Functional Analysis 2

Elke Rosenberger

University of Potsdam

Lecture Notes

Summer Term 2020

Preface

These lecture notes result from the second part of a two term course in functional analysis. Therefore, there are sometimes references to subjects which were introduced and discussed in the first part of the course. In order to improve the readability, some of these results are shortly recapitulated.
These notes are mostly based on Dirk Werner:" Funktionalanalysis" [W] and Michael Reed and Barry Simon:"Functional Analysis" [RS].

The starting point in this course was the spectral theorem for bounded self-adjoint operators. The continuous functional calculus version was already proven in the end of Functional Analysis 1. After giving a functional calculus for measurable functions, we introduce spectral projections and projection valued measures. The last result given in this context is the fact that each bounded self-adjoint operator is unitary equivalent to a multiplication operator on some Hilbert space. A generalisation to normal bounded operators is given in the appendix. It is based on a talk of Jan Möhring, who agreed to add his handout to this text.

The second chapter gives definitions and properties of unbounded operators on Hilbert spaces. In particular, we discuss the notion of symmetric, closed and (essentially) self-adjoint operators, give criteria for a symmetric operator to be (essentially) self-adjoint and define the Friedrichs extension. In the next section, the different versions of the spectral theorem described above are generalized to unbounded self-adjoint operators. In the third section, strongly continuous one-parameter semigroups of operators on a Banach space and in particular contraction semigroups are introduced and the theorems of Hille-Yosida and Lumer-Phillips are proven. These theorems give criteria for an operator to be the infinitesimal generator of a strongly continuous semigroup. Then we consider continuous unitary groups of operators on a Hilbert space and prove Stone's Theorem. In the last part of this chapter, we discuss the notion of commuting unbounded operators and Trotter's product formula.

The third chapter includes definitions and properties of locally convex spaces and Fréchet spaces, in the end, we briefly introduce distributions and the Fourier transform. The appendix includes the handouts for talks given by Jan Möhring and Pushya Mitra on the spectral theorem for normal operators and on the Gelfand-Naimark-Theorem respectively at the end of the term.

Exercises are given at the end of each section. For some of them Jan Möhring provided solutions which are given in the appendix.

I am pleased to thank the students who took part in the course and helped to improve these notes with their questions and comments. Especially I want to thank Jan Möhring for his constructive remarks, numerous suggestions and corrections and his consent to make some of his solutions for exercises available. Moreover I thank Jan Möhring and Pushya Mitra for their approval to add to this script the handouts of their talks.

Contents

Chapter 1

Bounded operators on Hilbert spaces

1.1 Spectral Theorem for bounded operators

1.1.1 Results from the previous Semester

We continue our analysis of the spectrum of self-adjoint (or normal) operators started in "Functional analysis 1" and give a recap on definitions and results we discussed there.

We start with a short reminder on the spectrum of operators.

Definition 1.1 (Spectrum of a bounded operator)
Let X be a Banach space over $\mathbb{K}$ and $T \in \mathcal{L}(X)$ a bounded linear operator on X.

i) A complex number $\lambda \in \mathbb{C}$ is said to be in the **resolvent set** *$\rho(T)$ of*

T $:\Leftrightarrow \lambda \operatorname{Id} - T$ is a bijection (with bounded inverse).

ii) $R_\lambda(T) := (\lambda \operatorname{Id} - T)^{-1}$ is called the **resolvent** *of T at $\lambda \in \rho(T)$.*

iii) If $\lambda \notin \rho(T)$, then λ is said to be in the **spectrum** *$\sigma(T) = \mathbb{C} \setminus \rho(T)$ of T.*

iv) A vector $x \in X$, $x \neq 0$, is called **eigenvector** *of T $:\Leftrightarrow Tx = \lambda x$ for some $\lambda \in \mathbb{C}$.*
In this case, λ is called corresponding **eigenvalue**. *(Then $\lambda \operatorname{Id} - T$ is not injective and thus in particular $\lambda \in \sigma(T)$). The set of all eigenvalues is called* **point spectrum** *$\sigma_p(T)$ of T.*

v) If λ is not an eigenvalue (i.e. $\lambda \operatorname{Id} - T$ is injective), but $\lambda \operatorname{Id} - T$ is not surjective and the range $\operatorname{Ran}(\lambda \operatorname{Id} - T) \subset X$ is dense, then λ is said to be in the **continuous spectrum** *$\sigma_c(T)$ of T.*

vi) If λ is not an eigenvalue and $\lambda \operatorname{Id} - T$ is not surjective, but $\operatorname{Ran}(\lambda \operatorname{Id} - T)$ is not dense in X, then λ is said to be in the **residual spectrum** *$\sigma_r(T)$ of T.*

In Theorem 80, we proved that the resolvent set $\rho(T) \subset \mathbb{C}$ is open and the map

$$R_\cdot(T) : \rho(T) \longrightarrow \mathcal{L}(X), \quad \lambda \mapsto R_\lambda(T)$$

is analytic.
Moreover, for any $\lambda, \mu \in \rho(T)$ the operators $R_\lambda(T)$ and $R_\mu(T)$ commute and

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda) R_\lambda(T) R_\mu(T) \qquad \textbf{First Resolvent Formula} \tag{1.1}$$

Lemma 79 told us that if $\sum_{n=0}^{\infty} T^n$ converges in $\mathcal{L}(X)$ (in particular if $\|T\| < 1$), then $(\mathrm{Id} - T)$ is invertible and

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n. \tag{1.2}$$

These two results imply (Corollary 81) that $\sigma(T)$ is compact, $|\lambda| \leq \|T\|$ for all $\lambda \in \sigma(T)$ and if $\mathbb{K} = \mathbb{C}$, then $\sigma(T) \neq \emptyset$. In particular, for $|\lambda| > \|T\|$

$$(\lambda \,\mathrm{Id} - T)^{-1} = \frac{1}{\lambda} \sum_{n=0}^{\infty} \Big(\frac{T}{\lambda}\Big)^n \qquad \textbf{Von Neumann Series} \tag{1.3}$$

In Theorem 84 we proved for the **spectral radius** of T given by $r(T) := \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$ that

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n} \tag{1.4}$$

$$|\lambda| \leq r(T) \quad \text{ for any } \quad \lambda \in \sigma(T) \tag{1.5}$$

$$\exists \lambda \in \sigma(T) \,:\, |\lambda| = r(T) \quad \text{if} \quad \mathbb{K} = \mathbb{C}. \tag{1.6}$$

If X is a Hilbert space and T is self-adjoint, then $r(T) = \|T\|$ (Prop.85). This implies that if $\mathbb{K} = \mathbb{C}$, there exists $\lambda \in \sigma(T)$ such that $|\lambda| = \|T\|$.

For a Banach space X with dual space $X^* = \mathcal{L}(X, \mathbb{K})$ and $T \in \mathcal{L}(X)$, the adjoint operator (or "Banach space adjoint") $T' \in \mathcal{L}(X^*)$ is defined by the relation

$$(T'\ell)(x) = \ell(Tx)\,, \qquad \ell \in X^*,\, x \in X\,.$$

The map $T \mapsto T'$ is an isometric isomorphism onto its range (in general not surjective).

Then Proposition 86 and 87 showed

$$\sigma(T) = \sigma(T'), \qquad R_\lambda(T') = R_\lambda(T)', \tag{1.7}$$
$$\sigma_r(T) \subset \sigma_p(T') \quad \text{and} \quad \sigma_p(T) \subset \big(\sigma_p(T') \cup \sigma_r(T')\big).$$

If $\mathscr{H}$ is a Hilbert space, the Riesz Lemma shows that the map

$$C : \mathscr{H} \to \mathscr{H}^*, \qquad y \mapsto C(y) = \langle y, \cdot \rangle$$

is isometric, surjective and conjugate linear.
It allows to define for $T \in \mathcal{L}(\mathscr{H})$ the adjoint operator (or "Hilbert space adjoint") $T^* \in \mathcal{L}(\mathscr{H})$ by $T^* = C^{-1}T'C$. Then T^* satisfies the relation

$$\langle x, Ty \rangle = \langle T^*x, y \rangle, \qquad x, y \in \mathscr{H}$$

and the map $T \mapsto T^*$ is a conjugate linear (i.e. $\alpha T \mapsto \bar{\alpha} T^*$) isometric isomorphism on $\mathcal{L}(\mathscr{H})$.

Then by Proposition 86 we have

$$\sigma(T^*) = \{\lambda \,|\, \bar{\lambda} \in \sigma(T)\} \quad \text{and} \quad R_\lambda(T^*) = R_{\bar{\lambda}}(T)^*. \tag{1.8}$$

In Theorem 88 we considered a bounded self-adjoint operator A on a Hilbert space $\mathscr{H}$. We saw that

$$\sigma_r(A) = \emptyset \qquad \text{and} \qquad \sigma(A) \subset \mathbb{R} \tag{1.9}$$

and that eigenvectors associated to different eigenvalues are orthogonal.

We already gave the following version of the Spectral Theorem:

Theorem 1.2 (Continuous Functional Calculus)
Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. Then there exists a unique map $\Phi_A : \mathcal{C}(\sigma(A)) \to \mathcal{L}(\mathscr{H})$ with the following properties:

i) $\Phi_A(1) = \mathrm{Id}$ *and* $\Phi_A(x) = A$ *(here* 1 *denotes the function* $f(x) = 1$ *for all* x *and* x *denotes the function* $f(x) = x$*).*

ii) Φ_A *is an algebraic* $*$*-homomorphism (with respect to multiplication in* $\mathcal{C}(\sigma(A))$ *and composition in* $\mathcal{L}(\mathscr{H})$*), i.e. for all* $f, g \in \mathcal{C}(\sigma(A))$ *and* $\lambda \in \sigma(A)$

$$\Phi_A(f+g) = \Phi_A(f) + \Phi_A(g) \quad \text{and} \quad \Phi_A(\lambda f) = \lambda \Phi_A(f) \qquad \textit{linear}$$

$$\Phi_A(f \cdot g) = \Phi_A(f) \circ \Phi_A(g) \qquad \textit{multiplicative}$$

$$\Phi_A(\bar{f}) = \Phi_A(f)^* \qquad \textit{involutive}$$

iii) Φ_A *is continuous.*

Moreover, Φ_A *has the following additional properties:*

iv) If $A\psi = \lambda\psi$*, then* $\Phi_A(f)\psi = f(\lambda)\psi$*.*

v) Spectral Mapping Theorem: $\sigma(\Phi_A(f)) = \{f(\lambda) \,|\, \lambda \in \sigma(A)\}$*.*

vi) Φ_A *is positive (preserving), i.e.* $f \geq 0$ *implies* $\Phi_A(f) \geq 0$ *.*

vii) Φ_A *is isometric, i.e.* $\|\Phi_A(f)\|_{\mathcal{L}} = \|f\|_\infty$*.*

Here we used for $T \in \mathcal{L}(\mathscr{H})$ the notation $T \geq 0$ if $\langle x, Tx \rangle \geq 0$ for all $x \in \mathscr{H}$.

We will sometimes write $\Phi_A(f) =: f(A)$. Then $v)$ is in short notation $\sigma(f(A)) = f(\sigma(A))$.

For compact self-adjoint operators, we have a more explicit form of the Spectral Theorem. To derive it, we start with a reminder on the Riesz-Schauder-Theorem (101) and the Hilbert-Schmidt-Theorem (102).

Theorem 1.3 (Riesz-Schauder-Theorem)
Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{K}(\mathscr{H})$ a self-adjoint compact operator.

Then $\sigma(A)$ is a discrete set having no limit points except perhaps 0. Furthermore, any non-zero $\lambda \in \sigma(A)$ is an eigenvalue of finite multiplicity (i.e. the corresponding eigenspace E_λ spanned by the eigenvectors is finite dimensional).

We remark that if $\mathbb{K} = \mathbb{C}$, then

Theorem 1.4 (Hilbert-Schmidt-Theorem)
Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{K}(\mathscr{H})$ a self-adjoint compact operator. Then there exist a complete orthonormal system $\{\phi_\alpha\}_{\alpha \in I}$ for $\mathscr{H}$, a countable or finite subset $I_0 = \{\alpha_n\} \subset I$ and a null sequence $(\lambda_{\alpha_n})_{n \in \mathbb{N}}$ in $\mathbb{K} \setminus \{0\}$ such that

$$A\phi_{\alpha_n} = \lambda_{\alpha_n}\phi_{\alpha_n} \qquad \text{and} \qquad A\phi_\beta = 0 \quad \text{if} \quad \beta \in I \setminus I_0 \,.$$

These theorems allow to show:

Theorem 1.5 (Spectral Theorem for compact Operators)
Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{K}(\mathscr{H})$ a self-adjoint compact operator.

Then there exists an orthonormal system $\{\phi_n\}$ *and a null sequence* (λ_n) *in* $\mathbb{K}\setminus\{0\}$ *(both maybe finite) such that*

$$\mathscr{H} = \ker A \oplus \overline{\operatorname{span}}\{\phi_n\} \qquad \textit{and} \qquad A\psi = \sum_n \lambda_n \langle \phi_n, \psi\rangle \phi_n, \tag{1.10}$$

in particular $A\phi_n = \lambda_n\phi_n$ *and* $\|A\| = \max_n |\lambda_n|$.
Denoting by Π_k *the orthogonal projection to the eigenspace*

$$.E_{\lambda_k} := \ker(\lambda_k \operatorname{Id} - A)$$

of λ_k *and* $\lambda_0 = 0$ *we get*

$$\mathscr{H} = \bigoplus_{n=0}^{\infty} E_{\lambda_n} \qquad \textit{and} \qquad A = \sum_{n=1}^{\infty} \lambda_n \Pi_n \tag{1.11}$$

and the sum converges in operator norm.

Before we start to prove Theorem 1.5, we give a short reminder on orthogonal projections:
In Theorem 40 (the Projection Theorem) we have seen that if $\mathscr{H}_1$ is a closed subspace of the Hilbert space $\mathscr{H}$, then every $\psi \in \mathscr{H}$ can uniquely be written as $\psi = \psi_1 + \psi_2$ where $\psi_1 \in \mathscr{H}_1$ and $\psi_2 \in \mathscr{H}_1^{\perp}$ (here $\mathscr{H}_1^{\perp} = \{\phi \in \mathscr{H} \mid \forall\varphi \in \mathscr{H}_1 : \langle \phi, \varphi\rangle = 0\}$ denotes the orthogonal complement on $\mathscr{H}_1$). Then we called the linear maps

$$\Pi_{\mathscr{H}_1} : \mathscr{H} \to \mathscr{H}_1\,, \quad \Pi_{\mathscr{H}_1}\psi = \psi_1 \quad \text{and}$$
$$\Pi^{\perp}_{\mathscr{H}_1} = \Pi_{\mathscr{H}_1^{\perp}} : \mathscr{H} \to \mathscr{H}_1^{\perp}\,, \quad \Pi^{\perp}_{\mathscr{H}_1}\psi = \psi_2$$

orthogonal projection to $\mathscr{H}_1$ and $\mathscr{H}_1^{\perp}$ respectively and proved $\|\Pi_{\mathscr{H}_1}\| = 1$.
If $\{e_\alpha\}$ is an orthonormal basis of $\mathscr{H}_1$, then $\Pi_{\mathscr{H}_1} = \Pi$ where

$$\Pi\psi := \sum_\alpha \langle e_\alpha, \psi\rangle e_\alpha\,, \qquad \psi \in \mathscr{H}. \tag{1.12}$$

To see this, it suffices to prove that $\psi - \Pi\psi \in \mathscr{H}_1^{\perp}$, because the partition of ψ given above is unique. But for any e_β

$$\left\langle e_\beta, \psi - \sum_\alpha \langle e_\alpha, \psi\rangle e_\alpha \right\rangle = \langle e_\beta, \psi\rangle - \sum_\alpha \langle e_\alpha, \psi\rangle\langle e_\beta, e_\alpha\rangle = 0$$

(since $\langle e_\beta, e_\alpha\rangle = \delta_{\alpha,\beta}$).

Proposition 70 told us that for each orthogonal projection Π on a closed subspace we have $\Pi^2 = \Pi$ and $\Pi^* = \Pi$. Moreover we even have the reverse, i.e. if $\Pi \in \mathcal{L}(\mathscr{H})$ such that $\Pi^2 = \Pi$ and $\Pi^* = \Pi$ holds, then the range of Π is a closed subspace of $\mathscr{H}$ and Π is the orthogonal projection its range.

Proof. Consider the basis $\{\phi_\alpha\}_{\alpha\in I}$ for $\mathscr{H}$ described in Theorem 1.4. Set $\phi_n := \phi_{\alpha_n}$ and $\lambda_n := \lambda_{\alpha_n} \neq 0$ for $\alpha_n \in I_0$, then (λ_n) is a null sequence in $\mathbb{K} \setminus \{0\}$ and the representation of $\mathscr{H}$ in (1.10) follows immediately from the fact that $\phi_\alpha \in \ker A$ for $\alpha \in I \setminus I_0$.
Moreover, by Theorem 96, each $\psi \in \mathscr{H}$ can be written as $\psi = \sum_{\alpha\in I}\langle \phi_\alpha, \psi\rangle\phi_\alpha$ and the summands are non-zero only for a countable subset of I. But since $A\phi_\alpha = 0$ for $\alpha \notin I_0$, it follows that

$$A\psi = A\sum_{\alpha\in I}\langle \phi_\alpha, \psi\rangle\phi_\alpha = \sum_{\alpha_n\in I_0}\langle \phi_{\alpha_n}, \psi\rangle A\phi_{\alpha_n} = \sum_n \lambda_n\langle \phi_n, \psi\rangle\phi_n\,.$$

Moreover this gives the estimate

$$\|A\psi\| = \left\|\sum_n \lambda_n\langle \phi_n, \psi\rangle\phi_n\right\| \leq \max_n |\lambda_n| \left\|\sum_n \langle \phi_n, \psi\rangle\phi_n\right\| \leq \max_n |\lambda_n|\|\psi\|$$

proving that $\|A\| \leq \max_n |\lambda_n|$. Since $|\lambda| \leq \|A\|$ for each $\lambda \in \sigma(A)$ we get equality.

The representation of $\mathscr{H}$ in (1.11) follows at once from (1.10) and the fact that the kernel of $\lambda_k \operatorname{Id} - A$ is spanned by the eigenvalues associated to λ_k (by Theorem 1.3 these eigenspaces are finite dimensional if $\lambda_k \neq 0$).
The last representation of A follows from (1.12) together with (1.10) by combining all summands belonging to the same value of λ.
To show that the sum converges in operator norm we use that $\|\sum_{n\geq 0} \Pi_n\| = 1$ to write

$$\left\| A - \sum_{n=1}^{N} \lambda_n \Pi_n \right\| = \left\| \sum_{n=N+1}^{\infty} \lambda_n \Pi_n \right\| \leq \max_{n>N} |\lambda_n| \longrightarrow 0 \qquad (N \to \infty).$$

□

Since for $A \in \mathcal{L}(\mathscr{H})$ compact and self-adjoint and $f \in \mathcal{C}(\sigma(A))$, the map $f \mapsto \sum_{n=1}^{\infty} f(\lambda_n)\Pi_n$ has the properties i),ii) and iii) from Theorem 1.2 (Exercise 1.32), it describes the map Φ_A, i.e. we have $f(A) = \sum_{n=1}^{\infty} f(\lambda_n)\Pi_n$.

Thus the Continuous Functional Calculus allows to regain the orthogonal projections Π_k as images of A under suitable characteristic functions. In fact, defining for $k \geq 1$ the continuous functions

$$f_k : \sigma(A) = \{0\} \cup \{\lambda_1, \lambda_2, \ldots\} \longrightarrow \mathbb{R} \quad \text{given by} \quad f_k(t) = \begin{cases} 1 & \text{if } t = \lambda_k \\ 0 & \text{otherwise} \end{cases}$$

These functions are continuous because by Theorem 1.3 all elements of the spectrum of A except 0 are isolated points. Then we get

$$f_j(A) = \Phi_A(f_j) = \sum_{n=1}^{\infty} f_j(\lambda_n)\Pi_n = \Pi_j.$$

1.1.2 Functional Calculus Form

To get a similar procedure for a bounded (non-compact) self-adjoint operator A, we have to define an operator $f(A)$ for some $\{0,1\}$-valued functions f on $\sigma(A)$. But in general, if the points of the spectrum are not isolated, these functions will not be continuous. Thus it will be necessary to get a Functional Calculus for the class $\mathcal{B}(\sigma(A))$ of Borel-measurable bounded functions on the (compact) set $\sigma(A)$.

To this end, we first introduce the notion of a spectral measure using the Riesz-Markov-Theorem, which we state here in two versions without proof.

We start with a short reminder on complex measures.

Remark 1.6

A **complex measure** *on a σ-algebra $(X, \mathcal{F})$ is a function $\mu : \mathcal{F} \to \mathbb{C}$ such that for any disjoint partition of $E \in \mathcal{F}$ (i.e. $E = \bigcup_j E_j$ and $E_j \cap E_k = \emptyset$ for $j \neq k$) the equality*

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E_j)$$

holds (in particular the series converges absolutely). We define the **total variation** *of μ by*

$$|\mu|(E) = \sup\Big\{ \sum_{j=1}^{\infty} |\mu(E_j)| \mid \{E_j\} \text{ is a disjoint partition of } E \Big\}.$$

Then $|\mu|$ is the smallest positive measure dominating μ and is in particular finite, i.e. $\|\mu\| := |\mu|(X) < \infty$.

The Radon-Nikodym-Theorem states that if μ is absolutely continuous with respect to a positive σ-finite measure λ (i.e. if $\mu(E) = 0$ whenever $\lambda(E) = 0$, *we write $\mu \ll \lambda$), then there exists a unique $h \in L^1(X)$ such that*

$$\mu(E) = \int_E h \, d\lambda \,, \qquad E \in \mathcal{F}. \tag{1.13}$$

Since by definition $\mu \ll |\mu|$, there exists a unique $h \in L^1(X)$ such that (1.13) *holds for $\lambda = |\mu|$. Moreover it can be shown that in this case $|h(x)| = 1$ for all $x \in X$.*

This suggests to define integration with respect to the complex measure μ by the formula

$$\int f \, d\mu := \int f h \, d|\mu| \,.$$

Then $\int f d\,(\mu_1 + \mu_2) = \int f \, d\mu_1 + \int f \, d\mu_2$ for any two complex measures μ_1, μ_2.

Now let X be a locally compact Hausdorffspace and $\mathcal{F}$ the Borel-σ-algebra. We call a complex Borel measure μ **regular***, if $|\mu|$ is regular i.e. if for all $E \in \mathcal{F}$*

$$\begin{aligned} \inf\{|\mu|(V) \,|\, E \subset V \text{ and } V \text{ is open}\} &= |\mu|(E) \\ &= \sup\{|\mu|(K) \,|\, K \subset E \text{ and } K \text{ is compact}\} \,. \end{aligned}$$

The first equality defines **outer regularity***, the second* **inner regularity** *of $|\mu|$, a regular measure is outer and inner regular.*

Theorem 1.7 (Riesz-Markov-Theorem)
Let X be a Hausdorff space.

i) If X is compact, then for any positive bounded linear functional ℓ on $\mathcal{C}(X)$ there is a unique regular Borel measure $\mu \in \mathscr{M}(X)$ on X with

$$\ell(f) = \int_X f \, d\mu \,, \qquad (f \in \mathcal{C}(X)). \tag{1.14}$$

ii) If X is locally compact, then for any bounded linear functional ℓ on $\mathcal{C}_0(X)$ there is a unique complex regular Borel measure $\mu \in \mathscr{M}_{\mathbb{C}}(X)$ on X with

$$\ell(f) = \int_X f \, d\mu \,, \qquad (f \in \mathcal{C}_0(X)). \tag{1.15}$$

Moreover, in both cases the operator-norm of ℓ is equal to the total variation $|\mu|$ of μ.

Remark 1.8

We remark that for a given Borel measure $\mu \in \mathscr{M}(X)$, the integral on the right hand side of (1.14) *defines a continuous positive linear functional and analog for complex measures. Moreover each positive linear functional is bounded with $\|\ell\| \leq \ell(1)$ (Exercise 1.33).*

We come back to our goal to extend the functional calculus for self-adjoint operators from continuous to bounded measurable functions.

Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ be self-adjoint. Let $\Phi_A : \mathcal{C}(\sigma(A)) \to \mathcal{L}(\mathscr{H})$ be the unique map given in Theorem 1.2. Then, for any fixed $\psi, \varphi \in \mathscr{H}$, the mapping

$$\mathcal{C}(\sigma(A)) \ni f \mapsto \ell_{\psi,\varphi}(f) := \langle \psi, \Phi_A(f)\varphi \rangle \in \mathbb{C}$$

is a complex valued linear functional $\ell_{\psi,\varphi}$ on $\mathcal{C}(\sigma(A))$ and

$$|\ell_{\psi,\varphi}(f)| \leq \|\Phi_A(f)\|\|\psi\|\|\varphi\| \leq \|f\|_\infty \|\psi\|\|\varphi\|$$

and therefore $\|\ell_{\psi,\varphi}\| \leq \|\psi\|\|\varphi\|$.

Thus by the Riesz-Markov-Theorem, there exists a unique regular complex Borel measure $\mu_{\psi,\varphi}$ on $\sigma(A)$ such that

$$\ell_{\psi,\varphi}(f) = \langle \psi, \Phi_A(f)\varphi \rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_{\psi,\varphi}(\lambda)\,. \tag{1.16}$$

This measure is called **spectral measure** of A associated with the vectors ψ and φ.

Moreover the map

$$\mathscr{H} \times \mathscr{H} \ni (\psi, \varphi) \;\mapsto\; b(\psi, \varphi) := \mu_{\psi,\varphi} \in \mathscr{M}_{\mathbb{C}}(\sigma(A)) \tag{1.17}$$

and for fixed $f \in \mathcal{C}(\sigma(A))$ the map

$$\mathscr{H} \times \mathscr{H} \ni (\psi, \varphi) \;\mapsto\; B_f(\psi, \varphi) := \ell_{\psi,\varphi}(f) \in \mathbb{C} \tag{1.18}$$

are sesquilinear[1] (since $(\psi, \varphi) \mapsto \langle \psi, \Phi_A(f)\varphi \rangle$ is sesquilinear) and bounded with

$$\|b(\psi, \varphi)\| = \|\mu_{\psi,\varphi}\| = \|\ell_{\psi,\varphi}\| \leq \|\psi\|\|\varphi\| \quad \text{and} \quad |B_f(\psi, \varphi)| \leq \|f\|_\infty \|\psi\|\|\varphi\|\,. \tag{1.19}$$

Then by a Corollary of the Riesz Lemma[2] (Corollary 45 in the previous course), there exists a unique bounded operator $\tilde{\Phi}_A(f)$ on $\mathscr{H}$ such that

$$B_f(\psi, \varphi) = \langle \psi, \tilde{\Phi}_A(f)\varphi \rangle\,.$$

[1] A map B on $\mathscr{H} \times \mathscr{H}$ is called **sesquilinear**, if for all $a, b \in \mathbb{C}$ and $x, y, z \in \mathscr{H}$ we have

$$B(x, ay + bz) = aB(x, y) + bB(x, z) \quad \text{and} \quad B(ax + by, z) = \bar{a}B(x, z) + \bar{b}B(y, z).$$

[2]

Lemma 1.9 (Riesz, (Thm 43))

For each bounded linear functional $\ell \in \mathscr{H}^*$ on a Hilbert space $\mathscr{H}$ there is unique vector $y \in \mathscr{H}$ such that $\ell(x) = \langle y, x \rangle$ for all $x \in \mathscr{H}$ and $\|\ell\| = \|y\|$.

By (1.16), we immediately see that $\tilde{\Phi}_A(f) = \Phi_A(f)$.

The procedure by which we regained the operator $\Phi_A(f)$ now allows to extend the Continuous Functional Calculus (Theorem 1.2) to the bounded Borel measurable functions $\mathcal{B}(\sigma(A))$ on $\sigma(A)$:

The integral on the right hand side of (1.16) makes sense, if the continuous function f is replaced by a measurable function $g \in \mathcal{B}(\sigma(A))$. This allows to define for any fixed $g \in \mathcal{B}(\sigma(A))$ a sesquilinear form $\tilde{B}_g$ on $\mathscr{H}$ by

$$\tilde{B}_g(\psi, \varphi) := \int_{\sigma(A)} g(\lambda)\, d\mu_{\psi,\varphi}(\lambda)\,. \tag{1.20}$$

As before, using Corollary 1.10, there exists a bounded operator $\tilde{\Phi}_A(g) \in \mathcal{L}(\mathscr{H})$ such that

$$\tilde{B}_g(\psi, \varphi) = \langle \psi, \tilde{\Phi}_A(g)\varphi \rangle \tag{1.21}$$

holds for all $\psi, \varphi \in \mathscr{H}$. By this procedure, we have constructed a map $\tilde{\Phi}_A$ from $\mathcal{B}(\sigma(A))$ to $\mathcal{L}(\mathscr{H})$, extending Φ_A given in Theorem 1.2.

The properties of this map are given in the following Theorem.

Corollary 1.10

For any sesquilinear bounded form B on $\mathscr{H}$ (i.e. sesquilinear bounded map from $\mathscr{H} \times \mathscr{H}$ to $\mathbb{C}$) there is a unique operator $A \in \mathcal{L}(\mathscr{H})$ such that $B(x, y) = \langle y, Ax \rangle$ holds for all $x, y \in \mathscr{H}$. The norm of A is the smallest constant $C > 0$ so that $B(x, y) \leq C\|x\|\|y\|$.

Theorem 1.11 (Spectral Theorem - Functional Calculus Form) *Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. Then there exists a unique map $\tilde{\Phi}_A : \mathcal{B}(\sigma(A)) \to \mathcal{L}(\mathscr{H})$ with the following properties:*

i) $\tilde{\Phi}_A(1) = \mathrm{Id}$ *and* $\tilde{\Phi}_A(\mathrm{id}) = A$ *where* $1(x) = 1$ *and* $\mathrm{id}(x) = x$ *for all* $x \in \sigma(A)$.

ii) $\tilde{\Phi}_A$ *is an algebraic $*$-homomorphism (with respect to multiplication in* $\mathcal{B}(\sigma(A))$ *and composition in* $\mathcal{L}(\mathscr{H})$*)*

iii) $\tilde{\Phi}_A$ *is norm continuous:* $\|\tilde{\Phi}_A(f)\|_{\mathcal{L}} \le \|f\|_\infty$.

iv) Suppose $f_n(x) \to f(x)$ *for each* $x \in \sigma(A)$ *and* $\sup_n \|f_n\|_\infty$ *is bounded. Then* $\tilde{\Phi}_A(f_n) \to \tilde{\Phi}_A(f)$ *weakly, i.e.* $\langle \psi, \tilde{\Phi}_A(f_n)\varphi\rangle \longrightarrow \langle \psi, \tilde{\Phi}_A(f)\varphi\rangle$ *for all* $\psi, \varphi \in \mathscr{H}$.

Moreover, $\tilde{\Phi}_A$ *has the following additional properties:*

v) If $A\psi = \lambda\psi$, *then* $\tilde{\Phi}_A(f)\psi = f(\lambda)\psi$.

vi) $\tilde{\Phi}_A$ *is positive, i.e.* $f \ge 0$ *implies* $\tilde{\Phi}_A(f) \ge 0$.

vii) If $BA = AB$, *then* $\tilde{\Phi}_A(f)B = B\tilde{\Phi}_A(f)$.

viii) If f is real-valued, then $\tilde{\Phi}_A(f)$ *is self-adjoint.*

Proof. Step 1: Uniqueness
From Theorem 1.2 we know that on $\mathcal{C}(\sigma(A))$ the map $\tilde{\Phi}_A$ is already uniquely determined by i), ii) and iii). We then use iv) to get uniqueness on $\mathcal{B}(\sigma(A))$. To this end, we first need the following lemma, which proves that the set

$\mathcal{B}(\sigma(A))$ of bounded measurable functions is the smallest family closed under limits of the form iv) (pointwise limits of uniform bounded sequences) containing all of $\mathcal{C}(\sigma(A))^3$:

Lemma 1.12

Let $M \subset \mathbb{C}$ be compact. Let $U \subset \mathcal{B}(M)$ be such that for any sequence (f_n) in U

$$\Big(\sup_n \|f\|_\infty < \infty \text{ and } f(t) := \lim_{n\to\infty} f_n(t) \text{ exists for all } t \in M\Big)$$
$$\implies f \in U \tag{1.22}$$

holds. Then $\mathcal{C}(M) \subset U$ implies $U = \mathcal{B}(M)$.

Proof of Lemma 1.12. Consider the system $\mathcal{S}$ of all sets S of functions such that

$$\mathcal{C}(M) \subset S \subset \mathcal{B}(M) \quad \text{and} \quad (1.22) \text{ holds for all sequences } (f_n) \text{ in } S\,. \tag{1.23}$$

Then $\mathcal{S} \neq \emptyset$, because in particular $\mathcal{B}(M) \in \mathcal{S}$.
We set $V := \bigcap_{S\in\mathcal{S}} S$. Then by definition $\mathcal{C}(M) \subset V$.

[3]This does not imply that each bounded measurable function is a pointwise limit of continuous functions. A counterexample is the function which is 1 on the rationals and zero otherwise. The continuous functions are called Baire functions of class zero. Then all functions which occur as pointwise limits of continuous functions are called of Baire functions of class 1. Pointwise limits of these lead to functions of Baire class 2. In this way one can define Baire functions of class 1, 2, 3, ...The numbers describing the classes are ordinal numbers and thus we can define functions of class ω, the first non-finite ordinal, as the pointwise limit of a sequence of functions each of which belongs to some finite class. From this one can proceed with $\omega+1, \omega+2, \dots \omega^2+1, \dots$. It can be shown that this process stops when one reaches the first non-contable ordinal ω_1.

Moreover V is a vector space: To $f \in V$ set $V_f := \{g \in \mathcal{B}(M) \mid f + g \in V\}$. Now let $f \in \mathcal{C}(M)$. Then $V_f \in \mathcal{S}$ and thus $V \subset V_f$. Thus we have seen that

$$f \in \mathcal{C}(M) \quad \text{and} \quad g \in V \quad \Longrightarrow \quad f + g \in V.$$

But this implies on the other hand that $\mathcal{C}(M) \subset V_g$ for $g \in V$ and since (1.22) holds in V_g this implies $V \subset V_g$.
Thus $f + g \in V$ for all $f, g \in V$.
Moreover $\alpha g \in V$ for all $\alpha \in \mathbb{C}$ and $g \in V$ (Exercise 2.6), thus V is a vector space.
In the next step we want to show that characteristic functions of Borel sets are in V. Let $\mathscr{B}_M$ denote the Borel sets in M and consider the set $\Delta = \{E \in \mathscr{B}_M \mid \chi_E \in V\}$, where χ_E denotes the characteristic function on E.
Then Δ includes all open sets, because the characteristic functions of open sets are pointwise limits of continuous functions (Urysohn's Lemma, see e.g. [R2]). Thus Δ includes a generator of $\mathscr{B}$ which is stable under (countable) intersections.
Therefore by some measure theoretic result (see e.g. [Ba], Dynkin systems), to show $\Delta = \mathscr{B}$ it suffices to prove

i) If $E, F \in \Delta$ and $E \subset F$, then $F \setminus E \in \Delta$.

ii) If $E_1, E_2, \ldots \in \Delta$ are pairwise disjoint, then $E := \bigcup E_n \in \Delta$.

But i) follows from $\chi_{F \setminus E} = \chi_F - \chi_E$ and the fact that V is a vector space and ii) follows from the fact that $\chi_E = \sum_n \chi_{E_n}$ (with pointwise convergence) together with (1.22).
Thus all characteristic functions of Borel sets are in V and since V is a

vector space this implies that all simple functions are in V. But since all measurable functions are pointwise limits of simple functions (even norm-limits), it follows that $V = \mathcal{B}(M)$.

□

We come back to the proof of Theorem 1.11.
Set

$$U_1 := \{f \in \mathcal{B}(\sigma(A)) \,|\, \tilde{\Phi}_A(f) \text{ is uniquely determined by i)-iv) }\},$$

then we have $\mathcal{C}(\sigma(A)) \subset U_1$ as mentioned above. Moreover, if (f_n) is a uniformly bounded sequence in U_1 converging pointwise to $f \in \mathcal{B}(\sigma(A))$, then $\tilde{\Phi}_A(f_n)$ converges weakly to $\tilde{\Phi}_A(f)$ by iv). This determines $\tilde{\Phi}_A(f)$ uniquely and thus $f \in U_1$. It follows that $U_1 \in \mathcal{S}$ and by Lemma 1.12 we can conclude $U_1 = \mathcal{B}(\sigma(A))$.

Step 2: Existence
Construct the bounded sesquilinear form $\tilde{B}_g$ given in (1.20) for $g \in \mathcal{B}(\sigma(A))$ as above, then there exists a unique bounded linear operator $\tilde{\Phi}_A(g)$, bounded by $\|g\|_\infty$ such that (1.21) holds. We have to show that this operator has the properties i)-iv).

Since the functions 1 and id are continuous, i) is already shown. Property iii) follows at once from (1.21), since $|B_g(\psi, \varphi)| \leq \|g\|\|\psi\|\|\varphi\|$. The convergence in iv) follows from the dominated convergence theorem for Lebesgue-integrals. Thus it remains to prove that $\tilde{\Phi}_A$ is an algebraic *-homomorphism. (Exercise 2.19)

Now we come to the proof of the additional properties v) - viii).

v): Assume that $A\psi = \lambda\psi$ holds and set $U_2 := \{f \in \mathcal{B}(\sigma(A)) \,|\, \tilde{\Phi}_A(f)\psi = f(\lambda)\psi\}$. Then $\mathcal{C}(\sigma(A)) \subset U_2$ by Theorem 1.2. Now choose a uniformly bounded sequence (f_n) in U_2 such that $f_n(x) \to g(x)$ for all $x \in \sigma(A)$. Then

$$\tilde{\Phi}_A(f_n)\psi = f_n(\lambda)\psi \longrightarrow g(\lambda)\psi\,.$$

Thus to see $\tilde{\Phi}_A(g)\psi = g(\lambda)\psi$ (which implies $U_2 \in \mathcal{S}$ for $\mathcal{S}$ as given in the proof of Lemma 1.12) we have show that iv) in fact implies strong convergence of $\tilde{\Phi}_A(f_n)$ to $\tilde{\Phi}_A(g)$. This can be seen as follows.

Let h_n be a sequence in $\mathcal{B}$ as given in iv) converging pointwise to $h \in \mathcal{B}$. Then $\bar{h}_n h_n$ converges pointwise to $\bar{h}h$ and $\|\bar{h}_n h_n\|_\infty$ is bounded.
Thus for any $\phi \in \mathscr{H}$ by ii) and iv)

$$\begin{aligned}\|\tilde{\Phi}_A(h_n)\phi\|^2 &= \langle \tilde{\Phi}_A(h_n)\phi, \tilde{\Phi}_A(h_n)\phi\rangle = \langle \tilde{\Phi}_A(h_n)^*\tilde{\Phi}_A(h_n)\phi, \phi\rangle \\ &= \langle \tilde{\Phi}_A(\bar{h}_n h_n)\phi, \phi\rangle \longrightarrow \langle \tilde{\Phi}_A(\bar{h}h)\phi, \phi\rangle = \|\tilde{\Phi}_A(h)\phi\|^2. \end{aligned} \tag{1.24}$$

But in a Hilbert space, we have the general equivalence (for $n \to \infty$)

$$\|\phi_n - \phi\| \to 0 \quad \Longleftrightarrow \quad \Big[\|\phi_n\| \to \|\phi\| \quad \text{and} \quad \langle \phi_n - \phi, \psi\rangle \to 0 \text{ for all } \psi \in \mathscr{H}\Big]. \tag{1.25}$$

This proves that weak convergence of $f_n(A)$ combined with (1.24) implies strong convergence.

Concerning (1.25), the implication " $\Rightarrow$ " follows from the triangle- and the Cauchy-inequality. To get the implication " $\Leftarrow$ ", we write

$$\begin{aligned}\|\phi_n - \phi\|^2 &= \langle \phi_n + \phi - 2\phi, \phi_n - \phi\rangle = \mathrm{Re}\langle \phi_n + \phi, \phi_n - \phi\rangle - 2\,\mathrm{Re}\langle \phi, \phi_n - \phi\rangle \\ &\le \|\phi_n\|^2 - \|\phi\|^2 + 2|\langle \phi, \phi_n - \phi\rangle|\end{aligned}$$

where the last estimate follows from the Polarization Identity.

It follows that $U_2 \in \mathcal{S}$ and thus $U_2 = \mathcal{B}(\sigma(A))$ by Lemma 1.12.

vi) Set $U_3 := \{f \in \mathcal{B}(\sigma(A)) \,|\, f \geq 0 \;\Rightarrow\; \tilde{\Phi}_A(f) \geq 0\}$, then $\mathcal{C}(\sigma(A)) \subset U_3$ by Theorem 1.2. Consider a uniformly bounded sequence (f_n) of positive functions in U_3 converging pointwise to $g \in \mathcal{B}(\sigma(A))$. Then $g(x) \geq 0$ and for each $\psi \in \mathscr{H}$ the sequence $\langle \psi, \tilde{\Phi}_A(f_n)\psi \rangle$ is non-negative and by iv) converges to $\langle \psi, \tilde{\Phi}_A(g)\psi \rangle$, which thus is non-negative. This shows that $U_3 \subset \mathcal{S}$ and again by Lemma 1.12 we get that $\tilde{\Phi}_A(g)$ is a positve operator for any positive $g \in \mathcal{B}(\sigma(A))$.

vii) If $B \in \mathcal{L}(\mathscr{H})$ commutes with A, then by i) and ii) it commutes with $\tilde{\Phi}_A(p)$ for any polynomial p on $\sigma(A)$. Then the assertion follows using approximation arguments (Exercise 1.35).

viii) Exercise 1.35

□

Remark 1.13 *i) We will sometimes use the notation* $\tilde{\Phi}_A(f) = f(A)$.

ii) By the construction given above Theorem 1.11, the operator $\tilde{\Phi}_A(f)$ *only depends on the values of* f *on the spectrum* $\sigma(A)$*, i.e.* $\tilde{\Phi}_A(f) = \tilde{\Phi}_A(\chi_{\sigma(A)}f)$*, where* χ_Ω *denotes the characteristic function on the Borel set* Ω.

iii) For $\Omega \subset \mathbb{R}$*, we always consider* $\mathcal{B}(\Omega)$ *as the subset of functions in* $\mathcal{M}(\mathbb{R})$*, the set of measurable functions on* $\mathbb{R}$*, which are bounded on* Ω.

1.1.3 Projection-valued Measures

We now come to the most important class of functions gained in passing from the continuous functional calculus to the Borel functional calculus, that is, the characteristic functions of Borel sets. We have the following Lemma and Definition.

Definition and Lemma 1.14 (Spectral Projection)
Let A be a bounded self-adjoint operator on a Hilbert space $\mathscr{H}$ and let $\Omega \subset \mathbb{R}$ be a Borel set and χ_Ω the characteristic function of Ω.

Then $\Pi_\Omega := \chi_\Omega(A)(= \tilde{\Phi}_A(\chi_\Omega))$ is an orthogonal projection, called a **spectral projection** *of A.*

Proof. Since χ_Ω is real and $\chi_\Omega^2 = \chi_\Omega$ it follows from the fact that $\tilde{\Phi}_A$ is a $*$-homomorphism, that we have $\Pi_\Omega^2 = \Pi_\Omega$ and $\Pi_\Omega^* = \Pi_\Omega$. □

The following proposition states some properties of the spectral projections of A.

Proposition 1.15 (Properties of Spectral Projections)
Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ be self-adjoint. Let $\mathscr{B}$ be the Borel-σ-algebra on $\mathbb{R}$, then the family of spectral projections $\{\Pi_\Omega \,|\, \Omega \in \mathscr{B}\}$ of A has the following properties:

i) $\Pi_\emptyset = 0$ *and* $\Pi_K = \mathrm{Id}$ *for some compact set* $K \subset \mathbb{R}$.

ii) If $\Omega = \bigcup_{n=1}^\infty \Omega_n$ *and* $\Omega_n \cap \Omega_m = \emptyset$ *for all* $n \neq m$, *then*

$$\sum_{n=1}^{N} \Pi_{\Omega_n} \longrightarrow \Pi_\Omega \quad \textit{strongly as} \quad N \to \infty.$$

iii) $\Pi_{\Omega_1}\Pi_{\Omega_2} = \Pi_{\Omega_1\cap\Omega_2}$.

Proof. i) follows from Theorem 1.11 i), since $\sigma(A) \subset K$ for some compact set $K \subset \mathbb{R}$. ii) follows from Theorem 1.11 iv), taking $f_N = \sum_{n=1}^{N} \chi_{\Omega_n}$, iii) follows from Theorem 1.11ii) together with the fact that $\chi_{\Omega_n}\chi_{\Omega_m} = \chi_{\Omega_n\cap\Omega_m}$. □

Properties i) and ii) remind of the properties of a measure. This makes it reasonable to define:

Definition 1.16 (Projection-valued Measure (Spektralmass))
A family $\Pi := \{\Pi_\Omega \,|\, \Omega \in \mathscr{B}\}$ *of orthogonal projections on* $\mathscr{H}$ *obeying property i) and ii) of Proposition 1.15 is called a* **bounded (or compactly supported) projection-valued measure** *(p.v.m.) (kompakt getragenes Spektralmaß).*
The smallest compact set such that Prop. 1.15,i) holds is called support of the p.v.m (we write $K = \operatorname{supp}\Pi$*).*

From the properties of orthogonal projections, in particular from Proposition 46[4] it follows that each bounded projection valued measure Π has the property iii) of Prop. 1.15.

We will now see, how it is possible to integrate a bounded measurable function with respect to a projection valued measure.

[4]Proposition 46: Let $\mathscr{H}$ be a Hilbert space and A, B be subspaces of $\mathscr{H}$. Let P_A and P_B denote the orthogonal projections on A and B respectively. Then $A \subset B$ if and only if $\|P_A x\| \leq \|P_B x\|$ for all $x \in \mathscr{H}$. In this case $P_A P_B = P_B P_A = P_A$.

The idea is, that if Π is a p.v.m. and $\psi, \phi \in \mathscr{H}$, then

$$\langle \psi, \Pi_{\cdot}\, \phi \rangle \,:\, \mathscr{B} \to \mathbb{C}\,, \quad \Omega \mapsto \langle \psi, \Pi_\Omega \phi \rangle$$

is an complex measure on $\mathscr{B}$.

Thus we can integrate a function $f \in \mathcal{B}(\mathbb{R})$ with respect to this measure and the map

$$\mathcal{B}(\mathbb{R}) \ni f \mapsto \ell_{\psi,\phi}(f) = \int_{\mathbb{R}} f(\lambda)\, d\langle \psi, \Pi_\lambda \phi \rangle \tag{1.26}$$

is a linear functional on $\mathcal{B}(\mathbb{R})$. Since as above (see (1.17)) the map $(\psi, \phi) \mapsto \langle \psi, \Pi_\lambda \phi \rangle$ is sesqulinear, we can, for fixed f, define the sesquilinear form $B_f(\psi, \phi) := \ell_{\psi,\phi}(f)$ on $\mathscr{H}$. Thus again by use of Corollary 1.10, there is a unique bounded self-adjoint linear operator $T(f)$ on $\mathscr{H}$ such that

$$\langle \psi, T(f)\phi \rangle = \int_{\mathbb{R}} f(\lambda)\, d\langle \psi, \Pi_\lambda \phi \rangle \tag{1.27}$$

In fact, the measure $\langle \phi, \Pi \phi \rangle$ is just the measure $\mu_{\psi,\phi}$ associated with $T(f)$ as constructed in (1.16).

To define $\int f d\Pi$ as bounded linear operator on $\mathscr{H}$ for any $f \in \mathcal{B}(\mathbb{R})$, we proceed in three steps (similar to the introduction of the Lebesgue-integral):

Step 1: Let f be a characteristic function, i.e. $f = \chi_\Omega$ for some $\Omega \in \mathscr{B}$. Then we set $\int_{\mathbb{R}} f\, d\Pi = \Pi_\Omega \in \mathcal{L}(\mathscr{H})$.

Step 2: Let f be a simple function, i.e. $f = \sum_{k=1}^n a_k \chi_{\Omega_k}$ for some $a_k \in \mathbb{C}$ and $\Omega_k \in \mathscr{B}$. Then we set

$$\int_{\mathbb{R}} f\, d\Pi = \sum_{k=1}^n a_k \Pi_{\Omega_k} \in \mathcal{L}(\mathscr{H})\,. \tag{1.28}$$

(and this definition is independent of the representation $\sum_{k=1}^n a_k \chi_{\Omega_k}$).

Step 3: Let f be bounded and measurable, i.e. $f \in \mathcal{B}(\mathbb{R})$, then there exists a sequence (f_n) of simple functions converging uniformly to f. Now we use that for any simple function $s = \sum_{k=1}^{n} a_k \chi_{\Omega_k}$, we have $\| \int s \, d\Pi \|_{\mathcal{L}} \leq \|s\|_\infty$. In fact, let $\phi \in \mathscr{H}$ and assume without loss of generality that the sets Ω_k are disjoint, then

$$\begin{aligned}\Big\| \Big(\int s \, d\Pi \Big)(\phi) \Big\|^2 &= \Big\| \sum a_k \Pi_{\Omega_k}(\phi) \Big\|^2 = \sum \|a_k \Pi_{\Omega_k}(\phi)\|^2 \\ &= \sum |a_k|^2 \, \|\Pi_{\Omega_k}(\phi)\|^2 \leq \sup_k |a_k|^2 \sum \|\Pi_{\Omega_k}(\phi)\|^2 \\ &= \|s\|_\infty \Big\| \sum \Pi_{\Omega_k}(\phi) \Big\|^2 = \|s\|_\infty \, \big\| \Pi_{\bigcup \Omega_k}(\phi) \big\|^2 \\ &\leq \|s\|_\infty \|\phi\|^2.\end{aligned}$$

By Theorem 1.11iv) and since (f_n) is a Cauchy sequence w.r.t. $\|.\|_\infty$, this implies that $(\int f_n \, d\Pi)$ is a Cauchy sequence in $\mathcal{L}(\mathscr{H})$ and thus has a limit

$$\int f(\lambda) \, d\Pi_\lambda := \int f \, d\Pi = \lim_{n\to\infty} \int f_n \, d\Pi \tag{1.29}$$

which is independent of the approximating sequence (f_n).

If $\Pi_K = \mathrm{Id}$ for some compact set K and $f \in \mathcal{B}(K)$, we set $\int f \, d\Pi = \int \chi_K f \, d\Pi$ (since $\Pi_A = 0$ whenever $A \cap K = \emptyset$, this definition is independent of the choice of K).

We have proven the following

Theorem 1.17 (Integration w.r.t. projection valued measures) *If* $\Pi = \{\Pi_\Omega \mid \Omega \in \mathscr{B}\}$ *is a bounded projection-valued measure with* $\operatorname{supp} \Pi = K$ *and* $f \in \mathcal{B}(K)$, *then there is a unique operator* $\int f \, d\Pi \in \mathcal{L}(\mathscr{H})$.

The map $f \mapsto \int f\, d\Pi$ *is linear and continuous with* $\| \int f\, d\Pi\|_{\mathcal{L}} \leq \|f\|_\infty$. *Moreover, if* f *is real-valued, then* $\int f\, d\Pi$ *is self-adjoint.*

1.1.4 Projection-valued Measure form

Since restricted to compact sets, the polynomials are bounded functions, we can associate to any given bounded p.v.m. $\{\Pi_\Omega\}$ the self-adjoint operator $T := \int \lambda\, d\Pi_\lambda \in \mathcal{L}(\mathscr{H})$.

Thus on the one hand we have associated a bounded projection-valued measure to a bounded self-adjoint operator. On the other hand, we can associate a bounded self-adjoint operator to a bounded projection-valued measure.

The central point of this section is to see that these two operations are in fact inverse to each other.

Theorem 1.18 (Spectral Theorem - p.v.m. form)
There is a one-one correspondence between bounded self-adjoint operators $A \in \mathcal{L}(\mathscr{H})$ *and bounded projection valued measures* $\Pi = \{\Pi_\Omega \,|\, \Omega \in \mathscr{B}\}$ *given by*

$$A \mapsto \{\Pi_\Omega\} = \{\chi_\Omega(A)\} \tag{1.30}$$

$$\{\Pi_\Omega\} \mapsto A = \int \lambda\, d\Pi_\lambda\,. \tag{1.31}$$

In more detail, if Π *is a bounded p.v.m. on* $\mathbb{R}$ *and* $A \in \mathcal{L}(\mathscr{H})$ *the self-adjoint operator given by* $A = \int \lambda\, d\Pi_\lambda$, *then the unique map* $\tilde{\Phi}_A$ *defined in Theorem*

1.11 is given by

$$\tilde{\Phi}_A(f) = \int f(\lambda)\, d\Pi_\lambda \ , \qquad f \in \mathcal{B}(\sigma(A)). \tag{1.32}$$

If on the other hand $A \in \mathcal{L}(\mathscr{H})$ *is self-adjoint, then there exists a unique bounded projection-valued measure* Π *such that*

$$A = \int_{\sigma(A)} \lambda\, d\Pi_\lambda \tag{1.33}$$

and (1.32) *holds. In this case,* $f(A) = \tilde{\Phi}_A(f)$ *for any* $f \in \mathcal{B}(\sigma(A))$ *is determined by*

$$\langle \psi, f(A)\phi \rangle = \int_{\sigma(A)} f(\lambda)\, d\langle \psi, \Pi_\lambda \phi \rangle \,. \tag{1.34}$$

The formula (1.33) is the generalization of the formula $A = \sum_{n=0}^{\infty} \lambda_n \Pi_n$ holding for compact operators (see Theorem 1.5). As before, A is composed by orthogonal projections Π_Ω for $\Omega \subset \sigma(A)$, but now in a "continuous" way, i.e. the sum is replaced by an integral.

Proof. Step 1:

Let Π be a bounded p.v.m. with $\operatorname{supp}\Pi = K \subset \mathbb{R}$ compact and define $A := \int \lambda\, d\Pi_\lambda$. We have to show that (1.32) holds.

We extend any $f \in \mathcal{B}(\sigma(A))$ to a function on $\mathbb{R}$ by setting $f = 0$ on $\mathbb{R} \setminus \sigma(A)$. We then define the map

$$\Psi : \mathcal{B}(\mathbb{R}) \to \mathcal{L}(\mathscr{H})\,, \quad \Psi(f) := \int f(\lambda)\, d\Pi_\lambda\,, \tag{1.35}$$

and have to show that

$$\tilde{\Phi}_A(f) = \Psi(f) \quad \text{for all} \quad f \in \mathcal{B}(\sigma(A))\,.$$

This equality follows from the uniqueness statement in Theorem 1.5, if Ψ satisfies i)- iv).

By Theorem 1.17, Ψ is linear and continuous.

To see that Ψ is multiplicative, we first consider simple functions $f = \sum_{k=1}^{n} a_k \chi_{\Omega_k}$ and $g = \sum_{j=1}^{m} b_j \chi_{\Omega_j}$. Then by (1.28) and (1.35) we have

$$\Psi(f) \circ \Psi(g) = \sum_{k=1}^{n} a_k \Pi_{\Omega_k} \Big(\sum_{j=1}^{m} b_j \Pi_{\Omega_j} \Big) = \sum_{k=1}^{n} a_k \sum_{j=1}^{m} b_j \Pi_{\Omega_k} \circ \Pi_{\Omega_j}$$

and since $\chi_A \chi_B = \chi_{A\cap B}$

$$\Psi(fg) = \Psi\Big(\sum_{k=1}^{n} a_k \sum_{j=1}^{m} b_j \chi_{\Omega_k \cap \Omega_j} \Big) = \sum_{k=1}^{n} a_k \sum_{j=1}^{m} b_j \Pi_{\Omega_k \cap \Omega_j} .$$

Thus using $\Pi_A \circ \Pi_B = \Pi_{A\cap B}$ (which holds for any p.v.m. as mentioned below Definition 1.16) we have $\Psi(f) \circ \Psi(g) = \Psi(fg)$. For general $f, g \in \mathcal{B}(\mathbb{R})$ we use approximation with step functions and the continuity of Ψ.

That Ψ is involutive follows in a similar way (approximate f by simple functions and use $(a\Pi_\Omega)^* = \bar{a}\Pi_\Omega$).

For the convergence property in iv) use (1.27) and dominated convergence for the integral.

Thus it remains to prove i). Since we consider 1 as $\chi_{\sigma(A)}$, by definition $\Psi(1) = \mathrm{Id}$ is equivalent to $\Pi_{\sigma(A)} = \mathrm{Id}$ and this immediately implies $\Psi(\mathrm{id}) = A$ by the definition of A above.

To show $\Pi_{\sigma(A)} = \mathrm{Id}$ consider an interval $(a, b]$ in $\mathbb{R}$ such that $\Pi_{(a,b]} = \mathrm{Id}$ (this exists by Proposition 1.15) and let $\mu \in \rho(A)$. The aim is now to show that $\Pi_U = 0$ for some neighbourhood U of μ. For this next step we need the following lemma.

Lemma 1.19

Let X be a Banach space. Then the set $\mathcal{I}(X)$ of invertible bounded linear operators on X is an open subset of $\mathcal{L}(X)$.

Proof. Assume that $T \in \mathcal{I}(X)$. We show that $B \in \mathcal{I}(X)$ whenever $\|A-B\| < \delta$ for some $\delta > 0$ sufficiently small. We write

$$B = T - (T - B) = T\big(\mathrm{Id} - T^{-1}(T - B)\big)\,.$$

Since T is invertible, it suffices to show that $\mathrm{Id} - T^{-1}(T-B)$ is invertible. By Lemma 79 this holds if $\|T^{-1}(T-B)\| < 1$ which is true if $\|T-B\| < \|T^{-1}\|^{-1}$. □

We come back to the proof of Theorem 1.18. Since μ was assumed to be in the resolvent set, $(\mu \,\mathrm{Id} - A)$ is invertible. Thus by Lemma 1.19, there exists some $\delta > 0$ such that for any $S \in \mathcal{L}(\mathscr{H})$

$$\|S - (\mu\,\mathrm{Id} - A)\| \le \delta \implies \quad S \text{ is invertible and } \|S^{-1}\| \le C := \|(\mu\,\mathrm{Id} - A)^{-1}\| + 1. \tag{1.36}$$

(for the last estimate choose $\delta < \|T^{-1}\|^{-1}/2$ in the proof of Lemma 1.19).

We can assume $\delta = \frac{b-a}{N}$ and $\delta < \frac{1}{C}$ and we set $a_k := a + k\delta$ for $k = 0, \ldots, N$ to get a disjoint partition of $(a, b]$. Then the step function $s := \sum_{k=1}^{N} a_k \chi_{(a_{k-1}, a_k]}$ approximates $\mathrm{id}(x) = x$ on $(a, b]$ in the sense that $\|s - \mathrm{id}\|_\infty = \delta$. Then by (1.28) and Theorem 1.17

$$\big\|A - \sum_{k=1}^{N} a_k \Pi_{(a_{k-1},a_k]}\big\| = \big\|A - \int s\, d\Pi\big\| = \big\|\int (\mathrm{id} - s)\, d\Pi\big\| \le \delta\,. \tag{1.37}$$

Since by the assumption on the interval and Proposition 1.15 we have

$$\sum_{k=1}^{N} \Pi_{(a_{k-1},a_k]} = \Pi_{(a.b]} = \mathrm{Id}\ .$$

Thus (1.37) implies (by adding zero)

$$\big\|(\mu\,\mathrm{Id}-A)-\sum_{k=1}^{N}(\mu-a_k)\Pi_{(a_{k-1},a_k]}\big\|\le\delta\,.$$

From (1.36) it follows that $S:=\sum_{k=1}^{N}(\mu-a_k)\Pi_{(a_{k-1},a_k]}$ is invertible and $\|S^{-1}\|\le C$.

But on the other hand $\|S\phi\|\ge\|\phi\|\inf\{|\mu-a_k|\,|\,\Pi_{(a_{k-1},a_k]}\ne 0\}$ and thus[5]

$$\|S^{-1}\|\le\sup\{|\mu-a_k|^{-1}\,|\,\Pi_{(a_{k-1},a_k]}\ne 0\}.$$

This implies $\Pi_{(a_{k-1},a_k]}=0$ for $|\mu-a_k|<\frac{1}{C}$. Therefore we have shown that for any $\mu\in\rho(A)$ there exists a neighbourhood U of μ such that $\Pi_U=0$.

Consider now a compact set $K\subset\rho(A)$. As shown above, for each $\mu\in K$ there exists a neighbourhood U_μ such that $\Pi_{U_\mu}=0$. These neighbourhoods are a covering of K and since K is compact, there exists a finite subcover $K\subset\bigcup_{k=1}^{m}U_{\mu_k}$. Thus $\Pi_K=0$ by Proposition 1.15. Since for each $\psi\in\mathscr{H}$ the map $\ell_{\psi,\psi}$ defined in (1.26) is a positive linear functional on $\mathcal{C}_0(\mathbb{R})$, it follows from Theorem 1.7, i) that the (positive) measure $\Omega\mapsto\langle\psi,\Pi_\Omega\psi\rangle$ is regular. Thus

$$\langle\psi,\Pi_{\rho(A)}\psi\rangle=\sup\{\langle\psi,\Pi_K\psi\rangle\,|\,K\subset\rho(A)\text{ compact }\}=0\,,\quad\psi\in\mathscr{H}.\quad(1.38)$$

[5]Let X,Y be normed spaces, $S:X\to Y$ be a linear map so that there exists a constant $m>0$ with $\|Sx\|\ge m\|x\|$ for all $x\in X$. Then $S:X\to\operatorname{Ran}S$ is bijective and the inverse $S^{-1}:\operatorname{Ran}S\to X$ is bounded with $\|S^{-1}\|\le m^{-1}$, since for all $y=Sx\in\operatorname{Ran}S$

$$\frac{\|S^{-1}y\|}{\|y\|}=\frac{\|S^{-1}Sx\|}{\|Sx\|}\le\frac{\|S^{-1}Sx\|}{m\|x\|}=\frac{1}{m}.$$

Since by assumption Π_Ω is self-adjoint for each $\Omega \in \mathscr{B}$, it follows from Lemma 69[6] that $\Pi_{\rho(A)} = 0$ and therefore $\Pi_{\sigma(A)} = \mathrm{Id}$.

Step 2:

Let $A \in \mathcal{L}(\mathscr{H})$ be self-adjoint and let Π be the family of spectral projections of A given by $\Pi_\Omega = \tilde{\Phi}_A(\chi_\Omega)$ as in Definition 1.14. We then have to show that the operator $B := \int \lambda d\Pi_\lambda$ is equal to A.

By Proposition 1.15 the projection valued measure Π given by the family of spectral projections is compactly supported (with $\operatorname{supp}\Pi \subset \sigma(A)$). Fix $\varepsilon > 0$ and choose a simple function $s = \sum_{k=1}^m c_k \chi_{\Omega_k}$ on $\sigma(A)$ such that $\|\operatorname{id} - s\|_\infty \leq \varepsilon$. Then for Ψ given in (1.35) and $\tilde{\Phi}_A$ given by Theorem 1.11 it follows from Theorem 1.11, Theorem 1.17 and the definition of Π that

$$\begin{aligned} \|A - B\| &\leq \|A - \tilde{\Phi}_A(s)\| + \|\tilde{\Phi}_A(s) - \Psi(s)\| + \|\Psi(s) - B\| \\ &\leq \|\operatorname{id} - s\|_\infty + \left\| \sum_{k=1}^m c_k \big(\tilde{\Phi}_A(\chi_{\Omega_k}) - \Pi_{\Omega_k}\big) \right\| + \|\operatorname{id} - s\|_\infty \\ &\leq \varepsilon + 0 + \varepsilon. \end{aligned}$$

Since ε is arbitrarily small, this proves the assertion.

□

As we have seen in the above proof, the spectral projection of the resolvent set is zero, which implies that the support is a subset of the spectrum. We even have the following:

[6] Lemma 69: Let $\mathscr{H}$ be a Hilbert space and $T \in \mathcal{L}(\mathscr{H})$ self-adjoint, then $\|T\| = \sup_{\|x\|=1} \langle x, Tx \rangle$.

Corollary 1.20

Let $A \in \mathcal{L}(\mathscr{H})$ *be self-adjoint and let* Π *be the family of spectral projections of* A*, then* $\operatorname{supp}\Pi = \sigma(A)$.

Thus $\lambda \in \rho(A)$ if and only if there exists some neighbourhood U of λ such that $\Pi_U = 0$.

Proof. "$\subset$": This follows at once from (1.38).
"$\supset$": Let U be a neighbourhood of λ such that $\Pi_U = 0$. Set

$$f(t) := \begin{cases} (\lambda - t)^{-1}\,, & \text{if} \quad t \notin U \\ 0\,, & \text{otherwise.} \end{cases}$$

Then f and $g(t) := (\lambda - t)$ are measurable and bounded on $\sigma(A)$ and $fg = \chi_{U^c}$. Thus

$$f(A)(\lambda \operatorname{Id} - A) = f(A)g(A) = (fg)(A) = \chi_{U^c}(A) = \Pi_{U^c} = \operatorname{Id}$$

where the last equality holds because $\Pi_U = 0$. Similar it can be shown that $(\lambda \operatorname{Id} - A)f(A) = \operatorname{Id}$, thus $(\lambda \operatorname{Id} - A)$ is invertible, proving $\lambda \in \rho(A)$. ☐

Example 1.21

On $\mathscr{H} = L^2([0,1])$ *consider for* $f \in \mathcal{C}([0,1],\mathbb{R})$ *the self-adjoint multiplication operator* $\mathcal{M}_f x(t) = f(t)\cdot x(t)$*. Then the spectrum* $\sigma(\mathcal{M}_f)$ *is given by* $f([0,1])$*, the range of* f*.*

In fact, if $\mu \notin f([0,1])$*, then* $\mu - f(t) \neq 0$ *for all* $t \in [0,1]$ *and the inverse operator* $\mathcal{M}_f^{-1}$ *is given by* $\mathcal{M}_g$ *with* $g(t) = (\mu - f(t))^{-1}$*.*

On the other hand, if $\mu \in f([0,1])$ *then the set* $\Omega_\mu := \{t \in [0,1] \mid f(t) = \mu\}$ *is not empty. Moreover,* χ_{Ω_μ}*, the characteristic set of* Ω_μ*, satisfies the* *equation*

$$\mathcal{M}_f \chi_{\Omega_\mu}(t) = f(t)\chi_{\Omega_\mu}(t) = \mu\chi_{\Omega_\mu}.$$

Thus μ is an eigenvalue of $\mathcal{M}_f$ with eigenfunction χ_{Ω_μ}, if χ_{Ω_μ} is not the zero function in L^2, i.e. if the Lebesgue-measure of Ω_μ is not zero.

Now assume that $\mu \in f([0,1])$ is not an eigenvalue of $\mathcal{M}_f$ and thus $(\mu \operatorname{Id} - \mathcal{M}_f)$ is injective. If $\mu \notin \sigma(\mathcal{M}_f)$, then $(\mu \operatorname{Id} - \mathcal{M}_f)$ is bijective, i.e. for any $y \in L^2([0,1])$ there exists $x \in L^2([0,1])$ such that

$$(\mu \operatorname{Id} - \mathcal{M}_f)x = y \quad \textit{and thus} \quad x(t) = \frac{1}{\mu - f(t)} y(t).$$

Since f is continuous we have $|f(t) - f(t_0)| < \varepsilon$ for $|t - t_0| < \delta$. Since the term $(\mu - f(t))^{-1}$ converges to ∞ as $t \to t_0$ it follows that x in not in L^2 for any $y \in L^2$.

Therefore we have shown that $\sigma(\mathcal{M}_f) = f([0,1])$ and

$$\sigma_p(\mathcal{M}_f) = \{\mu \in f([0,1]) \,|\, \Omega_\mu \textit{ has positive Lebesgue-measure}\}.$$

Since $\mathcal{M}_f$ is self-adjoint, it has no residual spectrum.

Now consider the operator $\mathcal{M}_{\mathrm{id}}$ (with $\mathrm{id}(t) = t$), then

$$\sigma(\mathcal{M}_{\mathrm{id}}) = \sigma_c(\mathcal{M}_{\mathrm{id}}) = [0,1]$$

and

$$\langle x, \mathcal{M}_{\mathrm{id}} y\rangle = \int_0^1 \bar{x}(t) t y(t)\, dt = \int_0^1 \lambda \, d\langle x, \Pi_\lambda y\rangle$$

where Π denotes the spectral projection of $\mathcal{M}_{\mathrm{id}}$. Thus $\Pi_\Omega x = \chi_{\Omega \cap [0,1]} x$ and the measure $d\langle x, \Pi_\lambda y\rangle$ is absolutely continuous with respect to Lebesgue-measure with Radon density $\bar{x}y$ (see (1.13)).

If, more general, f is differentiable and $f'(t) > 0$, then f is invertible, $f([0,1])$ is an interval and $\mathcal{M}_f$ has no eigenvalues. We then have (using the substitution $\lambda = f(t)$)

$$\begin{aligned}\langle x, \mathcal{M}_f y\rangle &= \int_0^1 \bar{x}(t) f(t) y(t)\, dt = \int_{f(0)}^{f(1)} \lambda \cdot \big(\bar{x}y\big) \circ f^{-1}(\lambda) (f^{-1})'(\lambda)\, d\lambda \\ &= \int_{f(0)}^{f(1)} \lambda\, d\langle x, \Pi_\lambda y\rangle\,.\end{aligned}$$

1.1.5 Multiplication Operator Form

In this section, we will show that each bounded self-adjoint operator on a Hilbert space is unitary equivalent to a multiplication operator.

We start with the notion of cyclic vectors.

Definition 1.22

Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. A vector $\psi \in \mathscr{H}$ is called a **cyclic vector** *for $A :\Longleftrightarrow$ Finite linear combinations of the elements $\{A^n \psi\}_{n=0}^\infty$ are dense in $\mathscr{H}$.*

If an operator A has a cyclic vector, then it is unitary equivalent to $\mathcal{M}_{\text{id}}$ introduced in Example 1.21 on some appropriate Hilbert space, as is shown in the following lemma.

Lemma 1.23

Let $A \in \mathcal{L}(\mathscr{H})$ be a self-adjoint operator with cyclic vector $\psi \in \mathscr{H}$. We denote by μ_ψ the spectral measure of A associated to ψ, i.e. such that

$$\langle \psi, \tilde{\Phi}_A(f)\psi\rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_\psi(\lambda)\,, \qquad f \in \mathcal{B}(\sigma(A)), \tag{1.39}$$

(in the notation introduced in (1.16) *we thus set* $\mu_\psi := \mu_{\psi,\psi}$*). Then there exists a unitary operator*

$$U : \mathscr{H} \longrightarrow L^2(\sigma(A), d\mu_\psi) \quad \text{such that} \quad \big(UAU^{-1}f\big)(\lambda) = \lambda f(\lambda) \quad \mu_\psi - a.e. \tag{1.40}$$

Proof. For $f \in \mathcal{C}(\sigma(A))$ and Φ_A given in Theorem 1.2, we define U on the vectors $\Phi_A(f)\psi \in \mathscr{H}$ by

$$U\big(\Phi_A(f)\psi\big) = f.$$

To see that U is well-defined we remark that

$$\begin{aligned}\|\Phi_A(f)\psi\|^2 &= \langle\psi, (\Phi_A(f))^*\Phi_A(f)\psi\rangle = \langle\psi, \Phi_A(\bar{f}f)\psi\rangle \\ &= \int_{\sigma(A)} |f|^2 \, d\mu_\psi \,.\end{aligned}$$

Thus $\Phi_A(f)\psi = \Phi_A(g)\psi$ if and only if $f = g$ μ_ψ-almost everywhere. This shows that

$$U: \; X := \{\Phi_A(f)\psi \,|\, f \in \mathcal{C}(\sigma(A))\} \longrightarrow L^2(\sigma(A), d\mu_\psi)$$

is well-defined and norm-preserving. Moreover, U is linear since Φ_A is linear. Since ψ was assumed to be cyclic, $\overline{X} = \mathscr{H}$. Therefore using the BLT-Theorem, we can extend U to an isometric linear map of $\mathscr{H}$ into $L^2(\sigma(A), d\mu_\psi)$.

Since $\mathcal{C}(\sigma(A))$ is dense in $L^2(\sigma(A), d\mu_\psi)$ and the range of an isometry is closed, it follows that U is surjective and thus unitary.

Finally, if $f \in \mathcal{C}(\sigma(A))$, we get using $\Phi_A(\mathrm{id}) = A$, the multiplicativity of Φ_A and the definition of U

$$\big(U \circ A \circ U^{-1} f\big)(\lambda) = \big(U \circ A \circ \Phi_A(f)\psi\big)(\lambda) = \big(U \circ \Phi_A(\mathrm{id} \cdot f)\psi\big)(\lambda)$$

$$= \big(\mathrm{id} \cdot f\big)(\lambda) = \lambda \cdot f(\lambda).$$

By continuity this equality extends to $L^2(\sigma(A), d\mu_\psi)$ and thus proves (1.40).

□

Unfortunately, not all self-adjoint operators have cyclic vectors (Exercise 1.38). But the following lemma shows that there exist cyclic vectors for subspaces and $\mathscr{H}$ is a direct sum[7] of these.

Lemma 1.24

Let $\mathscr{H}$ be a separable Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. Then there exists a direct sum decomposition $\mathscr{H} = \bigoplus_{n=1}^{N} \mathscr{H}_n$[8] with $N \in \mathbb{N}$ or $N = \infty$ so that

i) A leaves each $\mathscr{H}_n$ invariant, that is, $\psi \in \mathscr{H}_n$ implies $A\psi \in \mathscr{H}_n$.

---

[7]Suppose $\mathscr{H}_1$ and $\mathscr{H}_2$ are Hilbert spaces. Then the set of pairs (x, y) with $x \in \mathscr{H}_1$ and $y \in \mathscr{H}_2$ is a Hilbert space with the inner product

$$\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, y_1 \rangle_{\mathscr{H}_1} + \langle x_2, y_2 \rangle_{\mathscr{H}_2}.$$

This space is called **direct sum** of the Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ and is denoted by $\mathscr{H}_1 \oplus \mathscr{H}_2$. For a given Hilbert space $\mathscr{H}$ with closed subspace M, we have the decomposition $\mathscr{H} = M \oplus M^\perp$.

Analogously to the direct sum of two spaces, we can construct countable direct sums of Hilbert spaces, starting with a sequence $(\mathscr{H}_n)_{n=1}^\infty$. Then the direct sum $\bigoplus_{n=1}^\infty \mathscr{H}_n$ is given by the set of sequences $(x_n)_{n=1}^\infty$ such that $x_n \in \mathscr{H}_n$ and $\sum_{n=1}^\infty \|x_n\|^2_{\mathscr{H}_n} < \infty$.

[8]This means that each $x \in \mathscr{H}$ can be written uniquely as $x = \sum_{n=1}^N x_n$ where $x_n \in \mathscr{H}_n$ and $\|x\|^2 = \sum \|x_n\|^2$. In other words, $\mathscr{H}_n$ are pairwise orthogonal closed subspaces and the span of these is dense in $\mathscr{H}$. We identify the elements $x \in \mathscr{H}$ with the sequences $(x_n)_{n=1}^N$.

ii) For each n, there is a $\varphi_n \in \mathscr{H}_n$, which is cyclic for $A|_{\mathscr{H}_n}$, i.e.

$$\mathscr{H}_n = \overline{\{\Phi_A(f)\varphi_n \mid f \in \mathcal{C}(\sigma(A))\}}.$$

Proof. We use the Lemma of Zorn[9]. Let H denote the set of at most countable families $(\mathscr{H}_i)$ of pairwise orthogonal closed subspaces $\mathscr{H}_i$ of $\mathscr{H}$ satisfying i) and ii). Then H is not empty since $(\{0\}) \in H$. Moreover, H is partially ordered by inclusion. Let C be a chain in H and set $h_0 := \bigcup_{k\in C} k$. Then h_0 is an upper bound of C. To use Zorn's Lemma, we have to show that $h_0 \in H$.

If h_0 would contain a non-countable number of different $\mathscr{H}_i$, then (since the $\mathscr{H}_i$ where assumed to be pairwise orthogonal) $\mathscr{H}$ would have a non-countable orthonormal system, which would contradict the separability of $\mathscr{H}$. Therefore h_0 contains only countable many elements and is thus an element of H. This shows that C has an upper bound in H.

By Zorn's Lemma, this implies that H has a maximal element $h = (\mathscr{H}_i) \in H$, which includes $\{0\}$. Now let $\tilde{\mathscr{H}} := \overline{\operatorname{span}} \bigcup_{\mathscr{H}_i \in h} \mathscr{H}_i$ and assume that $\tilde{\mathscr{H}} \neq \mathscr{H}$. Then there exists some $x \in \tilde{\mathscr{H}}^\perp \setminus \{0\} \subset \mathscr{H}$. Set $V = \overline{\operatorname{span}}\{A^n x \mid n \geq 0\}$, then V is invariant under A and x is cyclic for $A|_V$. But this implies $h \subset h \cup \{V\}$ and by the maximality of h we get $V = \{0\}$. Then $x = 0$ in

[9]A set P together with a relation $\leq$ is called partially ordered set, if for all $x, y, z \in P$ we have a) $x \leq x$ (reflexive), b) $x \leq y$ and $y \leq z$ imply $x \leq z$ (transitive), c) $x \leq y$ and $y \leq x$ imply $x = y$ (antisymmetric). A subset $C \subset P$ is called chain, if C is linearly ordered, i.e. $x \leq y$ or $y \leq x$ for all $x, y \in C$.

Zorn's Lemma (1935): If every non-empty chain in a non-empty partially ordered set $(P, \leq)$ has an upper bound in P, then P has a maximal element a, i.e. for all $x \in P$ we have $a \leq x$ implies $a = x$.

contradiction to the assumption. Thus $\tilde{\mathscr{H}} = \mathscr{H}$, which proves the existence of the decomposition. □

Lemma 1.23 and Lemma 1.24 can be combined to get the following multiplication operator form of the spectral theorem.

Theorem 1.25

Let $\mathscr{H}$ be a separable Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. Then there exist measures $\{\mu_n\}_{n=1}^N$ with $N \in \mathbb{N}$ or $N = \infty$ on $\sigma(A)$ and a unitary operator

$$U : \mathscr{H} \longrightarrow \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n)$$

such that for any $g = (g_1, \dots g_N) \in \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n)$

$$\big(UAU^{-1}g\big)_n(\lambda) = \lambda g_n(\lambda) \quad \mu_n - a.e. \quad \textit{for all} \quad n = 1, \dots, N\,. \tag{1.41}$$

This realization of A is called **spectral representation** *of A.*

Proof. By Lemma 1.24 we get a direct sum decomposition $\mathscr{H} = \bigoplus_{n=1}^{N} \mathscr{H}_n$ into A-invariant subspaces $\mathscr{H}_n$ on which A has a cyclic vector φ_n. Thus by Lemma 1.23 there exist for each $n = 1, \dots, N$ unitary maps

$$U_n : \ \mathscr{H}_n \longrightarrow L^2(\mathbb{R}, d\mu_{\varphi_n}) \quad \text{such that} \quad U_n A U_n^{-1} g_n(\lambda) = \lambda g_n(\lambda) \quad \mu_{\varphi_n} - \text{a.e.}$$

(here we use that $U_n^{-1} g_n \in \mathscr{H}_n$ by definition and $AU_n^{-1} g_n \in \mathscr{H}_n$ since $\mathscr{H}_n$ is A-invariant).

We identify $\phi \in \mathscr{H}$ with the tuple $(\phi_1, \dots, \phi_N) \in \bigoplus_{n=1}^{N} \mathscr{H}_n$ and set $\mu_n := \mu_{\varphi_n}$ and

$$U : \mathscr{H} \longrightarrow \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n)\,, \quad U\phi = U(\phi_1, \dots, \phi_N) := (U_1\phi_1, \dots, U_N\phi_N).$$

Then

$$U^{-1}g = \big(U_1^{-1}g_1, \ldots, U_N^{-1}g_N\big)$$

for any $g = (g_1, \ldots g_N) \in \bigoplus_{n=1}^N L^2(\mathbb{R}, d\mu_n)$ and

$$\begin{aligned}\big(UAU^{-1}g\big)_n(\lambda) &= \big(UA(U_1^{-1}g_1, \ldots, U_N^{-1}g_N)\big)_n(\lambda)\\ &= \Big(\big(U_1AU_1^{-1}g_1, \ldots, U_NAU_N^{-1}g_N\big)\Big)_n(\lambda)\\ &= U_nAU_n^{-1}g_n\big(\lambda\big) = \lambda g_n(\lambda) \qquad \mu_n - \text{a.e.}\end{aligned}$$

□

Remark that each measure μ_n has support on the spectrum of $A_n := A|_{\mathscr{H}_n}$, thus $L^2(R, d\mu_n) = L^2(\sigma(A_n), d\mu_n)$ (two functions are equal in $L^2(\mu)$ if they only differ on a set of measure zero, i.e. in our case only outside of $\sigma(A_n)$).

This theorem shows that every bounded self-adjoint operator is in fact a multiplication operator on a suitable measure space.

Corollary 1.26

Let $\mathscr{H}$ be a separable Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. Then there exists a finite measure space (M, μ), a bounded function F on M and a unitary map

$$U : \mathscr{H} \longrightarrow L^2(M, d\mu) \quad \textit{so that} \quad \big(UAU^{-1}g\big)(m) = F(m)g(m)\,. \tag{1.42}$$

Proof. Choose a decomposition $\mathscr{H} = \bigoplus_{n=1}^N \mathscr{H}_n$ in invariant subspaces as given in Lemma 1.24 and cyclic vectors $\varphi_n \in \mathscr{H}_n$ such that $\|\varphi_n\| = 2^{-n}$. Let $M = \dot{\bigcup}_{n=1}^N \mathbb{R}$ be the disjoint union of N copies of $\mathbb{R}$ (this can be realized as a subspace of $\mathbb{R}^2$ given by $M = \bigcup_n (\mathbb{R} \times \{n\})$, we denote the elements of M by

$m = (\lambda, n)$ for $\lambda \in \mathbb{R}$ and $n \in \{1, \dots, N\}$). Let $\Sigma := \{\Omega \subset M \,|\, \Omega \cap \mathbb{R} \text{ is Borel}\}$ and define the measure $\mu : \Sigma \to [0, \infty]$ by

$$\mu(A) := \sum_{n=1}^{N} \mu_n(A \cap \mathbb{R})$$

where $\mu_n, n = 1, \dots, N$ are the measures given in Theorem 1.25, i.e. the restriction of μ to the n-th copy of $\mathbb{R}$ is just μ_n. Then μ is finite, because $\mu_n = \mu_{\varphi_n}$ was determined by (1.39) and thus (for $f(t) = 1$, using $\Phi_A(1) = \mathrm{Id}$)

$$\begin{aligned}\mu(M) &= \sum_{n=1}^{N} \mu_n(\mathbb{R}) = \sum_{n=1}^{N} \int_{\mathbb{R}} d\mu_{\varphi_n} = \sum_{n=1}^{N} \langle \varphi_n, \varphi_n \rangle \\ &= \sum_{n=1}^{N} \|\varphi_n\|^2 = \sum_{n=1}^{N} 2^{-2n} < \infty \,.\end{aligned}$$

Then the map $V : \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n) \longrightarrow L^2(M, d\mu)$ given by

$$\big(Vg\big)(\lambda, n) = \big(V(g_1, \dots, g_N)\big)(\lambda, n) := g_n(\lambda)$$

is unitary (Exercise 1.39). Thus if we denote the unitary map given in Theorem 1.25 by $\tilde{U}$, the map $U := V \circ \tilde{U}$ satisfies (1.42) for $F(\lambda, n) = \lambda$ for all $n = 1, \dots, N$ (Exercise 1.40). □

Remark 1.27

Corollary 1.26 essentially is a rigorous form of the Dirac notation used in physics. Using the unitary transform U, we get for $\psi, \varphi \in L^2(M, d\mu)$

$$\begin{aligned}\langle \psi, \varphi \rangle &= \sum_n \int \overline{\psi(\lambda, n)} \varphi(\lambda, n) \, d\mu_n \\ \langle \psi, A\varphi \rangle &= \sum_n \int \lambda \overline{\psi(\lambda, n)} \varphi(\lambda, n) \, d\mu_n \,.\end{aligned}$$

Example 1.28 *i) Let A be compact and self-adjoint, then by the Hilbert-Schmidt-Theorem 1.4 there is a complete set of eigenvectors $\{\phi_j\}_{j=1}^{\infty}$ with $A\phi_j = \lambda_j \phi_j$. In this case $\mu = \sum_{j=1}^{\infty} 2^{-j}\delta(t - \lambda_j)$ works as a spectral measure if there is no repeated eigenvalue.*

ii) Let $\mathscr{H} = \ell^2(\mathbb{Z}, \mathbb{C})$ and $A = R+L$, where L is the left-shift operator (i.e. $(La)_n = a_{n+1}$) and $R = L^*$ is the right-shift operator (with $(Ra)_n = a_{n-1}$).

To represent A as multiplication operator, use

$$U : \ell^2(\mathbb{Z}, \mathbb{C}) \to L^2([0,1]) \quad \text{given by } U(a_n) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi int} .$$

Then ULU^{-1} is multiplication by $e^{-2\pi it}$ and URU^{-1} is multiplication by $e^{+2\pi it}$. Therefore UAU^{-1} is multiplication by $2\cos(2\pi t)$. A can be represented as a multiplication operator by t on $L^2(\mathbb{R}, d\mu_1) \oplus L^2(\mathbb{R}, d\mu_2)$, where μ_1, μ_2 have support in $[-2, 2]$ (Exercise 1.41).

We now introduce another decomposition of the spectrum of an operator additional to the one given in Definition 1.1, using the notion of spectral projections.

In Corollary 1.20 we already saw that the spectrum of a self-adjoint operator A is equal to the support of the associated family of spectral projections and that λ is in the resolvent set of A, if and only if $\Pi_U = 0$ for some neighbourhood U of λ. This implies that $\lambda \in \sigma(A)$ if and only if $\Pi_U \neq 0$ for all neighbourhoods of λ.

Definition 1.29

Let $\mathscr{H}$ be a Hilbert space, $A \in \mathcal{L}(\mathscr{H})$ self-adjoint and $\Pi = \{\Pi_\Omega\}$ the associated family of spectral projections.

i) We say $\lambda \in \sigma(A)$ *is in the* **essential spectrum** $\sigma_{\text{ess}}(A)$ *of* A $:\Longleftrightarrow$

$$\dim \operatorname{Ran} \Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)} = \infty \qquad \text{for all} \quad \varepsilon > 0.$$

ii) We say $\lambda \in \sigma(A)$ *is in the* **discrete spectrum** $\sigma_{\text{disc}}(A)$ *of* A $:\Longleftrightarrow$

$$\dim \operatorname{Ran} \Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)} < \infty \qquad \text{for some} \quad \varepsilon > 0.$$

Theorem 1.30

The essential spectrum of a self-adjoint bounded operator is always closed.

Proof. Let $\lambda_n \to \lambda$ with $\lambda_n \in \sigma_{\text{ess}}(A)$. Let $\varepsilon > 0$ given, then $(\lambda_n - \delta, \lambda_n + \delta) \subset (\lambda - \varepsilon, \lambda + \varepsilon)$ for some $n \in \mathbb{N}$ and $\delta > 0$. Thus the range of $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)}$ is infinite dimensional. This shows that λ is an element of the essential spectrum. □

Theorem 1.31

$\lambda \in \sigma_{\text{disc}}(A)$ *if and only if* λ *is an isolated point of* $\sigma(A)$ *and* λ *is an eigenvalue of finite multiplicity.*

Proof. $\Rightarrow$: If $\lambda \in \sigma_{\text{disc}}(A)$, then there exists some ε_0 such that $\Pi_{(\lambda-\varepsilon_0,\lambda+\varepsilon_0)}$ is a projector of finite range independent of ε for all $\varepsilon < \varepsilon_0$. This is actually the projector $\Pi_{\{\lambda\}}$ and we observe that $\Pi_{(\lambda,\lambda+\varepsilon_0)} = 0$ and $\Pi_{(\lambda-\varepsilon_0,\lambda)} = 0$. This shows that λ is an isolated point in $\sigma(A)$.

If $\lambda \in \sigma(A)$ is an isolated point, the spectral representation of A shows that if $x = \Pi_{\{\lambda\}} x$ for some $x \neq 0$, then $Ax = \lambda x$, thus λ is an eigenvalue of finite multiplicity.

$\Leftarrow$: If λ is isolated, then $\Pi_{(\lambda-\varepsilon,\lambda+\varepsilon)} = \Pi_{\{\lambda\}}$ for ε small enough. Since λ is of finite multiplicity, the dimension of $\Pi_{\{\lambda\}}$ is finite. □

1.1.6 Exercises

Exercise 1.32 (Continuous Functional Calculus for compact operators)

Let $\mathscr{H}$ be a Hilbert space, $A \in \mathcal{K}(\mathscr{H})$ be a compact self-adjoint operator.

Let $(\lambda_n)_{n\geq 1}$ denote the non-zero eigenvalues of A and Π_n the orthogonal projections on the associated eigenspaces. Denote by Π_0 the orthogonal projection to the kernel of A (i.e. to the eigenspace ot $\lambda_0 = 0$).

Show that the map

$$\mathcal{C}(\sigma(A)) \ni f \mapsto \Psi(f) := \sum_{n=0}^{\infty} f(\lambda_n)\Pi_n$$

has the properties i), ii) and iii) of the map Φ_A given in Theorem 1.2.

Exercise 1.33 (Positive linear functional)

Let X, Y be compact topological spaces and let $Z := L(\mathcal{C}(X), \mathcal{C}(Y))$ be the set of linear (not necessarily bounded) operators from $\mathcal{C}(X)$ to $\mathcal{C}(Y)$. Here $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ are the sets of continuous complex valued functions on X and Y respectively.

We call $T \in Z$ positivity preserving, if T maps positive functions to positive functions, i.e.

$$f(x) \geq 0 \text{ for all } x \in X \qquad \Longrightarrow \qquad Tf(y) \geq 0 \text{ for all } y \in Y.$$

i) Prove that each positivity preserving operator $T \in Z$ is automatically continuous and $\|T\| = \|T1\|_\infty$, where 1 is the constant function with value 1 (i.e. $1(x) = 1$).

ii) Let $(S_n)_{n\in\mathbb{N}}$ be a family of linear operators in Z such that $S_{n+1} - S_n$ is positivity preserving for each $n \in \mathbb{N}$.

Prove that $(S_n)_{n\in\mathbb{N}}$ converges in operator norm if and only if $(S_n 1)_{n\in\mathbb{N}}$ converges in sup-norm (here again 1 is the constant function $1(x) = 1$).

Exercise 1.34 (Proof Lemma 1.12)
Let M be a compact subset of $\mathbb{C}$ and consider the system $\mathcal{S}$ of all sets S such that

i) $\mathcal{C}(M) \subset S \subset \mathcal{B}(M)$

ii) for any sequence (f_n) in S

$$\Big(\sup_n \|f\|_\infty < \infty \quad \text{and} \quad f(t) := \lim_{n\to\infty} f_n(t) \text{ exists for all } t \in M\Big)$$
$$\implies \quad f \in S$$

Set $V := \bigcap_{S\in\mathcal{S}} S$.
Show that $\alpha g \in V$ for all $\alpha \in \mathbb{C}$ and $g \in V$.

Exercise 1.35 (Spectral Theorem - Functional Calculus form)
Let $\mathscr{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint. Let $\tilde{\Phi}_A : \mathcal{B}(\sigma(A)) \to \mathcal{L}(\mathscr{H})$ be determined by

$$\mathcal{B}(\sigma(A)) \ni g \mapsto \langle \psi, \tilde{\Phi}_A(g)\varphi\rangle = \int_{\sigma(A)} g(\lambda)\, d\mu_{\psi,\varphi}(\lambda)$$

using Corollary 1.10 (or the Riesz Lemma 1.9).
Here for any $\psi \in \mathscr{H}$ fixed, the measure $\mu_{\psi,\varphi}$ is determined via the Riesz-Markov-Theorem (Theorem 1.7) by the bounded linear form on $\mathcal{C}(\sigma(A))$ given by

$$\ell_{\psi,\varphi}(f) = \langle \psi, \Phi_A(f)\varphi\rangle = \int_{\sigma(A)} f(\lambda)\, d\mu_{\psi,\varphi}(\lambda)\,.$$

Show the following statements.

i) $\tilde{\Phi}_A$ *is an algebraic *-homomorphism.*

ii) *If* $BA = AB$, *then* $\tilde{\Phi}_A(f)B = B\tilde{\Phi}_A(f)$.

iii) *If* f *is real-valued, then* $\tilde{\Phi}_A(f)$ *is self-adjoint.*

Exercise 1.36 (Invertible Operators)
Let X *be a Banach space. Show the following statements.*

i) *Let* $S \in \mathcal{L}(X)$ *and* $\mu \in \rho(S)$, *then*

$$\|(\mu \operatorname{Id} - S)^{-1}\| \geq \frac{1}{\operatorname{dist}(\mu, \sigma(S))} \tag{1.43}$$

ii) *If* X *is a Hilbert space and* S *in i) is self-adjoint, then equality holds in* (1.43).

iii) *The set* $\mathcal{I}(X)$ *of invertible bounded linear operators on* X *is an open subset of* $\mathcal{L}(X)$.

Exercise 1.37 (Analytical Functional Calculus)
Let X *be a Banach space and* $S \in \mathcal{L}(X)$. *Consider the power series* $f(z) = \sum_{n=0}^{\infty} a_n z^n$ *and assume that the radius of convergence is larger than the spectral radius* $r(S)$ *of* S. *Show the following statements.*

i) *The sequence* $f(S) := \sum_{n=0}^{\infty} a_n S^n$ *converges in* $\mathcal{L}(\mathscr{H})$.

ii) *If* $g(z) = \sum_{n=0}^{\infty} b_n z^n$ *is another power series with radius of convergence* $> r(S)$, *then* $(fg)(S) = f(S)g(S)$.

iii) *The spectral mapping theorem holds, i.e.* $\sigma(f(S)) = f(\sigma(S))$.

iv) If S is self-adjoint and X is a Hilbert space, then operator $f(S)$ defined in i) and $\Phi_S(f)$ coincide.

Exercise 1.38 (Cyclic vectors)

Prove that a self-adjoint operator on a finite dimensional space has a cyclic vector if and only if it has no repeated eigenvalue. Construct in this case the unitary map given in Lemma 1.23.

Exercise 1.39 (Multiplication operator)

Let μ_n, $n = 1, \dots N$ (for $N \in \mathbb{N}$ or $N = \infty$) be finite regular Borel-measures on $\mathbb{R}$. Let M be given by $M = \bigcup_{n=1}^{N} (\mathbb{R} \times \{n\}))$. We denote the elements of M by $m = (\lambda, n)$ for $\lambda \in \mathbb{R}$ and $n \in \{1, \dots, N\}$. Set $\Sigma := \{\Omega \subset M \,|\, \Omega \cap \mathbb{R} \times \{n\}$ is Borel for each $n \in \{1, \dots N\}\}$ and define the measure $\mu : \Sigma \to [0, \infty]$ by $\mu(A) := \sum_{n=1}^{N} \mu_n(A_n)$ where we set $A_n = A \cap (\mathbb{R} \times \{n\})$. Show that the map $V : \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n) \longrightarrow L^2(M, d\mu)$ given by

$$\big(Vg\big)(\lambda, n) = \big(V(g_1, \dots, g_N)\big)(\lambda, n) := g_n(\lambda)$$

is bijective and isometric.

Exercise 1.40 (Multiplication operator)

Let $\mathscr{H}$ be a separable Hilbert space and $A \in \mathcal{L}(\mathscr{H})$ self-adjoint and let $U : \mathscr{H} \to \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n)$ be a unitary operator as described in Theorem 1.22, i.e. such that

$$\big(UAU^{-1}g\big)_n(\lambda) = \lambda g_n(\lambda) \quad \mu_n - a.e. \quad \text{for all} \quad n = 1, \dots, N$$

for any $g = (g_1, \dots g_N) \in \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n)$.
Let $V : \bigoplus_{n=1}^{N} L^2(\mathbb{R}, d\mu_n) \longrightarrow L^2(M, d\mu)$ be the unitary map given in Exercise

1.39. Show that for any $\tilde{g} \in L^2(M, d\mu)$ (using the notation $m = (\lambda, n) \in M$)

$$\big(VUAU^{-1}V^{-1}\tilde{g}\big)(\lambda, n) = \lambda\tilde{g}(\lambda, n) \quad \mu - a.e.\,.$$

Exercise 1.41 (Right and left shift)

Let $\mathscr{H} = \ell^2(\mathbb{Z}, \mathbb{C})$ and $A = R + L$, where L is the left-shift operator (i.e. $(La)_n = a_{n+1}$) and R is the right-shift operator (with $(Ra)_n = a_{n-1}$). Set

$$U : \ell^2(\mathbb{Z}, \mathbb{C}) \to L^2([0,1]) \quad \text{given by } U(a_n) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n t}\,.$$

i) Show that $R = L^*$ and thus A is self-adjoint.

ii) Show that U is a unitary operator with inverse

$$\big(U^{-1}f\big)_n = \int_0^1 e^{-2\pi i n t} f(t)\, dt\,.$$

iii) Show that ULU^{-1} and URU^{-1} are multiplication by $e^{-2\pi i t}$ and $e^{+2\pi i t}$ respectively.

iv) Conclude that UAU^{-1} is multiplication by $2\cos(2\pi t)$.

v) Find measures μ_1, μ_2 on $\mathbb{R}$ supported in $[-2.2]$ and a unitary map

$$V : L^2([0,1]) \to L^2(\mathbb{R}, d\mu_1) \oplus L^2(\mathbb{R}, d\mu_2)$$

such that $\tilde{U} := V \circ U : \mathscr{H} \to L^2(\mathbb{R}, d\mu_1) \oplus L^2(\mathbb{R}, d\mu_2)$ satisfies

$$\Big(\tilde{U}A\tilde{U}^{-1}(g_1, g_2)\Big)(s) = s \cdot (g_1, g_2)(s)\,.$$

Exercise 1.42 (Spectrum of the adjoint operator)

Let X be a Banach space and $T \in \mathcal{L}(X)$ with adjoint T'. Show *i)* $\overline{\operatorname{Ran} T} = (\operatorname{Ker} T')_\perp$, *where we set for any* $Y \subset X^*$

$$Y_\perp := \{x \in X \,|\, \forall \ell \in Y \,:\, \ell(x) = 0\}\,.$$

ii) $\sigma(T) = \sigma(T')$

iii) $R_\lambda(T') = R_\lambda(T)'$ *for all* $\lambda \in \rho(T)$.

Let $\mathscr{H}$ *be a Hilbert space and* $T \in \mathcal{L}(\mathscr{H})$ *with adjoint* T^*. *Show*

i) $\operatorname{Ker} T^* = (\operatorname{Ran} T)^\perp$ *and* $\operatorname{Ker} T = (\operatorname{Ran} T^*)^\perp$

ii) $\sigma(T^*) = \{\lambda \in \mathbb{C} \,|\, \bar{\lambda} \in \sigma(T)\}$

iii) $R_\lambda(T^*) = R_{\bar{\lambda}}(T)^*$.

Chapter 2

Unbounded operators on Hilbert spaces

2.1 Domains, graphs, adjoints and spectrum

Many of the most important operators which occure in applications are not bounded. We already saw (Hellinger-Toeplitz-Theorem) that an operator A on a Hilbert space $\mathscr{H}$, which satisfies the relation $\langle x, Ay\rangle = \langle Ax, y\rangle$ and has domain $\mathscr{H}$ is necessarily bounded.

This already suggests, that unbounded operators are usually not defined everywhere on $\mathscr{H}$, but only on some linear subset of $\mathscr{H}$. To identify an unbounded operator, we have to specify both its domain and how it acts on that subspace.

2.1.1 Symmetric and closed operators, extensions

Definition 2.1

Let $\mathscr{H}$ be a Hilbert space.

i) *A (linear) operator $A : \mathscr{H} \supset \mathscr{D}(A) \longrightarrow \mathscr{H}$ is a linear map defined on $\mathscr{D}(A)$, its domain, which is a (possibly not closed) subspace of $\mathscr{H}$. We say that A is* **densely defined** *in $\mathscr{H}$, if $\mathscr{D}(A)$ is dense in $\mathscr{H}$.*

ii) *An operator $S : \mathscr{D}(S) \longrightarrow \mathscr{H}$ is called* **extension** *of $A : \mathscr{D}(A) \longrightarrow \mathscr{H}$, if $\mathscr{D}(A) \subset \mathscr{D}(S)$ and $Sx = Ax$ for all $x \in \mathscr{D}(A)$. We write $A \subset S$.*

iii) *Two operators A and S are equal, $A = B$, if $A \subset S$ and $S \subset A$.*

iv) *An operator $A : \mathscr{H} \supset \mathscr{D}(A) \longrightarrow \mathscr{H}$ is called* **symmetric** $:\Longleftrightarrow$

$$\langle Ax, y\rangle = \langle x, Ay\rangle \qquad \text{for all} \quad x, y \in \mathscr{D}(A)\,.$$

Example 2.2 *i)* **Position operator***:*

In classical mechanics, a particle is represented by its position *$x(t) \in \mathbb{R}^3$ at time $t \in \mathbb{R}$ and its momentum or impulse $\xi(t) = m\dot{x}(t)$ (where m denotes the mass of the particle and $\dot{x}(t) = \frac{d}{dt}x(t)$). A classical* observable *is then a real smooth function on the phase space $\mathbb{R}^3 \times \mathbb{R}^3$ and its value at the point $(x(t), \xi(t))$ gives information about the particle at time t.*

In quantum mechanics, the state of a particle *in space at time $t \in \mathbb{R}$ is described by its* wave function *$\psi_t \in \mathscr{H} = L^2(\mathbb{R}^3)$, which is normalized (i.e. $\|\psi_t\|_{L^2} = 1$). Here $|\psi_t(x)|^2$ is interpreted as a probability density: the probability of the presence of the particle at the point x at time t.*

The average position of the particle at the time t is then the expectation value of the random variable x with respect to this density:

$$\langle x\rangle_{\psi_t} := \langle \psi_t, x\psi_t\rangle_{L^2(\mathbb{R}^3)} = \big(\langle \psi_t, x_j\psi_t\rangle_{L^2(\mathbb{R}^3)}\big)_{j=1,2,3}.$$

So let $\mathscr{H} = L^2(\mathbb{R}^3)$ and $\mathscr{D}(T) := \{\psi \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} x^2|\psi(x)|^2\,dx < \infty\}$. For $\psi \in \mathscr{D}(T)$ define $(T\psi)(x) = x\psi(x)$.

It is clear that T is unbounded (choose $\psi_n = \chi_{[n,n+1]}$, then $\|\psi_n\| = 1$ and we see that $\|T\psi_n\| \geq n$). Moreover T is symmetric.

If we take $\phi \notin \mathscr{D}(T)$, then $x\phi(x)$ also has sense as a function, but this function is not in $\mathscr{H}$ anymore. The chosen domain is the largest one which is possible to get an operator with values in $\mathscr{H}$.

ii) **Momentum operator***:*

To define the average momentum as described above for the position, we need to use an analogy with plane waves in optics given by functions of the type

$$\varphi(t,x) := Ae^{i(k\cdot x-\omega t)}$$

where $\nu := \frac{\omega}{2\pi} \in \mathbb{R}$ describes the frequency of the wave and $k \in \mathbb{R}^3$ is the so called wave vector*, giving the direction in which the wave propagates (φ is independent of x on any plane, on which $x\cdot k$ is constant). Thus the momentum should have the same direction as k. Using the wave-particle-duality and the De Broglie relation, which relates the momentum of a particle to its wavelength, gives $\xi = \hbar k$, where $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant. Since*

$$\nabla_x\varphi(t,x) = ik\varphi(t,x) \quad \text{and} \quad \overline{\varphi(t,x)} = \overline{A}e^{-i(k\cdot x-\omega t)} = \frac{|A|^2}{\varphi(x,t)}$$

we therefore get

$$\xi = \frac{\hbar}{i}\big(\nabla_x \varphi(t,x)\big)\frac{\overline{\varphi(t,x)}}{|A|^2}\,.$$

This relation provides a way to get by analogy the average impulse of the quantum particle described by the wave function ψ_t*: Viewing* $|A|^2$ *as a normalization factor, we set*

$$\langle \xi \rangle_{\psi_t} := \left\langle \psi_t, \frac{\hbar}{i}\nabla_x \psi_t \right\rangle_{L^2(\mathbb{R}^3)} = \left(\left\langle \psi_t, \frac{\hbar}{i}\frac{\partial}{\partial x_k}\psi_t \right\rangle_{L^2(\mathbb{R}^3)} \right)_{k=1,2,3}.$$

As above we consider the Hilbert space $\mathscr{H} = L^2(\mathbb{R}^3)$*. The operator* M *given by* $M\phi(x) = i\nabla_x \phi(x)$ *on its domain* $\mathscr{D}(M) := \{\phi \in \mathcal{C}^1(\mathbb{R}^3) \,|\, |\nabla \phi| \in L^2(\mathbb{R}^3)\}$ *is then a symmetric unbounded operator.*

iii) *The operator* S *on* $L^2([0,1])$ *with* $\mathscr{D}(S) = \{\phi \in \mathcal{C}^1([0,1]) \,|\, \phi(1) = \phi(0)\}$ *and* $S\phi = i\frac{d}{dt}\phi$ *is an unbounded operator and an extension of* T *with* $\mathscr{D}(T) = \{\phi \in \mathcal{C}^1([0,1]) \,|\, \phi(1) = \phi(0) = 0\}$ *and* $T\phi = i\frac{d}{dt}\phi$*.*

Both operators are symmetric, which follows from integration by parts, since the boundary terms cancel.

We now come to the definition of a closed operator. The notion of a graph of a map $T : X \to Y$ for normed vector spaces X, Y was already used in Theorem 57 (Closed Graph Theorem)[1]

Definition 2.3

Let $\mathscr{H}$ *be a Hilbert space and* A *be an operator in* $\mathscr{H}$ *with domain* $\mathscr{D}(A)$*.*

[1] Let X, Y be normed vector spaces and $T : X \to Y$, then the graph of T is defined as $\Gamma(T) := \{(x,y) \in X \times Y \,|\, y = Tx\}$.
Closed Graph Theorem: Let $T : X \to Y$ be linear and X, Y be Banach spaces, then T is continuous if and only if $\Gamma(T)$ is closed.

i) The **graph** *of A is the set of pairs $\Gamma(T) := \{(x, Ax) \,|\, x \in \mathscr{D}(A)\}$. It is a subset of $\mathscr{H} \times \mathscr{H}$, which is a Hilbert space with inner product*

$$\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle. \tag{2.1}$$

ii) A is called a **closed operator** *if $\Gamma(A)$ is a closed subset of $\mathscr{H} \times \mathscr{H}$.*

iii) A is called **closable** *if A has a closed extension. In this case, A has a smallest closed extension, called its* **closure***, which is denoted by $\overline{A}$.*

Using the graph, we have for two operators A, B on $\mathscr{H}$

$$A \subset B \quad \Longleftrightarrow \quad \Gamma(A) \subset \Gamma(B)\,.$$

Remark 2.4

The definition of a closed operator can be written using sequences: A is closed if and only if for any sequence $(x_n)_{n=1}^{\infty}$ in $\mathscr{D}(A)$

$$x_n \longrightarrow x \quad \text{and} \quad Tx_n \longrightarrow y \quad \Longrightarrow \quad x \in \mathscr{D}(A) \quad \text{and} \quad Tx = y\,.$$

A natural way to find a closed extension of a given operator A seems to take the closure of its graph in $\mathscr{H} \times \mathscr{H}$. The problem is, that $\overline{\Gamma(A)}$ may not be the graph of an operator. But if A is closable, this procedure gives the closure of A.

Proposition 2.5

Let T be a closable operator on a Hilbert space $\mathscr{H}$, then $\Gamma(\overline{T}) = \overline{\Gamma(T)}$.

Proof. Suppose S is a closed extension of T, then $\mathscr{D}(T) \subset \mathscr{D}(S)$ and $\Gamma(S)$ is closed. Therefore $\overline{\Gamma(T)} \subset \Gamma(S)$, so if $(0, \phi) \in \overline{\Gamma(T)}$, then $\phi = S(0) = 0$. Define the operator R on $\mathscr{H}$ by

$$\mathscr{D}(R) = \{\psi \in \mathscr{H} \,|\, (\psi, \phi) \in \overline{\Gamma(T)} \text{ for some } \phi \in \mathscr{H}\}$$

$$R\psi = \phi \text{ where } \phi \text{ is the unique vector so that } (\psi, \phi) \in \overline{\Gamma(T)}.$$

Then $\Gamma(R) = \overline{\Gamma(T)}$ and thus R is a closed extension of T. But $R \subset S$, which is any closed extension, thus $R = \overline{T}$. □

Example 2.6

Let T_0, T_1 be operators on $\mathscr{H} = L^2(\mathbb{R})$ given by $T_k\phi(x) = i\frac{d}{dx}\phi(x)$ for $\phi \in \mathscr{D}(T_k)$ where $\mathscr{D}(T_0) = \mathcal{C}_0^\infty(\mathbb{R})$ and $\mathscr{D}(T_1) = \mathcal{C}_0^1(\mathbb{R})$. Then $T_0 \subset T_1$.

We will show that $\Gamma(T_1) \subset \overline{\Gamma(T_0)}$. To do this, we need the following

Definition 2.7 (Approximate Identity)

Let $j \in \mathcal{C}_0^\infty(\mathbb{R})$ be non-negative with $\operatorname{supp} j \subset (-1, 1)$ and $\int_{\mathbb{R}} j(x)\,dx = 1$. Then the family $\{j_\epsilon\}_{\epsilon\in(0,1)}$ where $j_\epsilon(x) := \epsilon^{-1} j(\frac{x}{\epsilon})$ is called **approximate identity** *or* **mollifier**.

We remark that if $\{j_\epsilon\}$ is an approximate identity, then $\operatorname{supp} j_\epsilon \subset (-\epsilon, \epsilon)$ (since $-1 < \frac{x}{\epsilon} < 1$ implies $-\epsilon < x < \epsilon$) and

$$\int_{\mathbb{R}} j_\epsilon(x)\,dx = \int_{\mathbb{R}} \epsilon^{-1} j\left(\frac{x}{\epsilon}\right)\,dx = \int_{\mathbb{R}} j(z)\,dz = 1.$$

Let $\{j_\epsilon\}$ be an approximate identity, $\phi \in \mathscr{D}(T_1)$ and set

$$\phi_\epsilon(x) := \big(\phi * j_\epsilon\big)(x) := \int_R f(y) j_\epsilon(x-y)\,dy.$$

Then

$$\begin{aligned}
|\phi_\epsilon(x) - \phi(x)| &\le \int_{\mathbb{R}} j_\epsilon(x-t)|\phi(t) - \phi(x)|\,dt \\
&\le \Big(\sup_{\{t \,|\, |x-t|\le\epsilon\}} |\phi(t)-\phi(x)|\Big)\int_{\mathbb{R}} j_\epsilon(x-t)\,dt \\
&= \sup_{\{t \,|\, |x-t|\le\epsilon\}} |\phi(t)-\phi(x)|\,.
\end{aligned}$$

Since ϕ has compact support, it is uniformly continuous, thus

$$\|\phi_\epsilon - \phi\|_\infty \longrightarrow 0 \qquad \text{as} \quad \epsilon \to 0\,. \tag{2.2}$$

Moreover, since the whole family $\{\phi_\epsilon\}$ is supported in a fixed compact set, this implies that $\phi_\epsilon \to \phi$ in $L^2(\mathbb{R})$.

Similarly, we get

$$\begin{aligned} i\frac{d}{dx}\phi_\epsilon(x) &= \int_{\mathbb{R}} i\frac{d}{dx} j_\epsilon(x-y)\phi(y)\,dy \\ &= \int_{\mathbb{R}} (-i)\big(\frac{d}{dy} j_\epsilon(x-y)\big)\phi(y)\,dy \\ &= \int_{\mathbb{R}} j_\epsilon(x-y) i\big(\frac{d}{dy}\phi(y)\big)\,dy \ \xrightarrow{L^2(\mathbb{R})}\ i\frac{d}{dx}\phi(x)\,. \end{aligned}$$

Since j_ϵ has compact support and is infinitely differentiable, $\phi_\epsilon \in \mathcal{C}_0^\infty(\mathbb{R})$.

Thus $\phi_\epsilon \in \mathscr{D}(T_0)$ for each $\epsilon > 0$. Since we have proven that

$$\phi_\epsilon \xrightarrow{L^2(\mathbb{R})} \phi \quad \text{and} \quad T_0\phi_\epsilon \xrightarrow{L^2(\mathbb{R})} T_1\phi$$

for any $\phi \in \mathscr{D}(T_1)$, the closure of $\Gamma(T_0)$ contains $\Gamma(T_1)$.

We remark that if $\{j_\epsilon\}$ is an approximate identity, for any $f \in L^p(\mathbb{R})$, $1 \le p < \infty$, the convolution $f_\epsilon := f * j_\epsilon$ is smooth and f_ϵ converges in L^p to f (see e.g. [R2]), which explains the name.

2.1.2 Definition and Properties of the adjoint operator

We come to the definition of the Hilbert space adjoint.

Definition 2.8 (Adjoint operator and self-adjointness)
Let T be a densely defined linear operator on a Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(T)$. Let $\mathscr{D}(T^*)$ be the set of $\phi \in \mathscr{H}$ for which there is an $\eta \in \mathscr{H}$ with

$$\langle \phi, T\psi \rangle = \langle \eta, \psi \rangle \quad \text{for all} \quad \psi \in \mathscr{D}(T). \tag{2.3}$$

For $\phi \in \mathscr{D}(T^*)$ we define $T^*\phi = \eta$. The operator T^* is called the **adjoint operator** *of T. If $T = T^*$ then T is called* **self-adjoint**.

Remark that the self-adjointness implies $\mathscr{D}(T) = \mathscr{D}(T^*)$.

The Riesz Lemma gives another possible characterisation of the domain of T^*:

$$\phi \in \mathscr{D}(T^*) \quad \Longleftrightarrow \quad \text{the map } \mathscr{D}(T) \ni \psi \mapsto \langle \phi, T\psi \rangle \text{ is bounded.} \tag{2.4}$$

In fact if, for $\phi \in \mathscr{H}$ fixed, the linear map

$$\ell_\phi^T \;:\; \mathscr{D}(T) \to \mathbb{C} \quad \text{given by} \quad \ell_\phi^T(\psi) := \langle \phi, T\psi \rangle \tag{2.5}$$

is bounded, then (since $\mathscr{D}(T)$ is dense in $\mathscr{H}$) it has a unique extension to a bounded linear functional $\tilde{\ell}_\phi^T$ on $\mathscr{H}$. The Riesz Lemma then shows that there exists a unique $\eta \in \mathscr{H}$ such that $\tilde{\ell}_\phi^T(\psi) = \langle \eta, \psi \rangle$. Thus if ℓ_ϕ^T is bounded and $\psi \in \mathscr{D}(T)$, then there exists $\eta \in \mathscr{H}$ such that equation (2.3) holds.

If T is not densely defined, the vector η (and thus the operator T^*) is not uniquely determined by (2.3).

Even if we assume that T is densely defined, it can occur that the domain of T^* is not dense and even that $\mathscr{D}(T^*) = \{0\}$.

Example 2.9
Take $\mathscr{H} = L^2(\mathbb{R})$ and suppose that $f \in \mathcal{B}(\mathbb{R})$ but $f \notin L^2(\mathbb{R})$. Let $\phi_0 \in \mathscr{H}$

be some fixed vector and define

$$\mathscr{D}(T) = \{\psi \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} |f(x)\psi(x)|\,dx < \infty\} \quad \textit{and}$$
$$T\psi = \langle f, \psi\rangle \phi_0 \textit{ for } \psi \in \mathscr{D}(T).$$

Since $\mathscr{D}(T)$ *contains all* L^2*-functions with compact support, it is dense in* $\mathscr{H}$.

Now suppose that $\phi \in \mathscr{D}(T^*)$, *then for all* $\psi \in \mathscr{D}(T)$

$$\begin{aligned}\langle \psi, T^*\phi\rangle &= \langle T\psi, \phi\rangle = \langle \langle f, \psi\rangle \phi_0, \phi\rangle \\ &= \overline{\langle f, \psi\rangle}\langle \phi_0, \psi\rangle = \langle \psi, \langle \phi_0, \phi\rangle f\rangle\,.\end{aligned}$$

Thus $T^*\phi = \langle \phi_0, \phi\rangle f$. *Since* f *is not in* $L^2(\mathbb{R})$ *it follows that* $\langle \phi_0, \phi\rangle = 0$. *Therefore* $\mathscr{D}(T^*) = \mathrm{span}\{\phi_0\}^\perp$, *which is not dense in* $\mathscr{H}$, *and* $T^*\phi = 0$ *for any* $\phi \in \mathscr{D}(T^*)$.

Example 2.10

Let $\{e_n\}_{n=1}^\infty$ *be an orthonormal basis of* $\mathscr{H} = L^2(\mathbb{R})$ *and set*

$$\mathscr{D}(A) = \mathcal{C}_0^\infty(\mathbb{R}) \quad \textit{and} \quad A\psi = \sum_{n=1}^\infty \psi(n) e_n\,.$$

Since ψ *is compactly supported, the sum is in fact finite, thus the operator is well-defined.*

We will show that for any $\phi \in \mathscr{H}$, $\phi \neq 0$, *the map* ℓ_ϕ^A *defined in* (2.5) *is not continuous and therefore* $\mathscr{D}(A^*) = \{0\}$ *by* (2.4).

Let $\phi \in \mathscr{H}$, $\phi \neq 0$, *be fixed, then* $\langle e_{n_0}, \phi\rangle \neq 0$ *for some* n_0. *Choose a sequence* (ψ_n) *in* $\mathscr{D}(A)$ *so that for all* $k \in \mathbb{N}$

$$\operatorname{supp}\psi_k \subset [n_0 - \tfrac{1}{2}, n_0 + \tfrac{1}{2}]\,, \quad \psi_k(n_0) = 1 \quad \textit{and} \quad \|\psi_k\|_{L^2} \longrightarrow 0.$$

Then for each $k \in \mathbb{N}$

$$\ell_\phi^A(\psi_k) = \langle \phi, A\psi_k \rangle = \sum_{n=1}^{\infty} \psi_k(n)\langle \phi, e_n \rangle = \langle \phi, e_{n_0} \rangle \neq 0$$

although $\psi_k \to 0$. Thus ℓ_ϕ^A is not continuous.

In the unbounded case, it is important to distinguish between symmetric and self-adjoint operators. Clearly each self-adjoint operator is symmetric and if T is symmetric and densely defined then $T \subset T^*$. In particular, T^* is densely defined in this case and it is possible to define $T^{**} = (T^*)^*$. The following Theorem gives a relationship between the notions of adjoint and closure.

Theorem 2.11
Let T be a densely defined linear operator on a Hilbert space $\mathscr{H}$, then

i) T^* is closed.

ii) T is closable if and only if $\mathscr{D}(T^*)$ is dense in $\mathscr{H}$. In this case $\overline{T} = T^{**}$.

iii) If T is closable, then $(\overline{T})^* = T^*$.

Proof. i) We use the characterisation given in Remark 2.4.
Let (y_n) be a sequence in $\mathscr{D}(T^*)$ such that $y_n \to y \in \mathscr{H}$ and $T^*y_n \to z \in \mathscr{H}$ as $n \to \infty$. We have to show that $y \in \mathscr{D}(T^*)$ and $T^*y = z$.

We have for any $x \in \mathscr{D}(T)$

$$\langle y, Tx \rangle = \lim_{n\to\infty} \langle y_n, Tx \rangle = \lim_{n\to\infty} \langle T^*y_n, x \rangle = \langle z, x \rangle.$$

Therefore ℓ_y^T given in (2.5) is bounded and it follows from (2.4) that $y \in \mathscr{D}(T^*)$. Moreover, $T^*y = z$ by Definition 2.8.

ii) "$\Longleftarrow$": Assume that $\mathscr{D}(T^*)$ is dense in $\mathscr{H}$. We have to show that T has a closed extension.

If $x \in \mathscr{D}(T)$, the linear map $\ell_x^{T^*}$ defined in (2.5) is bounded on $\mathscr{D}(T^*)$ and thus $x \in \mathscr{D}(T^{**})$ by (2.4). Moreover, for any $y \in \mathscr{D}(T^*)$

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = \langle T^{**}x, y\rangle\,.$$

Since by assumption $\mathscr{D}(T^*)$ is dense in $\mathscr{H}$, this implies $Tx = T^{**}x$ for all $x \in \mathscr{D}(T) \subset \mathscr{D}(T^{**})$ and thus $T \subset T^{**}$. Since T^{**} is closed by i), this shows that T has a closed extension and thus is closable.

Next we show that if T is closable then $\overline{T} = T^{**}$:

"$\overline{T} \subset T^{**}$" : Since T^{**} is closed by i), this follows from $T \subset T^{**}$ and the definition of the closure as smallest closed extension.

"$\overline{T} \supset T^{**}$" : Since the graphs of both operators are closed, it suffices to prove $\Gamma(T)^{\perp} \subset \Gamma(T^{**})^{\perp}$, where the orthogonal complement is taken with respect to the inner product (2.1) in $\mathscr{H} \times \mathscr{H}$. Let $(u, v) \in \Gamma(T)^{\perp}$, then

$$0 = \langle (u, v), (x, Tx)\rangle = \langle u, x\rangle + \langle v, Tx\rangle \quad \text{for all} \quad x \in \mathscr{D}(T). \tag{2.6}$$

This implies that ℓ_v^T is bounded on $\mathscr{D}(T)$. Thus from (2.4) and (2.6) we can deduce

$$(u, v) \in \Gamma(T)^{\perp} \qquad \Longrightarrow \qquad v \in \mathscr{D}(T^*) \quad \text{and} \quad T^*v = -u\,. \tag{2.7}$$

For any $(z, T^{**}z) \in \Gamma(T^{**})$ we therefore get

$$\begin{aligned} \langle (u, v), (z, T^{**}z)\rangle &= \langle u, z\rangle + \langle v, T^{**}z\rangle = \langle u, z\rangle + \langle T^*v, z\rangle \\ &= \langle u + T^*v, z\rangle = 0 \end{aligned}$$

and thus $(u, v) \in \Gamma(T^{**})^{\perp}$.

"$\Longrightarrow$": We use contraposition.

First we introduce the operator V on $\mathscr{H} \times \mathscr{H}$ given by $V(x, y) = (-y, x)$. Then V is unitary, since

$$\langle V(x, y), V(u, v)\rangle = \langle(-y, x), (-v, u)\rangle = \langle -y, -v\rangle + \langle x, u\rangle = \langle(x, y), (u, v)\rangle.$$

Thus $V\big(E^{\perp}\big) = \big(V(E)\big)^{\perp}$ for any subspace E of $\mathscr{H} \times \mathscr{H}$ and

$$\begin{aligned} (x, y) \in \big(V(\Gamma(T))\big)^{\perp} &\iff \langle(x, y), (-Tz, z)\rangle = 0 \quad \text{for all} \quad z \in \mathscr{D}(T) \\ &\iff \langle x, Tz\rangle = \langle y, z\rangle \quad \text{for all} \quad z \in \mathscr{D}(T) \\ &\iff (x, y) \in \Gamma(T^*). \end{aligned}$$

Thus

$$\Gamma(T^*) = \big(V(\Gamma(T))\big)^{\perp} \tag{2.8}$$

and

$$\overline{\Gamma(T)} = \big(\Gamma(T)^{\perp}\big)^{\perp} = \big(V^2(\Gamma(T)^{\perp})\big)^{\perp} = \Big(V\big(V(\Gamma(T))\big)^{\perp}\Big)^{\perp} = \Big(V(\Gamma(T^*))\Big)^{\perp}. \tag{2.9}$$

Now assume that $\mathscr{D}(T^*)$ is not dense. Then there exists $z \in \mathscr{D}(T^*)^{\perp}$ with $z \neq 0$. This implies that $(z, 0) \in \Gamma(T^*)^{\perp}$, because

$$\langle(z, 0), (x, T^*x)\rangle = \langle z, x\rangle + \langle 0, T^*x\rangle = 0 \quad \text{for all} \quad x \in \mathscr{D}(T^*).$$

But then $(0, z) \in V(\Gamma(T^*)^{\perp})$ and therefore $(0, z) \in \overline{\Gamma(T)}$.

If $(0, z)$ would be an element of the graph of some linear operator S, this would imply that $S0 = z \neq 0$, which is a contradiction to the linearity of S. Thus there exists no operator $\overline{T}$ such that $\Gamma(\overline{T}) = \overline{\Gamma(T)}$. Therefore by Proposition 2.5, T is not closable.

iii) If T is closable, then

$$T^* \overset{i)}{=} \overline{T^*} \overset{ii)}{=} T^{***} \overset{ii)}{=} \overline{T}^* .$$

□

We could have used the unitary operator V already for the first part of the proof (which then were a bit less direct):
Equation (2.8) implies that the graph of T^* is closed, since the orthogonal complement of a subspace is always closed.
If T^* is densely defined, (2.8) together with (2.9) give $\overline{\Gamma(T)} = \Gamma(T^{**})$, thus proving $T = T^{**}$ is this case.

2.1.3 Symmetric operators, essential self-adjointness

Theorem 2.11 leads at once to the following corollary.

Corollary 2.12
Let T be a densely defined linear operator on a Hilbert space $\mathscr{H}$, then

i) T is symmetric if and only if $T \subset T^*$. In this case T^{**} is symmetric as well and

$$T \subset T^{**} \subset T^* = T^{***} .$$

ii) T is closed and symmetric if and only if

$$T = T^{**} \subset T^* .$$

iii) T is self-adjoint if and only if

$$T = T^{**} = T^* .$$

Between i) and iii) there are operators with the following property.

Definition 2.13 (Essential Self-adjointness, Core)
Let T be a symmetric, densely defined operator on a Hilbert space $\mathscr{H}$.

i) T is called **essentially self-adjoint** $:\iff$ *$\overline{T}$ is self-adjoint.*

ii) If T is closed, a subset $C \subset \mathscr{D}(T)$ is called **core** *for T* $:\iff$ $\overline{T|_C} = T$.

Proposition 2.14
Let T, S be densely defined operators on a Hilbert space $\mathscr{H}$. Then the following statements hold.

i) If $T \subset S$ then $S^* \subset T^*$.

ii) If T is essentially self-adjoint, then it has a unique self-adjoint extension.

iii) If T is self-adjoint, then T has no proper symmetric extension.

iv) T is essentially self-adjoint if and only if

$$T \subset T^{**} = T^* .$$

Proof. Exercise 2.33 □

The importance of essential self-adjointness lies in the uniqueness of the self-adjoint extension. This allows to determine a self-adjoint operator A uniquely without giving the exact domain of A, which often is difficult. Instead it suffices to give a core for A.

2.1.4 Resolvent set and Spectrum for unbounded operators

Before we give criteria for (essential) self-adjointness of operators, we define the resolvent set in the case of closed unbounded operators.

DEFINITION 2.15 (RESOLVENT SET)
Let $\mathscr{H}$ be a Hilbert space and T a closed operator on $\mathscr{H}$ with domain $\mathscr{D}(T)$. A complex number $\lambda \in \mathbb{C}$ is in the resolvent set $\rho(T)$ of T :$\Longleftrightarrow$
$(\lambda \operatorname{Id} - T)$ is a bijection of $\mathscr{D}(T)$ onto $\mathscr{H}$ with bounded inverse.
If $\lambda \in \rho(T)$, then the operator $R_\lambda(T) = (\lambda \operatorname{Id} - T)^{-1}$ is called resolvent of T at λ.
The definitions of spectrum, point spectrum, continuous spectrum and residual spectrum are exactly as given in Definition 1.1 for bounded operators.

Since T is assumed to be closed, it follows from the closed graph theorem that if $(\lambda \operatorname{Id} - T)$ is a bijection of $\mathscr{D}(T)$ onto $\mathscr{H}$, then its inverse is bounded. If we speak of the spectrum of a closable operator, we always mean the spectrum of its closure.

As in the case of bounded operators (Theorem 80), we have the following Theorem.

THEOREM 2.16
Let T be a closed densely defined linear operator on a Hilbert space $\mathscr{H}$. Then the resolvent set of T is an open subset of $\mathbb{C}$ on which the resolvent is an analytic operator-valued function. Furthermore

$$\{R_\lambda(T) \mid \lambda \in \rho(T)\}$$

is a commuting family of bounded operators satisfying

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda) R_\lambda(T) R_\mu(T) \qquad \textbf{First Resolvent Formula} \tag{2.10}$$

The proof is exactly as in the case of bounded operators.

The following example will show that the spectrum of an unbounded operator depends on the choice of the domain. Before that we give for completeness the definition of absolute continuity.

Definition 2.17 (Absolute Continuity)
Let $J \subset \mathbb{R}$ be a (possibly unbounded) interval. Then a function $f : J \to \mathbb{C}$ is called **absolutely continuous**: $\Longleftrightarrow$
For each $\epsilon > 0$ there exists $\delta > 0$ such that for every finite collection of disjoint intervals $[a_k, b_k] \subset J$, $k = 1, \dots n$, $n \in \mathbb{N}$

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

Then the Fundamental Theorem of Calculus holds:

Theorem 2.18 (Fundamental Theorem of Calculus)
Let $J \subset \mathbb{R}$ be a (possibly unbounded) interval.

i) If $f : J \to \mathbb{C}$ is absolutely continuous, then f is differentiable almost everywhere, the derivative $f' \in L^1(J)$ and

$$f(t) - f(t_0) = \int_{t_0}^{t} f'(s)\, ds \qquad \textit{for all} \quad t_0, t \in J\,.$$

ii) If $g : J \to \mathbb{C}$ is integrable over compact subintervals and

$$f(t) := \int_a^t g(s)\, ds$$

for an arbitrary $a \in J$, then f is absolutely continuous, f' exists almost everywhere and $f' = g$ almost everywhere.

Example 2.19

We consider operators given by $i\frac{d}{dt}$ on different domains in $\mathscr{H} = L^2([0,1])$.

We already introduced in Example 2.2iii) the operator T acting on its domain $\mathscr{D}(T) = \{\phi \in \mathcal{C}^1([0,1]) \,|\, \phi(1) = \phi(0) = 0\}$ as $T\phi = i\frac{d}{dt}\phi$. T is symmetric but not closed and thus not self-adjoint. To see that it is not closed, consider the sequence

$$\phi_n(t) = \left((t-\tfrac{1}{2})^2 + \tfrac{1}{n}\right)^{1/2} - \left(\tfrac{1}{4} + \tfrac{1}{n}\right)^{1/2}, \qquad t \in [0,1].$$

Then $\phi_n \in \mathcal{C}^1([0,1])$ and $\phi_n(0) = 0 = \phi_n(1)$, thus $\phi_n \in \mathscr{D}(T)$ for all $n \in \mathbb{N}$. Moreover

$$\phi_n(t) \longrightarrow \phi(t) = |t - \tfrac{1}{2}| - \tfrac{1}{2}, \qquad t \in [0,1],$$

and the convergence is uniform and thus in $L^2([0,1])$:

Setting $t - \frac{1}{2} = s$ and using that $|s| \le \frac{1}{2}$ gives for any $t \in [0,1]$

$$\begin{aligned}
|\phi_n(t) - \phi(t)| &= \left|\left(s^2 + \tfrac{1}{n}\right)^{1/2} - \left(\tfrac{1}{4} + \tfrac{1}{n}\right)^{1/2} - \left(s^2\right)^{1/2} + \tfrac{1}{2}\right| \\
&= \left(\tfrac{1}{4} + \tfrac{1}{n}\right)^{1/2} - \tfrac{1}{2} + \left(s^2\right)^{1/2} - \left(s^2 + \tfrac{1}{n}\right)^{1/2} \\
&= \frac{\frac{1}{4} + \frac{1}{n} - \frac{1}{4}}{\left(\frac{1}{4} + \frac{1}{n}\right)^{1/2} + \frac{1}{2}} + \frac{s^2 - s^2 + \frac{1}{n}}{\left(s^2\right)^{1/2} + \left(s^2 + \frac{1}{n}\right)^{1/2}} \\
&\le \frac{1}{n} + \frac{1/n}{1/\sqrt{n}} \longrightarrow 0 \qquad \text{as} \quad n \to \infty\,.
\end{aligned}$$

We have for any $t \in [0,1]$

$$T\phi_n(t) = i\frac{d}{dt}\phi_n(t) = i\frac{t-\frac{1}{2}}{\left((t-\frac{1}{2})^2+\frac{1}{n}\right)^{1/2}} \longrightarrow \psi(t) := \begin{cases} -i & \text{if } t < \frac{1}{2} \\ 0 & \text{if } t = 0 \\ i & \text{if } t > \frac{1}{2} \end{cases}$$

and this convergence is in fact in L^2*:*

With the substitution $t - \frac{1}{2} = s$

$$\begin{aligned} &\int_0^1 \left|T\phi_n(t) - \psi(t)\right|^2 dt \\ &\quad = \int_{-1/2}^0 \left|\frac{s}{\sqrt{s^2+\frac{1}{n}}}+1\right|^2 ds + \int_0^{1/2} \left|\frac{s}{\sqrt{s^2+\frac{1}{n}}}-1\right|^2 ds \\ &\quad = 2\int_0^{1/2} \left(1-\frac{s}{\sqrt{s^2+\frac{1}{n}}}\right)^2 ds \\ &\quad = 2\int_0^{1/2} \left(1+\frac{s^2}{s^2+\frac{1}{n}}-\frac{2s}{\sqrt{s^2+\frac{1}{n}}}\right) ds \\ &\quad = 2\left[s+s-\frac{1}{\sqrt{n}}\arctan\left(\sqrt{n}s\right)-2\left(s^2+\frac{1}{n}\right)^{1/2}\right]_0^{1/2} \\ &\quad = 2\left(1-\frac{1}{\sqrt{n}}\arctan\left(\sqrt{n}s\right)-2\left(\frac{1}{4}+\frac{1}{n}\right)^{1/2}+\frac{2}{\sqrt{n}}\right) \\ &\quad \longrightarrow 0 \text{ as } n \to \infty\,. \end{aligned}$$

Thus we found a sequence in the graph $\Gamma(T)$ *converging to some element* $(\phi,\psi) \in \mathscr{H} \times \mathscr{H}$*, but since* $\phi \notin \mathscr{D}(T)$ *(since the absolute value is not differentiable),* $\Gamma(T)$ *is not closed and therefore* T *is not closed.*

Now we set

$$\mathrm{AC}[0,1] := \{\phi \in \mathcal{C}([0,1]) \,|\, \phi \textit{ is absolutely continuous with } \tfrac{d}{dt}\phi \in L^2([0,1])\}$$

and define the operators T_1 *and* T_2 *by setting* $T_k\phi = i\frac{d}{dt}\phi$ *for* $k = 1, 2$ *and*

$$\mathscr{D}(T_1) = \mathrm{AC}[0,1] \quad \textit{and} \quad \mathscr{D}(T_2) = \{\phi \in \mathrm{AC}[0,1] \,|\, \phi(0) = 0\}\,.$$

Then both domains are dense in $\mathscr{H}$ *and both operators are closed.*

To see that T_1 *is closed, let* (x_n) *be a sequence in* $\mathscr{D}(T_1)$ *such that*

$$x_n \to x \in \mathscr{H} \qquad \textit{and} \qquad T_1 x_n \to y \in \mathscr{H} \qquad \textit{in} \qquad L^2([0,1]). \tag{2.11}$$

Since x_n *is absolutely continuous, it follows from Theorem 2.18 that*

$$x_n(t) = x_n(0) + \tfrac{1}{i}\int_0^t T_1 x_n(s)\,ds \qquad \textit{for all} \quad t \in [0,1]\,. \tag{2.12}$$

We will use (2.12) *to show that* $x_n(t)$ *converges uniformly.*

We first observe that by the Hölder inequality for any $t \in [0,1]$

$$\begin{aligned}\Big|\int_0^t T_1 x_n(s)\,ds - \int_0^t y(s)\,ds\Big| &\le \int_0^t |T_1 x_n(s) - y(s)|\,ds \\ \le \int_0^1 |T_1 x_n(s) - y(s)|\,ds &\le \Big(\int_0^1 |T_1 x_n(s) - y(s)|^2\,ds\Big)^{1/2}\Big(\int_0^1 1^2\,ds\Big)^{1/2} \\ &= \|T_1 x_n - y\|_{L^2} \longrightarrow 0 \qquad \textit{as } n \to \infty\,.\end{aligned}$$

where for the convergence we used (2.11)*. Thus*

$$\int_0^t T_1 x_n(s)\,ds \longrightarrow_{n\to\infty} \int_0^t y(s)\,ds \quad \textit{uniformly.} \tag{2.13}$$

Moreover, using (2.12) *again, we have for any* $n, m \in \mathbb{N}$

$$x_n(0) - x_m(0) = x_n(t) - x_m(t) + \tfrac{1}{i}\int_0^t (T_1 x_m(s) - T_1 x_n(s))\,ds$$

and therefore

$$
\begin{aligned}
&|x_n(0) - x_m(0)| = \Big(\int_0^1 |x_n(0) - x_m(0)|^2 \, dt \Big)^{1/2} \\
&\le \Big(\int_0^1 |x_n(t) - x_m(t)|^2 \, dt \Big)^{1/2} + \Big(\int_0^1 \Big| \int_0^t (T_1 x_m(s) - T_1 x_n(s)) \, ds^2 \Big| \, dt \Big)^{1/2} .
\end{aligned}
\tag{2.14}
$$

Since by (2.11) *and* (2.13) *both summands on the right hand side of* (2.14) *converge to zero, it follows that* $(x_n(0))$ *is a Cauchy sequence and thus convergent. Set*

$$
\lim_{n\to\infty} x_n(0) = a \quad \textit{and} \quad z(t) = a + \tfrac{1}{i} \int_0^t y(s) \, ds \, ,
$$

then combining (2.12) *with* (2.13) *and the definition of* a *shows that* $x_n \to z$ *uniformly. Moreover,* z *is absolutely continuous by Theorem 2.18, since* $y \in L^2([0,1]) \subset L^1([0,1])$*, and* $\frac{d}{dt} z(t) = \frac{1}{i} y(t)$ *almost everywhere. This shows that*

$$
z \in \mathscr{D}(T_1) \qquad \textit{and} \qquad T_1 z = y \, .
$$

Since $x_n \to x$ *in* L^2 *and* $x_n \to z$ *uniformly, it follows that* $x(t) = z(t)$ *almost everywhere and thus* $x = z$ *in* $L^2([0,1])$*. Thus* T_1 *is closed.*

Similarly, it can be proven that T_2 *is closed (each sequence* x_n *in* $\mathscr{D}(T_2)$ *has the fixed value* $x_n(0) = 0$*, thus* $a = 0$ *in the above proof).*

But although T_1 *is an extension of* T_2 *and both operators are closed, the spectra of* T_1 *and* T_2 *are not equal. In fact we have*

$$
\sigma(T_1) = \mathbb{C} \qquad \textit{and} \qquad \sigma(T_2) = \emptyset \, . \tag{2.15}
$$

To see that $\sigma(T_1) = \mathbb{C}$, *observe that for all* $\lambda \in \mathbb{C}$

$$e^{-i\lambda t} \in \mathscr{D}(T_1) \qquad \textit{and} \qquad T_1 e^{-i\lambda t} = \lambda e^{-i\lambda t} .$$

To see that $\sigma(T_2) = \emptyset$, *we show that* $\lambda \operatorname{Id} - T_2$ *has an inverse for each* $\lambda \in \mathbb{C}$ *and thus* $\rho(T_2) = \mathbb{C}$. *In fact, define the operator* S_λ *on* $\mathscr{H}$ *given by*

$$S_\lambda x(t) = i \int_0^t e^{-i\lambda(t-s)} x(s)\, ds .$$

Then $S_\lambda x \in \mathscr{D}(T_2)$ *for any* $x \in \mathscr{H}$ *and by product rule and Theorem 2.18*

$$\begin{aligned} T_2 S_\lambda x(t) &= i \frac{d}{dt}\Big(i e^{-i\lambda t} \int_0^t e^{i\lambda s} x(s)\, ds \Big) \\ &= \lambda i e^{-i\lambda t} \int_0^t e^{i\lambda s} x(s)\, ds - e^{-i\lambda t} e^{i\lambda t} x(t) \\ &= \lambda S_\lambda x(t) - x(t) \end{aligned}$$

and therefore

$$(\lambda \operatorname{Id} - T_2) S_\lambda = \operatorname{Id} \qquad \textit{on} \quad \mathscr{H} .$$

On the other hand, for any $x \in \mathscr{D}(T_2)$ *we get by integration by parts, using* $x(0) = 0$,

$$\begin{aligned} S_\lambda T_2 x(t) &= i e^{-i\lambda t} \int_0^t e^{i\lambda s} i \frac{d}{ds} x(s)\, ds \\ &= -e^{-i\lambda t} \Big(\big[e^{i\lambda s} x(s) \big]_0^t - \int_0^t i\lambda e^{i\lambda s} x(s)\, ds \Big) \\ &= -x(t) + S_\lambda \lambda x(t) . \end{aligned}$$

This shows for each $\lambda \in \mathbb{C}$

$$S_\lambda (\lambda \operatorname{Id} - T_2) = \operatorname{Id} \qquad \textit{on} \quad \mathscr{D}(T_2) .$$

Therefore $(\lambda \operatorname{Id} - T_2)$ *is invertible on* $\mathscr{D}(T_2)$ *for any* $\lambda \in \mathbb{C}$ *and thus* $\rho(T_2) = \mathbb{C}$ *by Definition 2.15. This shows* $\sigma(T_2) = \emptyset$.

Now we will show that $T^* = T_1$.
Let $x \in \mathscr{D}(T^*)$ *and set*

$$y = T^*x \qquad \textit{and} \qquad F(t) = \int_0^t y(s)\, ds\,.$$

Since $y \in L^2([0,1]) \subset L^1([0,1])$ *by the definition of the domain of* T^*, *it follows from Theorem 2.18 that*

$$F \textit{ is absolutely continuous and } \quad F' = y \textit{ almost everywhere.} \tag{2.16}$$

Thus for any $z \in \mathscr{D}(T)$, *integration by parts yields*

$$\langle x, Tz\rangle = \langle y, z\rangle = \langle F', z\rangle = -\langle F, z'\rangle = \langle -iF, Tz\rangle\,.$$

Therefore

$$x + iF \in \operatorname{Ran}(T)^\perp. \tag{2.17}$$

Moreover, it follows from the definition of $\mathscr{D}(T)$ *that for each* $x \in \mathscr{D}(T)$

$$Tx \in \mathcal{C}([0,1]) \quad \textit{and} \quad \int_0^1 Tx(s)\, ds = i\int_0^1 x'(s)\, ds = i(x(1) - x(0)) = 0\,.$$

Since the closure of $\mathcal{C}([0,1])$ *with respect to the* L^2*-norm is* $L^2([0,1])$, *this shows*

$$\overline{\operatorname{Ran}(T)} = \Big\{ w \in L^2([0,1]) \,\big|\, \int_0^1 w(s)\, ds = 0 \Big\} = \{1\}^\perp\,.$$

Therefore $\operatorname{Ran}(T)^\perp = \{1\}^{\perp\perp} = \operatorname{span}\{1\}$ *and by* (2.17)

$$x + iF \in \operatorname{span}\{1\} \qquad \textit{and} \qquad x = -iF + \alpha 1 \quad \textit{for all} \quad x \in \mathscr{D}(T^*)\,.$$

Thus $x \in \mathscr{D}(T^*)$ *is absolutely continuous and by* (2.16)

$$L^2([0,1[) \ni T^*x = y = F' = i\frac{d}{dt}x = T_1x \quad \textit{and thus} \quad x \in \mathrm{AC}[0,1] = \mathscr{D}(T_1)\,,$$

proving $T^* \subset T_1$. *On the other hand, assume* $x \in \mathrm{AC}[0,1]$ *and set*

$$y := i\frac{d}{dt}x \in L^2([0,1])\,.$$

Then as above integration by parts yields

$$\langle x, Tz\rangle = \langle y, z\rangle \qquad \textit{for all} \quad z \in \mathscr{D}(T)$$

and therefore $T^*x = y = T_1x$ *and* $x \in \mathscr{D}(T^*)$. *This proves* $T^* \subset T_1$ *and we thus have shown* $T^* = T_1$.

We remark that although T *is symmetric, its adjoint* T^* *is not symmetric. Since by Corollary 2.12 we know that* T^{**} *is also symmetric, it follows that* $T^{**} \neq T^*$ *and therefore* T *is not essentially self-adjoint.*

Moreover, by Theorem 2.11, Corollary 2.12 and Proposition 2.14, if T *has any self-adjoint extension* S, *then*

$$T \subset \overline{T} = T^{**} \subset S = S^* \subset T^* = T^{***}\,. \tag{2.18}$$

In the next step, we will determine T^{**}. *We already know from above that* $T^* = T_1$, *i.e.* $\mathscr{D}(T^*) = \mathrm{AC}[0,1]$ *and* $T^*x(t) = i\frac{d}{dt}x(t)$. *Thus by* (2.18)

$$x \in \mathscr{D}(T^{**}) \quad \Longrightarrow \quad x \in \mathrm{AC}[0,1] \quad \textit{and} \quad T^{**}x(t) = i\frac{d}{dt}x(t)\,.$$

Let $x \in \mathscr{D}(T^{**})$, *then for any* $z \in \mathscr{D}(T^*)$ *we have on one hand*

$$\langle x, T^*z\rangle = \langle T^{**}x, z\rangle = \int_0^1 \big(-i\tfrac{d}{dt}\overline{x(t)}\big)\cdot z(t)\,dt \tag{2.19}$$

and on the other hand by integration by parts

$$\begin{aligned}\langle x, T^* z\rangle &= \int_0^1 \overline{x(t)} \cdot \big(i\tfrac{d}{dt} z(t)\big)\, dt \\ &= i\big(\bar{x}(1)z(1) - \bar{x}(0)z(0)\big) + \int_0^1 \big(-i\tfrac{d}{dt}\overline{x(t)}\big) \cdot z(t)\, dt\,.\end{aligned} \tag{2.20}$$

Combining (2.19) *and* (2.20) *shows that*

$$\bar{x}(1)z(1) - \bar{x}(0)z(0) = 0 \qquad \textit{for all} \quad z \in \mathrm{AC}[0,1]$$

and therefore $x(1) = x(0) = 0$*. It follows that*

$$T^{**} = i\frac{d}{dt} \qquad \textit{with} \qquad \mathscr{D}(T^{**}) = \{x \in \mathrm{AC}[0,1] \,|\, x(0) = x(1) = 0\}\,.$$

We remark that T_2 *is an extension of* T^{**}*, thus the spectrum of* T^{**} *is empty as well.*

Since T^{**} *is not equal to* $T^* = T^{***}$*, it follows that* T^{**} *is an extension of* T *which is closed and symmetric, but not self-adjoint. Since* T *is not essentially self-adjoint, there might be no or many self-adjoint extensions of* T*.*

In our example, T *has infinitely many self-adjoint extensions. In fact, consider for any* $\lambda \in \mathbb{C}$ *with* $|\lambda| = 1$ *the operators* S_λ *on* $\mathscr{H}$ *acting on their domain*

$$\mathscr{D}(S_\lambda) = \{x \in \mathrm{AC}[0,1] \,|\, x(0) = \lambda x(1)\} \qquad \textit{as} \qquad S_\lambda x(t) = i\frac{d}{dt}x(t)\,. \tag{2.21}$$

Then $T \subset S_\lambda \subset T^*$ *and* S_λ *is symmetric, since for all* $x, y \in \mathscr{D}(S_\lambda)$

$$\bar{x}(1)z(1) - \bar{x}(0)z(0) = \bar{x}(1)z(1) - \bar{\lambda}\bar{x}(1)\lambda z(1) = \big(1 - |\lambda|^2\big)\bar{x}(1)z(1) = 0 \tag{2.22}$$

by the assumption $|\lambda|^2 = 1$ *and therefore*

$$\langle x, S_\lambda y\rangle = i\int_0^1 \bar{x}(t)\big(\tfrac{d}{dt}y(t)\big)\, dt \tag{2.23}$$

$$= i\big(\bar{x}(1)z(1) - \bar{x}(0)z(0)\big) - i\int_0^1 \big(\tfrac{d}{dt}\overline{x(t)}\big)\cdot y(t)\,dt$$
$$= \langle S_\lambda x, y\rangle\,.$$

Thus $S_\lambda \subset S_\lambda^*$ by Corollary 2.12.

To see that $S_\lambda^* \subset S_\lambda$, we can proceed using arguments similar to those given above:

From Proposition 2.14 and $T \subset S_\lambda$ it follows that $S_\lambda \subset T^*$ and thus

$$x \in \mathscr{D}(S_\lambda^*) \quad \Longrightarrow \quad x \in \mathrm{AC}[0,1] \quad \textit{and} \quad S_\lambda^* x(t) = i\frac{d}{dt}x(t)\,. \tag{2.24}$$

Let $x \in \mathscr{D}(S_\lambda^*)$, then for any $z \in \mathscr{D}(S_\lambda)$ we have on one hand

$$\langle x, S_\lambda z\rangle = \langle S_\lambda^* x, z\rangle = \int_0^1 \big(-i\tfrac{d}{dt}\overline{x(t)}\big)\cdot z(t)\,dt \tag{2.25}$$

and on the other hand integration by parts shows

$$\langle x, S_\lambda z\rangle = \int_0^1 \overline{x(t)}\cdot\big(i\tfrac{d}{dt}z(t)\big)\,dt \tag{2.26}$$
$$= i\big(\bar{x}(1)z(1) - \bar{x}(0)z(0)\big) + \int_0^1 \big(-i\tfrac{d}{dt}\overline{x(t)}\big)\cdot z(t)\,dt\,.$$

Combining (2.25) and (2.26) shows that

$$\bar{x}(1)z(1) - \bar{x}(0)z(0) = (\bar{x}(1) - \lambda\bar{x}(0))z(1) = 0 \qquad \textit{for all} \quad z \in \mathscr{D}(S_\lambda)$$

and therefore $x(1) = \bar{\lambda}x(0)$. Since $\bar{\lambda} = \lambda^{-1}$ for $|\lambda| = 1$, it follows that

$$\mathscr{D}(S_\lambda^*) \subset \{x \in \mathrm{AC}[0,1]\,|\,x(0) = \lambda x(1)\} = \mathscr{D}(S_\lambda)$$

and therefore $S_\lambda^* \subset S_\lambda$ by (2.24).

Thus we have proven that for any $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ the operator S_λ defined in (2.21) is a self-adjoint extension of T (and of T^{**}). Thus T has in fact infinitely many self-adjoint extensions.

2.1.5 Criteria for (essential) self-adjointness

We now want to get criteria to decide, whether an operator on a Hilbert space is (essentially) self-adjoint.

We first observe that if T is a self-adjoint operator and if there exists $x \in \mathscr{D}(T^*) = \mathscr{D}(T)$ such that $T^*x = ix$, then $Tx = ix$ and thus

$$-i\langle x, x\rangle = \langle ix, x\rangle = \langle Tx, x\rangle = \langle x, T^*x\rangle = \langle x, Tx\rangle = i\langle x, x\rangle\,,$$

and therefore $x = 0$. A similar computation shows that $T^*x = -ix$ has no non-zero solutions.

Thus if T is self-adjoint, then $\operatorname{Ker}(T^* \pm i\operatorname{Id}) = \{0\}$.

The converse statement is one of the the basic criteria for self-adjointness.

We start with a lemma, giving the relation between the kernel of an operator and the range of its adjoint (see also Exercise 1.42). We set $\mathscr{D}(\lambda\operatorname{Id} - T) = \mathscr{D}(T)$ for $\lambda \in \mathbb{C}$.

Lemma 2.20

Let $\mathscr{H}$ be a Hilbert space and $T : \mathscr{H} \supset \mathscr{D}(T) \longrightarrow \mathscr{H}$ a densely defined linear operator. Then

i) $\operatorname{Ker}(T^* \mp i\operatorname{Id}) = \operatorname{Ran}(T \pm i\operatorname{Id})^\perp$ *and in particular*

$$\operatorname{Ker}(T^* \mp i\operatorname{Id}) = \{0\} \quad \Longleftrightarrow \quad \operatorname{Ran}(T \pm i\operatorname{Id}) \quad \textit{is dense in } \mathscr{H}\,.$$

ii) if T is closed and symmetric, $\operatorname{Ran}(T \pm i\operatorname{Id})$ *are closed.*

Proof. We only prove the case $T + i$, the case $T - i$ is similar.

i) First remark that $(T + i)^* = T^* - i$.

$$\text{"} \supset \text{"}: \quad y \in \operatorname{Ran}(T+i)^\perp \Longrightarrow \langle y, (T+i)z\rangle = 0 \quad \text{for all} \quad z \in \mathscr{D}(T)$$

$$\Longrightarrow y \in \mathscr{D}(T^*) \text{ and } \langle (T^* - i)y, z\rangle = 0 \quad \text{for all} \quad z \in \mathscr{D}(T)$$
$$\Longrightarrow y \in \operatorname{Ker}(T^* - i)\,.$$

" $\subset$ " : $\quad y \in \operatorname{Ker}(T^* - i) \Longrightarrow y \in \mathscr{D}(T^*)$ and $\langle (T^* - i)y, z\rangle = 0\ \forall z \in \mathscr{H}$

$$\Longrightarrow \quad \langle y, (T+i)z\rangle = 0 \quad \text{for all} \quad z \in \mathscr{D}(T)$$
$$\Longrightarrow \quad y \in \operatorname{Ran}(T+i)^{\perp}\,.$$

ii) Since T is symmetric, $\langle x, Tx\rangle \in \mathbb{R}$ and thus for all $x \in \mathscr{D}(T)$

$$\|(T+i)x\|^2 = \|Tx\|^2 + \|x\|^2 + 2\operatorname{Re}\langle ix, Tx\rangle = \|Tx\|^2 + \|x\|^2 \geq \|x\|^2\,. \quad (2.27)$$

Therefore $(T+i)^{-1} : \operatorname{Ran}(T+i) \to \mathscr{D}(T)$ exists and is continuous[5].

Let (x_n) be a sequence in $\mathscr{D}(T)$ such that $(T+i)x_n \to y \in \overline{\operatorname{Ran}(T+i)}$. Then $((T+i)x_n)$ is a Cauchy sequence in $\operatorname{Ran}(T+i)$ and therefore (x_n) is a Cauchy sequence in $\mathscr{D}(T)$. Thus there exists $x \in \mathscr{H}$ with $x = \lim_{n\to\infty} x_n$ and moreover $Tx_n \to y - ix$. Since by assumption T is closed, it follows that $x \in \mathscr{D}(T)$ and $y = (T+i)x \in \operatorname{Ran}(T+i)$.

□

Theorem 2.21 (Basic Criterion: Self-Adjointness)
Let T be a densely defined symmetric operator on a Hilbert space $\mathscr{H}$. Then the following three statements are equivalent:

i) T *is self-adjoint.*

ii) T *is closed and* $\operatorname{Ker}(T^* \pm i\operatorname{Id}) = \{0\}$.

iii) $\operatorname{Ran}(T \pm i\operatorname{Id}) = \mathscr{H}$.

Proof. "i) $\Rightarrow$ ii)" : If T is self-adjoint, then $T = T^*$ is closed by Theorem 2.11i). Assume $(T^* + i)x = 0$. Since T^* is symmetric, it follows from (2.27) that $x = 0$ and thus $\mathrm{Ker}(T^* + i) = \{0\}$ and analogously $\mathrm{Ker}(T^* - i) = \{0\}$.

"ii)$\Rightarrow$ iii)" : If T is closed and symmetric, $\mathrm{Ran}(T \pm i)$ is closed by Lemma 2.20ii). Since $\mathrm{Ker}(T^* \mp i) = \{0\}$ it follows from Lemma 2.20i) that $\mathrm{Ran}(T \pm i)$ is dense in $\mathscr{H}$. Thus $\mathrm{Ran}(T \pm i)$ is closed and dense in $\mathscr{H}$ showing $\mathrm{Ran}(T \pm i) = \mathscr{H}$.

"iii)$\Rightarrow$i)" : Since T is symmetric, it follows from Corollary 2.12 that $T \subset T^*$, thus it suffices to show $\mathscr{D}(T^*) \subset \mathscr{D}(T)$. Let $y \in \mathscr{D}(T^*)$. Since by assumption $\mathrm{Ran}(T \pm i) = \mathscr{H}$, there exists some $x \in \mathscr{D}(T)$ such that $(T^* - i)y = (T - i)x$. But since $T \subset T^*$ this implies $(T^* - i)y = (T^* - i)x$. Therefore, $(y - x) \in \mathrm{Ker}(T^* - i)$ and since $\mathrm{Ker}(T^* - i) = \{0\}$ by iii) together with Lemma 2.20i), it follows that $y = x \in \mathscr{D}(T)$.

□

There are similar criteria for essential self-adjointness.

Corollary 2.22 (Basic Criterion: Essential Self-Adjointness) *Let T be a densely defined symmetric operator on a Hilbert space $\mathscr{H}$. Then the following three statements are equivalent:*

i) T is essentially self-adjoint.

ii) $\mathrm{Ker}(T^* \pm i\,\mathrm{Id}) = \{0\}$.

iii) $\mathrm{Ran}(T \pm i\,\mathrm{Id})$ *is dense in* $\mathscr{H}$.

Proof. "i) $\Leftrightarrow$ ii)" : T is essentially self-adjoint if and only if T^{**} is self-adjoint, thus this statement follows from Theorem 2.21 applied to T^{**} and $T^* = T^{***}$.

"ii)⇔ iii)" : Lemma 2.20.

□

2.1.6 Deficiency Indices, existence of self-adjoint extensions

DEFINITION 2.23

Let T be a symmetric densely defined operator on a Hilbert space $\mathscr{H}$. The numbers

$$n_{\pm} := \dim \operatorname{Ker}(T^* \pm i)$$

are called **deficiency indices** *of T.*

Here $\dim M$ denotes the cardinality of the basis of the subspace $M \in \mathscr{H}$.

In the next theorem we give a criterion to decide wether a symmetric operator has a self-adjoint extension. As we saw in Example 2.19, this does not imply that T is essentially self-adjoint (or that this extension is unique).

THEOREM 2.24

Let T be a densely defined symmetric operator on a Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(T)$. Then T has a self-adjoint extension if and only if its deficiency indices are equal, i.e.

$$\exists S \,:\, T \subset S = S^* \quad \Longleftrightarrow \quad \dim \operatorname{Ker}(T^* + i) = \dim \operatorname{Ker}(T^* - i)\,.$$

Proof. Since T is symmetric, it is closable (since $T \subset T^*$ by Corollary 2.12 and T^* is closed by Theorem 2.11). Moreover by Proposition 2.14, if S is a self-adjoint extension of T, then $T \subset \overline{T} \subset S = S^* \subset T^*$. Thus T has a self-adjoint extension if and only if $\overline{T}$ has a self-adjoint extension. Thus we can assume without loss of generality that T is closed.

"$\Longleftarrow$": We assume that $n_+ = n_-$. By equation (2.27)

$$\|(T+i)x\| = \|(T-i)x\| \qquad \text{for all} \quad x \in \mathscr{D}(T)\,,$$

and by Lemma 2.20 $\mathrm{Ran}(T \pm i)$ is closed. Thus the operator

$$U \,:\, \mathrm{Ran}(T-i) \longrightarrow \mathrm{Ran}(T+i), \quad (T-i)x \mapsto (T+i)x$$

is a well-defined, isometric, surjective operator ($U = (T+i)(T-i)^{-1}$ defined on $\mathrm{Ran}(T-i)$ is called **Cayley-transform** of T). In $\mathscr{H}$ we choose an orthonormal basis $\{e_i\}$ such that the subset $\{e_{i_k}\}$ is a basis for $\mathrm{Ran}(T-i)^{\perp}$ and an orthonormal basis $\{f_j\}$ such that the subset $\{f_{j_k}\}$ is a basis for $\mathrm{Ran}(T+i)^{\perp}$. Since by assumption $n_+ = n_-$, it follows from Lemma 2.20i) that

$$|\{f_{j_k}\}| = \dim \mathrm{Ran}(T+i)^{\perp} = \dim \mathrm{Ran}(T-i)^{\perp} = |\{e_{i_k}\}|\,,$$

where $|M|$ denotes the cardinality of the set M. Thus there is a one-to-one correspondence between $\{e_{i_k}\}$ and $\{f_{j_k}\}$, which allows us to extend U to a unitary map $V \,:\, \mathscr{H} \longrightarrow \mathscr{H}$. In the first step, we show that $V - \mathrm{Id}$ is injective: Assume that $y \in \mathrm{Ker}(V - \mathrm{Id})$, then $Vy = y$ and since V is unitary (and thus $V^*V = \mathrm{Id}$) this implies $y = V^*y$. Then for any $x \in \mathscr{D}(T)$, using $V(T-i)x = (T+i)x$ by the definition of V,

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

But this implies that $y \in \mathscr{D}(T)^{\perp}$ and since T is densely defined $y = 0$.

Thus we can define the operator

$$S = i(V+\mathrm{Id})(V-\mathrm{Id})^{-1} \,:\, \mathrm{Ran}(V-\mathrm{Id}) \longrightarrow \mathscr{H}\,, \quad Vz - z \mapsto i(Vz+z)\,.$$

The goal is to show that S is in fact a self-adjoint extension of T.

Since for any $x \in \mathscr{D}(T)$ we have

$$(V-\mathrm{Id})(T-i)x = (T+i)x - (T-i)x = 2ix \tag{2.28}$$

it follows that $x \in \mathrm{Ran}(V-\mathrm{Id}) = \mathscr{D}(S)$ and thus $\mathscr{D}(T) \subset \mathscr{D}(S)$. Moreover for any $x \in \mathscr{D}(T)$ it follows from (2.28) and the definition of S and V that

$$Sx = \frac{1}{2i}S(V-\mathrm{Id})(T-i)x = \frac{1}{2}(V+\mathrm{Id})(T-i)x = \frac{1}{2}\big((T+i)x+(T-i)x\big) = Tx\,.$$

This shows that $T \subset S$.

Moreover, S is symmetric, since for any $x = (V-\mathrm{id})y \in \mathscr{D}(S)$

$$\begin{aligned}
\langle x, Sx\rangle &= \langle (V-\mathrm{Id})y, i(V+\mathrm{Id})y\rangle \\
&= i\big(\langle Vy,Vy\rangle - \langle y,Vy\rangle + \langle Vy,y\rangle - \langle y,y\rangle\big) \\
&= i\big(\langle Vy,y\rangle - \langle y,Vy\rangle\big) \\
&= -2\,\mathrm{Im}\langle y,Vy\rangle \in \mathbb{R}
\end{aligned}$$

where we used that V is unitary. Thus S is symmetric by Exercise 2.35.

To finally show the self-adjointness of S, we use that by Theorem 2.21 it suffices to show $\mathrm{Ran}(S \pm i) = \mathscr{H}$. But since for any $z \in \mathscr{H}$

$$(S-i)(V-\mathrm{Id})z = i(V+\mathrm{Id})z - i(V-\mathrm{Id})z = 2iz$$

and

$$(S+i)(V-\mathrm{Id})V^*z = i(V+\mathrm{Id})V^*z + i(V-\mathrm{Id})V^*z = iz + iV^*z + iz - iV^*z = 2iz\,,$$

each element of $\mathscr{H}$ is in the range of both $(T-i)$ and $(T+i)$, showing that S is self-adjoint.

"$\Longrightarrow$": Assume that T has a self-adjoint extension S. Thus by Theorem 2.21 we can define the map

$$V=(S+i)(S-i)^{-1} \quad \text{with domain} \quad \mathscr{D}(V)=\operatorname{Ran}(S-i)=\mathscr{H}\,.$$

Then V is surjective, because $\operatorname{Ran} V=\operatorname{Ran}(S+i)=\mathscr{H}$ by Theorem 2.21. Moreover, by (2.27)

$$\|V(S-i)x\|=\|(S+i)x\|=\|(S-i)x\| \qquad \text{for all} \quad x\in\mathscr{D}(S)\,,$$

and thus V is isometric. Thus V is unitary. Moreover, we set

$$U=(T+i)(T-i)^{-1} \quad \text{with domain} \quad \mathscr{D}(U)=\operatorname{Ran}(T-i)\,.$$

Since $T\subset S$ by assumption it follows that $U\subset V$.

By construction $V=U$ on $\mathscr{D}(U)$ yielding

$$V|_{\operatorname{Ran}(T-i)} : \operatorname{Ran}(T-i)\to\operatorname{Ran}(T+i) \quad \text{and}$$
$$V|_{\operatorname{Ran}(T-i)^\perp} : \operatorname{Ran}(T-i)^\perp\to\operatorname{Ran}(T+i)^\perp\,,$$

where the second statement holds since V is unitary and thus preserves the inner product. Therefore, V maps $\operatorname{Ker}(T^*-i)$ unitarily onto $\operatorname{Ker}(T^*+i)$, proving that the dimension of these two spaces has to be equal.

□

Example 2.25

We set $\mathbb{R}^+=(0,\infty)$. On the Hilbert space $L^2(\mathbb{R}^+)$ we consider the operator

$$T_3x(t)=i\frac{d}{dt}x(t) \quad \textit{with domain} \quad \mathscr{D}(T_3)=\mathcal{C}_0^\infty(\mathbb{R}^+)$$

(the infinitely differentiable functions with compact support on $\mathbb{R}^+$*). Then* T_3 *is symmetric (the boundary values in the integration by parts formula vanish).*

Similar to Example 2.19 it can be shown (Exercise 2.38) that $T_3^* y = i\frac{d}{dt}y$ *on the domain*

$$\mathscr{D}(T_3^*) = \Big\{ x \in \mathscr{H} \mid \tfrac{d}{dt}x \in \mathscr{H} \text{ and } \forall I \subset \mathbb{R}^+ \text{ compact } : x|_I \text{ is absolutely continuous} \Big\} . \quad (2.29)$$

Thus T_3^* *is not symmetric (and thus in particular not self-adjoint). We compute* $\operatorname{Ker}(T_3^* \pm i)$*:*
$(T_3^* \pm i)y = 0$ *holds if and only if* y *is a solution of the differential equation*

$$y' = \mp y \quad \text{i.e. if} \quad y(t) = a_\pm e^{\mp t} \quad \text{for some} \quad a_\pm \in \mathbb{C} .$$

But since the function $t \mapsto e^t$ *is not in* $L^2(\mathbb{R}^+)$ *while* $t \mapsto e^{-t}$ *is, we get*

$$\operatorname{Ker}(T_3^* - i) = \{0\} \qquad \text{and} \qquad \operatorname{Ker}(T_3^* + i) = \operatorname{span}\{e^{-t}\} \neq 0 .$$

Thus T_3 *is not essentially self-adjoint by Corollary 2.22 and has no self-adjoint extensions by Theorem 2.24.*

2.1.7 The Friedrichs Extension

We come to an important class of operators with self-adjoint extensions.

DEFINITION 2.26 (SEMIBOUNDED OPERATORS)
A densely defined operator A *on a Hilbert space* $\mathscr{H}$ *with domain* $\mathscr{D}(A)$ *is called* **semibounded from below** $: \Longleftrightarrow$

$$\exists C \in \mathbb{R} \; : \quad \langle x, Ax \rangle \geq C\|x\|^2 \qquad \text{for all} \quad x \in \mathscr{D}(A).$$

A *is called* **semibounded from above** *if* $-A$ *is semibounded from below.*

Remark 2.27

The condition given in Definition 2.26 implies that $\langle x, Ax\rangle \in \mathbb{R}$ *(since in* $\mathbb{C}$ *we have no order relation). Thus by Exercise 2.35, each semibounded operator is in particular symmetric.*

Example 2.28

Examples for semibounded operators are the Laplace operator (or free Hamiltonian) H_0 *on* $\mathscr{H} = L^2(\mathbb{R}^n)$ *given by*

$$H_0\phi(x) = -\Delta\phi(x) = -\sum_{k=1}^{n} \frac{\partial^2}{\partial_k^2}\phi(x) \quad \text{with domain} \quad \mathscr{D}(H_0) = \mathcal{C}_0^\infty(\mathbb{R}^n) \tag{2.30}$$

or more general a Schrödinger operator $H = H_0 + V$ *on the same domain, where* V *is a multiplication operator such that there exists a constant* $C \in \mathbb{R}$ *such that* $V(x) \geq C$ *for all* x.

In fact, for any $\phi \in \mathscr{D}(H)$*, integration by parts yields*

$$\langle \phi, H\phi\rangle = \sum_k \left\langle \frac{\partial}{\partial_k}\phi, \frac{\partial}{\partial_k}\phi \right\rangle + \langle \phi, V\phi\rangle = \sum_k \left\| \frac{\partial}{\partial_k}\phi \right\|^2 + \langle \phi, V\phi\rangle \geq C\|\phi\|^2 .$$

Theorem 2.29 (Friedrichs Extension)

Each densely defined semibounded operator A *on a Hilbert space* $\mathscr{H}$ *admits a self-adjoint extension* S *and* S *is semibounded with the same bound.*

To prove Theorem 2.29 we need the following Lemma:

Lemma 2.30

Let $\mathscr{H}$ *and* $\mathscr{K}$ *be Hilbert spaces and* $J \in \mathcal{L}(\mathscr{K}, \mathscr{H})$ *be injective with dense range.*

Then $JJ^* \in \mathcal{L}(\mathscr{H})$ *is an injective operator with dense range und the inverse operator* $S : \mathrm{Ran}(JJ^*) \longrightarrow \mathscr{H}$ *is self-adjoint.*

Here the adjoint operator J^* is determined by the equation

$$\langle x, Jy\rangle_{\mathscr{H}} = \langle J^* x, y\rangle_{\mathscr{K}} \qquad \text{for all} \quad x \in \mathscr{H},\, y \in \mathscr{K}\,. \tag{2.31}$$

Proof. Since $\operatorname{Ker} J^* = (\operatorname{Ran} J)^{\perp}$, the assumption that the range of J is dense implies that J^* is injective. Since J is injective this yields that JJ^* is injective. Thus the self-adjoint operator JJ^* has dense range and S is therefore densely defined. Since S is symmetric, we can use Theorem 2.21 to show that S is self-adjoint. To see that $\operatorname{Ran}(S \pm i) = \mathscr{H}$, let $y \in \mathscr{H}$ be given. In order to find a pre-image of y, the equation

$$(S \pm i)x = y \quad \text{ or equivalently } \quad (\operatorname{Id} \pm iJJ^*)x = JJ^* y \tag{2.32}$$

has to be solved in $\mathscr{D}(S)$. But since JJ^* is self-adjoint, $i \in \rho(JJ^*)$ and thus $(i\operatorname{Id} - JJ^*)$ is invertible, giving the solution

$$x = i(i\operatorname{Id} \mp JJ^*)^{-1} JJ^* y$$

in $\mathscr{H}$. But since $x = JJ^*(y \mp ix)$ by (2.32) it follows that $x \in \operatorname{Ran}(JJ^*) = \mathscr{D}(S)$.

□

Proof of Theorem 2.29. Assume without loss of generality that A is semi-bounded from below.

Moreover, we can consider $A + c\operatorname{Id}$ instead of A for some $c \in \mathbb{R}$, since both operators have the same domain and A admits a self-adjoint extension if and only if $A + c\operatorname{Id}$ does. Thus by choosing c appropriately we can assume without loss of generality that

$$\langle x, Ax\rangle \geq \|x\|^2 \qquad \text{for all} \quad x \in \mathscr{D}(A)\,. \tag{2.33}$$

Then the map $(x, y) \mapsto \langle x, Ay \rangle$ is a positive definite sesquilinear form on $\mathscr{D}(A)$. Since A is symmetric by Remark 2.27, the map is conjugate symmetric. We thus can define a new inner product on $\mathscr{D}(A)$ by

$$\mathscr{D}(A) \times \mathscr{D}(A) \ni (x, y) \mapsto \langle\langle x, y \rangle\rangle := \langle x, Ay \rangle . \tag{2.34}$$

Then $(\mathscr{D}(A), \langle\langle \cdot, \cdot \rangle\rangle)$ is a pre-Hilbert space, and its completion with respect to the induced norm $|||\cdot|||$ is a Hilbert space $\mathscr{K} \subset \mathscr{H}$ [2]. From (2.33) and the construction of the norm in $\mathscr{K}$ it follows that

$$|||x||| \geq \|x\| \quad \text{for all} \quad x \in \mathscr{K} .$$

Therefore, the map

$$J : (\mathscr{D}(A), \langle\langle \cdot, \cdot \rangle\rangle) \ni x \mapsto x \in \mathscr{H}$$

is a linear contraction and by the BLT-Theorem it can be extended to a contraction $J : \mathscr{K} \longrightarrow \mathscr{H}$. Moreover, for any $x \in \mathscr{K}$ there is a sequence (x_n) in $\mathscr{D}(A)$ such that $x_n \to x$ and thus by continuity of the inner product and J, (2.34) and since $Jx = x$ for $x \in \mathscr{D}(A)$

$$\langle\langle x, y \rangle\rangle = \lim_{n \to \infty} \langle\langle x_n, y \rangle\rangle = \lim_{n \to \infty} \langle Jx_n, Ay \rangle = \langle Jx, Ay \rangle \tag{2.35}$$

for all $y \in \mathscr{D}(A)$ and $x \in \mathscr{K}$. If $Jx = 0$, then it follows from (2.35) that x is orthogonal to all $y \in \mathscr{D}(A)$ and since $\mathscr{D}(A)$ is dense this implies $x = 0$. Therefore, J is injective.

[2] From (2.33) we get

$$|||x|||^2 = \langle\langle x, x \rangle\rangle = \langle x, Ax \rangle \geq \|x\|^2 \quad \text{for all} \quad x \in \mathscr{D}(A)$$

and therefore each Cauchy sequence (x_n) in $\mathscr{D}(A)$ with respect to $|||\cdot|||$ is a Cauchy sequence in $\mathscr{H}$ (i.e. with respect to $\|\cdot\|$). Thus $x \in \mathscr{H}$ is an element in $\mathscr{K}$ if there is a $|||\cdot|||$-Cauchy sequence with $x_n \to x$ in $\mathscr{H}$. This gives a norm on $\mathscr{K}$ as limit $|||x||| = \lim |||x_n|||$.

Since $\mathscr{D}(A) \subset \operatorname{Ran} J$, the map $J \in \mathcal{L}(\mathscr{K}, \mathscr{H})$ is injective with dense range. Thus by Lemma 2.30 the operator $JJ^* \in \mathcal{L}(\mathscr{H})$ is injective with dense range and its inverse operator

$$S : \operatorname{Ran}(JJ^*) \longrightarrow \mathscr{H}$$

is self-adjoint.

We will show that S is an extension of A and satisfies the estimate

$$\langle x, Sx \rangle \geq \|x\|^2 \qquad \text{for all} \quad x \in \mathscr{D}(S) = \operatorname{Ran}(JJ^*)\,. \tag{2.36}$$

Let $x \in \mathscr{D}(A)$ and $y \in \mathscr{K}$, then by (2.35)

$$\langle\langle y, x \rangle\rangle = \langle Jy, Ax \rangle = \langle\langle y, J^*Ax \rangle\rangle \tag{2.37}$$

where we used (2.31) (the definition of the adjoint) for the second equality. Using $Jx = x$ for $x \in \mathscr{D}(A)$, equation (2.37) yields

$$J^*Ax = x = Jx = JJ^*Ax \quad \text{for all} \quad x \in \mathscr{D}(A)\,.$$

Therefore each $x \in \mathscr{D}(A)$ is in the range of JJ^* and thus in the domain of S, i.e. $\mathscr{D}(A) \subset \mathscr{D}(S)$.

Since S is the inverse of JJ^*,

$$JJ^*Ax = x = JJ^*Sx \quad \text{for all} \quad x \in \mathscr{D}(A)$$

and this implies $Ax = Sx$ for $x \in \mathscr{D}(A)$, because JJ^* is injective. Therefore $A \subset S$.

Thus S is a self-adjoint extension of A.

To see that (2.36) holds (i.e. S is a semibounded operator with the same constant as A), we use that J is a contraction, yielding for all $x = JJ^*z \in \mathscr{D}(S)$, using (2.31)

$$\langle x, Sx \rangle = \langle JJ^*z, z \rangle = \langle\langle J^*z, J^*z \rangle\rangle = |||J^*z|||^2 \geq \|JJ^*z\|^2 = \|x\|^2 .$$

□

REMARK 2.31

The self-adjoint extension constructed above is called **Friedrichs Extension**. *It could also be constructed starting directly from a semibounded sesquilinear form defined on a dense subspace of* $\mathscr{H}$.

COROLLARY 2.32

Let A *be a densely defined, closed operator on a Hilbert space* $\mathscr{H}$. *Then the operator*

$$A^*A \quad \text{with domain} \quad \mathscr{D}(A^*A) = \{x \in \mathscr{D}(A) \mid Ax \in \mathscr{D}(A^*)\}$$

is densely defined and self-adjoint.

Proof. It is clear that A^*A is symmetric on its domain. To see that $\mathscr{D}(A^*A)$ is dense in $\mathscr{H}$, we define on $\mathscr{D}(A)$ the inner product

$$\langle x, y \rangle_A := \langle x, y \rangle + \langle Ax, Ay \rangle . \tag{2.38}$$

Since A is closed and $\|x\|_A^2 := \langle x, x \rangle_A$ is the graph norm, $(\mathscr{D}(A), \langle \cdot, \cdot \rangle_A)$ is a Hilbert space, which we denote by $\mathscr{K}$. Let $J \in \mathcal{L}(\mathscr{K}, \mathscr{H})$ be the inclusion operator (i.e. $Jx = x$ for all $x \in \mathscr{K}$). Since $\mathscr{D}(A)$ is dense and $\|x\|_A \geq \|x\|$, J

is an injective contraction with dense range (its range is equal to the domain of A) and $\operatorname{Ran} JJ^*$ is dense by Lemma 2.30. It therefore suffices to show that

$$\operatorname{Ran}(JJ^*) \subset \mathscr{D}(A^*A) .$$

Let $x = JJ^*y \in \operatorname{Ran}(JJ^*)$, then x is in the range of J^*, because $JJ^*y = J^*y$ and therefore $x \in \mathscr{K} = \mathscr{D}(A)$. To see that $Ax \in \mathscr{D}(A^*)$, we have to show that the linear map $z \mapsto \langle Ax, Az\rangle$ is bounded for all $z \in \mathscr{D}(A)$. In fact, by (2.38) we have for all $z \in \mathscr{D}(A)$

$$\begin{aligned} \langle Ax, Az\rangle &= \langle x, z\rangle_A - \langle x, z\rangle = \langle J^*y, z\rangle_A - \langle x, z\rangle \\ &= \langle y, Jz\rangle - \langle x, z\rangle = \langle y - x, z\rangle \leq C\|z\|^2 \end{aligned}$$

for some $C > 0$. Thus $x \in \mathscr{D}(A^*A)$ and $\mathscr{D}(A^*A)$ is dense in $\mathscr{H}$.

The densely defined operator

$$T := \operatorname{Id} + A^*A \quad \text{obeys} \quad \langle x, Tx\rangle = \|x\|_A = \|x\|^2 + \|Ax\|^2 \geq \|x\|^2$$

for all $x \in \mathscr{D}(A^*A)$. Thus T is semibounded and its Friedrichs extension S can be constructed exactly as in the proof of Theorem 2.29, where J and $\mathscr{K}$ are as above. Thus S is defined on $\mathscr{D}(S) = \operatorname{Ran}(JJ^*) \subset \mathscr{D}(A^*A) = \mathscr{D}(T)$ and therefore $S = T$. This shows that T is and $A^*A = T - \operatorname{Id}$ are self-adjoint.

□

2.1.8 Exercises

Exercise 2.33 (Essential Self-adjointness)

Prove Proposition 2.14:

Let T, S be densely defined operators on a Hilbert space $\mathscr{H}$. Show that the following statements hold.

i) If $T \subset S$ then $S^* \subset T^*$.

ii) If T is essentially self-adjoint, then it has a unique self-adjoint extension.

iii) If T is self-adjoint, then T has no proper symmetric extension.

iv) T is essentially self-adjoint if and only if

$$T \subset T^{**} = T^* \,.$$

Exercise 2.34 (Density of function spaces)

Show that the space $\mathcal{C}_0^\infty(\mathbb{R}^n)$ of smooth compactly supported functions is dense in the space $\mathcal{C}_\infty(\mathbb{R}^n)$ of continuous functions vanishing at infinity.

Exercise 2.35 (Symmetric operators)

Let T be a densely defined linear operator on a complex HIlbert space $\mathscr{H}$ with domain $\mathscr{D}(T)$. Show that T is symmetric if and only if $\langle x, Tx\rangle \in \mathbb{R}$ for all $x \in \mathscr{D}(T)$.

Hint: Consider $\langle x+y, T(x+y)\rangle$.

Exercise 2.36 (Adjoint operators)

Let $\mathscr{H}$ be a Hilbert space. For densely defined linear operators S, T on $\mathscr{H}$ with domains $\mathscr{D}(S)$ and $\mathscr{D}(T)$, we set

$$\mathscr{D}(S \circ T) := \{x \in \mathscr{H} \mid x \in \mathscr{D}(T) \text{ and } Tx \in \mathscr{D}(S)\} \,.$$

Show that if $S \circ T$ is densely defined, then $T^* \circ S^* \subset (S \circ T)^*$ and if $S \in \mathcal{L}(\mathscr{H})$, then $T^* \circ S^* = (S \circ T)^*$.

Exercise 2.37 (Resolvent set)

Let T be a densely operator on a Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(T)$. Assume that there exists $\lambda \in \mathbb{C}$ such that $(\lambda\,\mathrm{Id} - T) : \mathscr{D}(T) \to \mathscr{H}$ is bijective and has a bounded inverse. Show that T is closed.

Rem.: This shows that the definition of resolvent set and spectrum for non-closed operators is useless (the spectrum would always be $\mathbb{C}$).

Exercise 2.38 (Adjoint operator)

On the Hilbert space in $\mathscr{H} = L^2(\mathbb{R}^+)$, consider the operator T_3 with domain $\mathscr{D}(T_3) = \mathcal{C}_0^\infty(\mathbb{R}^+)$, acting as $T_3\phi = i\frac{d}{dt}\phi$ for $\phi \in \mathscr{D}(T_3)$.

Show that the adjoint is given by $T_3^* y = i\frac{d}{dt}y$ on the domain

$$\mathscr{D}(T_3^*) = \{x \in \mathscr{H} \mid x|_I \text{ is absolutely continuous for all compact intervals } I \subset \mathbb{R}^+ \text{ and } \frac{d}{dt}x \in \mathscr{H}\}\,.$$

Exercise 2.39 (Multiplication operator)

Let (Ω, Σ, μ) be a measure space and $f : \Omega \to \mathbb{R}$ measurable. Define the operator M_f on $\mathscr{H} = L^2(\Omega, \Sigma, \mu)$ by setting

$$\mathscr{D}(M_f) = \{x \in \mathscr{H} \mid f \cdot x \in \mathscr{H}\} \qquad \text{and} \qquad M_f x = f \cdot x \quad \text{for } x \in \mathscr{D}(M_f)\,.$$

Show that

i) M_f *is densely defined*

Hint: Consider the sets $\Omega_n := \{\omega \,|\, |f(\omega)| \leq n\} \subset \Omega$

ii) M_f *is self-adjoint.*

iii) $\sigma(M_f) = \Big\{\lambda \in \mathbb{R} \,|\, \forall \epsilon > 0 \,:\, \mu\Big(f^{-1}[\lambda - \epsilon, \lambda + \epsilon]\Big) > 0\Big\}.$

Hint: To see " $\supset$ ", show that if $h : \Omega \to \mathbb{R}$ is measurable and the multiplication operator M_h is bounded, then $|h(t)| \leq \|M_h\|$ almost everywhere (consider for $a > 1$ a set $E \in \Sigma$ with $\mu(E) < \infty$ and $E \subset \{t \,|\, |h(t)| \geq a\|M_h\|\}$).

Exercise 2.40 (Spectrum of self-adjoint operator)
Let A be a densely defined, closed, symmetric operator on a Hilbert space $\mathscr{H}$. Show that A is self-adjoint if $\sigma(A) \subset \mathbb{R}$.

2.2 Spectral Theorem for unbounded operators

As in the case of self-adjoint bounded operators, there are several versions of the spectral theorem for self-adjoint unbounded operators.

2.2.1 Multiplication operator form

In this section, we will give an extension of Theorem 1.25 and Corollary 1.26.

From Exercise 2.39 we already know that if (Ω, Σ, μ) is a measure space and f a real-valued measurable function, then the operator M_f on $\mathscr{H} = L^2(\Omega, \Sigma, \mu)$ given by

$$\mathscr{D}(M_f) = \{x \in \mathscr{H} \,|\, f \cdot x \in \mathscr{H}\} \qquad \text{and} \qquad M_f x = f \cdot x \quad \text{for } x \in \mathscr{D}(M_f) \tag{2.39}$$

is self-adjoint. Moreover, the spectrum $\sigma(M_f)$ is equal to the essential range, where $\lambda \in \mathbb{R}$ is said to be in the **essential range** of f if and only of

$$\mu(\{m \in M \,|\, \lambda - \epsilon < f(m) < \lambda + \epsilon\}) > 0 \quad \text{for all} \quad \epsilon > 0\,.$$

Similar to the case of bounded operators, the following Lemma holds.

Lemma 2.41

Let T be a self-adjoint operator on a complex Hilbert space $\mathscr{H}$, then $\sigma(T) \subset \mathbb{R}$ and $\sigma(T) \neq \emptyset$.

Proof. Let $z = \lambda + i\eta \in \mathbb{C}$ with $\eta \neq 0$ and set $S = \frac{1}{\eta}(T - \lambda \operatorname{Id})$ on $\mathscr{D}(S) = \mathscr{D}(T)$. Then S is self-adjoint and it follows from $z - T = z - \eta S - \lambda = \eta(i - S)$ together with equation (2.27), that for any $x \in \mathscr{D}(T)$

$$\|(z - T)x\|^2 = \eta^2 \|(i - S)x\|^2 \geq \eta^2 \|x\|^2 .$$

Therefore there exists a continuous[5] map

$$(z - T)^{-1} : \operatorname{Ran}(z - T) \longrightarrow \mathscr{D}(T) .$$

But since $\operatorname{Ran}(z - T) = \operatorname{Ran}(i - S) = \mathscr{H}$ by Theorem 2.21, it follows that $z \in \rho(T)$. This shows that $\sigma(T) \subset \mathbb{R}$.

To see $\sigma(T) \neq \emptyset$, we argue with contradiction: Assume $\rho(T) = \mathbb{C}$, then for any $\lambda \neq 0$, the operator $(\lambda^{-1} \operatorname{Id} - T)^{-1} : \mathscr{H} \to \mathscr{D}(T)$ exists and is bounded. Since

$$(T^{-1} - \lambda \operatorname{Id}) T \lambda^{-1} (\lambda^{-1} \operatorname{Id} - T)^{-1} = (\operatorname{Id} - \lambda T)(\operatorname{Id} - \lambda T)^{-1} = \operatorname{Id} ,$$

the operator $(T^{-1} - \lambda \operatorname{Id})$ is invertible with inverse $T\lambda^{-1}(\lambda^{-1} \operatorname{Id} - T)^{-1} : \mathscr{H} \to \mathscr{H}$. This shows that any $\lambda \neq 0$ is in the resolvent set of T^{-1} and thus the spectrum of T^{-1} is a subset of $\{0\}$. Therefore $T^{-1} = 0$, contradicting $T^{-1}T = \operatorname{Id}$. Thus $\sigma(T) \neq \emptyset$.

□

Before we state the theorem, we remark that Theorem 1.25 and Corollary 2.22 can be extended to normal bounded operators.

The idea is to write a normal bounded operator $T \in \mathcal{L}(\mathscr{H})$ for a separable Hilbert space $\mathscr{H}$ as

$$T = A + iB \quad \text{where} \quad A := \frac{T + T^*}{2} \quad \text{and} \quad B := \frac{T - T^*}{2}$$

are self-adjoint and commute.

One possible way to prove the assertion is described in Appendix A.1, written (in german) by Jan Möhring. It is based on a talk he gave at the end of this lecture.

Another possibe approach uses the families of spectral projections Π^A and Π^B of A and B respectively. It it follows from Theorem 1.11 that $\Pi^A_{\Omega_1}$ and $\Pi^B_{\Omega_2}$ commute for all $\Omega_1, \Omega_2 \in \mathscr{B}(\mathbb{R})$. If we define the map

$$\mathbb{R}^2 \supset \Omega_1 \times \Omega_2 \mapsto P(\Omega_1 \times \Omega_2) = \Pi^A_{\Omega_1} \Pi^B_{\Omega_2}$$

then $P(\Omega_1 \times \Omega_2)$ is an orthogonal projection.

Let $f = \sum_i a_i \chi_{C_i}$ be a simple function on the rectangles (i.e. $C_i = \Omega_{i,1} \times \Omega_{i,2} \in \mathbb{R}^2$ and $C_i \cap C_k = \emptyset$ for $i \neq k$), then we set

$$f(A, B) := \sum_i a_i P(C_i) = \sum_i a_i \Pi^A_{\Omega_{i,1}} \Pi^B_{\Omega_{i,2}} .$$

Using the BLT-Theorem, this allows to construct a continuous and a measurable functional calculus, i.e. a unique map $\tilde{\Phi}_T : \mathcal{B}(\sigma(T)) \longrightarrow \mathcal{L}(\mathscr{H})$ with the properties as in Theorem 1.11. The analog of the complex measure

$$\mathscr{B}(\mathbb{R}) \ni \Omega \mapsto \mu_{\phi,\psi}(\Omega) = \langle \phi, \Pi_\Omega \psi \rangle \in \mathbb{C}$$

constructed in the case of self-adjoint operators for fixed $\phi, \psi \in \mathscr{H}$ is then given by

$$\mathscr{B}(\mathbb{C}) \ni \Omega_1 + i\Omega_2 \mapsto \tilde{\mu}_{\phi,\psi}(\Omega_1 + i\Omega_2) = \langle \phi, P(\Omega_1 \times \Omega_2)\psi \rangle \in \mathbb{C}.$$

Then we have for any $f \in \mathcal{B}(\sigma(T))$

$$\langle \psi, \tilde{\Phi}_T(f)\phi \rangle = \int_{\mathbb{C}} f(\lambda + i\eta)\, d\langle \psi, P(\lambda, \eta)\phi \rangle \,. \tag{2.40}$$

The construction of the multiplication operator can then be done similar to the case of self-adjoint operators.

Thus we get the following extension of Theorem 1.25 and Corollary 1.26 from self-adjoint bounded operators to normal bounded operators.

Proposition 2.42

For each normal operator $T \in \mathcal{L}(\mathscr{H})$ there exists a unique bounded projection valued measure $P = \{P_\Omega \,|\, \Omega \in \mathscr{B}(\mathbb{C})\}$ on the Borel-σ-algebra $\mathscr{B}(\mathbb{C})$ such that

$$T = \int_{\sigma(T)} \lambda \, dP_\lambda \,.$$

There is a unique norm continuous algebraic $*$-homomorphism $\tilde{\Phi}_T : \mathcal{B}(\mathbb{C}) \to \mathcal{L}(\mathscr{H})$ such that $\tilde{\Phi}_T(1) = \mathrm{Id}$ and $\tilde{\Phi}_T(\mathrm{id}) = T$ (here $\mathrm{id}(t) = t$), given by

$$\tilde{\Phi}_T(f) = \int_{\sigma(T)} f(\lambda) dP_\lambda \,.$$

Moreover, if $\mathscr{H}$ is separable, T is unitary equivalent to a multiplication operator M_f on $L^2(\Omega, \Sigma, \mu)$ for some finite measure space (Ω, Σ, μ), where $f \in L^\infty(\Omega)$ is in general complex-valued.

We now come to the formulation of the spectral theorem for unbounded operators.

Theorem 2.43 (Spectral Theorem - Multiplication operator) *Let A be a self-adjoint operator on a separable Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(A)$. Then there exists a finite measure space (Ω, Σ, μ), a unitary operator $U : \mathscr{H} \longrightarrow L^2(\Omega, \Sigma, \mu)$ and real-valued measurable function f on Ω, which is finite almost everywhere so that*

i) $x \in \mathscr{D}(A)$ *if and only if* $f \cdot Ux \in L^2(\Omega, \Sigma, \mu)$.

ii) If $\phi \in U[\mathscr{D}(A)]$, *then*

$$UAU^{-1}\phi = f \cdot \phi \qquad \mu - a.e.$$

where $U[\mathscr{D}(A)] = \{\phi \in L^2(\Omega, \mu) \,|\, f \cdot \phi \in L^2(\Omega, \mu)\}$.

Proof. Since A is self-adjoint, $\pm i \in \rho(A)$ by Lemma 2.41 and thus the maps $(A \pm i\,\mathrm{Id})^{-1} : \mathscr{H} \to \mathscr{D}(A)$ exist and are bounded. Now let $x, y \in \mathscr{H}$. Since $\mathrm{Ran}(A \pm i) = \mathscr{H}$ by Theorem 2.21, it follows that there are $u, v \in \mathscr{D}(A)$ satisfying $(A + i\,\mathrm{Id})u = x$ and $(A - i\,\mathrm{Id})v = y$. Therefore, setting $R := (A + i\,\mathrm{Id})^{-1}$,

$$\langle y, Rx \rangle = \langle (A - i)v, u \rangle = \langle v, (A + i)u \rangle = \langle (A - i)^{-1} y, x \rangle\,,$$

proving that

$$R^* = [(A + i)^{-1}]^* = (A - i)^{-1}\,.$$

Since by Theorem 2.16 the operators $R = (A + i\,\mathrm{Id})^{-1}$ and $R^* = (A - i\,\mathrm{Id})^{-1}$ commute, it follows that $R \in \mathcal{L}(\mathscr{H})$ is normal. Thus by Proposition 2.42, there is a finite measure space (Ω, Σ, μ), a unitary operator $U : \mathscr{H} \to L^2(\Omega, \Sigma, \mu)$ and a measurable, bounded, complex-valued function g so that

$$URU^{-1}\phi = g \cdot \phi =: M_g \phi \quad \mu - \text{a.e.} \quad \text{for all} \quad \phi \in L^2(\Omega, \Sigma, \mu)\,. \tag{2.41}$$

Since R is injective, $\operatorname{Ker} R = \{0\}$. Thus $g \cdot \phi = 0$ implies that $\phi = 0$ a.e. by (2.41) and therefore $g \neq 0$ almost everywhere. Thus we can define the measurable function

$$f(\omega) = \frac{1}{g(\omega)} - i$$

and f is finite almost everywhere.

Proof of i): Let $x \in \mathscr{D}(A)$, then there is some $y \in \mathscr{H}$ such that $x = Ry$ and from (2.41) it follows that

$$Ux = URU^{-1}Uy = g \cdot Uy \qquad \mu - \text{a.e.}$$

g is bounded and therefore $f \cdot g = 1 - ig$ is bounded, showing

$$M_f Ux = f \cdot g \cdot Uy \in L^2(\Omega, \Sigma, \mu)\,.$$

On the other hand, if $M_f Ux \in L^2(\Omega, \mu)$, then there exists some $y \in \mathscr{H}$ such that $Uy = (f + i) \cdot Ux$. This implies

$$g \cdot Uy = g(f + i) \cdot Ux = Ux \qquad \mu - \text{a.e.} \tag{2.42}$$

and by (2.41) we get

$$x = U^{-1} M_g Uy = U^{-1} URU^{-1} Uy = Ry \in \mathscr{D}(A)\,.$$

Proof of ii): We first remark that the equality

$$U[\mathscr{D}(A)] = \{\phi \in L^2(\Omega, \mu) \,|\, f \cdot \phi \in L^2(\Omega, \mu)\} =: \mathscr{D}(M_f)$$

follows at once from i).

Let $U^{-1}\phi = x \in \mathscr{D}(A)$, then $Rx = (A + i)^{-1} x = y$ for some $y \in \mathscr{H}$. Therefore $y = (A + i)x$ and thus $Ax = y - ix$, which by (2.42) proves

$$UAU^{-1}\phi = UAx = Uy - iUx = \frac{1}{g} \cdot Ux - iUx = M_f Ux = M_f \phi \qquad \mu - \text{a.e.} \tag{2.43}$$

f real-valued: Since U is unitary, it follows from (2.43) that

$$\langle x, Ay \rangle_{\mathscr{H}} = \langle Ux, M_f Uy \rangle_{L^2}$$

and since A is symmetric, if follows that M_f has to be symmetric. Thus by Exercise 2.35

$$\mathbb{R} \ni \langle \phi, M_f \phi \rangle = \int_\Omega f(\omega) |\phi(\omega)|^2 \, d\mu(\omega) \qquad \text{for all} \quad \phi \in L^2(\Omega, \Sigma, \mu)$$

showing that f is real-valued almost everywhere. □

2.2.2 Functional Calculus form

The above possibility to represent any self-adjoint operator as a multiplication operator on an appropriate space provides us with a natural procedure to define functions of self-adjoint operators.

Let (Ω, Σ, μ) be a measure space, f a real-valued measurable function on Ω and let M_f denote the self-adjoint operator on $L^2(\Omega, \Sigma, \mu)$ defined in (2.39).

If $h \in \mathcal{B}(\mathbb{R})$ (i.e. h is bounded, Borel measurable and complex-valued), then the bounded operator

$$h(M_f) := \tilde{\Phi}_{M_f}(h) := M_{h \circ f} \in \mathcal{L}(L^2(\Omega, \Sigma, \mu)) \tag{2.44}$$

is normal, since $M^*_{h \circ f} = M_{\overline{h \circ f}}$ and all multiplication operators commute.

For an arbitrary self-adjoint operator A on a Hilbert space $\mathscr{H}$, we then set

$$h(A) := \tilde{\Phi}_A(h) := U^{-1} M_{h \circ f} U \tag{2.45}$$

where U is a unitary transform and f is a real-valued measurable function associated to A as described in Theorem 2.43.

We then have

THEOREM 2.44 (SPECTRAL THEOREM - FUNCTIONAL CALCULUS FORM) *Let A be a self-adjoint operator on a separable Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(A)$. Then there exists a unique map $\tilde{\Phi}_A : \mathcal{B}(\mathbb{R}) \to \mathcal{L}(\mathscr{H})$ such that*

i) *$\tilde{\Phi}_A$ is an algebraic $*$-homomorphism.*

ii) *$\tilde{\Phi}_A$ is norm continuous, i.e. $\|\tilde{\Phi}_A(h)\|_{\mathcal{L}} \leq \|h\|_\infty$.*

iii) *Let $(h_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{B}(\mathbb{R})$ converging pointwise to $\mathrm{id}(t) = t$ and satisfying $|h_n(t)| \leq |t|$ for all $t \in \mathbb{R}$ and $n \in \mathbb{N}$, then $\tilde{\Phi}_A(h_n)$ converges strongly to A, i.e. $\lim_{n\to\infty} \tilde{\Phi}_A(h_n)x = Ax$ for all $x \in \mathscr{H}$.*

iv) *Let $(h_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{B}(\mathbb{R})$ converging pointwise to h and if $\|h_n\|_\infty$ is a bounded sequence, then $\tilde{\Phi}_A(h_n) \to \tilde{\Phi}_A(h)$ strongly.*

In addition:

v) *If $Ax = \lambda x$ then $\tilde{\Phi}_A(h)x = h(\lambda)x$.*

vi) *If $h \geq 0$, then $\tilde{\Phi}_A(h) \geq 0$.*

Proof. Exercise 2.48 □

The functional calculus allows us to define the exponential e^{itA}. The direct definition for bounded operators using the power series is not applicable for unbounded A.

2.2.3 Projection valued measure form

The functional calculus introduced above can be used to define projection valued measures and thus to give a spectral decomposition of a self-adjoint operator.

As first step, we introduce the family of spectral projections as in the bounded case. As above we denote by $\mathscr{B}(\mathbb{R})$ be the Borel-σ-algebra on $\mathbb{R}$.

Definition and Lemma 2.45 (Spectral Projection)
Let A be a self-adjoint operator on a Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(A)$. Then for any $M \in \mathscr{B}(\mathbb{R})$ we set

$$\Pi_M^A := \chi_M(A)(= \tilde{\Phi}_A(\chi_M))\,.$$

(here χ_M denotes the characteristic function of M). Then

i) Π_M^A *is an orthogonal projection.*

ii) $\Pi_\emptyset^A = 0$ *and* $\Pi_\mathbb{R}^A = \mathrm{Id}$.

iii) If $M = \bigcup_{n=1}^\infty M_n$ *and* $M_n \cap M_m = \emptyset$ *for all* $n \neq m$*, then*

$$\sum_{n=1}^{N} \Pi_{M_n}^A \longrightarrow \Pi_M^A \quad \textit{strongly as} \quad N \to \infty.$$

iv) $\Pi_{M_1}^A \Pi_{M_2}^A = \Pi_{M_1 \cap M_2}^A$.

We call $\Pi^A = \{\Pi_M^A \,|\, M \in \mathscr{B}(\mathbb{R})\}$ *the family of* **spectral projections** *of* A *or* **projection valued measure (p.v.m)**.

The proof is similar to the bounded case. The important difference is, that in the unbounded case, the projection valued measure does not have a compact support (i.e. is not bounded).

In the special case of a multiplication operator M_f on L^2 with a real-valued measurable function f as defined in (2.39), we get from (2.44)

$$\Pi_B^{M_f} = \chi_B(M_f) = M_{\chi_B \circ f} = M_{\chi_{f^{-1}(B)}} \quad \text{for all} \quad B \in \mathscr{B}(\mathbb{R})\,. \tag{2.46}$$

and if $M_f = UAU^{-1}$ on $\mathscr{D}(M_f)$ for some self-adjoint operator A on a Hilbert space $\mathscr{H}$, where $U : \mathscr{H} \to L^2$ is a unitary transform, then it follows from (2.45) that for any $B \in \mathscr{B}(\mathbb{R})$

$$\Pi_B^A = \chi_B(A) = U^{-1} M_{\chi_B \circ f} U = U^{-1} \Pi_B^{M_f} U\,. \tag{2.47}$$

Given the spectral projection Π^A of a self-adjoint operator A, we can define for any $x, y \in \mathscr{H}$ the complex Borel measure

$$\mathscr{B}(\mathbb{R}) \ni M \mapsto \langle x, \Pi_M^A y \rangle \in \mathbb{C}$$

and the positive measure

$$\mathscr{B}(\mathbb{R}) \ni M \mapsto \langle x, \Pi_M^A x \rangle \in [0, \infty)$$

as described in Section 1.1.3. Then similar to Section 1.1.4 for $h \in \mathcal{B}(\mathbb{R})$ we can define $h(A)$ by setting

$$\langle x, h(A) y \rangle = \int_{\mathbb{R}} h(\lambda)\, d\langle x, \Pi_\lambda^A y \rangle \tag{2.48}$$

and as in the case of bounded operators it can be shown that the map $h \mapsto h(A)$ has properties i)-iv) in Theorem 2.44. Therefore, by the uniqueness of $\tilde{\Phi}_A$, the operator $h(A)$ coincides with the operator constructed there.

Now suppose that $g : \mathbb{R} \to \mathbb{C}$ is Borel measurable, but not necessarily bounded. Let Π^A denote the spectral projection of a self-adjoint operator A

on a Hilbert space $\mathscr{H}$, then we set

$$\mathscr{D}_g^A := \{x \in \mathscr{H} \mid \int_{\mathbb{R}} |g(\lambda)|^2 \, d\langle x, \Pi_\lambda^A x\rangle < \infty\} . \tag{2.49}$$

We claim that $\mathscr{D}_g^A$ is dense in $\mathscr{H}$.

To see this, we use that by Theorem 2.43 A is unitary equivalent to a multiplication operator M_f on $L^2(\Omega, \Sigma, \mu)$ for some real-valued measurable function f.

Let U be the unitary transform $U : \mathscr{H} \to L^2(\Omega, \Sigma, \mu)$ such that

$$UAU^{-1}\varphi = M_f\varphi \quad \text{for all} \quad \varphi \in U[\mathscr{D}(A)] .$$

Then writing $\phi = Ux$ and $\psi = Uy$ for $x, y \in \mathscr{H}$ we have for any $B \in \mathscr{B}(\mathbb{R})$ by (2.46) and (2.47)

$$\begin{aligned} \langle x, \Pi_B^A y\rangle_{\mathscr{H}} &= \langle Ux, U\Pi_B^A U^{-1} Uy\rangle_{L^2} = \langle \phi, \Pi_B^{M_f} \psi\rangle_{L^2} \\ &= \int_\Omega \bar{\phi}(\omega)\chi_{f^{-1}(B)}(\omega)\psi(\omega)\, d\mu(\omega) = \int_{f^{-1}(M)} \bar{\phi}\psi \, d\mu . \end{aligned} \tag{2.50}$$

We define the complex measure on (Ω, Σ)

$$\Sigma \ni M \quad \mapsto \quad \nu_{\phi,\psi}(M) := \int_M \bar{\phi}\psi \, d\mu .$$

Then for any integrable function k on Ω

$$\int_\Omega k \, d\nu_{\phi,\psi} = \int_\Omega k(\omega)\bar{\phi}\psi(\omega)\, d\mu(\omega) \tag{2.51}$$

and by (2.50) for any $B \in \mathscr{B}(\mathbb{R})$

$$\mu_{x,y}(B) := \langle x, \Pi_B^A y\rangle = \nu_{\phi,\psi}(f^{-1}(B)) = f_*\nu_{\phi,\psi}(B) .$$

Thus $\mu_{x,y}$ is the pushforward measure[3] of $\nu_{\phi,\psi}$ under $f : \Omega \to \mathbb{R}$.

For any $\mu_{x,y}$-integrable function $h : \mathbb{R} \to \mathbb{C}$, the Transformation Formula[3] together with (2.51) therefore yields

$$\int_{\mathbb{R}} g(\lambda)\, d\mu_{x,y}(\lambda) = \int_{\Omega} g \circ f(\omega)\, d\nu_{\phi,\psi}(\omega) = \int_{\Omega} (g \circ f) \cdot \bar{\phi}\psi(\omega)\, d\mu(\omega)\,. \quad (2.53)$$

By (2.53) we get

$$x \in \mathscr{D}_g^A \qquad \Longleftrightarrow \qquad Ux \in \mathscr{D} := \{\phi \in L^2(\mu) \mid \int_{\Omega} |g \circ f|^2 |\phi|^2\, d\mu < \infty\}\,. \quad (2.54)$$

But in Exercise 2.39 we have shown that $\mathscr{D} = \mathscr{D}(M_{g \circ f})$ is dense in $L^2(\mu)$. Since U is unitary, this implies that $\mathscr{D}_g^A \subset \mathscr{H}$ is dense.

For $x = U\phi \in \mathscr{H}$ and $y = U\psi \in \mathscr{D}_g^A$, equation (2.53) together with the Hölder inequality yields

$$\begin{aligned}\int_{\mathbb{R}} |g(\lambda)|\, d\langle x, \Pi_\lambda^A y\rangle &= \int_{\Omega} |(g \circ f)\bar{\phi}\psi|\, d\mu(\omega) \leq \left(\int_{\Omega} |g \circ f|^2 |\psi|^2\, d\mu(\omega)\right)^{\frac{1}{2}} \|\phi\| \\ &= \left(\int_{\mathbb{R}} |g(\lambda)|^2\, d\langle y, \Pi_\lambda^A y\rangle\right)^{\frac{1}{2}} \|x\| < \infty\,.\end{aligned}$$

---

[3] Let (Ω, Σ, ν) be a measure space, (Ω', Σ') measurable space and $f : \Omega \to \Omega'$ be $\Sigma - \Sigma'$-measurable. Then

$$f_*\mu \,:\, \Sigma' \longrightarrow [0, \infty], \quad B \mapsto f_*\mu(B) = \mu(f^{-1}(B))$$

is a measure on (Ω', Σ'), called **pushforward measure** of μ under the map f (Bildmaß).

Proposition 2.46 (Transformation formula)
In the above setting let $h : \Omega' \to \mathbb{R}$ be Borel measurable function. Then h is $f_*\mu$-integrable if and only if $h \circ f$ is μ-integrable and in this case

$$\int_{f^{-1}(M)} h \circ f\, d\mu = \int_M h\, df_*\mu \qquad \text{for any} \quad M \in \Sigma'\,. \quad (2.52)$$

The last estimate follows from (2.54).

Thus for any $y \in \mathscr{D}_g^A$ the map $\mathscr{H} \ni x \mapsto \int_{\mathbb{R}} g(\lambda)\langle x, \Pi_\lambda^A y\rangle$ is a bounded linear form and by the Riesz-Lemma there exists some element in $\mathscr{H}$, which we call $g(A)y \in \mathscr{H}$, such that

$$\langle x, g(A)y\rangle = \int_{\mathbb{R}} g(\lambda)\langle x, \Pi_\lambda^A y\rangle \qquad \text{for all} \quad x \in \mathscr{H}\,. \tag{2.55}$$

This allows to define the operator $g(A) : \mathscr{D}_g^A \longrightarrow \mathscr{H}$, which we symbolically write as $g(A) = \int g(\lambda) d\Pi_\lambda^A$.

Unlike the case of bounded functions as given in (2.48), this integral does not exist in the sense of Theorem 1.17), but only as described in (2.55).

If we consider the special case $g(t) = \mathrm{id}(t) = t$, we get from (2.54) for any $y = U^{-1}\psi \in \mathscr{D}(A)$ and $x = U^{-1}\phi \in \mathscr{H}$

$$\int_{\mathbb{R}} \lambda\langle x, \Pi_\lambda^A y\rangle = \int_\Omega f\bar{\phi}\psi\, d\mu = \langle \phi, M_f\psi\rangle_{L^2} = \langle U^* M_f U x, y\rangle_{\mathscr{H}} = \langle Ax, y\rangle_{\mathscr{H}}\,. \tag{2.56}$$

Thus $\mathscr{D}(A) = \mathscr{D}_{\mathrm{id}}^A$ and $A = \int \lambda\, d\Pi_\lambda^A$. Similarly, (2.53) and (2.55) yield for $y \in \mathscr{D}_g^A$ and $x \in \mathscr{H}$

$$\begin{aligned}\langle x, g(A)y\rangle &= \int_{\mathbb{R}} g(\lambda)\langle x, \Pi_\lambda^A y\rangle = \int_\Omega (g \circ f)\cdot \bar{\phi}\psi(\omega)\, d\mu(\omega)\\ &= \langle \phi, M_{g\circ f}\psi\rangle\,,\end{aligned}$$

thus the definition of $g(A)$ for an unbounded function g is consistent with the definition (2.45) in the case of bounded functions.

The results given above can be summarized to the following extension of Theorem 1.18 to the case of unbounded operators and unbounded measurable functions.

THEOREM 2.47 (SPECTRAL THEOREM - P.V.M FORM (SPECTRAL DECOMPOSITION))
There is a one-to-one correspondence between self-adjoint operators A and projection-valued measures $\{\Pi_B \,|\, B \in \mathscr{B}(\mathbb{R})\}$ on a separable Hilbert space $\mathscr{H}$, given by

$$\langle x, Ay \rangle = \int_{\mathbb{R}} \lambda \, d\langle x, \Pi_\lambda y \rangle \,, \qquad \textit{for all} \quad x \in \mathscr{H}, y \in \mathscr{D}(A) \,.$$

If $g : \mathbb{R} \to \mathbb{R}$ is Borel-measurable and $\mathscr{D}_g^A := \{x \in \mathscr{H} \mid \int_{\mathbb{R}} |g(\lambda)|^2 \, d\langle x, \Pi_\lambda^A x \rangle < \infty\}$, then there is a self-adjoint operator $g(A)$ with domain $\mathscr{D}(g(A)) = \mathscr{D}_g^A$ defined by

$$\langle x, g(A)y \rangle = \int_{\mathbb{R}} g(\lambda) \, d\langle x, \Pi_\lambda y \rangle \,, \qquad \textit{for all} \quad x \in \mathscr{H}, y \in \mathscr{D}(g(A)) \,.$$

If g is bounded, $g(A)$ coincides with $\tilde{\Phi}_A(g)$ given in Theorem 2.44.

2.2.4 Exercises

EXERCISE 2.48 (FUNCTIONAL CALCULUS)
Let A be a self-adjoint operator A on a Hilbert space $\mathscr{H}$ with domain $\mathscr{D}(A)$. Let $U : \mathscr{H} \to L^2(\Omega, \Sigma, \mu)$ and f measurable real-valued as in Theorem 2.43. We set

$$\tilde{\Phi}_A \,:\, \mathcal{B}(\mathbb{R}) \longrightarrow \mathcal{L}(\mathscr{H}) \,, \qquad \tilde{\Phi}_A(h) = U^{-1} M_{h \circ f} U \,. \tag{2.57}$$

Show that

i) $\tilde{\Phi}_A$ *is an algebraic $*$-homomorphism with* $\tilde{\Phi}_A(1) = \mathrm{Id}$

ii) $\|\tilde{\Phi}_A\| \leq \|h\|$

iii) *if* $(h_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{B}(\mathbb{R})$ *converging pointwise to* $\mathrm{id}(t) = t$ *and satisfying* $|h_n(t)| \leq |t|$ *for all* $t \in \mathbb{R}$ *and* $n \in \mathbb{N}$, *then* $\tilde{\Phi}_A(h_n)$ *converges strongly to* A.

iv) *let* $(h_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{B}(\mathbb{R})$ *converging pointwise to* h *and if* $\|h_n\|_\infty$ *is a bounded sequence, then* $\tilde{\Phi}_A(h_n) \to \tilde{\Phi}_A(h)$ *strongly.*

v) *if* $Ax = \lambda x$ *then* $\tilde{\Phi}_A(h)x = h(\lambda)x$.

vi) *if* $h \geq 0$, *then* $\tilde{\Phi}_A(h) \geq 0$.

Show that the function $\tilde{\Phi}_A$ *given in* (2.57) *is uniquely determined by properties i)-iv).*

2.3 Semigroups of operators

In this section, we will introduce the notion of semigroups of operators. An important application are differential equations of the form

$$x'(t) = Ax(t), \qquad x(0) = x_0, \tag{2.58}$$

where A is a linear operator. In the finite dimensional case, i.e. if A is a $(n \times n)$-matrix and x is a curve in $\mathbb{R}^n$, (2.58) is a system of ordinary linear differential equations, which is solved by $x(t) = e^{tA}x_0$. If A is an unbounded operator on a Hilbert or Banach space, (2.58) can describe a partial differential equation. Thus we are interested to define the exponential e^{tA} and e^{itA} for classes of operators. From Section 2.2 and in particular Theorem 2.47 it follows that we can define e^{tA} if A is self-adjoint, $\sigma(A)$ has an upper bound and $t \geq 0$.

2.3.1 Definition and Properties of Semigroups

Definition 2.49 (Strongly continuous Semigroup)
Let X be a complex Banach space and $T_t : X \longrightarrow X$, $t \geq 0$, a family of bounded linear operators.
$\{T_t\}_{t\geq 0}$ *is called* **strongly continuous one-parameter semigroup** $:\Longleftrightarrow$

a) $T_0 = \mathrm{Id}$.

b) $T_{s+t} = T_s T_t$ *for all* $s, t \geq 0$.

c) $\lim_{t\to 0} T_t x = x$ *for all* $x \in X$

If instead of c) the family has the stronger property

c)' $\lim_{t\to 0} \|T_t - \mathrm{Id}\,\| = 0$

then it is called **norm-continuous** *semigroup.*

Example 2.50
(a) If $A \in \mathcal{L}(X)$*, then the* **exponential** $\{T_t = e^{tA}\}_{t\geq 0}$ *is a norm-continuous one-parameter semigroup (for* $t \in \mathbb{R}$*, this would even be a group).*
(b) The **translation semigroup** $T_t f(x) = f(x+t)$ *builds a strongly continuous semigroup on* $\mathcal{C}_\infty(\mathbb{R})$ *(the space of continuous functions vanishing at* ∞*). Here a) and b) are obvious. For c) we remark that* $f \in \mathcal{C}_\infty$ *is uniformly continuous. Let* $\epsilon > 0$*, then there exists some* $\delta > 0$ *such that* $|f(x) - f(y)| < \epsilon$ *whenever* $|x - y| < \delta$*. This implies*

$$0 < t < \delta \quad \Longrightarrow \quad \|T_t f - f\|_\infty = \sup_{x\in\mathbb{R}} |f(x+t) - f(x)| < \epsilon.$$

Since $\mathcal{C}_\infty$ is dense in L^p for $1 \le p < \infty$ and $\|T_t\| < \infty$ it can be shown by an $\epsilon/3$-argument that T_t is a strongly continuous semigroup on $L^p(\mathbb{R})$. The family $\{T_t\}_{t\in\mathbb{R}}$ is even a group of operators.

As above it can be shown that $\{T_t\}_{t\ge 0}$ is a strongly continuous semigroup on $\mathcal{C}_\infty([0,\infty))$ and $L^p([0,\infty))$.

(c) The **heat diffusion semigroup** *on $L^p(\mathbb{R}^n)$ ($1 \le p < \infty$) is given by $T_0 = \mathrm{Id}$ and*

$$\big(T_t f\big)(x) := \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} f(y)\, dy \tag{2.59}$$

for $t > 0$. If we define the **heat kernel**

$$\gamma_t(x) := \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}\,, \qquad \text{for} \quad x \in \mathbb{R}^n,\, t > 0, \tag{2.60}$$

then we can define T_t as convolution

$$T_t f = \gamma_t * f\,.$$

Since $\gamma_t \in L^1(\mathbb{R}^n)$, with $\|\gamma_t\|_1 = 1$, if follows from Young's inequality (Exercise 2.84) that

$$\|T_t f\|_p = \|\gamma_t * f\|_p \le \|\gamma_t\|_1 \|f\|_p = \|f\|_p\,,$$

thus $T_t \in \mathcal{L}(L^p(\mathbb{R}^n))$ is a contraction. Property c) can be shown with methods similar to those in Example 2.6. In order to show b), it suffices to prove

$$\gamma_{t+s} = \gamma_t * \gamma_s\,,$$

because the convolution is associative. (Exercise 2.84).

We now state two elementary properties of strongly continuous semigroups.

Lemma 2.51

Let $\{T_t\}_{t\geq 0}$ be a strongly continuous one-parameter semigroup on a Banach space X. Then the following holds.

i) There exists constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that

$$\|T_t\| \leq M e^{\omega t} \qquad \text{for all} \quad t \geq 0 \,. \tag{2.61}$$

ii) The map

$$[0,\infty) \times X \ni (t,x) \quad \mapsto \quad T_t(x) \in X$$

is continuous.

With respect to t it is uniformly continuous on compact subsets of $[0,\infty)$.

Proof. i): First we show that there exists some $r > 0$ such that

$$K := \sup_{0\leq t\leq r} \|T_t\| < \infty \,. \tag{2.62}$$

We show this by contradiction: Assume that (2.62) does not hold. Then there exists some null sequence $(t_n)_{n\in\mathbb{N}}$ such that $\|T_{t_n}\| \to \infty$.

From Exercise 2.82 (or the Principle of Uniform Boundedness, Theorem 3.31) we can conclude that there exists some $x \in X$ so that $\|T_{t_n}x\| \to \infty$. But this is false by property c).

Now let K and r be such that (2.62) holds. We write any $t \geq 0$ as

$$t = nr + s \qquad \text{where} \quad n \in \mathbb{N} \quad \text{and} \quad 0 \leq s < r \,.$$

Using that $n \leq \frac{t}{r}$ and $K \geq \|T_0\| = 1$, it follows from property b) that

$$\|T_t\| = \|T_r^n T_s\| \leq \|T_r\|^n \|T_s\| \leq K^{n+1} \leq K(K^{\frac{1}{r}})^t \,.$$

Thus (2.61) holds for

$$M = K \qquad \text{and} \qquad \omega = \frac{\log K}{r}\,.$$

ii): Let $\epsilon > 0$ and fix $x \in X$ and $t_0 > 0$. Then we have to show that there are $\delta > 0$ and $h_0 > 0$ such that

$$\|x-y\| \le \delta \quad \wedge \quad 0 \le s \le t \le t_0 \quad \wedge \quad t-s \le h_0 \quad \Longrightarrow \quad \|T_t x - T_s y\| \le \epsilon\,. \tag{2.63}$$

We first remark that by property c) for $\delta > 0$ given we can choose $h_0 > 0$ such that

$$\|T_h x - x\| \le \delta \qquad \text{for all} \quad 0 \le h \le h_0\,. \tag{2.64}$$

Now let M, ω be such that (2.61) holds. Then, using property b)

$$\begin{aligned} \|T_t x - T_s y\| &\le \|T_t x - T_s x\| + \|T_s x - T_s y\| \\ &\le \|T_s\| \|T_{t-s} x - x\| + \|T_s\| \|x - y\| \\ &\le M e^{\omega s}\delta + M e^{\omega t}\delta, \end{aligned}$$

where for the last step we used (2.64).

Thus (2.63) holds, if we choose $\delta \le \frac{\epsilon}{2M}$ if $\omega \le 0$ and $\delta \le \frac{\epsilon}{2Me^{\omega t_0}}$ if $\omega > 0$.

□

2.3.2 Infinitesimal Generators

We can associate to each strongly continuous semigroup an operator, the so called generator. In example 2.50(a) this would be the operator A. To regain A from $T_t = e^{tA}$, we have to differentiate.

Definition 2.52 (Generator of a Semigroup)

Let $\{T_t\}_{t\geq 0}$ *be a strongly continuous one-parameter semigroup on a Banach space* X.

The **(infinitesimal) generator** *of the semigroup is the operator*

$$Ax := \lim_{h\searrow 0} \frac{T_h x - x}{h} \tag{2.65}$$

with domain

$$\mathscr{D}(A) = \{x \in X \mid \lim_{h\searrow 0} \frac{T_h x - x}{h} \quad \textit{exists}\ \}.$$

Thus A is the right derivative of the vector-valued function

$$T_{(\cdot)}x : [0,\infty) \to X \quad \text{with} \quad t \mapsto T_t x$$

at the point $t = 0$.

For the semigroups given in Example 2.50, we get the following.

Example 2.53

(a) As already mentioned above, for $T_t = e^{tA}$ *with* A *bounded, the generator is* A *itself. In fact, using* $e^{hA} = \sum_{n=0}^{\infty} \frac{1}{n!}(hA)^n$, *we get*

$$\left\| \frac{e^{hA} - \mathrm{Id}}{h} - A \right\| = \left\| \sum_{n=2}^{\infty} \frac{1}{n!} h^{n-1} A^n \right\| \leq \sum_{n=2}^{\infty} \frac{1}{n!} h^{n-1} \|A\|^n \leq hA^2 e^{h\|A\|} \to 0$$

as $h \searrow 0$. *In this case, the domain of the generator is* X.

(b) In the case of the translation semigroup on $\mathcal{C}_\infty(\mathbb{R})$, *we have pointwise*

$$\lim_{h\searrow 0} \frac{T_h f(x) - f(x)}{h} = \lim_{h\searrow 0} \frac{f(x+h) - f(x)}{h} =: f'_+(x)\,,$$

where f'_+ *denotes the right derivative of* f. *Thus we would guess that the generator* A *is given by the right derivative and* $f \in \mathcal{C}_\infty$ *is in the domain of*

A if $f'_+(x)$ exists for all $x \in \mathbb{R}$ and $f'_+ \in \mathcal{C}_\infty(\mathbb{R})$. But then f'_+ is uniformly continuous and this implies that for any $\epsilon > 0$ there exists $h_0 > 0$ such that for all $0 \leq h \leq h_0$ with $y = x - h$)

$$\left|\frac{f(x-h)-f(x)}{(-h)} - f'_+(x)\right| \leq \left|\frac{f(y)-f(y+h)}{(-h)} - f'_+(y)\right| + |f'_+(y) - f'_+(x)|$$
$$= \left|\frac{f(y+h)-f(y)}{h} - f'_+(y)\right| + |f'_+(y) - f'_+(x)| \leq \epsilon\,.$$

This shows that the left derivative f'_- exists and is equal to the right derivative and therefore f is differentiable. Thus

$$Af = f' \quad \text{and} \quad \mathscr{D}(A) = \{f \in \mathcal{C}_\infty(\mathbb{R}) \,|\, f' \text{ exists and } f' \in \mathcal{C}_\infty(\mathbb{R})\}\,.$$

(c) To derive the generator of the heat diffusion semigroup, we need the theory of Fourier transformation. Thus we discuss this example later.

In the following we will discuss some of the properties of the generator of a semigroup. One goal is to show that it is densely defined and closed.

First remark that the notion of a Riemannian integral can be extended to continuous functions on $\mathbb{R}$ with values in a Banach space X (taking the limit of Riemannian sums). Then the usual computation rules hold (as linearity) and the Fundamental Theorem

$$\lim_{h\to 0} \frac{1}{h} \int_t^{t+h} u(s)\,ds = u(t)\,. \tag{2.66}$$

holds for $u \in \mathcal{C}(\mathbb{R}, X)$. Moreover, if $T \in \mathcal{L}(X)$, then

$$T\left(\int_a^b u(s)\,ds\right) = \int_a^b T\big(u(s)\big)\,ds\,. \tag{2.67}$$

Lemma 2.54

Let A be the generator of a strongly continuous semigroup $\{T_t\}_{t\geq 0}$ on a Banach space X and let $t > 0$, then *i) for all* $x \in X$

$$\int_0^t T_s x\, ds \in \mathscr{D}(A) \quad \textit{and} \quad A\left(\int_0^t T_s x\, ds\right) = T_t x - x\,.$$

ii) $T_t(\mathscr{D}(A)) \subset \mathscr{D}(A)$.

iii) $T_t A x = A T_t x$ *for all* $x \in \mathscr{D}(A)$.

iv) $T_t x - x = \int_0^t T_s A x\, ds$ *for all* $x \in \mathscr{D}(A)$.

Proof. Remark that the map $t \mapsto T_t x$ is continuous for each $x \in X$ (Lemma 2.51).

i): Using (2.67) and the semigroup properties, we can write

$$\begin{aligned} \frac{1}{h}\left(T_h\Big(\int_0^t T_s x\, ds\Big) - \int_0^t T_s x\, ds\right) &= \frac{1}{h}\left(\int_0^t T_{s+h} x\, ds - \int_0^t T_s x\, ds\right) \qquad (2.68)\\ &= \frac{1}{h}\left(\int_h^{h+t} T_s x\, ds - \int_0^t T_s x\, ds\right)\,. \end{aligned}$$

Writing $\int_h^{h+t} - \int_0^t = \int_h^t + \int_t^{h+t} - \int_0^h - \int_h^t = \int_t^{t+h} - \int_0^h$, we get by (2.65) and (2.66)

$$\begin{aligned} A\left(\int_0^t T_s x\, ds\right) &= \lim_{h\to 0} \mathrm{lhs}(2.68) = \lim_{h\to 0} \frac{1}{h}\left(\int_t^{t+h} T_s x\, ds - \int_0^h T_s x\, ds\right)\\ &= T_t x - x\,. \end{aligned}$$

ii) and iii): Let $x \in \mathscr{D}(A)$, then by the continuity of T_t

$$\frac{1}{h}\Big(T_h(T_t x) - T_t x\Big) = T_t \frac{T_h x - x}{h} \longrightarrow T_t A x \quad \text{as} \quad h \to 0\,.$$

Thus $T_t x \in \mathscr{D}(A)$ and $A T_t x = T_t A x$.

iv) Let $x \in \mathscr{D}(A)$, then by i)

$$T_t x - x = A\left(\int_0^t T_s x\, ds\right) = \lim_{h\to 0} \frac{1}{h}\left(T_h\Big(\int_0^t T_s x\, ds\Big) - \int_0^t T_s x\, ds\right)$$

$$= \lim_{h \to 0} \int_0^t T_s \Big(\frac{T_h x - x}{h} \Big) \, ds \,.$$

Since by Lemma 2.51 the integrand converges uniformly on $[0,t]$ to $T_s A x$ as $h \to 0$, iv) follows.

□

PROPOSITION 2.55

The generator of a strongly continuous one-parameter semigroup on a Banach space is densely defined and closed.

Proof. Let A be the generator of the strongly continuous semigroup $\{T_t\}_{t \geq 0}$ on a Banach space X. For $x \in X$ and $t > 0$ we set

$$x_t := \frac{1}{t} \int_0^t T_s x \, ds \,.$$

Then $x_t \in \mathscr{D}(A)$ by Lemma 2.54 and $\lim_{t \to 0} x_t = x$ by (2.66). Thus $\mathscr{D}(A)$ is dense in X.

Now consider a sequence $(x_n)_{n \in N}$ in $\mathscr{D}(A)$ such that $x_n \to x$ and $Ax_n \to y$.

Since T_t is continuous, we have by Lemma 2.54

$$\frac{T_h x - x}{h} = \lim_{n \to \infty} \frac{T_h x_n - x_n}{h} = \lim_{n \to \infty} \frac{1}{h} \int_0^h T_s A x_n \, ds \,. \tag{2.69}$$

From Lemma 2.51 it follows that the convergence $T_s A x_n \to T_s y$ is uniform in $[0,t]$, thus

$$\text{rhs}(2.69) = \frac{1}{h} \int_0^h T_s y \, ds \longrightarrow T_0 y = y \quad \text{as} \quad h \to 0 \,.$$

This proves that $x \in \mathscr{D}(A)$ and $Ax = y$. Thus A is closed.

□

2.3.3 Application to a Cauchy problem

Lemma 2.54 gives us information about the solutions of an abstract Cauchy problem

$$u'(t) = Au(t), \qquad u(0) = x_0 \,. \tag{2.70}$$

PROPOSITION 2.56

Let A be the generator of the strongly continuous semigroup $\{T_t\}_{t\geq 0}$ on a Banach space X and let $x_0 \in \mathscr{D}(A)$.

Then the map

$$u : [0,\infty) \longrightarrow X \,, \qquad u(t) = T_t x_0$$

is continuously differentiable, $\mathscr{D}(A)$-valued and solves (2.70)*.*

Moreover, u is the unique solution of (2.70) *with these properties and the map $x_0 \mapsto u(t)$ is continuous.*

Proof. From Lemma 2.54 it follows that $T_t x_0 \in \mathscr{D}(A)$ for all $t > 0$ and therefore $A(u(t))$ is well defined. So see that $u' = Au$, we compute the right and left derivative of u separately.

We use that $t \mapsto T_t x$ is continuous by Lemma 2.51 and $T_t A x = A T_t x$ by Lemma 2.54. Then the right derivative of u is given by

$$\begin{aligned} \lim_{h\searrow 0} \frac{u(t+h)-u(t)}{h} &= \lim_{h\searrow 0} \frac{T_{t+h}x_0 - T_t x_0}{h} = \lim_{h\searrow 0} T_t \left(\frac{T_h x_0 - x_0}{h} \right) \\ &= T_t A x_0 = A T_t x_0 = Au(t) \,. \end{aligned} \tag{2.71}$$

To determine the left derivative, we first notice that

$$\lim_{h\searrow 0} \frac{u(t-h)-u(t)}{(-h)} = \lim_{h\searrow 0} \frac{T_{t-h}x_0 - T_t x_0}{(-h)} = \lim_{h\searrow 0} T_{t-h} \left(\frac{T_h x_0 - x_0}{h} \right) \,. \tag{2.72}$$

We set $A_h := \frac{1}{h}(T_h - \mathrm{Id})$, then the right hand side of (2.72) can be written as $\lim T_{t-h}A_h x_0$. We have

$$\begin{aligned}\|T_{t-h}A_h x_0 - T_t A x_0\| &\le \|T_{t-h}A_h x_0 - T_{t-h}Ax_0\| + \|T_{t-h}Ax_0 - T_t A x_0\| \\ &\le \|T_{t-h}\|\|A_h x_0 - Ax_0\| + \|T_{t-h}Ax_0 - T_t A x_0\| \longrightarrow 0 \quad \text{as} \quad h \searrow 0, \quad (2.73)\end{aligned}$$

where for the limit we used $\sup_{s\le t}\|T_s\| < \infty$, the definition of A and the continuity of T_t. By (2.73) we get

$$\mathrm{rhs}(2.72) = T_t A x_0 = A T_t x_0 = Au(t)\,. \qquad (2.74)$$

Combining equations (2.72), (2.73) and (2.74) shows that $u' = Au$ or more explicitly

$$\frac{d}{ds}T_s x_0 = u'(s) = Au(s) = AT_s x_0\,. \qquad (2.75)$$

Since $u'(t) = AT_t x_0 = T_t A x_0$, the continuity of T_t implies that u' is continuous.

For the uniqueness, let v be another solution of (2.70). Then by the product rule and (2.75)

$$\frac{d}{dt}T_{s-t}v(t) = (-1)AT_{s-t}v(t) + T_{s-t}v'(t) = -T_{s-t}Av(t) + T_{s-t}Av(t) = 0\,.$$

Therefore the function

$$F : [0,s] \to X\,, \qquad F(t) = T_{s-t}v(t)$$

is constant, because for any functional $\ell \in X^*$

$$\frac{d}{dt}(\ell \circ F) = \ell \circ \frac{d}{dt}F = 0 \quad \text{and thus} \quad \ell(F(0)) = \ell(F(t)) \quad (t \in [0,s])\,.$$

The Hahn-Banach-Theorem (or Corollary 62 from the previous semester) thus shows $F(0) = F(s)$ and thus

$$u(s) = T_s x_0 = F(0) = F(s) = T_{s-s} v(s) = v(s)\,.$$

Since s was arbitrary, this shows the uniqueness of the solution.

The fact that $x_0 \mapsto u(t) = T_t x_0$ continuous follows from Lemma 2.51.

□

COROLLARY 2.57

Two strongly continuous one-parameter semigroups with the same infinitesimal generator are equal.

Proof. Let $(S_t)_{t\geq 0}$ and $(T_t)_{t\geq 0}$ be strongly continuous one-parameter semigroups with the (same) generator A, then the maps

$$t \mapsto S_t x \qquad \text{and} \qquad t \mapsto T_t x$$

both solve the initial value problem

$$u'(t) = Au(t)\,, \qquad u(0) = x \in \mathscr{D}(A)\,.$$

The uniqueness of the solution (Proposition 2.56) then implies $S_t|_{\mathscr{D}(A)} = T_t|_{\mathscr{D}(A)}$ for all $t \geq 0$. Since S_t and T_t are continuous by Lemma 2.51 and $\mathscr{D}(A)$ is dense by Proposition 2.55, this implies $T_t = S_t$.

□

In order to use Proposition 2.56 for a Cauchy problem with a differential operator A, we need criteria to decide whether A is the generator of a strongly continuous semigroup.

First, consider a norm-continuous semigroup $(T_t)_{t\geq 0}$. Since for any $0 \leq s \leq t \leq t_0$ we have

$$\|T_t - T_s\| = \|T_s(T_{t-s} - \mathrm{Id})\| \leq \|T_s\|\|T_{t-s} - \mathrm{Id}\,\| \leq C\|T_{t-s} - \mathrm{Id}\,\|\,,$$

it follows from property $c)'$ that $t \mapsto T_t$ is continuous with respect to operator norm. Therefore, the Riemannian integral $\int_0^t T_s\,ds$ converges in operator norm and we can define the operators

$$M_t = \frac{1}{t}\int_0^t T_s\,ds \qquad \text{for all} \quad t > 0\,. \tag{2.76}$$

Since, moreover, the map $T \mapsto Tx$ is a continuous linear operator from $\mathcal{L}(X)$ to X, it follows that

$$M_t x = \frac{1}{t}\int_0^t T_s x\,ds \qquad \text{for all} \quad x \in X \quad \text{and} \quad t > 0\,.$$

Proposition 2.58

Let A be the generator of a strongly continuous semigroup $\{T_t\}_{t\geq 0}$ on a Banach space X. Then the following three statements are equivalent:

i) *$\{T_t\}_{t\geq 0}$ is norm-continuous.*

ii) *A is continuous.*

iii) *$\mathscr{D}(A) = X$.*

If these assumptions hold, then $T_t = e^{tA}$ for all $t \geq 0$.

Proof. "$i) \Rightarrow iii)$" : If T_t is norm-continuous, then $t \mapsto T_t$ is continuous and thus $\|M_t - \mathrm{Id}\,\| \to 0$ by (2.66). Thus by Lemma 79 from the previous semester (see (1.2)) it follows that M_r is invertible and thus in particular

surjective for r sufficiently small. Since $\operatorname{Ran} M_r \subset \mathscr{D}(A)$ by Lemma 2.54 it follows that $\mathscr{D}(A) = X$.
"*iii)* $\Rightarrow$ *ii)*" : It follows from Proposition 2.55 that A is closed. Since by assumption the domain of A is a Banach space, it follows from the Closed Graph Theorem that A is continuous.
"*ii)* $\Rightarrow$ *i)*" : If A is bounded, the semigroup $S_t = e^{tA}$ is norm-continuous and has A as generator (see example 2.50 and 2.53). Thus $T_t = S_t$ by Corollary 2.57.

□

2.3.4 Theorem of Hille-Yosida for Contraction Semigroups

First remark that Definition 2.15 (Resolvent set and spectrum of unbounded operators on a Hilbert space) and Theorem 2.16 (Properties of the resolvent) can be extended to the case of unbounded operators on Banach spaces.

Definition 2.59 (Contraction Semigroup)
A strongly continuous one-parameter semigroup $(T_t)_{t\geq 0}$ on a Banach space X is called contraction semigroup $:\Longleftrightarrow$ $\|T_t\| \leq 1$ *for all* $t \geq 0$.

Proposition 2.60
Let A be the infinitesimal generator of the contraction semigroup $(T_t)_{t\geq 0}$ on a Banach space X, then

i) $\{\lambda = \eta + i\nu \in \mathbb{C} \,|\, \eta > 0\} \subset \rho(A)$.

ii) For all $\lambda = \eta + i\nu$ *with* $\eta > 0$

$$(\lambda - A)^{-1}x = \int_0^\infty e^{-\lambda s} T_s x \, ds \,. \tag{2.77}$$

iii) $\|\eta(\lambda - A)^{-1}\| \leq 1$ *for all* $\lambda = \eta + i\nu$ *with* $\eta > 0$.

Proof. Let $\lambda = \eta + i\nu \in \mathbb{C}$ with $\eta > 0$, then

$$\lim_{t\to\infty} \|e^{-\lambda t} T_t\| \leq \lim_{t\to\infty} |e^{-\lambda t}| = \lim_{t\to\infty} e^{-\eta t} = 0 \,.$$

Moreover, the semigroup $(e^{-\lambda t} T_t)_{t\geq 0}$ has the generator $A - \lambda$ with domain $\mathscr{D}(A)$. Thus by Lemma 2.54 (applied to this semigroup)

$$e^{-\lambda t} T_t x - x = \begin{cases} (A - \lambda) \int_0^t e^{-\lambda s} T_s x \, ds & \text{for all} \quad x \in X \\ \int_0^t e^{-\lambda s} T_s (A - \lambda) x \, ds & \text{for all} \quad x \in \mathscr{D}(A) \end{cases}$$

and in the limit $t \to \infty$ this shows

$$x = \begin{cases} (\lambda - A) \int_0^\infty e^{-\lambda s} T_s x \, ds & \text{for all} \quad x \in X \\ \int_0^\infty e^{-\lambda s} T_s (\lambda - A) x \, ds & \text{for all} \quad x \in \mathscr{D}(A) \end{cases} .$$

Therefore, $(\lambda - A) : \mathscr{D}(A) \to X$ is bijective and thus $\lambda \in \rho(A)$ and the inverse of $(\lambda - A)$ is given by $\int_0^\infty e^{-\lambda s} T_s(\cdot) \, ds$, proving i) and ii).

In order to see iii), we write

$$\|(\lambda - A)^{-1}x\| \leq \int_0^\infty e^{-\eta s} \|T_s\| \|x\| \, ds \leq \int_0^\infty e^{-\eta s} \, ds \cdot \|x\| = \frac{\|x\|}{\eta} \,.$$

□

The converse statement is given in the following theorem.

Theorem 2.61 (Hille-Yosida for Contraction Semigroups)
Let A be an operator on a Banach space X with domain $\mathscr{D}(A)$. Then A is the infinitesimal generator of a contraction semigroup if and only if

i) *A is densely defined and closed,*

ii) $(0,\infty) \subset \rho(A)$,

iii)

$$\|\lambda(\lambda - A)^{-1}\| \leq 1 \qquad \textit{for all} \quad \lambda > 0\,. \tag{2.78}$$

Proof. " $\Rightarrow$ " : This follows from Proposition 2.55 and Proposition 2.60.
" $\Leftarrow$ " : For $\lambda > 0$ define the bounded operator (called Yosida approximation)

$$A_\lambda := \lambda A(\lambda - A)^{-1} = \lambda^2(\lambda - A)^{-1} - \lambda \in \mathcal{L}(X)\,, \tag{2.79}$$

where for the second equation we used that

$$(\lambda - A)(\lambda - A)^{-1} = \mathrm{Id} \quad \text{and thus} \quad A(\lambda - A)^{-1} = \lambda(\lambda - A)^{-1} - \mathrm{Id}\,. \tag{2.80}$$

Remark that

$$A_\lambda x := \lambda(\lambda - A)^{-1}Ax \qquad \text{for all} \quad x \in \mathscr{D}(A)\,. \tag{2.81}$$

We consider the associated norm-continuous semigroup $(e^{tA_\lambda})_{t\geq 0}$, which is contractive since by (2.79) and (2.78)

$$\|e^{tA_\lambda}\| \leq e^{-\lambda t}\|e^{\lambda^2(\lambda-A)^{-1}t}\| \leq e^{-\lambda t}e^{\|\lambda^2(\lambda-A)^{-1}\|t} \leq e^{-\lambda t}e^{\lambda t} = 1\,. \tag{2.82}$$

Since by (2.78) together with (2.81)

$$\|A(\lambda - A)^{-1}y\| = \|(\lambda - A)^{-1}Ay\| \leq \frac{\|Ay\|}{\lambda} \longrightarrow 0 \quad \text{as } \lambda \to \infty$$

for all $y \in \mathscr{D}(A)$, it follows from (2.80) that for $y \in \mathscr{D}(A)$

$$\lim_{\lambda\to\infty} \lambda(\lambda - A)^{-1}y = \lim_{\lambda\to\infty} (y + A(\lambda - A)^{-1}y) = y\,. \tag{2.83}$$

Since again by (2.78) the family of operators $\lambda(\lambda - A)^{-1}$ is bounded and $\mathscr{D}(A)$ is dense by assumption i), it follows with an $\epsilon/3$-argument that (2.83) holds for all $y \in X$. Inserting $y = Ax$ and using (2.81) shows

$$\lim_{\lambda\to\infty} A_\lambda x = Ax \qquad \text{for all} \quad x \in \mathscr{D}(A)\,. \tag{2.84}$$

In the next step we show that for $x \in X$ and $t \geq 0$ the limit $\lim_{\lambda\to\infty} e^{tA_\lambda}x$ exists. First we notice that

$$\frac{d}{ds}e^{s(A_\lambda - A_\eta)}x = e^{s(A_\lambda - A_\eta)}(A_\lambda - A_\eta)x\,.$$

Integration from 0 to t gives

$$e^{t(A_\lambda - A_\eta)}x - x = \int_0^t e^{s(A_\lambda - A_\eta)}(A_\lambda - A_\eta)x\,ds$$

and by multiplication with e^{tA_η} (using that A_η and A_λ commute) we get

$$e^{tA_\lambda}x - e^{tA_\eta}x = \int_0^t e^{sA_\lambda}e^{(t-s)A_\eta}(A_\lambda - A_\eta)x\,ds\,.$$

Using that $\|e^{tA_\eta}\| \leq 1$ by (2.82), equation (2.84) yields for $x \in \mathscr{D}(A)$

$$\begin{aligned}\left\|e^{tA_\lambda}x - e^{tA_\eta}x\right\| &\leq \int_0^t \left\|e^{sA_\lambda}\right\| \left\|e^{(t-s)A_\eta}\right\| \|(A_\lambda - A_\eta)x\|\,ds \\ &\leq t\,\|(A_\lambda - A_\eta)x\| \longrightarrow 0 \qquad \text{as} \quad \lambda, \eta \to \infty\,.\end{aligned}$$

Since e^{tA_η} is bounded by (2.82), again with an $\epsilon/3$-argument we can conclude that

$$\lim_{\lambda\to\infty} e^{tA_\lambda}x \qquad \text{exists for all} \quad x \in X\,.$$

Moreover, the convergence is uniform in t on bounded intervals.

Thus for each $t \geq 0$ and $x \in X$ we can define

$$T_t x := \lim_{\lambda \to \infty} e^{tA_\lambda} x \tag{2.85}$$

and $(T_t)_{t\geq 0}$ is a contraction semigroup:

In fact, T_t is is a linear contraction since each e^{tA_λ} is. Properties a) and b) of a semigroup are obvious by the properties of the exponential function. The strong continuity follows from the fact, that the convergence in (2.85) is uniform on bounded intervals.

It remains to show that A is the generator of the contraction semigroup $(T_t)_{t\geq 0}$.

We denote the generator of $(T_t)_{t\geq 0}$ by B.

First we claim that $A \subset B$. Let $x \in \mathscr{D}(A)$, then by (2.84) and (2.85)

$$\begin{aligned}
&\left\| \int_0^t e^{sA_\lambda} A_\lambda x \, ds - \int_0^t T_s A x \, ds \right\| \\
&\leq \int_0^t \left\| e^{sA_\lambda} \right\| \|A_\lambda x - Ax\| \; ds + \int_0^t \left\| e^{sA_\lambda} Ax - T_s Ax \right\| \; ds \\
&\qquad \longrightarrow 0 \quad \text{as} \quad \lambda \to \infty \, .
\end{aligned}$$

Since as above we have

$$e^{tA_\lambda} x - x = \int_0^t e^{sA_\lambda} A_\lambda x \, ds$$

we therefore get

$$T_t x - x = \lim_{\lambda \to \infty} e^{tA_\lambda} x - x = \lim_{\lambda \to \infty} \int_0^t e^{sA_\lambda} A_\lambda x \, ds = \int_0^t T_s A x \, ds \, .$$

Thus

$$B_t x := \frac{T_t x - x}{t} \longrightarrow Ax \qquad \text{as} \quad t \to 0$$

proving that $x \in \mathscr{D}(B)$ and $Bx = \lim B_t x = Ax$ for $x \in \mathscr{D}(A)$.

Since $1 \in \rho(A)$ by hypothesis ii) and $1 \in \rho(B)$ by Proposition 2.60 it follows from $A \subset B$ that for all $x \in X$

$$x = (\mathrm{Id} - B)(\mathrm{Id} - A)^{-1}x \quad \text{and thus} \quad (\mathrm{Id} - B)^{-1}x = (\mathrm{Id} - A)^{-1}x$$

proving that $\mathscr{D}(A) = \mathscr{D}(B)$. □

2.3.5 Theorem of Hille-Yosida for general semigroups

We will now extend the Theorem of Hille-Yosida to general strongly continuous semigroups. We start with the analog to Proposition 2.60.

Proposition 2.62

Let $(T_t)_{t\geq 0}$ be a strongly continuous semigroup with generator A and let M, ω be as in Lemma 2.51, i.e. $\|T_t\| \leq Me^{\omega t}$ for all $t \geq 0$. Then

i) $\{\lambda = \eta + i\nu \in \mathbb{C} \mid \eta > \omega\} \subset \rho(A)$.

ii) for all $\lambda = \eta + i\nu$ with $\eta > \omega$.

$$(\lambda - A)^{-1}x = \int_0^\infty e^{-\lambda s} T_s x \, ds \,. \tag{2.86}$$

iii) $\|(\eta - \omega)^n (\lambda - A)^{-n}\| \leq M$ *for all $\lambda = \eta + i\nu$ with $\eta > \omega$.*

Proof. Case 1: $\omega = 0$

In this case, $\|T_t\| \leq M$ for all $t \geq 0$. Then we can define a norm on X by setting

$$|||x||| := \sup_{t\geq 0} \|T_t x\|$$

and since

$$\|x\| = \|T_0 x\| \leq |||x||| \leq M\|x\| \tag{2.87}$$

it is equivalent to the original norm. Thus replacing $\|\cdot\|$ by $|||\cdot|||$ does not change the convergence properties of sequences and their limits and therefore the strong continuity of (T_t) and the generator A do not change, if we replace $\|\cdot\|$ by $|||\cdot|||$.

To $|||\cdot|||$, we can associate an operator norm

$$|||S||| = \sup_{|||x|||=1} |||Sx|||$$

and with respect to the new norm, (T_t) is a contraction semigroup, since for any $s \geq 0$

$$|||T_s x||| = \sup_{t\geq 0} \|T_s T_t x\| = \sup_{t\geq s} \|T_t x\| \leq |||x|||\,.$$

Since the generator A is unchanged as mentioned above, it follows from Proposition 2.60 that i) and ii) hold in this case. Moreover, for any $\lambda = \eta + i\nu$ with $\eta > 0$ and $n \in \mathbb{N}$

$$\left|\left|\left|\eta^n(\lambda - A)^{-n}\right|\right|\right| \leq \left|\left|\left|\eta(\lambda - A)^{-1}\right|\right|\right|^n \leq 1\,. \tag{2.88}$$

Since by (2.87) we have for any $S \in \mathcal{L}(X)$

$$\|S\| = \sup_{x\in X} \frac{\|Sx\|}{\|x\|} \leq \sup_{x\in X} M\frac{\|Sx\|}{|||x|||} \leq M \sup_{x\in X} \frac{|||Sx|||}{|||x|||} = M|||S|||\,,$$

iii) follows at once from (2.88).

Case 2: general $\omega \in \mathbb{R}$

We consider the semigroup

$$S_t := e^{-\omega t} T_t \qquad \text{with generator} \qquad A - \omega\,.$$

Then $\|S_t\| \leq M$ for all $t \geq 0$ and thus by the arguments above it follows that

$$\{\mu \in \mathbb{C} \,|\, \operatorname{Re}\mu > 0\} \subset \rho(A-\omega) = \{\lambda - \omega \,|\, \lambda \in \rho(A)\}$$

and for μ with $\operatorname{Re}\mu > 0$ we have for all $n \in \mathbb{N}$

$$\|(\operatorname{Re}\mu)^n(\mu - (A-\omega))^{-n}\| \leq M\,.$$

If we write $\mu = \lambda - \omega$, we get $\operatorname{Re}\mu > 0$ if and only of $\operatorname{Re}\lambda > \omega$, this yields i) and iii). Property ii) follows as in the proof of Proposition 2.60.

□

Theorem 2.63 (Hille-Yosida (strongly continuous semigroups)) *Let A be an operator on a Banach space X with domain $\mathscr{D}(A)$. Then A is the infinitesimal generator of a strongly continuous semigroup if and only if A is densely defined and closed and there exists constants $\omega \in \mathbb{R}$ and $M \geq 1$ such that*

i) $(\omega, \infty) \subset \rho(A)$

ii)

$$\|(\lambda-\omega)^n(\lambda - A)^{-n}\| \leq M \quad \text{for all} \quad \lambda > \omega \text{ and } n \in \mathbb{N}\,. \tag{2.89}$$

In this case, the generated semigroup satisfies the estimate $\|T_t\| \leq Me^{\omega t}$ for all $t \geq 0$.

Proof. " $\Rightarrow$ " : This follows from Proposition 2.55 and Proposition 2.62.
" $\Leftarrow$ " : Case 1: $\omega = 0$
For $\mu > 0$ we define the norm on X

$$\|x\|_\mu := \sup_{n \geq 0} \|\mu^n(\mu - A)^{-n}x\|\,. \tag{2.90}$$

Then $\|\cdot\|_\mu$ is equivalent to our original norm, since by (2.89) we have

$$\|x\| = \|\mu^0(\mu - A)^0 x\| \le \|x\|_\mu \le M\|x\| \,. \tag{2.91}$$

Moreover, we get by the definition (2.90) for all $x \in X$

$$\begin{aligned} \|\mu(\mu - A)^{-1}x\|_\mu &= \sup_{n\ge 0} \|\mu(\mu - A)^{-1}\mu^n(\mu - A)^{-n}x\| \\ &= \sup_{m\ge 1} \|\mu^m(\mu - A)^{-m}x\| \le \|x\|_\mu \end{aligned}$$

and thus, denoting the associated operator norm by $\|.\|_\mu$ as well

$$\|\mu(\mu - A)^{-1}\|_\mu \le 1 \,. \tag{2.92}$$

We will show that $\|x\|_\mu$ is monotonically increasing with μ.

By the Resolvent equation (2.10) and (2.92) it follows that for all $0 < \lambda \le \mu$

$$\begin{aligned} \|(\lambda - A)^{-1}\|_\mu &= \|(\mu - A)^{-1} + (\lambda - \mu)(\mu - A)^{-1}(\lambda - A)^{-1}\|_\mu \\ &\le \frac{1}{\mu} + \frac{\mu - \lambda}{\mu}\|(\lambda - A)^{-1}\|_\mu \\ &= \|(\lambda - A)^{-1}\|_\mu + \frac{1}{\mu}\big(1 - \|\lambda(\lambda - A)^{-1}\|_\mu\big) \,. \end{aligned}$$

This shows

$$\|\lambda(\lambda - A)^{-1}\|_\mu \le 1 \qquad \text{for all} \quad 0 < \lambda \le \mu \,. \tag{2.93}$$

Equation (2.93) implies together with (2.91) for all $0 < \lambda \le \mu$ and $n \in \mathbb{N}$

$$\|\lambda^n(\lambda - A)^{-n}x\| \le \|\lambda^n(\lambda - A)^{-n}x\|_\mu \le \|\lambda(\lambda - A)^{-1}\|_\mu^n \, \|x\|_\mu \le \|x\|_\mu$$

and therefore by taking the supremum over all $n \in \mathbb{N}$

$$\|x\|_\lambda \le \|x\|_\mu \qquad \text{for all} \quad 0 < \lambda \le \mu$$

implying the monotonicity.

This allows to define the norm

$$|||x||| := \lim_{\mu\to\infty} \|x\|_\mu \,,$$

which by (2.91) fulfills

$$\|x\| \leq |||x||| \leq M\|x\| \qquad \text{for all} \quad x \in X \tag{2.94}$$

and therefore is equivalent to the original norm. Moreover, (2.93) yields for all $\lambda > 0$

$$\left|\left|\left|\lambda(\lambda - A)^{-1}x\right|\right|\right| = \lim_{\mu\to\infty} \|\lambda(\lambda - A)^{-1}x\|_\mu \leq \lim_{\mu\to\infty} \|x\|_\mu = |||x||| \,.$$

Thus Hypothesis iii) in Theorem 2.61 holds with respect to $|||\cdot|||$ and since by assumption A is closed and $(0,\infty) \subset \rho(A)$, it follows from Theorem 2.61 that A is the generator of a strongly continuous semigroup satisfying $|||T_t||| \leq 1$. This implies by (2.94) for all $x \in X$

$$\|T_t x\| \leq |||T_t x||| \leq |||x||| \leq M\|x\|$$

and therefore $\|T_t\| \leq M$, proving the statement in the case $\omega = 0$.

Case 2: general $\omega \in \mathbb{R}$

Consider the operator $B = A - \omega$, then B satisfies the given assumptions for $\omega = 0$, and thus, by Case 1, it generates a strongly continuous semigroup S_t satisfying $\|S_t\| \leq M$. This implies that the operator A generates the semigroup

$$T_t = e^{\omega t} S_t \quad \text{which satisfies} \quad \|T_t\| \leq e^{\omega t} M \,.$$

This proves the theorem.

□

2.3.6 Accretive operators, Lumer-Phillips-Theorem

In order to be able to apply the Theorem of Hille-Yosida, it is necessary to know the resolvent of an operator. In the following we give a criterion which only uses information about the operator itself. In order to get an idea, we consider the Hilbert space case first:

Let T_t be a contraction semigroup with generator A. The fact that $\|T_t x\|^2 \leq \|x\|^2$ for all $t \geq 0$ implies that

$$\frac{d}{dt}\|T_t x\|^2|_{t=0} \leq 0\,.$$

On the other hand

$$\frac{d}{dt}\|T_t x\|^2|_{t=0} = \langle Ax, x\rangle + \langle x, Ax\rangle = 2\,\mathrm{Re}\langle x, Ax\rangle\,,$$

thus we can conclude that

$$\mathrm{Re}\langle x, Ax\rangle \leq 0\,.$$

This condition will be generalized to the Banach space case.

We first need the following definition.

Definition 2.64 (Accretive and Dissipative operators)
Let X be a Banach space and A a densely defined linear operator on X .

i) Let $x \in X$. An element $\ell \in X^*$ that satisfies

$$\|\ell\| = \|x\| \qquad \textit{and} \qquad \ell(x) = \|x\|^2$$

is called a **normalized tangent functional** *to x.*

The map $J : X \to \mathscr{P}(X^*)$ assigning to each $x \in X$ the set of all normalized tangent functionals to x is called **duality map** *on X.*

ii) A is called **dissipative** $:\iff$ *for all* $x \in \mathscr{D}(A)$ *there exists a normalized tangent vector* $\ell \in J(x)$ *such that*

$$\operatorname{Re} \ell(Ax) \leq 0 \, .$$

iii) A is called **accretive** $:\iff$ $-A$ *is dissipative.*

Remark 2.65

The Hahn-Banach-Theorem implies that each $x \in X$ *has at least one normalized tangent functional.*

If $X = \mathscr{H}$ *is a Hilbert space, it follows from the Riesz-Lemma that the only normalized tangent functional to* $x \in \mathscr{H}$ *is* $\langle x, \cdot \rangle$ *itself, i.e.* $J(x) = \{\langle x, \cdot \rangle\}$ *(and this functional is identified with* x *in the usual way). In this case, an operator* A *is dissipative if and only if* $\operatorname{Re}\langle x, Ax \rangle \leq 0$ *on* $\mathscr{D}(A)$*. In particular this holds for self-adjoint operators which have no spectrum on the positive real line.*

If $X = L^p(\Omega, \mathcal{F}, \mu)$ *with* $1 < p < \infty$ *we have* $J(f) \subset L^q \cong (L^p)^*$ *for* $\frac{1}{p} + \frac{1}{q} = 1$ *(see Example 58.1 in the previous semester) and there is exactly one normalized tangent functional to* f *given by*

$$\ell_g(f) = \int_\Omega g(\omega) f(\omega) \, d\mu(\omega)$$

where

$$g(\omega) = \begin{cases} \|f\|_p^{2-p} f(\omega) |f(\omega)|^{p-2} & , \quad \text{if} \quad f(\omega) \neq 0 \\ 0 & , \quad \text{if} \quad f(\omega) = 0 \end{cases} .$$

For $X = \mathcal{C}([0,1])$*,* $J(1)$ *is the set of all probability measures on* $[0,1]$ *(Exercise 2.84).*

Example 2.66 (Laplace operator)
On $X = \mathcal{C}_\infty(\mathbb{R}^n)$ consider the Laplace operator

$$\Delta f(x) = \sum_{j=1}^{n} \partial_{x_j}^2 f(x) \quad \textit{with domain} \quad \mathscr{D}(\Delta) = \mathscr{S}(\mathbb{R}^n)^4 .$$

Then Δ is dissipative: For each $f \in \mathscr{S}$ there exists some x_0 such that $|f(x_0)| = \|f\|_\infty$. Set $a = \overline{f(x_0)}$ and consider the functional $\ell(f) = a\delta_{x_0}(f) = af(x_0)$. Then $\ell \in J(f)$ and

$$\operatorname{Re} \ell(\Delta f) = \operatorname{Re} a(\Delta f)(x_0) \leq 0$$

since the real valued function $\operatorname{Re} af$ takes its maximum at x_0.

Proposition 2.67
A linear operator A on a Banach space X is dissipative if and only if

$$\|(\lambda - A)x\| \geq \lambda \|x\| \qquad \textit{for all} \quad \lambda > 0 \textit{ and } x \in \mathscr{D}(A) . \tag{2.95}$$

Proof. " $\Rightarrow$ ": We assume that A is dissipative. For $x \in \mathscr{D}(A)$ we choose $\ell \in J(x)$ such that $\operatorname{Re} \ell(Ax) \leq 0$. Then for all $\lambda > 0$

$$\|\ell\| \|(\lambda - A)x\| \geq |\ell\big((\lambda - A)x\big)| \geq \operatorname{Re} \ell\big((\lambda - A)x\big)$$

[4] The **Schwartz space** $\mathscr{S}(\mathbb{R}^n)$ is the set of infinitely differentiable complex-valued functions f on $\mathbb{R}^n$ which are rapidly decreasing, i.e. such that

$$\|f\|_{\alpha,\beta} := \sup_{x \in \mathbb{R}^n} \left| x^\alpha D^\beta f(x) \right| < \infty \qquad \text{for all} \quad \alpha, \beta \in \mathbb{N}^n .$$

For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ we set $|\alpha| = \sum_i \alpha_i$,

$$x^\alpha := \prod_{j=1}^{n} x_j^{\alpha_j} \qquad \text{and} \qquad D^\beta := \frac{\partial^{|\beta|}}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n} .$$

$$\geq \lambda \operatorname{Re} \ell(x) = \lambda \|x\|^2$$

and since $\|\ell\| = \|x\|$ this shows (2.95).

" $\Leftarrow$ ": We assume that (2.95) holds. For $x \in \mathscr{D}(A)$ and $\lambda > 0$ choose any $\ell_\lambda \in J((\lambda - A)x)$. Then

$$\|\ell_\lambda\| = \|(\lambda - A)x\| \qquad \text{and} \qquad \ell_\lambda((\lambda - A)x) = \|(\lambda - A)x\|^2$$

and for the normalized functional $\tilde{\ell}_\lambda = \frac{1}{\|\ell_\lambda\|}\ell_\lambda$ it follows from (2.95) that

$$\lambda\|x\| \leq \|(\lambda - A)x\| = \tilde{\ell}_\lambda(\lambda x - Ax) = \lambda \operatorname{Re} \tilde{\ell}_\lambda(x) - \operatorname{Re} \tilde{\ell}_\lambda(Ax)\,.$$

Since by Corollary 62b (previous semester) $\|x\| = \sup_{\ell \in X^*, \|\ell\|=1} |\ell(x)| \geq |\tilde{\ell}_\lambda(x)|$, this estimate implies the two inequalities

$$|\ell_\lambda(x)| \leq \|x\| \leq \operatorname{Re} \tilde{\ell}_\lambda(x) + \frac{\|Ax\|}{\lambda} \quad \text{and} \quad \operatorname{Re} \tilde{\ell}_\lambda(Ax) \leq 0 \tag{2.96}$$

for all $\lambda > 0$. Now let

$$E = \operatorname{span}\{x, Ax\} \quad \text{and} \quad \ell'_n = \tilde{\ell}_n|_E\,, \quad n \in \mathbb{N}\,.$$

Then $\|\ell'_n\| = 1$ for all $n \in \mathbb{N}$ (since $\ell'_n((n - A)x) = \|(n - A)x\|$ by construction). Therefore, the sequence $(\ell'_n)_{n\in\mathbb{N}}$ in the finite dimensional space E^* is bounded and thus has some accumulation point ℓ'. Since (2.96) holds for each ℓ'_n, it follows that

$$\|\ell'\| = 1 \quad \text{and} \quad |\ell'(x)| \leq \|x\| \leq \operatorname{Re} \ell'(x) \quad \text{and} \operatorname{Re} \ell'(Ax) \leq 0\,. \tag{2.97}$$

Thus $\ell'(x) = \|x\|$ and by the Hahn-Banach-Theorem (Corollary 61), there is an extension $\tilde{\ell}$ of ℓ' to X^* satisfying (2.97).

Therefore, the linear functional $\ell := \|x\| \cdot \tilde{\ell}$ is a normalized tangential functional satisfying $\operatorname{Re} \ell(Ax) \leq 0$. This shows that A is dissipative. $\square$

Now we come to the fundamental criterion for an operator to be the generator of a contraction semigroup.

Theorem 2.68 (Lumer-Phillips-Theorem)
Let A be a densely defined linear operator on a Banach space X. Then A is the generator of a contraction semigroup if and only if A is dissipative and $\lambda_0 - A$ is surjective for some $\lambda_0 > 0$.

Remark 2.69
In [RS], the definition of the generator of a semigroup is given with a minus, i.e. A is the generator of the semigroup e^{-tA}. With this change of sign, the theorem of Lumer-Phillips is formulated with "dissipative" replaced by "accretive".

Proof. " $\Rightarrow$ " : If A generates a contraction semigroup, then $(0, \infty) \subset \rho(A)$ by Theorem 2.61, showing the surjectivity of $\lambda - A$ for any $\lambda > 0$. Moreover, the Hille-Yosida-condition $\|\lambda(\lambda - A)^{-1}\| \leq 1$ for all $\lambda > 0$ implies (2.95), showing that A is dissipative by Proposition 2.67.

" $\Leftarrow$ ": Assume that A is dissipative and $\lambda_0 - A$ is surjective. Then (2.78) follows at once from (2.95). Thus by Theorem 2.61 it suffices to show that A is closed and $(0, \infty) \subset \rho(A)$.

By assumption $\lambda_0 - A$ is injective with bounded inverse by (2.95). In particular, $(\lambda_0 - A)^{-1}$ is closed and this shows that $\lambda_0 - A$ and A are closed (Exercise 2.37).

To see that $(0, \infty) \subset \rho(A)$, we set

$$\Lambda = \{\lambda \in (0, \infty) \,|\, \lambda - A \ \text{ is surjective } \} = (0, \infty) \cap \rho(A)\,.$$

Since $\rho(A)$ is open, $\Lambda \subset (0, \infty)$ is open and $\Lambda \neq \emptyset$ since $\lambda_0 \in \Lambda$.

Let (λ_n) be a sequence in Λ converging to $\lambda \in (0,\infty)$. We use the following estimate, which we proved for bounded linear operators on X in Exercise 1.36 and which also holds in the unbounded case:

$$\|(\mu - A)^{-1}\| \geq \frac{1}{\operatorname{dist}(\mu, \sigma(A))}, \qquad \text{for all} \quad \mu \in \rho(A).$$

Together with (2.95) this implies for all λ_n

$$\operatorname{dist}(\lambda_n, \sigma(A)) \geq \frac{1}{\|(\lambda_n - A)^{-1}\|} \geq \lambda_n$$

and therefore

$$\operatorname{dist}(\lambda, \sigma(A)) \geq \lambda > 0.$$

Therefore $\lambda \in \rho(A) \cap (0,\infty) = \Lambda$ and thus Λ is closed in the relative topology of $(0,\infty)$. Since Λ is both open and closed in $(0,\infty)$ it follows that $\Lambda = (0,\infty)$. This shows that $\rho(A) \supset (0,\infty)$.

□

Example 2.70

Consider the following initial- and boundary-value problem:

$$\begin{aligned} \frac{\partial}{\partial t} v(t,x) &= \frac{\partial^2}{\partial x^2} v(x,t), && \textit{for} \quad t \geq 0,\ 0 \leq x \leq 1, \\ v(0,x) &= f_0(x) && \textit{for} \quad 0 \leq x \leq 1, \\ v(t,0) &= v(t,1) = 0 && \textit{for} \quad t \geq 0. \end{aligned} \tag{2.98}$$

This problem can be translated in an abstract Cauchy problem as follows: Set $X = \{f \in \mathcal{C}([0,1]) \mid f(0) = f(1) = 0\}$ with supremum-norm,

$$Af = f'' \quad \textit{with domain} \quad \mathscr{D}(A) = \{f \in X \mid f \in \mathcal{C}^2([0,1]) \textit{ and } f'' \in X\}$$

and $(u(t))(x) = v(t,x)$*. Then* (2.98) *means that we try to find*

$$u : [0,\infty) \to X \quad \text{such that} \quad u' = Au\,, \qquad u(0) = f_0\,. \tag{2.99}$$

By Proposition 2.56 such a solution exists, if A *generates a strongly continuous semigroup* $(T_t)_{t\geq 0}$ *on* X *and* $f_0 \in \mathscr{D}(A)$*. In this case* $u(t) = T_t f_0$*.*

First we remark that $\mathscr{D}(A)$ *is dense in* X *since it contains the compactly supported smooth functions, which are dense in* X*. Similar to Example 2.66 it can be shown that* A *is dissipative.*

Since the boundary value problem

$$(\mathrm{Id} - A)f = f - f'' = g \qquad f(0) = f(1) = 0$$

is uniquely solvable for each $g \in X$ *(see e.g. [Wa]), it follows that* $(\mathrm{Id} - A)$ *is surjective. Thus by Theorem 2.68,* A *generates a contraction semigroup.*

Corollary 2.71

Let A *be a densely defined closed operator on a Banach space* X*. Then* A *generates a contraction semigroup if both* A *and the adjoint operator* A' *are dissipative.*

Remark 2.72

Here the adjoint operator A' *is defined similar to the Hilbert space case. We set*

$$\mathscr{D}(A') = \{\ell \in X^* \,|\, \forall x \in \mathscr{D}(A) :\, x \mapsto \ell(Ax) \;\; \text{is bounded}\,\}$$
$$A'\ell(x) = \ell(Ax) \qquad \text{for all} \quad \ell \in \mathscr{D}(A'),\; x \in \mathscr{D}(A)\,.$$

Proof. Since we assume that A is closed and dissipative, by the Lumer-Phillips-Theorem 2.68 it suffices to show that $\mathrm{Ran}(\lambda_0 - A) = X$ for some $\lambda_0 > 0$.

Since A is dissipative, it follows from (2.95) that $(\lambda - A)$ is injective for all $\lambda > 0$. Thus if (y_n) is a sequence in $\operatorname{Ran}(\lambda - A)$ converging to $y \in X$, then there exists a sequence (x_n) in $\mathscr{D}(A)$ such that $y_n = (\lambda - A)x_n$. Since (y_n) is a Cauchy sequence, it follows form (2.95) that (x_n) is a Cauchy sequence and thus converges to some $x \in X$. But by assumption A is closed and thus $x \in \mathscr{D}(A)$ and $y = (\lambda - A)x \in \operatorname{Ran}(\lambda - A)$. This shows that $\operatorname{Ran}(\lambda - A)$ is closed.

Now suppose that $\operatorname{Ran}(\operatorname{Id} - A)$ is not dense. Then by Corollary 63 of the Hahn-Banach-Theorem there exists an $\ell \in X^*$ so that $\|\ell\| > 0$ and $\ell((\operatorname{Id} - A)x) = 0$ for all $x \in \mathscr{D}(A)$. Therefore $\ell \in \mathscr{D}(A')$ and $(\operatorname{Id} - A')\ell = 0$. This implies, that any normalized tangent functional $\nu \in X^{**}$ fulfils $\nu(A'\ell) = \nu(\ell) = \|\ell\|^2 \geq 0$. This is a contradiction to the assumption that A' is dissipative. Thus $\operatorname{Ran}(\operatorname{Id} - A)$ is dense in X. □

Corollary 2.73

Any non-positive self-adjoint operator B on a Hilbert space $\mathscr{H}$ (i.e. satisfying $\langle x, Bx \rangle \leq 0$ for all $x \in \mathscr{D}(B)$) is the generator of a contraction semigroup.

Proof. The self-adjointness of B implies that B is closed (Corollary 2.12). Moreover, since B and $B^* = B$ are non-positive, they are dissipative by Remark 2.65. □

2.3.7 Unitary groups and Stone's Theorem

We will come to the case of strongly continuous unitary groups, i.e. groups consisting of unitary operators on a Hilbert space.

Definition 2.74 (Strongly continuous unitary group)
Let $\mathscr{H}$ be a complex Hilbert space and $U_t \in \mathcal{L}(\mathscr{H})$, $t \in \mathbb{R}$, a family of unitary operators. Then $(U_t)_{t\in\mathbb{R}}$ is called **strongly continuous one-parameter unitary group** $:\Longleftrightarrow$

a) $U_{s+t} = U_s U_t$ *for all* $s, t \in \mathbb{R}$.

b) $\lim_{t\to t_0} U_t x = U_{t_0} x$ *for all* $x \in \mathscr{H}$.

If A is a self-adjoint operator on a Hilbert space, the functional calculus (Theorem 2.44) allows to define e^{iAt}.

Theorem 2.75
Let $\mathscr{H}$ be a Hilbert space and A be a self-adjoint operator on $\mathscr{H}$ with domain $\mathscr{D}(A)$. We set

$$U_t := e^{iAt} \qquad \text{for all} \quad t \in \mathbb{R}\,.$$

Then

i) the family $(U_t)_{t\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group.

ii) for $x \in \mathscr{D}(A)$

$$\frac{U_t x - x}{t} \longrightarrow iA \quad \text{as} \quad t \to 0\,.$$

iii) if $\lim_{t\to 0} \frac{1}{t}(U_t x - x)$ exists, then $x \in \mathscr{D}(A)$.

Remark 2.76
Below we give a proof of Theorem 2.75 using functional calculus and spectral decomposition introduced in Section 2.2. Another possible way would be to use the Hille-Yosida-Theorem 2.61 for contraction semigroups as follows.

Since A is self-adjoint, it follows from Exercise 1.36 that for any $\mu \in \mathbb{C}$ with $\operatorname{Im}\mu \neq 0$ (which implies $\mu \in \rho(A)$ by Lemma 2.41)

$$\|(\mu - A)^{-1}\| = \frac{1}{\operatorname{dist}(\mu, \sigma(A))} \leq \frac{1}{|\operatorname{Im}\mu|} . \tag{2.100}$$

Thus if we define $B_\pm := \pm iA$, both operators satisfy the conditions for an operator to generate a semigroup as given in Theorem 2.61:

They are closed since A is closed by Corollary 2.12 and $(0,\infty) \subset \rho(B_\pm)$ since $\sigma(B_\pm) \subset i\mathbb{R}$ by Lemma 2.41. Moreover, for any $\lambda > 0$, using $\mp iB = -i^2 A = A$ and (2.100),

$$\|\lambda(\lambda - B_\pm)^{-1}\| = \|(\mp i\lambda)((\mp i\lambda) - A)^{-1}\| \leq 1 .$$

Thus the operators B_+ and B_- generate contraction semigroups $(T_t^\pm)_{t\geq 0}$. Moreover, $T_t^+ = e^{itA}$ for $t \geq 0$ and $T_t^- = e^{itA}$ for $t < 0$. Therefore, the properties given in Theorem 2.75 follow from the respective properties for the semigroups.

Proof. i) a) If $\tilde{\Phi}_A$ denotes the algebraic $*$-homomorphism given in Theorem 2.44, then the multiplicativity of $\tilde{\Phi}_A$ gives

$$U_{t+s} = \tilde{\Phi}_A(e^{i(t+s)\lambda}) = \tilde{\Phi}_A(e^{it\lambda}e^{is\lambda}) = \tilde{\Phi}_A(e^{it\lambda}) \circ \tilde{\Phi}_A(e^{is\lambda}) = U_t U_s .$$

The unitarity of U_t then follows at once from

$$U_t^* = \tilde{\Phi}_A(e^{it\lambda})^* = \tilde{\Phi}_A(\overline{e^{it\lambda}}) = \tilde{\Phi}_A(e^{-it\lambda}) = U_{-t} \quad \text{and} \quad U_{-t}U_t = \mathrm{Id} .$$

b) The strong continuity follows by use of the spectral decomposition given in Theorem 2.47 together with the fact that $\tilde{\Phi}_A(f)^* = \tilde{\Phi}_A(\bar{f})$: observe that

$$\|e^{itA}x - x\|^2 = \int_{\mathbb{R}} |e^{it\lambda} - 1|^2 d\langle x, \Pi_\lambda x\rangle \qquad \text{for all} \quad x \in \mathscr{H} .$$

For each $t \in \mathbb{R}$ the integrand $|e^{it\lambda} - 1|^2$ is dominated by 4. But the constant function $g(\lambda) = 4$ is integrable with respect to $\langle x, \Pi_\lambda x\rangle$ since

$$\langle x, x\rangle = \langle x, \mathrm{Id}\, x\rangle = \int_{\mathbb{R}} 1\, d\langle x, \Pi_\lambda x\rangle < \infty\,.$$

Since moreover

$$|e^{it\lambda} - 1|^2 \longrightarrow 0 \quad \text{as} \quad t \to 0 \quad \text{for all} \quad \lambda \in \mathbb{R}$$

we can conclude by the dominated-convergence theorem that

$$\|U_t x - x\|^2 \longrightarrow 0 \qquad \text{as} \quad t \to 0 \quad \text{for all} \quad x \in \mathscr{H}\,.$$

Thus (U_t) is strongly continuous at $t = 0$ and by the group property this implies property b).

ii) Similar to above we write

$$\|\tfrac{1}{t}(e^{itA}x - x) - iAx\|^2 = \int_{\mathbb{R}} \Big|\tfrac{1}{t}\big(e^{it\lambda} - 1\big) - i\lambda\Big|^2 d\langle x, \Pi_\lambda x\rangle \qquad \text{for all} \quad x \in \mathscr{D}(A)\,.$$

Since for each $t \in \mathbb{R}$ the integrand $\Big|\tfrac{1}{t}\big(e^{it\lambda} - 1\big) - i\lambda\Big|^2$ is dominated by the function $g(\lambda) = 4|\lambda|^2$, which is integrable, because

$$\int_{\mathbb{R}} |\lambda|^2\, d\langle x, \Pi_\lambda x\rangle = \|Ax\|^2 < \infty \qquad \text{for} \quad x \in \mathscr{D}(A)$$

and since

$$\Big|\tfrac{1}{t}\big(e^{it\lambda} - 1\big) - i\lambda\Big| \longrightarrow 0 \quad \text{as} \quad t \to 0 \quad \text{for all} \quad \lambda \in \mathbb{R}$$

we can conclude by the dominated-convergence theorem that

$$\lim_{t \to 0} \frac{U_t x - x}{t} = iAx \qquad \text{for all} \quad x \in \mathscr{D}(A)\,.$$

iii) Define the operator

$$Bx = \lim_{t\to 0} \frac{U_t x - x}{it} \text{ with domain } \mathscr{D}(B) = \left\{ x \in \mathscr{H} \mid \lim_{t\to 0} \frac{U_t x - x}{it} \text{ exists } \right\}.$$

Then $A \subset B$ and B is symmetric since for all $x, y \in \mathscr{D}(B)$, using $U_t^* = U_{-t}$,

$$\begin{aligned} \langle x, By \rangle &= \lim_{t\to 0} \left\langle x, \frac{(U_t - 1)y}{it} \right\rangle = \lim_{t\to 0} \left\langle \frac{(U_{-t} - 1)x}{-it}, y \right\rangle \\ &= \lim_{s=(-t)\to 0} \left\langle \frac{(U_s - 1)x}{is}, y \right\rangle = \langle Bx, y \rangle . \end{aligned}$$

Since A is self-adjoint and has no proper symmetric extension by Proposition 2.14iii), it follows that $A = B$. □

The following theorem tells us that each strongly continuous unitary group arises as exponential of a self-adjoint operator.

Theorem 2.77 (Stone's Theorem)

Let $(U_t)_{t\in\mathbb{R}}$ be a strongly continuous one-parameter unitary group on a Hilbert space $\mathscr{H}$. Then there is a self-adjoint operator A on $\mathscr{H}$ such that

$$U_t = e^{itA} \qquad \text{for all} \quad t \in \mathbb{R}.$$

Proof. For $f \in \mathcal{C}_0^\infty(\mathbb{R})$ we define

$$\phi_f := \int_{\mathbb{R}} f(t) U_t \phi \, dt , \qquad \text{for all} \quad \phi \in \mathscr{H} ,$$

where the integral is a Riemann integral (this is possible since U_t is strongly continuous).

Let D be the set of finite linear combinations of all such ϕ_f for $\phi \in \mathscr{H}$ and $f \in \mathcal{C}_0^\infty(\mathbb{R})$. If $j_\epsilon \in \mathcal{C}_0^\infty(\mathbb{R})$ denotes the approximate identity introduced in Definition 2.7 in Example 2.6, then

$$\|\phi_{j_\epsilon} - \phi\| = \left\| \int_{\mathbb{R}} j_\epsilon(t)(U_t\phi - \phi)\, dt \right\| \le \left(\int_{\mathbb{R}} j_\epsilon(t)\, dt \right) \sup_{t\in[-\epsilon,\epsilon]} \|U_t\phi - \phi\| .$$

Thus the strong continuity of U_t implies that D is dense in $\mathscr{H}$.

For $\phi_f \in D$ we have (see (2.67))

$$\begin{aligned}\left(\frac{U_s - \mathrm{Id}}{s}\right)\phi_f &= \int_{\mathbb{R}} f(t)\left(\frac{U_{t+s} - U_t}{s}\right)\phi\, dt \\ &= \int_{\mathbb{R}}\left(\frac{f(\tau - s) - f(\tau)}{s}\right)U_\tau \phi\, d\tau \longrightarrow -\int_{\mathbb{R}} f'(\tau)U_\tau\phi\, d\tau = \phi_{-f'}\end{aligned}$$

since $\frac{f(\tau-s)-f(\tau)}{s}$ converges uniformly to $-f'(\tau)$.

We define the operator $A : D \to D$ by

$$A\phi_f := i^{-1}\phi_{-f'} = \lim_{s\to 0}\left(\frac{U_s - \mathrm{Id}}{is}\right)\phi_f \qquad \text{for} \quad \phi_f \in D\,. \tag{2.101}$$

We notice that $U : D \to D$ since

$$U_s\phi_f = \int_{\mathbb{R}} f(t)U_sU_t\phi\, dt = (U_s\phi)_f \in D$$

and

$$U_sA\phi_f = i^{-1}\int_{\mathbb{R}}(-f')(t)U_sU_t\phi\, dt = i^{-1}(U_s\phi)_{-f'} = AU_s\phi_f\,.$$

Furthermore, if $\phi_f, \phi_g \in D$, we have by (2.101) and since $U_t^* = U_{-t}$

$$\begin{aligned}\langle A\phi_f, \phi_g\rangle &= \lim_{s\to 0}\left\langle\left(\frac{U_s - \mathrm{Id}}{is}\right)\phi_f, \phi_g\right\rangle = \lim_{s\to 0}\left\langle\phi_f, \left(\frac{\mathrm{Id} - U_{-s}}{is}\right)\phi_g\right\rangle \\ &= \langle\phi_f, i^{-1}\phi_{-g'}\rangle = \langle\phi_f, A\phi_g\rangle\,.\end{aligned}$$

showing that A is symmetric.

Now we show that A is essentially self-adjoint. Suppose there is $\psi \in \mathscr{D}(A^*)$ such that $A^*\psi = i\psi$ (i.e. $\psi \in \mathrm{Ker}(i\,\mathrm{Id} - A^*)$). Then for each $\phi \in D = \mathscr{D}(A)$

$$\frac{d}{dt}\langle U_t\phi, \psi\rangle = \lim_{h\to 0}\langle\left(\frac{U_{t+h} - U_t}{h}\right)\phi, \psi\rangle = \langle iAU_t\phi, \psi\rangle$$

$$= -i\langle U_t\phi, A^*\psi\rangle = -i\langle U_t\phi, i\psi\rangle = \langle U_t\phi, \psi\rangle\,.$$

Thus the complex-valued function $h(t) := \langle U_t\phi, \psi\rangle$ satisfies the ordinary differential equation $h' = h$ and thus is given by

$$h(t) = e^t h(0)\,.$$

But since on the other hand U_t is unitary and therefore $|h(t)|$ is bounded, it follows that $h(0) = \langle\phi, \psi\rangle = 0$. Since D is dense, this implies $\psi = 0$. Therefore $\operatorname{Ker}(i\operatorname{Id} - A^*) = \{0\}$.

A similar proof shows that $\operatorname{Ker}(i\operatorname{Id} + A^*) = \{0\}$. By Corollary 2.22 this shows that A is essentially self-adjoint.

Set $V_t = e^{it\bar{A}}$, then it remains to show that $U_t = V_t$.

Let $\phi \in D$, then $\phi \in \mathscr{D}(\bar{A})$ and by Theorem 2.75 iii) we have

$$V_t\phi \in \mathscr{D}(\bar{A}) \quad \text{and} \quad \frac{d}{dt}V_t\phi = i\bar{A}V_t\phi\,.$$

Let $w(t) = U_t\phi - V_t\phi$, then since $U_t\phi \in \mathscr{D}(A)$ and $A = \bar{A}$ on $\mathscr{D}(A)$,

$$\frac{d}{dt}w(t) = iAU_t\phi - i\bar{A}V_t\phi = i\bar{A}w(t)$$

and therefore

$$\frac{d}{dt}\|w(t)\|^2 = -i\langle\bar{A}w(t), w(t)\rangle + i\langle w(t), \bar{A}w(t)\rangle = 0\,.$$

Since $w(0) = 0$, this implies $w(t) = 0$ for all $t \in \mathbb{R}$ and thus $V_t\phi = U_t\phi$ for all $t \in \mathbb{R}$ and $\phi \in D$. Since D is dense, this shows $V_t = U_t$.

□

Definition 2.78 (Infinitesimal Generator)

Let $(U_t)_{t\in\mathbb{R}}$ be a strongly continuous one-parameter unitary group on a Hilbert *space $\mathscr{H}$, then the self-adjoint operator A with $U_t = e^{itA}$ is called* **infinitesimal generator** *of $(U_t)_{t\in\mathbb{R}}$.*

Remark 2.79

If a unitary group $(U_t)_{t\in\mathbb{R}}$ is weakly continuous, i.e. if

$$\langle (U_t - \mathrm{Id})\phi, \psi\rangle \longrightarrow 0 \quad \text{as} \quad t \to 0 \quad \text{for all} \quad \phi, \psi \in \mathscr{H}$$

then the group is in fact strongly continuous, since we can conclude, using that U_t is unitary,

$$\|U_t\phi - \phi\|^2 = \|U_t\phi\|^2 - \langle U_t\phi, \phi\rangle - \langle \phi, U_t\phi\rangle + \|\phi\|^2 \longrightarrow 2\|\phi\|^2 - 2\|\phi\|^2 = 0$$

as $t \to 0$.

Corollary 2.80

Suppose $(U_t)_{t\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group on a Hilbert space $\mathscr{H}$. Let $D \subset \mathscr{H}$ be a dense domain, which is invariant under $(U_t)_{t\in\mathbb{R}}$ and on which U_t is strongly differentiable at $t = 0$ with strong derivative iA.

Then A is essentially self-adjoint on D and its closure $\bar{A}$ is the infinitesimal generator of $(U_t)_{t\in\mathbb{R}}$.

Corollary 2.81

Let A be a self-adjoint operator on a Hilbert space $\mathscr{H}$ and let $D \subset \mathscr{H}$ be a dense linear set contained in $\mathscr{D}(A)$. If

$$e^{itA} : D \longrightarrow D \qquad \text{for all} \quad t \in \mathbb{R}$$

then D is a core for A (i.e. $\overline{A|_D} = A$).

2.3.8 Exercises

Exercise 2.82 (Condensation of Singularities, Principle of Uniform Boundedness)

Let X, Y be Banach spaces and $\{T_{jk}\}_{j,k\in\mathbb{N}}$ a family of linear maps from X to Y. Assume that for any $k \in \mathbb{N}$ there exists $x \in X$ such that $\sup_{j\in\mathbb{N}} \|T_{jk}x\| = \infty$.

Prove that there exists $x \in X$ such that $\sup_{j\in\mathbb{N}} \|T_{jk}x\| = \infty$ for all $k \in \mathbb{N}$.

Rem.: See Theorem 3.31

Exercise 2.83 (Semigroup)

Let $q \in \mathcal{C}(\mathbb{R}^n)$ be real-valued and bounded from above and set

$$T_t f(x) := e^{tq(x)} f(x) \qquad \text{for} \quad t \geq 0\,.$$

Show that $(T_t)_{t\geq 0}$ is a strongly continuous one-parameter semigroup on X for $X = \mathcal{C}_\infty(\mathbb{R}^n)$ and for $X = L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$ and determine its infinitesimal generator.

Under which additional assumptions is $(T_t)_{t\geq 0}$ a strongly continuous semigroup on $X = L^\infty(\mathbb{R}^n)$?

Exercise 2.84 (Convolution and Young inequality)

Let $1 \leq p, q \leq \infty$ and $\frac{1}{r} := \frac{1}{p} + \frac{1}{q} - 1 \geq 0$. Let $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$.

i) Show that

$$f * g \in L^r(\mathbb{R}^n) \qquad \text{and} \qquad \|f * g\|_{L^r} \leq \|f\|_{L^p}\, \|g\|_{L^q}\,.$$

Here $f * g$ denotes the convolution of f and g given by

$$(f * g)(x) := \int_{\mathbb{R}^n} f(y) g(x-y)\, dy\,.$$

ii) Show the the convolution is associative and commutative.

Exercise 2.85 (Semigroup-property, Heat kernel)

We define the **heat kernel**

$$\gamma_t(x) := \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}, \qquad \textit{for} \quad x \in \mathbb{R}^n, t > 0$$

and set $T_0 = \mathrm{Id}$ *and for* $t > 0$

$$T_t f(x) = \gamma_t * f(x).$$

Show that $(T_t)_{t\geq 0}$ *is a strongly continuous contraction semigroup on* $L^p(\mathbb{R}^n)$ *for* $1 \leq p < \infty$*, i.e.* $T_t \in \mathcal{L}(L^p(\mathbb{R}^d))$ *such that*

i) $T_{s+t} = T_s T_t$ *for all* $s, t \geq 0$.

ii) $\lim_{t\to 0} T_t x = x$ *for all* $x \in X$.

iii) $\|T_t\| \leq 1$ *for all* $t \geq 0$.

Exercise 2.86 (Normalized Tangent Functionals)

Show that the set of normalized tangent functionals on the Banach space $\mathcal{C}([0,1])$ *to the function* $f(x) = 1$ *is given by the set of all probability measures on* $[0,1]$.

Exercise 2.87 (Contraction Semigroup)

On the Banach space $X = \{f \in \mathcal{C}([0,1]) \mid f(0) = f(1) = 0\}$ *with supremums-norm, we consider the operator (see Example 2.70)*

$$Af = f'' \quad \textit{with domain} \quad \mathscr{D}(A) = \{f \in X \mid f \in \mathcal{C}^2([0,1]) \textit{ and } f'' \in X\}.$$

Show that A *generates a contraction semigroup.*

2.4 Commutation relations

Besides the fact that in the case of unbounded operators we have to take care about the domain, another important difference to bounded operators lays in the fact that formal calculations can be misleading. In the following this will be illustrated using the notion of "commuting operators".

Two bounded self-adjoint operators $A, B \in \mathcal{L}(\mathscr{H})$ on a Hilbert space $\mathscr{H}$ are said to commute, if

$$ABx = BAx \qquad \text{for all} \quad x \in \mathscr{H} \,. \tag{2.102}$$

If A and B are unbounded, (2.102) might not make sense for any $x \in \mathscr{H}$, for example if $\operatorname{Ran} A \cap \mathscr{D}(B) = \{0\}$. This suggests that we need to find an alternative way to define commutativity. To do this, we will use the spectral theorem.

For bounded operators A, B, it follows from Theorem 1.11 together with Theorem 1.18 that (2.102) holds if and only if all spectral projection $\{\Pi_\Omega^A\}$ and $\{\Pi_\Omega^B\}$ commute.

Definition 2.88 (Commuting operators)
Two self-adjoint operators A and B on a Hilbert space $\mathscr{H}$ are said to **commute** : $\Longleftrightarrow$
All projections in the families of spectral projections

$$\{\Pi_\Omega^A \,|\, \Omega \in \mathscr{B}(\mathbb{R})\} \text{ and } \{\Pi_\Omega^B \,|\, \Omega \in \mathscr{B}(\mathbb{R})\}$$

of A and B commute.

We then have

THEOREM 2.89

Let A and B be self-adjoint operators on a Hilbert space $\mathscr{H}$. Then the following statements are equivalent.

i) A and B commute.

ii) If $\operatorname{Im}\lambda \neq 0$ and $\operatorname{Im}\mu \neq 0$, then

$$R_\lambda(A)R_\mu(B) = R_\mu(B)R_\lambda(A)\,.$$

iii) For all $s,t \in \mathbb{R}$

$$e^{itA}e^{isB} = e^{isB}e^{itA}\,.$$

Proof. "$i) \Rightarrow ii)$" and "$i) \Rightarrow iii)$": These two implications follow immediately from Theorem 2.47 (the spectral decomposition and the functional calculus for unbounded operators), since the resolvent and the exponential are bounded functions.

"$ii) \Rightarrow i)$": We will use **Stone's Formula**.

For $\epsilon > 0$ and $a, b \in \mathbb{R}$, $a < b$, set

$$f_\epsilon(x) := \frac{1}{2\pi i}\int_a^b \left(\frac{1}{x-\lambda-i\epsilon} - \frac{1}{x-\lambda+i\epsilon}\right) d\lambda\,.$$

If we denote by Γ_ϵ the closed curve in $\mathbb{C}$

given by the composition of $\gamma_{k,\epsilon}, k = 1,\dots,4$[5] (see figure above) then

$$f_\epsilon(x) = \frac{1}{2\pi i}\int_{\Gamma_\epsilon}\frac{1}{x-z}\,dz - \frac{1}{2\pi i}\int_{\gamma_{2,\epsilon}}\frac{1}{x-z}\,dz - \frac{1}{2\pi i}\int_{\gamma_{4,\epsilon}}\frac{1}{x-z}\,dz\,.$$

5

$$\gamma_{1,\epsilon} : [0,1] \to \mathbb{C}, \quad \gamma_{1,\epsilon}(t) = (1-t)a + i\epsilon + tb,$$
$$\gamma_{2,\epsilon} : [-1,1] \to \mathbb{C}, \quad \gamma_{2,\epsilon}(t) = b - ti\epsilon,$$

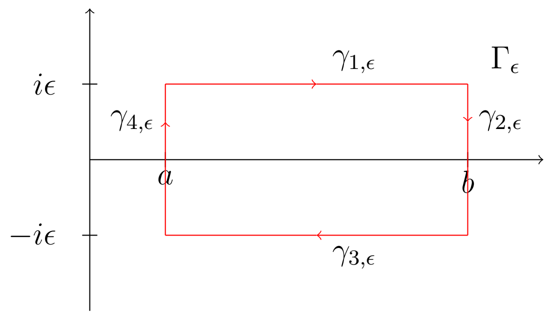

Figure 2.1: The integral contour Γ_ϵ

$$\gamma_{3,\epsilon} : [0,1] \to \mathbb{C}, \quad \gamma_{3,\epsilon}(t) = ta - i\epsilon + (1-t)b,$$

$$\gamma_{4,\epsilon} : [-1,1] \to \mathbb{C}, \quad \gamma_{4,\epsilon}(t) = a + ti\epsilon,$$

Using the Residual Theorem[6], it follows that

$$\frac{1}{2\pi i}\int_{\Gamma_\epsilon}\frac{1}{x-z}\,dz=\begin{cases}0 \quad \text{for} \quad x\notin[a,b]\\ \operatorname{Res}|_x\frac{1}{x-z}=1 \quad \text{for} \quad x\in(a,b)\\ \frac{1}{2}\operatorname{Res}|_x\frac{1}{x-z}=\frac{1}{2} \quad \text{for} \quad x\in\{a,b\}\end{cases}$$

where in the last case, this is the principal value of the integral. Since in the limit $\epsilon\to 0$ the integrals over γ_2 and γ_4 converge to zero it follows that

$$f_\epsilon(x)\longrightarrow\begin{cases}0 \quad \text{for} \quad x\notin[a,b]\\ 1 \quad \text{for} \quad x\in(a,b)\\ \frac{1}{2} \quad \text{for} \quad x\in\{a,b\}\end{cases} = \frac{1}{2}\big(\chi_{[a,b]}+\chi_{(a,b)}\big) \quad \text{as} \quad \epsilon\to 0\,.$$

[6]

Theorem 2.90 (Residual Theorem)
Let $\Omega\subset\mathbb{C}$ be a domain, $g:\Omega\to\mathbb{C}$ be meromorphic. Assume that $G\subset\Omega$ is compact and such that there are no singularities of g on the boundary ∂G. Then

$$\int_{\partial G} g(z)\,dz = 2\pi i\sum_{z\in G}\operatorname{Res}|_z g\,,$$

where $\operatorname{Res}|_z g$ denotes the residue of g at the point z.

If g has at the point z_0 a pole of order k, then

$$\operatorname{Res}|_{z_0} g = \frac{1}{(k-1)!}\partial_z^{(k-1)}|_{z_0}\Big((z-z_0)^k g(z)\Big)\,.$$

If in the setting of Theorem 2.90 there are at most finitely many poles of order 1 on the boundary ∂G, then

$$\mathcal{P}\int_{\partial G} g(z)\,dz = 2\pi i\sum_{z\in\mathring{G}}\operatorname{Res}|_z g \;+\pi i\sum_{z\in\partial G}\operatorname{Res}|_z g\,,$$

where $\mathcal{P}\int_{\partial G} g(z)\,dz$ denotes the **principal value** of the integral

Moreover, $|f_\epsilon|$ is bounded uniformly in ϵ. Thus by the functional calculus (Theorem 2.44)

$$\tilde{\Phi}_A(f_\epsilon) \xrightarrow{\text{strongly}} \tilde{\Phi}_A\Big(\frac{1}{2}(\chi_{[a,b]} + \chi_{(a,b)})\Big) \quad \text{as} \quad \epsilon \to 0$$

or more explicit

$$\frac{1}{2\pi i}\int_a^b \left(\frac{1}{A-\lambda-i\epsilon} - \frac{1}{A-\lambda+i\epsilon}\right) d\lambda \xrightarrow[\epsilon\to 0]{\text{strongly}} \frac{1}{2}(\Pi_{[a,b]} + \Pi_{(a,b)})\,. \quad (2.103)$$

This equation is called **Stone's Formula**.

Since, moreover,

$$g_\epsilon(x) = \frac{i\epsilon}{a+i\epsilon-x} \longrightarrow \begin{cases} 1\,, & \text{if} \quad a = x \\ 0\,, & \text{if} \quad a \neq x \end{cases} \qquad \text{as} \quad \epsilon \to 0$$

it follows that

$$i\epsilon R_{a+i\epsilon}(A) \xrightarrow[\epsilon\to 0]{\text{strongly}} \Pi^A_{\{a\}}\,.$$

Thus the spectral projections can be expressed as strong limits of resolvents.

"*iii*) $\Rightarrow$ *i*)": For this point we use the notion of Fourier transform and some facts, which will be proved later.

Definition 2.91 (Fourier transform)
Suppose $f \in \mathscr{S}(\mathbb{R}^n)^{2.66}$. *The* **Fourier transform** *of* f *is the function* $\hat{f}$ *given by*

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,dx \qquad (2.104)$$

where $x\cdot\xi = \sum_{j=1}^n x_j\xi_j$. *The* **inverse Fourier transform** *of* f, *denoted by* $\check{f}$, *is the function*

$$\check{f}(\xi) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} e^{ix\cdot\xi} f(x)\,dx\,. \qquad (2.105)$$

We sometimes write $\hat{f} = Ff$ *and* $\check{f} = F^{-1}f$.

Let $f \in \mathscr{S}(\mathbb{R})$, then by Fubini's theorem

$$\int_{\mathbb{R}} f(t)\langle e^{itA}\phi, \psi\rangle \, dt = \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} e^{-it\lambda} \, d\langle \Pi_\lambda^A \phi, \psi\rangle \right) dt$$
$$= \sqrt{2\pi} \int_{\mathbb{R}} \hat{f}(\lambda) d\langle \Pi_\lambda^A \phi, \psi\rangle = \sqrt{2\pi}\langle \phi, \hat{f}(A)\psi\rangle \,.$$

Using iii) and Fubini's theorem again,

$$\langle \phi, \hat{f}(A)\hat{g}(B)\psi\rangle = \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)g(s)\langle \phi, e^{-itA}e^{-isB}\psi\rangle \, ds \, dt = \langle \phi, \hat{g}(B)\hat{f}(A)\psi\rangle \,.$$

Thus we have shown that

$$\hat{f}(A)\hat{g}(B) - \hat{g}(B)\hat{f}(A) = 0 \qquad \text{for all} \quad f, g \in \mathscr{S}(\mathbb{R}) \,.$$

Now we use the following

Theorem 2.92 (Fourier Inversion Theorem)
The Fourier transform is a linear bicontinuous bijection from $\mathscr{S}(\mathbb{R}^n)$ onto $\mathscr{S}(\mathbb{R}^n)$. Its inverse map is the inverse Fourier transform, i.e.

$$\check{\hat{f}} = f = \hat{\check{f}} \qquad \textit{for all} \quad f \in \mathscr{S}(\mathbb{R}^n) \,.$$

Moreover, setting $p_\alpha(x) = (ix)^\alpha$ for all $\alpha \in \mathbb{N}^n$,

$$\left(p_\alpha D^\beta \hat{f}\right)(\xi) = D^\alpha \widehat{(-1)^{|\beta|} p_\beta f}(\xi) \,. \tag{2.106}$$

This shows that $f(A)g(B) - g(B)f(A) = 0$ for all $f, g \in \mathscr{S}(\mathbb{R})$.

Now we use that the Schwartz functions are dense in the set of bounded measurable functions $\mathcal{B}(\mathbb{R})$. Thus there are uniformly bounded sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in $\mathscr{S}(\mathbb{R})$ such that

$$f_n(x) \to \chi_{(a,b)}(x) \qquad \text{and} \qquad g_n(x) \to \chi_{(c,d)}(x) \qquad \text{for all} \quad x \in \mathbb{R} \,.$$

By Theorem 2.44 (Functional Calculus), this implies

$$f_n(A) \xrightarrow{\text{strongly}} \Pi^A_{(a,b)} \quad \text{and} \quad g_n(A) \xrightarrow{\text{strongly}} \Pi^B_{(c,d)}\,.$$

Since the sequences are uniformly bounded and

$$f_n(A)g_n(B) = g_n(B)f_n(A) \qquad \text{for all} \quad n \in \mathbb{N}$$

it follows that $\Pi^A_{(a,b)}$ and $\Pi^B_{(c,d)}$ commute. □

It is not always easy to deal with the above definition of commuting operators. A and B are often given only on sets of essential self-adjointness and it may be difficult to construct the spectral projections, resolvents or unitary groups generated by $\bar{A}$ and $\bar{B}$. It would be nice to have a criterion in terms of the operators themselves. This is not as easy.

The following two conjectures, which seem reasonable, are in fact **false**.

1) Let $D \in \mathscr{H}$ be dense and $D \subset \mathscr{D}(A) \cap \mathscr{D}(B)$. Moreover, assume that D is invariant under A and B. Then $(AB - BA)\phi = 0$ for all $\phi \in D$ implies that A and B commute.

2) Let D be a dense domain of essential self-adjointness for A and B, which is invariant under A and B. Then $(AB - BA)\phi = 0$ for all $\phi \in D$ implies that A and B commute.

Both of these statements are false, the hypotheses are not sufficient for commutativity. Although for all $n, m \in \mathbb{N}$ if follows from 1) and 2) that $A^m B^n \phi = B^n A^m \phi$ for all $\phi \in D$, we can not conclude that the unitary groups e^{itA} and e^{isB} commute on D. In the case of unbounded operators, these operators are not given by the the power series.

This can be seen by the following example of Nelson.

Example 2.93

Let M denote the Riemann surface of $\sqrt{z}$ and $\mathscr{H} = L^2(M)$ with Lebesgue measure (locally).

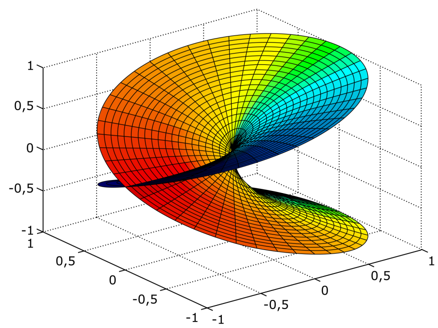

Figure 2.2: The Riemann surface of $\sqrt{z}$. Horizontal axes: real and imaginary parts of z. Vertical axis: real part of $\sqrt{z}$. Colors: imaginary part of $\sqrt{z}$, Leonid 2/CC BY-SA (https://creativecommons.org/licenses/by-sa/3.0)

Let D denote the set of all infinitely differentiable functions with compact support not containing 0. *Set $\mathscr{D}(A) = D = \mathscr{D}(B)$ with*

$$A = -i\frac{\partial}{\partial x} \qquad \text{and} \qquad B = -i\frac{\partial}{\partial y}\,.$$

Then

i) A and B are essentially self-adjoint on D.

ii) D is invariant for A and B, i.e. $A : D \to D$ and $B : D \to D$.

iii) $AB\phi = BA\phi$ for all $\phi \in D$.

iv) $e^{it\bar{A}}$ and $e^{is\bar{B}}$ do not commute.

The proofs of ii) and iii) are obvious by the definition of D.

To see i), we first remark that A and B are symmetric (by integration by parts).
Now let $D_x \subset D$ denote those functions in D, whose support does not contain the x-axis on either sheet. Then D_x is dense in $L^2(M)$ and we define the family of translation operators

$$\big(U_t\phi\big)(x,y) = \phi(x+t,y) \qquad \text{for all} \quad \phi \in D_x\,.$$

Then for each $t \in \mathbb{R}$ the operator U_t is norm-preserving with dense range D_x and extends to a unitary operator on $L^2(M)$. Since (U_t) is strongly continuous on D_x, the family of unitary operators on $L^2(M)$ is strongly continuous. Moreover, since for $\phi \in D_x$

$$\lim_{h\to 0}\frac{1}{h}(U_h\phi - \phi) = \lim_{h\to 0}\frac{1}{h}(\phi(x+h,y) - \phi(x,y)) = \frac{\partial}{\partial x}\phi(x,y)$$

it follows that (U_t) is strongly differentiable on D_x with strong derivative iA. Thus by Corollary 2.80, A is essentially self-adjoint on D_x and thus on D (this follows from Corollary 2.22 since if $\mathrm{Ran}(A \pm i\,\mathrm{Id})|_{D_x}$ is dense in $\mathscr{H}$, then $\mathrm{Ran}(A \pm i\,\mathrm{Id})$ is dense) and $\bar{A}$ generates (U_t). Similar arguments show that B is the infinitesimal generator of $V_t\phi(x,y) = \phi(x,y+t)$ defined on $D_y \subset D$ and is thus essential self-adjoint on D.

For iv), let $\phi \in D$ be supported in a small ball around the point $(-\frac{1}{2},-\frac{1}{2})$ on the first sheet. Then

$$U_1V_1\phi \neq V_1U_1\phi\,.$$

There is a similar effect in the following case.

Definition 2.94 (Canonical commutation relation, Weyl relation)
Let $\mathscr{H}$ be a Hilbert space.

i) A pair of self-adjoint operators P and Q on $\mathscr{H}$ is said to satisfy the **canonical commutation relation** $:\Longleftrightarrow$

There is a dense domain $D \subset \mathscr{D}(A) \cap \mathscr{D}(B)$, which is invariant under A and B, such that $A|_D$ and $B|_D$ are essentially self-adjoint and

$$PQ\phi - QP\phi = -i\phi \qquad \text{for all} \quad \phi \in D\,. \tag{2.107}$$

ii) A pair of continuous one-parameter unitary groups $(U_t)_{t\in\mathbb{R}}$ and $(V_t)_{t\in\mathbb{R}}$ on $\mathscr{H}$ is said to satisfy the **Weyl relation** $:\Longleftrightarrow$

$$U_t V_s = e^{ist} V_s U_t \qquad \text{for all} \quad s, t \in \mathbb{R}\,. \tag{2.108}$$

As shown in Exercise 2.99, P and Q cannot both be bounded, if they satisfy the canonical commutation relation.

The standard realisation used in quantum mechanics is the **Schrödinger representation**, where $\mathscr{H} = L^2(\mathbb{R})$ and P and Q are the closures of

$$P_0\phi(x) = \frac{1}{i}\frac{d}{dx}\phi(x) \qquad \text{and} \qquad Q_0\phi(x) = x\phi(x)$$

with domain $\mathscr{D}(P_0) = \mathscr{S}(\mathbb{R}) = \mathscr{D}(Q_0)$ (Exercise 2.99).

Then the groups e^{itP} and e^{isQ} satisfy the Weyl relation (Exercise 2.99). The following theorem, which we will not prove here, tells us that in some sense, these are the only such groups.

Theorem 2.95 (von Neumann's Theorem)
Let $(U_t)_{t\in\mathbb{R}}$ and $(V_t)_{t\in\mathbb{R}}$ be continuous one-parameter unitary groups on a separable Hilbert space $\mathscr{H}$ satisfying the Weyl relation.

Then there are closed subspaces $\mathscr{H}_k$ so that

i) $\mathscr{H} = \bigoplus_{k=1}^{N} \mathscr{H}_k$ *(here $N \in \mathbb{N}$ or $N = \infty$).*

ii) $U_t : \mathscr{H}_k \to \mathscr{H}_k$ *and* $V_s : \mathscr{H}_k \to \mathscr{H}_k$ *for all $s, t \in \mathbb{R}$ and for all k.*

iii) For each k there is a unitary operator $T_k : \mathscr{H}_k \to L^2(\mathbb{R})$ such that

$$T_k U_t T_k^{-1}\phi(x) = \phi(x-t) \quad \textit{and} \quad T_k V_s T_k^{-1}\phi(x) = e^{isx}\phi(x)\,.$$

Corollary 2.96
Let $(U_t)_{t\in\mathbb{R}}$ and $(V_t)_{t\in\mathbb{R}}$ be continuous one-parameter unitary groups on a separable Hilbert space $\mathscr{H}$ satisfying the Weyl relation. Let P be the generator of U_t and Q the generator of V_t. Then P and Q satisfy the canonical commutation relation.

Thus any solution of the Weyl relation has infinitesimal generators satisfying the canonical commutation relation. The converse of this statement is not true, as can be seen by the following example.

Example 2.97
In the setting of Example 2.93, let

$$P = A = \frac{1}{i}\frac{\partial}{\partial x} \qquad \textit{and} \qquad Q = M_x + B = x + \frac{1}{i}\frac{\partial}{\partial y}$$

with domain D given there. Then P and Q satisfy the canonical commutation relation (the proof of the self-adjointness is as in Example 2.93). But the groups they generate do not satisfy the Weyl relation (Exercise 2.98).

2.4.1 Exercises

Exercise 2.98 (Canonical commutation relation)

Let X be a normed linear space and $S, T \in L(X)$ with $ST - TS = \mathrm{Id}$. Show that

$$ST^{n+1} - T^{n+1}S = (n+1)T^n$$

holds for all $n \in \mathbb{N}_0$ and conclude that S and T can not both be bounded.

Exercise 2.99 (Canonical commutation relation - Weyl relation)

A pair of self-adjoint operators P, Q is said to satisfy the **canonical commutation relation** *if*

$$PQ - QP = -i\,\mathrm{Id}\ . \tag{2.109}$$

Consider on $\mathscr{H} = L^2(\mathbb{R})$ the operators

$$P_0\phi(x) = \frac{1}{i}\frac{d}{dx}\phi(x) \qquad \text{and} \qquad Q_0\phi(x) = x\phi(x)$$

with domain $\mathscr{D}(P_0) = \mathscr{S}(\mathbb{R}) = \mathscr{D}(Q_0)$.

i) Show that P_0 and Q_0 are essentially self-adjoint and that their self-adjoint extensions P and Q satisfy (2.109) *on $\mathscr{S}(\mathbb{R})$.*

ii) Show that the generated unitary groups $U_t = e^{itP}$ and $V_t = e^{itQ}$ satisfy the **Weyl relation***, i.e.*

$$U_tV_s = e^{its}V_sU_t \qquad \text{for all} \quad t, s \in \mathbb{R}\,. \tag{2.110}$$

Exercise 2.100 (Canonical commutation relation - Weyl relation)

Work out the details of Example 2.97:

Let M denote the Riemann surface of $\sqrt{z}$ and $\mathscr{H} = L^2(M)$ with Lebesgue measure (locally). Set

$$P = A = \frac{1}{i}\frac{\partial}{\partial x} \qquad \text{and} \qquad Q = M_x + B = x + \frac{1}{i}\frac{\partial}{\partial y}$$

with domain D given by the set of all infinitely differentiable functions with compact support not containing 0*.*

i) Show that P and Q satisfy the canonical commutation relation (2.109)

ii) Show that the generated unitary groups $U_t = e^{itP}$ and $V_t = e^{itQ}$ do not satisfy the Weyl relation (2.110)*.*

2.5 Trotters product formula

In this section, we give an approximation theorem for $e^{it(A+B)}$ in terms of e^{itA} and e^{itB}, where A and B are (essentially) self-adjoint operators.

We start with the version for finite-dimensional matrices.

Theorem 2.101 (Lie product formula)
Let A and B be finite-dimensional matrices. Then

$$e^{A+B} = \lim_{n\to\infty}\left(e^{\frac{1}{n}A}e^{\frac{1}{n}B}\right)^n.$$

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

$$S_n = e^{\frac{1}{n}(A+B)} \quad \text{and} \quad T_n = e^{\frac{1}{n}A}e^{\frac{1}{n}B}$$

then

$$S_n^n - T_n^n = \sum_{m=0}^{n-1} S_n^m(S_n - T_n)T_n^{n-1-m} \tag{2.111}$$

and therefore

$$\|S_n^n - T_n^n\| \le n\big(\max\{\|S_n\|, \|T_n\|\}\big)^{n-1}\|S_n - T_n\| \le n\|S_n - T_n\| e^{\|A\|+\|B\|}\,.$$

Since by the definition of S_n and T_n

$$\begin{aligned}\|S_n - T_n\| &= \left\| \sum_{m=0}^{\infty} \frac{1}{m!}\left(\frac{A+B}{n}\right)^m - \left(\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{A}{n}\right)^k\right)\left(\sum_{\ell=0}^{\infty}\frac{1}{\ell!}\left(\frac{B}{n}\right)^\ell\right)\right\| \\ &\le \frac{C}{n^2}\,,\end{aligned}$$

where C depends on $\|A\|$ and $\|B\|$, it follows that $\|S_n^n - T_n^n\| \to 0$ as $n \to \infty$.

□

This theorem can be extended to the case of contraction semigroups on Banach spaces. We give the proof only in the following case.

Theorem 2.102

Let A and B be self-adjoint operators on a Hilbert space $\mathscr{H}$ and suppose that $A + B$ is self-adjoint on $D = \mathscr{D}(A) \cap \mathscr{D}(B)$. Then

$$\left(e^{\frac{it}{n}A}e^{\frac{it}{n}B}\right)^n \xrightarrow{\text{strongly}} e^{it(A+B)} \qquad \textit{as} \quad n \to \infty\,.$$

Proof. Similar to the proof above, we set

$$S_n(t) = e^{\frac{it}{n}(A+B)} = S_1\left(\frac{t}{n}\right) \quad \text{and} \quad T_n(t) = e^{\frac{it}{n}A}e^{\frac{it}{n}B} = T_1\left(\frac{t}{n}\right)\,.$$

Let $\phi \in D$, then by the properties of continuous unitary groups (see Theorem 2.75)

$$\begin{aligned}\frac{1}{t}\big(T_1(t) - \mathrm{Id}\big)\phi &= \frac{1}{t}\big(e^{itA}e^{itB} - \mathrm{Id}\big)\phi = \frac{1}{t}\big(e^{itA} - \mathrm{Id}\big)\phi + \frac{1}{t}e^{itA}\big(e^{itB} - \mathrm{Id}\big)\phi \\ &\longrightarrow iA\phi + iB\phi \qquad \text{as} \quad t \to 0 \end{aligned} \tag{2.112}$$

and

$$\frac{1}{t}\big(S_1(t) - \mathrm{Id}\big)\phi = \frac{1}{t}\big(e^{it(A+B)} - \mathrm{Id}\big)\phi \longrightarrow i(A+B)\phi \qquad \text{as} \quad t \to 0\,. \tag{2.113}$$

Setting (with $s = \frac{t}{n}$)

$$K(s) := \frac{1}{s}\big(e^{isA}e^{isB} - e^{is(A+B)}\big) = \frac{1}{s}\Big(T_1(s) - S_1(s)\Big)\,,$$

it follows from (2.112) together with (2.113) that

$$K(s)\phi \longrightarrow 0 \qquad \text{as} \quad s \to 0 \quad \text{for any} \quad \phi \in D\,. \tag{2.114}$$

Since $A+B$ is self-adjoint on D, it is closed and thus D is a Banach space with respect to the graph norm

$$\|\phi\|_{A+B} := \|(A+B)\phi\| + \|\phi\|\,.$$

We denote this Banach space by $\tilde{D}$. Then for each fixed $s \in \mathbb{R}$, the map $K(s) : \tilde{D} \to \mathscr{H}$ is bounded ($\|K(s)\phi\| \leq \frac{2}{s}\|\phi\|$ for $s > 0$ and $\|K(0)\phi\| = 0$). Thus we have a family of bounded operators $\{K(s)\}_{s\in\mathbb{R}}$ on the Banach space $\tilde{D}$. Since, moreover, for each fixed $\phi \in \tilde{D}$ the set $\{\|K(s)\phi\| \,|\, s \in \mathbb{R}\}$ is bounded, it follows from the Principle of uniform boundedness (Theorem 3.31) that the family $\{\|K(s)\|\}_{s\in\mathbb{R}}$ is uniformly bounded, i.e. there is a constant C so that

$$\|K(s)\phi\| \leq C\|\phi\|_{A+B} \qquad \text{for all} \quad s \in \mathbb{R}\,,\ \phi \in D\,.$$

Thus on sets in $\tilde{D}$ which are compact (with respect to $\|\cdot\|_{A+B}$), the convergence $K(s)\phi \to 0$ is uniformly[7]. Since $\tilde{D} = \mathscr{D}(A+B)$, it follows from

[7] Let $\epsilon > 0$ and $M \subset \tilde{D}$ be compact. Cover M with balls of radius $\frac{\epsilon}{2C}$ and choose some

Lemma 2.54 that $e^{it(A+B)}\phi \in \tilde{D}$ if $\phi \in \tilde{D}$. Moreover by Lemma 2.51 the map $s \mapsto S_1(s)\phi = e^{is(A+B)}\phi$ is a continuous map from $\mathbb{R}$ into $\tilde{D}$. Thus for each fixed $\phi \in \tilde{D}$

$$\{S_1(t)\phi = e^{it(A+B)}\phi \,|\, t \in [-1,1]\} \subset \tilde{D} \quad \text{is compact}\,.$$

This shows that for any $\phi \in \tilde{D}$, uniformly for $s \in [-1,1]$,

$$K(t)S_1(s)\phi \longrightarrow 0 \qquad \text{as} \quad t \to 0\,. \tag{2.115}$$

Now we can proceed as in the proof of Theorem 2.101. The aim is to show that $\|(S_n(t))^n - (T_n(t))^n\| \to 0$ as $n \to \infty$ for any fixed $t \in \mathbb{R}$. Remark that $T_n(t) = T_1\left(\frac{t}{n}\right)$ and $S_n(t) = S_1\left(\frac{t}{n}\right)$ and therefore

$$\big(T_n(t) - S_n(t)\big) = \frac{t}{n}K\left(\frac{t}{n}\right)\,.$$

We write for $\psi \in \tilde{D}$ as in (2.111)

$$\Big(\big(S_n(t)\big)^n - \big(T_n(t)\big)^n\Big)\psi = \sum_{k=0}^{n-1}\big(T_n(t)\big)^k\frac{t}{n}K\Big(\frac{t}{n}\Big)\big(S_n(t)\big)^{n-1-k}\psi\,.$$

To estimate the norm of this term, we use that $T_n(t)$ is unitary,

$$\big(S_n(t)\big)^{n-1-k} = e^{i\frac{n-1-k}{n}t(A+B)}$$

finite subcover with balls around the points $\phi_1, \dots \phi_m$. Then by (2.114) there is some s_0 such that $\|K(s)\phi_\ell\| < \frac{\epsilon}{2}$ for all $s < s_0$ and $1 \le \ell \le m$. Let $\psi \in \tilde{D}$ be arbitrary. Then there is ϕ_ℓ such that $\|\psi - \phi_\ell\|_{A+B} < \frac{\epsilon}{2C}$ and thus for all $s < s_0$

$$\|K(s)\psi\| \le \|K(s)(\psi - \phi_\ell)\| + \|K(s)\phi_\ell\| < C\frac{\epsilon}{2C} + \frac{\epsilon}{2} = \epsilon\,.$$

and $s = \frac{n-1-k}{n} t \leq |t|$. Thus by (2.115) for any fixed $t \in \mathbb{R}$ and $\phi \in \tilde{D}$

$$\left\| \big(S_n(t)\big)^n - \big(T_n(t)\big)^n \psi \right\| \leq |t| \max_{s \leq |t|} \left\| K\left(\frac{t}{n}\right) \big(S_1(s)\big) \right\| \longrightarrow 0$$

as $n \to \infty$. Since D is dense and the operators are bounded by one, this proves the theorem. □

The following result, which we will not prove here, is stronger, since it only requires essential self-adjointness of $A + B$ on $\mathscr{D}(A) \cap \mathscr{D}(B)$.

Theorem 2.103 (Trotter product formula)
Let A and B be self-adjoint operators on a Hilbert space $\mathscr{H}$ and suppose that $A + B$ is essentially self-adjoint on $D = \mathscr{D}(A) \cap \mathscr{D}(B)$. Then

$$\left(e^{\frac{it}{n} A} e^{\frac{it}{n} B}\right)^n \xrightarrow{\text{strongly}} e^{it(A+B)} \qquad \text{as} \quad n \to \infty\,.$$

A similar result (which we also will not prove) holds in the case of generators of contraction semigroups.

Theorem 2.104 (Trotter product formula)
Let A and B be generators of contraction semigroups $(T_t)_{t \geq 0}$ and $(S_t)_{t \geq 0}$ on a Banach space X. Suppose that the closure of $A + B$ restricted to $D = \mathscr{D}(A) \cap \mathscr{D}(B)$ generates a contraction semigroup $(U_t)_{t \geq 0}$ on X. Then for all $\phi \in X$

$$\lim_{n \to \infty} \left(T_{t/n} S_{t/n}\right)^n \phi = U_t \phi\,.$$

If $(A + B)|_D$ is closed, the proof is exactly as the proof of Theorem 2.102.

Chapter 3

Locally convex spaces, Distributions, Fourier transform

3.1 Locally convex spaces

The idea behind locally convex spaces is that instead of a norm the topology is given by a family of seminorms.

3.1.1 Topology generated by families of seminorms

Definition 3.1 (Seminorm)
A **seminorm** *on a vector space X is a map $\rho : X \mapsto [0, \infty)$ obeying for all $x, y \in X$ and $\alpha \in \mathbb{K}$*

a) $\rho(x + y) \leq \rho(x) + \rho(y)$.

b) $\rho(\alpha x) = |\alpha|\rho(x)$.

A family of seminorms $P := \{\rho_\alpha\}_{\alpha \in A}$*, where* A *is any index set, is said to* **separate points** $:\iff$ $\rho_\alpha(x) = 0$ *for all* $\alpha \in A$ *implies* $x = 0$.

If $B \subset A$ *is finite and* $\epsilon > 0$ *we set*

$$U_{B,\epsilon} = \{x \in X \,|\, \forall \alpha \in B \;:\; \rho_\alpha(x) < \epsilon\} \quad \textit{and} \tag{3.1}$$
$$\mathscr{U} = \{U_{B,\epsilon} \,|\, B \subset A \textit{ finite},\ \epsilon > 0\}\,.$$

The system $\mathscr{U}$ substitutes the set of all balls at 0 in a normed vector space. It has the following properties:

1. $0 \in U$ for all $U \in \mathscr{U}$.

2. For any $U_{B_1,\epsilon_1}, U_{B_2,\epsilon_2} \in \mathscr{U}$ there exists $U \in \mathscr{U}$ such that $U \subset U_1 \cap U_2$ (in fact we can take $U_{B,\epsilon}$ with $B = B_1 \cup B_2$ and $\epsilon = \min\{\epsilon_1, \epsilon_2\}$).

3. For any $U \in \mathscr{U}$ there is $V \in \mathscr{U}$ such that $V + V \subset U$ (for $U = U_{B,\epsilon}$ take $V = U_{B,\epsilon/2}$).

4. Any $U \in \mathscr{U}$ is **absorbing**, i.e. the **Minkowski functional**

   $$p_U : X \to [0,\infty]\,, \quad p_U(x) := \inf\{\lambda > 0 \,|\, x \in \lambda U\} \tag{3.2}$$

   is finite for all $x \in X$.

   To see this, take $U = U_{B,\epsilon}$ and $x \in U$. Then $x \in \lambda U$ if $\max_{\alpha \in B} \rho_\alpha(x) < \lambda\epsilon$, so $p_U(x) \leq \frac{1}{\epsilon} \max_{\alpha \in B} \rho_\alpha(x) < \infty$.

5. For any $U \in \mathscr{U}$ and $\lambda > 0$ there is $V \in \mathscr{U}$ such that $\lambda V \subset U$ (if $U = U_{B,\epsilon}$ and $V = U_{B,\epsilon/\lambda}$, then $\lambda V = U$).

6. Each $U \in \mathscr{U}$ is **balanced**, i.e. $\gamma U \subset U$ for all $\gamma \in \mathbb{K}$ with $|\gamma| \leq 1$.

7. Each $U \in \mathscr{U}$ is **absolutely convex**, i.e. U is balanced and convex (i.e. $tU + (1-t)U \subset U$ for all $0 \leq t \leq 1$).

A set A is absolutely convex if for all $x, y \in A$ and $\lambda, \mu \in \mathbb{C}$ such that $|\lambda| + |\mu| \leq 1$ it holds that $\lambda x + \mu y \in A$.

Now let $\mathscr{U}$ be any system of sets in X obeying properties 1. to 6. Then it is possible to define a topology[1] on X by setting

$$M \subset X \text{ is open} \quad :\Longleftrightarrow \quad \forall\, x \in M\ \exists U \in \mathscr{U}\ : x + U \subset M\,. \tag{3.3}$$

This is in fact a topology:

i) It is clear that $\emptyset$ and X are open.

ii) Let M_1, M_2 be open and $x \in M_1 \cap M_2$. Then there exist $U_i \in \mathscr{U}$ such that $x + U_i \subset M_i$ for $i = 1, 2$. By 2. there is some $V \in \mathscr{U}$ such that $V \subset U_1 \cap U_2$ and therefore $x + V \subset M_1 \cap M_2$. Thus $M_1 \cap M_2$ is open.

[1] A **topology** τ on a set X is a system of subsets of X satisfying

i) $\emptyset \in \tau$ and $X \in \tau$.

ii) If $O_1, O_2 \in \tau$, then $O_1 \cap O_2 \in \tau$.

iii) If I is any index set and $O_i \in \tau$ for all $i \in I$, then $\bigcup_{i \in I} O_I \in \tau$.

The elements of τ are called **open** sets. A **neighbourhood** of a point $x \in X$ is a set $U \subset X$ such that there exists some $O \in \tau$ with $x \in O \subset U$. A **neighbourhood base** at a point $x \in X$ is a system $\mathscr{U}_x$ of neighbourhoods U of x, so that for each neighbourhood V of x there is some $U \in \mathscr{U}_x$ with $U \subset V$. The **boundary** ∂M of a set $M \subset X$ consists of all points $m \in X$ for which $U \cap M \neq \emptyset$ and $U \cap M^c \neq \emptyset$ for each $U \in \mathscr{U}_m$. A sequence (x_n) in X is said to **converge** to $x \in X$, if for all $U \in \mathscr{U}_x$ there is some $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. Such a limit does not need to be unique.

iii) Let M_i, $i \in I$, be open and $x \in \bigcup_{i\in I} M_i$, take $x \in M_{i_0}$. Then there exists some $U \in \mathscr{U}$ such that $x + U \subset M_{i_0} \subset \bigcup_{i\in I} M_i$. Thus $\bigcup_{i\in I} M_i$ is open.

By construction, $\mathscr{U}$ is a neighbourhood base (or local base) at 0 of the topology.

Lemma 3.2

In the topology given in (3.3)*, addition and scalar multiplication are continuous*[2] *with respect to the product topology on* $X \times X$ *and* $\mathbb{C} \times X$*.*

Proof. Let $M \subset X$ be open. It has to be shown that

$$\tilde{M} = \{(x,y) \,|\, x + y \in M\} \quad \text{and} \quad M' = \{(\lambda, x) \,|\, \lambda x \in M\} \quad \text{are open.}$$

If $x+y \in M$, then there exists $U \in \mathscr{U}$ such that $x+y+U \subset M$. By 3. there is some $V \in \mathscr{U}$ such that $V+V \subset U$. But this implies $(x+V)+(y+V) \subset M$ and thus $(x+V, y+V) \subset \tilde{M}$. Therefore $\tilde{M}$ is open.

If $\lambda x \in M$, then $\lambda x + U \subset M$ for some $U \in \mathscr{U}$. Choose again $V \in \mathscr{U}$ according to 3, i.e. such that $V + V \subset U$. Then by 4. there is some $\epsilon > 0$ such that $\epsilon x \in V$. Since V is balanced, it follows that

$$(\mu - \lambda)x \in V \quad \text{if} \quad |\mu - \lambda| < \epsilon\,.$$

[2] If X_1 and X_2 are topological spaces, then a map $f : X_1 \to X_2$ is called continuous at x_0, if for each neighbourhood V of $f(x_0)$ the pre-image $f^{-1}(V)$ is a neighbourhood of x_0. If $\mathscr{U}_{x_0}$ and $\mathscr{V}_{f(x_0)}$ are neighbourhood bases at x_0 and $f(x_0)$ respectively, then f is continuous at x_0 if for any $V \in \mathscr{V}_{f(x_0)}$ there is some $U \in \mathscr{U}_{x_0}$ such that $f(U) \subset V$. The function f is called continuous if it is continuous at all $x \in X$. This holds if and only if the pre-images of all open sets in X_2 are open in X_1.

Now by 5. and 6. there is $W \in \mathscr{U}$ such that $\mu W \subset V$ if $|\mu| < |\lambda| + \epsilon$. Thus if $|\mu - \lambda| < \epsilon$ and $w \in W$

$$\mu(x + w) - \lambda x = (\mu - \lambda)x + \mu w \in V + V \subset U$$

proving that $\{\mu \,|\, |\mu - \lambda| < \epsilon\} \times (x + W) \subset M'$. □

3.1.2 Definition, examples and fundamental properties

Definition 3.3 (Locally convex space)
Let X be a vector space and τ a topology on X.

i) (X, τ) *is called* **topological vector space** $:\iff$ *addition and scalar multiplication are continuous.*

ii) Let $P = \{\rho_\alpha\}_{\alpha \in A}$ be a family of seminorms on X and τ the associated topology defined in (3.3)*. Then (X, τ) is called* **locally convex topological vector space** *or simply* **locally convex space***.*

It follows from Lemma 3.2 that a locally convex space is in fact a topological vector space.

Example 3.4 *i) Let S be a set and $X = \mathbb{C}^S$ the space of functions $f : S \to \mathbb{C}$, then the family $P = \{p_t\}_{t \in S}$ of seminorms with $p_t(f) := |f(t)|$ generates the locally convex topology of pointwise convergence on X.*

ii) Let (S, τ) be a topological space and $X = \mathcal{C}(S, \mathbb{C})$. For any compact set $K \subset S$ we define the seminorm

$$p_K(f) := \sup_{t \in K} |f(t)|\,.$$

Then the family $P = \{p_K \,|\, K \subset S \text{ compact}\}$ *generates the locally convex topology of uniform convergence on compact sets.*

iii) *On* $X = \mathcal{C}^\infty(\Omega)$*, where* $\Omega = \mathbb{R}^n$ *or* $\Omega \subset \mathbb{R}^n$ *open, consider the seminorms*

$$p_{K,\alpha}(f) := \sup_{x\in K} |D^\alpha f(x)| \quad \textit{where} \quad K \subset \Omega \ \ \textit{compact}\,, \alpha \in \mathbb{N}^n\,.$$

The locally convex space generated by the family

$$P = \{p_{K,\alpha} \,|\, \alpha \in \mathbb{N}^n,\ K \subset \Omega \ \ \textit{compact}\ \}$$

is often denoted by $\mathscr{E}(\Omega)$*.*

iv) *The Schwartz space* $\mathscr{S}(\mathbb{R}^n)^{2.66}$ *(infinitely differentiable functions of rapid decrease) together with the family of seminorms* $\|\cdot\|_{\alpha\beta}$ *given by*

$$\|f\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} \left|x^\alpha D^\beta f(x)\right| < \infty \qquad \textit{for all} \quad \alpha, \beta \in \mathbb{N}^n \tag{3.4}$$

is a locally convex space (Exercise 3.35).

v) *For* $\Omega \subset \mathbb{R}^n$ *open and* $K \subset \Omega$ *compact set*

$$\mathscr{D}_K(\Omega) := \{f \in \mathcal{C}^\infty(\Omega) \,|\, \operatorname{supp} f \subset K\} \quad \textit{and} \quad p_\alpha(f) = \sup_{x\in\Omega} |D^\alpha f(x)|\,.$$

Then $\mathscr{D}_K(\Omega)$ *together with the family of seminorms* $P = \{p_\alpha\}_{\alpha\in\mathbb{N}^n}$ *is a locally convex space.*

vi) *For* $\mathscr{D}_K(\Omega)$ *as given in v) we denote the topology by* τ_K*. Set* $\mathscr{D}(\Omega) = \bigcup_K \mathscr{D}_K(\Omega)$ *and let* P *be the family of all seminorms* p *on* $\mathscr{D}(\Omega)$ *such that all restrictions* $p|_{\mathscr{D}_K}$ *are continuous with respect to* τ_K*.*

vii) If X is a vector space with norm $\|\cdot\|$, then the locally convex topology generated by $P = \{\|\cdot\|\}$ is the norm topology.

viii) On a normed vector space X with dual space X^*, the locally convex topology generated by the seminorms

$$p_\ell(x) := |\ell(x)|\,, \qquad \ell \in X^*,$$

is the weak topology $\sigma(X, X^*)$ on X.

ix) On the dual space X^* of a normed vector space X, the locally convex topology generated by the seminorms

$$p_x(\ell) := |\ell(x)|\,, \qquad x \in X$$

is the weak- topology on X^*.*

x) If X and Y are normed vector spaces, then on $\mathcal{L}(X, Y)$ there are three topologies: the norm topology generated by the operator norm, the strong operator topology generated by the seminorms

$$p_x(T) := \|Tx\|_Y\,, \qquad x \in X$$

and the weak operator topology generated by the seminorms

$$p_{x,\ell}(T) := |\ell(Tx)|\,, \qquad \ell \in Y^*\,,\; x \in X\,.$$

xi) On any vector space X the seminorm $p(x) = 0$ generates the chaotic topology, in which only X and $\emptyset$ are open. This "family of seminorms" does obviously not separate points.

Lemma 3.5

Let X be a locally convex space, where the topology τ is generated by the family of seminorms $P = \{p_\alpha\}_{\alpha \in A}$. Then the following statements are equivalent:

i) (X, τ) is a Hausdorff space[3].

ii) P separates points.

iii) There exists a neighbourhood base $\mathscr{U}$ at 0 such that $\bigcap_{U \in \mathscr{U}} U = \{0\}$.

Proof. $i) \Rightarrow ii)$: Let $x \neq 0$ and let U and V be neighbourhoods of 0 satisfying $(x+U) \cap V = \emptyset$. Since τ is generated by P there are $B \subset A$ finite and $\epsilon > 0$ such that $V = U_{B,\epsilon} = \{x \in X \,|\, \forall \alpha \in B : \; p_\alpha(x) < \epsilon\}$. Therefore $x \notin V$ implies that $p_\alpha(x) \geq \epsilon$ for some $\alpha \in B$.
$ii) \Rightarrow iii)$: This follows from the fact that

$$x \in \bigcap_{B,\epsilon} U_{B,\epsilon} \quad \Longleftrightarrow \quad p_\alpha(x) = 0 \quad \text{for all} \quad \alpha \in A$$

and if P separates points, the last statement is equivalent to $x = 0$.
$iii) \Rightarrow i)$: Let $x \neq y$, then by assumption there exists some $U = U_{B,\epsilon} \in \mathscr{U}$ such that $x - y \neq U$, i.e $p_\alpha(x-y) \geq \epsilon$ for all $\alpha \in B$. Let $V = U_{B,\epsilon/4}$, then $(x+V) \cap (y+V) = \emptyset$. In fact, assume that $x + v = y + w$ for some $v, w \in V$. It follows from the triangle inequality that

$$p_\alpha(x-y) = p_\alpha(x-v-y-w+v+w) \leq p_\alpha((x+v)-(y+w)) + p_\alpha(v) + p_\alpha(w)$$

and therefore for all $\alpha \in B$ by the definition of U and V

$$p_\alpha((x+v)-(y+w)) \geq p_\alpha(x-y) - p_\alpha(v) - p_\alpha(w) \geq \epsilon - \frac{\epsilon}{2} > 0$$

which contradicts the assumption. □

[3]Two different points have disjoint neighbourhoods

This shows that all examples except the last are Hausdorff spaces. The following proposition explains, why these spaces are called locally convex.

We start with a lemma.

Lemma 3.6
Let X be a vector space and $U \subset X$ a subset. Then the Minkowski functional p_U defined in (3.2) *is*

i) sublinear[4] *if U is convex and absorbing.*

ii) a seminorm, if U is absorbing, convex and balanced.

Proof. i) Since U is absorbing, p_U is finite by definition.

From the definition of p_U it follows by setting $\mu = t\lambda$ that

$$tp_U(x) = t\inf\{\lambda > 0 \,|\, \frac{x}{\lambda} \in U\} = \inf\{\mu > 0 \,|\, \frac{tx}{\mu} \in U\} = p_U(tx) \tag{3.5}$$

for any $t \geq 0$ and $x \in X$. In order to see the inequality

$$p_U(x+y) \leq p_U(x) + p_U(y) \qquad \text{for all} \quad x, y \in X \tag{3.6}$$

let $x, y \in X$ be fixed. Choose any $\epsilon > 0$. Since the Minkowski functional $p_U(x)$ is the infimum over all $t > 0$ such that $\frac{x}{t} \in U$, it follows that there are $\lambda, \mu > 0$ such that

$$\lambda \leq p_U(x) + \epsilon \quad \text{and} \quad \mu \leq p_U(y) + \epsilon \quad \text{and} \quad \frac{x}{\lambda}, \frac{y}{\mu} \in U\,.$$

[4] If V is a vector space, a map $g : X \to \mathbb{R}$ is called **sublinear** if

(a) $p(\lambda v) = \lambda p(v)$ for all $\lambda \geq 0$ and $v \in V$.

(b) $p(v + w) \leq p(v) + p(w)$ for all $v, w \in V$.

The convexity of U implies

$$\frac{\lambda}{\lambda+\mu}\,\frac{x}{\lambda}+\frac{\mu}{\lambda+\mu}\,\frac{y}{\mu}=\frac{x+y}{\lambda+\mu}\in U\,.$$

Therefore

$$p_U(x+y)\leq \lambda+\mu\leq p_U(x)+p_U(y)+2\epsilon$$

and since $\epsilon>0$ was arbitrary, this proves (3.6).
ii) Since U is balanced, i.e. $\gamma U\subset U$ for all $|\gamma|\leq 1$, it follows that $U=\lambda U$ for $|\lambda|=1$ and thus

$$p_U(\lambda x)=p_{\lambda U}(\lambda x)=p_U(x)=|\lambda|p_U(x)\qquad \text{for all}\quad |\lambda|=1\,,\ x\in X\,.$$

Combining this with (3.5) shows that

$$|\lambda|p_U(x)=p_U(\lambda x)\qquad \text{for any}\quad \lambda\in\mathbb{K}\,,\ x\in X\,.$$

Together with (3.6) and the fact that p_U is finite, this shows that p_U is a seminorm. □

Proposition 3.7

A topological vector space (X,τ) is locally convex if and only if it has a neighbourhood base at 0 consisting of convex, balanced and absorbing sets.

Proof. " $\Rightarrow$ ": This follows at once from the construction of a locally convex space.
" $\Leftarrow$ ": Assume that τ has a neighbourhood base $\mathscr{U}$ and that any $U\in\mathscr{U}$ is convex, balanced and absorbing. Then by Lemma 3.6, each p_U is a seminorm.

Consider now the family of seminorms $P=\{p_U\}_{U\in\mathscr{U}}$. This family generates a topology $\tilde{\tau}$ on X with the neighbourhood base at 0 given by

$$\tilde{\mathscr{U}}=\{U_{\mathscr{V},\epsilon}\,|\,\epsilon>0,\ \mathscr{V}\subset\mathscr{U}\ \text{ finite }\}$$

$$U_{\mathscr{V},\epsilon} = \{x \in X \,|\, \forall U \in \mathscr{V} \,:\, p_U(x) < \epsilon\}\,.$$

We show that the topologies τ and $\tilde{\tau}$ are equal:

In order to see $\tau \subset \tilde{\tau}$, we observe that

$$U_{\{U\},1} = \{x \in X \,|\, p_U(x) < 1\} = U \qquad \text{if} \quad 0 \in U \subset X \text{ open and convex}\,. \tag{3.7}$$

In fact, if $p_U(x) < 1$, then there exists $\lambda < 1$ such that $\frac{x}{\lambda} \in U$. Since $0 \in U$, it follows from the convexity that

$$x = \lambda \frac{x}{\lambda} + (1-\lambda)0 \in U\,.$$

Thus $U_{\{U\},1} \subset U$.

On the other hand, if $p_U(x) \geq 1$, then $\frac{x}{\lambda} \neq U$ for all $\lambda < 1$. Since U^c is closed, it follows that

$$x = \lim_{\lambda \nearrow 1} \frac{x}{\lambda} \in U^c\,.$$

This shows that $U^c_{\{U\},1} \subset U^c$. Thus $U = U_{\{U\},1} \in \tilde{\mathscr{U}}$ for any $U \in \mathscr{U}$.

The inclusion $\tilde{\tau} \subset \tau$ holds, since for any finite $\mathscr{V} \subset \mathscr{U}$ and $\epsilon > 0$

$$U_{\mathscr{V},\epsilon} = \bigcap_{U \in \mathscr{V}} \epsilon U \in \tau\,.$$

In fact, $x \in U_{\mathscr{V},\epsilon}$ if and only if $p_U(x) < \epsilon$ which is equivalent to $x \in \epsilon U$ for all $U \in \mathscr{V}$. $\square$

This Proposition shows, that locally convex spaces can be determined either by a family of seminorms or by an absolutely convex neighbourhood base at 0.

3.1.3 Continuous linear maps

The equivalence of continuity and boundedness for linear maps given in normed vector spaces can be generalized to locally convex spaces. We start with criteria for the continuity of a seminorm.

Lemma 3.8

Let X be a locally convex space with topology τ generated by the family of seminorms $P = \{p_\alpha\}_{\alpha \in A}$.

i) Let $q : X \to [0, \infty)$ be a seminorm. Then the following statements are equivalent.

  (a) q is continuous.

  (b) q is continuous at 0.

  (c) $\{x \in X \,|\, q(x) < 1\}$ is a neighbourhood of 0.

ii) p_α is continuous for all $\alpha \in A$.

iii) A seminorm q on X is continuous if and only if

$$\exists\, M > 0 \; \exists B \subset A \text{ finite} : \quad q(x) < M \max_{\alpha \in B} p_\alpha(x) \qquad \text{for all} \quad x \in X\,. \tag{3.8}$$

Proof. $i)$: The implication $(a) \Rightarrow (b)$ is clear. But q is continuous at 0 if and only if for each neighbourhood V of $q(0) = 0$, the pre-image is a neighbourhood of 0. Thus $(b) \Rightarrow (c)$ is clear.

$(c) \Rightarrow (a)$: Assume that (c) holds. Let $x \in X$ and $\epsilon > 0$ be given. We have to show that the pre-image of the ϵ-interval around $q(x)$ is open in X.

Set
$$U := \epsilon\{y \in X \mid q(y) < 1\} = \{y \in X \mid q(y) < \epsilon\}\,.$$
Then for any $y \in U$
$$|q(x+y) - q(x)| \leq q((x+y) - x) = q(y) < \epsilon$$
and therefore
$$q(x+U) \subset \{\lambda \in \mathbb{R} \mid |\lambda - q(x)| < \epsilon\}\,.$$
This proves the continuity of q.

ii): By the construction of τ the sets $\{x \in X \mid p_\alpha(x) < 1\}$ are neighbourhoods of 0 for all $\alpha \in A$. The continuity of all seminorms in P thus follows from i).

iii): By *i)* it suffices to show that (3.8) is equivalent to the fact that $N = \{x \in X \mid q(x) < 1\}$ is a neighbourhood of 0.

N is a neighbourhood of 0 if and only if there exist $\epsilon > 0$ and $B \subset A$ finite such that $U_{B,\epsilon} \subset N$. But this is equivalent to (3.8) for $M = \frac{1}{\epsilon}$ and the same finite $B \subset A$, since $\frac{1}{\epsilon}\max_{\alpha\in B} p_\alpha(x) < 1$ for any $x \in U_{B,\epsilon}$. □

It can be interesting to compare two different topologies (or two different families of seminorms) on the same vector space X.

Definition 3.9

On the vector space X consider the families of seminorms P and Q generating the topologies τ_P and τ_Q respectively. Then P and Q are called **equivalent** $:\iff \tau_P = \tau_Q$.

Corollary 3.10

On the vector space X consider the families of seminorms P and Q generating the topologies τ_P and τ_Q respectively.

i) $\tau_P = \tau_Q$ *holds if*

$$P \subset Q \subset \{q \ \ \textit{seminorm} \ | \, q \textit{ is continuous with respect to } \tau_P\}\,.$$

ii) P *and* Q *are equivalent if and only if each* $p \in P$ *is continuous with respect to* τ_Q *and each* $q \in Q$ *is continuous with respect to* τ_P*.*

This lemma allows to get the following characterisation of a convergent net.

Corollary 3.11

Under the assumptions of Lemma 3.8 a net[5] $(x_\gamma)_{\gamma \in I}$ *converges to* $x \in X$ *if and only if*

$$p_\alpha(x_\gamma - x) \longrightarrow 0\,, \qquad \textit{for all} \quad \alpha \in A\,.$$

Proof. We have $x_\gamma \to x$ if and only if $(x_\gamma - x) \to 0$, thus it suffices to consider $x = 0$. " $\Rightarrow$ ": If $x_\gamma \to 0$ it follows from Lemma 3.8ii) together with

[5] A **directed system** is an index set I together with an ordering $\prec$ wich satisfies

i) If $\alpha, \beta \in I$ then there exists $\gamma \in I$ such that $\gamma \succ \alpha$ and $\gamma \succ \beta$.

ii) $\prec$ is a partial ordering, i.e. $\alpha \prec \alpha$ for all $\alpha \in I$ and $\alpha \prec \beta$ and $\beta \prec \gamma$ implies $\alpha \prec \gamma$ for all $\alpha, \beta, \gamma \in I$.

A **net** is a mapping from a directed system I to X and is denoted by $(x_\alpha)_{\alpha \in I}$. A net $(x_\alpha)_{\alpha \in I}$ is said to converge to $x \in X$ (written $x_\alpha \to x$) if for any neighbourhood U of x there is some $\beta \in I$ so that $x_\alpha \in U$ for all $\alpha \succ \beta$.

Proposition 3.12

Let X *and* Y *be topological spaces. A function* $f : X \to Y$ *is continuous at* x_0 *if and only if for every convergent net* $(x_\alpha)_{\alpha \in I}$ *in* X *with* $x_\alpha \to x$*, the net* $\big(f(x_\alpha)\big)_{\alpha \in I}$ *converges in* Y *to* $f(x)$*.*

Proposition 3.12 that $p_\alpha(x_\gamma) \to 0$ for all $\alpha \in A$.

"$\Leftarrow$": Let $U \in \mathscr{U}$ be a neighbourhood of 0, then by construction $U = U_{B,\epsilon}$ for some $\epsilon > 0$ and $B \subset A$ finite. By assumption, for any $\alpha \in A$ there is $\gamma_\alpha \in I$ such that

$$p_\alpha(x_\gamma) < \epsilon \qquad \text{for all} \quad \gamma \succ \gamma_\alpha \,.$$

Since I is directed and B is finite, it follows that there is a $\beta \in I$ such that $\beta \succ \gamma_\alpha$ for all $\alpha \in B$. This implies

$$p_\alpha(x_\gamma) < \epsilon \qquad \text{for all} \quad \gamma \succ \beta \text{ and } \alpha \in B$$

and thus $x_\gamma \in U_{B,\epsilon}$ for all $\gamma \succ \beta$. This shows that $x_\gamma \to 0$. □

Definition 3.13 (Complete locally convex space)

A net $(x_\gamma)_{\gamma \in I}$ in a locally convex space X generated by the family of seminorms $P = \{p_\alpha\}_{\alpha \in A}$ is called **Cauchy** $:\Longleftrightarrow$

$$\forall \epsilon > 0 \; \forall \alpha \in A \; \exists \beta_0 \in I \; \forall \, \beta, \gamma \succ \beta_0 \; : \; p_\alpha(x_\beta - x_\gamma) < \epsilon \,.$$

X is called **complete** *if every Cauchy net converges.*

We now consider linear maps between locally convex spaces. The following proposition generalizes the fact that linear maps between normed vector spaces are continuous if and only if they are bounded.

Proposition 3.14

Let (X, τ_X) and (Y, τ_Y) be locally convex spaces where τ_X and τ_Y are generated by the families P and Q of seminorms respectively. Let $T : X \to Y$ be linear. Then the following statements 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 any $q \in Q$ there is a finite $F \subset P$ and $M > 0$ such that

$$q(Tx) < M \max_{p \in F} p(x) \qquad \text{for all} \quad x \in X\,. \tag{3.9}$$

Proof. $i) \Rightarrow ii)$: clear

$ii) \Rightarrow i)$: Since T is continuous at 0, for each neighbourhood $V \subset Y$ of 0 there exists a neighbourhood $U \subset X$ of 0 such that $T(U) \subset V$. But then for any $x \in X$ by the linearity of T we get $T(x+U) = T(x) + T(U) \subset T(x) + V$.

$ii) \Rightarrow iii)$: This follows from Lemma 3.8i), since the composition of continuous maps is continuous.

$iii) \Rightarrow iv)$: If $q \in Q$, then q is continuous by Lemma 3.8ii). Thus by iii) it follows that $q \circ T$ is a continuous seminorm on X. Lemma 3.8iii) implies that iv) holds.

$iv) \Rightarrow ii)$: Let $V \subset Y$ be a neighbourhood of 0. Without loss of generality we can assume that $V = \{y \in Y \,|\, q_i(y) < \epsilon\}$ for some ϵ, where $q_1, \ldots, q_n \in Q$. For each $i = 1, \ldots n$ choose M_i and F_i according to (2.95). If we set

$$F := \bigcup_{i=1}^{n} F_i \qquad \text{and} \qquad M := \max_{i=1,\ldots n} M_i\,,$$

it follows that

$$\max_{i=1,\ldots n} q_i(Tx) \leq M \max_{p \in F} p(x) \qquad \text{for all} \quad x \in X$$

and thus $T(U_{F,\epsilon/M}) \subset V$. □

In the special case of a linear functional, this implies:

Corollary 3.15
Let X be a locally convex space with topology generated by the family P of seminorms. A linear map $\ell : X \to \mathbb{K}$ is continuous if and only if there are finitely many seminorms $p_1, \dots, p_n \in P$ and $M > 0$ such that

$$|\ell(x)| < M \max_{i=1,\dots n} p_i(x)\,, \qquad \textit{for all} \quad x \in X\,.$$

Definition 3.16 (Dual space)
Let X, Y be locally convex spaces.

i) $\mathcal{L}(X, Y)$ denotes the set of all continuous linear maps from X to Y.

ii) $X' := \mathcal{L}(X, \mathbb{K})$ is called **topological dual space** *of X.*

It follows from Proposition 3.14 that X' and $\mathcal{L}(X, Y)$ are vector spaces.

3.1.4 Hahn-Banach and Separating Hyperplane

Theorem 3.17 (Hahn-Banach-Theorem)
Let X be a locally convex space, $U \subset X$ a subspace and $\ell \in U'$. Then there exists an extension $\Lambda \in X'$ of ℓ.

Proof. Let $P = \{p_\alpha\}_{\alpha \in A}$ be the family of seminorms generating the topology τ on X. Then the relative topology[6] on U is generated by the family $\{p_\alpha|_U\}_{\alpha \in A}$.

[6] The relative topology on U is given by $\tau_U = \{O \cap U \,|\, O \in \tau\}$.

By Corollary 3.15 there exists a continuous seminorm p on X (we can take $p := M \max_{i=1,\dots n} p_i$) such that

$$|\ell(x)| \leq p(x) \qquad \text{for all} \quad x \in U\,.$$

Thus it follows from the general versions of the Hahn-Banach-Theorem (Theorem 59 or Theorem 60) that there exists a linear extension extension $\Lambda : X \to \mathbb{K}$ of ℓ such that $|\Lambda(x)| \leq p(x)$ for all $x \in X$. Thus $\Lambda \in X'$ by Corollary 3.15. □

In the following we will use this theorem to show a geometric statement about a separating hyperplane.

Definition 3.18 (Hyperplane)

Let X be a locally convex space.

i) A **hyperplane** *is the set of points $x \in X$ where $\operatorname{Re} \ell(x) = a$ for some $\ell \in X'$ and $a \in \mathbb{R}$.*

ii) Two sets $A, B \subset X$ are said to be **separated by a hyperplane** $:\Longleftrightarrow$

$$\exists \ell \in X' \, \exists a \in \mathbb{R} \,:\, \forall x \in A \,:\, \operatorname{Re} \ell(x) \leq a \text{ and } \forall x \in B \,:\, \operatorname{Re} \ell(x) \geq a\,. \tag{3.10}$$

iii) If the inequalities in (3.10) are strict, A and B are called **strictly separated**.

Before we state the Theorem, we give two lemmata used in the proof.

Lemma 3.19

Let W be an open, absorbing, balanced, convex neighbourhood of 0 in a locally convex space X. Then the Minkowski functional p_W is a continuous seminorm.

Proof. By Lemma 3.6, p_W is a seminorm. The continuity follows from Lemma 3.8i) together with (3.7). □

THEOREM 3.20 (SEPARATING HYPERPLANE THEOREM)
Let X be a locally convex space and let A and B be disjoint convex sets in X.

i) If A is open and $0 \notin A$, there exists $\ell \in X'$ such that

$$\operatorname{Re} \ell(a) < 0 \qquad \textit{for all} \quad a \in A\,.$$

ii) If A is open, then A and B can be separated by a hyperplane.

iii) If A and B are open, then they can be strictly separated by a hyperplane.

iv) If A is compact and B is closed, they can be strictly separated by a hyperplane.

Proof. i): Let $x_0 \in A$, $y_0 = -x_0$ and $U = y_0 + A$. Then U is open and convex, $y_0 \notin U$ (since $y_0 - y_0 = 0 \notin A$) and $0 \in U$ (since $0 - y_0 = x_0 \in A$).

Consider the Minkowski functional p_U. Since U is open, there exists a convex, balanced and absorbing neighbourhood $W \subset U$ of 0. This shows that p_U is finite, since $p_U \leq p_W$ and W is absorbing. By Lemma 3.6, the convexity of U implies that p_U is sublinear[4]. Moreover $p_U(y_0) \geq 1$ by (3.7) since $y_0 \notin U$.

Set $Y = \{ty_0 \,|\, t \in \mathbb{R}\}$ and define the $\mathbb{R}$-linear functional $\lambda : Y \to \mathbb{R}$ by $\lambda(ty_0) = tp_U(y_0)$. Then

$$\lambda(y) \leq p_U(y) \qquad \text{for all} \quad y = ty_0 \in Y$$

because

$$t \leq 0 \;\Rightarrow\; \lambda(ty_0) \leq 0 \leq p_U(ty_0) \quad \text{and}$$
$$t > 0 \;\Rightarrow\; \lambda(ty_0) = tp_U(y_0) = p_U(ty_0)\,.$$

Thus by the Hahn-Banach-Theorem (real version) (Theorem 59), there exists an $\mathbb{R}$-linear functional $\Lambda : X \to \mathbb{R}$ such that

$$\Lambda(y) = \lambda(y) \text{ for all } y \in Y \quad \text{and} \quad \Lambda(x) \leq p_U(x) \text{ for all } x \in X\,.$$

Setting

$$\ell : X \to \mathbb{C}\,, \qquad \ell(x) := \Lambda(x) - i\Lambda(ix)\,,$$

it follows that ℓ is $\mathbb{C}$-linear[7] and

$$\operatorname{Re} \ell(x) = \Lambda(x) \leq p_U(x) \qquad \text{for all} \quad x \in X\,. \tag{3.11}$$

Since any $a \in A$ can be written as $a = u - y_0$ for some $u \in U$, it follows from $p_U(y_0) \geq 1$ and $p_U(u) < 1$ for all $u \in U$ (see (3.7)) that

$$\operatorname{Re} \ell(a) \leq p_U(a) \leq p_U(u) - p_U(y_0) \leq p_U(u) - 1 < 0\,.$$

The continuity of ℓ can be seen as follows: By (3.27) it follows that $\operatorname{Re} \ell(x) \leq p_W(x)$, since $W \subset U$. Since p_W is a seminorm, we have $p_W(\lambda x) = p_W(x)$ for all $|\lambda| = 1$. Thus choosing $\lambda \in \mathbb{C}, |\lambda| = 1$, so that $\lambda\ell(x) = |\ell(x)| \in \mathbb{R}$, it follows that

$$|\ell(x)| = \lambda\ell(x) = \ell(\lambda x) = \operatorname{Re} \ell(\lambda x) \leq p_w(\lambda x) = p_W(x)\,.$$

[7] For any $z = a + ib \in \mathbb{C}$ and $x \in X$, it follows from the $\mathbb{R}$-linearity of Λ that $\ell(zx) = \Lambda((a+ib)x) - i\Lambda(i(a+ib)x) = a\Lambda(x) + b\Lambda(ix) - ia\Lambda(ix) + ib\Lambda(x) = (a+ib)\Lambda(x) - i(a+ib)\Lambda(ix) = z\ell(x)$.

Thus by Proposition 3.14, ℓ is continuous and thus $\ell \in X'$.
ii): Since A and B are disjoint and A is open, the set $A - B$ is open and convex and does not contain 0. Thus by i) there exists $\ell \in X'$ such that

$$\operatorname{Re}\ell(x) = \operatorname{Re}\ell(a-b) = \operatorname{Re}\ell(a) - \operatorname{Re}\ell(b) < 0 \quad \text{for all} \quad x = a - b \in A - B\,. \tag{3.12}$$

This implies that for some $c \in \mathbb{R}$

$$\sup_{a\in A} \operatorname{Re}\ell(a) \le c \le \inf_{b\in B} \operatorname{Re}\ell(b)$$

and thus A and B are separated by the hyperplane $\ell(x) = c$.
iii) Let $\ell \in X'$ be such that (3.12) holds. Since ℓ is continuous and non-zero, it maps open sets to open sets (i.e. ℓ is open)[8]. Thus $\operatorname{Re}\ell(A)$ and $\operatorname{Re}\ell(B)$ are open in $\mathbb{R}$, proving the strict inequality.
iv) Let $\mathscr{U}$ denote the absolutely convex neighbourhood base at 0 generating the topology on X. We start showing that there exists $V \in \mathscr{U}$ such that

$$(A + V) \cap B = \emptyset\,. \tag{3.13}$$

If $A = \emptyset$, then $A + V = \emptyset$ and (3.13) is trivial. We therefore can assume that $A \neq \emptyset$. Consider $a \in A$, then $a \notin B$ and since B is closed there is $U \in \mathscr{U}$ such that $(a + U) \cap B = \emptyset$.

By property 3. of $\mathscr{U}$, there exists $V_a \in \mathscr{U}$ such that $V_a + V_a \subset U$ it follows that

$$(a + V_a + V_a) \cap B = \emptyset\,. \tag{3.14}$$

[8] In fact let $V \subset X$ be open and let $z \in \ell(V)$. Then there exists $v \in V$ such that $\ell(v) = z$. Since V is open, there exists $U \in \mathscr{U}$ such that $v + U \subset V$. Since U is a balanced neighbourhood of $0 \in X$, the linearity of ℓ implies that $\ell(U)$ is a balanced neighbourhood of $0 \in \mathbb{C}$, i.e. $\ell(U)$ contains some open ball K at 0 and $z + K \subset \ell(v + U) =\subset \ell(V)$

We assumed A to be compact, thus the open covering $A \subset \bigcup a \in A(a + V_a)$ has a finite subcovering, i.e. there are points $a_1, \dots, a_n \in A$ such that

$$A \subset \bigcup_{k=1}^{n} (a_k + V_{a_k}) .$$

Setting $V := \bigcap_{k=1}^{n} V_{a_k}$ it follows that

$$A + V \subset \bigcup_{k=1}^{n} (a_k + V_{a_k} + V) \bigcup_{k=1}^{n} (a_k + V_{a_k} + V_{a_k})$$

and since (3.14) holds for each a_k and V_{a_k}, this proves (3.13).

Since (3.13) holds, the sets B and $A + V$, which is open, are separated by a hyperplane by ii) given by some $\ell \in X'$. But $\ell(A)$ is a compact subset of $\ell(A + V$, because ℓ is continuous. This shows that A and B are strictly separated by a hyperplane. □

Corollary 3.21

If X is a locally convex space, then X' separates points, i.e. for any $x \neq y$ there exists some $\ell \in X'$ such that $\ell(x) \neq \ell(y)$.

Proof. Apply Theorem 3.20 iii) with $A = \{x\}$ and $B = \{y\}$. □

Corollary 3.22

Suppose M is a subspace of a locally convex space X and $x_0 \in X$. If x_0 is not in the closure of M, then there exists $\ell \in X'$ such that $\ell(x_0) = 1$ and $\ell(m) = 0$ for every $m \in M$.

Proof. Apply Theorem 3.20 iii) to $A = \{x_0\}$ and $B = \overline{M}$. Then there exists $\ell' \in X'$ such that $\operatorname{Re} \ell'(x_0)$ and $\operatorname{Re} \ell(M)$ are disjoint. Thus $\operatorname{Re} \ell'(M)$ is a proper subset of $\mathbb{R}$. Since ℓ' is linear and M is a subspace, this implies $\ell'(M) = 0$. The functional ℓ is then given by $\ell'/\ell'(x_0)$. □

3.1.5 Weak topologies

In Example 1viii) and ix) we introduced the weak and the weak-* topologies on a normed vector space X. More general, we can consider two vector spaces X and Y and a bilinear map $B : X \times Y \to \mathbb{K}$ (i.e. $B_x : Y \to \mathbb{K}$ given by $B_x(y) = B(x, y)$ and $B_y : X \to \mathbb{K}$ given by $B_y(x) = B(x, y)$ are both linear).

Definition 3.23 (Dual pair)
For vector spaces X and Y and a bilinear map $B : X \times Y \to \mathbb{K}$, the tuple (X, Y, B) is called **dual pair** *with respect to $B :\Longleftrightarrow$*

$$\begin{aligned} &\forall x \in X \setminus \{0\}\ \exists y \in Y\ : B(x, y) \neq 0 \quad \textit{and} \\ &\forall y \in Y \setminus \{0\}\ \exists x \in X\ : B(x, y) \neq 0\,. \end{aligned}$$

For a given pair (X, Y, B) of vector spaces, the bilinear map B can often be chosen in a canonical way. Moreover, the spaces X (and Y) can be identified with a subspace ot the dual space of Y (and X respectively) separating points, because the mappings $x \mapsto B(x, \cdot)$ and $y \mapsto B(\cdot, y)$ are injective.

Example 3.24 *(a) Let X be a locally convex space and X' the dual space, then it follows from Corollary 3.21 that (X, X', B) is a dual pair with respect to the canonical bilinear map $B : X \times X' \to \mathbb{K}$ given by $B(x, \ell) = \ell(x)$. Similarly $(X', X, \tilde{B})$ is a dual pair, where $\tilde{B} : X' \times X \to \mathbb{K}$ is given by $\tilde{B}(\ell, x) = \ell(x)$.*

(b) The set X of bounded continuous functions on $\mathbb{R}$ and the set Y of regular finite signed or complex Borel-measures on $\mathscr{B}(\mathbb{R})$ are a dual pair with respect to $B(f, \mu) = \int_{\mathbb{R}} f\, d\mu$.

(c) *Let* $X = \mathbb{R}^{\mathbb{R}}$ *be the set of functions* $f : \mathbb{R} \to \mathbb{R}$ *and denote by* $\delta_t : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}$ *the pointwise evaluation, i.e.* $\delta_t(f) = f(t)$*. Let* $Y = \operatorname{span}\{\delta_t \,|\, t \in \mathbb{R}\}$*, then* (X, Y, B) *is a dual pair for*

$$B\Big(f, \sum_{j=1}^{n} \lambda_j \delta_{t_j}\Big) = \sum_{j=1}^{n} \lambda_j f(t_j)\,.$$

Definition 3.25

Let (X, Y, B) *be a dual pair. The* $\sigma(X, Y)$ *(or* $\sigma(Y, X)$*) topology on* X *(or* Y *respectively) is the locally convex topology generated by the family of seminorms* $P = \{p_y\}_{y \in Y}$ *where* $p_y(x) = |B(x, y)|$ *(or* $P = \{p_x\}_{x \in X}$ *where* $p_x(y) = |B(x, y)|$ *respectively).*

Since by Definition 3.23 the family of seminorms P separates points (in both cases), if follows from Lemma 3.5 that the $\sigma(X, Y)$ topology for a dual pair is always Hausdorff.

In Example 3.24(a), the $\sigma(X, X')$ topology is the weak topology and the $\sigma(X', X)$ topology is the weak-* topology. In (c) $\sigma(X, Y)$ is the topology of pointwise convergence given in Example 1i).

In order to determine the topological dual space of $(X, \sigma(X, Y))$, we need the following lemma.

Lemma 3.26

Let X *be a vector space and* $\ell, \ell_1, \ldots, \ell_n : X \to \mathbb{K}$ *linear. Set*

$$N = \{x \in X \,|\, \forall\, j = 1, \ldots n \;:\; \ell_j(x) = 0\}\,.$$

Then the following statements are equivalent.

i) $\ell \in \operatorname{span}\{\ell_j \,|\, j \in \{1, \ldots n\}\}$.

ii) $\exists\, M \geq 0\, \forall\, x \in X\ :\ |\ell(x)| \leq M \max_{1\leq j\leq n} |\ell_j(x)|$.

iii) $\ell = 0$ *for all* $x \in N$*, i.e.* $\bigcap_{j=1}^n \mathrm{Ker}(\ell_j) \subset \mathrm{Ker}(\ell)$.

Proof. The implications $i) \Rightarrow ii) \Rightarrow iii)$ are clear.
$iii) \Rightarrow i)$: Let

$$V = \{\vec{\ell} = (\ell_1(x), \dots \ell_n(x)) \in \mathbb{K}^n \,|\, x \in X\}\,,$$

then by iii) the map $\phi : V \to \mathbb{K}$ given by $\phi\big(\vec{\ell}(x))\big) = \ell(x)$ is well defined and linear on V. Thus there exists a linear extension $\tilde{\phi} : \mathbb{K}^n \to \mathbb{K}$ with $\tilde{\phi}(\vec{\xi}) = \sum_{j=1}^n \alpha_j \xi_j$ for some $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Thus $\ell = \sum_{j=1}^n \alpha_j \ell_j \in \mathrm{span}\{\ell_1, \dots \ell_n\}$.
□

Corollary 3.27
A functional ℓ *on* $(X, \sigma(X,Y))$ *is continuous, if and only if* $\ell(x) = B(x,y)$ *for some* $y \in Y$*, i.e. the topological dual space of* X *with the* $\sigma(X,Y)$ *topology is equal to* Y.

Proof. Exercise 2.100 □

3.1.6 Fréchet spaces

Under some additional assumptions, locally convex spaces are metrizable, i.e. there is a metric on X which is compatible with the topology, i.e. such that the balls of radius $1/n$ at 0 build a neighbourhood base. Here it is not necessary that the topologies are given by a norm.

Theorem 3.28
Let (X, τ) *be a locally convex space. Then the following statements are equivalent.*

i) X is metrizable.

ii) 0 has a countable neighbourhood base.

iii) τ is generated by a countable family of seminorms.

Proof. $i) \Rightarrow ii)$: If there is a metric on X which is compatible with τ, then the balls with radius $\frac{1}{n}$ centred at 0 form a countable local neighbourhood base at 0.
$ii) \Rightarrow iii)$: If 0 has a countable neighbourhood base, there is a neighbourhood base of open, absorbing, balanced convex sets. Thus by Lemma 3.19, the associated family of Minkowski functionals is a countable family of continuous seminorms generating τ (see the proof of Proposition 3.7).
$iii) \Rightarrow i)$: Let $\{p_n\}_{n\in\mathbb{N}}$ be a countable family of seminorms generating τ. Set ρ on $X \times X$ by

$$\rho(x,y) := \sum_{n\in\mathbb{N}} 2^{-n} \frac{p_n(x-y)}{1+p_n(x-y)} \,. \tag{3.15}$$

Then $\rho(x,y) < \infty$ and ρ is a translation invariant metric (compare Exercise 3.35). Moreover, the topology $\tilde{\tau}$ for which the open balls of radius $1/n$ at 0 are a neighbourhood base, is equal to τ. In fact, let

$$\begin{aligned} U_{M,1/n} &= \{x \in X \,|\, \forall\, n \leq M \in \mathbb{N} : \ p_n(x) < 1/n\,\} \subset \mathscr{U} \quad \text{and} \\ B_{1/n} &= \{x \in X \,|\, \rho(0,x) < 1/n\,\}\,. \end{aligned}$$

Remark that for any $M \in \mathbb{N}$

$$\frac{p_1(x)}{2(1+p_1(x))} \leq \rho(0,x) \leq \sum_{k=1}^{M} 2^{-k} \frac{p_k(x)}{1+p_k(x)} + \sum_{k\geq M+1} 2^{-k} \leq 2 \max_{k\leq M} p_k(x) + 2^{-M}\,.$$

Since $2^{-n} \leq 1/n$ for all $n \in \mathbb{N}$, this implies that $U_{n,1/(4n)} \subset B_{1/n}$ and that $B_{1/n+2} \subset U_{1,1/n}$ for any $n \in \mathbb{N}$. Thus $\tau = \tilde{\tau}$. □

Proposition 3.29

Let (X, τ) be a locally convex space generated by a countable family of seminorms $P = \{p_n\}_{n\in\mathbb{N}}$. Let ρ be the metric given by (3.15). Then a net $(x_\alpha)_{\alpha\in I}$ is Cauchy with respect to τ if and only if it is Cauchy with respect to ρ.

Proof. $\Rightarrow$: Assume that $(x_\alpha)_{\alpha\in I}$ is Cauchy with respect to τ. Let $\epsilon > 0$, then by Definition 3.13 there is $\beta \in I$ such that

$$p_n(x_\gamma - x_\delta) < \frac{\epsilon}{2} \qquad \text{for all} \quad \gamma, \delta \succ \beta,\ n \in \mathbb{N}\,.$$

By (3.15) it follows that $\rho(x_\gamma, x_\delta) < \epsilon$.

$\Leftarrow$: Assume that there exists some $\epsilon > 0$ such that for any $\beta \in I$ there exist $\gamma, \delta \succ \beta$ and $n \in \mathbb{N}$ such that $p(x_\gamma - x_\delta) > \epsilon$. Then $\rho(x_\gamma, x_\delta) > 2^{-n}\frac{\epsilon}{1+\epsilon}$. ☐

Definition 3.30 (Fréchet space)

A complete metrisable locally convex space is called a **Fréchet space**.

Since any Fréchet space is a complete metric space, it obeys the Baire category theorem (Theorem 51), which allows to derive analogues of some of its consequences. In particular, the following analogues to the principle of uniform boundedness (Banach-Steinhaus-Theorem 52) and to the open mapping theorem (Theorem 54) on Banach spaces hold.

Theorem 3.31 (Principle of uniform boundedness)

Let X, Y be Fréchet spaces and let $\mathcal{F}$ be a family of continuous linear maps from X to Y. Assume that for each continuous seminorm q in Y and for every $x \in X$ the set $\{q(Tx) \,|\, T \in \mathcal{F}\}$ is bounded. Then for each continuous seminorm q in Y there is a continuous seminorm p in X and a constant $C > 0$ so that

$$q(Tx) \le Cp(x) \qquad \textit{for all} \quad x \in X\,,\ T \in \mathcal{F}\,.$$

THEOREM 3.32 (OPEN MAPPING AND INVERSE MAPPING THEOREM ON FRÉCHET SPACES)
Let X, Y be Fréchet spaces and $T : X \to Y$ a linear continuous map.

i) If T is surjective, then T is open.

ii) If T is bijective, then T has a continuous inverse.

For a proof of these theorems see e.g. [R].

EXAMPLE 3.33 *i) The most important example for a Fréchet space is the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ of functions decreasing to 0 as $|x| \to \infty$ faster than any polynomial (see Exercise 3.35).*

ii) Another example of a Fréchet space is $\mathscr{D}_K(\Omega)$ of infinitely differentiable functions, supported in a compact set $K \subset \Omega \subset \mathbb{R}^n$ introduced in Example 1 v).

3.1.7 Exercises

EXERCISE 3.34 (WEAK TOPOLGY)
Let (X, Y, B) be a dual pair. Show that a functional ℓ on $(X, \sigma(X, Y))$ is continuous, if and only if $\ell(x) = B(x, y)$ for some $y \in Y$.

EXERCISE 3.35 (SCHWARTZ SPACE)
The **Schwartz space** *$\mathscr{S}(\mathbb{R}^n)$ is the set of infinitely differentiable complex-valued functions f on $\mathbb{R}^n$ which are rapidly decreasing, i.e. such that*

$$\|f\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} \left|x^\alpha D^\beta f(x)\right| < \infty \qquad \text{for all} \quad \alpha, \beta \in \mathbb{N}^n .$$

For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ *we set* $|\alpha| = \sum_i \alpha_i$,

$$x^\alpha := \prod_{j=1}^{n} x_j^{\alpha_j} \qquad \text{and} \qquad D^\beta := \frac{\partial^{|\beta|}}{\partial^{\alpha_1} x_1 \dots \partial^{\alpha_n} x_n}\,.$$

Show the following statements.

i) $\|\cdot\|_{\alpha,\beta}$ *is a seminorm for any* $\alpha, \beta \in \mathbb{N}^n$.

ii) The family $(\|\cdot\|_{\alpha,\beta})$ *of seminorms can be used to define a metric* d *on* $\mathscr{S}$.

iii) The metric space $(\mathscr{S}, d)$ *is complete.*

Rem.: This shows that $\mathscr{S}$ *is a Fréchet space.*

3.2 Generalized functions or Distributions

The theory of distributions allows to enlarge the set of functions, on which several operators, in particular differential operators can be defined. Then e.g. the derivative of functions, which are not differentiable, or even of more general objects can be defined in a reasonable way.

In order to be useful, such an extension should have the following properties:

- it should include all continuous functions.
- on the subset of differentiable functions the definition of the derivative should coincide with the usual definition.
- the usual formal rules of calculus should hold.

- there should be convergence theorems allowing to handle limit processes.

The main idea in the theory of distributions (or generalized functions) is to identify functions f with linear functionals $\ell_f : X \to \mathbb{K}$ on a suitable function space X via

$$\ell_f(\phi) := \int f(x)\phi(x)\, dx\,. \tag{3.16}$$

Then the derivative of the function f (or the functional ℓ_f) is defined by integrating by parts, if ϕ is differentiable.

The set of functions, such that the integral in (3.16) exists, depends on the choice of X. If X includes only compactly supported functions, there are no restrictions on the behaviour of f at infinity.

Of course, not for every functional ℓ exists some f such that $\ell = \ell_f$ given by (3.16). However, even for the δ_x-functional, given by $\delta_x(\phi) = \phi(x)$, it is common in physics to formally calculate with some δ_x-function, which is assumed to satisfy $\phi(x) = \int_{\mathbb{R}} \delta_x(y)\phi(y)\, dy$. Nevertheless, the definition of the derivative of a functional by use of integration by parts can be extended to larger class of functionals than those given by (3.16) for some f.

In order to be able to treat as many distributions as possible, the space X' should be large. Thus X should be small (the integral in (3.16) should exist for many functions f if $\phi \in X$) and its topology should be fine (then it is easier for a functional to be continuous).

3.2.1 Definition of distribution spaces

There are three choices of (complete) locally convex function spaces on an open set $\Omega \subset \mathbb{R}^n$ or on $\mathbb{R}^n$ commonly used in this context:

i) The largest space is the space $\mathscr{E}(\Omega)$ of infinitely differentiable functions on Ω defined in Example 1 iii).

ii) The Schwartz space $\mathscr{S}(\mathbb{R}^n)$ defined in Footnote 2.66 and Example 1 iv).

iii) The smallest space is $\mathscr{D}(\Omega)$ defined in Example 1 vi) (with elements in $\mathcal{C}_0^\infty(\mathbb{R}^n)$).

Definition 3.36 (Distribution spaces)
The elements of the space $\mathscr{D}(\Omega)$ are called **test functions**. *The elements of its topological dual space $\mathscr{D}'(\Omega)$ are called* **distributions**. *The topological dual space $\mathscr{S}'(\mathbb{R}^n)$ of the Schwartz space is called the space of* **tempered distributions**. *The elements of $\mathscr{E}'(\Omega)$ are called* **distributions of compact support**. *A distribution ℓ, for which there exists a functions f such such $\ell = \ell_f$ as defined in (3.16) is called* **regular distribution**.

We start, discussing the topology on the space $\mathscr{D}(\Omega)$ of test functions in more detail. Explicitly, we have

$$\mathscr{D}(\Omega) = \bigcup_{K \subset \Omega \,\text{compact}} \mathscr{D}_K(\Omega)\,, \tag{3.17}$$

where

$$\mathscr{D}_K(\Omega) := \{ f \in \mathcal{C}^\infty(\Omega) \mid \operatorname{supp} f \subset K \} \quad \text{and} \quad p_{\alpha,K}(f) = \sup_{x\in\Omega} |D^\alpha f(x)|\,. \tag{3.18}$$

Then the spaces $\mathscr{D}_K(\Omega)$ are locally convex spaces with topology τ_K generated by $P_K = \{p_{\alpha,K}\}_{\alpha\in\mathbb{N}^n}$ or equivalently by the seminorms

$$p_m(f) := \sup_{|\alpha|\leq m} p_\alpha(f)\,.$$

The topology τ on $\mathscr{D}(\Omega)$ is generated by be the family P of all seminorms p on $\mathscr{D}(\Omega)$ such that all restrictions $p|_{\mathscr{D}_K}$ are continuous with respect to τ_K, i.e. such that

$$\forall\, K \subset \Omega \text{ compact } \exists\, c, m \geq 0\; \forall\, \phi \in \mathscr{D}_K(\Omega) \;:\; p(\phi) \leq c\, p_m(\phi)\,. \tag{3.19}$$

3.2.2 Properties of $\mathscr{D}(\Omega)$

Lemma 3.37

For an open set $\Omega \subset \mathbb{R}^n$ let $\mathscr{D}(\Omega)$ be the locally convex space defined in (3.17)*. Then the following holds.*

i) The relative topology of τ on $\mathscr{D}_K(\Omega)$ is equal to τ_K.

ii) $\mathscr{D}_K(\Omega)$ is τ-closed in $\mathscr{D}(\Omega)$.

iii) $\mathscr{D}(\Omega)$ is a Hausdorff space.

iv) Let Y is a locally convex space and $L : \mathscr{D}(\Omega) \to Y$ a linear map. L is τ-continuous if and only if for all compact $K \subset \Omega$ the restrictions $L|_{\mathscr{D}_K(\Omega)}$ are τ_K-continuous.

Proof. *i*) : The relative topology on $\mathscr{D}_K(\Omega)$ is generated by the family Q_K given by the restrictions $p|_{\mathscr{D}_K(\Omega)}$ of seminorms $p \in P$, which by definition are continuous with respect to τ_K. Since on the other hand, the family P_K is a

subset of Q_K by Lemma 3.8, the assertion follows from Corollary 3.10.
ii) : For $x \in \Omega$, consider the seminorm $q_x(\phi) := |\phi(x)|$. Then for all $K \subset \Omega$ compact and $\phi \in \mathscr{D}_K(\Omega)$ we have

$$q_x(\phi) \leq p_{0,K}(\phi) \qquad \text{for all} \quad x \in \Omega$$

and thus $q_x \in P$ by (3.19). This implies by Lemma 3.8 that q_x is continuous with respect to τ. Since for any $K \subset \Omega$ compact, we have

$$x \notin K \quad \to \quad \forall \phi \in \mathscr{D}_K(\Omega) \; : \; q_x(\phi) = 0$$

it follows that

$$\mathscr{D}_K(\Omega) \subset \bigcap_{x \in K^c} \operatorname{Ker} q_x \, .$$

Since on the other hand

$$\Big(\forall\, x \in K^c \; : \; q_x(\phi) = 0\Big) \;\Rightarrow\; \Big(\forall\, x \in K^c \; : \; |\phi(x)| = 0\Big) \;\Rightarrow\; \phi \in \mathscr{D}_K(\Omega)$$

we have

$$\mathscr{D}_K(\Omega) = \bigcap_{x \in K^c} \operatorname{Ker} q_x$$

and since intersections of closed sets are closed, it follows that $\mathscr{D}_K(\Omega)$ is closed (with respect to τ).
iii) : If $\phi \neq 0$, then $\phi(x) \neq 0$ for some $x \in \Omega$ and thus $q_x(\phi) \neq 0$. Since $q_x \in P$, it follows from Lemma 3.5 that $(\mathscr{D}(\Omega), \tau)$ is a Hausdorff space.
iv) $\Rightarrow$: If $L : \mathscr{D}(\Omega) \to Y$ is τ-continuous, then it follows from i) that all restrictions $L|_{\mathscr{D}_K(\Omega)}$ are τ_K-continuous.
$\Leftarrow$: Let q be any continuous seminorm in Y. Then by Proposition 3.14 and the assumption, $q \circ L|_{\mathscr{D}_K(\Omega)}$ is a continuous seminorm on $\mathscr{D}_K(\Omega)$ for all compact sets $K \subset \Omega$. By (3.19) this implies $q \circ L \in P$ and in particular

$q \circ L$ is τ-continuous by Lemma 3.5. Thus again by Proposition 3.14 we can conclude that L is τ-continuous. $\square$

The following proposition is about convergence of sequences in $\mathscr{D}(\Omega)$.

Proposition 3.38

Let $(\phi_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{D}(\Omega)$. Then the following statements are equivalent.

i) $\phi_n \to 0$ as $n \to \infty$ with respect to τ.

ii) There exists a compact set $K \subset \Omega$ such that $\phi_n \in \mathscr{D}_K(\Omega)$ for all $n \in \mathbb{N}$ and $\phi_n \to 0$ as $n \to \infty$ with respect to τ_K.

iii) There exists a compact set $K \subset \Omega$ such that $\operatorname{supp} \phi_n \subset K$ for all $n \in \mathbb{N}$ and for all $\alpha \in \mathbb{N}^n$ the sequence $(D^\alpha \phi_n)_{n\in\mathbb{N}}$ converges uniformly to 0.

Proof. By the construction of $\mathscr{D}(\Omega)$ and the definition of the seminorms in (3.17) and (3.18), the statement in *iii*) is only another way to formulate *ii*).

The implication *ii*) $\Rightarrow$ *i*) is clear by (3.19). Thus it remains to prove *i*) $\Rightarrow$ *ii*):

Assume that $\phi_n \to 0$ with respect to τ. If there exists some K such that $\phi_n \in \mathscr{D}_K(\Omega)$ for all $n \in \mathbb{N}$, then the statement follows from Lemma 3.37 i). We prove by contradiction.

Assume that there is no such compact set $K \subset \Omega$. Then there exists a sequence of compact sets $K_1 \subset K_2 \subset K_3 \ldots \subset \Omega$ with

$$\Omega = \bigcup_{n\in\mathbb{N}} \mathring{K}_n \tag{3.20}$$

and a subsequence (ψ_n) such that

$$\psi_n \in \mathscr{D}_{K_n}(\Omega) \quad \text{but} \quad \psi \notin \mathscr{D}_{K_{n-1}}(\Omega)\,.$$

For each $n \in \mathbb{N}$ choose $x_n \in K_n \setminus K_{n-1}$ such that $a_n := |\psi_n(x_n)| > 0$. It follows from the proof of Lemma 3.37 that the seminorms

$$q_n : \mathscr{D}(\Omega) \to [0, \infty)\,, \quad \phi \mapsto a_n^{-1}|\phi(x_n)|$$

are τ-continuous. Set $q := \sum_{n=1}^{\infty} q_n$. Since by (3.20) each compact set $K \subset \Omega$ is a subset of some K_n, it follows from (3.17) that $\mathscr{D}(\Omega) = \bigcup_{n\in\mathbb{N}} \mathscr{D}_{K_n}(\Omega)$. Since $q_n|_{\mathscr{D}_{K_m}(\Omega)} = 0$ for $n > m$, it follows that $q(\phi)$ is a finite sum for any $\phi \in \mathscr{D}(\Omega)$. Thus the seminorm q is well-defined and

$$q|_{\mathscr{D}_K(\Omega)} = \sum_{n=1}^{N} q_n|_{\mathscr{D}_K(\Omega)}\,, \quad \text{if} \quad K \subset K_N\,.$$

Thus $q \in P$ and τ-continuous. This implies that $q(\psi_n) \to 0$ as $n \to \infty$, since $\psi_n \to 0$ with respect to τ by assumption. But on the other hand, it follows from the definition of q and q_n that $q(\psi_n) \geq q_n(\psi_n) = 1$ for all $n \in \mathbb{N}$ in contradiction to the convergence property. This shows ii).

□

The construction of $\mathscr{D}(\Omega)$ is an example of a strict inductive limit of Fréchet spaces.

Definition 3.39 (Strict inductive limit topology)
Let X be a vector space, $\{X_n\}_{n\in\mathbb{N}}$ a family of subspaces such that $X_n \subset X_{n+1}$ for all $n \in \mathbb{N}$ and $X = \bigcup_n X_n$. Suppose that each X_n has a locally convex topology τ_n and $\tau_{n+1}|_{X_n} = \tau_n$. Let $\mathscr{U}$ be the collection of convex, absorbing,

balanced sets $U \subset X$ *for which* $U \cap X_n \in \tau_n$ *for each* $n \in \mathbb{N}$ *and let* τ *be the collection of all unions of sets of the form* $x + U$ *with* $x \in X$ *and* $U \in \mathscr{U}$*. Then* (X, τ) *is (a locally convex space) called* **strict inductive limit** *of the spaces* X_n*.*

It can be shown that the inductive limit of complete locally convex spaces is complete.

3.2.3 Characterisation of distributions and examples

By the above results, distributions, i.e. elements of the topological dual space $\mathscr{D}'(\Omega)$ of $\mathscr{D}(\Omega)$, can be characterised as follows.

Proposition 3.40

Let $\ell : \mathscr{D}(\Omega) \to \mathbb{C}$ *be a linear map. Then the following statements are equivalent.*

i) ℓ *is a distribution, i.e.* $\ell \in \mathscr{D}'(\Omega)$*.*

ii) $\ell|_{\mathscr{D}_K(\Omega)} \in \mathscr{D}'_K(\Omega)$ *for all compact* $K \subset \Omega$*.*

iii) For all compact $K \subset \Omega$ *there exist* $m \in \mathbb{N}_0$ *and* $c \geq 0$ *such that*

$$|\ell(\phi)| \leq c p_m(\phi) = c \sup_{|\alpha| \leq m} \sup_{x \in \Omega} |D^\alpha \phi(x)| \,, \qquad \textit{for all} \quad \phi \in \mathscr{D}_K(\Omega)\,.$$

iv) If $\phi_n \to 0$ *in* $\mathscr{D}(\Omega)$*, then* $\ell(\phi_n) \to 0$ *in* $\mathbb{C}$*.*

Proof. *i)* $\Leftrightarrow$ *ii)*: This is a special case of Lemma 3.37 iv).
ii) $\Leftrightarrow$ *iii)*: This follows from Corollary 3.15.

i) $\Rightarrow$ *iv)*: This holds by definition.

iv) $\Rightarrow$ *ii)*: Since τ_K is generated by a countable family of seminorms, $\mathscr{D}_K(\Omega)$ is metrizable by Theorem 3.28. iv) implies that $\ell(\phi_n) \to 0$ whenever $\phi_n \to 0$ in $\mathscr{D}_K(\Omega)$ for any compact $K \subset \Omega$. But on metric spaces, this is equivalent to the continuity of $\ell|_{\mathscr{D}_K(\Omega)}$ at 0 and by Proposition 3.14 it follows that $\ell|_{\mathscr{D}_K(\Omega)} \in \mathscr{D}'_K(\Omega)$. □

EXAMPLE 3.41 *i) Let $g \in \mathscr{S}(\mathbb{R})$ and define the functional ℓ_g on $\mathscr{S}$ by*

$$(g, \phi) := \ell_g(\phi) = \int_{\mathbb{R}} g(x)\phi(x)\, dx \qquad \text{for all} \quad \phi \in \mathscr{S}(\mathbb{R})\,. \tag{3.21}$$

Then ℓ_g is a linear functional and $\ell_g \in \mathscr{S}'(\mathbb{R})$ since

$$|\ell_g(\phi)| \leq \|g\|_{L^1}\, \|\phi\|_{0,0}\,.$$

Moreover, if $g_1 \neq g_2$ in $\mathscr{S}(\mathbb{R})$, then $g_1 \neq g_2$ in $\mathscr{S}'(\mathbb{R})$. Thus $\mathscr{S}(\mathbb{R})$ is naturally embedded in $\mathscr{S}'(\mathbb{R})$, i.e. each Schwartz functions can be identified with a regular and tempered distribution.

ii) $\mathscr{S}(\mathbb{R})$ is a subset of each $L^p(\mathbb{R})$ and the identity mapping of $\mathscr{S}$ into L^p is continuous. This can be seen as follows: For $p = 1$, we write for any $\phi \in \mathscr{S}(\mathbb{R})$

$$\begin{aligned}\|\phi\|_{L^1} &= \int_{-\infty}^{\infty} \frac{1}{1+x^2}\big((1+x^2)|\phi(x)|\big)\, dx \\ &\leq \big(\|\phi\|_{0,0} + \|\phi\|_{2,0}\big) \int_{-\infty}^{\infty} \frac{1}{1+x^2}\, dx = \pi\big(\|\phi\|_{0,0} + \|\phi\|_{2,0}\big)\,.\end{aligned}$$

For general p, notice that with $\frac{1}{p} + \frac{1}{q} = 1$ we can write

$$\|\phi\|_{L^p} \leq \||\phi|^{1/p}|\phi|^{1/q}\|_{L^p} \leq \|\phi\|_{L^1}^{1/p}\, \|\phi\|_{0,0}^{1/q}\,.$$

Now for $g \in L^q(\mathbb{R})$ *let* ℓ_g *be defined as in* (3.21), *then for any* $\phi \in \mathscr{S}(\mathbb{R})$ *by Hölder's inequality*

$$|\ell_g(\phi)| \le \|g\|_{L^q} \|\phi\|_{L^p} \,,$$

thus $\ell_g \in \mathscr{S}'(\mathbb{R})$. *This defines a continuous embedding of* $L^q(\mathbb{R})$ *in* $\mathscr{S}'(\mathbb{R})$. *Thus each* $g \in L^q(\mathbb{R})$ *can be seen as a regular tempered distribution.*

iii) Let $g \in L^1_{\mathrm{loc}}(\Omega)$[9] *and define* ℓ_g *on* $\mathscr{D}(\Omega)$ *by*

$$(g,\phi) := \ell_g(\phi) = \int_{\mathbb{R}} g(x)\phi(x)\,dx \qquad \textit{for all} \quad \phi \in \mathscr{D}(\Omega)\,. \tag{3.22}$$

Then ℓ_g *is linear and well-defined, since* g *is locally integrable. Since for any* $K \subset \Omega$ *compact,*

$$|\ell_g(\phi)| \le \int_K |g(x)|\,dx\, p_0(\phi) \qquad \textit{for all} \quad \phi \in \mathscr{D}_K(\Omega)$$

it follows that $\ell_g \in \mathscr{D}'(\Omega)$. ℓ_g *is the regular distribution associated to* g. *Since the mapping* $g \mapsto \ell_g$ *is injective*[10], *the function* $g \in L^1_{\mathrm{loc}}(\Omega)$ *can be identified with the distribution* $\ell_g \in \mathscr{D}'(\Omega)$.

iv) For some $a \in \Omega$ *fixed, define the functional* $\delta_a(\phi) := \phi(a)$. *Since*

$$|\delta_a(\phi)| = |\phi(a)| \le p_0(\phi) \qquad \textit{for any} \quad \phi \in \mathscr{D}_K(\Omega)\,,$$

[9] A measurable function $f : \Omega \to \mathbb{C}$ is called locally integrable, if $\int_K |f(x)|\,dx$ is finite for any compact $K \subset \Omega$. Two locally integrable functions are said to be equivalent if they are equal almost everywhere. Then $L^1_{\mathrm{loc}}(\Omega)$ is the space of all equivalence classes of locally integrable functions.

[10] Let $g \in L^1_{\mathrm{loc}}(\Omega)$ and ℓ_g as given in (3.22). If $\ell_g(\phi) = 0$ for all test functions ϕ, then $g = 0$ almost everywhere.

δ_a is continuous on all $\mathscr{D}_K(\Omega)$, thus δ is a distribution (i.e. $\delta_a \in \mathscr{D}'(\Omega)$) by Proposition 3.40. For $a = 0$, we usually write $\delta_0 = \delta$.

Moreover, for $\phi \in \mathscr{S}(\mathbb{R}^n)$ we have $|\delta_a(\phi)| \leq \|\phi\|_{0,0}$, thus δ_a is a tempered distribution (i.e. $\delta_a \in \mathscr{S}'(\mathbb{R}^n)$).

But δ_a is no regular distribution. This can be seen as follows: Set $\Omega_a := \Omega \setminus \{a\}$, then $\delta|_{\mathscr{D}(\Omega_a)} = 0$. If there would be some function f such that $\delta_a = \ell_f$ as given in (3.21), it would follow from the injectivity of $f \mapsto \ell_f$ that $f|_{\Omega_a} = 0$ almost everywhere and thus $f = 0$ almost everywhere. But this would imply $\delta_a = 0$, which is not the case.

v) Suppose that μ is a Borel measure on Ω and define the linear functional

$$\ell_\mu(\phi) := \int_\Omega \phi \, d\mu \, .$$

If $|\mu|(K) < \infty$ for all compact sets $K \subset \Omega$, then ℓ_μ is a distribution. If μ is finite and $\Omega = \mathbb{R}^n$, it is a tempered distribution (here it suffices if μ can be estimated by some polynomial). The map $\mu \mapsto \ell_\mu$ is injective.

3.2.4 Adjoint operator and derivative of distributions

Before we can define the derivative of a distribution, we need the following definition.

Definition 3.42 (Adjoint operator)
Let X, Y be locally convex spaces and $L \in \mathcal{L}(X, Y)$. Then the linear map

$$L' : Y' \to X' \, , \qquad y' \mapsto L'(y') = y' \circ L$$

is called adjoint of L.

As already explained above, the idea is to extend the derivative of regular distributions, which can be computed using integration by parts, to general distributions. If $f \in \mathbb{C}^1(\mathbb{R})$, then

$$\ell_{f'}(\phi) = \int_{\mathbb{R}} f'(x)\phi(x)\,dx = -\int_{\mathbb{R}} f(x)\phi'(x)\,dx = -\ell_f(\phi') \quad \text{for all} \quad \phi \in \mathscr{D}(\mathbb{R})$$

and analogue for higher derivatives and partial derivatives.

Definition 3.43
For $\ell \in \mathscr{D}'(\Omega)$ and $\alpha \in \mathbb{N}^n$ set

$$\big(\tilde{D}^\alpha \ell\big)(\phi) = (-1)^{|\alpha|}\ell\big(D^\alpha \phi\big)\,.$$

If $\ell = \ell_f$ for some $f \in L^1_{\text{loc}}(\Omega)$, then $\tilde{D}^\alpha \ell_f$ is called αth **distribution derivative** *or* **weak derivative** *of f.*

Lemma 3.44
The map $\tilde{D}^\alpha : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$ given in Definition 3.43 is well-defined and $\sigma(\mathscr{D}', \mathscr{D})$-continuous.

Proof. By definition, $\tilde{D}^\alpha = (-1)^{|\alpha|}\big(D^\alpha\big)'$, i.e. the adjoint operator except for the sign. Thus it suffices to verify that $D^\alpha : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is continuous. By Lemma 3.37, it suffices to show that $D^\alpha : \mathscr{D}_K(\Omega) \to \mathscr{D}_K(\Omega)$ is continuous for all $K \subset \Omega$ compact, but since

$$p_m(D^\alpha \phi) \le p_{m+|\alpha|}(\phi)\,,$$

this holds. □

Example 3.45 *i) If $f \in \mathcal{C}^k(\Omega)$ and $|\alpha| \le k$, then $\tilde{D}^\alpha \ell_f = \ell_{D^\alpha f}$, i.e. the operator $\tilde{D}^\alpha$ extends the classical differential operator D^α.*

ii) The derivative ot the δ-distribution $\delta_a(\phi) = \phi(a)$ is given by

$$\delta_a'(\phi) = -\delta_a(\phi') = -\phi'(a)\,.$$

iii) The Heavyside function $H : \mathbb{R} \to \mathbb{C}$ is the characteristic function of $[0, \infty)$, i.e. $H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ if $x < 0$. Then H is not differentiable in the usual sense, but we can differentiate in the sense of distributions: For any $\phi \in \mathscr{D}(\mathbb{R})$

$$\begin{aligned}\big(\tilde{D}\ell_H\big)(\phi) &= -\ell_H(D\phi) = -\int_{\mathbb{R}} H(x)\phi'(x)\,dx \\ &= -\int_0^\infty \phi'(x)\,dx = [-\phi(x)]_0^\infty = \phi(0) = \delta_0(\phi)\,.\end{aligned}$$

Thus the delta-distribution is the distributional derivative of the Heavyside function.

3.2.5 Multiplication operator

This idea of extending a differential operator on $\mathscr{D}(\Omega)$ or $\mathscr{S}(\mathbb{R}^n)$ to the dual spaces can be generalised to a greater class of operators.

The general philosophy is as follows:
We consider the locally convex space X, which is equal to $\mathscr{D}(\Omega)$ or $\mathscr{S}(\mathbb{R}^n)$, via the identification $f \sim \ell_f$ as subspace of its topological dual space X' of distributions. Then to a continuous linear operator $L : X \to X$ we have the adjoint $L' : X' \to X'$. If $L'|_X : X \to X$, then we set for $\ell \in X'$ and $\phi \in X$

$$\big(L\ell\big)(\phi) = \ell(L'\phi)\,.$$

Let $f \in \mathcal{C}^\infty(\Omega)$ and $\ell \in \mathscr{D}'(\Omega)$ a distribution. Let M_f denote the multiplication operator acting as $M_f\phi(x) = f(x)\phi(x)$. Then it follows from the

Leibniz formula that $M_f : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is continuous. We set for $\ell \in \mathscr{D}'(\Omega)$

$$\big(\tilde{M}_f \ell\big)(\phi) := \ell(M_f \phi) \qquad \text{for all} \quad \phi \in \mathscr{D}(\Omega)\,. \tag{3.23}$$

Then $\tilde{M}_f \ell$ is a distribution, i.e. the multiplication operator extends to a linear continuous operator $\tilde{M}_f : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$ and a direct calculation shows that the Leibniz formula holds, i.e.

$$\tilde{D}^\alpha \tilde{M}_f \ell = \sum_{\beta \le \alpha} c_{\alpha\beta} \tilde{M}_{D^{\alpha-\beta} f} \tilde{D}^\beta \ell \qquad \text{for all} \quad \ell \in \mathscr{D}'(\Omega)\,. \tag{3.24}$$

3.3 The Fourier Transform

We already gave the definition of the Fourier transform F on the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ in Definition 2.91 and we stated the following theorem:

Theorem 3.46 (Fourier Inversion Theorem)
The Fourier transform is a linear bicontinuous bijection

$$F : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)\,, \quad (Ff)(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx$$

with inverse

$$(F^{-1} f)(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(x)\, dx\,.$$

Moreover, setting $p_\alpha(x) := (ix)^\alpha$ for $\alpha \in \mathbb{N}^n$,

$$\Big(p_\alpha D^\beta F f\Big)(\xi) = \big(F D^\alpha (-1)^{|\beta|} p_\beta f\big)(\xi) \tag{3.25}$$

and

$$\int_{\mathbb{R}^n} |f(x)|^2\, dx = \int_{\mathbb{R}^n} |Ff(k)|^2\, dk \qquad \textit{for all} \quad f \in \mathscr{S}(\mathbb{R}^n)\,. \tag{3.26}$$

Proof. The linearity of the maps $\hat{\cdot}$ and $\check{\cdot}$ is clear. For any $f \in \mathscr{S}(\mathbb{R}^n)$

$$\begin{aligned}\Big(p_\alpha D^\beta F f\Big)(\xi) = (i\xi)^\alpha D_\xi^\beta \hat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \xi^\alpha D_\xi^\beta e^{-ix\cdot\xi} f(x)\,dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} (i\xi)^\alpha (-ix)^\beta e^{-ix\cdot\xi} f(x)\,dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} (-1)^{|\alpha|} \Big(D^\alpha e^{-ix\cdot\xi}\Big)(-ix)^\beta f(x)\,dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \Big(D^\alpha (-ix)^\beta f(x)\Big)\,dx \\ &= \big(F D^\alpha (-1)^{|\beta|} p_\beta f\big)(\xi)\end{aligned}$$

where we used integration by parts. This proves (3.25). We conclude from these equations that for any $f \in \mathscr{S}(\mathbb{R}^n)$

$$\|Ff\|_{\alpha\beta} = \sup_{\xi\in\mathbb{R}^n} |\xi^\alpha D^\beta F f(\xi)| \le \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \Big| D^\alpha (-ix)^\beta f(x)\Big|\,dx < \infty$$

and thus $Ff \in \mathscr{S}(\mathbb{R}^n)$. Furthermore, since $\int (1+x^2)^{-k}\,dx < \infty$ for k sufficiently large, it follows that

$$\begin{aligned}\|Ff\|_{\alpha\beta} &\le \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{(1+x^2)^k}{(1+x^2)^k} \Big| D^\alpha (-ix)^\beta f(x)\Big|\,dx \\ &\le \sup_{x\in\mathbb{R}^n} \Big((1+x^2)^k \big|D^\alpha(-ix)^\beta f(x)\big|\Big) \int_{\mathbb{R}^n} (1+x^2)^{-k}\,dx\,.\end{aligned}$$

Using Leibniz's rule, it follows from Proposition 3.14 that F is continuous. The proof for F^{-1} is similar.

Next we will prove that $F^{-1}Ff = f$, i.e. F^{-1} is the left inverse of F. Since F and F^{-1} are continuous linear maps on $\mathscr{S}(\mathbb{R}^n)$, it suffices to prove $F^{-1}F = \mathrm{Id}$ on the dense subspace $\mathcal{C}_0^\infty(\mathbb{R}^n)$. Let $f \in \mathcal{C}_0^\infty$ be given. For any $\epsilon > 0$ let C_ϵ be a cube around $0 \in \mathbb{R}^n$ of volume $\left(\frac{2}{\epsilon}\right)^n$. Choose ϵ such that $\operatorname{supp} f \subset C_\epsilon$ and set

$$K_\epsilon = \{k \in \mathbb{R}^n \,|\, \forall j = 1, \dots n \,:\, \frac{k_j}{\pi\epsilon} \in \mathbb{Z}\}\,.$$

Then the Fourier series of f is given by

$$f(x) = \sum_{k\in K_\epsilon} \left\langle \left(\frac{\epsilon}{2}\right)^{n/2} e^{ik\cdot(\,)}, f \right\rangle \left(\frac{\epsilon}{2}\right)^{n/2} e^{ik\cdot x} = \sum_{k\in K_\epsilon} \frac{Ff(k)e^{ik\cdot x}}{(2\pi)^{n/2}} (\pi\epsilon)^n \,. \quad (3.27)$$

Since f is continuously differentiable, the series converges uniformly in C_ϵ. $\mathbb{R}^n$ is the disjoint union of the cubes of volume $(\pi\epsilon)^n$ centered about the points in K_ϵ, thus the right hand side of (3.27) is the Riemann sum for the integral of the function $\frac{Ff(k)e^{ik\cdot x}}{(2\pi)^{n/2}} \in \mathscr{S}(\mathbb{R}^n)$ and converges to the integral as $\epsilon \to 0$. This shows $F^{-1}Ff = f$.

The proof that $FF^{-1}f = f$ is similar.

In order to show (3.26), assume that $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, then for $\epsilon > 0$ small enough, f is given by the Fourier series (3.27). Using that $\{\left(\frac{\epsilon}{2}\right)^{n/2} e^{ik\cdot x}\}_{k\in K_\epsilon}$ is an orthonormal basis for $L^2(C_\epsilon)$, it follows that

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x)|^2\,dx &= \int_{C_\epsilon} |f(x)|^2\,dx = \sum_{k\in K_\epsilon} \left|\left\langle \left(\frac{\epsilon}{2}\right)^{n/2} e^{ik\cdot(\,)}, f \right\rangle\right|^2 \\ &= \sum_{k\in K_\epsilon} |Ff(k)|^2 (\pi\epsilon)^n \to \int_{\mathbb{R}^n} |Ff(k)|^2\,dk \quad \text{as} \quad \epsilon \to 0\,.\end{aligned}$$

Since F and $\|\cdot\|_{L^2}$ are continuous on $\mathscr{S}$ and $\mathcal{C}_0^\infty$ is dense, (3.26) holds on $\mathscr{S}(\mathbb{R}^n)$. □

The Fourier transform on the space of tempered distributions $\mathscr{S}'(\mathbb{R}^n)$ can be defined as follows.

Definition 3.47

Let $\ell \in \mathscr{S}'(\mathbb{R}^n)$, then the Fourier transform $F\ell$ of ℓ is the tempered distribution defined by

$$F\ell(\phi) = \ell(F\phi)\,.$$

Appendix A

Appendix

A.1 Spectral Theorem for normal operators - by Jan Möhring

In der Vorlesung wurden verschiedene Versionen des Spektralsatzes für beschränkte selbstadjungierte Operatoren aus dem stetigen Funktionalkalkül (Theorem 1.2) hergeleitet. Thema dieser Note ist es, eine analoge Version des Funktionalkalküls für normale Operatoren aus der Version für selbstadjungierte Operatoren herzuleiten. Hat man einen solchen Funktionalkalkül zur Verfügung, so lässt sich der Spektralsatz für normale Operatoren fast wörtlich wie für selbstadjungierte Operatoren formulieren und es ist in gewisser Weise auch eine Ausdehnung auf unbeschränkte normale Operatoren möglich. Die folgende Argumentation basiert auf dem Artikel von Whitley [1]. Im Text werden einige Resultate herangezogen, die im Anhang (A.1.2) gezeigt werden. Außerdem werden einige einfache Rechnungen interessierten Lesenden als Übungsaufgaben (Abschnitt (A.3)) überlassen. Der Autor dieser Arbeit empfiehlt die Auseinandersetzung mit den Aufgaben ausdrücklich.

A.1.1 Der Beweis

Im Folgenden ist $\mathscr{H}$ immer ein komplexer Hilbertraum. Ein Operator ist ein Element von $\mathcal{L}(\mathscr{H})$, also ein beschränkter linearer Operator $\mathscr{H} \longrightarrow \mathscr{H}$. Mit N bezeichnen wir einen normalen Operator auf $\mathscr{H}$.

Die Idee

Der Beweis des Spektralsatzes für einen selbstadjungierten Operator A basierte auf dem Funktionalkalkül Ψ_A, der ein isometrischer Isomorphismus von $\mathcal{C}_{\mathbb{R}}(\sigma(A))$ nach $\mathcal{L}(\mathscr{H})$ war, hierbei bezeichnet $\mathcal{C}_{\mathbb{R}}(\sigma(A))$ die Menge aller stetigen Funktionen $\sigma(A) \longrightarrow \mathbb{C}$, wobei das "$\mathbb{R}$" von der bekannten Tatsache $\sigma(A) \subseteq \mathbb{R}$ kommt. Das Bild von Ψ_A war dabei die von A und der Identität I erzeugte abgeschlossene Unteralgebra von $\mathcal{L}(\mathscr{H})$ und Ψ_A hat ein reelles Polynom p dem Operator $p(A)$ zugeordnet. Es soll jetzt zunächst betrachtet werden, wie der Funktionalkalkül für normale Operatoren aussehen sollte.

Als ersten wesentlichen Unterschied zwischen dem selbstadjungierten und dem normalen Fall bemerkt man, dass das Spektrum eines normalen Operators im Allgemeinen eine kompakte Teilmenge von $\mathbb{C}$ ist, sodass das Identitäts-Polynom $p(z) = z$, das man mit einem normalen Operator N assoziieren möchte, eine Abbildung $\mathbb{C} \longrightarrow \mathbb{C}$ ist. Wir müssen also komplexe Polynome und den Raum $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ betrachten.

Hierbei entsteht jetzt ein Problem. Man erinnert sich daran, dass der selbstadjungierte Fall bewiesen wurde, indem zunächst Polynome betrachtet wurden und anschließend der Satz von Stone-Weierstraß in Kombination mit dem Satz über die lineare Fortsetzung (BLT-Theorem) angewendet wurde. Möchte man diese Argumentation jetzt übertragen, so benötigt man die komplexe Version des Satzes von Stone-Weierstraß, die als eine der Forderungen an die betrachtete Unteralgebra voraussetzt, dass die Unteralgebra selbstadjungiert ist. Das bedeutet, dass man, wenn man es wieder mit Polynomen versuchen will, mit einem Polynom $p(\cdot)$ auch das Polynom $\bar{p}(\cdot)$ in der Unteralgebra enthalten haben möchte. Es zeigt sich allerdings, dass die komplexen Polynome $p(\cdot)$ im Allgemeinen nicht dicht in $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ sind.

Um das einzusehen, betrachten wir den Einheitskreis S^1 und den zugehörigen Funktionenraum $\mathcal{C}_{\mathbb{C}}(S^1)$ (das ist keine besonders künstliche Wahl — siehe Aufgabe (A.40)). Bekanntlich gilt auf dem Einheitskreis $\bar{z} = z^{-1}$. Es folgt deswegen unmittelbar, dass mit $p(z)$ auf dem Einheitskreis $\bar{p}(z)$ im Allgemeinen kein Polynom, geschweige denn, überhaupt definiert ist. Der Satz von Stone-Weierstraß lässt sich in dieser Form also nicht anwenden. Da es sich dabei nur um ein hinreichendes Kriterium handelt, kann man sich dann fragen, ob die komplexen Polynome nicht trotzdem noch dicht in $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ liegen. Diese Frage muss spätestens mit dem folgenden Argument endgültig verneint werden. Es liegt $\bar{z}$ nicht im uniformen Abschluss der Polynome: Man weiß nämlich aus der Funktionentheorie, dass Polynome holomorph auf dem Einheitskreis sind, dass der gleichmäßige Limes holomorpher Funktionen holomorph ist, und dass $\bar{z}$ nicht holomorph ist.

Man kann also nicht wie bisher vorgehen und einem Polynom p den Operator $p(N)$ zuordnen und diese Abbildung dann nach $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ fortsetzen. Einstweilen sei nun daran erinnert, dass nach dem Satz von Stone-Weierstraß eine Unteralgebra dicht in $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ liegt, wenn sie die konstanten Funktionen enthält, die Punkte von $\sigma(N)$ trennt und unter Konjugation abgeschlossen ist. Ist $p(z) = z = x + \mathrm{i}y$ das Identitäts-Polynom assoziiert mit dem Operator N, so würde man analog das Polynom $\bar{p}(z) = \bar{z} = x - \mathrm{i}y$ mit dem Operator N^* assoziieren wollen. Dies suggeriert, Polynome der Form

$$p(z, \bar{z}) = \sum a_{ij} z^i \bar{z}^j, \qquad a_{ij} \in \mathbb{C},$$

zu betrachten und ein solches Polynom mit dem Operator

$$p(N, N^*) = \sum a_{ij} N^i (N^*)^j$$

zu assoziieren. Zu zeigen ist jetzt, dass die Abbildung Ψ_N, die $p(z, \bar{z})$ auf den Operator $p(N, N^*)$ abbildet, ein isometrischer Isomorphismus vom Raum $\mathcal{P}$ der Polynome der Form $p(z, \bar{z})$ in eine geeignete Teilmenge von $\mathcal{L}(\mathscr{H})$ ist.

Sei dazu also $\mathcal{P}$ der Raum aller komplexen Polynome der Form $p(z, \bar{z}) = \sum a_{ij} z^i \bar{z}^j$ mit $a_{ij} \in \mathbb{C}$, ausgestattet mit der Supremumsnorm

$$\|p\|_\infty = \sup\{|p(z, \bar{z})| : z \in \sigma(N)\}.$$

Es ist dann klar, dass $\mathcal{P}$ eine Algebra bezüglich der üblichen punktweisen Operationen ist und weil $\mathcal{P}$ offenbar die Punkte von $\sigma(N)$ trennt und abgeschlossen unter Konjugation ist, ist $\mathcal{P}$ dicht in $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ (Aufgabe (A.41)). Sei nun $\mathcal{A}$ die abgeschlossene Unteralgebra von $\mathcal{L}(\mathscr{H})$, die von N, N^* und I erzeugt wird und definiere die Abbildung $\Psi_N : \mathcal{P} \longrightarrow \mathcal{A}$ mit

$$\Psi_N : p(z, \bar{z}) \longmapsto p(N, N^*).$$

Dann ist $\Psi_N(\mathcal{P})$ dicht in $\mathcal{A}$ wegen des Satzes von Stone-Weierstraß und wegen der Stetigkeit von Ψ_N (Aufgaben (A.42) und (A.43)). Weiterhin ist klar, dass Ψ_N ein Algebrenhomomorphismus ist (Aufgabe (A.43)). Wir wollen jetzt zeigen, dass Ψ_N eine Isometrie ist, das heißt,

$$\|p(N, N^*)\| = \|p\|_\infty,$$

und können dann Ψ_N zu einem isometrischen Homomorphismus von $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ nach $\mathcal{A}$ fortsetzen. Es wird sich zeigen, dass der Nachweis der Isometrieeigenschaft der wesentliche Teil der jetzt folgenden Betrachtungen ist.

Reduktion von Operatoren durch Unterräume

Definition A.1

Sei M ein abgeschlossener Unterraum von $\mathscr{H}$. Es reduziert M einen Operator $S \in \mathcal{L}(\mathscr{H})$ (oder S wird von M reduziert) genau dann, wenn sowohl M als auch $M^\perp$ invariant unter S sind.

Remark A.2

Mit der Notation aus Definition (A.1) *sind die folgenden Aussagen äquivalent.*

i) M *reduziert* S,

ii) $M^\perp$ *reduziert* S,

iii) M *reduziert* S^*,

iv) M *ist invariant unter* S *und* S^*.

Für den Beweis beachte man einfach die Gültigkeit der folgenden zwei Aussagen und die Aufgabe (1.36)*:*

(a) $M^{\perp\perp} = M$,

(b) $S(M) \subset M$ *genau dann, wenn* $S^*(M^\perp) \subset M^\perp$.

Dabei wurde (a) schon im letzten Semester gezeigt und (b) folgt aus

$$\begin{aligned} u \in M, v \in M^\perp \quad &\Longrightarrow \quad Su \in M \\ &\Longrightarrow \quad \langle u, S^*v\rangle = \langle Su, v\rangle = 0 \\ &\Longrightarrow \quad S^*v \in M^\perp. \end{aligned}$$

NOTATION A.3

Ist $S \in \mathcal{L}(\mathscr{H})$ *und* M *ein abgeschlossener Unterraum von* $\mathscr{H}$*, so bezeichnet* $S|_M$ *die Einschränkung von* S *auf* M*.*

LEMMA A.4

Wenn M *den Operator* S *reduziert, dann gilt* $(S|_M)^* = S^*|_M$.

Proof. Setze $U = S|_M$ und $V = S^*|_M$. Dann gilt $U, V \in \mathcal{L}(M)$ wegen Bemerkung (A.2). Sind $x, y \in M$, so gilt nun

$$\langle U^*x, y\rangle = \langle x, Uy\rangle = \langle x, Sy\rangle = \langle S^*x, y\rangle = \langle Vx, y\rangle.$$

Das zeigt $U^* = V$ wegen der Abgeschlossenheit von M. □

COROLLARY A.5

Reduziert M *den Operator* S *und ist* S *normal, so ist* $S|_M$ *normal.*

Proof. Wir wenden Lemma (A.4) an und erhalten

$$\begin{aligned} (S^*S)|_M &= (S^*|_M)(S|_M) = (S|_M)^*(S|_M) \\ &= (SS^*|_M) = (S|_M)(S^*|_M) = (S|_M)(S|_M)^*. \end{aligned}$$

□

Lemma A.6

Sei N ein normaler Operator mit $0 \in \sigma(N)$ und $\varepsilon > 0$. Dann gibt es einen abgeschlossenen Unterraum $M \neq \{0\}$ derart, dass jeder Operator, der mit NN^* kommutiert, von M reduziert wird und $\|N|_M\| \leq \varepsilon$ gilt.

Proof. Setze $A = NN^*$. Wegen $0 \in \sigma(N)$ gibt es wegen des Satzes (A.15) aus dem Anhang eine Folge (x_k) in $\mathscr{H}$ mit $\|x_k\| = 1$ derart, dass $\|Nx_k\| \longrightarrow 0$. Daraus folgt $Ax_k \longrightarrow 0$ und der selbstadjungierte Operator A hat $0 \in \sigma(A)$ (für ein alternatives Argument vergleiche Aufgabe (A.46)).

Zu $\varepsilon > 0$ betrachten wir jetzt die stetige (!) Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R},\ t \longmapsto \begin{cases} 1, & |t| \leq \varepsilon/2, \\ 2(1 - |t/\varepsilon|), & \varepsilon/2 < |t| < \varepsilon, \\ 0, & |t| \geq \varepsilon. \end{cases}$$

Weil A selbstadjungiert ist, kann der stetige Funktionalkalkül aus Theorem 1.2 angewendet werden, um $f(A)$ zu definieren.

Sei nun M der abgeschlossene Unterraum $M = \{x \in \mathscr{H} : f(A)x = x\}$ von $\mathscr{H}$ (siehe auch Aufgabe (A.47)) und sei B ein Operator, der mit A kommutiert. Einstweilen sei daran erinnert, dass dann B auch mit $f(A)$ kommutiert. Es folgt also für alle $x \in M$, dass

$$Bx = Bf(A)x = f(A)Bx,$$

weswegen M invariant unter B ist. Da B^* auch mit A kommutiert, ist M auch invariant unter B^* und es folgt mit Bemerkung (A.2), dass M den Operator B reduziert.

Weil der Funktionalkalkül von A eine Isometrie ist, gilt für $x \in M$ mit $\|x\| = 1$, dass

$$\|Ax\| = \|Af(A)x\| \leq \|Af(A)\| = \sup\{|tf(t)| : t \in \sigma(A)\} \leq \varepsilon.$$

Mit $\|x\| = 1$ folgt dann

$$\|Nx\|^2 = \langle Nx, Nx\rangle = \langle N^*Nx, x\rangle = \langle Ax, x\rangle \leq \|Ax\| \leq \varepsilon,$$

sodass sich $\|N|_M\| \leq \sqrt{\varepsilon}$ ergibt.

Schließlich muss noch $M \neq \{0\}$ gezeigt werden. Man beachte dazu, dass $f(t) = 1$ gilt, falls $f(2t) \neq 0$. Diese Aussage impliziert

$$\|(I - f(A))(f(2A))\| = \sup\{|1 - f(t)||f(2t)| : t \in \sigma(A)\} = 0.$$

Es folgt, dass jedes Element im Bild des Operators $f(2A)$ ein Element von M ist, und dieses Bild ist nicht $\{0\}$ wegen

$$\|f(2A)\| = \sup\{|f(2t)| : t \in \sigma(A)\} \geq |f(0)| = 1.$$

□

Der Spektralabbildungssatz und eine Folgerung

Als nächstes zeigen wir eine Version des Spektralabbildungssatzes für Polynome in zwei Variablen und einen normalen Operator.

Theorem A.7

Sei $p(s,t)$ ein komplexes Polynom in zwei Variablen. Dann gilt

$$\sigma(p(N, N^*)) = \{p(z, \bar{z}) : z \in \sigma(N)\}.$$

Proof. Sei $p(s,t) = \sum a_{ij}s^i t^j$ und $\lambda \in \sigma(N)$. Wegen des Satzes (A.15) aus dem Anhang gibt es eine Folge (x_k) in $\mathscr{H}$ mit $\|x_k\| = 1$ so, dass $\|(\lambda - N)x_k\| \longrightarrow 0$ gilt. Lemma (A.14) besagt, dass in diesem Fall auch $\|(\bar{\lambda} - N^*)x_k\| \longrightarrow 0$ gilt, weil mit N auch $(\lambda - N)$ normal ist (Aufgabe (A.45)). Nun hat man

$$\begin{aligned}(p(N,N^*) - p(\lambda,\bar{\lambda}))x_k &= \sum a_{ij}(N^i(N^*)^j - \lambda^i\bar{\lambda}^j)x_k \\ &= \sum a_{ij}(N^i((N^*)^j - \bar{\lambda}^j)x_k + \bar{\lambda}^j(N^i - \lambda^i)x_k) \\ &= \sum a_{ij}\Big(N^i((N^*)^{(j-1)} + \cdots + \bar{\lambda}^{j-1})(N^* - \bar{\lambda}) \\ &\quad + \bar{\lambda}^j(N^{i-1} + \cdots + \lambda^{i-1})(N - \lambda)\Big)x_k \\ &\longrightarrow 0.\end{aligned}$$

Daraus folgt

$$p(\lambda,\bar{\lambda}) \in \sigma_{\mathrm{ap}}(p(N,N^*)) \subseteq \sigma(p(N,N^*)),$$

wobei σ_{ap} das approximierende Punktspektrum (Definition (A.11) im Anhang) bezeichnet. Das zeigt "$\supseteq$".

Sei jetzt $\mu \in \sigma(p(N,N^*))$. Der Operator $B = p(N,N^*) - \mu$ ist normal und hat $0 \in \sigma(B)$. Wegen Lemma (A.6) gibt es für jedes $n \in \mathbb{N}$ einen abgeschlossenen Unterraum $M_n \neq \{0\}$, der B reduziert und $\|B|_{M_n}\| \leq 1/n$ erfüllt. Da N mit B^*B kommutiert, reduziert jedes M_n auch N. Mit Korollar (A.5) folgt, dass $N|_{M_n}$ normal ist. Sei $\lambda_n \in \sigma(N|_{M_n})$ (das existiert!). Dann gibt es eine Folge (y_n) in M_n mit $\|y_n\| = 1$ so, dass $\|(\lambda_n - N)y_n\| \leq 1/n$ gilt. Die Folge (λ_n) ist durch $\|N\|$ beschränkt und hat daher eine konvergente Teilfolge, die hier auch mit (λ_n) bezeichnet wird. Sei $\lambda \in \mathbb{C}$ der Grenzwert der Teilfolge. Dann gilt $\lambda \in \sigma(N)$ wegen

$$\|(\lambda - N)y_n\| \leq \|(\lambda - \lambda_n)y_n\| + \|(\lambda_n - N)y_n\| \leq |\lambda_n - \lambda| + \|(\lambda_n - N)y_n\| \longrightarrow 0.$$

Im ersten Teil des Beweises wurde allerdings gezeigt, dass im Falle $(\lambda - N)y_n \longrightarrow 0$ folgt

$$(p(N, N^*) - p(\lambda, \bar{\lambda}))y_n \longrightarrow 0.$$

Andererseits ist $y_n \in M_n$, also gilt wegen $\|B|_{M_n}\| \leq 1/n$ auch

$$\|By_n\| = \|(p(N, N^*) - \mu)y_n\| \leq 1/n.$$

Wegen der Eindeutigkeit des Grenzwertes schließt man $\mu = p(\lambda, \bar{\lambda})$. Das zeigt "$\subseteq$". □

Das Ergebnis

Mit Satz (A.7) folgt nun die Isometrie genau wie im selbstadjungierten Fall.

Corollary A.8
Ist $N \in \mathcal{L}(\mathscr{H})$ normal, so gilt

$$\|p(N, N^*)\| = \sup\{|p(z, \bar{z})| : z \in \sigma(N)\}.$$

Proof. Wir benutzen, dass $(\bar{p}p)(N, N^*)$ normal ist (Aufgabe (A.48)) und folgern mit Satz (A.7), dass

$$\begin{aligned} \|p(N, N^*)\|^2 &= \|p(N, N^*)^* p(N, N^*)\| \\ &= \|(\bar{p}p)(N, N^*)\| \\ &= \sup_{z \in \sigma((\bar{p}p)(N, N^*))} |z| \end{aligned}$$

$$= \sup_{z\in\sigma(N)} |(\bar{p}p)(z,\bar{z})|$$
$$= \left(\sup_{z\in\sigma(N)} |p(z,\bar{z})| \right)^2 .$$

□

Ist nun $\Psi_N : \mathcal{P} \longrightarrow \mathcal{L}(\mathscr{H})$ der Funktionalkalkül definiert über

$$\Psi_N p = p(N, N^*),$$

dann ist Ψ_N ein isometrischer algebraischer $*$-Homomorphismus von $\mathcal{P}$ nach $\mathcal{L}(\mathscr{H})$. Weil $\mathcal{P}$ dicht in $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ ist, kann Ψ_N zu einem isometrischen algebraischen $*$-Homomorphismus von $\mathcal{C}_{\mathbb{C}}(\sigma(N))$ nach $\mathcal{L}(\mathscr{H})$ fortgesetzt werden, wobei die Fortsetzung auch Ψ_N heiße. Genau wie im selbstadjungierten Fall folgert man dann weiter leicht, dass Ψ_N eindeutig ist, dass $\Psi_N(\bar{f}) = \Psi_N(f)^*$ für jedes $f \in \mathcal{C}_{\mathbb{C}}(\sigma(N))$ gilt und dass jeder mit N und N^* kommutierende Operator auch mit $\Psi_N f$ für jedes $f \in \mathcal{C}_{\mathbb{C}}(\sigma(N))$ kommutiert (Aufgabe (A.49)). Insgesamt haben wir daher den folgenden Satz gezeigt.

Theorem A.9

Sei N ein normaler Operator auf $\mathscr{H}$ und $f, g \in \mathcal{C}_{\mathbb{C}}(\sigma(N))$. Dann gibt es eine eindeutige Abbildung $\Psi_N : \mathcal{C}_{\mathbb{C}}(\sigma(N)) \longrightarrow \mathcal{L}(\mathscr{H})$ mit

(a) *Ψ_N ist linear und es gilt $\Psi_N(fg) = \Psi_N(f)\Psi_N(g)$.*

(b) *$\|\Psi_N(f)\| = \|f\|_\infty$.*

(c) *$\Psi_N(\bar{f}) = \Psi_N(f)^*$.*

(d) *Es gilt $\sigma(\Psi_N(f)) = \{f(\lambda) : \lambda \in \sigma(N)\}$.*

(e) *Kommutiert* $B \in \mathcal{L}(\mathscr{H})$ *mit* N *und* N^*, *so kommutiert* B *auch mit jedem* $\Psi_N(f)$.

Remark A.10 *i) Es kann (e) in Satz* (A.9) *verbessert werden, wenn man den Satz von Fuglede anwendet: Sind* $T, N \in \mathcal{L}(\mathscr{H})$ *mit* N *normal, so gilt: Aus* $TN = NT$ *folgt* $TN^* = N^*T$.

ii) Eine ähnliche Konstruktion des Funktionalkalküls kann auch mit direkten maßtheoretischen Argumenten vollzogen werden. Eine Anleitung für den Beweis steht in [W] und in [Ha] ist der Beweis etwas detaillierter beschrieben. Außerdem ist es auch möglich, den Satz aus der Theorie von C^**-Algebren zu erhalten, eine solche Vorgehensweise wird beispielsweise bei [R] benutzt. In [Ha] ist außerdem beschrieben, wie man vom stetigen Funktionalkalkül für normale Operatoren mit relativ wenig Aufwand zu einer Multiplikationsoperator-Version für unbeschränkte normale Operatoren kommt.*

A.1.2 Das approximierende Punktspektrum

Im Text wird mehrmals eine Aussage benötigt, die aus einer Betrachtung über das approximierende Punktspektrum für normale Operatoren folgt. Da diese Ergebnisse von eigenem Interesse sind, sammeln wir sie hier im Anhang.

Definition A.11
Sei $A \in \mathcal{L}(\mathscr{H})$.

(a) Ein $\lambda \in \mathbb{C}$ *heißt genau dann approximierender Eigenwert von* A, *wenn es für jedes* $\varepsilon > 0$ *ein* $x \in \mathscr{H}$ *mit* $\|x\| = 1$ *so gibt, dass* $\|Ax - \lambda x\| < \varepsilon$ *gilt.*

(b) *Das approximierende Punktspektrum $\sigma_{ap}(A)$ ist die Menge aller approximierenden Eigenwerte von A.*

REMARK A.12

Die Werte im approximierenden Punktspektrum sind jene, für welche die Eigenwertgleichung $Ax = \lambda x$ nur approximativ gilt, das heißt, im Grenzwert. Folgende Aussagen sind zu (a) aus Definition (A.11) *äquivalent:*

- *Für jedes $\varepsilon > 0$ gibt es ein $x \neq 0$ mit $\|Ax - \lambda x\| < \varepsilon \|x\|$.*
- *Es gibt eine Folge (x_n) von Vektoren aus $\mathscr{H}$ mit $\|x_n\| = 1$ für alle n derart, dass $\lambda x_n - Ax_n \longrightarrow 0$ für $n \uparrow \infty$ gilt. Das ist äquivalent zu $\|Ax_n - \lambda x_n\| \longrightarrow 0$ für $n \uparrow \infty$.*

THEOREM A.13

Für jedes $A \in \mathcal{L}(\mathscr{H})$ gilt $\sigma_{\mathrm{ap}}(A) \subseteq \sigma(A)$.

Proof. Ist $\lambda \notin \sigma(A)$, so ist $A - \lambda$ invertierbar und es folgt

$$\|x\| = \|(A-\lambda)^{-1}(A-\lambda)x\| \leq \|(A-\lambda)^{-1}\|\|Ax - \lambda x\|$$

für jedes $x \in \mathscr{H}$. Daraus folgt $\|Ax - \lambda x\| \geq \varepsilon \|x\|$ mit $\varepsilon = 1/\|(A-\lambda)^{-1}\|$ für jedes x, also $\lambda \notin \sigma_{\mathrm{ap}}(A)$. □

Für den Beweis des nachfolgenden Satzes erinnern wir noch an das folgende einfache Lemma ([W], Lemma V.5.10).

LEMMA A.14

Ist $N \in \mathcal{L}(\mathscr{H})$ normal, so gilt für jedes $x \in \mathscr{H}$, dass

$$\|Nx\| = \|N^*x\|.$$

Der folgende Satz besagt nun, dass für normale Operatoren das Spektrum mit dem approximierenden Punktspektrum zusammenfällt.

THEOREM A.15
Ist $N \in \mathcal{L}(\mathscr{H})$ normal, so gilt $\sigma_{\mathrm{ap}}(N) = \sigma(N)$.

Proof. Wegen Satz (A.13) reicht es aus, "$\supseteq$" zu zeigen. Sei $\lambda \notin \sigma_{\mathrm{ap}}(N)$. Dann gibt es ein $\varepsilon > 0$ mit

$$\|Ny - \lambda y\| \geq \varepsilon\|y\|, \qquad y \in \mathscr{H}. \tag{$*$}$$

Weil mit N auch $N-\lambda$ normal ist (Aufgabe (1.37)) und weiterhin $(N-\lambda)^* = N^* - \bar{\lambda}$ gilt, folgt mit Lemma (A.14), dass

$$\|N^*y - \bar{\lambda}y\| \geq \varepsilon\|y\|, \qquad y \in \mathscr{H}. \tag{$**$}$$

Um $\lambda \notin \sigma(N)$ zu zeigen, also die Invertierbarkeit von $N - \lambda$, reicht es wegen ($*$) zu zeigen, dass das Bild von $N - \lambda$ dicht ist. Äquivalent dazu ist, dass das orthogonale Komplement des Bildes von $N - \lambda$ nur aus der 0 besteht. Sei also y orthogonal zum Bild von $N - \lambda$, dann gilt

$$0 = \langle (N - \lambda)x, y\rangle = \langle x, (N^* - \bar{\lambda})y\rangle$$

für alle $x \in \mathscr{H}$. Daraus folgt $N^*y - \bar{\lambda}y = 0$. Wegen ($**$) ergibt sich schließlich $y = 0$. ☐

A.1.3 Aufgaben

AUFGABE A.16
Einstweilen sei daran erinnert, dass ein $U \in \mathcal{L}(\mathscr{H})$ genau dann unitär heißt, wenn $U^* = U^{-1}$ gilt.

Zeigen Sie, dass $\sigma(U) \subseteq S^1$ für jeden unitären Operator U gilt.

Zeigen Sie weiter, dass die Umkehrung der obigen Aussage falsch ist, das heißt, aus $\sigma(U) \subseteq S^1$ folgt im Allgemeinen nicht, dass U unitär ist.

Beweisen Sie schließlich, dass ein normales $U \in \mathcal{L}(\mathscr{H})$ genau dann unitär ist, wenn $\sigma(U) \subseteq S^1$ gilt.

Aufgabe A.17

Zeigen Sie, dass der Raum $\mathcal{P}$ bezüglich der üblichen punktweisen Addition und Multiplikation eine Algebra ist. Weisen Sie weiterhin nach, dass $\mathcal{P}$ bezüglich $\|\cdot\|_\infty$ eine normierte Algebra ist, begründen Sie dabei insbesondere die Wohldefiniertheit der Supremumsnorm auf $\mathcal{P}$ bezüglich $\sigma(N)$.

Aufgabe A.18

Seien S und T topologische Räume, $f : S \longrightarrow T$ stetig und $E \subseteq S$ dicht in S. Zeigen Sie, dass $f(E)$ dicht in $f(S)$ ist. Untersuchen Sie weiter, ob $f(E)$ stets auch dicht in T ist.

Aufgabe A.19

Zeigen Sie, dass die Abbildung $\Psi_N : \mathcal{P} \longrightarrow \mathcal{L}(\mathscr{H})$ ein stetiger Algebrenhomomorphismus ist.

Aufgabe A.20

Geben Sie die Details des Beweises von Bemerkung (A.2).

Aufgabe A.21

Zeigen Sie, dass mit $N \in \mathcal{L}(\mathscr{H})$ normal auch $(\lambda - N)$ für jedes $\lambda \in \mathbb{C}$ normal ist.

Aufgabe A.22

Zeigen Sie mit einem anderen Argument als im Text: Ist N normal mit $0 \in \sigma(N)$, so gilt auch $0 \in \sigma(A)$ für $A = NN^*$.

Aufgabe A.23

Zeigen Sie, dass $M = \{x \in \mathscr{H} : f(A)x = x\}$ ein abgeschlossener Unterraum von $\mathscr{H}$ ist.

Aufgabe A.24

Zeigen Sie, dass $(\bar{p}p)(N, N^*)$ selbstadjungiert ist.

Aufgabe A.25

Folgern Sie analog zum selbstadjungierten Fall die restlichen Eigenschaften von Ψ_N aus Satz (A.9).

A.2 The Gelfand-Naimark-Theorem by Pushya Mitra

A.2.1 Introduction

Formulation and interpretation of quantum mechanics in some sense is difficult if one treats quantum mechanics completely through the notion of a Hilbert space. Hilbert spaces do not have any analogy in the classical theory and cannot be directly accessed by "measurements". In the meta-sense, measurements can only be associated with observables of the theory and only observables have counterparts in classical mechanics. Additionally, the notion of states in a Hilbert space possess few mathematical difficulties as we

shall see in the succeeding section where we try to set up the mathematical identity required for the existence of hidden variables in quantum mechanics.
The empiricists approach, where they do not mathematically clarify what they mean by state and what they mean by observables, forces a lot of trouble and with that approach for example one then has to start admitting almost everything as definitions and notions like quasi states, contextual, non-contextual, etc. which creates immense amount of confusions and sometime may mislead to the conclusions which are not consistent with the theory. This shouts for the need to formulate Quantum mechanics based on the Algebra of observables. In quantum mechanics, the most general form for an algebra of observables is a C^*-algebra. The dynamics of the quantum mechanical system is described through an element of a C^*-algebra.

In this talk, we discuss an important theorem which helps one to a deeper understanding of the structure of a Banach-Algebra. This is the Gelfand-Naimark-Theorem, which (in one of its forms) says that every commutative C^*-algebra $\mathfrak{A}$ is isometrically isomorphic to the space of continuous complex-valued functions on its spectrum. In the classical reinterpretation of quantum mechanics, this theorem helps one to formulate the theorem of Kochen and Specker, which rules out the existence of hidden variables in quantum mechanics if the dimension of the Hilbert space is at least 3.

The following work is largely adapted from the book [HS] of Hirzebruch and Scharlau, few intermediate results are from the books [**?**] of Reed and Simon and [BR] of Bratelli and Robinson.

A.2.2 Properties of an algebra

Definition A.26 (Algebra)

Let $\mathfrak{A}$ be a $\mathbb{C}$-vector space. The space $\mathfrak{A}$ is called a complex **algebra** *if it is equipped with a multiplication law which associates the product to each pair $a, b \in \mathfrak{A}$. The product is assumed to be associative and distributive, i.e. $\forall a, b, c \in \mathfrak{A}$, $\forall \alpha, \beta \in \mathbb{C}$*

i) $a(bc) = (ab)c$

ii) $a(b + c) = ab + ac$ *and* $(a + b)c = ac + bc$

iii) $\alpha(ab) = (\alpha a)b = a(\alpha b)$

A subspace $\mathscr{B}$ of $\mathfrak{A}$ which is also an algebra with respect to the operations of $\mathfrak{A}$ is called **sub-algebra***.*

An algebra is **commutative***, or abelian if the product is commutative, i.e. if $ab = ba$ for all $a, b \in \mathfrak{A}$.*

An algebra $\mathfrak{A}$ is a **normed algebra***, if to each element $a \in \mathfrak{A}$ there is associated a real number $\|a\|$, the norm of a, satisfying the following requirements for all $a, b \in \mathfrak{A}$ and $\alpha \in \mathbb{C}$:*

i) $\|a\| \geq 0$ *and* $\|a\| = 0$ *iff* $a = 0$

ii) $\|\alpha a\| = |\alpha|\|a\|$

iii) $\|a + b\| \leq \|a\| + \|b\|$*- triangle inequality*

iv) $\|ab\| \leq \|a\|\|b\|$*.*

If a normed algebra $\mathfrak{A}$ is ***complete****, it is called* ***Banach algebra****.*
A Banach algebra is called **unital** *if there exists $1_\mathfrak{A} \in \mathfrak{A}$ such that $a1_\mathfrak{A} = 1_\mathfrak{A} a = a$, $\forall a \in \mathfrak{A}$ and $\|1_\mathfrak{A}\| = 1$. We call $1_\mathfrak{A}$* **unit***.*

The norm defines a **metric topology** on $\mathfrak{A}$ also known as **uniform topology**. A special (open) neighbourhood of an element $a \in \mathfrak{A}$ in this topology is given by $\mathfrak{U}(a; \epsilon) = \{b \in \mathscr{A} \mid \|b - a\| < \epsilon\}$, where $\epsilon > 0$; the open ball around a with radius ϵ.

If there exists a unit then it is unique.

Definition A.27 (Homomorphism, Inverse)
Let $\mathfrak{A}$, $\mathfrak{B}$ be complex unital algebras.
A linear map $\Phi : \mathfrak{A} \to \mathfrak{B}$ is called **(algebra) homomorphism***, if $\forall a, b \in \mathfrak{A}$*

$$\Phi(ab) = \Phi(a)\Phi(b)\,.$$

A linear functional ℓ on $\mathfrak{A}$, which is not identically zero, is called **complex homomorphism***, if $\ell(ab) = \ell(a)\ell(b)$ $\forall a, b \in \mathfrak{A}$.*
An element $a \in \mathfrak{A}$ is said to be **invertible***, if it has an* **inverse***, i.e., if there exists an element $a^{-1} \in \mathfrak{A}$ such that*

$$aa^{-1} = a^{-1}a = 1_\mathfrak{A}\,.$$

If follows that $\ell(1_\mathfrak{A}) = 1$ for any complex homomorphism and if a has an inverse, than $\ell(a)\ell(a^{-1}) = \ell(1_\mathfrak{A}) = 1$ and thus $\ell(a) \neq 0$.

The following discussion involves the spectrum of $a \in \mathfrak{A}$.

Definition A.28 (Spectrum and Resolvent)
Let $\mathfrak{A}$ be a unital Banach algebra and $a \in \mathfrak{A}$.

i) A complex number λ is said to be in **resolvent set** $\rho(a)$ *of a, if $\lambda 1_{\mathfrak{A}} - a$ has an inverse $\big(\lambda 1_{\mathfrak{A}} - a\big)^{-1} \in \mathfrak{A}$.*

ii) If $\lambda \notin \rho(a)$, then λ is said to be in **spectrum** $\sigma(a) := \mathbb{C} \setminus \rho(a)$ *of a.*

iii) The **spectral radius** *of a is given by $r(a) := \sup_{\lambda \in \sigma(a)} |\lambda|$.*

Theorem A.29

Let $\mathfrak{A}$ be a unital Banach algebra. Then every $a \in \mathfrak{A}$ has non-empty spectrum and $|\lambda| \leq \|a\|$ for any $\lambda \in \sigma(a)$. Moreover $r(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}}$.

Proof. For $|\lambda| > \|a\|$, let us do the following computation:
First remark that $\|a^n\| \leq \|a\|^n$ and $\frac{\|a\|}{|\lambda|} < 1$. Then

$$S_N(a) := 1_{\mathfrak{A}} + \sum_{n=1}^{N} \left(\frac{a}{\lambda}\right)^n$$

form a Cauchy sequence. Since $\mathfrak{A}$ is complete, there is some limit element $S(a) \in \mathfrak{A}$. Since $\left(\frac{a}{\lambda}\right)^n \to 0$ as $n \to \infty$ and

$$\lambda^{-1} S_N(a)\big(\lambda 1_{\mathfrak{A}} - a\big) = 1_{\mathfrak{A}} - \left(\frac{a}{\lambda}\right)^{N+1} = \big(\lambda 1_{\mathfrak{A}} - a\big)\lambda^{-1} S_N(a)$$

it follows that $R_a(\lambda) := \lambda^{-1} S(a)$ is the inverse of a and thus $\lambda \in \rho(a)$. Moreover, $\|R_a(\lambda)\| \to 0$ as $|\lambda| \to \infty$. If $\sigma(a)$ were empty, R_a would be an entire bounded analytic function of λ with values in $\mathfrak{A}$. By Liouville's[1] theorem, R_a would be constant and and thus zero everywhere, which is a contradiction. Thus $\sigma(a)$ is non-empty.
Since R_a is analytic on $\{\lambda > r(a)\}$, the series converges uniformly on every circle Γ_r around zero with radius $r > r(a)$. Term by term integration gives

$$a^n = \frac{1}{2\pi i} \int_{\Gamma_r} \lambda^n R_a(\lambda)\, d\lambda\,.$$

[1] Liouville's theorem says that every bounded entire function must be constant

By the continuity of R_a we have

$$M(r) := \max_{\theta\in[0,2\pi]} \left\|R_a(re^{i\theta})\right\| < \infty\,, \quad r > r(a)\,.$$

Thus $\|a^n\| \leq r^{n+1}M(r)$, giving $\limsup_{n\to\infty} \|a^n\|^{1/n} \leq r$ for all $r > r(a)$ and thus

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq r(a)\,. \tag{A.1}$$

If $\lambda \in \sigma(a)$, then $\lambda^n 1_{\mathfrak{A}} - a^n = (\lambda 1_{\mathfrak{A}} - a)(\lambda^{n-1}1_{\mathfrak{A}} + \ldots + a^{n-1})$ and thus $\lambda^n 1_{\mathfrak{A}} - a^n$ is not invertible, i.e. $\lambda^n \in \sigma(a^n)$ and $|\lambda^n| \leq \|a^n\|$ for $n \in \mathbb{N}$. This shows

$$r(a) \leq \inf_{n\in\mathbb{N}} \|a^n\|^{1/n}\,. \tag{A.2}$$

Combining (A.1) and (A.2) proves the third statement. □

We remark that $\sigma(a)$ is always closed (as the complement of the resolvent set $\rho(a)$ which is always open).

Lemma A.30

Let $\ell : \mathfrak{A} \to \mathbb{C}$ be an arbitrary complex homomorphism. Then $\ell(a) \in \sigma(a)$ for any $a \in \mathfrak{A}$ and ℓ is bounded with $\|\ell\| = 1$.

Proof. For $a \in \mathfrak{A}$, set $\ell(a) = \lambda$, then $\ell(a - \lambda 1_{\mathfrak{A}}) = 0$. This implies that $(a - \lambda 1_{\mathfrak{A}})$ is not invertible which then further implies that $\lambda \in \sigma(a)$. Thus, $|\ell(a)| \leq \|a\|$ for any $a \in \mathfrak{A}$, which gives $\|\ell\| \leq 1$. On the other side, $\ell(1_{\mathfrak{A}}) = 1$ implies $\|\ell\| \geq 1$ □

Theorem A.31 (Theorem of Gelfand-Mazur)

Let $\mathfrak{A}$ be a unital commutative Banach algebra in which every non-zero element is invertible. Then $\Psi : \mathbb{C} \to \mathfrak{A}$, $\lambda \mapsto \lambda 1_{\mathfrak{A}}$ is an isomorphism (i.e. $\mathfrak{A}$ is isometrically isomorphic to $\mathbb{C}$).

Remark The condition is equivalent to the unital Banach algebra being a field. Typically in the notion of the field, the commutativity is contained. So it stresses once-more the fact that the algebra is commutative.

Proof. In order to prove that Ψ is an isomorphism, we go through the standard procedure. It is clear that Ψ is an injective homomorphism. Thus it remains to show surjectivity. Let $a \in \mathfrak{A}$, $a \neq 0$, then by Theorem A.19, it's spectrum is non-empty. Let $\lambda \in \mathbb{C}$ be a spectral value of a, then $a - \lambda 1_{\mathfrak{A}}$ is non-invertible by the definition of the spectrum. By assumption $a - \lambda 1_{\mathfrak{A}}$ must be the zero element, which then implies $a = \lambda 1_{\mathfrak{A}}$. Hence $\mathfrak{A} = \mathbb{C}1_{\mathfrak{A}}$. □

A.2.3 C^*-algebras

Definition A.32 (Involution, C^*-algebra)
Let $\mathfrak{A}$ be a unital Banach-Algebra. An involution in $\mathfrak{A}$ is a map $* : \mathfrak{A} \to \mathfrak{A}$, $a \mapsto a^*$ such that $\forall a, b \in \mathfrak{A}$ and $\lambda \in \mathbb{C}$, the following holds:

i) $(a^*)^* = a$

ii) $(a + b)^* = a^* + b^*$

iii) $(\lambda a)^* = \overline{\lambda} a^*$

iv) $(ab)^* = b^* a^*$

A C^*-algebra is a unital Banach-algebra equipped with an involution $*$ such that the C^*-identity

$$\|a^* a\| = \|a\|^2$$

is satisfied for all $a \in \mathfrak{A}$.

There are two important examples of C^*-algebras:

i) Let X be a compact Hausdorff space. The space $\mathcal{C}(X)$ of all complex valued continuous functions over X is a C^*-algebra with point-wise addition and multiplication, where for $f \in \mathcal{C}(X)$ the adjoint is given by $f^*(x) = \overline{f(x)}$ for $x \in X$ and the norm is given by $\|f\|_\infty = \sup_{x\in X} |f(x)|$.

ii) Let $\mathscr{H}$ be a separable Hilbert space and denote by $\mathcal{L}(\mathscr{H})$ the set of all bounded operators on $\mathscr{H}$. We define the sum of elements of $\mathcal{L}(\mathscr{H})$ pointwise, the product is given by the composition. The Hilbert space adjoint defines an involution on $\mathcal{L}(\mathscr{H})$. With respect to these operations and the operator norm, $\mathcal{L}(\mathscr{H})$ is a C^*-algebra (non-commutative for $\dim \mathscr{H} > 1$). To see this we verify the most important C^*-identity: For a given $a \in \mathcal{L}(\mathscr{H})$, we have

$$\begin{aligned}
\|a\|^2 &= \sup_{\|\psi\|=1} \|a\psi\|^2 = \sup_{\|\psi\|=1} \langle a\psi, a\psi\rangle \\
&= \sup_{\|\psi\|=1} \langle \psi, a^*a\psi\rangle \\
&\leq \sup_{\|\psi\|=1} \|\psi\| \cdot \|a^*a\psi\| = \|a^*a\| \\
&\leq \|a\| \cdot \|a^*\| = \|a\|^2
\end{aligned}$$

Hence $\|a\|^2 \leq \|a^*a\| \leq \|a\|^2$. This implies $\|a^*a\| = \|a\|^2$.

Let $\mathfrak{A}$ be a C^*-algebra and denote by $\mathcal{L}(\mathfrak{A})$ the Banach-algebra of continuous linear maps from $\mathfrak{A}$ into itself. For any $a \in \mathfrak{A}$, we set $T_a : \mathfrak{A} \to \mathfrak{A}$ by $T_a(x) = ax$. Then $\mathfrak{A} \to L(\mathfrak{A})$, $a \mapsto T_a$ is an algebraic homomorphism, also known as Canonical Regular Representation.

Furthermore $\|T_a(x)\| \leq \|a\|\|x\|$ and $\|T_a(1_\mathfrak{A})\| = \|a\| = \|a\|\|1_\mathfrak{A}\|$ which then imply $\|T_a\| = \|a\|$.

Moreover, $\sigma(a) = \sigma(T_a)$: if T_a is invertible and T_a^{-1} is its inverse map, then, clearly $aT_a^{-1}(1_{\mathfrak{A}}) = T_a T_a^{-1}(1_{\mathfrak{A}}) = 1_{\mathfrak{A}}$, which implies that a is also invertible and $T_a^{-1} = T_{a^{-1}}$.

A.2.4 Commutative Banach-Algebras

From now on we consider only commutative unital Banach algebras

Definition A.33 (Ideal)

An **ideal** Σ *of a complex commutative algebra* $\mathfrak{A}$ *is a vector subspace with the additional property that* $xy \in \Sigma$ *for all* $x \in \Sigma$ *and* $y \in \mathfrak{A}$.

Lemma A.34

Let Σ *be an ideal of a commutative algebra* $\mathfrak{A}$. *Then the closure* $\overline{\Sigma}$ *of* Σ *is also an ideal of* $\mathfrak{A}$. *If* Σ *contains any invertible element of* $\mathfrak{A}$, *then* $\Sigma = \mathfrak{A}$.

Proof. Let $a, b \in \overline{\Sigma}$ and $\{a_n\}$, $\{b_n\}$ be sequences in Σ, which converge to a and b respectively. Then clearly $\{a_n + b_n\}$ converges to $a + b$. Which means $a + b \in \overline{\Sigma}$. If $x \in \mathfrak{A}$ then $\{a_n x\}$ is a sequence in Σ which converges to ax, which implies $ax \in \overline{\Sigma}$.

We assume that there exists $a \in \Sigma$ such that a^{-1} exists. Then $a^{-1}a = 1_{\mathfrak{A}} \in \Sigma$, which implies $x = x1_{\mathfrak{A}} \in \Sigma$ for all $x \in \mathfrak{A}$ and thus $\Sigma = \mathfrak{A}$. □

Definition A.35 (Maximal Ideal)

Let $\mathfrak{A}$ *be a commutative algebra and* M *be an ideal. Then* M *is called* **proper** *if* $M \neq \mathfrak{A}$. *A* **maximal ideal** *is a proper ideal* M *such that* $M = J$ *for any other proper ideal* J *with* $M \subseteq J$.

Corollary A.36

Every maximal ideal of a commutative algebra is closed.

Proof. Let M be a maximal ideal of $\mathfrak{A}$ then $\overline{M}$ is also an ideal and $M \subseteq \overline{M}$. As M is maximal, we must have either $\overline{M} = M$ or $\overline{M} = \mathfrak{A}$. But since M contains no invertible elements and the set of invertible elements in $\mathfrak{A}$ is open, the same is true for $\overline{M}$, in particular $1_{\mathfrak{A}} \notin \overline{M}$ and thus $\overline{M} \neq \mathfrak{A}$. □

Note that from now onward all our Banach algebras will be unital.

Lemma A.37

Let $\mathfrak{A}$ be a commutative Banach algebra and Σ be any closed ideal. Then the quotient space $\mathfrak{A}/\Sigma$ is in a canonical way a Banach algebra.

Proof. By former results we know that $\mathfrak{A}/\Sigma$ is a Banach space (denoting for $a \in \mathfrak{A}$ the equivalence class by $\bar{a} = a + \Sigma \in \mathfrak{A}/\Sigma$, we set $\bar{a} + \bar{b} = \overline{a+b}$ and $\lambda\bar{a} = \overline{\lambda a}$) with respect to the quotient norm

$$\|\bar{a}\| = \inf_{s\in\Sigma} \|a - s\| . \tag{A.3}$$

For $a - a' \in \Sigma$ and $b - b' \in \Sigma$ it follows that $(a'b' - ab) = (a' - a)b' + a(b' - b) \in \Sigma$ (since Σ is an ideal). Therefore $\overline{ab} = \overline{a'b'}$, i.e. multiplication is well-defined by $\bar{a}\bar{b} = \overline{ab}$. It is quite clear that $\mathfrak{A}/\Sigma$ is a complex algebra and the map $\pi : \mathfrak{A} \to \mathfrak{A}/\Sigma$ is a homomorphism. Since $\|\bar{a}\| \leq \|a\|$ by (A.3), π is continuous. And now we have to show the norm properties with respect to the multiplicative structure: For any $x_j \in \mathfrak{A}$, $j = 1, 2$, and $\delta > 0$ by (A.3) there exists $s_j \in \Sigma$ such that

$$\|x_j + s_j\| \leq \|\bar{x_j}\| + \delta . \tag{A.4}$$

Since $(x_1+s_1)(x_2+s_2)\in\overline{x_1x_2}$ it follows that

$$\|\bar{x}\bar{y}\|=\|\overline{xy}\|\le\|(x_1+s_1)(x_2+s_2)\|\le\|(x_1+s_1)\|\|(x_2+s_2)\|\,.$$

Thus by (A.4) we get

$$\|\bar{x}\bar{y}\|\le\|\overline{x}\|\|\overline{y}\|\,. \tag{A.5}$$

It remains to show that $\mathfrak{A}/\Sigma$ has a unit. Clearly, $\overline{1_{\mathfrak{A}}}$ is a unit in $\mathfrak{A}/\Sigma$ by the definition of the product. To see that the norm is one, first remark that $\left\|\overline{1_{\mathfrak{A}}}\right\|\le\|1_{\mathfrak{A}}\|=1$. Moreover, $\overline{1_{\mathfrak{A}}}\neq 0$ since $1_{\mathfrak{A}}\notin\Sigma$ (otherwise $\Sigma=\mathfrak{A}$). Thus by (A.5) we have on the other hand

$$\left\|\overline{1_{\mathfrak{A}}}\right\|^2\ge\left\|\overline{1_{\mathfrak{A}}}^2\right\|=\left\|\overline{1_{\mathfrak{A}}}\right\|$$

it follows that $\left\|\overline{1_{\mathfrak{A}}}\right\|\ge 1$ and thus $\left\|\overline{1_{\mathfrak{A}}}\right\|=1$. □

If M is a maximal ideal in a commutative Banach algebra $\mathfrak{A}$, then by Corollary A.26 it is closed and thus $\mathfrak{A}/M$ is a Banach algebra by Lemma A.27. Choose any $a\in\mathfrak{A}$, $a\notin M$ and put $J_a=\{ab+m\,|\,b\in\mathfrak{A},\,m\in M\}$. Then J_a is an ideal and J_a is larger than M since $a\in J$. Thus $J_a=\mathfrak{A}$ by the maximality of M, showing that there exist $b\in\mathfrak{A}$, $m\in M$ such that $ab+m=1_{\mathfrak{A}}$. It follows that $\overline{ab}=\overline{1_{\mathfrak{A}}}$, therefore each nonzero element of $\mathfrak{A}/M$ is invertible. Hence from the Gelfand-Mazur-Theorem (Theorem A.21), it follows that $\mathfrak{A}/M$ is isometrically isomorphic to the field $\mathbb{C}$. If we denote this isomorphism by $j:\mathfrak{A}/M\to\mathbb{C}$, the map $h_M=j\circ\pi:\mathfrak{A}\to\mathbb{C}$ is a complex homomorphism and its kernel is M. Thus each maximal ideal M is the kernel of some complex homomorphism h.

If on the other hand h is any complex homomorphism on $\mathfrak{A}$, then its kernel $h^{-1}(0)$ is an ideal. Since it has codimension 1, it is maximal.

The above computations have also shown that an element $a \in \mathfrak{A}$ is invertible if and only if a lies in no proper ideal of $\mathfrak{A}$.

Definition A.38 (Spectrum of $\mathfrak{A}$, Gelfand transform)
Let $\mathfrak{A}$ be a commutative complex Banach algebra. Then the set of all maximal ideals of $\mathfrak{A}$ is said to be the **spectrum** *of $\mathfrak{A}$, written as $\sigma(\mathfrak{A})$.*

We denote by $\Delta_{\mathfrak{A}}$ the set of all complex homomorphisms of $\mathfrak{A}$. Then for each $a \in \mathfrak{A}$, the function

$$\hat{a} : \Delta_{\mathfrak{A}} \to \mathbb{C}\,, \quad \hat{a}(\ell) := \ell(a)$$

is called **Gelfand transform** *of a. We denote by $\widehat{\mathfrak{A}}$ the set of all Gelfand transforms.*

The **Gelfand topology** *of $\Delta_{\mathfrak{A}}$ is the weakest topology such that every $\hat{a}$ is continuous.*

There is a one-to-one correspondence between $\Delta_{\mathfrak{A}}$ and $\sigma(\mathfrak{A})$:
As described above, every maximal ideal M defines in a canonical way the continuous complex algebra homomorphism

$$h_M : \mathfrak{A} \to \mathfrak{A}/M \to \mathbb{C}$$

with kernel M and for any $h \in \Delta_{\mathfrak{A}}$, the kernel of h is a maximal ideal. Thus we also can equip $\sigma(\mathfrak{A})$ with the Gelfand topology.
It follows that $\widehat{\mathfrak{A}}$ is a subset of $\mathcal{C}(\sigma(\mathfrak{A}))$, the set of complex-valued continuous functions on the spectrum $\sigma(\mathfrak{A})$.

By Lemma A.20, we may consider $\sigma(\mathfrak{A})$ as a subset of the unit sphere in the (Banach space) dual $\mathfrak{A}'$. The Gelfand topology is the restriction of the weak*-topology of $\mathfrak{A}'$ to $\Delta_{\mathfrak{A}}$ (or $\sigma(\mathfrak{A})$).

Theorem A.39

Let $\mathfrak{A}$ be a commutative Banach algebra. Then $\hat{a}(\sigma(\mathfrak{A})) = \sigma(a)$.

Proof. Since $\ell(a) \in \sigma(a)$ for any $\ell \in \Delta_{\mathfrak{A}}$, it follows that $\lambda \in \sigma(a)$ if $\lambda = \ell(a)$ for some $\ell \in \Delta_{\mathfrak{A}}$. If on the other hand $\lambda \in \sigma(a)$, then $a - \lambda 1_{\mathfrak{A}}$ is not invertible and thus the set $\{x(a - \lambda 1_{\mathfrak{A}}) \,|\, x \in \mathfrak{A}\}$ does not contain $1_{\mathfrak{A}}$. It therefore is a proper ideal, which lies in a maximal ideal M. Thus it is an element of the kernel of h_M constructed above, proving that $\lambda = \ell(a)$ for some $\ell \in \Delta_{\mathfrak{A}}$. It follows that for each $a \in \mathfrak{A}$, the range of $\hat{a}$ is the spectrum of a. □

Theorem A.40

Let $\mathfrak{A}$ be a commutative complex Banach algebra. Then $\sigma(\mathfrak{A})$ is closed and thus a compact subset of the norm-closed unit sphere B in the dual space $\mathfrak{A}'$ with respect to the weak* $-$ topology.

Proof. By the Banach-Alaoglu-Theorem, B is weak*-compact. Since $\sigma(\mathfrak{A}) \subset B$, it suffices to show that it is weak*-closed. Let $\{\ell_n\}_{n\in\mathbb{N}}$ be a sequence in $\Delta_{\mathfrak{A}}$ so that $\{\ell_n x\}$ converges for all $x \in \mathfrak{A}$. Then, setting $\ell(x) = \lim_{n\to\infty} \ell_n(x)$, it remains to show that ℓ is a complex homomorphism, i.e. $\ell(xy) = \ell(x)\ell(y)$ and $\ell(1_{\mathfrak{A}}) = 1$ for all $x, y \in \mathfrak{A}$.
For $x, y \in \mathfrak{A}$ and $\epsilon > 0$ fixed, set

$$U := \{f \in \mathfrak{A}' \,|\, |f(a) - \ell(a)| < \epsilon \text{ for } a \in \{1_{\mathfrak{A}}, x, y, xy\}\}.$$

Then U is a weak*-neighbourhood of ℓ, containing some $f \in \Delta_{\mathfrak{A}}$. By assumption $|f(1_{\mathfrak{A}}) - \ell(1_{\mathfrak{A}})| = |1 - \ell(1_{\mathfrak{A}})| < \epsilon$, giving the second equation. To see the first one, write

$$\ell(xy) - \ell(x)\ell(y) = (\ell(xy) - f(xy)) + (f(x)f(y) - \ell(x)\ell(y))$$

$$= (\ell(xy) - f(xy)) + (f(y) - \ell(y))f(x) + (f(x) - \ell(x))\ell(y)$$

which gives

$$|\ell(xy) - \ell(x)\ell(y)| < (1 + \|x\| + |\ell(y)|)\epsilon\,.$$

□

Thus, since $\sigma(\mathfrak{A})$ with the topology described above is a compact Hausdorff space, one obtains by standard arguments

Corollary A.41

The algebra $\mathcal{C}(\sigma(\mathfrak{A}))$ of complex-valued continuous function on the topological space $\sigma(\mathfrak{A})$ is a Banach-algebra.

Theorem A.42

Let $\mathfrak{A}$ be a commutative complex Banach algebra. Then the Gelfand transform

$$\mathcal{G} : \mathfrak{A} \to \mathcal{C}\big(\sigma(\mathfrak{A})\big)\,, \quad \mathcal{G}(a)(M) = \hat{a}(h_M) = h_M(a)\,, \quad M \in \sigma(\mathfrak{A}) \tag{A.6}$$

is a continuous homomorphism and $\|\mathcal{G}(a)\| = \|\hat{a}\| = r(a) \leq \|a\|$ for any $a \in \mathfrak{A}$, where $r(a)$ is the spectral radius of a.
The kernel of $\mathcal{G}$ is the intersection of all maximal ideals of $\mathfrak{A}$ (which is said to be the radical rad $\mathfrak{A}$ of $\mathfrak{A}$).

Proof. Since $h_M \in \Delta_{\mathfrak{A}}$, it is straightforward by the definition that $\mathcal{G}$ is an algebra homomorphism, e.g.

$$\mathcal{G}(ab)(M) = \widehat{ab}(h_M) = h_M(ab) = h_M(a)h_M(b) = \big(\mathcal{G}(a)\mathcal{G}(b)\big)(M)\,.$$

$\mathcal{G}(a)$ is continuous by the definition of the Gelfand topology. Moreover, from Lemma A.20 it follows that $\|h_M\| = 1$ for all $M \in \sigma(\mathfrak{A})$ and thus $\|\mathcal{G}\| \leq 1$.

The statement on the spectral radius follows from Theorem A.29.

The kernel of $\mathcal{G}$ consists of those $a \in \mathfrak{A}$ for which $h = 0$ for all $h \in \Delta_{\mathfrak{A}}$. Since the kernel of each $h \in \Delta_{\mathfrak{A}}$ is a maximal ideal (see the discussion above the definition of the Gelfand transform), the kernel of $\mathcal{G}$ is the intersection of all these maximal ideals. □

A.2.5 The Gelfand-Naimark-Theorem

THEOREM A.43 (GELFAND-NAIMARK-THEOREM)
Let $\mathfrak{A}$ be a commutative C^*-algebra. Then the Gelfand transform $\mathcal{G} : \mathfrak{A} \to \mathcal{C}\big(\sigma(\mathfrak{A})\big)$ is an isometric isomorphism and $\overline{\mathcal{G}(a)} = \mathcal{G}(a^*)$ for all $a \in \mathfrak{A}$, i.e. $\mathcal{G}$ is a homomorphism of commutative C^*-algebras. In particular $a \in \mathfrak{A}$ is self-adjoint (i.e. $a^* = a$) if and only if $\hat{a}$ is real-valued.

Proof. We start proving that $\mathcal{G}(a^*)(M) = \overline{\mathcal{G}(a)(M)}$ for all $a \in \mathfrak{A}$ and $M \in \sigma(\mathfrak{A})$.

Let $\mathcal{G}(a)(M) = \alpha + i\beta$ and $\mathcal{G}(a^*)(M) = \gamma + i\delta$ with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. In order to get a contradiction, we now assume that $\beta + \delta \neq 0$ and set $c = \frac{a + a^* - (\alpha+\gamma)1_{\mathfrak{A}}}{\beta + \delta}$. It is clear that $c = c^*$ and $\mathcal{G}(c)(M) = i$.

For all $\lambda \in \mathbb{R}$, $\mathcal{G}(c + i\lambda 1_{\mathfrak{A}})(M) = i(1 + \lambda)$. Hence, $|1 + \lambda| \leq \|c + i\lambda 1_{\mathfrak{A}}\|$ and

$$\begin{aligned}(1+\lambda)^2 \leq \|c + i\lambda 1_{\mathfrak{A}}\|^2 &= \|(c + i\lambda 1_{\mathfrak{A}})(c + i\lambda 1_{\mathfrak{A}})^*\| \\ &\leq \left\|c^2 + \lambda^2 1_{\mathfrak{A}}\right\| \leq \left\|c^2\right\| + \lambda^2 .\end{aligned}$$

But for significant large λ the above inequality does not hold. This contradicts our assumption that $\beta + \delta \neq 0$ and thus shows $\beta + \delta = 0$.

We apply the same line of argument on the elements ia, $(ia)^*$ and the result follows.

Since $\mathfrak{A}$ is commutative, each $a \in \mathfrak{A}$ is normal, i.e. $a^*a = aa^*$, and thus $\|a\|^2 = \|a^2\|$. It follows that $\|a\| = r(a)$ and thus, by Theorem A.32, the homomorphism $\mathcal{G}$ is an isometry.

Now it is to be shown that $\mathcal{G}$ is surjective. We do this via the Stone-Weierstrass Theorem (Theorem A.37 below). Since $\mathcal{G}$ is an isometry, the set $\widehat{\mathfrak{A}} \subset \mathcal{C}\big(\sigma(\mathfrak{A})\big)$ of all Gelfand transforms is closed. Clearly $\widehat{\mathfrak{A}}$ is a subalgebra separating points in $\sigma(\mathfrak{A})$ and id $\in \widehat{\mathfrak{A}}$, thus the prerequisites for the application of the Stone-Weierstrass Theorem are given. Therefore $\mathcal{G}(\mathfrak{A})$ is dense in $\mathcal{C}\big(\sigma(\mathfrak{A})\big)$, proving the surjectivity. □

Stone-Weierstrass Theorem

This part is taken from the book of Folland [F].

Let X be a compact Hausdorff space and $\mathcal{C}(X,\mathbb{R})$ be a space of real-valued continuous functions on X equipped with uniform metric. A subset $\mathscr{A}$ of $\mathcal{C}(X,\mathbb{R})$ is said to separate points, if for every $x, y \in X$ with $x \neq y$ there exists $f \in \mathscr{A}$ such that $f(x) \neq f(y)$.

If $\mathscr{A} \subset \mathcal{C}(X,\mathbb{R})$, $\mathscr{A}$ is called **lattice** if $\max(f,g)$ and $\min(f,g)$ are in $\mathscr{A}$ whenever $f, g \in \mathscr{A}$. Since lattice operations are continuous, one easily sees that if $\mathscr{A}$ is an lattice, so it's closure $\overline{\mathscr{A}}$ in the uniform metric. Before we discuss the Stone-Weierstrass theorem, we shall see a few ingredients needed for it's proof in form of the following lemmas (the proofs can be found in [F]).

Lemma A.44

Consider $\mathbb{R}^2$ *as an algebra under coordinate-wise addition and multiplication. The only sub-algebras of* $\mathbb{R}^2$ *are* $\mathbb{R}^2$, $\{(0,0)\}$, *and linear spans of* $(1,0)$, $(0,1)$

and $(1,1)$.

Lemma A.45

If $\mathscr{A}$ *is a closed sub-algebra of* $\mathcal{C}(X,\mathbb{R})$*, then* $|f| \in \mathscr{A}$ *if* $f \in \mathscr{A}$ *and* $\mathscr{A}$ *is a lattice.*

Lemma A.46

Suppose $\mathscr{A}$ *is a closed lattice in* $\mathcal{C}(X,\mathbb{R})$ *and* $f \in \mathcal{C}(X,\mathbb{R})$*. If for every* $x, y \in X$ *there exists* $g_{xy} \in \mathscr{A}$ *such that* $g_{xy}(x) = f(x)$ *and* $g_{xy}(y) = f(y)$*, then* $f \in \mathscr{A}$*.*

Theorem A.47 (Stone-Weierstrass-Theorem)

Let X *be a compact Hausdorff space. If* $\mathscr{A}$ *is a closed sub-algebra of* $\mathcal{C}(X,\mathbb{R})$ *that separates points, then either* $\mathscr{A} = \mathcal{C}(X,\mathbb{R})$ *or* $\mathscr{A} = \{f \in \mathcal{C}(X,\mathbb{R}) : f(x_0) = 0\}$ *for some* $x_0 \in X$*. The first alternative holds if and only if* $\mathscr{A}$ *contains the constant functions.*

Proof. Given $x \neq y \in X$, let $\mathscr{A}_{xy} = \{(f(x), f(y)) : f \in \mathscr{A}\}$. Then $\mathscr{A}_{xy}$ is a sub-algebra of $\mathbb{R}^2$ in the sense of Lemma A.34, because $f \mapsto (f(x), f(y))$ is an algebra homomorphism. If $\mathscr{A}_{xy} = \mathbb{R}^2$ for all x, y then Lemma A.35 and Lemma A.36 imply that $\mathscr{A} = C(X,\mathbb{R})$. Otherwise, there exists x, y for which $\mathscr{A}_{xy}$ is a proper sub-algebra of $\mathbb{R}^2$. It cannot be $\{(0,0)\}$ or the linear span of $(1,1)$ because $\mathscr{A}$ separates points, so by Lemma A.34 $\mathscr{A}_{xy}$ is the linear span of $(1,0)$ or $(0,1)$. In either case there exists $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in \mathscr{A}$. There is only one such x_0, since $\mathscr{A}$ separates points, so if neither x nor y is x_0 we have $\mathscr{A}_{xy} = \mathbb{R}^2$. Lemma A.35 and Lemma A.36 now imply that $\mathscr{A} = \{f \in \mathcal{C}(X,\mathbb{R}) : f(x_0) = 0\}$. Finally, if $\mathscr{A}$ contains the constant functions, there is no x_0 such that $f(x_0) = 0$ for all $f \in \mathscr{A}$, so $\mathscr{A}$ must equal $\mathcal{C}(X, \mathbb{R})$. $\square$

The above theorem is quite general but for the application purposes one typically encounters a sub-algebra $\mathscr{B}$ of $\mathcal{C}(X, \mathbb{R})$ that is not closed, and one applies the theorem to $\overline{\mathscr{A}} = \mathscr{B}$

Corollary A.48
Suppose $\mathscr{B}$ is a sub-algebra of $\mathcal{C}(X, \mathbb{R})$ that separates points. If there exists $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in \mathscr{B}$, then $\mathscr{B}$ is dense in $\{f \in \mathcal{C}(X, \mathbb{R}) : f(x_0) = 0\}$. Otherwise, $\mathscr{B}$ is dense in $\mathcal{C}(X, \mathbb{R})$.

This could also be extended to a complex sub-algebra. The following theorem allows us to do so.

Theorem A.49
Let X be a compact Hausdorff space. If $\mathscr{A}$ is closed complex sub-algebra of $\mathcal{C}(X)$ that separates points and is closed under complex conjugation, then either $\mathscr{A} = \mathcal{C}(X)$ or $\mathscr{A} = \{f \in \mathcal{C}(X) : f(x_0) = 0\}$ for some $x_0 \in X$.

Proof. Since, $\operatorname{Re} f = (f + \overline{f})/2$ and $\operatorname{Im} f = (f - \overline{f})/2i$, the set $\mathscr{A}_{\mathbb{R}}$ of real and imaginary parts of functions in $\mathscr{A}$ is a sub-algebra of $\mathcal{C}(X, \mathbb{R})$ to which the Stone-Weierstrass Theorem applies. Since $\mathscr{A} = \{f + ig : f, g \in \mathscr{A}\}$, the desired result follows. $\square$

A.3 Solutions for some Exercises

Exercise 1.32: The first properties given by (i) are easily checked. We have (for $1(x) = 1$)

$$\Psi(1) = \sum_{n=0}^{\infty} 1\Pi_n = \sum_{n=0}^{\infty} \Pi_n = \mathrm{Id} \tag{A.7}$$

since H is the direct sum of the eigenspaces by Theorem 1.5. Furthermore, for $f(x) = x$ one has

$$\Psi(f) = \sum_{n=0}^{\infty} \lambda_n \Pi_n = A \tag{A.8}$$

again by Theorem 1.5.

Now we come to linearity. Let $f, g \in \mathcal{C}(\sigma(A))$ and $\lambda \in \sigma(A)$, then

$$\begin{aligned}
\Psi(f+g) &= \sum_{n=0}^{\infty}(f+g)(\lambda_n)\Pi_n = \sum_{n=0}^{\infty}(f(\lambda_n)\Pi_n + g(\lambda_n)\Pi_n) \\
&= \sum_{n=0}^{\infty} f(\lambda_n)\Pi_n + \sum_{n=0}^{\infty} g(\lambda_n)\Pi_n = \Psi(f) + \Psi(g)
\end{aligned}$$

by convergence of all series and

$$\Psi(\lambda f) = \sum_{n=0}^{\infty}(\lambda f)(\lambda_n)\Pi_n = \lambda \sum_{n=0}^{\infty} f(\lambda_n)\Pi_n = \lambda\Psi(f).$$

This shows linearity. Next, we prove multiplicativity. This holds by

$$\begin{aligned}
\Psi(fg) &= \sum_n (fg)(\lambda_n)\Pi_n = \sum_n f(\lambda_n)g(\lambda_n)\Pi_n \\
&= \sum_n f(\lambda_n) \sum_k g(\lambda_k)\delta_{nk}\Pi_n = \sum_n f(\lambda_n) \sum_k g(\lambda_k)\Pi_n \circ \Pi_k \\
&= \sum_n f(\lambda_n)\Pi_n \left(\sum_k g(\lambda_k)\Pi_k \right) = \sum_n f(\lambda_n)\Pi_n \circ \sum_n g(\lambda_n)\Pi_n \\
&= \Psi(f) \circ \Psi(g).
\end{aligned}$$

The involution property is true by

$$\Psi(\overline{f}) = \sum_n \overline{f}(\lambda_n)\Pi_n = \sum_n (f(\lambda_n)\Pi_n)^*$$

$$= \left(\sum_n f(\lambda_n) \Pi_n \right)^* = \Psi(f)^*.$$

Here, we have used that the map $T \longmapsto T^*$ is conjugate linear.

Finally, we show continuity (iii). This follows by

$$\begin{aligned} \|\Psi(f)\| &= \| \sum_n f(\lambda_n) \Pi_n \| \leq \sum_n \| f(\lambda_n) \Pi_n \| \\ &\leq \sum_n \|f\|_\infty \|\Pi_n\| = \|f\|_\infty \sum_n \|\Pi_n\| \leq \|f\|_\infty. \end{aligned}$$

Exercise 1.34:

i) We start by showing linearity. This is easy, since the properties follow directly from the linearity of the integral. Let $f, g \in \mathcal{B}(\sigma(A))$ and $\alpha \in \sigma(A)$. Then we have

$$\begin{aligned} \tilde{\Phi}_A(f+g) &= \langle \psi, \tilde{\Phi}_A(f+g)\varphi \rangle = \int (f+g) \, \mathrm{d}\mu_{\psi,\varphi} \\ &= \int f \, \mathrm{d}\mu_{\psi,\varphi} + \int_{\sigma(A)} g \, \mathrm{d}\mu_{\psi,\varphi} = \langle \psi, \tilde{\Phi}_A(f)\varphi \rangle + \langle \psi, \tilde{\Phi}_A(g)\varphi \rangle \\ &= \tilde{\Phi}_A(f) + \tilde{\Phi}_A(g). \end{aligned}$$

Furthermore, one has

$$\begin{aligned} \tilde{\Phi}_A(\alpha f) &= \langle \psi, \tilde{\Phi}_A(\alpha f)\varphi \rangle = \int (\alpha f) \, \mathrm{d}\mu_{\psi,\varphi} \\ &= \alpha \int f \, \mathrm{d}\mu_{\psi,\varphi} = \alpha \langle \psi, \tilde{\Phi}_A(f)\varphi \rangle = \alpha \tilde{\Phi}_A(f). \end{aligned}$$

Linearity is proven.

Next, we consider multiplicativity. Therefore, we need some properties of the measures $\mu_{\psi,\varphi}$ which we first want to provide.

The first fact is from basic measure theory. We want to recall that if m is a measure and f is a measurable function, $f \cdot m$ is the unique measure μ such that

$$\mu(E) = \int_E f \,\mathrm{d}m \tag{A.9}$$

holds for all measurable sets E. In other words, μ is the unique measure such that $\mathrm{d}\mu \,/\, \mathrm{d}m = f$ (Radon-Nikodym-Theorem).

Now we need the

Lemma.

(a) For all $g \in \mathcal{C}(\sigma(A))$ we have $g \cdot \mu_{\psi,\varphi} = \mu_{\psi,\Phi_A(g)\varphi}$, where $\Phi_A(g)$ is the continuous functional calculus of g.

(b) For all $f \in \mathcal{B}(\sigma(A))$ we have $f \cdot \mu_{\psi,\varphi} = \mu_{\tilde{\Phi}_A(f)^*\psi,\varphi}$.

Proof.

(a) Let $f \in \mathcal{C}(\sigma(A))$. Then

$$\begin{aligned}\int f \,\mathrm{d}g \cdot \mu_{\psi,\varphi} &= \int (fg) \,\mathrm{d}\mu_{\psi,\varphi} = \langle \psi, \Phi_A(fg)\varphi \rangle \\ &= \langle \psi, \Phi_A(f)\Phi_A(g)\varphi \rangle = \int f \,\mathrm{d}\mu_{\psi,\Phi_A(g)\varphi} \,.\end{aligned}$$

(b) Let $h \in \mathcal{C}(\sigma(A))$. Then

$$\begin{aligned}\int h \,\mathrm{d}f \cdot \mu_{\psi,\varphi} &= \int (hf) \,\mathrm{d}\mu_{\psi,\varphi} = \int f \,\mathrm{d}h \cdot \mu_{\psi,\varphi} \\ &\overset{\text{(a)}}{=} \int f \,\mathrm{d}\mu_{\psi,\Phi_A(h)\varphi} = \langle \psi, \tilde{\Phi}_A(f)\Phi_A(h)\varphi \rangle \\ &= \langle \tilde{\Phi}_A(f)^*\psi, \Phi_A(h)\varphi \rangle = \int h \,\mathrm{d}\mu_{\tilde{\Phi}_A(f)^*\psi,\varphi} \,.\end{aligned}$$

□

Using the Lemma, we can now prove the multiplicativity of $\tilde{\Phi}_A$. Let therefore $f, g \in \mathcal{B}(\sigma(A))$. Then it holds that

$$\begin{aligned}\langle \psi, \tilde{\Phi}_A(fg)\varphi \rangle &= \int (fg)\,\mathrm{d}\mu_{\psi,\varphi} = \int g\,\mathrm{d}f \cdot \mu_{\psi,\varphi} \\ &\overset{\text{(b)}}{=} \int g\,\mathrm{d}\mu_{\tilde{\Phi}_A(f)^*\psi,\varphi} = \langle \tilde{\Phi}_A(f)^*\psi, \tilde{\Phi}_A(g)\varphi \rangle \\ &= \langle \psi, \tilde{\Phi}_A(f)\tilde{\Phi}_A(g)\varphi \rangle.\end{aligned}$$

This shows $\tilde{\Phi}_A(fg) = \tilde{\Phi}_A(f)\tilde{\Phi}_A(g)$.

Finally we prove the involution property. To that end we first assume $f \in \mathcal{B}(\sigma(A))$ to be real. Then we have

$$\begin{aligned}\langle \psi, \tilde{\Phi}_A(f)^*\varphi \rangle &= \langle \tilde{\Phi}_A(f)\psi, \varphi \rangle = \overline{\langle \varphi, \tilde{\Phi}_A(f)\psi \rangle} \\ &= \overline{\int f\,\mathrm{d}\mu_{\psi,\varphi}} = \int f\,\mathrm{d}\mu_{\psi,\varphi} = \langle \psi, \tilde{\Phi}_A(f)\varphi \rangle.\end{aligned}$$

We have therefore shown viii) from Theorem 1.8: If f is real valued, then $\tilde{\Phi}_A(f)$ is self-adjoint. Now consider $f = u + \mathrm{i}v$ to be complex valued with real functions u and v. By viii), the usual adjoint properties and linearity the calculation

$$\begin{aligned}\langle \psi, \tilde{\Phi}_A(f)^*\varphi \rangle &= \langle \psi, \tilde{\Phi}_A(u)^*\varphi \rangle + \langle \psi, \tilde{\Phi}_A(\mathrm{i}v)^*\varphi \rangle \\ &= \langle \psi, \tilde{\Phi}_A(u)^*\varphi \rangle - \mathrm{i}\langle \psi, \tilde{\Phi}_A(v)^*\varphi \rangle \\ &= \langle \psi, \tilde{\Phi}_A(u)\varphi \rangle - \mathrm{i}\langle \psi, \tilde{\Phi}_A(v)\varphi \rangle \\ &= \langle \psi, (\tilde{\Phi}_A(u) - \mathrm{i}\tilde{\Phi}_A(v))\varphi \rangle\end{aligned}$$

$$= \langle \psi, \tilde{\Phi}_A(u - \mathrm{i}v)\varphi \rangle = \langle \psi, \tilde{\Phi}_A(\bar{f})\varphi \rangle$$

reveals $\tilde{\Phi}_A(\bar{f}) = \tilde{\Phi}_A(f)^*$.

ii) Let $f \in \mathcal{B}(\mathbb{R})$ and μ be a Borel measure. Then there is a sequence of continuous functions (f_n) such that $\int f_n \,\mathrm{d}\mu \longrightarrow \int f \,\mathrm{d}\mu$. The assumption $AB = BA$ implies $\tilde{\Psi}_A(f_n)B = B\tilde{\Psi}_A(f_n)$, since the claim is true for polynomials and therefore for continuous functions (which are uniform limits of polynomials) and since $\tilde{\Psi}_A$ is continuous in $\|\cdot\|_\infty$.

Now let $x \in H$, μ_x the corresponding spectral measure and f_n as above. Then we have

$$\langle x, \tilde{\Psi}_A(f_n)x \rangle = \int f_n \,\mathrm{d}\mu_x \longrightarrow \int f \,\mathrm{d}\mu_x = \langle x, \tilde{\Psi}_A(f)x \rangle.$$

Since H is a complex Hilbert space, the Polarization formula yields

$$\langle y, \tilde{\Psi}_A(f_n)x \rangle \longrightarrow \langle y, \tilde{\Psi}_A(f)x \rangle$$

for all $x, y \in H$. We conclude that $\tilde{\Psi}_A(f_n) \xrightarrow{\mathrm{w}} \tilde{\Psi}_A(f)$. Using this reveals

$$\begin{aligned} \langle y, \tilde{\Psi}_A(f_n)Bx \rangle &= \langle y, B\tilde{\Psi}_A(f_n)x \rangle = \langle B^*y, \tilde{\Psi}_A(f_n)x \rangle \\ &\longrightarrow \langle B^*y, \tilde{\Psi}_A(f)x \rangle = \langle y, B\tilde{\Psi}_A(f)x \rangle. \end{aligned}$$

On the other hand, $\langle y, \tilde{\Psi}_A(f_n)Bx \rangle \longrightarrow \langle y, \tilde{\Psi}_A(f)Bx \rangle$. Thus we have shown

$$\tilde{\Psi}_A(f_n)B = B\tilde{\Psi}_A(f_n) \xrightarrow{\mathrm{w}} \tilde{\Psi}_A(f)B,$$
$$\tilde{\Psi}_A(f_n)B = B\tilde{\Psi}_A(f_n) \xrightarrow{\mathrm{w}} B\tilde{\Psi}_A(f).$$

The claim now follows since weak limits are unique.

iii) Let $f \in \mathcal{B}(\sigma(A))$ be real. We have

$$\begin{aligned}\langle \psi, \tilde{\Phi}_A(f)^* \varphi\rangle &= \langle \tilde{\Phi}_A(f)\psi, \varphi\rangle = \overline{\langle \varphi, \tilde{\Phi}_A(f)\psi\rangle} \\ &= \overline{\int f \,\mathrm{d}\mu_{\psi,\varphi}} = \int f \,\mathrm{d}\mu_{\psi,\varphi} = \langle \psi, \tilde{\Phi}_A(f)\varphi\rangle.\end{aligned}$$

Exercise 1.36:

i) Let d be the distance in question. Since $\sigma(S)$ is closed, there exists a $\lambda_0 \in \sigma(S)$ such that $d = |\mu - \lambda_0|$. Assume now to the contrary that $\|\mu \operatorname{id} - S^{-1}\| < 1/d$. Then

$$\begin{aligned}\|\operatorname{id} - (S-\mu)^{-1}(S-\lambda_0)\| &= \|(S-\mu)^{-1}[(S-\mu) - S - \lambda_0)]\| \\ &\leq \|(S-\mu)^{-1}\| \|S - \mu - S + \lambda_0\| \\ &< \frac{1}{d} \cdot d = 1.\end{aligned}$$

By Lemma 79 it follows that $\operatorname{id} - (\operatorname{id} - (S-\mu)^{-1}(S-\lambda_0)) = (S - \mu)^{-1}(S-\lambda_0)$ is invertible. But then $(S - \lambda_0)$ must be invertible, as it is the composition of invertible operators:

$$(S - \lambda_0) = (S-\mu)((S-\mu)^{-1}(S-\lambda_0)).$$

This is a contradiction because $\lambda_0 \in \sigma(S)$.

ii) For $\mu \in \rho(S)$ define the function $r_\mu(s) := (\mu - s)^{-1}$, where $s \in \sigma(S)$. Then $r_\mu \in \mathcal{C}(\sigma(S))$ and for $r_\mu(S) \in \mathcal{L}(H)$ we obtain

$$(\mu - S) r_\mu(S) = ((\mu - \operatorname{id}) r_\mu)(S) = r_\mu(S)(\mu - S)$$

using the functional calculus for self-adjoint operators. This implies $r_\mu(S) = (\mu \operatorname{Id} - S)^{-1}$. Using the functional calculus one now concludes

$$\|(\mu \operatorname{Id} - S)^{-1}\| = \sup_{s\in\sigma(S)} |r_\lambda(s)| = \frac{1}{\inf_{s\in\sigma(S)}|\lambda - s|} = \frac{1}{\operatorname{dist}(\mu, \sigma(S))}.$$

iii) We denote with $\mathrm{GL}(X, Y)$ the set of continuous (and therefore bounded) bijective linear maps from X to Y. First we remember the following (Lemma 79):

If $A \in L(X)$ such that $\|A\| < 1$, then $I - A \in \mathrm{GL}(X)$ (where I denotes the identity on X) and $(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$.

Proof By norm properties, one has $\|A^k\| \leq \|A\|^k$ and therefore

$$\sum_{k=0}^{\infty} \|A\|^k \overset{\|A\|<1}{=} \frac{1}{1 - \|A\|}$$

with the geometric series. Using Weierstraß-M-test, it follows that $\sum_{k=0}^{\infty} A_k$ is an absolutely convergent series in the Banach space $L(X)$ and in particular, $B := \sum_{k=0}^{\infty} A^k$ is well-defined in $L(X)$. One easily sees $AB = BA = \sum_{k=1}^{\infty} A^k$, therefore we also have $(I - A)B = B(I - A) = I$ and this is equivalent to $(I - A)^{-1} = B$. □

Now we prove our desired result:

The set $\mathrm{GL}(X, Y)$ is open in $L(X, Y)$.

Proof. Let $T_0 \in \mathrm{GL}(X, Y)$ and $T \in L(X, Y)$ such that $\|T_0 - T\| < \|T_0^{-1}\|^{-1}$. We have

$$T = T_0 + T - T_0 = T_0[I + T_0^{-1}(T - T_0)].$$

Now we show: $I + T_0^{-1}(T - T_0) \in \mathrm{GL}(X)$. This follows from the estimate

$$\| - T_0^{-1}(T - T_0)\| \leq \|T_0^{-1}\| \|T - T_0\| < 1$$

by our preceding result. All in all, we have shown that the open ball in $L(X, Y)$ with center T_0 and radius $\|T_0^{-1}\|^{-1}$ is an element of $\mathrm{GL}(X, Y)$, so the proof is complete. □

EXERCISE 1.38: We first do some basic observations. Let $\mathscr{H}$ be a finite dimensional Hilbert space an A be a self-adjoint operator on $\mathscr{H}$. Since $\mathscr{H}$ is the only dense subspace, the condition of the existence of a cyclic vector $x \in \mathscr{H}$ for A is equivalent to

$$\mathscr{H} = \{A^n x : n = 0, 1, 2, \dots\}.$$

This expression can be also written as

$$\mathscr{H} = \{p(A)x : p \in \mathbb{C}[t]\}. \tag{A.10}$$

Fix an orthonormal basis for H, then there exists a unitary matrix V and a diagonal matrix D such that $A = VDV^\dagger$. Since we have $p(A) = Vp(D)V^\dagger$, (A.10) is equivalent to

$$\mathscr{H} = \{p(D)y : p \in \mathbb{C}[t]\}, \tag{A.11}$$

where $y = V^\dagger x$. Note that $\{p(D) : p \in \mathbb{C}[t]\}$ is a vector space (for $\mathbb{C}[t]$ is) and furthermore,

$$\dim\{p(D) : p \in \mathbb{C}[t]\} = \text{number of distinct eigenvalues of } A. \tag{A.12}$$

To see this, we remark that $p(D)$ is a diagonal matrix with diagonal elements $p(D_11), \dots, p(D_{nn})$, and that given $\lambda_1, \dots, \lambda_k$ distinct, the set $\{(p(\lambda_1), \dots, p(\lambda_k)) : p \in \mathbb{C}[t]\}$ equals $\mathbb{C}^k$.

"$\Longrightarrow$": Assume that equality holds in (A.10). Then the set on the right hand side of this equation has dimension n and by (A.12) A has n distinct eigenvalues.

"$\Longleftarrow$": If A has n distinct eigenvalues, choose y as the vector with all coordinates equal to 1 (the sum of all unit eigenvectors of A) and the set on the right hand side of (A.10) has dimension n, which implies that it is equal to $\mathscr{H}$.

To construct a measure μ on the spectrum of A and a unitary map $U : \mathscr{H} \to L^2(\sigma(A), d\mu)$ such that

$$UAU^{-1}f(\lambda) = \lambda f(\lambda)\,, \qquad \text{for all} \quad \lambda \in \sigma(A), \tag{A.13}$$

we set $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$,

$$\mu = \sum_{j=1}^{n} \delta(x - \lambda_j)\,.$$

(compare Example 1.28) and we identify functions $f \in L^2(\sigma(A), d\mu)$ with vectors $f = (f(\lambda_1), \ldots, f(\lambda_n)) \in \mathbb{C}^n$. Choosing an orthonormal basis of associated eigenvectors $w_1, \ldots w_n \in \mathscr{H}$, we define

$$U : \mathscr{H} \to L^2(\sigma(A), d\mu), \quad Uv(\lambda_j) = \langle w_j, v\rangle_{\mathscr{H}} \in \mathbb{C}\,.$$

Then U is bijective with $U^{-1} : \mathbb{C}^n \to \mathscr{H}$ given by $U^{-1}(a_1, \ldots, a_n) = \sum_{\ell=1}^{n} a_\ell w_\ell \in \mathscr{H}$. Moreover, U is unitary, since

$$\begin{aligned}\langle Ux, Uy\rangle_{L^2} &= \sum_{j=1}^{n} \overline{Ux(\lambda_j)}Uy(\lambda_j) = \sum_{j=1}^{n} \overline{\langle w_j, x\rangle_{\mathscr{H}}}\langle w_j, y\rangle_{\mathscr{H}} \\ &= \left\langle \sum_{j=1}^{n} \langle w_j, x\rangle_{\mathscr{H}} w_j, y \right\rangle_{\mathscr{H}} = \langle x, y\rangle_{\mathscr{H}}\,.\end{aligned}$$

To show (A.13), we write for $f \in L^2(\sigma(A), d\mu)$, using the spectral decomposition of A, i.e. $Ax = \sum_k \lambda_k \langle w_k, x\rangle w_j$,

$$\left(UAU^{-1}f\right)(\lambda_j) = \langle w_j, AU^{-1}f\rangle_{\mathscr{H}} = \langle w_j, \sum_k \lambda_k \langle w_k, U^{-1}f\rangle w_j\rangle_{\mathscr{H}}$$

$$= \sum_k \lambda_k \langle w_k, U^{-1} f \rangle \langle w_j, w_k \rangle_{\mathscr{H}} = \lambda_j \left\langle w_j, \sum_\ell f(\lambda_\ell) w_\ell \right\rangle_{\mathscr{H}}$$
$$= \sum_\ell \lambda_j f(\lambda_\ell) \langle w_j, w_e ll \rangle_{\mathscr{H}} = \lambda_j f(\lambda_j)\,.$$

Exercise 2.34: We prove the statement in the following for the case $n = 1$. After doing this, we explain how the argument can be easily extended for the case of general n.

We begin with a

Lemma A.50

Every $f \in \mathcal{C}_\infty(\mathbb{R})$ is a uniform limit of continuous functions with compact support.

Proof. Let $f \in \mathcal{C}_\infty(\mathbb{R})$. Given $n \in \mathbb{N}$, there exists an $N > 0$ with $|f(x)| < 1/n$ when $|x| > N$. Take g_n continuous with $0 \leq g_n \leq 1$ such that $g_n(x) = 1$ when $|x| \leq N$ and $g_n = 0$ when $|x| > N + 1$ (recall that this exists by Urysohn's lemma). Then $f g_n \in \mathcal{C}_c(\mathbb{R})$ and

$$|f - f g_n| = |f(1 - g_n)| < 1/n \tag{A.14}$$

implying $f g_n \longrightarrow f$ uniformly. □

Now let $f \in \mathcal{C}_0(\mathbb{R})$. By the previous lemma we can assume without loss of generality that f is continuous with compact support. We now have to show that there is a sequence of functions in $\mathcal{C}_0^\infty(\mathbb{R})$ which converges uniformly to f. Therefore, we start with

$$h_0(x) = \begin{cases} \mathrm{e}^{-1/x^2}, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{A.15}$$

Now notice that $h(x) = h_0(x)h_0(1-x) \in \mathcal{C}_0^\infty$ (induction on the order of the derivatives) with support in $[0,1]$. One can normalize h such that $\int_{\mathbb{R}} h = 1$. For $\varepsilon > 0$ define the mollifiers

$$h_\varepsilon(x) = \varepsilon^{-1} h(x\varepsilon^{-1}) \tag{A.16}$$

and recall $\int_{\mathbb{R}} h_\varepsilon = \int_{\mathbb{R}} h = 1$. Now we form the convolutions

$$f_\varepsilon(x) = \int_{\mathbb{R}} f(t) h_\varepsilon(x-t)\,\mathrm{d}t\,. \tag{A.17}$$

Because $h_\varepsilon \in \mathcal{C}^\infty$ it easily follows that $f_\varepsilon \in \mathcal{C}^\infty$, and since f and h_ε have compact support, so does f_ε.

Given $\varepsilon > 0$, as f is continuous with compact support, it is uniformly continuous. So given $c > 0$, there exists a $\delta > 0$ such that $|f(x) - f(y)| < c$ if $|x - y| < \delta$. Then

$$\begin{aligned}
|f_\varepsilon(x) - f(x)| &= \left| \int_{\mathbb{R}} [f(x-t) - f(x)] h_\varepsilon(t)\,\mathrm{d}t \right| && \text{(A.18)}\\
&\le \int_{\mathbb{R}} |f(x-t) - f(x)| h_\varepsilon(t)\,\mathrm{d}t && \text{(A.19)}\\
&\le c \int_{|t|<\delta} h_\varepsilon(t)\,\mathrm{d}t + 2\|f\|_\infty \int_{|t|\ge\delta} h_\varepsilon(t)\,\mathrm{d}t && \text{(A.20)}\\
&\le c + 2\|f\|_\infty \int_{|t|\ge\delta} h_\varepsilon(t)\,\mathrm{d}t\,. && \text{(A.21)}
\end{aligned}$$

With δ fixed, the last integral goes to zero as $\varepsilon \downarrow 0$ (because the support of h_ε is contained in $[0,\varepsilon]$). Thus

$$\limsup_{\varepsilon\to 0} |f_\varepsilon(x) - f(x)| \le c \tag{A.22}$$

for all x and all $c > 0$. This shows $\|f_\varepsilon - f\|_\infty \longrightarrow 0$.

Finally, we want to mention how one can extend the above arguments for $\mathbb{R}^n$. This is easily done, since

$$h_n(x_1, \ldots, x_n) = h(x_1) \cdots h(x_n) \tag{A.23}$$

can be taken.

Exercise 2.35: "$\Longrightarrow$":

$$\langle x, Tx \rangle = \langle Tx, x \rangle = \overline{\langle x, Tx \rangle}. \tag{A.24}$$

"$\Longleftarrow$": Consider for $\lambda \in \mathbb{C}$ the real number

$$\langle x + \lambda y, T(x + \lambda y) \rangle = \langle x, Tx \rangle + \lambda \langle y, Tx \rangle + \overline{\lambda} \langle x, Ty \rangle + |\lambda|^2 \langle y, Ty \rangle.$$

Taking the complex conjugate yields

$$\langle x + \lambda y, T(x + \lambda y) \rangle = \langle Tx, Tx \rangle + \overline{\lambda} \langle Tx, y \rangle + \lambda \langle Ty, x \rangle + |\lambda|^2 \langle Ty, y \rangle.$$

Setting $\lambda = 1$ and $\lambda = -\mathrm{i}$ implies

$$\begin{aligned} \langle y, Tx \rangle + \langle x, Ty \rangle &= \langle Tx, y \rangle + \langle Ty, x \rangle, \\ \langle y, Tx \rangle - \langle x, Ty \rangle &= -\langle Tx, y \rangle + \langle Ty, x \rangle. \end{aligned}$$

Adding both equations reveals $\langle Ty, x \rangle = \langle y, Tx \rangle$.

Exercise 2.36: Let $y \in \mathscr{D}(T^* S^*)$ and let $x \in \mathscr{D}(ST)$. Then $S^* y \in \mathscr{D}(T^*)$ and $x \in \mathscr{D}(T)$, so

$$\langle Tx, S^* y \rangle = \langle x, T^* S^* y \rangle. \tag{A.25}$$

On the other hand, $y \in \mathscr{D}(S^*)$, so

$$\langle STx, y \rangle = \langle Tx, S^* y \rangle. \tag{A.26}$$

Hence

$$\langle STx, y\rangle = \langle x, T^*S^*y\rangle, \tag{A.27}$$

which implies that $(ST)^*y = T^*S^*y$ for each $y \in \mathscr{D}(T^*S^*)$, that is, $T^*S^* \subset (ST)^*$.

Suppose now that $S \in \mathcal{L}(\mathscr{H})$, hence $S^* \in \mathcal{L}(\mathscr{H})$, for which $\mathscr{D}(S^*) = \mathscr{H}$. Let $y \in \mathscr{D}((ST)^*)$. For $x \in \mathscr{D}(ST)$,

$$\langle Tx, S^*y\rangle = \langle STx, y\rangle = \langle x, (ST)^*y\rangle. \tag{A.28}$$

This implies that $S^*y \in \mathscr{D}(T^*)$ and hence $y \in \mathscr{D}(T^*S^*)$, showing

$$\mathscr{D}((ST)^*) = \mathscr{D}(T^*S^*). \tag{A.29}$$

Exercise 2.37: We start showing:

Claim: Let $A : X \supset \mathscr{D}(A) \longrightarrow Y$ (with X, Y Banach spaces) be closed and $B \in \mathcal{L}(X, Y)$. Then $A + B$ with $\mathscr{D}(A + B) = \mathscr{D}(A)$ is closed.
To see this, let $x_n \in \mathscr{D}(A + B), n \in \mathbb{N}$, and $x \in X, y \in Y$ such that $x_n \longrightarrow x$ in X and $(A + B)x_n = Ax_n + Bx_n \longrightarrow y$ in Y for $n \longrightarrow \infty$. Since B is bounded, $Bx = \lim Bx_n$ exists, implying $Ax_n \longrightarrow y - Bx$ for $n \longrightarrow \infty$. Because A is closed, it follows that $x \in \mathscr{D}(A) = \mathscr{D}(A+B)$ and $Ax = y - Bx$, or $Ax + Bx = y$.

Now let $\lambda \in \mathbb{C}$ be such that $(\lambda - T) : \mathscr{D}(T) \to \mathscr{H}$ is bijective and has bounded inverse $(\lambda - T)^{-1} : \mathscr{H} \to \mathscr{D}(T) \subset \mathscr{H}$. Then by the Closed Graph Theorem[1], the graph $\Gamma((\lambda - T)^{-1})$ is closed.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{H}$ such that $x_n \to x \in \mathscr{H}$ and $y_n := (\lambda - T)^{-1}x_n \to y$. Then $(\lambda - T)^{-1}x = y$ and thus $y_n, y \in \operatorname{Ran}(\lambda - T)^{-1} =$

$\mathscr{D}(\lambda - T)$ and $(\lambda - T)y = x$. Therefore, for any sequence (y_n) in $\mathscr{D}(\lambda - T)$ such that $y_n \to y$ and $(\lambda - T)y_n = x_n \to x$, it follows that $y \in \mathscr{D}(\lambda - T)$ and $(\lambda - T)y = x$. This shows that $(\lambda - T)$ is closed. Using the above claim for $B = -\lambda\,\mathrm{Id}$ shows that T is closed.

Exercise 2.40: Let A be a densely defined, closed, symmetric operator on $\mathscr{H}$ with $\sigma(A) \subset \mathbb{R}$. Then $\pm i \in \rho(A)$ and thus $(A \mp i\,\mathrm{Id}) : \mathscr{D}(A) \to \mathscr{H}$ is bijective with bounded inverse. This implies $\mathrm{Ran}(A \mp i\,\mathrm{Id}) = \mathscr{H}$. It follows from Theorem 2.21 that A is self-adjoint.

Exercise 2.85:

i) We want to show that $T_{s+t}f = \gamma_{s+t} * f = \gamma_s * (\gamma_t * f) = T_sT_tf$ holds for all f and $s, t > 0$. Since convolution is associative, it therefore suffices to show that $\gamma_{s+t} = \gamma_s * \gamma_t$ holds for all $s, t > 0$. We remark that this identity can be obtained by elementary multivariable integral calculus. We however give here a slightly other proof using facts from the theory of partial differential equations.

   Recall first that γ is a so-called fundamental solution of the heat-equation. This means that given f bounded and continuous on $\mathbb{R}^n$, the function

$$u(t, x) = \gamma_t * f(x)$$

   is $\mathcal{C}^\infty$ in $\mathbb{R}^+ \times \mathbb{R}^n$, satisfies the heat equation

$$\partial_t u = \Delta u$$

and the initial data $u|_{t=0} = f$ in the sense that

$$u(t,x) \longrightarrow f(x) \quad \text{as} \quad t \longrightarrow 0$$

locally uniformly in x (in the distributional sense, the convolution ist δ_0). We now use this in order to show $\gamma_{t-s}*\gamma_s = \gamma_t$ for all $0 < s < t$. Let f be continuous in $\mathbb{R}^n$ with compact support. Then $u_t = \gamma_t * f$ solves the bounded Cauchy problem with the initial function f. Consider now the Cauchy problem with the initial function u_s. Then u_t gives the bounded solution to this problem a time $t-s$. On the other hand, the solution at time $t-s$ is given by $\gamma_{t-s} * u_s$. Hence, we obtain the identity

$$u_t = \gamma_{t-s} * u_s,$$

that is

$$\gamma_t * f = \gamma_{t-s} * (\gamma_s * f).$$

Because convolution is associative, we have

$$\gamma_{t-s} * (\gamma_s * f) = (\gamma_{t-s} * \gamma_s) * f,$$

whence

$$\gamma_t * f = (\gamma_{t-s} * \gamma_s) * f.$$

Since this is true for all f, we conclude $\gamma_t = \gamma_{t-s} * \gamma_s$.

ii) Let

$$\gamma(x,y,t) = \frac{1}{(4\pi t)^{d/2}} \cdot \mathrm{e}^{-|x-y|^2/(4t)}$$

and set

$$u(x,t) = \int_{\mathbb{R}^d} \gamma(x,y,t) f(y) \, \mathrm{d}y \, .$$

Then $(T_t f)(x) = u(x,t)$. The statement $\lim_{t\to 0} T_t f = f$ follows if we show that $\lim_{t\to 0} u(x,t) = f(x)$ for every $x \in \mathbb{R}^d$ and $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ bounded and continuous.

To this end, we use $\int_{\mathbb{R}^d} \gamma(x,y,t)\,\mathrm{d}y = 1$ and calculate

$$\begin{aligned}
|f(x) - u(x,t)| &= \left| f(x) - \int_{\mathbb{R}^d} \gamma(x,y,t) f(y)\,\mathrm{d}y \right| \\
&= \left| \int_{\mathbb{R}^d} \gamma(x,y,t)(f(x) - f(y))\,\mathrm{d}y \right| \\
&= \left| \frac{1}{(4\pi t)^{d/2}} \int_0^\infty \mathrm{e}^{-r^2/(4t)} r^{d-1} \int_{S^{d-1}} \big(f(x) - f(x + r\xi)\big)\,\mathrm{d}o(\xi)\,\mathrm{d}r \right| \\
&= \left| \frac{1}{\pi^{d/2}} \int_0^\infty \mathrm{e}^{-s^2} s^{d-1} \int_{S^{d-1}} \big(f(x) - f(x + 2\sqrt{t} s\xi)\big)\,\mathrm{d}o(\xi)\,\mathrm{d}s \right| \\
&= \left| \cdots \int_0^M \cdots + \cdots \int_M^\infty \cdots \right| \\
&\le \sup_{y \in B(x, 2\sqrt{t}M)} |f(x) - f(y)| + 2 \sup_{\mathbb{R}^d} |f| \frac{d\omega_d}{\pi^{d/2}} \int_M^\infty \mathrm{e}^{-s^2} s^{d-1}\,\mathrm{d}s\,.
\end{aligned}$$

Given $\varepsilon > 0$, we first choose M so large that the second summand is less then $\varepsilon/2$ and we then choose $t_0 > 0$ so small that for all t with $0 < t < t_0$, the first summand is less than $\varepsilon/2$ as well. This implies the continuity.

iii) We recall first that in Analysis it was shown that $\|\gamma_t\| = 1$ (use polar coordinates and Fubini). Then an application of Young's inequality with $r = p$ and $q = 1$ shows

$$\|T_t f\|_p = \|\gamma_t * f\| \le \|\gamma_t\|_1 \|f\|_p = \|f\|_p,$$

implying $\|T_t\| \le 1$.

Exercise 2.87: We want to apply the theorem of Lumer-Philipps (Theorem 2.68). First of all, A is densely defined since $\mathscr{D}(A)$ contains $D(0,1)$ and

$D(0,1)$ is dense in X (here, $D(0,1)$ is the test function space on $(0,1)$ and the density follows by a similar argument as given in Exercise 2.34).

Next, we want to show that A is dissipative. Take $f \in \mathscr{D}(A)$. Then f is continuous on a compact set and hence it exists some x_0 such that $|f(x_0)| = \|f\|_\infty$. Set $a = \overline{f(x_0)}$ and consider the functional $\ell(f) = a\delta_{x_0}(f) = af(x_0)$. Then $l \in J(f)$ and

$$\operatorname{Re} \ell(f'') = \operatorname{Re} a(f'')(x_0) \leq 0$$

since the real valued function $\operatorname{Re} af$ takes its maximum at x_0.

Finally, we want to show that $(\mathrm{Id} - A)$ is surjective. This is equivalent to the statement that the boundary value problem

$$f - f'' = g, \qquad f(0) = f(1) = 0$$

is solvable for each $g \in X$. This is proven e.g. in [Wa].

Bibliography

[Ba] Heinz Bauer: *Measure Theory* de Gruyter, Berlin (1992)

[B] Adam Bobrowski: *Functional Analysis for probability and stochastic Processes* Cambridge University Press (2007)

[BR] O. Bratelli, D.W. Robinson: *Operator Algebras and Quantum Statistical Mechanics Vol. 1: C^*- and W^*-Algebras, Symmetry Groups, Decomposition of States* Springer Science Business Media (2012)

[F] G.B. Folland: *Real Analysis: Modern Techniques and their Application*, John Wiley (1999)

[Ha] M. Haase: *Lectures on Functional Calculus. 21st International Internet Seminar* Universität Kiel (2018)

[H] Bernard Helffer: *Spectral Theory and its Applications)*, Cambridge University Press (2013).

[HS] F. Hirzebruch, W. Scharlau: *Einführung in die Funktionalanalysis* BI-Wissenschaftsverlag (1971)

[RS] Micheal Reed, Barry Simon: *Methods of Modern Mathematical Physics 1 (Functional Analysis) and 2 (Fourier Analysis, Self-Adjointness)*, Academic Press, London (1980).

[R] Walter Rudin: *Functional Analysis*, McGraw-Hill, 2.ed. (1991)

[R2] Walter Rudin: *Real and complex Analysis* McGraw-Hill, 3.ed. (1987)

[Wa] Wolfgang Walter: *Gewöhnliche Differentialgleichungen* Springer, 7.ed, Berlin, (2000)

[W] Dirk Werner: *Funktionalanalysis*, Springer, 5.ed, Berlin (2005).

[1] R. Whitley: *The Spectral Theorem for a Normal Operator* The American Mathematical Monthly, 75(8), pp. 856-861 (1968). doi:10.2307/2314336