Technical University Berlin

Lecture Notes

Functional Analysis II

Spectral Theory, Fourier-Analysis, Distributions and regularisation of inverse problems

Based on lecture notes and homework exercises by Prof. Dr. Gitta Kutyniok from the winter term 2018/2019 and more

Contents

List of Figures — Page i

1 Spectral theory for self-adjoint operators on Hilbert spaces — Page 1

1.1 Recapitulation of Spectral Theory . . . 1
1.2 Functional Calculus . . . 3
1.3 Continuous Functional Calculus . . . 11
1.4 Measurable Functional Calculus . . . 17
1.5 Spectral theorem for bounded self-adjoint operators . . . 28
1.6 Applications of the spectral theorem . . . 32
1.7 Spectral theorem for unbounded self-adjoint operators . . . 34
1.8 Spectral properties of unbounded operators . . . 38

2 Fourier Analysis, Distributions and Sobolev spaces — Page 40

2.1 The Fourier transform on $L^1(\mathbb{R}^n)$ . . . 40
2.2 The Uncertainty Principle . . . 47
2.3 Fourier transform on $L^1(G)$ and $L^2(G)$ . . . 50
2.4 The Fourier Transform on $\mathcal{S}(\mathbb{R}^n)$ . . . 54
2.5 Tempered distributions . . . 58
2.6 Sobolev spaces . . . 67

3 Linear Ill-Posed Inverse Problems — Page 70

3.1 The Moore-Penrose Pseudoinverse . . . 71
3.2 Singular Value Decomposition for Compact Operators . . . 77
3.3 Regularisation of ill-posed problems . . . 81

4 Spectral regularisation — Page 84

4.1 Discretisation as Regularisation . . . 90

Appendices — Page 92

A Additional proofs — Page 92

A.1 Spectral theory . . . 93

A.2 Unbounded Operators 94
A.3 FOURIER transforms 95

INDEX — **PAGE 97**

By Viktor Glombik.

Edited on April 13, 2022.

List of Figures

1 The numerical range of a normal matrix with eigenvalues -1, $1 \pm i$, $2 \pm i$ and a complex matrix. 3

2 The unit pulse, its translated version and their Fourier transforms. 40

3 The convolution of the unit pulse with itself yields the standard hat function. 42

4 The hat function φ in two dimensions. 43

5 The Fejér kernel and its Fourier transform for $\lambda = 5$. 44

6 An example of a mollifier ω, given by $\exp\left(\frac{1}{x^2-1}\right) \cdot \mathbb{1}_{|x|\leqslant 1}(x)$. 59

7 [10] 70

8 Visualisation of the proof of theorem 3.1.2. 75

9 The integration domain 79

1 Spectral theory for self-adjoint operators on Hilbert spaces

1.1 Recapitulation of Spectral Theory

From now on, let $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space over $\mathbb{C}$. For sake of brevity, we will write $T-\lambda$ instead of $T-\lambda I$.

Definition 1.1.1 (Spectrum, eigenvalue, resolvent)
For $T \in L(\mathscr{H})$

- the spectrum of T is defined by

$$\sigma(T) := \{\lambda \in \mathbb{C} : T-\lambda \text{ is not bijective}\}.$$

  Its complement $\rho(T)$ is the resolvent set of T. For $\lambda \in \rho(T)$ the operator $(T-\lambda)^{-1} \in L(\mathscr{H})$ is called resolvent of T in λ.
- $\lambda \in \mathbb{C}$ is called an eigenvalue of T if $\ker(T-\lambda) \neq \{0\}$.
- $r(T) := \sup_{\lambda\in\sigma(T)} |\lambda|$ is called spectral radius of T.

Definition 1.1.2 (Decomposition of the spectrum)
Let $T_\lambda := T-\lambda$. $\sigma(T)$ is the disjoint union of the

- point spectrum $\sigma_p(T) := \{\lambda \in \sigma(T) : T_\lambda \text{ not injective}\}$,

  $\lambda \in \sigma_p(T) \iff \lambda$ eigenvalue.
- continuous spectrum

  continuous spectrum

$$\sigma_c(T) := \{\lambda \in \sigma(T)\backslash\sigma_p(T) : \operatorname{ran}(T_\lambda) \subsetneq \mathscr{H} \text{ dense}\},$$

- residual spectrum $\sigma_r(T) := \sigma(T)\backslash(\sigma_p(t) \cup \sigma_c(T))$ or

  residual spectrum

$$\sigma_r(T) = \{\lambda \in \sigma(T) : T_\lambda \text{ injective}, \operatorname{ran}(T_\lambda) \subset \mathscr{H} \text{ not dense}\}.$$

Furthermore the approximate point spectrum is

approximate point spectrum

$$\sigma_{\text{app}}(T) := \{\lambda \in \mathbb{C} : \inf_{\|x\|=1} \|T_\lambda x\| = 0\} \supset \sigma_p(T).$$

We have $\sigma_c = \sigma_{\text{app}}\backslash(\sigma_p \cup \sigma_r)$.

Definition 1.1.3 (Self-adjoint, normal, unitary)
$T \in L(\mathscr{H})$ is called self-adjoint if $T = T^*$, normal if $TT^* = T^*T$, unitary if $TT^* = T^*T = I$ and positive ($T \geqslant 0$) if $\langle Tx, x\rangle \geqslant 0$ holds for all $x \in \mathscr{H}$.

Self-adjoint and unitary operators are normal.

Theorem 1.1.1: Spectral Theory from FA I

For $T \in L(\mathscr{H})$ the following statements hold:

- $r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leqslant \|T\|$
- $\sigma(T) \neq \varnothing$ is compact and $\rho(T)$ is open.
- $\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|$ for self-adjoint T.

Lemma 1.1.4 (Bound on resolvent operator)
For $T = T^* \in L(\mathscr{H})$ and $\lambda \in \mathbb{C}\setminus\mathbb{R}$, $\|(T-\lambda)^{-1}\| \leqslant |\Im(\lambda)|^{-1}$ holds.

Proof. As T is self-adjoint, $\langle Tx, x\rangle$ is real for all $x \in \mathscr{H}$. Thus

$$\Im(\langle (T-\lambda)x, x\rangle) = -\Im(\lambda)\|x\|^2$$

holds. For $A := T - \lambda$

$$|\Im(\lambda)|\|x\|^2 = |\Im(\langle Ax, x\rangle)| \leqslant |\langle Ax, x\rangle| \overset{\text{CS}}{\leqslant} \|Ax\|\|x\|$$

holds. For a unit vector $x \in \mathscr{H}$

$$|\Im(\lambda)|\|x\| \leqslant \|Ax\| \implies |\Im(\lambda)|\|A^{-1}x\| \leqslant \|x\| \implies \|A^{-1}x\| \leqslant \frac{1}{|\Im(\lambda)|}\|x\|,$$

holds, implying $\|A^{-1}\| \leqslant \frac{1}{|\Im(\lambda)|}$. □

Corollary 1.1.5 (GÜ2-1 Hanover WiSe11)
For $T = T^*$, $T_\pm := T \pm iI$ are boundedly invertible with $\|T_\pm^{-1}\| \leqslant 1$.

Lemma 1.1.6 (Unitary resolvent bound)
For unitary $U \in L(\mathscr{H})$ and $|\lambda| \neq 1$ it holds that $\|(U-\lambda)^{-1}\| \leqslant \frac{1}{||\lambda|-1|}$.

Proof. We have $\|U\| = \|U^{-1}\| = 1$. For $|\lambda| > 1$ we therefore have $\left\|\frac{U}{\lambda}\right\| < 1$. We have $(U-\lambda)^{-1} = -\frac{1}{\lambda}\left(1 - \frac{U}{\lambda}\right)^{-1}$. Thus

$$\begin{aligned}\|(U-\lambda)^{-1}\| &= \left\| -\frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{U}{\lambda}\right)^k \right\| \leqslant \frac{1}{|\lambda|}\sum_{k=0}^{\infty}\left\|\frac{U}{\lambda}\right\|^k \\ &= \frac{1}{|\lambda|\left(1 - \left\|\frac{U}{\lambda}\right\|\right)} = \frac{1}{|\lambda| - \|U\|} = \frac{1}{|\lambda| - 1}\end{aligned}$$

holds. For $|\lambda| < 1$ we have $(1 - U^{-1}\lambda)^{-1} = \sum_{k=0}^{\infty}(U^{-1}\lambda)^k$ and thus similarly as before it holds that

$$\|(U-\lambda)^{-1}\| = \left\| U^{-1}\sum_{k=0}^{\infty}(\lambda\cdot U^{-1})^k \right\| \leqslant \frac{\|U^{-1}\|}{1 - \|\lambda\cdot U^{-1}\|} = \frac{1}{1-|\lambda|}.$$ □

1.2 Functional Calculus

16.10.19

If $f:\mathbb{R}\to\mathbb{R}$ is a function and $T\in L(\mathscr{H})$ an operator, there is no reason why the expression

$$f(T)$$

should make sense (without changing the domain of f). The idea of a functional calculus is to create a *principled* approach to this kind of overloading of notation which we have already used plenty, e.g. T^2, where $f(x)=x^2$. [1]

Definition 1.2.1 (Numerical Range (O. Toelpitz, 1918))
For $T\in L(\mathscr{H})$, $W(T):=\{\langle Tx,x\rangle : \|x\|=1\}$ is called the numerical range of T (and $r(T):=\sup_{\lambda\in W(T)}|\lambda|$ its numerical radius).

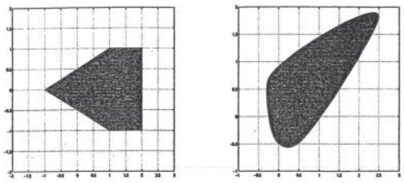

Figure 1: The numerical range of a normal matrix (left) with eigenvalues -1, $1\pm i$, $2\pm i$ and a complex matrix (right). [2]
More examples with pictures here and here.

In A.1.1 we collect properties of the numerical range and radius.

Lemma 1.2.2 (Spectral Inclusion Theorem (Wintner, 1929))
For $T\in L(\mathscr{H})$, $\sigma(T)\subset\overline{W(T)}$ holds.

Proof. Let $\lambda\notin\overline{W(T)}$ and define $d:=\operatorname{dist}(\lambda,W(T))>0$. For $x\in\mathscr{H}$ with $\|x\|=1$

$$d\leqslant|\lambda-\langle Tx,x\rangle|=|\lambda\|x\|^2-\langle Tx,x\rangle|=|\langle(\lambda-T)x,x\rangle|\overset{\text{CS}}{\leqslant}\|(\lambda-T)x\|$$

holds, implying the injectivity of $\lambda-T$ and the closedness of $\operatorname{ran}(\lambda-T)$ since $(\lambda-T)^{-1}:\operatorname{ran}(\lambda-T)\to\mathscr{H}$ exists and is bounded by $\frac{1}{d}$ (cf. [7]). Assume there exists a unit vector $x_0\in(\operatorname{ran}(\lambda-T))^{\perp}$. Then

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

holds, implying $\lambda\in W(T)$, a contradiction. Therefore $\operatorname{ran}(\lambda-T)=\mathscr{H}$ and $\lambda\in\rho(T)$.

We present an alternative more direct proof of lemma 1.2.2 using the decomposition of the spectrum, to better show why and when the closure of $W(T)$ is needed. [6]

As a direct consequence, if $\dim(\mathscr{H})<\infty$ we have $\sigma(T)\subset W(T)$.

Proof. Case 1: $\lambda\in\sigma_p(T)$. Then there exists a $v\in H$ such that $Tv=\lambda v$. Thus $\langle Tv,v\rangle=\lambda\|v\|^2$ holds, implying $\left\langle T\frac{v}{\|v\|},\frac{v}{\|v\|}\right\rangle=\lambda$, which means $\lambda\in W(T)$.

Case 2: $\lambda\in\sigma_r(T)$. Since the range is not dense, there exists a $v\in(\overline{\operatorname{ran}(T-\lambda)})^{\perp}$. Thus $0=\left\langle\frac{v}{\|v\|},(T-\lambda)\frac{v}{\|v\|}\right\rangle$ holds, implying $\lambda\in W(T)$ as above.

Case 3: $\lambda\in\sigma_c(T)$. There exist a sequence of unit vectors z_n with $(T-\lambda)z_n\to0$, otherwise $T-\lambda$ would be bounded from below and would necessarily have a closed range. Thus

$A\in L(\mathscr{H})$ is called bounded from below if there exists a $c>0$ such that $\|Ax\|\geqslant c\|x\|$ for all $x\in\mathscr{H}$.
$A\in L(\mathscr{H})$ is bijective $\iff$ $\operatorname{ran}(A)\subset\mathscr{H}$ is dense and A is bounded from below.

$$\frac{\langle z_n,Tz_n\rangle}{\|z_n\|^2}-\lambda=\frac{\langle z_n,(T-\lambda)z_n\rangle}{\|z_n\|^2}\to0,$$

holds, implying $\lambda\in\overline{W(T)}$. □

Example 1.2.3 (Numerical range of the left-shift operator)

1. Let $T_1\in L(\mathbb{C}^2)$ be given by the matrix $\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$.

   Let $x=(x_1,x_2)\in\mathbb{C}^2$ with $\|x\|=1$. Then

   $$\langle T_1x,x\rangle=\langle\left(\begin{smallmatrix}x_2\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)\rangle=\overline{x_1}x_2$$

   This shows that $W(T_1)=W\left(\left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)\right)=W\left(\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)\right)$. (HW 1-1 WiSe00).

   holds. Since $\|x\|=1$ we can write $x_1=e^{i\varphi_1}\cos(\theta)$ and $x_2:=e^{i\varphi_2}\sin(\theta)$ for $\theta\in\left[0,\frac{\pi}{2}\right]$ and $\varphi_{1,2}\in\mathbb{R}$. This implies

   $$\overline{x_1}x_2=e^{i(\varphi_2-\varphi_1)}\sin(\theta)\cos(\theta).$$

The range of $\sin(x)\cos(x) = \frac{\sin(2x)}{2}$ is $\left[-\frac{1}{2}, \frac{1}{2}\right]$ and the since image of $z \mapsto e^{iz}$ is the closed unit disk $\overline{\mathcal{D}}$, we deduce that $W(T_1) = \frac{\overline{\mathcal{D}}}{2}$.

If that argument doesn't seem rigorous enough, we can instead notice that $|x_1|^2 + |x_2|^2 = 1$ also allows $|x_1|^2 + |e^{i\theta}x_2|^2 = 1$, thus we have $e^{i\theta}W(T_1) = W(T_1)$ for $\theta \in [0, 2\pi)$, implying that $W(T_1)$ is circular around 0. Since $|\overline{x}_1 x_2| = \left|\overline{x}_1\sqrt{1-|x_1|^2}\right| = |x_1|\sqrt{1-|x_1|^2}$, we can define the radius of $W(T_1)$ as $f : [0,1] \to [0,1]$, $z \mapsto z\sqrt{1-z^2}$. We find $\max_{z\in[0,1]} f(z) = \frac{1}{2}$.

Since $\sigma(T_1) = \{0\}$ and $T_1^2 = 0$ we have $(T_1 - \lambda I)^{-1} = \sum_{k=0}^{\infty} \frac{T_1^k}{\lambda^{k+1}} = \frac{1}{\lambda} I - \frac{1}{\lambda^2} T_1$.

2 Let $T_2 \in L(\ell^2(\mathbb{N}))$ be the left shift operator which is given by $T_2((x_1, x_2, \ldots)) = (x_2, \ldots)$.

We have

$$W(T_2) = \left\{ \sum_{k=1}^{\infty} x_k \overline{x_{k+1}} : \|x\|_{\ell_2} = 1 \right\}$$

Noticing that T_2 is the infinite-dimensional analogon to T_1, we can, analogously to the above, see that $W(T_2)$ is circular disk around zero. Therefore, it suffices to find its radius $\sup_{\|x\|=1} |\langle T_2 x, x\rangle|$, which is the numerical radius.

From the exercises **(WE DO???)** we know that $\mathcal{D} = \sigma(T) \subset \overline{W(T)}$, where inclusion follows from lemma 1.5. That $\mathcal{D} \subset W(T)$ can be seen by considering

$$x^{(n)} := \Big(\underbrace{\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}}_{n \text{ times}}, 0, \ldots \Big)$$

for $n \in \mathbb{N}$ which has norm one. We have

$$\langle T_2 x^{(n)}, x^{(n)} \rangle = \sum_{k=1}^{\infty} x_k \overline{x_{k+1}} = \sum_{k=1}^{n-1} \frac{1}{n} = 1 - \frac{1}{n} \xrightarrow{n\to\infty} 1.$$

Furthermore

$$|\langle T_2 x, x\rangle| \overset{\text{CS}}{\leqslant} \|T_2\| \|x\|^2 = \|T_2\| = 1,$$

holds for all $x \in \ell_2$ with unit norm. This implies $\mathcal{D} \subset W(T_2) \subset \overline{\mathcal{D}}$. We now aim to show that the latter inclusion is strict, i.e. $W(T_2) = \mathcal{D}$.

Let $x \in \ell_2$ with $\|x\|_{\ell_2} = |\langle T_2 x, x\rangle| = 1$. By the Cauchy-Schwartz inequality (cf. above) we have that $\|T_2 x\| = \|x\|$, implying that $x_1 = 0$. Since the Cauchy-Schwartz inequality is an equality precisely when there exists a linear dependency there exists a $\lambda \in \mathbb{C}\setminus\{0\}$ such that $x_{k+1} = \lambda x_k$ for all $k \in \mathbb{N}$. Together with $x_1 = 0$ this implies $x = 0$ contradicting $\|x\|_{\ell_2} = 1$. ◇

Exercise, 21.10.19

Example 1.2.4 (Spectrum of the right-shift operator)
Define the (unitary!) right-shift operator by

$$T : \ell_2 \to \ell_2,\ (x_1, x_2, \ldots) \mapsto (0, x_1, x_2, \ldots).$$

Now for if $Tx = \lambda x$ for some $\lambda \in \mathbb{C}$ and some $x \in \ell_2$ we have $x = 0$ for $\lambda \neq 0$ and $\lambda = 0$, so $\sigma_p(T) = \varnothing$.

Now let $\lambda \in \mathbb{C}$ with $|\lambda| \leqslant 1$. Then $z_0 := (1, 0, \dots) \in \ell_2 \setminus \operatorname{ran}(\lambda - T)$: Assume there exists a $z \in \ell_2$ such that $(\lambda - T)z = z_0$. This implies $z_1 = \lambda^{-1}$ and $\lambda z_k - \lambda^{-1} = 0$ for all $k \geqslant 2$. Thus we have $z = \left(\lambda^{-k}\right)_{k \in \mathbb{N}_{>0}} \notin \ell_2$ as $|\lambda| \leqslant 1$. Thus $\sigma_r(T) = \{\lambda \in \mathbb{C} : |\lambda| \leqslant 1\}$ holds, as $\operatorname{ran}(T - \lambda) = \ell_2 \setminus \operatorname{span}(z_0)$, which is not dense.

If $|\lambda| < 1$ we have, as T is an isometry,

$$\|(T - \lambda)x\| \geqslant \|Tx\| - |\lambda|\|x\| = (1 - |\lambda|)\|x\|.$$

Hence $\lambda \notin \sigma_{\text{app}}(T)$.

Now let $\lambda \in \sigma_r(T)$. If $|\lambda| < 1$, let $x^{(n)} := \frac{1}{\sqrt{n}}\left(1, \lambda^{-1}, \dots \lambda^{1-n}, 0 \dots\right)$ for $n \in \mathbb{N}$. Then $\|x^{(n)}\| = 1$ but

$$\begin{aligned} \|(\lambda - T)x^{(n)}\| &= \left\| \left(\frac{\lambda}{\sqrt{n}}, \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n}}, \frac{1}{\lambda\sqrt{n}} - \frac{1}{\lambda\sqrt{n}}, \dots \right) \right\| \\ &= \left\| \left(\frac{\lambda}{\sqrt{n}}, 0, \dots \right) \right\| \xrightarrow{n \to \infty} 0 \end{aligned}$$

hence $\sigma_{\text{app}}(T) = \{\lambda \in \mathbb{C} : |\lambda| = 1\} = \sigma(T)$. ◇

This conclusion is also true for a more general class of operators. A unitary operator is normal. By a different version of the spectral theorem, a bounded operator on a Hilbert space is normal if and only if it is equivalent (after identification of $\mathscr{H}$ with an L^2 space) to a multiplication operator. It can be shown that the approximate point spectrum of a bounded multiplication operator equals its spectrum. By the theorem, U is unitarily equivalent to multiplication by a Borel-measurable f on $L^2(\mu)$, for some finite measure space (X, μ). Now $UU^* = I$ implies $|f(x)|^2 = 1$, μ-a.e. This shows that the essential range of f, therefore the spectrum of U, lies on the unit circle.

TODO

First, we need some properties of normal operators, some of which will be used in the proof of theorem 1.3.1:

Lemma 1.2.5 (Properties of normal operators)
First we note that T normal $\iff$ T^* normal $\iff$ $\|Tx\| = \|T^*x\|$ for all $x \in \mathscr{H}$. For normal T the following statements hold:

1. $\ker(T) = \ker(T^*)$, $\overline{\operatorname{ran}}(T) = \overline{\operatorname{ran}}(T^*)$ *and* $\mathscr{H} = \ker(T) \oplus \overline{\operatorname{ran}}(T)$.
2. *If $\alpha \neq \beta$ are eigenvalues,* $\ker(T - \alpha) \perp \ker(T - \beta)$.
3. $\sigma_r(T) = \varnothing$
4. $r(T) = \|T\|$, *which follows from* $\|T^*T\| = \|T\|^2 = \|T^2\|$.
5. $\sigma_{\text{app}}(T) = \sigma(T)$ *(Weyl's criterion)*

We show $\operatorname{ran}(T) = \operatorname{ran}(T^*)$ in A.1.2.

Proof. (HW) The first equivalence is clear. For $x \in \mathscr{H}$

$$\langle Tx, Tx \rangle = \langle x, T^*Tx \rangle = \langle x, TT^*x \rangle = \langle T^*x, T^*x \rangle$$

holds. For " $\Longleftarrow$ " consider $L := TT^* - T^*T$, which is self-adjoint. Thus

$$\|L\| = \sup_{\|x\|=1} |\langle (TT^* - T^*T)x, x \rangle| = \sup_{\|x\|=1} |\langle T^*x, T^*x \rangle - \langle Tx, Tx \rangle| = 0$$

holds, implying $L = 0$, i.e. $TT^* = T^*T$.

1. We have $Tx = 0$ iff $\|Tx\| = \|T^*x\| = 0$ iff $T^*x = 0$. Therefore $\overline{\operatorname{ran}}(T) = \ker(T^*)^\perp = \ker(T)^\perp = \overline{\operatorname{ran}}(T^*)$ holds. As $\overline{\operatorname{ran}}(T)$ is closed, it holds that

$$\mathscr{H} = \overline{\operatorname{ran}}(T) \oplus \overline{\operatorname{ran}}(T)^\perp = \overline{\operatorname{ran}}(T) \oplus \ker(T^*) = \overline{\operatorname{ran}}(T) \oplus \ker(T).$$

2. A simple computation shows that $T - \alpha$ is normal. Therefore $\ker(T - \alpha)^\perp = \overline{\operatorname{ran}}(T - \alpha)$ follows from 1. Let x_α be an eigenvector

of T with eigenvalue α and x_β be chosen analogously. Then $x_\alpha \in \ker(T-\alpha)$ and $x_\beta \in \overline{\operatorname{ran}}(T-\alpha)$ hold as

$$(T-\alpha)\frac{x_\beta}{\beta-\alpha} = \frac{\beta x_\beta - \alpha x_\beta}{\beta-\alpha} = x_\beta.$$

3 By 1 we have that $\ker(T) = \{0\} \iff \overline{\operatorname{ran}}(T) = \mathscr{H}$, implying $\sigma_r(T) = \varnothing$.

4 As T^*T is self-adjoint we have

$$\|T^*T\| = \sup_{\|x\|=1} \langle T^*Tx, x\rangle = \sup_{\|x\|=1} \|Tx\|^2 = \|T\|^2.$$

Furthermore,

$$\|T^2x\|^2 = \langle T^*T^*TTx, x\rangle = \langle T^*Tx, T^*Tx\rangle = \|TT^*x\|^2.$$

By taking $\sup_{\|x\|=1}$ and the square root, we can conclude $\|T^2\| = \|T\|^2$. We can inductively conclude $\|T^{2^k}\| = \|T\|^{2^k}$ for $k \in \mathbb{N}$, therefore

$$\lim_{k\to\infty} \|T^{2^k}\|^{\frac{1}{2^k}} = \lim_{k\to\infty} \|T\|^{\frac{2^k}{2^k}} = \|T\|, \qquad \square$$

because if the limit of a sequence exists, then the limit of every subsequence is the equal to that limit.

5 "$\subset$": Let $\lambda \in \sigma_{\text{app}}(T)$, i.e. $\inf_{\|x\|=1} \|(T-\lambda)x\| = 0$. Thus there exists a sequence of unit vectors $(x_n)_{n\in\mathbb{N}} \subset \mathscr{H}$ such that $(T-\lambda)x_n \to 0$, as the norm is continuous. If $\lambda \in \rho(T)$, $T-\lambda$ is invertible, so $x_n \to (T-\lambda)^{-1}\cdot 0 = 0$, which is contradiction to $\|x_n\| = 1$ together with the continuity of the norm.

"$\supset$": Let $\lambda \in \sigma(T)$ and $\ker(T-\lambda) \neq \{0\}$. Then there exists a unit vector $v \in \mathscr{H}$ such that $(T-\lambda)v = 0$, so $\lambda \in \sigma_{\text{app}}(T)$ is immediate by choosing $(x_n := v)_{n\in\mathbb{N}}$.

Now assume $\ker(T-\lambda) = \{0\}$. Then $\overline{\operatorname{ran}}(T-\lambda) = \mathscr{H}$ by 1. If $\operatorname{ran}(T-\lambda)$ were closed, then $T-\lambda I$ were bijective, implying $\lambda \in \rho(T)$; a contradiction.

Thus, $T-\lambda$ is injective with a dense, non-closed range and an unbounded inverse on $\operatorname{ran}(T-\lambda)$. Hence, there exists a sequence of unit vectors $(x_n)_{n\in\mathbb{N}} \subset \operatorname{ran}(T-\lambda)$ such that $\|y_n\| \to \infty$, where $y_n := .(T-\lambda)^{-1}x_n$. Normalising the tail of this sequence gives a sequence of unit vectors $\left(z_n := \frac{1}{\|y_n\|}y_n\right) \subset \mathscr{H}$ such that $(T-\lambda)z_n \to 0$, implying $\lambda \in \sigma_{\text{app}}(T)$.

This result from example 1.2.4 can be generalised: The spectrum of unitary operators such as R lies on the unit circle:

Lemma 1.2.6 ($\sigma(U) \subset \{z \in \mathbb{C} : |z| = 1\}$ for unitary U)
Let $U \in L(\mathscr{H})$ be unitary. Then $\sigma(U) \subset \{z \in \mathbb{C} : |z| = 1\}$ holds.

Proof. (HW) As U is unitary, its spectrum is contained in the unit ball: we have $|\lambda| \leqslant \|U\|$ for all $\lambda \in \sigma(U)$ and

$$\|U\| = \sup_{\|x\|=1} \sqrt{\langle Ux, Ux\rangle} = \sup_{\|x\|=1} \sqrt{\langle x, U^*Ux\rangle} = 1.$$

Assume there exists a $\lambda \in \sigma(T)$ with $|\lambda| < 1$. Then

$$\|(\lambda - U)x\| \geqslant \|Ux\| - |\lambda|\|x\| \geqslant (1-\lambda)\|x\|$$

holds, implying that $U - \lambda$ is bounded from below and therefore $\lambda \notin \sigma_{\text{app}}(T)$. This is a contradiction to $\lambda \in \sigma(T)$ by lemma 1.2.5 5. □

Alternatively one can show that $U - \lambda I$ is bijective for $|\lambda| < 1$: As U^* is unitary, too, we have $\|\lambda U^*\| < 1$ so by Neumann $I - \lambda U^* = U^*(U - \lambda)$ is invertible: $(I - \lambda U^*)^{-1} = \sum_{k=0}^{\infty}(\lambda U^*)^k$. Therefore, U^* and $(I-\lambda U^*)^{-1}$ (and $U - \lambda I$) commute, so $(U-\lambda)^{-1} = U^*(I-\lambda U^*)^{-1}$ holds as $(U - \lambda I)U^* = I - \lambda U^*$.

TODO: Is this true for unitary operators on spaces without scalar product???

Remark 1.2.7 ([13]) If $U \in L(\mathscr{H})$ is unitary with $\sigma(U) = \{1\}$, we have $U = I$ as, because $U - I$ is normal we have $0 \leqslant \|U - I\| = \sup_{\lambda\in\sigma(U-I)} |\lambda| \leqslant \sup_{\lambda\in\sigma(U)\setminus\sigma(I)} |\lambda| = 0$.

Example 1.2.8 (Spectrum of the "ℓ^p-multiplication operator")
Let $a := (a_n)_{n\in\mathbb{N}} \in \ell_\infty(\mathbb{C})$, $p \in [1,\infty)$ and consider

$$A_p : \ell_p(\mathbb{C}) \to \ell_p(\mathbb{C}),\ (x_n)_{n\in\mathbb{N}} \mapsto (a_n x_n)_{n\in\mathbb{N}}$$

Then $A_p \in L(\ell_p(\mathbb{C}))$. From FA I (HW 4, Ex 3) we know that $\|A_p\| = \|a\|_\infty$ and $T_p^* = T_q$ for Hölder-conjugates $p, q \geqslant 1$.

To determine the spectrum of A_p we first check $\sigma_p(A_p)$. We have $A_p e_k = a_k e_k$, so each a_k is an eigenvalue with eigenvector e_k.

Since

$$A_2^* : \ell_2(\mathbb{C}) \to \ell_2(\mathbb{C}),\ (y_n)_{n\in\mathbb{N}} \mapsto (\overline{a_n} y_n)_{n\in\mathbb{N}},$$

we conclude that A_2 is normal, so $\sigma(A_2) = \sigma_{\text{app}}(A_2) = \overline{(a_n)_{n\in\mathbb{N}}}$, which is the set of all a_n and all limit points of convergent subsequences.

TODO[HW 6-1 WiSe01]Let a_n converge to zero. Show that A_p is compact and find $\sigma(A_p)$. Is zero an eigenvalue? Show that for each non-zero eigenvalue of A_p there are at most finitely many linearly independent eigenvectors.

Wie muss a_n gewählt werden, damit A_p kein endlichdimensionales Bild hat? ◇

Example 1.2.9 (Volterra integral operator on $\mathcal{C}$ (HW 2-1))
We have $\sigma_c(T) = \sigma_{\text{app}}(T) = \{0\}$ and $\sigma_p(T) = \sigma_r(T) = \varnothing$ for the Volterra integral operator

$$T : \mathcal{C}([0,1]) \to \mathcal{C}([0,1]),\ Tf(x) \mapsto \int_0^x f(t)\,\mathrm{d}t.$$

From FA I we know that that since $X := (\mathcal{C}([0,1]), \|\cdot\|_\infty)$ is a Banach space and T is compact, if $0 \neq \lambda \in \sigma(T)$, then λ is an eigenvalue. This implies that $\sigma(T)$ can only contain 0, as $\sigma_p(T) = \varnothing$: <u>Case 1:</u> Assume there exists an eigenvalue $\lambda \neq 0$ and corresponding eigenfunction $f \in X$ of T, i.e.

$$\int_0^x f(t)\,\mathrm{d}t = \lambda f(x) \quad \forall x \in [0,1]. \tag{1}$$

Differentiation yields (cf. fundamental theorem of calculus) yields

$$f(x) = \lambda f'(x) \implies f(x) = c \cdot e^{\frac{x}{\lambda}}$$

for $c \in \mathbb{R}$. Substituting $f(x) = c \cdot e^{\frac{x}{\lambda}}$ back into (1) yields

$$c\lambda e^{\frac{x}{\lambda}} - \lambda \overset{!}{=} \lambda \cdot c e^{\frac{x}{\lambda}} \implies \lambda = 0,$$

which is a contradiction. Case 2: If $\lambda = 0$ were an eigenvalue, there would have to exists an eigenfunction $f \not\equiv 0$ such that

$$\int_0^x f(t)\,\mathrm{d}t = 0 \quad \forall x \in [0,1]$$

but this implies that $f|_{[0,1]} \equiv 0$, a contradiction.

The theorem of Gelfand-Mazur states that the $\sigma(T) \neq \varnothing$, thus $\sigma(T) = \{0\}$ must hold. Since $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$ is a disjoint union by definition, we only have to show that $0 \in \sigma_c(T)$: Since T is injective (which can easily by differentiation and the fundamental theorem of calculus), we just have to show that $T^{-1} : \operatorname{ran}(T) \to \mathcal{C}([0,1])$ is unbounded. One can easily see that $T^{-1}(f) = f'$ (cf. fundamental theorem of calculus), which is unbounded: Consider $f_n : [0,1] \to [0,1],\ t \to t^n$, then we have $\|t^n\|_\infty = 1$ but $\|f_n'\|_\infty = \max_{x\in[0,1]} |nt^{n-1}| = n \xrightarrow{n\to\infty} \infty$.

$\sigma_{\text{app}}(T)$: As a consequence of the open mapping theorem and lemma 3.8 from FA I we see that the approximate spectrum is contained in the spectrum, as every continuous invertible linear operator is bounded from below, so we just show that $0 \in \sigma_{\text{app}}(T)$.

Consider $f_n : [0,1] \to [0,1],\ t \mapsto t^n$. Then for all $n \in \mathbb{N}$ we have $\|f_n\|_\infty = 1$ and because all f_n are monotonically increasing

$$\|Tf_n\|_\infty = \max_{x\in[0,1]} \left|\int_0^x f_n(t)\,\mathrm{d}t\right| = \max_{x\in[0,1]} \frac{|x|^{n+1}}{n+1} = \frac{1}{n+1}$$

holds, thus

$$0 \leqslant \inf_{\|f\|_\infty = 1} \|Tf_n\|_\infty \leqslant \inf_{n\in\mathbb{N}} \frac{1}{n+1} = 0$$

follows. ◇

Corollary 1.2.10 (Numerical range of self-adjoint operator)
For self-adjoint $T \in L(\mathscr{H})$

$$\sigma(T) \subset \big[\inf(W(T)), \sup(W(T))\big] = \big[\min(\sigma(T)), \max(\sigma(T))\big]$$

holds. Furthermore, $\sigma(T) \subset [0,\infty)$ holds for positive T.

Proof. For self-adjoint $T \in L(\mathscr{H})$ we have that $\sigma(T) \subset \mathbb{R}$ and

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|. \tag{2}$$

□

Counterexample: (2) doesn't hold for $S := \left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right)$ or $T := \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ (90°-rotation).

Remark 1.2.11 (reverse direction of second statement from Corollary 1.2.10)
Self-adjoint operators with $\sigma(T) \subset [0,\infty)$ are positive: Since $\|T\| = r(T)$ we have $\sigma(T) \subset [0, \|T\|]$, implying $\sigma\left(T - \frac{1}{2}\|T\|\right) \subset \left[-\frac{1}{2}\|T\|, \frac{1}{2}\|T\|\right]$, implying $\left\|T - \frac{1}{2}\|T\|\right\| \leqslant \frac{1}{2}\|T\|$. Consequently, $\left|\left\langle\left(T - \frac{1}{2}\|T\|\right)x, x\right\rangle\right| \leqslant \frac{1}{2}\|T\|\|x\|^2$, implying $0 \leqslant \langle Tx, x\rangle \leqslant \|T\|\|x\|^2$. [4]

Remark 1.2.12 A positive operator is self-adjoint, which one can see by comparing real and imaginary terms in the polarisation identity [8]

$$\langle Tf, g\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \langle T(f + i^k g), f + i^k g\rangle.$$

In some sense all self-adjoint operators are 'not far from being positive', in the sense that $T + \|T\|I$ is positive. Positivity defines an order relation on self-adjoint operators.

Lemma 1.2.13 (cf. Kreyzig 9.2-4)
For $T = T^* \in L(\mathscr{H})$ we have $\sigma_r(T) = \varnothing$.

Proof. Suppose the exists a $\lambda \in \sigma_r(T) \subset \mathbb{R}$. Then $T_\lambda^{-1} := (T-\lambda)^{-1}$ inverse exists but its domain, $\operatorname{dom}(T_\lambda^{-1})$ is not dense in $\mathscr{H}$. Then there exists a $y \in \mathscr{H}\backslash\{0\}$ which is orthogonal to $\operatorname{dom}(T_\lambda^{-1}) = \operatorname{ran}(T_\lambda)$. As T_λ is self-adjointness (as $\lambda \in \mathbb{R}$),

$$0 = \langle T_\lambda x, y\rangle = \langle x, T_\lambda y\rangle$$

holds for all $x \in \mathscr{H}$. With $x = T_\lambda y$ we get $\|T_\lambda y\| = 0$, i.e. $Ty = \lambda y$, which is a contradictions to $\sigma_r(T) \cap \sigma_p(T) = \varnothing$. □

We will later need the following

Lemma 1.2.14 (HW 2-1 WiSe18)
$T \in L(\mathscr{H})$ is self-adjoint if and only if $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathscr{H}$.

Proof. " $\Longrightarrow$ ": For $x \in \mathscr{H}$ and self-adjoint $T \in L(\mathscr{H})$ it holds that

$$\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}.$$

" $\Longleftarrow$ ": Let $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathscr{H}$. For $\lambda \in \mathbb{C}$ consider

$$\langle T(x+\lambda y), x+\lambda y\rangle = \langle Tx, x\rangle + \overline{\lambda}\langle Tx, y\rangle + \lambda\langle Ty, x\rangle + |\lambda|^2\langle Ty, y\rangle$$

As $A := \langle T(x+\lambda y), x+\lambda y\rangle \in \mathbb{R}$ we have

$$A = \overline{A} = \langle Tx, x\rangle + \lambda\langle Tx, y\rangle + \overline{\lambda}\langle Ty, x\rangle + |\lambda|^2\langle Ty, y\rangle,$$

which implies

$$\lambda\langle Tx, y\rangle + \overline{\lambda}\langle Ty, x\rangle = \overline{\lambda}\langle Tx, y\rangle + \lambda\langle Ty, x\rangle$$

For $\lambda = 1$ and $\lambda = -i$ we obtain

$$\begin{aligned} \langle Tx, y\rangle + \langle Ty, x\rangle &= \langle Tx, y\rangle + \langle Ty, x\rangle \qquad (3)\\ -i\langle Tx, y\rangle + i\langle Ty, x\rangle &= i\langle Tx, y\rangle - i\langle Ty, x\rangle \end{aligned}$$

The second line implies (via multiplication with i)

$$\langle Tx, y\rangle - \langle Ty, x\rangle = \langle Ty, x\rangle - \langle Tx, y\rangle \qquad (4)$$

Adding (3) and (4) yields $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathscr{H}$. □

Lemma 1.2.15 (HW 2-1 WiSe18)
$\langle Tx, x\rangle = 0$ for all $x \in \mathscr{H}$ implies $T = 0$.

Proof. For $x, y \in \mathscr{H}$ we have, motivated by the polarisation identity,

$$\begin{aligned}
0 = \quad & \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \\
& + i\langle T(x+iy), x+iy\rangle - i\langle T(x-iy), x-iy\rangle \\
= \quad & \cancel{\langle Tx, x\rangle} + \langle Tx, y\rangle + \langle Ty, x\rangle \cancel{+\langle Ty, y\rangle} \\
& \cancel{-\langle Tx, x\rangle} + \langle Tx, y\rangle + \langle Ty, x\rangle \cancel{-\langle Ty, y\rangle} \\
& + \cancel{i\langle Tx, x\rangle} + i\langle Tx, iy\rangle + i\langle Tiy, x\rangle \cancel{+i\langle Tiy, iy\rangle} \\
& \cancel{-i\langle Tx, x\rangle} + i\langle Tx, iy\rangle + i\langle Tiy, x\rangle \cancel{-i\langle Tiy, iy\rangle} \\
&= 2\langle Tx, y\rangle + 2i\langle Tx, iy\rangle + 2\langle Ty, x\rangle + 2i\langle Tiy, x\rangle \\
&= 2\left(\langle Tx, y\rangle + i(-i)\langle Tx, y\rangle\right) + 2\left(\langle Ty, x\rangle + i \cdot i\langle Ty, x\rangle\right) \\
&= 4\langle Tx, y\rangle,
\end{aligned}$$

□

Corollary 1.2.16 (GÜ 2-2 Hannover WiSe11)
$T \in L(\mathscr{H})$ is unitary if and only if $\operatorname{ran}(T) = \mathscr{H}$ and $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in \mathscr{H}$ if and only if $\operatorname{ran}(T) = \mathscr{H}$ and $\|Tx\| = \|x\|$ for all $x \in \mathscr{H}$.

Proof. 1 $\iff$ 2 $\implies$ 3 are clear.

We show 3 $\implies$ 2. From $\|Tx\| = \|x\|$ we have that $\langle T^*Tx, x\rangle = \langle x, x\rangle$ holds for all $x \in \mathscr{H}$.

For $A = T^*T$ and all $x, y \in \mathscr{H}$

$$\begin{aligned}
4\langle x, y\rangle = \langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle \\
+ i\langle A(x+iy), x+iy\rangle - i\langle A(x-iy), x-iy\rangle
\end{aligned}$$

holds, as the RHS is also equal to

$$\begin{aligned}
&\cancel{\|x\|^2} + \langle x, y\rangle + \cancel{\langle y, x\rangle} + \cancel{\|y\|^2} \cancel{-\|x\|^2} + \langle x, y\rangle + \cancel{\langle y, x\rangle} \cancel{-\|y\|^2} \\
&+ \cancel{i\|x\|^2} + \langle x, y\rangle \cancel{-\langle y, x\rangle} \cancel{-i\|y\|^2} \cancel{-i\|x\|^2} + \langle x, y\rangle \cancel{-\langle y, x\rangle} + \cancel{i\|y\|^2}.
\end{aligned}$$

By the proof of lemma 1.2.15 the RHS is also equal to $4\langle Ax, y\rangle$, showing that $A - I = 0$, i.e. $T^*T = I$ again by that lemma.

As the RHS is invariant under complex conjugation in the right sense, also $TT^* = I$ follows. □

Counterexample 1.2.17 ($\mathscr{r}(T) = r(T)$ for non-self-adjoint T [8])
Consider $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, which is not self-adjoint, as $T^* = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and $TT^* = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right) \neq \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) = T^*T$. Now, $T^n = \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$ and therefore

$$r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leqslant \lim_{n\to\infty} (\sqrt{n^2+1})^{\frac{1}{n}} = 1.$$

But

$$\mathscr{r}(T) = \sup_{x^2+y^2=1} x(x+y) + y^2 = \sup_{y>0} 1 - \sqrt{1-y^2}y = \frac{3}{2}. \qquad \diamond$$

1.3 Continuous Functional Calculus

17.10.19

For an operator $T \in L(\mathscr{H})$ and a polynomial $p(z) := \sum_{k=0}^{n} a_k z^k$ we would like to have $p(T) = \sum_{k=0}^{n} a_k T^k$.

This is called polynomial functional calculus, which is a homomorphism from the polynomial ring to the ring $\mathbb{C}^{n\times n}$ if $\mathscr{H} = \mathbb{C}^n$.

For $\mathscr{H} = \mathbb{C}^n$ and self-adjoint $T \in L(\mathscr{H})$ we consider the diagonalization

$$T = U^{-1}DU, \quad U \text{ unitary}, \, D := \operatorname{diag}((d_k)_{k=1}^{n}) \text{ contains eigenvalues},$$

which is equivalent to **todo: should u $= \mu$??**

$$T = \sum_{k=1}^{\ell} d_k E_k, \tag{5}$$

where μ_k are the distinct eigenvalues and E_k the corresponding projections onto the eigenspace of μ_k and we have $((U^{-1}DU)x)_k = d_k x_k$.

If f is an entire function with the MACLAURIN series $f(z) := \sum_{k=0}^{\infty} a_k z^k$

A function $f : \mathbb{C} \to \mathbb{C}$ is called entire if it is holomorphic everywhere.

$$\begin{aligned} f(T) &= \sum_{k=0}^{\infty} a_k T^k = \sum_{k=0}^{\infty} a_k (U^{-1}DU)^k = \sum_{k=0}^{\infty} a_k U^{-1} D^k U \\ &= U^{-1}\left(\sum_{k=0}^{\infty} D^k\right) U = U^{-1} f(D) U \end{aligned}$$

holds (at least formally), where $f(D) = \operatorname{diag}((f(d_k))_{k=1}^{n})$.

We see that to define $f(T)$, f only needs to be defined on $\sigma(T)$ and it must hold that "$f(\sigma(T)) = \sigma(f(T))$", which is called the Spectral Mapping Theorem. Now let $f : \sigma(T) \to \mathbb{C}$ and define $f(T) := ((U^{-1}f(D)U)x)_i = f(d_i)x_i$.

For compact self-adjoint operators $T \in L(\mathscr{H})$ we can write

$$T = \sum_{j=0}^{\infty} \mu_j E_j,$$

where $(\mu_j)_{j=0}^{\infty}$ are the distinct spectral values of T with $\mu_0 = 0$ and $(\mu_j)_{j=0}^{\infty}$ denotes the eigenvalues do T and $(E_j)_{j=0}^{\infty}$ the corresponding orthogonal projections onto the eigenspaces. Then, for a continuous function $f : \sigma(T) \to \mathcal{C}$ we define $f(T) = \sum_{j=0}^{\infty} f(\mu_j)E_j$. Note that for $j \geqslant 1$ the function $f_j : \sigma(T) = \{0\} \cup (\mu_j)_{j=1}^{\infty} \to \mathbb{C}$ defined by $f_j(\mu_k) = \delta_{j,k}$ is continuous. Now we have $f_j(T) = E_j$, in other words: the recovery of the orthogonal projection onto the eigenspaces is possible.

Lemma 1.3.1 (Spectral Mapping Theorem for Polynomials)
Let $T \in L(X)$, X a complex BANACH *space and $f \in \mathbb{C}[x]$. Then*

$$\sigma(f(T)) = f(\sigma(T)) \quad \text{and} \quad \sigma(T^{-1}) = \sigma(T)^{-1} \text{ if } 0 \in \rho(T)$$

holds.

Proof. 1 Without loss of generality let f be non-constant.
"$\subset$": Suppose $\mu \in \sigma(f(T))$. We can then write

$$g(x) := f(x) - \mu = c \prod_{k=1}^{n} (x - x_k),$$

for some $x_k \in \mathbb{C}$ for $k \in \{1, \dots, n\}$ and also

$$g(T) = f(T) - \mu \,\mathrm{id} = c \prod_{k=1}^{n} (T - x_k \,\mathrm{id}).$$

If $x_k \in \rho(T)$ for all $k \in \{1, \dots, n\}$, $g(T)$ would be invertible, which is a contradiction to $\mu \in \sigma(f(T))$. Therefore there exists some $i \in \{1, \dots, n\}$ such that $x_i \in \sigma(T)$, meaning $T - x_i \,\mathrm{id}$ is not invertible, implying $f(x_i) - \mu = c \prod_{k=1}^{n} (x_i - x_k) = 0$, implying $f(x_i) = \mu$.

"$\supset$": Let $\lambda \in \sigma(T)$. Since λ is a root of $\Phi(z) := f(z) - f(\lambda)$, we can write

$$\Phi(z) = (z - \lambda)\Psi(z)$$

for some $\Psi \in \mathbb{C}[z]$. Thus

$$\Phi(T) = f(T) - f(\lambda) \,\mathrm{id} = (T - \lambda \,\mathrm{id})\Psi(T) = \Psi(T)(T - \lambda \,\mathrm{id})$$

holds, implying that $f(\lambda) \notin \rho(f(T))$: Assuming $f(T) - f(\lambda)I$ were invertible it would follow that $f(T) - f(\lambda)I$ is bijective. By the equations above we see that $(T - \lambda I)$ is bijective, as ($f \circ g$ surjective $\implies$ g surjective and $f \circ g$ injective $\implies$ f injective). Then by the bounded inverse Theorem, $(T - \lambda I)$ has a bounded inverse, a contradiction. If $f(\lambda)$ were in the resolvent set of $f(T)$ we would have that $(f(T) - f(\lambda) \,\mathrm{id})^{-1}\Psi(T)$ were a bounded inverse for $T - \lambda \,\mathrm{id}$. Therefore, $f(\lambda) \in \sigma(f(T))$.

2 "$\subset$": Let $\lambda \in \mathbb{C} \backslash \sigma(T)^{-1}$, i.e. $\lambda^{-1} \in \rho(T)$. Thus there exists an operator $S \in L(\mathscr{H})$ such that

$$(T - \lambda^{-1} \,\mathrm{id})S = S(T - \lambda^{-1}) = \mathrm{id},$$

which we can "refactor" as (from the equation we get that S and T commute)

$$(T^{-1} - \lambda \,\mathrm{id})(-\lambda^{-1} TS) = (-\lambda^{-1} TS)(T^{-1} - \lambda \,\mathrm{id}) = \mathrm{id},$$

yielding that $T^{-1} - \lambda \,\mathrm{id}$ is invertible, and $\lambda \in \rho(T^{-1})$.

"$\supset$": Let $\lambda \in \rho(T^{-1})$, the $T^{-1} - \lambda \,\mathrm{id}$ is invertible. Thus there exists an operator $S \in L(\mathscr{H})$ such that

$$(T - \lambda \,\mathrm{id})S = S(T^{-1} - \lambda \,\mathrm{id}) = \mathrm{id}\,.$$

We can factor those products differently to obtain

$$(T - \lambda^{-1} \,\mathrm{id})(-\lambda T^{-1} S) = (-\lambda T^{-1} S)(T - \lambda^{-1} \,\mathrm{id}) = \mathrm{id}\,.$$

Therefore, $T - \lambda^{-1} \,\mathrm{id}$ is invertible and $\lambda^{-1} \in \rho(T)$, i.e. $\lambda \in (\rho(T))^{-1}$.

2 **(Alternative von Mones.)** Let $0 \in \rho(T)$, i.e. $T^{-1} \in L(\mathscr{H})$. Then

$$T^{-1}(T - \lambda I) = I - \lambda T^{-1} = \lambda(\lambda^{-1} - T^{-1}) \tag{6}$$

holds.

If $\lambda \in \sigma(T)$, then $\lambda \neq 0$ and $T^{-1}(T - \lambda)$ is not bijective, implying $\lambda^{-1} - T^{-1}$ is not bijective.

If $\lambda^{-1} \in \sigma(T^{-1})$, then $\lambda \in \sigma(T)$ by (6). $\square$

This enables us to give an alternative proof of lemma 1.2.6, courtesy of this answer from Math.SE.

Proof. We have that $z \in \sigma(U)$ implies $\overline{z} \in \sigma(U^*) = \sigma(U^{-1})$. As U is invertible ($\implies 0 \in \rho(U)$), by lemma 1.3.1 we have that $\overline{z} \in \sigma(U^{-1})$ implies that $(\overline{z})^{-1} \in \sigma(U)$.

We know that $\sigma(U) \subset \{z \in \mathbb{C} : |z| \leqslant 1\}$ as $\|U\| = 1$. From the implication $z \in \sigma(U) \implies (\overline{z})^{-1} \in \sigma(U)$ we can conclude the statement as $|z| = 1$ implies $|(\overline{z})^{-1}| = 1$ and $|z| \in (0,1)$ implies $|(\overline{z})^{-1}| > 1$. □

Theorem 1.3.1: Continuous Functional Calculus

For all self-adjoint $T \in L(\mathscr{H})$ there exists an unique

$$\Phi_T : \mathcal{C}(\sigma(T)) \to L(\mathscr{H})$$

such that

1. $\Phi_T(\mathrm{id}) = T$, $\Phi_T(\mathbb{1}) = \mathrm{id}$,
2. Φ_T is an involutive homomorphism of algebras, i.e. Φ_T is
   - a. linear and multiplicative: $\Phi_T(fg) = \Phi_T(f) \circ \Phi_T(g)$
   - b. involutive: $(\Phi_T(f))^* = \Phi_T(\overline{f})$
3. Φ_T is continuous,

called continuous functional calculus of T. We denote $f(T) := \Phi_T(f)$ for $f \in \mathcal{C}(\sigma(T))$.

An algebra over a field is a vector space equipped with a bilinear product.

involutive

continuous functional calculus

A functional calculus therefore describes the assignment of operators from commutative algebras to functions defined on their spectra.

Proof. Uniqueness: By multiplicativity and 2b we have $\Phi_T(z \mapsto z^n) = T^n$, hence Φ_T is unique on all polynomials. Furthermore, $\sigma(T) \subset [m(T), M(T)]$ is compact and the polynomials dense in $\mathcal{C}(\sigma(T))$ by the Stone-Weierstrass-Theorem. Due to continuity, Φ_T is unique on $\mathcal{C}(\sigma(T))$.

Existence: Set $\Phi_T(f) = \sum_{k=0}^n a_k T^k$ for a polynomial $f(z) = \sum_{k=0}^n a_k z^k$. If we show continuity of Φ_T on polynomials, there would be an unique extension (cf. above) to $\mathcal{C}(\sigma(T))$, which we denote by Φ_T again.

By the spectral mapping theorem for polynomials f we obtain

$$\begin{aligned}
\|\Phi_T(f)\|^2 &\overset{1.2.5}{\underset{3}{=}} \|\Phi_T(f)^*\Phi(f)\| \overset{1.3.1}{\underset{2}{=}} \|\Phi_T(\overline{f}f)\| \\
&= \sup_{\lambda \in \sigma\left(\Phi_T(\overline{f}f)\right)} |\lambda| = \sup_{\lambda \in \sigma(T)} |(\overline{f}f)(\lambda)| \\
&= \sup_{\lambda \in \sigma(T)} |f(\lambda)|^2 = \|f\|_\infty^2
\end{aligned}$$

Now, all the required properties are proven by approximation: Assume that $f_n \to f$ and $g_n \to g$ for polynomials f_n and g_n and $f, g \in \mathcal{C}(\sigma(T))$. Then it holds that

$$\Phi_T(fg) \to \Phi_T(f_n g_n) = \Phi(f_n) \circ \Phi(g_n) \to \Phi(f) \circ \Phi(g). \qquad \square$$

Theorem 1.3.2: Properties of the CFC

Let $T = T^* \in L(\mathcal{H})$ and $f \mapsto f(T)$ be the continuous functional calculus for $f \in \mathcal{C}(\sigma(T))$. Then

1. $\|f(T)\| = \|f\|_\infty := \sup_{\lambda \in \sigma(T)} |f(\lambda)|$.
2. If $f|_{\sigma(T)} \geqslant 0$, then $f(T)$ is positive.
3. If $Tx = \lambda$ for some $x \in \mathcal{H}$, then $f(T)x = f(\lambda)x$.
4. The spectral mapping theorem holds for all $f \in \mathcal{C}(\sigma(T))$.
5. $\{f(T) : f \in \mathcal{C}(\sigma(T))\}$ is a commutative Banach algebra of operators.
6. All $f(T)$ are normal; if f is real, then $f(T)$ is self-adjoint.

Proof. 1 This was shown for polynomials in the above proof and is shown by approximation in the general case (cf. above).

2 Let $f \geqslant 0$ and $g \in \mathcal{C}(\sigma(T))$ with $g^2 = f$ and $g \geqslant 0$. Then

$$\begin{aligned} \langle f(T)x, x\rangle &= \langle g(T)x, g(T)^* x\rangle = \langle g(T)x, \overline{g}(T)x\rangle \\ &= \langle g(T)x, g(T)x\rangle = \|g(T)x\|^2 \geqslant 0. \end{aligned}$$

3 Follows through approximation like above.

4 "$\subset$": Let $\mu \notin f(\sigma(T))$. Then $g(x) := \frac{1}{f(x)-\mu} \in \mathcal{C}(\sigma(T))$ and $g(f-\mu) = (f-\mu)g = 1$. Hence we get

$$g(T)(f(T) - \mu\,\mathrm{id}) = (f(T) - \mu\,\mathrm{id})g(T) = \mathrm{id},$$

hence $\mu \in \rho(f(T))$.

"$\supset$": Let $\mu = f(\lambda)$ for some $\lambda \in \sigma(T)$ and choose polynomials f_n with $\|f_n - f\|_\infty \leqslant \frac{1}{n}$. Then

$$|f(\lambda) - f_n(\lambda)| \leqslant \frac{1}{n} \quad \text{and} \quad \|f(T) - f_n(T)\| \leqslant \frac{1}{n}.$$

We know that $f_n(\lambda) \in \sigma(f_n)T))$, i.e. there exists $x_n \in \mathcal{H}$ with $\|x_n\| = 1$ and $\|(f_n(T) - f_n(\lambda)\,\mathrm{id})x_n\| \leqslant \frac{1}{n}$. Thus,

$$\begin{aligned} \|(f(T) - \mu\,\mathrm{id})x_n\| &\leqslant \|(f(T) - f_n(T))x_n\| + \|(f_n(T) - f_n(\lambda)\,\mathrm{id})x_n\| \\ &\quad + \|(f_n(\lambda) - \mu\,\mathrm{id})x_n\| \\ &\leqslant \frac{1}{n} + \frac{1}{n} + \frac{1}{n} = \frac{3}{n} \end{aligned}$$

holds, implying that $(f(T) - \mu\,\mathrm{id})$ is not boundedly invertible, i.e. $\mu \in \sigma(f(T))$.

5 is clear and 6 follows from $f(T)^* = \overline{f}(T)$ and $f(T)^* f(T) = \overline{f} f(T) = f\overline{f}(T) = f(T)f(T)^*$. □

Applications of the continuous functional calculus

We can now give a completely different proof of lemma 1.1.4:

Proof. As $T = T^*$ we have $\sigma(T) \subset \mathbb{R}$, implying $\lambda \notin \sigma(T)$. Therefore,

$$f : \sigma(T) \to \mathbb{C},\ x \mapsto \frac{1}{x - \lambda}$$

is continuous and by the continuous functional calculus we have

$$\begin{aligned}
\|(T-\lambda I)^{-1}\| &= \|f(T)\| = \sup_{\mu\in\sigma(T)} |f(\mu)| = \sup_{\mu\in\sigma(T)} \frac{1}{|\mu-\lambda|} \\
&= \frac{1}{\inf_{\mu\in\sigma(T)} |\mu-\lambda|} = \frac{1}{\inf_{\mu\in\sigma(T)} \sqrt{(\mu-\Re(\lambda))^2+\Im(\lambda)^2}} \\
&\leqslant \frac{1}{\sqrt{\Im(\lambda)^2}} = \frac{1}{|\Im(\lambda)|},
\end{aligned}$$

where equality holds if $\Re(\lambda)\in\sigma(T)$. □

Lemma 1.3.2 (HW 1-3 WiSe00)
For $T=T^*\in L(H)$ there exist two positive self-adjoint operators $T_\pm$ such that

1. *$T=T_+-T_-$, $T\pm T_\mp\geqslant 0$ and $T_+T_-=T_-T_+=0$,*
2. *If T commutes with $B\in L(H)$, so do T_+ and T_-,*
3. *If T is compact, so are $T_\pm$.*

Proof. On $\sigma(T)$ define the continuous functions $f_+(x):=\max(x,0)=\frac{x+|x|}{2}$ and $f_-(x):=\max(-x,0)$. We show that $T_\pm:=f_\pm(T)$ fulfil the desired properties.

1. We have

$$T_+-T_- = \frac{T+|T|}{2}-\frac{|T|-T}{2} = T, \quad T+T_- = T+\frac{|T|-T}{2} = T_+,$$

and similarly $T-T_+=T_-$. It therefore suffices to show $T_\pm\geqslant 0$ which follows from theorem 1.3.2 (2) and that $f_+,f_-\geqslant 0$. By the properties of the continuous functional calculus we have

$$T_+\cdot T_- = f_+(T)\cdot f_-(T) = (f_+\cdot f_-)(T) = 0,$$

as $4(f_+\cdot f_-)(x) = x|x|-x^2+|x|^2-|x|x = 0$. Analogously, $T_-T_+=0$ follows.

2. If $TB=BT$, then also $f(T)B=Bf(T)$ for all polynomials f on $\sigma(T)$. By continuity of the functional calculus ((3)) we also have $T_\pm B = BT_\pm$.

3. The limit of compact operators is compact. Therefore this follows as above. □

Example 1.3.3 (Differential equation in $L(\mathscr{H})$ (HW 1-4 WiSe00))
Let $T=T^*\in L(\mathscr{H})$ and $X:[0,1]\to L(H)$ fulfil $X(0)=I$ and $\dot X(t)=TX(t)$ for $t\in(0,1)$, where $\dot X(t):=\lim_{t_0\to t}\frac{X(t)-X(t_0)}{t-t_0}$. We show that the only solution is $X(t)=e^{tT}$.

Using the continuous functional calculus we can define $X(t):=e^{tT}$ as T is self-adjoint and f is continuous. We have

$$\frac{X(t)-X(s)}{t-s} = \sum_{k=0}^{\infty}\frac{t^k-s^k}{t-s}\frac{T^k}{k!} = \sum_{k=1}^{\infty}\frac{t^k-s^k}{t-s}\frac{T^k}{k!}.$$

Now using that $\lim_{y\to x} \frac{x^n-y^n}{x-y} = n \cdot x^{n-1}$ for $n \in \mathbb{N}$, we have

The limit can be seen via $\frac{a^n-b^n}{a-b} = \sum_{k=0}^n a^{n-k} b^{k-1}$.

$$\begin{aligned}\dot{X}(t) &= \lim_{s\to t} \sum_{k=1}^{\infty} \frac{t^k - s^k}{t-s} \frac{T^k}{k!} = \sum_{k=1}^{\infty} \left(\lim_{s\to t} \frac{t^k-s^k}{t-s} \right) \frac{T^k}{k!} \\ &= \sum_{k=1}^{\infty} \left(k \cdot t^{k-1}\right) \frac{T^k}{k!} = \sum_{k=1}^{\infty} t^{k-1} \frac{T^k}{(k-1)!} = T \cdot e^{tT}.\end{aligned}$$ $\diamond$

Example 1.3.4 (Multiplication operator II (HW 1-5 WiSe00))
Recall the operator $T := T_2$ from example 1.2.8 with $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$.

In order to be able to explain $\log(T)$, $|T|$, T^{-1}, $\sqrt{T}$ or $\sin(T)$, the corresponding functions $x \mapsto \log(x)$, $x \mapsto |x|$ etc. have to be continuous on $\sigma(T) = \overline{(a_n)_{n\in\mathbb{N}}}$.

For the logarithm we need $\sigma(T) \subset (0,\infty)$, for the inverse inverse also $\sigma(T) \subset (-\infty, 0)$ is possible, for the square root we need $\sigma(T) \subset [0,\infty)$. We don't need any further conditions on the sequence to define $|T|$ or $\sin(T)$. $\diamond$

If we believe that the spectral theorem can also be extended to normal operators (cf. e.g. Werner), we can give an elegant proof of lemma 1.2.6, courtesy of a comment on this question from Math.SE.

Proof. Define $f(z) := \overline{z}z$. Then we have

$$f(\sigma(U)) = \sigma(f(U)) = \sigma(I) = \{1\}.$$

This implies that for all $\lambda \in \sigma(U)$ we have $1 = \overline{\lambda}\lambda = |\lambda|^2$. $\square$

1.4 Measurable Functional Calculus

We can write $T = T^* \in L(\mathbb{C}^n)$ as $T = \sum_{k=1}^n \lambda_k E_k$, where λ_k are the (mutually distinct) eigenvalues of T and E_k are the orthogonal projections onto $\ker(T - \lambda_k)$. As $\sigma(T) = (\lambda_k)_{k=1}^n$ is finite, every $f : \sigma(T) \to \mathbb{C}$ is continuous: $f \in \mathcal{C}(\sigma(T))$. By the continuous functional calculus we obtain

$$f(T) = \sum_{k=1}^{n} f(\lambda_k) E_k,$$

as we foreshadowed at the beginning of the last subsection.

If $T = T^* \in L(\mathscr{H})$ is not compact, $\sigma(T)$ is not necessarily countable anymore. Recall the spectral decomposition for compact self-adjoint operators from FA I is

$$Tx = \sum_{k=0}^{\infty} \lambda_k \langle x, x_k \rangle x_k,$$

where $E_k x := \langle x, x_k \rangle x_k$ is the projection of x onto $\operatorname{span}(x_k)$.

Given T, an interesting question is how to extract the projections E_k from T. This can be done by using $f_k(\lambda_j) = \delta_{j,k}$, i.e. $f_k = \mathbb{1}_{\{\lambda_k\}}$:

$$f_k(T)x = \sum_{n=0}^{\infty} f_k(\lambda_n) \langle x, x_n \rangle x_n = \langle x, x_k \rangle x_k = E_k x.$$

Unfortunately, $\{0, 1\}$-valued functions such as $\delta_{j,k}$ need no be continuous, take for example $\dim(\mathscr{H}) = \infty$ and a λ, which is not an isolated point of $\sigma(T)$ and $f := \mathbb{1}_{\{\lambda\}}$.

Therefore we need the Borel measurable functional calculus, whose goal it is to generalize $T = \sum_{k=0}^{\infty} \lambda_k E_k$ to non-compact operators in the sense of

$$T = \int_{\sigma(T)} \lambda \,\mathrm{d}E \quad \text{and} \quad f(T) = \int_{\sigma(T)} f(\lambda) \,\mathrm{d}E,$$

where $f : \sigma(T) \to \mathbb{C}$ is a Borel-measurable function.

Projections

A projection is a linear map $P : E \to E$ on a vector space E satisfying $P^2 = P$. P is a projection onto $Y \subset E$ if $\operatorname{ran}(P) = Y$ and $P|_Y = \operatorname{id}|_Y$. 28.10.2019

Lemma 1.4.1 (Algebraic properties of projections)

1. *If P is a projection, $\operatorname{id} - P$ is a projection with*

$$\ker(\operatorname{id} - P) = \operatorname{ran}(P) \quad \textit{and} \quad \operatorname{ran}(\operatorname{id} - P) = \ker(P)$$

Thus $E = \ker(P) \oplus \operatorname{ran}(P)$ (direct sum) holds.

2. *For all subspaces $E \subset Y$ there exists a projection onto Y.*

Proof. 1 We have $(\operatorname{id} - P)^2 = I - 2P + P^2 = I - P$. Let $y \in \operatorname{ran}(P)$, then there exists a $x \in E$ such that $Px = y$. Applying P to the equality yields $P^2 x = Px = Py$. Substituting back yields $Py = y$, i.e. $(\operatorname{id} - P)y = 0$.

For $x \in E$ we have $x = x - Px + Px$ with $x - Px \in \operatorname{ran}(I - P) = \ker(P)$ and $Px \in \operatorname{ran}(P)$.

(2) ZORNs lemma. □

We now investigate continuity of projections on a normed space E.

Lemma 1.4.2 (Sufficient conditions for continuity of projections) *Let $P \in L(E)$ be a projection. Then $\ker(P), \operatorname{ran}(P) \subset E$ are closed and $\|P\| \geqslant 1$ or $P \equiv 0$.*

Proof. If suffices to show that $\ker(P)$ is closed as $\operatorname{ran}(P) = \ker(I - P)$. Let $(x_n)_{n\in\mathbb{N}} \subset \ker(P)$ converge to $x \in E$. We have $x \in \ker(P)$ as by the continuity $(\star)$ of P we have $Px = P \lim_{n\to\infty} x_n \overset{(\star)}{=} \lim_{n\to\infty} Px_n = 0$. For $P \neq 0$ we have $\|P\| = \|P^2\| \leqslant \|P\|^2$, implying $1 \leqslant \|P\|$. □

Remark 1.4.3 There exist normed spaces E with closed subspaces $Y \subset E$ such that there is no continuous projection onto Y. (Ggbsp. $c_0 \subset \ell^\infty$ **TODO**??) In HILBERT spaces there always exists such an projection and one unique orthogonal projection (i.e. $\operatorname{ran}(P)^\perp = \ker(P)$).

Lemma 1.4.4 (Characterisation of orthogonal projections) *Let $P \in L(\mathcal{H})\backslash\{\mathbb{0}\}$ be a projection. Then the following are equivalent.*

(1) *P is orthogonal:* $\operatorname{ran}(P)^\perp = \ker(P)$ (2) $\|P\| = 1$.

(3) *P is self-adjoint.* (4) *P is normal.* (5) *P is positive.*

Proof. "(1) $\implies$ (2)": By lemma 1.4.2 we only need to show $\|P\| \leqslant 1$. Let $x = Py + z \in \mathcal{H}$ with $Py \in \operatorname{ran}(P)$ and $z \in \ker(P)$. As $Py \perp z$,

$$\|Px\|^2 = \|Py\|^2 \leqslant \|Py\|^2 + \|z\|^2 = \|x\|^2.$$

"(2) $\implies$ (1)": For $x \in \ker(P)$, $Py \in \operatorname{ran}(P)$ and all $\lambda \in \mathbb{C}$

$$\|\lambda Py\|^2 = \|P(\lambda Py + x)\|^2 \overset{(2)}{\leqslant} \|\lambda Py + x\|^2 = \|x\|^2 + 2\Re(\lambda)\langle x, y\rangle + \|\lambda Py\|^2$$

holds, implying $-2\Re(\lambda)\langle x, Py\rangle \leqslant \|x\|^2$. Thus $\langle x, Py\rangle = 0$ follows.

"(3) $\implies$ (4)": clear.

"(1) $\implies$ (3), (5)": For $x = Py + z \in \mathcal{H}$ with $Py \in \operatorname{ran}(P)$ and $z \in \ker(P)$

$$\langle Px, x\rangle = \langle Py, Py + z\rangle = \langle Py, Py\rangle + \langle Py, z\rangle = \|Py\|^2 \in \mathbb{R}_{\geqslant 0}\,.$$

"(3) $\implies$ (1)": For $y \in \mathcal{H}$ and $z \in \ker(P)$, $\langle Py, z\rangle = \langle y, Pz\rangle = 0$ holds.

"(5) $\implies$ (3)": If $\langle Px, x\rangle \in \mathbb{R}$, P is self-adjoint by lemma 1.2.14.

"(4) $\implies$ (1)": As P is normal, $\ker(P) = \ker(P^*)$ holds by lemma 1.2.5. Thus $\langle x, Py\rangle = \langle P^*x, y\rangle = 0$ holds for all $x \in \ker(P)$ and $y \in \mathcal{H}$.

"(1) $\implies$ (4)": We have $\ker(P) = \operatorname{ran}(P)^\perp = \ker(P^*)$. Let $x = y + z \in \ker(P) \oplus \operatorname{ran}(P) = \mathcal{H}$. Then $P^*Px = P^*z = Px = Pz = PP^*x$ holds, implying $P^*P = P = P^*P$. □

Lemma 1.4.5 (Properties of orthogonal projections)
For orthogonal projections $P_1, P_2 \in L(\mathscr{H})$ onto subspaces $S_1, S_2 \subset \mathscr{H}$, the following are equivalent:

1 $\operatorname{ran}(P_1) \subset \operatorname{ran}(P_2)$ 2 $\ker(P_2) \subset \ker(P_1)$

3 $P_1P_2 = P_2P_1 = P_1$ 4 $P_2 - P_1$ *is positive.*

Proof. "1 $\iff$ 2": The claim follows from $\ker(P_i) = \operatorname{ran}(P_i)^\perp$ for $i \in \{1, 2\}$ and $A \subset B \iff B^\perp \subset A^\perp$.

"1, 2 $\implies$ 3": Let $x = y + P_1z \in \ker(P_1) \oplus \operatorname{ran}(P_1) = \mathscr{H}$. By 1 there exists a $a \in \mathscr{H}$ such that $P_1y = P_2a$ thus

$$P_1x = P_1^2z = P_1z = P_2a \quad \text{and thus} \quad P_2P_1x = P_2P_2a = P_2a,$$

implying $P_1 = P_2P_1$.

Let $x = y + P_2z \in \ker(P_2) \oplus \operatorname{ran}(P_2) = \mathscr{H}$. As $\ker(P_2) \subset \ker(P_1)$ we have

$$P_1x = P_1P_2z \quad \text{and thus} \quad P_1P_2x = P_1P_2P_2z = P_1P_2z,$$

implying $P_1 = P_1P_2$.

"3 $\implies$ 1": We have $\operatorname{ran}(P_1) \overset{3}{=} \operatorname{ran}(P_2P_1) \subset \operatorname{ran}(P_2)$.

"3 $\implies$ 4": As P_1 and P_2 are orthogonal projections, they are positive by lemma 1.4.4 ($\star$) and self-adjoint ($\ddagger$). For $x = y + P_1z \in \ker(P_1) \oplus \operatorname{ran}(P_1)$

$$(P_2 - P_1)x = P_2y + P_2P_1z - P_1y - P_1^2z \overset{3}{=} P_2y + \cancel{P_1z} - 0 - \cancel{P_1z} = P_2y.$$

holds. Thus we have

$$\begin{aligned}\langle (P_2 - P_1)x, x\rangle &= \langle P_2y, y + P_1z\rangle = \langle P_2y, y\rangle + \langle P_2y, P_1z\rangle \\ &\overset{(\star)}{\geqslant} \langle P_2y, P_1z\rangle \overset{(\ddagger)}{=} \langle P_1P_2y, z\rangle \overset{3}{=} \langle P_1y, z\rangle = 0.\end{aligned}$$

"4 $\implies$ 2": For $x \in \ker(P_2)$

$$0 \leqslant \|P_1x\|^2 = \langle P_1x, P_1x\rangle \overset{(\ddagger)}{=} \langle P_1^2x, x\rangle = \langle P_1x, x\rangle \overset{4}{\leqslant} \langle P_2x, x\rangle = 0$$

holds, implying $x \in \ker(P_1)$. ☐

Lemma 1.4.6 (Self-adjoint operators and orthogonal projections)
$T = T^* \in L(\mathscr{H})$ *is an orthogonal projection iff* $\sigma(T) = \{0, 1\}$.

Proof. "$\implies$" Since $\mathscr{H} = \ker(P) \oplus \operatorname{ran}(P)$ and $Px = x$ for $x \in \operatorname{ran}(P)$ and $Px = 0$ for $x \in \ker(P)$ we have $\sigma(T) \subset \{0, 1\}$.

Notice that for $\implies$ we don't use the orthogonality or self-adjointness, $\sigma(T) = \{0, 1\}$ holds for all projections.

"$\impliedby$": We have $T \neq 0$ since otherwise $\sigma(T) = \{0\}$. We show $T^2 = T$. Since $T = T^*$ we can use the continuous functional calculus to construct $f_{1,2} : \sigma(T) \to \mathbb{C}$ with $z \mapsto z$ and $z \mapsto z^2$. Then we have $f_1(T) = T$ and $f_2(T) = T^2$. Since f_1 and f_2 coincide on the spectrum we have $f_1(T) = f_2(T)$. ☐

Remark 1.4.7 (Resolvent operator of a projection)
For $|\lambda| > 1$ we have

$$\begin{aligned}(\lambda - P)^{-1} &= \sum_{k=0}^{\infty} \frac{P^k}{\lambda^{k+1}} = \frac{1}{\lambda} + \sum_{k=1}^{\infty} \frac{P}{\lambda^{k+1}} \\ &= \frac{1}{\lambda} + P \sum_{k=2}^{\infty} \left(\frac{1}{\lambda}\right)^k = \frac{1}{\lambda} + \frac{P}{\lambda(\lambda-1)}.\end{aligned}$$

Corollary 1.4.8 (Point spectrum contains isolated spectral values)
Let $T = T^* \in L(\mathscr{H})$ and $\varepsilon > 0$ such that $\sigma(T) \cap (\lambda - \varepsilon, \lambda + \varepsilon) = \{\lambda\}$ ($\lambda \in \sigma(T)$ is called isolated). Then $\lambda \in \sigma_p(T)$ holds.

Proof. Consider $f : \sigma(T) \to \mathbb{C}$ given by $f := \mathbb{1}_{\{\lambda\}} \in \mathcal{C}(\sigma(T))$. Hence $f(T)$ is well defined. By the spectral mapping theorem we have $\sigma(f(T)) = f(\sigma(T)) = \{0, 1\}$. By the above lemma we know that **???**. Let $g(x) := (x - \lambda) f(x)$. Then we have $g|_{\sigma(T)} \equiv 0$. By the functional calculus we have $(T - \lambda \,\mathrm{id}) f(T) = 0$. This implies $\{0\} \neq \mathrm{ran}(f(T)) \subset \ker(T - \lambda \,\mathrm{id})$, where the first non-equality follows from $1 \in \sigma(f(T))$. □

RADON measures and duality

Let (K, τ) be a compact topological space and $\mathcal{B}$ the BOREL σ-algebra.

DEFINITION 1.4.9 ((COMPLEX) SIGNED RADON MEASURE)
A finite measure μ on $(K, \mathcal{B})$ is called RADON measure if

$$\mu(B) = \inf_{\substack{B \subset G \\ G \text{ open}}} \mu(G) \quad \forall B \in \mathcal{B} \qquad \text{(outer regularity)}$$

$$\mu(G) = \inf_{\substack{T \subset G \\ G \text{ comp.}}} \mu(T) \quad \forall G \text{ open} \qquad \text{(inner regularity)}$$

hold. A signed RADON measure μ is a σ-additive mapping on $\mathcal{B}$ with $\mu = \mu^+ - \mu^-$, where μ^+ and μ^- are RADON measures. RADON measure
A complex RADON measure μ has the form $\mu = \mu_1 + i\mu_2$, where μ_1 and μ_2 are signed RADON measures.

Remark 1.4.10 (HAHN-JORDAN decomposition)
This representation is not unique: If ν is another **(bounded? isn't it finite?)** positive (why?) we have $\mu = \mu^+ - \mu^- = (\mu^+ - \nu) - (\mu^- - \nu)$. But there always exists a unique "minimal", so called HAHN-JORDAN decomposition, in which μ^+ and μ^- are singular to each other.

Example 1.4.11 The DIRAC and LEBESGUE are RADON measures. ◇

DEFINITION 1.4.12 (TOTAL VARIATION (NORM) ON $\mathcal{M}(K)$)
By $|\mu| := \mu^+ + \mu^-$ we denote the total variation of μ, whose minimal decomposition is $\mu = \mu^+ - \mu^-$. In the complex case, $|\mu| = \Re(\mu) + \Im(\mu)$. $\mathcal{M}(K)$ are all complex RADON measures on K, which becomes a BANACH space with $\|\mu\|_{\mathcal{M}(K)} := |\mu|(K)$.

Equipped with new notation we can now state another Riesz representation theorem that, in short, states $\mathcal{C}(K)^* \cong \mathcal{M}(K)$:

Theorem 1.4.1: Riesz-Markov-Kakutani (1938)

For every functional $\varphi : \mathcal{C}(K) \to \mathbb{C}$ there exists a Radon measure $\mu \in \mathcal{M}(K)$ such that

$$\varphi(f) = \int_K f(x)\,\mathrm{d}\mu(x) \qquad \forall f \in \mathcal{C}(K).$$

The isometric isomorphism $\Phi : \mathcal{C}(K)^* \to \mathcal{M}(K)$ maps positive functionals to non-negative measures.

Proof. See e.g. J. Conway: A Course in Functional Analysis, Appendix C. □

Remark 1.4.13 (Mones) Converse of this theorem is true due to the spectral mapping theorem and $\sigma(f(T)) \subset \mathbb{R}$ for self-adjoint T.

Lemma 1.4.14 (TODO)

Let $Q : \mathscr{H} \to \mathbb{C}$ be a function. Then the following are equivalent.

1. *There exists a unique $A \in L(\mathscr{H})$ such that $Q(x) = \langle Ax, x\rangle$ for all $x \in \mathscr{H}$.*
2. *There exists a $C > 0$ such that for all $x, y \in \mathscr{H}$ and all $\lambda \in \mathbb{C}$*
   1. $|Q(x)| \leqslant C\|x\|^2$,
   2. $Q(x+y) + Q(x-y) = 2Q(x) + 2Q(y)$,
   3. $Q(\lambda x) = |\lambda|^2 Q(x)$

   hold.

Proof. "1 $\Longrightarrow$ 2": clear.

Existence: Define (cf. polarisation identity)

$$\Psi(x,y) := \frac{1}{4}\left(Q(x+y) - Q(x-y) + iQ(x+iy) - iQ(x-iy)\right)$$

and for an orthonormal basis $(e_j)_{j\in J}$ of $\mathscr{H}$

$$Ax := \sum_{j\in J} \Psi(x, e_j)e_j.$$

Then we have $\Psi(x, y+z) = \Psi(x,y) + \Psi(y,z)$, which can be proven analogously to lemma 7.5 in FA I.

We now show for $\Phi(x,y) = \langle Ax, y\rangle$ for all $x, y \in \mathscr{H}$. Let $y = \sum_{k\in J} y_k e_k$. As $(e_j)_{j\in J}$ is a ONB we have

$$\begin{aligned}\langle Ax, y\rangle &= \Big\langle \sum_{j\in J} \Psi(x,e_j)e_j, \sum_{k\in J} y_k e_k\Big\rangle = \sum_{j\in J}\Psi(x,e_j)\sum_{k\in J} y_k\langle e_j, e_k\rangle \\ &= \sum_{j\in J}\Psi(x,e_j)y_j = \Psi\Big(x, \sum_{j\in J} y_j e_j\Big) = \Psi(x,y),\end{aligned}$$

as **TODO: some continuity**.

Uniqueness: This follows from lemma 1.2.15 □

Remark 1.4.15 (credits to Daniel) One doesn't need property 3 in lemma 1.4.14.

From 2 we get
$$Q(2x) - Q(0) = 4Q(x).$$
By property 1 we get $|Q(0)| \leqslant 0$, i.e. $Q(0) = 0$, implying $Q(2x) = |2|^2 Q(x)$.

One can proceed with strong induction: for $n \in \mathbb{N}$ one has
$$Q(nx) + Q((n-1)x - x) = 2Q((n-1)x) + 2Q(x).$$
Using the induction hypothesis for $n-1$ and $n-2$ we have
$$Q(nx) = (2(n-1)^2 - (n-2)^2 + 2)Q(x) = n^2 Q(x)$$
Similarly it holds that
$$Q(x) + Q(-x) = 2Q(0) + 2Q(x) \implies Q(-x) = -Q(x),$$
so one can show 2 for all $\lambda \in \mathbb{Z}$. Furthermore, for $m \in \mathbb{Z}$ we have by the above
$$Q\left(\frac{x}{m}\right) = \frac{1}{m^2} Q\left(m \cdot \frac{x}{m}\right) = \frac{1}{m^2} Q(x),$$
proving 3 for $\lambda \in \mathbb{Q}$.

As the map
$$\lambda \mapsto \textbf{TODO}$$
is zero on $\mathbb{Q}$ and continuous, and $\mathbb{Q} \subset \mathbb{R}$ is dense, it is also zero of $\mathbb{R}$.

Theorem 1.4.2: Idea of Borel measurable Calculus

Let $T = T^* \in L(\mathscr{H})$.

1. For every $x \in \mathscr{H}$ there exists a non-negative Radon-measure E^x such that
$$\langle f(T)x, x\rangle = \int_{\sigma(T)} f \, \mathrm{d}E^x \quad \forall f \in \mathcal{C}(\sigma(T)).$$
2. For every on $\sigma(T)$ bounded Borel-measurable function g there exists a unique $G \in L(\mathscr{H})$ such that
$$\langle Gx, x\rangle = \int_{\sigma(T)} g \, \mathrm{d}E^x \quad \forall x \in \mathscr{H}.$$
If g is real (non-negative), G is self-adjoint (positive).

Proof. 1 follows directly from theorem 1.4.1. 2: Uniqueness: see above. Existence: Define $Q(x) := \int_{\sigma(T)} g \, \mathrm{d}E^x$. Then we have
$$\begin{aligned} |Q(x)| &\overset{\triangle\neq}{\leqslant} \int_{\sigma(T)} \|g\|_\infty \, \mathrm{d}E^x = \|g\|_\infty \int_{\sigma(T)} \mathbb{1} \, \mathrm{d}E^x \\ &= \|g\|_\infty \langle \mathbb{1}(T)x, x\rangle \overset{1.3.1}{=} \|g\|_\infty \|x\|^2. \end{aligned} \tag{7}$$
For every $f \in \mathcal{C}(\sigma(T))$
$$\underbrace{\int_{\sigma(T)} f \, \mathrm{d}E^{x+y}}_{\langle f(T)(x+y), x+y\rangle} + \underbrace{\int_{\sigma(T)} f \, \mathrm{d}E^{x-y}}_{\langle f(T)(x-y), x-y\rangle} = 2\underbrace{\int_{\sigma(T)} f \, \mathrm{d}E^{x}}_{\langle f(T)x, x\rangle} + 2\underbrace{\int_{\sigma(T)} f \, \mathrm{d}E^{y}}_{\langle f(T)y, y\rangle}$$
and
$$\underbrace{\int_{\sigma(T)} f \, \mathrm{d}E^{\lambda x}}_{\langle f(T)(\lambda x), \lambda x} = |\lambda|^2 \underbrace{\int_{\sigma(T)} f \, \mathrm{d}E^{x}}_{\langle f(T)x, x\rangle}$$
holds. Due to the uniqueness of theorem 1.4.1 we get $E^{x+y} + E^{x-y} = 2E^x + 2E^y$ and $E^{\lambda x} = |\lambda|^2 E^x$. $\square$

The set of bounded Borel-measurable functions is the smallest set including bounded continuous functions and is closed under pointwise limits of such functions:

Lemma 1.4.16 (cf. Werner Lemma VII.1.5)
Let $K \subset \mathbb{C}$ be compact and $(\mathcal{B}(K), \|\cdot\|_\infty)$ the Banach space of bounded Borel-measurable functions on K and $\mathcal{C}(K) \subset U \subset \mathcal{B}(K)$ a set of functions with the following property: for all $(f_n)_{n\in\mathbb{N}} \subset$ with $f(t) := \lim_{n\to\infty} f_n(t)$ existing everywhere and $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ implies that $f \in U$. Then $U = \mathcal{B}(K)$.

Theorem 1.4.3: Borel-measurable Calculus

For every self-adjoint $T \in L(\mathscr{H})$ there exists a unique mapping $\Phi_T : \mathcal{B}(\sigma(T)) \to L(\mathscr{H})$ such that

1. $\Phi_T(\mathrm{id}) = T$ and $\Phi_T(\mathbb{1}) = \mathrm{id}$,
2. Φ_T is an involutive homomorphism of algebras.
3. Φ_T is continuous
4. $f_n \in \mathcal{B}(\sigma(T))$ with $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ and $f_n(t) \to f(t)$ for every $t \in \sigma(T)$ implies

$$\langle \Phi_T(f_n)x, y\rangle \to \langle \Phi_T(f)x, y\rangle \quad \forall x, y \in \mathscr{H}.$$

Proof. Uniqueness. As before, the first three conditions imply uniqueness on $\mathcal{C}(\sigma(T))$ and lemma 1.4.16 implies uniqueness of $\mathcal{B}(\sigma(T))$.

Existence. For a measurable bounded function g define $\Phi_T(g) = G$, where G is the operator from theorem 1.4.2. As by theorem 1.4.2

$$\langle f(T)x, x\rangle = \int_{\sigma(T)} f \, \mathrm{d}E^x$$

holds for all $f \in \mathcal{C}(\sigma(T))$, i.e. $\Phi_T(f) = f(T)$ for all $f \in \mathcal{C}(\sigma(T))$, where $f(T)$ is the continuous functional calculus, property 1 is fulfilled as $z \mapsto z$ and $\mathbb{1}$ are continuous on $\sigma(T)$.

Case 1: g is real-valued. By theorem 1.4.2 2 G is self-adjoint, therefore

$$\|\Phi(g)\| = \|G\| = \sup_{\substack{x\in\mathscr{H} \\ \|x\|=1}} |\langle Gx, x\rangle| \overset{(7)}{\leqslant} \sup_{\substack{x\in\mathscr{H} \\ \|x\|=1}} \|g\|_\infty \|x\|^2 = \|g\|_\infty$$

holds.

Case 2: g is complex-valued. Split g into its real and imaginary part first.

The proof of 4 follows from

$$\langle \Phi(f_n)x, x\rangle = \int_{\sigma(T)} f_n \, \mathrm{d}E^x \to \int_{\sigma(T)} f \, \mathrm{d}E^x = \langle \Phi(f)x, x\rangle,$$

which holds by the Lebesgue dominated convergence theorem **TODO:Majorante?**, and the polarisation identity.

The property 2 follows from a limit procedure often used in measure theory (Prinzip der guten Mengen). For example lets show that $\Phi(fg) = \Phi(f) \circ \Phi(g)$.

This result is already known for $f, g \in \mathcal{C}(\sigma(T))$. If $g \in \mathcal{C}(\sigma(T))$ we set

$$U := \{f : \mathcal{B}(\sigma(T)) : \Phi_T(fg) = \Phi_T(f) \circ \Phi_T(g)\}.$$

As $\mathcal{C}(\sigma(T)) \subset U$ and U is closed under pointwise limits of uniformly bounded sequences, $U = \mathcal{B}(\sigma(T))$ follows by lemma 1.4.16.

If $f \in \mathcal{B}(\sigma(T))$ we set

$$V := \{g :\in \mathcal{B}(\sigma(T)) : \Phi_T(fg) = \Phi_T(f) \circ \Phi_T(g)\}.$$

We have shown $\mathcal{C}(\sigma(T)) \subset V$. As V is closed under pointwise limits of uniformly bounded sequences, $V = \mathcal{B}(\sigma(T))$ by lemma 1.4.16.

The other properties can be shown analogously. □

Remark 1.4.17 The last property in the above theorem can be improved to $f_n(T)x \to f(T)x$ for all $x \in \mathscr{H}$ as applying 4 to $\overline{f}f$ yields

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

As Hilbert spaces have the Radon-Riesz property, meaning that

$$\|x_n\| \to \|x\|, \quad x_n \xrightarrow{\text{weak}} x \implies x_n \to x,$$

Proof. $\|x_n - x\|^2 = \|x_n\|^2 - 2\langle x_n, x\rangle + \|x\|^2 \xrightarrow{n\to\infty} \|x\|^2 - 2\|x\|^2 + \|x\|^2 = 0.$ □

the claim follows.

Example 1.4.18 (Generalization of example 1.4.20)
For $\mathscr{H} := L^2(0,1)$ and $h \in L^\infty(0,1)$ define the multiplication operator

$$T_h : \mathscr{H} \to \mathscr{H}, \ f \mapsto hf.$$

We have $T_h \in L(\mathscr{H})$ and $T_h^* = T_{\overline{h}}$.

For Lipschitz-continuous $h : [0,1] \to \mathbb{R}$ we have $\sigma(T_h) = h([0,1])$. The measurable functional calculus of T_h is given by $\Psi_{T_h}(g) = T_{g\circ h}$ for $g \in \mathcal{B}(h([0,1])$. ◇

Proof. As $h \in L^\infty(0,1)$ there exists a constant $M > 0$ such that $|h| \leqslant M$ almost everywhere. Thus $\|hf\|_2^2 \leqslant M^2\|f\|_2^2$, holds, implying the well-definedness of T_h and $T_h \in L(\mathscr{H})$ as linearity is clear. For $g \in \mathscr{H}$

$$\langle T_h f, g\rangle = \int_0^1 h(x)f(x)\overline{g(x)}\,\mathrm{d}x = \int_0^1 f(x)\overline{\overline{h(x)}g(x)}\,\mathrm{d}x = \langle f, T_{\overline{h}}g\rangle \qquad (8)$$

holds, implying $T_h^* = T_{\overline{h}}$.

If h is real, T_h is self-adjoint, so we have $\sigma_r(T_h) = \varnothing$ by lemma 1.2.13. A quick inspection reveals $\sigma_p(T_h) = \varnothing$ as well.

TODO $\sigma_c(T_h) = h([0,1])$.

Since the functional calculus for self-adjoint operators is unique, it suffices to verify the properties:

1 It is clear that $\Psi_{T_h}(\mathrm{id}) = T_h$. Furthermore for $x \in [0,1]$,

$$\begin{aligned}\Psi_{T_h}(\mathbb{1}_{\sigma(T_h)})(f(x)) &= (T_{\mathbb{1}_{\sigma(T_h)}\circ h})(f(x)) = (\mathbb{1}_{\sigma(T_h)}\circ h)(x)\cdot f(x) \\ &= \mathbb{1}_{\sigma(T_h)}\underbrace{(h(x))}_{\in\sigma(T_h)}\cdot f(x) = 1\cdot f(x) = f(x)\end{aligned}$$

holds, implying $\Psi_{T_h}(\mathbb{1}_{\sigma(T_h)}) = \mathrm{id}$.

(2) The linearity of Ψ_{T_h} is clear. The multiplicativity of Ψ_{T_h} follows from $(f\cdot g)\circ h = (f\circ h)\cdot(g\circ h)$ for $f,g \in \mathcal{C}(\sigma(T_h))$. Analogously to (8) it follows that $(T_{g\circ h})^* = T_{\overline{g}\circ h}$, as h is real valued, but $g \in \mathcal{C}(\sigma(T_h))$ must not be. This implies $(\Psi_{T_h}(g))^* = \Psi_{T_h}(\overline{g})$.

(3) We have **IS THIS TRUE??**

$$
\begin{aligned}
\|\Psi_{T_h}\|^2 &= \sup_{\|g\|_\infty=1} \|\Psi_{T_h}(g)\|^2 = \sup_{\|g\|_\infty=1} \|T_{g\circ h}\|^2 \\
&= \sup_{\|g\|_\infty=1} \sup_{\|f\|_{L^2(0,1)}=1} \|T_{g\circ h}(f)\|^2_{L^2(0,1)} \\
&= \sup_{\|g\|_\infty=1} \sup_{\|f\|_2=1} \|(g\circ h)\cdot f\|_2^2 \\
&= \sup_{\|g\|_\infty=1} \sup_{\|f\|_2=1} \int_0^1 |g(h(x))|^2 |f(x)|^2 \,\mathrm{d}x \\
&\leqslant \sup_{\|f\|_2=1} \Big(\sup_{\|g\|_\infty=1} \sup_{x\in[0,1]} |g(h(x))|^2 \Big) \int_0^1 |f(x)|^2 \,\mathrm{d}x \\
&= \sup_{\|f\|_2=1} \int_0^1 |f(x)|^2 \,\mathrm{d}x = 1.
\end{aligned}
$$

Thus Ψ_{T_h} is continuous.

(4) Let $f_n \in \mathcal{B}(\sigma(T))$ with $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ and $f_n(t) \to f(t)$ for every $t \in \sigma(T)$. We have

$$
\langle \Phi_T(f_n)x, y\rangle \to \langle \Phi_T(f)x, y\rangle \quad \forall x,y \in \mathscr{H}. \qquad \square
$$

TODO: hier nur den relevanten (FC) Teil, der rest weiter nach vorne im Skript Note: polar decomposition of this operator.

Lemma 1.4.19 (Properties of the measurable functional calculus)
Let $T = T^* \in L(\mathscr{H})$ and $f \in \mathcal{B}(\sigma(T))$.

(1) $\|f(T)\| \leqslant \|f\|_\infty$ holds.

(2) If $f \geqslant 0$, then $f(T) \geqslant 0$.

(3) If $Tx = \lambda x$, then $f(T)x = f(\lambda)x$

(4) $\sigma(f(T)) \subset \overline{f(\sigma(T))}$.

Proof. (1) The functional calculus for real valued f is self-adjoint. Thus

$$
\begin{aligned}
\|\Phi_T(f)\| &= \sup_{\|x\|=1} \langle \Phi(f)x, x\rangle = \sup_{\|x\|=1} \left| \int_{\sigma(T)} f \,\mathrm{d}\mu_x \right| \\
&\leqslant \sup_{\|x\|=1} \|f\|_\infty \mu_x(\sigma(T)) \leqslant \|f\|_\infty.
\end{aligned}
$$

For general f the multiplicativity, involutivity and normality of the functional calculus yield

$$
\|\Phi_T(f)\|^2 = \|\Phi_T(f)\Phi_T(f)^*\| = \|\Phi_T(|f|^2)\| \leqslant \||f|^2\|_\infty = \|f\|_\infty^2,
$$

where we use the previously discussed case in the inequality.

(2) For $x \in \mathscr{H}$ we have

$$
\langle \Phi_T(f)x, x\rangle = \int_{\sigma(T)} f \,\mathrm{d}\mu_x \geqslant 0
$$

(3) Let $Tx = \lambda x$ and set

$$U := \{f \in \mathcal{B}(\sigma(T)) : \Phi_T(f)x = f(\lambda)x\}.$$

By theorem 1.3.2 we have $\mathcal{C}(\sigma(T)) \subset U$. Let $(f_n)_{n\in\mathbb{N}} \subset U$ such that $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ such that $f_n \to f$ pointwise. Then $\Phi_T(f_n)x \to \Phi_T(f)x$ implies $f \in U$. An application of lemma 1.4.16 yields the claim.

(4) Let $\mu \notin \overline{f(\sigma(T))}$. For all $\lambda \in \sigma(T)$

$$|\mu - f(\lambda)| \geqslant \operatorname{dist}(\mu, \overline{f(\sigma(T))}) =: d > 0 \tag{9}$$

holds. For $g : \sigma(T) \to \mathbb{C}, \lambda \mapsto (\mu - f(\lambda))^{-1}$ it holds that $\|g\|_\infty \leqslant \frac{1}{d}$ by (9) and thus $g \in \mathcal{B}(\sigma(T))$. As

$$(\mu - f(\lambda))g(\lambda) = g(\lambda(\mu - f(\lambda)) = 1$$

we have

$$(\mu - f(T))g(T) = g(T)(\mu - f(T)) = \mathrm{id},$$

implying $\mu \in \rho(f(T))$. □

TODO: Beispiel um die Unterschiede zum cts FC zu verdeutlichen

Application of the measurable functional calculus

Example 1.4.20 (TODO)
Let

$$T : L^2([0,1]) \to L^2([0,1]),\ f(x) \mapsto xf(x).$$

Then $T = T^* \in L(L^2([0,1]))$, and $\sigma(T) = \sigma_c(T) = [0,1]$:

(1) Linearity is clear. We have

$$\int_0^1 |xf(x)|^2 \,\mathrm{d}x = \int_0^1 |x|^2 |f(x)|^2 \,\mathrm{d}x \leqslant \int_0^1 |f(x)|^2 \,\mathrm{d}x,$$

implying $\|T\| \leqslant 1$.

(2) First, not that as T is self-adjoint, we have $\sigma_r(T) = \varnothing$ by lemma 1.2.13. To see that $[0,1] \subset \sigma(T)$, note that the constant function 1 is not in the range of $T - \lambda$ for any $\lambda \in [0,1]$ because $(x-\lambda)g(x) = 1$ would imply $g(x) = \frac{1}{x-\lambda}$ almost everywhere, which is not in $L^2([0,1])$ for such λ, as the singularity $x = \lambda$ is inside the integration domain. For any other λ, $T - \lambda$ is invertible.

As T is self-adjoint, it is normal, so $\sigma(T) = \sigma_{\mathrm{app}}(T)$ holds by lemma 1.2.5. For $\delta > 0$, consider $f_{\lambda,\delta} := \frac{1}{\sqrt{2\delta}} \mathbb{1}_{[\lambda-\delta,\lambda+\delta]}$.
For $\lambda \in [\delta, 1-\delta] \subset [0,1]$ we have

$$\|f_{\lambda,\delta}\|_2^2 = \int_{\lambda-\delta}^{\lambda+\delta} \frac{1}{2\delta} \,\mathrm{d}x = \frac{\lambda + \delta - (\lambda - \delta)}{2\delta} = 1.$$

We have

$$\|(T-\lambda)f_{\delta,\lambda}\|_2^2 = \frac{1}{2\delta} \int_{\lambda-\delta}^{\lambda+\delta} (x-\lambda)^2 \,\mathrm{d}x = \frac{1}{2\delta} \int_{-\delta}^{\delta} u^2 \,\mathrm{d}u = \frac{\delta^2}{3} \xrightarrow{\delta \to 0} 0.$$

TODO

Let $g : \sigma(T) \to \mathbb{C}$ be bounded and measurable. For $x \in [0,1]$ and $f \in L^2([0,1])$ we have $(g(T)f)(x) = g(x)f(x)$.

TODO $\diamond$

1.5 Spectral theorem for bounded self-adjoint operators

In view of the spectral theorem, which we will prove in this subsection, $f(T)$ is especially important for $f = \mathbb{1}_A$ for a BOREL measurable set (from now on: BOREL set) $A \subset \mathbb{R}$.

From now on let $T \in L(\mathcal{H})$ always be self-adjoint.

Lemma 1.5.1 ($\mathbb{1}_A(T)$ orthogonal projection)
Let Ψ_T be the measurable functional calculus of T. For a BOREL *set $A \subset \sigma(T)$ define*

$$A \mapsto \Psi_T(\mathbb{1}_A) = \mathbb{1}_A(T).E : \sigma(T) \to L(\mathcal{H}),$$

Each E_A is an orthogonal projection onto $\mathcal{H}_A := \operatorname{ran}(E_A)$, which is T-invariant, implying that $T_A := T|_{\mathcal{H}_A} \in L(\mathcal{H}_A)$ is self-adjoint and $\sigma(T_A) \subset \overline{A}$.

Proof. Due to $\mathbb{1}_A^2 = \mathbb{1}_A$ and the multiplicativity of the functional calculus we have $E_A^2 = E_A$ and due to $\overline{\mathbb{1}_A} = \mathbb{1}_A$ (as $\sigma(T) \subset \mathbb{R}$) and the involutivity of the functional calculus we have $E_A^* = E_A$. The claim now follows from lemma 1.4.4.

Therefore, $\mathcal{H}_A = \ker(E_A)^\perp$ holds. Let $y \in \mathcal{H}_A$. Then there exists a $x \in \mathcal{H}$ such that $E_A x = y$. We show that $Ty \in \ker(E_A)^\perp$.

By multiplicativity of Ψ_T we have $TE_A = E_A T$:

$$\Psi_T(\mathrm{id})\Psi_T(\mathbb{1}_A) = \Psi_T(\mathrm{id} \cdot \mathbb{1}_A) = \Psi_T(\mathbb{1}_A \cdot \mathrm{id}) = \Psi(\mathbb{1}_A)\Psi_T(\mathrm{id}).$$

Thus for $z \in \ker(E_A)$.

$$\langle Ty, z\rangle = \langle TE_A x, z\rangle = \langle E_A Tx, z\rangle = \langle Tx, E_A z\rangle = 0.$$

holds, implying the T-invariance.

We show that $\overline{A}^{\complement} := \sigma(T)\backslash A \subset \rho(T_A)$. For $\lambda_0 \in A$ define the function

$$f : \sigma(T) \to A, \; x \mapsto x \cdot \mathbb{1}_A(x) + \lambda_0 \mathbb{1}_{A^\complement}(x),$$

which is measurable due to A and $A^\complement$ being measurable, so $f \in \mathcal{B}(\sigma(T))$. By lemma **TODO** we have $\sigma(f(T)) \subset \overline{f(\sigma(T))}$, implying

$$\overline{A}^{\complement} = \overline{f(\sigma(T))}^{\complement} \subset \sigma(f(T))^\complement = \rho(f(T)). \tag{10}$$

By linearity and multiplicativity we have

$$\begin{aligned} f(T) &= \Psi_T(f) = \Psi_T(\mathrm{id} \cdot \mathbb{1}_A + \lambda_0 \mathbb{1}_{A^\complement}) \\ &= \Psi_T(\mathrm{id})\Psi_T(\mathbb{1}_A) + \lambda_0(\Psi_T(\mathbb{1}) - \Psi_T(\mathbb{1}_A)) = TE_A + \lambda_0(\mathrm{id} - E_A). \end{aligned}$$

We can finally show that $T_A - \lambda$ is bijective on $\mathcal{H}_A$.

Injectivity. Let $x \in \mathcal{H}_A$ such that $T_A x - \lambda x = 0$. Then we have

$$(f(T) - \lambda)x = TE_A x + \lambda_0(x - E_A x) - \lambda x = Tx - \lambda_0(x - x) - \lambda x = 0.$$

As $f(T) - \lambda I$ is bijective by (10), it follows that $x = 0$.

Surjectivity. Let $a \in \mathscr{H}_A$. As $f(T) - \lambda I$ is bijective, there exists a $x \in \mathscr{H}$ such that $f(T)x - \lambda x = a$. As $\mathscr{H}_A$ is closed there exists $y \in \mathscr{H}_A$ and $z \in \mathscr{H}_A^\perp$ such that $x = y + z$. Thus

$$\begin{aligned} a &= f(T)x - \lambda x = TE_A x + \lambda_0(x - E_A x) - \lambda x \\ &= Ty + \lambda_0(y + z - y) - \lambda x = \underbrace{Ty - \lambda y}_{\in \mathscr{H}_A} + \underbrace{(\lambda_0 - \lambda)z}_{\in \mathscr{H}_A^\perp} \end{aligned}$$

As $\lambda_0 \in A$ and $\lambda \in \overline{A}^\complement$ we have $\lambda \neq \lambda_0$, implying $z = 0$, because $a \in \mathscr{H}_A$. This implies $x \in \mathscr{H}_A$.

But for $x \in \mathscr{H}_A$ we have (cf. above) $(f(T) - \lambda\,\mathrm{id})x = (T - \lambda\,\mathrm{id})x$. Hence $T_A - \lambda\,\mathrm{id}$ is surjective. □

Gilt $E_A : \mathscr{H} \to \{0, 1\}$, $x \mapsto 1$ if $Tx \in A$ and 0 else?

Lemma 1.5.2 (Properties of E_A)

1. *We have $E_\varnothing = \mathbb{1}_\varnothing(T) = 0$ and $E_{\sigma(T)} = \mathbb{1}_{\sigma(T)}(T) = \mathrm{id}$.*
2. *For pairwise disjoint sets $(A_k \subset \sigma(T))_{k\in\mathbb{N}}$ and $x \in \mathscr{H}$*

$$\sum_{k\in\mathbb{N}} E_{A_k} x = E_{\bigcup_{k\in\mathbb{N}} A_k} x$$

holds.

One easily shows that $\mathbb{1}_A(T)\,\mathbb{1}_B(T) = \mathbb{1}_{A\cap B}(T)$ holds for Borel set $A, B \subset \sigma(T)$, implying $E_A E_B = E_B E_A = E_{A\cap B}$.

Proof. 1 follows from theorem 1.4.3 1, 3 from theorem 1.4.3 2 and 2 from remark 1.4.17 with $f_n := \sum_{k=1}^n \mathbb{1}_{A_k}$. □

Remark 1.5.3 Note that, in general,

$$\sum_{k\in\mathbb{N}} E_{A_k} \neq E_{\bigcup_{k\in\mathbb{N}} A_k},$$

since almost never $\|\,\mathbb{1}_{A_k(T)}\| \xrightarrow{k\to\infty} 0$ holds, (so the sum would diverge).

The two preceding lemmas show that the map

$$E : A \mapsto \mathbb{1}_{A\cap\sigma(T)}(T) \tag{11}$$

is a spectral measure in the sense of the following definition.

Definition 1.5.4 (Spectral measure)

Let Σ be the σ-Algebra of Borel sets on $\mathbb{R}$. A map

$$E : \Sigma \to L(\mathscr{H}),\ A \mapsto E_A$$

is called spectral measure if every E_A is an orthogonal projection the and properties from lemma 1.5.2 hold.

spectral measure

For 2 replace $\sigma(A)$ by Σ.

E has compact support if there exists a compact set $K \subset \mathbb{R}$ with $E_K = \mathrm{id}$.

Lemma 1.5.5 (Properties of spectral measures)

Let $A, B \subset \mathbb{R}$ be BOREL *sets.*

1. *We have $E_A + E_B = E_{A\cap B} + E_{A\cup B}$.*
2. *Let $A \subset B$. Then $E_{B\setminus A} = E_B - E_A$ is a projection, i.e. $E_B \geqslant E_A$. We have $E_B E_A = E_A E_B = E_A$.*
3. *Let $A \cap B = \varnothing$. Then $E_A E_B = 0$ holds.*
4. *We have $E_A E_B = E_{A\cap B}$.*
5. *Let $A \cap B = \varnothing$. Then $\langle E_A x, E_B y\rangle = 0$ for all $x, y \in \mathscr{H}$, i.e. E_A and E_B project onto mutually orthogonal subspaces*

Proof. 1 As $A = (A \cap B) \sqcup (A\setminus B)$ and $A \cup B = (A\setminus B) \sqcup B$, where $\sqcup$ denotes disjoint union, by lemma 1.5.2 2 we have

$$E_A + E_B = E_{A\cap B} + E_{A\setminus B} + E_B = E_{A\cap B} + E_{A\cup B}.$$

2 Let $A \subset B$. As $B = A \sqcup (B\setminus A)$ we have by lemma 1.5.2 2

$$E_B = E_A + E_{B\setminus A} \iff E_{B\setminus A} = E_B - E_A.$$

By lemma 1.4.5 (L) we have

$$E_{B\setminus A}^2 = E_B^2 - 2E_B E_A + E_A^2 = E_B - 2E_B E_A + E_A \overset{\text{L}}{=} E_B - 2E_A + E_A,$$

so $E_{B\setminus A}$ is a projection. Furthermore it is self-adjoint, and therefore positive by 1.4.4, i.e. $E_B \geqslant E_A$. The last equation is 1.4.5.

3 Let $A \cap B = \varnothing$. By lemma 1.5.2 we have

$$E_A + E_B = E_{A\cup B},$$

implying (by multiplication of E_A)

$$E_A^2 + E_A E_B = E_A E_{A\cup B} \overset{2}{=} E_A.$$

This yields

$$E_A + E_A E_B = E_A^2 + E_A E_B = E_A,$$

and therefore $E_A E_B = 0$.

4 By 2 and 3 we have

$$E_A E_B = E_A(E_{A\cap B} + E_{B\setminus A}) = E_A E_{A\cap B} + E_A E_{B\setminus A} = E_{A\cap B},$$

as $A \cap B \subset A$ and $A \cap (B\setminus A) = \varnothing$.

5 Let $A \cap B = \varnothing$ and $x, y \in \mathscr{H}$. We have

$$\langle E_A x, E_B y\rangle = \langle E_B^* E_A x, y\rangle = \langle E_B E_A x, y\rangle = \langle 0, y\rangle = 0. \qquad \square$$

Interpreting the aforementioned properties, in some sense, "E is a measure with values in $L(\mathscr{H})$ instead of $[0, \infty)$.”

We can now state the spectral theorem for bounded self-adjoint operators.

Theorem 1.5.1: Spectral theorem

There exists a unique compactly supported spectral measure E such that

$$T = \int_{\sigma(T)} \lambda \, \mathrm{d}E_\lambda$$

holds. The map

$$\Psi : \mathcal{B}(\sigma(T)) \to L(\mathcal{H}), \; f \mapsto f(T) = \int f(\lambda) \, \mathrm{d}E_\lambda$$

coincides with the Borel-measurable functional calculus and

$$\langle f(T)x, y\rangle = \int_{\sigma(T)} f(\lambda) \, \mathrm{d}\langle E_\lambda x, y\rangle$$

holds for all $x, y \in \mathcal{H}$, where $\langle E_\lambda x, y\rangle$ is the complex-valued measure $A \mapsto \langle E_A x, y\rangle$.

Remark 1.5.6 One can generalize the above theorem to normal operators, as a normal operator $T \in L(\mathcal{H})$ can be written as the sum of two self-adjoint operators

$$T = S_1 + iS_2 = \frac{T + T^*}{2} + i \cdot \frac{T - T^*}{2i}.$$

Then by the spectral theorem for self-adjoint operators there exist two spectral measures E_1 and E_2. Since T is normal, S_1 and S_2 commute, and therefore E_1 and E_2 commute, too. We can then identify $\mathbb{R}^2 \cong \mathbb{C}$ (cf. Werner for more details).

Generalization to non-normal is difficult as **TODO, SE**

Example 1.5.7 (Spectral decomposition of $\boldsymbol{f(x) \mapsto xf(x)}$) Consider the operator from example 1.4.20 and the spectral measure E from (11), where Σ is the set of Borel subsets of $\sigma(T)$, so E becomes compactly supported. Then E represents T:

$$\int \lambda d\langle E_\lambda x, y\rangle = \int \lambda x(\lambda)\overline{y(\lambda)} d\lambda = \langle Tx, y\rangle,$$

as

$$\langle E_A x, y\rangle = \int_{A \cap [0,1]} x(\lambda)\overline{y(\lambda)} d\lambda = \int_0^1 \mathbb{1}_A(\lambda) x(\lambda)\overline{y(\lambda)} d\lambda.$$

Thus, $d\langle E_\lambda x, y\rangle$ has the density $x\overline{y} \in L^1$ with respect to the Lebesgue measure. This shows $T = \int \lambda dE_\lambda$. ◇

1.6 Applications of the spectral theorem

Corollary 1.6.1 (Polar decomposition)
For $T \in L(\mathscr{H})$ define $|T| = (T^*T)^{\frac{1}{2}}$. There exists a partial isometry U with $T = U|T|$.

$U \in L(\mathscr{H})$ is an partial isometry if $U^*UU^*U = U^*U$.

Proof. The proof is analogous to the one for compact operators. As $(T^*T)^{\frac{1}{2}}$ is self-adjoint **(really??)**, for $x \in \mathscr{H}$

$$\||T|x\|^2 = \langle x, T^*Tx\rangle = \langle Tx, Tx\rangle = \|Tx\|^2$$

holds. Therefore, $U(|T|x) := Tx$ is an isometry of $\operatorname{ran}(|T|)$ to $\operatorname{ran}(T)$. This can be extended to $U : \overline{\operatorname{ran}}(|T|) \to \overline{\operatorname{ran}}(T)$. We put $U \equiv 0$ on $(\operatorname{ran}(|T|)^\perp = \ker(|T|) = \ker(T)$. □

Example 1.6.2 (Polar decomposition of the shift operators)
Recall the left- and right-shift operators L and R from examples 1.2.3 and 1.2.4. We have $L^* = R$ and $R^* = L$. For $x = (x_k)_{k=1}^\infty \in \ell^2$ we have

$$(R^*R)(x) = (LR)(x) = L(0, x_1, x_2, \dots) = x,$$

implying $R^*R = \mathrm{id}$ and thus $|R| = \mathrm{id}$. Therefore the polar decomposition of R is trivial, we can choose $U_R = R$, which is indeed unitary.

For $x = (x_k)_{k=1}^\infty \in \ell^2$ we have

$$(L^*L)x = (RL)x = R(x_2, x_3, \dots) = (0, x_2, x_3, \dots).$$

We further observe that $\sqrt{L^*L} = L^*L$, implying $|L| = L^*L$. We have

$$L(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots) \overset{!}{=} (U_L|L|)x = U(0, x_2, x_3, \dots),$$

implying $U_L = L$ and therefore $L = |L|L = RL^2$. ◇

Example 1.6.3 (Polar decomposition II)
Let $T = |\xi\rangle\langle\eta|$ for $\|\xi\| = \|\eta\| = 1$, i.e.

this is called Dirac notation

$$T : \mathscr{H} \to \mathscr{H}, \; x \mapsto \xi\langle\eta, x\rangle,$$

where we assume that the scalar product is conjugate-linear in the first variable. We have $T^* = |\eta\rangle\langle\xi|$, as for $x, y \in \mathscr{H}$

$$\langle T^*x, y\rangle = \langle \eta\langle\xi, x\rangle, y\rangle = \langle x, \xi\rangle\langle\eta, y\rangle = \langle x, \xi\langle\eta, y\rangle\rangle = \langle x, Ty\rangle.$$

holds. Thus, as $\|\xi\| = 1$ it holds that

$$(T^*T)(x) = T^*(\xi\langle\eta, x\rangle) = \eta\langle\xi, \xi\langle\eta, x\rangle\rangle = \eta\langle\eta, x\rangle.$$

Furthermore, we have $|T|(x) = \langle\eta, x\rangle$, as for $x \in \mathscr{H}$

$$(|T|^2)(x) = (|T|)(\langle\eta, x\rangle) = \langle\eta, \langle\eta, x\rangle\rangle = \eta\langle\eta, x\rangle = (T^*T)(x)$$

holds. **TODO** ◇

We can also use the spectral theorem to prove lemma 1.1.4:

Proof. (LMU Exam 24.07.09) As $(T^{-1})^* = (T^*)^{-1}$ we have

$$\begin{aligned}
\|(T-\lambda)^{-1}\| &= \sup_{\|x\|_{\mathscr{H}}=1} \langle x, (T-\overline{\lambda})^{-1}(T-\lambda)^{-1}x\rangle \\
&= \sup_{\|x\|=1} \int_{\sigma(T)} \mathrm{d}\mu_x(z) \frac{1}{|z-\lambda|^2} \leqslant \frac{1}{|\Im(\lambda)|^2} \sup_{\|x\|=1} \int_{\sigma(T)} \mathrm{d}\mu_x(z) \\
&= \frac{1}{|\Im(\lambda)|^2} \sup_{\|x\|=1} \|x\|^2 = \frac{1}{|\Im(\lambda)|^2},
\end{aligned}$$

where μ_x is the spectral measure of T relative to the vector x. **TODO**$\square$

1.7 Spectral theorem for unbounded self-adjoint operators

Motivation 1 (canonical commutation relations cannot be satisfied by bounded operators) Consider two densely defined operators P and Q on $\mathscr{H}$ such that $D := \operatorname{dom}(PQ) \cap \operatorname{dom}(QP)$ is dense in $\mathscr{H}$ and $QP - PQ = i\mathbb{1}$, called Heisenberg's Commutation Relation, holds on D. Similarly to an exercise from FA I (see Appendix), we can show that either P or Q must be unbounded. In quantum mechanics this relation appears frequently and therefore studying unbounded operators is essential. [9] ◇

Take $P := x_i$ and $Q := D_j$ (one-dimensional position and momentum operators up to constants). Then

$$[P, Q] := PQ - QP = \delta_{i,j}$$

(Max Born, 1925), implying the Heisenberg uncertainty principle (E. Kennard, 1927).

The eigenvalues of the position operator – when it is considered with its widest possible domains (Schwartz spaces) – are the possible position vectors of the particle.

Motivation 2 Let $T = T^* \in L(\mathscr{H})$. Then $\overline{\operatorname{ran}}(T) = \ker(T)^{\perp}$. If T is injective, $\operatorname{ran}(T) \subset \mathscr{H}$ is dense. Therefore, T^{-1} can only be defined on $\operatorname{ran}(T)$ and is unbounded if $\operatorname{ran}(T) \neq \mathscr{H}$.

Motivation 3 Many differential operators are unbounded, i.e. $T = \frac{\mathrm{d}}{\mathrm{d}t}$ with $T : \mathcal{C}([0,1]) \to L^2([0,1])$, which is not continuous with respect to $\|\cdot\|_2$. Thus we are interested in operators defined on a subspace of $\mathscr{H}$.

Definition 1.7.1 (densely defined, closed, closable)
Let $\operatorname{dom}(T) \subset \mathscr{H}$ be a linear subspace and $T : \operatorname{dom}(T) \to \mathscr{H}$.

1. T is densely defined if $\operatorname{dom}(T) \subset \mathscr{H}$ is dense. — densely defined
2. the graph of T is $\Gamma(T) := \{(Tx, x), x \in \operatorname{dom}(T)\} \subset \mathscr{H} \times \mathscr{H}$ where $\langle (u,v), (x,y) \rangle_{\mathscr{H}\times\mathscr{H}} := \langle u, x\rangle + \langle v, y\rangle$. — graph
3. T is called closed if $\Gamma(T) \subset \mathscr{H} \times \mathscr{H}$ is closed. — closed
4. T is called closable if $\overline{\Gamma(T)}$ is the graph of some linear operator T_0, called closure of T, denoted by $T_0 := \overline{T}$. — closable, closure of T
5. An operator $S : \mathscr{H} \supset \operatorname{dom}(S) \to \mathscr{H}$ is called extension of T and we write $T \subset S$ if $\operatorname{dom}(T) \subset \operatorname{dom}(S)$ and $S \equiv T$ on $\operatorname{dom}(T)$, or equivalently $\Gamma(T) \subset \Gamma(S)$. — extension of T
6. We say $T = S$ if $S \subset T \subset S$.

The domain of an operator is an essential part of its definition. For operators S and T on $\mathscr{H}$ we set $\operatorname{dom}(S+T) := \operatorname{dom}(S) \cap \operatorname{dom}(T)$, $(S+T)(x) = Sx + Tx$, $\operatorname{dom}(ST) := \{x \in \operatorname{dom}(T) : Tx \in \operatorname{dom}(S)\}$ and $(ST)(x) = S(T(x))$.

From now on, let $T : \mathscr{H} \supset \operatorname{dom}(T) \to \mathscr{H}$ be an operator.

Lemma 1.7.2 (Characterisation of closable operators)
Let T be linear. Then the following statements are equivalent:

1. *T is closable.*
2. *For all sequences $(x_n)_{n\in\mathbb{N}} \subset \operatorname{dom}(T)$ converging to zero such that $(Tx_n)_{n\in\mathbb{N}}$ converges in $\mathscr{H}$, $Tx_n \to 0$ holds.*
3. *T has a closed linear extension S.*

4 *$(y,0) \in \overline{\Gamma(T)}$ implies $y = 0$, i.e. $(\{0\} \times \mathscr{H}) \cap \overline{\Gamma(T)} = \{(0,0)\}$.*

Proof. 1 $\implies$ 2: Let $T_0 = \overline{T}$. Since $T \equiv T_0$ on $\operatorname{dom}(T)$ we have $(x_n, Tx_n) \in \Gamma(T_0) = \overline{\Gamma(T)}$. Thus, $z = \lim_{n\to\infty} Tx_n \in \Gamma(T_0)$, yielding $z = T_0 x = Tx = 0$.

1 $\implies$ 3: We have $\Gamma(T) \subset \overline{\Gamma(T)} = \Gamma(\overline{T})$ and therefore $T \subset \overline{T}$, latter of which is closed.

3 $\implies$ 4: If $(y,0) \in \overline{\Gamma(T)}$, then $y = S(0) = 0$, as $\overline{\Gamma(T)} \subset \overline{\Gamma(S)} = G(S)$ holds.

4 $\implies$ 2: trivial.

4 $\implies$ 1: Define $\mathcal{D} := \{x \in \mathscr{H} \mid \exists y \in \mathscr{H} : (x,y) \in \overline{\Gamma(T)}\}$. Let $x \in \mathcal{D}$ and $y, z \in \mathscr{H}$ such that $(x,y), (x,z) \in \overline{\Gamma(T)}$. Then also $(0, y-z) \in \overline{\Gamma(T)}$ and thus $y = z$ by 4. Hence the y in the definition of $\mathcal{D}$ is unique so we can define the sought after closure $S : \mathscr{H} \supset \mathcal{D} =: \operatorname{dom}(S) \to \mathscr{H}$ by $Sx = y$. □

Definition 1.7.3 (Symmetric operator)
An operator T is symmetric if $\langle Tx, y\rangle = \langle x, Ty\rangle$ holds for all $x, y \in \operatorname{dom}(T)$.

Everywhere defined means that $\operatorname{dom}(T) = \mathscr{H}$. Everywhere defined symmetric operators are self-adjoint.

Example 1.7.4 (Symmetric closed densely defined operator)
Consider the densely defined, closed and symmetric operator

$$T : \ell_2 \supset \operatorname{dom}(T) := \{(x_n)_{n\in\mathbb{N}} \in \ell_2 : (nx_n)_{n\in\mathbb{N}} \in \ell_2\} \to \ell_2$$

given by $T((x_n)_{n\in\mathbb{N}}) = (nx_n)_{n\in\mathbb{N}}$: for $x, y \in \operatorname{dom}(T)$ we have

$$\langle Tx, y\rangle_{\ell_2} = \sum_{k\in\mathbb{N}} nx_n\overline{y_n} = \sum_{k\in\mathbb{N}} x_n\overline{ny_n} = \langle x, Ty\rangle_{\ell_2}.$$

Let $(x^{(k)})_{k\in\mathbb{N}} \subset \operatorname{dom}(T)$ converge to $x \in \ell_2$ and $Tx^{(k)}$ to $y \in \ell_2$. Since convergence in ℓ^p implies coordinatewise convergence we have $y_n := \lim_{k\to\infty} nx_n^{(k)} = nx_n^{(k)}$, implying $\|y\|_{\ell_2}^2 = \sum_{n\in\mathbb{N}} n^2|x_n|^2 < \infty$. Density follows as usual by compactly supported sequences. ◇

Theorem 1.7.1: Hellinger-Toeplitz

An everywhere defined symmetric operator T is bounded.

Proof. (With CGT) Let $x_n \to 0$ and $Tx_n \to z \in \mathscr{H}$. By the closed graph theorem it suffices to show $z = 0$:

$$\langle z, z\rangle = \langle \lim_{n\to\infty} Tx_n, z\rangle = \lim_{n\to\infty}\langle Tx_n, z\rangle = \lim_{n\to\infty}\langle x_n, Tz\rangle = 0.$$ □

Proof. (With UBP) Consider $\{f_x : \mathscr{H} \to \mathbb{K},\ y \mapsto \langle Tx, y\rangle\}_{\|x\|\leqslant 1}$.

Let $y \in \mathscr{H}$ with $\|y\| \leqslant 1$. Then

$$\sup_{\|x\|\leqslant 1} |\langle Tx, y\rangle| = \sup_{\|x\|\leqslant 1} |\langle x, Ty\rangle| < \infty$$

as the inner product defines a linear functional. The Uniform Boundedness Principle implies $\sup_{\|x\|\leqslant 1}\sup_{\|y\|\leqslant 1} |\langle Tx, y\rangle| < \infty$.

Finally note that

$$\|T\|^2 = \sup_{\|x\|\leqslant 1} \|Tx\|^2 = \sup_{\|x\|\leqslant 1} |\langle Tx, Tx\rangle| \leqslant \sup_{\|x\|\leqslant 1} \sup_{\|y\|\leqslant 1} |\langle Tx, y\rangle| < \infty. \qquad \square$$

Lemma 1.7.5

Every closed, unbounded operator can be transformed into a bounded operator by modifying the scalar product of $\operatorname{dom}(T)$.

Proof. Let $T : \mathscr{H} \supset \operatorname{dom}(T) \to \mathscr{H}$ be closed. Define

$$\langle\cdot,\cdot\rangle_T : \operatorname{dom}(T)^2 \to \mathbb{K},\ (x, y) \mapsto \langle x, y\rangle + \langle Tx, Ty\rangle,$$

which is a scalar product on $\operatorname{dom}(T)$. Furthermore, $\mathscr{H}_T := (\operatorname{dom}(T), \langle\cdot,\cdot\rangle_T)$ is a HILBERT space as **TODO**

Finally, the operator $\hat{T} : \mathscr{H}_T \to \mathscr{H}, x \mapsto Tx$ is linear and bounded: **TODO** $\square$

DEFINITION 1.7.6 (ADJOINT OPERATOR)

Let T be densely defined and $\theta_y : \operatorname{dom}(T) \to \mathbb{C},\ x \mapsto \langle Tx, y\rangle$. Define $\operatorname{dom}(T^*) := \{y \in \mathscr{H} : \theta_y \text{ is continuous}\}$. According to RIESZ θ_y can be extended to a continuous map on $\mathscr{H}$ and be represented as $\langle\cdot, z\rangle$ for some $z \in \mathscr{H}$. We denote $z = T^*y$, which is unique due to the density of $\operatorname{dom}(T)$. The operator T^* is called the adjoint of T. If $T = T^*$, T is self-adjoint.

Corollary 1.7.7

We have $\langle Tx, y\rangle = \langle x, T^*y\rangle$ *for all* $x \in \operatorname{dom}(T)$ *and all* $y \in \operatorname{dom}(T^*)$.

Example 1.7.8 (symmetric, not self-adjoint operator) Define

$$T : L^2(\mathbb{R}) \supset \mathcal{C}_c(\mathbb{R}) \to L^2(\mathbb{R}),\ (Tf)(x) := e^x f(x)$$

T is densely defined as $\mathcal{C}_c^\infty(\mathbb{R}) \subset L^2(\mathbb{R})$ is dense and $\mathcal{C}_c^\infty(\mathbb{R}) \subset \mathbb{C}_c(\mathbb{R})$ holds. For $f, g \in \operatorname{dom}(T) = \mathbb{C}_c(\mathbb{R})$ we have

$$\langle Tf, g\rangle = \int_{\mathbb{R}} e^x f(x)\overline{g(x)}\,\mathrm{d}x = \int_{\mathbb{R}} f(x)\overline{e^x g(x)}\,\mathrm{d}x = \langle f, Tg\rangle.$$

Let $h : \mathbb{R} \to \mathbb{R},\ x \mapsto e^{-x^2}$. Then $h \in L^2(\mathbb{R}) \setminus \mathbb{C}_c(\mathbb{R})$ but $j := Th \in L^2(\mathbb{R})$. Thus for all $g \in \operatorname{dom}(T)$

$$\langle h, Ag\rangle = \int_{\mathbb{R}} e^{-x^2} e^x g(x)\,\mathrm{d}x = \langle j, g\rangle$$

holds, implying $h \in \operatorname{dom}(T^*) \setminus \operatorname{dom}(T)$, so $\operatorname{dom}(T) \subsetneq \operatorname{dom}(T^*)$, so T can't be self-adjoint. [LMU Exam 24.07.09] $\diamond$

Lemma 1.7.9 (Properties of the adjoint)

Let S and T be two densely defined operators on $\mathscr{H}$.

1. $(\alpha S)^* = \overline{\alpha} S^*$ *for all* $\alpha \in \mathcal{C}$.
2. *If $S + T$ is densely defined,* $S^* + T^* \subset (S + T)^*$, *with equality if* $S \in L(\mathscr{H})$.
3. *If ST is densely defined,* $S^*T^* \subset (TS)^*$ *with equality if* $S \in L(\mathscr{H})$.
4. *If* $S \subset T$ *then* $T^* \subset S^*$.

Proof. TODO □

If T and T^* are densely defined, one can define T^{**}.

Lemma 1.7.10 (The double adjoint: $T \subset T^{**} = \overline{T}$)
Let T be densely defined. The following statements hold:

1. *If T is symmetric, $T \subset T^*$.*
2. *If T^* is densely defined we have $T \subset T^{**}$ and* 3 *$T^{**} = \overline{T}$.*

Proof. 1 For $y \in \mathrm{dom}(T)$ let $\Phi_y : \mathrm{dom}(T) \to \mathbb{K}$ given by $x \mapsto \langle Tx, y\rangle$. We have $\|\Phi_y(x)\| \leqslant \|T\|\|y\|\|x\| = c\|x\|$, so Φ_y is bounded on $\mathrm{dom}(T)$, so $\mathrm{dom}(T) \subset \mathrm{dom}(T^*)$. Thus, for all $x, y \in \mathrm{dom}(T)$

$$\langle x, T^*y\rangle \overset{1.7.7}{=} \langle Tx, y\rangle \overset{1.7.3}{=} \langle x, Ty\rangle$$

holds, implying $T \equiv T^*$ on $\mathrm{dom}(T)$. Therefore, $T \subset T^*$.

2 Let $(y_n)_{n\in\mathbb{N}} \subset \mathrm{dom}(T^*)$ converging to $y \in \mathscr{H}$ and $T^*y_n \to z$. For $x \in \mathrm{dom}(T)$

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

holds. This implies that $x \mapsto \langle Tx, y\rangle$ is continuous on $\mathrm{dom}(T)$, thus $y \in \mathrm{dom}(T^*)$ holds. Furthermore, this implies $Ty^* = z$ **(sure?)**.

3

Example 1.7.11 (Not closable operator + its adjoint)
Let $D := \{y \in \ell^2 : \exists N \in \mathbb{N} : y_n = 0 \ \forall n > N\}$ and $x \in \ell^2 \setminus D$ and define

$$T : \mathrm{span}(x) + D =: \mathrm{dom}(T) \to \ell^2, \ cx + d \mapsto cx, \quad c \in \mathbb{C}, d \in D.$$

Then T is linear and densely defined but not closable (lemma 1.7.2): Consider $d_n := (-x_1, \ldots, -x_n, 0, \ldots) \in D$. Then $\hat{x} := (x - d_n)_{n\in\mathbb{N}}$ converges to zero but $T\hat{x}$ does not.

$\mathrm{dom}(T^*) = \{x\}^\perp$, $T^*|_{\mathrm{dom}(T^*)} \equiv 0$. ◇

TODO A densely defined operator with finite-dimensional codomain has a densely defined adjoint if and only if it is continuous, since the only dense subspace of a finite-dimensional Hausdorff TVS is the entire space, and every linear operator $T : \mathbb{K}^n \to V$, where V is a topological vector space, is continuous.

The inverse mapping theorem generalises to closed operators.

The spectral theorem above was proved by Hilbert (around 1906) for a self-adjoint operator (disguised as a bounded quadratic form in infinitely many variables), whereas the complete result for a normal (and possibly unbounded) operator is due to von Neumann. [11]

1.8 Spectral properties of unbounded operators

From now on let $T : \mathscr{H} \supset \operatorname{dom}(T) \subset \mathscr{H}$ be densely defined.

DEFINITION 1.8.1

1. The set $\rho(T)$ defined by

$$\{\lambda \in \mathbb{C} : \lambda I - T : \operatorname{dom}(T) \to \mathscr{H} \text{ has a bounded inverse in } \mathscr{H}\}$$

is the resolvent set and $\sigma(T) = \mathbb{C} \setminus \rho(T)$ the spectrum of T.

2. The mapping $R : \rho(T) \to L(H)$, $R_\lambda = R(\lambda) := (\lambda I - T)^{-1}$ is called the resolvent mapping.

resolvent set

resolvent mapping

THEOREM 1.8.1: PROPERTIES OF THE RESOLVENT

1. $\rho(T)$ is open and $\sigma(T)$ closed.
2. The resolvent mapping is analytic and the resolvent identity

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

holds.

resolvent identity

Proof. 1. If $\rho(T) = \varnothing$ it is open, so let $\lambda_0 \in \rho(T)$ and choose $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < \|R_{\lambda_0}\|^{-1}$. Then

$$\lambda I - T = (\lambda_0 I - T) + (\lambda - \lambda_0) I = (\lambda_0 I - T)\left[I - (\lambda - \lambda_0) R_{\lambda_0}\right]$$

holds and by the NEUMANN series we have

$$\left[I - (\lambda - \lambda_0)(\lambda_0 I - T)^{-1}\right]^{-1} = \sum_{n=0}^{\infty} ((\lambda - \lambda_0)(\lambda_0 I - T)^{-1})^n$$

Hence $\lambda I - T$ is invertible and analytic:

$$R_\lambda = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R_{\lambda_0}^{n+1}.$$

A function is analytic if $f(z) = \sum_{n=0}^{\infty} (z - z_0)^n x_n$.

2. A formal calculations yields

$$(\lambda I - T)(\mu I - T)\left[(\lambda I - T)^{-1} - (\mu I - T)^{-1}\right] = (\mu - \lambda) I$$

and a similar identity can be derived by multiplying from the right side. So, only the (easy) inspection of domains is missing. □

Corollary 1.8.2 (CAYLEY transform (HW10 Hannover WiSe11))
Let $T = T^* \in L(\mathscr{H})$ and $V := (T - i)(T + i)^{-1}$ its CAYLEY *transform. Then V is unitary and $T = i(1 + V)(1 - V)^{-1}$ holds.*

Proof. From $(T^{-1})^* = (T^*)^{-1} = T^{-1}$ it follows that

$$V V^* = (T - i)(T + i)^{-1}(T - i)^{-1}(T + i),$$

where $T + i$ is invertible due to corollary 1.1.5. By the resolvent identity we have

$$(T + i)^{-1}(T - i)^{-1} = \frac{(T - i)^{-1} - (T + i)^{-1}}{i - (-i)}$$

and thus

$$\begin{aligned}
VV^* &= \frac{1}{2i}(T-i)\left((T-i)^{-1}-(T+i)^{-1}\right)(T+i) \\
&= \frac{1}{2i}\left[(T-i)(T-i)^{-1}(T+i)-(T-i)(T+i)^{-1}(T+i)\right] \\
&= \frac{T+i-(T-i)}{2i} = 1.
\end{aligned}$$

It follows analogously that $V^*V = 1$.

Rearranging gives

$$V(T+i) = T-i \implies (V-1)T = -i(V+1),$$

implying $T = -i \cdot (V-1)^{-1}(V+i)$. Now, $V-1$ is invertible as by the resolvent identity we have

$$\begin{aligned}
V-1 &= (T-i)(T+i)^{-1} - 1 = (T-i)\left[(T+i)^{-1}-(T-i)^{-1}\right] \\
&= \frac{T-i}{-2i}(T+i)^{-1}(T-i)^{-1},
\end{aligned}$$

implying $(V-1)^{-1} = -\frac{i}{2}(T-i)(T+i)^{-1}(T-i)^{-1}$.

As T is self-adjoint and V is unitary (U) we have

$$\begin{aligned}
T &= T^* = (-i\cdot(V-1)^{-1}(V+1))^* \\
&= \overline{-i}\cdot(V^*+1)(V^*-1)^{-1} \overset{\text{U}}{=} i\cdot V^{-1}(1+V)(V^*-1)^{-1} \\
&\overset{\text{U}}{=} i(1+V)(1-V)^{-1}.
\end{aligned}$$

□

2 Fourier Analysis, Distributions and Sobolev spaces

2.1 The Fourier transform on $L^1(\mathbb{R}^n)$

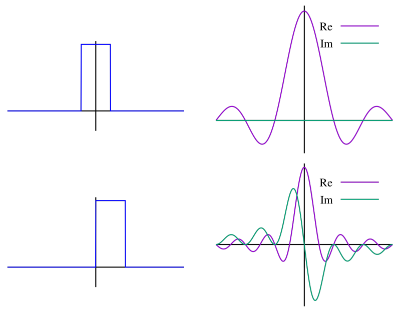

Figure 2: The unit pulse $\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}$, its translated version $\mathbb{1}_{[0,1]}$ and their Fourier transforms, $\sqrt{\frac{2}{\pi}}\frac{\sin(\frac{x}{2})}{x}$ and $\frac{1}{\sqrt{2\pi}}\left(\frac{\sin(x)}{x}+i\left(\frac{\cos(x)-1}{x}\right)\right)$, showing that translation (that is, "delay") in the time domain is interpreted as complex phase shifts in the frequency domain. [19]

Definition 2.1.1 (Fourier transform on $L^1(\mathbb{R}^n)$)
The Fourier transform of $f \in L^1(\mathbb{R}^n)$ is

$$\mathcal{F}f(t) := \hat{f}(t) := (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,t\rangle}\,\mathrm{d}x$$

for $t \in \mathbb{R}^n$.

Remark 2.1.2 In other contexts and fields other definitions of $\hat{f}$ are customary, i.e a factor of $(2\pi)^{-n}$ instead of $(2\pi)^{-\frac{n}{2}}$. One can interpret $|\hat{f}(t)|$ as the "strength with which the "frequency" t appears in f.

Lemma 2.1.3 ($\mathcal{F} \in L(L^1, L^\infty)$, uniform continuity)
The operator $\mathcal{F} : L^1(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ is linear and bounded. Furthermore, $\hat{f}$ is uniformly continuous for $f \in L^1(\mathbb{R})$.

Proof. The linearity is clear. For $t \in \mathbb{R}^d$ and $f \in L^1(\mathbb{R}^d)$

$$|\mathcal{F}f(t)| \overset{\triangle\neq}{\leqslant} (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} |f(x)|\,\mathrm{d}x = (2\pi)^{-\frac{n}{2}} \|f\|_1 < \infty$$

holds, as $|e^{ix}| = 1$ for all $x \in \mathbb{R}$. Thus $\mathcal{F}f \in L^\infty$ holds.

For $\varepsilon > 0$ and $x, y \in \mathbb{R}$

$$\begin{aligned}\left|\hat{f}(x) - \hat{f}(y)\right| &= \frac{1}{\sqrt{2\pi}} \left|\int_{\mathbb{R}} f(t)(e^{-ixt} - e^{-iyt})\,\mathrm{d}t\right| \\ &\overset{\triangle\neq}{\leqslant} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(t)||e^{-ixt} - e^{-iyt}|\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(t)||e^{-i(x-y)t} - 1|\,\mathrm{d}t \xrightarrow{|x-y|\to 0} 0\end{aligned}$$

holds by the theorem of Lebesgue as $|e^{-i(x-y)t} - 1| \overset{\triangle\neq}{\leqslant} 2$ and $f \in L^1(\mathbb{R})$ hold. Therefore there exists a $\delta_\varepsilon > 0$ such that $|x - y| \leqslant \delta_\varepsilon$ implies $\left|\hat{f}(x) - \hat{f}(y)\right| \leqslant \varepsilon\sqrt{2\pi}$. □

Lemma 2.1.4 (Compact support and Fourier transforms)
If $f \in L^2(\mathbb{R})\backslash\{0\}$ is compactly supported, $\hat{f}$ can't be. If in turn $\hat{f}$ is compactly supported, $f \in L^2(\mathbb{R})$ can't be.

Proof. Fourierreihe ist Polynom □

does this result also hold for other p and / or $\mathbb{R}^n$?

Remark 2.1.5 (HW 6-2) We have

$$\int_{\mathbb{R}^n} f(t)\,\mathrm{d}t = \int_{\mathbb{R}^n} f(t) e^{-i\langle t,0\rangle}\,\mathrm{d}t = (2\pi)^{n/2} \hat{f}(0).$$

Assuming that f is differentiable and $t \mapsto tf(t) \in L^1(\mathbb{R})$ we have

$$\int_{\mathbb{R}} tf(t)\,\mathrm{d}t = -\hat{f}'(0)$$

by partial integration and the same reasoning as above.

Definition 2.1.6 (Translation/Modulation/Dilation)
For $y \in \mathbb{R}^n$, $\lambda > 0$, $k \in \{1, \infty\}$ let $T_y, M_y, D_\lambda : L^k(\mathbb{R}^n) \to L^k(\mathbb{R}^n)$ be the translation, modulation and dilation operator given by

$$T_y f(\cdot) := f(\cdot - y), \quad M_y f(\cdot) = f(\cdot) e^{i\langle y, \cdot \rangle}, \quad D_\lambda f(\cdot) = f\left(\frac{\cdot}{\lambda}\right).$$

Lemma 2.1.7 (Properties of Translation/Modulation/Dilation)
For $f \in L^1(\mathbb{R}^n)$, $y \in \mathbb{R}$ and $\lambda > 0$

1. $\mathcal{F}(T_y(f)) = M_{-y}(\mathcal{F}(f))$,
2. $\mathcal{F}(M_y(f)) = T_y(\mathcal{F}(f))$,
3. $\mathcal{F}(D_\lambda(f)) = \lambda^n D_{\lambda^{-1}}(\mathcal{F}(f))$,
4. $g(x) = \overline{f(-x)}\ \forall x \in \mathbb{R}^n \implies \mathcal{F}(g) = \overline{\mathcal{F}(f)}$,
5. $g(x) = -ix_j f(x)$ *for some* $g \in L^1(\mathbb{R}^n)$ *implies the differentiability of* $\mathcal{F}(f)$ *and* $\partial_j(\mathcal{F}(f)) = \mathcal{F}(g)$.

hold.

Proof. 1 By the translation invariance of the Lebesgue measure it holds for all $\xi, y \in \mathbb{R}^n$

$$\begin{aligned}
\mathcal{F}(T_y(f))(\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x-y) e^{-i\langle x, \xi\rangle}\,\mathrm{d}x \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i\langle y, \xi\rangle} f(x-y) e^{-i\langle x-y, \xi\rangle}\,\mathrm{d}x \\
&\overset{(\mathrm{T})}{=} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\langle -y, \xi\rangle} f(x) e^{-i\langle x, \xi\rangle}\,\mathrm{d}x = M_{-y}\,\mathcal{F}(f)(\xi).
\end{aligned}$$

2 For $\xi, y \in \mathbb{R}^n$ we have

$$\begin{aligned}
\mathcal{F}(M_y(f(\xi))) &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x, \xi\rangle} e^{i\langle x, y\rangle}\,\mathrm{d}x \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x, \xi - y\rangle}\,\mathrm{d}x = T_y(\mathcal{F}(f(\xi))).
\end{aligned}$$

3 For $\xi \in \mathbb{R}^n$ and $\lambda > 0$ we have

$$\begin{aligned}
\mathcal{F}(D_\lambda(f(\xi))) &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f\left(\frac{x}{\lambda}\right) e^{-i\langle x, \xi\rangle}\,\mathrm{d}x \\
&= \frac{\lambda^n}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle \lambda x, \xi\rangle}\,\mathrm{d}x = \lambda^n D_{\lambda^{-1}}(\mathcal{F}(f(\xi)))
\end{aligned}$$

4 For $\xi \in \mathbb{R}^n$ we have

$$\begin{aligned}
\mathcal{F}(g(\xi)) &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} \overline{f(-x)} e^{-i\langle x, \xi\rangle}\,\mathrm{d}x \\
&= \overline{\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(-x) e^{-i\langle -x, \xi\rangle}\,\mathrm{d}x} = \overline{\mathcal{F}(f(\xi))}.
\end{aligned}$$

5 For $\xi \in \mathbb{R}^n$ and $j \in \{1, \dots, n\}$ we have

$$\begin{aligned}
\frac{\hat{f}(\xi + he_j) - \hat{f}(\xi)}{h} &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(x) \frac{e^{-i\langle x, \xi + he_j\rangle} - e^{-i\langle x,\xi\rangle}}{h} \,\mathrm{d}x \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,\xi\rangle} \frac{e^{-ihx_j} - 1}{h} \,\mathrm{d}x \\
&\xrightarrow[\text{Lebesgue L'Hôpital}]{h\to 0} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,\xi\rangle} (-ix_j) \,\mathrm{d}x. \qquad \square
\end{aligned}$$

Lemma 2.1.8 (Convolution lemma)
For $f, g \in L^1(\mathbb{R}^n)$, it holds that $f * g \in L^1(\mathbb{R}^n)$ and $\widehat{f * g} = (2\pi)^{\frac{n}{2}} \hat{f}\hat{g}$.

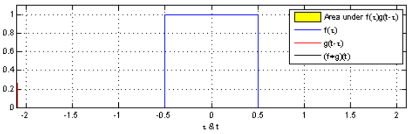

Figure 3: The convolution of the unit pulse (cf. Figure 2) with itself yields the standard hat function. [17]

Proof. For $\xi \in \mathbb{R}^n$ we have with Fubini's theorem

$$\begin{aligned}
\mathcal{F}((f * g)(\xi)) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) g(x-y) \,\mathrm{d}y \; e^{-i\langle x,\xi\rangle} \,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} f(y) \left(\int_{\mathbb{R}^n} g(x-y) e^{-i\langle x-y,\xi\rangle} \,\mathrm{d}x \right) e^{-i\langle y,\xi\rangle} \,\mathrm{d}y \\
&= (2\pi)^{\frac{n}{2}} \mathcal{F}(f(\xi)) \mathcal{F}(g(\xi)). \qquad \square
\end{aligned}$$

The first claim follows from

Lemma 2.1.9 (Young's inequality)
For $f \in L^1(\mathbb{R}^n)$, $g \in L^p(\mathbb{R}^n)$, $p \in [1, \infty]$ we have $\|f * g\|_p \leqslant \|f\|_1 \|g\|_p$.

Proof. Using the Minkowski inequality and that $f * g = g * f$ one obtains for $p \in [1, \infty)$

$$\begin{aligned}
\left(\int_{\mathbb{R}} |(g * f)(x)|^p \,\mathrm{d}x \right)^{\frac{1}{p}} &= \left(\int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(x-y) g(y) \,\mathrm{d}y \right|^p \mathrm{d}x \right)^{\frac{1}{p}} \\
&\leqslant \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y) g(y)|^p \,\mathrm{d}x \right)^{\frac{1}{p}} \mathrm{d}y \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y)|^p \,\mathrm{d}x \right)^{\frac{1}{p}} |g(y)| \,\mathrm{d}y \\
&= \|f\|_p \|g\|_1. \qquad \square
\end{aligned}$$

A longer proof is given in lemma A.3.1.

Theorem 2.1.1: Riemann-Lebesgue Lemma

For $f \in L^1(\mathbb{R}^n)$ we have $\hat{f} \in \mathcal{C}_0(\mathbb{R}^n)$, i.e. $\mathcal{F}(L^1(\mathbb{R}^n)) \subset \mathcal{C}_0(\mathbb{R}^n)$.

$\mathcal{C}_0(\mathbb{R}^n) := \{f \in \mathcal{C}(\mathbb{R}^n) : \lim_{\|x\|\to\infty} |f(x)| \to 0\}$

Proof. The continuity of $\hat{f}$ follows from the theorem of Lebesgue:

$$\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,\xi+h\rangle} \xrightarrow{h\to 0} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,\xi\rangle},$$

where a suitable majorant is $(2\pi)^{-\frac{n}{2}} \|f\|_1$.

Let $f := \mathbb{1}_{\prod_{k=1}^n [a_k, b_k]}$. Then for $\xi_k > 0$ for all $k \in \{1, \dots, n\}$

$$\begin{aligned}\int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} e^{-i\langle x,\xi\rangle} dx_1 \dots dx_n &= \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} \exp\left(-i\sum_{j=1}^n x_j\xi_j\right) dx_1 \dots dx_n \\ &= \prod_{k=1}^n \int_{a_k}^{b_k} e^{-ix_k\xi_k} dx_k \\ &= \prod_{k=1}^n \frac{i(e^{-ib_k\xi_k} - e^{-ia_k\xi_k})}{\xi_k}.\end{aligned}$$

holds, implying

$$\left|\hat{f}(\xi)\right| \leqslant \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{k=1}^n \frac{\left|e^{-ib_k\xi_k} - e^{-ia_k\xi_k}\right|}{|\xi_k|} \xrightarrow{\|\xi\|\to\infty} 0,$$

as

$$\left|e^{-ib\xi} - e^{-ia\xi}\right| = 2\left|\sin\left(\frac{a-b}{2}\cdot\xi\right)\right| \leqslant 2$$

is bounded. Then same conclusion can be drawn for finite sums of such functions. Let $f \in L^1(\mathbb{R}^n)$ then for $\varepsilon > 0$ there exists a step function h on $\mathbb{R}^n$ with $\|f - h\|_1 \leqslant \varepsilon$. By lemma 2.1.3 we have

$$\begin{aligned}|\mathcal{F}(f(\xi))| &\overset{\triangle\neq}{\leqslant} |\mathcal{F}(f-h)(\xi)| + |\mathcal{F}(h(\xi)| \leqslant \frac{\|f-h\|_1}{(2\pi)^{\frac{n}{2}}} + |\mathcal{F}(h(\xi)| \\ &\leqslant \varepsilon(2\pi)^{-\frac{n}{2}} + |\mathcal{F}(h(\xi)| \xrightarrow{|\xi|\to\infty} \varepsilon(2\pi)^{-\frac{n}{2}}.\end{aligned}$$

□

THEOREM 2.1.2: FOURIER INVERSION FORMULA

Let $f \in L^1(\mathbb{R}^n)$ such that $\hat{f} \in L^1(\mathbb{R}^n)$. Then

$$f(x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\langle x,\xi\rangle}\, \mathrm{d}\xi = \mathcal{F}^{-1}(\hat{f})(x)$$

holds almost everywhere.

Proof. Define the hat function

$$\varphi: \mathbb{R}^n \to \mathbb{R},\ x \mapsto \begin{cases} 1 - \|x\|, & \text{if } \|x\| < 1, \\ 0, & \text{else.}\end{cases}$$

and for $\lambda > 0$ the FEJÉR kernel

$$F_\lambda := \lambda D_\lambda F, \quad \text{where } F(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \varphi(t)e^{i\langle x,t\rangle}\,\mathrm{d}t$$

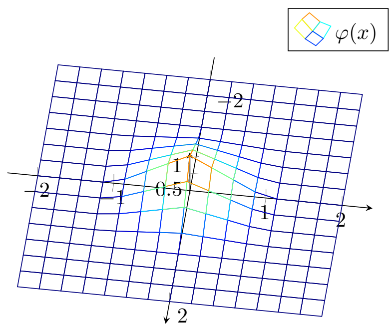

Figure 4: The hat function φ in two dimensions.

CAN WE SHOW $F_\lambda \in \mathcal{S}(\mathbb{R})$??

We will only prove the statement for $n = 1$, as the proof is very similar for $n > 1$.

With the substitution $t \to \lambda t$ we obtain

$$F_\lambda(x) = \frac{\lambda}{2\pi}\int_{-1}^{1}(1 - \lambda|t|)e^{ix\lambda t}\,\mathrm{d}t = \frac{1}{2\pi}\int_{-1}^{1}\left(1 - \frac{|t|}{\lambda}\right)e^{ixt}\,\mathrm{d}t,$$

implying $F_\lambda(-x) = \sqrt{2\pi}\widehat{D_\lambda\varphi}$.

We know that $\varphi(x) = (g * g)(x)$, where $g := \mathbb{1}_{[-1/2,1/2]}$ is the unit pulse (cf. Figure 3), whose Fourier transform is $\sqrt{\frac{2}{\pi}}\frac{\sin(x/2)}{x}$ (cf. Figure 2). By lemma 2.1.7 (4) and lemma 2.1.8

$$\begin{aligned} F_\lambda(-x) &= \sqrt{2\pi}\widehat{D_\lambda \varphi} = \sqrt{2\pi}\lambda D_{\lambda^{-1}}\hat{f}(x) = \lambda D_{\lambda^{-1}}\hat{g}^2(x) \\ &= \lambda^1 D_{\lambda^{-1}}\frac{2}{\pi}\cdot\frac{\sin^2\left(\frac{-x}{2}\right)}{(-x)^2} = \frac{2\sin^2\left(\frac{\lambda x}{2}\right)}{\lambda\pi x^2}. \end{aligned}$$

holds for $x \in \mathbb{R}$, as $\sin^2$ is an even function. Therefore F_λ is even as well. It is now easy to see that this kernel satisfies $F_\lambda \geqslant 0$, $F_\lambda(x) \xrightarrow{x\to\pm\infty} 0$ (as sin is bounded) and $\int_{\mathbb{R}} F_\lambda(x)\,\mathrm{d}x = 1$:

$$\begin{aligned} \int_{\mathbb{R}} F_\lambda(x)\,\mathrm{d}x &= \sqrt{2\pi}\int_{\mathbb{R}} \widehat{D_\lambda\varphi}(x)\,\mathrm{d}x \overset{(12)}{=} (D_\lambda\varphi)(0) = \varphi(0) = (g*g)(0) \\ &= \int_{\mathbb{R}} g(-y)g(y)\,\mathrm{d}y = \int_{\mathbb{R}} g(y)^2\,\mathrm{d}y = \int_{\mathbb{R}} g(y)\,\mathrm{d}y = 1. \end{aligned}$$

Using a standard mollification argument we now show

$$\|f - f * F_\lambda\|_1 \xrightarrow{\lambda\to\infty} 0 \quad \text{and} \quad (f * F_\lambda)(x) \to f(x) \text{ a.e.}$$

(1) As $\int_{\mathbb{R}} F_\lambda(x)\,\mathrm{d}x = 1$,

$$f(x) = \int_{\mathbb{R}} f(x)F_\lambda(x-y)\,\mathrm{d}y$$

holds, implying

$$f(x) - (f * F_\lambda)(x) = \int_{\mathbb{R}} \big(f(x) - f(y)\big)F_\lambda(x-y)\,\mathrm{d}y.$$

(2) First let f be continuous and compactly supported and thus uniformly continuous. For $\varepsilon > 0$

$$\begin{aligned} \int_{|x-y|<\delta} |f(x) - f(y)|F_\lambda(x-y)\,\mathrm{d}y &\leqslant \varepsilon\int_{|x-y|<\delta} F_\lambda(x-y) \leqslant \varepsilon, \\ \int_{|x-y|\geqslant\delta} |f(x) - f(y)|F_\lambda(x-y)\,\mathrm{d}y &\leqslant 2\|f\|_{L^\infty}\int_{|x-y|\geqslant\delta} F_\lambda(x-y) \end{aligned}$$

is possible due to the uniform continuity of f ($|x-y| < \delta \implies |f(x) - f(y)| < \varepsilon$).

Thus

$$|f(x) - (f * F_\lambda)(x)| \overset{\triangle\neq}{\leqslant} \varepsilon + 2\|f\|_{L^\infty}\int_{|x-y|\geqslant\delta} F_\lambda(x-y).$$

We can choose λ large enough such that $\int_{|x-y|\geqslant\delta} F_\lambda(x-y) \leqslant \varepsilon$ as $F_\lambda(t) \xrightarrow{\lambda\to\infty} 0$ holds for $t \neq 0$. Thus $f * F_\lambda$ converges uniformly to f.

(3) The substitution $t(y) := x - y$ yields

$$\begin{aligned} \|f - f * F_\lambda\|_1 &\overset{\triangle\neq}{\leqslant} \int_{\mathbb{R}\times\mathbb{R}} |f(x) - f(y)|F_\lambda(x-y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{\mathbb{R}\times\mathbb{R}} |f(x) - f(x-t)|F_\lambda(t)\,\mathrm{d}t\,\mathrm{d}x. \end{aligned}$$

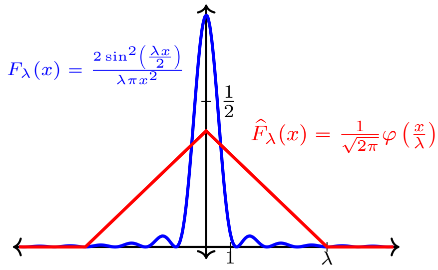

Figure 5: The Fejér kernel and its Fourier transform for $\lambda = 5$.

As above we split the inner integral into $|t| > \delta$ and $|t| \geqslant \delta$:

$$\begin{aligned}
&\int_{\mathbb{R}} \int_{-\delta}^{\delta} |f(x) - f(x-t)| F_\lambda(t) \,\mathrm{d}t \,\mathrm{d}x \\
&\qquad \overset{\text{(F)}}{=} \int_{-\delta}^{\delta} F_\lambda(t) \left(\int_{\mathbb{R}} |f(x) - f(x-t)| \,\mathrm{d}x \right) \mathrm{d}t \\
&\qquad \leqslant \sup_{|t|<\delta} \int_{\mathbb{R}} |f(x) - f(x-t) \,\mathrm{d}x,
\end{aligned}$$

which is small for sufficiently small $\delta > 0$.

Proceeding analogously, we obtain

$$\int_{|t|\geqslant\delta} F_\lambda(t) \int_{\mathbb{R}} |f(x) - f(x-t)| \,\mathrm{d}x \,\mathrm{d}t \leqslant 2\|f\|_1 \int_{|t|\geqslant\delta} F_\lambda(t) \,\mathrm{d}t,$$

which again can be made arbitrarily small for fixed $\delta > 0$ by choosing λ large enough. Thus $f * F_\lambda \xrightarrow{L^1} f$.

4 Now let $f \in L^1$ and $\varepsilon > 0$. As the continuous and compactly supported functions are dense in L^1, there exists a continuous compactly supported g such that $\|f - g\|_1 < \varepsilon$. We have

$$\|f - f * F_\lambda\|_1 \overset{\triangle\neq}{\leqslant} \|f - g\|_1 + \|g - g * F_\lambda g\|_1 + \|g * F_\lambda - f * F_\lambda\|_1.$$

We have dealt with the first two terms. As convolution is distributive

$$\begin{aligned}
\|g * F_\lambda - f * F_\lambda\|_1 &\overset{\triangle\neq}{\leqslant} \int_{\mathbb{R}\times\mathbb{R}} |g(y) - f(y)| F_\lambda(x-y) \,\mathrm{d}y \,\mathrm{d}x \\
&= \int_{\mathbb{R}} |g(y) - f(y)| \left(\int_{\mathbb{R}} F_\lambda(x-y) \,\mathrm{d}x \right) \mathrm{d}y \\
&= \|g - f\|_1 < \varepsilon,
\end{aligned}$$

implying $f * F_\lambda \xrightarrow{L^1} f$.

5 First we show that the pointwise convergence holds for all $x \in \mathbb{R}$ where f is continuous: **TODO**

6 We now consider all Lebesgue-points of $f - f * F_\lambda$, i.e. all $x \in \mathbb{R}$ such that Lebesgue-point

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} (f - f * F_\lambda)(y) \,\mathrm{d}y = (f - f * F_\lambda)(x),$$

as almost all $x \in \mathbb{R}$ are Lebesgue points for any L^1 function. **TODO**

Using the same substitution we obtain for $x \in \mathbb{R}$

$$\begin{aligned}(f * F_\lambda)(x) &= \int_{\mathbb{R}} f(t)\left[\lambda \int_{-1}^{1} (1-|\theta|)e^{i(x-t)\lambda\theta}\,\mathrm{d}\theta\right]\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} f(t)\left[\int_{-\lambda}^{\lambda}\left(1-\frac{|u|}{\lambda}\right)e^{iu(x-t)}\,\mathrm{d}u\right]\mathrm{d}t \\ &= \int_{-\lambda}^{\lambda}\left(1-\frac{|u|}{\lambda}\right)\left[\frac{1}{2\pi}\int_{\mathbb{R}} f(t)e^{-itu}\,\mathrm{d}t\right]e^{iux}\,\mathrm{d}u \\ &= \int_{|\theta|\leqslant\sqrt{\lambda}}\left(1-\frac{|u|}{\lambda}\right)\hat{f}(\theta)e^{iux}\,\mathrm{d}u \\ &\quad + \int_{\sqrt{\lambda}\leqslant|\theta|\leqslant\lambda}\left(1-\frac{|u|}{\lambda}\right)\hat{f}(\theta)e^{iux}\,\mathrm{d}u \\ &\xrightarrow{\lambda\to\infty} \int_{\mathbb{R}} \hat{f}(\theta)e^{iux}\,\mathrm{d}u\end{aligned}$$

□

Remark 2.1.10 The same reasoning as in remark 2.1.5 implies

$$\int_{\mathbb{R}^n} f(t)\,\mathrm{d}t = \int_{\mathbb{R}^n} f(t)e^{i\langle t,0\rangle}\,\mathrm{d}t = \check{f}(0). \tag{12}$$

Example 2.1.11 (HW 6-5, WiSe01) The operator

$$K\,\mathcal{F}^{-1}\,N\,\mathcal{F}\,M : L^2(\mathbb{R}) \to L^2(\mathbb{R})$$

is compact, where M, N and K are multiplication operators on $L^2(\mathbb{R})$ for $k, m, n \in \mathcal{C}_c(\mathbb{R})$, i.e. $Mf(x) := m(x)f(x)$ for $f \in L^2(\mathbb{R})$. ◇

Proof. We first show that $M : L^2(\mathbb{R}) \to L^1(\mathbb{R})$ (and therefore N and K) are well-defined. As $m \in \mathcal{C}_c(\mathbb{R})$, there exists a $M > 0$ such that $m(x) = 0$ for all $|x| > M$ and a $C > 0$ such that $|m(x)| \leqslant C$ As $L^2(I) \subset L^1(I)$ on compact intervals I, we have

$$\|Mf\|_1 = \int_{\mathbb{R}} |m(x)||f(x)|\,\mathrm{d}x \leqslant C\int_{-M}^{M} |f(x)|\,\mathrm{d}x \leqslant \tilde{C}\|f\|_{L^2([-M,M])} \leqslant \tilde{C}\|f\|_{L^2(\mathbb{R})} < \infty.$$

We now show that $\mathcal{F}\,M$ (and thus $\mathcal{F}^{-1}\,N$) is well-defined. Let $f \in L^2(\mathbb{R})$. Thus

$$(\mathcal{F}^{(-1)}\,Mf)(x) = \frac{1}{\sqrt{2\pi}}\int_{-M}^{M} m(\xi)e^{\pm i\xi x}\,\mathrm{d}\xi \leqslant \frac{C}{\sqrt{2\pi}}\int_{-M}^{M} e^{\pm i\xi x}\,\mathrm{d}\xi = \frac{C}{\sqrt{2\pi}}\frac{\sin(Mx)}{x}$$

and

$$\|\,\mathcal{F}^{(-1)}\,Mf\|_2 = \frac{C^2}{2\pi}\int_{\mathbb{R}} \frac{\sin^2(Mx)}{x^2}\,\mathrm{d}x = \frac{C^2}{2\pi}M\pi = \frac{1}{2}MC^2 < \infty$$

holds.

As the compact operators form a two sided ideal in $L(L^2(\mathbb{R}))$ it suffices to show that M is compact. Let $(f_n)_{n\in\mathbb{N}} \subset L^2(\mathbb{R})$ be bounded, i.e. there exists a $C > 0$ such that $\|f_n\| \leqslant C$ holds for all $n \in \mathbb{N}$. Then $\overline{M((f_n)_{n\in\mathbb{N}})}$ is compact: Let $(g_n)_{n\in\mathbb{N}} \subset \overline{M((f_n)_{n\in\mathbb{N}})}$, i.e. $g_n := m \cdot f_n$. We need to show that g_n has a convergent subsequence. □

2.2 The Uncertainty Principle

The idea of the uncertainty principle (UP) is that it is impossible for a function to both vanish outside some finite interval and have only frequency components smaller than some constant. [12]

Theorem 2.2.1: Heisenberg's UP I

For $f \in L^2(\mathbb{R})$ and $a, b \in \mathbb{R}$ it holds that

$$\frac{1}{2}\|f\|_2^2 \leqslant \left(\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} (\xi - b)^2 |\hat{f}(\xi)|^2 \,\mathrm{d}\xi\right)^{\frac{1}{2}}.$$

Proof. We first show that we can assume $a = b = 0$ without loss of generality. Let the statement hold with $a = b = 0$ and consider $g(x) := f(x+\alpha)e^{-ix\beta}$ with $g \in L^2(\mathbb{R})$. Then

$$\int_{\mathbb{R}} x^2 |g(x)|^2 \,\mathrm{d}x = \int_{\mathbb{R}} x^2 |e^{-ix\beta}||f(x+\alpha)|^2 \,\mathrm{d}x = \int_{\mathbb{R}} (x-\alpha)|f(x)|^2 \,\mathrm{d}x$$

and

$$\begin{aligned}
\sqrt{2\pi}\hat{g}(\xi) &= \int_{\mathbb{R}} f(x+\alpha)e^{-ix(\xi+\beta)} \,\mathrm{d}x = \int_{\mathbb{R}} f(y)e^{-i(y-\alpha)(\xi+\beta)} \,\mathrm{d}y \\
&= e^{i\alpha(\xi+\beta)} \int_{\mathbb{R}} f(y)e^{-iy(\xi+\beta)} \,\mathrm{d}y
\end{aligned}$$

and by lemma 2.1.7 2

$$\int_{\mathbb{R}} f(y)e^{-iy(\xi+\beta)} \,\mathrm{d}y = \int_{\mathbb{R}} f(y)e^{-iy\xi}e^{-iy\beta} \,\mathrm{d}y = \sqrt{2\pi}\hat{f}(\xi+\beta)$$

and therefore

$$\int_{\mathbb{R}} \xi^2 |\hat{g}(\xi)|^2 \,\mathrm{d}\xi = \int_{\mathbb{R}} \xi^2 |\hat{f}(\xi+\beta)|^2 \,\mathrm{d}\xi = \int_{\mathbb{R}} (\xi-\beta)^2 |\hat{f}(\xi)|^2 \,\mathrm{d}\xi.$$

Assume without loss of generality $\|f\|_2^2 = 1$. Let $f \in \mathcal{S}(\mathbb{R})$. Then with partial integration (boundary terms vanish) and $|f|^2 = f\overline{f}$ we have

$$\begin{aligned}
1 = \int_{\mathbb{R}} |f(x)|^2 \,\mathrm{d}x &= -\int_{\mathbb{R}} x \,\frac{\mathrm{d}}{\mathrm{d}x}|f(x)|^2 \,\mathrm{d}x \\
&= -\int_{\mathbb{R}} xf'(x)\overline{f(x)} + x\overline{f'(x)}f(x) \,\mathrm{d}x
\end{aligned}$$

Thus by Cauchy-Schwarz

$$1 \leqslant 2\int_{\mathbb{R}} |x||f'(x)||f(x)| \,\mathrm{d}x \leqslant 2\left(\int_{\mathbb{R}} x^2|f(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} |f'(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}}$$

By the Plancherel theorem we obtain

$$\left(\int_{\mathbb{R}} |f'(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} = \left(\int_{\mathbb{R}} |\mathcal{F}(f')(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} = \left(\int_{\mathbb{R}} x^2|\hat{f}(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}},$$

as by partial integration

$$\int_{\mathbb{R}} f'(x)e^{-ix\xi} \,\mathrm{d}x = i\xi \int_{\mathbb{R}} f(x)e^{-ix\xi} \,\mathrm{d}x \tag{13}$$

Now let $f \in L^2(\mathbb{R})$. Then there exists a sequence of Schwartz-functions $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R})$ such that $f_n \to f$ pointwise **TODO TODO**. Then $\|f_n\|_2^2 \to \|f\|_2^2$ and $\int_{\mathbb{R}} x^2|f_n(x)|^2 \,\mathrm{d}x \to \int_{\mathbb{R}} x^2|f(x)|^2 \,\mathrm{d}x$.

TODO und dann F-Koeffizienten durch L^1-Koeffizienten approximieren, die man mit Schwartz-Funktionen approximieren kann. □

Theorem 2.2.2: Heisenberg's UP II

For self-adjoint operators S, T on $\mathscr{H}$ and $a, b \in \mathbb{R}$

$$\|(S-aI)f\|\|(T-bI)\| \geqslant \frac{1}{2}|\langle [S,T]f, f\rangle|$$

holds for all $f \in \operatorname{dom}(ST) \cap \operatorname{dom}(TS)$, where $[S,T] := ST - TS$ is the commutator of S and T.

Proof. An easy calculation shows $[S-aI, T-bI] = [S,T]$. As S, T are self-adjoint, so are $S-aI$ and $T-bI$. Thus

$$\begin{aligned}\langle [S,T]f, f\rangle &= \langle (S-aI)(T-bI) - (T-bI)(S-aI)f, f\rangle \\ &= \langle (T-bI)f, (S-aI)f\rangle - \langle (S-aI)f, (T-bI)f\rangle \\ &= 2i \cdot \Im\left(\langle (T-bI)f, (S-aI)f\rangle\right)\end{aligned}$$

holds. With the Cauchy-Schwarz-inequality we have

$$|\langle [S,T]f, f\rangle| \leqslant 2\|(T-bI)f\|\|(S-aI)\|.$$ □

We can now give a second version of the proof of theorem 2.2.1:

Proof. Define

$$\begin{aligned}S &: L^2(\mathbb{R}^n) \to ??,\ f(x) \mapsto xf(x) \\ T &: L^2(\mathbb{R}^n) \cap \mathcal{D}^1(\mathbb{R}^n) \to ??,\ f(x) \mapsto if'(x)\end{aligned}$$

where $\mathcal{D}^1(\mathbb{R}^n)$ denotes the differentiable functions on $\mathbb{R}^n$. For $f \in \operatorname{dom}(ST) \cap \operatorname{dom}(TS)$ we have

$$\begin{aligned}([S,T]f)(x) &= ixf'(x) = i\frac{\mathrm{d}}{\mathrm{d}x}(x \cdot f(x)) \\ &= ixf'(x) - if(x) - ixf'(x) = -if(x)\end{aligned}$$

and by theorem 2.2.2

$$\frac{1}{2}\|f\|_2^2 = \frac{1}{2}|\langle -if(x), f(x)\rangle| \leqslant \|(S-aI)f\|_2\|(T-bI)f\|_2$$

By the Plancherel theorem we have

$$\|(T-bI)f\|_2 = \|\mathcal{F}((T-bI)f)\|_2 = \|(\xi-b)\mathcal{F}(f)\|_2,$$

where the last equality follows from (13). This yields

$$\begin{aligned}\frac{1}{2}\|f\|_2^2 &\leqslant \|(S-aI)f\|_2\|(\xi-b)\mathcal{F}(f)\|_2 \\ &= \left(\int_{\mathbb{R}} |xf(x)-af(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} (\xi-b)^2|\hat{f}(\xi)|^2 \,\mathrm{d}\xi\right)^{\frac{1}{2}},\end{aligned}$$

which is the statement of theorem 2.2.1. □

Remark 2.2.1 We can without loss of generality assume that both integrals on the right hand side of theorem 2.2.1 are finite as otherwise, there is nothing to show. Therefore we have $(x \mapsto x \cdot f(x)) \in L^2(\mathbb{R})$ and by Plancherel $f' \in L^2(\mathbb{R}^n)$.

2.3 Fourier transform on $L^1(G)$ and $L^2(G)$

The theory of Fourier transforms can be generalised to functions on locally compact (abelian) groups, leading to Fourier transforms on e.g. $L^2(\mathbb{T})$ and $\ell_2(\mathbb{Z})$, where result analogous to the Plancherel theorem hold.

Definition 2.3.1 (Topological group)
A topological group is a group G equipped with a topology such that the group operations $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are continuous.

topological group

From now on let G be a topological group.

Corollary 2.3.2 (Translations and inversion-invariance)
The topology is invariant under translation and inversion: if U is open, so are $xU := \{xy : y \in U\}$, $Ux := \{yx : y \in U\}$ and $U^{-1} := \{y^{-1} : y \in U\}$ for all $x \in G$.

Definition 2.3.3 (Locally compact, Hausdorff space)
A topology is locally compact if every point has a compact neighbourhood. In a Hausdorff space points can be separated by open sets.

locally compact

[Wiki:] A Hausdorff TVS is locally compact if and only if it is finite-dimensional.

Example 2.3.4 (Wiki) Every compact Hausdorff space is also locally compact, but not the other way around: $\mathbb{Q}$ with the topology from $\mathbb{R}$ is Hausdorff but not locally compact, since any neighbourhood contains a Cauchy sequence corresponding to an irrational number, which has no convergent subsequence in $\mathbb{Q}$. $\diamond$

Hausdorff space

Definition 2.3.5 (locally compact abelian group)
If the topology of G is locally compact and Hausdorff, G is a locally compact group. If G is also abelian, it is called locally compact abelian group (LCAG).

Definition 2.3.6 (Left/right Haar measure)
A left (right) Haar measure on a locally compact group G is a non-zero Radon measure μ satisfying $\mu(xE) = \mu(E)$ $(\mu(Ex) = \mu(E))$ for all Borel sets $E \subset G$ and all $x \in G$.

Haar measure

Example 2.3.7 (From König: Analysis III)
The Lebesgue integral $I : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathbb{R}$, $f \mapsto \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ is a Haar measure on $\mathbb{R}^d$, as $\int_{\mathbb{R}} f(x-h)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ holds for all $h \in \mathbb{R}^n$. $\diamond$

Theorem 2.3.1: Haar, 1933

Every locally compact group possesses a left (right) Haar measure uniquely determined up to rescaling by a positive number.

Proof. Similar to the proof of the existence of the Lebesgue measure, see [21]. $\square$

Although most of the following can be generalised to locally compact non-abelian groups, we will from now on require G to be a LCAG.

definition 2.3.8 (Pontryagin duality, 1934)
A continuous homomorphism $\gamma : G \to \mathbb{T}$ is a character of G. The dual group of G, $\widehat{G}$, is the set of all its characters.

character
dual group

The circle group $\mathbb{T}$ is isomorphic to, and can hence also be viewed as, $(\mathbb{R}/\mathbb{Z}, +)$ or $(S^1, \cdot)$, where $S^1 := \{z :\in \mathbb{C} : |z| = 1\}$ and $\cdot$ denotes multiplication.

This is a consequence of the first Isomorphism theorem, which states that $\operatorname{ran}(f) \cong G/\ker(f)$ for a homomorphism $f : G \to H$, as $f : \mathbb{R} \to \mathbb{T}$, $x \mapsto \exp(ix/2\pi)$ is surjective with $\ker(f) = \mathbb{Z}$.

TODO, WIKI Die harmonische Analyse betrachtet unitäre Darstellungen von lokalkompakten topologischen Gruppen, d. h. stetige Homomorphismen von der Gruppe in die unitäre Gruppe über einem Hilbertraum versehen mit der starken Operatortopologie. Aufbauend darauf wird die verallgemeinerte Fourier-Transformation von Funktionen auf der Gruppe mittels der irreduziblen Darstellungen der Gruppe definiert. Eine besondere Rolle spielen die eindimensionalen Darstellungen, d.h. Darstellungen in die Kreisgruppe, genannt Charaktere. Diese sind stets irreduzibel. Aus dem Lemma von Schur folgt umgekehrt, dass jede irreduzible, starkstetige unitäre Darstellung einer abelschen lokalkompakten topologischen Gruppe eindimensional, also ein Charakter ist. Für den abelschen Fall reduziert sich die Fourier-Transformation also auf ein Funktional auf den Charakteren.

Einerseits wird die Kreisgruppe zur Definition des Charakters verwendet, andererseits hat die Kreisgruppe auch Charaktere. Die Charaktere der Kreisgruppe $\mathbb{T}$ sind genau die stetigen Homomorphismen $\mathbb{T} \to \mathbb{T}$, und die kann man alle angeben. Jeder Charakter von $\mathbb{T}$ hat die Form $\chi_n(z) := z^n$ für $n \in \mathbb{Z}$. Daher kann man die Menge der Charaktere mit $\mathbb{Z}$ identifizieren. Dass die Menge der Charaktere wieder eine Gruppenstruktur trägt, ist kein Zufall; es handelt sich um einen Spezialfall der allgemeineren Pontrjagin-Dualität.

Remark 2.3.9 (Dual group is LCAG)
Interpreting $\mathbb{T}$ as S^1, $\widehat{G}$ is a group via the group operation

$$(\gamma\omega)(x) := \gamma(x)\omega(x) \in \mathbb{T},$$

unit element $1(x) := 1 \in \mathbb{T}$ and inverse $\gamma^{-1}(x) := \overline{\gamma(x)} = \gamma(x^{-1})$. Equipping $\widehat{G}$ with the topology of compact convergence, under which the group operations are continuous, $\widehat{G}$ becomes LCAG.

Theorem 2.3.2: Examples for dual groups

1. $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{T}$ and $\mathbb{Z}_k$ are LCAGs.
2. $\widehat{\mathbb{R}} \cong \mathbb{R}$ via $x \mapsto \gamma_x$, where $\gamma_x(y) := e^{ixy}$.
3. $\widehat{\mathbb{T}} \cong \mathbb{Z}$ via $m \mapsto \gamma_m$, where $\gamma_m(\theta) := \theta^m$.
4. $\widehat{\mathbb{Z}} \cong \mathbb{T}$ via $\theta \mapsto \gamma_\theta$, where $\gamma_\theta(m) := \theta^m$.
5. $\widehat{\mathbb{Z}}_k \cong \mathbb{Z}_k$ via $m \mapsto \gamma_m$, where $\gamma_m(n) := \exp\left(2\pi i \frac{mn}{k}\right)$.

Proof. TODO □

The cartesian product of two LCAGs is an LCAG.

Lemma 2.3.10

For LCAGs $(G_k)_{k=1}^n$, $\widehat{\bigotimes_{k=1}^n G_k} \cong \bigotimes_{k=1}^n \widehat{G}_k$ *holds.*

In particular, $\widehat{R}^n \cong \mathbb{R}^n$, $\widehat{T}^n \cong \mathbb{Z}^n$, $\widehat{Z}^n \cong \mathbb{T}^n$ and $\widehat{G} \cong G$ holds for any finite LCAG G, as any finite abelian group is the finite product of Z_m for $m \in \mathbb{Z}$.

Definition 2.3.11 (Fourier transform on $L^1(G)$)

Let $f \in L^1(G)$ and μ be the left-invariant Haar measure on G. For $\gamma \in \widehat{G}$ the Fourier transform of f is

$$\mathcal{F} f(\gamma) := \hat{f}(\gamma) := \int_G f(x)\overline{\gamma(x)}\, \mathrm{d}\mu(x).$$

Denote by $\mathcal{C}_0(\widehat{G})$ the set of continuous and bounded functions on $\widehat{G}$.

Theorem 2.3.3: TODO

We have $\mathcal{F} : L^1(G) \to \mathcal{C}_0(\widehat{G})$.

Theorem 2.3.4: Plancherel

The Fourier transform on $L^1(G) \cap L^2(G)$ uniquely extends to a unitary isomorphism from $L^2(G)$ to $L^2(\widehat{G})$.

Theorem 2.3.5: Pontryagin duality

The map $\Phi : G \to \widehat{\widehat{G}}$, $(\Phi(x))(\gamma) := \gamma(x)$ is an isomorphism of topological groups.

Theorem 2.3.6: Fourier inversion formula

Let $f \in L^1(G)$ such that $L^1(\widehat{G})$. Then $f(x) = \hat{\hat{f}}(x^{-1})$ holds for almost all $x \in G$, i.e.

$$f(x) = \int_{\widehat{G}} \hat{f}(\gamma(\gamma(x)\, \mathrm{d}\mu(\gamma) \quad \text{a.e. in } G,$$

where μ is the appropriately normalised left-invariant Haar invariant Haar measure on $\widehat{G}$.

If $f \in \mathcal{C}(G)$ these identities hold for all $x \in G$.

Definition 2.3.12 (Convolution on $L^1(G)$)

For $f, g \in L^1(G)$ the convolution of f and g is

$$(f * g)(x) := \int_G f(y)g(y^{-1}x)\, \mathrm{d}\mu(x) \quad \text{a.e. in } G$$

Lemma 2.3.13

For $f, g \in L^1(G))$ we have $\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1$. For $f, g \in L^2(G)$ we have $\widehat{fg} = \hat{f} * \hat{g}$.

The following example show that this theory is powerful enough to deliver properties of the usually separately treated FOURIER transform of sequences.

Example 2.3.14 (FOURIER transform on $\mathbb{Z}$)

Let $G := (\mathbb{Z}, +)$ with the standard topology $\mathcal{P}(\mathbb{Z})$ and the the counting measure μ. The FOURIER transform of $c := (c_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{Z}) = L^1(\mathbb{Z}, \mu)$ is $\hat{c}(x) = \sum_{n\in\mathbb{Z}} c_n x^{-n}$, where $x \in \mathbb{T}$.

By theorem 2.3.3 we have $\hat{c} \in \mathcal{C}(\mathbb{T})$. As $\mathbb{T}$ is compact, $\hat{c} \in \ell^1(\mathbb{T})$

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} \hat{c}(x) x^n \,\mathrm{d}x$$

holds by theorem 2.3.6. As $\ell^1(\mathbb{Z}) \subset \ell^2(\mathbb{Z})$ (cf. FA I, HW 02-04), $\|c\|_{\ell^2(\mathbb{Z})} = \|\hat{c}\|_{L^2(\mathbb{T})}$ holds by theorem 2.3.4. $\diamond$

2.4 The Fourier Transform on $\mathcal{S}(\mathbb{R}^n)$

We use multiindex notation and $|x| := \|x\|_2$ for $x \in \mathbb{R}^n$.

Definition 2.4.1 (Schwartz space and convergence)

The Schwartz space is

Schwartz space

$$\mathcal{S}(\mathbb{R}^n) := \{f \in \mathcal{C}^\infty(\mathbb{R}^n) : \|f\|_{(k,\ell)} < \infty \; \forall k, l \in \mathbb{N}_0\},$$

where

$$\|f\|_{(k,\ell)} := \sup_{x \in \mathbb{R}^n} (1 + |x|^2)^{\frac{k}{2}} \sum_{|\alpha| \leqslant \ell} |D^\alpha f(x)|.$$

$\|f\|_{(k,\ell)} < \infty$ means that all derivatives vanish very quickly (faster than any polynomial).

A sequence $(f_j)_{j \in \mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ converges to $f \in \mathcal{S}(\mathbb{R}^n)$ if

$$f_j \xrightarrow{\mathcal{S}} f \iff \|f_j - f\|_{(k,\ell)} \to 0 \; \forall k, \ell \in \mathbb{N}_0 .$$

Example 2.4.2 (Schwartz function)
For $a > 0$ the function $x \mapsto x^a e^{-a|x|^2}$ is a Schwartz function. ◇

Proof. TODO □

Remark 2.4.3 (Wiki, todo) We have $\mathcal{C}_c^\infty \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{C}^\infty$, where the first inclusion is dense. Furthermore, $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ is dense for $p \in [1, \infty]$ If $f \in \mathcal{S}(\mathbb{R}^n)$ then f is uniformly continuous. $\mathcal{S}(\mathbb{R}^n)$ is separable.

Lemma 2.4.4 (Alternative characterisation)

We have

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

where $\|f\|_{\alpha,\beta} := \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta f(x)|$ *is a norm on* $\mathcal{S}(\mathbb{R}^n)$.

Proof. To verify the norm properties it suffices to show positive definiteness, everything else is clear. **TODO** □

Lemma 2.4.5 (Norm and completing metric on $\mathcal{S}(\mathbb{R}^n)$)

For $N \in \mathbb{N}$

$$\|f\|_{(N)} := \max_{|\alpha| \leqslant N} \sup_{x \in \mathbb{R}^n} (1 + |x|^2)^{\frac{N}{2}} |D^\alpha f(x)|$$

defines a norm on $\mathcal{S}(\mathbb{R}^n)$ *and*

$$d(f, g) := \sum_{N=0}^{\infty} 2^{-N} \frac{\|f - g\|_{(N)}}{1 + \|f - g\|_{(N)}}$$

defines a metric making $\mathcal{S}(\mathbb{R}^n)$ *complete.*

Proof. TODO □

Remark 2.4.6 The Schwartz space is a vector space, but it is impossible to define a norm on it such that norm convergence is equal to $\mathcal{S}$-convergence. This follows from the following lemma **TODO**

Lemma 2.4.7 (Compactness result)
Let $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ such that for all $N \in \mathbb{N}_0$ there exists a constant C_N such that $\|f_n\|_{(N)} \leqslant C_N$. Then there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and a $f \in \mathcal{S}(\mathbb{R}^n)$ such that $f_{n_k} \xrightarrow[k\to\infty]{\mathcal{S}} f$.

Proof. TODO □

Remark 2.4.8 But $\mathcal{S}(\mathbb{R}^n)$ can be equipped with a topology such that convergence with respect to this topology equals $\mathcal{S}$-convergence, turning $\mathcal{S}(\mathbb{R}^n)$ into a topological vector space (TVS).

topological vector space

Theorem 2.4.1: Schwartz convergence

Let $(f_j)_{j\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ with $f_j \xrightarrow{\mathcal{S}} f \in \mathcal{S}(\mathbb{R}^n)$.

1. $f_j \to f$ in $L^p(\mathbb{R}^n)$ holds for all $p \in (0,\infty)$.
2. $D^\alpha f_j \xrightarrow{\mathcal{S}} D^\alpha f$ holds for all $\alpha \in \mathbb{N}_0^n$.
3. $(x \mapsto x^\alpha f_j(x)) \xrightarrow{\mathcal{S}} (x \mapsto x^\alpha f(x))$ holds for all $\alpha \in \mathbb{N}_0^n$.
4. Then $T_h f \xrightarrow[h\to 0]{\mathcal{S}} f$ holds.
5. For $g \in \mathcal{S}(\mathbb{R}^n)$ also $fg, f * g \in \mathcal{S}(\mathbb{R}^n)$.
6. $D^\alpha(f * g) = (D^\alpha f) * g = f * (D^\alpha g)$ holds for all $\alpha \in \mathbb{N}_0^n$.

Proof. 1. For $k \in \mathbb{N}_0$ and $x \in \mathbb{R}^n$ we have

$$|f_j(x) - f(x)| \leqslant (1+|x|^2)^{-\frac{k}{2}} \|f_j - f\|_{(k,0)}.$$

Choosing k large enough, we obtain for $p \in [1,\infty)$

$$\begin{aligned}\|f_j - f\|_p^p &= \int_{\mathbb{R}^n} |f_j(x) - f(x)|^p \,\mathrm{d}x \\ &\leqslant \|f_j - f\|_{(k,0)}^p \cdot \underbrace{\int_{\mathbb{R}^n} (1+|x|^2)^{-p\cdot\frac{k}{2}} \,\mathrm{d}x}_{<\infty \quad \text{for } k\cdot p>n} \xrightarrow{j\to\infty} 0.\end{aligned}$$

For $p = \infty$ notice that $\|f_j - f\|_{L^\infty(\mathbb{R}^n)} = \|f_j - f\|_{(0,0)}$, so the result follows immediately.

For $p < 1$ the standard L^p-norm doesn't fulfil the triangle inequality, hence we have to use the metric $d(f,g) := \int_{\mathbb{R}^n} |f(x) - g(x)|^p \,\mathrm{d}x$ but can use the same proof as above.

2. For $\ell \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$

$$\sum_{|\beta|\leqslant\ell} |D^\beta D^\alpha f| = \sum_{|\gamma|\leqslant \ell+|\alpha|} |D^\gamma f|$$

holds, which implies $\|D^\alpha f_j - D^\alpha f\|_{(k,\ell)} = \|f_j - f\|_{(k,\ell+|\alpha|)}$ and therefore the statement.

3. Define $g_j(x) := x^\alpha f_j(x)$ and g accordingly. For $k,\ell \in \mathbb{N}_0$ we have

$$\|g_j - g\|_{(k,\ell)} = \sup_{x\in\mathbb{R}^n} (1+|x|^2)^{\frac{k}{2}} \sum_{|\beta|\leqslant\ell} \Big| \underbrace{D^\beta x^\alpha \big(f_j(x) - f(x)\big)}_{=\sum_{0\leqslant i\leqslant\beta} \binom{\beta}{i} x^{\alpha-\beta+i} D^i \big(f_j(x)-f(x)\big)} \Big|$$

We have

$$D^\beta x^\alpha \big(f_j(x) - f(x)\big) = \sum_{0\leqslant i\leqslant\beta} \binom{\beta}{i} x^{\alpha-\beta+i} D^i \big(f_j(x) - f(x)\big)$$

TODO

4. Taylor??
5. Using that $\mathcal{S}(\mathbb{R}^n) \subsetneq L^p(\mathbb{R}^n)$ for all $p \in [1, \infty]$ **TODO**, we can use lemma 2.1.9 to show that the convolution is well-defined. $\square$

Theorem 2.4.2: Fourier transform on $\mathcal{S}(\mathbb{R}^n)$

Let $(f_j)_{j\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ Schwartz-converge to $f \in \mathcal{S}(\mathbb{R}^n)$.

1. It holds that $\hat{f}, \check{f} \in \mathcal{S}(\mathbb{R}^n)$.
2. For all $\alpha \in \mathbb{N}_0^n$ and $x \in \mathbb{R}^n$ we have
$$D^\alpha \hat{f}(\xi) = (-i)^{|\alpha|}\widehat{x^\alpha f}(\xi) \quad \text{and} \quad \xi^\alpha \hat{f}(\xi) = (-i)^{|\alpha|}\widehat{D^\alpha f}(\xi)$$
3. It holds that $\hat{f}_j \xrightarrow{\mathcal{S}} \hat{f}$ and $\check{f}_j \xrightarrow{\mathcal{S}} \check{f}$.

Proof. TODO $\square$

Theorem 2.4.3: TODO

1. Let $f \in \mathcal{S}(\mathbb{R}^n)$, $g \in \mathcal{S}(\mathbb{R}^m)$. Then $f \otimes g \in \mathcal{S}(\mathbb{R}^{n+m})$ and $\widehat{f \otimes g}(\xi, \eta) = \hat{f}(\xi)\hat{g}(\eta)$ holds for all $(\xi, \eta) \in \mathbb{R}^{n+m}$.
2. Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then $\widehat{fg} = (2\pi)^{-\frac{n}{2}} \hat{f} * \hat{g}$ holds.
3. We have $\mathcal{F}\left(\exp\left(-\frac{|x|^2}{2}\right)\right)(\xi) = \exp\left(-\frac{|\xi|^2}{2}\right)$ for all $\xi \in \mathbb{R}^n$.

$(f \otimes g)(x, y) := f(x)g(y)$.

Proof. TODO $\square$

Theorem 2.4.4: Fourier inversion formula

We have $\check{\hat{f}} = f = \hat{\check{f}}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$. Also, $\mathcal{F}$ and $\mathcal{F}^{-1}$ are bijective on $\mathcal{S}(\mathbb{R}^n)$.

Proof. For $\varepsilon > 0$ define
$$I_\varepsilon(x) := (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\langle x, \xi\rangle} e^{-\varepsilon^2 \frac{|\xi|^2}{2}} \, \mathrm{d}\xi.$$

With the theorem of Lebesgue **(Majorant???)** we see that I is continuous in ε, i.e $I_\varepsilon(x) \xrightarrow{\varepsilon \to 0} I_0(x) = \mathcal{F}^{-1}(\mathcal{F}(f))(x)$.

Define
$$g : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \exp\left(-\frac{|x|^2}{2}\right), \quad h(x) := g(\varepsilon x).$$

lemma 2.1.7 (3) and theorem 2.4.3 yield
$$\hat{h}(\xi) = \varepsilon^{-n}\hat{g}(\varepsilon^{-1}\xi) = \varepsilon^{-n} g(\varepsilon^{-1}\xi) = \varepsilon^{-n} \exp\left(-\frac{|\xi|^2}{2\varepsilon}\right)$$

By lemma 2.1.7 (1) we have
$$\int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\langle x, \xi\rangle} h(\xi) \, \mathrm{d}\xi = \int_{\mathbb{R}^n} M_x \hat{f}(\xi) h(\xi) \, \mathrm{d}\xi = \int_{\mathbb{R}^n} \mathcal{F}(T_{-x} f)(\xi) h(\xi) \, \mathrm{d}\xi$$

Thus by Fubini and the substitution $z := y\varepsilon^{-1}$ we have

$$\begin{aligned}
I(\varepsilon) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\langle x,\xi\rangle} h(\xi)\,\mathrm{d}\xi = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} T_{-x} f(\xi)\hat{h}(\xi)\,\mathrm{d}\xi \\
&= (2\pi\varepsilon^2)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x+y) \exp\left(-\frac{\left|\frac{y}{\varepsilon}\right|^2}{2}\right) \mathrm{d}y \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x+\varepsilon z) \exp\left(-\frac{|z|^2}{2}\right) \mathrm{d}z \\
&\xrightarrow{\varepsilon\to 0} f(x)(2\pi)^{-\frac{n}{2}} \underbrace{\int_{\mathbb{R}^n} \exp\left(-\frac{|z|^2}{2}\right) \mathrm{d}z}_{=(2\pi)^{\frac{n}{2}}} = f(x). \qquad \square
\end{aligned}$$

2.5 Tempered distributions

We introduce the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$, which is the continuous dual of the Schwartz space. We have $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$ (cf. subsubsection section 2.5.1). All functions in $\mathcal{S}'(\mathbb{R}^n)$ have a Fourier transform, whereas not all functions in $\mathcal{D}'(\mathbb{R}^n)$ have.

The derivative of a tempered distribution is again a tempered distribution. Tempered distributions T generalize the bounded (or slow-growing) locally integrable functions; meaning that each derivative of T grows at most as fast as some polynomial. This characterisation is dual to the rapidly falling behaviour of the derivatives of a function in the Schwartz space. An example of a rapidly falling function is $|x|^n \cdot \exp(-\lambda|x|^\beta)$ for $n, \lambda, \beta > 0$.

All distributions with compact support and all square-integrable functions are tempered distributions. More generally, all products of polynomials with elements of $L^p(\mathbb{R}^n)$ for $p \in [1, \infty]$ are tempered distributions. [16]

Definition 2.5.1 (Tempered distributions)
The tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ is the set of all continuous complex linear functionals on $\mathcal{S}(\mathbb{R}^n)$, i.e. all $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ such that

- $T(\lambda f + \mu g) = \lambda T(f) + \mu T(g)\ \forall \lambda, \mu \in \mathbb{C},\ f, g \in \mathcal{S}(\mathbb{R}^n)$,
- $f_k \xrightarrow{\mathcal{S}} f$ implies $Tf_k \to Tf$.

Remark 2.5.2 (Vector space structure of $\mathcal{S}'(\mathbb{R}^n)$)
By defining $(\lambda T + \mu S)(f) := \Lambda Tf + \mu Sf$ for all $\lambda, \mu \in \mathbb{C}$ and $f \in \mathcal{S}(\mathbb{R}^n)$, $\mathcal{S}'(\mathbb{R}^n)$ becomes a linear space. We equip $\mathcal{S}'(\mathbb{R}^n)$ with a weak convergence: $T_k \xrightarrow{\mathcal{S}} T$ if and only if $T_k f \to T$ for all $f \in \mathcal{S}(\mathbb{R}^n)$.

Theorem 2.5.1: Continuity and Boundedness

A linear functional $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is continuous (i.e. $T \in \mathcal{S}'(\mathbb{R}^n)$) if and only if there exists $c > 0$ and $k, \ell \in \mathbb{N}_0$ with

$$|Tf| \leqslant c\|f\|_{(k,\ell)}\ \forall f \in \mathcal{S}(\mathbb{R}^n).$$

Proof. " $\implies$ ": Towards contradiction assume that for all $c > 0$, $k, \ell \in \mathbb{N}_0$ there exist $f_{c,k,\ell} \in \mathcal{S}(\mathbb{R}^n)$ with

$$1 = |Tf_{c,k,\ell}| > c\|f_{c,k,\ell}\|. \tag{14}$$

Let $f_k := f_{k,k,k}$. Now (14) implies $\|f_k\|_{(k,k)} < \frac{1}{k}$ for all $k \in \mathbb{N}_0$. Thus **(TODO: REALLY? WIR BRAUCHEN DOCH ALLE (ℓ, k)???)** $f_k \xrightarrow{\mathcal{S}} 0$ and by continuity of T, $Tf_k \to 0$ holds, which contradicts $|Tf_k| = 1$.

" $\impliedby$ ": For $f_j \xrightarrow{\mathcal{S}} f$ the exist $c > 0$ and $(k, \ell) \in \mathbb{N}$ such that

$$0 \leqslant |T(f_j - f)| \leqslant c\|f_j - f\|_{(k,\ell)} \xrightarrow{j \to \infty} 0$$

holds, implying $Tf_j \to Tf$. □

Remark 2.5.3 (Regular tempered distributions T_f)
For $f \in L^1(\mathbb{R}^n)$ define

$$T_f : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C},\ g \mapsto \int_{\mathbb{R}^n} f(x)g(x)\,\mathrm{d}x,$$

which is linear and well-defined: for $g \in \mathcal{S}(\mathbb{R}^n)$

$$|T_f(g)| \overset{\triangle\neq}{\leqslant} \int_{\mathbb{R}^n} |f(x)||g(x)|\,\mathrm{d}x \leqslant \|g\|_\infty \|f\|_1 = \|g\|_{(0,0)} \|f\|_1$$

holds. Theorem 2.5.1 implies $T_f \in \mathcal{S}'(\mathbb{R}^n)$. This can by extended to $f \in L^p(\mathbb{R}^n)$ for $p \in [1,\infty]$ but not $p < 1$ by noting that $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for $p \in [1,\infty)$ and using Hölder's inequality.

Definition 2.5.4 (Regular / singular distribution)
A tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ is called regular distribution if there exists $f \in L^1(\mathbb{R}^n)$ such that $T = T_f$ holds, and singular distribution else.

regular distribution

WHY NOT ALSO L^p??

Lemma 2.5.5 (TODO)
Let $f, g \in L^1(\mathbb{R}^n)$ such that $T_f(\varphi) = T_g(\varphi)$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then $f = g$ holds almost everywhere.

Proof. 1 It suffices to show that $T_f(g) = 0$ for a $f \in L^1(\mathbb{R}^n)$ and all $g \in \mathcal{S}(\mathbb{R}^n)$ implies $f = 0$ almost everywhere.

2 The following proof uses a standard technique called mollification. Let $\omega \in \mathcal{S}(\mathbb{R}^n)$ be a non-negative symmetric function supported in the unit ball integrating to unity. For $\varepsilon > 0$ define $\omega_\varepsilon(x) := \varepsilon^{-n}\omega(\varepsilon^{-1}x)$, which still integrates to unity.

mollification

TODO PLOT

Figure 6: An example of a mollifier ω, given by $\exp\left(\frac{1}{x^2-1}\right) \cdot \mathbb{1}_{|x|\leqslant 1}(x)$.

The convolution $f * \omega_\varepsilon$ is continuous by the dominated convergence theorem and $|f * \omega_\varepsilon(x)| \leqslant \|f\|_1$ for all $x \in \mathbb{R}^n$.

3 By theorem 2.4.1 4 we have $\omega_\varepsilon * g \in \mathcal{S}(\mathbb{R}^n)$ for $g \in \mathcal{S}(\mathbb{R}^n)$. Thus

$$\begin{aligned}
0 = T_f(\omega_\varepsilon * g) &= \int_{\mathbb{R}^n} f(x)(\omega_\varepsilon * g)(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^n} \omega_\varepsilon(x-y)g(y)\,\mathrm{d}y\,\mathrm{d}x \\
&\overset{(\star)}{=} \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^n} \omega_\varepsilon(y-x)g(y)\,\mathrm{d}y\,\mathrm{d}x \\
&\overset{(\mathrm{F})}{=} \int_{\mathbb{R}^n} g(y) \int_{\mathbb{R}^n} f(x)\omega_\varepsilon(y-x)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{\mathbb{R}^n} g(y)(f * \omega_\varepsilon)(y)\,\mathrm{d}y = T_g(f * \omega_\varepsilon)
\end{aligned}$$

where in $(\star)$ we use the symmetry of ω (and thus of ω_ε), (F) denotes Fubini's theorem and that, by a simple substitution,

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)\,\mathrm{d}y = \int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y$$

holds.

4 As $f * \omega_\varepsilon = \omega_\varepsilon * f$ is continuous (as above) for each $x \in \mathbb{R}^n$ with $(\omega_\varepsilon * f)(x) > 0$ there exists a neighbourhood U_x of x such that $(\omega_\varepsilon * f)(y) > 0$ for all $y \in U_x$.

For each such x, consider a non-negative non-zero function $g_x \in \mathcal{S}(\mathbb{R}^n)$ supported in U_x. From

$$0 = T_f(\omega_\varepsilon * g_x) = T_{g_x}(f * \omega_\varepsilon)$$

it follows that $(\omega_\varepsilon * f)(x) = 0$ for all $x \in \mathbb{R}^n$.

5 As the continuous, compactly supported functions are dense in L^1 we can write, for every $t > 0$, $f \in L^1(\mathbb{R}^n)$ can be written as $f = f_1 + f_2$, where f_1 is continuous and compactly supported and $\|h_2\|_1 \leqslant t$. Since convolution is distributive,

$$\begin{aligned} \|f\|_1 &\overset{4}{=} \|f - \omega_\varepsilon * f\|_1 = \|f_1 + f_2 - \omega_\varepsilon * f_1 - \omega_\varepsilon * f_2\|_1 \\ &\overset{\triangle\neq}{\leqslant} \|f_1 - \omega_\varepsilon * f_1\|_1 + \underbrace{\|f_2 - \omega_\varepsilon * f_2\|_1}_{\overset{\triangle\neq}{\leqslant} 2t}. \end{aligned}$$

The first term can be estimated as follows

$$\begin{aligned} \|f_1 - \omega_\varepsilon * f_1\|_1 &= \int_{\mathbb{R}^n} |f_1(x) - (\omega_\varepsilon * f_1)(x)| \,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} \left| f_1(x) - \int_{\mathbb{R}^n} f_1(x-y)\omega_\varepsilon(y)\,\mathrm{d}y \right| \mathrm{d}x \\ &\overset{(\ddagger)}{=} \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \big[f_1(x) - f_1(x-y)\big]\omega_\varepsilon(y)\,\mathrm{d}y \right| \mathrm{d}x \\ &\leqslant \int_{\mathbb{R}^n} \omega_\varepsilon(y) \int_{\mathbb{R}^n} |f_1(x) - f_1(x-y)| \,\mathrm{d}x\,\mathrm{d}y \\ &\overset{(\ddagger)}{\leqslant} \sup_{|y|\leqslant\varepsilon} \|h_1 - h_1(\cdot - y)\|_1, \end{aligned}$$

where in ($\ddagger$) we use that $\int_{\mathbb{R}^n} \omega_\varepsilon(x)\,\mathrm{d}x = 1$.

Due to the bounded support of f_1 and its uniform continuity the above expression converge to zero for $\varepsilon \to 0$. Thus by choosing ε, t small enough, $\|f_1\|$ is arbitrarily small and therefore equal to zero, implying $f = 0$ almost everywhere. □

continuous functions with compact support are uniformly continuous.

Example 2.5.6 (Borel measures are tempered distributions)
A finite Borel measure μ on $\mathbb{R}^n$ is a tempered distribution via

$$\mu(f) := \int_{\mathbb{R}^n} f(x)\,\mathrm{d}\mu(x).$$

TODO ◇

Remark 2.5.7 (Extending convolution to measures)
First let $f, g, h \in \mathcal{S}(\mathbb{R}^n)$ and μ, λ be measures on $\mathbb{R}^n$ with density f and g with respect to the Lebesgue measure, respectively. Then

$$\begin{aligned} (f * g)(h) &:= \int_{\mathbb{R}^n} (f * g)(x)h(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} h(x) \int_{\mathbb{R}^n} f(x-y)h(y)\,\mathrm{d}y\,\mathrm{d}x \\ &\overset{(\mathrm{F})}{=} \int_{\mathbb{R}^n \times \mathbb{R}^n} h(z+y)f(z)h(y)\,\mathrm{d}z\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} h(z+y)\,\mathrm{d}\mu(z)\,\mathrm{d}\lambda(y). \end{aligned}$$

The last expression makes sense for all Borel measures μ and λ and $f \in \mathcal{C}_0(\mathbb{R}^n)$.

Using the Riesz representation theorem we can define $\mu * \lambda$ for two finite Borel measures μ, λ on $\mathbb{R}^n$ to be the unique measure on $\mathbb{R}^n$ satisfying

$$\int_{\mathbb{R}^n} f(x)\,\mathrm{d}(\mu * \lambda)(x) = \int_{\mathbb{R}^n \times \mathbb{R}^n} f(x+y)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y) \quad \forall f \in \mathcal{C}_0(\mathbb{R}^n). \tag{15}$$

Using standard arguments **TODO** this also holds for all bounded Borel measurable functions, in particular for $f_\xi(x) := e^{i\langle x,\xi\rangle}$. Thus

$$\mathcal{F}(\mu * \lambda)(\xi) = (2\pi)^{n/2}\hat{\mu}(\xi)\hat{\lambda}(\xi)$$

holds for all $\xi \in \mathbb{R}^n$, where $\hat{\mu}(\xi) := \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle}\,\mathrm{d}\mu(x)$

Recalling the total variation norm, we have $\|\mu * \lambda\| \leqslant \|\mu\| \cdot \|\lambda\|$.

For every Borel measurable set $E \subset \mathbb{R}^n$ (15) implies that we have

$$(\mu * \lambda)(E) = (\mu \otimes \lambda)\big(\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x + y \in E\}\big).$$

Definition 2.5.8 (Derivative, F-transform in $\mathcal{S}'(\mathbb{R}^n)$)
For $T \in \mathcal{S}'(\mathbb{R}^n)$, $f, g \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ we define

$$\begin{aligned} (D^\alpha T)(f) &:= (-1)^{|\alpha|}T(D^\alpha f), & (\mathcal{F}\,T)(f) &:= T(\mathcal{F}\,f), \\ (\mathcal{F}^{-1}\,T)(f) &:= T(\mathcal{F}^{-1}\,f), & (fT)(g) &:= T(fg). \end{aligned}$$

Theorem 2.5.1 implies that $D^\alpha T$, fT, $\hat{T}$, $\check{T} \in \mathcal{S}'(\mathbb{R}^n)$.

Example 2.5.9 (Fourier transform of tempered distributions)
For $f \in \mathcal{S}(\mathbb{R}^n)$ define

1. $T_1(f) := \int_{\mathbb{R}^n} x^\alpha f(x)\,\mathrm{d}x$, $\alpha \in \mathbb{N}_0^n$.
2. $T_2(f) := \int_{\mathbb{R}^n} \sin(\langle a, x\rangle_{\mathbb{R}^n}) f(x)\,\mathrm{d}x$, $a \in \mathbb{R}^n$.
3. $T_3(f) := \int_{\mathbb{R}} \mathbb{1}_{[-1,1]}(x) f(x)\,\mathrm{d}x$.

Their Fourier transforms are as follows.

1. By definition 2.5.8 and theorem 2.4.2 ($\star$)

$$\begin{aligned} \mathcal{F}(T_1(f)) = T_1(\mathcal{F}(f)) = \int_{\mathbb{R}^n} x^\alpha \hat{f}(x)\,\mathrm{d}x &\overset{(\star)}{=} i^{|\alpha|}\int_{\mathbb{R}^n} \mathcal{F}(D^\alpha f)(x)\,\mathrm{d}x \\ &\overset{(12)}{=} i^{|\alpha|}(2\pi)^{n/2}(\cancel{\mathcal{F}^{-1}\mathcal{F}}D^\alpha f)(0) = i^{|\alpha|}(2\pi)^{n/2}D^\alpha f(0) \end{aligned}$$

holds.

2. By lemma 2.1.7

$$\int_{\mathbb{R}^n} e^{i\langle a,x\rangle}\mathcal{F}\varphi(x)\,\mathrm{d}x \overset{(12)}{=} (2\pi)^{n/2}\varphi(a).$$

holds. Now, $2i\sin(t) = e^{it} - e^{-it}$ implies

$$\begin{aligned} \mathcal{F}(T_2(f)) = T_2(\hat{f}) &= \int_{\mathbb{R}^n} \sin(\langle a, x\rangle)\hat{f}(x)\,\mathrm{d}x \\ &= (2\pi)^{n/2}\frac{\varphi(a) - \varphi(-a)}{2i} \end{aligned}$$

③ FUBINI's theorem allows use to slide over the FOURIER transform inside the integral:

$$\begin{aligned}\mathcal{F}(T_3(f)) &= \int_{\mathbb{R}} \mathbb{1}_{[-1,1]}(x)\hat{f}(x)\,\mathrm{d}x = \int_{\mathbb{R}} \widehat{\mathbb{1}_{[-1,1]}}(x)f(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}} f(x)\cdot(2\pi)^{-1/2}\int_{\mathbb{R}} e^{ixt}\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_{\mathbb{R}} \sqrt{\frac{2}{\pi}}\frac{\sin(x)}{x}f(x)\,\mathrm{d}x.\end{aligned}$$

◇

DEFINITION 2.5.10 (PRINCIPLE VALUE INTEGRAL)
For $\delta > 0$ let $I_\delta := \mathbb{R}\setminus[-\delta,\delta]$. If $f \in L^1(I_\delta)$ for every $\delta > 0$, its principle value integral is

$$\text{p. V.}\int_{\mathbb{R}} f(x)\,\mathrm{d}x := \lim_{\delta\searrow 0}\int_{I_\delta} f(x)\,\mathrm{d}x.$$

Example 2.5.11 Define $T\colon \mathcal{S}(\mathbb{R}) \to \mathbb{C}$, $f \mapsto \int_{\mathbb{R}} f(x)\ln(|x|)\,\mathrm{d}x$. Then $T \in \mathcal{S}'(\mathbb{R})$ and

$$\begin{aligned}T'(f) &= \text{p. V.}(x^{-1})(f) := \text{p. V.}\int_{\mathbb{R}} \frac{f(x)}{x}\,\mathrm{d}x \text{ and} \\ T''(f) &= -\text{p. V.}(x^{-2})(f) := \text{p. V.}\int_{\mathbb{R}} \frac{f(x)-f(0)}{x^2}\,\mathrm{d}x\end{aligned}$$

hold. ◇

Proof. T is linear, so by theorem 2.5.1 we only need to show it is bounded. For $f \in \mathcal{S}(\mathbb{R})$ there exist $C > 0$ and $k \in \mathbb{N}$ such that $|f(x)| \leqslant C(1+|x|^2)^{-k}$ holds for all $x \in \mathbb{R}$. Thus by symmetry

$$\begin{aligned}|Tf| &\overset{\triangle\neq}{\leqslant} \int_{\mathbb{R}} |f(x)||\ln(|x|)|\,\mathrm{d}x \leqslant C\int_{\mathbb{R}} \frac{|\ln(|x|)|}{(1+|x|^2)^k}\,\mathrm{d}x \\ &= 2C\int_0^\infty \frac{|\ln(x)|}{(1+x^2)^k}\,\mathrm{d}x = \int_0^1 \frac{-2C\ln(x)}{(1+x^2)^k}\,\mathrm{d}x + \int_1^\infty \frac{2C\ln(x)}{(1+x^2)^k}\,\mathrm{d}x\end{aligned}$$

The second integral can be estimated brutally and integrated by parts (assuming $k \geqslant 1$):

$$\begin{aligned}\int_1^\infty \frac{\ln(x)}{(1+x^2)^k}\,\mathrm{d}x &\leqslant \int_1^\infty \frac{\ln(x)}{1+x^2}\,\mathrm{d}x \leqslant \int_1^\infty \frac{\ln(x)}{x^2}\,\mathrm{d}x \\ &= \left[-\frac{\ln(x)}{x}\right]_{x=1}^\infty + \int_1^\infty \frac{1}{x^2}\,\mathrm{d}x = 0+1 = 1.\end{aligned}$$

The other integral is handled similarly (we assume $k \geqslant 2$):

$$\int_0^1 \frac{\ln(x)}{(1+x^2)^k}\,\mathrm{d}x \leqslant \int_0^1 \frac{\ln(x)}{(1+x^2)^2}\,\mathrm{d}x$$

We first compute a related integral using the substitution $x = \tan(u)$, which is invertible on $(0,\pi/4) \supset (0,1)$:

$$\begin{aligned}\int \frac{1}{(1+x^2)^2}\,\mathrm{d}x &= \int \cos^4(u)\cdot\frac{\mathrm{d}u}{\cos^2(u)} = \int \cos^2(u)\,\mathrm{d}u \\ &= \frac{u+\sin(u)\cos(u)}{2} = \frac{\tan^{-1}(x)+\frac{x}{1+x^2}}{2}\end{aligned}$$

Integrating by parts, where $\mathfrak{C} := \int_0^1 \frac{\tan^{-1}(x)}{x}\,\mathrm{d}x$ is Catalan's constant,

$$\begin{aligned}\int_0^1 \frac{\ln(x)}{(1+x^2)^2}\,\mathrm{d}x &= \left[\ln(x)\cdot\frac{\tan^{-1}(x)+\frac{x}{1+x^2}}{2}\right]_{x=0}^{1} \\ &\quad - \frac{1}{2}\int_0^1 \frac{1}{x}\left(\tan^{-1}(x)+\frac{x}{1+x^2}\right)\mathrm{d}x \\ &= (0-0) - \frac{\mathfrak{C}}{2} - \frac{\tan^{-1}(1)-\tan^{-1}(0)}{2} = -\frac{1}{2}\left(\mathfrak{C}+\frac{\pi}{4}\right),\end{aligned}$$

holds, implying in summary $|Tf| \leqslant C\left(2+\mathfrak{C}+\frac{\pi}{4}\right)$.

By definition 2.5.8

$$\begin{aligned}T'(f) &= -T(f') = -\int_{\mathbb{R}} f'(x)\ln(|x|)\,\mathrm{d}x \\ &= \lim_{\delta\searrow 0} -\int_{-\infty}^{-\delta} f'(x)\ln(|x|)\,\mathrm{d}x - \int_{\delta}^{\infty} f'(x)\ln(|x|)\,\mathrm{d}x\end{aligned}$$

From the substitution $y=-x$ it follows that $\int_a^b f(x)\,\mathrm{d}x = \int_{-b}^{-a} f(-x)\,\mathrm{d}x$. Thus,

$$\begin{aligned}\left(\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}\right) f'(x)\ln(|x|)\,\mathrm{d}x &= \int_{\delta}^{\infty}\big[f'(x)+f'(-x)\big]\ln(|x|)\,\mathrm{d}x \\ &= \Big[\ln(|x|)\big[f(x)+f(-x)\big]\Big]_{x=\delta}^{\infty} \\ &\quad - \int_{\delta}^{\infty}\frac{f(x)-f(-x)}{x}\,\mathrm{d}x\end{aligned}$$

holds. As $f\in\mathcal{S}(\mathbb{R})$,

$$\lim_{x\to\infty}\ln(|x|)f(\pm x) = 0 \quad\text{and}\quad \lim_{x\to 0}\ln(|x|)\big[f(x)-f(-x)\big] \overset{\text{L'H}}{=} 0$$

With L'Hôpitals rule we get a factor of $x\log^2(|x|)$ instead of $\log(|x|)$, which approaches 0 ($2\times$ L'H), as does $f'(x)-f'(-x)$.

hold. Using the same substitution backwards we obtain

$$\begin{aligned}\lim_{\delta\searrow 0}\int_{\delta}^{\infty}\frac{f(x)-f(-x)}{x}\,\mathrm{d}x &= \lim_{\delta\searrow 0}\left(\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}\right) - \frac{f(x)}{x}\,\mathrm{d}x \\ &= -\,\mathrm{p.\,V.}\int_{\mathbb{R}}\frac{f(x)}{x}\,\mathrm{d}x\end{aligned}$$

Thus,

$$T'(f) = 0 - -\,\mathrm{p.\,V.}\int_{\mathbb{R}}\frac{f(x)}{x}\,\mathrm{d}x = \mathrm{p.\,V.}\int_{\mathbb{R}}\frac{f(x)}{x}.$$

Repeating the above with f' instead of f yields

$$T''(f) = T(f'') = \int_{\mathbb{R}} f''(x)\ln(|x|)\,\mathrm{d}x = \mathrm{p.\,V.}\int_{\mathbb{R}}\frac{f'(x)}{x}\,\mathrm{d}x.$$

As before by partial integration

$$\begin{aligned}\left(\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}\right)\frac{f'(x)}{x}\,\mathrm{d}x &= \int_{\delta}^{\infty}\frac{f'(x)-f'(-x)}{x}\,\mathrm{d}x \\ &= \left[\frac{f(x)+f(-x)}{x}\right]_{x=\delta}^{\infty} + \int_{\delta}^{\infty}\frac{f(x)+f(-x)}{x^2}\,\mathrm{d}x\end{aligned}$$

holds. As above, $\frac{f(x)+f(-x)}{x} \xrightarrow{x\to\infty} 0$ and

$$\text{p. V.} \int_{\mathbb{R}} \frac{f'(x)}{x} \, \mathrm{d}x = \lim_{\delta \searrow 0} \int_{\delta}^{\infty} \frac{f(x)+f(-x)}{x^2} \, \mathrm{d}x$$

hold. Furthermore,

$$\begin{aligned}
\int_{\delta}^{\infty} \frac{f(x)+f(-x)}{x^2} \, \mathrm{d}x &= \int_{\delta}^{\infty} \frac{f(x)+f(-x)}{x^2} \, \mathrm{d}x + \underbrace{\frac{2f(0)}{\delta} - \int_{\delta}^{\infty} \frac{2f(0)}{x^2} \, \mathrm{d}x}_{=0} \\
&= \int_{\delta}^{\infty} \frac{f(x)-f(0)+f(-x)-f(0)}{x^2} \, \mathrm{d}x + \frac{2f(0)}{\delta} \\
&= \left(\int_{-\infty}^{-\delta} + \int_{\delta}^{\infty} \right) \frac{f(x)-f(0)}{x^2} \, \mathrm{d}x + \frac{2f(0)}{\delta}.
\end{aligned}$$

Finally by L'Hôpitals rule,

$$\lim_{\delta \to 0} \frac{f(\delta)+f(-\delta)-2f(\delta)}{\delta} = \lim_{\delta \to 0} f'(\delta) - f'(-\delta) = 0,$$

yielding

$$T''(f) = \lim_{\delta \searrow 0} \left(\int_{-\infty}^{-\delta} + \int_{\delta}^{\infty} \right) \frac{f(x)-f(0)}{x^2} \, \mathrm{d}x = \text{p. V.} \int_{\mathbb{R}} \frac{f(x)-f(0)}{x^2} \, \mathrm{d}x.$$

□

Theorem 2.5.2: TODO

For $T \in \mathcal{S}'(\mathbb{R}^n)$

1. $\mathcal{F}^{-1} \mathcal{F} T = \mathcal{F} \mathcal{F}^{-1} T = T$ holds.
2. $\mathcal{F}^{(-1)}$ map $\mathcal{S}'(\mathbb{R}^n)$ bijectively and continuously onto itself.
3. $\mathcal{F}(D^\alpha T) = i^{|\alpha|} x^\alpha \mathcal{F} T$ and $\mathcal{F}(x^\alpha T) = i^{|\alpha|} D^\alpha \mathcal{F} T$ holds.
4. For $\varepsilon > 0$ let $T_\varepsilon(f) := T(\varepsilon^{-n} f(\varepsilon^{-1} \cdot))$ for $f \in \mathcal{S}(\mathbb{R}^n)$ be the dilation of T. Then $\mathcal{F} T_\varepsilon = \varepsilon^{-n} \mathcal{F}(T)(\varepsilon^{-1} \cdot)$ holds.
5. For $h \in \mathbb{R}^n$ and $f \in \mathcal{S}(\mathbb{R}^n)$, the translation of T is $(\tau_h T)(f) := T(f(\cdot + h))$ and $\mathcal{F}(\tau_h T) = e^{-i\langle h, \cdot \rangle} \mathcal{F} T$ holds.
6. The modulation of T is $(M_h T)(f) := T(e^{i\langle h, \cdot \rangle} f)$ and $\mathcal{F}(M_h T) = \tau_h(\mathcal{F} T)$ holds.

Proof. 1. By definition 2.5.8

$$\begin{aligned}
\mathcal{F}^{-1}(\mathcal{F} T)(f) &= (\mathcal{F} T)(\mathcal{F}^{-1} f) = T(\mathcal{F}(\mathcal{F}^{-1} f)) = Tf \\
\mathcal{F}(\mathcal{F}^{-1} T)(f) &= (\mathcal{F}^{-1} T)(\mathcal{F} f) = T(\mathcal{F}^{-1}(\mathcal{F} f)) = Tf
\end{aligned}$$

holds for all $f \in \mathcal{S}(\mathbb{R}^n)$.

2. **TODO**

3. By definition 2.5.8 and theorem 2.4.2

$$\begin{aligned}
\mathcal{F}(D^\alpha T)(f) = (D^\alpha T)(\mathcal{F} f) &= (-1)^{|\alpha|} T(\mathcal{D}^\alpha \mathcal{F} f) \\
&= (-1)^{|\alpha|} T((-i)^{|\alpha|} \mathcal{F}(x^\alpha f)) \\
&= i^{|\alpha|} T(\mathcal{F}(x^\alpha f)) = i^{|\alpha|} \mathcal{F} T(x^\alpha f) \\
&= i^{|\alpha|} \mathcal{F}(x^\alpha T(f)) = i^{|\alpha|} x^\alpha T(\mathcal{F} f) \\
&= i^{|\alpha|} x^\alpha \mathcal{F}(Tf)
\end{aligned}$$

and

$$\begin{aligned}\mathcal{F}(x^\alpha T) &= x^\alpha T(\mathcal{F} f) = T(x^\alpha \mathcal{F} f) = T((-i)^{|\alpha|} \mathcal{F}(\mathcal{D}^\alpha f)) \\ &= (-i)^{|\alpha|} T(\mathcal{F}(\mathcal{D}^\alpha f) = (-i)^{|\alpha|}(\mathcal{F} T)(\mathcal{D}^\alpha f) \\ &= (-i)^{|\alpha|}(-1)^{|\alpha|} \mathcal{D}^\alpha(\mathcal{F} T)(f) = i^{|\alpha|} \mathcal{D}^\alpha(\mathcal{F} T)(f)\end{aligned}$$

hold for all $f \in \mathcal{S}(\mathbb{R}^n)$.

4 Often calculations become clearer if one introduces an operator instead of writing for example $\mathcal{F}(f(\varepsilon x))$. Therefore, define the operator σ_ε by $(\sigma_\varepsilon f)(x) = f(\varepsilon x)$ when f is an ordinary function. If T is an ordinary function we have by $y = \varepsilon x$

$$\begin{aligned}(\sigma_\varepsilon T)(f) &= \int_{\mathbb{R}^n} \sigma_\varepsilon T(x) f(x) \, \mathrm{d}x = \int_{\mathbb{R}^n} T(\varepsilon x) f(x) \, \mathrm{d}x \\ &= \varepsilon^{-n} \int_{\mathbb{R}^n} T(y) f(\varepsilon^{-1} y) \, \mathrm{d}y = \varepsilon^{-n} \int_{\mathbb{R}^n} T(y) \sigma_{\varepsilon^{-1}} f(y) \, \mathrm{d}y \\ &= \varepsilon^{-n} T(\sigma_{\varepsilon^{-1}} f)\end{aligned}$$

holds. By lemma 2.1.7 $\sigma_\varepsilon \mathcal{F} f(\xi) = \varepsilon^{-n} \mathcal{F}(\sigma_{\varepsilon^{-1}} f)$ holds (with $D_\lambda = \sigma\varepsilon^{-1}$). Thus

$$\begin{aligned}\mathcal{F}(\sigma_\varepsilon T)(f) &= \sigma_\varepsilon T(\mathcal{F} f) = \varepsilon^{-n} T(\sigma_{\varepsilon^{-1}}(\mathcal{F} f)) = T(\mathcal{F}(\sigma_\varepsilon f)) \\ &= \mathcal{F} T(\sigma_\varepsilon f) = \varepsilon^{-n} \sigma_{\varepsilon^{-1}} \mathcal{F} T(f),\end{aligned}$$

holds, i.e. $\mathcal{F}(\sigma_\varepsilon T) = \varepsilon^{-n} \sigma_{\varepsilon^{-1}} \mathcal{F} T$.

5 For ordinary functions we define τ_h (cf. T_y in lemma 2.1.7) so that $\tau_h f$ is f translated to the right with a distance h. For distributions we do the same. But since distributions are not defined pointwise, we cannot just say $(\tau_h T)(x) = T(x-h)$. Instead we have to do the best we can: we transform the test function in such a way that we get the correct result when T is an ordinary function and $T(f)$ is defined as $\int_{\mathbb{R}^n} T(x) f(x) \, \mathrm{d}x$: with $y = x - h$ we obtain

$$\int_{\mathbb{R}^n} T(x-h) f(x) \, \mathrm{d}x = \int_{\mathbb{R}^n} T(y) f(y+h) \, \mathrm{d}y,$$

i.e. $(\tau_h T)(f) = T(\tau f)$. By definition 2.5.8 and lemma 2.1.7

$$\begin{aligned}\mathcal{F}(\tau_h T)(f(x)) &= (\tau_h T)(\mathcal{F} f(x)) = T(\tau_{-h}(\mathcal{F} f)) = T((\mathcal{F} f)(\xi + h)) \\ &= T(\mathcal{F}(e^{-i\langle h,x\rangle} f(x))) = \mathcal{F} T(e^{-i\langle h,x\rangle} f(x)) \\ &= e^{-i\langle h,x\rangle} \mathcal{F} T(f(x))\end{aligned}$$

holds for all $f \in \mathcal{S}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$.

6 Analogously to the above,

$$\begin{aligned}\mathcal{F}(M_h T)(f(x)) &= (M_h T)(\mathcal{F} f)(x) = T(M_{-h} \mathcal{F} f)(x) \\ &= T(e^{-i\langle h,x\rangle} \mathcal{F} f)(x) = T \mathcal{F} f(x+h) \\ &= \tau_h(\mathcal{F} T f)(x)\end{aligned}$$

holds for all $f \in \mathcal{S}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$. □

The spaces $\mathcal{D}(\Omega)$ and $\mathcal{D}'(\Omega)$

Remark 2.5.12 $S(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$ are not well-suited for partial differential equations on domains. This gave rise to the following definitions, which are historically older than those of the Schwartz spaces and tempered distributions.

Definition 2.5.13 (Test functions, distributions)
Let $\Omega \subset \mathbb{R}^n$ be open.

- The space of test functions is

$$\mathcal{D}(\Omega) := \{f \in \mathcal{C}^\infty(\Omega) : \operatorname{supp}(f) \subset \Omega \text{ is compact}\}.$$

- A sequence $(f_k)_{k\in\mathbb{N}} \subset \mathcal{D}(\Omega)$ converges to $f \in \mathcal{D}(\Omega)$ in $\mathcal{D}(\Omega)$ if there exists a compact set $K \subset \Omega$ with $\operatorname{supp}(f_k) \subset K$ for all $k \in \mathbb{N}$ and $D^\alpha f_k \to D^\alpha f$ holds for all $\alpha \in \mathbb{N}_0$. We then write $f_k \xrightarrow{\mathcal{D}} f$.
- The space of distributions $\mathcal{D}'(\Omega)$ is the set of complex-valued linear continuous functions on $\mathcal{D}(\Omega)$. $T : \mathcal{D}(\Omega) \to \mathbb{C}$ belongs to $\mathcal{D}'(\Omega)$ if
  - $T(\lambda f + \mu g) = \lambda T(f) + \mu T(g)\ \forall f, g \in \mathcal{D}(\Omega),\ \lambda, \mu \in \mathbb{C}$,
  - $f_j \xrightarrow{\mathcal{D}} f$ implies $Tf_j \to T_f$.

Remark 2.5.14 $\mathcal{D}'(\Omega)$ can be a vector space analogously to $\mathcal{S}'(\mathbb{R}^n)$. We say $T_k \to T$ in $\mathcal{D}'(\Omega)$ if $T_k f \to Tf$ holds for all $f \in \mathcal{D}(\Omega)$.

Remark 2.5.15 The spaces $\mathcal{D}(\Omega)$ are much more flexible due to the inclusion of the compact support in their definition. However, a lot of algebraic structure is lost, e.g. Fourier transforms or convolutions do not have an easy counterpart on $\mathcal{D}(\Omega)$.

2.6 SOBOLEV spaces

We would like to extend the definition of the $\mathcal{C}^k$ spaces to functions which are differentiable in a weaker sense.

DEFINITION 2.6.1 (WEAK DERIVATIVE IN L^1_{loc})
A function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, which is defined as

$$\left\{u : u \text{ is measurable, } \int_K |u| \,\mathrm{d}x < \infty \;\forall K \subset \mathbb{R}^n \text{ compact.}\right\},$$

has a weak derivative of order $\alpha \in \mathbb{N}_0^n$ if there exists a function $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ such that for all $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$

weak derivative

$$\int_{\mathbb{R}^n} f(x) D^\alpha f(x) \,\mathrm{d}x = (-1)^{|\alpha|} \int_{\mathbb{R}^n} w(x) f(x) \,\mathrm{d}x$$

holds and we write $D^\alpha f := w$.

Remark 2.6.2 By the fundamental lemma of variational calculus a weak derivative if unique up to sets of zero measure.

DEFINITION 2.6.3 (CLASSICAL SOBOLEV SPACE)
For $k \in \mathbb{N}_0$ and $p \in [1, \infty)$ the classical SOBOLEV space is

$$W_p^k(\mathbb{R}^n) := \{f : L^p(\mathbb{R}^n) : D^\alpha f \in L^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \;\forall |\alpha| \leqslant k\}.$$

Remark 2.6.4 We can interpret L^p functions as tempered distributions if they are regular, as there is a one-to-one corresponds between L^p functions $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{C}$ and their corresponding tempered distribution T_f defined by $\mathcal{S}(\mathbb{R}^n) \ni \varphi \mapsto \int_{\mathbb{R}^n} f(x)\varphi(x) \,\mathrm{d}x \in \mathbb{C}$.

THEOREM 2.6.1

The space $W_k^p(\mathbb{R}^n)$ is complete with respect to

$$\|f\|_{W_k^p(\mathbb{R}^n)} := \|f\|_{k,p} := \left(\sum_{|\alpha| \leqslant k} \|D^\alpha f\|_p^p\right)^{\frac{1}{p}}.$$

We have that

$$\mathcal{S}(\mathbb{R}^n) \subset W_k^p(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n),$$

where the first inclusion and $\mathcal{D}(\mathbb{R}^n) \subset W_k^p(\mathbb{R}^n)$ are dense.

Proof. TODO □

Remark 2.6.5 The space $W_2^k(\mathbb{R}^n)$ are HILBERT space when equipped with

$$\langle f, g \rangle_{W_2^k(\mathbb{R}^n)} := \sum_{|\alpha| \leqslant k} \int_{\mathbb{R}^n} D^\alpha f(x) \overline{D^\alpha g(x)} \,\mathrm{d}x.$$

Our next goal is to describe the SOBOLEV spaces in term of the FOURIER Transform.

Definition 2.6.6

Let $n \in \mathbb{N}$ and w a continuous positive function on $\mathbb{R}^n$. The weighted L^2-space is

$$L^2(\mathbb{R}^n, w) := \{f \in L^1_{\text{loc}}(\mathbb{R}^n) : wf \in L^2(\mathbb{R}^n)\}.$$

Remark 2.6.7 $L^2(\mathbb{R}^n, w)$ equipped with

$$\langle f, g \rangle_{L^2(\mathbb{R}^n, w)} := \langle wf, wg \rangle_{L^2(\mathbb{R}^n)}$$

becomes a Hilbert space. Also,

$$\Phi : L^2(\mathbb{R}^n, w) \to L^2(\mathbb{R}^n),\ f \mapsto wf$$

is unitary.

Example 2.6.8 (Weights) An important instance of weights are

$$w_s : \mathbb{R}^n \to [0, \infty),\ x \mapsto \langle x \rangle := (1 + |x|^2)^{\frac{s}{2}}, \qquad s \in \mathbb{R}\,. \qquad \diamond$$

Theorem 2.6.2: Inclusions of $L^2(\mathbb{R}^n, w_s)$

For all $s \in \mathbb{R}$ we have

$$\mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n, w_s) \subset \mathcal{S}'(\mathbb{R}^n).$$

Both $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are dense in $L^2(\mathbb{R}^n, w_s)$.

Proof. TODO ☐

Theorem 2.6.3: todo

$\mathcal{F}$ and $\mathcal{F}^{-1}$ generate unitary maps of $W_2^k(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n, w_k)$ and vice versa:

$$\mathcal{F}(W_2^k(\mathbb{R}^n)) = \mathcal{F}^{-1}(W_2^k(\mathbb{R}^n)) = L^2(\mathbb{R}^n, w_k). \tag{16}$$

Proof. TODO ☐

Remark 2.6.9 (Alternative Definition of $W_2^k(\mathbb{R}^n)$)
One can rewrite (16) as

$$W_2^k(\mathbb{R}^n) = \mathcal{F}(L^2(\mathbb{R}^n, w_k)) = \mathcal{F}^{-1}(L^2(\mathbb{R}^n, w_k)),$$

which gives rise to an alternative definition of $W_2^k(\mathbb{R}^n)$.

Definition 2.6.10 (Bessel potential spaces)

For $s \in \mathbb{R}$

$$H^s(\mathbb{R}^n) := \{f \in \mathcal{S}'(\mathbb{R}^n) : w_s\hat{f} \in L^2(\mathbb{R}^n)\}$$

is the Sobolev space of order s.

Remark 2.6.11 We can replace $\hat{}$ by $\check{}$ in the above definition.

Remark 2.6.12 We have $H^k = W_2^k(\mathbb{R}^n)$ by theorem 2.6.3, i.e. $(H^s)_{s\in\mathbb{R}}$ are a natural extension of $(W_2^k)_{k\in\mathbb{N}_0}$. It is also possible to define H_p^s, extending W_p^k.

Lemma 2.6.13

Equipped with

$$\langle f, g\rangle_{H^s(\mathbb{R}^n)} := \int_{\mathbb{R}^n} w_s(x)\hat{f}(x)\overline{w_s(x)\hat{g}(x)}\,\mathrm{d}x,$$

$H^s(\mathbb{R}^n)$ becomes a HILBERT *space. We have*

$$\mathcal{S}(\mathbb{R}^n) \subset H^s(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n),$$

where the first inclusion is dense.

Remark 2.6.14 By definition, for all $0 \leqslant s \leqslant t \leqslant \infty$ we have

$$H^s(\mathbb{R}^n) \subset H^t(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$$

Remark 2.6.15 Interpreting fractional smoothness in terms of differences

$$(\nabla_h f)(x) = f(x+h) - f(x),$$

where $h, x \in \mathbb{R}^n$, we can define the space $W_2^s(\mathbb{R}^n)$ for $s = k + \sigma$, where $k \in \mathbb{N}_0$, $\sigma \in (0,1)$ without the FOURIER transform, yielding the same spaces we defined.

This leads to a SOBOLEV embedding theorem.:

$$W_2^s(\mathbb{R}^n) \hookrightarrow \mathcal{C}^\ell(\mathbb{R}^n), \quad \ell \in \mathbb{N}_0,\ s > \ell + \frac{n}{2}.$$

Example 2.6.16 (Regularisation of the numerical differentiation operator)
Let

$$K : L_0^2(0,1) \to L^2(0,1),\ (Kx)(t) := \int_0^t x(s)\,\mathrm{d}s,$$

where

$$L_0^2(0,1) := \left\{ z \in L^2(0,1) : \int_0^1 z(s)\,\mathrm{d}s = 0 \right\}$$

is a closed subspace and for $a > 0$ let

$$f_a : \mathbb{R} \to \mathbb{R},\ t \mapsto \frac{1}{a\sqrt{a}} \exp\left(-\frac{t^2}{a^2}\right).$$

Then there exists a constant $C > 0$ independent on a such that $\|f_a * z - z\|_{L^2(\mathbb{R})} \leqslant aC\|z'\|_{L^2(0,1)}$ holds for all $z \in H^1(0,1)$ with $z(0) = z(1) = 0$. In particular this implies $\|f_a * z - z\|_{L^2(\mathbb{R})} \xrightarrow{a\to 0} 0$.

Note that $f_a \in \mathcal{S}(\mathbb{R})$ and $\int_{\mathbb{R}} f_a(x)\,\mathrm{d}x = 1$ holds for all $a > 0$. ◇

3 Linear Ill-Posed Inverse Problems

For an operator $T : X \to Y$ between BANACH spaces, we consider equations of the form

$$Tx = y. \tag{17}$$

Given $y \in Y$, we want to find a solution $x \in X$ fulfilling (17).

DEFINITION 3.0.1 (WELL-POSED PROBLEM (HADAMARD))
The inverse problem (17) is well-posed if for all $y \in Y$ well-posed

1. There exists a $x \in X$ with $Tx = y$. (surjectivity of T)
2. The solution is unique. (injectivity of T)
3. The solution depends continuously on y.

If at least one condition is not fulfilled, (17) is called ill-posed. ill-posed

If T is a linear map between finite dimensional spaces, 3 is always satisfied: Let $y_n := Tx_n \to y = Tx$. Then $x_n \to x$, as ?? (in finite dimensions linear operators are continuous??)

Figure 7: [10]

Example 3.0.2 inpainting und so **TODO** ◇

Example 3.0.3 (Numerical differentiation operator)
Consider $Y := L^\infty([0,1])$ and $y \in Y$ with $y(0) = 0$. We aim to finds its derivative $x := y'$, which we can rewrite as

$$y(t) = \int_0^t x(s)\,\mathrm{d}s = \int_0^1 k(s,t)x(s)\,\mathrm{d}s, \qquad k(s,t) := \begin{cases} 1, & \text{if } s \leqslant t, \\ 0, & \text{else}, \end{cases}$$

which in turn can be written as $y = Kx$, where

$$(Kx)(t) := \int_0^1 k(s,t)x(s)\,\mathrm{d}s.$$

Assume we only know perturbed data $y = \tilde{y} + \eta$, where $\tilde{y} \in \mathcal{C}^1([0,1])$ and $\eta \in L^\infty([0,1])$ is a noise or perturbation of $\tilde{y}$.

The solution exists only if η is differentiable but even then the problem is ill-posed: Consider a zero sequence $(\delta_n)_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ and choose $\eta_n(t) := \delta_n \sin\left(\frac{kt}{\delta_n}\right)$ for $k > 0$. Then $\eta_n \in \mathcal{C}^1([0,1]) \hookrightarrow L^\infty([0,1])$ and $\|\eta_n\|_\infty = \delta_n \to 0$ and

$$x_n(t) = y_n'(t) = \tilde{y}'(t) + k\cos\left(\frac{kt}{\delta_n}\right)$$

for $n \in \mathbb{N}$. Setting $\tilde{x}(t) := \tilde{y}'(t)$ we obtain

$$\|\tilde{x} - x_n\|_{L^\infty} = \left\| k\cos\left(\frac{kt}{\delta_n}\right)\right\|_{L^\infty} = k \neq 0$$

If $Y = \mathcal{C}^1([0,1])$ the problem becomes well-posed. **TODO** ◇

3.1 The MOORE-PENROSE Pseudoinverse

Notation

We write $\mathcal{R} :=$ ran, $\mathcal{N} :=$ ker and $\mathcal{D} :=$ dom.

From now on let X and Y be Hilbert spaces and $T \in L(X, Y)$.

Remark 3.1.1 In this new setting consider (17). If $y \notin \mathcal{R}(T)$, there does not exists a solution, so we weaken the definition of solution: we choose $x \in X$ such that $\|Tx - y\|$ is minimal. If $\mathcal{N}(T) \neq \{0\}$, there exists infinitely many solutions so out of those x we pick the one with $\|x\|$ minimal as well.

DEFINITION 3.1.2 (LS AND MINIMAL NORM SOLUTION)

Let $y \in Y$. Then $x \in X$ is a

- least-square-solution of (17) if $x = \arg\min_{z \in X} \|Tz - y\|$. least-square-solution
- minimal norm solution of (17) if x is a least-square solution $x = \arg\min_{z \in X} \|z\|$. minimal norm solution

DEFINITION 3.1.3 (MOORE-PENROSE PSEUDOINVERSE)

Set $\tilde{T} := T|_{\mathcal{N}(T)^\perp} : \mathcal{N}(T)^\perp \to \mathcal{R}(T)$. The MOORE-PENROSE Pseudoinverse T^+ is the unique linear extension of $\tilde{T}^{-1}$ with $\mathcal{D}(T^+) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ and $\mathcal{N}(T^+) = \mathcal{R}(T)^\perp$.

Remark 3.1.4 The extension is unique as

$$\overline{\mathcal{D}(T^+)} = \overline{\mathcal{R}(T)} \oplus \mathcal{R}(T)^\perp = Y.$$

Therefore $\mathcal{R}(T)$ is closed if and only if $\mathcal{D}(T^+) = Y$.
Then second condition $\mathcal{N}(T^+) = \mathcal{R}(T)^\perp$ ensures that T^+ is extended by zero outside of $\mathcal{R}(T)$.

Proof. (Well-definedness of T^+) As $\mathcal{R}(T) = \mathcal{N}(T)^\perp$ holds in HILBERT spaces, $\tilde{T}$ is bijective, hence $\tilde{T}^{-1}$ exists,and thus T^+ is well-defined on $\mathcal{R}(T)$.

We can uniquely decompose $y \in \mathcal{D}(T^+)$ as $y = x + z$ with $x \in \mathcal{R}(T)$ and $z \in \mathcal{R}(T)^\perp$. Since $\mathcal{R}(T)^\perp = \mathcal{N}(T^+)$

$$T^+ y = T^+ x + T^+ z = T^+ x = \tilde{T}^{-1} x \tag{18}$$

holds, implying the well-definedness of T^+ on $\mathcal{D}(T^+)$. □

Example 3.1.5 (Pseudoinverse of $\int_a^b$ (HA 7-2, [15]))
For $a < b \in \mathbb{R}$ let

$$T : L^2(a, b) \to \mathbb{R}, \; f \mapsto \int_a^b f(x)\,\mathrm{d}x.$$

We have

$$\mathcal{N}(T) = \left\{ f \in L^2 : \int_a^b f(x) \cdot 1\,\mathrm{d}x = 0 \right\} = \{\mathbb{1}_{(a,b)}\}^\perp$$

thus

$$\mathcal{N}(T)^\perp = \{\mathbb{1}_{(a,b)}\}^{\perp\perp} = \{f : (a, b) \to \mathbb{K}, f(x) = c \in \mathbb{R} \;\; \forall x \in (a, b)\},$$

i.e. $\mathcal{N}(T)^\perp$ consists of all constant functions on (a, b). We have $\mathcal{R}(T) = \mathbb{R}$ and thus $\mathcal{R}(T)^\perp = \{0\}$. Therefore

$$\widetilde{T} : \mathcal{N}(T)^\perp \to \mathbb{R},\ \big(f(x) = c\big) \mapsto \int_a^b f(x)\,\mathrm{d}x = (b-a)c,$$

implying

$$T^+ = \widetilde{T}^{-1} : \mathbb{R} \to \mathcal{N}(T)^\perp,\ c \mapsto \left(f(x) := \frac{c}{b-a}\right).$$

Now let $w := (w_i)_{i=1}^n \subset L^2(a, b)$ be an orthonormal system and consider

$$T : L^2(a, b) \to \mathbb{R}^n,\ f \mapsto \left(\int_a^b f(x) w_i(x)\,\mathrm{d}x\right)_{i=1}^n.$$

As above, $\mathcal{R}(T) = \mathbb{R}^n$. Furthermore

$$\mathcal{N}(T) = \{f \in L^2(a, b) : \langle f, w_i\rangle = 0\ \forall i \in \{1, \ldots, n\}\} = w^\perp$$

and thus $\mathcal{N}(T)^\perp = w^{\perp\perp} = \overline{w} = \operatorname{span}(w)$. Then

$$\begin{aligned}\widetilde{T} : \operatorname{span}(w) \to \mathbb{R}^n,\ \sum_{k=1}^n a_k w_k(x) &\mapsto \left(\int_a^b \sum_{k=1}^n a_k w_k(x) w_i(x)\,\mathrm{d}x\right)_{i=1}^n \\ &= \left(\sum_{k=1}^n a_k \langle w_k, w_i\rangle\right)_{i=1}^n = (a_i)_{i=1}^n,\end{aligned}$$

holds as $\langle w_i, w_j\rangle = \delta_{i,j}$ since w is orthonormal. Thus

$$T^+ = \widetilde{T}^{-1} : \mathbb{R}^n \to \operatorname{span}(w),\ (a_i)_{i=1}^n \mapsto \sum_{k=1}^n a_k w_k(x) \qquad \diamond$$

Example 3.1.6 (Pseudoinverse of example 1.4.18 (HA 7-2, [15]))
For $0 \neq g \in \mathcal{C}(a, b)$ let

$$T : L^2(a, b) \to L^2(a, b),\ f \mapsto gf$$

T is unbounded, as gf need not be in $L^2(a, b)$, i.e. pick $(a, b) := (0, 1)$, $f := g := x^{-1/4}$. Then $\|f\|_{L^2(0,1)} = \|g\|_{L^2(0,1)} = \int_0^1 x^{-1/2}\,\mathrm{d}x = 2$ but $\|fg\|_{L^2(0,1)} = \int_0^1 x^{-1}\,\mathrm{d}x = \infty$.
Let $G := \{x \in (a, b) : g(x) \neq 0\}$. Then $\mathcal{N}(T) = \{f \in L^2(a, b) : f|_G = 0\}$ holds, and thus

$$\begin{aligned}\mathcal{N}(T)^\perp &= \{h \in L^2(a, b) : \int_a^b h(x) j(x)\,\mathrm{d}x = 0\ \forall j \in \mathcal{N}(T)\} \\ &= \{h \in L^2(a, b) : h(x) j(x) = 0\ \forall j \in \mathcal{N}(T)\} \\ &= \{h \in L^2(a, b) : h|_{G^\complement} = 0\}.\end{aligned}$$

We have $\mathcal{R}(T) = \{fg : f \in L^2(a, b)\}$ and therefore

$$\begin{aligned}\mathcal{R}(T)^\perp &= \{h \in L^2(a, b) : \int_a^b h(x) f(x) g(x)\,\mathrm{d}x = 0 : fg \in L^2(a, b)\} \\ &= \{h \in L^2(a, b) : h|_G = 0\} = \mathcal{N}(T).\end{aligned}$$

Thus

$$\widetilde{T} : \{f \in L^2(a,b) : f|_{G^\complement} = 0\} \to \{fg : f \in L^2(a,b)\},\ f \mapsto fg.$$

and

$$\widetilde{T}^{-1} : \{fg : f \in L^2(a,b)\} \to \{f \in L^2(a,b) : f|_{G^\complement} = 0\}, f \mapsto \frac{f}{g}$$

and therefore

$$T^+ : \mathcal{R}(T) \oplus \mathcal{N}(T) \to \mathcal{N}^\perp, f \mapsto \frac{f}{g}.$$ ◇

Lemma 3.1.7 (Moore-Penrose Equations)
T^+ satisfies $\mathcal{R}(T^+) = \mathcal{N}(T)^\perp$ and we have

The four equations characterise T^+ uniquely.

1. $TT^+T = T$
2. $T^+TT^+ = T^+$
3. $T^+T = \operatorname{id} - P_{\mathcal{N}(T)} = P_{\mathcal{N}(T)^\perp}$
4. $TT^+ = (P_{\overline{\mathcal{R}(T)}})|_{\mathcal{D}(T^+)}$

hold.

Proof. "$\subset$": By definition of T^+ and (18) for all $y \in \mathcal{D}(T^+)$

We can't write $P_{\mathcal{R}(T)}$ as $\mathcal{R}(T)$ might not be closed.

$$T^+y \overset{(18)}{=} \tilde{T}^{-1}P_{\overline{\mathcal{R}(T)}}y = T^+P_{\overline{\mathcal{R}(T)}}y \tag{19}$$

holds, as $\tilde{T}^{-1}$ and T^+ coincide on $\mathcal{R}(T)$. Hence $T^+y \in \mathcal{R}(\tilde{T}^{-1}) = \mathcal{N}(T)^\perp$.
"$\supset$": For all $x \in \mathcal{N}(T)^\perp$

$$T^+T = \tilde{T}^{-1}\tilde{T}x = x$$

holds as T and $\tilde{T}$ coincide on $\mathcal{N}(T)^\perp$ $\tilde{T}^{-1}$ and T^+ coincide on $\mathcal{R}(T)$. This implies $x \in \mathcal{R}(T^+)$.

4. For $y \in \mathcal{D}(T^+)$

$$TT^+y \overset{(19)}{=} TT^+P_{\overline{\mathcal{R}(T)}}y \overset{(\star)}{=} T\tilde{T}^{-1}P_{\overline{\mathcal{R}(T)}}y \overset{(\ddagger)}{=} \tilde{T}\tilde{T}^{-1}P_{\overline{\mathcal{R}(T)}}y = P_{\overline{\mathcal{R}(T)}}y$$

holds, as $\tilde{T}^{-1}$ and T^+ coincide on $\mathcal{R}(T)$ $(\star)$ and as T and $\tilde{T}$ coincide on $\mathcal{N}(T)^\perp = \mathcal{R}(T^+)$ $(\ddagger)$.

3. For all $x \in X$ we have

$$T^+Tx = \tilde{T}^{-1}Tx \quad \text{and} \quad x = P_{\mathcal{N}(T)}x + (\operatorname{id} - P_{\mathcal{N}(T)})x.$$

Hence

$$T^+Tx = \tilde{T}^{-1}\underbrace{TP_{\mathcal{N}(T)}x}_{=0} + \tilde{T}^{-1}\tilde{T}(\operatorname{id} - P_{\mathcal{N}(T)})x = (\operatorname{id} - P_{\mathcal{N}(T)})x$$

holds. As $I - P_{\mathcal{N}(T)}$ is self-adjoint

$$\mathcal{N}(T) = \operatorname{ran}(P_{\mathcal{N}(T)}) = \ker(I - P_{\mathcal{N}(T)}) = \operatorname{ran}(I - P_{\mathcal{N}(T)})^\perp,$$

implying $I - P_{\mathcal{N}(T)} = P_{\mathcal{N}(T)^\perp}$.

2. For $y \in \mathcal{D}(T^+)$

$$T^+y \overset{(19)}{=} T^+P_{\overline{\mathcal{R}(T)}}y \overset{4}{=} T^+TT^+y$$

holds.

1. For $x \in X$

$$T(T^+T)x \overset{3}{=} T(\mathrm{id} - P_{\mathcal{N}(T)})x = Tx - TP_{\mathcal{N}(T)}x = Tx \qquad \square$$

holds.

THEOREM 3.1.1: THE CASE $y \in \mathcal{D}(T^+)$

If $y \in \mathcal{D}(T^+)$, then (17) has a unique minimal norm solution $x^+ = T^+y$. The set of all least squares solutions is given by $x^+ + \mathcal{N}(T)$.

Proof. 1 We first prove the existence of a least-square solutions. Consider

$$S := \{z \in X : Tz = P_{\overline{\mathcal{R}(T)}}y\}$$

and the set Sol of all least square solutions to (17). We show that $S = \mathsf{Sol}$.

"$\subset$": As $y \in \mathcal{D}(T^+)$ by the fourth MOORE-PENROSE equation $z = T^+y \in S$ holds, implying $S \neq \varnothing$.

For all $z \in S$ and $x \in X$

$$\|Tz - y\| = \|P_{\mathcal{R}(T)}y - y\| = \min_{w \in \overline{\mathcal{R}(T)}} \|w - y\| \leqslant \|Tx - y\|$$

holds because $P_{\mathcal{R}(T)}$ is an orthogonal projection (O).

"$\supset$": Similarly, for all $s \in \mathsf{Sol}$

$$\begin{aligned} \|P_{\overline{\mathcal{R}(T)}}y - y\| &\overset{(O)}{\leqslant} \|Ts - y\| = \min_{x \in X} \|Tx - y\| \\ &= \min_{w \in \mathcal{R}(T)} \|w - y\| \leqslant \|P_{\overline{\mathcal{R}(T)}}y - y\| \end{aligned}$$

holds, implying $Tz = P_{\overline{\mathcal{R}(T)}}$.

2 We now show the existence of a unique minimal norm solution.

From 1 we now that every least square solution x to (17) solves $Tx = P_{\overline{\mathcal{R}(T)}}y$. Each such solution can be written as $x = a + b$, where $a \in \mathcal{N}(T)^\perp$ and $b \in \mathcal{N}(T)$, implying

$$\|x\|^2 = \|a\|^2 + 2\langle a, z\rangle + \|b\|^2 = \|a\|^2 + \|b\|^2 \geqslant \|a\|^2.$$

As T is injective on $\mathcal{N}(T)^\perp$, x is independent of a: If $\bar{x} \in S$ such that $\bar{x} = x_1 + x_2$, where $x_1 \in \mathcal{N}(T)^\perp \backslash \{a\}$ and $x_1 \in \mathcal{N}(T)$, it holds that $T\bar{x} = Tx_1 \neq Ta = P_{\overline{\mathcal{R}(T)}}y$, which is a contradiction.

Therefore, $x^+ = a$ is the unique minimal norm solution to (17). By the MOORE-PENROSE equations (M) we have

$$\begin{aligned} x^+ &\overset{(M)}{=} P_{\mathcal{N}(T)^\perp}x^+ = (\mathrm{id} - P_{\mathcal{N}(T)})x^+ \overset{(M)}{=} T^+Tx^+ = T^+P_{\overline{\mathcal{R}(T)}}y \\ &\overset{(M)}{=} T^+TT^+y \overset{(M)}{=} T^+y. \end{aligned} \qquad \square$$

Corollary 3.1.8
For $y \in \mathcal{R}(T)^\perp = \mathcal{N}(T^+)$ the minimum norm solution is $x^+ = 0$.

Theorem 3.1.2: Normal equation

For $y \in \mathcal{D}(T^+)$, $x \in X$ is a least-square solution of (17) if and only if $x \in X$ satisfies the normal equation $T^*Tx = T^*y$.
If in addition $x \in \mathcal{N}(T)^\perp$, we have $x = x^+$.

Proof. By the proof of the previous theorem, (1) is equivalent to

$$Tx = P_{\overline{\mathcal{R}(T)}}y.$$

By the properties of orthogonal projections this is equivalent to $Tx \in \overline{\mathcal{R}(T)}$ and $Tx - y \in \overline{\mathcal{R}(T)}^\perp = \mathcal{N}(T^*)$, i.e. $T^*(Tx - y) = 0$. □

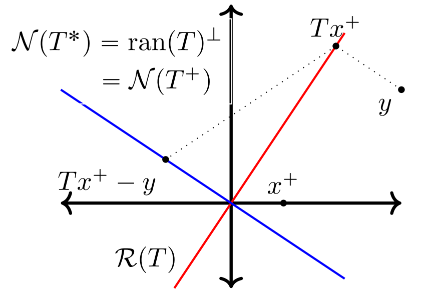

Figure 8: Visualisation of the proof of theorem 3.1.2.

Example 3.1.9 Consider the singular linear operator $T := \left(\begin{smallmatrix} 2 & 0 \\ 3 & 0 \end{smallmatrix}\right)$ on $X := \mathbb{R}^2$ in Figure 8. Consider

$$y := (3, 2) \in \mathcal{D}(T^+) = X$$

and Tx^+, its projection onto $\mathcal{R}(T) = \operatorname{span}((2, 3))$, where

$$x^+ = \frac{12}{13}(1, 1) \notin \mathcal{N}(T)^\perp = \operatorname{span}((1, 0))$$

is the minimal solution. We have

$$T^+ = \frac{1}{13}T^* = (T^*T)^+T^*$$

with $\mathcal{R}(T^+) = \operatorname{span}((0, 1))$.

We can further observe that $Tx^+ - y = \frac{12}{13}(2, 3) - (3, 2) = \frac{5}{13}(-3, 2) \in \mathcal{N}(T^*) = \operatorname{span}((-3, 2))$. ◇

Corollary 3.1.10 ($x^+ = (T^*T)^+T^*y$)
The minimal norm solution satisfies $x^+ = (T^*T)^+T^*y$.

Proof. Multiply $(T^*T)^+$ on both sides of the normal equation. □

Theorem 3.1.3: Continuity of T^+

$T^+ \in L(\mathcal{D}(T^+), X)$ holds if and only if $\mathcal{R}(T)$ is closed.

Recall $\overline{\mathcal{R}(T)}$ closed $\iff \mathcal{D}(T^+) = Y$.

Proof. " $\Longrightarrow$ ": As $\mathcal{D}(T^+) \subset Y$ is dense, T^+ can be uniquely and continuously extended to Y by $\overline{T^+} \in L(Y, X)$ defined by $\overline{T^+}y := \lim_{n\to\infty} T^+y_n$ for some sequence $(y_n)_{n\in\mathbb{N}} \subset \mathcal{D}(T^+)$ converging to $y \in Y$.

Let $(y_n)_{n\in\mathbb{N}} \subset \mathcal{R}(T)$ be a sequence converging to $y \in \overline{\mathcal{R}(T)}$. By the fourth Moore-Penrose equation and the continuity of T

$$y = P_{\overline{\mathcal{R}(T)}}y = \lim_{n\to\infty} P_{\overline{\mathcal{R}(T)}}y_n = \lim_{n\to\infty} TT^+y_n = TT^+y \in \mathcal{R}(T),$$

hence $\mathcal{R}(T) = \overline{\mathcal{R}(T)}$.

" $\Longleftarrow$ ": First we show that T^+ is closed.

Let $(y_n)_{n\in\mathbb{N}} \subset \mathcal{D}(T^+)$ converge to $y \in Y$ with $T^+y_n \to x \in X$. By the fourth Moore-Penrose formula

$$TT^+y_n = P_{\overline{\mathcal{R}(T)}}y_n \to P_{\overline{\mathcal{R}(T)}}y$$

holds by the continuity of orthogonal projections. Since T is continuous,

$$P_{\overline{\mathcal{R}(T)}}y = \lim_{n\to\infty} P_{\overline{\mathcal{R}(T)}}y_n = \lim_{n\to\infty} TT^+y_n = Tx,$$

implying that x is a least-square solution to (17).

As $T^+y_n \in \mathcal{R}(T^+) = \mathcal{N}(T)^\perp$, which is closed, holds for all $n \in \mathbb{N}$ we have that $T^+y_n \to x \in \mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$. Similarly to the proof of theorem 3.1.1, x is a minimal norm solution to (17), so T^+ is closed. The closed graph theorem now implies 1. □

Corollary 3.1.11 (Compact operators)
Let $K \in \mathcal{K}(X, Y)$ with $\dim(\mathcal{R}(K)) = \infty$. Then K^+ is not continuous.

Proof. Towards contradiction assume that K^+ is continuous. By theorem 3.1.3 $\mathcal{R}(K)$ is closed. Let $\tilde{K} := K|_{\mathcal{N}(K)^\perp} : \mathcal{N}(K)^\perp \to \mathcal{R}(K)$, which is bijective. Then $\tilde{K}^{-1} \in L(\mathcal{R}(K), \mathcal{N}(K)^\perp)$ holds by the inverse mapping theorem.

As K is compact, so is $K \circ \tilde{K}^{-1}$, which is the identity on $\mathcal{R}(K)$. In FA I this way shown to imply $\dim(\mathcal{R}(K)) < \infty$, a contradiction. □

3.2 Singular Value Decomposition for Compact Operators

From now on let $K \in \mathcal{K}(X, Y)$ and ONB, ONS be the shorthand for orthonormal basis and orthonormal system, respectively.

Definition 3.2.1 (Singular Value Decomposition)

A sequence $((\sigma_n, u_n, v_n))_{n\in\mathbb{N}}$ is the singular value decomposition of K if $(\sigma_n)_n \subset \mathbb{R}^+$ is a decreasing sequence converging to 0, $(u_n)_{n\in\mathbb{N}} \subset Y$ an ONB of $\overline{\mathcal{R}(K)}$ and $(v_n)_{n\in\mathbb{N}} \subset X$ an ONB of $\overline{\mathcal{R}(K^*)}$ such that

1. $Kv_n = \sigma_n v_n$ and $K^* u_n = \sigma_n v_n$ holds for all $n \in \mathbb{N}$
2. $Kx = \sum_{n\in\mathbb{N}} \sigma_n \langle x, v_n\rangle u_n$ holds for all $x \in X$.

Theorem 3.2.1: Singular value decomposition

For $K \in \mathcal{K}(X, Y)$ there exists a singular value decomposition.

Proof. As $K^*K : X \to X$ is compact and self-adjoint, there exists a in absolute value decreasing sequence $(\lambda_n)_{n\in\mathbb{N}} \in \mathbb{R}\setminus\{0\}$ converging to zero and a ONS $(v_n)_{n\in\mathbb{N}} \subset X$ such that

$$K^*Kx = \sum_{n\in\mathbb{N}} \lambda_n \langle x, v_n\rangle v_n$$

holds for all $x \in X$. We have

$$\lambda_n = \lambda\|v_n\|^2 = \langle \lambda_n v_n, v_n\rangle = \langle K^*Kv_n, v_n\rangle = \langle Kv_n, Kv_n\rangle = \|Kv_n\|^2 > 0$$

for all $n \in \mathbb{N}$. Set $\sigma_n := \sqrt{\lambda_n} > 0$ and $u_n := \sigma_n^{-1} K v_n \in Y$ for all $n \in \mathbb{N}$. Then $(u_n)_{n\in\mathbb{N}} \subset Y$ is an ONS since

$$\langle u_i, u_j\rangle = \frac{1}{\sigma_i\sigma_j}\langle Kv_i, Kv_j\rangle = \frac{1}{\sigma_i\sigma_j}\langle K^*Kv_i, v_j\rangle = \frac{\lambda_i}{\sigma_i\sigma_j}\langle v_i, v_j\rangle = \delta_{i,j}$$

holds for all $i, j \in \mathbb{N}$. Furthermore

$$K^*u_n = \sigma_n^{-1}K^*Kv_n = \sigma_n^{-1}\lambda_n v_n = \sigma_n v_n \tag{20}$$

holds for all $n \in \mathbb{N}$.

By the spectral theorem, $(v_n)_{n\in\mathbb{N}}$ is an ONB for $\overline{\mathcal{R}(K^*K)} = \overline{\mathcal{R}(K)}$ The equality can be seen like this: for $x \in \overline{\mathcal{R}(K^*)}\setminus\{0\}$ there exists a sequence $(y_n)_{n\in\mathbb{N}} \subset Y$ such that $K^*y_n \to x$. Without loss of generality assume $(y_n)_{n\in\mathbb{N}} \subset \mathcal{N}(K^*)^\perp = \overline{\mathcal{R}(K)}$. Using a diagonal argument we have $x \in \overline{\mathcal{R}(K^*K)}$. The other inclusion is clear.

Hence $(v_n)_{n\in\mathbb{N}}$ can be extended to an ONB V for X, as the remaining elements must be in $\mathcal{N}(K) = \overline{\mathcal{R}(K^*)}^\perp$. Thus

$$\begin{aligned} Kx &= \sum_{v\in V}\langle x, v\rangle Kv = \sum_{n\in\mathbb{N}}\langle x, v_n\rangle Kv_n \overset{1}{=} \sum_{n\in\mathbb{N}}\langle x, v_n\rangle \sigma_n u_n \\ &\overset{(20)}{=} \sum_{n\in\mathbb{N}}\langle x, K^*u_n\rangle u_n = \sum_{n\in\mathbb{N}}\langle Kx, u_n\rangle u_n \end{aligned}$$

holds for all $x \in X$, implying that $(u_n)_{n\in\mathbb{N}}$ is an ONB for $\overline{\mathcal{R}(K)}$. $\square$

Theorem 3.2.2: Picard condition

Let $((\sigma_n, u_n, v_n))_{n\in\mathbb{N}}$ be a singular system for K and $y \in \overline{\mathcal{R}(K)}$. Then $y \in \mathcal{R}(K)$ holds if and only if the Picard condition

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2} |\langle y, u_n\rangle|^2 < \infty \tag{21}$$

is satisfied. In this case we have

$$K^+ y = \sum_{n\in\mathbb{N}} \sigma_n^{-1} \langle y, u_n\rangle v_n. \tag{22}$$

Proof. " $\Longrightarrow$ ": Let $y \in \mathcal{R}(K)$. Then there exists an $x \in X$ such that $Kx = y$. Thus for all $n \in \mathbb{N}$

$$\langle y, u_n\rangle = \langle Kx, u_n\rangle = \langle x, K^* u_n\rangle = \sigma_n \langle x, v_n\rangle$$

holds. The Bessel inequality implies

$$\sum_{n\in\mathbb{N}} \sigma_n^{-2} |\langle y, u_n\rangle|^2 = \sum_{n\in\mathbb{N}} |\langle x, v_n\rangle|^2 \leqslant \|x\|^2 < \infty.$$

" $\Longleftarrow$ ": Let $y \in \overline{\mathcal{R}(K)}$. By (21) $\left(\sum_{n=1}^{N} \sigma_n^{-2} |\langle y, u_n\rangle|^2\right)_{N\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Also $\left(X_N := \sum_{n=1}^{N} \sigma_n^{-1} \langle y, u_n\rangle v_n\right)_{N\in\mathbb{N}}$ is a Cauchy sequence in X: for $M \geqslant N$

$$\|X_N - X_M\|^2 \leqslant \left\| \sum_{n=N}^{M} \sigma_n^{-1} \langle y, u_n\rangle v_n \right\|^2 = \sum_{n=N}^{M} \sigma_n^{-2} |\langle y, u_n\rangle|^2 \xrightarrow{N,M\to\infty} 0$$

holds by Parseval's identity. As $(v_n)_{n\in\mathbb{N}} \subset \overline{\mathcal{R}(K^*)}$, $(X_N)_{N\in\mathbb{N}} \subset \overline{\mathcal{R}(K^*)}$ converges to some $x := \sum_{n\in\mathbb{N}} \sigma_n^{-1} \langle y, u_n\rangle v_n \in X$ As $\overline{\mathcal{R}(K^*)}$ is closed, $x \in \overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$. Thus

$$Kx = \sum_{n\in\mathbb{N}} \sigma_n^{-1} \langle y, u_n\rangle K v_n = \sum_{n\in\mathbb{N}} \sigma_n^{-1} \langle y, u_n\rangle u_n = P_{\overline{\mathcal{R}(K)}} = y,$$

holds, implying $y \in \mathcal{R}(K)$.

By theorem 3.1.2 $Kx = P_{\overline{\mathcal{R}(K)}} y$ and $x \in \mathcal{N}(K)^\perp$ is equal to $K^+ y = x$. □

Remark 3.2.2 (Interpretation of theorem 3.2.2)
By the Picard condition, a minimal norm solution can only exists if the "Fourier coefficients" $(\langle y, u_n\rangle)_{n\in\mathbb{N}}$ decay rapidly compared to the singular values $(\sigma_n)_{n\in\mathbb{N}}$.

Example 3.2.3 (Perturbations)
Consider (22). If $y_n^{(\delta)} := y + \delta u_n$ for some $\delta > 0$,

$$\|K^+ y_n^{(\delta)} - K^+ y\| = \delta \|K^+ u_n\| = \delta \sigma_n^{-1} \xrightarrow{n\to\infty} \infty$$

holds. This shows how perturbations of y affect perturbations of x^+: the faster $(\sigma_n)_{n\in\mathbb{N}}$ decay, the worse the amplification of the error is. ◇

The previous remark motivates classifying ill-conditioned problems.

Definition 3.2.4 (Types of ill-conditionedness)

1. (17) is moderately ill-conditioned if the decay of the singular values is at most polynomial, i.e there exist $c, r > 0$ such that $\sigma_n \geqslant cn^{-r}$ for all $n \in \mathbb{N}$.
2. If 1 is not the case, (17) is strongly ill-conditioned.
3. (17) is called exponentially ill-conditioned if there exists $c, r > 0$ such that $\sigma_n \leqslant ce^{-nr}$ for all $n \in \mathbb{N}$.

moderately ill-conditioned

strongly ill-conditioned

exponentially ill-conditioned

Example 3.2.5 (SVD of the Volterra operator (HW 10-1))
Consider Volterra integral operator from example 3.0.3 with $X := L^2([0,1])$. Then (17) is moderately ill-conditioned.

The adjoint of K is given by

$$K^* : L^2([0,1]) \to L^2([0,1]), \ (K^*x)(t) := \int_t^1 x(s)\,\mathrm{d}s$$

as using Fubini it holds that

$$\langle Kx, y\rangle = \int_0^1 \int_0^t x(s)\,\mathrm{d}s y(t)\,\mathrm{d}t = \int_0^1 x(s) \int_s^1 y(t)\,\mathrm{d}t\,\mathrm{d}s = \langle x, K^*y\rangle,$$

as we are integrating over a triangular region, see figure 9. Thus

$$(Tx)(\xi) := (K^*Kx)(\xi) = \int_\xi^1 \int_0^t x(s)\,\mathrm{d}s\,\mathrm{d}t$$

holds. In order for $\lambda \neq 0$ to be an eigenvalue of K^*K, $(Tx)(\xi) = \lambda x(\xi)$, i.e

$$\int_\xi^1 \int_0^t x(s)\,\mathrm{d}s\,\mathrm{d}t = \lambda x(\xi)$$

must hold. This yields $x(1) = 0$. Differentiating with respect to ξ yields

$$\lambda x'(\xi) = -\int_0^\xi x(s)\,\mathrm{d}s,$$

implying $x'(0) = 0$ and $\lambda x''(\xi) = -x(\xi)$. We have reduced the eigenvalue problem to a second order ordinary differential equation in x:

$$\begin{cases} x''(\xi) + \frac{1}{\lambda}\cdot x(\xi) = 0, \\ x(1) = 0, \\ x'(0) = 0. \end{cases} \tag{23}$$

The same argument shows that for well-behaved kernel $k(s,t)$ (in this example $k(s,t) := \mathbb{1}_{\{s\leqslant t\}}(s,t)$ we have $(K^*x)(t) = \int_0^1 k(t,s)x(s)\,\mathrm{d}s$.

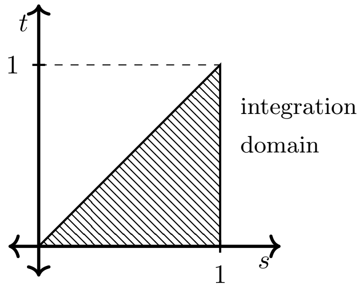

Figure 9: The integration domain

The solution of (23) is given by

$$x(\xi) = c_1\cos(t\xi) + c_2\sin(t\xi), \qquad \text{where } t := \frac{1}{\sqrt{\lambda}}.$$

Plugging in the second boundary condition into $x'(\xi) = c_2 t\cos(t\xi) - c_1 t\sin(t\xi)$ yields $x'(0) = c_2 t$, implying $c_2 = 0$. Plugging in the first boundary condition yields

$$x(1) = c_1\cos(t) + c_2\sin(t) = c_1\cos(t) \overset{!}{=} 0 \implies \frac{1}{\sqrt{\lambda}} = t = \pi\left(n - \frac{1}{2}\right)$$

for all $n \in \mathbb{Z}$. The roots of the eigenvalues λ of K^*K are the singular values of K. Therefore the singular values of K are given by

$$\left(\sigma_n := \frac{1}{\pi\left(n - \frac{1}{2}\right)}\right)_{n\in\mathbb{N}_{>0}}$$

Therefore K is moderately ill-conditioned as $\sigma_n \geqslant \frac{1}{\pi}\cdot n^{-1}$. ◇

Remark 3.2.6 (Functional calculus with SVD)
The singular value decomposition allows us to define functions of compact operators: Let $f : [0, \infty) \to \mathbb{R}$ be a piecewise continuous (locally bounded?) function. For $K \in \mathcal{K}(X, Y)$ with singular system $((\sigma_n, u_n, v_n))_{n\in\mathbb{N}}$ and $x \in X$ define

$$f(K^*K) : X \to X,\ x \mapsto \sum_{n\in\mathbb{N}} f(\sigma_n^2)\langle x, v_n\rangle v_n + f(0)P_{\mathcal{N}(K)}x$$

This series converges in X, as f is evaluated on the compact interval $[0, \sigma_1^2] = [0, \|K\|^2]$. We have $f(K^*K) \in L(X)$:

$$\|f(K^*K)\| = \sup_{n\in\mathbb{N}} |f(\sigma_n^2)| \leqslant \sup_{\lambda\in[0,\sigma_1^2]} |f(\lambda)| < \infty.$$

Example 3.2.7 Let $f = \sqrt{\cdot}$. The absolute value of K is absolute value

$$|K| := f(K^*K) = \sum_{n\in\mathbb{N}} \sigma_n\langle\cdot, v_n\rangle v_n. \qquad \diamond$$

3.3 Regularisation of ill-posed problems

Consider (17) to be ill-posed. For $y \in \mathcal{D}(T^+)$ there exists a unique minimal norm solution $x^+ = T^+y$ by theorem 3.1.1. Often, one only knows a perturbation $y^{(\delta)} \in Y$ with $\|y - y^{(\delta)}\| \leqslant \delta$ for some $\delta > 0$. In general, T^+ need not be continuous, so $T^+y^{(\delta)}$ need not be close to x^+. If even $y^{(\delta)} \notin \mathcal{D}(T^+)$, it is impossible to calculate $T^+y^{(\delta)}$.

We will deal with the former problem as follows: we find an approximation $x_a^{(\delta)}$ which depends continuously on $y^{(\delta)}$ and hence on δ. We then employ regularisation parameters $a = a(\delta) > 0$ such that $x_{a(\delta)}^{(\delta)} \xrightarrow{\delta\to 0} x^+$.

To achieve this we need (at least) a pointwise approximation of the pseudoinverse by continuous operators:

DEFINITION 3.3.1 (REGULARISATION OF T^+)
A family $(R_a)_{a>0} \subset L(Y, X)$ is called regularisation of T^+ if $R_a y \xrightarrow{a\to 0} T^+y$ holds for all $y \in \mathcal{D}(T^+)$.

regularisation of T^+

THEOREM 3.3.1: REGULARISATION & CONTINUITY I

Let $(R_a)_{a>0}$ be a regularisation of T^+. If T^+ is not continuous, then $(R_a)_{a>0}$ is not uniformly bounded.

Proof. Assume that $(R_a)_{a>0}$ is uniformly bounded, i.e. there exists an $M > 0$ such that $\|R_a\| \leqslant M$ holds for all $a > 0$. Since $R_a \to T^+$ converges pointwise on $\mathcal{D}(T^+)$, this convergence also holds on $\overline{\mathcal{D}(T^+)} = Y$ by the Banach-Steinhaus theorem. Thus $T^+ \in L(Y, X)$. □

Remark 3.3.2 If $(R_a)_{a>0}$ is not uniformly bounded, there exists a $y \in Y$ such that $\|R_a y\| \xrightarrow{a\to 0} \infty$: Assume the opposite, then then $(R_a)_{a>0}$ would be pointwise bounded, which contradicts the uniform boundedness principle.

THEOREM 3.3.2: REGULARISATION & CONTINUITY II

Let T^+ be discontinuous and $(R_a)_{a>0}$ be a regularisation of T^+. If $\sup_{a>0} \|TR_a\| < \infty$, then $\|R_a y\| \xrightarrow{a\to 0} \infty$ for all $y \notin \mathcal{D}(T^+)$.

Proof. Let $y \in Y$. Assume that there exists a zero sequence $(a_n)_{n\in\mathbb{N}}$ such that $(R_{a_n} y)_{n\in\mathbb{N}}$ is bounded. Since X is reflexive there exists a weakly convergent subsequence $\left(x_k := R_{a_{n_k}} y\right)_{k\in\mathbb{N}}$ with $x_k \rightharpoonup x \in X$.

As every $T \in L(X, Y)$ is weakly continuous, $Tx_k \rightharpoonup Tx$ holds. By the continuity of T and the pointwise convergence $R_a \to T^+$ we have

$$TR_a z \to TT^+z = P_{\overline{\mathcal{R}(T)}} z, \quad z \in \mathcal{D}(T^+).$$

Since $\sup_{a>0} \|TR_a\| < \infty$, $TR_a z$ converges for all $z \in Y$. As $Tx_k \to P_{\overline{\mathbb{R}(T)}} y$ **(for all y or only on $\mathcal{D}(T^+)$?? Can it extended by continuity??)** and $Tx_k \rightharpoonup Tx$, we have $Tx = P_{\overline{\mathcal{R}(T)}} y$, which implies $y \in \mathcal{D}(T^+)$ **(WHY???)**. □

Parameter choice rules

We now investigate the case $y \notin \mathcal{D}(T^+)$. Consider the total error

$$\|R_a y^{(\delta)} - T^+ y\| \leqslant \|R_a y^{(\delta)} - R_a y\| + \|R_a y - T^+ y\|.$$

Then second term is called the error of the regularisation method converges to 0 for $a \to 0$. Then first term is called the data error and is unbounded for $\delta > 0$ if T^+ is discontinuous.

Definition 3.3.3 (Parameter choice rules)

A function $a : \mathbb{R}_{>0} \times Y \to \mathbb{R}_{>0}$ is a parameter choice rule or strategy. We distinguish between

parameter choice rule

- a priori strategies, where a depends only on δ.
- a posteriori strategies, where a depends on δ and $y^{(\delta)}$.
- heuristic strategies, where a depends only on $y^{(\delta)}$.

If $(R_a)_{a>0}$ is a regularisation of T^+ and a a strategy, then (R_a, a) is called (convergent) regularisation method

regularisation method

$$\sup_{\|y^{(\delta)}-y\|} \|R_{a(\delta, y^{(\delta)})} y^{(\delta)} - T^+ y\| \xrightarrow{\delta \to 0} 0$$

holds for all $y \in \mathcal{D}(T^+)$. We then assume that for all $y \in Y$ we have

$$\lim_{\delta \to 0} \sup\{a(\delta, y^{(\delta)}) : \|y^{(\delta)} - y\| \leqslant \delta\} = 0.$$

From now let $(R_a)_{a>0}$ be a regularisation of T^+.

Theorem 3.3.3: Existence of a priori strategy

For every $(R_a)_{a>0}$ there exists an a priori parameter strategy such that (R_a, a) is a regularisation method.

Proof. TODO $\square$

Theorem 3.3.4

$(R_a)_{a>0}$ and an a priori strategy $a : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ form a regularisation method if and only if $a(\delta)$, $\delta \|R_{a(\delta)}\| \xrightarrow{\delta \to 0} 0$.

Proof. TODO $\square$

bla bla zwei remarks

Theorem 3.3.5: TODO

If there exists a heuristic rule a such that (R_a, a) is a regularisation method, T^+ is continuous.

bla bla remark

Convergence rates

bla bla remark

definition 3.3.4 (Worst case error)

The worst case error is

$$\varepsilon(y,\delta) := \sup_{\|y^{(\delta)}-y\|\leqslant\delta} \|R_{a(\delta,y^{(\delta)})}y^{(\delta)} - T^+y\|.$$

Theorem 3.3.6: TODO

Let (R_a, a) be a regularisation method. IF there exists a function $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ with $f(t) \xrightarrow{t\to 0} 0$ and $\sup_{y\in\mathcal{D}(T^+)\cap B_1(0_Y)} \varepsilon(y,\delta) \leqslant f(\delta)$, then T^+ is continuous.

Proof. TODO □

bla bla remark

4 Spectral regularisation

Remark 4.0.1 For compact operators $K \in \mathcal{K}(X,Y)$ regularisation can be constructed using the singular value decomposition of K. Recall that $K^+y = (K^*K)^+K^*y$ for $y \in \mathcal{D}(K^+)$ by theorem 3.1.2.

Let $((\sigma_n, u_n, v_n))_{n\in\mathbb{N}}$ be a singular system for K. Then $((\sigma_n^2, v_n, v_n))_{n\in\mathbb{N}}$ is a singular system for K^*K. Thus

$$\begin{aligned}(K^*K)^+K^*y &= \sum_{n\in\mathbb{N}} \sigma_n^{-2}\langle K^*y, v_n\rangle v_n = \sum_{n\in\mathbb{N}} \sigma_n^{-2}\langle y, Kv_n\rangle v_n \\ &= \sum_{n\in\mathbb{N}} \sigma_n^{-2}\sigma_n\langle y, u_n\rangle v_n = \sum_{n\in\mathbb{N}} f(\sigma_n^2)\sigma_n\langle y, u_n\rangle v_n,\end{aligned}$$

where $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $x \mapsto x^{-1}$.

The discontinuity of K^+ is related to the fact that $f|_{(0,\|K^*K\|]}$ is unbounded because $\sigma_n \to 0$. To regularise this problem we replace f by a family of bounded functions converging pointwisely to f.

Let $\kappa := \|K\|^2 = \|K^*K\| = \sigma_1^2$.

Definition 4.0.2 (Regularising filter)
A set of piecewise continuous bounded functions $(f_a : (0,\kappa] \to \mathbb{R})_{a>0}$ is called regularising filter if

regularising filter

1. $f_a(x) \xrightarrow{a\to 0} f(x) = x^{-1}$ holds for all $x \in (0,\kappa]$.
2. there exists a constant $C_f > 0$ such that $x|f_a(x)| \leqslant C_f$ holds for all $x \in (0,\kappa]$ and all $a > 0$.

Example 4.0.3 (Truncated SVD as regularisation filter)
The set $(f_a(x) := x^{-1}\mathbb{1}_{\{x\geqslant a\}})_{a>0}$ is regularisation filter, as $|f_a(x)| \leqslant a^{-1}$ and f_a is piecewise continuous for $a > 0$, so $C_f = 1$. The associated regularisation operator is given by

$$f_a(K^*K)K^*y = \sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\sigma_n\langle y, u_n\rangle v_n = \sum_{\sigma_n^2\geqslant a} \frac{\langle y, u_n\rangle v_n}{\sigma_n}. \qquad \diamond$$

Example 4.0.4 (Tikhonov regularisation)
The set $(f_a(x) := (x+a)^{-1})_{a>0}$ is a regularisation filter as $|f_a(x)| \leqslant x^{-1}$ and f_a is continuous for $\alpha > 0$, so $C_f = 1$. The associated regularisation operator is given by

$$f_a(K^*K)K^*y = \sum_{n\in\mathbb{N}} \frac{\sigma_n}{\sigma_n^2 + a}\langle y, u_n\rangle v_n$$

The main reason for the popularity of this regularisation is that the regularisation operator can be computed without knowledge of a singular system for K. $\diamond$

Lemma 4.0.5 ($\|Kf_a(K^*K)K^*\|$ bound)
For a regularisation filter $(f_a)_{a>0}$ and all $a > 0$ we have

$$\|Kf_a(K^*K)K^*\| \leqslant \sup_{n\in\mathbb{N}} |f_a(\sigma_n^2)\sigma_n^2 \leqslant C_f.$$

Proof. For $y \in Y$ and $a > 0$ (cf. remark 3.2.6) **(WAS IST MIT $f(0)P_{\mathcal{N}(K)}$???)**

$$\begin{aligned} Kf_a(K^*K)K^*y &= K\Big[\sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\langle K^*y, v_n\rangle v_n\Big] \\ &= K\Big[\sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\sigma_n\langle y, u_n\rangle v_n\Big] \\ &= \sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\sigma_n\langle y, u_n\rangle Kv_n \\ &= \sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\sigma_n^2\langle y, u_n\rangle u_n \end{aligned}$$

holds, hence

$$\begin{aligned} \|Kf_a(K^*K)K^*y\|^2 &= \Big\|\sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\sigma_n^2\langle y, u_n\rangle u_n\Big\|^2 \\ &\leqslant \sup_{n\in\mathbb{N}}(|f_a(\sigma_n^2)|\sigma_n^2)^2 \sum_{n\in\mathbb{N}} |\langle y, u_n\rangle|^2 \\ &\leqslant \sup_{n\in\mathbb{N}}(|f_a(\sigma_n^2)|\sigma_n^2)^2\|y\|^2, \end{aligned}$$

using Parseval's identity and the Bessel inequality in that order. Taking square roots yields the claims since $0 < \sigma_n^2 \leqslant \sigma_1^2 = \kappa$ and $(f_a)_{a>0}$ is a regularisation filter. □

Lemma 4.0.6 (Regularisation operator bound)

For a regularisation filter $(f_a)_{a>0}$ and all $\alpha > 0$ we have

$$\|f_a(K^*K)K^*\|^2 \leqslant C_f \cdot \sup_{x\in(0,\kappa]} |f_a(x)| < \infty.$$

Proof. For $y \in Y$ and $\alpha > 0$ as $K^*u_n = \sigma_n v_n$

$$\begin{aligned} \|f_a(K^*K)K^*y\|^2 &= \langle f_a(K^*K)K^*y, f_a(K^*K)K^*y\rangle \\ &= \langle f_a(K^*K)K^*y, \sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\sigma_n\langle y, u_n\rangle v_n\rangle \\ &= \sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\langle y, u_n\rangle\langle f_a(K^*K)K^*y, \sigma_n v_n\rangle \\ &= \sum_{n\in\mathbb{N}} f_a(\sigma_n^2)\langle y, u_n\rangle\langle Kf_a(K^*K)K^*y, u_n\rangle \\ &\leqslant \sup_{n\in\mathbb{N}} |f_a(\sigma_n^2)|\langle Kf_a(K^*K)K^*y, \sum_{n\in\mathbb{N}}\langle y, u_n\rangle u_n\rangle \\ &\overset{\text{CS}}{\leqslant} \sup_{n\in\mathbb{N}} |f_a(\sigma_n^2)| \cdot \|Kf_a(K^*K)K^*y\| \cdot \|P_{\overline{\mathcal{R}(K)}}y\| \\ &\leqslant \sup_{n\in\mathbb{N}} |f_a(\sigma_n^2)| \cdot C_f \cdot \|y\|^2, \end{aligned}$$

where the last inequality follows from the proof of lemma 4.0.5. **(GILT $\|P_{\overline{\mathcal{R}(K)}}y\| \leqslant \|y\|$???)**. The claim now follows from the boundedness of f_a. □

Theorem 4.0.1: TODO

Let $(f_a)_{a>0}$ be a regularisation filter. For

- all $y \in \mathcal{D}(K^+)$, $f_a(K^*K)K^*y \to K^+y$ holds.
- discontinuous K^+, $\|f_a(K^*K)K^*y\| \to \infty$ for $y \notin \mathcal{D}(K^+)$.

Proof. TODO $\square$

We fix the regularisation filter $(f_a)_{a>0}$ as the Tikhonov regularisation.

Tikhonov regularisation

Theorem 4.0.2: Tikhonov regularisation

For $y \in Y$ and $a > 0$, $x_a := R_a y := f_a(K^*K)K^*y$ is the unique solution of

$$(K^*K + a\,\mathrm{id})x = K^*y.$$

Proof. Let $((\sigma_n, u_n, v_n))_{n\in\mathbb{N}}$ be a singular system for K. Then

$$ax_a = \sum_{n\in\mathbb{N}} \frac{a\sigma_n}{\sigma_n^2 + a}\langle y, u_n\rangle v_n$$

holds. Now $K^*Kv_n = K^*\sigma_n u_n = \sigma_n^2 v_n$ implies

$$K^*Kx_a = \sum_{n\in\mathbb{N}} \frac{\sigma_n}{\sigma_n^2 + a}\langle y, u_n\rangle K^*Kv_n = \sum_{n\in\mathbb{N}} \frac{\sigma_n^3}{\sigma_n^2 + a}\langle y, u_n\rangle v_n.$$

Hence

$$(K^*K + a\,\mathrm{id})x_a = \sum_{n\in\mathbb{N}} \frac{a\sigma_n + \sigma_n^3}{\sigma_n^2 + a}\langle y, u_n\rangle v_n = \sum_{n\in\mathbb{N}} \sigma_n\langle y, u_n\rangle v_n = K^*y$$

holds. The uniqueness follows from the injectivity of $K^*K + a\,\mathrm{id}$:

$$\langle K^*Kx + ax, x\rangle \geqslant \|Kx\|^2 + a\|x\|^2 \geqslant a\|x\|^2$$ $\square$

Theorem 4.0.3: Tikhonov regularisation II

$R_a y$, $y \in Y$ is the unique minimiser of the Tikhonov functional

$$J_a(x) := \frac{1}{2}\|Kx - y\|^2 + \frac{a}{2}\|x\|^2.$$

Proof. For $x \in X$ **HIER NICHT BETRAGSSTRICHE?**

$$\begin{aligned}
2J_a(x) - 2J_a(x_a) &= \langle Kx - y, Kx - y\rangle + a\langle x, x\rangle \\
&\quad - \langle Kx_a - y, Kx_a - y\rangle - a\langle x_a, x_a\rangle \\
&\overset{(\star)}{=} \|Kx - Kx_a\|^2 + a\|x - x_a\|^2 \\
&\quad + \langle K^*(Kx_a - y) + ax_a, x - x_a\rangle \\
&= \|Kx - Kx_a\|^2 + a\|x - x_a\|^2 \geqslant 0,
\end{aligned}$$

holds, where $\langle K^*(Kx_a - y) + ax_a, x - x_a\rangle$ vanishes due to theorem 4.0.2. Hence x_a is minimal.

Let $\bar{x}$ be another minimiser. Then

$$0 = 2J_a(\bar{x}) - 2J_a(x_a) \geqslant \|\bar{x} - x_a\|^2 \geqslant 0$$

holds with equality if and only if $x_a = \bar{x}$.

We now explain ($\star$) in more detail:

$$
\begin{aligned}
&\langle Kx - y, Kx - y\rangle + a\|x\|^2 - \langle Kx_a - y, Kx_a - y\rangle - a\|x_a\|^2 \\
=&\langle K^*(Kx - y), x\rangle - \langle Kx - y, y\rangle + a\|x\|^2 - \langle K^*(Kx_a - y), x_a\rangle + \langle Kx_a - y, y\rangle - a\|x_a\|^2 \\
=&\langle K^*(Kx - y), x\rangle + a\|x\|^2 - \langle K^*(Kx_a - y), x_a\rangle + \langle Kx_a - Kx, y\rangle - a\|x_a\|^2 \\
=&\langle K^*(Kx - y) + ax, x\rangle - \langle K^*(Kx_a - y) + ax_a, x_a\rangle + \langle Kx_a - Kx, y\rangle \\
=&\underbrace{\langle K^*Kx - K^*y + ax, x\rangle}_{=0 \quad (\ddagger)} - \langle K^*K(x_a - x) + a(x_a - x), x_a\rangle + \langle x_a - x, K^*y\rangle \\
=& - \langle K^*K(x_a - x) + a(x_a - x), x_a\rangle + \langle x_a - x, K^*Kx + ax\rangle \\
=& - \langle K^*K(x_a - x), x_a\rangle - a\langle x_a - x, x_a\rangle + \langle x_a - x, K^*Kx\rangle + a\langle x_a - x, x\rangle \\
=&\langle K^*K(x_a - x), -x_a\rangle + \langle x_a - x, K^*Kx\rangle - a\|x_a - x\|^2 \\
=&\langle K^*K(x_a - x), x - x_a\rangle - a\|x_a - x\|^2 = -\|Kx - Kx_a\|^2 - a\|x - x_a\|^2.
\end{aligned}
$$

where in ($\ddagger$) we use $K^*y = K^*Kx + ax$ from theorem 4.0.2.

TODO SOMETHING IS WRONG HERE □

Remark 4.0.7 (Interpretation of the TIHONOV functional) The TIHONOV functional can be interpreted as the sum of the so-called residual or data fidelity term $\frac{1}{2}\|Kx - y\|$ measuring how well x "relates" to y and the penalty or regularisation term $\frac{a}{2}\|x\|^2$. The regularisation parameter a balances the terms. The smaller the noise, the more emphasis can be put on minimising the residual term, hence the smaller a can be chosen. A too large a puts more emphasis on minimising the regularisation term without appropriately minimising the residual term.

Remark 4.0.8 (Outlook) The exists more data adapted extension of the TIHONOV functional, where $\|x\|$ is replaced by a more general penalty term $P(x)$ ensuring continuous dependence of the data and incorporating properties of the solution.

In one approach [14], $(f_i)_{i\in I} \subset X$ is an ONB for X and $(w_i)_{i\in I} \subset \mathbb{R}_{>0}$ a sequence of weights. For $p \in [1, 2]$ consider the penalty term $\sum_{i\in I} w_i|\langle x, f\rangle|^p$. If the penalty term is small for a solution x_w, the coefficients $(\langle x, f_i\rangle)_{i\in I}$ decay rapidly. Then e.g. choose $w_i = 1$ and $p = 1$.

An extension is the notion of frames which generalises the concept of orthonormal bases. For this, see FA III.

LANDWEBER Regularisation

Recall that by theorem 3.1.2 for $y \in \mathcal{D}(K^+)$, $x \in X$ is a least square solution if and only if $K^*Kx = K^*y$ holds. If $x \in \mathcal{N}(K)^\perp$, then $x = x^+$. This can be reformulated as a fixed point equation

$$x = x - \omega(K^*Kx - K^*y)$$

for $\omega > 0$. The associated RICHARDSON iteration for $x_0 \in X$ is

$$x_{n+1} = x_n + \omega(K^*y - K^*Kx_n) = (\operatorname{id} - \omega K^*K)x_n + \omega K^*y. \tag{24}$$

By the BANACH fixed point theorem $(x_n)_{n\in\mathbb{N}}$ converges to the solution of the normal equation if $\|\operatorname{id} - \omega K^*K\| < 1$ and $y \in \mathcal{R}(K)$ as

$$\|a + \omega(\cancel{K^*y} - K^*Ka) - b - \omega(\cancel{K^*y} - K^*Kb)\| \leqslant \|\operatorname{id} - \omega K^*K\|\|a - b\|.$$

If $x_0 \in \mathcal{R}(K^*) \subset (\mathcal{N}(K))^\perp$, then $x_n \to x^+$.

If $y^{(\delta)} \notin \mathcal{R}(K)$, the convergence need not hold. The idea is then to stop after a certain number of iterations (cf. chapter 5 from [15] for details).

Lemma 4.0.9 (Closed form of Richardson iteration)

For $x_0 := 0$ and $m \in \mathbb{N}$ we have

$$x_m = \omega \sum_{n=0}^{m-1} (\mathrm{id} - \omega K^* K)^n K^* y$$

Proof. We proceed by induction over m. For $m = 1$ we have $x_1 = \omega K^* y$. Assume the statement is true for $m \in \mathbb{N}$ then by (24)

$$\begin{aligned} x_{m+1} = (\mathrm{id} - \omega K^* K) x_m + \omega K^* y &= \omega \left[\sum_{n=0}^{m-1} (\mathrm{id} - \omega K^* K)^{n+1} K^* y + K^* y \right] \\ &= \omega \left[\sum_{n=1}^{m} (\mathrm{id} - \omega K^* K)^n K^* y + K^* y \right] = \omega \sum_{n=0}^{m} (\mathrm{id} - \omega K^* K)^n K^* y. \end{aligned}$$

holds. □

Theorem 4.0.4: Closed form II

For $\omega \in (0, 2\kappa^{-1})$ we have

$$x_m = f_m(K^* K) K^* y, \qquad f_m(x) := \frac{1 - (1 - \omega x)^m}{x}$$

Furthermore, $(f_m(K^* K) K^*)_{m \in \mathbb{N}}$ is a regularisation method.

Proof. By the formula for a truncated geometric series

$$f_m(x) = \omega \cdot \frac{1 - (1 - \omega x)^m}{1 - (1 - \omega x)} = \omega \sum_{n=0}^{m-1} (1 - \omega x)^n$$

holds, implying

$$f_m(K^* K) K^* y = \omega \sum_{n=0}^{m-1} (\mathrm{id} - \omega K^* K)^n K^* y = x_m$$

by lemma 4.0.9. Note that $\omega \in (0, 2\kappa^{-1})$ implies $|1 - \omega x| < 1$ for all $x \in (0, \kappa]$, ensuring that $(1 - \omega x)^m \xrightarrow{m \to \infty} 0$ hence $f_m(x) \to x^{-1}$. Thus

$$x |f_m(x)| = |1 - (1 - \omega x)^m| \leqslant 2$$

holds for all $m \in \mathbb{N}$ and $x \in (0, \kappa]$. Thus $(f_m)_{m \in \mathbb{N}}$ is a regularisation filter. By theorem 4.0.1 $(f_m(K^* K) K^*)_{m \in \mathbb{N}}$ is a regularisation method. □

Remark 4.0.10 (Decay of Least-squares functional [15])
Under those conditions on ω the least-squares functional decays:

$$
\begin{aligned}
\|Kx_{m+1} - y\|^2 &= \|Kx_m - y\|^2 + \omega^2 \|KK^*(Kx_m - y)\|^2 \\
&\quad - 2\omega\langle Kx^k - y, KK^*(Kx_m - y)\rangle \\
&= \|Kx_m - y\|^2 \\
&\quad + \omega \left(\|KK^*(Kx_m - y)\|^2 - 2\|K^*(Kx_m - y)\|^2\right) \\
&\leqslant \|Kx_m - y\|^2 + \omega \cdot \|KK^*(Kx_m - y)\|^2 \underbrace{(\omega\|K\|^2 - 2)}_{\leqslant 0 \text{ as } \omega \in (0, 2\kappa^{-1})} \\
&\leqslant \|Kx_m - y\|^2.
\end{aligned}
$$

4.1 Discretisation as Regularisation

By corollary 3.1.11 the pseudoinverse of $K \in \mathcal{K}(X,Y)$ with infinite dimensional range is discontinuous. We construct a sequence $(K_n)_{n\in\mathbb{N}} \subset \mathcal{K}(X,Y)$ with $\dim(\mathcal{R}(K_n)) < \infty$ such that the continuous pseudoinverses $(K_m^+)_{m\in\mathbb{N}}$ approximate K^+ pointwise.

We will consider two approaches

1. Least-squares projection: Let $X_n \subset X$ be finite-dimensional subspaces and define $K_n := K|_{X_n}$. — Least-squares projection:
2. Dual least-squares projection: Let $Y_n \subset Y$ be finite-dimensional subspaces and define $K_n : X \to Y_n$ by e.g. $P_{Y_n}K$. — Dual least-squares projection:

Least squares projection

Definition 4.1.1 (Least squares projection)
Let $T \in L(X,Y)$ and $(X_n \subset X)_{n\in\mathbb{N}}$ be n-dimensional subspaces with $X_1 \subset X_2 \subset \dots$ and $\bigcup_{n\in\mathbb{N}} X_n = X$.

Set $P_n := P_{X_n}$ and $T_n := TP_n \in L(X,Y)$. For $y \in Y$ set $x_n := T_n^+ y$. This procedure is called least squares projection. — least squares projection

Remark 4.1.2 As $\mathcal{R}(T_n)$ is finite-dimensional, T_n^+ is continuous. x_n is the minimum-norm solution of $T_n x = y$.

Lemma 4.1.3 (TODO)
We have $x_n \rightharpoonup x^+$ if and only if $(\|x_n\|)_{n\in\mathbb{N}}$ is bounded.

Proof. " $\implies$ ": This follows immediately from the fact that weakly convergent sequences are bounded.

" $\impliedby$ ": **TODO** □

Corollary 4.1.4
We have $x_n \to x^+$ if and only if $\limsup_{n\to\infty} \|x_n\| \leqslant \|x^+\|$.

Proof. " $\implies$ ": For all $n \in \mathbb{N}$ we have

$$\|x_n\| \overset{\triangle\neq}{\leqslant} \|x_n - x^+\| + \|x^+\| \xrightarrow{n\to\infty} \|x^+\|.$$

" $\impliedby$ ": As $(\|x_n\|)_{n\in\mathbb{N}}$ is bounded, thus by lemma 4.1.3 $x_n \rightharpoonup x^+$ holds. Thus

$$\|x^+\|^2 = \langle x^+, x^+\rangle = \lim_{n\to\infty}\langle x_n, x^+\rangle \leqslant \liminf_{n\to\infty}\|x_n\|\|x^+\|.$$

Therefore

$$\|x^+\| \leqslant \liminf_{n\to\infty}\|x_n\| \leqslant \limsup_{n\to\infty}\|x_n\| \leqslant \|x^+\|$$

holds, implying $\|x_n\| \to \|x^+\|$. As $x_n \rightharpoonup x^+$ and Hilbert spaces have the Radon-Riesz property $x_n \to x^+$. □

Theorem 4.1.1: TODO

If $y \in \mathcal{D}(T^+)$ and

$$\limsup_{n\to\infty} \|(T_n^*)^+ x_n\| = \limsup_{n\to\infty} \|(T_n^+)^* x_n\| < \infty,$$

$x_n \to x^+$ holds.

Proof. TODO □

Theorem 4.1.2: TODO

If $K \in \mathcal{K}(X, Y)$ and

$$\limsup_{n\to\infty} \|(K_n^*)^+ x_n\| = \limsup_{n\to\infty} \|(K_n^+)^* x_n\| < \infty,$$

$x^+ \in \mathcal{R}(K^*)$ holds.

Proof. TODO □

Theorem 4.1.3: TODO

If the requirements of theorem 4.1.2 are fulfilled, there exists a constant $C > 0$ with

$$\|x_n - x^+\| \leqslant C \|(\mathrm{id} - P_n) K^*\|$$

Proof. TODO □

Dual least squares projection

Definition 4.1.5 (Dual least squares projection)
Let $T \in L(X;Y)$ and $(Y_n \subset \overline{\mathcal{R}(T)})_{n\in\mathbb{N}}$ n-dimensional subspaces with $Y_1 \subset Y_2 \subset \ldots$ and $\overline{\bigcup_{n\in\mathbb{N}} Y_n} = \mathcal{N}(T^*)^\perp$.

Let $Q_n := P_{Y_n}$ and $T_n := Q_n T \in L(X, Y_n)$. This procedure is called dual least squares projection.

Remark 4.1.6 Then T_n^*, T_n^+ and $T_n^+ Q_n$ are continuous and $x_n := T_n^+ Q_n y$ is the minimum norm solution of $T_n x = Q_n y$.

We investigate under which circumstances $x_n \to x^+$ holds.

Lemma 4.1.7 (TODO)
Set $X_n := T^+ Y_n$ and $P_n := P_{X_n}$. For $y \in \mathcal{D}(T^+)$, $x_n = P_n x^+$ holds.

Proof. TODO □

Lemma 4.1.8
Let $y \in \mathcal{D}(T^+)$, Then $x_n \to x^+$.

Proof. TODO □

Appendices

A Additional proofs

TODO: We will see later that the numerical range connects algebraic information with properties of the norm.

A.1 Spectral theory

Remark A.1.1 (Properties of the numerical range / radius)

- We have $W(\mu T + \lambda I) = \mu W(T) + \lambda$ for $T \in L(\mathscr{H})$ and $\mu, \lambda \in \mathbb{C}$.
- The Toeplitz-Hausdorff theorem states that $W(T)$ is convex (and compact in finite dimensions) for any linear operator T on a inner product space H.
- This holds in finite dimensions, but in infinite?? Let E be the set of Eigenvalues of T. Then $\operatorname{conv}(E) \subset W(T)$ with equality if T is normal.
- For every $\lambda \in W(T)$ with $|\lambda| = \|T\|$ we have $\lambda \in E$. [3]
- r is a norm on $L(\mathscr{H})$ and continuous in the uniform operator topology ($T_n \to T :\iff \|T_n - T\| \to 0$) but not (weakly) pointwise. [5]
- $r(A^n) \leqslant r(A)^n$ for $n \in \mathbb{N}$. [5]
- Since every linear continuous operator T can be decomposed as $T = X + iY$, where $X := \frac{1}{2}(T + T^*)$ and $Y := \frac{i}{2}(T^* - T)$ are self-adjoint, we have

  $$\frac{1}{2}\|T\| \leqslant r(T) \leqslant \|T\| \leqslant 2r(T) \quad \text{and} \quad \rho(T) \leqslant r(T).$$

  Alternatively, this can be proven with the polarisation identity and CS, $\triangle \neq$.
- Do we have $W(T^*) = W(T)$ (true for self-adjoint obviously) and $W(S+T) \subset W(S) + W(T)$?

Lemma A.1.2
If $T \in L(\mathscr{H})$ is normal, $\operatorname{ran}(T) = \operatorname{ran}(T^*)$.

Proof. ([18]) As $\ker(T) = \ker(T^*)$ and $\overline{\operatorname{ran}}(T) = \overline{\operatorname{ran}}(T^*)$, we only need to consider $\ker(T) = \{0\}$, i.e. $\overline{\operatorname{ran}}(T) = \overline{\operatorname{ran}(T^*)} = \mathscr{H}$. **WHYYY?**

The map $U := T^*T^{-1} : \operatorname{ran}(T) \to \operatorname{ran}(T^*)$ is a isometry: for $x \in \mathcal{R}(T)$

$$\begin{aligned}\|Ux\|^2 &= \langle T^*T^{-1}x, T^*T^{-1}x\rangle = \langle T^{-1}x, TT^*T^{-1}x\rangle \\ &= \langle T^{-1}x, T^*\cancel{TT^{-1}}x\rangle = \langle \cancel{TT^{-1}}x, x\rangle = \|x\|^2\end{aligned}$$

holds. Furthermore, U is unitary:

$$\begin{aligned}(T^*T^{-1})^*T^*T^{-1} &= (T^*)^{-1}TT^*T^{-1} = (T^*)^{-1}T^*TT^{-1} = I, \\ T^*T^{-1}(T^*T^{-1})^* &= T^*T^{-1}(T^*)^{-1}T = T^*(T^*T)^{-1}T = T^*(TT^*)^{-1}T \\ &= (T^*)^{-1}T^*TT^{-1} = I.\end{aligned}$$

□

By continuity and density there exists an isometric extension $U : \mathscr{H} \to \mathscr{H}$, which is unitary because $\operatorname{ran}(U) = \operatorname{ran}(T^*) \subset \mathscr{H}$ is dense. Hence, $UT = T^*$, which gives $T^*U^* = T$ by applying the adjoint. Because U is unitary, $T^* = TU$ holds. The last two identities together imply $\mathcal{R}(T) = \mathcal{R}(T^*)$. **WHYYY?**

A.2 Unbounded Operators

Lemma A.2.1 (Canonical commutation relation)
Let X be a non-trivial normed space and $P, Q : X \to X$ linear such that $PQ - QP = \mathrm{id}_X$, Then not both P and Q can be bounded.

Proof. First we show by induction, that

$$PQ^n - Q^n P = nQ^{n-1} \tag{25}$$

for all $n \in \mathbb{N}$.

The bases cases: $n = 1$. This is fulfilled by definition.

Induction hypothesis: The equation (25) is fulfilled for some $n \in \mathbb{N}$.

Induction step $n \to n + 1$.

$$\begin{aligned} PQ^{n+1} - Q^{n+1}P &= PQQ^n - QQ^nP \\ &= PQQ^n + Q(PQ^n - Q^nP) - QPQ^n \\ &= Q(PQ^n - Q^nP) + (PQ - QP)Q^n \\ &\overset{\text{IH}}{=} Q(nQ^{n-1}) + \mathrm{id}_X Q^n = (n+1)Q^n. \end{aligned}$$

From this follows that we have $n \leqslant 2\|P\|\|Q\|$ for all $n \in \mathbb{N}$:

$$n\|Q^{n-1}\| = \|nQ^{n-1}\| = \|PQ^n - Q^nP\| \leqslant \|P\|\|Q^n\| + \|Q^n\|\|P\| = 2\|P\|\|Q^n\|.$$

Now dividing by $\|Q^{n-1}\|$ yields the claim since $\|AB\| \leqslant \|A\|\|B\|$ for all A, B linear operators and therefore

$$\frac{\sup\limits_{\|x\|=1} \|Q^n x\|}{\sup\limits_{\|x\|=1} \|Q^{n-1}x\|} \leqslant \|Q\| \cdot \frac{\sup\limits_{\|x\|=1} \|Q^{n-1}x\|}{\sup\limits_{\|x\|=1} \|Q^{n-1}x\|} = \|Q\|.$$

Therefore, either P or Q is unbounded. $\square$

A.3 Fourier transforms

Lemma A.3.1
For $f \in L^1(\mathbb{R}^n)$, $g \in L^p(\mathbb{R}^n)$, $p \in [1,\infty]$ *we have* $\|f * g\|_p \leqslant \|f\|_1 \|g\|_p$.

Proof. For $p \in (1,\infty)$ and its Hölder-conjugate $q \in (1,\infty)$

$$\begin{aligned}
|(f \circ g)(x)| &= \left| \int_{\mathbb{R}^n} f(x-y) g(y) \,\mathrm{d}y \right| \overset{\triangle\neq}{\leqslant} \int_{\mathbb{R}^n} |f(x-y)||g(y)| \,\mathrm{d}y \\
&= \int_{\mathbb{R}^n} |f(x-y)|^{\frac{1}{p}} |g(y)| |f(x-y)|^{\frac{1}{q}} \,\mathrm{d}y \\
&\overset{(H)}{\leqslant} \left(\int_{\mathbb{R}^n} |g(y)|^p |f(x-y)| \,\mathrm{d}y \right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n} |f(x-y)| \,\mathrm{d}y \right)^{\frac{1}{q}} \\
&\overset{(\star)}{=} \left(\int_{\mathbb{R}^n} |g(y)|^p |f(x-y)| \,\mathrm{d}y \right)^{\frac{1}{p}} \|f\|_1^{\frac{1}{q}}
\end{aligned}$$

holds, where $(\star)$ is the translational invariance of the Lebesgue integral, (H) is Hölder's inequality and (F) is Fubini's theorem. Thus

$$\begin{aligned}
\|f * g\|_p^p &\leqslant \int_{\mathbb{R}^n} \|f\|_1^{\frac{p}{q}} \int_{\mathbb{R}^n} |g(y)|^p |f(x-y)| \,\mathrm{d}y \,\mathrm{d}x \\
&\overset{(F)}{=} \|f\|_1^{\frac{p}{q}} \int_{\mathbb{R}^n} |g(y)|^p \left(\int_{\mathbb{R}^n} |f(x-y)| \,\mathrm{d}x \right) \mathrm{d}y \\
&\overset{(\star)}{=} \|f\|_1^{\frac{p}{q}+1} \int_{\mathbb{R}^n} |g(y)|^p \,\mathrm{d}y = \|f\|_1^p \|g\|_p^p
\end{aligned}$$

holds.

For $p = \infty$ we have

$$\begin{aligned}
\|f * g\|_\infty &= \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x-y) g(y) \,\mathrm{d}y \right| \overset{\triangle\neq}{\leqslant} \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y)| \,\mathrm{d}y \|g\|_\infty \\
&\overset{(\star)}{=} \|g\|_\infty \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} |f(y)| \,\mathrm{d}y = \|f\|_1 \|g\|_\infty .
\end{aligned}$$

$\square$

Lemma A.3.2

$$\int_0^\infty \frac{\ln(x)}{(1+x^2)^n}\,\mathrm{d}x = -\frac{(\gamma + 2\ln(2)) + \Gamma(n-1/2)\psi(n-1/2)}{4\Gamma(n)}, n > 0$$

Proof. ([20]) Define

$$I(m,n) := \int_0^\infty \frac{x^m}{(1+x^2)^n}\,\mathrm{d}x \tag{26}$$

Differentiating yields

$$\frac{d}{\mathrm{d}m} I(m,n) = \int_0^\infty \frac{x^m \ln(x)}{(1+x^2)^n}\,\mathrm{d}x \tag{27}$$

Substituting $x = \sqrt{t}$ in (26) yields

$$I(m,n) = \frac{1}{2}\int_0^\infty \frac{t^{\frac{m-1}{2}}}{(1+t)^n}\,\mathrm{d}t$$

Using the definition of the Beta function:

$$B(p,q) = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}}\,\mathrm{d}t = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

we obtain

$$I(m,n) = B\left(\frac{m+1}{2}, n - \frac{(m+1)}{2}\right)) = \frac{\Gamma(\frac{m+1}{2})\Gamma(n - \frac{(m+1)}{2})}{\Gamma(n)}.$$

Thus

$$\int_0^\infty \frac{\ln(x)}{(1+x^2)^n}\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}m} I(m,n) = \frac{1}{2\Gamma(n)}\left(\Gamma'(m+\frac{1}{2}) - \Gamma'(n - \frac{m+1}{2}\right)$$

Use the definition of the Digamma function, $\Gamma'(x) = \psi(x)\Gamma(x)$, yields for $m = 0$

$$\int_0^\infty \frac{\ln(x)}{(1+x^2)^n}\,\mathrm{d}x = \frac{\Gamma(1/2)\psi(1/2) - \Gamma(n-1/2)\psi(n-1/2)}{4\Gamma(n)}$$

Using $\psi(1/2) = -\gamma - 2\ln(2)$ and $\Gamma(1/2) = \sqrt{\pi}$ yields the result. $\square$

References

[1] Functional Calculus on Wikipedia.

[2] Numerical Range: (in) a Matrix Nutshell, Part 1 by Panayiotis J. Psarrakos and Michael J. Tsatsomeros at Washington State University

[3] German Wikipedia page for the numerical range in context of Hilbert spaces.

[4] Answer by @DisintegratingByParts on Math.SE.

[5] "An elementary proof of the power inequality for the numerical radius" by Carl Pearcy in 1966

[6] Answer by @s.harp on Math.SE.

[7] Answer by @mechanodroid on Math.SE.

[8] Lecture Notes for an Analysis III course at ANU by Scott Morrison.

[9] Functional Analysis II, Winter Semester 11/12 at LMU Munich, Problem set 13.

[10] `techxplore.com`.

[11] Anaylsis Now, Pedersen, 2nd Edition, Springer.

[12] FOURIER ANALYSIS AND THE UNCERTAINTY PRINCIPLE by WENFEI DU.

[13] this question on Math.SE

[14] Daubechies, Defrise, DeMol; 2008.

[15] LMU Inverse Problems Lecture Notes by Martin Burger.

[16] Wikipedia page for distributions.

[17] Wikipedia page for convolution.

[18] This answer on Math.SE

[19] Wikipedia page for the Fourier transform

[20] This answer on the the art of problem solving website.

[21] Folland - A Course in Harmonic Analysis

Index

Symbols

LEBESGUE-point 45
TIKHONOV regularisation 86

A

absolute value 80
approximate point spectrum .. 1

B

bounded from below 3

C

character 51
closable 34
closed 34
continuous functional calculus 13

D

densely defined 34
dual group 51
Dual least-squares projection: 90

E

exponentially ill-conditioned . 79

G

graph 34

H

HAAR measure 50

I

ill-posed 70
involutive 13

L

least squares projection 90
least-square-solution 71
Least-squares projection: 90

M

minimal norm solution 71
moderately ill-conditioned ... 79
mollification 59

N

numerical range 3

P

parameter choice rule 82

R

RADON measure 20
regular distribution 59
regularisation method 82
regularisation of T^+ 81
regularising filter 84
resolvent identity 38
resolvent mapping 38
resolvent set 38

S

SCHWARTZ space 54
spectral measure 29
spectrum
- continuous 1
- point 1
- residual 1

strongly ill-conditioned 79

T

topological group 50
topological vector space 55

W

weak derivative 67
well-posed 70