Functional Analysis

Karlsruher Institut für Technologie (KIT)

Prof. Dr. Tobias Lamm

Winter term 2016/17

Contents

1 Introduction

We are interested in solutions of linear equations $Lu = f$, where $L\colon X \to Y$ is an operator between vector spaces over $\mathbb{R}$ or $\mathbb{C}$. In the following examples we illustrate how various results which were shown to be true on finite-dimensional vector spaces in Linear Algebra are no longer true on infinite-dimensional spaces.

Example: Let $L = \Delta$; $X = C^2(\mathbb{R}^n)$, $Y = C^0(\mathbb{R}^n)$. Then

$$L\colon X \to Y \text{is linear:}$$
$$L(\lambda x + \mu y) = \lambda Lx + \mu Ly$$

In general we are interested in the case $\dim X = \dim Y = \infty$.

Example: Let $X = \{x = (x^1, x^2, ...)\colon x^i \in \mathbb{R}\}$ be the sequence space and let $A\colon X \to X$

$$A((x^1, x^2, ...)) = (0, x^1, x^2, ...)$$

be the shift-map. Then A is injective but not surjective.

Now, let $B\colon X \to X$ be defined by

$$B((x^1, x^2, ...)) = (x^2, x^3, ...).$$

Then B is surjective but not injective.

Example: Let $X = C^0([-\pi, \pi])$ and consider the map $A\colon X \to X$,

$$(Au)(t) = \sin(t)u(t)$$

for all $t \in [-\pi, \pi]$. Assume that A has an eigenvalue $\lambda \in \mathbb{R}$ (or $\lambda \in \mathbb{C}$). Hence there exists $u \in C^0([-\pi, \pi])$ with $u \neq 0$, so that

$$sin(t)u(t) = \lambda u(t)\ \forall t \in [-\pi, \pi].$$

But then $\{t\colon u(t) \neq 0\} \subset \{t\colon \sin t = \lambda\}$ which gives the contradiction $u \equiv 0$.

Central example: The Dirichlet principle for elliptic boundary value problems. Let

$\Omega \subset \mathbb{R}^n$ be open, connected, bounded and assume that $\partial\Omega$ is smooth.

Given $f \in C^\infty(\overline{\Omega})$ we want to find a solution $u \in C_0^\infty(\overline{\Omega})$ of the elliptic boundary value problem

$$\begin{cases} -\Delta u = f \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

This can be done by using a variational approach and for this one considers the functional

$$\mathcal{F}(v) := \frac{1}{2}\int_\Omega |Dv|^2\, d\lambda - \int_\Omega fv\, d\lambda,$$

where $\lambda = n$-dimensional Lebesgue measure.

Lemma 1.1. *Let $u \in C_0^\infty(\overline{\Omega})$ satisfy*

$$\mathcal{F}(u) \leq \mathcal{F}(v)\ \forall v \in C_0^\infty(\overline{\Omega}),$$

then $-\Delta u = f$ in Ω.

Proof. Let $\eta \in C_c^\infty(\overline{\Omega})$, i.e. $\mathrm{spt}_\eta \subset\subset \Omega$. We estimate for every $\varepsilon \in \mathbb{R}$

$$\begin{aligned} \mathcal{F}(u) &\leq \mathcal{F}(u + \epsilon\eta) \\ &= \frac{1}{2}\int |Du|^2 d\lambda + \varepsilon \int \langle Du, D\eta\rangle d\lambda + \frac{\varepsilon^2}{2}\int |D\eta|^2 d\lambda - \int fu d\lambda - \varepsilon \int f\eta d\lambda \\ &= \mathcal{F}(u) + \varepsilon\left(\int \langle Du, D\eta\rangle d\lambda - \int f\eta d\lambda\right) + \frac{\epsilon^2}{2}\int |D\eta|^2 d\lambda. \end{aligned}$$

Since this has to be true for every ε, we conclude that $\forall \eta \in C_c^\infty(\Omega)$ we have

$$0 = \int (\langle Du, D\eta\rangle - f\eta) d\lambda = -\int (\Delta u + f)\eta d\lambda$$

and the fundamental theorem of the calculus of variations then implies that $-\Delta u = f$in Ω. □

I Minimize $\mathcal{F}(v)$ in $C_0^\infty(\overline{\Omega})$.
Formally this is similar to the problem

II Minimize $F(x) = \frac{1}{2}|x|^2 - \langle a, x\rangle$ in $\mathbb{R}^n$, $a, x \in \mathbb{R}^n$.

Theorem 1.2. *On a finite dimensional vector space V all norms are equivalent. For $\|\cdot\|_1$, $\|\cdot\|_2$ on V there exists $0 < m < M < \infty$, such that*

$$m\|v\|_1 \leq \|v\|_2 \leq M\|v\|_1\ \forall v \in V.$$

Proof. 1) Without loss of generality we can assume that $V = \mathbb{R}^n$. Otherwise choose basis $\{v_1, ..., v_n\}$ of V and define

$$f\colon \mathbb{R}^n \to V,\ f(x) = \sum_{i=1}^{n} x^i v_i$$

Then we look at the new norms

$$\|x\|_{1,2} = \|f(x)\|_{1,2}$$

on $\mathbb{R}^n$.

2) Next we can always assume that $\|\cdot\|_2 = \|\cdot\|_\infty$ and hence it remains to show:

$$m\|\cdot\|_\infty \leq \|\cdot\| \leq M\|\cdot\|_\infty.$$

For this we note that

$$\begin{aligned}\|x\| = \|\sum_{i=1}^{n} x^i e_i\| &\leq \sum_{i=1}^{n} |x^i|\|e_i\| \\ &\leq \|x\|_\infty \sum_{i=1}^{n} \|e_i\| =: M\|x\|_\infty\end{aligned}$$

Hence $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous with constant M with respect to $\|\cdot\|_\infty$:

$$|\|x\| - \|y\|| \leq \|x - y\| \leq M\|x - y\|_\infty.$$

We also know that $\{x \in \mathbb{R}^n\colon \|x\|_\infty = 1\}$ is compact and therefore

$$\|\cdot\|\colon \{x \in \mathbb{R}^n\colon \|x\|_\infty = 1\} \to \mathbb{R}$$

attains its infimum. Let $m := \inf\{\|x\|\colon \|x\|_\infty = 1\} \Rightarrow \exists x_0 \in \mathbb{R}^n$ such that $\|x_0\|_\infty = 1$, $\|x_0\| = m \Rightarrow m > 0$
$\Rightarrow \forall x \in \mathbb{R}^n$:

$$\|x\| = \|x\|_\infty \|\frac{x}{\|x\|_\infty}\| \geq m\|x\|_\infty.$$

□

<u>Solution of II:</u> Choose a minimising sequence $x_j \in \mathbb{R}^n$:

$$\begin{aligned}&F(x_j) \to \inf\{F(x)\colon x \in \mathbb{R}^n\} \\ \Rightarrow \frac{1}{2}|x_j - a|^2 &= \frac{1}{2}(|x_j|^2 - 2\langle a, x_j\rangle + |a|^2) \\ &= F(x_j) + \frac{1}{2}|a|^2 \leq \text{const.}\end{aligned}$$

$\Rightarrow$ (x_j) is bounded. Bolzano-Weierstrass now implies that up to a subsequence $x_j \to x_0 \in \mathbb{R}^n$.

$$\Rightarrow F(x_0) = \lim_{j\to\infty} F(x_j) = \inf_{x\in\mathbb{R}^n} F(x)$$

and hence x_0 is a minimiser.

Problems when copying this approach for I

1. On $C^0(I)$, $I = [0,1]$ not all norms are equivalent: Take

$$\|u\|_{L^2(I)} := \left(\int_I u^2 d\lambda\right)^{1/2}$$
$$\|u\|_{C^0(I)} := \max\{|u|(x) \colon x \in I\}.$$

and a sequence

$$u_n(x) = \begin{cases} 1 - nx, \ 0 \le x \le 1/n \\ 0, \ \text{otherwise.} \end{cases}$$

Then $\|u_n\|_{C^0} = 1$ for all $n \in \mathbb{N}$ but $\|u_n\|_{L^2} \le 1/\sqrt{n} \to 0$ as $n \to \infty$ and so the norms are not equivalent.

$\rightsquigarrow$ minimising $\mathcal{F}$ in $C_0^\infty(\overline{\Omega})$ will only give a bound for the quantity

$$\|u_j\|_{L^2(\Omega)} + \|Du_j\|_{L^2(\Omega)} \le C$$

and this will eventually force us to study Sobolev spaces.

2. The Bolzano-Weierstrass theorem is not true in infinite dimensional vector spaces. For this we consider

$$l^p(\mathbb{R}) = \{x = (x^i)_{i\in\mathbb{N}} \colon x^i \in \mathbb{R}, \|x\|_p = \left(\sum_{i=1}^{\infty} (x^i)^p\right)^{1/p} < \infty\}$$
$$l^\infty(\mathbb{R}) = \{\text{same but with } \|x\|_\infty := \sup |x^i| < \infty\}.$$

Now look at sequence $x_k \in l^p(\mathbb{R})$, $1 \le p \le \infty$ defined by $x_k = (0, ..., 0, 1, 0, ...)$. We have that $x_k^i = \delta_{ik}$, $\|x_k\|_{l^p} = 1$ but this sequence has no converging subsequence, since

$$\|x_j - x_k\|_p = 2^{1/p} \ge 1, \ j \ne k.$$

2 Metric and normed spaces

Definition 2.1. *Let X be a set. A map $d\colon X \times X \to [0, \infty)$ is called a **metric** (and (X, d) a metric space) if for all $x, y, z \in X$ there holds*

1. $d(x, y) = 0 \Leftrightarrow x = y$.

2. $d(x, y) = d(y, x)$.

3. $d(x, z) \leq d(x, y) + d(y, z)$.

*The metric space (X, d) is called **complete**, if every Cauchy sequence converges in X. ((x_k) is a Cauchy sequence, $\Leftrightarrow$ $d(x_k, x_l) < \varepsilon$ for all k, l large enough).*

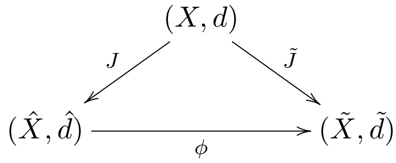
In the next theorem we show that for every metric space there exists a completion.

Theorem 2.2. *For every metric space (X, d) there exists a complete metric space $(\hat{X}, \hat{d})$ and an isometric map*

$$J\colon (X, d) \to (\hat{X}, \hat{d})$$

such that $J(X)$ is dense in $\hat{X}$.

Uniqueness: If $(\tilde{X}, \tilde{d})$ is a complete metric space and $J\colon X \to \tilde{X}$ is isometric such that $\tilde{J}(X)$ is dense in $\tilde{X}$, then there exists a unique isometry $\phi\colon \hat{X} \to \tilde{X}$ with $\tilde{J} = \phi \circ J$.



Proof. We start with the existence of $(\hat{X}, \hat{d})$. For this we define $\mathrm{CS}(X) = \{x = (x^i)\colon x \text{ is Cauchy sequence in } X\}$ and we introduce the equivalence relation $x \sim y$ $\Leftrightarrow$ $\lim\limits_{i \to \infty} d(x^i, y^i) = 0$. We now define $\hat{X} := \mathrm{CS}(X)/\sim$.

Claim 1: $\lim\limits_{i \to \infty} d(x^i, y^i)$ exists

For i, j large enough we have

$$d(x^i, x^j), d(y^i, y^j) < \frac{\varepsilon}{2}$$

and hence

$$\begin{aligned} d(x^i, y^i) &\leq d(x^i, x^j) + d(x^j, y^j) + d(y^i, y^j) \\ &< \varepsilon + d(x^j, y^j) \end{aligned}$$

which implies that $|d(x^i, y^i) - d(x^j, y^j)| < \varepsilon$ and hence $(d(x^i, y^i))$ is a Cauchy sequence in $\mathbb{R}$. Additionally we remark that for a Cauchy sequence (x^i) and a subsequence (x^{i_k}) we have that $(x^i) \sim (x^{i_k})$ since $d(x^k, x^{i_k}) \to 0$.

Claim 2: The function $\hat{d}([x], [y]) := \lim_{i \to \infty} d(x^i, y^i)$ where $(x^i) \in [x]$, $(y^i) \in [y]$ is a norm on $\hat{X}$.

We note that $\hat{d}$ is well-defined since for $(x^i), (\tilde{x}^i) \in [x]$ and $(y^i), (\tilde{y}^i) \in [y]$ we have

$$d(\tilde{x}^i, \tilde{y}^i) \leq d(x^i, \tilde{x}^i) + d(x^i, y^i) + d(y^i, \tilde{y}^i)$$

and hence $|d(\tilde{x}^i, \tilde{y}^i) - d(x^i, y^i)| \to 0$ as $i \to \infty$.

Now for $\hat{d}([x], [y]) = 0$ we conclude $\lim_{i \to \infty} d(x^i, y^i) = 0$ and therefore $x \sim y$ which is equivalent to $[x] = [y]$. The other two properties of a metric are easy to verify.

Claim 3: The map $J \colon X \to \hat{X}$, $J(x) = [(x, x, x, ...)]$ is isometric and $J(X)$ is dense in $\hat{X}$.

By definition we have $\hat{d}(J(x), J(y)) = d(x, y)$ and hence J is isometric. For $[x] \in \hat{X}$ we choose $(x^i) \in [x]$ and then we conclude for k large enough

$$\hat{d}([x], J(x^k)) = \lim_{i \to \infty} d(x^i, x^k) < \varepsilon.$$

Therefore $J(X)$ is dense in $\hat{X}$.

Claim 4: The metric space $(\hat{X}, \hat{d})$ is complete.

We let $[x_k]$ be a Cauchy sequence in $\hat{X}$, i.e. $\lim_{i \to \infty} d(x_k^i, x_l^i) < \varepsilon$ for all $k, l \geq k(\varepsilon)$. Without loss of generality we can assume that

$$d(x_k^i, x_k^j) < \frac{1}{k} \; \forall i, j \in \mathbb{N}$$

Now we look at the diagonal sequence $(y^k) = (x_k^k)$. For $k, l \geq k(\varepsilon)$ we have

$$\begin{aligned} d(y^k, y^l) &\leq d(x_k^k, x_k^i) + d(x_k^i, x_l^i) + d(x_l^i, x_l^l) \\ &< \frac{1}{k} + \varepsilon + \frac{1}{l} \end{aligned}$$

and hence (y^k) is a Cauchy sequence in X. We have to show that $[x_k] \to [y]$. In order to do this we note that

$$\begin{aligned} d(x_k^i, y^i) &\leq d(x_k^i, x_k^j) + d(x_k^j, x_i^j) + d(x_i^j, x_i^i) \\ &< \frac{1}{k} + d(x_k^j, x_i^j) + \frac{1}{i} \end{aligned}$$

and for $j \to \infty$ this implies

$$d(x_k^i, y^i) < \frac{1}{k} + \varepsilon + \frac{1}{i}.$$

Therefore we get $\hat{d}([x_k], [y]) = \lim_{i \to \infty} d(x_j^i, y^i) \leq \frac{1}{k} + \varepsilon$ which shows that $[x_k] \to [y]$.

In order to show the uniqueness statement of the theorem we let $(\hat{X}, \hat{d})$ and $(\tilde{X}, \tilde{d})$ be complete spaces with two isometric (and hence injective) maps $\hat{J}\colon X \to \hat{X}$ and $\tilde{J}\colon X \to \tilde{X}$. Therefore the map

$$\phi\colon \hat{J}(X) \to \tilde{X},\ \phi(\hat{J}(x)) = \tilde{J}(x)$$

is well defined and we have

$$\begin{aligned} \tilde{d}(\phi(\hat{J}(x)), \phi(\hat{J}(y))) &= \tilde{d}(\tilde{J}(x), \tilde{J}(y) \\ &= d(x, y) \\ &= \hat{d}(\hat{J}(x), \hat{J}(y)). \end{aligned}$$

This shows that $\phi\colon \hat{J}(X) \to \tilde{X}$ is isometric and therefore $\phi|_{\hat{J}(X)}$ is uniformly continuous. Since $\hat{J}(X)$ is dense in $\hat{X}$ and $\tilde{X}$ is complete, we can use exercise 1 from the exercise sheet 1 in order to conclude that there exists a unique continuous extension (still denoted by ϕ) $\phi\colon \hat{X} \to \tilde{X}$ which is isometric. Moreover $\phi(\hat{X}) \supset \phi(\hat{J}(X)) = \tilde{J}(X)$ is dense in $\tilde{X}$ by definition of $\tilde{J}$. And $\phi(\hat{X})$ is complete with respect to $\tilde{d}$.

Now exercise 4 from sheet 1 says that if we have a complete space (X, d) and if $Y \subset X$ is also complete, then Y is closed subspace. Hence $\phi(\hat{X})$ is closed in $\tilde{X}$. Altogether this shows that $\phi(\hat{X}) = \tilde{X}$, and hence $\phi\colon \hat{X} \to \tilde{X}$ is an isometry. □

Definition 2.3. *Let X be an $\mathbb{R}$- or $\mathbb{C}$ vector space. A function $\|\cdot\|\colon X \to [0, \infty)$ is called a* ***norm****, if the following three properties are satisfied:*

1. $\|x\| = 0 \Leftrightarrow x = 0$ *(definite).*

2. $\|\lambda x\| = |\lambda|\|x\|$ *for all* $\lambda \in \mathbb{K}$, $x \in X$.

3. $\|x + y\| \le \|x\| + \|y\|$ *for all* $x, y \in X$.

On X *we have the induced metric* $d(x, y) = \|x - y\|$. *Is* (X, d) *complete, then* $(X, \|\cdot\|)$ *is called a* ***Banach space****.*

Definition 2.4. *Let* $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ *be normed spaces. A linear map* $A\colon X \to Y$ *is called* ***bounded****, if*

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Ax\|_Y < \infty.$$

$\|A\|$ *is called the* ***operator norm*** *of* A*. The space*

$$L(X, Y) := \{A\colon X \to Y | A \text{ linear}, \|A\| < \infty\}$$

is a normed space with the operator norm.

Lemma 2.5. *For a linear map* $A\colon (X, \|\cdot\|) \to (Y, \|\cdot\|)$ *the following statements are equivalent.*

1. A *is bounded.*

2. A *is continuous.*

3. A *is continuous in* 0.

Proof. 1. $\Rightarrow$ 2. For all $x, y \in X$ we estimate

$$\|Ax - Ay\| = \|A(x - y)\| \le \|A\|\|x - y\|$$

and therefore A is even Lipschitz continuous.

3. $\Rightarrow$ 1. For $\varepsilon = 1$ $\exists \delta > 0$ such that

$$\|Ax\| = \|Ax - 0\| \le 1 \; \forall \|x\| \le \delta$$

Hence we conclude for $x \neq 0$ that

$$\|Ax\| = \frac{\|x\|}{\delta} \|A\left(\delta \frac{x}{\|x\|}\right)\| \le \frac{\|x\|}{\delta}$$

and therefore $\|A\| \le \frac{1}{\delta}$. □

Theorem 2.6. *Let $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ be normed spaces and assume that Y is a Banach space. Then $L(X,Y)$ with the operator norm is a Banach space.*

Proof. Let (A_i) be a Cauchy sequence in $L(X,Y)$. This implies that for every $x \in X$ we have

$$\|A_i x - A_j x\| \leq \|A_i - A_j\| \|x\| < \varepsilon \|x\| \ \forall i, j \in I(\varepsilon)$$

and hence the sequence $(A_i x)$ is a Cauchy sequence in Y for all $x \in X$. Since Y is a Banach space, this shows that the limit

$$Ax := \lim_{i \to \infty} A_i x$$

exists for all $x \in X$. It follows right away that A is linear. Next we estimate

$$|\|A_i\| - \|A_j\|| \leq \|A_i - A_j\| < \varepsilon \ \forall i, j \geq I(\varepsilon)$$

and this shows that $(\|A_i\|)$ is a Cauchy sequence in $\mathbb{R}$. Again we conclude that the limit $\Lambda := \lim_{i \to \infty} \|A_i\|$ exists and we get

$$\|Ax\| = \lim_{i \to \infty} \|A_i x\| \leq \lim_{i \to \infty} \|A_i\| \|x\| = \Lambda \|x\|$$

and therefore $\|A\| \leq \Lambda$ which implies that $A \in L(X,Y)$. Finally we observe that for every $x \in X$

$$\|A_i x - Ax\| = \lim_{j \to \infty} \|A_i x - A_j x\| \leq \varepsilon \|x\| \ \forall i \geq I(\varepsilon)$$

and hence

$$\|A_i - A\| < \varepsilon \ \forall i \geq I(\varepsilon),$$

which shows that $A_i \to A$ in the operator norm. □

Definition 2.7. *Let X be a normed space. The Banach space $X' := L(X, \mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) with the operator norm is called the* ***dual space*** *of X.*

Theorem 2.8. *Let X be a Banach space and let $V \subset X$ be a closed subspace. Then X/V with the norm*

$$\|[x]\| := \inf_{v \in V} \|x + v\|$$

is also a Banach space.

Proof. 1. We first show that $\|\cdot\|$ is indeed a norm on X/V.

(i) $\|[x]\| \geq 0$ and $\|[x]\| = 0$ iff $[x] = 0$

The non-negativity is obvious and for the second claim we note that if $\|[x]\| = 0$ then there exists a sequence $v_k \in V$ so that $x + v_k \to 0$ and hence $v_k \to -x \in V$ since V is closed. hence $[x] = 0$.

(ii) For $\lambda \in \mathbb{K}$ we have $\|\lambda[x]\| = |\lambda|\|[x]\|$.

Since this is obvious for $\lambda = 0$ we assume that $\lambda \in \mathbb{K}\backslash\{0\}$ and we calculate

$$\begin{aligned}\|\lambda[x]\| &= \|[\lambda x]\| = \inf_{v\in V} \|\lambda x + v\| \\ &= |\lambda| \inf_{v\in V} \left\| x + \frac{v}{\lambda} \right\| = |\lambda|\|[x]\|.\end{aligned}$$

(iii) Triangle inequality

For this we estimate

$$\begin{aligned}\|[x_1] + [x_2]\| &= \|[x_1 + x_2]\| \\ &= \inf_{v_1,v_2\in V} \|x_1 + x_2 + v_1 + v_2\| \\ &\le \inf_{v_1\in V} \|x_1 + v_1\| + \inf_{v_2\in V} \|x_2 + v_2\| \\ &= \|[x_1]\| + \|[x_2]\|.\end{aligned}$$

2. Next we show that X/V with this norm is complete.

In order to do this we note that

$$\|[y] - [x]\| = \inf_{v\in V} \|y - x + v\|$$

implies that if $\|[y] - [x]\| < \varepsilon$, then there exists $\tilde{y} \in [y]$ so that

$$\|\tilde{y} - x\| < \varepsilon.$$

Now let $[x_i]$ be a Cauchy sequence in X/V. Without loss of generality we assume that

$$\|[x_i + 1] - [x_i]\| < 2^{-i} \; \forall i \in \mathbb{N}\,.$$

We choose inductively $\tilde{x}_i \in [x_i]$ so that

$$\|\tilde{x}_{i+1} - \tilde{x}_i\| < 2^{-i}.$$

This can be done as follows: Set $\tilde{x}_1 = x_1$ and assume that $\tilde{x}_i$ has already been found for some $i \in \mathbb{N}$. Then we know that

$$\|[x_{i+1}] - [\tilde{x}_i]\| = \|[x_{i+1}] - [x_i]\| < 2^{-i}$$

and hence it follows from the above remark that there exists $\tilde{x}_{i+1} \in [x_{i+1}]$ so that

$$\|\tilde{x}_{i+1} - \tilde{x}_i\| < 2^{-i}.$$

By this construction $(\tilde{x}_i)$ is a Cauchy sequence in X and therefore $\tilde{x} := \lim_{i\to\infty} \tilde{x}_i$ exists. Moreover

$$\|[x_i] - [\tilde{x}]\| = \|[\tilde{x}_i] - [\tilde{x}]\| = \|[\tilde{x}_i - \tilde{x}]\| \leq \|\tilde{x}_i - \tilde{x}\| \to 0$$

as $i \to \infty$ and this finishes the proof.

□

Definition 2.9. *Let X be a $\mathbb{K}$-Vector space. The function*

$$\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$$

*is called a **scalar product**, if the following three properties are satisfied:*

1. *For all $\lambda, \mu \in \mathbb{K}$ we have*

$$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, u \rangle$$
$$\langle x, \lambda y + \mu z \rangle = \bar{\lambda} \langle x, y \rangle + \bar{\mu} \langle x, z \rangle.$$

2. $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

3. $\langle x, x \rangle \geq 0$ *for all* $x \in X$ *and* $\langle x, x \rangle = 0$ *iff* $x = 0$.

For every scalar product one can define the so called Euclidean norm by $\|x\| := \sqrt{\langle x, x \rangle}$.

Theorem 2.10. *Let $(X, \langle \cdot, \cdot \rangle)$ be a scalar product space. Then we have for all $x, y \in X$:*

1. $|\langle x, y \rangle| \leq \|x\| \|y\|$.

2. $\|x + y\| \leq \|x\| + \|y\|$.

Proof. 1. Without loss of generality we can assume that $\|x\| = 1 = \|y\|$ and by replacing x by $xe^{i\theta}$ for some appropriate θ, we can also assume that $\langle x, y \rangle \geq 0$.

Under these conditions we calculate

$$\begin{aligned}\|x\|\|y\| - \langle x, y\rangle &= \frac{1}{2}(\|x\|^2 + \|y\|^2 - 2\langle x, y\rangle) \\ &= \frac{1}{2}\|x-y\|^2 \geq 0.\end{aligned}$$

Statement 2 follows from the first one as follows:

$$\|x+y\|^2 = \|x\|^2 + 2Re(\langle x, y\rangle) + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

□

Remark 1. *If $\|\cdot\|$ is the euclidean norm of a scalar product, then*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)\ \forall x, y \in X.$$

Vice versa a norm satisfying this equation defines a scalar product by

- $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$, $\mathbb{K} = \mathbb{R}$.
- $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) + \frac{i}{4}(\|x+iy\|^2 - \|x-iy\|^2)$, $\mathbb{K} = \mathbb{C}$.

Definition 2.11. *A **Hilbert space** is a scalar product space $(X, \langle\cdot,\cdot\rangle)$ which is complete with respect to the induced norm $\|x\| = \sqrt{\langle x, x\rangle}$.*

Example: For $\Omega \subset \mathbb{R}^n$ the space $X = L^2(\Omega)$ with the scalar product

$$\langle f, g\rangle = \int_\Omega fg d\lambda^n$$

and the Lebesgue measure λ^n is complete by the Fischer-Riesz theorem.

3 Compactness (in metric spaces)

Definition 3.1. *A set $K \subset X$ of a metric space (X, d) is called* ***compact****, if the following holds:*

Every family U_λ, $\lambda \in \Lambda$, of open sets with $K \subset \bigcup_{\lambda\in\Lambda} U_\lambda$ has a ***finite*** *subfamily U_{λ_i}, $1 \leq i \leq N$ with $K \subset \bigcup_{i=1}^{N} U_{\lambda_i}$.*

Theorem 3.2. *Let (X, d) be a metric space. For $K \subset X$ the following are equivalent:*

1. *K is compact.*

2. *K is complete and K is* ***precompact*** *(i.e. for all $\varrho > 0$ K is covered by finitely many balls $B_\varrho(x_i)$, $1 \leq i \leq N$, $x_i \in K$).*

3. *K is* ***sequentially compact*** *(i.e every sequence $(x_i) \subset K$ has a convergent subsequence).*

Proof. 3. $\Rightarrow$ 2.: The completeness of K is clear since every Cauchy sequence which has a converging subsequence must converge. Now assume that K is not precompact. Then we can choose $x_1 \in K$ and inductively $x_k \in K$, $k \in \mathbb{N}$, such that $x_k \notin B_\varrho(x_1) \cup \cdots \cup B_\varrho(x_{k-1})$. The sequence $(x_k) \subset K$ has no converging subsequence, since $d(x_k, x_l) \geq \varrho$ for all $k \neq l$ and this contradicts the assumption in 3.

2. $\Rightarrow$ 1. We assume that $K \subset \bigcup_{\lambda\in\Lambda} U_\lambda$ but K is not covered by finitely many of the sets U_λ. Then we construct balls $B_k = B_{2^{-k}}(x_k)$, $x_k \in K$, $k \in \mathbb{N}$ so that $B_k \cap B_{k-1} \neq \emptyset$ and $B_k \cap K$ is not covered by finitely many of the sets U_λ. The construction is done as follows: First we let $B_0 := X$ and then we assume that $B_1, ..., B_{k-1}$ have already been found. We cover $B_{k-1} \cap K$ by finitely many balls

$$B_{2^{-k}}(y_i),\ y_i \in K \text{ so that} B_{2^{-k}}(y_i) \cap B_{k-1} \neq \emptyset.$$

This is possible since K is assumed to be precompact. Now at least one of the balls $B_{2^{-k}}(y_i)$ is not covered by finitely many of the sets U_λ and we let B_k be one of these balls. The sequence of centres (x_k) is a Cauchy sequence in K since for $l > k$ large enough

$$d(x_k, x_l) \leq d(x_k, x_{k+1}) + \cdots + d(x_{l-1}, x_l) \leq \sum_{i=k}^{l} 2^{-i} < \varepsilon.$$

Since K is assumed to be complete the limit $x = \lim x_k$ exists. Since $K \subset \bigcup_{\lambda\in\Lambda} U_\lambda$ there has to exist a $\lambda_0 \in \Lambda$ so that $x \in U_{\lambda_0}$ and therefore $B_k \subset U_{\lambda_0}$ for k large enough, which is a contradiction to the fact that B_k is not covered by finitely many of the sets U_λ.

1. $\Rightarrow$ 3.: We assume that there exists a sequence $(x_k) \subset K$ which has no accumulation point in K. Hence, for every $x \in K$ there exists a radius $\varrho_x > 0$ so that $x_k \in B_{\varrho_x}(x)$ for at most finitely many $k \in \mathbb{N}$. Now $K \subset \bigcup_{x\in K} B_{\rho_x}(x)$ and since K is compact this implies that $K \subset \bigcup_{i=1}^{N} B_{\rho_{x_i}}(x_i)$ for some $N \in \mathbb{N}$ and $x_i \in K$, $1 \leq i \leq N$. Hence $x_k \in K$ for at most finitely many k which is a contradiction. □

Lemma 3.3. *Let $(X, \|\cdot\|)$ be a normed space and let $V \subsetneq X$ be a closed subspace. For all $\theta < 1$ there exists $x_\theta \in X$ such that $\|x_\theta\| = 1$ and $\operatorname{dist}(x_\theta, V) \geq \theta$.*

Proof. By our assumptions there exists an element $y \in X\backslash V$ with $\operatorname{dist}(y, V) > 0$. For every $\theta < 1$ there exists an element $v_\theta \in V$ with $\|y - v_\theta\| \leq \frac{1}{\theta}\operatorname{dist}(y, V)$. We define $x_\theta := \frac{y-v_\theta}{\|y-v_\theta\|}$ and we note that for every $v \in V$

$$\|x_\theta - v\| = \|y - v_\theta\|^{-1}\|y - (v_\theta - \|y - v_\theta\|v)\| \geq \|y - v_\theta\|^{-1}\operatorname{dist}(y, V) \geq \theta,$$

where we used that $v_\theta - \|y - v_\theta\|v \in V$. □

Theorem 3.4. *Let $(X, \|\cdot\|)$ be a normed space. Then $\overline{B_1(0)} \subset X$ is compact if and only if $\dim X < \infty$.*

Proof. $\Leftarrow$ This follows immediately from Theorem 1.2, the Bolzano-Weierstrass theorem and Theorem 2.2.

$\Rightarrow$ We assume that $\dim X = \infty$. Then we can inductively choose a sequence $x_k \in X$, $k \in \mathbb{N}$, with $\|x_k\| = 1$ and $\operatorname{dist}(x_k, \operatorname{Span}\{x_1, \ldots, x_{k-1}\}) \geq \frac{1}{2}$. For this construction we use Lemma 3.3 and the fact that finite-dimensional subspaces of a normed space are always closed.

The sequence (x_k) has no converging subsequence since by construction $\|x_k - x_j\| \geq \frac{1}{2}$ for all $k \neq j$ and hence it follows from Theorem 3.2 that $\overline{B_1(0)}$ is not compact. □

Definition 3.5. *Let X, Y be metric spaces. We define the spaces of bounded respec-*

tively continuous functions between X and Y by

$$B(X,Y) := \{f \colon X \to Y \colon f(X) \text{ is bounded}\}$$
$$C^0(X,Y) := \{f \colon X \to Y \colon f \text{ is continuous}\}.$$

Theorem 3.6. *Let X, Y be metric spaces and assume that Y is complete. Then we have that*

1. $d_B(f,g) := \sup\{d(f(x), g(x)) \colon x \in X\}$ *is a complete metric on* $B(X,Y)$.

2. $(C^0 \cap B(X,Y))$ *is a closed subset of* $(B(X,Y), d_B)$ *and hence complete as well.*

Proof. 1. It is clear that d_B is a metric. Now let $f_k \in B(X,Y)$ be a Cauchy sequence. It follows that

$$d(f_k(x), f_l(x)) < \varepsilon$$

for all $x \in X$ and $k, l \geq k(\varepsilon)$. In particular the sequence $(f_k(x))$ is a Cauchy sequence in Y for all $x \in X$. Since Y is complete the limit

$$f(x) := \lim_{h \to \infty} f_h(x)$$

exists and if we let $l \to \infty$ in the above estimate, we conclude for all $x \in X$ and all $k \geq k(\varepsilon)$

$$d(f_k(x), f(x) \leq \varepsilon.$$

This implies that

$$d_B(f_k, f) \to 0$$

as $k \to \infty$. Now it remains to show that $f \in B(X,Y)$ but this follows from the estimate

$$d(f(x), f(x_0)) \leq d(f(x), f_k(x)) + d(f_k(x), f_k(x_0)) + d(f_k(x_0), f(x_0)) < \infty,$$

where x_0 is a fixed point in X and k is chosen large enough.

2. Now we let $f_k \in (C^0 \cap B)(X,Y)$ be a Cauchy sequence. As in part 1. we show that f_k converges to some $f \in B(X,Y)$ with respect to the d_B-norm. It remains to show that $f \in C^0(X,Y)$. For this we let $\varepsilon > 0$ and we choose $k(\varepsilon)$ so that

$$d(f_k(x), f(x) \leq \varepsilon$$

for all $k \geq k(\varepsilon)$ and all $x \in X$. Since f_k is continuous, there exists $\delta > 0$ such that

$$d(f_k(x), f_k(x_0)) < \varepsilon$$

for all $x, x_0 \in X$ with $d(x, x_0) < \delta$. Arguing as above we get

$$d(f(x), f(x_0)) \leq d(f(x), f_k(x)) + d(f_k(x), f_k(x_0)) + d(f_k(x_0), f(x_0)) < 3\varepsilon,$$

for all $k \geq k(\varepsilon)$ and all $x, x_0 \in X$ with $d(x, x_0) < \delta$. □

Definition 3.7. *Let (X, d) and (Y, d) be two metric spaces. The **oscillation** of $f\colon X \to Y$ is the function*

$$\omega_f\colon (0, \infty) \to [0, \infty],\ \omega_f(\delta) = \sup_{d(x_1,x_2)<\delta} d(f(x_1), f(x_2)).$$

*If $\omega(\delta) \to 0$ as $\delta \to 0$, then f is called **uniformly continuous**.*

*A family $\mathcal{F}$ of maps $f\colon X \to Y$ is called **equicontinuous** if*

$$\sup_{f\in\mathcal{F}} \omega_f(\delta) \to 0 \text{ as } \delta \to 0.$$

Example: Let $0 < \alpha \leq 1$. The **α-Hölder constant** of $f\colon (X, d) \to (Y, d)$ is defined by

$$[f]_\alpha := \sup_{x\neq y} \frac{d(f(x), f(y))}{d(x, y)^\alpha}$$

and f is called **α-Hölder continuous** if $[f]_\alpha < \infty$. We estimate for $x \neq y$

$$d(f(x), f(y)) \leq \frac{d(f(x), f(y))}{d(x, y)^\alpha} d(x, y)^\alpha$$

and hence

$$\omega_f(\delta) \leq [f]_\alpha \delta^\alpha.$$

Therefore every α-Hölder continuous function is uniformly continuous and the family of functions $\mathcal{F} = \{f\colon X \to Y : [f]_\alpha \leq \Lambda\}$ is equicontinuous.

Theorem 3.8 (Theorem of Arzela-Ascoli)**.** *Let X, Y be metric spaces and assume that X is compact and Y is complete. For a family $\mathcal{F} \subset C^0(X, Y)$ the following three statements are equivalent:*

1. *$\mathcal{F}$ is relatively compact in $(C^0(X, Y), d_B(\cdot, \cdot))$ (i.e. $\overline{\mathcal{F}}$ is compact).*

2. *Every sequence $f_k \in \mathcal{F}$ has a subsequence f_{k_j} which converges to $f \in C^0(X, Y)$ (with respect to the metric d_B).*

3. *$\mathcal{F}$ is equicontinuous and for all $f \in \mathcal{F}$ the set $\{f(x) \colon x \in X\}$ is relatively compact in Y.*

Proof. 3. $\Rightarrow$ 2.: Since X is compact there exists a countably dense subset $D \subset X$. For example we can cover X by finitely many balls with radius $1/\nu$ for every $\nu \in \mathbb{N}$ and then $D = \{\text{centres of all these balls}\}$ is countable and dense in X. By the assumptions, after choosing a subsequence, for every $x \in D$ the limit $f(x) \colon \lim_{k\to\infty} f_k(x)$ exists and with a diagonal sequence argument we obtain the existence of the limit $f(x) \colon \lim_{k\to\infty} f_k(x)$ for every $x \in D$.

Now we let $x_1, x_2 \in D$, such that $d(x_1, x_2) \leq \delta$ and we obtain again by the assumption in 3.

$$\begin{aligned} d(f(x_1), f(x_2)) &= \lim_{k\to\infty} d(f_k(x_1), f_k(x_2)) \\ &\leq \limsup_{k\to\infty} \omega_{f_k}(\delta), \\ &\leq \sup_{f\in\mathcal{F}} \omega_f(\delta) \to 0 \end{aligned}$$

as $\delta \to 0$. Therefore $f|_D$ is uniformly continuous. By exercise 4 on the exercise sheet 1 there exists a unique continuous extension (still denoted by f) $f \colon X \to Y$. It remains to show that f_k converges uniformly to f.

For this we let $\varepsilon > 0$ and we choose $\delta > 0$ such that $\omega_f(\delta) < \varepsilon$ for all $f \in \mathcal{F}$. Next we note that $\{B_\delta(x) \colon x \in D\}$ is an open cover of X and since X is compact we choose a finite subcover $\{B_\delta(x) \colon x \in \mathcal{D}_\delta\}$ where $D_\delta \subset D$ is a finite set. Moreover, we choose $k_0 \in \mathbb{N}$ so that for all $k \geq k_0$

$$\max_{x'\in\mathcal{D}_\delta} d(f_k(x'), f(x')) \leq \varepsilon.$$

Now we let $x \in X$, $k \geq k_0$ and we choose $x' \in \mathcal{D}_\delta$ so that $d(x, x') \leq \delta$. This yields

$$\begin{aligned} d(f_k(x), f(x)) &\leq d(f_k(x), f_k(x')) + d(f_k(x'), f(x')) + d(f(x'), f(x)) \\ &\leq \omega_{f_k}(\delta) + \varepsilon + \omega_f(\delta) \leq 3\varepsilon. \end{aligned}$$

Taking the supremum over all $x \in X$ we conclude that $d_B(f_k, f) \to 0$ as $k \to \infty$.

2. $\Rightarrow$ 1.: By assumption 2. the set $\overline{\mathcal{F}}$ is sequentially compact and hence it follows from Theorem 3.2 that $\overline{\mathcal{F}}$ is compact.

1. $\Rightarrow$ 3.: We show first that $\mathcal{F}$ is equicontinuous. By Theorem 3.2 the set $\overline{\mathcal{F}}$ is

precompact and hence for every $\varepsilon > 0$ there exist $\{f_1, \dots, f_N\} \subset \overline{\mathcal{F}}$ so that

$$\overline{\mathcal{F}} \subset \cup_{i=1}^N B_\varepsilon(f_i).$$

This shows that for every $f \in \mathcal{F}$ there exists $i \in \{1, \dots, N\}$ so that $d_B(f, f_i) \leq \varepsilon$ and therefore we conclude for $x, y \in X$ with $d(x, y) \leq \delta$

$$\begin{aligned} d(f(x), f(y)) &\leq d(f(x), f_i(x)) + d(f_i(x), f_i(y)) + d(f_i(y), f(y)) \\ &\leq 2\varepsilon + \max_{1\leq i\leq N} \omega_{f_i}(\delta). \end{aligned}$$

Taking the supremum over all $x, y \in X$ with $d(x, y) \leq \delta$ and all $f \in \mathcal{F}$ yields

$$\sup_{f\in\mathcal{F}} \omega_f(\delta) \leq 2\varepsilon + \max_{1\leq i\leq N} \omega_{f_i}(\delta)$$

Since $f_i \in \overline{\mathcal{F}}$ is continuous and hence uniformly continuous since X is compact, it follows that

$$\lim_{\delta\to 0} \max_{1\leq i\leq N} \omega_{f_i}(\delta) = 0$$

and hence we conclude that

$$\limsup_{\delta\to 0} (\sup_{f\in\mathcal{F}} \omega_f(\delta)) \leq 2\varepsilon$$

which implies that $\mathcal{F}$ is equicontinuous.

Finally we look at the sequence $z_k = f_k(x)$ with $f_k \in \mathcal{F}$. After choosing a subsequence we have that $f_k \to f \in C^0(X, Y)$ uniformly, in particular $z_k = f_k(x) \to f(x)$ which finishes the proof. □

Next we let $(X, d(\cdot, \cdot))$ be a metric space and let $: X \to \mathbb{R}$. Then we define the $C^{0,\alpha}$-norm of u by

$$\|u\|_{C^{0,\alpha}(X)} := \|u\|_{C^0(X)} + [u]_{\alpha,X} = \sup_{x\in X} |u(x)| + \sup_{x\neq y\in X} \frac{|u(x) - u(y)|}{d(x-y)^\alpha}.$$

Theorem 3.9. *The space $C^{0,\alpha}(X) = \{u \in C^0(X) : \|u\|_{C^{0,\alpha}(X)} < \infty\}$ with the $C^{0,\alpha}$-norm is a Banach space.*

Proof. Let $u_k \in C^{0,\alpha}(X)$ be a Cauchy sequence. Then $u_k \in (C^0 \cap B)(X, \mathbb{R})$ and u_k is a Cauchy sequence with respect to $\|\cdot\|_{C^0(X)}$. Hence it follows from Theorem 2.6 that u_k converges to $u \in C^0(X)$ with respect to $\|\cdot\|_{C^0(X)}$. It remains to show that $u \in C^{0,\alpha}(X)$ and that $\|u_k - u\|_{C^{0,\alpha}(X)} \to 0$.

In order to show this we let $x, y \in X$, $x \neq y$ and we estimate

$$\frac{|u(x) - u(y)}{d(x,y)^\alpha} = \lim_{k\to\infty} \frac{|u_k(x) - u_k(y)|}{d(x,y)^\alpha} \leq \lim_{k\to\infty} \|u_k\|_{C^{0,\alpha}(X)} < \infty$$

which shows that $u \in C^{0,\alpha}(X)$. Finally,

$$\begin{aligned}\frac{|(u-u_k)(x) - (u-u_k)(y)|}{d(x,y)^\alpha} &= \lim_{l\to\infty} \frac{|(u_l-u_k)(x) - (u_l-u_k)(y)|}{d(x,y)^\alpha} \\ &\leq \limsup_{l\to\infty} \|u_l - u_k\|_{C^{0,\alpha}(X)} \to 0\end{aligned}$$

as $k \to \infty$ and thus $u_k \to u$ in $C^{0,\alpha}(X)$. □

Theorem 3.10. *Let X be a compact metric space and let $u_k \in C^{0,\alpha}(X)$, $0 < \alpha \leq 1$, with $\|u_k\|_{C^{0,\alpha}(X)} \leq \Lambda < \infty$ for all $k \in \mathbb{N}$. Then there exists $u \in C^{0,\alpha}(X)$, so that up to a subsequence $u_k \to u$ in $C^{0,\beta}(X)$ for all $0 \leq \beta < \alpha$.*

Proof. It follows from example before Theorem 3.8 that for all $\delta > 0$ and all $k \in \mathbb{N}$

$$\omega_{u_k}(\delta) \leq [u_k]_{\alpha,X}\delta^\alpha \leq \Lambda\delta^\alpha.$$

Hence the family of functions $\{u_k\}$ is equicontinuous. Additionally $\{u_k(x)\}$ is uniformly bounded for every $x \in X$ and therefore $\{u_k(x)\}$ is relatively compact for all $x \in X$. By Theorem 3.8 there exists $u \in C^0(X)$ so that up to a subsequence $u_k \to u$ in $C^0(X)$. Now $u \in C^{0,\alpha}(X)$ since

$$\frac{|u(x) - u(y)|}{d(x,y)^\alpha} \leq \lim_{l\to\infty} \frac{|u_l(x) - u_l(y)}{d(x,y)^\alpha} \leq \Lambda.$$

In order to show that $u_k \to u$ in $C^{0,\beta}(X)$ we fix $\delta > 0$ and we look at two cases:

1. $d(x,y) \leq \delta$

$$\begin{aligned}\frac{|(u-u_k)(x) - (u-u_k)(y)|}{d(x,y)^\beta} &= d(x,y)^{\alpha-\beta} \lim_{l\to\infty} \frac{|(u_l-u_k)(x) - (u_k-u_l)(y)}{d(x,y)^\alpha} \\ &\leq 2\Lambda\delta^{\alpha-\beta},\end{aligned}$$

2. $d(x,y) > \delta$

$$\frac{(u-u_k)(x) - (u-u_k)(y)}{d(x,y)^\beta} \leq 2\delta^{-\beta}\|u - u_k\|_{C^0(X)}.$$

For both cases it follows that

$$\limsup_{k\to\infty}[u-u_k]_{\beta,X} \leq 2\Lambda\delta^{\alpha-\beta} \to 0$$

as $\delta \to 0$. □

In the following we let $\Omega \subset \mathbb{R}^n$ be open and bounded, $k \in \mathbb{N}_0$, $0 < \alpha \leq 1$. For $u \in C^k(\Omega)$ we define the norms

$$\|u\|_{C^k(\Omega)} := \sum_{0\leq|\gamma|\leq k} \|D^\gamma u\|_{C^0(\Omega)}$$
$$\|u\|_{C^{k,\alpha}(\Omega)} := \|u\|_{C^k(\Omega)} + \sum_{|\gamma|=k} [D^\gamma u]_{\alpha,\Omega}$$

and the sets

$$C^k(\overline{\Omega}) := \{u \in C^k(\Omega) \colon D^\gamma u \text{ is continously extendable to } \overline{\Omega} \quad \forall |\gamma| \leq k\}$$
$$C^{k,\alpha}(\overline{\Omega}) := \{u \in C^k(\bar{\Omega}) \colon \|u\|_{C^{k,\alpha}(\Omega)} < \infty\}.$$

Theorem 3.11. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then*

1. *$(C^k(\bar{\Omega}), \|u\|_{C^k(\Omega)})$ is a Banach space.*
2. *$(C^{k,\alpha}(\overline{\Omega}), \|u\|_{C^{k,\alpha}(\Omega)})$ is a Banach space.*

Proof. 1. The case $k = 0$ follows from Theorem 3.6. For the case $k = 1$ we argue as follows: Let $u_j \in C^1(\bar{\overline{\Omega}})$ be a Cauchy sequence with respect to $\|u\|_{C^1(\Omega)}$. Then it follows that $u_j \to u$ in $C^0(\overline{\Omega})$ and $\partial_i u_j \to v_i$ in $C^0(\overline{\Omega})$ for all $1 \leq i \leq n$. Hence we have to show that $\partial_i u = v_i$ since then $u_i \to u$ in $C^1(\overline{\Omega})$. In order to see this we choose $x_0 \in \Omega$ and we let $s \in \mathbb{R}$ be small. Then

$$u_j(x_0 + se_i) = u_j(x_0) + \int_0^s \partial_i u_j(x_0 + te_i)dt$$

and the uniform convergence of u_j resp. $\partial_i u_j$ implies that

$$u(x_0 + se_i) = u(x_0) + \int_0^s v_i(x_0 + te_i)dt.$$

Hence it follows from the fundamental theorem of calculus that $u \in C^1(\overline{\Omega})$ with $\partial_i u(x_0) = v_i(x_0)$ for all $1 \leq i \leq n$. The general case now follows by induction using the same argument.

2. Let $u_j \in C^{k,\alpha}(\overline{\Omega})$ be a Cauchy sequence with respect to the norm $\|\cdot\|_{C^{k,\alpha}(\Omega)}$. It follows from 1. that $u_j \to u$ in $C^k(\overline{\Omega})$. For $|\gamma| = k$ we have $D^\gamma u_j \to v^\gamma$ in $C^{0,\alpha}(\overline{\Omega})$

by Theorem 3.9 and hence we conclude that $v^\gamma = D^\gamma u$ which directly implies that $u_j \to u \in C^{k,\alpha}(\overline{\Omega})$. □

Definition 3.12. *We say that $\Omega \subset \mathbb{R}^n$ satisfies a* ***chord-arc*** *condition with constant $\kappa \in [1, \infty)$, if for all $x, y \in \Omega$ there exists a path $\gamma \in C^1([0,1], \Omega)$ with $\gamma(0) = x$, $\gamma(1) = y$ and $L(\gamma) \leq \kappa|x-y|$.*

Theorem 3.13. *Let $\Omega \subset \mathbb{R}^n$ be open, bounded and assume that it satisfies a chord-arc condition with constant κ. Moreover, let $k, l \in \mathbb{N}_0$, $\alpha, \beta \in [0,1]$, with $k+\alpha > l+\beta$. Is $u_j \in C^{k,\alpha}(\overline{\Omega})$ a sequence with $\|u_j\|_{C^{k,\alpha}(\Omega)} \leq \Lambda$ for all $j \in \mathbb{N}$, then there exists a subsequence $u_j \to u$ in $C^{l,\beta}(\overline{\Omega})$.*

Proof. Step 1: $k = l = 0$:

From the assumptions it follows that $\alpha > \beta$ and the result is then a consequence of Theorem 3.10.

Step 2: We have $\|u\|_{C^{0,1}(\Omega)} \leq \kappa \|u\|_{C^1(\Omega)}$.

In order to show this, let $x, y \in \Omega$, $\gamma \in C^1([0,1], \Omega)$, $\gamma(0) = x$, $\gamma(1) = y$, $L(\gamma) \leq \kappa|x-y|$. Then we estimate

$$\begin{aligned} |u(x) - u(y)| &= \left| \int_0^1 \frac{d}{dt} u(\gamma(t)) dt \right| \\ &= \left| \int_0^1 Du(\gamma(t)) \gamma'(t) dt \right| \\ &\leq \|Du\|_{C^0(\Omega)} L(\gamma) \\ &\leq \kappa \|Du\|_{C^0(\Omega)} |x-y| \end{aligned}$$

and hence it follows that

$$\|u\|_{C^{0,1}(\Omega)} = \|u\|_{C^0(\Omega)} + [u]_{1,\Omega} \leq \|u\|_{C^0(\Omega)} + \kappa \|Du\|_{C^0(\Omega)} \leq \kappa \|u\|_{C^1(\Omega)}.$$

Step 3: $k = l \geq 1$:

In this case we have again $1 \geq \alpha > \beta \geq 0$ and Step 1, 2 imply inductively that $u_j \to u$ in $C^k(\overline{\Omega})$. Step 1 yields that $D^\gamma u_j \to v^\gamma$ in $C^{0,\beta}(\overline{\Omega})$ for all $|\gamma| = k$. As in the proof of Theorem 3.11 one then obtains $D^\gamma u = v^\gamma$ for all $|\gamma| = k$ and hence $u_j \to u$ in $C^{k,\beta}(\bar{\Omega})$.

Step 4: $k > l \geq 0$:

By step 2 we know that $\|u_j\|_{C^{l,1}(\Omega)} \leq C \|u_j\|_{C^{k,\alpha}(\Omega)}$. Now we consider two subcases

1. $\alpha > 0$: In this case it follows from step 3 that $u_j \to u$ in $C^k(\overline{\Omega})$ and then step 2 implies that $u_j \to u$ in $C^{l,1}(\overline{\Omega})$. Applying step 3 again yields $u_j \to u$ in $C^{l,\beta}(\overline{\Omega})$.

2. $\alpha = 0$, $\beta = 1$: Here we conclude that $k \geq l+2$ and step 2 implies

$$\|u_j\|_{C^{k-1,1}(\Omega)} \leq C\|u_j\|_{C^k(\Omega)} \leq C\Lambda.$$

   Now step 3 implies $u_j \to u$ in $C^{k-1}(\overline{\Omega})$ and hence also in $C^{l,\beta}(\overline{\Omega})$.

$\square$

Next we apply these results in order to study the kernel of a linear differential operator of second order. We let $\Omega \subset \mathbb{R}^n$ be a bounded C^∞-domain and we introduce the elliptic boundary value problem (BVP)

$$(*) \begin{cases} \sum_{i,j=1}^n a_{ij}(x)\partial^2_{i,j}u(x) + \sum_{i=1}^n b_i(x)\partial_i u(x) + c(x)u(x) = f(x) \quad \forall x \in \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

We assume that there exist constants $0 < \lambda \leq \Lambda < \infty$ such that

1. $\max_{1\leq i,j\leq n} \|a_{ij}\|_{C^{0,\alpha}(\Omega)}$, $\max_{1\leq i\leq n} \|b_i\|_{C^{0,\alpha}(\Omega)}$, $\|c\|_{C^{0,\alpha}(\Omega)} \leq \Lambda$.

2. $\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2$ for all $x \in \Omega$, $\xi \in \mathbb{R}^n$.

We define the operator $L\colon C^2_0(\overline{\Omega}) \to C^0(\overline{\Omega})$, where $C^2_0(\overline{\Omega}) = \{u \in C^2(\overline{\Omega})\colon u = 0 \text{ on } \partial\Omega\}$ by

$$Lu := \sum_{i,j=1}^n a_{ij}\partial^2_{ij}u + \sum_{i=1}^n b_i\partial_i u + cu.$$

It follows from the above assumptions that

$$\|Lu\|_{C^0(\Omega)} \leq \Lambda\left(\sum_{i,j=1}^n \|\partial^2_{ij}u\|_{C^0(\Omega)} + \sum_{i=1}^n \|\partial_i u\|_{C^0(\Omega)} + \|u\|_{C^0(\Omega)}\right) = \Lambda\|u\|_{C^2(\Omega)}$$

and therefore

$$\|L\| \leq \Lambda$$

which shows that $L \in L(C^2_0(\overline{\Omega}), C^0(\overline{\Omega}))$.

In order to proceed we need the so called Schauder estimates (see the book of Gilbarg-Trudinger for a proof).

Theorem 3.14 (Schauder estimates)**.** *Let Ω and L be as above and let $u \in C^2_0(\overline{\Omega})$ solve $Lu = f$. If $f \in C^{0,\alpha}(\overline{\Omega})$, for $\alpha > 0$, then $u \in C^{2,\alpha}(\overline{\Omega})$ with*

$$\|u\|_{C^{2,\alpha}(\Omega)} \leq C(\|f\|_{C^{0,\alpha}(\Omega)} + \|u\|_{C^0(\Omega)})$$

where $C = C(n, \alpha, \lambda, \Lambda, \Omega)$.

As an application we show that the kernel of L is always finite-dimensional.

Theorem 3.15. *The kernel of $L\colon C^2_0(\overline{\Omega}) \to C^0(\overline{\Omega})$ is finite dimensional.*

Proof. On $\ker(L)$ we choose the $\|u\|_{C^0(\Omega)}$-norm. Let $u_k \subset \ker(L)$ be a sequence with $\|u_k\|_{C^0(\Omega)} \leq 1$ for all $k \in \mathbb{N}$. Since u_k is a solution of $Lu_k = 0$ we can use Theorem 3.14 to conclude

$$\|u_k\|_{C^{2,\alpha}(\Omega)} \leq C\|u_k\|_{C^0(\Omega)} \leq C.$$

Hence it follows from Theorem 3.13 that there exists a subsequence u_{k_j} so that $u_{k_j} \to u \in C^2(\overline{\Omega})$ with respect to the C^2-norm. We conclude that $u \in \ker(L)$ with $\|u\|_{C^0(\Omega)} \leq 1$ and therefore $\overline{B_1^{C^0(\overline{\Omega})}(0)} \cap \ker L$ is sequentially compact which, by Theorem 3.4, shows that $\ker L$ has to be finite dimensional. □

4 Extension of linear functionals

In this chapter we show how to extend linear functionals in infinite-dimensional spaces.

Theorem 4.1 (Hahn-Banach)**.** *Let X be a $\mathbb{R}$-vector space and let $p\colon X \to \mathbb{R}$ be a sublinear function, i.e.*

$$\begin{aligned} &p(x+y) \leq p(x) + p(y) \\ &p(\lambda x) = \lambda p(x), \ \forall\ x, y \in X, \ \lambda \geq 0. \end{aligned}$$

Moreover, let $V \subset X$ be a linear subspace and let $\phi\colon V \to \mathbb{R}$ be linear with $\phi(v) \leq p(v)$ for all $v \in V$. Then there exists a linear map $\Phi\colon X \to \mathbb{R}$ with $\Phi|_V = \phi$ and $\Phi(x) \leq p(x)$ for all $x \in X$.

We show this theorem in two steps.

Proof. Step 1: Let $x \notin V \Rightarrow$ there exists an extension Φ of ϕ on $V \oplus \mathbb{R}\, x$ with $\Phi \leq p$.

In order to show this, we make the ansatz $\Phi(v + \alpha x) := \phi(v) + \alpha s$ for all $\alpha \in \mathbb{R}$ with $s = \Phi(x)$. Hence Φ is linear and $\Phi|_V = \phi$. Moreover $\Phi \leq p$ if and only if there exists a number $s \in \mathbb{R}$ so that $\phi(v) + \alpha s \leq p(v + \alpha x)$ for all $v \in V$ and $\alpha \in \mathbb{R}$. But this is equivalent to

$$\begin{cases} s \leq \frac{p(v+\alpha x)-\phi(v)}{\alpha}, \ \forall \alpha \geq 0, \ v \in V \\ s \geq \frac{\phi(v')-p(v'-\alpha' x)}{\alpha'}, \ \forall \alpha' > 0, \ v' \in V. \end{cases}$$

and by setting $w = v/\alpha$ this is in turn equivalent to

$$\begin{cases} s \leq p(w+x) - \phi(w), \ \forall w \in V \\ s \geq \phi(w') - p(w'-x), \ \forall w' \in V \end{cases}$$

We can choose s as we wish if and only if $\phi(w') - p(w'-x) \leq p(w+x) - \phi(w) \Leftrightarrow \phi(w+w') \leq p(w+x) + p(w'-x)$ for all $w, w' \in V$. But since p is assumed to be sublinear it follows that

$$p(w+x) + p(w'-x) \geq p(w+w') \geq \phi(w+w')$$

for all $w, w' \in V$ and hence we conclude that the desired value of s exists.

Step 2: In order to show the result in full generality we need to use Zorn's Lemma. For this we recall a few definitions. A set E with an ordering $\leq$ is called **partially ordered** if the following properties are satisfied for all $a, b, c \in E$

- $a \leq a$,
- $a \leq b$ and $b \leq a$ implies $a = b$ and
- $a \leq b$, $b \leq c$ implies $a \leq c$.

A subset $M \subset E$ is called **totally ordered** if for all $a, b \in M$ we have either $a \leq b$ or $b \leq a$. An element $b \in E$ is called **upper bound** of M if $a \leq b$ for all $a \in M$ and an element $s \in E$ is called **maximal element** if $s \leq a$ for some $a \in E$ implies $a = s$.

Finally, we call $(E, \leq)$ **inductively ordered** if every totally ordered subset $M \subset E$ has an upper bound. With all these definitions we are now able to formulate the

Lemma of Zorn: Let $(E, \leq)$ be inductively ordered, then E has a maximal element.

In order to apply this result in our situation we let E be the set of pairs (W, ψ), where W is a linear subspace of X with $V \subset W$ and $\psi \colon W \to \mathbb{R}$ is linear with $\psi|_V = \phi$ and $\psi(w) \leq p(w)$ for all $w \in W$. Moreover we define the ordering $\leq$ by

$$(W_1, \psi_1) \leq (W_2, \psi_2) \Leftrightarrow W_1 \subset W_2 \text{ and } \psi_2|_{W_1} = \psi_1.$$

It is easy to check that with these definitions $(E, \leq)$ is partially ordered.

Now we claim that $(E, \leq)$ is inductively ordered. For this we let $M = (W_i, \psi_i)_{i \in J} \subset E$ be totally ordered ($W_i \subset W_j$ for all $j \geq i$) and we define $W = \bigcup_{i \in J} W_i$, $\psi \colon W \to \mathbb{R}$ with $\psi|_{W_i} = \psi_i$. We note that

- W is a linear subspace with $V \subset W$. In order to see this let $w_1, w_2 \in W$, thus there exist $i_1, i_2 \in J$ so that $w_1 \in W_{i_1}$, $w_2 \in W_{i_2}$. Without loss of generality we assume $i_1 \leq i_2$ and therefore $w_1, w_2 \in W_{i_2}$. Thus $\lambda w_1 + \mu w_2 \in W_{i_2} \subset W$ for all $\lambda, \mu \in \mathbb{R}$.
- The function ψ is well-defined since for $w \in W_{i_1} \cap W_{i_2}$ with $i_1 < i_2$ we have $w \in W_{i_1}$ and hence $\psi|_{W_{i_2}}(w) = \psi|_{W_{i_1}}(w)$.
- Using the same argument it also follows that ψ is linear.

- It follows from the definition that $\psi \le p$ on W.

We conclude that (W, ψ) is an upper bound for M and therefore $(E, \le)$ is inductively ordered. By the Lemma of Zorn there exists a maximal element (W, Φ) of E. If $W \neq X$, then we use step 1 in order to extend Φ to $W \oplus \mathbb{R}x$ where $x \in X \backslash W$. This gives a contradiction to maximality of (W, Φ) and hence $W = X$ and Φ is the extension we were looking for. □

Theorem 4.2. *Let X be a normed $\mathbb{K}$-vector space ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) and let $V \subset X$ be a linear subspace with the induced norm. For every $\phi \in V'$ there exists a $\Phi \in X'$ with $\Phi|_V = \phi$ and $\|\Phi\|_{X'} = \|\phi\|_{V'}$.*

Proof. We split the proof into two parts:

1. $\mathbb{K} = \mathbb{R}$.

   Here we define the sublinear function $p(x) := \|\phi\|_{V'}\|x\|$ for all $x \in X$. For all $v \in V$ we conclude $\phi(v) \le \|\phi\|_{V'}\|v\| = p(v)$ and hence we can apply Theorem 4.1 in order to conclude that there exists a linear function $\Phi \colon X \to \mathbb{R}$ with $\Phi|_V = \phi$ and $\Phi(x) \le p(x) = \|\phi\|_{V'}\|x\|$ for all $x \in X$. This shows that $\|Phi\|_{X'} \le \|\phi\|_{V'}$ and the reverse inequality is obvious since Φ is an extension of ϕ.

2. $\mathbb{K} = \mathbb{C}$.

   In this case V is a complex subspace and ϕ is $\mathbb{C}$-linear. Hence for $\phi_1 := Re(\phi)$, $\phi_2 := Im(\phi)$ we get

$$
\begin{aligned}
i\phi(x) &= \phi(ix), \ \forall x \in V \\
\Leftrightarrow (-\phi_2 + i\phi_1)(x) &= \phi_1(ix) + i\phi_2(ix), \ \forall x \in V \\
\Leftrightarrow \phi_2(x) &= -\phi_1(ix), \ \forall x \in V \\
\Leftrightarrow \phi(x) &= \phi_1(x) - i\phi_1(ix) \ \forall x \in V.
\end{aligned}
$$

   Now we consider X, V as $\mathbb{R}$-vector spaces. Since ϕ is $\mathbb{C}$-linear it follows that ϕ_1 is $\mathbb{R}$-linear and it follows from the definition of the operator norm that $\|\phi_1\|_{V'} \le \|\phi\|_{V'}$. By applying case 1 we get the existence $\Phi_1 \in X'_{\mathbb{R}}$ with $\Phi_1|_V = \phi_1$ and $\|\Phi\|_{X'} = \|\phi_1\|_{V'} \le \|\phi\|_{V'}$. Now we define

$$
\Phi(x) := \Phi_1(x) - i\Phi_1(ix)
$$

   and it follows from the above argument that Φ is $\mathbb{C}$-linear with $\Phi|_V = \phi$. It remains to show that $\|\Phi\|_{X'} = \|\phi\|_{V'}$. In order to show this, we let $x \in X$ and we express Φ in polar coordinates $\Phi(x) = re^{i\theta}$. We conclude

$$\begin{aligned} |\Phi(x)\| = r &= Re(e^{-i\theta}\Phi(x)) \\ &= Re(\Phi(e^{-i\theta}x)) \\ &= \Phi_1(e^{-i\theta}x) \\ &\leq \|\phi\|_{V'}\|x\| \end{aligned}$$

and therefore $\|\Phi\|_{X'} \leq \|\phi\|_{V'}$. The reverse inequality is again trivial.

□

Theorem 4.3. *Let $(X, \|\cdot\|)$ be a normed vector space and let $V \subset X$ be a subspace. For every $x_0 \in X$ with $\text{dist}(x_0, V) > 0$ there exists a $\phi \in X'$ with $\phi|_V \equiv 0$, $\|\phi\| = 1$ and $\phi(x_0) = \text{dist}(x_0, V)$.*

Proof. We start by defining a functional ϕ on $V \oplus \mathbb{K}x_0$ by

$$\phi(v + \alpha x_0) = \alpha \text{dist}(x_0, V)$$

for all $v \in V$ and all $\alpha \in \mathbb{K}$. Hence $\phi|V = 0$ and $\phi(x_0) = \text{dist}(x_0, V)$. We claim that $\|\phi\|_{(V\oplus\mathbb{K}x_0)'} = 1$. In order to show this we note that without loss of generality we can assume $\alpha \neq 0$ and then we estimate

$$\begin{aligned} |\phi(v + \alpha x_0)| &\leq |\alpha|\text{dist}(x_0, v) \\ &\leq |\alpha| \left\| x_0 - \left(-\frac{1}{\alpha}v\right) \right\| \\ &= \|v + \alpha x_0\|. \end{aligned}$$

Hence we get

$$\|\phi\|_{(V\oplus\mathbb{K}x_0)'} \leq 1$$

and therefore $\phi \in (V \oplus \mathbb{K}x_0)'$.

In order to show the reverse inequality we note that since $\text{dist}(x_0, V) > 0$ for every $\varepsilon > 0$ there exists $v_\varepsilon \in V$ so that $\|x_0 - v_\varepsilon\| \leq (1+\varepsilon)\text{dist}(x_0, V)$ and hence $\phi(x_0 - v_\varepsilon) = \text{dist}(x_0, V) \geq \frac{\|x_0 - v_\varepsilon\|}{1+\varepsilon}$ which yields

$$\|\phi\|_{(V\oplus\mathbb{K}x_0)'} \geq \frac{1}{1+\varepsilon} \; \forall \varepsilon > 0.$$

Therefore the claim is proved and we finish the proof by using Theorem 4.2 in order to extend ϕ to all of X. □

Lemma 4.4. *Let X be a normed space.*

1. *For all* $x_0 \in X$ *there exists* $\phi \in X'$ *so that* $\|\phi\| = 1$ *and* $\phi(x_0) = \|x_0\|$.

2. *If* $\phi(x_0) = 0$ *for all* $\phi \in X'$ *then* $x_0 = 0$.

Proof. Claim 1 follows from Theorem 4.3 applied to $V = \{0\}$ since then $\operatorname{dist}(x_0, V) = \|x_0\| \Rightarrow$ and claim 2 follows from Claim 1. □

Theorem 4.5. *Let* $(X, \|\cdot\|)$ *be a normed space. The canonical map* $J\colon X \to X'' := (X')'$

$$(Jx)(\phi) = \phi(x), \ \forall x \in X, \ \phi \in X'$$

is an isometric embedding.

Proof. We have to show that $\|Jx\|_{X''} = \|x\|_X$ and we split this into two parts:

(i) Let $\phi \in X'$ with $\|\phi\|_{X'} \leq 1$. Then we have

$$|(Jx)(\phi)| = |\phi(x)\| \leq \|\phi\|_{X'}\|x\|_X \leq \|x\|_x$$

which implies $\|Jx\|_{X''} \leq \|x\|_X$.

(ii) Let $x \neq 0$ and choose ϕ as in Lemma 4.4, 1. Then $\|\phi\|_{X'} = 1$ and

$$Jx(\phi) = \phi(x) = \|x\|_X$$

which implies

$$\|Jx\|_{X''} = \sup_{\|\phi\|_{X'} \leq 1} |Jx(\phi)| \geq \|x\|_X.$$

□

Remark: A Banach space X is called reflexive if the canonical map $J : X \to X''$ is additionally surjective. Examples of reflexive spaces are all finite dimensional spaces, all Hilbert spaces, $l^p(\mathbb{R})$ for $1 < p < \infty$ and $L^p(\mu)$ again for $1 < p < \infty$. On the other hand, the spaces $l^1(\mathbb{R})$, $l^\infty(\mathbb{R})$, $L^1(\mu)$, $L^\infty(\mu)$ and $C^0(\bar{\Omega})$ are not reflexive.

Theorem 4.6. *Let* $(X, \|\cdot\|)$ *be a normed space. If the dual space* X' *is separable then* X *is also separable (recall that a metric space is called separable if there exists a countably dense subset).*

Proof. By the assumption of the theorem the set $\{\phi \in X' \colon \|\phi\|_{X'} = 1\}$ has a countably dense subset $\{\phi_k\}_{k\in\mathbb{N}}$. Now we choose $x_k \in X$ so that $\|x_k\|_X = 1$ and $\phi_k(x_k) \geq 1/2$ for all $k \in \mathbb{N}$.

We claim that $\operatorname{span}\{x_k\}$ is dense in X, i.e. $\overline{\operatorname{span}\{x_k\}} = X$. In order to see this we argue by contradiction and we assume that there exists some $x_0 \in X$ with $\operatorname{dist}(x_0, V) > 0$. By Theorem 4.3 there exists $\phi \in X'$ with $\|\phi\| = 1$, $\phi(x_0) = \operatorname{dist}(x_0, V)$ and $\phi|_V = 0$. Since $\{\phi_k\}$ is dense in $\{\phi \in X' \colon \|\phi\|_{X'} = 1\}$ we get

$$0 = \phi(x_k) = \phi_k(x_k) - (\phi_k - \phi)(x_k) \geq \frac{1}{2} - \frac{1}{4} > 0$$

if k is chosen appropriately. □

Now we come to a geometric version of the Hahn-Banach theorem.

Definition 4.7. *Let $(X, \|\cdot\|)$ be a normed space. Two sets $A, B \subset X$ are* ***separated*** *by $\phi \in X'$ if*

$$\sup_{x \in A} \phi(x) \leq \inf_{y \in B} \phi(y).$$

Lemma 4.8. *Let $(X, \|\cdot\|)$ be a normed space and let $K \subset X$ be open and convex with $0 \in K$. Then the function $p(x) := \inf\{t > 0 \colon x/t \in K\}$ is sublinear and*

$$K = \{x \in X \colon p(x) < 1\}.$$

Proof. Since K is open we can choose $\rho > 0$ so that $B_{2\rho}(0) \subset K$ and therefore $\rho \frac{x}{\|x\|} \in K$ for all $x \neq 0$ and hence $p(x) \leq \frac{\|x\|}{\rho} < \infty$ for all $x \neq 0$.

Now we claim that $x/t \in K$ is equivalent to the fact that $p(x) < t$. In order to see this we note that $x/t \in K$ implies that $x/s \in K$ for some $s < t$ since K is open. But now it follows from the definition of p that $p(x) \leq s < t$. On the other hand, if $p(x) < t$ then there exists $s \in (p(x), t)$. As K is convex, K is also starshaped with respect to $0 \in K$ and since $x/s \in K$ we conclude that also $x/t = \frac{s}{t}\frac{x}{s} \in K$.

It remains to show that p is sublinear. For this we let $\lambda > 0$ and we calculate

$$\begin{aligned} p(\lambda x) &= \inf\{t > 0 \colon (\lambda x)/t \in K\} \\ &= \lambda \inf\{t/\lambda > 0 \colon x/(t/\lambda) \in K\} \\ &= \lambda p(x). \end{aligned}$$

Next, we let $\lambda > p(x)$ and $\mu > p(y)$ which implies x/λ, $y/\mu \in K$ and therefore

$$\frac{\lambda}{\lambda + \mu}\frac{x}{\lambda} + \frac{\mu}{\lambda + \mu}\frac{y}{\mu} = \frac{x + y}{\lambda + \mu} \in K.$$

Hence

$$p(\frac{x + y}{\lambda + \mu}) < 1 \ \Leftrightarrow p(x + y) < \lambda + \mu.$$

Finally, we let $\lambda \searrow p(x)$ and $\mu \searrow p(y)$ and we get

$$p(x+y) \le p(x) + p(y).$$

□

Lemma 4.9. *Let $(X, \|\cdot\|)$ be a normed space and let $K \subset X$ be open and convex. Then for $x_0 \notin K$ there exists $\phi \in X'$ so that $\phi(x) < \phi(x_0)$ for all $x \in K$.*

Proof. Without loss of generality we assume that $0 \in K$ and we let p be as in Lemma 4.8. We define $\varphi(tx_0) := t$ on $\mathbb{R}x_0$ and we claim that $\varphi \le p$ on $\mathbb{R}\,x_0$. For $t > 0$ we have $\varphi(tx_0) = t$ and $\frac{tx_0}{t} = x_0 \notin K$ which yields $p(tx_0) \ge t = \varphi(tx_0)$. For $t \le 0$ we have $\varphi(tx_0) = t \le 0 \le p(tx_0)$ and hence the claim is proved.

By Theorem 3.1 there exists $\phi\colon X \to \mathbb{R}$ linear with $\phi(x) \le p(x)$ for all $x \in X$ and $\phi|_{\mathbb{R}\,x_0} = \varphi$. Since by Lemma 3.8 we know that $p(x) < 1$ for all $x \in K$ we conclude that $\phi(x) < 1 = \varphi(x_0) = \phi(x_0)$ for all $x \in K$. Since $0 \in K$ and K is open, there exists $\rho > 0$ so that $B_{2\rho}(0) \subset K$ and hence for all $x \ne 0$ we have

$$\phi(x) \le p(x) = \frac{\|x\|}{\rho} p\left(\rho \frac{x}{\|x\|}\right) < \frac{\|x\|}{\rho}$$

which implies

$$\|\phi\| \le \frac{1}{\rho}$$

and hence $\phi \in X'$. □

Theorem 4.10 (Hahn-Banach for convex sets)**.** *Let $(X, \|\cdot\|)$ be a normed space and let $A, B \subset X$ be convex> Moreover, we assume that A is open and $A \cap B = \emptyset$. Then there exists $\phi \in X'$ which separates A and B.*

Proof. We define the set

$$\begin{aligned} K &:= \{x - y\colon x \in A, y \in B\} \\ &= \bigcup_{Y \in B} \{x - y\colon x \in A\} \end{aligned}$$

and we note that K is open and convex:

$$\lambda(x_1 - y_1) + \mu(x_2 - y_2) = \lambda x_1 + \mu x_2 - (\lambda y_1 + \mu y_2) \in K$$

for all $x_1, x_2 \in A$, $y_1, y_2 \in B$, $\lambda, \mu \in [0,1]$ with $\lambda + \mu = 1$. Since $A \cap B = \emptyset$ we have that $0 \notin K$ and by Lemma 4.9 we get the existence of a $\phi \in X'$ with

$$\phi(z) < \phi(0) = 0$$

for all $z \in K$, which implies that

$$\phi(x) < \phi(y)$$

for all $x \in A$, $y \in B$. Taking the supremum over all $x \in A$ and the infimum over all $y \in B$ finishes the proof. □

Next we show that under certain assumptions two convex sets can be strictly separated.

Theorem 4.11. *Let $(X, \|\cdot\|)$ be a normed vector space and let $A, B \subset X$ be convex with $A \cap B = \emptyset$. Moreover, we assume that A is closed and B is compact. Then A and B can be strictly separated, i.e. there exists $\phi \in X'$ with*

$$\sup_{x \in A} \phi(x) < \inf_{y \in B} \phi(y).$$

Proof. For $\rho > 0$ we define the sets

$$\begin{aligned} A_\rho &:= A + B_\rho(0) = \{x + z \colon x \in A, z \in B_\rho(0)\}, \\ B_\rho &:= B + B_\rho(0) = \{y + z \colon y \in B, z \in B_\rho(0)\} \end{aligned}$$

and we note that A_ρ and B_ρ are open and convex. Moreover, $A_\rho \cap B_\rho = \emptyset$, if ρ is small enough, since the assumptions of the theorem imply that $\operatorname{dist}(A, B) > 0$. Now we can apply Theorem 3.10 to A_ρ and B_ρ and hence there exists $\phi \in X'$ with

$$\phi(x) + \rho\phi(z) = \phi(x + \rho z) \leq \phi(y + \rho z') \leq \phi(y) + \rho\phi(z')$$

for all $x \in A$, $z, z' \in B_1(0)$ and $y \in B$. Next we take the supremum over all $z \in B_1(0)$ and the infimum over all $z' \in B_1(0)$ to get

$$\phi(x) + \rho\|\phi\| \leq \phi(y) - \rho\|\phi\|$$

for all $x \in A$, $y \in B$. Hence

$$\sup_{x \in A} \phi(x) + 2\rho\|\phi\| \leq \inf_{y \in B} \phi(y)$$

and since $2\rho\|\phi\| > 0$ this shows the result. □

Lemma 4.12. *Let $(X, \|\cdot\|)$ be a normed vector space and let $K \subset X$ be closed and convex with $0 \notin K$. Then there exists $\phi \in X'$ with $\|\phi\| = 1$ and*

$$\sup_{x \in K} \phi(x) \leq -\operatorname{dist}(0, K).$$

Proof. We set $R := \operatorname{dist}(0, K) > 0$ and we note that by Theorem 4.10 there exists $\phi \in X'$ separating the sets K and $B_R(0)$. We normalise ϕ so that $\|\phi\| = 1$ and we estimate

$$\sup_{x \in K} \phi(x) \leq \inf_{y \in B_R(0)} \phi(y) = R \inf_{y \in B_1(0)} \phi(y) = -R\|\phi\| = -R.$$

□

<u>Remark:</u> We can not drop the assumption that the set A is open in Theorem 3.10. Namely, if we let A be a dense subspace of a Banach space X and if $B = \{x_0\}$, $x_0 \notin A$ then the assumption that there exists a $\phi \in X'$ with $\phi(x) < \phi(x_0)$ for all $x \in A$ gives a contradiction as follows: Since A is dense in X and ϕ is continuous we have $\phi(x) \leq \phi(x_0)$ for all $x \in X$ and therefore $\phi \equiv 0$.

Lemma 4.13. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded satisfying a chord-arc condition with constant κ. For $u, v \in C^{k,\alpha}(\overline{\Omega})$ we have $uv \in C^{k,\alpha}(\overline{\Omega})$ and*

$$\|uv\|_{C^{k,\alpha}(\Omega)} \leq C(k, \Omega)\|u\|_{C^{k,\alpha}(\Omega)}\|v\|_{C^{k,\alpha}(\Omega)}.$$

Proof. Sheet 4, exercise 4. □

Next we look again at the elliptic differential operator

$$L \colon C_0^{2,\alpha}(\overline{\Omega}) \to C^{0,\alpha}(\overline{\Omega}),$$
$$Lu = \sum_{i,j=1}^{n} a_{i,j}\partial_{i,j}^2 u + \sum_{i=1}^{n} b_i \partial_i u + cu$$

with the same assumptions as in chapter 3. It follows from Lemma 4.13 that

$$\|Lu\|_{C^{0,\alpha}(\Omega)} \leq C\|u\|_{C^{2,\alpha}(\Omega)}$$

for all $u \in C^{2,\alpha}(\overline{\Omega})$.

Theorem 4.14. *Let Ω and L be as above, then the image of the operator $L \colon C_0^{2,\alpha}(\overline{\Omega}) \to C^{0,\alpha}(\overline{\Omega})$ is closed in $C^{0,\alpha}(\overline{\Omega})$.*

Proof. It follows from Theorem 3.15 that the kernel of L is finite-dimensional. By exercise 2 on sheet 4 there exists $X \subset C_0^{2,\alpha}(\overline{\Omega})$ closed so that

$$C_0^{2,\alpha}(\overline{\Omega}) = \ker(L) \oplus X$$

We claim that there exists a $\mu > 0$ so that $\|Lu\|_{C^{0,\alpha}(\Omega)} \geq \mu\|u\|_{C^{2,\alpha}(\Omega)}$ for all $u \in X$.

Assuming that the claim is wrong we get the existence of a sequence $u_k \in X$ with

$$\|Lu_k\|_{C^{0,\alpha}(\Omega)} \leq \frac{1}{k}\|u_k\|_{C^{2,\alpha}(\Omega)}.$$

By replacing u_k with $\frac{u_k}{\|u_k\|_{C^{2,\alpha}(\Omega)}}$ we can assume that $\|u_k\|_{C^{2,\alpha}(\Omega)} = 1$. By Theorem 3.13 there exists a subsequence so that $u_k \to u$ in $C^0(\overline{\Omega})$ and Theorem 3.14 implies

$$\|u_k - u_l\|_{C^{2,\alpha}(\Omega)} \leq C(\|Lu_k - Lu_l\|_{C^{0,\alpha}(\Omega)} + \|u_k - u_l\|_{C^0}) < \varepsilon$$

for all k, l large enough. Hence (u_k) is a Cauchy sequence in $C_0^{2,\alpha}(\overline{\Omega})$ and by Theorem 3.11 this implies that $u_k \to u$ in $C^{2,\alpha}(\overline{\Omega})$ and $u|_{\partial\Omega} = 0$. Now

$$\|Lu - Lu_k\|_{C^{0,\alpha}(\Omega)} \leq \|u - u_k\|_{C^{2,\alpha}(\Omega)} \to 0$$

as $k \to \infty$ which gives $Lu = 0$ since $\|Lu_k\|_{C^{0,\alpha}(\Omega)} \leq \frac{1}{k} \to 0$ as $k \to \infty$. Hence $u \in \ker(L)$ but since X is closed we also have $u \in X$ which gives $u \equiv 0$ which contradicts the fact that $1 = \|u_k\|_{C^{2,\alpha}(\Omega)} \to \|u\|_{C^{2,\alpha}(\Omega)}$. So the claim is true.

Now we let $f_k \in \text{Image}(L)$ be a sequence with $f_k \to f$ in $C^{0,\alpha}(\overline{\Omega})$. By definition there exists a $u_k \in X$ such that $Lu_k = f_k$ and the claim implies that

$$\|u_k - u_l\|_{C^{2,\alpha}(\Omega)} \leq \frac{1}{\mu}\|Lu_k - Lu_l\|_{C^{0,\alpha}(\Omega)} = \frac{1}{\mu}\|f_k - f_l\|_{C^{0,\alpha}(\Omega)} \to 0$$

as $k, l \to \infty$ and hence (u_k) is a Cauchy sequence. By Theorem 3.11 we conclude that $u_k \to u$ in $C^{2,\alpha}(\overline{\Omega})$ with $u|_{\partial\Omega} = 0$ and hence $Lu_k \to Lu$ in $C^{0,\alpha}(\overline{\Omega})$. We conclude that $Lu = f$ and $f \in \text{Image}(L)$. □

5 Uniform boundedness principle

Definition 5.1. *Let X, Y be Banach spaces. A map $T \in L(X,Y)$ is called* ***invertible*** *if T is bijective and $T^{-1} \in L(Y,X)$.*

Theorem 5.2 (Neumann series)**.** *Let X and Y be two Banach spaces. Then we have:*

1. *If $T \in L(X,Y)$ is invertible and if $S \in L(X,Y)$ satisfies $\|S-T\| < \|T^{-1}\|^{-1}$, then S is also invertible.*

2. *The set of invertible maps in $L(X,Y)$ is open.*

Proof. Statement 2 follows directly from statement 1. In order to show statement 1, we let $Q \in L(X,X)$, $Q = T^{-1}(T-S)$ and we calculate

$$S = T - (T-S) = T(\mathrm{id}_X - T^{-1}(T-S)) =: T(\mathrm{id}_X - Q).$$

Since $\|Q\| \le \|T^{-1}(T-S)\| \le \|T^{-1}\|\|T-S\| < 1$ it remains to show that the operator $\mathrm{id}_X - Q$ is invertible. For this we define $A_N := \sum_{n=0}^{N} Q^n$, $N \in \mathbb{N}$, and we estimate

$$\left\|\sum_{n=0}^{N} Q^n\right\| \le \sum_{n=0}^{N} \|Q\|^n \le \sum_{n=0}^{\infty} \|Q\|^n = \frac{1}{1-\|Q\|} < \infty$$

and hence the operator $A := \sum_{n=0}^{\infty} Q^n \in L(X,X)$ is well-defined. We get

$$A_N(\mathrm{id}_X - Q) = \sum_{n=0}^{N} (Q^n - Q^{n+1}) = \mathrm{id}_X - Q^{N+1} \longrightarrow \mathrm{id}_X$$

as $N \to \infty$. Similarly one shows that $(\mathrm{id}_X - Q)A = \mathrm{id}_X$. □

Theorem 5.3 (Continuity principle)**.** *Let X, Y be Banach spaces and let $L(\cdot) : [0,1] \to L(X,Y)$ be a continuous family of operators. Assume that there exists a constant $c > 0$, such that*

$$\|x\| \le c\|L(t)x\|,$$

for all $t \in [0,1]$ and all $x \in X$. Then $L(1)$ is invertible if $L(0)$ is invertible.

Proof. Assume that $L(t)$ is invertible for some $t \in [0,1]$. Then we can use the assumption of the Theorem with $x = L(t)^{-1}y$ in order to get

$$\|L^{-1}(t)\| \leq c.$$

Theorem 4.2 then implies that $L(s)$ is invertible for all $s \in [0,1]$ for which

$$\|L(s) - L(t)\| < \frac{1}{c}.$$

Since $L(\cdot)$ is uniformly continuous on $[0,1]$ we can choose $\tau > 0$ so that $\|L(s) - L(t)\| < 1/c$ for all $s, t \in [0,1]$ with $|s-t| < 2\tau$. Next we decompose $[0,1] = [0,\tau] \cup [\tau, 2\tau] \cup \cdots \cup [N\tau, 1]$ and since $L(0)$ is invertible it follows that $L(\tau)$, ..., $L(1)$ is invertible. ☐

Definition 5.4. *Let (X,d) be a metric space. A subset $S \subset X$ is called* ***nowhere dense*** *if $(\overline{S})^\circ = \emptyset$. Moreover, S is of* ***second category*** *if S is not a countable union of nowhere dense subsets.*

Theorem 5.5 (Baire category theorem)**.** *Let $X \neq \emptyset$ be a complete metric space. Then X is of second category, i.e. if $X = \bigcup_{n=1}^\infty A_n$ with closed sets A_n, then there exists at least one $n_0 \in \mathbb{N}$ so that $\mathring{A}_{n_0} \neq \emptyset$.*

Proof. We argue by contradiction and we assume that $X = \bigcup_{n=1}^\infty A_n$ where all A_n are closed and $\mathring{A}_n = \emptyset$. Now we construct closed balls $\overline{B}_n := \overline{B_{r_n}(x_n)}$, $n \in \mathbb{N}$ so that

- $r_n \in (0, 1/n]$.
- $\overline{B}_n \cap A_n = \emptyset$.
- $\overline{B}_{n+1} \subset B_n$.

This is done inductively as follows: First the set $X \backslash A_1 \neq \emptyset$ is open and hence the existence of B_1 follows. Now we assume that $B_1, ..., B_{n-1}$ have been constructed and we note that $B_{n-1} \backslash A_n \neq \emptyset$ is again open which implies the existence of B_n.

For $m \geq n$ we have $d(x_n, x_m) \leq 2/n \to 0$ as $n \to \infty$ and hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X which implies that $\lim x_n = x \in X$ exists. Now by construction $x \in \bigcap_{n=1}^\infty \overline{B}_n$ and therefore $x \notin A_n$ for all $n \in \mathbb{N}$. But this contradicts the assumption that $X = \bigcup_{n=1}^\infty A_n$. ☐

Lemma 5.6. *Let (X,d) be a complete metric space with $X \neq \emptyset$ and let $(Y, \|\cdot\|)$ be a normed space. Assume that the family $\mathcal{F} \subset C^0(X,Y)$ is pointwise uniformly bounded, i.e.*

$$S(x) = \sup_{f \in \mathcal{F}} \|f(x)\| < \infty \ \forall x \in X,$$

then there exists $x_0 \in X$ *and* $\rho > 0$ *so that*

$$\sup_{d(x,x_0)\leq\rho} S(x) < \infty.$$

Proof. We define the closed sets

$$A_n := \bigcap_{f\in\mathcal{F}} \{x \in X \colon \|f(x)\| \leq n\}$$

and we note that $x \in A_n$ if and only if $S(x) \leq n$. Hence $X = \bigcup_{n=1}^{\infty} A_n$ and by Theorem 5.5 there exists $\overline{B_\rho(x_0)} \subset A_n$ for some $n \in \mathbb{N}$. This implies $S(x) \leq n$ for all $x \in \overline{B_\rho(x_0)}$and therefore

$$\sup_{x\in\overline{B_\rho(x_0)}} S(x) \leq n.$$

□

Theorem 5.7 (Banach-Steinhaus, uniform boundedness principle)**.** *Let X be a Banach space and let Y be a normed space. If the family* $\mathcal{F} \subset L(X, Y)$ *is pointwise uniformly bounded, then*

$$\sup_{T\in\mathcal{F}} \|T\| < \infty$$

and hence $\mathcal{F}$ *is uniformly bounded.*

Proof. It follows from Lemma 5.6 that there exist $x_0 \in X$, $\rho > 0$ and $c < \infty$ so that

$$\|Tx\| \leq c,\ \forall x \in \overline{B_\rho(x_0)},\ \forall T \in \mathcal{F}.$$

Now let $x \in X$ be arbitrary, then we have

$$Tx = \frac{\|x\|}{\rho}\left(T\left(x_0 + \rho\frac{x}{\|x\|}\right) - T(x_0)\right)$$

and hence

$$\|Tx\| \leq \frac{\|x\|}{\rho}2c.$$

This shows that

$$\|T\| \leq \frac{2c}{\rho}$$

for all $T \in \mathcal{F}$. □

Lemma 5.8. *Let X be a Banach space and let* $\phi_k \in X'$ *with* $\phi_k(x) \to \phi(x)$ *for all* $x \in X$*. Then*

$$\sup_k \|\phi_k\| < \infty.$$

Proof. We note that it follows from the convergence assumption that $\sup_k |\phi_k(x)| < \infty$ for all $x \in X$. Hence Theorem 5.7 implies

$$\sup_k \|\phi_k\| < \infty.$$

□

Theorem 5.9 (Open mapping theorem)**.** *Let X and Y be Banach spaces and let $T \in L(X, Y)$ be surjective. Then T is open, i.e. $T(\Omega)$ is open for all $\Omega \subset X$ open.*

Remark: If T is open then there exists $\rho > 0$ so that $T(B_1(0)) \supset B_\rho(0)$ and hence $T(B_R(0)) \supset B_{R\rho}(0)$ for all $R > 0$ which implies that T surjective.

Proof. We split the proof into three steps.

Step 1: We have that $\overline{T(B_1(0))} \supset B_\delta(0)$ for some $\delta > 0$.

We have that $Y = \bigcup_{k\in\mathbb{N}} \overline{T(B_k(0))}$ since T is surjective. Hence we can apply Theorem 5.5 in order to get that $\overline{T(B_k(0))} \supset B_\varepsilon(y_0)$ for some $y_0 \in Y$, $\varepsilon > 0$ and $k \in \mathbb{N}$. Thus for all $\eta \in Y$ with $\|\eta\| < \varepsilon$ there exists a sequence $x_j \in B_k(0)$ so that

$$Tx_j \to y_0 + \eta.$$

In particular, there exists a sequence $\xi_j \in B_k(0)$ with $T\xi_j \to y_0$ and therefore

$$T\left(\frac{x_j - \xi_j}{2k}\right) \to \frac{1}{2k}(y_0 + \eta - y_0) = \frac{\eta}{2k}.$$

Since $\|\frac{x_j - \xi_j}{2k}\| < 1$ this shows that

$$\overline{T(B_1(0))} \supset B_\delta(0),$$

with $\delta := \frac{\epsilon}{2k}$.

Step 2: We have that $B_\delta(0) \subset T(\overline{B_2(0)}) \subset T(B_3(0))$.

We let $y \in B_\delta(0)$ and we note that from the Step 1 we get the existence of a point $x_0 \in B_1(0)$ so that

$$\|y - Tx_0\| < \delta/2 \quad \Leftrightarrow \quad 2(y - Tx_0) \in B_\delta(0).$$

Inductively there exist points $x_k \in B_1(0)$ with

$$y_{k+1} := 2(y_k - Tx_k) \in B_\delta(0),\ y_0 := y.$$

Now for $N \in \mathbb{N}$ we have

$$T\left(\sum_{k=0}^{N} 2^{-k}x_k\right) = \sum_{k=0}^{N} 2^{-k}\left(y_k - \frac{y_{k+1}}{2}\right) = y - 2^{-N-1}y_{N+1} \overset{N\to\infty}{\longrightarrow} y$$

and since $\|x_k\| \leq 1$ it follows that $\sum_{k=0}^{\infty} 2^{-k}\|x_k\| \leq 2$. Therefore $\left(\sum_{k=0}^{n} 2^{-k}x_k\right)$ is a Cauchy sequence in X and the limit $x := \sum_{k=0}^{\infty} 2^{-k}x_k \in X$ exists and satisfies $\|x\| \leq 2$ and $Tx = y$.

Step 3: Let $\Omega \subset X$ be open and let $x_0 \in \Omega$ and $y_0 := Tx_0 \in Y$. There exists $s > 0$ so that $B_{3s}(x_0) \subset \Omega$ and hence it follows from Step 2 that

$$B_{\delta s}(y_0) = y_0 + B_{\delta s}(0) \subset Tx_0 + T(B_{3s}(0)) = T(B_{3s}(x_0)) \subset T(\Omega).$$

□

Theorem 5.10 (Inverse mapping theorem)**.** *Let X, Y be Banach spaces and assume that $T \in L(X,Y)$ is bijective. Then $T^{-1} \in L(Y,X)$.*

Proof. The open mapping theorem, Theorem 5.9, implies that there exists $\rho > 0$ with $B_\rho(0) \subset T(B_1(0))$. Hence $T^{-1}(B_\rho(0)) \subset B_1(0)$ or $T^{-1}(B_1(0)) \subset B_{1/\rho}(0)$ which implies $\|T^{-1}\| \leq 1/\rho$. □

Example: Let $\|\cdot\|_1$, $\|\cdot\|_2$ be two complete norms on the vector space X. Assume that there exists a constant $c < \infty$ with $\|x\|_2 \leq c\|x\|_1$ for all $x \in X$. Then there also exists a constant $c' > 0$ with

$$c'\|x\|_1 \leq \|x\|_2 \leq c\|x\|_1$$

for all $x \in X$, since by the assumption the map $\mathrm{id}\colon (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is continuous and Theorem 4.10 then implies that the inverse map $\mathrm{id}\colon (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is also continuous.

Theorem 5.11 (Closed graph theorem)**.** *Let X, Y be Banach spaces and let $T\colon X \to Y$ be linear. The following two statements are equivalent:*

1. *The graph of T, $G(T) := \{(x, Tx)\colon x \in X\} \subset X \times Y$, is a closed subspace of $X \times Y$.*
2. *The map T is continuous.*

Proof. We note that $X \times Y$ with the norm $\|(x,y)\| = \max\{\|x\|_X, \|y\|_Y\}$ is a Banach space (see the exercises) and the maps

$$\begin{aligned} P_X &: X \times Y \to X,\ P_X(x,y) = x, \\ P_Y &: X \times Y \to Y,\ P_Y(x,y) = y \end{aligned}$$

are continuous since $\|P_X\|, \|P_Y\| \leq 1$.

2. $\Rightarrow$ 1.: We let $(x_n, Tx_n) \to (x_0, y_0) \in X \times Y$ and hence $y_0 = \lim_{n\to\infty} Tx_n = Tx_0$ since T is continuous. Therefore $(x_0, y_0) \in G(T)$ and $G(T)$ is closed.

1. $\Rightarrow$ 2. Since $G(T)$ is closed $(G(T), \|\cdot\|_{X\times Y})$ is a Banach space (see exercise sheet 1). Now the map $P_X|_{G(T)} : G(T) \to X$ is continuous and bijective and hence it follows from Theorem 5.10 that $(P_X|_{G(T)})^{-1} : X \to G(T)$ is also continuous. Then T is also continuous since $T = P_Y \circ (P_X|_{G(T)})^{-1} : X \to Y$. □

<u>Example:</u> (Theorem of Hellinger-Toeplitz) Let H be a Hilbert space with a scalar product $\langle \cdot, \cdot \rangle$ and let $A : H \to H$ be linear with

$$\langle Ax, y \rangle = \langle x, Ay \rangle \ \forall x, y \in H.$$

Then A is continuous.

We will show this using the closed graph theorem and hence we let $(x_n, Ax_n) \in G(A)$ with $(x_n, Ax_n) \to (x_0, y_0)$. Now we have for all $z \in H$

$$\begin{array}{ccc} \langle Ax_n, z \rangle & \xrightarrow{=} & \langle x_n, Az \rangle \\ \downarrow & & \downarrow \\ \langle y_0, z \rangle & & \langle x_0, Az \rangle = \langle Ax_0, z \rangle \end{array}$$

which implies that $Ax_0 = y_0$ and therefore $(x_0, y_0) \in G(A)$.

Lemma 5.12. *Let Y be a closed subspace of the Banach space X. Let Z be an algebraic complement of Y, i.e. $X = Y \oplus Z$ and let $P_Y : X \to Y$ be the linear projection onto the first summand. Then P_Y is continuous if and only if Z closed.*

Proof. "$\Rightarrow$": This is obvious since $Z = \ker(P_Y)$ is closed.

"$\Leftarrow$": Define the map $T : (Y \times Z, \|(y,z)\|_{Y\times Z} = \max\{\|y\|_Y, \|z\|_Z\}) \to (X, \|\cdot\|_X)$ between the two banach spaces by $T(y,z) = y + z$ and note that $\|T\| \leq 2$ and T is bijective by the assumption of the Lemma. It follows from Theorem 5.10 that T^{-1} is also continuous. Next we let $\pi_1 : Y \times Z \to Y$, $\pi_1(y,z) = y$ and we note that π_1 is continuous since $\|\pi\| \leq 1$. Hence it follows that $P_Y = \pi_1 \circ T^{-1}$ is the composition of two continuous maps and therefore it is continuous as well. $\square$

6 L^p-spaces

Let μ be a measure on X and let $f\colon X \to \mathbb{R}$ be a μ-measurable function. Then the L^p-norm of f, $1 \le p \le \infty$, is defined by

$$\|f\|_{L^p(\mu)} = \begin{cases} \left(\int |f|^p d\mu\right)^{1/p}, & 1 \le p < \infty \\ \operatorname{ess\,sup}|f| = \inf\{s > 0\colon |f(x)| \le s, \mu\text{-a.e., } x \in X\}, & p = \infty \end{cases}.$$

The space $\{f : X \to \mathbb{R}\colon \|f\|_{L^p} < \infty\}$ is a vector space since the map $t \mapsto t^p$ is convex and therefore $|f+g|^p = 2^p|(f+g)/2|^p \le 2^{p-1}(|f|^p + |g|^p)$. We also define the set

$$\mathcal{N} = \{f\colon X \to \mathbb{R}\colon f(x) = 0, \mu\text{-a. e.}, x \in X\}.$$

Definition 6.1. *The L^p-space is defined by*

$$L^p(\mu) = \{f\colon X \to \mathbb{R}\colon f \text{ measurable}, \ \|f\|_{L^p(\mu)} < \infty\}/\mathcal{N}$$

and $\|\cdot\|_{L^p}$ is a norm on $L^p(\mu)$. Moreover the normed space $(L^p(\mu), \|\cdot\|_{L^p(\mu)})$ is complete by the theorem of Fischer-Riesz.

Theorem 6.2. *Let $\mu = \mathcal{L}^n$ on $\mathbb{R}^n$. Then $L^p(\mathbb{R}^n)$ is separable for $1 \le p < \infty$.*

Proof. It is enough to show this result for the space $C_c^0(\mathbb{R}^n)$ since $C_c^0(\mathbb{R}^n)$ is dense in $L^p(\mathcal{L}^n)$. Let $B_\nu(0)$ be the ν-Ball in $\mathbb{R}^n$ and for every $\nu \in \mathbb{N}$ we cover the compact set $\overline{B_\nu(0)}$ by finitely many balls $B_{1/\nu}(x_{\nu,j})$, $1 \le j \le j_\nu$. For $1 \le j \le j_\nu$ we define the functions

$$\chi_{\nu,j}(x) := \begin{cases} 1 - \nu d(x, x_{\nu,j}), & \text{if } d(x, x_{\nu,j}) < 1/\nu \\ 0, & \text{if } d(x, x_{\nu,j}) \ge 1/\nu \end{cases} \in C_c^0(\mathbb{R}^n). \quad \text{and}$$

$$\chi_{\nu,0}(x) := \operatorname{dist}(x, B_\nu(0)).$$

Moreover, we let $\eta_{\nu,j}(x) := \frac{\chi_{\nu,j}(x)}{\sum_{k=0}^{j_\nu} \chi_{\nu,k}(x)}$ for $j = 0, ..., j_\nu$. It follows that

- $\eta_{\nu,j} \in C_c^0(\mathbb{R}^n)$, for all $j = 1, ..., j_\nu$ and
- $\sum_{j=0}^{j_\nu} \eta_{\nu,j}(x) = 1$ for all $x \in \mathbb{R}^n$.

Now we let $u \in C_c^0(\mathbb{R}^n)$ and we assume that $u \equiv 0$ on $\mathbb{R}^n \setminus B_\nu(0)$. We let $\alpha_j := u(x_{\nu,j})$ for $1 \leq j \leq j_\nu$ and we note that for all $x \in \mathbb{R}^n$

$$\begin{aligned}\left|u(x) - \sum_{j=1}^{j_\nu} \alpha_j \eta_{\nu,j}(x)\right| &= \left|\sum_{j=1}^{j_\nu} (u(x) - \alpha_j)\eta_{\nu,j}(x)\right| \\ &\leq \sup_{d(x,y)\leq 1/\nu} |u(x) - u(y)| \sum_{j=1}^{j_\nu} \eta_{\nu,j}(x) \\ &\leq \omega_u\left(\frac{1}{\nu}\right) \to 0\end{aligned}$$

as $\nu \to \infty$. Note that the first equality holds since for $x \notin B_\nu(0)$ we have $u(x) = 0$ and for $x \in B_\nu(0)$ it follows that $\eta_{\nu,0} \equiv 0$. Hence we have shown that span $\{\eta_{\nu,j} \colon \nu \in \mathbb{N}, 1 \leq j \leq j_\nu\}$ is dense in $C_c^0(\mathbb{R}^n)$. Now we just look linear combinations of these functions with rational coefficients and this is a countably dense subset of $C_c^0(\mathbb{R}^n)$. □

We remark that the same argument works for a Radon measure μ on a σ-finte metric space X since in this case $C_c^0(X)$ is also dense in $L^p(\mu)$.

Next we want to characterise the dual spaces of the L^p-spaces. For this we define for all $1 \leq q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ the map

$$J \colon L^q(\mu) \to L^p(\mu)', \quad (Jv)(u) := \langle Jv, u\rangle = \int_X uv d\mu$$

and we note that the Hölder inequality yields $|(Jv)(u)| \leq \|u\|_{L^p}\|v\|_{L^q}$. Hence $\|Jv\| = \sup_{\|u\|_{L^p}\leq 1} |(Jv)(u)| \leq \|v\|_{L^q}$ and therefore

$$\|J\| = \sup_{\|v\|_{L^q}\leq 1} \|Jv\| \leq 1,$$

which implies that J is continuous.

Lemma 6.3. *Let $p, q \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$ and assume that the measure μ is σ-finite if $q = \infty$. Then the map $J \colon L^q(\mu) \to L^p(\mu)$ is an isometry, i.e. $\|Jv\| = \|v\|_{L^q}$ for all $v \in L^q(\mu)$.*

Proof. Without loss of generality we assume that $\|v\|_{L^q} = 1$ and we split the proof into two parts.

1. $q < \infty$

Here we define the map

$$\mathcal{D}_q\colon \{v \in L^q(\mu)\colon \|v\|_{L^q} = 1\} \to \{u \in L^p(\mu)\colon \|u\|_{L^p} = 1\}, \quad \mathcal{D}_q(v) = |v|^{q-2}v$$

and we calculate

$$\|\mathcal{D}_q(v)\|_{L^p} = \left(\int |v|^{(q-1)p} d\mu\right)^{1/p} = \left(\int |v|^q d\mu\right)^{1/p} = 1$$

which shows that $\mathcal{D}_q$ is well-defined. Now

$$(Jv)(\mathcal{D}_q(v)) = \int \mathcal{D}_q(v) v d\mu = \int |v|^q d\mu = 1 = \|\mathcal{D}_q(v)\|_{L^p}$$

and therefore

$$\|Jv\| \geq 1 = \|v\|_{L^q}$$

which, together with the argument before the lemma, finishes the proof for $q < \infty$.

2. $q = \infty$

In this case $p = 1$ and $\|v\|_{L^\infty} = 1$. By the assumption of the lemma there exists sets $E_1 \subset E_2 \subset \ldots$ so that E_j is μ-measurable for all $j \in \mathbb{N}$ with $\mu(E_j) < \infty$ and $X = \bigcup_{j=1}^\infty E_j$. We let

$$E_{j,\delta} := \{x \in E_j\colon |v(x)| \geq 1 - \delta\}$$

and we note that $\|v\|_{L^\infty} = 1$ implies that $\mu(E_{j,\delta}) > 0$ if j is large enough. Next we let $u = (\text{sign} v)\chi_{E_{j,\delta}} \in L^1(\mu)$ and we calculate

$$(Jv)(u) = \int uv d\mu = \int_{E_{j,\delta}} |v| d\mu \geq (1-\delta)\mu(E_{j,\delta}) = (1-\delta)\|u\|_{L^1}$$

which implies

$$\|Jv\| \geq (1-\delta)$$

for all $\delta > 0$ and therefore

$$\|Jv\| \geq 1 = \|v\|_{L^\infty}.$$

□

Theorem 6.4 (Riesz)**.** *Let $1 < q \leq \infty$, $1 \leq p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ and let μ be a σ-finite measure on X in the case $p = 1 \Leftrightarrow q = \infty$. Then the map*

$$J\colon L^q(\mu) \to L^p(\mu)',\ (Jv)(u) = \int_X uv d\mu$$

is a surjective isometry.

Proof. It follows from Lemma 6.3 that we only have to show that J is surjective.

We first consider the case $p > 1$:

For this we let $\phi \in L^p(\mu)'$ and we assume without loss of generality that $\|\phi\| = 1$. Our goal is to find a function $v \in L^q(\mu)$ with $\phi = Jv$. We let

$$S := \{u \in L^p(\mu)\colon \|u\|_{L^p} = 1\}$$

and we note that $\sup_{u\in S} \phi(u) = 1$. Now if we assume there exists a function $v_0 \in L^q(\mu)$ with $Jv_0 = \phi$, then it follows from Lemma 6.3 that $\|v_0\|_{L^q} = \|Jv_0\| = \|\phi\| = 1$. Therefore $\|\mathcal{D}_q(v_0)\|_{L^p} = 1$ and if we define $\mathcal{D}_q(v_0) = |v_0|^{q-2}v_0 =: u_0$ it follows that $u_0 \in S$ with

$$\phi(u_0) = \phi(\mathcal{D}_q(v_0)) = Jv_0(\mathcal{D}_q(v_0)) = \int_X |v_0|^q d\mu = 1$$

which means that ϕ attains its maximum on S in u_0. Therefore our idea is now to show the existence of a maximum $u_0 \in S$ of $\phi|_S$. In order to do this we let $u_k \in S$ be a minimising sequence with $\phi(u_k) \to 1 = \sup\{\phi(u)\colon u \in S\}$. Next we use the **uniform convexity** of the L^p-spaces for $1 < p < \infty$ which we will show in Lemma 6.5 below. This implies that for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $\|u\|_{L^p} = \|v\|_{L^p} = 1$ with $\left\|\frac{u+v}{2}\right\|_{L^p} \geq 1 - \delta$ it holds that

$$\|u - v\|_{L^p} < \varepsilon.$$

Now for every $\delta > 0$ and k, l large enough we have that

$$1 - \delta \leq \frac{1}{2}(\phi(u_k) + \phi(u_l)) = \phi\left(\frac{u_k + u_l}{2}\right) \leq \left\|\frac{u_k + u_l}{2}\right\|_{L^p}$$

and therefore

$$\|u_k - u_l\|_{L^p} < \varepsilon$$

which shows that (u_k) is a Cauchy sequence in the complete space $L^p(\mu)$ and hence $u_k \to u_0 \in S$ in $L^p(\mu)$ and $\phi(u_0) = 1$.

Next we show that $v_0 := \mathcal{D}_p(u_0) \in L^q(\mu)$ satisfies $Jv_0 = \phi$. For this we define $G\colon L^p(\mu) \to \mathbb{R}$, $G(u) := \int_X |u|^p d\mu$ and we let $|t| < 1$ and $u \in L^p(\mu)$ be arbitrary. Then

$$\partial_t |u_0 + tu|^p = |u_0 + tu|^{p-2}(u_0 + tu)up$$

and this function is in $L^1(\mu)$ since

$$|\partial_t |u_0 + tu|^p| \leq (|u_0 + |u|)^{p-1}|u|p \leq 2^p p(|u_0|^p + |u|^p).$$

Hence we can apply the theorems on parameter differentiation of integrals and we get

$$\begin{aligned}\frac{d}{dt}\|u_0 + tu\|_{L^p}^{-1}|_{t=0} &= \frac{d}{dt}|_{t=0} G(u_0 + tu)^{1/p} \\ &= -\frac{1}{p} G(u_0)^{-1/p-1} \frac{d}{dt} G(u_0 + tu)|_{t=0} \\ &= -\frac{1}{p} p \int_X |u_0|^{p-2} u_0 u d\mu \\ &= -(J(\mathcal{D}_p(u_0)))(u).\end{aligned}$$

Since $u_0 \in S$ is a maximum of $\phi|_S$ this yields

$$\begin{aligned}0 &= \frac{d}{dt}\phi\left(\frac{u_0 + tu}{\|u_0 + tu}\|_{L^p}\right)|_{t=0} \\ &= \frac{d}{dt}|_{t=0}(\|u_0 + tu\|_{L^p}^{-1}(\phi(u_0) + t\phi(u))) \\ &= \phi(u) - \phi(u_0)(J(\mathcal{D}_p(u_0)))(u_0) \\ &= \phi(u) - J(\mathcal{D}_p(u_0))\end{aligned}$$

from which the desired result follows.

Next we consider the case $p = \infty$:

Here we let $\phi \in L^1(\mu)'$ and for $A \subset X$ measurable with $\mu(A) < \infty$ and $p \geq 1$ we define

$$\phi_p \colon L^p(\mu) \to \mathbb{R},\ \phi_p(u) = \phi(\chi_A u)$$

and we note that the Hölder inequality implies

$$|\phi_p(u)| \leq \|\phi\| \|\chi_A u\|_{L^1} \leq \|\phi\| \mu(A)^{1/q} \|u\|_{L^p}$$

for $\frac{1}{p} + \frac{1}{q} = 1$ and hence $\phi_p \in L^p(\mu)'$. Therefore we can use the result we just showed for $p > 1$ and there exists $v_p \in L^q(\mu)$ such that for all $u \in L^p(\mu)$

$$(Jv_p)(u) = \int uv_p d\mu = \phi_p(u) = \phi(\chi_A u).$$

Next we note that

$$\begin{aligned}(J(\chi_{X\setminus A}v_p))(u) &= \int u\chi_{X\setminus A}v_p d\mu \\ &= (Jv_p)(\chi_{X\setminus A}u) \\ &= \phi(\chi_A \chi_{X\setminus A}u) = 0\end{aligned}$$

and therefore $J(\chi_{X\setminus A}v_p) = 0$ which by Lemma 6.3 implies $v_p = 0$ almost everywhere on $X\setminus A$.

Next we let $1 < p < p'$ and $u \in L^{p'}(\mu)$. This implies $q > q'$ and since $v_p = 0$ a.e. on $X\setminus A$ we get $v_p \in L^{q'}(\mu)$ and hence

$$\begin{aligned}(Jv_{p'})(u) = \int uv_{p'}d\mu = \phi_{p'}(\mu) &= \phi(\chi_A\chi_A u) \\ &= \phi_p(\chi_A u) \\ &= \int u\chi_A v_p d\mu \\ &= \int uv_p d\mu = (Jv_p)(u)\end{aligned}$$

where we used that $\chi_A u \in L^p(\mu)$. Therefore $J(v_{p'} - v_p) = 0$ and we conclude again from Lemma 6.3 that $v_{p'} = v_p$ almost everywhere. Thus we define the function $v := v_p$ for some $p > 1$ and we note that for all $q < \infty$

$$\|v\|_{L^q} = \|Jv\| = \|\phi_p\| \leq \|\phi\|\mu(A)^{1/q}.$$

Letting $q \to \infty$ we conclude with the help of exercise 2, sheet 6 that

$$\|v\|_{L^\infty} \leq \|\phi\|.$$

Therefore $v \in L^\infty(\mu)$ with $\int uvd\mu = \phi(\chi_A u)$ for all $u \in L^p(\mu)$, $p > 1$ and by density als for all $u \in L^1(\mu)$.

Now we let $X = \bigcup_{k\in\mathbb{N}} A_k$, $\mu(A_k) < \infty$, A_k measurable, $A_1 \subset A_2 \subset \ldots$ and we let v_k be constructed as above with A replaced by A_k. Hence $\|v_k\|_{L^\infty} \leq \|\phi\|$ and for $k < k'$, $u \in L^1(\mu)$ we conclude

$$\begin{aligned}\int u(\chi_{A_k}v_{k'})d\mu &= \int (\chi_{A_k}u)v_{k'}d\mu \\ &= \phi(\chi_{A'_k}u\chi_{A_k}) \\ &= \phi(\chi_{A_k}u) = \int uv_k d\mu\end{aligned}$$

which shows that $v_{k'}|_{A_k} = v_k$ almost everywhere. There the function $v \in L^\infty(\mu)$, $\|v\|_{L^\infty} \leq \|\phi\|$ with $v|_{A_k} = v_k$ is well-defined and by the Lebesgue convergence theorem we get

$$(Jv)(u) = \int uv d\mu = \lim_{k\to\infty} \int u\chi_{A_k} v d\mu = \lim_{k\to\infty} \phi(\chi_{A_k} u) = \phi(u).$$

□

Lemma 6.5 (Uniform convexity of L^p, $1 < p < \infty$)**.** *For $1 < p < \infty$ there exists $C = C(p) > 0$ such that for all $u, v \in L^p(\mu)$ we have*

$$\frac{1}{2}(\|u\|_{L^p}^p + \|v\|_{L^p}^p) - \left\|\frac{u+v}{2}\right\|_{L^p}^p \geq \begin{cases} C\|u-v\|_{L^p}^p, & p \geq 2 \\ C(\|u\|_{L^p}^{p-2} + \|v\|_{L^p}^{p-2})\|u-v\|_{L^p}^2, & 1 < p \leq 2 \end{cases}.$$

Proof. The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|^p$ satisfies $f'(x) = p|x|^{p-2}x$ for all $x \in \mathbb{R}$ and

$$f''(x) = p(p-1)|x|^{p-2}, \text{ for all } \; x \neq 0 \text{ if } p < 2.$$

Hence f is convex and it follows that for all $x_0, x_1 \in \mathbb{R}$

$$\frac{1}{2}(|x_0|^p + |x_1|^p) - \left|\frac{x_0+x_1}{2}\right|^p > 0.$$

Now we claim that we can improve this estimate to

$$\frac{1}{2}(|x_0|^p + |x_1|^p) - \left\|\frac{x_0+x_1}{2}\right\| \geq \begin{cases} c(p)|x_0-x_1|^p, & p \geq 2 \\ c(p)(|x_0| + |x_1|)^{p-2}|x_0-x_1|^2, & 1 < p \leq 2 \end{cases}. \tag{6.1}$$

In order to show this we note that without loss of generality we can assume that $x_1 = 1$ and $x_0 = x \in [-1, 1]$. Next we denote the left hand side of (6.1) by

$$\sigma(x) := \frac{1}{2}(1 + |x|^p) - \left|\frac{1+x}{2}\right|^p$$

and the right hand side by

$$\tau(x) := \begin{cases} c(p)|1-x|^p, & p \geq 2 \\ c(p)(1+|x|)^{p-2}(1-x)^2, & 1 < p \leq 2 \end{cases}$$

We calculate

$$\begin{aligned} \sigma'(x) &= \frac{p}{2}\left(|x|^{p-2}x - \left(\frac{1+x}{2}\right)^{p-1}\right) < 0 \text{ for } x < 1, \\ \sigma(1) &= 0 = \sigma'(1), \\ \sigma''(1) &= \frac{p}{2}\left((p-1) - \frac{(p-1)}{2}\right) = \frac{p(p-1)}{4} > 0 \end{aligned}$$

and

$$\tau(1) = \tau'(1) = 0$$
$$\tau''(1) = \begin{cases} 0, & p > 2 \\ c(p)2^{p-1}, & 1 < p \leq 2 \end{cases}$$

Hence there exists $\mu = \mu(p) > 0$ so that $\sigma''(1) - \mu\tau''(1) > 0$ and a $\delta = \delta(p) > 0$ such that $\sigma(x) - \mu\tau(x) > 0$ for all $1 - \delta \leq x \leq 1$. For all $-1 \leq x \leq 1 - \delta$ we estimate

$$\sigma(x) \geq \sigma(1-\delta) \geq \frac{\sigma(1-\delta)}{c(p)2^p}\tau(x)$$

since $\tau(x) \leq c(p)2^p$ for all $x \in [-1, 1]$. Therefore $\sigma(x) \geq \tilde{c}(p)\tau(x)$ for $\tilde{c}(p) := \min\left\{\mu(p), \frac{\sigma(1-\delta)}{c(p)2^p}\right\}$ and hence the claim is proved.

The claim directly yields the statement of the lemma for $p \geq 2$ by integration. For $1 \leq p < 2$ we use the claim and the Hölder inequality with the dual exponents $\frac{2}{2-p}$ and $\frac{2}{p}$ in order to get

$$\begin{aligned} \|u-v\|_{L^p}^2 &= \left(\int (|u|+|v|)^{\frac{p}{2}(2-p)}(|u|+|v|)^{\frac{p}{2}(p-2)}|u-v|^p d\mu\right)^{2/p} \\ &\leq \left(\int (|u|+|v|)^p\right)^{(2-p)/p} \left(\int (|u|+|v|)^{p-2}|u-v|^2 d\mu\right) \\ &\leq c(p)(\|u\|_{L^p} + \|v\|_{L^p})^{2-p}\left(\frac{1}{2}\|u\|_{L^p}^p + \|v\|_{L^p}^p - \left\|\frac{u+v}{2}\right\|_{L^p}^p\right). \end{aligned}$$

□

7 The dual space of $C^0(X)$

In this chapter we want to characterise the dual space of $C^0(X)$ for a compact metric space X.

Definition 7.1. *An outer measure μ on a metric space X is called **Borel regular**, if*

1. *every Borel set B is μ-measurable.*

2. *for all $S \subset X$ there exists a Borel set $B \supset S$ with $\mu(B) = \mu(S)$.*

*The measure μ ist called a **Radon measure**, if $\mu(K) < \infty$ for all $K \subset X$ compact.*

Lemma 7.2. *Let μ be a Borel regular measure on the metric space X and let $E \subset X$. Then the measure $(\mu \llcorner E)(S) := \mu(E \cap S)$ for all $S \subset X$ is again Borel regular, if one of the following two conditions is satisfied.*

1. *E is a Borel set.*

2. *E is σ-finite, i.e. $E = \bigcup_{j\in\mathbb{N}} E_j$ for E_j μ-measurable and $\mu(E_j) < \infty$ for all j.*

Proof. We first show that every Borel set B is $\mu \llcorner E$-measurable. This follows from the fact that for all $S \subset X$ we have

$$\begin{aligned}(\mu \llcorner E)(S) &= \mu(E \cap S)\\ &= \mu((E \cap S) \cap B) + \mu((E \cap S)\backslash B)\\ &= \mu(E \cap (S \cap B)) + \mu(E \cap (S \backslash B))\\ &= (\mu \llcorner E)(S \cap B) + (\mu \llcorner E)(S \backslash B).\end{aligned}$$

We split the rest of the proof into three parts.

1. E is a Borel set.

   For $S \subset X$ we choose a Borel set $B_1 \supset E \cap S$ with $\mu(B_1) = \mu(E \cap S)$. Then

$B = B_1 \cup (X \backslash E)$ is also a Borel set with $B \supset S$ and

$$\begin{aligned}(\mu \llcorner E)(S) &\leq (\mu \llcorner E)(B) = \mu(E \cap B) \\ &= \mu(B_1 \cap E) \leq \mu(B_1) \\ &= \mu(E \cap S) = (\mu \llcorner E)(S).\end{aligned}$$

Therefore we conclude that $(\mu \llcorner E)(S) = (\mu \llcorner E)(B)$.

2. E is μ-measurable with $\mu(E) < \infty$.

   Choose a Borel set $B \supset E$ with $\mu(B) = \mu(E)$. For every $S \subset X$ we have

$$\begin{aligned}(\mu \llcorner B)(S) &= \mu(B \cap S) \leq \mu(E \cap S) + \mu(B \backslash E) \\ &= \mu(E \cap S) + \mu(B) - \mu(E) = (\mu \llcorner E)(S)\end{aligned}$$

   and hence $\mu \llcorner B = \mu \llcorner E$ and we are back in case 1.

3. $E = \bigcup_{j=1}^{\infty} E_j$ with E_j μ-measurable and $\mu(E_j) < \infty$, $E_j \subset E_{j+1}$ for all j.

   For $S \subset X$ we use step 2 in order to conclude that there exist Borel sets $B_j \supset S$ with $(\mu \llcorner E_j)(B_j) = (\mu \llcorner E_j)(S)$. The set $B := \bigcap_{j=1}^{\infty} B_j$ is then again a Borel set with $B \supset S$ and

$$\begin{aligned}(\mu \llcorner E)(B) = \mu\left(\bigcup_{j=1}^{\infty} (E_j \cap B)\right) &= \lim_{j \to \infty} \mu(E_j \cap B) \\ &\leq \limsup_{j \to \infty} \mu(E_j \cap B_j) \leq \limsup_{j \to \infty} \mu(E_j \cap S) \\ &\leq \mu(E \cap S) = (\mu \llcorner E)(S).\end{aligned}$$

□

Theorem 7.3 (Caratheodory criterion)**.** *Let μ be an outer measure on the metric space X with the property that if A and B are subsets of X with* $\mathrm{dist}(A, B) > 0$ *then*

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Then all Borel sets are μ-measurable.

Proof. We show that under the assumption every closed set is μ-measurable and since the Borel σ-algebra is generated by the closed sets this implies the full statement. Hence we let $C \subset X$ be closed and we have to show that for every $S \subset X$

$$\mu(S) \geq \mu(S \cap C) + \mu(S \backslash C).$$

We assume without loss of generality that $\mu(S) < \infty$ and we define

$C_k := \left\{x \in X : \operatorname{dist}(x, C) \leq \frac{1}{k}\right\}$ for $k \in \mathbb{N}$. Since $\operatorname{dist}(S \cap C, S \backslash C_k) \geq \frac{1}{k} > 0$ the assumption of the theorem implies

$$\mu(S \cap C) + \mu(S \backslash C_k) = \mu((S \cap C) \cup (S \backslash C_k)) \leq \mu(S).$$

Hence it remains to show that

$$\mu(S \backslash C) = \lim_{k \to \infty} \mu(S \backslash C_k).$$

For this we define for every $k \in \mathbb{N}$

$$S_k := \left\{x \in S \colon \frac{1}{k+1} < \operatorname{dist}(x, C) \leq \frac{1}{k}\right\}$$

and we have $S \backslash C = (S \backslash C_k) \cup \bigcup_{j=k}^{\infty} S_j$ since C is closed. We estimate

$$\mu(S \backslash C) \leq \mu(S \backslash C_k) + \sum_{j=k}^{\infty} \mu(S_j)$$

and thus it remains to show that that the series $\sum_{j=1}^{\infty} \mu(S_j)$ converges. Since for $j \geq i+2$

$$\operatorname{dist}(S_i, S_j) \geq \frac{1}{i+1} - \frac{1}{j} > 0,$$

we get by induction and by using the assumption that for all $N \in \mathbb{N}$

$$\sum_{i=1}^{N} \mu(S_{2i}) = \mu\left(\bigcup_{i=1}^{N} S_{2i}\right) \leq \mu(S) < \infty,$$
$$\sum_{i=1}^{N} \mu(S_{2i-1}) = \mu\left(\bigcup_{i=1}^{N} S_{2i-1}\right) \leq \mu(S) < \infty.$$

Therefore

$$\mu(S \backslash C) \leq \lim_{k \to \infty} \mu(S \backslash C_k) + \lim_{k \to \infty} \sum_{j=k}^{\infty} \mu(S_j) = \lim_{k \to \infty} \mu(S \backslash C_k) \leq \mu(S \backslash C)$$

which proves the theorem. □

Theorem 7.4. *Let μ be a Borel regular measure on the metric space (X, d) and assume that $X = \bigcup_{j=1}^{\infty} X_j$ with $\mu(X_j) < \infty$ and X_j open for all $j \in \mathbb{N}$. Then we have*

1. $\mu(A) = \inf_{A\subset U \text{ open}} \mu(U)$ *for all* $A \subset X$*.*

2. $\mu(A) = \sup_{C\subset A \text{ closed}} \mu(C)$ *for all* $A \subset X$ μ*-measurable.*

Proof. We show the result again in three steps.

1. Proof of 1. for $\mu(X) < \infty$.

   We claim that

   $$\mathcal{A} := \{A \subset X \colon A \text{ Borel, 1. holds}\}$$

   is the Borel σ-Algebra. Then it follows from Definition 6.1 that 1. is true for all $A \subset X$. We show that $\mathcal{A}$ is closed under countable intersections and unions. For this we let $A_j \in \mathcal{A}$ and hence there exist $U_j \supset A_j$ open with

   $$\mu(U_j \backslash A_j) = \mu(U_j) - \mu(A_j) < 2^{-j}\varepsilon,$$

   where we used that A_j is μ-measurable. It follows that

   $$\mu\left(\bigcap_{j=1}^{\infty} U_j\right) - \mu\left(\bigcap_{j=1}^{\infty} A_j\right) = \mu\left(\bigcap_{j=1}^{\infty} U_j \backslash \bigcap_{j=1}^{\infty} A_j\right) \le \mu\left(\bigcup_{j=1}^{\infty} (U_j\backslash A_j)\right) < \varepsilon.$$

   Since $\mu(X) < \infty$ we have $\lim_{k\to\infty} \mu\left(\bigcap_{j=1}^{\infty} U_j\right) = \mu\left(\bigcap_{j=1}^{\infty} U_j\right)$ and thus for every $\varepsilon > 0$ there exists a $k \in \mathbb{N}$ so that

   $$\mu\left(\bigcap_{j=1}^{k} U_j\right) < \mu\left(\bigcap_{j=1}^{\infty} A_j\right) + \varepsilon.$$

   and hence

   $$\bigcap_{j=1}^{\infty} A_j \in \mathcal{A}.$$

   Additionally, we have

   $$\begin{aligned} \mu\left(\bigcap_{j=1}^{\infty} A_j\right) &= \mu\left(\bigcup_{j=1}^{\infty} U_j\right) - \mu\left(\bigcup_{j=1}^{\infty} U_j \backslash \bigcup_{j=1}^{\infty} A_j\right) \\ &\ge \mu\left(\bigcup_{j=1}^{\infty} U_j\right) - \mu\left(\bigcup_{j=1}^{\infty} (U_j\backslash A_j)\right) \ge \mu\left(\bigcup_{j=1}^{\infty} U_j\right) - \varepsilon \end{aligned}$$

   and thus

   $$\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}.$$

Next we define the system $\mathcal{A}' := \{A \in \mathcal{A} \colon X \backslash A \in \mathcal{A}\}$ and we note that

$$\mathcal{A}' \subset \mathcal{A} \subset \text{Borel } \sigma\text{-algebra}.$$

The system $\mathcal{A}'$ is by definition closed under taking complements and it is also closed under taking countable unions since

$$X \backslash \bigcup_{j=1}^{\infty} A_j = \bigcap_{j=1}^{\infty} (X \backslash A_j).$$

Moreover, $\mathcal{A}'$ contains all open sets since for $C \subset X$ closed we have

$$C = \bigcap_{j=1}^{\infty} \{x \in X \colon \operatorname{dist}(x, C_j) > 1/j\} \in \mathcal{A}$$

and hence $\mathcal{A}'$ and $\mathcal{A}$ are the Borel σ-algebra.

2. Proof of 1. for $X = \bigcup_{j=1}^{\infty} X_j$ with X_j open and $\mu(X_j) < \infty$.

   It follows from Lemma 7.2 that $\mu \llcorner X_j$ is a Borel regular measure. Step 1 then implies that for all $\varepsilon > 0$ there exists $U_j \supset A$ open with

$$(\mu \llcorner X_j)(U_j) = (\mu \llcorner X_j)(A) + 2^{-j}\varepsilon$$

   for all $A \subset X$. We can assume that A is a Borel set since this does not change the measure as μ and $\mu \llcorner X_j$ are both Borel regular. Hence A is $\mu \llcorner X_j$-measurable by Lemma 7.2 and

$$(\mu \llcorner X_j)(U_j \backslash A) = (\mu \llcorner X_j)(U_j) - (\mu \llcorner X_j)(A) \leq 2^{-j}\varepsilon.$$

   This implies

$$\begin{aligned}
\mu\left(\bigcup_{j=1}^{\infty} (U_j \cap X_j)\right) &= \mu\left(\bigcup_{j=1}^{\infty} (U_j \cap X_j) \cap A\right) + \mu\left(\bigcup_{j=1}^{\infty} (U_j \cap X_j) \backslash A\right) \\
&= \mu(A) + \mu\left(\bigcup_{j=1}^{\infty} X_j \cap (U_j \backslash A)\right) \\
&\leq \mu(A) + \varepsilon
\end{aligned}$$

   and hence statement 1. is shown to be true

3. Proof of 2.

   We start again with the case $\mu(X) < \infty$ and we note that statement 1. implies that for all $\varepsilon > 0$ there exists $U \supset X \backslash A$ open with $\mu(U) < \mu(X \backslash A) + \varepsilon$. We note that $C := X\backslash U \subset A$ is closed and

$$\mu(C) = \mu(X\backslash U) = \mu(X) - \mu(U) > \mu(X) - \mu(X\backslash A) - \varepsilon = \mu(A) - \varepsilon,$$

where we used that A is μ-measurable in the last equality sign.

Now if $X = \bigcup_{j=1}^\infty X_j$ with X_j open and $\mu(X_j) < \infty$, there exists $C_j \subset A$ closed so that

$$(\mu \llcorner X_j)(C_j) \geq (\mu \llcorner X_j)(A) - \varepsilon.$$

Without loss of generality we can assume that $C_1 \subset C_2 \subset \ldots$. Therefore

$$\mu(A) = \lim_{j\to\infty} \mu(X_j \cap A) \leq \lim_{j\to\infty} \mu(C_j) + \varepsilon.$$

Altogether this finishes the proof of statement 2.

□

For the rest of this chapter we assume that the metric space (X, d) is σ-compact, i.e. for all $x \in X$ and $\rho > 0$ we have that $\overline{B_\rho(x)}$ is compact.

Lemma 7.5. *Let (X, d) be a σ-compact metric space and let μ be Radon measure on X. Then we have*

1. $\mu(A) = \inf_{A\subset U \text{ open}} \mu(U)$ *for all $A \subset X$.*

2. $\mu(A) = \sup_{K\subset A \text{ compact}} \mu(K)$ *for all $A \subset X$ which are μ-measurable.*

Proof. Since μ is assumed to be a Radon measure we have that $\mu(\overline{B_\rho(x)}) < \infty$ for all $x \in X$ and all $\rho > 0$. Moreover, $X = \bigcup_{j=1}^\infty B_j(0)$ and hence statement 1. follows from statement 1. of Theorem 7.4. Next, for every closed set $C \subset X$ and every $j \in \mathbb{N}$ the set $C \cap \overline{B_j(0)}$ is compact and

$$\lim_{j\to\infty} \mu(C \cap \overline{B_j(0)}) = \mu\left(\bigcup_{j=1}^\infty (C \cap \overline{B_j(0)})\right) = \mu(C),$$

thus statement 2. follows from statement 2. in Theorem 7.4. □

Now we come to the representation problem. For this we let μ be a Radon measure on (X, d) and we let $\eta\colon X \to \mathbb{R}^k$ be μ-measurable with $|\eta(x)| = 1$ for μ-a.e. $x \in X$.

We define the linear form

$$\phi\colon C_c^0(X,\mathbb{R}^k) \to \mathbb{R},\ \phi(f) = \int_X \langle f, \eta\rangle d\mu$$

and we note that

$$|\phi(f)| \le C(K)\|f\|_{C^0(X)} \text{ for } \operatorname{spt} f \subset K \tag{7.1}$$

with $C(K) = \mu(K)$. Hence, for every compact K the linear form ϕ is continuous on the subspace of functions $f \in C_c^0(X)$ whose support is contained in K. If X is compact, then ϕ is continuous on all of $C^0(X)$. We call a linear form ϕ satisfying (7.1) a linear functional on $C_c^0(X)$ and our goal is to show that every linear functional on $C_c^0(X)$ has an integral representation as above. The measure μ will be the following one.

Definition 7.6. *Let $\phi \in C_c^0(X,\mathbb{R}^k) \to \mathbb{R}$ be a linear functional. Define the associated* ***variational measure*** *$|\phi|\colon \mathcal{P}(X) \to [0,\infty]$ in two steps:*

1. *$|\phi|(U) = \sup\{\phi(f)\colon |f| \le 1,\ \operatorname{spt} f \subset U\}$ for $U \subset X$ open,*
2. *$|\phi|(E) = \inf\{|\phi|(U)\colon U \supset E,\ U \text{ open}\}$ for $E \subset X$ arbitrary.*

Note that the two steps are consistent since one has in step 1 that $|\phi|(U) \le |\phi|(V)$ for $U \subset V$.

Theorem 7.7. *The variational measure $|\phi|$ is a Radon measure.*

Proof. We split the proof into four steps.

1. $|\phi|$ is an outer measure.

   For $u = \emptyset$ only the function $f \equiv 0$ is admissible in the definition of $|\phi$ and hence $|\phi|(\emptyset) = 0$. Next we let U_j, $j \in \mathbb{N}$, be open and we let $f \in C_c^0(X)$ with $|f| \le 1$ and $\operatorname{spt} f \subset \bigcup_{j=1}^\infty U_j$. By Theorem 3.2 there exists $N \in \mathbb{N}$ so that

   $$\operatorname{spt} f \subset \bigcup_{j=1}^N U_j.$$

   Next we choose a partition of unity $\chi_j \in C_c^0(X,[0,1])$ with $\operatorname{spt}\chi_j \subset U_j$ and $\sum_{j=1}^N \chi_j = 1$ on $\operatorname{spt} f$ (see the remark after this theorem for the existence of such a partition of unity). For $f_j := \chi_j f \in C_c^0(X)$ we conclude $\operatorname{spt} f_j \subset U_j$,

$|f_j| \le 1$ and $f = \sum_{j=1}^N f_j$. Hence we get

$$\phi(f) = \sum_{j=1}^N \phi(f_j) \le \sum_{j=1}^N |\phi|(U_j) \le \sum_{j=1}^\infty |\phi|(U_j).$$

Taking the supremum over all such functions f, we have

$$|\phi| \left(\bigcup_{j=1}^\infty U_j \right) \le \sum_{j=1}^\infty |\phi|(U_j).$$

Now we let $E \subset \bigcup_{j=1}^\infty E_j$ with E, E_j arbitrary. For $\varepsilon > 0$ we choose open sets $U_j \supset E_j$ with $|\phi|(U_j) < |\phi|(E_j) + 2^{-j}\varepsilon$, which is possible by Definition 7.6. It follows that $E \subset \bigcup_{j=1}^\infty U_j$ and

$$|\phi|(E) \le |\phi| \left(\bigcup_{j=1}^\infty U_j \right) \le \sum_{j=1}^\infty |\phi|(U_j) < \sum_{j=1}^\infty |\phi|(E_j) + \varepsilon.$$

The subadditivity of $|\phi|$ follows from letting $\varepsilon \searrow 0$.

2. Borel sets are $|\phi|$-measurable.

   Let $A, B \subset X$ with $\operatorname{dist}(A, B) > 0$. By the Caratheodory criterion, Theorem 7.3, we have to show that

$$|\phi|(W) \ge |\phi|(A) + |\phi|(B), \ \forall W \supset (A \cup B) \text{ open}.$$

   For $\delta > 0$ sufficiently small the sets $U := B_\delta(A) \cap W$ and $V := B_\delta(B) \cap W$ are disjoint. Let $f, g \in C_c^0(X, \mathbb{R}^k)$ with $\operatorname{spt} f \subset U$, $\operatorname{spt} g \subset V$ and $|f| \le 1$, $|g| \le 1$. Then $\operatorname{spt}(f+g) \subset (\operatorname{spt} f \cup \operatorname{spt} g) \subset W$ and $|f+g| \le 1$ on X. This implies

$$\phi(f) + \phi(g) = \phi(f+g) \le |\phi|(W).$$

   Taking the supremum over all such functions f and g implies

$$|\phi|(U) + |\phi|(V) \le |\phi(W)|.$$

3. $|\phi$ is Borel regular

   For $E \subset X$ with $|\phi|(E) < \infty$ we choose $U_j \supset E$ open with $|\phi|(U_j) \le |\phi|(E) + \frac{1}{j}$ and without loss of generality we assume $U_1 \supset U_2 \supset \ldots$. Then the set $B := \bigcap_{j=1}^\infty U_j \supset E$ is a Borel set and

$$|\phi|(B) = \lim_{j \to \infty} |\phi|(U_j) \le |\phi|(E).$$

4. $|\phi|$ is a Radon measure

   Let $K \subset X$ be compact. Since X is assumed to be σ-compact there exists $U \supset K$ open with $\overline{U}$ compact and it follows from (7.1) that

$$|\phi|(K) \leq |\phi|(U) \leq C(\overline{U}) < \infty.$$

□

Remark on the existence of the partition of unity:
The partition of unity used in the above proof can be constructed as follows: Let $K \subset X$ be compact and assume that $K \subset \bigcup_{\lambda \in \Lambda} U_\lambda$ where U_λ is open for all $\lambda \in \Lambda$. For $x \in K$ choose $r(x) > 0$ and $\lambda(x) \in \Lambda$ so that

$$\overline{B_{2r(x)}(x)} \subset U_{\lambda(x)}.$$

By Theorem 3.2 it follows that there exist points $x_1, \ldots, x_N$ with $K \subset \bigcup_{j=1}^N B_{r_j}(x_j)$ with $r_j := r(x_j)$. Next we choose a non-negative function $\tilde{x}_j \in C_c^0(X)$ with

$$\tilde{\chi}_j = \left\{ \begin{array}{ll} 1, & \text{on } B_{r_j}(x_j) \\ 0, & \text{on } X \backslash B_{2r_j}(x_j) \end{array} \right\}$$

and we define $\chi := \sum_{j=1}^N \tilde{\chi}_j$. Then we have $\chi \geq 1$ on K, hence $\chi > 1/2$ on an open neighbourhood of K. Now we choose $\eta \in C_c^0(X)$,with spt$\eta \subset U$ and $\eta|_K = 1$ and we finally define

$$\chi_j := \frac{\eta \tilde{\chi}_j}{\chi} \in C_c^0(X).$$

By construction we have spt$\chi_j \subset \overline{B_{2r_j}(x_j)} \subset U_{\lambda(x_j)}$ and $\sum_{j=1}^N \chi_j = 1$ on K.

Before we come to the representation theorem we need the following result of Lusin.

Theorem 7.8 (Lusin)**.** *Let μ be a Radon measure on the metric space (X, d) and let $A \subset X$ with $\mu(A) < \infty$ and let $\varepsilon > 0$. If the function $g \colon X \to \mathbb{R}$ is μ-measurable, then there exists another function $\tilde{g} \in C^0(X)$ so that*

$$\mu(\{x \in A \colon \tilde{g}(x) \neq g(x)\}) < \varepsilon \quad \textit{and} \quad \|\tilde{g}\|_{C^0(X)} \leq \sup_{x \in A} |g(x)|.$$

Proof. For $j \in \mathbb{N}$, $k \in \mathbb{Z}$ we consider the sets

$$A_{j,k} = \{x \in A \colon \frac{k}{j} \leq f(x) \leq \frac{k+1}{j}\}.$$

Since we know from Lemma 7.2 that $\mu \llcorner A$ is a Radon measure, we get that for every $\varepsilon > 0$ there exist compact sets $K_{j,k} \subset A_{j,k}$ with $\mu(A_{j,k} \backslash K_{j,k}) < 2^{-j-|k|}\varepsilon/3$. Therefore

$$\lim_{N\to\infty} \mu\left(A \backslash \bigcup_{k=-N}^{N} K_{j,k}\right) = \mu\left(A \backslash \bigcup_{k=-\infty}^{\infty} K_{j,k}\right) < \sum_{k=-\infty}^{\infty} 2^{-j-|k|}\varepsilon/3 < 2^{-j}\varepsilon.$$

For N_j sufficiently large we then also get $\mu(K_j) < 2^{-j}\varepsilon$, where $K_j = \bigcup_{k=-N_j}^{N_j} K_{j,k}$. Letting $K = \bigcap_{j=1}^{\infty} K_j$ we conclude

$$\mu(A \backslash K) \leq \mu\left(\bigcup_{j=1}^{\infty} A \backslash K_j\right) \leq \sum_{j=1}^{\infty} \mu(A \backslash K_j) < \varepsilon.$$

Next we consider the functions $f_j : A \to \mathbb{R}$, $f_j(x) = \frac{k}{j}$ for $x \in A_{j,k}$. Since the sets $K_{j,k} \subset A_{j,k}$ are compact they have positive distance from each other and thus f_j is locally constant and hence continuous on $K_j \supset K$. Moreover, for all $x \in A_{j,k}$ we have $|f(x) - f_j(x)| \leq \frac{1}{j} \to 0$ as $j \to \infty$. Therefore $f|_K$ is the uniform limit of a sequence of continuous functions and hence continuous itself. The result now follows from the Tietze extension theorem, which says that, on every metric space X a continuous function on a closed set $C \subset X$, $f : C \to \mathbb{R}$, can be extended to a function $\tilde{f} \in C^0(X)$ with $\|\tilde{f}\|_{C^0(X)} = \sup_{x\in C} |f(x)|$. □

Theorem 7.9 (Representation Theorem of Riesz-Markow-Radon)**.** *Let (X, d) be a σ-compact metric space. Then for every linear functional ϕ on $C^0_c(X, \mathbb{R}^k)$ there exists a Radon measure μ and a μ-measurable function $\eta\colon X \to \mathbb{R}^k$ with $|\eta| = 1$ μ-a.e., so that*

$$\phi(f) = \int_X \langle f, \eta\rangle d\mu \;\forall f \in C^0_c(X, \mathbb{R}^k).$$

The pair μ, η is uniquely determined and we have $\mu = |\phi|$.

Proof. We start with the proof of the uniqueness statement. For this we assume that there exists another Radon measure λ and a λ-measurable function $\xi\colon X \to \mathbb{R}^k$ with $|\xi = 1$ λ-a.e. and so that

$$\int_X \langle f, \eta\rangle d\mu = \int_X \langle f, \xi\rangle d\lambda$$

for all $f \in C^0_c(X)$. Let $U \subset X$ be open with $\mathrm{spt} f \subset U$ and $|f| \leq 1$. Then we have $\phi(f) \leq \lambda(U)$ and hence, by taking the supremum over such functions f, $\mu(U) = |\phi|(U) \leq \lambda(U)$ for all $U \subset X$ open. Statement 1 of Lemma 7.5 then implies that $\mu \leq \lambda$. Next we let $K \subset X$ be compact. For $U \supset K$ open with $\bar{U}$ compact and

$\varepsilon > 0$ we can use Lusin's theorem to obtain a function $\tilde{\xi} \in C^0(X, \mathbb{R}^k)$ such that

$$\lambda(E) < \varepsilon \quad \text{where} \quad E = \{x \in U : \tilde{\xi}(x) \neq \xi(x)\} \quad \text{and}$$
$$\|\tilde{\xi}\|_{C^0(X)} \leq \sup_{x \in U} |\xi(x)\| \leq 1.$$

Next we choose a function $\chi \in C^0_c(X)$ with $0 \leq \chi \leq 1$, spt$\chi \subset U$ and $\chi \equiv 1$ on K. We obtain

$$\begin{aligned}\mu(U) &\geq \phi(\chi\tilde{\xi}) \\ &= \int_U \langle \chi\tilde{\xi}, \xi \rangle d\lambda \\ &= \int_U \chi d\lambda + \int_U \chi(\langle \tilde{\xi}, \xi \rangle - 1) d\lambda \\ &\geq \lambda(K) - 2\lambda(E) \\ &\geq \lambda(K) - 2\varepsilon.\end{aligned}$$

Letting $\varepsilon \to 0$ and $U \searrow K$ we conclude $\mu(K) \geq \lambda(K)$ for all K compact and statement 2 of Lemma 7.5 then implies $\mu(B) \geq \lambda(B)$ for all $B \subset X$ Borel and thus $\mu = \lambda$ on Borel sets.

For $E \subset X$ arbitrary there exist Borel sets $B, B' \supset E$ with

$$\mu(B) = \mu(E),\ \lambda(B') = \lambda(E).$$

Without loss of generality we assume $B = B'$ since otherwise we look at $B \cap B'$, and this shows that $\lambda = \mu$.

Finally, we fix $v \in \mathbb{R}^k$ and for $f \in C^0_c(X)$ we have $fv \in C^0_c(X, \mathbb{R}^k)$. Hence

$$\int_X \langle fv, \eta \rangle d\mu = \int_X \langle fv, \xi \rangle d\mu,\ \forall f \in C^0_c(X).$$

By exercise sheet 7 we know that $C^0_c(X)$ is dense in $L^1(\mu)$ and therefeore

$$\int_X \langle fv, \eta \rangle d\mu = \int_X \langle fv, \xi \rangle d\mu,\ \forall f \in L^1(\mu),$$

which by Lemma 6.3 implies $\langle v, \xi \rangle = \langle v, \eta \rangle$ μ-almost everywhere. Inserting $v = e_1, ..., e_n$ we finally get $\xi = \eta$ μ-almost everywhere and this finishes the proof of the uniqueness part of the statement.

In order to show the existence part we let $\mu = |\phi|$. For $v \in \mathbb{R}^k$ with $|v| = 1$ we define

$$\phi_v : C^0_c(X) \to \mathbb{R},\ \phi_v(f) = \phi(fv).$$

Our goal is to show that ϕ_v can be extended as a linear functional to $L^1(\mu)$. In order to estimate ϕ_v we consider the functional

$$\varphi\colon C_c^0(X,\mathbb{R}_0^+)\to\mathbb{R}_0^+,\quad \varphi(f)=\sup\{\phi(g)\colon g\in C_c^0(X,\mathbb{R}^k),|g|\le f\}.$$

We claim that for $U\subset X$ open we have

$$\mu(U)=\sup\{\varphi(\chi)\colon \chi\in C_c^0(X,\mathbb{R}_0^+),\ \mathrm{spt}\chi\subset U,\chi\le 1\}. \tag{7.2}$$

This is true since for $g\in C_c^0(X,\mathbb{R}^k)$ with $\mathrm{spt}g\subset U$ and $|g|\le 1$ we have

$$\phi(g)\le\varphi(|g|)\le\sup\{\varphi(\chi)\colon \chi\in C_c^0(X,\mathbb{R}^k),\ \mathrm{spt}\chi\subset U,\ \chi\le 1\}.$$

and by taking the supremum over all such functions g it follows that

$$\mu(U)\le\sup\{\varphi(\chi)\colon \chi\in C_c^0(X,\mathbb{R}_0^+),\ \mathrm{spt}\chi\subset U,\ \chi\le 1\}$$

On the other hand, we let $\chi\in C_c^0(X,\mathbb{R}_0^+)$ with $\chi\le 1$ and $\mathrm{spt}\chi\subset U$ and obtain

$$\varphi(\chi)=\sup\{\phi(g)\colon g\in C_c^0(X,\mathbb{R}^k),\ |g|\le\chi\}\le\mu(U)$$

and these two inequalities imply the claim.

We split the rest of the proof into four steps.

Step 1: φ is a semilinear functional on $C_c^0(X)$, i.e.

(i) $\varphi(\alpha f)=\alpha\varphi(f)$ for all $f\in C_c^0(X,\mathbb{R}_0^+)$, $\alpha\ge 0$.

(ii) $\varphi(f_1+f_2)=\varphi(f_1)+\varphi(f_2)$, for all $f_1,f_2\in C_c^0(X,\mathbb{R}_0^+)$.

In order to show this we note that (i) follows directly from the definition of φ. For (ii) we choose functions $g_1,g_2\in C_c^0(X,\mathbb{R}^k)$ with $|g_i|\le f_i$ and $\phi(g_1)\ge\varphi(f_i)-\varepsilon$ It follows from an appropriate choice of the sign

$$\varphi(f_1)+\varphi(f_2)-2\varepsilon\le|\phi(g_1)|+|\phi(g_2)|=|\phi(g_1\pm g_2)|\le\varphi(f_1+f_2),$$

and with $\varepsilon\searrow 0$ we get $\varphi(f_1)+\varphi(f_2)\le\varphi)f_1+f_2)$. For the other inequality we let $g\in C_c^0(X,\mathbb{R}^k)$ with $|g|\le f_1+f_2$ and we define

$$g_i:=\begin{cases}\frac{f_i}{f_1+f_2}g, & f_1+f_2>0\\ 0, & f_1+f_2=0\end{cases}$$

It follows that $|g_i| \le f_i$ and in particular $g_i \in C_c^0(X, \mathbb{R}^k)$. Since $g = g_1 + g_2$ we get

$$|\phi(g)| \le |\phi(g_1)| + \phi(g_2)| \le \varphi(f_1) + \varphi(f_2)$$

and thus

$$\varphi(f_1 + f_2) \le \varphi(f_1) + \varphi(f_2).$$

Step 2: We have $\varphi(f) = \int_X f d\mu$ for all $f \in C_c^0(X, \mathbb{R}_0^+)$.

For $\varepsilon > 0$ we choose points $0 = t_0 < \ldots < t_N < \infty$ with $|t_i - t_{i-1}| < \varepsilon$, $t_N > \max f$ and

$$\mu(f^{-1}t_i) = 0, \ \forall 1 \le i \le N.$$

Note that this is possible since the set of all $t > 0$ so that $\mu(f^{-1}\{t\}) > 0$ is countable since $\mu(\text{spt} f) < \infty$. Since f is continuous we have that these sets $U_i := f^{-1}((t_{i-1}, t_i))$ are open.

Now we choose a function $\chi_i \in C_c^0(X, \mathbb{R}_0^+)$ with $\text{spt}\chi_i \subset U_i$ and $\chi_i \le 1$ for all $1 \le i \le N$. It follows that $\sum_{i=1}^N t_{i-1}\chi_i \le f$ and since φ is monotone by definition, we obtain from step 1

$$\sum_{i=1}^N t_{i-1}\varphi(\chi_i) = \varphi\left(\sum_{i=1}^N t_{i-1}\chi_i\right) \le \varphi(f).$$

Taking the supremum over the functions χ_i and using (7.2) we get $\sum_{i=1}^N t_{i-1}\mu(U_i) \le \varphi(f)$ and

$$\int_X f d\mu \le \sum_{i=1}^N t_i\mu(U_i) \le \sum_{i=1}^N (t_{i-1} + \varepsilon)\mu(U_i) \le \varphi(f) + \varepsilon\mu(\text{spt} f),$$

which as $\varepsilon \searrow 0$ implies $\int_X f d\mu \le \varphi(f)$.

In order to show the reverse inequality we choose $V_i \supset \overline{U}_i \cup f^{-1}\{t_i\}$ open for all $1 \le i \le N$ with

$$\mu(V_i) \le \mu(\overline{U}_i) + \varepsilon/N.$$

There exist $\chi_i \in C_c^0(V_i)$ with $\text{spt}\chi_i \subset V_i$, $0 \le \chi_i \le 1$ and $\chi_i \equiv 1$ on $\overline{U}_i \cup f^{-1}\{t_i\}$.

This gives $f \le \sum_{i=1}^N t_i\chi_i$ and then

$$\begin{aligned}
\varphi(f) &\le \sum_{i=1}^N t_i\varphi(\chi_i)\\
&\le \sum_{i=1}^N (t_{i-1}+\varepsilon)\mu(V_i)\\
&\le \sum_{i=2}^N (t_{i-1}+\varepsilon)\left(\mu(U_i)+\frac{\varepsilon}{N}\right)+\varepsilon\left(\mu(\overline{U}_1)+\frac{\varepsilon}{N}\right)\\
&\le \int_X f d\mu + 2\varepsilon\mu(\mathrm{spt} f)+\varepsilon\|f\|_{C^0}+C\varepsilon
\end{aligned}$$

where we used again (7.2). This shows the reverse inequality after letting $\varepsilon \searrow 0$.

Step 3: There exists a μ-measurable function $\eta : X \to \mathbb{R}^k$ with $\Phi(g) = \int_X \langle g, \eta\rangle d\mu$ for all $g \in C_c^0(X, \mathbb{R}^k)$.

For $f \in C_c^0(X)$ we have $\phi_v(f) = \phi_v(f^+) - \phi_v(f^-)$ and hence by step 2

$$|\phi_v(f)| \le |\phi_v(f^+)| + |\phi_v(f^-)| \le \varphi(f^+) + \varphi(f^-) = \int_X |f| d\mu.$$

It follows from the exercises that $C_c^0(X)$ is dense in $(L^1(\mu), \|\cdot\|_{L^1})$ and hence there exists a unique continuous extension $\phi_v \colon L^1(\mu) \to \mathbb{R}$ with $|\phi_v(f)| \le \|f\|_{L^1(\mu)}$, i.e. $\phi_v \in L^1(\mu)'$. By Theorem 6.4 there exists a function $\eta_v \in L^\infty(\mu)$ so that for all $f \in L^1(\mu)$ we have

$$\phi_v(f) = \int_f \eta_v d\mu.$$

If we let $v = e_i$, $1 \le i \le n$, and if we define

$$\eta\colon X \to \mathbb{R}^k, \quad \eta = \sum_{i=1}^k \eta_{e_i} e_i,$$

then we obtain for all $g \in C_c^0(X, \mathbb{R}^k)$

$$\begin{aligned}
\phi(g) = \sum_{i=1}^k \phi_{e_i}(g_i) &= \sum_{i=1}^k \int_X g_i\eta_{e_i} d\mu\\
&= \int_X \langle g, \eta\rangle.
\end{aligned}$$

Step 4: $|\eta| = 1$ μ-a.e.

Let $U \subset X$ be open with $\mu(U) < \infty$. Then

$$\mu(U) = \sup\{\phi(g) \colon g \in C^0_c(X, \mathbb{R}^k),\ \mathrm{spt} g \subset U, |g| \leq 1\} \leq \int_U |\eta| d\mu.$$

We define the μ-measurable function $\tilde{\eta} : X \to \mathbb{R}$ by

$$\tilde{\eta}(x) := \begin{cases} \frac{\eta(x)}{|\eta(x)|}, & \eta(x) \neq 0 \\ 0, & \eta(x) = 0, \end{cases}$$

and we note that Lusin's theorem implies that there exists $K \subset U$ compact with $\mu(U \backslash K) < \varepsilon$ and a function $\tilde{\tilde{\eta}} \in C^0_c(X, \mathbb{R}^k)$ with $\tilde{\tilde{\eta}}|_K = \tilde{\eta}$. We choose $\chi \in C^0_c(U)$ with $\chi \equiv 1$ on K, $0 \leq \chi \leq 1$ and we define $g := \chi \tilde{\tilde{\eta}}$ and we note that $|g| \leq 1$ and $\mathrm{spt} g \subset U$. Therefore

$$\begin{aligned} \mu(U) \geq \phi(g) &= \int_U \langle g, \eta \rangle d\mu = \int_K \langle g, \eta \rangle d\mu + \int_{U \backslash K} \langle g, \eta \rangle d\mu \\ &= \int_K |\eta| d\mu + \int_{U \backslash K} \langle g, \eta \rangle d\mu \\ &= \int_U |\eta| d\mu - \int_{U \backslash K} |\eta| d\mu + \int_{U \backslash K} \langle g, \eta \rangle d\mu \\ &\to \int_U |\eta| d\mu \end{aligned}$$

as $\varepsilon \searrow 0$. Hence

$$\mu(U) = \int_U |\eta| d\mu$$

for all $U \subset X$ open.

Next we let $A \subset X$ be a Borel set. Then it follows from Lemma 7.5 that there exists an open set $U \supset A$ with $\mu(U \backslash A) < \varepsilon$. Now

$$\left| \mu(A) - \int_A |\eta| d\mu \right| \leq \mu(U) - \mu(A) + \left| \mu(U) - \int_U |\eta| d\mu \right| + \int_U |\eta| d\mu - \int_A |\eta| d\mu \leq c\varepsilon$$

and therefore $\mu(A) = \int_A |\eta| d\mu$ for all Borel sets $A \subset X$.

Now we assume that there exists a set $E \subset X$ with $0 < \mu(E) < \infty$ and so that $|\eta| > 1$ on E. By Lusin's theorem there exist a compact set $K \subset E$ and a function $\hat{\eta} \in C^0(X, \mathbb{R}^k)$ with

$$\mu(E \backslash K) < \varepsilon \quad \text{and} \quad \hat{\eta}|_K = \eta$$

which implies $|\hat{\eta}| > 1$ on K. But then we conclude

$$0 < \int_K (|\hat{\eta}| - 1) d\mu = \int_K |\eta| d\mu - \mu(K) = 0,$$

which gives a contradiction. The same argument shows that there can't exists a set $E \subset X$ with $0 < \mu(E) < \infty$ and $|\eta| < 1$ on E. Since

$$X = \bigcup_{i=1}^{\infty} \overline{B_i(0)} \quad \text{with} \quad \mu(\overline{B_i(0)}) < \infty$$

it follows that $|\eta| = 1$ μ-a.e. $\square$

8 Weak convergence

Definition 8.1. *Let X be a Banach space.*

1. *A sequence x_k converges* ***weakly*** *to x in X ($x_k \rightharpoonup x$) if $\varphi(x_k) \to \varphi(x)$ for all $\varphi \in X'$.*

2. *A sequence φ_k converges* ***weakly*** *$*$ to φ in X' ($\varphi_k \overset{*}{\rightharpoonup} \varphi$) if $\varphi_k(x) \to \varphi(x)$ for all $x \in X$.*

Example 1.
1. *$f_k \rightharpoonup f$ in $L^p(\mu)$ for $1 \leq p < \infty \Leftrightarrow \int f_k g d\mu \to \int f g d\mu$ for all $g \in L^q(\mu)$ with $\frac{1}{p} + \frac{1}{q} = 1$.*

2. *$f_k \overset{*}{\rightharpoonup} f$ in $L^p(\mu) = L^q(\mu)'$ for $1 < p \leq \infty \Leftrightarrow \int f_k g d\mu \to \int f g d\mu$ for all $g \in L^q(\mu)$ with $\frac{1}{p} + \frac{1}{q} = 1$.*

Therefore weak and weak $*$ convergence agree if $1 < p < \infty$.

Theorem 8.2. *Let X and Y be Banach spaces. Then the following statements hold*

1. *Weak and weak $*$ limits are unique.*

2. *Norm convergence implies weak convergence.*

3. *Weakly and weakly $*$ converging sequences are bounded.*

4. *If $x_k \rightharpoonup x$ then $\|x\| \leq \liminf_{k\to\infty} \|x_k\|$ and if $\varphi_k \overset{*}{\rightharpoonup} \varphi$ then $\|\varphi\| \leq \liminf_{k\to\infty} \|\varphi_k\|$.*

5. *If $x_k \to x$ and $\varphi_k \overset{*}{\rightharpoonup} \varphi$ then $\varphi_k(x_k) \to \varphi(x)$. The same is true if $x_k \rightharpoonup x$ and $\varphi_k \to \varphi$.*

6. *For $T \in L(X, Y)$ and $x_k \rightharpoonup x$ in X it follows that $Tx_k \rightharpoonup Tx$ in Y.*

Proof.
1. The statement is clear for weak $*$ convergence since this is just pointwise convergence. Now let $x_k \rightharpoonup x$ and $x_k \rightharpoonup y$. For all $\varphi \in X'$ this implies

$$\varphi(x - y) = \varphi(x) - \varphi(y) = \lim \varphi(x_k) - \lim \varphi(x_k) = 0.$$

It follows from Lemma 4.4 that $x = y$.

2. Since $|\varphi(x_k) - \varphi(x)| \leq \|\varphi\|\|x_k - x\| \to 0$ for all $\varphi \in X'$, we get that norm convergence implies weak convergence. Similarly

$$|\varphi(x) - \varphi_k(x)\| \leq \|\varphi - \varphi_k\|\|x\| \to 0$$

for all $x \in X$.

3. The statement for the weak * convergence follows directly from Lemma 5.8. Now we let $x_k \rightharpoonup x$ in X and we let $J : X \to X''$ be the canonical embedding. Then

$$(Jx_k)(\varphi) = \varphi(x_k) \to \varphi(x) = (Jx)(\varphi)$$

for all $\varphi \in X'$ and therefore $Jx_k \overset{*}{\rightharpoonup} Jx$ in $X'' = (X')'$. Thus Jx_k is uniformly bounded in X'' and since we have shown in Theorem 4.5 that J is an isometry, it follows that x_k is also uniformly bounded in X.

4. Let $\varphi_k \overset{*}{\rightharpoonup} \varphi$ in X'. By choosing a subsequence we can assume that $\lim\limits_{k\to\infty} \|\varphi_k\|$ exists and is finite. For $\|x\| \leq 1$ we obtain

$$|\varphi(x)| = \lim_{k\to\infty} |\varphi_k(x)| \leq \lim_{k\to\infty} \|\varphi\| = \liminf_{k\to\infty} \|\varphi_k\|.$$

For $x_k \rightharpoonup x$ in X we conclude as in the statement above that $Jx_k \overset{*}{\rightharpoonup} Jx$ in X'' and hence $\|Jx\| \leq \liminf \|Jx_k\|$. Again by Theorem 4.5 this shows that

$$\|x\| \leq \liminf \|x_k\|.$$

5. Assume that $x_k \rightharpoonup x$ in X and $\varphi_k \to \varphi$ in X'. Then we estimate

$$\begin{aligned} |\varphi_k(x_k) - \varphi(x)| &\leq |\varphi_k(x_k) - \varphi(x_k) + |\varphi(x_k) - \varphi(x)| \\ &\leq \|\varphi_k - \varphi\|\|x_k\| + |\varphi(x_k) - \varphi(x)| \to 0 \end{aligned}$$

since $\|x_k\|$ is uniformly bounded by statement 3. The other case follows similarly.

6. Let $\psi \in Y'$, $T \in L(X,Y)$ and assume that $x_k \rightharpoonup x$ in X. Then

$$\psi(Tx_k) = (\psi \circ T)(x_k) \to (\psi \circ T)(x) = \psi(Tx)$$

as $\psi \circ T \in X'$. Since this is true for all $\psi \in Y'$, we get $Tx_k \rightharpoonup Tx$ in Y.

□

Theorem 8.3. *Let X be a separable Banach space. Then every uniformly bounded sequence $\varphi_k \in X'$ has a subsequence which converges weakly $*$ to a $\varphi \in X'$, i.e. the set $\{\varphi \in X' \colon \|\varphi\| \leq R\}$ is weakly $*$ sequentially compact for every $R < \infty$.*

Proof. Let $\{x_n \colon n \in \mathbb{N}\}$ be dense in X. Let $\|\varphi_k\| \leq c < \infty$ for all $k \in \mathbb{N}$ and let $\{x_n \colon n \in \mathbb{N}\}$ be dense in X. Then we have for all $k, n \in \mathbb{N}$

$$|\varphi_k(x_n)| \leq \|\varphi_k\| \|x_n\| \leq c\|x_n\|$$

and a diagonal sequence argument as in the proof of the Arzela-Ascoli theorem then implies that there exists a subsequence (wlog φ_k itself) so that $\lim \varphi_k(x_n) = \varphi(x_n)$ for all $n \in \mathbb{N}$. Thus the map $\varphi(x) = \lim_{k\to\infty} \varphi_k(x)$ is well-defined on $Y = \operatorname{span}\{x_1, ...\}$. We have that $\varphi \in Y'$ since for all $x \in Y$ we have

$$|\varphi(x)| = \lim_{k\to\infty} |\varphi_k(x)| \leq c\|x\|.$$

It follows from the exercise sheet 1 that there exists a unique continuous extension $\varphi \in X'$ with $\|\varphi\| \leq c$. It remains to show that $\varphi_k(x) \to \varphi(x)$ for all $x \in X$, but this follows from the estimate

$$|\varphi(x) - \varphi_k(x)| \leq |\varphi(x) - \varphi(x_k)| + |\varphi(x_n) - \varphi_k(x_n)| + |\varphi_k(x_n) - \varphi_k(x)| \leq \varepsilon,$$

where we first choose $n \in \mathbb{N}$ so that $\|x - x_n\| < \varepsilon/(3c)$ and then we choose $k_0 \in \mathbb{N}$ such that $|\varphi(x_n) - \varphi_k(x_n)| \leq \varepsilon/3$ for all $k \geq k_0$. $\square$

Next we recall that a Banach space X is called reflexive if the canonical map $J_X \colon X \to X''$, $(J_X x)(\varphi) = \varphi(x)$, for all $\varphi \in X'$, is surjective. It will follow from the results below and in the next chapter that all Hilbert spaces and all L^p-spaces with $1 < p < \infty$ are reflexive.

Lemma 8.4. *Let X be a Banach space.*

1. *If X is reflexive, then every closed subspace $Y \subset X$ is also reflexive.*

2. *X is reflexive if and only if X' is reflexive.*

Proof. 1. Let $\lambda \in Y''$ and define $\Lambda \colon X' \to \mathbb{R}$ by $\Lambda(\varphi) = \lambda(\varphi|_Y)$ for all $\varphi \in X'$. Since

$$|\Lambda(\varphi)| = |\lambda(\varphi|_Y)| \leq \|\lambda\| \|\varphi|_Y\| \leq \|\lambda\| \|\varphi\|$$

it follows that $\Lambda \in X''$. By the assumption of the lemma there exists $x \in X$ so that $\Lambda = J_X x$. Next we assume $x \notin Y$, i.e. $\operatorname{dist}(x, y) > 0$. By Theorem 4.3

there exists $\varphi \in X'$ with $\varphi|_Y = 0$ and $\varphi(x) = \operatorname{dist}(x, Y) > 0$ and we get the contradiction

$$0 = \lambda(\varphi|_Y) = \Lambda(\varphi) = (J_X x)(\varphi) = \varphi(x) > 0.$$

Hence $x \in Y$ and for an arbitrary $\varphi \in Y'$ we use Theorem 4.2 in order to obtain $\tilde{\varphi} \in X'$ with $\tilde{\varphi}|_Y = \varphi$. We get

$$(J_Y x)(\varphi) = \varphi(x) = \tilde{\varphi}(x) = (J_X x)(\tilde{\varphi}) = \Lambda(\tilde{\varphi}) = \lambda(\tilde{\varphi}|_Y) = \lambda(\varphi)$$

and therefore $J_Y x = \lambda$.

2. We start with the implication "$\Rightarrow$". For this we let $\phi \in X'''$ and we note that $\phi \circ J_X \in X'$. For $\Lambda \in X''$ we calculate

$$\begin{aligned}(J_{X'}(\phi \circ J_X))(\Lambda) =& \Lambda(\phi \circ J_X)\\ =& (J_X(J_X^{-1}\Lambda))(\phi \circ J_X)\\ =& (\phi \circ J_X)(J_X^{-1}\Lambda) = \phi(\Lambda),\end{aligned}$$

where we used that J_X is invertible by Theorem 4.5 since X is assumed to be reflexive. Therefore we conclude $J_{X'}(\phi \circ J_X) = \phi$.

Next we show the implication "$\Leftarrow$". From what we just proved it follows that if X' reflexive then also $X'' = (X')'$ is reflexive and we claim that $J_X(X)$ is a closed subspace of X''. For this we let $J_X x_k$ be a Cauchy sequence in $J_X(X)$ and since J_X is an isometry this implies that x_k is a Cauchy sequence in the Banach space X. Hence $x_k \to x \in X$ and $J_X x_k \to J_X x$. Thus it follows from the first part of this lemma that $J_X(X)$ is reflexive. It follows that X is reflexive since we have the general

Fact: Let X and Y be Banach spaces and let $T \in L(X, Y)$ be invertible. Then

$$X \text{ is reflexive } \Leftrightarrow Y \text{ is reflexive.}$$

In order to show this we look at the following diagram

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ {\scriptstyle J_X}\downarrow & & \downarrow{\scriptstyle J_Y} \\ X'' & \xrightarrow{T''} & Y'' \end{array}$$

For $\Lambda \in X''$ and $\psi \in Y'$ we define

$$(T''\Lambda)(\psi) = \Lambda(\psi \circ T) \in X'.$$

Moreover, for $\tilde{\Lambda} \in Y''$ and $\varphi \in X'$, we define the map $\tilde{T''}\colon Y'' \to x''$ by

$$(\tilde{T''}\tilde{\Lambda})\varphi = \tilde{\Lambda}(\varphi \circ T^{-1}).$$

We calculate

$$(T'' \circ \tilde{T''})(\tilde{\Lambda})(\psi) = T''(\tilde{T''}(\tilde{\Lambda}))(\psi) = (\tilde{T''}\tilde{\Lambda})(\psi \circ T) = \tilde{\Lambda}(\psi)$$

and thus $T'' \circ \tilde{T''} = \mathrm{id}_{Y''}$. Similarly, one shows that $\tilde{T''} \circ T'' = \mathrm{id}_{X''}$ and hence $T'' \in L(X'', Y'')$ is invertible.

Now we assume that X is reflexive, i.e. J_X is surjective, and we claim that $J_Y = T'' \circ J_X \circ T^{-1}$, which implies that also J_Y is surjective and Y is reflexive. In order to see this we let $y \in Y$, $\psi \in Y'$ and we calculate

$$\begin{aligned}(T''J_XT^{-1}y)(\psi) =&(J_XT^{-1}y)(\psi \circ T)\\ =&(\psi \circ T)(T^{-1}y)\\ =&\psi(y) = (J_Y y)(\psi).\end{aligned}$$

□

Theorem 8.5. *Let X be a reflexive Banach space. Then every bounded sequence in X has a weakly converging subsequence in X, i.e. the set $\{x \in X\colon \|x\| \le R\}$ is weakly sequentially compact for all $R < \infty$.*

Proof. Let (x_k) be a sequence in X with $\|x_k\| \le c$ for all $k \in \mathbb{N}$ and define the space $Y := \overline{\operatorname{span}\{x_k\colon k \in \mathbb{N}\}}$. By definition Y is separable and closed. Hence it follows from Lemma 8.4 that Y is reflexive and moreover $Y'' = J_Y Y$ is separable, which, by Theorem 4.6, implies that Y' is separable. Now $(J_Y x_k)_{k\in\mathbb{N}} \subset Y'' = (Y')'$ with $\|J_Y x_k\| \le c$ and by Theorem 8.3 there exists $\Lambda \in Y''$, so that for all $\varphi \in Y'$

$$\varphi(x_k) = (J_Y x_k)(\varphi) \to \Lambda(\varphi)$$

up to a subsequence. Define $x := J_Y^{-1}(\Lambda) \in Y$ and observe that for all $\varphi \in Y'$

$$\varphi(x_k) \to \Lambda(\varphi) = (J_Y x)(\varphi) = \varphi(x).$$

It remains to show that this convergence property holds for all $\varphi \in X'$:

$$\varphi(x_k) = (\varphi|_Y)(x_k) \to (\varphi|_Y)(x) = \varphi(x).$$

□

Lemma 8.6. *The space $L^p(\mu)$ is reflexive for $1 < p < \infty$.*

Proof. Let $\phi \in L^p(\mu)''$ and thus $\phi\colon L^p(\mu)' \to \mathbb{R}$ is linear and bounded. Let

$$J_q : L^q(\mu) \to L^p(\mu)'$$

be the duality map of Theorem 6.4 and hence $\phi \circ J_q \in L^q(\mu)'$. By Theorem 6.4 there exists $f \in L^p(\mu)$ such that for all $g \in L^q(\mu)$

$$(\phi \circ J_q)(g) = \int_X fg d\mu.$$

Let $J_{L^p}\colon L^p(\mu) \to L^p(\mu)''$ be the canonical embedding and note that for every $g \in L^q(\mu)$

$$(J_{L^p} f)(J_q g) = (J_q g)(f) = \int_X fg d\mu = (\phi \circ J_q)(g) = \phi(J_q g).$$

As J_q is surjective, Theorem 6.4 gives us $J_{L^p} f = \phi$. □

Application: Assume that $\|u_k\|_{L^p} \le C$ for all $k \in \mathbb{N}$.

Case 1: If $1 < p < \infty$ then $L^p(\mu)$ reflexive and Theorem 8.5 implies that there exists a subsequence $u_k \rightharpoonup u$ in $L^p(\mu) \Leftrightarrow \int_X u_k v d\mu \to \int_X uv d\mu$ for all $v \in L^q(\mu)$ with $p^{-1} + q^{-1} = 1$.

Case 2: If $p = \infty$ then $L^\infty = (L^1)'$, L^1 is separable and it follows from Theorem 8.3 that there exists a subsequence $u_k \overset{*}{\rightharpoonup} u$ in $L^\infty(\mu) \Leftrightarrow \int u_k v d\mu \to \int uv d\mu$ for all $v \in L^1(\mu)$.

Theorem 8.7. *Let X be a Banach space and let $K \subset X$ be convex and closed. Then for every sequence $x_k \in K$ with $x_k \rightharpoonup x \in X$ we have $x \in K$, i.e. K is weakly closed.*

Proof. Without loss of generality we can assume that $x = 0$. If $0 \notin K$, then $\operatorname{dist}(0, K) =: \rho > 0$ and by Lemma 4.12 there exists $\varphi \in X'$ with $\|\varphi\| = 1$ and $\varphi(y) \le -\rho$ for all $y \in K$ and we get the contradiction

$$0 = \varphi(0) = \lim_{k\to\infty} \varphi(x_k) \le -\rho < 0.$$

□

Lemma 8.8. *If X is a reflexive Banach space and $K \subset X$ is closed and convex,* *then for every* $x_0 \in X$ *there exists* $x \in K$ *with*

$$\|x - x_0\| = \operatorname{dist}(x_0, K).$$

Proof. Without loss of generality we can assume that $x_0 = 0$. By choosing a minimizing sequence $x_k \in K$, i.e. $\|x_k\| \to \operatorname{dist}(0, K)$, we can use Theorem 8.5 in order to conclude that $x_k \rightharpoonup x$ up to a subsequence and by Theorem 8.7 it follows that $x \in K$. Finally, by Theorem 8.2 we get

$$\operatorname{dist}(0, K) \leq \|x\| \leq \liminf \|x_k\| = \operatorname{dist}(0, K).$$

□

Definition 8.9. *A normed space* $(X, \|\cdot\|)$ *is called* ***uniformly convex*** *if for all* $\varepsilon > 0$ *there exists* $\delta > 0$ *so that the following implication is true*

$$\left(\|x\| = 1 = \|y\|, \left\|\frac{x+y}{2}\right\| \geq 1 - \delta\right) \Rightarrow \|x - y\| < \varepsilon, \forall x, y \in X.$$

We have seen in Lemma 6.5 that all L^p-spaces with $1 < p < \infty$ are uniformly convex. Moreover, all Hilbert spaces H are uniformly convex, since for all $x, y \in H$ with $\|x\|^2 = 1 = \|y\|^2$ and $\left\|\frac{x+y}{2}\right\| \geq 1 - \delta$ it follows that

$$4 = \|x + y\|^2 + \|x - y\|^2 \geq 4(1 - \delta)^2 + \|x - y\|^2$$

and hence $\|x - y\| \leq 2\sqrt{2\delta} =: \varepsilon/2$.

Lemma 8.10. *Let* $(X, \|\cdot\|)$ *be uniformly convex and let* x_n, y_n *be sequences in* X *with* $\limsup \|x_n\| \leq 1$, $\limsup \|y_n\| \leq 1$ *and*

$$\lim \left\|\frac{x_n + y_n}{2}\right\| = 1.$$

Then $\|x_n - y_n\| \to 0$.

Proof. We note that $\rho_n := \|x_n\| \to 1$, $\sigma_n := \|y_n\| \to 1$ and we define the new sequences $\xi_n := \frac{x_n}{\rho_n}$, $\eta_n := \frac{y_n}{\sigma_n}$. It follows that

$$\begin{aligned}\frac{1}{2}\|\xi_n + \eta_n\| =& \frac{1}{2}\left\|\frac{x_n}{\rho_n} + \frac{y_n}{\rho_n} - \frac{y_n}{\rho_n} + \frac{y_n}{\sigma_n}\right\| \\ \geq& \frac{1}{2}\left(\frac{1}{\rho_n}\|x_n + y_n\| - \left|\frac{1}{\sigma_n} - \frac{1}{\rho_n}\right| \|y_n\|\right) \to 1.\end{aligned}$$

Hence we can use the uniform convexity of X to get $\|\xi_n - \eta_n\| \to 0$ which shows that

$$\begin{aligned}\|x_n - y_n\| =& \|\rho_n \xi_n - \rho_n \eta_n + \rho_n \eta_n - \sigma_n \eta_n\| \\ =& |\rho_n| \|\xi_n - \eta_n\| + \|\eta_n\| |\rho_n - \sigma_n\| \to 0.\end{aligned}$$

□

Theorem 8.11. *Let $(X, \|\cdot\|)$ be uniformly convex. Then the following two statements are equivalent.*

(i) $x_n \to x$ *with respect to the norm* $\|\cdot\|$.

(ii) $x_n \rightharpoonup x$ *and* $\limsup_{n\to\infty} \|x_n\| \leq \|x\|$.

Proof. It is obvious that statement (i) implies statement (ii).

In order to show that (ii) implies (i) we assume without loss of generality that $x \neq 0$. Then we get that $\rho_n := \max\{\|x_n\|, \|x\|\} > 0$ and $\rho_n \to \|x\|$ by Theorem 8.2 and the assumption. Now

$$\xi_n := \frac{x_n}{\rho_n} \rightharpoonup \frac{x}{\|x\|} =: \xi$$

since $\varphi(x_n/\rho_n) = \frac{1}{\rho_n}\varphi(x_n) \to \frac{1}{\|x\|}\varphi(x) = \varphi(\xi)$ for all $\varphi \in X'$. Theorem 8.2 thus yields

$$1 = \|\xi\| = \|\xi_n\| \leq \liminf \left\| \frac{1}{2}(\xi + \xi_n) \right\|$$

and by Lemma 8.10 we conclude $\|\xi_n - \xi\| \to 0$ or equivalently

$$\|x_n - x\| \to 0.$$

□

9 Hilbert spaces

Recall that $(X, \langle \cdot, \cdot \rangle)$ is a Hilbert space if the norm $\|x\| = \sqrt{\langle x, x \rangle}$ is complete.

Theorem 9.1 (Projection Theorem)**.** *Let X be a Hilbert space and let $K \subset X$ be closed and convex. Then for every $x \in X$ there exists a unique $x_0 \in K$ with*

1. $\|x - x_0\| = \mathrm{dist}(x, K) = \inf\{\|x - z\| \colon z \in K\}$.

2. *Moreover, $Re\langle x - x_0, z - x_0 \rangle \leq 0$ for all $z \in K$ and x_0 is the unique point with this property.*

   We write $x_0 = P_K(x)$ with $P_K \colon X \to K$ and we call P_K the nearest point projection.

Proof. In order to show statement 1 we choose a minimising sequence $x_k \in K$ and hence

$$\|x - x_k\| \to d := \mathrm{dist}(x, K).$$

Next we use the parallelogram identity

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

in order to get

$$\|x_k - x_l\|^2 = 2(\|x_k - x\|^2 + \|x_l - x\|^2) - 4 \left\| \frac{x_k + x_l}{2} - x \right\|^2 \to 0$$

as $k, l \to \infty$, where we used that $\frac{x_k + x_l}{2}$ and therefore $\left\| \frac{x_k+x_l}{2} - x \right\|^2 \geq d^2$. Hence (x_k) is a Cauchy sequence and thus $x_k \to x_0$ with $x_0 \in K$ and $\|x - x_0\| = d$ since K is closed.

Now let $x_0, x_1 \in K$ be two such points as in statement 1. It follows again from the parallelogram identity that

$$\|x_0 - x_1\|^2 = 2(\|x_0 - x\|^2 + \|x_1 - x\|^2) - 4 \left\| \frac{x_0 - x_1}{2} - x \right\|^2 \leq 2(d^2 + d^2) - 4d^2 = 0$$

and thus $x_0 = x_1$.

In order to show statement 2 we let $z \in K$ and $0 \leq \varepsilon \leq 1$. It follows from statement 1 that

$$\begin{aligned} d^2 \leq & \|x - ((1-\varepsilon)x_0 + \varepsilon z)\|^2 = \|x - x_0 - \varepsilon(z - x_0)\|^2 \\ = & \|x - x_0\|^2 + \varepsilon^2\|z - x_0\|^2 - 2\varepsilon Re\langle x - x_0, z - x_0\rangle \\ = & d^2 + \varepsilon^2\|z - x_0\|^2 - 2\varepsilon Re\langle x - x_0, z - x_0\rangle \end{aligned}$$

Hence

$$Re\langle x - x_0, z - x_0\rangle \leq \frac{\varepsilon}{2}\|z - x_0\|^2$$

and for $\varepsilon \searrow 0$ we get

$$Re\langle x - x_0, z - x_0\rangle \leq 0.$$

Finally, we let $\tilde{x}_0 \in K$ be another point so that $Re\langle x - \tilde{x}_0, y - \tilde{x}_0\rangle \leq 0$ for all $y \in K$. Inserting $y = x_0$ and $z = \tilde{x}_0$ in the above inequality for x_0 get

$$Re\langle x - x_0, \tilde{x}_0 - x_0\rangle \leq 0 \quad \text{and} \quad Re\langle x - \tilde{x}_0, x_0 - \tilde{x}_0\rangle \leq 0$$

Adding these two inequalities yields

$$\|\tilde{x}_0 - x_0\|^2 = Re\|\tilde{x}_0 - x_0\|^2 = Re\langle \tilde{x}_0 - x_0, \tilde{x}_0 - x_0\rangle \leq 0$$

and therefore

$$\tilde{x}_0 = x_0.$$

□

Lemma 9.2. *Let X be a Hilbert space and let $K \subset X$ be closed and convex. Then the nearest point projection $P_K\colon X \to K$ is Lipschitz continuous with constant $L = 1$.*

Proof. Let $x, y \in X$ with $x \neq y$, and recall from Theorem 9.1 that for all $z_1 \in K$ resp. $z_2 \in K$

$$\mathrm{Re}\,\langle x - P_K(x), z_1 - P_K(x)\rangle \leq 0 \quad \mathrm{Re}\,\langle y - P_K(y), z_2 - P_K(Y)\rangle \leq 0.$$

Now we insert $z_2 := P_K(x)$ and $z_1 := P_K(y)$ and we add the two inequalities in order to get

$$Re\langle y - P_K(y) + P_K(x) - x, P_K(x) - P_K(y)\rangle \leq 0.$$

Therefore

$$\|P_K(x) - P_K(y)\|^2 \leq Re\langle x - y, P_K(x) - P_K(y)\rangle \leq \|x - y\|\|P_K(x) - P_K(y)\|,$$

which implies

$$\|P_K(x) - P_K(y)\| \leq \|x - y\|.$$

□

Definition 9.3. *Let X be a Hilbert space and let $M \subset X$. Then the closed subspace*

$$M^\perp := \{x \in X \colon \langle x, y\rangle = 0 \; \forall y \in M\}$$

*is the **orthogonal complement** of M.*

Lemma 9.4. *Let X be a Hilbert space and let $Y \subset X$ be a closed subspace. Then $P_Y \colon X \to Y$ is linear and $\ker(P_y) = Y^\perp$. Moreover, we have that $X = Y \oplus Y^\perp$.*

Proof. Recall from Theorem 9.1 that $P_Y(x) = x_0$ if and only if $Re\langle x - x_0, z - x_0\rangle \leq 0$ for all $z \in Y$. Now we replace $z \in Y$ by $x_0 + \lambda y \in Y$ for all $y \in Y$ and $\lambda \in \mathbb{K}$ This yields

$$Re\langle x - x_0, \lambda y\rangle \leq 0$$

for all $\lambda \in \mathbb{K}$, $y \in Y$. Choosing $\lambda = \langle x - x_0, y\rangle$ then gives that $P_Y(x) = x_0$ if and only if

$$\langle x - x_0, y\rangle = 0$$

for all $y \in Y$.

Next, we let $P_Y(x) = x_0$, $P_Y(z) = z_0$, $\lambda, \mu \in \mathbb{K}$ and we note that $\lambda x_0 + \mu z_0 \in Y$. Moreover,

$$\langle(\lambda x + \mu z) - (\lambda x_0 + \mu z_0), y\rangle = 0$$

for all $y \in Y$ and hence

$$P_Y(\lambda x + \mu z) = \lambda x_0 + \mu z_0 = \lambda P_Y(x) + \mu P_Y(z)$$

which implies that P_Y is linear and continuous with norm 1 (see Lemma 9.2).

We also not that $P_Y(x) = 0$ if and only if $\langle x, y\rangle = 0$ for all $y \in Y$ and thus if and only if $x \in Y^\perp$.

Finally, we let $x \in X$ be arbitrary and we note that

$$P_Y(x - P_Y(x)) = P_Y(x) - P_Y^2(x) = 0$$

since $P_Y|_Y = \mathrm{id}$ and therefore $x = P_Y(x) + (x - P_Y(x))$ with $P_Y(x) \in Y$, $(x - P_Y(x)) \in Y^\perp$ and hence $X = Y \oplus Y^\perp$. □

Theorem 9.5 (Riesz representation for Hilbert spaces)**.** *Let X be a Hilbert space.*

Then the map

$$R\colon X \to X',\ (Ry)(x) = \langle x, y\rangle$$

is a surjective, conjugate linear isometry.

Proof. 1. R is surjective.

Let $\varphi \in X'$ with $\varphi \neq 0$. Then $\ker\varphi = \varphi^{-1}(0)$ is a closed subspace and Lemma 9.4 implies that there exists $v \in (\ker(\varphi))^{\perp}$ with $\|v\| = 1$ and $\varphi(v) \neq 0$. For $x \in X$ we have

$$x = x - \frac{\varphi(x)}{\varphi(v)}v + \frac{\varphi(x)}{\varphi(v)}v \in (\ker\varphi) \oplus (\ker\varphi)^{\perp}$$

and when we take the scalar product with v we obtain

$$\langle x, v\rangle = \frac{\varphi(x)}{\varphi(v)}$$

and thus

$$\varphi(x) = \langle x, \overline{\varphi(v)}v\rangle,$$

i.e. $\varphi = R(\overline{\varphi(v)}v)$.

2. $R(y) \in X'$.

For this we not that for all $\alpha, \beta \in \mathbb{K}$ and $x_1, x_2, y \in X$ we have

$$\langle \alpha x_1 + \beta x_2, y\rangle = \alpha\langle x_1, y\rangle + \beta\langle x_2, y\rangle$$

and therefore

$$(Ry)(\alpha x_1 + \beta x_2) = \alpha(Ry)(x_1) + \beta(Ry)(x_2).$$

Moreover, we estimate with the help of the Cauchy-Schwarz inequality

$$|(Ry)(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$$

and hence $\|Ry\| \leq \|y\|$.

3. R is conjugate linear.

We again let $\alpha, \beta \in \mathbb{K}$ and $x, y_1, y_2 \in X$. Then

$$\begin{aligned}(R(\alpha y_1 + \beta y_2))(x) =& \langle x, \alpha y_1 + \beta y_2 \rangle \\ =& \bar{\alpha}\langle x, y_1 \rangle + \bar{\beta}\langle x, y_2 \rangle \\ =& \bar{\alpha}(Ry_1)(x) + \bar{\beta}(Ry_2)(x).\end{aligned}$$

4. R is an isometry.

   For $y \in X$ we choose $x = y/\|y\|$ and conclude

$$(Ry)(x) = \left\langle \frac{y}{\|y\|}, y \right\rangle = \|y\|.$$

   Together with step 2 of the proof this shows that

$$\|Ry\| = \|y\|.$$

□

Theorem 9.6. *Hilbert spaces are reflexive.*

Proof. Let X be a Hilbert space and let $x'' \in X''$. Define $\varphi \colon X \to \mathbb{K}$ by

$$\varphi(y) = \overline{x''(Ry)},$$

where R is as in Theorem 9.5. Then φ is linear and

$$|\varphi(y)| \leq \|x''\| \|Ry\| \leq \|x''\| \|y\|$$

and therefore $\|\varphi\| \leq \|x''\|$ which shows that $\varphi \in X'$. Next we define $x := R^{-1}\varphi \in X$ and it follows that

$$x''(Ry) = \overline{\varphi(y)} = \overline{\langle y, x \rangle} = \langle x, y \rangle = (Ry)(x) = (Jx)(Ry).$$

Thus $x'' = Jx$ and the canonical embedding J is surjective. □

Definition 9.7. *Let X, Y be Banach spaces and let $T \in L(X, Y)$. The (Banach)* ***adjoint*** *is defined by*

$$T' \colon Y' \to X', \ T'\psi := \psi \circ T.$$

We have that $\|T'\| \leq \|T\|$ since for all $\psi \in Y'$ with $\|\psi\| \leq 1$ we estimate

$$\begin{aligned} \|T'\psi\| = \sup_{\|x\|\leq 1} |T'\psi(x)| &= \sup_{\|x\|\leq 1} |\psi(Tx)| \\ &\leq \|\psi\| \sup_{\|x\|\leq 1} |Tx| \leq \|T\|. \end{aligned}$$

Theorem 9.8. *Let X, Y be Hilbert spaces over $\mathbb{K}$. For $T \in L(X,Y)$ there exists a unique map $T^* \in L(Y,X)$ with*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \ \forall x \in X, y \in Y.$$

for all $x \in X$ and $y \in Y$.

Proof. 1. Uniqueness:

Assume that there exist two maps $S_1, S_2 \in L(Y,X)$ with

$$\langle Tx, y\rangle = \langle x, S_1 y\rangle \quad \text{and} \quad \langle Tx, y\rangle = \langle x, S_2 y\rangle$$

for all $x \in X$ and $y \in Y$. Subtracting these equations yields

$$0 = \langle x, (S_1 - S_2)y\rangle$$

for all $x \in X$ and $y \in Y$ and hence

$$S_1 y = S_2 y$$

for all $y \in Y$, which shows that $S_1 \equiv S_2$.

2. Existence:

We look at the diagram

$$\begin{array}{ccc} X & \xrightarrow{\ T\ } & Y \\ {\scriptstyle R_X}\downarrow & & \downarrow{\scriptstyle R_Y} \\ X' & \xrightarrow[\ T'\ ]{} & Y' \end{array}$$

where $R_X : X \to X'$ and $R_Y : Y \to Y'$ are the maps from Theorem 9.5 and we claim that the map $T^* = R_X^{-1} T' R_Y \in L(Y,X)$ satisfies the desired property.

For this we calculate for all $x \in X$ and $y \in Y$

$$\begin{aligned}\langle x, T^* y\rangle =& \langle x, R_X^{-1} T' R_Y y\rangle \\ =& (T' R_Y y)(x) \\ =& (R_Y y)(Tx) \\ =& \langle Tx, y\rangle.\end{aligned}$$

□

It follows from Theorem 9.5 that and the remark after Definition 9.7 that $\|T^*\| \leq \|T\|$. Indeed, we have equality in this estimate (see the exercises).

Lemma 9.9. *Let $B\colon X \times X \to \mathbb{K}$ be a bounded sesquilinear form on a Hilbert space X, i.e. for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{K}$ we have*

$$\begin{aligned}\|B\| :=& \sup_{\|x\|,\|y\|\leq 1} |B(x,y)| < \infty, \\ B(\alpha x + \beta y, z) =& \alpha B(x,z) + \beta B(y,z), \\ B(z, \alpha x + \beta y) =& \bar{\alpha} B(z,x) + \bar{\beta} B(z,y).\end{aligned}$$

Then there exists a unique map $T = T_B \in L(X,X)$ with $\|T\| = \|B\|$ and so that for all $x, y \in X$

$$B(x,y) = \langle Tx, y\rangle \ \forall x, y \in X.$$

Proof. The uniqueness statement follows in the same way as we have shown uniqueness in the proof of Theorem 9.8.

In order to show the existence of T we let $x \in X$ be fixed and we define

$$\varphi_x(y) := \overline{B(x,y)}.$$

It follows from the properties of B that φ_x is linear, $\|\varphi_x\| \leq \|B\|\|x\| < \infty$ and hence $\varphi_x \in X'$. By Theorem 9.5 there exists a map $T\colon X \to X$, $T := R^{-1}(\varphi_x)$ with

$$\overline{B(x,y)} = \langle y, Tx\rangle$$

or equivalently

$$B(x,y) = \langle Tx, y\rangle.$$

For all $\alpha, \beta \in \mathbb{K}$ and $x, y, z \in X$ we have

$$\begin{aligned}\langle T(\alpha x + \beta y), z\rangle &= B(\alpha x + \beta y, z)\\ &= \alpha B(x, z) + \beta B(y, z)\\ &= \alpha\langle Tx, z\rangle + \beta\langle Ty, z\rangle\\ &= \langle \alpha Tx + \beta Ty, z\rangle\end{aligned}$$

which shows that T is linear. Choosing $y = Tx$ for all $x \in X$ with $\|x\| \le 1$, we get

$$\|Tx\|^2 = \langle Tx, Tx\rangle = B(x, Tx) \le \|B\|\|Tx\|$$

and therefore

$$\|T\| \le \|B\|.$$

The estimate $\|B\| \le \|T\|$ follows from

$$|B(x, y)| = |\langle Tx, y\rangle| \le \|Tx\|\|y\| \le \|T\|\|x\|\|y\|.$$

□

Theorem 9.10 (Lax-Milgram)**.** *Let $B\colon X \times X \to \mathbb{R}$ be a bounded bilinear form on the $\mathbb{R}$-Hilbert space X and let B be* ***coercive****, i.e.*

$$B(x, x) \ge \lambda\|x\|^2$$

for all $x \in X$ and for some $\lambda > 0$. Then the map $R_B\colon X \to X'$, $(R_B y)(x) = B(x, y)$ is invertible with

$$\|R_B^{-1}\| \le \frac{1}{\lambda}.$$

Additionally, if B is symmetric, then $R_B^{-1}\varphi = y_0$ is the unique minimum of the quadratic functional

$$Q\colon X \to \mathbb{R},\ Q(y) = \frac{1}{2}B(y, y) - \varphi(y).$$

Proof. It follows from Theorem 9.8 and Lemma 9.9 that there exists a map $S \in L(X, X)$ so that $B(x, y) = \langle x, Sy\rangle$ for all $x, y \in X$.

1. S is injective and $im(S)$ is closed.

   For $x \neq 0$ it follows from the coercivity of B that

   $$0 < \lambda\|x\|^2 \le B(x, x) = \langle x, Sx\rangle \le \|x\|\|Sx\|$$

which implies

$$\|Sx\| \geq \lambda \|x\| \tag{9.1}$$

and hence S is injective.

Next, we let $y_k = Sx_k \in im(S)$ be a Cauchy sequence. It follows from (9.1) that $\|x_k - x_l\| \leq \lambda^{-1}\|y_k - y_l\| \to 0$ as $k, l \to \infty$ and hence x_k is also a Cauchy sequence. Therefore $x_k \to x$ and $Sx_k \to Sx$ which shows that $im(S)$ is closed.

2. S is surjective.

   Let $y \in (im(S))^\perp$. Then we conclude

   $$\lambda \|y\|^2 \leq B(y, y) = \langle y, Sy \rangle = 0.$$

   Hence $(im(S))^\perp = \{0\}$ and it follows from Lemma 9.4 that $X = im(S) \oplus (im(S))^\perp = im(S)$.

3. R_B is invertible and $\|R_B^{-1}\| \leq \lambda^{-1}$.

   For this we show that $R_B = RS$, where $R : X \to X'$ is the map from Theorem 8.5, and then R_B is the composition of two invertible maps. For every $x, y \in X$ we have

   $$(RSy)(x) = \langle x, Sy \rangle = B(x, y) = (R_B y)(x).$$

   Moreover, we have that

   $$\|R_B^{-1}\| = \|S^{-1}R^{-1}\| = \|S^{-1}\| \leq \frac{1}{\lambda}$$

   where we have used (9.1) and the fact that R is an isometry. Note that as S is surjective, for every $y \in X$ there exists $x \in X$ so that $Sx = y$ and hence one applies (9.1) with $x = S^{-1}y$ in order to get $\|S^{-1}y\| \leq \frac{1}{\lambda}\|y\|$.

4. The minimising property.

   Now we assume that B is symmetric and we let $y_0 = R_B^{-1}\varphi$ for $\varphi \in X'$.

Moreover, we let $y = y_0 + \eta \in X$. Then we calculate

$$\begin{aligned}Q(y) =& \frac{1}{2}B(y_0+\eta, y_0+\eta) - \varphi(y_0+\eta)\\ =& \frac{1}{2}B(y_0,y_0) + B(y_0,\eta) + \frac{1}{2}B(\eta,\eta) - \varphi(y_0) - \varphi(\eta)\\ =& Q(y_0) + \frac{1}{2}B(\eta,\eta) + B(\eta,y_0) - \varphi(\eta)\\ \geq& Q(y_0) + \frac{\lambda}{2}\|\eta\|^2,\end{aligned}$$

where we used that $\varphi(\eta) = (R_B y_0)(\eta) = B(\eta, y_0)$. Hence we get that y_0 is the unique minimiser of the functional Q.

□

Definition 9.11. *A Hilbert space X is called **Hilbert sum** of closed subspaces E_j, $j \in J$ if and only if*

1. *$E_i \perp E_j$ for $i \neq j$.*
2. *$\bigoplus_{j\in J} E_j$ is dense in x. Here we only look at finite linear combinations of the spaces E_j, $j \in J$.*

Theorem 9.12. *Let $\{E_j, j \in \mathbb{N}_0\}$ be a system of closed, pairwise orthogonal subspaces of the Hilbert space X and let P_j be the corresponding nearest point projections onto E_j. Then the following four statements are equivalent.*

1. *$\{E_j\}$ is maximal: If $x \perp E_j$ for all j then $x = 0$.*
2. *X is the Hilbert sum of the sets $\{E_j\}$.*
3. *$x = \sum_{j=0}^{\infty} P_j x$ for all $x \in X$ (Fourier expansion).*
4. *$\|x\|^2 = \sum_{j=0}^{\infty} \|P_j x\|^2$ (Parseval identity).*

Proof. 1. $\Rightarrow$ 2.: Let $V := \overline{\bigoplus_{j\in\mathbb{N}_0} E_j}$. By the assumption we have $V^\perp = \{0\}$ and hence it follows from Lemma 9.4 that $X = V \oplus V^\perp = V$.

2. $\Rightarrow$ 3. Let $y \in \bigoplus_{j=0}^{N} E_j$ for some $N \in N$. Then

$$\begin{aligned}\|x-y\|^2 &= \left\| x - \sum_{j=0}^{N} P_j x + \sum_{j=0}^{N} P_j x - y \right\|^2\\ &= \left\| x - \sum_{j=0}^{N} P_j x \right\|^2 + \left\| \sum_{j=0}^{N} P_j x - y \right\|^2,\end{aligned}$$

where we used that $x - \sum_{j=0}^{N} P_j x \in \left(\oplus_{j=0}^{N} E_j\right)^{\perp}$ and $\sum_{j=0}^{N} P_j x - y \in \oplus_{j=0}^{N} E_j$. Hence

$$\left\| x - \sum_{j=0}^{N} P_j x \right\| = \operatorname{dist}\left(x, \bigoplus_{j=0}^{N} E_j \right) \to 0$$

as $N \to \infty$ because of the assumption.

3. $\Rightarrow$ 4. We calculate

$$\begin{aligned} \|x\|^2 =& \langle x, x \rangle = \lim_{N\to\infty} \left\langle \sum_{j=0}^{N} P_j x, \sum_{j=0}^{N} P_j x \right\rangle \\ &= \lim_{N\to\infty} \sum_{i,j=0}^{N} \langle P_i x, P_j x \rangle \\ &= \lim_{N\to\infty} \sum_{i=0}^{N} \|P_i x\|^2. \end{aligned}$$

4. $\Rightarrow$ 1. If $x \perp E_j$ for all $j \in \mathbb{N}_0$ then Lemma 9.4 implies that $P_j x = 0$ for all $j \in \mathbb{N}_0$ and hence

$$\|x\|^2 = \sum_{j=0}^{\infty} \|P_j x\|^2 = 0.$$

□

10 Sobolev spaces and elliptic boundary value problems

Definition 10.1. *Let $\Omega \subset \mathbb{R}^n$ be open and let $u \in L^1_{\text{loc}}(\Omega)$. A function $g \in L^1_{\text{loc}}(\Omega)$ is called weak derivative of u with respect to x_i, $i = 1, ..., n$, if*

$$\int_\Omega u \partial_i \eta \, dx = - \int_\Omega g \eta \, dx,$$

for all $\eta \in C_c^\infty(\Omega)$. We use the notation $\partial_i u = g$ weakly.

<u>Remarks:</u>

1. The weak derivative is unique: If $g_1, g_2 \in L^1_{\text{loc}}(\Omega)$ are two weak derivatives, then $g := (g_1 - g_2)$ satisfies

   $$\int_\Omega g \eta dx = 0$$

   for all $\eta \in C_c^\infty(\Omega)$. By the Fundamental Lemma of the Calculus of Variations it then follows that $g = 0$ almost everywhere.

2. If $\partial_i u = g$, $\partial_i v = g$ and $\alpha, \beta \in \mathbb{R}$, then $\partial_1(\alpha u + \beta v) = \alpha g + \beta v$.

3. For $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$ the α-weak derivative $D^\alpha u = g$ is defined by

   $$\int_\Omega u D^\alpha \eta \, dx = (-1)^{|\alpha|} \int_\Omega g \eta \, dx$$

   for all $\eta \in C_c^\infty(\Omega)$.

4. If $D^\alpha u = v$ and $D^\beta v = g$ weakly, where $\alpha, \beta \in \mathbb{N}_0^n$ then $D^{\alpha+\beta} u = g$ weakly since

   $$\begin{aligned}\int_\Omega u D^{\alpha+\beta} \eta \, dx &= \int_\Omega u D^\alpha (D^\beta) \, dx \\ &= (-1)^{|\alpha|} \int_\Omega v D^\beta \eta \, dx = (-1)^{|\alpha|+|\beta|} \int_\Omega g \eta \, dx\end{aligned}$$

   for all $\eta \in C_c^\infty(\Omega)$.

Example: Let $u(x) = |x|^\alpha$. For which $\alpha \in \mathbb{R}$ exists a weak derivative in $L^1_{\text{loc}}(\mathbb{R}^n)$?

For $x \neq 0$ the function u is clasically differentiable and $\partial_i u(x) = \alpha |x|^{\alpha-2} x_i$. Hence the only candidate for the weak derivative is $g_i(x) = \alpha |x|^{\alpha-2} x_i$. Now $g_i \in L^1_{\text{loc}}(\mathbb{R}^n)$ if and only if $\alpha - 1 > -n \Leftrightarrow \alpha > 1 - n$. In order to show that g_i is indeed the weak derivative for $\alpha > 1 - n$ we let $\eta \in C_c^\infty(\Omega)$ be arbitrary and we calculate

$$\begin{aligned}
\int u(x)\partial_i\eta(x)\,dx &= \lim_{\rho\to 0} \int_{\mathbb{R}^n \setminus B_\rho(0)} u(x)\partial_i\eta(x)\,dx \\
&= \lim_{\rho\to 0} \left(\int_{\mathbb{R}^n \setminus B_\rho(0)} \partial_i(u(x)\eta(x))\,dx - \int_{\mathbb{R}^n \setminus B_\rho(0)} g_i(x)\eta(x)\,dx \right) \\
&= \lim_{\rho\to 0} \left(-\int_{\partial B_\rho(0)} u(x)\eta(x)\frac{x_i}{\rho}\,d\mu_{\partial B_\rho(0)}(x) - \int_{\mathbb{R}^n \setminus B_\rho(0)} g_i(x)\eta(x)\,dx \right) \\
&= -\int_{\mathbb{R}^n} g_i\eta\,dx
\end{aligned}$$

where we used that

$$\int_{\partial B_\rho(0)} |u(x)\eta(x)|\frac{x_i}{\rho}|\,d\mu_{\partial B_\rho(0)}(x) \leq C\varrho^{n-1+\alpha} \to 0$$

as $\varrho \to 0$.

Definition 10.2. *Let $\Omega \subset \mathbb{R}^n$ be open and let $1 \leq p \leq \infty$. Then we define the Sobolev space*

$$W^{1,p}(\Omega) := \{u \in L^p(\Omega) \colon \partial_i u \in L^p(\Omega),\ \textit{for all } 1 \leq i \leq n\}.$$

Moreover, we define the $W^{1,p}$-norm by

$$\|u\|_{W^{1,p}} := \|u\|_{L^p} + \sum_{i=1}^n \|\partial_i u\|_{L^p}.$$

Similarly, for $k \in \mathbb{N}$ and $1 \leq p \leq \infty$, we let $u \in W^{k,p}(\Omega) \Leftrightarrow D^\alpha u \in L^p(\Omega)$ for all $|\alpha| \leq k$.

Theorem 10.3. *The normed space $(W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}})$ is a Banach space.*

Proof. Let (u_k) be a Cauchy sequence in $W^{1,p}(\Omega)$. Then it follows that both (u_k) and $(\partial_i u_k)$, for all $1 \leq i \leq n$, are Cauchy sequences in $L^p(\Omega)$. Since $L^p(\Omega)$ is complete (Fischer-Riesz) we conclude that that there exist functions $u, v_i \in L^p(\Omega)$, $1 \leq i \leq n$ with $u_k \to u$, $\partial_i u_k \to v_i$ in $L^p(\Omega)$. It remains to show that $v_i = \partial_i u$ weakly. Let $\eta \in C_c^\infty(\Omega)$. We calculate with the help of the dominated convergence theorem

$$\begin{aligned}\int u\partial_i\eta\,dx &= \lim_{k\to\infty}\int u_k\partial_i\eta\,dx\\ &= -\lim_{k\to\infty}\int(\partial_i u_k)\eta\,dx\\ &= -\int v_i\eta dx.\end{aligned}$$

Thus $u_k \to u$ in $W^{1,p}(\Omega)$. □

<u>Remark:</u> It was shown in the exercises that

- $W^{1,p}(\Omega)$, $1 \le p < \infty$, is separable.
- $W^{1,p}(\Omega)$, $1 < p < \infty$, is reflexive.

Theorem 10.4. *Let* $\eta \in C_c^\infty(\mathbb{R}^n)$, $\eta \ge 0$, $\operatorname{spt}\eta \subset \overline{B_1(0)}$ *and*

$$\int_{\mathbb{R}^n}\eta dx = 1.$$

Let $\eta_\rho(x) := \rho^{-n}\eta\left(\frac{x}{\rho}\right)$ *for* $\rho > 0$ *and define*

$$u_\rho(x) := (\eta_\rho * u)(x) = \int_{\mathbb{R}^n}\eta_\rho(x-y)u(y)dy$$

for $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. *Then we have*

1. $u_\rho \in C^\infty(\mathbb{R}^n)$ *for all* $\rho > 0$.
2. $\|u_\rho\|_{L^p} \le \|u\|_{L^p}$ *for all* $1 \le p \le \infty$.
3. *If* $u \in L^p(\mathbb{R}^n)$ *with* $1 \le p < \infty$ *then* $\|u_\rho - u\|_{L^p} \to 0$ *as* $\rho \to 0$.
4. *If* $u \in L^\infty(\mathbb{R}^n)$ *then* $u_\rho \overset{*}{\rightharpoonup} u$ *in* $L^\infty(\mathbb{R}^n)$ *and there exists a subsequence* $\rho_i \to 0$ *so that* $u_{\rho_i} \to u$ *pointwise almost everywhere.*

Example.

$$\eta(x) = \begin{cases} c\exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1\\ 0, & |x| \ge 1\end{cases}.$$

Proof. Statement 1. follows from differentiation results of parameter integrals.

For statement 2 we estimate for all $1 \leq p \leq \infty$

$$\int |u_\rho(x)|^p \, dx = \int \left| \int \eta_\rho(x-y)u(y)dy \right|^p \, dx$$
$$\overset{\text{Hölder}}{\leq} \int \left(\int \eta_\rho(x-y)|u(y)|^p \, dy \right) \left(\int \eta_\rho(x-y) \, dy \right)^{p-1} \, dx.$$
$$\overset{\text{Fubini}}{=} \int \left(|u(y)|^p \int \eta_\rho(x-y) \, dx \right) \, dy = \|u\|_{L^p}^p.$$

In order to show statement 3. we define the function $\tau_z(x) = x + z$ for $z \in \mathbb{R}^n$. We first show that for all $1 \leq p < \infty$ we have

$$\|u \circ \tau_z - u\|_{L^p} \to 0$$

as $z \to 0$. Since $C_c^0(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ we choose $v \in C_c^0(\mathbb{R}^n)$ so that $\|v-u\|_{L^p} < \frac{\varepsilon}{3}$. Then

$$\|u \circ \tau_z - u\|_{L^p} \leq \|(u-v) \circ \tau_z\|_{L^p} + \|v \circ \tau_z - v\|_{L^p} + \|v - u\|_{L^p} \leq \varepsilon$$

for $|z|$ small, since $v \circ \tau_z \to v$ uniformly as $|z| \to 0$. In order to show the general statement we observe that

$$\int |u_\rho(x) - u(x)|^p \, dx = \int \left| \int \eta_\rho(x-y)(u(y) - u(x)) \, dy \right|^p \, dx$$
$$\overset{\text{Hölder}}{\leq} \int \int \eta_\rho(x-y)|u(y) - u(x)|^p \, dy \, dx$$
$$= \int \int \eta(z)|u(x - \rho z) - u(x)|^p \, dz \, dx \to 0$$

as $\rho \to 0$.

For statement 4. we let $u \in L^\infty(\mathbb{R}^n)$ and this implies $u \in L^p_{\text{loc}}(\mathbb{R}^n)$ for all $p \in [1, \infty)$. Statement 3. then implies that there exists a subsequence $u_{\rho_i} \to u$ in $L^p_{\text{loc}}(\mathbb{R}^n)$ and hence also pointwise almost everywhere. Finally, we let $v \in L^1(\mathbb{R}^n)$, $\overline{\eta}(x) := \eta(-x)$ and we conclude that

$$\int u_\rho(x)v(x) \, dy = \int \int \eta_\rho(x-y)u(y)v(x) \, dy \, dx$$
$$= \int u(y) \left(\int \eta_\rho(x-y)v(x) \, dx \right) \, dy$$
$$= \int u(y) \left(\int \overline{\eta_\rho}(y-x)v(x) \, dx \right) \, dy$$
$$= \int u(y)\bar{\eta}_\rho * v(y) \, dy$$
$$\to \int u(y)v(y) \, dy$$

as $\rho \to 0$ by statement 3. □

Lemma 10.5. *Let $\Omega \subset \mathbb{R}^n$ be open, let $u \in L^p_{\text{loc}}(\Omega)$ with $p \in [1, \infty)$ and let η be as in Theorem 10.4. Then the function u_ρ is well defined on the set $\Omega_\rho := \{x \in \Omega \colon \text{dist}(x, \partial\Omega) > \rho\}$ and*

$$\|u_\rho - u\|_{L^p(\Omega')} \to 0,$$

for all $\Omega' \subset\subset \Omega$ as $\rho \to 0$.

Proof. We note that

$$\begin{aligned} u_\rho(x) &= \int_{\mathbb{R}^n} \eta_\rho(x-y)u(y)dy \\ &= \int_{B_\rho(0)} \eta_\rho(y)u(x-y)dy \end{aligned}$$

and hence we see that this function is well-defined for all $x \in \Omega_\rho$. Without loss of generality we let $\Omega' = \Omega_\sigma$ for some $\sigma > 0$. Let

$$\tilde{u}(x) = \begin{cases} u(x), & \text{if } x \in \Omega_{\sigma/2} \\ 0, & \text{otherwise} \end{cases}.$$

Then $\tilde{u} \in L^p(\mathbb{R}^n)$ and for all $x \in \Omega_\sigma$ with $\rho < \frac{\sigma}{2}$ we have

$$\begin{aligned} u_\rho(x) = (\eta_\rho * u)(x) &= \int_{B_\rho(0)} \eta_\rho(y)u(x-y)\,dy \\ &= \int_{B_1(0)} \eta(z)u(x-\rho z)\,dz = \int_{B_1(0)} \eta(z)\tilde{u}(x-\rho z)\,dz \\ &= \tilde{u}_\rho(x). \end{aligned}$$

Thus we conclude with the help of Theorem 10.4 that

$$\|u_\rho - u\|_{L^p(\Omega')} \le \|\tilde{u}_\rho - \tilde{u}\|_{L^p(\mathbb{R}^n)} \to 0$$

as $\rho \to 0$. □

Lemma 10.6 (Fundamental Lemma of the Calculus of Variations)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and let $u \in L^1_{\text{loc}}(\Omega)$ with $\int_\Omega u\varphi dx \ge 0$ for all $\varphi \in C^\infty_c(\Omega)$ with $\varphi \ge 0$. Then $u(x) \ge 0$ for almost every $x \in \Omega$.*

Proof. Let $0 < \rho < \sigma$ so that u_ρ is well defined on Ω_σ, and let $\varphi \in C^\infty_c(\Omega_\sigma)$ with

$\varphi \geq 0$. Then we have

$$\begin{aligned}\int u_\rho \varphi \, dx &= \int \int \eta_\rho(x-y) u(y) \varphi(x) \, dy \, dx \\ &= \int u(y) \left(\int \eta_\rho(x-y) \varphi(x) \, dx \right) \, dy \\ &= \int u(y) (\overline{\eta}_\rho * \varphi)(y) \, dy \geq 0\end{aligned}$$

since $\overline{\eta}_\rho * \varphi \in C_c^\infty(\Omega, \mathbb{R}_0^+)$. Since u_ρ is continuous a standard argument yields that $u_\rho \geq 0$ on Ω_σ. Since $u_{\rho_i} \to u$ pointwise almost everywhere it follows that $u \geq 0$ almost everywhere. □

Lemma 10.7. *Let $\Omega \subset \mathbb{R}^n$ be open, let $u \in W^{1,p}(\Omega)$ and let η be as in Theorem 10.4. Then*

1. *$\partial_i u_\rho = (\partial_i u)_\rho$ on Ω_ρ.*

2. *For $1 \leq p < \infty$ and $\Omega' \subset\subset \Omega$ we have $u_\rho \to u$ in $W^{1,p}(\Omega')$.*

Proof. Statement 1. implies statement 2. by Lemma 10.5. In order to show statement 1. we calculate for all $x \in \Omega_\rho$

$$\begin{aligned}\partial_i u_\rho = \partial_i(\eta_\rho * u)(x) &= \int \left(\frac{\partial}{\partial x_i} \eta_\rho(x-y) \right) u(y) \, dy \\ &= -\int \frac{\partial}{\partial y_i} \eta_\rho(x-y) u(y) \, dy \\ &= \int \eta_\rho(x-y) \partial_i u(y) \, dy \\ &= (\eta_\rho * \partial_i u)(x) = (\partial_i u)_\rho(x),\end{aligned}$$

where we used the definition of the weak derivative in the third line. □

<u>Remark:</u> If $u \in W^{1,p}(\mathbb{R}^n)$ then $u_\rho \to u$ in $W^{1,p}(\mathbb{R}^n)$ by Theorem 10.4.

Lemma 10.8. *Let $\Omega \subset \mathbb{R}^n$ be open and let $u, v \in W^{1,p} \cap L^\infty(\Omega)$. Then $uv \in W^{1,p} \cap L^\infty(\Omega)$ with*

$$\partial_i(uv) = (\partial_i u)v + u(\partial_i v).$$

Proof. Let $\varphi \in C_c^\infty(\Omega)$ and note that it follows from Lemma 10.5, Lemma 10.7 and the dominated convergence theorem that

$$\begin{aligned}\int_\Omega uv\partial_i\varphi &= \lim_{\rho\to 0}\int u(\eta_\rho * v)\partial_i\varphi \\ &= \lim_{\rho\to 0}\int u\partial_i\left[(\eta_\rho * v)\varphi\right] - u\varphi\partial_i(\eta_\rho * v) \\ &\overset{\text{Lemma 9.7}}{=} -\lim_{\rho\to 0}\int \left[\partial_i u(\eta_\rho * v) + u\varphi\eta\rho * \partial_i v\right] \\ &= -\int (u\partial_i v\varphi + \partial_i uv\varphi).\end{aligned}$$

□

Remark: The same proof implies that if $u \in W^{1,p}(\mathbb{R}^n)$ and $v \in C_c^\infty(\mathbb{R}^n)$ then $uv \in W^{1,p}(\mathbb{R}^n)$ and $\partial_i(uv) = \partial_i uv + u\partial_i v$.

Theorem 10.9 (Meyers-Serrin)**.** *We have that*

1. *$C_c^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$ for all $1 \le p < \infty$.*

2. *$W^{1,p} \cap C^\infty(\Omega)$ is dense in $W^{1,p}(\Omega)$ for all $1 \le p < \infty$ and all $\Omega \subset \mathbb{R}^n$ open.*

Proof. 1. We know from Lemma 10.7 that $W^{1,p} \cap C^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$. Hence we let $u \in W^{1,p} \cap C^\infty(\mathbb{R}^n)$ and we choose $\varphi \in C_c^\infty(\mathbb{R}^n, [0,1])$ with

$$\varphi(x) = \begin{cases} 1, & |x| \le 1 \\ 0, & |x| \ge 2 \end{cases}$$

Set $\varphi_R(x) := \varphi(x/R)$ and $u_R(x) = \varphi_R(x)u(x)$. Then $u_R \in C_c^\infty(\mathbb{R}^n)$ and Lemma 10.8 yields

$$\partial_i u_R(x) = \varphi_R(x)\partial_i u(x) + \frac{1}{R}\partial_i\varphi\left(\frac{x}{R}\right)u(x)$$

and hence

$$\begin{aligned}\|u_R - u\|_{W^{1,p}} =& \|u_R - u\|_{L^p} + \sum_{i=1}^n \|\partial_i(u_R - u)\|_{L^p} \\ \le& \left(\int\limits_{\mathbb{R}^n\setminus B_R(0)} |u|^p\right)^{1/p} + \sum_{i=1}^n \left(\int\limits_{\mathbb{R}^n\setminus B_R(0)} |\partial_i u|^p\right)^{1/p} + \sum_{i=1}^n \frac{1}{R}\|\partial_i\varphi\left(\frac{\cdot}{R}\right)u\|_{L^p} \\ \to& 0\end{aligned}$$

as $R \to \infty$.

2. Let $U_k = \Omega_{1/k} \cap B_k(0)$ for $k \in \mathbb{N}$ and let $U_0 := \emptyset$. Moreover, we define $V_k :=$

$U_{k+1}\backslash\overline{U}_{k-1}$ for all $k \in \mathbb{N}$ and we choose a partition of unity

$$\varphi_k \in C_c^\infty(V_k),\ \varphi_k \geq 0,\ \sum_{k=1}^\infty \varphi_k \equiv 1 \text{ on } \Omega.$$

For $\varepsilon > 0$ there exists $\delta_k > 0$ so that for $u_k = \eta_{\delta_k} * (\varphi_k u)$ we have

$$\|u_k - \varphi_k u\|_{W^{1,p}(\Omega)} \leq 2^{-k}\varepsilon$$

by Lemma 10.7. For $v := \sum_{k=1}^\infty u_k \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$ we conclude

$$\|v-u\|_{W^{1,p}} \leq \sum_{k=1}^\infty \|u_k - \varphi_k u\|_{W^{1,p}} < \varepsilon.$$

□

Remarks:

1. Let $H^{1,p}(\Omega)$ be the completion of the space

$$X = \{u \in C^\infty(\Omega) \colon \|u\|_{W^{1,p}(\Omega)} < \infty\}$$

with respect to the $W^{1,p}$-norm. Every element $u \in H^{1,p}(\Omega)$ is represented by a Cauchy sequence $u_k \in X$ and we obtain an isometric embedding $H^{1,p}(\Omega) \to W^{1,p}(\Omega)$, $u \mapsto \lim\limits_{k\to 0} u_k$. By Theorem 10.9 this map is surjective and therefore $H^{1,p}(\Omega) \cong W^{1,p}(\Omega)$ isometrically.

2. For $\Omega \subset \mathbb{R}^n$ open, bounded we have that $C_c^\infty(\Omega)$ is not dense in $W^{1,p}(\Omega)$. For this we let $u \in C_c^\infty(\Omega)$ and we note that

$$0 = \int_\Omega \partial_i(ux_i) = \int_\Omega (x_i\partial_i x_i + u.$$

If $C_c^\infty(\Omega)$ would be dense in $W^{1,p}(\Omega)$, then this equality would be true for all $u \in W^{1,p}(\Omega)$ but for $u \equiv 1$ we obtain the contradiction $0 = \mathcal{L}^n(\Omega)$.

3. In general we have that $C^\infty(\overline{\Omega})$ is not dense in $W^{1,p}(\Omega)$.

4. $W^{1,\infty} \cap C^\infty(\Omega)$ is not dense in $W^{1,\infty}(\Omega)$ since this would imply that $\partial_i u \in C^0(\Omega)$ for all $u \in W^{1,\infty}(\Omega)$ as it is the uniform limit of a sequence of continuous functions. But $|x| \in W^{1,\infty}(B_1(0))$ with

$$\partial_i|x| = \frac{x_i}{|x|} \notin C^0(B_1(0)).$$

Definition 10.10. *For $\Omega \subset \mathbb{R}^n$ open and $1 \leq p < \infty$ we denote by $W_0^{1,p}(\Omega)$ the* *closure of* $C_c^\infty(\Omega)$ *in* $W^{1,p}(\Omega)$.

Theorem 10.11. *Let* $\Omega \subset \mathbb{R}^n$ *be open and let* $u \in W^{1,1}_{\text{loc}}(\Omega)$, $f \in C^1(\mathbb{R})$ *with* $\|f'\|_{C^0} =: L < \infty$. *Then we have*

$$D(f \circ u) = (f' \circ u) Du \in L^1_{\text{loc}}(\Omega).$$

Proof. Let $\varphi \in C_c^\infty(\Omega_\sigma)$ for some $\sigma > 0$ and let $0 < \rho < \sigma$. Then it follows that

$$\partial_i(f \circ u_\rho) = (f' \circ u_\rho)\partial_i u_\rho \text{ on } \Omega_\sigma$$

and thus

$$-\int f \circ u_\rho \partial_i \varphi = \int (f' \circ u_\rho)\partial_i u_\rho \varphi.$$

Next we let $\rho \searrow 0$ and we note that $u_\rho \to u$, $Du_\rho \to Du$ in $L^1(\Omega_\sigma)$ by Lemma 10.7. Choosing a subsequence implies that $u_\rho \to u$ also converges pointwise almost everywhere in Ω_σ. Hence we conclude that for $\rho \searrow 0$

$$\begin{aligned}
\left|\int (f \circ u_\rho - f \circ u)\partial_i \varphi\right| &\le L \|\partial_i \varphi\|_{L^\infty} \int_{\Omega_\sigma} |u_\rho - u| \to 0,\\
\left|\int (f' \circ u_\rho)(\partial_i u_\rho)\varphi - (f' \circ u)(\partial_i u)\varphi\right| &\le \int |f' \circ u_\rho||\partial_i u_\rho - \partial_i u||\varphi|\\
&\quad + \int |f' \circ u_\rho - f' \circ u||\partial_i u||\varphi| \to 0.
\end{aligned}$$

Together, this shows that

$$-\int f \circ u \partial_i \varphi = \int (f' \circ u)\partial_i u \varphi.$$

□

In the following we are interested in solving elliptic boundary value problems. More precisely, for $a \in L^\infty(\Omega, M_n(\mathbb{R}))$, $\Omega \in \mathbb{R}^n$ open and bounded we study the operator L given by

$$Lv := -div(aDv) = -\sum_{\alpha,\beta=1}^n \partial_\alpha(a_{\alpha\beta}\partial_\beta v).$$

The Dirichlet boundary value problem for L is given by

$$\begin{cases} Lv = f & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega, \end{cases}.$$

Definition 10.12. *The bilinear form associated to* L *on* $W_0^{1,2}(\Omega)$ *is given by*

$$B(u,v) = \int_\Omega \langle Du, aDv \rangle = \sum_{\alpha,\beta=1}^n \int a_{\alpha\beta}\partial_\alpha u \partial_\beta v.$$

It follows that B is bounded since

$$|B(u,v)| \leq \|a\|_{L^\infty}\|Du\|_{L^2}\|Dv\|_{L^2} \leq \|a\|_{L^\infty}\|u\|_{W^{1,2}}\|v\|_{W^{1,2}}$$

and therefore $\|B\| \leq \|a\|_{L^\infty}$.

From now on we consider L as an operator

$$\begin{aligned} L\colon W_0^{1,2}(\Omega) &\to (W_0^{1,2}(\Omega))', \\ (Lv)(u) &= \sum_{\alpha,\beta=1}^{n} \int a_{\alpha\beta}\partial_\alpha u \partial_\beta v = B(u,v). \end{aligned}$$

It follows that L is continuous with $\|L\| \leq \|a\|_{L^\infty}$ and we have that $L = R_B$ (compare Theorem 9.10). It follows from Theorem 9.10 that L is surjective if the bilinear form is coercive and this is what we want to verify in the following.

Theorem 10.13 (Poincare inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. For $1 \leq p \leq \infty$ and all $u \in W_0^{1,2}(\Omega)$ we have*

$$\|u\|_{L^p} \leq d\|Du\|_{L^p},$$

where $d = \mathrm{diam}(\Omega)$.

Proof. Without loss of generality we assume that $u \in C_c^\infty(\Omega)$ and $\Omega \subset \{x\colon 0 \leq x^n \leq d\}$. Set $u(x) = 0$ for all $x \notin \Omega$ and let $x = (\xi, x^n) \in \mathbb{R}^{n-1} \times \mathbb{R} = \mathbb{R}^n$, $0 \leq x^n \leq d$. Then we estimate

$$\begin{aligned} |u(\xi,x^n)|^p =& \left| \int_0^{x^n} (\partial_n u)(\xi,s)\,ds \right|^p \\ \leq& d^{p-1} \int_0^d |\partial_n u(\xi,s)|^p\,ds, \end{aligned}$$

where we used that $u(\cdot,0) = 0$ and Hölders inequality. By Fubini's theorem we now get

$$\begin{aligned} \int_\Omega |u(x)|^p\,dx =& \int_{\mathbb{R}^{n-1}} \int_0^d |u(\xi,x^n)|^p\,dx^n\,d\xi \\ \leq& d^{p-1} \int_{\mathbb{R}^{n-1}} \int_0^d \int_0^d |\partial_n u(\xi,s)|^p\,ds\,dx^n\,d\xi \\ =& d^p \int_{\mathbb{R}^{n-1}} \int_0^d |\partial_n u(\xi,s)|^p ds d\xi \\ =& d^p \|Du\|_{L^p}^p. \end{aligned}$$

□

Lemma 10.14. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and let L be as above. Moreover,* *we assume that L is elliptic with constant $\mu > 0$, i.e.*

$$\langle \xi, a(x)\xi \rangle \geq \mu |\xi|^2$$

for all $x \in \Omega$, $\xi \in \mathbb{R}^n$. Then the associated bilinear form B is coercive on $W_0^{1,2}(\Omega)$ with

$$B(u,u) \geq \lambda \|u\|^2_{W^{1,2}},$$

where $\lambda = \frac{\mu}{d^2+1}$ and $d = \operatorname{diam}(\Omega)$.

Proof. It follows from the ellipticity assumption that

$$B(u,u) = \int \langle aDu, Du \rangle \geq \mu \int |Du|^2.$$

By Theorem 9.13 we get

$$\begin{aligned} \|u\|^2_{W^{1,2}} &= \|u\|^2_{L^2} + \|Du\|^2_{L^2} \\ &\leq (d^2+1)\|Du\|^2_{L^2} \end{aligned}$$

and hence

$$B(u,u) \geq \frac{\mu}{d^2+1} \|u\|^2_{W^{1,2}}.$$

□

Theorem 10.15. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $a \in L^\infty(\Omega, M_n(\mathbb{R}))$ be elliptic with constant $\mu > 0$. Then*

$$\begin{aligned} L\colon W_0^{1,2}(\Omega) &\Rightarrow (W_0^{1,2}(\Omega))', \\ (Lv)(u) = B(u,v) &= \sum_{\alpha,\beta=1}^{n} \int a_{\alpha\beta} \partial_\alpha u \partial_\beta v \end{aligned}$$

is invertible and

$$\|L^{-1}\| \leq \frac{d^2+1}{\mu}.$$

Moreover, if a is symmetric on Ω and $\varphi \in (W_0^{1,2}(\Omega))'$, then the solution $u \in W_0^{1,2}(\Omega)$ of $Lu = \varphi$ is the unique minimum of the quadratic functional

$$Q(v) = \frac{1}{2}B(v,v) - \varphi(v) = \frac{1}{2}\int \langle aDv, Dv \rangle - \varphi(v).$$

Proof. This follows immediately from Theorem 9.10 and Lemma 10.14. □

Special cases:

1. The Dirichlet problem with a right hand side in L^2.

   We let $f \in L^2(\Omega)$ and we want to solve the problem

$$\begin{cases} -div(aDv) = f, & \text{in } \Omega \\ v = 0, & \text{on } \partial\Omega \end{cases}$$

   where $a \in L^\infty(\Omega, M_n(\mathbb{R}))$ is elliptic with constant $\mu > 0$. This is the so called classical formulation. In order to get the weak formulation of this problem, we multiply the partial differential equation by $u \in W_0^{1,2}(\Omega)$ and we formally integrate by parts in order to get

$$\begin{cases} \int_\Omega \langle aDv, Du\rangle = \int_\Omega fu \\ \quad v \in W_0^{1,2}(\Omega) \end{cases}$$

   Next we define the map $L^2(\Omega) \to (W_0^{1,2}(\Omega))'$, $f \mapsto \varphi_f$, $\varphi_f(u) = \int_\Omega fu$ for all $u \in W_0^{1,2}(\Omega)$. We note that

$$|\varphi_f(u)| \leq \|f\|_{L^2}\|u\|_{L^2} \leq \|f\|_{L^2}\|u\|_{W^{1,2}}, \quad \text{i.e.} \quad \|\varphi_f\| \leq \|f\|_{L^2}.$$

   In particular we can restate the weak formulation as follows:

$$\begin{cases} Lv = \varphi_f, & \text{in } (W_0^{1,2}(\Omega))' \\ v \in W_0^{1,2}(\Omega) \end{cases} . \tag{10.1}$$

   The following Lemma is a direct consequence of Theorem 10.15.
   Lemma 10.16. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and let $f \in L^2(\Omega)$. Then the problem* (10.1) *has a unique solution $v \in W_0^{1,2}(\Omega)$ with*

$$\|v\|_{W^{1,2}} \leq \frac{d^2+1}{\mu}\|f\|_{L^2}.$$

2. The Dirichlet problem with a right hand side in divergence form.

   We let $a \in L^\infty(\Omega, M_n(\mathbb{R}))$ be elliptic with constant $\mu > 0$ and we let $F \in L^2(\Omega, \mathbb{R}^n)$. We want to solve the problem

$$\begin{cases} -div(aDv) = divF, & \text{in } \Omega \\ v = 0, & \text{on } \partial\Omega \end{cases} ,$$

   Here the weak formulation is given by

$$\begin{cases} \int_\Omega \langle aDv, Du\rangle = -\int_\Omega \langle F, Du\rangle \\ \quad v \in W_0^{1,2}(\Omega) \end{cases} .$$

for all $u \in W_0^{1,2}(\Omega)$. We define the map $L^2(\Omega) \to (W_0^{1,2}(\Omega))'$, $F \to \Psi_F$ with $\Psi_F(u) = -\int \langle F, Du \rangle$ and we note that

$$|\Psi_F(u)| \leq \|F\|_{L^2} \|Du\|_{L^2} \leq \|F\|_{L^2} \|u\|_{W^{1,2}}$$

and hence

$$\|\Psi_F\| \leq \|F\|_{L^2}.$$

Therefore we can again restate the weak formulation as

$$\begin{cases} Lv = \Psi_F, & \text{in } (W_0^{1,2}(\Omega))' \\ v \in W_0^{1,2}(\Omega). \end{cases} \tag{10.2}$$

and the following Lemma is again a direct consequence of Theorem 10.15.

Lemma 10.17. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and let $F \in L^2(\Omega, \mathbb{R}^n)$. Then the problem* (10.2) *has a unique solution $v \in W_0^{1,2}(\Omega)$ with*

$$\|v\|_{W^{1,2}} \leq \frac{d^2+1}{\mu} \|F\|_{L^2}.$$

3. The Dirichlet problem with a nonzero boundary condition.

We let $a \in L^\infty(\Omega, M_n(\mathbb{R}))$ be elliptic with constant $\mu > 0$ and we let $f \in L^2(\Omega)$, $v_0 \in W^{1,2}(\Omega)$. Now we want to solve the problem

$$\begin{cases} -div(aDv) = f, & \text{in } \Omega \\ v - v_0 \in W_0^{1,2}(\Omega) \end{cases}.$$

For this we make the ansatz $v = v_0 + w$ with $w \in W_0^{1,2}(\Omega)$ and we note that w is then a solution of the problem

$$\begin{cases} -div(aDw) = f + div(aDv_0), & \text{in } \Omega \\ w \in W_0^{1,2}(\Omega) \end{cases}.$$

It the follows from Lemma 10.16 and Lemma 10.17 that there exists a unique solution $w \in W_0^{1,2}(\Omega)$ of this problem which satisfies the estimate

$$\|w\|_{W^{1,2}} \leq \frac{d^2+1}{\mu} (\|f\|_{L^2} + \|a\|_{L^\infty} \|Dv_0\|_{L^2}).$$

Hence we conclude that $v = v_0 + w$ is the unique solution of our original problem and it satisfies the estimate

$$\|v\|_{W^{1,2}} \leq c(d, \mu, a)(\|f\|_{L^2} + \|v_0\|_{W^{1,2}}).$$

11 Compact and Fredholm operators

One of our motivations for studying compact and Fredholm operators comes again from the theory of elliptic boundary value problems. Namely, we have seen in Theorem 10.15 that the operator

$$L_0 : W_0^{1,2}(\Omega) \to W_0^{1,2}(\Omega)', \quad (L_0 v)(u) = \int_\Omega \langle aDv, Du \rangle,$$

where $a \in L^\infty(\Omega, M_n(\mathbb{R}))$ is elliptic, is an isomorphism. But the second order elliptic operator

$$L = L_0 + K\colon W_0^{1,2}(\Omega) \to W_0^{1,2}(\Omega)',$$

with

$$Kv = -div(bv) + \langle c, Dv \rangle + qv,$$

and $b, c\colon \Omega \to \mathbb{R}^n$, $q\colon \Omega \to \mathbb{R}$ or in the weak formulation

$$(Kv)(u) = \int (\langle bv, Du \rangle + \langle c, Dv \rangle + quv)$$

is in general **not** an isomorphism.

Example: Let $Lu = -u'' - u$, hence $L_0(u) = u''$ and $Ku = -u$ on $\Omega = (-\pi/2, \pi/2)$.

1. The function $v_0(x) = \cos(x)$ is in $W_0^{1,2}(\Omega)$.

   For every $0 < \delta < \frac{1}{2}$ we let $\eta_\delta : \Omega \to [0,1]$ be a smooth function with $\mathrm{spt}\eta_\delta \subset (-\frac{\pi}{2} + \delta, \frac{\pi}{2} - \delta$, $\eta_\delta(x) = 1$ for all $x \in (-\frac{\pi}{2} + 2\delta, \frac{\pi}{2} - 2\delta)$ and $\|D\eta_\delta\|_{L^\infty} \leq C\delta^{-1}$. Then $\eta_\delta v_0 \in C_c^\infty(\Omega)$ and

   $$\|\eta_\delta v_0 - v_0\|_{L^2}^2 \leq 2 \int_{\Omega \setminus (-\frac{\pi}{2} + 2\delta, \frac{\pi}{2} - 2\delta)} |v_0|^2 \to 0$$

   as $\delta \to 0$. Moreover

   $$\|D(\eta_\delta v_0 - v_0)\|_{L^2}^2 \leq 2 \int_{\Omega \setminus (-\frac{\pi}{2} + 2\delta, \frac{\pi}{2} - 2\delta)} (|Dv_0|^2 + |D\eta_\delta|^2 |v_0|^2) \to 0$$

   as $\delta \to 0$ since $|v_0(x)| \leq c\delta$ for all $x \in \Omega \setminus (-\frac{\pi}{2} + 2\delta, \frac{\pi}{2} - 2\delta)$.

2. L is **not injective**.

For this we let $v_0(x) = \cos(x)$ be as above and we note that for all $u \in C_c^\infty(\Omega)$

$$(Lv_0)(u) = [uv_0']_{-\pi/2}^{\pi/2} - \int_{-\pi/2}^{\pi/2} (v_0'' u + v_0 u) = 0.$$

Since $C_c^\infty(\Omega)$ is dense in $W_0^{1,2}(\Omega)$ this remains to be true for all $u \in W_0^{1,2}(\Omega)$ and therefore $Lv_0 \equiv 0$.

3. L is **not surjective**.

   Let $f \in W_0^{1,2}(\Omega)'$ and let $v \in W_0^{1,2}(\Omega)$ be a solution of $Lv = f$ in the weak sense. Then we get

$$\int_{-\pi/2}^{\pi/2} f v_0 = \varphi_f(v_0) = (Lv)(v_0) = B(v_0, v) = B(v, v_0) = (Lv_0)(v) = 0$$

   where $v_0(x) = \cos(x)$ as above. Therefore we must have that $f \perp_{L^2} v_0$ and all functions which are not L^2-orthogonal to v_0 are not in the image of L.

We want to consider the operator K as a perturbation of the isomorphism L_0 and this is what naturally leads us to the notion of compact and Fredholm operators.

Definition 11.1. *Let X, Y be Banach spaces. A map $K \in L(X, Y)$ is called* ***compact****, if for every bounded sequence $\{x_n\} \subset X$ the image sequence $\{Kx_n\} \subset Y$ has a converging subsequence. We define*

$$K(X, Y) := \{K \in L(X, Y) \colon K \text{ is compact}\}.$$

Lemma 11.2. *Let X and Y be Banach spaces. For $K \in L(X, Y)$ the following statements are equivalent*

1. *$K \in K(X, Y)$.*

2. *$K(M)$ is relatively compact in Y for all $M \subset X$ bounded (indeed $M = B_1(0)$ is sufficient).*

3. *If X is additionally reflexive then $x_n \rightharpoonup x$ implies $Kx_n \to Kx$.*

Proof. 1. $\Rightarrow$ 2. We show that $\overline{K(M)}$ is sequentially compact for all $M \subset X$ bounded and by Theorem 3.2 this implies the claim. We let $y_n \in \overline{K(M)}$ be a sequence and it follows that there exists a sequence $x_n \in M$ so that $\|y_n - Kx_n\| < \frac{1}{n}$ for all $n \in \mathbb{N}$. As K is compact we conclude $Kx_n \to y$ up to a subsequence and therefore $y \in \overline{K(M)}$ and $y_n \to y$.

2. $\Rightarrow$ 1. Let $\{x_n\} \subset X$ be a bounded sequence, i.e. we assume that there exists

$R < \infty$ with $\|x_n\| \leq R$ for all $n \in \mathbb{N}$. Hence $\{Kx_n\} \subset \overline{K(B_R(0))} = R\overline{K(B_1(0))}$ which is compact by the assumption. Thus $Kx_n \to y$ up to a subsequence.

1. $\Rightarrow$ 3. Assume that x_n is a sequence in X with $x_n \rightharpoonup x$. It follows from Theorem 8.2 that $Kx_n \rightharpoonup Kx$ in Y and $\|x_n\| \leq c$ for some $c < \infty$. Since $K \in K(X,Y)$ we conclude that there exists a subsequence $Kx_{n_j} \to y$ and it follows again from Theorem 8.2 that $y = Kx$. We claim that this already implies that $Kx_n \to Kx$.

Indeed, this follows from the following general fact: If $\{x_n\}$ is a sequence in a metric space X such that every subsequence has a further subsequence which converges to the same limit y for all subsequences, then $x_n \to y$. Namely, if we assume that x_n does not converge to y, then there exists at least one subsequence x_{n_j} which does not converge to y and this contradicts the assumption.

3. $\Rightarrow$ 1. Let $\{x_n\} \subset X$ be bounded. As X is reflexive, it follows from Theorem 8.5 that $x_n \rightharpoonup x$ up to a subsequence and by the assumption we conclude $Kx_n \to Kx$. □

Examples:

1. $K \in L(X,Y)$ is compact if $\dim \mathrm{Image}(K) < \infty$, since by Theorem 1.2 all norms on $\mathrm{Image}(K)$ are equivalent and we can then use the Bolzano-Weierstrass theorem.

2. Let $\Omega \subset \mathbb{R}^n$ be open and bounded and let $0 \leq \beta < \alpha \leq 1$ then the inclusion map $C^{0,\alpha}(\overline{\Omega}) \subset C^{0,\beta}(\overline{\Omega})$ is compact by Theorem 3.10.

3. $\mathrm{id}\colon X \to X$ is compact if and only if $\dim X < \infty$ by Theorem 3.4 and Lemma 11.2.

Lemma 11.3. *Let X and Y be Banach spaces. Then $K(X,Y)$ is a closed subspace of $L(X,Y)$.*

Proof. Let $T \in \overline{K(X,Y)}$ and let $\{x_n\} \subset X$ be bounded, i.e. $\|x_n\| \leq c$ for all $n \in \mathbb{N}$ and some $c < \infty$. There exist maps $K_j \in K(X,Y)$ so that $\|T - K_j\| < 1/j$ and by the diagonal sequence argument we get that $\{K_j x_n\}_{n\in\mathbb{N}}$ converges for all $j \in \mathbb{N}$ up to a subsequence. Then we estimate

$$\begin{aligned}\|Tx_n - Tx_m\| \leq& \|Tx_n - K_j x_n\| + \|K_j x_n - K_j x_m\| + \|K_j x_m - Tx_m\| \\ \leq& \|T - K_j\|(\|x_n\| + \|x_m\|) + \|K_j x_n - K_j x_m\| \\ \leq& \frac{2c}{j} + \|K_j x_n - K_j x_m\| < \varepsilon,\end{aligned}$$

if j, m, n are sufficiently large. Therefore the sequence Tx_n is a Cauchy sequence and hence it converges $Tx_n \to y$ which shows that $T \in K(X, Y)$. □

Theorem 11.4 (Schauder)**.** *Let X, Y, Z be Banach spaces and let $S \in L(Y, Z)$, $T \in K(X, Y)$. Then $ST \in K(X, Z)$. The same is true if $S \in K(Y, Z)$ and $T \in L(X, Y)$. Moreover, if $K \in K(X, Y)$, then the Banach adjoint K' is also compact, i.e. $K' \in K(Y', X')$.*

Proof. In order to show the first statement we let $\{x_n\} \subset X$ be bounded. Then $Tx_n \to y \in Y$ up to a subsequence and hence $S(Tx_n) \to Sy$ up to a subsequence. The other case is treated similarly.

For the proof of the second statement we use Lemma 11.2 and we let $M := \overline{K(B_1(0))} \subset Y$ be compact. Let $\psi_n \in Y'$ be a sequence with $\|\psi_n\| \leq c$ for all $n \in \mathbb{N}$. We have to show that $K'\psi_n = \psi_n \circ K \in X'$ has a converging subsequence.

Since M is compact we have that $M \subset B_R(0)$ for some $R < \infty$ and hence

$$\sup_{y \in M} |\psi_n(y)| \leq \|\psi_n\| R \leq cR$$

which implies that $\{\psi_n|_M\}$ is uniformly bounded. Moreover, for all $y_1, y_2 \in M$ we have

$$|\psi_n(y_1) - \psi_n(y_2) \leq \|\psi_n\| \|y_1 - y_2\| \leq C\|y_1 - y_2\|$$

and thus $\{\psi_n|_M\}$ is uniformly Lipschitz continuous and therefore also equicontinuous. Theorem 3.8 (Arzela-Ascoli) then implies that $\psi_n|_M$ is a Cauchy sequence in $C^0(M)$, at least up to the choice of a subsequence. From this we conclude that for all $x \in B_1(0)$ and all n, m large enough $|\psi_n(Kx) - \psi_m(Kx)| < \varepsilon$ and thus

$$\|\psi_n \circ K - \psi_m \circ K\| \leq \varepsilon$$

which shows that $\{K'\psi_n\}$ is a Cauchy sequence in X'. □

Definition 11.5. *Let X and Y be Banach spaces. A map $T \in K(X, Y)$ is called a* ***Fredholm operator****, and we use the notation $T \in \mathrm{Fred}(X, Y)$, if and only if* $\mathrm{Image}(T) \subset Y$ *is closed,* $\dim(\ker T) < \infty$ *and* $\dim \operatorname{coker}(T) < \infty$*, where* $\operatorname{coker}(T) := Y/\operatorname{Image}(T)$*.*

Moreover, we define the index of T by $\operatorname{ind}(T) := \dim(\ker T) - \dim(\operatorname{coker} T) \in \mathbb{Z}$*.*

Remark: If X and Y are finite dimensional and $T \in L(X,Y)$ then

$$\operatorname{ind}(T) = \dim(\ker T) - \dim(Y) + \dim(\operatorname{Image} T) = \dim X - \dim Y$$

is independent of T.

Examples:

1. If T is an isomorphism then $\operatorname{ind}(T) = 0$.
2. - The map $l^p \to l^p$, $(x_1, x_2, ...) \mapsto (0, x_1, x_2, ...)$ has ind $= -1$.
   - The map $l^p \to l^p$, $(x_1, x_2, ...) \mapsto (x_2, x_3, ...)$ has ind $= +1$.

Theorem 11.6. *Let X and Y be Banach spaces and assume that the map $T \in L(X,Y)$ has a closed image. Then* $\operatorname{Image}(T')$ *is closed as well and we have*

1. $\ker(T') \cong (\operatorname{coker} T)'$.

2. $\operatorname{coker}(T') \cong (\ker T)'$.

Proof. Since the image of T is closed we can use Theorem 2.8 in order to conclude that $\operatorname{coker} T = Y/\operatorname{Image} T$ is a Banach space with norm

$$\|[y]\| = \inf_{x\in X} \|y + Tx\|$$

and we define the projection map

$$\pi\colon Y \to Y/\operatorname{Image} T,\ \pi(y) = [y].$$

In order to show claim 1 we prove that the adjoint map

$$\pi'\colon (\operatorname{coker} T)' \to \ker(T') \subset Y'$$

is an isomorphism. We note that $T' \circ \pi' = (\pi \circ T)' = 0$ and hence the image of π' is indeed contained in $\ker(T')$. Now we define the map

$$\rho\colon \ \ker(T') \to (\operatorname{coker} T)', \quad (\rho\psi)[y] := \psi(y)$$

for all $\psi \in \ker(T')$ and all $y \in Y$. Note that the map $\rho\psi$ is well-defined since $\psi(Tx) = T'(\psi x) = 0$ as $\psi \in \ker(T')$. Moreover, we have for all $x \in X$ that

$$|\rho\psi[y]| = |\psi(y + Tx)| \leq \|\psi\|\|y + Tx\|$$

and by taking the infimum over all $x \in X$ we get

$$|\rho\psi[y]| \leq \|\psi\|\|[y]\|$$

and thus $\rho\psi \in (\operatorname{coker} T)'$.

We calculate

$$(\pi'(\rho\psi))(y) = (\rho\psi)(\pi y) = \rho\psi([y]) = \psi(y)$$

and thus

$$\pi' \circ \rho = \mathrm{id}_{\ker(T')}.$$

On the other hand, for $g \in (\operatorname{coker} T)'$, we have

$$(\rho\pi' g)([y]) = (\pi' g)(y) = g(\pi(y)) = g([y]),$$

which shows that

$$\rho \circ \pi' = \mathrm{id}_{(\operatorname{coker} T)'}$$

and π' is an isomorphism.

In order to show claim 2. we consider the map

$$\sigma\colon X' \to (\ker T)', \quad \sigma\varphi = \varphi|_{\ker T},$$

for all $\varphi \in X'$ and we note that σ is continuous with $\|\sigma\| \leq 1$.

We claim that

$$\mathrm{Image}(T') = \ker(\sigma)$$

and we note that this implies that $\mathrm{Image}(T')$ is closed.

(i) First we observe that

$$\sigma(T'\psi) = \sigma(\psi \circ T) = (\psi \circ T)|_{\ker(T)} = 0$$

for all $\psi \in Y'$ and therefore $\mathrm{Image}(T') \subset \ker(\sigma)$.

(ii) For the other inclusion we define the map $\overline{T}\colon X/\ker(T) \xrightarrow{\sim} \mathrm{Image}(T)$, $\overline{T}([x]) = Tx$ and we note that $\overline{T}$ is well-defined, bijective and continuous. By Theorem 5.10 there exists $S \in L(\mathrm{Image}(T), X/\ker(T))$ with $S \circ \overline{T} = \mathrm{id}_{X/\ker(T)}$. Let now $\varphi \in \ker(\sigma) \subset X'$ ($\Leftrightarrow \varphi|_{\ker(T)} = 0$) and define

$$\overline{\varphi} \in (X/\ker T)', \quad \overline{\varphi}([x]) = \varphi(x).$$

Moreover, we let $\psi = \overline{\varphi} \circ S \in (\operatorname{Image}(T))'$ and we extend ψ to $\tilde{\psi} \in Y'$ with the help of Theorem 4.2. It follows that

$$\begin{aligned}(T'\tilde{\psi})(x) = \tilde{\psi}(Tx) = \psi(Tx) =&(\overline{\varphi} \circ S \circ \overline{T})([x]) \\ =&\overline{\varphi}([x]) \\ =&\varphi(x).\end{aligned}$$

Hence $T'\tilde{\psi} = \varphi$, i.e. $\ker(\sigma) \subset \operatorname{Image}(T')$, and the claim is proved.

Finally we show that the map

$$\overline{\sigma}\colon X'/\operatorname{Image}(T') = \operatorname{coker}(T') \to (\ker T)', \quad \overline{\sigma}([\varphi]) := \varphi|_{\ker T} = \sigma(\varphi)$$

is an isomorphism.

First we note that it follows from the claim that $\overline{\sigma}$ is well-defined, continuous and injective. Moreover, $\overline{\sigma}$ is also surjective since every $\lambda \in (\ker T)'$ can be extended to $\Lambda \in X'$ with the help of Theorem 4.2 and then we calculate

$$\overline{\sigma}([\Lambda]) = \Lambda|_{\ker T} = \lambda.$$

Hence it follows from Theorem 5.10 that $\overline{\sigma}$ is an isomorphism. □

Lemma 11.7. *Let X and Y be Banach spaces and let $T \in \operatorname{Fred}(X, Y)$. Then $T' \in \operatorname{Fred}(Y', X')$ and $\operatorname{ind}(T') = -\operatorname{ind}(T)$.*

Proof. This result follows directly from Theorem 11.6. □

Lemma 11.8. *Let X and Y be Banach spaces and let $T \in L(X, Y)$ with $\operatorname{Image}(T)$ closed. Then the equation $Tx = y$ is solvable if and only if $\psi(y) = 0$ for all $\psi \in \ker(T')$.*

Proof. We note that $y \in \operatorname{Image} T$ is equivalent to $\pi(y) = [y] = 0$, where $\pi\colon Y \to Y/\operatorname{Image} T = \operatorname{coker} T$ is the projection defined in the proof of Theorem 11.6. It follows from Lemma 4.4. that this is then equivalent to the fact that $g([y]) = 0$ for all $g \in (\operatorname{coker} T)'$ and this in turn is equivalent to the fact that $(\rho\psi)([y]) = 0$ for all $\psi \in \ker(T')$, where $\rho\colon \ker(T') \overset{\cong}{\to} (\operatorname{coker} T)'$, $(\rho\psi)[y] = \psi(y)$ is the isomorphism constructed in the proof of Theorem 11.6. □

Lemma 11.9. *Let V be a subspace of the Banach space X. If either*

1. $\dim V < \infty$*, or*
2. *V is closed and* $\dim(X/V) \leq \infty$*,*

then V has a closed complement, i.e. there exists $W \subset X$ closed with $X = V \oplus W$.

Proof. We first assume that $\dim V < \infty$. Then the space $(V, \|\cdot\|_X)$ is complete (see Theorem 1.2) and thus V is closed (see exercise sheet 1). Next we choose a basis $\{v_1, ..., v_n\}$ of V and a dual basis $\{\varphi_1, ..., \varphi_n\}$ of V', i.e. $\varphi_i(v_j) = \delta_{ij}$ for all $1 \leq i, j \leq n$. We extend φ_i to $\tilde{\varphi}_i \in X'$ with the help of Theorem 4.2 and we define the map $P : X \to V$ by

$$Px := \sum_{i=1}^{n} \tilde{\varphi}_i(x) v_i.$$

It follows that

$$\|Px\| \leq n(\sup_{i=1,...,n} \|\bar{\varphi}_i\| \|v_i\|) \|x\|$$

and hence $P \in L(X, V)$. Moreover, we have

$$Pv_j = \sum_{i=1}^{n} \tilde{\varphi}_i(v_j) v_i = v_j$$

and therefore

$$P|_V = \mathrm{id}_V \quad \text{and} \quad P^2 = P.$$

Since for every $x \in X$ we can write

$$x = Px + (x - Px) \in V \oplus \ker P$$

we have that $\ker P$ is a closed complement of V.

In the case $\dim(X/V) < \infty$ we choose a basis $\{[x_1], ..., [x_n]\}$ of X/V and we note that $\operatorname{span}\{x_1, ..., x_n\}$ is a closed complement of V. □

Lemma 11.10. *Let V be a closed subspace of the Banach space X and let $W \subset X$ a finite-dimensional subspace. Then $V + W$ is a closed subspace of X.*

Proof. Without loss of generality we assume that $V \cap W = \{0\}$. Let $x_k = v_k + w_k \in V + W$ and assume that $x_k \to x \in X$. Then we claim that $\|w_k\| \leq c$ for all $k \in \mathbb{N}$ and some $c < \infty$. Namely, if $\|w_k\| =: R_k \to \infty$, then $x_k/R_k \to 0$, $w_k/R_k \to w \in W$ with $\|w\| = 1$ and therefore $V \ni v \leftarrow v_k/R_k = x_k/R_k - w_k/R_k \to -w \in W$ contradicting the fact that $V \cap W = \{0\}$. Since w_k is bounded and x_k converges, we conclude that v_k is bounded and we can extract subsequences so that $w_k \to w \in W$ and $v_k \to v \in V$, which implies $x = v + w \in V + W$. □

Theorem 11.11. *Let X and Y be Banach spaces. Then* $\mathrm{Fred}(X, Y) \subset L(X, Y)$ *is open and the index is locally constant on* $\mathrm{Fred}(X, Y)$.

Proof. Let $T \in \mathrm{Fred}(X,Y)$. Then it follows from Lemma 11.9 that there exist closed subspaces $V \subset X$ and $W \subset Y$ with $\dim W < \infty$ and

$$X = V \oplus \ker T \quad \text{and}$$
$$Y = W \oplus \mathrm{Image}(T).$$

Since V and W are closed they are also complete and hence the space $(V \times W, \|(v,w)\| := \|v\|_X + \|w\|_Y)$ is a Banach space. Next we define the map

$$L(X,Y) \to L(V \times W, Y),\ S \mapsto \tilde{S} \quad \text{with} \quad \tilde{S}(v,w) = Sv + w$$

and we note that this map is continuous. We split the rest of the proof into four steps.

- Step 1: $\tilde{T}$ is invertible.

  For this we note that $T|_V \colon V \to \mathrm{Image}\, T$ is bijective and continuous and it follows from Theorem 5.10 that $T|_V$ is invertible. We note that $\tilde{T}$ can be expressed as

  $$\tilde{T} = \begin{pmatrix} T|_V & 0 \\ 0 & \mathrm{id}_W \end{pmatrix}$$

  where the first row corresponds to elements in V and the second row to elements in W. Similarly the first line represents elements in $\mathrm{Image}(T)$ and the second line in W. From this form it is easy to see that the inverse can be represented as follows

  $$\tilde{T}^{-1} = \begin{pmatrix} (T|_V)^{-1} & 0 \\ 0 & \mathrm{id}_W \end{pmatrix}.$$

- Step 2: If $\|S - T\| < \|\tilde{T}^{-1}\|^{-1}$ then the map $\tilde{S}$ is invertible.

  We note that

  $$(\tilde{S} - \tilde{T})(v,w) = (Sv + w) - (Tv + w) = (S - T)v$$

  and therefore
  $$\|\tilde{S} - \tilde{T}\| \le \|S - T\| < \|\tilde{T}^{-1}\|^{-1}.$$

  The claim now follows from Theorem 5.2.

- Step 3: If $\tilde{S}$ is invertible then $S \in \mathrm{Fred}(X,Y)$.

If $\tilde{S}$ is invertible then we know that $\ker S \cap V = \{0\}$. Thus the projection map $\pi\colon X \to X/V$ maps $\ker S \subset X$ injectively onto $\ker S \subset X/V$ and we conclude that

$$\dim \ker S \leq \dim X/V = \dim \ker T < \infty.$$

By Lemma 11.10 we get that $V \oplus \ker S \subset X$ is a closed subspace with finite codimension and Lemma 11.9 then implies that that there exists a closed subset $U \subset X$ with $X = V \oplus \ker S \oplus U$ and $\dim U < \infty$. Next, we know that $S(V) = \tilde{S}(V \times \{0\})$ is closed, since $\tilde{S}^{-1}(\tilde{S}(V \times \{0\})) = V \times \{0\}$ and $\tilde{S}^{-1}$ is continuous by Theorem 5.10 and $V \times \{0\}$ is closed. Therefore

$$\operatorname{Image} S = S(X) = S(V) \oplus S(U)$$

is closed by Lemma 11.10. It remains to show that $\dim(\operatorname{coker} S) < \infty$.

Since $\tilde{S}$ is invertible we know that $Y = S(V) \oplus W$ (note that $S(V) \cap W = \{0\}$ since otherwise $\tilde{S}$ would not be injective). We consider the projection map

$$p\colon Y/S(V) \to Y/S(X),\ p(y + S(V)) = y + S(X)$$

and we note that p is well-defined since $S(V) \subset S(X)$. Since p is surjective it follows that

$$\dim(\operatorname{coker}(S)) = \dim(Y/S(X)) \leq \dim(Y/S(V)) = \dim W < \infty.$$

Note that so far we shown that $\operatorname{Fred}(X, Y)$ is open in $L(X, Y)$.

- Step 4: If $\tilde{S}$ is invertible then $\operatorname{ind} S = \operatorname{ind} T$.

  We recall that

$$X = V \oplus \ker T = V \oplus \ker S \oplus U$$

  and therefore

$$\dim(\ker S) = \dim(\ker T) - \dim U.$$

  Additionally, we have that

$$\dim(Y/S(V)) = \dim W = \dim(Y/T(X)) = \dim(\operatorname{coker} T).$$

  Consider again the projection $p\colon Y/S(V) \to Y/S(X)$ defined in step 3. It follows that $\ker(p) = S(X)/S(V) \subset Y/S(V)$ and (note that all vector spaces considered here are finite-dimensional)

$$\begin{aligned}\dim(\operatorname{coker} S) = \dim(\operatorname{Image}(p)) &= \dim(Y/S(X)) \\ &= \dim(Y/S(V)) - \dim(\ker(p)) \\ &= \dim(\operatorname{coker} T) - \dim(S(U)) \\ &= \dim(\operatorname{coker} T) - \dim U\end{aligned}$$

as $\ker(S) \cap U = \{0\}$. Hence we conclude

$$\operatorname{ind} S = \operatorname{ind} T.$$

□

The following important result of Riesz combines the theory of compact and Fredholm operators.

Theorem 11.12 (on compact operators by Riesz)**.** *Let X be a Banach space and let $K \in K(X,X)$. Then $T := \mathrm{id}_X - K \in \operatorname{Fred}(X,X)$ with $\operatorname{ind}(T) = 0$.*

Proof. We show this result in five steps.

1. We have $\dim(\ker(T)) < \infty$.

   Let $x_j \in \ker(T)$ with $\|x_j\| \leq 1$. Since $K \in K(X,X)$ it follows that

   $$x_j = Kx_j \to y \in X$$

   up to a subsequence. As $\ker(T)$ is closed w get $y \in \ker T$ with $\|y\| \leq 1$ and hence the set $\{x \in \ker(T) \colon \|x\| \leq 1\}$ is sequentially compact. By Theorem 3.2 and Theorem 3.4 it follows that $\dim(\ker(T)) < \infty$.

2. By Lemma 11.9 there exists a closed set $V \subset X$ so that $X = V \oplus \ker T$. We claim that there exists a constant $m > 0$ so that $\|Tv\| \geq m\|v\|$ for all $v \in V$.

   We show this estimate by contradiction. Thus we assume that there exists a sequence $v_j \in V$ with $\|v_j\| = 1$ and so that

   $$\|Tv_j\| \leq \frac{1}{j}\|v_j\|$$

   for all $j \in \mathbb{N}$. As the map K is compact it follows that $Kv_j \to y \in X$ up to the choice of a subsequence. Hence, as V is closed, $v_j = Tv_j + Kv_j \to y \in V$

and $\|y\| = 1$. Moreover, we have that

$$Ty = \lim_{j\to\infty} Tv_j = 0$$

and hence $y \in \ker T$. But this contradicts the fact that $V \cap \ker(T) = \{0\}$.

3. We claim that Image(T) is complete and therefore closed.

   We let $(y_j) \in \operatorname{Image} T$ be a Cauchy sequence with $y_j = Tx_j$. Without loss of generality we assume that $x_j \in V$. It follows from Step 2. that

$$\|x_j - x_k\| \leq \frac{1}{m}\|Tx_j - Tx_k\| = \frac{1}{m}\|y_j - y_k\| < \varepsilon$$

   for j, k large enough. Hence (x_j) is a Cauchy sequence in X and $x_j \to x \in X$ which gives $y_j = Tx_j \to Tx =: y$.

4. We have $\dim \operatorname{coker} T < \infty$.

   Consider the adjoint map $T' = (\mathrm{id}_X - K)' = \mathrm{id}_{X'} - K'$. From Theorem 11.4 we know that $K' \in K(X', X')$ and thus we can apply Step 1. and Theorem 11.6 in order to conclude that

$$\dim(\operatorname{coker}(T)) = \dim(\operatorname{coker}(T))' = \dim(\ker(T')) < \infty.$$

5. We have $\operatorname{ind} T = 0$.

   Let $T(t) = \mathrm{id}_X - tK$ with $0 \leq t \leq 1$ be a continuous curve in $L(X, X)$. By the first four steps we get that $T(t) \in \operatorname{Fred}(X, X)$ for all $t \in [0, 1]$ and hence $\operatorname{ind}(T(t)) \in \mathbb{Z}$ is independent of t by Theorem 11.11. Therefore we get

$$\operatorname{ind} T = \operatorname{ind}(T(1)) = \operatorname{ind}(T(0)) = \operatorname{ind}(\mathrm{id}_X) = 0.$$

□

Lemma 11.13. *Let X and Y be Banach spaces. Then $T \in \operatorname{Fred}(X, Y)$ if and only if there exist $S_1, S_2 \in L(Y, X)$ with*

$$\mathrm{id}_X - S_1 T \in K(X, X), \quad \textit{and} \quad \mathrm{id}_Y - TS_2 \in K(Y, Y).$$

Additionally, one can choose $S = S_1 = S_2$ and such a map $S \in L(Y, X)$ is called a ***parametrix*** *for T.*

Proof. "$\Rightarrow$":

Let $T \in \operatorname{Fred}(X,Y)$. It follows from Lemma 10.9 that there exist closed subsets $V \subset X$ and $W \subset Y$ so that

$$X = V \oplus \ker(T) \quad \text{and} \quad Y = \operatorname{Image}(T) \oplus W.$$

As in the proof of Theorem 11.11 we consider the map

$$\tilde{T} \colon (V \times W, \|(v,w)\| = \|v\|_X + \|w\|_Y) \to Y, \quad \tilde{T}(v,w) = Tv + w.$$

The argument from step 1 in proof of Theorem 11.11 implies that $\tilde{T}$ is invertible and we let $\tilde{S} = \tilde{T}^{-1} \colon Y \to V \times W$ be the inverse map and let $P_V : V \times W$, $P_V(v,w) = v$ be the projection onto the first factor. We define $S \in L(Y,X)$ by $S = P_V \tilde{S} \colon Y \to V \subset X$ and we note that for $v \in V$ and $x \in \ker T$ we have

$$(\mathrm{id}_X - ST)(v + x) = v + x - P_V \tilde{S}\tilde{T}(v,0) = v + x - v = x.$$

Thus the map $\mathrm{id}_X - ST$ is a projection onto the finite-dimensional space $\ker T$ and hence $\mathrm{id}_X - ST \in K(X,X)$. Next, we let $v \in V$, $w \in W$ and we calculate

$$(\mathrm{id}_Y - TS)(Tv + w) = Tv + w - TP_V\tilde{S}\tilde{T}(v,w) = Tv + w - Tv = w$$

and therefore $\mathrm{id}_Y - TS$ is a projection onto the finite-dimensional space W and we conclude that $\mathrm{id}_Y - TS \in K(Y,Y)$.

"$\Leftarrow$":

We assume that $S_1 T = \mathrm{id}_X - K_1$, where $K_1 \in K(X,X)$. It follows from Theorem 11.12 that $S_1 T \in \operatorname{Fred}(X,X)$ and hence $\ker(T) \subset \ker(S_1 T)$ is finite-dimensional. Additionally, we let $TS_2 = \mathrm{id} - K_2$ with $K_2 \in K(Y,Y)$ and Theorem 11.12 again implies that $TS_2 \in \operatorname{Fred}(Y,Y)$ which yields that $\operatorname{Image}(TS_2)$ is closed with finite codimension. By Lemma 11.9 there exists a finite-dimensional closed subspace $Z \subset Y$ with

$$\operatorname{Image}(T) = \operatorname{Image}(TS_2) \oplus Z$$

and thus it follows from Lemma 11.10 that $\operatorname{Image}(T)$ is closed. Finally, we note that

$$p \colon Y/\operatorname{Image}(TS_2) \to Y/\operatorname{Image}(T)$$

is a projection and hence $\dim(\operatorname{coker}(T)) \leq \dim(\operatorname{coker}(TS_2)) < \infty$. □

In the next Lemma we strengthen the Theorem of Riesz for compact operators.

Lemma 11.14. *Let X and Y be Banach spaces and let $T_0 \in \operatorname{Fred}(X,Y)$, $K \in K(X,Y)$. Then $T := T_0 + K \in \operatorname{Fred}(X,Y)$ with $\operatorname{ind} T = \operatorname{ind} T_0$.*

Proof. Let $S_0 \in L(Y,X)$ be a parametrix for T_0, i.e. we have that there exist $K_X \in K(X,X)$ and $K_Y \in K(Y,Y)$ with

$$\mathrm{id}_X - S_0T_0 = K_X \quad \text{and} \quad \mathrm{id}_Y - T_0S_0 = K_Y \in K(Y,Y)$$

We conclude that

$$\begin{aligned}\mathrm{id}_X - S_0T &= \mathrm{id}_X - S_0T_0 - S_0K \in K(X,X) \quad \text{and}\\ \mathrm{id}_Y - TS_0 &= \mathrm{id}_Y - T_0S_0 - KS_0 \in K(Y,Y).\end{aligned}$$

where we used Theorem 11.4. Hence S_0 is also a parametrix for T and it follows from Lemma 10.13 that $T \in \operatorname{Fred}(X,Y)$. The claim about the index follows as in the proof of step 5 in Theorem 11.12. □

Theorem 11.15. *Let X, Y, Z be Banach spaces and let $T_1 \in \operatorname{Fred}(X,Y)$, $T_2 \in \operatorname{Fred}(Y,Z)$. Then $T_2T_1 \in \operatorname{Fred}(X,Z)$ and $\operatorname{ind}(T_2T_1) = \operatorname{ind}(T_1) + \operatorname{ind}(T_2)$.*

Proof. We first show that $T := T_2T_1 \in \operatorname{Fred}(X,Z)$. By Lemma 11.13 there exist $S_1 \in L(Y,X)$ and $S_2 \in L(Z,Y)$ so that

$$\begin{aligned}\mathrm{id}_X - S_1T_1 = K_1 \in K(X,X),\ \mathrm{id}_Y - T_1S_1 = \tilde{K}_1 \in K(Y,Y)\\ \mathrm{id}_Y - S_2T_2 = K_2 \in K(Y,Y),\ \mathrm{id}_Z - T_2S_2 = \tilde{K}_2 \in K(Z,Z).\end{aligned}$$

For $S := S_1S_2$ we calculate

$$\begin{aligned}\mathrm{id}_X - ST &= \mathrm{id}_X - S_1S_2T_2T_1\\ &= \mathrm{id}_X - S_1(\mathrm{id}_Y - K_2)T_1\\ &= \mathrm{id}_X - S_1T_1 + S_1K_2T_1\\ &= K_1 + S_1K_2T_1 \in K(X,X),\end{aligned}$$

where we used Theorem 11.4, and

$$\begin{aligned}\mathrm{id}_Z - TS &= \mathrm{id}_Z - T_2T_1S_1S_2\\ &= \mathrm{id}_Z - T_2(\mathrm{id}_Y - \tilde{K}_1)S_2\\ &= \mathrm{id}_Z - T_2S_2 + T_2\tilde{K}_1S_2\\ &= \tilde{K}_2 + T_2\tilde{K}_1S_2 \in K(Z,Z)\end{aligned}$$

again by Theorem 11.4. Therefore S is a parametrix for T and hence Lemma 11.13 implies that $T \in \operatorname{Fred}(X,Z)$.

In order to show the claim about the index of T, we look at the following sequences between finite-dimensional vector spaces

$$\{0\} \longrightarrow \ker(T_1) \subset \ker(T_2T_1) \xrightarrow{T_1} \ker(T_2) \cap \mathrm{Image}(T_1) \longrightarrow \{0\}\ .$$

$$\{0\} \longrightarrow \frac{\mathrm{Image}(T_2)}{\mathrm{Image}(T_2T_1)} \subset \frac{Z}{\mathrm{Image}(T_2T_1)} \xrightarrow{p} \frac{Z}{\mathrm{Image}(T_2)} \longrightarrow \{0\}$$

$$\{0\} \longrightarrow \frac{\mathrm{Image}(T_1)+\ker(T_2)}{\mathrm{Image}(T_1)} \subset \frac{Y}{\mathrm{Image}(T_1)} \xrightarrow{T_2} \frac{\mathrm{Image}(T_2)}{\mathrm{Image}(T_2T_1)} \longrightarrow \{0\}$$

$$\{0\} \longrightarrow \ker(T_2) \cap \mathrm{Image}(T_1) \subset \ker(T_2) \xrightarrow{\pi} \frac{\mathrm{Image}(T_1)+\ker(T_2)}{\mathrm{Image}(T_1)} \longrightarrow \{0\}$$

These four sequences are exact since at each point we have ker = Image. Recall that a sequence

$$\{0\} \longrightarrow \{0\} \xrightarrow{f_0} V_1 \xrightarrow{f_2} V_2 \longrightarrow \ldots \xrightarrow{f_m} V_{m+1} = \{0\} \longrightarrow \{0\}$$

is called **exact**, if Image $f_j = \ker f_{j+1}$ for all $j = 1, ..., m$. It then follows that

$$\sum_{j=1}^{m} (-1)^i \dim V_i = 0$$

since $\dim V_j = \dim(\ker(f_j)) + \dim(\mathrm{Image}(f_j)) = \dim(\ker(f_j)) + \dim(\ker(f_{j+1}))$.

In our situation we have that

$$\begin{aligned}
0 &= \dim(\ker(T_1)) - \dim(\ker(T_2T_1)) + \dim(\ker(T_2) \cap \mathrm{Image}(T_1)) \\
0 &= -\dim(\frac{\mathrm{Image}(T_2)}{\mathrm{Image}(T_2T_1)}) + \dim(\frac{Z}{\mathrm{Image}(T_2T_1)}) - \dim(\frac{Z}{\mathrm{Image}(T_2)}) \\
0 &= \dim(\frac{\mathrm{Image}(T_1) + \ker(T_2)}{\mathrm{Image}(T_1)}) - \dim(\frac{Y}{\mathrm{Image}(T_1)}) + \dim(\frac{\mathrm{Image}(T_2)}{\mathrm{Image}(T_2T_1)}) \\
0 &= -\dim(\ker(T_2) \cap \mathrm{Image}(T_1)) + \dim(\ker(T_2)) - \dim(\frac{\mathrm{Image}(T_1) + \ker(T_2)}{\mathrm{Image}(T_1)}),
\end{aligned}$$

where we note that we are always free to add the trivial vector space to each sequence and hence we can decide with which sign we start the dimension formula for the exact sequence. Adding these for equations yields that

$$\begin{aligned}
\dim(\ker(T_1)) - &\dim(\mathrm{coker}(T_1)) + \dim(\ker(T_2)) - \dim(\mathrm{coker}(T_2)) \\
&= \dim(\ker(T_2T_1)) - \dim(\mathrm{coker}(T_2T_1))
\end{aligned}$$

and hence

$$\operatorname{ind} T_1 + \operatorname{ind} T_2 = \operatorname{ind} T_2 T_1$$

as claimed. □

Next we are coming back to elliptic boundary value problems and we want to use the abstract results we just obtained in order to study the operator $L = L_0 + K$, which we introduced at the beginning of this chapter. Recall that for $\Omega \subset \mathbb{R}^n$ open and bounded the operator $L_0 : W_0^{1,2}(\Omega) \to W_0^{1,2}(\Omega)'$ is defined by

$$L_0 v = -div(aDv),$$

where $a \in L^\infty(\Omega, M_n(\mathbb{R}))$ is elliptic with constant $\mu > 0$. By Theorem 10.15 L_0 is an isomorphism. Moreover, the operator $K : W_0^{1,2}(\Omega) \to W_0^{1,2}(\Omega)'$ is defined by

$$Kv = -div(bv) + \langle c, Dv \rangle + qv,$$

$q \in L^\infty(\Omega)$ and $b, c \in L^\infty(\Omega, \mathbb{R}^n)$. Our goal is to show that K is compact since Lemma 11.14 then implies that $L \in \operatorname{Fred}(W_0^{1,2}(\Omega), W_0^{1,2}(\Omega)')$ with $\operatorname{ind} L = 0$.

We start with a Lemma.

Lemma 11.16. *Let $1 \leq p < \infty$ and let $u \in W^{1,p}(\mathbb{R}^n)$. Then we have that*

1. *$\|u \circ \tau_h - u\|_{L^p} \leq |h| \|Du\|_{L^p}$ for all $h \in \mathbb{R}^n$ and where $\tau_h(x) = x + h$ for all $x, h \in \mathbb{R}^n$.*

2. *$\|u - u_\rho\|_{L^p} \leq \rho \|Du\|_{L^p}$, where u_ρ is defined as in Theorem 10.4.*

Proof. It follows from Theorem 10.9 that we can assume without loss of generality that $u \in C_c^\infty(\mathbb{R}^n)$.

1. We estimate for all $h \in \mathbb{R}^n$

$$\begin{aligned}\int_{\mathbb{R}^n} |u(x+h) - u(x)|^p \, dx &= \int_{\mathbb{R}^n} \left| \int_0^1 \frac{d}{ds} u(x+sh) \, ds \right|^p \, dx \\ &= \int_{\mathbb{R}^n} \left| \int_0^1 \langle Du(x+sh), h \rangle \right|^p \, dx \\ &\leq |h|^p \int_{\mathbb{R}^n} \int_0^1 |Du(x+sh)|^p \, ds \, dx \\ &= |h|^p \|Du\|_{L^p}^p,\end{aligned}$$

which shows statement 1.

2. For every $\rho > 0$ we get with the help of the transformation formula, Hölder's inequality and the Theorem of Fubini that

$$\begin{aligned}\int_{\mathbb{R}^n} |u(x) - u_\rho(x)|^p \, dx =& \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \eta_\rho(x-y)(u(x)-u(y)) \, dy \right|^p \, dx \\ =& \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \eta(z)(u(x)-u(x-\rho z)) \, dz \right|^p \, dx \\ \leq& \int_{\mathbb{R}^n} \left(\int \eta(z)|u(x)-u(x-\rho z)|^p \, dz \right) dx \\ =& \int_{B_1(0)} \eta(z) \left(\int_{\mathbb{R}^n} |u(x)-u(x-\rho z)|^p \, dx \right) dz \\ \leq& \rho^p \|Du\|^p_{L^p}\end{aligned}$$

by statement 1.

□

Theorem 11.17 (Rellich embedding theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded and let $1 \leq p < \infty$. Then the inclusion $W_0^{1,p}(\Omega) \subset L^p(\Omega)$ is compact.*

Proof. We have to show that every sequence $u_k \in W_0^{1,p}(\Omega)$ with $\|u_k\|_{W^{1,p}} \leq C$ has a subsequence such that $u_k \to u$ in $L^p(\Omega)$. Without loss of generality we assume again that $u_k \in C_c^\infty(\Omega)$. Let η be as in Theorem 10.4 and extend u_k by 0 to all of $\mathbb{R}^n$.

For $j \geq 0$ we have

$$\begin{aligned}D^j(\eta_\rho * u_k)(x) =& \int D_x^j \eta_r ho(x-y) u_k(y) \, dy \\ =& \rho^{-(j+n)} \int (D_x^j \eta) \left(\frac{x-y}{\rho} \right) u_k(y) \, dy.\end{aligned}$$

Therefore

$$\|D^j(\eta_\rho * u_k)\|_{L^\infty} \leq c(j) \rho^{-j-n} \rho^{n-\frac{n}{p}} \|u_k\|_{L^p} \leq c(j,n) \rho^{-(j+n/p)}$$

and

$$\operatorname{spt}(\eta_\rho * u_k) \subset \{x \colon \operatorname{dist}(\Omega, x) \leq \rho\}.$$

For $\rho > 0$ we conclude, using Theorem 3.8, that there exists a subsequence such that $(eta_\rho * u_k) \to u^\rho$ in $C^1(\mathbb{R}^n)$. Next we choose a sequence $\rho_i \to 0$ and successively subsequences $\{k_j^1\} \supset \{k_j^2\} \supset \ldots$ for ρ_1, ρ_2,... such that the C^1-convergence remains true.

The diagonal sequence satisfies $(\eta_\rho * u_k) \to u^\rho$ in $C^1(\mathbb{R}^n)$ for all $\rho \in \{\rho_1, \rho_2, ...\}$.

Moreover, we have that $\operatorname{spt}(u^\rho) \subset \{x\colon \operatorname{dist}(x,\Omega) \le \rho\}$ and

$$\|u^\rho\|_{L^p} = \lim_{k\to\infty} \|\eta_\rho * u_k\|_{L^p} \le C \quad \text{and}$$
$$\|Du^\rho\|_{L^p} = \lim_{k\to\infty} \|\eta_\rho * Du_k\|_{L^p} \le C.$$

For $\rho, \rho' \in \{\rho_1, \rho_2, ...\}$ we get

$$\begin{aligned}\|u^\rho - u^{\rho'}\|_{L^p} &= \lim_{k\to\infty} \|\eta_\rho * u_k - \eta_{\rho'} * u_k\|_{L^p} \\ &\le \limsup_{k\to\infty} (\|\eta_\rho * u_k - u_k\|_{L^p} + \|\eta_{\rho'} * u_k - u_k\|_{L^p}) \\ &\le \limsup_{k\to\infty} (\rho\|Du_k\|_{L^p} + \rho'\|Du_k\|_{L^p}) \le C(\rho+\rho'),\end{aligned}$$

where we used Lemma 11.16. Hence $\{u^\rho\}$ is a Cauchy sequence in $L^p(\mathbb{R}^n)$ and $u^\rho \to u \in L^p(\mathbb{R}^n)$. We have that $\operatorname{spt}(u) \subset \overline{\Omega}$ and $\|u - u^{\rho_i}\|_{L^p} \le C\rho_i$. Therefore we get

$$\begin{aligned}\|u - u_k\|_{L^p} \le& \|u - u^{\rho_i}\|_{L^p} + \|u^{\rho_i} - \eta_{\rho_i} * u_k\|_{L^p} + \|\eta_{\rho_i} * u_k - u_k\|_{L^p} \\ \le& C\rho_i + \|u^{\rho_i} - \eta_{\rho_i} * u_k\|_{L^p} \to 0\end{aligned}$$

as $i, k \to \infty$. □

<u>Remark:</u> If $\Omega \subset \mathbb{R}^n$ is open and bounded with C^1-boundary, then the inclusion $W^{1,p}(\Omega) \subset L^p(\Omega)$ is also compact.

Lemma 11.18. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then the operator $K\colon W_0^{1,2}(\Omega) \to (W_0^{1,2}(\Omega))'$,*

$$(Kv)(u) = \int_\Omega (\langle Du, bv\rangle + u(\langle c, Dv\rangle + qv)),$$

with $b, c \in L^\infty(\Omega, \mathbb{R}^n)$ and $q \in L^\infty(\Omega)$ is compact.

Proof. We denote by $I\colon W_0^{1,2}(\Omega) \to L^2(\Omega)$ the inclusion map and we note that I is compact by Theorem 11.17. Moreover, we define the map

$$J\colon L^2(\Omega) \to (L^2(\Omega))',\ (Jv)(u) = \int uv$$

and we recall that J is continuous with $\|J\| = 1$. It follows from Theorem 11.4 that

$$I' \circ J\colon L^2(\Omega) \to (W_0^{1,2}(\Omega))'$$

is compact. Next we define the operators

$$\begin{aligned} L^2(\Omega) \to& W_0^{1,2}(\Omega)', \ (Bv)(u) = \int \langle b, Du\rangle v \\ W_0^{1,2}(\Omega) \to& L^2(\Omega), \ Cv = \langle c, Dv\rangle \\ L^2(\Omega) \to& L^2(\Omega), \ Qv = qv \end{aligned}$$

and we observe that all of them are continuous. Moreover, we have that

$$K = B \circ I + I' \circ J \circ C + I' \circ J \circ Q \circ I$$

and it follows again from Theorem 11.4 that K is compact. □

Our main result for elliptic boundary value problems is the following

Theorem 11.19. *Let $L = L_0 + K : W_0^{1,2}(\Omega) \to (W_0^{1,2}(\Omega))'$ be as above,. Then $L \in \mathrm{Fred}(W_0^{1,2}(\Omega), W_0^{1,2}(\Omega)')$ and $\operatorname{ind} L = 0$. Thus the so called* ***Fredholm alternative*** *holds: L is injective if and only if L is surjective.*

Proof. We know from Theorem 10.15 that L_0 and from Lemma 11.18 that K is compact. Hence the result is a direct consequence of Lemma 10.14. □

Next we define the so called formal adjoint operator of L. For this we consider the diagram

$$\begin{array}{ccc} W_0^{1,2}(\Omega) & & \\ {\scriptstyle J}\downarrow & \searrow^{L^* := L' \circ J} & \\ W_0^{1,2}(\Omega)'' & \xrightarrow[L']{} & W_0^{1,2}(\Omega)' \end{array},$$

where $J : W_0^{1,2}(\Omega) \to W_0^{1,2}(\Omega)''$ is the canonical embedding. We have that

$$(L^*v)(u) = (L'Jv)(u) = (Jv)(Lu) = (Lu)(v)$$

for all $u, v \in W_0^{1,2}(\Omega)$. Moreover,

$$(Lu)(v) = \int_\Omega (\langle Dv, aDu\rangle + \langle Dv, bu\rangle + v\langle c, Du\rangle + qvu)$$

and thus

$$(L^*v)(u) = \int_\Omega (\langle Du, a^*Dv\rangle + \langle Du, cv\rangle + u\langle b, Dv\rangle + quv),$$

or

$$L^*v = -div(a^*Dv) - div(cv) + \langle b, Dv\rangle + qv.$$

The following theorem is a variant of Lemma 11.8 for the special case of operators considered here.

Theorem 11.20. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let L, L^* be as above and let $\varphi \in W_0^{1,2}(\Omega)'$. Then the following two statements are equivalent.*

1. *$Lv = \varphi$ has a solution $v \in W_0^{1,2}(\Omega)$.*

2. *$\varphi(u) = 0$ for all $u \in \ker(L^*)$.*

Proof. See the exercises. □

12 Spectral theory for compact operators

In this chapter we let X be a Banach space over $\mathbb{K}$ and we mostly consider the case $\mathbb{K} = \mathbb{C}$.

Definition 12.1. *Let $A \in L(X, X)$.*

1. *The* ***resolvent set*** *is given by*

$$\varrho(A) = \{\lambda \in \mathbb{K} \colon \lambda \operatorname{id} - A \textit{ is invertible}\}$$

   $\lambda \in \varrho(A)$ is called regular with respect to A.

2. *The set $\mathbb{K} \setminus \varrho(A) =: \sigma(A)$ is called the* ***spectrum*** *of A and $\lambda \in \sigma(A)$ is called* ***spectral value****. $\sigma(A)$ is decomposed into three parts:*

   a) *The* ***point spectrum****:*

   $$\sigma_p(A) = \{\lambda \in \sigma(A) \colon \lambda \operatorname{id} - A \textit{ is not injective}\}.$$

   b) *The* ***continuous spectrum****:*

   $$\sigma_c(A) = \{\lambda \in \sigma(A) \colon \lambda \operatorname{id} - A \textit{ is injective but not surjective, } \operatorname{Image}(\lambda \operatorname{id} - A) \textit{ is dense in } X\}.$$

   c) *The* ***residual spectrum****:*

   $$\sigma_r(A) = \{\lambda \in \sigma(A) \colon \lambda \operatorname{id} - A \textit{ is injective but not surjective, } \operatorname{Image}(\lambda \operatorname{id} - A) \textit{ is not dense in } X\}.$$

   It follows from Theorem 5.10 that if $\lambda \in \mathbb{K} \setminus \sigma(A)$, then $\lambda \operatorname{id} - A$ is bijective and hence $\lambda \in \varrho(A)$.

3. *The* ***resolvent function*** *$R_A \colon \varrho(A) \to L(X, X)$ is defined by*

$$R_A(\lambda) = (\lambda \operatorname{id} - A)^{-1}.$$

In the next theorem we derive some general facts about the spectrum and the resolvent set.

Theorem 12.2. *Let $X \neq \{0\}$ be a Banach space over $\mathbb{C}$ and let $A \in L(X,X)$. Then*

1. $\operatorname{dist}(\lambda, \sigma(A)) \geq \|R_A(\lambda)\|^{-1}$ *for all $\lambda \in \varrho(A)$ and thus $\varrho(A)$ is open. Moreover, for all $\Lambda \in L(X,X)'$ the map $\lambda \in \varrho(A) \subset \mathbb{C} \mapsto \Lambda R_A(\lambda) \in \mathbb{C}$ is holomorphic.*

2. $\sup\{|\lambda| \colon \lambda \in \sigma(A)\} = \lim_{n\to\infty} \|A^n\|^{1/n} \leq \|A\| < \infty$ *which implies that $\sigma(A) \neq \emptyset$ is compact. The number $\sup\{|\lambda| \colon \lambda \in \sigma(A)\}$ is called the **spectral radius** of A.*

Proof. 1. Let $\lambda_0 \in \varrho(A)$, i.e. $\lambda_0 \operatorname{id} - A$ is invertible. Then it follows from Theorem 5.2 that for $\lambda \in \mathbb{C}$ the map

$$\lambda \operatorname{id} - A = \lambda_0 \operatorname{id} - A - (\lambda_0 - \lambda)\operatorname{id} = (\lambda_0 \operatorname{id} - A)(\operatorname{id} - (\lambda_0 - \lambda) R_A(\lambda_0))$$

is invertible if $|\lambda - \lambda_0| < \|R_A(\lambda_0)\|^{-1}$, and therefore $\operatorname{dist}(\lambda_0, \sigma(A)) \geq \|R_A(\lambda_0)\|^{-1}$. Additionally, for $|\lambda - \lambda_0| < \|R_A(\lambda_0)\|^{-1}$, we get

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

Hence we get for all $\Lambda \in L(X,X)'$ that

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

and hence this function is holomorphic for all $\Lambda \in L(X,X)'$ and all $\lambda \in \varrho(A)$.

2. We define $R := \sup\{|\lambda| \colon \lambda \in \sigma(A)\}$ and we note that for $|\lambda| > \|A\|$ we have that the map

$$\lambda \operatorname{id} - A = \lambda \left(\operatorname{id} - \frac{A}{\lambda} \right)$$

is invertible by Theorem 5.2 and thus

$$R_A(\lambda) = \left(\frac{1}{\lambda}\right) \sum_{n=0}^{\infty} \lambda^{-n} A^n. \tag{12.1}$$

This shows that $\sigma(A) \subset \{\lambda \colon |\lambda \leq \|A\|\}$.

Next we claim that $\lambda \in \sigma(A)$ implies $\lambda^n \in \sigma(A^n)$. This follows from the fact that

$$\lambda^n \operatorname{id} - A^n = \sum_{j=1}^{n} (\lambda^{n-j+1} A^{j-1} - \lambda^{n-j} A^j) = \begin{cases} (\lambda \operatorname{id} - A) \sum_{j=1}^{n} \lambda^{n-j} A^{j-1} \\ \left(\sum_{j=1}^{n} \lambda^{n-j} A^{j-1} \right) (\lambda \operatorname{id} - A) \end{cases}$$

Now, if $\lambda \operatorname{id} - A$ is not surjective we use the first inequality, and if $\lambda \operatorname{id} - A$ is not injective we use the second equality in order to get that $\lambda^n \in \sigma(A^n)$.

Hence it follows from the above estimate that $\lambda \in \sigma(A)$ implies $\lambda^n \in \sigma(A^n)$ and therefore $|\lambda|^n \leq \|A^n\|$ which gives

$$|\lambda| \leq \liminf_{n\to\infty} \|A^n\|^{1/n}.$$

For $s > R$ and an arbitrary $\Lambda \in L(X,X)'$ we now consider the holomorphic function $\Lambda(R_A(\lambda))$ and we note that the integral

$$-\frac{1}{2\pi i}\int_{\partial B_s(0)} \lambda^\mu \Lambda(R_A(\lambda))d\lambda.$$

is independent of s as long as $s > R$ by Cauchy's integral theorem. Next, we let $s > \|A\|$ and we calculate with the help of (12.1)

$$-\frac{1}{2\pi i}\int_{\partial B_s(0)} \lambda^\mu \Lambda(R_A(\lambda))d\lambda = \sum_{n=0}^{\infty} \Lambda(A^n)\frac{1}{2\pi i}\int_{\partial B_s(0)} \frac{d\lambda}{\lambda^{n+1-\mu}} = \Lambda(A^\mu).$$

On the other hand we have for all $s > R$

$$\left|\frac{1}{2\pi i}\int_{\partial B_s(0)} \lambda^\mu \Lambda(R_A(\lambda))d\lambda\right| \leq s^{\mu+1}\|\Lambda\| \sup_{|\lambda|=s} |R_A(\lambda)|.$$

Together these estimates show that for all $s > R$ and all $\Lambda \in L(X,X)'$ we have

$$|\Lambda(A^\mu)| \leq s^{\mu+1}\|\Lambda\| \sup_{|\lambda|=s} |R_A(\lambda)|.$$

By Lemma 4.4 there exists $\Lambda \in L(X,X)'$ such that $\Lambda(A^\mu) = \|A^\mu\|$ and $\|\Lambda\| = 1$. Hence we get for all $s > R$

$$\|A^\mu\| \leq s^{\mu+1} \sup_{|\lambda|=s} |R_A(\lambda)|$$

and therefore

$$\limsup_{\mu\to\infty} \|A^\mu\|^{1/\mu} \leq R.$$

In particular we conclude that

$$R = \lim_{n\to\infty} \|A^n\|^{1/n} \leq \|A\|.$$

If we now assume that $\sigma(A) = \emptyset$ then $\rho(A) = \mathbb{C}$ and thus $\lambda \operatorname{id} - A$ is invertible for all $\lambda \in \mathbb{C}$. Therefore

$$-\frac{1}{2\pi i}\int_{\partial B_s(0)} \lambda^\mu \Lambda(R_A(\lambda))d\lambda = 0$$

for all $s \geq 0$ and by using the above argument with $\mu = 1$ we get $A = 0$. But $\ker(A) = \{0\}$ as $A - 0\,\mathrm{id}$ is invertible and this gives a contradiction. □

Theorem 12.3 (Spectral theorem for compact operators)**.** *Let X be a Banach space over $\mathbb{C}$ and let $K \in L(X, X)$ compact. Then we have that*

1. $\sigma(K)\backslash\{0\} = \sigma_p(K)\backslash\{0\}$.
2. *$\sigma(K)$ is a compact, at most countable subset of $\mathbb{C}$. The only possible accumulation point is 0.*
3. *If $\dim X = \infty$ then $0 \in \sigma(K)$.*

Proof. 1. Let $\lambda \in \sigma(K)\backslash\{0\}$. Then it follows from Theorem 11.12 that

$$\lambda\,\mathrm{id} - K = \lambda\left(\mathrm{id} - \frac{K}{\lambda}\right) \in \mathrm{Fred}(X, X)$$

with $\mathrm{ind}(\lambda\,\mathrm{id} - K) = 0$, which shows that $\lambda \in \sigma_p(K)$.

2. Assume that $\mu \neq 0$ is an accumulation point of $\sigma(K)$. Then it follows that there exists a sequence $\lambda_n \in \sigma(K)\backslash\{0\}$ with $\lambda_n \to \mu$ and we assume without loss of generality that $\lambda_n \neq \lambda_m$ for all $m \neq n$. By statement 1. there exists $x_n \in X\backslash\{0\}$ with $Kx_n = \lambda_n x_n$ and we define for all $n \in \mathbb{N}$

$$X_n := \mathrm{span}\,\{x_1, ..., x_n\}.$$

It follows that X_n is closed and invariant under K, i.e. $KX_n \subset X_n$. Moreover, $\dim X_n = n$ since if we assume that $x_n = \sum_{i=1}^{n-1} \alpha_i x_i$, then $0 = Kx_n - \lambda_n x_n = \sum_{i=1}^{n-1} \alpha_i(\lambda_i - \lambda_n)$ and thus by induction $\alpha_i(\lambda_i - \lambda_n) = 0$ which implies $\alpha_i = 0$. By Lemma 3.3 there exists $z_n \in X_n$ with $\|z_n\| = 1$ and $\mathrm{dist}(z_n, X_{n-1}) \geq \frac{1}{2}$. We have $z_n = \alpha_n x_n + y_n$ for some $y_n \in X_{n-1}$ and some $\alpha_n \in \mathbb{C}$ and we note that

$$\left(\mathrm{id} - \frac{1}{\lambda_n}K\right) z_n = \left(\mathrm{id} - \frac{1}{\lambda_n}K\right) y_n \in X_{n-1}$$

and therefore

$$\left(\mathrm{id} - \frac{1}{\lambda_n}K\right) z_n + \frac{1}{\lambda_n}Kz_i \in X_{n-1}$$

for all $1 \leq i \leq n-1$. Thus

$$\left\|z_n - \left(\mathrm{id} - \frac{1}{\lambda_n}K\right) z_n - \frac{1}{\lambda_n}Kz_i\right\| \geq \frac{1}{2}$$

for all $1 \leq i \leq n-1$ and we conclude that

$$\|Kz_n - Kz_i\| \geq \frac{|\lambda_n|}{2} \geq \frac{|\mu|}{4} > 0$$

for all $1 \leq i \leq n-1$ and all n large enough. This shows that $\{Kz_n\}$ has no converging subsequence which contradicts the fact that $K \in K(X,X)$. Hence the sets $\{\lambda \in \sigma(K) \colon |\lambda| \geq 1/N\}$ are finite for all $N \in \mathbb{N}$.

3. Assume that $0 \notin \sigma(K)$ which implies that K is invertible by Theorem 5.10. Choose a sequence $\{x_i\} \subset X$ with $\|x_i\| \leq 1$ for all $i \in \mathbb{N}$. Then $Kx_i \to y$ up to a subsequence and therefore $x_i \to K^{-1}y$. This shows that the set $\{x \in X \colon \|x\| \leq 1\}$ is sequentially compact and we conclude from Theorem 3.4 that $\dim X < \infty$. □

Theorem 12.4 (Normal form for compact operators)**.** *Let X be a Banach space over $\mathbb{C}$ and let $K \in K(X,X)$. Then for $\lambda \in \sigma(K)\backslash\{0\}$ there exists a* ***unique*** *decomposition $X = N(\lambda) \oplus F(\lambda)$ so that*

1. *$N(\lambda), F(\lambda)$ are closed subspaces of X which are invariant under K with $\dim N(\lambda) = \operatorname{codim} F(\lambda) < \infty$.*

2. *$\lambda \operatorname{id} - K \colon N(\lambda) \to N(\lambda)$ is nilpotent, i.e. $(\lambda \operatorname{id} - K)^n|_{N(\lambda)} = 0$ for some $n \in \mathbb{N}$, and $\lambda \operatorname{id} - K \colon F(\lambda) \to F(\lambda)$ is an isomorphism.*

   Moreover, we have that

3. *$\ker(K - \lambda \operatorname{id}) \subset N(\lambda)$ and*

4. *$N(\lambda) \subset F(\mu)$ for $\lambda \neq \mu \in \sigma(K)\backslash\{0\}$.*

Proof. Step 1: Construction.

For $\lambda \in \sigma(K)\backslash\{0\}$ and all $j \in \mathbb{N}$ we define

$$N_j(\lambda) := \ker(\lambda \operatorname{id} - K)^j \quad \text{and} \quad F_j(\lambda) := \operatorname{Image}(\lambda \operatorname{id} - K)^j.$$

These spaces are invariant under K since

$$\begin{aligned} K(\lambda \operatorname{id} - K)^j &= -(\lambda \operatorname{id} - K)^{j+1} + \lambda(\lambda \operatorname{id} - K)^j \\ &= -(\lambda \operatorname{id} - K)^j(\lambda \operatorname{id} - K) + (\lambda \operatorname{id} - K)^j \lambda \operatorname{id} \\ &= (K - \lambda \operatorname{id})^j K. \end{aligned}$$

Moreover, we have by Theorem 11.12 that

$$(\lambda \operatorname{id} - K)^j = \lambda^j (\operatorname{id} + \text{ compact}) \in \operatorname{Fred}(X, X)$$

and $\operatorname{ind}(\lambda \operatorname{id} - K)^j = 0$. Hence, $F_j(\lambda)$ is closed for all $j \in \mathbb{N}$ and $\dim N_j(\lambda) = \operatorname{codim} F_j(\lambda)$. From the definition and Theorem 12.3 it follows that

$$\{0\} \subsetneq N_1(\lambda) \subset N_2(\lambda), \dots .$$

and if we assume that $N_j(\lambda) \subsetneq N_{j+1}(\lambda)$ for all $j \in \mathbb{N}$ it follows from Lemma 3.3 that there exists a sequence $z_j \in N_j(\lambda)$ with $\|z_j\| = 1$ and $\operatorname{dist}(z_j, N_{j-1}(\lambda)) \geq 1/2$. Thus

$$\frac{1}{\lambda}((\lambda \operatorname{id} - K)z_j + Kz_i) \in N_{j-1}(\lambda)$$

for all $i \leq j - 1$ and hence

$$\left\| \frac{1}{\lambda}((\lambda \operatorname{id} - K)z_j + Kz_i) - z_j \right\| \geq \frac{1}{2}$$

for all $1 \leq i \leq j - 1$. But this implies that $\|Kz_j - Kz_i\| \geq \frac{|\lambda|}{2} > 0$ for all $i < j$ and is therefore a contradiction to $K \in K(X, X)$. Thus there exists a smallest $n \in \mathbb{N}$ so that $N_n(\lambda) = N_{n+1}(\lambda)$. Now we assume that $N_{n+j}(\lambda) \subsetneq N_{n+j+1}(\lambda)$ for some $j \in \mathbb{N}$. Then there exists $x \in X$ with

$$(\lambda \operatorname{id} - K)^{n+1}(\lambda \operatorname{id} - K)^j x = 0 \quad \text{and} \quad (\lambda \operatorname{id} - K)^n(\lambda \operatorname{id} - K)^j x \neq 0.$$

But this means $(\lambda \operatorname{id} - K)^j x \in N_{n+1}(\lambda) \backslash N_n(\lambda) = \emptyset$ and we get another contradiction. Since $\operatorname{codim}(F_j(\lambda)) = \dim N_j(\lambda)$ we have

$$X \supsetneq F_1(\lambda) \supsetneq F_2(\lambda) \supsetneq \dots \supsetneq F_n(\lambda) = F_{n+1}(\lambda) = \dots$$

Therefore the spaces $N(\lambda) := N_n(\lambda)$ and $F(\lambda) := F_n(\lambda)$ are well-defined. Now assume that there exists $x \in N(\lambda) \cap F(\lambda)$ and $x \neq 0$. Then there exists $y \in X$ so that $x = (K - \lambda \operatorname{id})^n y$ and $0 = (K - \lambda \operatorname{id})^n x = (K - \lambda \operatorname{id})^{2n} y$. Since $N_{2n}(\lambda) = N_n(\lambda) = N(\lambda)$ it follows that $0 = (K - \lambda \operatorname{id})^n y = x$ and we get a contradiction. Finally, we note that the projection

$$N(\lambda) \to X/F(\lambda)$$

is injective and hence surjective because of dimensional reasons. This shows that

$$X = N(\lambda) \oplus F(\lambda)$$

and we have proved statement 1.

By definition we have that

$$(\lambda\,\mathrm{id} - K)^n|_{N(\lambda)} = 0.$$

Moreover, $\lambda\,\mathrm{id} - K\colon F(\lambda) \to F(\lambda)$ is injective since $N_1(\lambda) \cap F(\lambda) = \{0\}$ because of $N_1(\lambda) \subset N(\lambda)$. Next, we let $y \in F(\lambda) = F_n(\lambda) = F_{n+1}(\lambda)$ and hence there exists $x \in X$ so that

$$y = (\lambda\,\mathrm{id} - K)^{n+1}x = (\lambda\,\mathrm{id} - K)(\lambda\,\mathrm{id} - K)^n x \in (\lambda\,\mathrm{id} - K)(F(\lambda))$$

and therefore $(\lambda\,\mathrm{id} - K\colon F(\lambda) \to F(\lambda)$ is also surjective, which shows statement 2.

Step 2: Uniqueness.

Let $X = N \oplus F$ be another decomposition so that N and F satisfy statements 1. and 2. By statement 2. we get that $(\lambda\,\mathrm{id} - K)|_N$ is nilpotent which implies by the definition of $N(\lambda)$ that $N \subset N(\lambda)$. Next we let $z \in F$ and we note that $(\lambda\,\mathrm{id} - K)^n z \in F(\lambda)$ and therefore it follows from statement 2. that

$$F = (\lambda\,\mathrm{id} - K)^n F \subset F(\lambda).$$

Since $X = N \oplus F$ it follows that $N = N(\lambda)$ and $F = F(\lambda)$.

Step 3: Proof of statement 4.

Let $\lambda, \mu \in \sigma(K)\backslash\{0\}$ with $\lambda \neq \mu$ and let $x \in N(\lambda)$ with $x = y + z$, where $y \in N(\mu)$ and $z \in F(\mu)$. Here we assume that $N(\mu) = N_m(\mu)$ and $F(\mu) = F_m(\mu)$ for some $m \in \mathbb{N}$. Then

$$(\mu\,\mathrm{id} - K)^m x = (\mu\,\mathrm{id} - K)^m z \in F(\mu)$$

and therefore

$$(\mu\,\mathrm{id} - K)^m N(\lambda) \subset F(\mu).$$

Next we claim that $(\mu\,\mathrm{id} - K)^m N(\lambda) = N(\lambda)$ which will finish the proof. Assume that there exists $x \in N(\lambda)\backslash\{0\}$ with $(\mu\,\mathrm{id} - K)x = 0$. Then

$$(\lambda\,\mathrm{id} - K)x = (\mu\,\mathrm{id} - K)x + (\lambda - \mu)x = (\lambda - \mu)x$$

and hence $\lambda - \mu \neq 0$ is an eigenvalue of the map $(\lambda\,\mathrm{id} - K)|_{N(\lambda)}$. But this contradicts the fact that $(\lambda\,\mathrm{id} - K)|_{N(\lambda)}$ is nilpotent and therefore $(\mu\,\mathrm{id} - K)|_{N(\lambda)}$ is injective. As $\dim N(\lambda) < \infty$ it follows that $(\mu\,\mathrm{id} - K)|_{N(\lambda)}$ is also surjective and this shows the claim and therefore also statement 4. □

Example: A compact operator with no non-zero eigenvalues.

Let $X = C^0([0,1],\mathbb{C})$ and note that $K \in L(X,X)$, $(Kf)(x) := \int_0^x f(y)dy$ is compact (see the exercises). Assume that there exists $\lambda \neq 0$ and $f \in X\backslash\{0\}$ with

$$\lambda f(x) = \int_0^x f(y)dy.$$

It follows from the Fundamental Theorem of Calculus that $f \in C^1((0,1))$ and $\lambda f'(x) = f(x)$. Hence $f(x) = c\exp\left(\frac{x}{\lambda}\right)$ for some $c \neq 0$ but this contradicts that fact that $f(0) = 0$. Therefore we conclude that $\sigma(K) = \{0\}$.

Definition 12.5. *Let X be a Hilbert space over $\mathbb{C}$. A map $A \in L(X,X)$ is called*

- ***normal*** *if $[A, A^*] := AA^* - A^*A = 0$.*
- ***hermitian*** *or* ***self-adjoint*** *if $A^* = A$.*
- ***unitary*** *if $A^*A = AA^* = \mathrm{id}$.*

Lemma 12.6. *Let X be a Hilbert space over $\mathbb{K}$ and let $A \in L(X,X)$ be normal. Then*

1. *$\ker A = \ker A^*$ and $\|Ax\| = \|A^*x\|$ for all $x \in X$.*
2. *If $\mathbb{K} = \mathbb{C}$ then $\max_{\lambda\in\sigma(A)} |\lambda| = \|A\|$.*

Proof. 1. For all $x \in X$ we have

$$\|Ax\|^2 = \langle x, A^*Ax\rangle = \overline{\langle A^*Ax, x\rangle} = \overline{\langle AA^*x, x\rangle} = \|A^*x\|^2$$

which finishes the proof of statement 1. 2. By Theorem 12.2 we know that

$$\max_{\lambda\in\sigma(A)} |\lambda| = \lim_{n\to\infty} \|A^n\|^{1/n} \leq \|A\|.$$

Now we claim that for a normal operator A we have $\|A^n\| = \|A\|^n$ for all $n \in \mathbb{N}$. We show this by induction and we note that the case $n = 1$ is obvious $n = 1$. For $n \hookrightarrow n+1$ we have

$$\begin{aligned}
\|A^n x\|^2 =& \langle A^n x, A^n x\rangle \\
=& \langle A^{n-1}x, A^*A^n x\rangle \\
\leq& \|A^*A^n x\|\|A^{n-1}x\| \\
=& \|A^{n+1}x\|\|A^{n-1}x\| \\
\leq& \|A^{n+1}x\|\|A^{n-1}\|\|x\| \\
\leq& \|A^{n+1}x\|\|A\|^{n-1}\|x\|
\end{aligned}$$

where we used statement 1. in the fourth line and the induction hypotheses in the last line. By taking the supremum over $\|x\| \leq 1$ we obtain

$$\|A\|^{2n} = \|A^n\|^2 \leq \|A^{n+1}\| \|A\|^{n-1}$$

and hence

$$\|A\|^{n+1} \leq \|A^{n+1}\|.$$

□

Theorem 12.7 (Spectral theorem for compact and normal operators)**.** *Let X be a Hilbert space over $\mathbb{C}$ with $\dim X = \infty$ and let $K \in K(X, X)$ be normal. Then*

1. *$\sigma(K)$ is compact and 0 is the only possible accumulation point, $0 \in \sigma(K)$ and $\max\{|\lambda|, \lambda \in \sigma(K)\} = \|K\|$.*

2. *For $\lambda \in \sigma(K)\backslash\{0\}$ the eigenspace $E_\lambda(K) := \ker(\lambda \operatorname{id} - K)$ is non-trivial, finite-dimensional and $E_\lambda(K) \perp E_\mu(K)$ for $\lambda \neq \mu$.*

3. *X is the Hilbert sum of the $E_\lambda(K)$ with $\lambda \in \sigma(K)$ and $Kx = \sum_{\lambda \in \sigma(K)\backslash\{0\}} \lambda P_\lambda(x)$ where $P_\lambda : X \to E_\lambda(K)$ is the orthogonal projection onto $E_\lambda(K)$.*

Proof. Statement 1. follows directly from Theorem 12.2, Theorem 12.3 and Lemma 12.6.
The first part of Statement 2. follows from Theorems 12.3 and 12.4. since $\lambda \operatorname{id} - K$ is a normal operator as $(\lambda \operatorname{id} - K)^* = \overline{\lambda} \operatorname{id} - K^*$. By Lemma 12.6 we have that

$$E_\lambda(K) = \ker(\lambda \operatorname{id} - K) = \ker(\overline{\lambda} \operatorname{id} - K^*) = E_{\overline{\lambda}}(K^*)$$

and for $x \in E_\lambda(K)$, $y \in E_\mu(K)$ we get

$$\begin{aligned}(\lambda - \mu)\langle x, y\rangle =& \langle \lambda x, y\rangle - \langle x, \overline{\mu} y\rangle \\ =& \langle Kx, y\rangle - \langle x, K^* y\rangle = 0\end{aligned}$$

and therefore $\langle x, y\rangle = 0$, which shows that $E_\lambda(K) \perp E_\mu(K)$.

3. We use Theorem 9.12 and hence we have to show the maximality of the sets $\{E_\lambda(K)\}_{\lambda \in \sigma(K)}$. We let

$$V = \left(\bigoplus_{\lambda \in \sigma(K)} E_\lambda(K) \right)^\perp$$

and our goal is to show that $V = \{0\}$. First we observe that V is invariant under

K since for $v \in V$ and $x \in E_\lambda(K)$, $\lambda \in \sigma(K)$, we have

$$\langle Kv, x\rangle = \langle v, K^*x\rangle = \langle v, \overline{\lambda}x\rangle = \lambda\langle v, x\rangle = 0$$

and hence $Kv \in V$. Moreover, V is a complex Hilbert space, $K|_V \in L(V,V)$ and $(K|_V)^* = K^*|_V$, which shows that $K|_V$ is normal. The map $K|_V$ is also compact and by the definition of V it follows that $K|_V$ has no eigenvalue. Hence we conclude from statement 2. that $\sigma(K|_V) = \{0\}$. Statement 1. then yields that $\|K|_V\| = 0$, i.e. $K|_V \equiv 0$, and thus $V \subset \ker(K) = E_0(K)$. By the definition of V we also have that $V \subset E_0(K)^\perp$ and therefore we conclude $V = \{0\}$. The rest of the proof now follows from Theorem 9.12. □

Theorem 12.8 (Spectral theorem for compact and selfadjoint operators)**.** *Let X be a Hilbert space over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and let $K \in K(X,X)$ with $K^* = K$. Then $\sigma(K) \subset \mathbb{R}$ and all three statements from Theorem 12.7 remain true.*

Proof. Case 1: $\mathbb{K} = \mathbb{C}$

Let $\lambda \in \sigma(K)\backslash\{0\}$. Then we know from Theorem 12.7 that there exists an eigenvector $x \neq 0$, so that

$$\begin{aligned}\lambda\|x\|^2 &= \langle\lambda x, x\rangle = \langle Kx, x\rangle \\ &= \langle x, Kx\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\|x\|^2,\end{aligned}$$

which implies that $\lambda = \overline{\lambda}$ and therefore $\lambda \in \mathbb{R}$.

Step 2: $\mathbb{K} = \mathbb{R}$

We complexify the real vector space X and we get a complex vector space $X^{\mathbb{C}} := X \times X$ with the complex multiplication $i(x,y) = (-y,x)$ for all $x,y \in X$. Hence we get for all $a,b \in \mathbb{R}$ and all $x,y \in X$

$$(a+ib)(x,y) = (ax - by, ay + bx)$$

and one concludes that $X^{\mathbb{C}}$ is a complex vector space with $X \subset X^{\mathbb{C}}$ via the map $x \mapsto (x,0)$. We also have that $\overline{(x,y)} = (x,-y)$. Next we define

$$\langle (x_1,y_1),(x_2,y_2)\rangle := \langle x_1,x_2\rangle + \langle y_1,y_2\rangle - i(-\langle x_1,y_2\rangle - \langle y_1,x_2\rangle)$$

and we note that

$$\langle (x_1,y_1),(x_2,y_2)\rangle = \overline{\langle (x_2,y_2),(x_1,y_1)\rangle}$$

and

$$\|(x,y)\|^2 = \|x\|^2 + \|y\|^2$$

Moreover, we have

$$\begin{aligned}\langle i(x_1,y_1),(x_2,y_2)\rangle =&\langle(-y_1,x_1),(x_2,y_2)\rangle\\ =&-\langle y_1,x_2\rangle+\langle x_1,y_2\rangle+i(\langle y_1,y_2\rangle+\langle x_1,x_2\rangle)\\ =&i\langle(x_1,y_1),(x_2,y_2)\rangle\end{aligned}$$

and all of these properties combined show that $(X^{\mathbb{C}}=X\times X,\langle\cdot,\cdot\rangle)$ is a complex Hilbert space.

Now we define the map $\tilde{K}\colon X^{\mathbb{C}}\to X^{\mathbb{C}}$ by $\tilde{K}(x,y)=(Kx,Ky)$ and we note that $\tilde{K}$ is complex linear since

$$\tilde{K}(i(x,y))=\tilde{K}(-y,x)=(-Ky,Kx)=i(Kx,Ky)=i\tilde{K}(x,y).$$

Moreover, we have that $(\tilde{K})^*=\widetilde{(K^*)}$ and hence we can apply the results from step 1 to the triple $(X^{\mathbb{C}},\langle\cdot,\cdot\rangle,\tilde{K})$. We get that $\sigma(\tilde{K})\subset\mathbb{R}$, $\dim E_\lambda(\tilde{K})<\infty$ if $\lambda\neq 0$ and

$$X^{\mathbb{C}}=\overline{\bigoplus_{\lambda\in\sigma(\tilde{K})}E_\lambda(\tilde{K})}.$$

Since $\tilde{K}(x,y)=(Kx,Ky)$ we conclude that $\sigma(\tilde{K})=\sigma(K)$ and

$$E_\lambda(\tilde{K})=E_\lambda(K)\times E_\lambda(K)\subset X\times X=X^{\mathbb{C}}.$$

It remains to show that

$$X=\overline{\bigoplus_{\lambda\in\sigma(K)}E_\lambda(K)}.$$

For this we let $x\in X$ with $x\perp E_\lambda(K)$ for all $\lambda\in\sigma(K)$ and we note that this implies $(x,0)\perp E_\lambda(\tilde{K})$ for all $\lambda\in\sigma(\tilde{K})$. Since the set of $E_\lambda(\tilde{K})$'s is maximal, we conclude that $(x,0)=0\in X^{\mathbb{C}}$ and thus $x=0$. This shows that the set of $E_\lambda(K)$, $\lambda\in\sigma(K)$, are also maximal and this finishes the proof by Theorem 9.12. □

As an application of this result we obtain the

Theorem 12.9 (Spectral theorem for the Dirichlet problem)**.** *Let $\Omega\in\mathbb{R}^n$ be open and bounded and let $L:W_0^{1,2}(\Omega)\to W_0^{1,2}(\Omega)$ be defined by*

$$Lv:=-div(aDv)+qv,$$

where $a\in L^\infty(\Omega,M_n(\mathbb{R}))$ is symmetric and elliptic with constant $\mu>0$ and $q\in L^\infty(\Omega)$.

Then there exists a sequence $\{\lambda_j\}\subset\mathbb{R}$ with $\lambda_1<\lambda_2<\cdots\to\infty$, so that

1. *the eigenspace* $E_\lambda(L) := \{v \in W_0^{1,2}(\Omega)\colon Lv = \lambda v\}$ *is finite-dimensional and it is non-trivial if and only if* $\lambda \in \{\lambda_j\}$.

2. *The space* $L^2(\Omega)$ *is the Hilbert sum of the sets* $E_{\lambda_j}(L)$, $j \in \mathbb{N}$.

Proof. Without loss of generality we assume that $q \geq 0$ since otherwise look at the operator $\tilde{L} := L + \|q\|_{L^\infty}$ id and this does not change the eigenspaces. We consider the bilinear form $B : W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega) \to \mathbb{R}$ associated to L

$$B(u, v) = \int_\Omega \langle Du, aDv\rangle + quv = (Lv)(u)$$

and we note that B is symmetric since a is symmetric. Moreover, it follows from the ellipticity of a and the Poincare inequality (Theorem 10.13) that

$$B(v, v) \geq \mu \int |Dv|^2 \geq c\|v\|_{W^{1,2}}.$$

Hence we can apply the Lax-Milgram Theorem (Theorem 9.11) in order to conclude that L is invertible with $\|L^{-1}\| \leq c$. Now the map

$$I\colon W_0^{1,2}(\Omega) \to L^2(\Omega),\ v \mapsto v$$

is compact by Theorem 11.16 and the adjoint map is given by

$$I'\colon L^2(\Omega) \to (W_0^{1,2}(\Omega))',\ (I'f)(u) = \int_\Omega fu.$$

We define $G\colon L^2(\Omega) \to L^2(\Omega)$ by $G := I \circ L^{-1} \circ I'$ and we note that for $f \in L^2(\Omega)$ the map $Gf \in W_0^{1,2}(\Omega)$ is the unique weak solution $v \in W_0^{1,2}(\Omega)$ of $Lv = f$. For $f_1, f_2 \in L^2(\Omega)$ we let $Gf_i = v_i$, $i = 1, 2$, i.e. $Lv_i = f_i$ and we calculate

$$\begin{aligned}\langle Gf_1, f_2\rangle_{L^2} =& \langle v_1, Lv_2\rangle \\ =& (Lv_2)(v_1) \\ =& B(v_1, v_2) \\ =& B(v_2, v_1) \\ =& \langle f_1, Gf_2\rangle\end{aligned}$$

which shows that G is self-adjoint. Moreover, it follows from Theorem 11.4 that G is compact since it is the composition of a compact map with continuous maps. Hence we can apply Theorem 12.8 to G and we conclude that $\sigma(G) = \{\mu_1, \mu_2, \ldots\} \subset \mathbb{R}$ with $\mu_j \to 0$ as $j \to \infty$. Note that we have infinitely many eigenvalues since $\dim(E_{\mu_j}(G)) < \infty$ but $\dim(L^2(\Omega)) = \infty$. Now for $\lambda, \mu \in \mathbb{R}$ we have

1. If there exists $v \in W_0^{1,2}(\Omega)\setminus\{0\}$ with $Lv = \lambda v$ then we have

$$\lambda\|v\|^2 = (Lv)(v) = B(v,v) \geq c\|v\|^2_{W^{1,2}} > 0.$$

   This shows that $\lambda > 0$ and hence $L(v/\lambda) = v$ which is equivalent to $Gv = \frac{v}{\lambda}$.

2. If $Gv = \mu v$ for some $v \in L^2(\Omega)\setminus\{0\}$ then $\mu v \in W_0^{1,2}(\Omega)$ and $L(\mu v) = v$. Hence $\mu \neq 0$, $v \in W_0^{1,2}(\Omega)$ and $Lv = \frac{v}{\mu}$.

Therefore we conclude that $\sigma(G) = \{\mu_1, \mu_2, \dots\} \subset \mathbb{R}^+$ and $\sigma(L) = \{\frac{1}{\mu_1}, \frac{1}{\mu_2}, \dots\}$ from which the claim follows. □

13 Semigroups

Motivation: For an open subset $\Omega \subset \mathbb{R}^n$ and a given initial data $u_0 \in W^{2,2} \cap W_0^{1,2}(\Omega)$ we want to construct a solution of the heat equation

$$\begin{aligned} \partial_t u =& \Delta u \text{ in } \Omega \times (0, \infty) \\ u =& 0 \text{ on } \partial\Omega \times (0, \infty) \\ u =& u_0 \text{ on } \Omega \times \{0\}, \\ u \colon & \Omega \times [0, \infty) \to \mathbb{R} \end{aligned}$$

The idea is to consider this PDE as an ordinary differential equation in a Banach space

$$u \colon [0, \infty) \to W^{2,2} \cap W_0^{1,2}(\Omega)$$

so that

$$\begin{cases} \frac{d}{dt} u = & \Delta u \\ u(0) = & u_0 \end{cases}$$

We have the **formal** solution $u(t) = e^{t\Delta} u_0$. In the following we consider the time evolution $u(0) \to u(t)$ as a time dependent family of operators and this is what leads us to C^0-semigroups.

Definition 13.1. *Let X be a Banach space. A map $T \colon [0, \infty) \to L(X, X)$ is called C^0**-semigroup** if and only if*

1. $T(t_1 + t_2) = T(t_1)T(t_2)$ *for all* $t_1, t_2 \geq 0$.
2. $T(0) = \mathrm{id}$.
3. *the map* $T_x \colon [0, \infty) \to X$, $T_x(t) = T(t)x$ *is continuous for all* $x \in X$.

Examples:

1. Let $A \in L(X, X)$ and consider for all $t \geq 0$ the map

$$T(t) = \exp(tA) := \sum_{k=0}^{\infty} \frac{(tA)^k}{k!}.$$

This series converges absolutely in the operator norm and the first two properties of a C^0-semigroup are obvious. Moreover, the map $T\colon \mathbb{R} \to L(X,X)$ is continuous.

2. Let $X = L^2(\mathbb{R})$ and let $(T(t)u)(x) = u(x+t)$. Then T is a C^0-semigroup since $\|T(t)u - u\|_{L^2} \to 0$ as $t \to 0$, a fact which follows from the proof of Theorem 10.4. Note that $T : \mathbb{R} \to L(X,X)$ is not continuous since $\|T(t_1) - T(t_2)\| = 2$ for $t_1 \neq t_2$.

3. Let $X = L^\infty(\mathbb{R})$ and $(T(t)u)(x) = u(x+t)$. Then properties 1 and 2 of the definition of a C^0-semigroup hold and $\|T(t)\| = 1$ but T is not a C^0-semigroup.

Lemma 13.2. *Let X be a Banach space and let T be a C^0-semigroup on X. Then*

1. *the limit*

$$\lim_{t\to\infty}\left(\frac{\log\|T(t)\|}{t}\right) = \omega_0 \in [-\infty,\infty).$$

exists and

2. *for all $\omega > \omega_0$ there exists $c = c(\omega) < \infty$ so that for all $t \geq 0$*

$$\|T(t)\| \leq ce^{\omega t}.$$

Proof. For all $t > 0$ we define

$$\omega(t) := \frac{\log\|T(t)\|}{t}$$

and we let $\tau > 0$. For every $t > 0$ there exists $n \in \mathbb{N}_0$ and $t_0 \in [0,\tau)$ so that $t = n\tau + t_0$. It follows from the definition of a C^0-semigroup that

$$\begin{aligned}\omega(t) =&\omega(nt+t_0) = \log\left(\frac{\|T(n\tau+t_0)\|}{t}\right)\\ =&\log\left(\frac{\|T(\tau)^nT(t_0)\|}{t}\right)\\ \leq&\frac{n\log\|T(\tau)\| + \log\|T(t_0)\|}{t}\end{aligned}$$

and therefore, by letting $t \nearrow \infty$, we get

$$\limsup_{t\nearrow\infty}\omega(t) \leq \frac{\log\|T(\tau)\|^t}{\tau} = \omega(\tau) < \infty.$$

Next we choose a sequence $\tau_k \nearrow \infty$ so that $\omega(\tau_k) \to \liminf_{t\to\infty}\omega(t)$ and we conclude that the limit $\lim\limits_{t\to\infty}\omega(t) \in [-\infty,\infty)$ exists.

In order to show statement 2. we note that there exists $t_1 > 0$ so that $\omega(t) \le \omega$ for all $t \ge t_1$ by statement 1. Hence it follows that

$$\|T(t)\| \le e^{\omega t}$$

for all $t \ge t_1$. Next, we recall that the map $[0, t_1] \to X$, $t \mapsto T(t)x$ is continuous for all $x \in X$ and therefore

$$\sup_{0 \le t \le t_1} \|T(t)x\| < \infty$$

for all $x \in X$. By Theorem 5.7 there exists a constant $c < \infty$ so that

$$\sup_{0 \le t \le t_1} \|T(t)\| \le c < \infty$$

and thus

$$\|T(t)\| \le c \max\{1, e^{-\omega t_1}\} e^{\omega t}$$

for all $t \ge 0$. □

Examples:

1. Let $X = L^2(\mathbb{R})$ and $(T(t)u)(x) = u(x+t)$. Then we know that $\|T(t)\| = 1$ for all $t \ge 0$ and thus $\omega_0 = 0$.

2. Let $I = (0,1)$ and $X = L^2(I)$. Then the map

$$(T(t)u)(x) = \begin{cases} u(x+t), & \text{if} \quad x+t < 1 \\ 0, & \text{otherwise} \end{cases}$$

   is a C^0-semigroup with $T(t) = 0$ for all $t \ge 1$ and therefore $\omega_0 = -\infty$.

3. Let X be a Hilbert space over $\mathbb{C}$ and let $K \in K(X, X)$ be normal with eigenvalues λ_j, $j \in \mathbb{N} \cup \{0\}$. Then $T(t)x := e^{tK}x = \sum_{j=1}^{\infty} e^{t\lambda_j} P_j x$ (see Theorem 12.7) is a C^0-semigroup with

$$\begin{aligned} \|T(t)x\| &= \left\| \sum_{j=1}^{\infty} e^{t\lambda_j} P_j x \right\| \\ &= \left(\sum_{j=1}^{\infty} e^{2tRe(\lambda_j)} \|P_j x\|^2 \right)^{1/2} \\ &\le e^{t \max_j (Re\lambda_j)} \|x\| \end{aligned}$$

   and therefore $\omega_0 = \max_j(Re(\lambda_j))$.

Definition 13.3. *Let X be a Banach space and let T be a C^0-semigroup on X. We* *define*

$$Ax := \lim_{h \searrow 0} \frac{T(h)x - x}{h}$$

if the limit exists. Moreover, we define the ***domain*** *of A by*

$$D(A) = \{x \in X : \ Ax \text{ exists}\}$$

and we note that it is a linear space and $A\colon D(A) \to X$ is a linear map. The map A is called the ***infinitesimal generator*** *of the C^0-semigroup T.*

Example: Let $X = L^2(\mathbb{R})$ and let $(T(t)u)(x) = u(x+t)$. Then

$$\frac{(T(h) - \mathrm{id})u}{h} = \Delta_h u := \frac{u(x+h) - u(x)}{h}$$

is the so called difference quotient of u. Let $u \in D(A)$, i.e. $Au = \lim_{h\to 0} \Delta_h u$ exists, and let $\varphi \in C_c^\infty(\mathbb{R})$. Then we calculate

$$\int (\Delta_h u)\varphi = -\int u(\Delta_{-h}\varphi) \to -\int u\varphi'$$

as $h \searrow 0$. Hence we get that

$$\int (Au)\varphi = -\int u\varphi'$$

and therefore $u \in W^{1,2}(\mathbb{R})$ and $Au = \frac{d}{dx}u$.

Conversely, if we assume that $u \in W^{1,2}(\mathbb{R})$, we get

$$\left\| \Delta_h u - \frac{du}{dx} \right\|_{L^2} \leq \sup_{0 \leq t \leq h} \|u'(\cdot + t) - u'(\cdot)\|_{L^2} \to 0$$

as $h \to 0$ by the proof of Theorem 10.4. Hence $u \in D(A)$ and $A(u) = \frac{d}{dx}u$.

We therefore conclude that

$$A = \frac{d}{dx} \colon W^{1,2}(\mathbb{R}) = D(A) \to L^2(\mathbb{R}).$$

Lemma 13.4. *Let X be a Banach space and let T be a C^0-semigroup on X. Let A be an ininitesimal generator of $T(t)$. Then we have for all $t > 0$*

1. *$T(t)x = \lim_{h\to 0} \frac{1}{h}\int_t^{t+h} T(s)x\,ds$.*

2. *$\int_0^t T(s)x\,ds \in D(A)$ and $A\int_0^t T(s)x\,ds = T(t)x - x$ for all $x \in X$.*

3. *$T(t)x \in D(A)$ and $AT(t)x = T(t)Ax$ if $x \in D(A)$.*

4. *$\frac{d}{dt}T(t)x = T(t)Ax$ for all $x \in D(A)$.*

Proof. 1. This follows from the Fundamental Theorem of Calculus.

2. We note that for all $h > 0$ we have

$$\begin{aligned}\frac{T(h)-\mathrm{id}}{h}\int_0^t T(s)x\,ds =& \frac{1}{h}\int_0^t (T(s+h)-T(s))x\,ds\\ =&\frac{1}{h}\left(\int_t^{t+h} T(s)x\,ds - \int_0^h T(s)x\,ds\right) \to T(t)x - x\end{aligned}$$

as $h \searrow 0$ since the map $T(\cdot)x$ is continuous by assumption.

3. Let $x \in D(A)$ and $h > 0$, then

$$\frac{T(h)-\mathrm{id}}{h}T(t)x = T(t)\frac{T(h)-\mathrm{id}}{h}x \to T(t)Ax$$

as $h \searrow 0$ and thus $T(t)x \in D(A)$ and $AT(t)x = T(t)Ax$.

4. Let $h > 0$ and $x \in D(A)$. Then

$$\frac{T(t+h)x - T(t)x}{h} = T(t)\frac{T(h)-\mathrm{id}}{h}x \to T(t)Ax$$

as $h \searrow 0$ and

$$\begin{aligned}\frac{T(t)x - T(t-h)x}{h} =& T(t-h)\frac{T(h)-\mathrm{id}}{h}x\\ =& T(t-h)Ax + T(t-h)\left(\frac{T(h)-\mathrm{id}}{h} - A\right)x\\ \to& T(t)Ax\end{aligned}$$

as $h \searrow 0$. Hence the claim follows.

□

Theorem 13.5. *Let X be a Banach space and let T be a C^0-semigroup on X with infinitesimal generator A. Then $D(A)$ is dense in X and A is closed, i.e. the graph of A, $G(A) \subset X \times X$, is closed.*

Proof. By Lemma 13.4, statement 1, we get for every $x \in X$

$$x = T(0)x = \lim_{h\to 0}\frac{1}{h}\int_0^h T(s)x\,ds$$

and the right hand side is in $D(A)$ by Lemma 13.4, statement 2. Hence $D(A)$ is dense in X.

Next we let x_n be a sequence in $D(A)$ with $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$. In order to show that $G(A)$ is closed we have to verify that $x \in D(A)$ and $y = Ax$. By Lemma 13.4, statement 4, we have for every $h > 0$

$$T(h)x_n - x_n = \int_0^h \frac{d}{dt} T(t)x_n dt = \int_0^h T(t)Ax_n\, dt$$

and by Lemma 13.2 we conclude that

$$\left\| \int_0^h (T(t)Ax_n - T(t)y)\, dt \right\| \leq \sup_{0\leq t\leq h} \|T(t)\| \|Ax_n - y\| h \leq Ch\|Ax_n - y\| \to 0$$

as $n \to \infty$. This implies that $T(h)x - x = \int_0^h T(t)y\, dt$ and thus

$$\frac{T(h) - \mathrm{id}}{h} x = \frac{1}{h} \int_0^h T(t)y dt \to y$$

as $h \searrow 0$. Therefore we conclude that $x \in D(A)$ and $Ax = y$. □

Definition 13.6. *Let X be a Banach space and let $A : D(A) \to X$ be linear. Then we define the **resolvent set** of A by*

$$\varrho(A) = \{\lambda \in \mathbb{K} \colon \lambda \,\mathrm{id} - A : D(A) \to X \quad \text{is bijective and} \quad (\lambda \,\mathrm{id} - A)^{-1} \quad \text{is bounded}\}.$$

*Moreover, the **resolvent function** $R_A : \varrho(A) \to L(D(A), X)$ is defined by*

$$R_A(\lambda) = (\lambda \,\mathrm{id} - A)^{-1}.$$

Note that here we do not assume that A is bounded.

Lemma 13.7. *Let X be a Banach space and let T be a C^0-semigroup on X infinitesimal generator A and let $\omega_0 = \lim_{t\to\infty} \left(\frac{\log \|T(t)\|}{t}\right)$. Then we have for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\,(\lambda) > \omega_0$ that*

1. *$\lambda \in \varrho(A)$ and $R_A(\lambda)x = \int_0^\infty e^{-\lambda t} T(t)x\, dt$ for all $x \in D(A)$.*
2. *if $x \in D(A)$ then $R_A(\lambda)x \in D(A)$ and $AR_A(\lambda)x = R_A(\lambda)Ax$.*

Proof. Let $\omega_0 < \omega < \mathrm{Re}\,(\lambda)$ and note that by Lemma 13.2 we get the estimate $\|T(t)\| \leq ce^{\omega t}$ for all $t > 0$. Define

$$I(\lambda)x = \int_0^\infty e^{-\lambda t} T(t)x\, dt$$

and note that

$$\|I(\lambda)\| \leq c\int_0^\infty e^{\omega-\mathrm{Re}\,(\lambda)t}\,dt \leq \frac{c}{Re(\lambda)-\omega}. \tag{13.1}$$

Now let $x \in D(A)$ and calculate with the help of Lemma 13.4

$$\begin{aligned} I(\lambda)Ax =& T\int_0^\infty e^{-\lambda t}T(t)Ax\,dt \\ =& \int_0^\infty e^{-\lambda t}\frac{d}{dt}T(t)x\,dt \\ =& [e^{-\lambda t}T(t)x]_{t=0}^{t=\infty} + \lambda\int_0^\infty e^{-\lambda t}T(t)x\,dt \\ =& -x + \lambda I(\lambda)x \end{aligned}$$

and therefore

$$I(\lambda)(\lambda\,\mathrm{id}-A)x = x \quad \forall x \in D(A). \tag{13.2}$$

This implies that $\lambda\,$id-A is injective and $I(\lambda)$ is a left inverse of $\lambda\,\mathrm{id}-A$. Next, for $x \in X$ and $h > 0$ we calculate

$$\begin{aligned} \frac{T(h)-\mathrm{id}}{h}I(\lambda)x =& \frac{1}{h}\left(\int_0^\infty e^{-\lambda t}T(t+h)x\,dt - \int_0^\infty e^{-\lambda t}T(t)x\,dt\right) \\ =& \frac{e^{\lambda h}}{h}\int_h^\infty e^{-\lambda s}T(s)x\,ds - \frac{1}{h}\int_0^\infty e^{-\lambda s}T(s)x\,ds \\ =& \frac{e^{\lambda h}-1}{h}I(\lambda)x - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda s}T(s)x\,ds \\ \to& \lambda I(\lambda)x - x \end{aligned}$$

as $h \searrow 0$ since the function $s \mapsto e^{-\lambda s}T(s)x$ is continuous for all $x \in X$. Hence we get that

$$(\lambda\,\mathrm{id}-A)I(\lambda)x = x. \tag{13.3}$$

By (13.1)-(13.3) it follows that for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\,(\lambda) > \omega$ we have $\lambda \in \rho(A)$ and $R_A(\lambda) = I(\lambda)$, which proves statement 1. Moreover, we have that $R_A(\lambda)x \in D(A)$ for all $x \in D(A)$.

If $x \in D(A)$ it follows from (13.2) and (13.3) that for all $x \in D(A)$

$$(\lambda\,\mathrm{id}-A)R_A(\lambda)x = R_A(\lambda)(\lambda\,\mathrm{id}-A)x$$

and therefore

$$AR_A(\lambda)x = R_A(\lambda)Ax$$

□

Lemma 13.8. *Let X be a Banach space and let $A\colon D(A) \to X$ be a linear operator. For $\lambda, \mu \in \varrho(A)$ we have*

1. $R_A(\lambda) - R_A(\mu) = (\lambda - \mu) R_A(\lambda) R_A(\mu)$.

2. $R_A(\lambda) R_A(\mu) = R_A(\mu) R_A(\lambda)$.

Proof. First we note that statement 2. follows from statement 1. and hence this is the only thing to prove. For $\mu, \lambda \in \varrho(A)$ and $y \in X$ there exists $x \in D(A)$ with

$$y = (\mu \operatorname{id} - A)x = (\lambda \operatorname{id} - A)x + (\mu - \lambda)x$$

and hence

$$(\mu - \lambda) R_A(\lambda) R_A(\mu) y = (\mu - \lambda) R_A(\lambda) x.$$

Using these facts we get

$$\begin{aligned}(R_A(\lambda) - R_A(\mu))y =& x + (\mu - \lambda) R_A(\lambda) x - x \\ =& (\mu - \lambda) R_A(\lambda) R_A(\lambda) R_A(\mu) y.\end{aligned}$$

□

Theorem 13.9 (Hille-Yosida)**.** *Let X be a Banach space and let $A : D(A) \to X$ be a closed and linear operator so that $D(A)$ is dense in X and let $\omega \in \mathbb{R}$, $M > 0$. Then the following two statements are equivalent:*

1. *A is the infinitesimal generator of a C^0-semigroup T on X with $\|T(t)\| \leq M e^{\omega t}$ for all $t > 0$.*

2. *$(\omega, \infty) \subset \varrho(A)$ and $\|R_A(\lambda)^n\| \leq \frac{M}{(\lambda - \omega)^n}$ for all $n \in \mathbb{N}$ and all $\lambda > \omega$.*

Proof. <u>$1 \Rightarrow 2$:</u>

For $n \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ with $\operatorname{Re}(\lambda) > \omega$ define

$$I_n(\lambda)x := \frac{1}{(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} T(t) x \, dt.$$

As in the proof of Lemma 13.7 one shows that the integral exists and we have the estimate

$$\|I_n(\lambda)\| \leq \frac{M}{(n-1)!} \int_0^\infty t^{n-1} e^{(\omega - \operatorname{Re}(\lambda))t} \, dt \leq \frac{M}{(\operatorname{Re}(\lambda) - \omega)^n}.$$

Now we show by induction that $I_n(\lambda) = R_A(\lambda)^n$. The case $n = 1$ can be found in

Lemma 13.7. Next we let $n \geq 2$ and we show that the formula for $(n-1)$ implies the corresponding one for n. For this we let $x \in D(A)$ and using Lemma 13.4, statement 4, we calculate

$$\begin{aligned} I_n(\lambda)Ax =& \frac{1}{(n-1)!}\int_0^\infty t^{n-1}e^{-\lambda t}T(t)Ax\,dt \\ =& -\frac{1}{(n-2)!}\int_0^\infty t^{n-2}e^{-\lambda t}T(t)x\,dt + \frac{\lambda}{(n-1)!}\int_0^\infty t^{n-1}e^{-\lambda t}T(t)x\,dt \\ =& -I_{n-1}(\lambda)x + \lambda I_n(\lambda)x. \end{aligned}$$

Hence it follows that

$$I_n(\lambda)(\lambda\,\mathrm{id} - A)x = I_{n-1}(\lambda)x = R_A(\lambda)^{n-1}x = R_A(\lambda)^n(\lambda\,\mathrm{id} - A)x.$$

By Lemma 13.7 we know that $(\omega, \infty) \subset \varrho(A)$ and hence the operator $(\lambda\,\mathrm{id} - A) : D(A) \to X$ is surjective and therefore it follows that $I_n(\lambda) = R_A(\lambda)^n$.

<u>2 ⇒ 1:</u>

We let $\lambda > \omega$ and start with a formal computation:

$$R_A(\lambda) = \frac{1}{\lambda}\left(\mathrm{id} - \frac{1}{\lambda}A\right)^{-1} \approx \frac{1}{\lambda}\left(\mathrm{id} + \frac{A}{\lambda} + \cdots\right)$$

and hence one expects that

$$A \approx \lambda(\lambda R_A(\lambda) - \mathrm{id})$$

for λ large enough. Therefore we define $A(\lambda) := \lambda(\lambda R_A(\lambda) - \mathrm{id}) \in L(X, X)$ and we note that for $x \in D(A)$ and $\lambda > \omega$ we have

$$\lambda R_A(\lambda)x - x = R_A(\lambda)(\lambda x - \lambda x + Ax) = R_A(\lambda)Ax$$

and thus

$$\|\lambda R_A(\lambda)x - x\| \leq \|R_A(\lambda)\|\|Ax\| \leq \frac{M}{\mathrm{Re}\,(\lambda) - \omega}\|Ax\| \to 0\lambda \to \infty.$$

as $\lambda \to \infty$. For an arbitrary $x \in X$ we choose a sequence $x_n \in D(A)$ with $x_n \to x$ and we estimate

$$\begin{aligned} \|\lambda R_A(\lambda)x - x\| \leq& \|\lambda R_A(\lambda)(x - x_n)\| + \|\lambda R_A(\lambda)x_n - x_n\| + \|x_n - x\| \\ \leq& \left(\frac{M\lambda}{\lambda - \omega} + 1\right)\|x_n - x\| + \frac{M}{\lambda - \omega}\|Ax_n\| \end{aligned}$$

and therefore

$$\limsup_{\lambda\to\infty} \|\lambda R_A(\lambda)x - x\| \le (M+1)\|x_n - x\| \to 0$$

as $n \to \infty$.

We thus conclude that for $x \in D(A)$ we have

$$A(\lambda)x = \lambda R_A(\lambda)Ax \to Ax,$$

as $\lambda \to \infty$. As $R_A(\lambda)$ is a bounded operator, we can define for all $t \ge 0$

$$T_\lambda(t) = e^{tA(\lambda)} = e^{-\lambda t} e^{\lambda^2 R_A(\lambda) t}$$

and we note that T_λ is a C^0-semigroup with infinitesimal generator $A(\lambda)$. We estimate for all $t \ge 0$

$$\begin{aligned}\|T_\lambda(t)\| \le & e^{-\lambda t} \sum_{n=0}^{\infty} \frac{1}{n!} (\lambda^2 t)^n \|R_A(\lambda)^n\| \\ \le & M e^{-\lambda t} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\lambda^2 t}{\lambda - \omega} \right)^n \\ = & M e^{-\lambda t} e^{\frac{\lambda^2 t}{\lambda - \omega}}\end{aligned}$$

and thus

$$\|T_\lambda(t)\| \le M e^{\left(\frac{\omega}{1-\frac{\omega}{\lambda}}\right) t}. \tag{13.4}$$

By Lemma 13.8 we know that $R_A(\lambda)R_A(\mu) = R_A(\mu)R_A(\lambda)$ for all $\lambda, \mu > \omega$ and therefore we also get $A(\lambda)A(\mu) = A(\mu)A(\lambda)$ which in turn implies that

$$A(\mu)T_\lambda(t) = T_\lambda(t)A(\mu).$$

for all $\lambda, \mu > \omega$. Using this property we calculate with the help of Lemma 13.4

$$\begin{aligned}T_\lambda(t) - T_\mu(t) = & \int_0^t \frac{d}{ds} (T_\mu(t-s)T_\lambda(s))\, ds \\ = & \int_0^t T_\mu(t-s)T_\lambda(s)(A(\lambda) - A(\mu))\, ds\end{aligned}$$

In the following we let $\lambda, \mu \ge 2\omega$ since then $\frac{\omega}{1-\frac{\omega}{\lambda}}, \frac{\omega}{1-\frac{\omega}{\mu}} \le 2\omega$. Hence we use the estimate (13.4) in order to conclude for all $x \in D(A)$ that

$$\begin{aligned}\|T_\lambda(t)x - T_\mu(t)x\| \le & \int_0^t M^2 e^{2\omega(t-s)+2\omega s} \|A(\lambda)x - A(\mu)x\|\, ds \\ \le & M^2 t e^{2\omega t} \|A(\lambda)x - A(\mu)x\| \to 0\end{aligned}$$

as $\lambda, \mu \to \infty$. Note that the convergence is uniform for $t \in [0, t_0]$ for all $t_0 < \infty$. For $x \in X$ arbitrary we choose again a sequence $x_n \in D(A)$ with $x_n \to x$ and we get with the help of (13.4)

$$\begin{aligned}\|T_\lambda(t)x - T_\mu(t)x\| \leq &\|T_\lambda(t)(x - x_n)\| + \|T_\lambda(t)x_n - T_\mu(t)x_n\| + \|T_\mu(t)(x - x_n)\| \\ \leq &2Me^{\omega t}\|x - x_n\| + \|T_\lambda(t)x_n - T_\mu(t)x_n\| \to 0\end{aligned}$$

as $n, \lambda, \mu \to \infty$. This convergence is again uniform if $t \in [0, t_0]$.

Now we define $T(t)x := \lim_{\lambda\to\infty} T_\lambda(t)x$ for all $x \in X$ and we note that this limit exists continuous in t. We have that $T(0) = \mathrm{id}$ and

$$\begin{aligned}T(t_1 + t_2)x \leftarrow &T_\lambda(t_1 + t_2)x = T_\lambda(t_1)T_\lambda(t_2)x \\ = &T_\lambda(t_1)T(t_2)x + T_\lambda(t_1)\left(T_\lambda(t_2) - T(t_2)\right)x \\ \to &0\end{aligned}$$

as $\lambda \to \infty$ since $\|T_\lambda(t_1)\| \leq Me^{2\omega t_1}$ for λ large enough. Moreover, it follows from (13.4) that

$$\|T(t)\| \leq Me^{\omega t}$$

and thus T is a C^0-semigroup on X. Next we show that A is an infinitesimal generator of T. It follows from Lemma 13.4 that

$$T_\lambda(t)x - x = A(\lambda)\int_0^t T_\lambda(s)x\,ds = \int_0^t T_\lambda(s)A(\lambda)x\,ds.$$

For $x \in D(A)$, λ large enough and $s \in [0, t]$ we estimate

$$\begin{aligned}\|T_\lambda(s)A(\lambda)x - T(s)Ax\| \leq &\|T_\lambda(s)(A(\lambda)x - Ax)\| + \|T_\lambda(s)Ax - T(s)Ax\| \\ \leq &Me^{2\omega t}\|A(\lambda)x - Ax\| + \|(T_\lambda(s) - T(s))Ax\| \\ \to &0\end{aligned}$$

as $\lambda \to \infty$ uniformly in $s \in [0, t]$. Hence, for $x \in D(A)$, we obtain

$$T(t)x - x = \int_0^t T(s)Ax\,ds$$

and therefore we get for all $x \in D(A)$ that

$$\lim_{t\searrow 0}\frac{T(t)x - x}{t} = \lim_{t\searrow 0}\frac{1}{t}\int_0^t T(s)Ax\,ds = Ax$$

since the function in the integral is continuous in s. Altogether we have shown that A is an infinitesimal generator of the C^0-semigroup T.

Finally, assume that $\tilde{A} : D(\tilde{A}) \to X$ is also an infinitesimal generator of the C^0-

semigroup T on X with $D(A) \subset D(\tilde{A})$. Then it follows from Lemma 13.7 and our assumption that $(\omega, \infty) \subset \varrho(\tilde{A})$ and $(\omega, \infty) \subset \varrho(A)$. Hence for $\lambda \in (\omega, \infty)$ we have that the operator $(\lambda \operatorname{id} - \tilde{A})$ is surjective and

$$(\lambda \operatorname{id} - \tilde{A}) D(A) = X$$

since $\tilde{A} = A$ on $D(A)$. But the map $(\lambda \operatorname{id} - \tilde{A}) : D(\tilde{A}) \to X$ is also injective and hence we conclude that $D(A) = D(\tilde{A})$ and $\tilde{A} = A$. □

Remarks:

1. Under the assumptions of Theorem 13.9 we have that $\{\lambda \in \mathbb{C} \colon \mathrm{Re}\,(\lambda) > \omega\} \subset \varrho(A)$ and $\|R_A(\lambda)^n\| \leq \frac{M}{(\mathrm{Re}\,(\lambda) - \omega)^n}$.

2. A C^0-semigroup T on X is uniquely determined by its infinitesimal generator.

   In order to see this, we assume that T_1, T_2 are two C^0-semigroups with infinitesimal generator A. It follows from Lemma 13.2 that there exist $\omega_1, \omega_2 \in [-\infty, \infty)$ so that for $\omega_0 := \max\{\omega_1, \omega_2\}$ we have for $i = 1, 2$, all $t \geq 0$ and all $\omega > \omega_0$

   $$\|T_i(t)\| \leq c e^{\omega t}.$$

   Thus, we can apply Lemma 13.7 and we get for all $\lambda > \omega_0$ that

   $$R_A(\lambda) x = \int_0^\infty e^{-\lambda t} T_i(t) x \, dt.$$

   By letting $T(t) = T_1(t) - T_2(t)$ we conclude that

   $$\int_0^\infty e^{-\lambda t} T(t) x \, dt = 0$$

   for all $\lambda > \omega$. For $\phi \in X'$ and $\omega > \omega_0$ we define $f(t) = e^{-\omega t} \phi(T(t) x$ and we note that $f \in L^1(0, \infty)$ and

   $$\int_0^\infty e^{-\mu t} f(t) \, dt = 0$$

   for all $\mu \geq 0$. Next we do the substitution $t = -\log u$ and we get

   $$\int_0^1 u^\mu \frac{f(-\log u)}{u} \, du = 0$$

   for all $\mu \geq 0$. The function $g : (0,1) \to \mathbb{R}$, $g(u) = \frac{f(-\log u)}{u}$ is in $L^1(0,1)$ since $\int_0^1 |g(u)| \, du = \int_0^\infty |f(t)| \, dt$. Moreover, by the Theorem of Weierstrass, we know that the polynomials u^μ are dense in $C^0(0,1)$ and hence it follows from the Fundamental Theorem of the Calculus of Variations that $g \equiv 0$ which yields

$f \equiv 0$. By Lemma 4.4 this finally shows that $T \equiv 0$.

Example:

We look again at the heat equation on a smooth domain $\Omega \subset \mathbb{R}^n$, i.e. we want to solve the PDE

$$\begin{cases} \partial_t u &= \Delta u, & \text{in } \Omega \times (0,\infty) \\ u &= 0, & \text{on } \partial\Omega \times (0,\infty) \\ u &= u_0, & \text{on } \Omega \times \{0\} \end{cases}$$

with $u_0 \in W^{2,2} \cap W_0^{1,2}(\Omega)$. We want to use the Theorem of Hille-Yosida and for this we let $X = L^2(\Omega)$, $A = \Delta$ and $D(A) = W^{2,2} \cap W_0^{1,2}(\Omega)$. Since $C_c^\infty(\Omega)$ is dense in X and in $D(A)$ it follows that $D(A)$ is dense in X.

First we claim that $(0,\infty) \subset \varrho(A)$, which is equivalent to the fact that for all $\lambda \in (0,\infty)$ we have that the operator $\lambda \operatorname{id} - \Delta\colon W^{2,2} \cap W_0^{1,2}(\Omega) \to L^(\Omega)$ is bijective. Or in other words, we have to show that for every $\lambda > 0$ and every $f \in L^2(\Omega)$ there exists a unique solution $u \in W^{2,2} \cap W_0^{1,2}(\Omega)$ of the PDE

$$-\Delta u + \lambda u = f. \tag{13.5}$$

In order to see this, we note that the associated bilinear form $B : W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega) \to \mathbb{R}$

$$B(u,v) = \int (\langle Du, Dv \rangle + \lambda uv)$$

is coercive with the estimate

$$B(u,u) \geq \min(1,\lambda) \|u\|^2_{W^{1,2}}.$$

Hence the existence of a unique solution $u \in W_0^{1,2}(\Omega)$ of the above PDE follows from the Lax-Milgram Theorem (Theorem 9.10). The fact that u is additionally in $W^{2,2}(\Omega)$ is shown in standard PDE courses.

Next we have to show that for all $\lambda > 0$ we have the estimate $\|R_A(\lambda)^n\| \leq \lambda^{-n}$. In order to do this, we let $f \in L^2(\Omega)$ and we let $u \in W^{2,2} \cap W_0^{1,2}(\Omega)$ be the unique solution of (13.5) which is equivalent to the fact that $u = R_A(\lambda) f$. We get

$$\int |\nabla u|^2 + \lambda u^2 = B(u,u) = \int fu \leq \frac{1}{2\lambda} \int |f|^2 + \frac{\lambda}{2} \int u^2$$

and therefore

$$\frac{\lambda}{2} \int u^2 \leq \frac{1}{2\lambda} \int f^2$$

which yields

$$\|u\|_{L^2} \leq \lambda^{-1} \|f\|_{L^2}$$

or

$$\|R_A(\lambda)\| \leq \lambda^{-1}.$$

It remains to show that A is closed. For this we let $x_n \in D(A)$ be a sequence with $x_n \to x$ in $D(A)$ and $Ax_n \to y$ in X. Let $\lambda > 0$ so that the operator $(\lambda \operatorname{id} - A) : D(A) \to X$ is bijective. Hence, for every $\lambda > 0$ there exists $z \in D(A)$ with $(\lambda \operatorname{id} - A)z = \lambda x - y$. Using the fact that $\|R_A(\lambda)\| \leq \lambda^{-1}$ we estimate

$$\|x_n - z\| \leq \lambda^{-1}\|(\lambda \operatorname{id} - A)(x_n - z)\| \leq \lambda^{-1}\left(\lambda\|x_n - x\| + \|Ax_n - y)\|\right) \to 0$$

as $n \to \infty$. Hence we conclude that $z = x \in D(A)$ and $Ax = y$.

Finally, Theorem 13.9 implies that A generates a C^0-semigroup T on X with $\|T(t)\| \leq 1$ for all $t \geq 0$ and by Lemma 13.4 it follows that $u(t) := T(t)u_0$ is a solution of our problem.