# Spectral Theory
# Summer Term 2011

Peer Christian Kunstmann
Karlsruhe Institute of Technology (KIT)
Institute for Analysis
Kaiserstr. 89, D – 76128 Karlsruhe, Germany
e-mail: peer.kunstmann@kit.edu

These lecture notes contain a summary of the course. They are intended to be used parallel to the lecture. They are **not** meant to be used as a text book or for self studies. Attending the lectures cannot be substituted by reading these lecture notes.

# 1 The spectrum of closed operators

**Notation:** i) When nothing else is said, $X$, $Y$, and $Z$ are complex Banach spaces.

ii) $\mathcal{L}(X,Y) := \{T : X \to Y : \; T \text{ is linear and bounded}\}$. For $T \in \mathcal{L}(X,Y)$ we have the *operator norm*

$$\|T\| := \sup_{\|x\|_X \leq 1} \|Tx\|_Y,$$

and $(\mathcal{L}(X,Y), \|\cdot\|)$ is a Banach space.

iii) A *linear operator* $A$ from $X$ to $Y$ is a linear operator $A : D(A) \to Y$ where $D(A)$, the *domain* of $A$, is a linear subspace of $X$. We also write $A : X \supseteq D(A) \to Y$. Further we use

$$\begin{aligned} R(A) &:= \{Ax : x \in D(A)\} \text{ \textit{range} of } A, \\ N(A) &:= \{x \in D(A) : Ax = 0\} \text{ \textit{kernel} (\textit{null space}) of } A. \end{aligned}$$

We recall the notion of a closed operator.

**1.1. Definition:** A linear operator $A$ fro $X$ to $Y$ is called *closed* if its graph

$$\mathrm{gr}(A) := \{(x, Ax) : x \in D(A)\} \subseteq X \times Y$$

is a closed subspace of $X \times Y$.

**Recall:** Since $A$ is linear, its graph $\mathrm{gr}(A)$ is a linear subspace of $X \times Y$. The space $X \times Y$, equipped with the norm given by $\|(x,y)\| := \|x\|_X + \|y\|_Y$ is a Banach space.

**Closed Graph Theorem:** Let $A : X \to Y$ be a linear operator. Then

$$A \text{ is closed} \quad \Longleftrightarrow \quad A \text{ is bounded.}$$

**1.2. Lemma:** Let $A : X \supseteq D(A) \to Y$ be a linear operator. The following are equivalent:

(i) $A$ is closed,

(ii) $D(A)$ is a Banach space for the *graph norm* given by

$$\|x\|_A := \|x\|_X + \|Ax\|_Y.$$

(iii) For every sequence $(x_n)_{n\in\mathbb{N}}$ in $D(A)$ and $x \in X$, $y \in Y$ such that $x_n \to x$ in $X$ and $Ax_n \to y$ in $Y$ one has $x \in D(A)$ and $Ax = y$.

**Notation:** We write $[D(A)] := (D(A), \|\cdot\|_A)$. Then obviously $A \in \mathcal{L}([D(A)], Y)$.

*Proof.* By definition, closedness of $A$ is equivalent to closedness of $\mathrm{gr}(A)$, and since $X \times Y$ is a Banach space, this is equivalent to completeness of $\mathrm{gr}(A)$. Since the map

$$J : [D(A)] \to \mathrm{gr}(A), x \mapsto (x, Ax),$$

is linear, bijective, and isometric, the latter is equivalent to $[D(A)]$ being a Banach space. On the other hand, closedness of $\mathrm{gr}(A)$ means

> for any sequence $(x_n, Ax_n)$ in $\mathrm{gr}(A)$ and $(x, y) \in X \times Y$ such that $(x_n, Ax_n) \to (x, y)$ one has $(x, y) \in \mathrm{gr}(A)$,

which clearly is equivalent to (iii). □

**Examples:** 1) Let $X = Y = (C[0,1], \|\cdot\|_\infty)$, $A := \frac{d}{dx}$ with $D(A) := C^1[0,1]$. Then $A$ is closed: If $f_n \to f$ and $f_n' \to g$ with respect to $\|\cdot\|_\infty$, then $f \in C^1[0,1]$ and $f' = g$ [cf. Calculus I, for a proof write, for $t \in [0,1]$ fixed,

$$f_n(t) = f_n(0) + \int_0^t f_n'(s)\,ds;$$

by assumption, $f_n(t) \to f(t)$, $f_n(0) \to f(0)$, and (by uniform convergence $f_n' \to g$) also $\int_0^t f_n'(s)\,ds \to \int_0^t g(s)\,ds$; hence

$$f(t) = f(0) + \int_0^t g(s)\,ds, \quad t \in [0,1],$$

which implies $f \in C^1$, $f' = g$.]

2) Let $X = Y = L^1[0,1]$, $A := \frac{d}{dx}$ with $D(A) = C^1[0,1]$. Then $A$ is *not* closed: Take $g \in L^1[0,1] \setminus C[0,1]$, say $g = 1_{[1/2,1]} - 1_{[0,1/2)}$, and approximate $g$ in $\|\cdot\|_1$ by a sequence $(g_n)$ in $C[0,1]$. Then let

$$f(t) := |t - \frac{1}{2}| \quad \text{and} \quad f_n(t) := \frac{1}{2} + \int_0^t g_n(s)\,ds.$$

We obtain $f_n \in C^1[0,1]$ and $f_n' = g_n \to g$ in $\|\cdot\|_1$, $f_n \to f$ in $\|\cdot\|_\infty$, hence in particular $f_n \to f$ in $\|\cdot\|_1$. But $f \notin D(A)$. [As a concrete $g_n$ one can take

$$g_n(t) := \begin{cases} -1 & ,t \in [0, \frac{1}{2} - \frac{1}{n}) \\ n(t - \frac{1}{2} & ,|t - \frac{1}{2}| \leq \frac{1}{n} \\ 1 & ,t \in (\frac{1}{2} + \frac{1}{n}, 1] \end{cases}, \quad n \geq 3.]$$

3) Let $X = L^1[0,1]$, $Y = \mathbb{C}$, $Af := f(0)$ with $D(A) = C[0,1]$. Then $A$ is not closed: Take $f_n(t) := (1 - nt)1_{[0,1/n]}(t)$. Then $f_n \to 0 =: f$ in $\|\cdot\|_1$ and $Af_n = f_n(0) = 1 \to 1$. Here $f \in D(A)$, but $Af = f(0) = 0 \neq 1$.

**Comment:** $L^1[0,1]$ is a Banach space whose elements are equivalence classes of functions that coincide almost everywhere (a.e.). Strictly speaking, $C[0,1]$ should be considered here as the set of equivalence classes in $L^1[0,1]$ that contain a continuous function. This continuous function is then unique within its class. $A$ evaluates the unique continuous function in the class at 0. In fact, a similar interpretation has to be given to $C^1[0,1]$ as a subspace of $L^1[0,1]$ in Example 2) above.

4) Let $p \in [1,\infty]$, $X = Y = L^p(\Omega,\mu)$ where $(\Omega,\mu)$ is a $\sigma$-finite measure space. Let $m : \Omega \to \mathbb{C}$ be measureable and

$$Af := mf \text{ with } D(A) := \{f \in L^p(\Omega,\mu) : mf \in L^p(\Omega,\mu)\}.$$

Then $A$ is closed: If $f_n \to f$ and $mf_n \to g$ with respect to $\|\cdot\|_p$, then we find subsequences such that $f_{k(n)} \to f$ and $mf_{k(l(n))} \to g$ almost everywhere. Hence $mf = g$ a.e., and by $g \in L^p$ we have $f \in D(A)$, $Af = g$.

**Notation:** Let $A$ and $B$ be linear operators form $X$ to $Y$ and $C$ be a linear operator from $Y$ to $Z$. We define

$$\begin{aligned} A+B \quad &\text{by} \quad (A+B)x := Ax + Bx \text{ for } x \in D(A+B) := D(A) \cap D(B), \\ CA \quad &\text{by} \quad (CA)x := C(Ax) \text{ for } x \in D(CA) := \{\tilde{x} \in D(A) : A\tilde{x} \in D(C)\}. \end{aligned}$$

$A+B$ is a linear operator from $X$ to $Y$, and $CA$ is a linear operator from $X$ to $Z$.

Sum and product of linear operators are associative, but not distributive in general.

**1.3. Lemma (Properties of closed operators):** Let $A$ be a closed linear operator from $X$ to $Y$, $T \in \mathcal{L}(X,Y)$, and $S \in \mathcal{L}(Z,X)$. Then:

(a) $B := A + T$ is closed (here $D(B) = D(A)$).
(b) $C := AS$ is closed (here $D(C) = \{z \in Z : Sz \in D(A)\}$).
(c) If $A$ is injective then $A^{-1}$ is closed (here $D(A^{-1}) = R(A)$).
(d) If $R$ is injective and closed from $Y$ to $Z$ such that $R^{-1} \in \mathcal{L}(Z,Y)$ then $D := RA$ is closed (here $D(D) = \{x \in D(A) : Ax \in D(R)\}$).

*Proof.* (a) Take $(x_n)$ in $D(A)$ such that $x_n \to x$, $Bx_n \to y$. Since $T$ is bounded we have $Tx_n \to Tx$. Now $(A+T)x_n \to y$ implies that $Ax_n \to y - Tx$. By closedness of $A$ we obtain $x \in D(A)$ and $Ax = y - Tx$, i.e. $Bx = y$.

(b) Let $(z_n)$ be a sequence in $D(C)$ such that $z_n \to z$, $Cz_n \to y$. Since $S$ is bounded we have $Sz_n \to Sz$, $A(Sz_n) \to y$. But $(Sz_n)$ is a sequence in $D(A)$, thus $Sz \in D(A)$, $A(Sz) = y$ by closedness of $A$. We have shown $z \in D(C)$, $Cz = y$.

(c) follows from $\mathrm{gr}(A^{-1}) = \{(y,x) : (x,y) \in \mathrm{gr}(A)\}$. End Lect.1

(d) Let $(x_n)$ be a sequence in $D(D)$ such that $x_n \to x$ and $Dx_n \to z$. By $R^{-1} \in \mathcal{L}(Z, Y)$ we have $Ax_n = R^{-1}Dx_n \to R^{-1}z$. Since $A$ is closed and $D(D) \subset D(A)$, we obtain $x \in D(A)$ and $Ax = R^{-1}z \in R(R^{-1}) = D(R)$. Hence $x \in D(D)$ and $Dx = z$. □

In general, sums or products of closed operators are not closed.

**Example:** Take $X = l^1$ and, for $x = (x_n)$ define $A$ by $(Ax)_n = nx_{n-1}$ if $n$ is even and $(Ax)_n = 0$ if $n$ is odd, and let $D(A) := \{x \in l^1 : Ax \in l^1\}$.

Then $A$ is closed: If $x^{(n)}$ is a sequence in $D(A)$ such that $x^{(n)} \to x$ and $Ax^{(n)} \to y$, then $x_k = \lim_n x_k^{(n)}$ for each $k$, and $y_k = \lim_n kx_{k-1}^{(n)}$ for even $k$, $y_k = 0$ for odd $k$. We conclude that $y_k = kx_{k-1}$ for even $k$. Hence we have $Ax = y$, and $x \in D(A)$ by $y \in l^1$.

However, $B := A + (-A) = 0$ with $D(B) = D(A)$ and $C := AA = 0$ with $D(C) = D(A)$ are not closed (otherwise $D(A)$ would be a closed subspace of $l^1$, but $D(A) \neq l^1$ is dense in $l^1$).

**1.4. Definition (spectrum and resolvent):** Let $A : X \supseteq D(A) \to X$ be a linear operator. Then

$$\rho(A) := \{\lambda \in \mathbb{C} : \lambda I - A : D(A) \to X \text{ is bijective and } (\lambda I - A)^{-1} \in \mathcal{L}(X)\}$$

is called the *resolvent set* of $A$ and

$$\sigma(A) := \mathbb{C} \setminus \rho(A)$$

is the *spectrum* of $A$. The map

$$\rho(A) \to \mathcal{L}(X), \quad \lambda \mapsto (\lambda I - A)^{-1},$$

is called the *resolvent* of $A$, and for $\lambda \in \rho(A)$, the operator

$$R(\lambda, A) := (\lambda I - A)^{-1}$$

is called the *resolvent operator* (at $\lambda$).

**Remark:** It is common to write $\lambda - A$, $(\lambda - A)^{-1}$ instead of $\lambda I - A$, $(\lambda I - A)^{-1}$. For $\lambda \in \rho(A)$ we have

$$(\lambda - A)R(\lambda, A) = I_X, \quad R(\lambda, A)(\lambda - A) = I_{D(A)}.$$

**Further Remarks:** (a) If $\rho(A) \neq \emptyset$ then $A$ is closed.

(b) If $A$ is closed then $\rho(A) = \{\lambda \in \mathbb{C} : \lambda I - A : D(A) \to X \text{ is bijective}\}$.

*Proof.* (a) We find $\lambda_0 \in \rho(A)$. Then $-R(\lambda_0, A) = -(\lambda_0 - A)^{-1}$ is closed. By 1.3(c), $A - \lambda_0 I$ is closed, and by 1.3(a), $A = (A - \lambda_0 I) + \lambda_0 I$ is closed.

(b) "$\subseteq$" is clear. For "$\supseteq$" assume that $\lambda - A : D(A) \to X$ is bijective. By 1.3(a), $\lambda - A$ is closed, and by 1.3(c), $(\lambda - A)^{-1} : X \to X$ is closed. By the Closed Graph Theorem we obtain $(\lambda - A)^{-1} \in \mathcal{L}(X)$, i.e. $\lambda \in \rho(A)$. $\square$

**Examples:** 1) $X = C[0,1]$, $A = \frac{d}{dx}$, $D(A) = C^1[0,1]$. Let $\lambda \in \mathbb{C}$. Then $f := e^{\lambda(\cdot)} \in D(A)$ and $Af = f' = \lambda e^{\lambda(\cdot)} = \lambda f$, hence $\lambda - A : D(A) \to X$ is not injective. We have shown $\sigma(A) = \mathbb{C}$.

2) $X = C[0,1]$, $A_0 = \frac{d}{dx}$, $D(A_0) = \{f \in C^1[0,1] : f(0) = 0\}$. Let $\lambda \in \mathbb{C}$, $g \in C[0,1]$. By ODE-Theory there is a unique solution $f \in C^1[0,1]$ to the initial value problem

$$\lambda f - f' = g \text{ in } [0,1]; \quad f(0) = 0,$$

given by $f(t) := -\int_0^t e^{\lambda(t-s)} g(s)\, ds$, $t \in [0,1]$. Thus $\lambda - A_0 : D(A_0) \to X$ is bijective. Since $A_0$ is closed (!), we have shown $\sigma(A_0) = \emptyset$, $\rho(A) = \mathbb{C}$. Moreover, we have

$$|f(t)| \le \int_0^t e^{Re\,\lambda(t-s)} |g(s)|\, ds \le \int_0^1 e^{Re\,\lambda s}\, ds\, \|g\|_\infty = \begin{cases} \frac{e^{Re\,\lambda} - 1}{Re\,\lambda} \|g\|_\infty & , Re\,\lambda \neq 0 \\ \|g\|_\infty & , Re\,\lambda = 0 \end{cases},$$

which means

$$\|R(\lambda, A_0)\| \le \begin{cases} \frac{e^{Re\,\lambda} - 1}{Re\,\lambda} & , Re\,\lambda \neq 0 \\ 1 & , Re\,\lambda = 0 \end{cases}.$$

Notice that this implies $\sigma(A_0) = \emptyset$ without using closedness of $A_0$.

**1.5. Proposition (Properties of the resolvent):** Let $A$ be a closed linear operator in $X$. Then:

(a) For all $\lambda, \mu \in \rho(A)$:

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A) \quad (\textit{resolvent equation}).$$

In particular, $R(\lambda, A)R(\mu, A) = R(\mu, A)R(\lambda, A)$ for all $\lambda, \mu \in \rho(A)$.

(b) If $\lambda \in \rho(A)$ then $B(\lambda, 1/\|R(\lambda, A)\|) \subset \rho(A)$ and

$$R(\mu, A) = \sum_{k=0}^{\infty} (-1)^k R(\lambda, A)^{k+1} (\mu - \lambda)^k, \quad \mu \in B(\lambda, 1/\|R(\lambda, A)\|),$$

where the series converges in operator norm. In particular, $\rho(A)$ is open and $\sigma(A)$ is closed, and, for any $\lambda \in \rho(A)$,

$$\|R(\lambda, A)\| \ge \frac{1}{d(\lambda, \sigma(A))}.$$

*Proof.* (a) Write

$$R(\lambda, A) - R(\mu, A) = R(\lambda, A)\underbrace{(\mu - A)R(\mu, A)}_{=I_X} - \underbrace{R(\lambda, A)(\lambda - A)}_{=I_{D(A)}}R(\mu, A) = (\mu - \lambda)R(\lambda, A)R(\mu, A).$$

(b) For $\mu \in \mathbb{C}$ we have

$$\mu - A = \mu - \lambda + \lambda - A = ((\mu - \lambda)R(\lambda, A) + I)(\lambda - A).$$

For $|\mu - \lambda| < 1/\|R(\lambda, A)\|$ the operator

$$I + (\mu - \lambda)R(\lambda, A) \in \mathcal{L}(X)$$

is invertible in $\mathcal{L}(X)$ by a Neumann series and

$$(I + (\mu - \lambda)R(\lambda, A))^{-1} = \sum_{k=0}^{\infty} (-1)^k R(\lambda, A)^k (\mu - \lambda)^k.$$

We obtain for these $\mu$ that $\mu - A : D(A) \to X$ is bijective and

$$(\mu - A)^{-1} = R(\lambda, A)(I + (\mu - \lambda)R(\lambda, A))^{-1} = \sum_{k=0}^{\infty} (-1)^k R(\lambda, A)^{k+1} (\mu - \lambda)^k$$

as claimed. □

End Lect.2

— — **Detour: Analyticity** — —

**Definition:** Let $\Omega \subset \mathbb{C}$ be open and $X$ a complex Banach space. A function $f : \Omega \to X$ is called *analytic* (or *holomorphic*) in $\Omega$ if, for every $z_0 \in \Omega$ the limit

$$f'(z_0) := \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists. In this case, the function $f' : \Omega \to X$ is called the (complex) *derivative* of $f$.

**Remark:** A holomorphic function is continuous.

**Power series:** A (formal) *power series* in $X$ has the form

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

where $(a_k)_{k \in \mathbb{N}_0}$ is a sequence in $X$ and $z_0 \in \mathbb{C}$.

**Radius of convergence:** Define $R \in [0, \infty]$ by $\frac{1}{R} := \limsup_{k \to \infty} \|a_k\|_X^{1/k}$. Then the power series above converges absolutely and uniformly on compact subsets of $B(z_0, R)$ and defines a holomorphic function $f$ in $B(z_0, R)$ which satisfies

$$f^{(k)}(z_0) = k! a_k, \quad k \in \mathbb{N}_0.$$

Moreover, for $\tilde{R} > R$, $f$ cannot be extended to a holomorphic function on $B(z_0, \tilde{R})$.

*Proof.* Let $\varepsilon \in (0, R/2)$ and choose $k_0 \in \mathbb{N}$ such that

$$\|a_k\|^{1/k} \le \frac{1}{\tilde{R} - \varepsilon} \text{ for all } k \ge k_0.$$

For $|z - z_0| \le \tilde{R} - 2\varepsilon$ we then have

$$\sum_{k=k_0}^{\infty} \|a_k\| |z - z_0|^k \le \sum_{k=k_0}^{\infty} \Big( \frac{\tilde{R} - 2\varepsilon}{\tilde{R} - \varepsilon} \Big)^k < \infty.$$

Differentiability follows as for scalar-valued power series.

Now let $\tilde{R} \ge R$ such that $f$ has an analytic extension $g : B(z_0, \tilde{R}) \to X$. For arbitrary $\varphi \in X'$, the function $\varphi \circ g : B(z_0, \tilde{R}) \to \mathbb{C}$ is analytic and

$$\sum_{k=0}^{\infty} \varphi(a_k)(z - z_0)^k$$

converges for all $|z - z_0| < \tilde{R}$. Hence for any $\delta \in (0, \tilde{R})$,

$$\sup_k |\varphi(a_k)|(\tilde{R} - \delta)^k < \infty.$$

By the Uniform Boundedness Principle we conclude that

$$C_\delta := \sup_k \|a_k\|(\tilde{R} - \delta)^k < \infty \text{ for any } \delta \in (0, \tilde{R}),$$

which implies

$$\|a_k\|^{1/k} \le \frac{C_\delta^{1/k}}{\tilde{R} - \delta} \text{ for all } \delta \in (0, \tilde{R}),\ k \in \mathbb{N},$$

and

$$\frac{1}{R} = \limsup_k \|a_k\|^{1/k} \le \frac{1}{\tilde{R} - \delta} \text{ for all } \delta \in (0, \tilde{R}).$$

Letting $\delta \to 0$, we obtain $\tilde{R} \le R$. □

**Remark:** The proof indicates the basic two ways of approaching holomorphic functions with values in a Banach space. The first is to proceed as in the complex valued case, and the second is to apply linear functionals and resort to results on complex valued functions.

The following will be an exercise:

**Proposition:** A function $f : \Omega \to X$ is analytic if and only if it is *weakly analytic*, i.e. $\varphi \circ f : \Omega \to \mathbb{C}$ is analytic for every $\varphi \in X'$.

**— — End of the Detour — —**

For the applications we need another lemma.

**1.6. Lemma:** Let $(c_n)$ be a real sequence such that $0 \le c_{n+m} \le c_n c_m$ for all $n, m \in \mathbb{N}$. Then $\sqrt[n]{c_n} \to c := \inf_k \sqrt[k]{c_k}$ as $n \to \infty$.

*Proof.* Let $\varepsilon > 0$ and choose $m$ such that $\sqrt[m]{c_m} < c + \varepsilon$. Let $b := \max\{c_1, \dots, c_m\}$ and write $n > m$ as $n = km + r$ with $k \in \mathbb{N}$ and $r \in \{1, \dots, m\}$. Then

$$c_n^{1/n} = (c_{km+r})^{1/n} \le (c_m^k c_r)^{1/n} \le (c+\varepsilon)^{km/n} b^{1/n} = (c+\varepsilon)(c+\varepsilon)^{-r/n} b^{1/n} \le c + 2\varepsilon$$

for large $n$, since $(c+\varepsilon)^{-r/n} b^{1/n} \to 1$ as $n \to \infty$ (here we assumed $b > 0$ without loss of generality). □

**1.7. Corollary:** If $A$ is closed linear operator in $X$, then its resolvent $\rho(A) \to \mathcal{L}(X)$, $\lambda \mapsto R(\lambda, A)$, is an analytic function and

$$\frac{d^k}{d\lambda^k} R(\lambda, A) = (-1)^k k! R(\lambda, A)^{k+1}, \quad k \in \mathbb{N}_0.$$

For $\lambda \in \rho(A)$ we have

$$d(\lambda, \sigma(A)) = \frac{1}{\inf_k \|R(\lambda, A)^k\|^{1/k}}. \qquad (+)$$

*Proof.* The first part is clear from 1.5(b) and analyticity of power series. Applying 1.6 to $c_k := \|R(\lambda, A)^k\|$ we see that the right hand side of (+) is the radius of convergence $R$ for the power series in 1.5(b). By the above, $\partial B(\lambda, R) \cap \sigma(A) \ne \emptyset$ (otherwise the resolvent would be holomorphic on a strictly larger ball which is impossible). □

We give an application to bounded operators.

**1.8. Proposition and Definition:** Let $T \in \mathcal{L}(X)$ and $X \ne \{0\}$. Then $\sigma(T) \ne \emptyset$, and the *spectral radius of* $T$,

$$r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

satisfies

$$r(T) = \lim_{k\to\infty} \|T^k\|^{1/k} = \inf_k \|T^k\|^{1/k}.$$

In particular, $\sigma(T)$ is a non-empty, compact subset of $B(0, \|T\|)$.

*Proof.* We begin with the observation that, for $|\lambda| > \|T\|$,

$$\lambda I - T = \lambda\Big(I - \frac{T}{\lambda}\Big),$$

which is invertible by a Neumann series

$$(\lambda I - T)^{-1} = \lambda^{-1} \sum_{k=0}^{\infty} \frac{T^k}{\lambda^k} = \sum_{k=0}^{\infty} T^k \lambda^{-(k+1)}.$$

The latter is a power series in $\lambda^{-1}$ with radius of convergence $R := 1/\inf_k \|T^k\|^{1/k}$. Hence it converges for $|\lambda| > 1/R$, and $\sigma(T) \subset \overline{B}(0, 1/R)$, i.e. $r(T) \leq 1/R$. The above on power series also shows that $r(T) = 1/R$ if $1/R > 0$.

We now show that $\sigma(T) \neq 0$ (then $\sigma(T) = \{0\}$ in case $1/R = 0$). From the series representation above we also obtain that $R(\lambda, T) \to 0$ as $|\lambda| \to \infty$. If $\sigma(T) = \emptyset$ then, for any $\varphi \in \mathcal{L}(X)'$, $\lambda \to \varphi \circ R(\lambda, T)$ is an entire function, which is bounded (since it tends to 0 as $|\lambda| \to \infty$). By Liouville's Theorem, this function is constant. We conclude $\varphi \circ R(\lambda, T) = 0$ for any $\lambda$ and any $\varphi$. But then $R(\lambda, T) = 0$ which implies $X = \{0\}$, and this case was excluded. □

End Lect.3

**1.9. Definition (Fine structure of the spectrum):** Let $A$ be a closed linear operator in $X$. We define

$\sigma_p(A) := \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not injective}\} = \{\lambda \in \mathbb{C} : N(\lambda I - A) \neq \{0\}\}$ (*point spectrum*).

Any $\lambda \in \sigma_p(A)$ is called an *eigenvalue of* $A$ and any $x \in N(\lambda I - A) \setminus \{0\}$ is called an *eigenvector* for the eigenvalue $\lambda$. We further define

$$\begin{aligned}
\sigma_r(A) &:= \{\lambda \in \mathbb{C} : R(\lambda I - A) \text{ is not dense in } X\} \ (\textit{residual spectrum}) \\
\sigma_c(A) &:= \{\lambda \in \mathbb{C} : R(\lambda I - A) \text{ is dense but not closed in } X\} \ (\textit{continuous spectrum}) \\
\sigma_{ap}(A) &:= \{\lambda \in \mathbb{C} : \text{ there exists a sequence } (x_n) \text{ in } D(A) \text{ with } \|x_n\| = 1 \text{ for all } n \\
&\qquad \text{and } (\lambda I - A)x_n \to 0\} \ (\textit{approximate point spectrum}).
\end{aligned}$$

A sequence $(x_n)$ as in the definition of the approximate point spectrum is sometimes called an *approximate eigenvector*.

**Remark:** $\sigma_p(A) \subset \sigma_{ap}(A)$ (take $x_n = x$ eigenvector).

**Examples:** 1) If $\dim X < \infty$ and $A \in \mathcal{L}(X)$ then $\sigma(A) = \sigma_p(A) = \sigma_r(A)$ ($A$ is a finite-dimensional matrix for which injectivity is equivalent to surjectivity).

2) Let $X = C[0, 1]$, $A = \frac{d}{dx}$ with $D(A) = C^1[0, 1]$. We saw above that $\sigma(A) = \sigma_p(A) = \mathbb{C}$. On the other hand, for any $\lambda \in \mathbb{C}$ the ordinary differential equation $\lambda f - f' = g$ has a (non-unique) solution $f \in C^1[0, 1]$ for each right hand side $g \in C[0, 1]$. Hence $R(\lambda I - A) = X$ for each $\lambda \in \mathbb{C}$ and $\sigma_r(A) = \sigma_c(A) = \emptyset$.

**1.10. Proposition:** Let $A$ be a closed operator in $X$. Then:

(a) $\sigma_{ap}(A) = \sigma_p(A) \cup \{\lambda \in \mathbb{C} : \lambda I - A \text{ is injective and } (\lambda I - A)^{-1} \text{ is not bounded}\}$.

(b) $\sigma(A) = \sigma_{ap}(A) \cup \sigma_r(A)$.

(c) $\partial\sigma(A) \subseteq \sigma_{ap}(A)$.

*Proof.* (a) Let $\lambda \in \mathbb{C}$ and let $\lambda - A$ be injective, $Y := R(\lambda - A)$. Then:

$$\begin{aligned} & (\lambda - A)^{-1} : Y \to X \text{ is not bounded} \\ \Leftrightarrow\ & \text{there exists } (y_n) \text{ in } Y \text{ such that } \|y_n\| = 1 \text{ and } \|(\lambda - A)^{-1} y_n\| \to \infty \\ \Leftrightarrow\ & \text{there exists } (z_n) \text{ in } Y \text{ such that } z_n \to 0 \text{ and } \|(\lambda - A)^{-1} z_n\| = 1 \\ \Leftrightarrow\ & \text{there exists } (x_n) \text{ in } D(A) \text{ s.t. } \|x_n\| = 1 \text{ and } (\lambda - A)^{-1} x_n \to 0 \end{aligned}$$

To see this put $z_n = y_n / \|(\lambda - A)^{-1} y_n\|$, $y_n = z_n / \|z_n\|$, and $x_n = (\lambda - A)^{-1} z_n$, $z_n = (\lambda - A) x_n$, respectively. Thus we have shown

$$\sigma_{ap}(A) \setminus \sigma_p(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ is injective and } (\lambda I - A)^{-1} \text{ is not bounded}\}.$$

(b) "$\supseteq$" is clear. For the proof of "$\subseteq$" let $\lambda \in \sigma(A) \setminus \sigma_{ap}(A)$. By (a), $\lambda - A$ is injective and $(\lambda - A)^{-1}$ is bounded. In particular, $R(\lambda - A)$ is closed in $X$. Hence $\overline{R(\lambda - A)} = R(\lambda - A) \neq X$ (otherwise $\lambda \notin \sigma(A)$!). We have shown $\lambda \in \sigma_r(A)$.

(c) Let $\lambda \in \partial\sigma(A)$. We find a sequence $(\lambda_n)$ in $\rho(A)$ with $\lambda_n \to \lambda$. By 1.5(b) we have $\|R(\lambda_n, A)\| \to \infty$. By the Uniform Boundedness Principle we find $y \in X$ such that $\alpha_n := \|R(\lambda_n, A) y\| \to \infty$ (observe that we have $\alpha_n > 0$ for any $n$). Let $x_n := R(\lambda_n, A) y / \alpha_n$, $n \in \mathbb{N}$. Then $(x_n)$ is a sequence in $D(A)$ with $\|x_n\| = 1$ for all $n$. Moreover,

$$(\lambda - A) x_n = (\lambda - \lambda_n) x_n + (\lambda_n - A) x_n = (\lambda - \lambda_n) x_n + y/\alpha_n \to 0 \quad (n \to \infty).$$

□

**Example:** Let $X = l^2 = \{(x_n)_{n\in\mathbb{N}} : \|(x_n)\| l^2 := \Big(\sum_n |x_n|^2\Big)^{1/2} < \infty\}$ and let $L \in \mathcal{L}(l^2)$ be the *left shift* given by $L(x_n) := (x_{n+1})$, i.e. $L(x_n) = (x_2, x_3, x_4, \ldots)$. Clearly, we have $\|L\| = 1$, $\|L^k\| = 1$ for every $k \in \mathbb{N}$, and hence $r(L) = 1$. For $\lambda \in \mathbb{C}$ with $|\lambda| \le 1$ we have

$$(\lambda - L)(x_n) = 0 \Leftrightarrow \forall n : \lambda x_n - x_{n+1} = 0 \Leftrightarrow \forall n : x_{n+1} = \lambda^n x_1.$$

For $|\lambda| < 1$ we have $(1, \lambda, \lambda^2, \ldots) \in l^2$, hence $\lambda \in \sigma_p(L)$. For $|\lambda| = 1$ we have $(1, \lambda, \lambda^2, \ldots) \notin l^2$, and $\lambda \notin \sigma_p(L)$.

We conclude $\sigma_p(L) = \{|\lambda| < 1\}$, $\sigma(L) = \sigma_{ap}(L) = \{|\lambda| \le 1\}$, and $\sigma_{ap}(L) \setminus \sigma_p(L) = \{|\lambda| = 1\} = \partial\sigma(L)$.

**1.11. Dual operators:** Recall that, for $T \in \mathcal{L}(X, Y)$, its *dual* (or *adjoint*)[1] *operator* $T' \in \mathcal{L}(X', Y')$ is given by

$$T'\phi := \phi \circ T, \quad \phi \in Y'.$$

[1] In this course we will reserve the term "adjoint" for the *Hilbert space adjoint* and thus speak of "dual operators" here.

Using *duality brackets*

$$\langle y, \phi\rangle_{Y\times Y'} := \phi(y), \quad \text{for all } y \in Y,\, \phi \in Y',$$

this can be written as

$$\langle x, T'\phi\rangle_{X\times X'} = \langle Tx, \phi\rangle_{Y\times Y'} \quad \text{for all } x \in X,\, \phi \in Y',$$

**Example:** Let $X = Y = l^2$ and $L$ be the left shift. We calculate $L'$. For $X = l^2$ we have $X' = l^2$ where the duality bracket is given by

$$\langle (x_n), (y_n)\rangle_{l^2\times l^2} = \sum_n x_n y_n.$$

For $(x_n), (y_n) \in l^2$ we then have

$$\langle L(x_n), (y_n)\rangle = \sum_{n=1}^{\infty} x_{n+1}y_n = \sum_{n=2}^{\infty} x_n y_{n-1} = \langle (x_n), (y_{n-1})\rangle$$

if we let $y_0 := 0$. Hence $L' = R$ where $R$ is the *right shift* given by

$$R(y_n) := (0, y_1, y_2, y_3, \ldots).$$

Clearly, we have $\|R(y_n)\|_{l^2} = \|(y_n)\|_{l^2}$, $\|R\| = 1$, $r(T) = 1$, and $\sigma(R) \subseteq \{|\lambda| \le 1\}$. For $|\lambda| \le 1$ we have

$$(\lambda - R)(y_n) = 0 \Leftrightarrow \lambda y_1 = 0, \forall n \ge 2 : \lambda y_n = y_{n-1} \Leftrightarrow (y_n) = 0.$$

Hence $\sigma_p(R) = \emptyset$.

**Rules for dual operators:** The following rules are easily verified:

$$(I_X)' = I_{X'}, \quad (S+T)' = S' + T', \quad (\alpha T)' = \alpha T', \quad (ST)' = T'S'$$

where $T \in \mathcal{L}(X,Y)$, $\alpha \in \mathbb{C}$, $S \in \mathcal{L}(X,Y)$ for the sum and $S \in \mathcal{L}(Y,Z)$ for the product, respectively.

**Remark:** For $T \in \mathcal{L}(X,Y)$ we have

$$T : X \to Y \text{ is an isomorphism} \Longleftrightarrow T' : Y' \to X' \text{ is an isomorphism}.$$

Moreover, we have $(T')^{-1} = (T^{-1})'$ in this case.

*Proof.* "$\Rightarrow$": We have

$$T'(T^{-1})' = (T^{-1}T)' = (I_X)' = I_{X'}, \quad (T^{-1})'T' = (TT^{-1})' = (I_Y)' = I_{Y'},$$

which proves invertibility of $T'$ and $(T')^{-1} = (T^{-1})'$.

"$\Leftarrow$": Suppose that $T'$ has an inverse $S : X' \to Y'$. Then $T'' := (T')'$ is an isomorphism $X'' \to Y''$ by the first part of the proof. Since $X$ is a closed subspace of $X''$, $T''(X)$ is a closed subspace of $Y''$. But $T''(X) = T(X) \subseteq Y$, so $T(X)$ is a closed subspace of $Y$.

If $\phi \in X'$ satisfies $\phi|_{T(X)} = 0$ then

$$\langle x, T'\phi\rangle = \langle Tx, \phi\rangle = 0, \quad x \in X,$$

i.e. $T'\phi = 0$. Since $T'$ is injective we conclude $\phi = 0$. By Hahn-Banach we thus have shown that $T(X)$ is dense in $Y$. Since $T(X)$ is also closed in $Y$ we obtain $T(X) = Y$, i.e. $T$ is surjective.

Since $T''$ is injective, $T = T''|_X$ is injective. We have shown that $T : X \to Y$ is bijective. By the Open Mapping Theorem, $T : X \to Y$ is an isomorphism. □

End Lect.4

**Recall for the previous proof:** The *bidual* $X'' = (X')' = \mathcal{L}(X', \mathbb{C})$ of $X$ is a Banach space, and the map

$$J : X \to X'',\ x \mapsto \delta_x \quad \text{where } \delta_x : X' \to \mathbb{C}, \phi \mapsto \delta_x(\phi) := \phi(x)$$

is (by Hahn-Banach) an isometric injection. Written with duality brackets we have

$$\langle \phi, \delta_x\rangle_{X' \times X''} = \langle x, \phi\rangle_{X \times X'}, \quad x \in X, \phi \in X'.$$

It is common to identify $X$ with the closed subspace $J(X)$ of $X''$. The space $X$ is called *reflexive* if $J(X) = X''$.

If $T \in \mathcal{L}(X, Y)$ then $T'' := (T')' \in \mathcal{L}(X'', Y'')$ and $T = T''|_X$, since for $x \in X$, $\phi \in Y'$ we have

$$\langle \phi, T''\delta_x\rangle = \langle T'\phi, \delta_x\rangle = \langle x, T'\phi\rangle = \langle Tx, \phi\rangle = \langle \phi, \delta_{Tx}\rangle,$$

which means $T''\delta_x = \delta_{Tx}$ for all $x \in X$, i.e. $T''x = Tx$, $x \in X$, if we identify $X$ and $J(X)$.

**1.12. Definition:** Let $A$ be a linear operator from $X$ to $Y$ with $D(A)$ dense in $X$. We define the linear operator $A' : Y' \supset D(A') \to X'$ by letting for $\phi \in Y'$, $\psi \in X'$:

$$\phi \in D(A'), A'\phi = \psi :\Longleftrightarrow \forall x \in D(A) : \langle Ax, \phi\rangle_{Y \times Y'} = \langle x, \psi\rangle_{X \times X'}$$

Then $D(A')$ is the set of all $\phi \in Y'$ such that the map

$$D(A) \to \mathbb{C},\ x \mapsto \langle Ax, \phi\rangle,$$

has a continuous extension $\psi \in X'$. By denseness of $D(A)$ in $X$, this extension is uniquely determined (if it exists). Then $A'\phi$ is this unique extension $\psi$.

**Remark:** $A'$ is a **always** closed: Let $(\phi_n)$ be a sequence in $D(A')$ such that $\phi_n \to \phi$ in $Y'$ and $A'\phi_n \to \psi$ in $X'$. Then we have, for any $x \in D(A)$:

$$\langle Ax, y\rangle = \lim_n \langle Ax, \phi_n\rangle = \lim_n \langle x, A'\phi_n\rangle = \langle x, \psi\rangle,$$

which means $\phi \in D(A')$ and $A'\phi = \psi$.

**Example:** $X = L^2(0,1)$, $A = \frac{d}{dx}$, $D(A) = C_0^1[0,1] := \{\varphi \in C^1[0,1] : \varphi(0) = \varphi(1) = 0\}$. Then

$$D(A') = \{f \in L^2(0,1) : \exists g \in L^2(0,1) \forall \varphi \in C_0^1[0,1] : \int_0^1 f\varphi'\,dx = \int_0^1 g\varphi\,dx\}.$$

From the integration-by-parts formula we see that $C^1[0,1] \subset D(A')$ and $A'f = -f$ for $f \in C^1[0,1]$: If $\varphi \in C_0^1[0,1]$ then

$$\int_0^1 f\varphi'\,dx = f\varphi\Big|_0^1 - \int_0^1 f'\varphi\,dx = -\int_0^1 f'\varphi\,dx.$$

**Rules:** $(\lambda - A)' = \lambda - A'$ for $\lambda \in \mathbb{C}$ and $X = Y$ (more general one has $(A+T)' = A' + T'$ if $T \in \mathfrak{L}(X,Y)$). If $B$ is an extension of $A$ then $A'$ is an extension of $B'$: $A \subseteq B \Rightarrow B' \subseteq A'$.

**1.13. Proposition:** Let $A$ be a closed linear operator in $X$ which is densely defined. Then $\sigma(A') = \sigma(A)$ and $\sigma_p(A') = \sigma_r(A)$.

*Proof.* We start with the second assertion. For $\lambda \in \mathbb{C}$ we have

$$\begin{aligned} &\lambda \in \sigma_r(A) \;\Leftrightarrow\; \overline{(\lambda - A)X} \neq y \;\Leftrightarrow\; \exists \phi \in Y' \setminus \{0\} \forall x \in D(A) : \langle(\lambda - A)x, \phi\rangle = 0 \\ \Leftrightarrow\;& \exists \phi \in D(A') \setminus \{0\} : (\lambda - A')\phi = 0 \;\Leftrightarrow\; \lambda \in \sigma_p(A'). \end{aligned}$$

$\sigma(A') \subseteq \sigma(A)$: Let $\lambda \in \rho(A)$. By what we have just proved, $\lambda \notin \sigma_r(A) = \sigma_p(A')$ and $\lambda - A'$ is injective. Now let $\psi \in X'$ and set $\phi := R(\lambda, A)'\psi$. For $x \in D(A)$ we then have

$$\langle(\lambda - A)x, \phi\rangle = \langle R(\lambda, A)(\lambda - A)x, \psi\rangle = \langle x, \psi\rangle.$$

This means $\phi \in D((\lambda - A)') = D(A')$ and $(\lambda - A')\phi = \psi$. Hence $\lambda - A'$ is also surjective, and $\lambda \in \rho(A)$.

$\sigma(A) \subseteq \sigma(A')$: Let $\lambda \in \rho(A')$. Then $\lambda \notin \sigma_p(A') = \sigma_r(A)$, and thus $Y := (\lambda - A)(D(A))$ is dense in $X$.

We let $S := R(\lambda, A') \in \mathfrak{L}(X')$. Then $S' \in \mathfrak{L}(X'')$ and we define $R := S'|_X$. We claim that $R \in \mathfrak{L}(X)$ is the inverse of $\lambda - A$. For $x \in D(A)$ and $\phi \in X'$ we have

$$\langle \phi, S'\delta_{(\lambda - A)x}\rangle = \langle S\phi, \delta_{(\lambda - A)x}\rangle = \langle(\lambda - A)x, S\phi\rangle.$$

Since $S\phi \in D(A')$ we can continue

$$\langle (\lambda - A)x, S\phi \rangle = \langle x, (\lambda - A')S\phi \rangle = \langle x, \phi \rangle = \langle \phi, \delta_x \rangle.$$

We thus have shown $S'\delta_{(\lambda-A)x} = \delta_x$ for all $x \in D(A)$. In particular $R$ maps the dense range of $\lambda - A$ into $X$, hence $R \in \mathcal{L}(X)$. Moreover, $R(\lambda - A)x = x$ for all $x \in D(A)$. Hence $\lambda - A$ is injective and $Ry = (\lambda - A)^{-1}y$ for all $y \in Y$. Thus we see that $(\lambda - A)^{-1}$ is a bounded operator. Since it is also closed, its domain $Y$ has to be a closed subspace of $X$. By denseness of $Y$ we thus obtain $Y = X$, $\lambda \in \rho(A)$ and $R = R(\lambda, A)$. □

**Example:** Let $X = l^2$, $L$ be the left shift, and $R$ be the right shift. Since $L' = R$, we obtain $\sigma_r(L) = \sigma_p(R) = \emptyset$.

The spectrum of an operator is related to the spectrum of its resolvents. This is called *spectral mapping*.

**1.14. Proposition:** Let $A$ be a closed linear operator in $X$ and $\lambda_0 \in \rho(A)$. Then

$$\sigma(R(\lambda_0, A)) \setminus \{0\} = \left\{ \frac{1}{\lambda_0 - z} : z \in \sigma(A) \right\}$$

and $0 \in \sigma(R(\lambda_0, A))$ if and only if $A \notin \mathcal{L}(X)$.

*Proof.* We have

$$0 \notin \sigma(R(\lambda_0, A)) \;\Leftrightarrow\; R(\lambda_0, A) : X \to X \text{ surjective} \;\Leftrightarrow\; D(A) = X \;\Leftrightarrow\; A \in \mathcal{L}(X),$$

where we used the Closed Graph Theorem for the last equivalence. Now let $z \neq \lambda_0$. Then

$$z - A = z - \lambda_0 + \lambda_0 - A = ((z - \lambda_0)R(\lambda_0, A) + I)(\lambda_0 - A) = \Big(R(\lambda_0, A) - \frac{1}{\lambda_0 - z}\Big)(z - \lambda_0)(\lambda_0 - A),$$

where $(z - \lambda_0)(\lambda_0 - A) : D(A) \to X$ is bijective. Hence $z - A : D(A) \to X$ is bijective if and only if $R(\lambda_0, A) - (\lambda_0 - z)^{-1} : X \to X$ is bijective. We obtain that $z \in \sigma(A)$ is equivalent to $(\lambda_0 - z)^{-1} \in \sigma(R(\lambda_0, A))$. □

End Lect.5

**Remark:** From the proof we see that, for $\lambda_0, z \in \rho(A)$ with $z \neq \lambda_0$:

$$\Big(\frac{1}{\lambda_0 - z} - R(\lambda_0, A)\Big)^{-1} = \frac{1}{\lambda_0 - z}(\lambda_0 - A)R(z, A).$$

What can be done if a linear operator $A$ from $X$ to $Y$ is not closed? Of course, one can consider the closure $\overline{\text{gr}(A)}$ in $X \times Y$. But this need not be the graph of an operator.

**1.15. Definition:** A linear operator $A$ from $X$ to $Y$ is called *closable* if it has a closed linear extension $B : X \supset D(B) \to Y$. In this case, $A$ has a smallest closed linear extension $\overline{A}$ which is called the *closure* of $A$.

By Exercise 7 the following properties are equivalent:

- $A$ is closable,
- $\overline{\mathrm{gr}(A)}$ is the graph of a function,
- if $(x_n)$ is a sequence in $D(A)$ such that $x_n \to 0$ in $X$ and $Ax_n \to y$ in $Y$ then $y = 0$.

In this case, one has $\mathrm{gr}(\overline{A}) = \overline{\mathrm{gr}(A)}$.

**1.16. Proposition:** Let $A$ be a densely defined linear operator in $X$. If $D(A')$ is dense in $X'$ then $A$ is closable. The converse holds if $X$ is reflexive.

*Proof.* If $D(A')$ is dense in $X'$ then $A'' := (A')'$ is a closed linear operator in $X''$. We show that $A''$ is an extension of $A$: For $x \in D(A)$, $\phi \in D(A')$ we have

$$\langle Ax, \phi \rangle = \langle x, A'\phi \rangle = \langle A'\phi, \delta_x \rangle,$$

which shows that $\delta_x \in D(A'')$ and $A''\delta_x = \delta_{Ax}$ for all $x \in D(A)$.

Now let $X$ be reflexive and assume that $D(A')$ is not dense in $X'$. By Hahn-Banach we find $y \in X'' = X$, $y \neq 0$, such that

$$\langle y, \phi \rangle = 0 = \langle 0, A'\phi \rangle \quad \text{for all } \phi \in D(A').$$

We claim that $(0, y) \in \overline{\mathrm{gr}(A)}$. If $(\psi, \phi) \in X' \times X'$ such that $(\psi, \phi)|_{\mathrm{gr}(A)} = 0$. Then $\phi \in D(A')$ and $A'\phi = -\psi$. From the property of $y$ above we get $(\psi, \phi)(0, y) = 0$. By Hahn-Banach we conclude that $(0, y) \in \overline{\mathrm{gr}(A)}$ as claimed. By $y \neq 0$, $A$ is not closable. □

# 2 Fredholm operators and the spectral theory of compact operators

Again, $X$ and $Y$ are Banach spaces if nothing else is said.

Recall from Linear Algebra: If $\dim X = \dim \mathrm{Y} < \infty$ and $T \in \mathcal{L}(X,Y)$ then

$$\dim N(T) = \underbrace{\operatorname{codim} R(T)}_{=:\dim Y/R(T)} < \infty.$$

**2.1. Definition:** An operator $T \in \mathcal{L}(X,Y)$ is called a *Fredholm operator* if

$$\alpha(T) := \dim N(T) < \infty \text{ and } R(T) \text{ is closed with } \beta(T) := \operatorname{codim} R(T) < \infty.$$

In this case, the *index of* $T$ is defined by

$$\operatorname{ind} T := \alpha(T) - \beta(T).$$

The set of all Fredholm operators from $X$ to $Y$ is denoted by $\Phi(X,Y)$.

**Remark:** If $\dim X = \dim Y < \infty$, any $T \in \mathcal{L}(X,Y)$ is a Fredholm operator of index 0. If $X$ and $Y$ are arbitrary and $T \in \mathcal{L}(X,Y)$ is an isomorphism from $X$ to $Y$ then $\alpha(T) = \beta(T) = 0$ and $T$ is a Fredholm operator of index 0.

The following proposition shows that closedness of $R(T)$ can be omitted from the definition.

**2.2. Proposition:** Let $T \in \mathcal{L}(X,Y)$. If $\operatorname{codim} R(T) < \infty$ then $R(T)$ is closed.

*Proof.* We find a basis $[y_1], \ldots, [y_n]$ of $Y/R(T)$ and let $H := \operatorname{span}\{y_1, \ldots, y_n\}$. Then $Y = R(T) + H$ and $R(T) \cap H = \{0\}$. Since $H$ is complete, $X \times H$ is a Banach space (e.g. for the sum-norm) and the map

$$\widetilde{T} : X \times H \to Y,\ (x,y) \mapsto Tx + y$$

is continuous and surjective. By the Open Mapping Theorem we have

$$\gamma := \inf \left\{ \frac{\|\widetilde{T}(x,y)\|_Y}{d((x,y), N(T))} : (x,y) \notin N(\widetilde{T}) \right\} > 0.$$

Since

$$N(\widetilde{T}) = \{(x,y) \in X \times H : Tx = -y\} = \{(x,0) : Tx = 0\} = N(T) \times \{0\}$$

we have for $x \notin N(T)$:

$$\|Tx\|_Y = \|\widetilde{T}(x,0)\|_Y \geq \gamma d((x,0), N(\widetilde{T})) = d(x, N(T)).$$

From this we obtain closedness of $R(T)$. □

**Remark:** We recall that, if $Y$ is a Banach space and $Z$ is a closed linear subspace, then the *quotient space* $Y/Z$ is a Banach space defined by

$$Y/Z := \{y + Z : y \in Y\}, \|y + Z\|_{Y/Z} := \inf\{\|y - z\|_Y : z \in Z\} = d(y, Z).$$

If $Z$ is clear from the context, it is common to write $[y]$ instead of $y + Z$ and $q$ for the *quotient map* $Y \to Y/Z$, $y \mapsto [y]$. We also recall that an operator $T \in \mathcal{L}(X, Y)$ has a unique factorization $T = S \circ q$ where $S \in \mathcal{L}(X/N(T), Y)$ is injective and $q : X \to X/N(T)$ is the quotient map. $S$ is given by $S[x] = Tx$.

**Remark:** (a) Suppose that $W$ and $Z$ are subspaces of a Banach space $X$ such that $W \cap Z = \{0\}$ and $X = W + Z$. We write $X = W \oplus Z$ if both $W$ and $Z$ are closed subspaces of $X$. Notice that $X = W \oplus Z$ if and only if $W \times Z \to X$, $(w, z) \mapsto w + z$, is an isomorphsm if and only if there exists a projection $P \in \mathcal{L}(X)$ such that $P(X) = Z$ and $N(P) = W$.

A closed subspace $Z$ is called *complemented* in $X$ if there is a (closed) subspace $W$ of $X$ such that $X = W \oplus Z$.

(b) Any finite-dimensional subspace $Z$ of $X$ is complemented in $X$: Choose a basis $z_1, \ldots, z_n$ of $Z$ and linear functionals $\psi_1, \ldots, \psi_n \in Z'$ such that $\psi_j(x_k) = \delta_{jk}$. Extend the $\psi_k$ by Hahn-Banach and obtain $\phi_1, \ldots, \phi_n \in X'$. Then $W := \bigcap_{k=1}^n N(\phi_k)$ is a complement of $Z$.

**2.3. Theorem:** $\Phi(X, Y)$ is open in $\mathcal{L}(X, Y)$ and $\mathrm{ind} : \Phi(X, Y) \to \mathbb{Z}$ is continuous.

End Lect.6

*Proof.* Let $T \in \mathcal{L}(X, Y)$. We shall show the existence of $\varepsilon > 0$ such that, for $S \in \mathcal{L}(X, Y)$, we have

$$\|S - T\| < \varepsilon \implies S \in \Phi(X, Y) \text{ and } \mathrm{ind}\, S = \mathrm{ind}\, T.$$

First we find closed subspaces $G$ of $X$ and $H$ of $Y$ such that

$$X = N(T) \oplus G,\ Y = R(T) \oplus H,\ \dim H < \infty.$$

Notice that $R(T) = T(X) = T(G)$. For any $S \in \mathcal{L}(X, Y)$ we define

$$\widehat{S} : G \times H \to Y, \quad (g, h) \mapsto Tg + h.$$

Then we have, for $T_1, T_2 \in \mathcal{L}(X, Y)$:

$$\|\widehat{S}_1 - \widehat{S}_2\| = \sup_{\|(g,h)\|=1} \|S_1 g + h - (S_2 g + h)\| \le \|S_1 - S_2\|.$$

We claim that $\widehat{T} : G \times H \to Y$ is an isomorphism. Indeed, $\widehat{T}$ is surjective by $R(T) = T(G)$ and the choice of $H$, and $\widehat{T}$ is injective, since $\widehat{T}(g, h) = 0$ implies $Tg = -h \in R(T) \cap H = \{0\}$, and $T|_G$ is injective. Hence we find $\varepsilon > 0$ such that, for $\tilde{S} \in \mathcal{L}(G \times H, Y)$, $\|\tilde{S} - \widehat{T}\| < \varepsilon$ implies that $\tilde{S} : G \times H \to Y$ is an isomorphism.

Now let $S \in \mathcal{L}(X, Y)$ with $\|S - T\| < \varepsilon$. Then $\|\widehat{S} - \widehat{T}\| < \varepsilon$, and $\widehat{S} : G \times H \to Y$ is an isomorphism. We now show

(i) $\alpha(S) \leq \alpha(T) < \infty$,

(ii) $\beta(S) \leq \beta(T) < \infty$,

(iii) $\mathrm{ind}\,(S) = \mathrm{ind}\,(T)$.

**(i):** If $g \in N(S) \cap G$ then $\widehat{S}(g, 0) = 0$, and $g = 0$. Hence $N(S) \cap G = \{0\}$ and

$$\alpha(S) = \dim N(S) \leq \dim N(T) = \alpha(T) < \infty.$$

**(ii):** $G \times \{0\}$ is closed in $G \times H$, so $S(G) = \widehat{S}(G \times \{0\})$ is closed in $Y$. Since $\widehat{S}$ is surjective we have $Y = S(G) + H$. If $y = S(g) \in S(G) \cap H$ then $\widehat{S}(g, -y) = S(g) - S(g) = 0$ and $g = 0$, $y = 0$, since $\widehat{S}$ is injective. Hence $S(G) \oplus H = Y = T(G) \oplus H$ and

$$\beta(S) = \mathrm{codim}\, R(S) \leq \mathrm{codim}\, S(G) = \mathrm{codim}\, T(G) = \mathrm{codim}\, R(T) = \beta(T) < \infty.$$

**(iii):** By what we have already shown we know that $N(S) \oplus G$ is a closed subspace of $X$ of finite codimension. We thus find a subspace $W$ of $X$ with $\dim W < \infty$ such that $X = W \oplus N(S) \oplus G$. We then have

$$\alpha(T) = \dim N(T) = \dim N(S) \oplus \dim W = \alpha(S) + \dim W.$$

On the other hand, we have

$$S(G) \subseteq S(X) = S(G \oplus W) = S(G) \oplus S(W)$$

since both subspaces on the right hand side are closed and their intersection is trivial (by injectivity of $S|_{G \oplus W}$). Hence

$$\beta(T) = \mathrm{codim}\, T(X) \overset{\text{(ii)}}{=} \mathrm{codim}\, S(G) = \mathrm{codim}\, S(X) + \dim S(W) = \beta(S) + \dim W,$$

and we conclude that

$$\mathrm{ind}\, T = \alpha(T) - \beta(T) = \alpha(S) + \dim W - (\beta(S) + \dim W) = \mathrm{ind}\, S,$$

which ends the proof. □

**2.4. Compact Operators:** Recall (cf. FA) that a linear operator $T : X \to Y$ is called *compact* if $\overline{T(B_X)}$ is compact in $Y$ where $B_X := \{x \in X : \|x\| \leq 1\}$ denotes the closed unit ball in $X$. The set of all compact linear operators from $X$ to $Y$ is denoted by $\mathcal{K}(X, Y)$. Properties:

- Any compact linear operator $T : X \to Y$ is bounded, i.e. $\mathcal{K}(X, Y) \subset \mathcal{L}(X, Y)$.
- $\dim T(X) < \infty \Longrightarrow T$ is compact.

- $I_X$ is compact $\iff \dim X < \infty$.
- $\mathcal{K}(X,Y)$ is a closed linear subspace of $\mathcal{L}(X,Y)$.
- If $W$, $Z$ are Banach spaces and $S \in \mathcal{L}(W,X)$, $R \in \mathcal{L}(X,Y)$ and $T \in \mathcal{K}(X,Y)$ then $RT \in \mathcal{K}(X,Z)$ and $TS \in \mathcal{K}(W,Y)$ (*ideal property*).

**2.5. Proposition:** Let $K \in \mathcal{K}(X)$. Then $I - K \in \Phi(X)$ and $\operatorname{ind}(I-K) = 0$.

*Proof.* On $N := N(I-K)$ we have $I_N = K|_N \in \mathcal{K}(N)$ and thus $\dim N < \infty$. We let $T := I - K$ and show that $R(T)$ is closed. To this end we factorize $T = S \circ q$ where $q : X \to X/N$ is the quotient map and $S : X/N \to X$. Then $R(T) = R(S)$ and $S$ is injective. We claim that there exists $\eta > 0$ such that $\|Sw\| \geq \eta \|w\|$ for all $w \in X/N$. From this, closedness of $R(S) = R(T)$ follows.

Assume that such an $\eta$ does not exist. Then we find a sequence $(w_n)$ such that $\|w_n\| = 1$ and $Sw_n \to 0$, and a sequence $(x_n)$ in $X$ such that $q(x_n) = w_n$, $\|x_n\| \leq 2$. Then $Tx_n = Sw_n \to 0$ and $d(x_n, N) = \|q(x_n)\| = 1$.

Since $K$ is compact we have $Kx_{k(n)} \to y \in X$ for a subsequence, which implies $x_n = Tx_n + Kx_n \to y$ and $Sw_n = Tx_n \to Ty$. So $Ty = 0$ and $y \in N$. However, $d(y, N) = 1$, a contradiction.

By Schauder's Theorem (cf. FA) $K' \in \mathcal{K}(X')$, hence $\dim N((I-K)') < \infty$. Since $R(I-K)$ is closed we obtain

$$\begin{aligned}\operatorname{codim} R(I-K) &= \dim X/R(I-K) = \dim \left(X/R(I-K)\right)' \\ &= \dim \{\phi \in X' : \phi|_{R(I-K)} = 0\} = \dim \left(N(I-K)'\right) < \infty.\end{aligned}$$

It remains to show $\operatorname{ind}(I-K) = 0$. We consider $\gamma : [0,1] \to \mathbb{Z}$, $t \mapsto \operatorname{ind}(I - tK) = 0$ (observe that $tK \in \mathcal{K}(X)$ and hence $I - tK \in \Phi(X)$ for any $t \in \mathbb{R}$). By 2.3, $\gamma$ is continuous, hence constant and $\operatorname{ind}(I-K) = \gamma(1) = \gamma(0) = \operatorname{ind} I = 0$. □

End Lect.7

**2.6. Corollary:** Let $K \in \mathcal{K}(X)$ and $\dim X = \infty$. Then $0 \in \sigma(K)$ and for $\lambda \in \sigma(K) \setminus \{0\}$ one has that $\lambda - K$ is a Fredholm operator of index 0. For $\lambda \in \mathbb{C} \setminus \{0\}$ one has in particular

$$\lambda - K \text{ is injective} \quad \iff \quad \lambda - K \text{ is surjective.}$$

and, more precisely, the following **Fredholm alternative**:

Either $(\lambda - K)x = 0$ has just the trivial solution; in this case $(\lambda - K)x = y$ has a unique solution for any $y \in X$,

or $(\lambda - K)x = 0$ has exactly $n \in \mathbb{N}$ linearly independent solutions and also the dual equation $(\lambda - K')\phi = 0$ has exactly $n$ linearly independent solutions; in this case $(\lambda - K)x = y$ has solutions if and only if $\phi(y) = 0$ for any $\phi \in N(\lambda - K')$.

**2.7. Corollary:** Let $T \in \mathcal{L}(X,Y)$ be an isomorphism from $X$ to $Y$ and $K \in \mathcal{K}(X,Y)$. Then $T-K \in \Phi(X,Y)$ and $\operatorname{ind}(T-K) = 0$.

*Proof.* We have $T^{-1}K \in \mathcal{K}(X)$, hence by Proposition 2.5, $T^{-1}(T-K) = I_X - T^{-1}K \in \Phi(X)$ and $\operatorname{ind} T^{-1}(T-K) = 0$. Since $T^{-1}$ is an isomorphism from $Y$ to $X$ we obtain the assertion. $\square$

The assertion in 2.7 is the keystone for the following powerful perturbation property of Fredholm operators.

**2.8. Theorem:** Let $T \in \Phi(X,Y)$ and $K \in \mathcal{K}(X,Y)$. Then $T+K \in \Phi(X,Y)$ and $\operatorname{ind}(T+K) = \operatorname{ind} T$.

*Proof.* We take up the construction and notation of the proof of 2.3 and write

$$X = N(T) \oplus G,\ Y = R(T) \oplus H,\ \dim H < \infty.$$

where $G$ is a closed subspace of $X$. We let $S := T - K$ and consider the operators $\widehat{T}, \widehat{S} : G \times H \to Y$. Again, $\widehat{T}$ is an isomorphism, and we have $\widehat{S} = \widehat{T} - \widetilde{K}$ where $\widetilde{K}(g,h) := Kg$. The operator $\widetilde{K}$ is compact, thus by 2.7 we have $\widehat{S} \in \Phi(G \times H, Y)$ and $\operatorname{ind} \widehat{S} = 0$, i.e. $\alpha(\widehat{S}) = \beta(\widehat{S})$. Clearly,

$$(N(S) \cap G) \times \{0\} = \{(g,0) \in G \times H : Sg = 0\} \subseteq N(\widehat{S}),$$

which implies

$$\alpha(S) = \dim N(S) \le \operatorname{codim} G + \alpha(\widehat{S}) = \alpha(T) + \alpha(\widehat{S}) < \infty.$$

Moreover, we have $\widehat{S}(G \times H) = S(G) + H$, and $S(G)$ has finite codimension and is thus closed by 2.2. We obtain

$$\beta(S) = \operatorname{codim} S(X) \le \operatorname{codim} S(G) \le \operatorname{codim} \widehat{S} + \dim H = \beta(\widehat{S}) + \beta(T) < \infty,$$

and $S(X)$ is closed by 2.2.

For the calculation of the index we have to look closer. We let $G_0 := N(S) \cap G$, find a complement $G_1$ of $G_0$ in $G$, a complement $W_0$ of $G_0$ in $N(S)$, and a complement $W_1$ of $N(S) + G = W_0 \oplus G_0 \oplus G_1$ in $X$. Then we have

$$\begin{aligned} \alpha(T) &= \dim W_1 + \dim W_0, \quad \beta(T) = \dim H, \quad \alpha(S) = \dim W_0 + \dim G_0, \\ \beta(S) &= \operatorname{codim} S(X) = \operatorname{codim} S(G_1) - \dim W_1, \end{aligned}$$

and

$$\begin{aligned} \beta(\widehat{S}) &= \operatorname{codim}(S(G) + H) = \operatorname{codim} S(G) - \dim H + \dim(S(G) \cap H) \\ &= \operatorname{codim} S(G) - \dim H + \dim(S(G_1) \cap H). \end{aligned}$$

Now

$$\dim(S(G_1)\cap H) = \dim\{g\in G_1 : Sg\in H\} = \dim \underbrace{\{(g,h)\in G_1\times H : Sg=-h\}}_{=:V},$$

and $V$ is a complement of $G_0\times\{0\}$ in $N(\widehat{S})$. We thus obtain

$$\alpha(\widehat{S}) = \dim N(\widehat{S}) = \dim G_0 + \dim V = \dim G_0 + \dim(S(G_1)\cap H).$$

Putting everything together, we have

$$\begin{aligned}
\operatorname{ind} S &= \alpha(S)-\beta(S)\\
&= \dim W_0 + \dim G_0 - \operatorname{codim} S(G_1) + \dim W_1\\
&= \alpha(T) - \operatorname{codim} S(G_1) + \dim G_0\\
&= \alpha(T) - \dim H = \alpha(T)-\beta(T) = \operatorname{ind} T,
\end{aligned}$$

which ends the proof. □

**2.9. Theorem:** Let $K\in\mathcal{K}(X)$ and $\lambda\in\sigma(K)\setminus\{0\}$, and denote, for $k\in\mathbb{N}$, $N_k := N((\lambda-K)^k)$ and $R_k := R((\lambda-K)^k)$. Then one has

(a) For each $k\in\mathbb{N}$, $N_k$ and $R_k$ are closed and $\dim N_k = \operatorname{codim} R_k < \infty$.

(b) There is a minimal $p\in\mathbb{N}$ such that $N_p = N_{p+1}$.

(c) $N_{p+k}=N_p$ and $R_{p+k}=R_p$ for each $k\in\mathbb{N}$.

(d) $X = N_p\oplus R_p$, $(\lambda-K)N_p\subset N_p$, $((\lambda-K)|_{N_p})^p = 0$ and $\lambda-K : R_p\to R_p$ is an isomorphism.

In particular, $\sigma(K)\setminus\{0\}\subseteq\sigma_p(K)$. Moreover, $\sigma(K)\setminus\{0\}$ is finite or $\sigma(K)\setminus\{0\} = \{\lambda_n : n\in\mathbb{N}\}$ where $\lambda_n\to 0$.

End Lect.8

*Proof.* (a) Writing $(\lambda-K)^k = \lambda^k(I-KM_k)$ where $M_k\in\mathcal{L}(X)$ we see by 2.5 that $(\lambda-K)^k\in\Phi(X)$ and $\operatorname{ind}(\lambda-K)^k = 0$.

(b) If not, we find a sequence $(x_n)$ in $X$ such that, for each $n$, $x_n\in N_{n+1}\setminus N_n$ and $1 = d(x_n,N)\le\|x_n\|\le 2$. For $n>m$ we then have

$$\|Kx_n - Kx_m\| = \|\lambda x_n - \underbrace{((\lambda-K)x_n + Kx_m)}_{=:y}\|$$

and

$$(\lambda-K)^n y = (\lambda-K)^{n+1}x_n + K(\lambda-K)^n x_m = 0,$$

i.e. $y \in N_n$. Hence

$$\|Kx_n - Kx_m\| = \|\lambda x_n - y\| \geq |\lambda| d(x_n, N_n) = |\lambda| > 0,$$

in contradiction with $K \in \mathcal{K}(X)$.

(c) If $(\lambda - K)^{p+2}x = 0$ then $(\lambda - K)x \in N_{p+1} = N_p$ and $(\lambda - K)^{p+1}x = 0$, i.e. $x \in N_{p+1} = N_p$. We have shown $N_{p+2} = N_p$. Now we iterate and use (a).

(d) For each $k$, $\lambda - K : R_k \to R_{k+1}$ is surjective. Hence $\lambda - K : R_p \to R_p$ is surjective. But by 2.5, $(\lambda - K)|_{R_p} \in \Phi(R_p)$ has index 0, so $\lambda \in \rho(K|_{R_p})$. On the other hand, $\lambda - K : N_{k+1} \to N_k$ for each $k$, hence $\lambda - K : N_p \to N_p$ and $((\lambda - K)|_{N_p})^p = 0$. Since $\dim N_p < \infty$ we obtain form linear algebra that $\sigma(K|_{N_p}) = \{\lambda\}$.

In particular, there exists $\varepsilon > 0$ such that $B(\lambda, \varepsilon) \setminus \{\lambda\} \subset \rho(A)$. This implies that the only possible accumulation point of $\sigma(K) \setminus \{0\}$ is 0. □

**Example:** Let $X = C[0,1]$ and define $V \in \mathcal{L}(X)$ by $Vf(t) := \int_0^t f(s)\,ds$. Then $V : X \to X$ is compact, but $\sigma_p(V) = \emptyset$, since $\lambda f = Vf$ implies $f(0) = 0$ and $\lambda f' = f$, hence $f = 0$. By 2.6, $\sigma(V) = \{0\}$.

Via $(V^n f)(t) = \int_0^t \frac{(t-s)^{n-1}}{(n-1)!} f(s)\,ds$ it is also possible to show $\|V^n\| \leq \frac{1}{n!}$ which implies $r(V) = 0$ for the spectral radius of $V$.

We want to have a corresponding result for unbounded operators, and this can be achieved via resolvents.

**2.10. Lemma:** Let $A$ be a closed operator in $X$ with non-empty resolvent set. The following are equivalent:

(i) $I : [D(A)] \to X$ is compact,

(ii) there exists $\lambda_0 \in \rho(A)$ such that $R(\lambda_0, A) \in \mathcal{K}(X)$,

(iii) $R(\lambda, A) \in \mathcal{K}(X)$ for all $\lambda \in \rho(A)$.

In this case we say that $A$ *has compact resolvents*.

*Proof.* (i) ⇒ (ii): We choose $\lambda_0 \in \rho(A)$ and factorize $R(\lambda_0, A) = R(\lambda_0, A) \circ I$ where $R(\lambda_0, A) \in \mathcal{L}(X, [D(A)])$ and $I \in \mathcal{K}([D(A)], X)$.

(ii) ⇒ (iii): Use the resolvent equation.

(iii) ⇒ (i): Choose $\lambda_0 \in \rho(A)$ and factorize $I = R(\lambda_0, A)(\lambda_0 - A)$ where $\lambda_0 - A \in \mathcal{L}([D(A)], X)$ and $R(\lambda_0, A) \in \mathcal{K}(X)$. □

**2.11. Lemma:** Let $A$ be a closed operator in $X$, $\lambda_0 \in \rho(A)$ and $z \in \mathbb{C} \setminus \{\lambda_0\}$. Then

$$N(z - A) = N((\lambda_0 - z)^{-1} - R(\lambda_0, A)), \quad R(z - A) = R((\lambda_0 - z)^{-1} - R(\lambda_0, A)),$$

In particular, $z - A \in \Phi([D(A)], X)$ if and only if $(\lambda_0 - z)^{-1} - R(\lambda_0, A) \in \Phi(X)$. In this case, one has $\operatorname{ind}(z - A) = \operatorname{ind}((\lambda_0 - z)^{-1} - R(\lambda_0, A))$.

*Proof.* is an exercise. □

Combination of 2.9, 2.10, and 2.11 yields the following version of Theorem 2.9 for unbounded operators.

**2.12. Theorem:** Let $A$ be a closed linear operator in $X$ with $\rho(A) \neq \emptyset$ and compact resolvents. Let $z \in \sigma(A)$ and denote, for $k \in \mathbb{N}$, $N_k := N((z - A)^k)$ and $R_k := R((z - A)^k)$. Then one has

(a) For each $k \in \mathbb{N}$, $N_k$ and $R_k$ are closed and $\dim N_k = \operatorname{codim} R_k < \infty$.

(b) There is a minimal $p \in \mathbb{N}$ such that $N_p = N_{p+1}$.

(c) $N_{p+k} = N_p$ and $R_{p+k} = R_p$ for each $k \in \mathbb{N}$.

(d) $X = N_p \oplus R_p$, $(z-A)N_p \subset N_p$, $((z-A)|_{N_p})^p = 0$ and $A|_{R_p}$ with $D(A|_{R_p}) = D(A) \cap R_p$ is a closed linear operator in $R_p$ with $z \in \rho(A|_{R_p})$.

In particular, $\sigma(A) = \sigma_p(A)$. Moreover, $\sigma(A)$ is finite or $\sigma(A) = \{z_n : n \in \mathbb{N}\}$ where $|z_n| \to \infty$.

Observe that, for a fixed $\lambda_0 \in \rho(A)$, the operator $(\lambda_0 - z)^{-1} - R(\lambda_0, A) : R_p \to R_p$ is by 2.11 and 2.9 an isomorphism and that $R(\lambda_0, A)(R_p) = R_p \cap D(A) = D(A|_{R_p})$. End Lect.9

**Example:** Recall Exercise 20 where $X = \{f \in C[0,1] : f(0) = f(1)\}$ and $A = \frac{d}{dx}$ with $D(A) = \{f \in X \cap C^1[0,1] : f' \in X\}$. The operator $A$ has compact resolvents and $\sigma(A) = \sigma_p(A) = 2\pi i\mathbb{Z}$. For $z \in \sigma(A)$ one has $\alpha(z - A) = 1 = \beta(z - A)$.

Regarding Fredholm operators as generalizations of isomorphisms leads to the following notion.

**2.13. Definition:** Let $A$ be a closed linear operator in $X$. We define the *essential spectrum* of $A$ by

$$\sigma_{\text{ess}}(A) := \{\lambda \in \mathbb{C} : \lambda - A \notin \Phi([D(A)], X)\}.$$

Clearly, $\sigma_{\text{ess}}(A) \subseteq \sigma(A)$. Any $\lambda \in \mathbb{C} \setminus \sigma_{\text{ess}}(A)$ is called a *Fredholm point for A*.

**Example:** If $\dim X < \infty$ then $\sigma_{\text{ess}}(A) = \emptyset$ for any $A \in \mathcal{L}(X)$. If $K \in \mathcal{K}(X)$ and $\dim X = \infty$ then $\sigma_{\text{ess}}(K) = \{0\}$ ($R(K)$ is either finite dimensional or not closed). If $A$ is a closed operator in $X$ with non-empty resolvent set and compact resolvents then $\sigma_{\text{ess}}(A) = \emptyset$.

**Remarks:** (a) Observe that $\sigma_c(A) \subseteq \sigma_{\text{ess}}(A)$.

(b) By 2.5, $\sigma_{\text{ess}}(A)$ is closed, $\mathbb{C} \setminus \sigma_{\text{ess}}(A)$ is open, and the map $\mathbb{C} \setminus \sigma_{\text{ess}}(A) \to \mathbb{Z}$, $\lambda \mapsto \operatorname{ind}(\lambda - A)$ is continuous.

(c) By 2.11 we have: If $\lambda_0 \in \rho(A)$ then

$$\sigma_{\text{ess}}(R(\lambda_0, A)) = \Big\{ \frac{1}{\lambda_0 - z} - R(\lambda_0, A) : z \in \sigma_{\text{ess}}(A) \Big\}.$$

We give two examples.

**Example:** (a) Let $X = l^2$ and $L$ be the left shift $L(x_n) = (x_{n+1})$. We know $\sigma(L) = \{|\lambda| \leq 1\}$ and $\sigma_p(L) = \{|\lambda| < 1\}$ with $\alpha(\lambda - L) = 1$ for $|\lambda| < 1$. If we let $X_0 := \{(x_n) \in l^2 : x_1 = 0\}$, then $L : X_0 \to X$ is bijective and even isometric with $(L|_{X_0})^{-1} = R : X \to X_0$, in other words $LR = I_X$. For $|\lambda| < 1$ one has absolute convergence of $\sum_{k=0}^{\infty} \lambda^k R^{k+1}$ in operator norm and

$$(\lambda - L)\Big( - \sum_{k=0}^{\infty} \lambda^k R^{k+1} \Big) = - \sum_{k=0}^{\infty} \lambda^{k+1} R^{k+1} + \sum_{k=0}^{\infty} \lambda^k R^k = I_X.$$

Hence $R(\lambda - L) = X$ and $\beta(\lambda - L) = 0$, $\lambda - L \in \Phi(X)$ and $\operatorname{ind}(\lambda - L) = 1$ for any $|\lambda| < 1$. Continuity of the index then yields that $\sigma_{\text{ess}}(L) = \{|\lambda| = 1\}$.

(b) Let $X = \{f \in C[0,\infty) \lim_{x\to\infty} f(x) = 0\}$ with sup-norm $\|\cdot\|_\infty$, and $A = \frac{d}{dx}$, $D(A) = \{f \in X \cap C^1[0,\infty) : f' \in X\}$. For $\lambda \in \mathbb{C}$ and $f \in D(A)$ we have $(\lambda - A)f = 0$ if and only if $f' = \lambda f$, i.e. $f = ce^{\lambda(\cdot)}$. We thus obtain $\sigma_p(A) = \{\operatorname{Re}\lambda < 0\}$.

For $\operatorname{Re}\lambda > 0$ and $g \in X$ the unique solution $f \in D(A)$ of $(\lambda - A)f = g$ is given by

$$f(x) = ce^{\lambda x} - e^{\lambda x} \int_0^x e^{-\lambda t} g(t)\, dt,$$

where $c = c(g)$ has to be chosen such that $\lim_{x\to\infty} f(x) = 0$, i.e.

$$c(g) = \int_0^\infty e^{-\lambda t} g(t)\, dt,$$

so that

$$(\lambda - A)^{-1} f(x) = \int_x^\infty e^{\lambda(x-t)} g(t)\, dt = \int_0^\infty e^{-\lambda s} g(x+s)\, ds.$$

We obtain $\{\operatorname{Re}\lambda > 0\} \subseteq \rho(A)$. For $\operatorname{Re}\lambda < 0$ and $g \in X$ the unique solution of $(\lambda - A)f = g$, $f(0) = 0$ is given by

$$f(x) = \int_0^x e^{\lambda(x-t)} g(t)\, dt.$$

Observe that $f \in D(A)$, which means that $R(\lambda - A) = X$, $\beta(\lambda - A) = 0$, $\lambda - A \in \Phi(X)$ and $\operatorname{ind}(\lambda - A) = 1$ for any $\operatorname{Re}\lambda < 0$. Arguing as in (a) we now obtain $\sigma_{\text{ess}}(A) = \{\operatorname{Re}\lambda = 0\}$.

**2.14. Lemma:** Let $T \in \Phi(X, Y)$. Then $T' \in \Phi(Y', X')$ and

$$\alpha(T') = \beta(T), \quad \beta(T') = \alpha(T), \quad \operatorname{ind} T' = -\operatorname{ind} T.$$

*Proof.* will be an exercise. □

From this lemma we obtain via spectral mapping:

**2.15. Corollary:** Let $A$ be a closed and densely defined linear operator in $X$ with $\rho(A) \neq \emptyset$. Then $\sigma_{\text{ess}}(A') = \sigma_{\text{ess}}(A)$ and, for $z \in \mathbb{C} \setminus \sigma_{\text{ess}}(A)$,

$$\alpha(z - A') = \beta(z - A), \quad \beta(z - A') = \alpha(z - A), \quad \operatorname{ind}(z - A') = -\operatorname{ind}(z - A).$$

**2.16. Proposition:** Let $A$ be a closed operator in $X$ and $K \in \mathcal{K}([D(A)], X)$. Then the Banach spaces $[D(A)]$ and $[D(A+K)]$ are isomorphic and one has

$$\sigma_{\text{ess}}(A + K) = \sigma_{\text{ess}}(A).$$

*Proof.* The first part is Exercise 24, and the second part uses 2.8. □

We give a somehow typical example.

**Example:** Let $X = \{f \in C[0, \infty) : \lim_{x\to\infty} f(x) = 0\}$ with sup-norm $\|\cdot\|_\infty$, and $A = \frac{d}{dx}$, $D(A) = \{f \in X \cap C^1[0, \infty) : f' \in X\}$. We know that $\sigma_{\text{ess}}(A) = i\mathbb{R}$.

Let $m \in C[0, \infty)$ such that $m(t) = 0$ for $t \geq a$ where $a > 0$. Then $f \mapsto mf$ is compact from $[D(A)]$ to $X$. Hence, if we define $Bf := f' + mf$ for $f \in D(B) := D(A)$, then $\sigma_{\text{ess}}(B) = i\mathbb{R}$.

By an approximation argument, the same holds for $m \in X$. End Lect.10

For the following example we recall the Fourier transform:

For $f \in L^1(\mathbb{R}^d)$ the Fourier transform $\hat{f} : \mathbb{R}^d \to \mathbb{C}$ is defined by

$$\mathscr{F} f(\xi) := \hat{f}(\xi) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\xi\cdot x} f(x)\, dx, \quad \xi \in \mathbb{R}^d,$$

and $\mathscr{F} : L^1(\mathbb{R}^d) \to C_0(\mathbb{R}^d) = \{f \in C(\mathbb{R}^d) : \lim_{|x|\to\infty} f(x) = 0\}$. For $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ one has $\mathscr{F} \in L^2(\mathbb{R}^d)$ and

$$\|\mathscr{F} f\|_{L^2} = \|f\|_{L^2}.$$

The Fourier transform $\mathscr{F}$ extends to an isometry $\mathscr{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ whose inverse is, for $g \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, given by

$$\mathscr{F}^{-1}g(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^d} e^{i\xi\cdot x} g(\xi)\, d\xi, \quad x \in \mathbb{R}^d.$$

For $s \geq 0$ we define the *Bessel potential spaces* $H^{s,2}(\mathbb{R}^d)$ by

$$H^{s,2}(\mathbb{R}^d) := \{f \in L^2(\mathbb{R}^d) : \xi \mapsto (1 + |\xi|^2)^{s/2} \mathscr{F} f(\xi) \in L^2(\mathbb{R}^d)\},$$

equipped with the norm

$$\|f\|_{H^{s,2}} := \|\xi \mapsto (1 + |\xi|^2)^{s/2}\mathscr{F} f(\xi)\|_{L^2} = \Big( \int_{\mathbb{R}^d} (1 + |\xi|^2)^s |\mathscr{F} f(\xi)|^2\, d\xi \Big)^{1/2}.$$

Clearly, $H^{0,2}(\mathbb{R}^d) = L^2(\mathbb{R}^d)$ and $H^{s,2}(\mathbb{R}^d)$ is a Hilbert space for the inner product

$$(f|g)_{H^{s,2}} = \int_{\mathbb{R}^d} (1 + |\xi|^2)^s \mathscr{F} f(\xi)\overline{\mathscr{F} g(\xi)}\, d\xi.$$

**Example:** In $X = L^2(\mathbb{R}^d)$ we consider $A = -\Delta$, defined on $D(A) = H^{2,2}(\mathbb{R}^d)$ by

$$-\Delta f := \mathscr{F}^{-1}(\xi \mapsto |\xi|^2 \mathscr{F} f(\xi)).$$

We have $\|f\|_{L^2}^2 + \|\Delta f\|_{L^2} = \|f\|_{H^{2,2}}^2$, i.e. the graph norm is equivalent to $\| \cdot \|_{H^{2,2}}$, and $A$ is closed.

For $\lambda \in \mathbb{C} \setminus [0, \infty)$ we have $\lambda \in \rho(A)$ and

$$R(\lambda, A)f = (\lambda + \Delta)^{-1} f = \mathscr{F}^{-1}(\xi \mapsto \underbrace{(\lambda - |\xi|^2)^{-1}}_{\in L^\infty} \mathscr{F} f(\xi)), \quad f \in L^2(\mathbb{R}^d).$$

For $\lambda \geq 0$, we have $N(\lambda - A) = \{0\}$ and $R(\lambda - A) = \mathscr{F}^{-1}((\lambda - |\cdot|^2)H^{2,2})$ which is dense (it contains the dense set

$$\mathscr{F} - 1(\{f \in L^2 : \exists \varepsilon > 0 : f(x) = 0 \text{ if } |x| \geq \tfrac{1}{\varepsilon} \text{ or } ||x| - \sqrt{\lambda}| < \varepsilon \}) )$$

but not closed in $L^2$ (since $(\lambda - A)^{-1}$ is unbounded). Hence $\sigma(A) = \sigma_{\text{ess}}(A) = [0, \infty)$.

We claim that, for $s > d/2$, $H^{s,2}(\mathbb{R}^d)$ embeds into the space of Hölder continuous functions

$$C^\gamma(\mathbb{R}^d) = \{f \in C(\mathbb{R}^d) : \|f\|_\infty < \infty, \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\gamma} < \infty\},$$

if $\gamma \in (0, s - d/2) \cap (0, 1)$. (The space $C^\delta(\mathbb{R}^d)$ is a Banach space for the norm

$$\|f\|_{C^\delta} := \|f\|_\infty + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\delta},$$

if $\delta \in (0,1]$.) For $f \in H^{s,2}(\mathbb{R}^d)$, we can write

$$\mathscr{F}f(\xi) = \underbrace{(1+|\xi|^2)^{-s/2}}_{\in L^2} \underbrace{(1+|\xi|^2)^{s/2}\mathscr{F}f(\xi)}_{\in L^2} \in L^1,$$

which yields by Fourier inversion $f \in C_0(\mathbb{R}^d)$ and

$$\|f\|_\infty \le C_1 \|f\|_{H^{s,2}(\mathbb{R}^d)}.$$

Moreover, we even have $|\xi|^\gamma \mathscr{F}f(\xi) \in L^1$ and

$$\|\xi \mapsto |\xi|^\gamma \mathscr{F}f(\xi)\|_{L^1(\mathbb{R}^d)} \le C_2 \|H^{s,2}(\mathbb{R}^d)\|.$$

Writing

$$f(x) - f(y) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} (e^{ix\cdot\xi} - e^{iy\cdot\xi}) \mathscr{F}(\xi)\, d\xi$$

we obtain by

$$|e^{ix\cdot\xi} - e^{iy\cdot\xi}| \le \min\{2, c|x-y||\xi|\} \le C|x-y|^\gamma |\xi|^\gamma$$

that

$$|f(x) - f(y)| \le C'|x-y|^\gamma \underbrace{\||\cdot|^\gamma \mathscr{F}f\|_{L^1}}_{\le C_2\|f\|_{H^{s,2}}}.$$

The continuous inclusion $H^{s,2}(\mathbb{R}^d) \subset C^\gamma(\mathbb{R}^d)$ for $\gamma \in (0, \min\{1, s-d/2\})$ is proved. In particular, for $d = 3$ and $\gamma \in (0, 1/2)$, we thus have the continuous inclusion $H^{2,2}(\mathbb{R}^3) \subset C^\gamma(\mathbb{R}^3)$.

If $m \in L^\infty(\mathbb{R}^3)$ with $m(x) = 0$ for $|x| \ge a$ and some $a > 0$ then $f \mapsto mf$ is compact $H^{2,2}(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ since we can factorize via the compact restriction operator $C^\gamma(\mathbb{R}^3) \to C(\overline{B(0,a)})$ where $\gamma \in (0, 1/2)$: The map $H^{2,2}(\mathbb{R}^3) \to C^\gamma(\mathbb{R}^3)$, $f \mapsto f$, is bounded, the restriction $C^\gamma(\mathbb{R}^3) \to C(\overline{B(0,a)})$, $f \mapsto f|_{\overline{B(0,a)}}$, is compact (by Arzelá-Ascoli), and the map $g \mapsto (m\cdot g)_0$, where $h_0$ denotes the extension of $h$ by 0 outside of $\overline{B(0,a)}$, is bounded $C(\overline{B(0,a)}) \to L^2(\mathbb{R}^3)$.

We conclude that $\sigma_{\text{ess}}(-\Delta + m) = \sigma_{\text{ess}}(-\Delta) = [0, \infty)$.

# 3 Operators in Hilbert space

When nothing else is said, $H$ denotes a complex Hilbert space with inner product $(\cdot|\cdot)$.

Recall the following properties of an inner (or *scalar*) product:

- $(\cdot|\cdot)$ is linear in the first component,
- $(y|x) = \overline{(x|y)}$ for all $x, y \in H$,
- these two properties imply that $(\cdot|\cdot)$ is *antilinear* in the second component, i.e. $(x|\alpha y + z) = \overline{\alpha}(x|y) + (x|z)$ for al $x, y, z \in H$, $\alpha \in \mathbb{C}$.
- $(x|x) \geq 0$ and $(x|x) = 0 \iff x = 0$ for all $x \in H$.

These properties imply that $x \mapsto \|x\| := \sqrt{(x|x)}$ defines a norm on $H$ and that the *Cauchy-Schwarz inequality*

$$|(x|y)| \leq \|x\|\|y\|$$

holds for all $x, y \in H$.

A space $H$ equipped with such an inner product $(\cdot|\cdot)$ is called a *Hilbert space* if it is complete for the norm $\|\dot{}\|$ associated with the inner product.

Let $H' := \mathcal{L}(H, \mathbb{C})$ denote the dual space of $H$. Then the map

$$J_H : H \mapsto H', \quad y \mapsto (\cdot|y)$$

is bijective, isometric, and antilinear.

**3.1. Definition:** Let $H_1, H_2$ be Hilbert spaces and $T \in \mathcal{L}(H_1, H_2)$. The *adjoint operator* $T^* \in \mathcal{L}(H_2, H_1)$ of $T$ is defined by

$$(x|T^*y)_{H_1} = (Tx|y)_{H_2} \quad \text{for all } x \in H_1,\, y \in H_2,$$

i.e. $T^* = J_{H_1}^{-1} T' J_{H_2}$ where $T' \in \mathcal{L}(H_2', H_1')$ denotes the dual operator of $T$. We clearly have

$$\|T^*\|_{\mathcal{L}(H_2,H_1)} = \|T\|_{\mathcal{L}(H_1,H_2)}.$$

In the same way we can define the adjoint operator $A^*$ as an operator from $H_2$ to $H_1$ for a densely defined operator $A : H_1 \supseteq D(A) \to H_2$. End Lect.11

An operator $S \in \mathcal{L}(H)$ is called *self adjoint* if $S^* = S$, *normal* if $SS^* = S^*S$, and *unitary* if $SS^* = S^*S = I$.

**Rules:** $(T^*)^* = T$, $(S+T)^* = S^* + T^*$, $(\alpha T)^* = \overline{\alpha} T^*$, $(ST)^* = T^*S^*$.

**Remark:** For $S \in \mathcal{L}(H)$ we have

$$\begin{aligned} S \text{ self adjoint} &\iff \forall x, y \in H : (Sx|y) = (x|Sy) \\ S \text{ normal} &\iff \forall x, y \in H : (Sx|Sy) = (S^*x|S^*y) \\ S \text{ unitary} &\iff S \text{ is bijective and isometric.} \end{aligned}$$

**3.2. Lemma:** Let $S \in \mathcal{L}(H)$.

(a) $\|S^*S\| = \|SS^*\| = \|S\|^2$ and $S^*S$, $SS^*$ are self adjoint.

(b) If $S$ is normal, then

$$\|S\| = r(S) = \max\{|\lambda| : \lambda \in \sigma(S)\}.$$

*Proof.* (a) We have

$$\|Sx\|^2 = |(Sx|Sx)| = |(x|S^*Sx)| \leq \|x\|\|S^*Sx\|,$$

hence

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

(b) We calculate $r(S)$. Using (a) repeatedly and normality of $S$ once we have

$$\|S^2\|^2 = \|S^2(S^2)^*\| = \|(SS^*)^2\| = \|(SS^*)(SS^*)^*\| = \|SS^*\|^2 = \|S\|^4.$$

Hence $\|S^2\| = \|S\|^2$. Iteration yields $\|S^{2^k}\| = \|S\|^{2^k}$ for any $k \in \mathbb{N}$, from which $r(S) = \|S\|$ follows. □

**3.3. Definition und Lemma:** Let $T \in \mathcal{L}(H)$ and

$$W(T) := \{(Tx|x) : \|x\| = 1\}$$

be the *numerical range* of $T$. Then $\sigma(T) \subseteq \overline{W(T)}$.

*Proof.* Let $\lambda \notin \overline{W(T)}$. Then $d := d(\lambda, \overline{W(T)}) > 0$ and for $\|x\| = 1$:

$$d \leq |\lambda - (Tx|x)| = |((\lambda - T)x|x)| \leq \|(\lambda - T)x\| \cdot \|x\| = \|(\lambda - T)x\|.$$

Hence $\lambda - T$ is injective, $(\lambda - T)^{-1} : R(\lambda - T) \to H$ is bounded, and $R(\lambda - T)$ is closed. Now $R(\lambda - T)$ is dense if and only if $(\lambda - T)^* = \overline{\lambda} - T^*$ is injective. Since $W(T^*) = \{\overline{\mu} : \mu \in W(T)\}$, we have $d(\overline{\lambda}, \overline{W(T^*)}) = d(\lambda, W(T)) = d > 0$. The previous argument shows that $\overline{\lambda} - T^*$ is injective. □

**3.4. Corollary:** If $S \in \mathcal{L}(H)$ is self adjoint then $W(S) \subset \mathbb{R}$ and

$$\sigma(S) \subseteq [m, M] \subseteq [-\|S\|, \|S\|],$$

where $m := \inf\{(Sx|x) : \|x\| = 1\}$ and $M := \sup\{(Sx|x) : \|x\| = 1\}$.

*Proof.* For $x \in H$ with $\|x\| = 1$ we have

$$(Sx|x) = (x|Sx) = \overline{(Sx|x)},$$

i.e. $(Sx|x) \in \mathbb{R}$, and

$$|(Sx|x)| \leq \|Sx\|\|x\| \leq \|S\|.$$

The rest follows from 3.3. □

The following is a first spectral theorem for bounded self adjoint operators and establishes a *functional calculus* for functions that are continuous on the spectrum. We write $p_j(\lambda) = \lambda^j$ for $j \in \mathbb{N}_0$, so $p_0 = 1_{\sigma(S)}$ and $p_1 = \mathrm{id}_{\sigma(S)}$.

**3.5. Theorem:** Let $S \in \mathcal{L}(H)$ be self adjoint. There exists a unique continuous linear map $\Phi : C(\sigma(S)) \to \mathcal{L}(H)$ such that

$$\Phi(p_0) = I, \quad \Phi(p_1) = S, \quad \Phi(f \cdot g) = \Phi(f)\Phi(g), \quad f, g \in C(\sigma(S)),$$

[$\Phi$ is multiplicative, i.e. an *algebra homomorphism*.]

For any $f \in C(\sigma(S))$ the operator $\Phi(f)$ is normal,

$$\Phi(f)^* = \Phi(\overline{f}),$$

and $\Phi(f)$ is self adjoint if and only if $f$ is real-valued.

Moreover, $\Phi$ is an isometry, i.e. $\|\Phi(f)\|_{\mathcal{L}(H)} = \|f\|_{\infty,\sigma(S)}$ for all $f \in C(\sigma(S))$.

*Proof.* For polynomials $p(\lambda) = \sum_{j=0}^n a_j \lambda^j$ we define $\Phi(p) := \sum_{j=0}^n a_j S^j$, and let $\mathcal{P}$ the algebra of all polynomial functions $p : \sigma(S) \to \mathbb{C}$. Then $\Phi : \mathcal{P} \to \mathcal{L}(H)$ is an algebra homomorphism, $\Phi(p)^* = \Phi(\overline{p})$ and $\Phi(p)$ is normal for each $p \in \mathcal{P}$. Moreover, $\Phi(p)$ is self adjoint if $p$ is real-valued. Clearly, $\Phi(p_0) = I$ and $\Phi(p_1) = S$.

By Weierstraß, $\mathcal{P}$ is dense in $C(\sigma(S))$ for the sup-norm $\|\cdot\|_\infty$. So we only have to check that $\Phi : \mathcal{P} \to \mathcal{L}(H)$ is an isometry. For $p \in \mathcal{P}$ we have

$$\begin{aligned} \|\Phi(p)\|^2 &= \|\Phi(p)^*\Phi(p)\| = \|\Phi(\overline{p})\Phi(p)\| = \|\Phi(\overline{p}p)\| = r(\Phi(\overline{p}p)) \\ &= \max\{|\lambda| : \lambda \in \sigma(\Phi(\overline{p}p))\} = \max\{|\overline{p}p(\mu)| : \mu \in \sigma(S)\} = \|p\|^2_{\infty,\sigma(S)}, \end{aligned}$$

where we used the polynomial spectral mapping result from Exercise 14. If $\Phi(p)$ is self adjoint, then we obtain $\Phi(\overline{p}) = \Phi(p)^* = \Phi(p)$, from which $\overline{p} = p$, i.e. real-valuedness of $p$, follows by injectivity of $\Phi$.

The properties of $\Phi$ are preserved when approximating continuous functions by polynomials. $\square$

End Lect.12

As a first application we show:

**3.6. Lemma:** Let $S \in \mathcal{L}(H)$ be self adjoint and $m, M$ be as in Corollary 3.4. Then $m, M \in \sigma(S)$.

*Proof.* We give the proof for $m$ and may assume $m = 0$ by considering $S - m$ otherwise. We find a sequence $(x_n)$ with $\|x_n\| = 1$ such that $(Sx_n|x_n) \to 0+$. By 3.4 we have $\sigma(S) \subseteq [0, M]$. The function $q(t) := \sqrt{t}$ thus belongs to $C(\sigma(S))$ and we let $\sqrt{S} := \Phi(q) \in \mathcal{L}(H)$ where $\Phi$ is the functional calculus from 3.5. Then $\sqrt{S}$ is self adjoint, $(\sqrt{S})^2 = S$ and

$$\|\sqrt{S}x_n\|^2 = (\sqrt{S}x_n|\sqrt{S}x_n) = (Sx_n|x_n) \to 0 \quad (n \to \infty).$$

By $\sqrt{S} \in \mathcal{L}(H)$ this implies

$$Sx_n = \sqrt{S}(\sqrt{S}x_n) \to 0 \quad (n \to \infty).$$

But this means $0 \in \sigma_{ap}(S)$. □

Another application shows "optimality" of self adjoint operators with respect to resolvent bounds.

**3.7. Corollary:** Let $S \in \mathcal{L}(H)$ be self adjoint. Then, for any $\lambda \in \rho(S)$,

$$R(\lambda, S) = \Phi((\lambda - (\cdot))^{-1}) \quad \text{and} \quad \|R(\lambda, S)\| = \frac{1}{d(\lambda, \sigma(S))}.$$

*Proof.* We have

$$(\lambda - S)\Phi((\lambda - (\cdot))^{-1}) = \Phi(\lambda - (\cdot))\Phi((\lambda - (\cdot))^{-1}) = \Phi(1_{\sigma(S)}) = I,$$

and $\Phi((\lambda - (\cdot))^{-1})(\lambda - S) = I$ is proved similarly. Recall that $\Phi$ is isometric and observe $\|(\lambda - (\cdot))^{-1}\|_{\infty,\sigma(S)} = d(\lambda, \sigma(S))^{-1}$. □

**3.8. Definition:** we call a self adjoint operator $S \in \mathcal{L}(H)$ *positive* and write $S \geq 0$ if $(Sx|x) \geq 0$ for all $x \in H$. If $T \in \mathcal{L}(H)$ is another self adjoint operator then $S \geq T$ means $S - T \geq 0$, i.e. $(Sx|x) \geq (Tx|x)$ for all $x \in H$.

**3.9. Lemma:** Let $S \in \mathcal{L}(H)$ be self adjoint and $\Phi : C(\sigma(S)) \to \mathcal{L}(H)$ be the functional calculus for $S$ from 3.5. If $f, g \in C(\sigma(S))$ are real-valued and $f \geq g$ then $\Phi(f) \geq \Phi(g)$ in the sense of 3.8.

*Proof.* It suffices to prove the case $g = 0$, i.e. $f \geq 0$. Then $\sqrt{f}$ is well defined, $\sqrt{f} \in C(\sigma(S))$, and $\Phi(f) = \Phi(\sqrt{f}\sqrt{f}) = \Phi(\sqrt{f})\Phi(\sqrt{f})$, where $\Phi(\sqrt{f})$ is self adjoint. This implies, for $x \in H$,

$$(\Phi(f)x|x) = (\Phi(\sqrt{f})\Phi(\sqrt{f})x|x) = (\Phi(\sqrt{f})x|\Phi(\sqrt{f})x) \geq 0.$$

□

Before turning to the study of self adjoint compact operators we give a lemma on general self adjoint operators. Recall

- $x \perp y :\iff (x|y) = 0$ for $x, y \in H$,
- $M \perp N :\iff \forall x \in M, y \in N$: $(x|y) = 0$ for linear subspaces $M, N \subset H$,

- $M^\perp := \{x \in H : x \perp M\} = \{x \in H : \forall y \in M : x \perp y\}$ for linear subspaces $M$ of $H$.

If $M$ is a linear subspace of $H$ then $M^\perp$ is always closed, $\overline{M}^\perp = M^\perp$, and $(M^\perp)^\perp = \overline{M}$. One has always $H = \overline{M} \oplus M^\perp$ where $\overline{M} \perp M^\perp$.

**3.10. Lemma:** Let $S \in \mathfrak{L}(H)$ be self adjoint. Then:

(a) For $\lambda, \mu \in \sigma(S)$ with $\lambda \neq \mu$: $N(\lambda - S) \perp N(\mu - S)$.

(b) For $\lambda \in \sigma(S)$: $N((\lambda - S)^2) = N(\lambda - S)$.

(c) The space $H$ is the orthogonal direct sum of $N(S)$ and $\overline{R(S)}$, i.e. $H = N(S) \oplus \overline{R(S)}$ and $N(S) \perp \overline{R(S)}$.

*Proof.* (a) For $x \in N(\lambda - S)$ and $y \in N(\mu - S)$ we have

$$\lambda(x|y) = (\lambda x|y) = (Sx|y) = (x|Sy) = \mu(x|y),$$

and $x \perp y$, since $\lambda \neq \mu$.

(b) Let $x \in N((\lambda - S)^2)$. Then

$$\|(\lambda - S)x\|^2 = ((\lambda - S)x|(\lambda - S)x) = (x|\underbrace{(\lambda - S)^2 x}_{=0}) = 0,$$

i.e. $(\lambda - S)x = 0$, and $x \in N(\lambda - S)$. The reverse inclusion is clear.

(c) Let $x \in N(S)$ and $Sy \in R(S)$. Then $(x|Sy) = (Sx|y) = 0$, hence $N(S) \perp \overline{R(S)}$. If, on the other hand, $x \in H$ such that $x \perp \overline{R(S)}$ then

$$\|Sx\|^2 = (Sx|Sx) = (x|\underbrace{S^2 x}_{\in R(S)}) = 0,$$

i.e. $Sx = 0$ and $x \in N(S)$. Hence $N(S) = \overline{R(S)}^\perp$. □

**3.11. Theorem:** Let $S \in \mathfrak{L}(H)$ be compact and self adjoint and $\dim H = \infty$. There exists a real sequence $(\lambda_n)_{n \in \mathbb{N}_0}$ with $\lambda_n \to 0$ and an orthonormal sequence $(e_n)_{n \in \mathbb{N}_0}$ in $H$ such that

$$S = \sum_{n=0}^{\infty} \lambda(\cdot|e_n)e_n,$$

where the series converges in operator norm.

*Proof.* We apply Theorem 2.9 to $S$. By 3.10(b) we know that $p = 1$ for every $\lambda \in \sigma(S)\backslash\{0\}$. We also know that $\sigma(S)$ consists of a null sequence. For any $\lambda \in \sigma(S)\backslash\{0\}$ we choose a finite orthonormal basis of $N(\lambda - S)$. We obtain thus an orthonormal sequence $(e_n)_{n \in N}$ where we assume that the corresponding eigenvalues are ordered such that $|\lambda_1| \geq |\lambda_2| \geq \ldots$. Here $N = \{0, \ldots, n_0\}$ is finite (if $\sigma(S) \setminus \{0\}$ is finite) or $N = \mathbb{N}_0$.

We let $P := \sum_{n \in N} (\cdot|e_n)e_n$. Then $P$ is the orthogonal projection onto $H_1 := \overline{\operatorname{span}\{e_n : n \in N\}}$. Letting $H_0 := H_1^\perp$ we have, for $x \in H_0$ and $n \in N$,

$$(Sx|e_n) = (x|Se_n) = \lambda_n(x|e_n) = 0.$$

i.e. $Sx \in H_0$. This means that $S_0 := S|_{H_0} \in \mathcal{L}(H_0)$. Clearly, $S_0$ is self adjoint and compact. By 2.9, $\sigma(S_0) \setminus \{0\} = \emptyset$, hence $S_0 = 0$ by 3.2(b), and $H_0 \subset N(S)$. On the other hand, $N(S)$ is by 3.10(a) orthogonal to each $N(\lambda_n - S)$, $n \in N$. Hence $N(S) \subset H_0$, and we obtain $H_0 = N(S)$, and by 3.10(c) also $H_1 = \overline{R(S)}$. End Lect.13

But then

$$Sx = SPx = \sum_{n \in N} \lambda_n (x|e_n)e_n \quad \text{for all } x \in H.$$

If $N = \{0, \ldots, n_0\}$ is finite, we set $\lambda_n = 0$ for $n > n_0$, and choose an orthonormal sequence $(e_n)_{n > n_0}$ in $N(S)$.

For $x \in H$ and $k \in \mathbb{N}$ we have, by Pythagoras and Bessel's inequality,

$$\|Sx - \sum_{n=0}^{k} \lambda_n (x|e_n)e_n\|^2 = \sum_{n>k} |\lambda_n (x|e_n)|^2 \leq \|x\|^2 (\sup_{n>k} |\lambda_n|)^2,$$

which proves convergence of the series in operator norm, since $(\lambda_n)$ is a null sequence. □

**3.12. Theorem:** Let $G$ be another Hilbert space with $\dim G = \infty$ and $T \in \mathcal{K}(H, G)$. Then there exists a decreasing null sequence $(s_n)_{n \in \mathbb{N}_0}$ in $[0, \infty)$ and orthonormal systems $(e_n)_{n \in \mathbb{N}_0}$ in $H$ and $(f_n)_{n \in \mathbb{N}_0}$ in $G$ such that

$$T = \sum_{n=0}^{\infty} s_n (\cdot|e_n) f_n,$$

where the series converges in operator norm.

*Proof.* The operator $T^*T \in \mathcal{L}(H)$ is compact and self adjoint. Moreover, $T^*T \geq 0$ in the sense of 3.8, and $\sigma(T^*T) \subset [0, \|T\|^2]$ by 3.4. By 3.11 we obtain a decreasing null sequence $(s_n)_{n \in \mathbb{N}_0}$ and an orthonormal system $(e_n)_{n \in \mathbb{N}_0}$ in $H$ such that

$$T^*T = \sum_{n=0}^{\infty} s_n^2 (\cdot|e_n)e_n.$$

For $n \in \mathbb{N}_0$ with $s_n > 0$ we let $f_n := s_n^{-1} T e_n$. For $n, m \in \mathbb{N}_0$ with $s_n s_m > 0$ we then have

$$(f_n|f_m) = (s_n s_m)^{-1} (Te_n|Te_m) = (s_n s_m)^{-1} (T^*Te_n|e_m) = \frac{s_n^2}{s_n s_m} (e_n|e_m) = \delta_{nm}.$$

If $N = \{s \in \mathbb{N}_0 : s_n > 0\}$ is finite then we extend $(f_n)_{n\in N}$ to an orthonormal sequence $(f_n)_{n\in\mathbb{N}_0}$ in $G$. If $y \perp e_n$ for all $n \in N$ then

$$\|Ty\|^2 = (T^*Ty|y) = 0$$

by the representation of $T^*T$. Hence, for any $x \in H$,

$$\begin{aligned} Tx &= T\Big(\underbrace{x - \sum_{n\in N}(x|e_n)e_n}_{\in N(T)}\Big) + T\Big(\sum_{n\in N}(x|e_n)e_n\Big) \\ &= \sum_{n\in N}(x|e_n)Te_n = \sum_{n\in N} s_n(x|e_n)f_n = \sum_{n=0}^{\infty} s_n(x|e_n)f_n. \end{aligned}$$

The proof for convergence in operator norm is the same as in 3.11. □

**3.13. Corollary:** Let $G$ be another Hilbert space and

$$\mathcal{F}(H,G) := \{T \in \mathcal{L}(H,G) : \dim R(T) < \infty\}$$

denote the space of *finite rank operators* from $H$ to $G$. Then $\mathcal{F}(H,G)$ is dense in $\mathcal{K}(H,G)$ with respect to operator norm.

**Remark:** This result is in general false in the Banach space context (by an example of Enflo 1973). The *rank* of an operator in the dimension of its range.

**3.14. Definition and Remark:** A representation of $T \in \mathcal{K}(H,G)$ as a series with the properties of 3.12 is called a *Schmidt representation* of $T$. The sequence $(s_n)_{n\in\mathbb{N}_0}$ is uniquely determined by $T$ (as the decreasing eigenvalue sequence of $T^*T$), but in general not the orthonormal systems $(e_n)$ and $(f_n)$. The sequence $(s_n)_{n\in\mathbb{N}_0}$ is called the sequence of *singular values* of the operator $T$ and denoted by $(s_n(T))_{n\in\mathbb{N}_0}$. The following gives a characterization of the singular values of an operator as *approximation numbers*:

For $T \in \mathcal{K}(H,G)$ and $n \in \mathbb{N}_0$ one has

$$s_n(T) = \inf\{\|T-U\| : U \in \mathcal{L}(H,G), \dim R(U) \le n\} =: \alpha_n(T).$$

Here $\alpha_n(T)$ measures how good $T$ can be approximated by operators of rank of at most $n$.

*Proof.* Let $\sum_{j=0}^{\infty} s_j(\cdot|e_j)f_j$ be a Schmidt representation of $T$. Then, for $n \in \mathbb{N}_0$ and $x \in H$, we have by Bessel's inequality

$$\|Tx - \sum_{j=0}^{n-1} s_j(x|e_j)f_j\|^2 \le \sum_{j=n}^{\infty} s_j^2|(x|e_j)|^2 \le s_n^2\|x\|^2.$$

This implies that $\alpha_n(T) \le s_n$.

Now let $U \in \mathcal{L}(H,G)$ with $\dim R(U) \le n$. Th restriction of $U$ to the $(n+1)$-dimensional space span $\{e_0, \ldots, e_n\}$ has non-trivial kernel, so we find $y = \sum_{j=0}^{n} \xi_j e_j$ with $\|y\| = 1$ and $Uy = 0$. By Pythagoras we thus obtain

$$\|T - U\|^2 \ge \|(T-U)y\|^2 = \|Ty\|^2 = \|\sum_{j=0}^{n} s_j \xi_j f_j\|^2 = \sum_{j=0}^{n} s_j^2 |\xi_j|^2 \ge s_n^2 \sum_{j=0}^{n} |\xi|^2 = s_n^2.$$

Hence $\alpha_n(T) \ge s_n$. □

**Remark:** For $1 \le p < \infty$ the *Schatten p-class* is defined by

$$S_p(H,G) := \{T \in \mathcal{K}(H,G) : (s_n(T))_{n \in \mathbb{N}_0} \in l^p\}$$

and $\nu_p(T) := (\sum_{n=0}^{\infty} s_n(T)^p)^{1/p}$ for $T \in S_p(H,G)$. Elements of $S_2(H,G)$ are called *Hilbert-Schmidt opertors* and elements of $S_1(H,G)$ are called *nuclear operators* or, for $G = H$, are said to be of *trace class*.

In many respects, the spaces $S_p(H,G)$ may be viewed as "non-commutative" analogs of the spaces $l^p$ (cf., e.g., §16 in Meise/Vogt "Introduction to Functional Analysis"). End Lect.14

**Remark:** If $H$ and $G$ are separable infinite-dimensional Hilbert spaces, then the sequences $(e_n)_{n \in \mathbb{N}_0}$ and $(f_n)_{n \in \mathbb{N}}$ in 3.11 and 3.12 can be chosen to be orthonormal bases of $H$ and $G$, respectively. This can be seen form the proof.

A related observation is that, for a compact operator $T \in \mathcal{K}(H,G)$, the space $\overline{R(T)}$ is always separable. Moreover, if $T \in \mathcal{K}(H)$ is self adjoint and injective, then the space $H$ is separable (since the space $H_0$ in the proof is trivial).

# 4 The spectral theorem for self adjoint operators

For a self adjoint operator $S \in \mathcal{L}(H)$, we want to extend the functional calculus of Theorem 3.5 to a larger class of functions than $C(\sigma(S))$.

To this end we first identify the dual space of $C[a,b]$ where $[a,b] \subset \mathbb{R}$ is a non-trivial compact interval.

**4.1. Functions of bounded variation and Stieltjes-Integral:** We call a function $g : [a,b] \to \mathbb{C}$ of *bounded variation* (and write $g \in BV[a,b]$) if

$$\|g\|_{BV} := \sup\Big\{\sum_{j=1}^{n} |g(t_j) - g(t_{j-1})| : n \in \mathbb{N}, a = t_0 < t_1 < \ldots < t_n = b\Big\} < \infty.$$

For $f \in C[a,b]$ and $g \in B[a,b]$ we define the *Stieltjes integral* of $f$ with respect to $g$ by

$$\int_a^b f(t)\, dg(t) = \lim \Big[\sum_{j=1}^{n} f(\xi_j)(g(t_j) - g(t_{j-1}))\Big],$$

where $a = t_0 < t_1 < \ldots < t_n = b$, $\xi_j \in [t_{j-1}, t_j]$ for $j = 1, \ldots, n$, and the limit is taken for $\max |t_j - t_{j-1}| \to 0$ (the argument for existence of the limit is similar to the case $g(t) = t$ which gives the Riemann integral). Clearly, one has

$$\Big|\int_a^b f(t)\, dg(t)\Big| \leq \|f\|_\infty \|g\|_{BV}.$$

**Remarks:** (a) Any monotone function $g : [a,b] \to \mathbb{R}$ is of bounded variation and $\|g\|_{BV} = |g(b) - g(a)|$.

(b) Any real-valued function of bounded variation is the difference of two increasing functions.

(c) If $g \in C^1[a,b]$ then $g$ is of bounded variation, $\|g\|_{BV} = \int_a^b |g'(t)|\, dt$, and

$$\int_a^b f(t)\, dg(t) = \int_a^b f(t) g'(t)\, dt.$$

This holds also for so-called *absolutely continuous* functions $g : [a,b] \to \mathbb{C}$.

(d) If we let $BV_0[a,b] := \{g \in BV[a,b] : g(a) = 0\}$, then $(BV_0[a,b], \|\cdot\|_{BV})$ is a Banach space. On $BV[a,b]$ however, $\|\cdot\|_{BV}$ is not a norm, since $\|1_[a,b]\|_{BV} = 0$.

(e) A function of bounded variation has one-sided limits at every $t \in [a,b]$. For any $g \in BV[a,b]$ there is a countable subset $M \subset [a,b]$ such that $g$ is continuous at every $t \in [a,b] \setminus M$.

**4.2. Theorem:** Let $[a,b] \subset \mathbb{R}$ be a non-trivial compact interval. Any $\phi \in (C[a,b])'$ has a representation as a Stieltjes integral

$$\phi(f) = \int_a^b f(t)\, dg(t), \quad f \in C[a,b],$$

where $g \in BV_0[a,b]$ and $\|\phi\| = \|g\|_{BV}$.

*Proof.* We denote by $B[a,b]$ the space of all bounded function on $[a,b]$, equipped with the sup-norm $\|\cdot\|_\infty$. $C[a,b]$ is a closed subspace of $B[a,b]$, and by Hahn-Banach $\phi$ has an extension $\psi \in B[a,b]'$ with $\|\psi\| = \|\phi\|$. We now define $g : [a,b] \to \mathbb{C}$ by $g(a) = 0$ and $g(t) := \psi(1_{[a,t]})$ for $t \in (a,b]$. Then we have, for $a = t_0 < t_1 < \ldots < t_n = b$ and with $\varepsilon_j = \operatorname{sgn}(g(t_j) - g(t_{j-1}))$, $j = 1, \ldots, n$,

$$\begin{aligned}
\sum_{j=1}^n |g(t_j) - g(t_{j-1})| &= \varepsilon_1 g(t_1) + \sum_{j=2}^n \varepsilon_j (g(t_j) - g(t_{j-1})) \\
&= \varepsilon_1 \psi(1_{[a,t_1]}) + \sum_j \varepsilon_{j=2}^n (\psi(1_{[a,t_j]}) - \psi(1_{[a,t_{j-1}]})) \\
&= \psi\Big(\varepsilon_1 1_{[a,t_1]} + \sum_{j=2}^n \varepsilon_j 1_{(t_{j-1},t_j]}\Big) \\
&\leq \|\psi\| \|\varepsilon_1 1_{[a,t_1]} + \sum_{j=2}^n \varepsilon_j 1_{(t_{j-1},t_j]}\|_\infty \leq \|\phi\|.
\end{aligned}$$

This means that $\|g\|_{BV} \leq \|\phi\|$.

Any $f \in C[a,b]$ is uniformly continuous and thus

$$f(a + \frac{b-a}{n}) 1_{[a,a+\frac{b-a}{n}]} + \sum_{k=2}^n f(a + k\frac{b-a}{n}) 1_{(a+(k-1)\frac{b-a}{n}, a+k\frac{b-a}{n}]}$$

converges to $f$ as $n \to \infty$ in $B[a,b]$. Hence

$$\begin{aligned}
\phi(f) &= \psi(f) = \lim_{n\to\infty} f(a + \frac{b-a}{n}) \psi(1_{[a,a+\frac{b-a}{n}]}) + \sum_{k=2}^n f(a + k\frac{b-a}{n}) \psi(1_{(a+(k-1)\frac{b-a}{n}, a+k\frac{b-a}{n}]}) \\
&= \lim_{n\to\infty} \sum_{k=1}^n f(a + k\frac{b-a}{n})(g(a + k\frac{b-a}{n}) - g(a + (k-1)\frac{b-a}{n})) \\
&= \int_a^b f(t)\, dg(t),
\end{aligned}$$

which proves the desired representation. Moreover, by the estimate in 4.1 this implies

$$|\phi(f)| \leq \|f\|_\infty \|g\|_{BV}, \quad f \in C[a,b],$$

i.e. $\|\phi\| \leq \|g\|_{BV}$. □

**4.3. Discussion:** The $g \in BV_0[a,b]$ in 4.2 is not unique. We can obtain uniqueness if we require in addition that $g$ is continuous from the right at avery point $t \in (a,b)$. This does not affect the integrals $\int_a^b f(t)\,dg(t)$ for continuous $f$, and in the situation of 4.2 it also does not affect the norm $\|g\|_{BV}$.

If we now define $\widetilde{g} : \mathbb{R} \to \mathbb{C}$ by $\widetilde{g}(t) = 0$ for $t < a$, $\widetilde{g}(a) = g(a+)$, $\widetilde{g} = g$ on $(a,b]$, and $\widetilde{g}(t) = g(b)$ for $t > b$ then $\widetilde{g}$ is of bounded variation on $\mathbb{R}$ and continuous from the right in every point.

Moreover, $\widetilde{g}$ is the distribution function of a countably additive complex measure $\mu$ on the Borel subsets of $\mathbb{R}$, the relation is given by $\mu((c,d]) = \widetilde{g}(d) - \widetilde{g}(c)$ for all $-\infty \le c < d \le \infty$, and $\mu$ can be written as

$$\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4),$$

where $\mu_j$, $j = 1,2,3,4$, are positive and finite measures on the Borel subsets of $\mathbb{R}$ that vanish outside $[a,b]$. We put $\|\mu\| := \|g\|_{BV}$ (observe that $\|g\|_{BV[a,b]} = \|\widetilde{g}\|_{BV(I)}$ for any compact interval $I \supset [a,b]$ with $\min I < a$).

Any bounded Borel measurable function $f : \mathbb{R} \to \mathbb{C}$ can be integrated with respect to $\mu$, and

$$\int f\,d\mu := \int f\,d\mu_1 - \int f\,d\mu_2 + i\Big(\int f\,d\mu_3 - \int f\,d\mu_4\Big).$$

The same holds for any bounded Borel measurable function $f : [a,b] \to \mathbb{C}$. In particular, any $f \in C[a,b]$ can be integrated with respect to $\mu$, and we have

$$\int_{[a,b]} f\,d\mu = \int_a^b f(t)dg(t)$$

where the right hand side is the Stieltjes integral from 4.1. End Lect.15

Moreover, by the dominated convergence theorem we have the following property for integration with respect to $\mu$ (since it holds for integration with respect to each of the $\mu_j$):

If $(f_n)$ is a bounded sequence of bounded Borel measurable functions (i.e. $\sup_n \|f_n\|_\infty < \infty$) that converges pointwise to a function $f$, then $\int f_n\,d\mu \to \int f\,d\mu$.

For convenience we recall the following definitions and facts:

- $\sigma$-algebra: Let $\Omega \neq \emptyset$ be a set. A *$\sigma$-algebra in* $\Omega$ is a system $\mathcal{A}$ of subsets of $\Omega$ having the properties: $\emptyset \in \mathcal{A}$, if $A \in \mathcal{A}$ then $\Omega \setminus A \in \mathcal{A}$, if $(A_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ then $\bigcup_{j\in\mathbb{N}} \in \mathcal{A}$.
- A (*countably additive*) *complex measure* is a map $\mu : \mathcal{A} \to \mathbb{C}$ defined on a $\sigma$-algebra $\mathcal{A}$ such that $\mu(\bigcup_{j\in\mathbb{N}} A_j) = \sum_{j\in\mathbb{N}} \mu(A_j)$ for any sequence $(A_j)_{j\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{A}$ (notice that the sequence converges unconditionally, hence absolutely).
- The *Borel $\sigma$-algebra* $\mathcal{B}$ in $\mathbb{R}$ is the smallest $\sigma$-algebra in $\mathbb{R}$ that contains all open subsets of $\mathbb{R}$. Equivalently, $\mathcal{B}$ is the smallest $\sigma$-algebra in $\mathbb{R}$ containing all open intervals (or: all closed intervals, or: all intervals of the form $(-\infty,b]$, $b \in \mathbb{R}$).

- For a set $M \in \mathcal{B}$, the system of Borel subsets of $M$, i.e. $\{B \in \mathcal{B} : B \subseteq M\}$, is a $\sigma$-algebra in $M$, which is the smallest $\sigma$-algebra in $M$ that contains all sets $(-\infty, b] \cap M$, $b \in \mathbb{R}$.
- A function $f : \mathbb{R} \to \mathbb{C}$ is *Borel measurable* if $f^{-1}(Q) \in \mathcal{B}$ for any open subset $Q \subset \mathbb{C}$ or, equivalently, if $\{t : \operatorname{Re} f(t) \le b\}, \{t : \operatorname{Im} f(t) \le b\} \in \mathcal{B}$ for any $b \in \mathbb{R}$.

The followiing extends 4.2, 4.3.

**4.4. Corollary:** Let $K \subset \mathbb{R}$ be non-empty and compact and $\phi \in C(K)'$ then there exists a unique complex measure $\mu$ on the Borel subsets of $\mathbb{R}$ that vanishes outside $K$ such that

$$\phi(f) = \int_K f \, d\mu, \quad f \in C(K).$$

This measure satisfies $\|\mu\| = \|\phi\|$.

*Proof.* Let $a := \min K$ and $b := \max K$. The restriction map $F \mapsto F|_K$ is continuous and surjective $C[a, b] \to C(K)$.[2] We can represent the functional $F \mapsto \phi(F|_K)$ by 4.2 and 4.3 as an integral with respect to a complex Borel measure $\mu$. This measure vanishes outside $K$ and satisfies $\|\mu\| = \|\phi\|$. Uniqueness is shown below. □

**4.5. Definition:** For a compact subset $K \subset \mathbb{R}$ we let

$$\mathcal{B}_b(K) := \{f : K \to \mathbb{C} : \ f \text{ is Borel measurable and bounded}\},$$

equipped with the norm $\|f\|_\infty := \sup_{t \in K} |f(t)|$ (observe that $\mathcal{B}_b(K)$ is really a space of functions, not a space of equivalence classes!). Clearly, $\mathcal{B}_b(K)$ is a Banach space.

If $(f_n)$ is a sequence of functions $M \to \mathbb{C}$ and $f : M \to \mathbb{C}$ is a function we write $f_n \to f$ *bounded pointwise*, if $f_n \to f$ pointwise and $\sup_n \|f_n\|_\infty < \infty$. Clearly, $\mathcal{B}_b(K)$ is closed under bounded pointwise convergence.

We call a map $\Psi : \mathcal{B}_b(K) \to \mathcal{L}(H)$ *$\sigma$-continuous* if $f_n \to f$ bounded pointwise implies $(\Psi(f_n)x|y) \to (\Psi(f)x|y)$ for all $x, y \in H$.

**Example:** If $\mu$ is as in 4.4 then $\mathcal{B}_b(K) \to \mathbb{C}$, $f \mapsto \int_K f \, d\mu$ is $\sigma$-continuous. Here, $H = \mathbb{C}$ and $\mathcal{L}(\mathbb{C})$ is identified with $\mathbb{C}$.

[2] The set $[a, b] \setminus K$ is the disjoint union of an at most countable family $(I_j)_{j \in N}$ of open intervals $I_j = (a_j, b_j)$. We extend a given $f \in C(K)$ to a function $F \in C[a, b]$ by letting $F(t) := \frac{t - a_j}{b_j - a_j} f(a_j) + \frac{b_j - t}{b_j - a_j} f(b_j)$ for $t \in I_j$ and $F(t) := f(t)$ for $t \in K$. Then $F|_K = f$, $\|F\|_{\infty,[a,b]} = \|f\|_{\infty,K}$ and $f \mapsto F$ is linear.

**4.6. Lemma:** Let $K \subset \mathbb{R}$ be compact. $\mathcal{B}_b(K)$ is the smallest subset $M$ of $\mathbb{C}^K$ such that

(1) $C(K) \subset M$,
(2) $M$ is closed under bounded pointwise convergence.

*Proof.* Let $M_0$ denote the smallest subset $M$ of $\mathbb{C}^K$ with (1) and (2), i.e. the intersection of all subsets $M$ of $\mathbb{C}^K$ with (1) and (2). Since $\mathcal{B}_b(K)$ satisfies (1) and (2), we have $M_0 \subseteq \mathcal{B}_b(K)$.

For $\lambda \in \mathbb{C} \setminus \{0\}$, $\frac{1}{\lambda}M_0$ satisfies (1) and (2), so $M_0 \subseteq \frac{1}{\lambda}M_0$, i.e. $\lambda M_0 \subseteq M_0$.

If $f \in C(K)$ then $M_0 - f$ satisfies (1), (2), so $M_0 \subseteq M_0 - f$, i.e. $f + M_0 \subseteq M_0$. Hence $C(K) + M_0 \subseteq M_0$. Let $S := \{f : f + M_0 \subseteq M_0\}$. We have just shown that $C(K) \subset S$. If $(f_n)$ is a sequence in $S$ with $f_n \to f$ bounded pointwise and $g \in M_0$, then $(f_n + g)$ is a sequence in $M_0$ ( by $f_n \in S$) and $f_n + g \to f + g$ bounded pointwise. By (2) for $M_0$ we have $f + g \in M_0$. Since $g \in M_0$ was arbitrary, we have $f \in S$, and (2) for $S$ is verified. We obtain $M_0 \subseteq S$, i.e. $M_0 + M_0 \subseteq M_0$.

We have thus shown that $M_0$ is a complex vector space. Moreover $M := \{g : |g| \in M_0\}$ satisfies (1), (2), hence $M_0 \subseteq M$, i.e. $f \in M_0$ implies $|f| \in M_0$. Since $M_0$ is vector space, we obtain

$$\max\{f, g\}, \min\{f, g\} = \frac{1}{2}((f + g) \pm |f - g|) \in M_0$$

for real-valued $f, g \in M_0$.

If we now let $\mathcal{F} := \{A \subset K : 1_A \in M_0\}$, then $\mathcal{F}$ is a $\sigma$-algebra in $K$: $\emptyset, K \in \mathcal{F}$ is clear by $0, 1_K \in C(K) \subset M_0$. If $A \in \mathcal{F}$, then $K \setminus A \in \mathcal{F}$, since $M_0$ is a vector space. If $(A_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{F}$, then $\bigcup_{j=1}^n A_j \in \mathcal{F}$ since the max of a finite family of functions in $M_0$ belongs to $M_0$. Finally, $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{F}$ by (2).

Again by (2), $\mathcal{F}$ contains all sets $(-\infty, b] \cap K$, $b \in \mathbb{R}$, hence all Borel subsets of $K$. Since each $f \in \mathcal{B}_b(K)$ can be uniformly approximated by a sequence $(f_n)$ of Borel step functions on $K$, we conclude that $\mathcal{B}_b(K) \subseteq M_0$. □

End Lect.16

**4.7. Lemma:** Let $K \subset \mathbb{R}$ be compact. If $\Phi, \Psi : \mathcal{B}_b(K) \to \mathcal{L}(H)$ are $\sigma$-continuous and coincide on $C(K)$ then $\Phi = \Psi$.

*Proof.* Let

$$M := \{f \in \mathcal{B}_b(K) : \Phi(f) = \Psi(f)\} = \{f \in \mathcal{B}_b(K) : \forall x, y \in H : (\Phi(f)x|y) = (\Psi(f)x|y)\}.$$

Then $M$ satisfies (1), (2) from 4.6, hence $M = \mathcal{B}_b(K)$, i.e. $\Phi = \Psi$. □

**Example:** $\mu$ in 4.4 is unique since, by 4.7, the $\sigma$-continuous map $\mathcal{B}_b(K) \to \mathbb{C}$, $f \mapsto \int_K f\, d\mu$ is uniquely determined by its restriction to $C(K)$ and this restriction is the given linear functional $\phi \in C(K)'$.

**Remark:** Let $\beta : H \times H \to \mathbb{C}$ be *sesquilinear* (i.e. linear in the first and antilinear in the second component) and $M \geq 0$ such that

$$|\beta(x,y)| \leq M\|x\|\|y\| \quad \text{for all } x,y \in H.$$

Then there exists a unique operator $B \in \mathcal{L}(H)$ such that

$$\beta(x,y) = (Bx|y) \quad \text{for all } x,y \in H,$$

and one has $\|B\| \leq M$.

*Proof.* For any $x \in H$, we have $\overline{\beta(x,\cdot)} \in H'$ and define $Bx := J_H^{-1}(\overline{\beta(x,\cdot)})$. Then $B$ is linear (since it is a composition of two antilinear maps), and $\|Bx\| \leq M\|x\|$ for all $x \in H$, i.e. $B \in \mathcal{L}(H)$ and $\|B\| \leq M$. For $x,y \in H$ we have

$$(Bx|y) = \overline{(y|Bx)} = \overline{(J_H(Bx))(y)} = \overline{\overline{\beta(x,\cdot)}(y)} = \beta(x,y).$$

Uniqueness of $B$ is clear. □

**4.8. Theorem:** Let $S \in \mathcal{L}(H)$ be self adjoint. Then there exists a unique $\sigma$-continuous extension $\Psi : \mathcal{B}_b(\sigma(S)) \to \mathcal{L}(H)$ of the functional calculus $\Phi : C(\sigma(S)) \to \mathcal{L}(H)$ from 3.5. The map $\Psi$ is linear and multiplicative, i.e. an algebra homomorphism. Moreover, for any $f \in \mathcal{B}_b(\sigma(S))$, the operator$\Psi(f)$ is normal and $\Psi(\overline{f}) = \Psi(f)^*$.

*Proof.* Uniqueness follows from 4.7. For the existence proof, we put, for $x,y \in H$,

$$\phi_{x,y} : C(\sigma(K)) \to \mathbb{C},\ f \mapsto \phi_{x,y}(f) := (\Phi(f)x|y).$$

Then $\phi_{x,y} \in C(\sigma(S))'$ and $\|\phi_{x,y}\| \leq \|x\|\|y\|$. By 4.4 there exists a unique complex measure $\mu_{x,y}$ such that

$$\widetilde{\mu}_{x,y} : \mathcal{B}_b(\sigma(S)) \to \mathbb{C},\ f \mapsto \int_K f\, d\mu_{x,y},$$

is an extension of $\phi_{x,y}$ with $\|\widetilde{\mu}_{x,y}\| = \|\phi_{x,y}\| \leq \|x\|\|y\|$. Clearly, $\widetilde{\mu}_{x,y}$ is $\sigma$-continuous.

For $f \in \mathcal{B}_b(\sigma(S))$ we define

$$\beta_f : H \times H \to \mathbb{C},\ (x,y) \mapsto \beta_f(x,y) := \widetilde{\mu}_{x,y}(f).$$

Then

$$|\beta_f(x,y)| = |\widetilde{\mu}_{x,y}(f)| \leq \|f\|_\infty \|x\|\|y\| \quad \text{for all } x,y \in H,\ f \in \mathcal{B}_b(\sigma(S)).$$

For $\alpha \in \mathbb{C}$ and $x,y,z \in H$ we have coincidence of the $\sigma$-continuous functionals $\widetilde{\mu}_{\alpha x+y,z}$, $\alpha\widetilde{\mu}_{x,z}+\widetilde{\mu}_{y,z}$ and $\widetilde{\mu}_{x,\alpha y+z}$, $\overline{\alpha}\widetilde{\mu}_{x,y}+\widetilde{\mu}_{x,z}$, respectively, on $C(\sigma(S))$. By 4.7, each $\beta_f$ is sesquilinear.

By the preceding remark there exists, for each $f \in \mathcal{B}_b(\sigma(S))$, a unique operator $\Psi(f) \in \mathcal{L}(H)$ satisfying

$$\widetilde{\mu}_{x,y}(f) = \beta_f(x,y) = (\Psi(f)x|y) \quad \text{for all } x, y \in H.$$

We have $\|\Psi(f)\| \le \|f\|_\infty$. Taking $\alpha \in \mathbb{C}$ and $f, g \in \mathcal{B}_b(\sigma(S))$, we have

$$\begin{aligned}((\alpha\Psi(f) + \Psi(g))x|y) &= \alpha(\Psi(f)x|y) + (\Psi(g)x|y) = \alpha\widetilde{\mu}_{x,y}(f) + \widetilde{\mu}_{x,y}(g) = \widetilde{\mu}_{x,y}(\alpha f + g) \\ &= (\Psi(\alpha f + g)x|y),\end{aligned}$$

for all $x, y \in H$. The uniqueness assertion in the preceding remark implies $\alpha\Psi(f) + \Psi(g) = \Psi(\alpha f + g)$, i.e. $\Psi : f \mapsto \Psi(f)$ is linear.

By 4.4, $\Psi$ is an extension of $\Phi$, and by 4.3 the map $\Psi$ is $\sigma$-continuous. Since the map $f \mapsto \Psi(\overline{f}) - \Psi(f)^*$ is $\sigma$-continuous on $\mathcal{B}_b(\sigma(S))$ and vanishes on $C(\sigma(S))$, it vanishes on $\mathcal{B}_b(K)$ by 4.7. Hence we have $\Psi(\overline{f}) = \Psi(f)^*$ for all $f \in \mathcal{B}_b(\sigma(S))$.

For $f \in C(\sigma(S))$ we have that $g \mapsto \Psi(fg) - \Psi(f)\Psi(g)$ is $\sigma$-continuous and vanishes on $C(\sigma(S))$. By 4.7 again, we obtain $\Psi(fg) = \Psi(f)\Psi(g)$ for all $g \in \mathcal{B}_b(\sigma(S))$. For fixed $g \in \mathcal{B}_b(\sigma(S))$ we thus know that $f \mapsto \Psi(fg) - \Psi(f)\Psi(g)$ vanishes on $C(\sigma(S))$. Since this map is $\sigma$-continuous, we obtain by 4.7 that $\Psi(fg) = \Psi(f)\Psi(g)$ for all $f, g \in \mathcal{B}_b(\sigma(S))$.

Finally, for any $f \in \mathcal{B}_b(\sigma(S))$, $\Psi(f)\Psi(f)^* = \Psi(f)\Psi(\overline{f}) = \Psi(|f|^2) = \Psi(\overline{f})\Psi(f) = \Psi(f)^*\Psi(f)$, so $\Psi(f)$ is normal. □

As in §3 we now can show that $\Psi$ from Theorem 4.8 has further properties.

**4.9. Corollary:** Let $S \in \mathcal{L}(H)$ be self-adjoint and let $\Psi$ be the functional calculus from 4.8. Let $f \in \mathcal{B}_b(\sigma(S))$.

(i) If $f$ is real-valued then $\Psi(f)$ is self-adjoint.

(ii) If $f \ge 0$ then $\Psi(f) \ge 0$ in the sense of 3.8.

*Proof.* For real-valued $f$ we have

$$\Psi(f)^* = \Psi(\overline{f}) = \Psi(f).$$

If $f \ge 0$ then $g := \sqrt{f} \in \mathcal{B}_b(\sigma(S))$ and, for $x \in H$,

$$(\Psi(f)x|x) = (\Psi(g^2)x|x) = (\Psi(g)\Psi(g)x|x) = (\Psi(g)x|\Psi(g)x) \ge 0,$$

where we have used that $\Psi(g)$ is self-adjoint. □

**Warning:** We know that $\|\Phi(f)\| = \|f\|_\infty$ for $f \in C(\sigma(S))$ and $\|\Psi(g)\| \le \|g\|_\infty$ for all $g \in \mathcal{B}_b(\sigma(S))$, but it may happen that $\|\Psi(g)\| < \|g\|_\infty$ for some $g \in \mathcal{B}_b(\sigma(S))$. For example, for $\lambda \in \sigma(S)$ one has $\Psi(1_{\{\lambda\}}) = 0$ if and only if $\lambda \notin \sigma_p(S)$ ($\to$ later and exercises).

**4.10. Definition:** Let $A$ be a linear operator in $H$, i.e. $A : H \supseteq D(A) \to H$ is linear.

(i) $A$ is called *symmetric* if $(Ax|y) = (x|Ay)$ for all $x, y \in D(A)$.

(ii) If $A$ is densely defined then $A$ is called self-adjoint if $A = A^*$.

Recall that, in the situation of (ii), the adjoint operator $A^*$ of $A$ is given by

$$x \in D(A^*) \text{ and } A^*x = y \iff \forall z \in D(A) : (Az|x) = (z|y).$$

**Remark:** If $A$ is densely defined in $H$ then $A$ is symmetric if and only if $A \subseteq A^*$. Here, $A \subseteq A^*$ means that $x \in D(A)$ implies $x \in D(A^*)$ and $A^*x = Ax$.

In particular, any self-adjoint operator is symmetric. Moreover, any self-adjoint operator is closed. End Lect.17

**Remark:** If $A$ is symmetric then $(Ax|x) \in \mathbb{R}$ for all $x \in D(A)$, since

$$(Ax|x) = (x|Ax) = \overline{(Ax|x)}.$$

**4.11. Lemma:** If $A$ is densely defined and symmetric in $H$, then $A$ is closable and its closure $\overline{A}$ is symmetric.

*Proof.* By the remark above we have $A \subseteq A^*$, and $A^*$ is always closed. Hence also $\overline{A} \subseteq A^* = (\overline{A})^*$ (for the last identity see below). □

**Recall:** In the situation of 4.11, we have by 1.16 and reflexivity of $H$:

$$A \text{ is closable} \iff A^* \text{ is densely defined.}$$

Actually, the proof shows that, in this case, $\overline{A} = (A^*)^*$. Applied to the closed operator $A^*$ in place of $A$, we get

$$(\overline{A})^* = ((A^*)^*)^* = A^*.$$

**4.12. Lemma:** Let $A$ be symmetric and closed. Then, for any $z \in \mathbb{C} \setminus \mathbb{R}$, the operator $z - A$ is injective and $R(z - A)$ is closed. If $R(z - A) = H$ then $\|R(z, A)\| \leq 1/|\mathrm{Im}\, z|$.

*Proof.* Since $A$ is symmetric, we have for $x \in D(A)$ and $\xi + i\eta \in \mathbb{C}$ with $\eta \neq 0$ by the preceding remark

$$\begin{aligned}
\|(\xi + i\eta - A)x\|^2 &= (\xi^2 + \eta^2)\|x\|^2 - 2\mathrm{Re}\,((\xi + i\eta)x|Ax) + \|Ax\|^2 \\
&= (\xi^2 + \eta^2)\|x\|^2 - 2\xi(x|Ax) + \|Ax\|^2 \\
&= \eta^2\|x\|^2 + \|(\xi - A)x\|^2 \geq \eta^2\|x\|^2.
\end{aligned}$$

Since $\xi + i\eta - A$ is closed, the assertion follows. □

**4.13. Proposition:** If $A$ is a self-adjoint operator in $H$ then $\sigma(A) \subset \mathbb{R}$, and $\|R(z,A)\| \le |\mathrm{Im}\, z|^{-1}$ for any $z \in \mathbb{C} \setminus \mathbb{R}$.

*Proof.* We have $(z-A)^* = \overline{z} - A$ for any $z \in \mathbb{C}$. For $z \in \mathbb{C} \setminus \mathbb{R}$ we have by 4.12 that $z - A$, $\overline{z} - A$ are injective and that $R(z-A)$ is closed. But

$$R(z-A)^\perp = \{y \in H : \forall x \in D(A) : ((z-A)x|y) = 0\} = N((z-A)^*) = N(\overline{z}-A) = \{0\},$$

i.e. $R(z-A) = H$, and we obtain $z \in \rho(A)$. The resolvent estimate is from 4.12. □

**4.14. Proposition:** For any densely defined operator in $A$ the following are equivalent:

(i) $A$ is closable and $\overline{A}$ is self-adjoint.

(ii) $A$ is symmetric and $\sigma(\overline{A}) \subset \mathbb{R}$.

(iii) $A$ is symmetric and there exists $z \in \mathbb{C} \setminus \mathbb{R}$ such that $z - A^*$ and $\overline{z} - A^*$ are injective.

*Proof.* (i) $\Longrightarrow$ (ii): If $\overline{A}$ is self-adjoint, then $\overline{A}$ and hence also $A$ are symmetric. $\sigma(\overline{A}) \subset \mathbb{R}$ holds by 4.13.

(ii) $\Longrightarrow$ (iii): By 4.11, $\overline{A}$ is symmetric. Moreover, $(\overline{A})^* = A^*$. For $z \in \mathbb{C} \setminus \mathbb{R}$ we thus have $z, \overline{z} \in \sigma(A^*)$, and $z - A^*$, $\overline{z} - A^*$ are injective.

(iii) $\Longrightarrow$ (i): By 4.11, $A$ is closable, and we have $\overline{A} \subset A^* = (\overline{A})^*$. By 4.12 (and its proof) the assumption implies $z, \overline{z} \in \rho(\overline{A})$, and this in turn implies $\overline{z}, z \in \rho(A^*)$. Hence we have $\rho(\overline{A}) \cap \rho(A^*) \neq \emptyset$ which by $\overline{A} \subset A^*$ is only possible if $A = A^*$. □

**Remark:** Here we have used the following: If $B, C$ are clsoed operators in a Banach space $X$ such that $B \subset C$ and there exists $\lambda \in \rho(B) \cap \rho(C)$, then $B = C$.

For the proof observe that $\lambda - C : D(C) \to X$ is bijective so that $\lambda - B : D(B) \to X$ cannot be surjective if $D(B) \overset{\subset}{\neq} D(C)$.

**4.15. Definition:** A densely defined symmetric operator $A$ in $H$ is called *essentially self-adjoint* if its closure $\overline{A}$ is self-adjoint in $H$.

**Remark:** By 4.14, a densely defined and symmetric operator $A$ is essentially self-adjoint if and only if the operators $i - A^*$ and $-i - A^*$ are both injective. We will get back to that later on.

We aim at an extension of the functional calculus from Theorem 4.8 to unbounded self-adjoint operators $A$. This will be done by an orthogonal decompopsition of $H$ with respect to an auxiliary bounded self-adjoint operator $B$ associated to $A$.

**4.16. Lemma:** Let $A$ be self-ajoint in $H$ and set

$$B := \frac{1}{2i}(R(i,A)^* - R(i,A)), \quad C := \frac{-1}{2}(R(i,A) + R(i,A)^*).$$

Then $B, C \in \mathcal{L}(H)$ are self-adjoint, $BA \subset AB = C$, $B$ is injective and $0 \le B \le I$ (in the sense of 3.8).

*Proof.* Self-adjointness of $B$ and $C$ is clear. By $R(i,A)^* = R(-i,A)$ we obtain easily $BA \subset AB$. Moreover, since $AR(\lambda, A) = \lambda R(\lambda, A) - I$, we have

$$AB = \frac{1}{2i}(AR(-i,A) - AR(i,A)) = \frac{1}{2i}(-iR(-i,A) - iR(i,A)) = C.$$

By 4.12, $\|R(\pm, A)\| \le 1$, and we obtain $B \le I$. For $x \in H$ and $y = R(i,A)x$ we have

$$(Bx|x) = \frac{1}{2i}((x|R(i,A)x) - (R(i,A)x|x)) = \mathrm{Im}\,(x|R(i,A)x) = \mathrm{Im}\,((i-A)y|y) = (y|y) \ge 0.$$

Hence $B \ge 0$. On the other hand, $Bx = 0$ implies $(i-A)x = y = 0$ and $x = 0$ by injectivity of $i - A$. Hence $B$ is injective. □

**Remark:** By the resolvent equation 1.5(a) we have

$$B = \frac{1}{2i}(R(-i,A) - R(i,A)) = \frac{1}{2i}((2i)R(i,A)R(-i,A)) = R(i,A)R(i,A)^*,$$

and the properties of $B$ also follow from this representation.

We shall use the functional calculus $\Psi$ from 4.8 for the operator $B$ and write $f(B) := \Psi(f)$ for $f \in \mathcal{B}_b(\sigma(B))$. If $f$ is defined on a superset of $\sigma(B)$ then we write $f(B) := (f|_{\sigma(B)})(B)$. Observe that $\sigma(B) \subseteq [0,1]$ by 4.16 and 3.3 and that $0 \notin \sigma_p(B)$ by 4.16.

**4.17. Proposition:** Let $A$ be a self-adjoint operator in $H$, and let $B$, $C$ be as in 4.16. For any $n \in \mathbb{N}$, define functions $\theta_n, s_n : \mathbb{R} \to \mathbb{R}$ by $\theta_n := 1_{(\frac{1}{n+1}, \frac{1}{n}]}$ and $s_n(t) := \frac{1}{t}\theta_n(t)$, and let $P_n := \theta_n(B)$. Then:

(a) For each $n \in \mathbb{N}$, $P_n$ is an orthogonal projection in $H$ and

$$P_n A \subset AP_n = s_n(B)C \in \mathcal{L}(H).$$

(b) With $H_n := R(P_n)$ we have $H_n \subseteq D(A)$, $A(H_n) \subseteq H_n$ and $H_n \perp H_k$ for all $n, k \in \mathbb{N}$ with $n \ne k$. Moreover, $A_n := A|_{H_n} \in \mathcal{L}(H_n)$ is self-adjoint in the Hilbert space $H_n$.

(c) We have

$$x = \sum_{n\in\mathbb{N}} P_n x \quad \text{for all } x \in H \text{ (convergence in } H)$$

and

$$\begin{aligned} D(A) &= \{x \in H : \sum_{n\in\mathbb{N}} \|AP_n x\|^2 < \infty\} \\ Ax &= \sum_{n\in\mathbb{N}} AP_n x \quad \text{for all } x \in D(A) \text{ (convergence in } H). \end{aligned}$$

End Lect.18

(d) For each $n \in \mathbb{N}$ we have

$$\sigma(A_n) \subseteq \sigma(A) \cap \left([-\sqrt{n}, -\sqrt{n-1}] \cup [\sqrt{n-1}, \sqrt{n}]\right).$$

*Proof.* (a) By $ts_n(t) = \theta_n(t)$, $t \in \mathbb{R}$, we get $Bs_n(B) = \theta_n(B) = P_n$. Hence, by 4.16,

$$AP_n = ABs_n(B) = Cs_n(B) \overset{(*)}{=} s_n(B)C \in \mathcal{L}(H)$$

(we say more on the equality $(*)$ below in the proof of (d)). Moreover, again by 4.16,

$$P_n A = Bs_n(B)A = s_n(B)BA \subseteq s_n(B)AB = s_n(B)C = AP_n.$$

By 4.8, we have $P_n^2 = P_n$ and $P_n^* = P_n$.

(b) Each $H_n$ is a closed subspace of $H$ and thus itself a Hilbert space. By (a), $AP_n \in \mathcal{L}(H)$, so $H_n \subset D(A)$ and, for $x \in H_n$:

$$P_n Ax = AP_n x = Ax, \quad \text{i.e. } Ax \in H_n.$$

Hence $A(H_n) \subseteq H_n$. For $x \in H_n$, $y \in H_k$ and $n \neq k$ we have by 4.8:

$$(x|y) = (P_n x|P_k y) = (\underbrace{P_k P_n}_{=0} x|y) = 0,$$

and thus $H_n \perp H_k$. Finally, $A_n \in \mathcal{L}(H_n)$ is self-adjoint in $H_n$, since it inherits symmetry from $A$.

(c) We let $P_0 := 1_{\{0\}}(B)$ and $H_0 := P_0(H) = R(P_0)$. $P_0$ is an orthogonal projection and $BP_0 = (0 \cdot 1_{\{0\}})(B) = 0$ by 4.8, so $H_0 \subseteq N(B) = \{0\}$ by 4.16. Hence $P_0 = 0$.

Using 4.8, we thus obtain

$$I_H = 1_{[0,1]}(B) = 1_{(0,1]}(B) + P_0 = 1_{(0,1]}(B).$$

We have

$$\sum_{n=1}^{m} \theta_n = 1_{(\frac{1}{m+1},1]} \to 1_{(0,1]} \quad \text{bounded pointwise as } m \to \infty.$$

By 4.8 we obtain, for any $x \in H$,

$$\begin{aligned} \|\sum_{n=1}^{m} P_n x - x\|^2 &= \|1_{(0,\frac{1}{m+1}]}(B)x\|^2 = (1_{(0,\frac{1}{m+1}]}(B)x|1_{(0,\frac{1}{m+1}]}(B)x) \\ &= (1_{(0,\frac{1}{m+1}]}(B)x|x) \to 0 \ (m \to \infty), \end{aligned}$$

since $1_{(0,\frac{1}{m+1}]} \to 0$ bounded pointwise as $m \to \infty$. Moreover, by orthogonality of the summands we have $\|x\|^2 = \sum_{n\in\mathbb{N}} \|P_n x\|^2$. For $x \in D(A)$ we thus have

$$Ax = \sum_{n\in\mathbb{N}} P_n Ax = \sum_{n\in\mathbb{N}} AP_n x \quad \text{and} \quad \sum_{n\in\mathbb{N}} \|AP_n x\|^2 = \|Ax\|^2 < \infty.$$

Now let $x \in H$ such that $\sum_{n\in\mathbb{N}} \|AP_n x\|^2 < \infty$. Then $\sum_{n\in\mathbb{N}} AP_n x$ converges in $H$, and by (b) we have that

$$x_m := \sum_{n=1}^{m} P_n x \in D(A), \quad Ax_m = \sum_{n=1}^{m} AP_n$$

for each $m \in \mathbb{N}$, and $x_m \to x$ in $H$, $Ax_m \to y := \sum_{n\in\mathbb{N}} AP_n x$ in $H$ as $m \to \infty$. By closedness of $A$ we obtain $x \in D(A)$ and $Ax = \sum_{n\in\mathbb{N}} AP_n x$.

(d) Let $n \in \mathbb{N}$. Then $H_n$ is invariant under $B$, since $BP_n = B\theta_n(B) = \theta_n(B)B$ by 4.8. We set $B_n := B|_{H_n}$. Since $B$ commutes with resolvents of $A$, the space $H_n$ is also invariant under resolvents of $A$.[3] Hence, for any $\lambda \in \rho(A)$, we have $\lambda \in \rho(A_n)$ and

$$R(\lambda, A_n) = R(\lambda, A)|_{H_n}$$

(in particular, we have $\sigma(A_n) \subset \sigma(A)$): Indeed, for $\lambda \in \rho(A)$ the operator $\lambda - A_n$, which is a restriction of $\lambda - A$, is clearly injective. Moreover, for $y \in H_n$ and $x := R(\lambda, A)$ we have $x \in H_n$ and $(\lambda - A_n)x = (\lambda - A)R(\lambda, A)y = y$, and $\lambda - A_n : H_n \to H_n$ is also surjective. Hence we have, for any $n \in \mathbb{N}$,

$$B_n = R(i, A_n)R(-i, A_n), \quad B_n^{-1} = (i - A_n)(-i - A_n) = 1 + A_n^2.$$

Since $\sigma(B_n) \subset [\frac{1}{n+1}, \frac{1}{n}]$ by 4.8 we thus have $\sigma((1 + A_n^2)^{-1}) \subset [\frac{1}{n+1}, \frac{1}{n}]$. By spectral mapping this implies $\sigma(A_n^2) \subset [n-1, n]$ and finally

$$\sigma(A_n) \subset [-\sqrt{n}, -\sqrt{n-1}] \cup [\sqrt{n-1}, \sqrt{n}],$$

so (d) is proved. □

[3]The following argument also proves (∗) in the proof of (a) above (and is a special case of Exercise 43): Let $T \in \mathcal{L}(H)$ such that $TB = BT$. Then $Tp(B) = p(B)T$ for all polynomials $p$. By Weierstraß, one can approximate continuous functions on $\sigma(B)$ by polynomials, and this gives $Tf(B) = f(B)T$ for all $f \in C(\sigma(B))$. Finally, application of 4.6 shows $Tf(B) = f(B)T$ for all $f \in \mathcal{B}_b(\sigma(B))$, where 4.6(2) is shown with the help of $\sigma$-continuity of the map $\Psi$ from 4.8.

Here and in (∗) this is applied to $f = \theta_n$ and $T = R(\lambda, A)$, $T = C$, respectively.

**Remark:** But by 4.17 and Exercise 48 we have

$$\sigma(A) = \overline{\bigcup_{n\in\mathbb{N}} \sigma(A_n)} = \bigcup_{n\in\mathbb{N}} \sigma(A_n)$$

where the last equality is due to the fact that by (d), for any Cauchy sequence in $\bigcup_n \sigma(A_n)$, there exists $n_0 \in \mathbb{N}$ and a convergent subsequence contained in $\sigma(A_{n_0})$. From this we obtain

$$\sigma(A) \cap (-\sqrt{n}, -\sqrt{n-1}) \cup (\sqrt{n-1}, \sqrt{n}) = \sigma(A_n) \cap (-\sqrt{n}, -\sqrt{n-1}) \cup (\sqrt{n-1}, \sqrt{n})$$

for each $n \in \mathbb{N}$. End Lect.19

In 4.17, we have established an orthogonal decomposition of the space $H$ into closed subspaces $H_n$, $n \in \mathbb{N}$, such that $H_n \subset D(A)$, $A_n := A|_{H_n} \in \mathcal{L}(H_n)$ and $A_n$ is self-adjoint in $H_n$ and corresponds to a bounded part of the spectrum of $A$. We now apply 4.8 to each of the parts $A_n$ and thus define a functional calculus for $A$ for bounded Borel measurable functions on the spectrum $\sigma(A)$.

**4.18. Theorem:** Let $A$ be a self-adjoint operator in $H$ and, for $n \in \mathbb{N}$, let $P_n$, $H_n$, $A_n$ be as in 4.17. For $f \in \mathcal{B}_b(\sigma(A))$ we define

$$f(A)x := \sum_{n\in\mathbb{N}} f(A_n)P_n x, \quad x \in H.$$

Then $f(A) \in \mathcal{L}(H)$ is well-defined, $\|f(A)\| \leq \|f\|_\infty$, and the map $\Psi : \mathcal{B}(\sigma(A)) \to \mathcal{L}(H)$ is an algebra homomorphism with $\Psi(1_{\sigma(A)}) = I$, $\Psi((z - (\cdot))^{-1}) = R(z, A)$ for $z \in \mathbb{C} \setminus \mathbb{R}$.

For $f \in \mathcal{B}_b(\sigma(A))$ one has $\Psi(f)^* = \Psi(\overline{f})$. In particular, each $f(A)$ is normal, $f(A)$ is self-adjoint for real-valued $f$, and $f(A) \geq 0$ if $f \geq 0$.

If $f_m \to f$ bounded pointwise as $m \to \infty$ then $f_m(A)x \to f(A)x$ for any $x \in H$.

*Proof.* Since $f(A_n) \in \mathcal{L}(H_n)$ with $\|f(A_n)\| \leq \|f\|_{\infty,\sigma(A_n)} \leq \|f\|_{\infty,\sigma(A)}$ we have

$$\sum_{n\in\mathbb{N}} \|f(A_n)P_n x\|^2 \leq \|f\|^2_{\infty,\sigma(A)} \sum_{n\in\mathbb{N}} \|P_n x\|^2 = \|f\|^2_{\infty,\sigma(A)} \|x\|^2, \quad x \in H.$$

Hence $\sum_n f(A_n)P_n x$ converges and

$$\|f(A)x\| = \Big(\sum_{n\in\mathbb{N}} \|f(A_n)P_n x\|^2\Big)^{1/2} \leq \|f\|_{\infty,\sigma(A)} \|x\|.$$

This implies that $\Psi : \mathcal{B}_b(\sigma(A)) \to \mathcal{L}(H)$ is linear and continuous with norm $\leq 1$. $\Psi(1_{\sigma(A)}) = I_H$ follows from 4.17(c). For $x \in H$ and $z \in \mathbb{C} \setminus \mathbb{R}$ we have (cp. 4.17(d) and its proof)

$$R(z, A)x = \sum_{n\in\mathbb{N}} R(z, A)P_n x = \sum_{n\in\mathbb{N}} R(z, A_n)P_n x.$$

By 3.7 we have $R(z, A_n) = (z - (\cdot))^{-1}(A_n)$ for any $n \in \mathbb{N}$, so $R(z, A) = \Psi((z - (\cdot))^{-1})$ is proved.

Notice that, for each $m \in \mathbb{N}$,

$$P_m f(A)x = \sum_{n\in\mathbb{N}} P_m f(A_n) P_n x = f(A_m) P_m x.$$

Hence, for $f, g \in \mathcal{B}_b(\sigma(A))$ and $x \in H$,

$$(gf)(A) = \sum_{n\in\mathbb{N}} (gf)(A_n) P_n x = \sum_{n\in\mathbb{N}} g(A_n) f(A_n) P_n x = \sum_{n\in\mathbb{N}} g(A_n) P_n f(A)x = g(A)(f(A)x),$$

and $\Psi$ is multiplicative. By orthogonality of the summands we have, for $f \in \mathcal{B}_b(\sigma(A))$ and $x, y \in H$,

$$\begin{aligned}(f(A)x|y) &= \sum_{n\in\mathbb{N}} (P_n f(A)x|P_n y) = \sum_{n\in\mathbb{N}} (f(A_n)P_n x|P_n y)_{H_n} \\ &= \sum_{n\in\mathbb{N}} (P_n x|f(A_n)^* P_n)_{H_n} = \sum_{n\in\mathbb{N}} (P_n x|\overline{f}(A_n)P_n(y))_{H_n} \\ &= (x|\overline{f}(A)y),\end{aligned}$$

which implies that $f(A)^* = \overline{f}(A)$.

The assertions on normality, self-adjointness and positivity follow as before. Let $(f_m)_{m\in\mathbb{N}}$ be a sequence in $\mathcal{B}_b(\sigma(A))$ that converges bounded pointwise to a function $f$. By linearity of $\Psi$ we may assume that $f = 0$. By Exercise 41, we have for the functional calculus from 4.8 that $f_m(A_n)x \to 0$ as $m \to \infty$ for all $n \in \mathbb{N}$ and $x \in H_n$.[4] Hence, for fixed $x \in H$, we obtain

$$f_m(A_n)P_n x \to 0 \quad \text{for each } n \in \mathbb{N}.$$

Letting $M := \sup_{m\in\mathbb{N}} \|f_m\|_\infty$ we have, for any $N \in \mathbb{N}$,

$$\|f_m(A)x\|^2 = \sum_{n\in\mathbb{N}} \|f_m(A_n)P_n x\|^2 \leq \sum_{n=1}^{N} \|f_m(A_n)P_n x\|^2 + M \sum_{n>N} \|P_n x\|^2.$$

For a given $\varepsilon > 0$ we now choose $N$ such that $M \sum_{n>N} \|P_n x\|^2 \leq \varepsilon/2$, and then $m_0$ such that $\sum_{n=1}^{N} \|f_m(A)P_n x\|^2 \leq \varepsilon/2$ for all $m \geq m_0$. Thus $f_m(A)x \to 0$ in $H$ is proved. □

**Remark:** One can show that $\Psi : \mathcal{B}_b(\sigma(A)) \to \mathcal{L}(H)$ with the stated properties is unique. Moreover, one can actually see that $P_n = (\theta_n \circ \psi)(A)$ where $\psi(z) = \frac{1}{1+z^2}$, so $P_n = \eta_n(A)$ where $\eta_n = 1_{(-\sqrt{n}, -\sqrt{n-1}] \cup [\sqrt{n-1}, \sqrt{n})}$.

[4] Simply write $\|f_m(A_n)x\|^2 = (f_m(A_n)^* f_m(A_n)x|x) = (|f_m|^2(A_n)x|x) \to 0$ since $|f_m|^2 \to 0$ bounded pointwise.

**4.19. Discussion:** Suppose that we are in the situation of 4.19. Let $\Psi_n : \mathcal{B}_b(\mathbb{R}) \to \mathcal{L}(H_n)$ denote the functional calculus from 4.8 for the operator $A_n$. Then

$$f(A)x = \sum_{n\in\mathbb{N}} \Psi_n(f) P_n x, \quad x \in H, f \in \mathcal{B}_b(\mathbb{R}),$$

and, recalling the proof of 4.8 and using the notation $\widetilde{\mu}_{n,x,y}$ for the complex measures associated to $A_n$, we have

$$(f(A)x|y) = \sum_{n\in\mathbb{N}} (\Psi_n(f) P_n x | P_n y) = \sum_{n\in\mathbb{N}} \widetilde{\mu}_{n,P_n x, P_n y}(f),$$

where $|\widetilde{\mu}_{n,P_n x,P_n y}(f)| \leq \|f\|_\infty \|P_n x\| \|P_n y\|$. Define $\widetilde{\mu}_{x,y}$ by

$$\widetilde{\mu}_{x,y}(f) := \sum_{n\in\mathbb{N}} \widetilde{\mu}_{n,P_n x,P_n y}(f), \quad f \in \mathcal{B}_b(\mathbb{R}).$$

Then $\widetilde{\mu}_{x,y}$ is a complex measure on the Borel subsets of $\mathbb{R}$, $\widetilde{\mu}_{x,y}$ vanishes outside $\sigma(A)$, $\|\widetilde{\mu}_{x,y}\| \leq \|x\|\|y\|$, and $\widetilde{\mu}_{x,x}$ is a positive measure for all $x \in H$:

$$\widetilde{\mu}_{x,x}(f) = (f(A)x|x) \geq 0 \text{ if } f \geq 0 \text{ and } \widetilde{\mu}_{x,x}(1_\mathbb{R}) = (1_\mathbb{R}(A)x|x) = (x|x) = \|x\|^2.$$

In particular, $\widetilde{\mu}_{x,x}$ is a probability measure for $x \in H$ with $\|x\| = 1$.

Moreover, if we let $E(M) := 1_M(A)$ for $M \in \mathcal{B}$ (Borel subsets of $\mathbb{R}$), then

$$\widetilde{\mu}_{x,y}(M) = (1_M(A)x|y) = (E(M)x|y),$$

so that, for any $f \in \mathcal{B}_b(\mathbb{R})$:

$$(f(A)x|y) = \widetilde{\mu}_{x,y}(f) = \int f(\lambda)\, d\widetilde{\mu}_{x,y}(\lambda) =: \int f(\lambda)\, d(E(\lambda)x|y).$$

This is written as

$$f(A) = \int f(\lambda)\, dE(\lambda),$$

and $E : \mathcal{B} \to \mathcal{L}(H)$ is called the *spectral measure* of $A$ (observe that $E(M)$ is an orthogonal projection for any $M \in \mathcal{B}$). End Lect.20

The spectral measure $E(\cdot)$ is uniquely determined by the function

$$F : \mathbb{R} \to \mathcal{L}(H), \quad b \mapsto F(b) := E((-\infty, b]) = 1_{(-\infty,b]}(A),$$

and $F$ is called the *resolution of the identity* associated to $A$, which has the following properties:

- each $F(b)$ is an orthogonal projection,
- if $b \geq c$, then $F(b) \geq F(c)$,

- for any $x \in H$, $b \mapsto F(b)x$ is right continuous with left hand limits at every point, and $\lim_{b\to-\infty} F(b)x = 0$, $\lim_{b\to\infty} F(b)x = x$.
- $F(b) - F(b-) = E(\{b\}) = 1_{\{b\}}(A) = N(b - A)$ for any $b \in \mathbb{R}$.

**Remark:** For $f \in \mathcal{B}_b(\mathbb{R})$ and $x \in H$ one has

$$\|f(A)x\|^2 = (f(A)x|f(A)x) = (|f|^2(A)x|x) = \int |f|^2 \, d(E(\cdot)x|x).$$

Actually, one can define $f(A)$ as an *unbounded* operator in $H$ for any Borel measurable function $f : \sigma(A) \to \mathbb{C}$ with

$$D(f(A)) = \{x \in H : \int |f|^2 \, d(E(\cdot)x|x) < \infty\},$$

see, e.g. Rudin, Functional Analysis.

We come back to the question of self-adjointness for symmetric operators and state the following in the situation of 4.12.

**4.20. Lemma:** Let $A$ be symmetric and closed. If $\lambda, \mu \in \mathbb{C} \setminus \mathbb{R}$ such that $\operatorname{Im}\lambda \cdot \operatorname{Im}\mu > 0$ then

$$\operatorname{codim} R(\lambda - A) = \operatorname{codim} R(\mu - A).$$

*Proof.* Let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. We show the assertion for $\mu \in \mathbb{C} \setminus \mathbb{R}$ with $|\lambda - \mu| < |\operatorname{Im}\lambda|$. The lemma follows by taking $\lambda = \pm in$ and $n \to \infty$. We let $X := [D(A)]$. By 4.12, $\lambda - A$ is injective and $Y := R(\lambda - A)$ is closed. Let $Z := X \times Y^\perp$ and consider

$$\widehat{z - A} : X \times Y^\perp \to H, \quad (x, y) \mapsto (z - A)x + y, \quad z \in \mathbb{C} \setminus \mathbb{R}.$$

For $z = \lambda$, this operator is bijective with inverse $Rh := ((\lambda - A)^{-1}Ph, (I - P)h)$, where $P$ denotes the orthogonal projection from $H$ onto $Y$. If $|\mu - \lambda| < |\operatorname{Im}\lambda|$, then

$$\widehat{\mu - A} = \widehat{\lambda - A} + (\mu - \lambda)\pi_1 = (I_H + (\mu - \lambda)\pi_1 R)\widehat{\lambda - A},$$

where $\pi_1 : X \times Y^\perp \to H$, $(x, y) \mapsto x$, and $\|\pi_1 R\| = \|(\lambda - A)^{-1}P\|_{\mathcal{L}(H)} \le 1/|\operatorname{Im}\lambda|$ by 4.12. Hence

$$(\widehat{\mu - A})^{-1} = R \sum_{k=0}^{\infty} (-\pi_1 R)^k (\mu - \lambda)^k$$

and $\widehat{\mu - A}$ is bijective $X \times Y^\perp \to H$. In particular, $R(\mu - A) + Y^\perp = H$. Moreover, if $x \in D(A)$ and $(\mu - A)x = -y \in Y^\perp$ then $\widehat{\mu - A}(x, y) = 0$, so $x = 0 = y$, which implies $R(\mu - A) \cap Y^\perp = \{0\}$. This means that $Y^\perp$ is a complement both for $R(\lambda - A)$ and $R(\mu - A)$, so $\operatorname{codim} R(\lambda - A) = \operatorname{codim} R(\mu - A)$. □

**Remark:** The proof is similar to what we have done for Fredholm operators. The numbers $n_\pm(A) := \operatorname{codim} R(\pm i, A)$ are called *defect indices*, and it can be shown that $A$ has a self-adjoint extension if and only if $n_+(A) = n_-(A)$.

If $A$ is in addition positive, i.e. $(Ax|x) \geq 0$ for all $x \in D(A)$, then one can show in a similar way that $n_+(A) = n_-(A) = \operatorname{codim}(\mu + A)$ for any $\mu > 0$, and $A$ has self-adjoint extensions.

# 5 Weak derivatives and Sobolev spaces

Let $\Omega \subseteq \mathbb{R}^d$ be open. We write $f \in L^1_{\text{loc}}(\Omega)$ if $f|_K \in L^1(K)$ for every $K \subset \Omega$ or, equivalently, if $f|_B \in L^1(B)$ for verey closed ball $B \subset \Omega$. The *support* $\operatorname{supp} f$ of such an $f$ is defined via

$$\Omega \setminus \operatorname{supp} f := \{x \in \Omega : \exists \varepsilon > 0 : f = 0 \text{ a.e. on } B(x, \varepsilon)\}.$$

For continuous $\varphi : \Omega \to \mathbb{C}$ this coincides with the usual definition

$$\operatorname{supp} \varphi := \overline{\{x \in \Omega : \varphi(x) \neq 0\}},$$

where the closure is taken in $\Omega$.

**5.1. Mollifiers:** (a) For $\varphi \in C_c^\infty(\mathbb{R})$ and $f \in L^1_{\text{loc}}(\mathbb{R})$ define

$$\varphi * f(x) := \int_{\mathbb{R}} \varphi(x - y) f(y)\, dy, \quad x \in \mathbb{R}^d.$$

Then $\varphi * f \in C^\infty(\mathbb{R})$ and

$$\partial_j(\varphi * f) = (\partial_j \varphi) * f \quad \text{for any } j \in \{1, \ldots, d\}, \text{ where } \partial_j := \tfrac{\partial}{\partial x_j}.$$

(b) There are $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\varphi \geq 0$, $\operatorname{supp} \varphi \subseteq \overline{B}(0, 1)$ and $\int \varphi\, dx = 1$.

(c) For $p \in [1, \infty)$ and $f \in L^p(\mathbb{R}^d)$ the map $\mathbb{R}^d \to L^p(\mathbb{R}^d)$, $y \mapsto \tau_y f$ is continuous.

(d) If $\varphi$ is as in (b) and $\varphi_n := n^d \varphi(n \cdot)$, $n \in \mathbb{N}$, then for $p \in [1, \infty)$ and $f \in L^p(\mathbb{R}^d)$ we have $\varphi_n * f \in L^p(\mathbb{R}^d)$, $\|\varphi_n * f\|_p \leq \|f\|_p$ and

$$\|\varphi_n * f - f\|_p \to 0 \quad (n \to \infty).$$

Moreover, $\operatorname{supp} (\varphi_n * f) \subseteq \operatorname{supp} f + \overline{B}(0, 1/n)$. End Lect.21

A sequence $(\varphi_n)$ as in (c) is called a *mollifier*, and the functions $\varphi_n * f$ are *mollified versions* of $f$.

**Remark:** Recall from Functional Analysis (Winter 2010/11) that $f \in L^1(\mathbb{R}^d)$ implies $\varphi * f \in L^1(\mathbb{R}^d)$ and $\|\varphi * f\|_1 \leq \|\varphi\|_1 \|f\|_1$.

*Proof.* (a) Fix $x \in \mathbb{R}^d$. Since $\varphi$ has compact support, we see from the definition that we can assume $f \in L^1(\mathbb{R}^d)$ without loss of generality. For small $h$ we then have

$$\begin{aligned} &\frac{1}{h}((\varphi * f)(x + he_j) - \varphi * f(x)) - (\partial_j \varphi) * f(x) \\ = \;&\int_{\mathbb{R}^d} \Big( \frac{1}{h}(\varphi(x + he_j - y) - \varphi(x - y)) - (\partial_j \varphi)(x - y) \Big) f(y)\, dy, \end{aligned}$$

and, by the mean value theorem, the modulus of the integrand is bounded by

$$|\partial_j\varphi(x+\xi(h,y)e_j-y)-\partial_j\varphi(x-y)||f(y)| \le |h|\|\partial_j^2\varphi\|_\infty|f(y)|.$$

The assertion follows by dominated convergence. (In fact, it would be sufficient that $\partial_j\varphi$ is uniformly continuous.)

(b) Define $\rho : \mathbb{R} \to [0,1]$ by $\rho(t) := e^{-1/t}$ if $t > 0$ and $\rho(t) := 0$ if $t \le 0$. Then $\rho \in C^\infty(\mathbb{R})$ and $\operatorname{supp}\rho = [0,\infty)$. Then let $\varphi(x) := c\rho(1-|x|^2)$ where $c > 0$ is suitably chosen.

(c) It suffices to show continuity at $y_0 = 0$ on the dense subset of step functions. By linearity it suffices to consider $f = 1_Q$ where $Q$ is a cube. There the assertion is easy:

$$\|\tau_h 1_Q - 1_Q\|_p \le |(Q+B(0,h)) \setminus Q|^{1/p} \le C_Q|h|^{1/p} \to 0 \quad (h \to 0).$$

(d) First notice that $\operatorname{supp}\varphi_n \subseteq \overline{B}(0,1/n)$ and that $\|\varphi_n\|_1 = \int \varphi_n\,dx = 1$. In particular,

$$\varphi_n * f(x) = \int_{\mathbb{R}^d} \varphi_n(y)f(x-y)\,dy = \int_{B(0,1/n)} \varphi_n(y)f(x-y)\,dy = 0,$$

if $(x-\overline{B}(0,1/n)) \cap \operatorname{supp} f) = \emptyset$. This proves $\operatorname{supp}(\varphi_n * f) \subseteq \operatorname{supp} f + \overline{B}(0,1)$. Now we write $\tau_y f := f(\cdot - y)$ and

$$\varphi_n * f = \int_{\mathbb{R}^d} \varphi_n(y)\tau_y f\,dy.$$

Observe that the function $y \mapsto \varphi_n(y)\tau_y f$ is a continuous function $\mathbb{R}^d \to L^p(\mathbb{R}^d)$ with compact support. The integral is a *Riemann integral for compactly supported continuous functions $g$ with values in a Banach space $X$*:

> If $g : \mathbb{R}^d \to X$ is such a function, then $g$ is uniformly continuous, and the Riemannian sums
>
> $$\sum_{k\in\mathbb{Z}^d} g(\xi_k^{(n)})|n^{-1}(k+[0,1]^d)|,$$
>
> where $\xi_k^{(n)} \in n^{-1}(k+[0,1]^d)$ for each $k$, converge to an element of $X$, which we call $\int_{\mathbb{R}^d} g(y)\,dy$ (observe that, due to compactness of the support, the sum is actually finite for each $n$). Clearly, $g \mapsto \int_{\mathbb{R}^d} g(y)\,dy$ is linear, and we have
>
> $$\| \int_{\mathbb{R}^d} g(y)\,dy\|_X \le \int_{\mathbb{R}^d} \|g(y)\|_X\,dy,$$
>
> as a "triangle inequality" for integrals (sometimes called *Minkowski's inequality*).

By this inequality we obtain

$$\|\varphi_n * f\|_p \le \int_{\mathbb{R}^d} |\varphi_n(y)| \underbrace{\|\tau_y f\|_p}_{=\|f\|_p} dy = \|f\|_p,$$

and, for any $\delta > 0$,

$$\begin{aligned}
\|\varphi_n * f - f\|_p &= \|\int_{\mathbb{R}^d} \varphi_n(y)(\tau_y f - f)\, dy\|_p \\
&= \int_{\mathbb{R}^d} \varphi_n(y)\|\tau_y f - f\|_p \, dy \\
&\le \sup_{|y|\le\delta} \|\tau_y f - f\|_p \underbrace{\int_{|y|\le\delta} \varphi_n(y)\, dy}_{\le 1} + 2\|f\|_p \int_{|y|\ge\delta} \varphi_n(y)\, dy.
\end{aligned}$$

For a given $\varepsilon > 0$ we can choose $\delta > 0$ by (c) such that the sup is $\le \varepsilon/2$. Since substitution $y = \eta/n$ gives

$$\int_{|y|\ge\delta} \varphi_n(y)\, dy = \int_{|\eta|\ge n\delta} \varphi(y)\, dy,$$

we can choose $n_0$ such that $2\|f\|_p \int_{|y|\ge\delta} \varphi_n(y)\, dy \le \varepsilon/2$ for $n \ge n_0$. □

**Corollary:** $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for $1 \le p < \infty$: The space of step functions with compact support in $\Omega$ is dense in $L^p(\Omega)$, and by (d) any such function can be approximated by a sequence in $C_c^\infty(\Omega)$ (one has to apply (d) to the extension of the function which is $= 0$ outside $\Omega$).

**5.2. Weak derivatives:** Let $\Omega \subset \mathbb{R}^d$ be open, $f \in L^1_{\text{loc}}(\Omega)$, and $j \in \{1, \ldots, d\}$. $f$ is said to have *a weak derivative in* $\Omega$ *with respect to* $x_j$ if there exists $g \in L^1_{\text{loc}}(\Omega)$ such that

$$-\int_\Omega f \partial_j \varphi \, dx = \int_\Omega g\varphi \, dx \quad \text{for all } \varphi \in C_c^\infty(\Omega).$$

If $g_1, g_2 \in L^1_{\text{loc}}(\Omega)$ have this property then $g_1 = g_2$ a.e.; notation is therefore $\partial_j f := g$ (in case of existence).

*Proof.* Let $g := g_1 - g_2$. Then $\int_\Omega g\varphi\, dx = 0$ for all $\varphi \in C_c^\infty(\Omega)$. Let $x_0 \in \Omega$ and choose $r > 0$ such that $B_2 := \overline{B}(x_0, 2r) \subset \Omega$. Let $B := \overline{B}(x_0, r)$ and $\widetilde{g} := g1_{B_2} \in L^1(\mathbb{R}^d)$. Then for $x \in B$ and $n \ge 1/r$:

$$\begin{aligned}
0 &= \int \varphi_n(x-y)g(y)\, dy = \int_{B(0,1/n)} \varphi_n(y)g(x-y)\, dy \\
&= \int_{B(0,1/n)} \varphi_n(y)(g1_{B_2})(x-y)\, dy \\
&= (\varphi_n * g1_{B_2})(x).
\end{aligned}$$

By 5.1(c) we obtain $g1_{B_2} = 0$ a.e. on $B$, which proves the assertion. □

For $p \in [1,\infty]$, we say that $f$ has *a weak derivative in $L^p(\Omega)$ w.r.t. $x_j$*, if $\partial_j f \in L^p(\Omega)$. We shall use $\partial_j f \in L^p(\Omega)$ as a short hand notation (which thus is used to comprise implicitely also existence of the corresponding weak derivative). End Lect.22

Higher order derivatives are defined in the usual way. We use multiindex notation: for $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ let

$$\partial^\alpha := \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_d^{\alpha_d}.$$

The order of $\partial^\alpha$ is $|\alpha| := \alpha_1 + \ldots + \alpha_d$.

**5.3. Sobolev spaces:** Let $\Omega \subset \mathbb{R}^d$ be open. For $k \in \mathbb{N}$, $p \in [1,\infty)$ we let

$$\begin{aligned} W^{k,p}(\Omega) &:= \{u \in L^p(\Omega) : u \text{ has weak derivatives in } L^p(\Omega) \text{ up to order } k\} \\ &= \{u \in L^p(\Omega) : \forall |\alpha| \le k : \partial^\alpha u \in L^p(\Omega)\}. \end{aligned}$$

**Proposition:** $W^{k,p}(\Omega)$ is a Banach space for $p \in [1,\infty)$ and a Hilbert space for $p = 2$, where

$$\begin{aligned} \|u\|_{W^{k,p}(\Omega)} &:= \Big( \sum_{|\alpha| \le k} \|\partial^\alpha u\|_p^p \Big)^{1/p}, \\ (u|v)_{W^{k,2}(\Omega)} &:= \sum_{|\alpha| \le k} \int_\Omega \partial^\alpha u \, \overline{\partial^\alpha v} \, dx. \end{aligned}$$

The space $W^{k,p}(\Omega)$ is separable for $p \in [1,\infty)$ and reflexive for $p \in (1,\infty)$.

*Proof.* Let $(u_n)$ be a Cauchy sequence in $W^{k,p}(\Omega)$. Then, for each $|\alpha| \le k$, $(\partial^\alpha u_n)$ is a Cauchy sequence in $L^p(\Omega)$ and thus converges in $\|\cdot\|_{L^p}$ to some $v_\alpha \in L^p(\Omega)$. We let $u := v_0$ and show by induction on $|\alpha|$ that $v_\alpha = \partial^\alpha u$ for all $|\alpha| \le k$. For simplicity of notation we just treat $\alpha = e_j$ and $v_j := v_{e_j}$. For any $\varphi \in C_c^\infty(\Omega)$ we have $\partial_j \varphi u_n \to \partial_j \varphi u i$ and $\varphi \partial_j u_n \to \varphi v_j$ in $\|\cdot\|_{L^1}$. Hence

$$-\int \varphi v_j \, dx = -\lim_{n\to\infty} \int \varphi \partial_j u_n \, dx = \lim_{n\to\infty} \int \partial_j \varphi u_n \, dx = \int \partial_j \varphi u \, dx,$$

so $\partial_j u = v_j$ by definition, and $\partial_j u \in L^p(\Omega)$.

The final assertions follow, since the map

$$W^{k,p}(\Omega) \to Y := \oplus_{|\lambda| \le k} L^p(\Omega), \quad u \mapsto (\partial^\alpha u)_{|\alpha| \le k},$$

is isometric, and thus $W^{k,p}(\Omega)$ is isomorphic to a closed subspace of $Y$, which is separable and reflexive, respectively, under the stated assumptions on $p$. □

**Examples:** (1) The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, has a weak derivative $g$ on $\mathbb{R}$ given by $g(x) := \operatorname{sgn}(x)$. If $\varphi \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}\varphi \subseteq [-a, a]$ then, integrating by parts,

$$\begin{aligned}
-\int |x|\varphi'(x)\,dx &= \int_{-a}^{0} x\varphi'(x)\,dx - \int_0^a x\varphi'(x)\,dx \\
&= x\varphi(x)\Big|_{-a}^{0} - \int_{-a}^{0}\varphi(x)\,dx - x\varphi(x)\Big|_0^a + \int_0^a \varphi(x)\,dx \\
&= \int g(x)\varphi(x)\,dx.
\end{aligned}$$

(2) The function $f := 1_{[0,\infty)}$ has no weak derivative. For $\varphi \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}\varphi \subseteq [-a, a]$ we have

$$-\int f(x)\varphi'(x)\,dx = -\int_0^a \varphi'(x)\,dx = -\varphi\Big|_0^a = \varphi(0).$$

Suppose that $g$ is a weak derivative of $f$. Then $\int g\varphi\,dx = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$ with $\varphi(0) = 0$. By uniqueness in 5.2 we have $g = 0$ a.e. on $(0, \infty)$ and $g = 0$ a.e. on $(-\infty, 0)$. By $g \in L^1_{\mathrm{loc}}(\mathbb{R})$ we thus have $g = 0$ a.e. on $\mathbb{R}$, so $\int g\varphi\,dx = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$. By the line above this implies $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$ and (by translation) $C_c^\infty(\mathbb{R}) = \{0\}$, a contradiction to 5.1(b).

(3) Consider $A := \frac{d}{dx}$ in $L^2(0, 1)$ with $D(A) = C_c^\infty(0, 1)$. Observe that $D(A)$ is dense in $L^2(0, 1)$ (cf. 5.1 Corollary). So $A$ has an adjoint operator $A^*$, given by

$$\begin{aligned}
u \in D(A^*), A^*u = g &\iff \forall\varphi \in D(A) : (A\varphi|u) = (\varphi|g) \\
&\iff \forall\varphi \in C_c^\infty(0, 1) : \int_0^1 \varphi'\overline{u}\,dx = \int_0^1 \varphi\overline{g}\,dx.
\end{aligned}$$

Hence we have for $u, g \in L^2(0, 1)$: $u \in D(A^*)$ and $A^*u = g$ if and only if $u$ has weak derivative $-g$. This means that $D(A^*) = W^{1,2}(0, 1)$ and $A^*u = -u'$ in the sense of a weak derivative.

**5.4. Weak derivatives and Fourier transform:** We restrict to $p = 2$. Let $f \in L^2(\mathbb{R}^d)$.

(a) If $\xi \mapsto i\xi_j\hat{f}(\xi) \in L^2(\mathbb{R}^d)$ then $\partial_j f \in L^2(\mathbb{R}^d)$.

(b) If $\partial_j f \in L^2(\mathbb{R}^d)$ then $\widehat{\partial_j f}(\xi) = i\xi_j\hat{f}(\xi)$ for a.e. $\xi \in \mathbb{R}^d$.

*Proof.* (a) Let $g(\xi) := \mathscr{F}^{-1}(\xi \mapsto i\xi_j\hat{f}(\xi))$. Then $g \in L^2(\mathbb{R}^d)$, and for $\varphi \in C_c^\infty(\mathbb{R}^d)$ we have by Plancherel:

$$\begin{aligned}
\int g\,\overline{\varphi} &= \int \hat{g}\,\overline{\hat{\varphi}}\,d\xi = \int i\xi_j\hat{f}\,\overline{\hat{\varphi}}\,d\xi \\
&= -\int \hat{f}\,\overline{i\xi_j\hat{\varphi}}\,d\xi = -\int \hat{f}\,\overline{\widehat{\partial_j\varphi}}\,d\xi \\
&= -\int f\,\overline{\partial_j\varphi}\,dx.
\end{aligned}$$

Thus $\partial_j f = g \in L^2(\mathbb{R}^d)$.

(b) For $\varphi \in C_c^\infty(\mathbb{R}^d)$ we have, again by Plancherel and the definition of a weak derivative,

$$\begin{aligned}\int \widehat{\partial_j f}\,\overline{\hat\varphi}\,d\xi &= \int \partial_j f\,\overline{\varphi}\,dx = -\int f\,\partial_j\overline{\varphi}\,dx \\ &= \int \hat f\,\overline{\widehat{\partial_j\varphi}}\,d\xi = \int \hat f\,\overline{i\xi_j\hat\varphi}\,d\xi \\ &= \int i\xi_j \hat f\,\overline{\hat\varphi}\,d\xi.\end{aligned}$$

Since $C_c^\infty$ is dense in $L^2$ and $\mathscr{F} : L^2 \to L^2$ is an isometry, the set $\widehat{C_c^\infty}$ is dense in $L^2$. Hence $\widehat{\partial_j f}(\xi) = i\xi_j \hat f(\xi)$ for a.e. $\xi \in \mathbb{R}^d$. □

**Corollary:** For any $k \in \mathbb{N}$ we have $W^{k,2}(\mathbb{R}^d) = H^{k,2}(\mathbb{R}^d)$.

*Proof.* By the properties just shown we see that $u \in W^{k,2}(\mathbb{R}^d)$ if and only if $\xi \mapsto (i\xi)^\alpha \hat u(\xi) \in L^2$ for all $|\alpha| \le k$ where

$$z^\alpha := z_1^{\alpha_1} \cdot z_2^{\alpha_2} \cdot \ldots \cdot z_d^{\alpha_d} \quad \text{for } z = (z_1, z_2, \ldots, z_d) \in \mathbb{C}^d \text{ and } \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{N}_0^d.$$

Since, for $|\alpha| \le k$,

$$|(i\xi)^\alpha| \le |\xi|^{|\alpha|} \le (1 + |\xi|^2)^{|\alpha|/2} \le (1 + |\xi|^2)^{k/2},$$

we thus have $H^{k,2} \subseteq W^{k,2}$. On the other hand,

$$(1 + |\xi|^2)^k = (1 + \sum_{j=1}^d \xi_j^2)^k = \sum_{|\alpha| \le k} \xi^{2\alpha},$$

so that $u \in W^{k,2}$ implies $u \in H^{k,2}$. □

**Example:** Recall that we had defined $\Delta u := \mathscr{F}^{-1}(\xi \mapsto -|\xi|^2 \hat u(\xi))$ for $u \in H^{2,2}(\mathbb{R}^d)$. Now we know that $H^{2,2}(\mathbb{R}^d) = W^{k,2}(\mathbb{R}^d)$, and for $u \in W^{k,2}(\mathbb{R}^d)$ we have, using weak derivatives,

$$\sum_{j=1}^d \widehat{\partial_j^2 u}(\xi) = \sum_{j=1}^d (i\xi_j)^2 \hat u(\xi) = -|\xi|^2 \hat u(\xi),$$

and the definition via Fourier transform coincides with the more familiar one

$$\Delta u := \sum_{j=1}^d \partial_j^2 u$$

where derivatives are understood in the weak sense. End Lect.23

**5.5. Elliptic operators with constant coefficients:** Let $m \in \mathbb{N}$ and consider, in $L^2(\mathbb{R}^d)$,

$$A := \sum_{|\alpha|=2m} a_\alpha \partial^\alpha \quad \text{with } D(A) := W^{2m,2}(\mathbb{R}^d),$$

where $a_\alpha \in \mathbb{C}$ for $|\alpha| = 2m$. For $u \in W^{2m,2}(\mathbb{R}^d)$ and $\xi \in \mathbb{R}^d$ we have

$$\widehat{Au}(\xi) = \sum_{|\alpha|=2m} a_\alpha \widehat{\partial^\alpha u}(\xi) = \sum_{|\alpha|=2m} a_\alpha (i\xi)^\alpha \hat{u}(\xi).$$

Observe that $(i\xi)^\alpha = i^{|\alpha|}\xi^\alpha$, so that

$$\widehat{Au}(\xi) = \underbrace{(-1)^m \sum_{|\alpha|=2m} a_\alpha \xi^\alpha}_{=:a(\xi)} \hat{u}(\xi), \quad \xi \in \mathbb{R}^d,$$

and

$$Au = \mathscr{F}^{-1}(\xi \mapsto a(\xi)\hat{u}(\xi)).$$

The function $a : \mathbb{R}^d \to \mathbb{C}$, $\xi \mapsto a(\xi)$, is called the *symbol* of the differential operator $A$. It is clear that

$$|a(\xi)| \le C|\xi|^{2m}, \quad \xi \in \mathbb{R}^d,$$

for $C := \sum_{|\alpha|=2m} |a_\alpha|$. The operator $A$ is called *elliptic* if, for some $\eta > 0$,

$$|a(\xi)| \ge \eta|\xi|^{2m}, \quad \xi \in \mathbb{R}^d.$$

This is the case if and only if $a(\xi) \neq 0$ for all $\xi \in \mathbb{R}^d \setminus \{0\}$ (one can take $\eta := \inf\{|a(\xi)| : |\xi| = 1\}$, then $|a(\xi)| = |\xi|^{2m}|a(\xi/|\xi|)| \ge \eta|\xi|^{2m}$ for $\xi \in \mathbb{R}^d \setminus \{0\}$).

If $A$ is elliptic then $A$ is a closed operator in $L^2(\mathbb{R}^d)$, since the square of the graph norm

$$\|u\|_2^2 + \|Au\|_2^2 = \|\hat{u}\|_2^2 + \|\widehat{Au}\|_2^2 = \int_{\mathbb{R}^d} (1 + |a(\xi)|^2)|\hat{u}(\xi)|^2 \, d\xi$$

is equivalent to

$$\int_{\mathbb{R}^d} (1 + |\xi|^{4m})|\hat{u}(\xi)|^2 \, d\xi, \text{ i.e. to } \int_{\mathbb{R}^d} (1 + |\xi|^2)^{2m}|\hat{u}(\xi)|^2 \, d\xi = \|u\|_{H^{2m,2}}^2.$$

The operator $A$ is unitarily equivalent to the multiplication operator $f \mapsto a \cdot f$ with domain $\{f \in L^2 : a \cdot f \in L^2\}$, and thus we have

$$\sigma(A) = \overline{\{a(\xi) : \xi \in \mathbb{R}^d\}}.$$

**5.6. Example:** Let $A\varphi := \varphi''$ in $L^2(0,1)$ with $D(A) = C_c^\infty(0,1)$. Then $A$ is densely defined and

$$\begin{aligned} u \in D(A^*), A^*u = g \quad &\iff \quad \forall \varphi \in D(A) : (A\varphi|u) = (\varphi|g) \\ &\iff \quad \forall \varphi \in C_c^\infty(0,1) : \int_0^1 \varphi'' u \, dx = \int_0^1 \varphi g \, dx. \end{aligned}$$

**Claim:** For $g \in L(0,1)$ the function $h : x \mapsto \int_0^x g \, dt$ is in $C[0,1]$, and for $\varphi \in C_c^\infty(0,1)$ we have

$$- \int_0^1 h\varphi' \, dx = \int_0^1 g\varphi \, dx,$$

i.e. $h \in W^{1,1}(0,1)$ and $h' = g$.

*Proof.* Take a mollifier $\rho$ and let $fg_n := \rho_n * g$, $h_n := \int_0^\cdot g_n \, dt$, $n \in \mathbb{N}$. Then $g_n \to g$ in $\|\cdot\|_1$, hence $h_n \to h$ in $\|\cdot\|_\infty$, and for $\varphi \in C_c^\infty(0,1)$ we have

$$- \int_0^1 h\varphi' \, dx = - \lim_{n\to\infty} \int_0^1 h_n\varphi' \, dx = \lim_{n\to\infty} \int_0^1 g_n\varphi \, dx = \int_0^1 g\varphi \, dx.$$

□

Now let $u \in D(A^*)$ and $g := A^*u$, and set $h := \int_0^\cdot g \, dt$. By the claim we have $h \in W^{1,2}(0,1)$ and

$$- \int_0^1 h\varphi' \, dx = \int_0^1 g\varphi \, dx = \int_0^1 \varphi'' u \, dx, \quad \varphi \in C_c^\infty(0,1).$$

We shall show that $u' = h + c$ for some constant $c \in \mathbb{C}$.

Take $\rho \in C_c^\infty(0,1)$ such that $\int_0^1 \rho \, dt = 1$. Let $\varphi \in C_c^\infty(0,1)$ and consider $\psi := \varphi - \rho \cdot \int_0^1 \varphi \, dt$. Then $\psi \in C_c^\infty(0,1)$ and

$$\int_0^1 \psi \, dx = \int_0^1 \varphi \, dx - \int_0^1 \rho \, dx \cdot \int_0^1 \varphi \, dt = 0,$$

and thus $\int_0^\cdot \psi \, dt \in C_c^\infty(0,1)$. Hence $- \int_0^1 h\psi \, dx = \int_0^1 \psi' u \, dx$ by the equation above. This means

$$\begin{aligned} - \int_0^1 u\varphi' \, dx \quad &= \quad \int_0^1 h\varphi \, dx - \int_0^1 h\rho \, dx \cdot \int_0^1 \varphi \, dt - \int_0^1 u\rho' \, dx \cdot \int_0^1 \varphi \, dt \\ &= \quad \int_0^1 \Big( h - \int_0^1 h\rho \, dt - \int_0^1 u\rho' \, dt \Big) \varphi \, dx. \end{aligned}$$

If we let $c := - \int_0^1 h\rho \, dt - \int_0^1 u\rho' \, dt$ then $u' = h + c$. We conclude that $(u')' = h' = g$ and $u \in W^{2,2}(0,1)$. We thus have shown $D(A^*) \subseteq W^{2,2}(0,1)$ and $A^*u = u''$ for $u \in D(A^*)$.

If, on the other hand, $u \in W^{2,2}(0,1)$ then

$$\int_0^1 u''\varphi\, dx = -\int_0^1 u'\varphi'\, dx = \int_0^1 u\varphi''\, dx, \quad \varphi \in C_c^\infty(0,1),$$

which implies $u \in D(A^*)$ and $A^*u = u''$. We thus have shown that $A^* = \frac{d^2}{dx^2}$ with $D(A^*) = W^{2,2}(0,1)$.

Clearly $A \subseteq A^*$, $A$ is symmetric and closable, $\overline{A} \subseteq A^*$. However, $\overline{A}$ is **not self-adjoint**: If $(\varphi_n)$ is a sequence in $C_c^\infty(0,1)$ such that $\varphi_n \to u$ in $\|\cdot\|_2$ and $\varphi_n'' \to v$ in $\|\cdot\|_2$ then $\varphi_n(x) := \int_0^x \int_0^t \varphi_n''(s)\, ds\, dt$ converges in $\|\cdot\|_\infty$ and thus $u(x) := \int_0^x \int_0^t v(s)\, ds\, dt$. We conclude that $u \in C^1[0,1]$ and $\varphi_n \to u$ in $\|\cdot\|_{C^1}$. In particular $u(0) = u(1) = u'(0) = u'(1) = 0$ which is not true for arbitrary $u \in W^{2,2}(0,1) = D(A^*)$. Hence $D(\overline{A})$ is strictly smaller than $D(A^*)$. Roughly speaking, $D(A)$ is too small and $D(A^*)$ is too large.

Before we proceed the notice that an exercise gives the continuos embedding $W^{1,2}(0,1) \to C^{1/2}[0,1]$:

$$|u(x) - u(y)| = |\int_x^y u'(t)\, dt| \le \int_x^y |u'(t)|\, dt \le \|u'\|_2 |x-y|^{1/2},$$

where we used the Cauchy-Schwarz inequality in the last step. Similarly, one has the continuous embedding $W^{2,2}(0,1) \to C^{1,1/2}[0,1] := \{u \in C^1[0,1] : u' \in C^{1/2}[0,1]\}$.

For $u, v \in W^{2,2}(0,1)$ one has the following formula (which can be proved by a mollifier argument):

$$\int_0^1 u''v\, dx = u'v\Big|_0^1 - \int_0^1 u'v'\, dx = u'v\Big|_0^1 - uv'\Big|_0^1 + \int_0^1 uv''\, dx.$$

For $B \in \{D, N\}$ we define $\Delta_B$ as the restriction of $\frac{d^2}{dx^2}$ to the domains

$$\begin{aligned} D(\Delta_D) &:= \{u \in W^{2,2}(0,1) : u(0) = u(1) = 0\} \quad (\textit{Dirichlet boundary conditions}), \\ D(\Delta_N) &:= \{u \in W^{2,2}(0,1) : u'(0) = u'(1) = 0\} \quad (\textit{Neumann boundary conditions}), \end{aligned}$$

respectively. By the formula above, $\Delta_{D/N}$ is symmetric, i.e. $\Delta_{D/N} \subseteq (\Delta_{D/N})^*$, and the operator is closed since its domain is a closed subspace of $W^{2,2}(0,1)$ (use the continuous embedding).

Since $\Delta_{D/N}$ is an extension of $A$, the operator $(\Delta_{D/N})^*$ is a restriction of $A^*$, i.e. $D((\Delta_{D/N})^*) \subseteq W^{2,2}(0,1)$ and $(\Delta_{D/N})^* v = v''$ for $v \in D((\Delta_{D/N})^*)$. End Lect.24

For $v \in W^{2,2}(0,1)$ we thus have

$$v \in D((\Delta_{D/N})^*) \iff \forall u \in D(\Delta_{D/N}) : (u''|v) = (u|v'') \iff \forall u \in D(\Delta_{D/N}) : u'v\Big|_0^1 = uv'\Big|_0^1.$$

For Dirichlet boundary conditions this means

$$v \in D(\Delta_D^*) \iff \forall u \in D(\Delta_D) : u'v\Big|_0^1 = 0 \iff v \in D(\Delta_D),$$

and for Neumann boundary conditions this means

$$v \in D(\Delta_N^*) \iff \forall u \in D(\Delta_N) : uv'\Big|_0^1 = 0 \iff v \in D(\Delta_N).$$

Hence $\Delta_{D/N}$ is self-adjoint.

**5.7. The Dirichlet form:** We first recall the *divergence theorem*: Let $\Omega \subset \mathbb{R}^d$ be open and bounded with $\partial\Omega \in C^1$. If $F \in C^1(\Omega) \cap C(\overline{\Omega})$ then

$$\int_\Omega \operatorname{div} F \, dx = \int_{\partial\Omega} \nu \cdot F \, d\sigma,$$

where $\nu : \partial\Omega \to \mathbb{R}^d$ denotes the outer normal unit.

**Consequence:** For $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ and $v \in C^1(\Omega) \cap C(\overline{\Omega})$ we have

$$\int_\Omega \nabla u \cdot \overline{\nabla v} \, dx = -\int_\Omega (\Delta u)\overline{v} \, dx + \int_{\partial\Omega} \frac{\partial u}{\partial \nu} \overline{v} \, d\sigma. \tag{(*)}$$

For $u, v \in W^{1,2}(\Omega)$ define

$$\mathfrak{a}(u, v) := \int_\Omega \nabla u \cdot \overline{\nabla v} \, dx.$$

Then $\mathfrak{a} : W^{1,2}(\Omega) \times W^{1,2}(\Omega) \to \mathbb{C}$ is sesquilinear, *symmetric* (i.e. $\mathfrak{a}(u, v) = \overline{\mathfrak{a}(v, u)}$), and *accretive* (i.e. $\operatorname{Re} \mathfrak{a}(u, u) \geq 0$, observe that, by symmetry, $\mathfrak{a}(u, u)$ is real in our case). We let $V_N := W^{1,2}(\Omega)$ and $V_D := W_0^{1,2}(\Omega)$, where $W^{1,2}(\Omega)$ denotes the closure of $C_c^\infty(\Omega)$ in $(W^{1,2}(\Omega), \| \cdot \|_{W^{1,2}})$ (functions $u \in W_0^{1,2}(\Omega)$ "vanish on the boundary $\partial\Omega$" in a certain sense). Then $V_{D/N}$ is dense in $L^2(\Omega)$ and $V_{D/N}$ is complete for the norm $\| \cdot \|_{\mathfrak{a}}$, given by

$$\|u\|_{\mathfrak{a}} := \Big( \|u\|_2^2 + \mathfrak{a}(u, u) \Big)^{1/2},$$

since this norm equals $\| \cdot \|_{W^{1,2}}$. We shall define Dirichlet and Neumann Laplacian on $\Omega$ via the following abstract result.

**5.8. Theorem:** Let $H$ be a Hilbert space and $V$ a dense linear subspace. Let $\mathfrak{a} : V \times V \to \mathbb{C}$ be sesquilinear, symmetric, accretive, and such that $V$ is a Hilbert space for the inner product

$$(u|v)_V := \mathfrak{a}(u, v) + (u|v)_H.$$

Define the operator $A$ in $H$ by

$$u \in D(A) \text{ and } Au = f \;:\iff\; u \in V \text{ and } \forall v \in V : \mathfrak{a}(u, v) = (f|v).$$

Then $A$ is a self-adjoint linear operator in $H$ with $(Au|u) \geq 0$ for all $u \in D(A)$.

*Proof.* We denote by $V^*$ the *anti-dual* of $V$, i.e. the space of all continuous antilinear functionals $\phi : V \to \mathbb{C}$. By the Riesz representation theorem, the operator

$$\mathcal{B} : V \to V^*, \quad u \mapsto \mathfrak{a}(u, \cdot) + (u|\cdot)_H$$

is an isomorphism and thus has a continuous inverse $\mathcal{R} : V^* \to V$.

We identify $H$ with its anti-dual $H^*$. Since $V$ is densely and continuously embedded in $H$ we have a continuous injection $H^* \to V^*$. Thus we have continuous embeddings $V \hookrightarrow H = H^* \hookrightarrow V^*$. This means that $u \in H$ is identified with $(u|\cdot)_H\Big|_V \in V^*$.

Now let $R := \mathcal{R}|_H$. Then $R \in \mathcal{L}(H)$ and $R$ is injective (since $\mathcal{R}$ is). We define $A := R^{-1} - I_H$. Then $A$ is a closed linear operator in $H$ and $(A+1)^{-1} = R \in \mathcal{L}(H)$, i.e. $-1 \in \rho(A)$. Moreover, we have, for $u, f \in H$:

$$\begin{aligned} u \in D(A),\ Au = f \iff & R(f+u) = u \iff u \in V \text{ and } \mathcal{R}(f+u) = u \\ \iff & u \in V \text{ and } \mathcal{B}u = u + f \\ \iff & u \in V \text{ and } \forall v \in V : \mathfrak{a}(u,v) + (u|v)_H = (u|v)_H + (f|v)_H \\ \iff & u \in V \text{ and } \forall v \in V : \mathfrak{a}(u,v) = (f|v)_H. \end{aligned}$$

$A$ is symmetric: For $u, v \in D(A)$ we have

$$(Au|v)_H = \mathfrak{a}(u,v) = \overline{\mathfrak{a}(v,u)} = \overline{(Av|u)_H} = (u|Av)_H.$$

We also have $(Au|u)_H = \mathfrak{a}(u,u) \geq 0$. Using 4.20 (including the remark) we see that $A$ is self-adjoint, since $-1 \in \rho(A)$ (cf. also 4.14). □

**5.9. Application:** Let $\Omega$ be bounded with $\partial\Omega \in C^1$. Define the negative of the Dirichlet Laplacian $-\Delta_D$ on $\Omega$ by applying 5.8 to the sesquilinear form $\mathfrak{a}$ from 5.7, defined on $V_D \times V_D$. If $u \in D(-\Delta_D)$ and $v \in V_D$ satisfy the smoothness assumptions for ((*)) then $u$ vanishes at the boundary and

$$\mathfrak{a}(u,v) = (-\Delta u|v),$$

since $v$ vanishes at the boundary. This justifies the term "Dirichlet Laplacian" (the classical Dirichlet boundary condition reads $u|_{\partial\Omega} = 0$). If, on the other hand, $u, v$ satisfy the regularity assumptions and $u|_{\partial\Omega} = 0$, $v|_{\partial\Omega} = 0$, then

$$\mathfrak{a}(u,v) = (-\Delta u|v).$$

The classical Neumann boundary condition reads: $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$. If $u$ satisfies this condition and $u, v$ satisfy the regularity assumptions for ((*)), then

$$\mathfrak{a}(u,v) = (-\Delta u|v).$$

On the other hand, if $u$ satisfies the regularity assumptions for ((*)) and we have

$$\mathfrak{a}(u, v) = (-\Delta u | v).$$

for all $v$ satisfying the regularity assumptions, then $\int_{\partial\Omega} \frac{\partial u}{\partial \nu} \overline{v} \, d\sigma = 0$ for all such $v$, which implies $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$.

**Remark:** Via Theorem 5.8 we can define Dirichlet and Neumann Laplacian as self-adjoint operators on *arbitrary* open subsets $\Omega$ of $\mathbb{R}^d$.

The standard reference for the method of sesquilinear forms is Kato's book: Perturbation Theory for Linear Operators. End Lect.25

# 6 Spectral projections and holomorphic functional calculus

When nothing els is said, $X$ denotes a Banach space and $A$ is a closed linear operator in $X$.

For self-adjoint operators in a Hilbert space we were able to define projection operators corresponding to arbitrary subsets of the spectrum. For arbitrary operators in Banach spaces we need a topological assumption on the part of the spectrum "on which" we want to project.

**6.1. Spectral projections:** Suppose that $\sigma(A) = \sigma_0 \cup \sigma_1$ where $\sigma_0 \cap \sigma_1 = \emptyset$, $\sigma_0$ is compact, and $\sigma_1$ is closed. Then we find finitely many closed piecewise $C^1$-curves $\Gamma$ in $\mathbb{C} \setminus \sigma(A)$ such that

$$n(z, \Gamma) = \begin{cases} 1 & , z \in \sigma_0 \\ 0 & , z \in \sigma_1 \end{cases} ,$$

where $n(z, \Gamma) := \frac{1}{2\pi i} \int_\Gamma \frac{d\zeta}{\zeta - z}$ denotes the *winding number*. Let

$$P := \frac{1}{2\pi i} \int_\Gamma R(\lambda, A)\, d\lambda = \frac{1}{2\pi i} \int_0^1 R(\gamma(t), A) \dot{\gamma}(t)\, dt$$

where $\gamma : [0, 1] \to \mathbb{C}$ is a parametrization of $\Gamma$ and the integral is a Riemann integral in the sense of §5.

The operator $P$ has the following properties:

(a) $P \in \mathcal{L}(X)$ and $P^2 = P$, i.e. $P$ is a projection.

(b) $P$ commutes with $A$, i.e. for $x \in D(A)$ we have $Px \in D(A)$ and $APx = PAx$. Moreover, $X_0 := R(P) \subseteq D(A)$.

(c) If we let $A_0 := A|_{X_0}$ then $A_0 \in \mathcal{L}(X_0)$ and $\sigma(A_0) = \sigma_0$.

(d) Let $X_1 := N(P)$ and define $A_1$ as the restriction of $A$ to $D(A_1) := D(A) \cap X_1$. Then $A_1$ is a closed linear operator in $X_1$ and $\sigma(A_1) = \sigma_1$.

For the proof we shall need the following.

**6.2. Holomorphic functions with values in a Banach space:** Cauchy's theorem and the Cauchy integral formula hold for $X$-valued holomorphic functions $f : \Omega \to X$:

$$\int_\Gamma f(z)\, dz = 0 \quad \text{and} \quad \frac{1}{2\pi i} \int_\Gamma \frac{f(z)}{z - z_0}\, dz = f(z_0)$$

if $\Gamma$ is a finite family of closed piecewise $C^1$-curves such that $n(z, \Gamma) = 0$ for all $z \notin \Omega$ and $n(z_0, \Gamma) = 1$.

*Proof of 6.1.* (a) $P \in \mathcal{L}(X)$ is well defined since $t \mapsto R(\gamma(t), A)\dot\gamma(t)$ is piecewise continuous with values in $\mathcal{L}(X)$. By Cauchy's theorem, $\Gamma$ does not depend on the special choice of $\Gamma$, so we choose another family $\Gamma'$ such that $n(\lambda, \Gamma') = 1$ for all $\lambda \in \Gamma$ ($\Gamma'$ encircles all points of $\Gamma$ once) and $n(\mu, \Gamma) = 0$ for all $\mu \in \Gamma'$. This means that $\Gamma$ is "inside" $\Gamma'$. Then, by the resolvent equation and 6.2,

$$\begin{aligned}
P^2 &= \Big(\frac{1}{2\pi i}\Big)^2 \int_\Gamma \int_{\Gamma'} R(\lambda, A)R(\mu, A)\, d\mu\, d\lambda \\
&= \Big(\frac{1}{2\pi i}\Big)^2 \int_\Gamma \int_{\Gamma'} \frac{R(\lambda, A) - R(\mu, A)}{\mu - \lambda}\, d\mu\, d\lambda \\
&= \frac{1}{2\pi i} \int_\Gamma \underbrace{\frac{1}{2\pi i} \int_{\Gamma'} \frac{1}{\mu - \lambda}\, d\mu}_{=1}\, R(\lambda, A)\, d\lambda + \Big(\frac{1}{2\pi i}\Big)^2 \int_{\Gamma'} \underbrace{\int_\Gamma \frac{R(\mu, A)}{\mu - \lambda}\, d\lambda}_{=0}\, d\mu \\
&= \frac{1}{2\pi i} \int_\Gamma R(\lambda, A)\, d\lambda = P.
\end{aligned}$$

(b) Actually, $t \mapsto R(\gamma(t), A)\dot\gamma(t)$ is piecewise continuous with values in $\mathcal{L}(X, [D(A)])$, so $P \in \mathcal{L}(X, [D(A)])$, in particular $R(P) \subseteq D(A)$. We thus have

$$AP = \frac{1}{2\pi i} \int_\Gamma AR(\lambda, A)\, d\lambda,$$

where $t \mapsto AR(\gamma(t), A)\dot\gamma(t)$ is pieceweise continuous with values in $\mathcal{L}(X)$. For $x \in D(A)$ we have $AR(\lambda, A)x = R(\lambda, A)Ax$, which implies $APx = PAx$.

(c) and (d): Since $P \in \mathcal{L}(X)$ is a projection, $X_0 = R(P)$ is a closed subspace of $X$. By (b) we have $A_0 : X_0 \to X_0$. By closedness of $A$, this operator is closed, hence $A_0 \in \mathcal{L}(X_0)$. Clearly, $P$ commutes with resolvents of $A$, which implies that $X_0$ is invariant under resolvents of $A$. By Exercise 51 we thus have $\sigma(A_0) \subseteq \sigma(A)$, and $R(\lambda, A_0) = R(\lambda, A)|_{X_0}$ for $\lambda \in \rho(A)$.

On the other hand, for $\mu\sigma\sigma_1$ let

$$R_\mu := \frac{1}{2\pi i} \int_\Gamma \frac{R(\lambda, A)}{\mu - \lambda}\, d\lambda.$$

Then $R_\mu|_{X_0} \in \mathcal{L}(X_0)$, and by $AR(\lambda, A) = \lambda R(\lambda, A) - I_X$ we have

$$(\mu - A_0)R_\mu|_{X_0} = \frac{1}{2\pi i} \int_\Gamma \frac{(\mu - A)R(\lambda, A)|_{x_0}}{\mu - \lambda}\, d\lambda = P|_{X_0} - \frac{1}{2\pi i} I_{X_0}\, d\lambda = I_{X_0}.$$

Clearly, $R_\mu|_{X_0}$ commutes with $\mu - A_0$, so also $R_\mu|_{X_0}(\mu - A_0) = I_{X_0}$. Hence $\mu \in \rho(A_0)$ and $R(\mu, A_0) = R_\mu|_{X_0}$. We thus have shown $\sigma(A_0) \subseteq \sigma_0$.

Similarly, $X_1$ is invariant under resolvents of $A$, and thus $\sigma(A_0) \subseteq \sigma(A)$. For $\mu \in \sigma_0$ let

$$R_\mu := \frac{1}{2\pi i} \int_\Gamma \frac{R(\lambda, A)}{\mu - \lambda}\, d\lambda.$$

Then $R_\mu \in \mathcal{L}(X, [D(A)])$ commutes with resolvents of $A$, hence with $P$, and $R_\mu|_{X_1} \in \mathcal{L}(X_1)$. For $x \in X_1$ we have

$$(\mu - A)R_\mu x = \underbrace{\frac{1}{2\pi i}\int_\Gamma R(\lambda, A)x\, d\lambda}_{=Px=0} + \frac{1}{2\pi i}\int_\Gamma \frac{x}{\mu - \lambda}\, d\lambda = -x.$$

Since $R_\mu$ commutes with $A$ we obtain $\mu \in \rho(A_1)$ and $R(\mu, A_1) = -R_\mu|_{X_1}$. This means that we have shown $\sigma(A_1) \subseteq \sigma_1$.

The decomposition $X = X_0 \oplus X_1$ leads to the decomposition of $A$ as $\begin{pmatrix} A_0 & 0 \\ 0 & A_1 \end{pmatrix}$ with domain $X_0 \oplus D(A_1)$. It is thus clear that $\sigma(A) = \sigma(A_0) \cup \sigma(A_1)$ (this union is closed since $\sigma(A_0)$ is compact). Since $\sigma(A_j) \subseteq \sigma_j$ $(j = 0, 1)$ and $\sigma_0 \cap \sigma_1 = \emptyset$ this is only possible if $\sigma(A_j) = \sigma_j$ for $j = 0, 1$. □

**6.3. The Dunford functional calculus:** Let $A$ be bounded in $X$. Then $\sigma(A)$ is compact.

(1) Let $U \subseteq \mathbb{C}$ be an open neighborhood of $\sigma(A)$. Then we find a finite family of closed piecewise $C^1$-curves $\Gamma$ such that $n(z, \Gamma) = 1$ if $z \in \sigma(A)$ and $n(z, \Gamma) = 0$ for $z \notin U$. Let $f : U \to \mathbb{C}$ be a holomorphic function. Define

$$f(A) := \frac{1}{2\pi i}\int_\Gamma f(\lambda)R(\lambda, A)\, d\lambda.$$

Then $f(A) \in \mathcal{L}(X)$ and the definition does not depend on the special choice of $\Gamma$.

(2) The map $\Psi : f \mapsto f(A)$ is linear and multiplicative, and

$$p_n(A) = A^n \quad \text{for all } n \in \mathbb{N}_0, \qquad \text{where } p_n(\lambda) := \lambda^n,$$

i.e. $\Psi$ is a functional calculus or the operator $A$ for functions that are holomorphic in a neighborhood of $\sigma(A)$.

*Proof.* (1) is clear. (2): Linearity is clear. The proof of multilicity is similar to what we have done in the proof of 6.1(a). This implies

$$p_n(A) = \frac{1}{2\pi i}\int_\Gamma \lambda^n R(\lambda, A), d\lambda = p_1(A)^n, \quad n \in \mathbb{N}_0.$$

Moreover, $Ap_0(A) = p_1(A)p_0(A)$. Thus it rests to show $p_0(A) = I_X$. Apply 6.1 with $\sigma_0 = \sigma(A)$ and $\sigma_1 = \emptyset$. Then $A_1 \in \mathcal{L}(X_1)$ (since $A \in \mathcal{L}(X)$), and $\sigma(A_1) = \sigma_1 = \emptyset$. Hence $X_1 = \{0\}$, $X_0 = X$, and $p_0(A) = P = I_X$. □

End
Lect.26

## Further aspects

**Sobolev embeddings:** (a) Let $\Omega \subseteq \mathbb{R}^d$ be open and $d \geq 2$. Then

$$W_0^{1,2}(\Omega) \hookrightarrow L^q(\Omega) \quad \text{for all} \begin{cases} q \in [2, \infty) & \text{, if } d = 2 \\ q \in [2, \frac{2d}{d-2}] & \text{, if } d \geq 3. \end{cases}$$

If $\Omega$ is bounded then the embedding $W_0^{1,2}(\Omega) \hookrightarrow L^2(\Omega)$ is compact.

(b) Let $\Omega \subset \mathbb{R}^d$ be bounded and $\partial\Omega \in C^1$. Then

$$W^{1,2}(\Omega) \hookrightarrow L^q(\Omega) \quad \text{for all} \begin{cases} q \in [2, \infty) & \text{, if } d = 2 \\ q \in [2, \frac{2d}{d-2}] & \text{, if } d \geq 3. \end{cases}$$

Moreover, the embedding $W^{1,2}(\Omega) \hookrightarrow L^q(\Omega)$ is compact.

Instead of a proof we look at the following scaling argument in $\mathbb{R}^d$. For this purpose we let $\|u\|_{\dot{W}^{1,2}} := \|\nabla u\|_{L^2(\mathbb{R}^d)}$. For any $\lambda > 0$ and $q \in [1, \infty]$ we have, by substituting $x = y/\lambda$,

$$\|u(\lambda\cdot)\|_{L^q(\mathbb{R}^d)} = \lambda^{-d/q}\|u\|_{L^q(\mathbb{R}^d)}.$$

On the other hand, for $u \in W^{1,2}(\mathbb{R}^d)$,

$$\|u(\lambda\cdot)\|_{\dot{W}^{1,2}} = \|\lambda(\nabla u)(\lambda\cdot)\|_{L^2(\mathbb{R}^d)} = \lambda^{1-d/2}\|u\|_{\dot{W}^{1,2}}.$$

This means that the existence of a constant $C$ with

$$\|u\|_{L^q(\mathbb{R}^d)} \leq C\|u\|_{\dot{W}^{1,2}} \quad \text{for all } u \in W^{1,2}(\mathbb{R}^d)$$

implies $\lambda^{-d/q} \leq C'\lambda^{1-d/2}$ for all $\lambda > 0$, whch in turn implies $-d/q = 1 - d/2$, i.e. $q = \infty$ if $d = 2$ and $q = 2d/(d-2)$ if $d \geq 3$. For this $q$, the above estimate holds for $d \geq 3$, but it does **not** hold for $d = 2$. For $d \geq 3$ one thus has

$$W^{1,2}(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d) \cap L^{2d/(d-2)}(\Omega),$$

and the latter space embeds into $L^q(\Omega)$ for $q \in (, 2d/(d-2))$.

**Consequences:** (a) The Dirichlet Laplacian on bounded $\Omega \subset \mathbb{R}^d$: The operator $-\Delta_D$ in $L^2(\Omega)$ is associated (via 5.8) to the sesquilinear form

$$\mathfrak{a}(u, v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\, dx,$$

defined on $V := W_0^{1,2}(\Omega)$ (cf. 5.9). In particular, we have

$$D(-\Delta_D) \subseteq W_0^{1,2}(\Omega) \subseteq L^q(\Omega) \subseteq L^2(\Omega)$$

for $q \in [2, \infty)$ (if $d = 2$) or $q \in [2, 2d/(d-2)]$ (if $d \geq 3$), respectively, and $L^q(\Omega)$ is invariant under the resolvents of $-\Delta_D$. By Exercise 51, the operator $-\Delta_{D,q}$ defined as the restriction of $-\Delta_D$ to

$$D(-\Delta_{D,q}) := \{u \in D(-\Delta_D) : -\Delta_D u \in L^q(\Omega)\}$$

is a closed linear operator in $L^q(\Omega)$, whose spectrum satisfies

$$\sigma(-\Delta_{D,q}; L^q(\Omega)) = \sigma(-\Delta_D; L^2(\Omega)).$$

Here we used $\sigma(A; X)$ to indicate that the spectrum is considered for $A$ as a linear operator in $X$. Moreover, for $\lambda$ in the resolvent set, $R(\lambda, -\Delta_{D,q}) \in \mathcal{L}(L^q(\Omega))$ is the restriction of $R(\lambda, -\Delta_D) \in \mathcal{L}(L^2(\Omega))$ to $L^q(\Omega)$. This is sometimes called *consistency* of resolvents.

By (a) above, $-\Delta_D$ has compact resolvents, so

$$\sigma(-\Delta_D; L^2(\Omega)) = \{\lambda_n : n \in \mathbb{N}\} \text{ where } 0 \leq \lambda_1 < \lambda_2 < \lambda_3 \ldots \uparrow \infty$$

and each $\lambda_n$ is an eigenvalue of finite multiplicity. One can show $\lambda_1 > 0$ ($Au = 0$ implies $\mathfrak{a}(u, u) = 0$, so $\nabla u = 0$ a.e. in $\Omega$; a closer study of weak derivatives shows that $u$ must then be constant on connected components of $\Omega$, but $u \in W_0^{1,2}(\Omega)$ then implies $u = 0$ a.e. in $\Omega$ – here one needs a better understanding of the sense in which functions in $W_0^{1,2}(\Omega)$ vanish on $\partial\Omega$).

We conclude that $\sigma(-\Delta_{D,q}) = \{\lambda_n : n \in \mathbb{N}\}$. Actually, also the corresponding eigenspaces coincide. To see this we define, for fixed $n_0 \in \mathbb{N}$, $P \in \mathcal{L}(L^2(\Omega))$ as the spectral projection for $-\Delta_D$ and $\sigma_0 := \{\lambda_{n_0}\}$. By consistency of resolvents, $P|_{L^q(\Omega)}$ is the corresponding spectral projection for $-\Delta_{D,q}$. The space $P(L^2(\Omega))$ is the eigenspace of $-\Delta_D$ for the eigenvalue $\lambda_{n_0}$ which has finite dimension. Since clearly $P(L^q(\Omega)) \subseteq P(L^2(\Omega))$, also $P(L^q(\Omega))$ consists of eigenfunctions and has finite dimensions, and $P(L^q(\Omega))$ is the eigenspace of $-\Delta_{D,q}$ as an operator in $L^q(\Omega)$. Since $L^q \cap L^2$ is dense in $L^2$ and in $L^q$, the space $P(L^2 \cap L^q)$ is dense in in $P(L^2)$ and dense in $P(L^q)$. We conclude that $P(L^2(\Omega)) = P(L^q(\Omega))$, i.e. the eigenspaces coincide.

(b) The Neumann Laplacian on bounded $\Omega \subset \mathbb{R}^d$ with $\partial\Omega \in C^1$: Same conclusions, but $\lambda_1 = 0$ and $1_\Omega$ is an eigenfunction.

In fact, the assertions on the spectrum in $L^q(\Omega)$ in (a) and (b) above even hold for any $q \in [1, \infty]$. This can be seen via the following steps: for $q = \infty$ one can use suitable versions of the *maximal principle*, for $q \in (2, \infty)$ one can use interpolation arguments, and the case $q \in [1, 2)$ then follows by considering adjoint operators.

For $\Omega$ as above, $-\Delta_{D/N}$ has compact resolvents and $\sigma_{\text{ess}}(-\Delta_{D/N}) = \emptyset$. For self-adjoint operators one can show

$$\sigma(A) = \sigma_{\text{ess}}(A) \,\dot{\cup}\, \sigma_d(A),$$

where $\sigma_d(A)$ is the *discrete spectrum of $A$*:

$$\sigma_d(A) := \{\lambda \in \mathbb{C} : \lambda \text{ is an isolated eigenvalue of finite multiplicity}\}.$$

**Warnings for the Neumann Laplacian:** (1) For any closed subset $M \subseteq [0,\infty)$ there exists a bounded domain $\Omega \subseteq \mathbb{R}^2$ such that $\sigma_{\text{ess}}(-\Delta_N) = M$ (cf. Hempel, Seco, Simon, Journal of Functional Analysis, 1991).

(2) By a result due to Liskevich and Perelmuter (1996) one always has, for any $q \in [1,\infty]$,

$$\sigma(-\Delta_{N,q}; L^q(\Omega)) \subseteq \Sigma(\theta_q) := \{z \in \mathbb{C} \setminus \{0\} : |\arg z| \leq \theta_q\} \cup \{0\},$$

where $\theta_q = \arcsin|1 - 2/q|$. There exists $\Omega \subset \mathbb{R}^2$ connected such that

$$\sigma(-\Delta_{N,q}; L^q(\Omega)) = \Sigma(\theta_q), \quad q \in [1,\infty],$$

cf. Kunstmann, Mathematische Zeitschrift, 2002. The domain $\Omega$ is unbounded there, but this should also be possible for bounded $\Omega$. An intermediate step in the proof is that, for

$$\Omega := \{(x,y) \in \mathbb{R}^2 : x > 0, |y| \leq e^{-x}\},$$

the essential spectrum $\sigma_{\text{ess}}(-\Delta_{N,q}; L^q(\Omega))$ is a parabolic region intersecting $\partial\Sigma(\theta_q)$ in two points. Then a scaling argument is used. On the technical side, perturbation arguments for Fredholm operators are crucial.

**Regularity:** In general, it is **not** true that $D(-\Delta_{D/N}) \subset W^{2,2}(\Omega)$ or $D(-\Delta_{D/N,q}) \subset W^{2,q}(\Omega)$! For bounded $\Omega$ this holds, e.g., if $\partial\Omega \in C^2$ (these question are strongly related to *elliptic regularity*).

**Non-constant coefficients:** Theorem 5.8 is also applicable, e.g., to the sesquilinear form

$$\mathfrak{a}(u,v) := \int_\Omega a(x)\nabla u \cdot \overline{\nabla v}\, dx,$$

where $a : \Omega \to \mathbb{R}^{d\times d}$ is bounded measurable and satisfies $a(x)^* = a(x)$ and

$$a(x)\xi \cdot \overline{\xi} \geq \eta|\xi|^2, \quad \xi \in \mathbb{C}^d,$$

for some $\eta > 0$ and a.e. $x \in \Omega$. The operator associated to $\mathfrak{a}$ via 5.8 is formally given by $A = -\text{div}\,(a(x)\nabla)$, and boundary conditions are determined by the form domain $V$.

**Perturbations:** Perturbation arguments can be used to include first and zero order terms.

**Sectorial operators:** Many elliptic operators (with lower order terms included and after adding a suitable constant) are *sectorial* in various function spaces. A linear operator $A$ in a Banach space $X$ is called *sectorial of angle* $\omega \in [0,\pi)$, if $\sigma(A) \subset \Sigma(\omega)$ and if, for any $\theta \in (\omega,\pi)$, the set

$$\{\lambda R(\lambda, A) : \lambda \notin \Sigma(\theta)\} \subset \mathcal{L}(X)$$

is bounded.

Observe that a self-adjoint operator $A$ with $A \geq 0$ is sectorial of angle $\omega = 0$: For $\lambda \in \mathbb{C} \setminus [0, \infty)$ one has by the spectral theorem

$$\|\lambda R(\lambda, A)\| = \frac{|\lambda|}{d(\lambda, \sigma(A))} \leq \frac{|\lambda|}{d(\lambda, [0, \infty))}.$$

If $\operatorname{Re} \lambda \leq 0$ then $d(\lambda, [0, \infty)) = |\lambda|$. If $\operatorname{Re} \lambda > 0$ then $d(\lambda, [0, \infty)) = |\operatorname{Im} \lambda|$. For $\operatorname{Re} \lambda > 0$ we have $\sin(\arg \lambda) = \operatorname{Im} \lambda / |\lambda|$. Thus we see that, for $\theta \in (0, \pi/2)$,

$$\sup\{\|\lambda R(\lambda, A)\| : \lambda \notin \Sigma(\theta)\} \leq \frac{1}{\sin \theta}.$$

More general, if $A$ is a linear operator in a Hilbert space with $\rho(A) \neq \emptyset$ whose numerical range $W(A)$ satisfies $W(A) \subseteq \Sigma(\omega)$ where $\omega \in (0, \pi)$, then similar arguments show that $A$ is a sectorial operator of angle $\omega$.

If $A$ is a sectorial operator of angle $\omega$ then $A$ has a functional calculus for functions $f$ that are holomorphic on a sector $\Sigma(\theta)$ with $\theta \in (\omega, \pi)$ and satisfy

$$|f(z)| \leq C \frac{|z|^\alpha}{1 + |z|^{2\alpha}}, \quad z \in \Sigma(\theta),$$

for some $C, \alpha > 0$. The operator $f(A)$ is given by the (absolutely convergent) Dunford type integral

$$f(A) := \frac{1}{2\pi i} \int_{\partial \Sigma(\sigma)} f(\lambda) R(\lambda, A) \, d\lambda,$$

where $\sigma \in (\omega, \theta)$ and $\partial \Sigma(\sigma)$ is parametrized by arc length in such a way that $\Sigma(\sigma)$ lies to the left.

It might even be that $A$ has a functional calculus for functions that are holomorphic and bounded on $\Sigma(\theta)$, i.e. an $H^\infty(\Sigma(\theta))$-functional calculus. For more on the $H^\infty$-functional calculus see, e.g., Kunstmann and Weis in Lecture Notes in Mathematics 1855 (2004). End Lect.27