes of infinite order at the boundary of its support, hence $\varphi^{(N)}(N) = 0$.

stemming from functions that are sufficiently often continuously differentiable, i.e., $\partial^\alpha T_f = T_{\partial^\alpha f}$ in these cases.

Examples: 1) The *Heaviside function* is $H := \chi_{[0,\infty[} \in L^\infty(\mathbb{R}) \subset L^1_{\text{loc}}(\mathbb{R}) \hookrightarrow \mathscr{D}'(\mathbb{R})$. We have

$$(T_H)'(\varphi) = -T_H(\varphi') = -\int_0^\infty \varphi'(x)\,dx = \varphi(0) = \delta_0(\varphi) \qquad (\varphi \in \mathscr{D}(\mathbb{R})).$$

Furthermore, $(T_H)''(\varphi) = T_H(\varphi'') = -\varphi'(0) = -\delta_0(\varphi') = \delta_0'(\varphi)$.

2) The (class of the) measurable function f, given by $f(x) := \log(|x|)$, belongs to $L^1_{\text{loc}}(\mathbb{R})$ and we obtain, for every $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\begin{aligned}
(T_f)'(\varphi) = -T_f(\varphi') &= -\int_{-\infty}^\infty \varphi'(x)\log(|x|)\,dx = -\int_0^\infty \big(\varphi'(x) + \varphi'(-x)\big)\log(x)\,dx \\
&= -\int_0^\infty (\varphi(x) - \varphi(-x))'\log(x)\,dx = -\lim_{\varepsilon>0,\varepsilon\to 0}\int_\varepsilon^\infty (\varphi(x) - \varphi(-x))'\log(x)\,dx \\
&= -\lim_{\varepsilon>0,\varepsilon\to 0}\big((\varphi(x) - \varphi(-x))\log x\big)\Big|_\varepsilon^\infty + \lim_{\varepsilon>0,\varepsilon\to 0}\int_\varepsilon^\infty \frac{\varphi(x) - \varphi(-x)}{x}\,dx \\
&= \lim_{\varepsilon>0,\varepsilon\to 0}\int_\varepsilon^\infty \frac{\varphi(x) - \varphi(-x)}{x}\,dx = \int_0^\infty \frac{\varphi(x) - \varphi(-x)}{x}\,dx,
\end{aligned}$$

since $\varphi(x) - \varphi(-x) = 2x\varphi'(0) + O(x^2)$ near $x = 0$ and φ has compact support. The distribution $\text{vp}(1/x) := (T_f)'$ is called the *Cauchy principal value of* $\frac{1}{x}$, since its action on a test function is equivalently given by

$$\text{vp}(\frac{1}{x})(\varphi) = \lim_{\varepsilon>0,\varepsilon\to 0}\int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\,dx.$$

Note that, with the classical pointwise derivative, we have $(\log(|x|))' = 1/x$ on $\mathbb{R}\setminus\{0\}$.

7.9. Fourier transform of temperate distributions: Finally, we will briefly discuss the extension of the Fourier transform beyond $L^1(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$. Our goal is to define it as the adjoint of the classical Fourier transform on an appropriate test function space. As it turns out, $\mathscr{D}(\mathbb{R}^n)$ is not suitable for that purpose, since the Fourier transform $\mathcal{F}\varphi$ of a compactly supported function φ is easily seen to define an entire function on $\mathbb{C}^n$ and thus $\mathcal{F}\varphi$ cannot have compact support too, unless $\varphi = 0$. However, a good alternative is the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ with seminorms $p_{\alpha,m}(\varphi) := \sup_{x\in\mathbb{R}^n}(1 + |x|^m)\,|\partial^\alpha\varphi(x)|$ $(m \in \mathbb{N}_0,\ \alpha \in \mathbb{N}_0^n)$. Its dual, the space of temperate distributions $\mathscr{S}'(\mathbb{R}^n)$, has been introduced already in Example 5.9.1), where we have also seen that $L^p(\mathbb{R}^n) \hookrightarrow \mathscr{S}'(\mathbb{R}^n)$ $(1 \le p \le \infty)$ via the map $f \mapsto T_f$, $T_f(\varphi) = \int_{\mathbb{R}^n} f\varphi$. It is not difficult to show (see [Wer18, Satz VIII.5.11]) that the identical embedding $\mathscr{D}(\mathbb{R}^n) \hookrightarrow \mathscr{S}(\mathbb{R}^n)$ is continuous with dense image and therefore has an injective adjoint $\mathscr{S}'(\mathbb{R}^n) \hookrightarrow \mathscr{D}'(\mathbb{R}^n)$. Thus, every temperate distribution is a distribution.

Recall (or take for granted or look-up in [Wer18, Section V.2] or [Con16, Theorem 6.1] or [Con10, Chapter X, §6]) that $\mathcal{F}$, defined by

$$\mathcal{F}\varphi(x) = \frac{1}{(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{-ix\xi}\, \varphi(x)\, dx \qquad (\varphi \in \mathscr{S}(\mathbb{R}^n), x \in \mathbb{R}^n),$$

is a bijective linear map $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ with inverse given by

$$(\mathcal{F}^{-1}\psi)(x) = \frac{1}{(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{ix\xi}\, \psi(\xi)\, d\xi = (\mathcal{F}\psi)(-x).$$

Moreover, we have the so-called "exchange formulae"

$$\partial^\alpha(\mathcal{F}\varphi) = (-i)^{|\alpha|}\mathcal{F}(x^\alpha\varphi) \quad \text{and} \quad \mathcal{F}(\partial^\alpha\varphi) = i^{|\alpha|}\xi^\alpha\mathcal{F}\varphi$$

and, by a simple application of Fubini's theorem, we obtain also

$$(*) \qquad \int\limits_{\mathbb{R}^n} (\mathcal{F}\psi)(\xi) \cdot \varphi(\xi)\, d\xi = \int\limits_{\mathbb{R}^n} \psi(x) \cdot (\mathcal{F}\varphi)(x)\, dx.$$

We note that an equivalent, technically convenient, family of seminorms on $\mathscr{S}(\mathbb{R}^n)$ is given by

$$q_{\alpha,Q}(\varphi) := \sup_{x\in\mathbb{R}^n} |Q(x)\partial^\alpha\varphi(x)|,$$

where Q is a polynomial function on $\mathbb{R}^n$.

Lemma: (a) For every $\alpha \in \mathbb{N}_0^n$, the assignments $\varphi \mapsto x^\alpha\varphi$ and $\varphi \mapsto \partial^\alpha\varphi$ define continuous linear maps $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$.

(b) There is a constant $c > 0$ such that

$$\forall\varphi \in \mathscr{S}(\mathbb{R}^n)\, \forall\xi \in \mathbb{R}^n\colon \quad |(\mathcal{F}\varphi)(\xi)| \le c\, p_{0,n+1}(\varphi).$$

Proof: (a): The Leiniz rule applied to $\partial^\beta(x^\alpha\varphi(x))$ yields an upper bound for $p_{\beta,m}(x^\alpha\varphi)$ in the form of a finite sum with terms $q_{\beta,Q_{m+|\alpha|-l}}$ $(0 \le l \le \max(|\beta|, m+|\alpha|))$, where $Q_{m+|\alpha|-l}$ is a polynomial function of order at most $m+|\alpha|-l$. The continuity of $\varphi \mapsto \partial^\alpha\varphi$ is even more obvious, since $p_{\beta,m}(\partial^\alpha\varphi) = p_{\beta+\alpha,m}(\varphi)$.

(b): $$(2\pi)^{n/2}|(\mathcal{F}\varphi)(\xi)| = |\int\limits_{\mathbb{R}^n} e^{-ix\xi}\, \varphi(x)\, dx| \le \int\limits_{\mathbb{R}^n} |\varphi(x)|\, dx \le \int\limits_{\mathbb{R}^n} \frac{dx}{1+|x|^{n+1}} \cdot p_{0,n+1}(\varphi). \qquad \square$$

By part (a) of this lemma and the embedding $\mathscr{S}'(\mathbb{R}^n) \hookrightarrow \mathscr{D}'(\mathbb{R}^n)$, we obtain differentiation as linear map $\mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$. Moreover, multiplication by polynomial functions is obtained as linear map $\mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ as the transpose of the map in (a). For the Fourier transform, we obtain the analogous result from the following

Theorem: Both $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous linear maps $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$.

Proof: For any $\alpha \in \mathbb{N}_0^n$ and any polynomial function $Q(\xi) = \sum_{|\gamma|\leq N} a_\gamma \xi^\gamma$, we have to give an upper bound of $q_{\alpha,Q}(\mathcal{F}\varphi) = \sup_{\xi\in\mathbb{R}^n} |Q(\xi)\partial^\alpha(\mathcal{F}\varphi)(\xi)|$ in terms of finitely many seminorms of φ. Let $Q(-i\partial) := \sum_{|\gamma|\leq N} a_\gamma(-i)^{|\gamma|}\partial^\gamma$, then we have, by (b) in the above lemma and the exchange formulae,

$$\forall \xi \in \mathbb{R}^n: \quad |Q(\xi)(\partial^\alpha \mathcal{F}\varphi)(\xi)| = |\mathcal{F}(Q(-i\partial)(x^\alpha\varphi))(\xi)| \leq c\, p_{0,n+1}(Q(-i\partial)(x^\alpha\varphi)).$$

By part (a) in the above lemma, we have an upper bound for $p_{0,n+1}(Q(-i\partial)(x^\alpha\varphi))$ in terms of $\max_{j=1}^N p_{\beta_j,m_j}(\varphi)$, which completes the proof of continuity of $\mathcal{F}$. Since $(\mathcal{F}^{-1}\varphi)(x) = (\mathcal{F}\varphi)(-x)$, the continuity of $\mathcal{F}^{-1}$ follows as well. □

Formula $(*)$ shows that on Schwartz functions, considered as regular temperate distributions, the Fourier transform is its own adjoint. Thus, we come up with the following

Definition: If $T \in \mathscr{S}'(\mathbb{R}^n)$, then its Fourier transform $\mathcal{F}T \in \mathscr{S}'(\mathbb{R}^n)$ is defined by

$$(\mathcal{F}T)(\varphi) := T(\mathcal{F}\varphi) \quad (\varphi \in \mathscr{S}(\mathbb{R}^n)),$$

i.e., as the adjoint of the Fourier transform on $\mathscr{S}(\mathbb{R}^n)$.

The Fourier transform is a bijective[5] linear map $\mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ and it is easily verified that the exchange formulae extend to $\mathscr{S}'(\mathbb{R}^n)$.

Examples: 1) Let $f \in L^1(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$, then applying Fubini's theorem we obtain

$$\begin{aligned}(\mathcal{F}T_f)(\varphi) = T_f(\mathcal{F}\varphi) &= \int_{\mathbb{R}^n} f(x)\cdot(\mathcal{F}\varphi)(x)\,dx = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(x)\int_{\mathbb{R}^n} e^{-ix\xi}\varphi(\xi)\,d\xi\,dx \\ &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-ix\xi} f(x)\,dx\,\varphi(\xi)\,d\xi = \int_{\mathbb{R}^n}(\mathcal{F}f)(\xi)\cdot\varphi(\xi)\,d\xi = (T_{\mathcal{F}f})(\varphi).\end{aligned}$$

(Recall that, by the lemma of Riemann-Lebesgue, $\mathcal{F}f$ is a continuous function [vanishing at infinity], if $f \in L^1(\mathbb{R}^n)$.) Similarly, one can show that also the classical Fourier-Plancherel transform on $L^2(\mathbb{R}^n)$ is consistently extended by the Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$.

2) $(\mathcal{F}\delta_0)(\varphi) = \delta_0(\mathcal{F}\varphi) = (\mathcal{F}\varphi)(0) = \dfrac{1}{(2\pi)^{n/2}}\displaystyle\int_{\mathbb{R}^n} 1\cdot\varphi(x)\,dx = \dfrac{1}{(2\pi)^{n/2}}\,T_1(\varphi)$.

3) $(\mathcal{F}T_1)(\varphi) = (2\pi)^{n/2}(\mathcal{F}(\mathcal{F}\delta_0))(\varphi) = (2\pi)^{n/2}\delta_0(\mathcal{F}\mathcal{F}\varphi)$ and applying $(\mathcal{F}\varphi)(x) = (\mathcal{F}^{-1}\varphi)(-x)$ we obtain

$$(\mathcal{F}T_1)(\varphi) = (2\pi)^{n/2}\delta_0(\mathcal{F}\mathcal{F}\varphi) = (2\pi)^{n/2}\varphi(-0) = (2\pi)^{n/2}\varphi(0) = (2\pi)^{n/2}\delta_0(\varphi).$$

[5]It is easy to see that the adjoint of a bijective linear continuous map is bijective.

Appendix

We restate and prove Lemma 0.13

Lemma: Let $\Omega \subset \mathbb{C}$ be compact and $(B_b(\Omega), \|.\|_\infty)$ be the Banach space of bounded Borel measurable functions $\Omega \to \mathbb{C}$. Suppose $U \subseteq B_b(\Omega)$ has the following properties:

(a) $C(\Omega) \subseteq U$,

(b) $f_n \in U$ $(n \in \mathbb{N})$, $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$, and $f(t) := \lim_{n\to\infty} f_n(t)$ exists for every $t \in \Omega$
$\implies \quad f \in U$.

Then $U = B_b(\Omega)$.

Proof: Let $\mathscr{S} := \{S \subseteq B_b(\Omega) \mid S \supseteq C(\Omega)$ and S satisfies (b)$\}$ and put $V := \bigcap_{S\in\mathscr{S}} S$. Clearly $B_b(\Omega) \in \mathscr{S}$, hence $\mathscr{S} \neq \emptyset$ and $C(\Omega) \subseteq V \subseteq U$. We will show that $V = B_b(\Omega)$.

Claim 1: V is a vector subspace.

For any $f_0 \in V$ put $V_{f_0} := \{g \in B_b(\Omega) \mid f_0 + g \in V\}$. Note that V_{f_0} possesses property (b).

If $f_0 \in C(\Omega)$, then $C(\Omega) \subseteq V_{f_0}$ and V_{f_0} satisfies (b); hence $V_{f_0} \in \mathscr{S}$ and $V \subseteq V_{f_0}$, i.e.,

$$f_0 \in C(\Omega), g \in V \quad \implies \quad f_0 + g \in V.$$

Thus, if $g_0 \in V$, then $f + g_0 \in V$ for every $f \in C(\Omega)$; therefore $C(\Omega) \subseteq V_{g_0}$ and, furthermore, $V_{g_0} \in \mathscr{S}$; this in turn implies $V \subseteq V_{g_0}$, i.e.,

$$g_0 \in V, g \in V \quad \Rightarrow \quad g \in V_{g_0} \quad \Rightarrow \quad g_0 + g \in V.$$

It remains to show that $\alpha g \in V$, if $\alpha \in \mathbb{C}$ and $g \in V$: Fix an arbitrary $\alpha \in \mathbb{C}$ and put $V_\alpha := \{g \in B_b(\Omega) \mid \alpha g \in V\}$. Then $C(\Omega) \subseteq V_\alpha$ and V_α satisfies (b), hence $V \subseteq V_\alpha$, i.e.

$$g \in V \quad \implies \quad \alpha g \in V.$$

Claim 2: The step functions are contained in V.

From Claim 1 we already know that V is a vector space, hence it suffices to show $\chi_E \in V$ for any Borel set $E \in \mathcal{B}(\Omega)$. Let $\Delta := \{E \in \mathcal{B}(\Omega) \mid \chi_E \in V\}$.

If $E \subseteq \Omega$ is open (in the subspace topology of Ω), then there is a sequence (f_n) of functions $f_n \in C(\Omega)$ with $0 \leq f_n \leq 1$ and converging pointwise to χ_E. (Choose a compact exhaustion

$E = \bigcup_{n\in\mathbb{N}} A_n$ and let f_n be an Urysohn function for the closed pair A_n and $\Omega \setminus E$.) Property (b) applied to (f_n) yields $\chi_E \in V$, thus $E \in \Delta$. Therefore Δ contains the topology of Ω, which is a generating system for the Borel sigma algebra $\mathcal{B}(\Omega)$ stable under finite intersections.

By the theory of Dynkin systems from measure theory ([Els11, Kapitel I, §6.2] or [Bau01, Chapter I, §2]), we obtain $\Delta = \mathcal{B}(\Omega)$, if we prove the following two properties:
(1) $E \in \Delta \Rightarrow \Omega \setminus E \in \Delta$,
(2) if (E_n) is a sequence of pairwise disjoint sets in Δ, then $E := \bigcup_{n\in\mathbb{N}} E_n \in \Delta$.

Property (1) holds, since 1 and χ_E belong to V and therefore also $\chi_{\Omega\setminus E} = 1 - \chi_E$. To show (2) we simply note that $\chi_E = \sum \chi_{E_n}$ in the sense of pointwise convergence, hence (b) implies $\chi_E \in V$.

It follows from Claims 1 and 2 that V is $\|.\|_\infty$-dense in $B_b(\Omega)$. By property (b), every $S \in \mathscr{S}$ is a closed subset of $B_b(\Omega)$, hence V is also closed and therefore $V = B_b(\Omega)$. □

Sketch of the construction in 1.2 for compact normal operators

The unitary equivalence with a multiplication operator can be constructed along the following lines: There is an orthonormal (possibly finite) sequence of eigenvectors of T, say $W = \{w_1, w_2, \ldots\} = \{w_j \mid j \in N\}$, where $N = \mathbb{N}$ or $N = \{1, \ldots, m\}$ for some $m \in \mathbb{N}$, corresponding to the sequence $d\colon N \to \mathbb{C}$ of all non-zero eigenvalues $d(1), d(2), \ldots$ (with multiplicities) of T, which is either finite or converges to 0 ([Wer18, Theorem VI.3.2]), such that $H = \ker(T) \oplus \overline{\operatorname{span}(W)}$ (orthogonal direct sum). If $\ker(T) \neq \{0\}$ extend W by an orthonormal system $K = \{y_r \mid r \in R\}$ of $\ker(T)$, where $R \cap N = \emptyset$, to obtain an orthonormal system $B = W \cup K$ of H and write $B = \{b_s \mid s \in S\}$, where $S = N \cup R$ and $b_s := w_j$, if $s = j \in N$, and $b_s := y_r$, if $s = r \in R$. Let $e_s \in l^2(S)$ be the complete orthonormal system given by $e_s(t) = 1$, if $t = s$, and $e_s(t) = 0$ otherwise, and define U by unique (uniformly continuous) extension of $Ue_s := b_s$ $(s \in S)$ to $l^2(S) = \overline{\operatorname{span}(\{e_s \mid s \in S)}$. Then $TUe_s = 0$, if $s \in R$, and $TUe_s = d(s)\, b_s$, if $s \in N$, and therefore we have for any $x \in l^2(S)$, $(U^{-1}TUx)(s) = 0 = 0\, x(s)$, if $s \in R$, and $(U^{-1}TUx)(s) = d(s)\, x(s)$, if $s \in N$; in other words, $(U^{-1}TUx)(s) = h(s)\, x(s)$, where $h \in l^\infty(S)$ is given by $h\,|_R = 0$ and $h\,|_N = d$.

We restate and prove Theorem 1.5

Theorem (Continuous functional calculus): If $T \in L(H)$ is self-adjoint, then there is a unique map $\Phi\colon C(\sigma(T)) \to L(H)$ with the following properties:

(a) $\Phi(\mathrm{id}) = T$ and $\Phi(1) = I$,

(b) Φ is an involutive algebra homomorphism, i.e.,

- Φ is $\mathbb{C}$-linear,
- $\forall f, g \in C(\sigma(T))$: $\Phi(f \cdot g) = \Phi(f) \cdot \Phi(g)$,
- $\Phi(\bar{f}) = \Phi(f)^*$,

(c) Φ is continuous with respect to the norm $\|.\|_\infty$ on $C(\sigma(T))$, in fact, isometric, i.e., $\|\Phi(f)\| = \|f\|_\infty$.

We write $f(T)$ instead of $\Phi(f)$ and call $f \mapsto f(T)$ the *continuous functional calculus* of T.

Proof: Uniqueness: Because of (c) and thanks to the Stone-Weierstraß theorem applied to the compact subset $\sigma(T) \subseteq \mathbb{R}$ (cf., e.g., [Wer18, Satz VIII.4.7]), Φ is determined by the values on polynomial functions; then by linearity, the knowledge of $\Phi(\mathrm{id}^n)$ suffices; finally, by multiplicativity according to (b), this boils down to fixing $\Phi(\mathrm{id})$ and $\Phi(1)$, which are determined by (a).

Existence: If $f \in C(\sigma(T))$ is the restriction of a polynomial function $p(t) = \sum_{k=0}^n a_k t^k$, then we put $\Phi_0(f) := p(T) = \sum_{k=0}^n a_k T^k$. Before we proceed, we have to show that Φ_0 is well-defined[6], i.e., if two polynomials p and q satisfy $p\,|_{\sigma(T)} = q\,|_{\sigma(T)}$, then $p(T) = q(T)$. This will be clarified once we have shown the following two claims for a polynomial p:

(i) $\sigma(p(T)) = \{p(\lambda) \mid \lambda \in \sigma(T)\} = p(\sigma(T))$,

(ii) $\|p(T)\| = \sup\limits_{\lambda \in \sigma(T)} |p(\lambda)| = \|p\,|_{\sigma(T)}\|_\infty$.

Proof of (i): This is obvious for constant polynomials, hence we may suppose $\deg(p) \geq 1$.

Let $\lambda \in \sigma(T)$. We may write $p(t) - p(\lambda) = (t - \lambda)h(t)$ with some polynomial $h \neq 0$ and obtain $p(T) - p(\lambda) = (T - \lambda)h(T)$. Due to the commutativity of $\lambda - T$ with $h(T)$ and $p(T) - p(\lambda)$, continuous invertibility of $p(T) - p(\lambda)$ would then imply the same for $T - \lambda$, hence $p(\lambda) \in \sigma(p(T))$. Thus, $p(\sigma(T)) \subseteq \sigma(p(T))$.

Let $\mu \in \sigma(p(T))$, then $\deg(p - \mu) \geq 1$ and we have a polynomial factorization $p(t) - \mu = c(t - \lambda_1) \cdots (t - \lambda_m)$ with $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ and $c \in \mathbb{C} \setminus \{0\}$, which yields $p(T) - \mu = c(T - \lambda_1) \cdots (T - \lambda_m)$. If λ_j belonged to $\rho(T)$ for every $j = 1, \ldots, m$, then $p(T) - \mu$ would be continuously invertible, a contradiction; hence $\lambda_j \in \sigma(T)$ for some j and therefore $\mu = p(\lambda_j) \in p(\sigma(T))$. Thus, $\sigma(p(T)) \subseteq p(\sigma(T))$.

Proof of (ii): Elementary calculations show that $p \mapsto p(T)$ is an involutive algebra homomorphism from polynomials into $L(H)$. Note that $p(T)$ is normal, since $p(T)^* p(T) = \overline{p}(T)p(T) = (\overline{p}p)(T) = (p\overline{p})(T) = p(T)p(T)^*$, and thus has norm equal to its spectral radius ([Wer18, Satz VI.1.7]). We obtain

$$\|p(T)\| = \sup\{|\mu| \mid \mu \in \sigma(p(T))\} = \sup\{|p(\lambda)| \mid \lambda \in \sigma(T)\} = \sup_{\lambda \in \sigma(T)} |p(\lambda)|.$$

We have now shown that the map Φ_0 is well-defined on the space V_0 of polynomial functions on $\sigma(T)$ and gives an involutive algebra homomorphism into $L(H)$ which is isometric due to Claim (ii), since $\|\Phi_0(f)\| = \|f\|_\infty$ for any $f \in V_0$; in particular, Φ_0 is continuous. Let Φ denote the unique continuous linear extension of Φ_0 to $C(\sigma(T))$ (making use of the density of V_0 due to the Stone-Weierstraß theorem). That Φ is isometric follows directly from the density of the polynomial functions and Claim (ii) above. It remains to show that Φ is involutive and multiplicative, which follows from routine arguments using limits of polynomials and the fact that Φ_0 is already known to be involutive and multiplicative. □

[6]Note that $\sigma(T)$ can be a finite discrete subset of $\mathbb{R}$.

Bibliography

[Ara99] H. Araki: *Mathematical Theory of Quantum Fields.*
Oxford University Press 1999.

[Bau90] H. Bauer: *Maß- und Integrationstheorie.*
De Gruyter 1990.

[Bau01] H. Bauer: *Measure and Integration Theory.*
De Gruyter 2001.

[Bog07] V. I. Bogachev: *Measure Theory. Volume I*
Springer-Verlag 2007.

[Coh80] D. L. Cohn: *Measure Theory.*
Birkhäuser 1980.

[Con16] A. Constantin: *Fourier Analysis. Part I – Theory.*
Cambridge University Press 2016.

[Con10] J. B. Conway: *A Course in Functional Analysis.*
Springer Science+Business Media, 2nd edition 2010.

[Els11] J. Elstrodt: *Maß- und Integrationstheorie.*
Springer-Verlag, 7. Auflage 2011.

[Fol99] G. B. Folland: *Real Analysis.*
Wiley, 2nd edition 1999.

[Hal13] B. C. Hall: *Quantum Theory for Mathematicians.*
Springer-Verlag 2013.

[Hoe21] G. Hörmann: *Funktionalanalysis.*
Vorlesung, Universität Wien, Sommersemester 2021.
`http://www.mat.univie.ac.at/~gue/material.html`

[Hoe20] G. Hörmann: *Grundbegriffe der Topologie.*
Vorlesung, Universität Wien, Wintersemester 2020/21.
`http://www.mat.univie.ac.at/~gue/material.html`

[Kab14] W. Kaballo: *Aufbaukurs Funktionalanalysis und Operatortheorie.*
Springer-Verlag 2014.

[KR] R. V. Kadison and J. R. Ringrose: *Fundamentals of the theory of operator algebras.* Volumes I and II.
Academic Press, 1983 and 1986.

[MV92] R. Meise and D. Vogt: *Einführung in die Funktionalanalysis.*
Vieweg Verlag 1992.

[RS75] M. Reed and B. Simon: *Methods of modern mathematical physics Volume II: Fourier analysis, self-adjointness.*
Academic Press 1975.

[Rud86] W. Rudin: *Real and Complex Analysis.*
McGraw-Hill, 3rd edition 1986.

[Sch71] H. Schaefer: *Topological Vector Spaces.*
Springer-Verlag 1971.

[Sch00] H. Schröder: *Funktionalanalysis.*
Harri Deutsch, 2. Auflage 2000.

[Tes14] G. Teschl: *Topics in Real and Functional Analysis.*
Lecture Notes, University of Vienna, winter term 2013/14.
`http://www.mat.univie.ac.at/~gerald/ftp/book-fa/index.html`

[Tes14b] G. Teschl: *Mathematical Methods in Quantum Mechanics. With Applications to Schrödinger Operators*
American Mathematical Society, 2nd edition 2014.

[Thi03] W. Thirring: *Quantum Mathematical Physics: Atoms, Molecules and Large Systems.*
Springer-Verlag, 2nd edition 2003.

[Tre67] F. Trèves: *Topological vector spaces, distributions and kernels.*
Academic Press 1967.

[Wei00] J. Weidmann: *Lineare Operatoren in Hilberträumen: Teil I: Grundlagen.*
Teubner-Verlag 2000.

[Wer18] D. Werner: *Funktionalanalysis.*
Springer-Verlag, 8. Auflage 2018.

[Wil70] S. Willard: *General Topology.*
Addison-Wesley 1970.