Functional Analysis I
Winter 2006/07

This course is based on the textbooks of Hans Wilhelm Alt [Alt02] and Michael Reed and Barry Simon [RS75] on Functional Analysis. The concepts and notation are based on the course "Einführung in die Funktionalanalysis" held in winter 2005/06.

Contents

1 Compact operators

In this section X, Y are Banach spaces over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with norm $\|\cdot\|$.

1.1 Definition and examples

Definition 1.1 *The set of the* **compact (linear) operators** *from X to Y is defined by*

$$K(X;Y) := \{T \in L(X;Y) \,|\, T(U_1(0)) \textit{ is totally bounded}\}.$$

Lemma 1.2 For $T \in L(X;Y)$ the following properties are equivalent:
(i) $T \in K(X;Y)$.
(ii) $\overline{T(U_1(0))}$ is compact in Y.
(iii) $T(M)$ is totally bounded for all bounded $M \subset X$.
(iv) For all bounded sequences $(x_n)_{n\in\mathbb{N}}$ in X the sequence $(Tx_n)_{n\in\mathbb{N}}$ has a convergent subsequence.

Proof: (i) is equivalent to (ii): Corollary E3.4(iii).

(ii) implies (iv): Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X. Then there exists $r > 0$ such that $\|x_n\| < r$ for all $n \in \mathbb{N}$. Set $u_n := x_n/r$, $n \in \mathbb{N}$. Then $(Tu_n)_{n\in\mathbb{N}}$ is a sequence in $T(U_1(0))$. Since $\overline{T(U_1(0))}$ is compact by (ii), $(Tu_n)_{n\in\mathbb{N}}$ has a convergent subsequence, i.e., $(Tx_{n_k}/r)_{k\in\mathbb{N}}$ is convergent for some subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$. But then also $(Tx_{n_k})_{k\in\mathbb{N}}$ is convergent.

(iv) implies (iii): Let $(y_n)_{n\in\mathbb{N}}$ be a sequence in $T(M)$. Then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in M such that $Tx_n = y_n$ for all $n \in \mathbb{N}$. Since M is bounded, also $(x_n)_{n\in\mathbb{N}}$ is bounded. Then (iv) implies that $(y_n)_{n\in\mathbb{N}}$ has a convergent subsequence. Thus, each sequence in $T(M)$ has a convergent subsequence. This implies that $T(M)$ is totally bounded (see the proof of Proposition E3.3)

(iii) implies (i): Obvious. ■

Example 1.3 (i) Let Y be finite dimensional. Then $K(X;Y) = L(X;Y)$.
(ii) Let $T \in L(X;Y)$ with $\dim\mathcal{R}(T) < \infty$ (**finite rank operators**). Then $T \in K(X;Y)$.
(iii) Let $k : [0,1] \times [0,1] \to \mathbb{K}$ be continuous. Then the linear mapping $T : C([0,1]) \to C([0,1])$ defined by

$$(Tf)(x) := \int_0^1 k(x,y) f(y)\, dy, \quad f \in C([0,1]),\ x \in [0,1],$$

is compact.
(iv) Let $\Omega_1 \subset \mathbb{R}^{d_1}$, $\Omega_2 \subset \mathbb{R}^{d_2}$ be open, $1 < p < \infty$, $1 < q < \infty$, $\frac{1}{p} + \frac{1}{p'} = 1$, and $K : \Omega_1 \times \Omega_2 \to \mathbb{K}$ measurable with

$$\|K\| := \left(\int_{\Omega_1} \left(\int_{\Omega_2} |K(x,y)|^{p'} \, dy \right)^{\frac{q}{p'}} \, dx \right)^{\frac{1}{q}} < \infty.$$

Then the linear mapping $T : L^p(\Omega_2) \to L^q(\Omega_1)$ defined by

$$(Tf)(x) := \int_{\Omega_2} K(x,y) f(y) \, dy, \quad f \in L^p(\Omega_2), \, x \in \Omega_1,$$

is bounded with $\|T\|_{L(L^p;L^q)} \leq \|K\|$. Furthermore one can show that T is compact. The function K is called the **integral kernel** corresponding to T.
(v) Let

$$D := \{f \in C^2([0,\pi]) | f(0) = f(\pi) = 0\} \subset L^2([0,\pi]).$$

Such boundary conditions are called **Dirichlet boundary condition**. We consider the linear mapping $L : D \to L^2([0,\pi])$ defined by

$$Lf := f'', \quad f \in D.$$

Then L is injective and $L^{-1} : \mathcal{R}(L) \to D$ extends to a self-adjoint, compact operator on $L^2([0,\pi])$. **Eigenfunctions** of L (and thus of L^{-1}) are given by

$$f_n := \sin(n\cdot), \quad n \in \mathbb{N},$$

with corresponding **eigenvalues** $-n^2$, $n \in \mathbb{N}$ $(-1/n^2, \, n \in \mathbb{N})$. Moreover, $(f_n)_{n\in\mathbb{N}}$ is an orthogonal basis of $L^2([0,\pi])$.

This statement generalizes to the **Laplace operator**

$$\Delta := \sum_{i=1}^{d} \partial_i^2$$

with Dirichlet boundary conditions for quite general bounded subsets $\Omega \subset \mathbb{R}^d$. Of course, with different eigenfunctions and eigenvalues. This can be shown by an application of the spectral theorem for compact operators, because Δ^{-1} is a self-adjoint, compact operator on $L^2(\Omega)$.

Proof: (i): $T \in L(X;Y)$ maps bounded sets to bounded sets. But bounded sets in finite dimensional spaces are totally bounded by Corollary E3.4(iv) (there exists $n \in \mathbb{N}$ such that Y is isometrically isomorph to $\mathbb{K}^n$ equipped with the norm induced by $\|\cdot\|$).

(ii): Since, in particular, $T \in L(X;\mathcal{R}(T))$, this follows immediately from (i).

(iii): See Exercise E4.3, E4.4.

(iv): Will be shown later.

(v): L is injective, because if $Lf = 0$ the integration by parts formula yields

$$
\begin{aligned}
0 = (Lf, f)_{L^2} &= \int_0^\pi f''(x)f(x)\,dx = -\int_0^\pi f'(x)f'(x)\,dx + f'f\Big|_0^\pi \\
&= -\int_0^\pi f'(x)f'(x)\,dx + f'(\pi)f(\pi) - f'(0)f(0) = -\int_0^\pi f'(x)f'(x)\,dx.
\end{aligned}
$$

Thus, $f' = 0$. This together with $f(0) = 0$ implies $f = 0$. Hence there exits $L^{-1} : \mathcal{R}(L) \to D$. Later on we will show that L^{-1} is bounded and $\overline{\mathcal{R}(L)} = L^2([0,\pi])$. Thus, L^{-1} extends to a bounded operator on $L^2([0,\pi])$, see Exercise 1.1. $L^{-1} \in K(L^2([0,\pi]))$ we will show later.

Since L is symmetric on D w.r.t. $(\cdot,\cdot)_{L^2}$, i.e.,

$$
\begin{aligned}
(Lf, g)_{L^2} &= \int_0^\pi f''(x)g(x)\,dx = -\int_0^\pi f'(x)g'(x)\,dx + f'g\Big|_0^\pi \\
&= -\int_0^\pi f'(x)g'(x)\,dx = (f, Lg)_{L^2}, \quad \text{for all } f, g \in D,
\end{aligned}
$$

L^{-1} is self-adjoint on $L^2([0,\pi])$.

The statement about eigenfunctions and eigenvalues is obvious, except for being a basis. This also will be shown later. ■

1.2 Elementary properties

Lemma 1.4 (i): $K(X;Y)$ is a closed, subspace of $L(X;Y)$.
(ii): If $T \in L(X;Y)$, $S \in L(Y;Z)$ with Z a Banach space and T or S compact, then also ST is compact.

Proof: (i): $K(X;Y)$ is a subspace, because if $T_1, T_2 \in K(X;Y)$ and $\alpha \in \mathbb{K}$, and if $(x_m)_{m\in\mathbb{N}}$ is a bounded sequence in X, then by Lemma 1.2 there exists a convergent subsequence $(T_1 x_{n_k})_{k\in\mathbb{N}}$. From this one can drop to a further convergent subsequence $(T_2 x_{n_{k_l}})_{l\in\mathbb{N}}$. Then also

$$((\alpha T_1 + T_2)x_{n_{k_l}})_{l\in\mathbb{N}}$$

is convergent. Thus, $\alpha T_1 + T_2$ is compact by Lemma 1.2.

For proving $K(X;Y)$ being closed, let $(T_n)_{n\in\mathbb{N}}$ be a sequence in $K(X;Y)$ which converges to $T \in L(X;Y)$. Let $\varepsilon > 0$ and choose $n_\varepsilon \in \mathbb{N}$ such that

$$\|T - T_{n_\varepsilon}\|_{L(X;Y)} < \frac{\varepsilon}{2}.$$

Since T_{n_ε} is compact, there exit balls $U_{\frac{\varepsilon}{2}}(y_i)$, $i = 1, \ldots, m_\varepsilon$, such that

$$T_{n_\varepsilon}(U_1(0)) \subset \bigcup_{i=1}^{m_\varepsilon} U_{\frac{\varepsilon}{2}}(y_i).$$

But then is

$$T(U_1(0)) \subset \bigcup_{i=1}^{m_\varepsilon} U_\varepsilon(y_i).$$

Thus, T is compact.

(ii): Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X. Since T is continuous also $(Tx_n)_{n\in\mathbb{N}}$ is bounded. If S is compact, then $(STx_n)_{n\in\mathbb{N}}$ has a convergent subsequence. If T is compact, there exists a convergent subsequence $(Tx_{n_k})_{k\in\mathbb{N}}$ and continuity of S implies convergence of $(STx_{n_k})_{k\in\mathbb{N}}$. So in both cases ST is compact. ■

Lemma 1.5 A projection $P \in P(X)$ is compact, iff $\dim\mathcal{R}(P) < \infty$.

Proof: Finite rank operators are compact by Example 1.3(ii). The fact that compact projections have a finite dimensional range we know from Exercise E5.3. ■

Lemma 1.6 Let Y be a Hilbert space and $T \in L(X;Y)$. Then T is compact, iff there exists a sequence of finite rank operators which converges to T.

Proof: If T is the limit of finite rank operators, then by Lemma 1.4(i) T is compact, because from Example 1.3(ii) we already know that finite rank operators are compact.

Now let $T \in K(X;Y)$ and $\varepsilon > 0$. Then there exist balls $U_\varepsilon(y_i)$, $i = 1, \ldots, m_\varepsilon$, such that

$$T(U_1(0)) \subset \bigcup_{i=1}^{m_\varepsilon} U_\varepsilon(y_i).$$

Set

$$Y_\varepsilon := \operatorname{span}\{y_1, \ldots, y_{m_\varepsilon}\}$$

and denote by P_ε the orthogonal projection on Y_ε (which exists due to Corollary E5.14). Then $\|Id - P_\varepsilon\|_{L(Y)} \leq 1$, because

$$\|y - P_\varepsilon y\|_Y^2 = (y - P_\varepsilon y, y - P_\varepsilon y)_Y = (y, y - P_\varepsilon y)_Y \leq \|y\| \, \|y - P_\varepsilon y\|$$

for all $y \in Y$ due to the properties of P_ε and Cauchy–Schwartz inequality. Note that

$$T_\varepsilon := P_\varepsilon T : X \to Y_\varepsilon$$

is a finite rank operator. Now for $x \in U_1(0)$ there exists $i_0 \in \{1, \ldots, m_\varepsilon\}$ such that $Tx \in U_\varepsilon(y_{i_0})$. Hence

$$(T - T_\varepsilon)x = (Id - P_\varepsilon)Tx = (Id - P_\varepsilon)(Tx - y_{i_0})$$

and therefore

$$\|(T - T_\varepsilon)x\| \leq \|Id - P_\varepsilon\|_{L(Y)} \|Tx - y_{i_0}\| < \varepsilon \quad \text{for all } x \in U_1(0).$$

Thus, $\|T - T_\varepsilon\|_{L(X;Y)} \leq \varepsilon$. ■

1.3 Spectrum and resolvent

Definition 1.7 *The* **resolvent set** *of $T \in L(X)$ is defined by*

$$\rho(T) := \big\{\lambda \in \mathbb{K} \,\big|\, \mathcal{N}(\lambda Id - T) = \{0\}, \\ \mathcal{R}(\lambda Id - T) = X \text{ and } (\lambda Id - T)^{-1} \in L(X)\big\}$$

and the **spectrum** *by*

$$\sigma(T) := \mathbb{K} \setminus \rho(T).$$

For $\lambda \in \rho(T)$ *the operator*

$$R(\lambda; T) := (\lambda Id - T)^{-1} \in L(X)$$

is called **resolvent** *of* T *at* λ *and the function*

$$\rho(T) \ni \lambda \mapsto R(\lambda; T) \in L(X)$$

is called **resolvent function**.

The spectrum can be decomposed into the **point spectrum**

$$\sigma_p(T) := \big\{\lambda \in \sigma(T) \,\big|\, \mathcal{N}(\lambda Id - T) \neq \{0\}\big\},$$

the **continuous spectrum**

$$\sigma_c(T) := \big\{\lambda \in \sigma(T) \,\big|\, \mathcal{N}(\lambda Id - T) = \{0\} \text{ and } \mathcal{R}(\lambda Id - T) \neq X, \text{but } \overline{\mathcal{R}(\lambda Id - T)} = X\big\},$$

and the **residual spectrum**

$$\sigma_r(T) := \big\{\lambda \in \sigma(T) \,\big|\, \mathcal{N}(\lambda Id - T) = \{0\} \text{ and } \overline{\mathcal{R}(\lambda Id - T)} \neq X\big\}.$$

Remark 1.8 (i) The condition $(\lambda Id - T)^{-1} \in L(X)$ in the definition of $\rho(T)$ is already implied by $(\lambda Id - T) \in L(X)$, $(\lambda Id - T)$ injective and surjective by the inverse mapping theorem, see Theorem 3.9 below. This we will prove later in this course.
(ii) $\lambda \in \sigma_p(T)$ is equivalent to the existence of an $0 \neq x \in X$ such that $Tx = \lambda x$. Then x is called **eigenvector** corresponding to the **eigenvalue** λ. The space $\mathcal{N}(\lambda Id - T)$ is called **eigenspace** of T to the eigenvalue λ. The eigenspace is a T-invariant subspace of X. A subspace $Y \subset X$ is called T**-invariant**, if $T(Y) \subset Y$.

Proposition 1.9 *Let* $T \in L(X)$. $\rho(T) \subset \mathbb{K}$ *is open and the resolvent function* $R(\cdot; T)$ *is a* $\mathbb{K}$*-analytic mapping from* $\rho(T)$ *to* $L(X)$. *Furthermore*

$$\|R(\lambda; T)\|_{L(X)}^{-1} \leq \operatorname{dist}(\lambda, \sigma(T)), \quad \lambda \in \rho(T).$$

Remark 1.10 A mapping $F : D \to Y$, $D \subset \mathbb{K}$ open, Y Banach space, is called $\mathbb{K}$-analytic, if for each $\lambda_0 \in D$ there exists a ball $U_{r_0}(\lambda_0) \subset D$, $r_0 > 0$ and a sequence $(y_n)_{n\in\mathbb{N}}$ in Y, such that

$$F(\lambda) = \sum_{n=1}^{\infty} y_n(\lambda - \lambda_0)^n, \quad \lambda \in U_{r_0}(\lambda_0).$$

$\mathbb{C}$-analytic mappings with values in Y are holomorphic and many results from Complex Analysis generalize to this infinite dimensional setting, see e.g. [Alt02, App. 8], [RS75, Chap. VI]. See also the proof of Lemma 6.8 below, where this will be shown exemplary by using the Hahn-Banach theorem

Proof of Proposition 1.9: Let $\lambda \in \rho(T)$. Then we have for all $\mu \in \mathbb{K}$:

$$(\lambda - \mu)Id - T = (\lambda Id - T) - \mu Id = (\lambda Id - T)(Id - \mu R(\lambda; T)).$$

The operator

$$S(\mu) := Id - \mu R(\lambda; T)$$

is continuously invertible for

$$|\mu| \|R(\lambda; T)\|_{L(X)} < 1$$

by Proposition E4.6. Then $\lambda - \mu \in \rho(T)$ with

$$R(\lambda - \mu; T) = S(\mu)^{-1} R(\lambda; T) = \sum_{k=0}^{\infty} \mu^k R(\lambda; T)^{k+1}$$

again by Proposition E4.6. Therefore, with $d := \|R(\lambda; T)\|_{L(X)}^{-1}$ we obtain

$$U_d(\lambda) \subset \rho(T),$$

i.e. $\text{dist}(\lambda, \sigma(T)) \geq d$. ■

Proposition 1.11 *Let $T \in L(X)$ and $\mathbb{K} = \mathbb{C}$. $\sigma(T) \subset \mathbb{C}$ is compact and non-empty (if $X \neq \{0\}$) with*

$$r(T) := \sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{m\to\infty} \|T^m\|_{L(X)}^{\frac{1}{m}} \leq \|T\|_{L(X)}.$$

$r(T)$ is called **spectral radius** *of T.*

Proof: Let $\lambda \neq 0$. By Proposition E4.6.

$$Id - \frac{T}{\lambda}$$

is continuously invertible, if

$$\left\| \frac{T}{\lambda} \right\|_{L(X)} < 1,$$

i.e. $|\lambda| > \|T\|_{L(X)}$. Then

$$R(\lambda; T) = \frac{1}{\lambda} \left(Id - \frac{T}{\lambda} \right)^{-1} = \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}. \tag{1.1}$$

Thus

$$r := \sup_{\lambda \in \sigma(T)} |\lambda| \leq \|T\|_{L(X)}.$$

Observe that

$$\lambda^m Id - T^m = (\lambda Id - T) p_m(T) = p_m(T)(\lambda Id - T)$$

where

$$p_m(T) = \sum_{k=0}^{m-1} \lambda^{m-1-k} T^k.$$

Hence $\lambda \in \sigma(T)$ implies $\lambda^m \in \sigma(T^m)$. Then as before

$$|\lambda^m| \leq \|T^m\|_{L(X)}$$

and therefore

$$|\lambda| \leq \|T^m\|_{L(X)}^{\frac{1}{m}}.$$

Thus

$$r \leq \liminf_{m \to \infty} \|T^m\|_{L(X)}^{\frac{1}{m}}.$$

Now it is left to show that

$$r \geq \limsup_{m\to\infty} \|T^m\|_{L(X)}^{\frac{1}{m}}.$$

Proposition 1.9 implies that $R(\cdot, T)$ is $\mathbb{C}$-analytic in $\mathbb{C}\backslash\overline{U_r(0)}$ ($\mathbb{C}$ if $\sigma(T) = \emptyset$). Therefore the integral

$$\int_{\partial U_s(0)} \lambda^m R(\lambda; T)\, d\lambda, \quad m \in \mathbb{N}_0,$$

for $s > r$ is independent of s. Hence together with (1.1) we obtain

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial U_s(0)} \lambda^m R(\lambda; T)\, d\lambda &= \frac{1}{2\pi i}\int_{\partial U_s(0)} \sum_{k=0}^{\infty} \lambda^{m-k-1} T^k\, d\lambda \\ &= \frac{1}{2\pi}\sum_{k=0}^{\infty} s^{m-k}\int_0^{2\pi} \exp(i\theta(m-k))\, d\theta\, T^k = \sum_{k=0}^{\infty} s^{m-k}\delta_{m,k} T^k = T^m.\end{aligned}$$

The exchange of infinite sum and integral is justified by the uniform convergence of the series on $\partial U_s(0)$. Hence we have for $m \in \mathbb{N}_0$ and $s > r$

$$\|T^m\|_{L(X)} = \frac{1}{2\pi}\left\|\int_{\partial U_s(0)} \lambda^m R(\lambda; T)\, d\lambda\right\|_{L(X)} \leq s^{m+1} \sup_{|\lambda|=s} \|R(\lambda; T)\|_{L(X)}. \quad (1.2)$$

Therefore we obtain for $s > r$

$$\limsup_{m\to\infty} \|T^m\|_{L(X)}^{\frac{1}{m}} \leq s \limsup_{m\to\infty} (s \sup_{|\lambda|=s} \|R(\lambda; T)\|_{L(X)})^{\frac{1}{m}} = s \text{ (or } 0).$$

Since this holds for all $s > r$ we obtain the desired inequality:

$$\limsup_{m\to\infty} \|T^m\|_{L(X)}^{\frac{1}{m}} \leq r.$$

Hence the statement concerning the spectral radius is proved. In the case when $\sigma(T) = \emptyset$, we get from (1.2) ($m = 0$)

$$\|Id\|_{L(X)} \leq s \sup_{|\lambda|\leq 1} \|R(\lambda; T)\|_{L(X)} \quad \text{for all } 0 < s \leq 1.$$

Since the resolvent in this case is $\mathbb{C}$-analytic on $\mathbb{C}$ we have

$$\sup_{|\lambda|\leq 1} \|R(\lambda; T)\|_{L(X)} < \infty.$$

Thus $\|Id\|_{L(X)} = 0$, i.e. $X = \{0\}$. ■

Analyzing the proof of Proposition 1.11 we obtain in the real case the following corollary.

Corollary 1.12 *Let $T \in L(X)$ and $\mathbb{K} = \mathbb{R}$. $\sigma(T) \subset \mathbb{R}$ is compact with*

$$r(T) = \sup_{\lambda \in \sigma(T)} |\lambda| \leq \|T^m\|_{L(X)}^{\frac{1}{m}} \leq \|T\|_{L(X)} \quad \textit{for all} \quad m \in \mathbb{N}.$$

Remark 1.13 (i) If $\dim X < \infty$, then $\sigma(T) = \sigma_p(T)$.
(ii) If $\dim X = \infty$ and $T \in K(X)$, then $0 \in \sigma(T)$. In general, however, 0 might not be an eigenvalue.

Proof: (i): If $\lambda \in \sigma(T)$, then $\lambda Id - T$ is not bijective. Since $\dim X < \infty$, this implies that $\lambda Id - T$ is not injective, i.e., $\lambda \in \sigma_p(T)$.

(ii): Let $T \in K(X)$ and $0 \in \rho(T)$. Then $T^{-1} \in L(X)$ and therefore also

$$Id = T^{-1}T \in K(X)$$

by Lemma 1.4(ii). Thus, X is finite dimensional by Theorem E3.8 (Heine–Borel). See Exercise 1.3(ii) for a compact operator not having 0 as an eigenvalue. ■

1.4 Fredholm operators

Definition 1.14 *A mapping $A \in L(X;Y)$ is called* **Fredholm operator***, iff:*
(i) $\dim \mathcal{N}(A) < \infty$,
(ii) $\mathcal{R}(A)$ *is closed,*
(iii) $\operatorname{codim} \mathcal{R}(A) < \infty$.

The **index** *of a Fredholm operator is defined by*

$$\operatorname{ind}(A) := \dim \mathcal{N}(A) - \operatorname{codim} \mathcal{R}(A).$$

Remark 1.15 One says a closed subset Y of a Banach space X has finite **codimension** ($\operatorname{codim} Y < \infty$), if

$$X = Y \oplus Z$$

and $\dim Z = n$ for some $n \in \mathbb{N}_0$. Then $\operatorname{codim} Y = n$ (codim$Y$ is independent of the choice of Z, see Corollary 5.5 below.

Proposition 1.16 *Let $T \in K(X)$. Then $A := Id - T$ is a Fredholm operator with index* 0.

Proof: Step 1: $\dim \mathcal{N}(A) < \infty$: Since $Ax = 0$ is equivalent to $x = Tx$, we have

$$U_1(0) \cap \mathcal{N}(A) \subset T(U_1(0)).$$

Thus the unit ball in $\mathcal{N}(A)$ is totally bounded. Therefore $\dim \mathcal{N}(A) < \infty$ by Theorem E3.8 (Heine–Borel).

Step 2: $\mathcal{R}(A)$ is closed: Let $x \in \overline{\mathcal{R}(A)}$ and $(x_n)_{n\in\mathbb{N}}$ a sequence in X such that

$$\lim_{n\to\infty} Ax_n = x.$$

W.l.o.g., we may assume that

$$\|x_n\| \le 2d_n \quad \text{with} \quad d_n := \text{dist}(x_n, \mathcal{N}(A)), \quad n \in \mathbb{N},$$

otherwise choose $(a_n)_{n\in\mathbb{N}}$ in $\mathcal{N}(A)$ such that

$$\|x_n - a_n\| \le 2\ \text{dist}(x_n, \mathcal{N}(A)), \quad n \in \mathbb{N},$$

and use the sequence $(\tilde{x}_n)_{n\in\mathbb{N}}$ with $\tilde{x}_n := x_n - a_n$, $n \in \mathbb{N}$. Note that

$$\text{dist}(\tilde{x}_n, \mathcal{N}(A)) = \ \text{dist}(x_n, \mathcal{N}(A)), \quad n \in \mathbb{N}.$$

Assume that $(d_n)_{n\in\mathbb{N}}$ is not bounded. Then there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} d_{n_k} = \infty$. Set

$$y_k := \frac{x_{n_k}}{d_{n_k}}, \quad k \in \mathbb{N}.$$

Then

$$\lim_{k\to\infty} Ay_k = \lim_{k\to\infty} \frac{Ax_{n_k}}{d_{n_k}} = 0.$$

Since $(y_k)_{k\in\mathbb{N}}$ is bounded and T compact, there exists a subsequence $(k_l)_{l\in\mathbb{N}}$ and $y \in X$ such that

$$\lim_{l\to\infty} Ty_{k_l} = y.$$

Hence

$$\lim_{l\to\infty} y_{k_l} = \lim_{l\to\infty} Ay_{k_l} + \lim_{l\to\infty} Ty_{k_l} = y. \tag{1.3}$$

Since A is continuous, it follows

$$Ay = \lim_{l\to\infty} Ay_{k_l} = 0.$$

Thus $y \in \mathcal{N}(A)$. This implies

$$\begin{aligned} \|y_{k_l} - y\| \geq & \ \operatorname{dist}(y_{k_l}, \mathcal{N}(A)) \\ = & \ \operatorname{dist}\left(\frac{x_{n_{k_l}}}{d_{n_{k_l}}}, \mathcal{N}(A)\right) = \frac{\operatorname{dist}(x_{n_{k_l}}, \mathcal{N}(A))}{d_{n_{k_l}}} = 1. \end{aligned}$$

But this contradicts (1.3). Hence, $(d_n)_{n\in\mathbb{N}}$ is bounded and therefore also $(x_n)_{n\in\mathbb{N}}$. Now, because T is compact, we can conclude the existence of a subsequence $(n_k)_{k\in\mathbb{N}}$ and $z \in X$ such that

$$\lim_{k\to\infty} Tx_{n_k} = z.$$

Thus

$$x = \lim_{k\to\infty} Ax_{n_k} = A(\lim_{k\to\infty} Ax_{n_k} + \lim_{k\to\infty} Tx_{n_k}) = A(x+z),$$

i.e., $x \in \mathcal{R}(A)$.

Step 3: $\mathcal{N}(A) = \{0\}$ implies $\mathcal{R}(A) = X$: Assume there exists $x \in X \setminus \mathcal{R}(A)$. Then

$$A^n x \in \mathcal{R}(A^n) \setminus \mathcal{R}(A^{n+1}) \quad \text{for all} \quad n \in \mathbb{N}.$$

Because if there would exist $y \in X$ such that $A^n x = A^{n+1} y$, then

$$A^n(x - Ay) = 0.$$

But then $\mathcal{N}(A) = \{0\}$ implies (inductively)

$$x - Ay = 0,$$

i.e. $x \in \mathcal{R}(A)$. Contradiction!

Furthermore, $\mathcal{R}(A^{n+1})$, $n \in \mathbb{N}$, is closed by Step 2, because

$$A^{n+1} = (Id - T)^{n+1} = Id + \sum_{k=1}^{n+1} \binom{n+1}{k} (-T)^k$$

and

$$\sum_{k=1}^{n+1} \binom{n+1}{k} (-T)^k$$

is compact by Lemma 1.4. Hence there exists $a_{n+1} \in \mathcal{R}(A^{n+1})$ with

$$0 < \|A^n x - a_{n+1}\| \le 2 \operatorname{dist}\big(A^n x, \mathcal{R}(A^{n+1})\big).$$

Now consider

$$x_n := \frac{A^n x - a_{n+1}}{\|A^n x - a_{n+1}\|} \in \mathcal{R}(A^n), \quad n \in \mathbb{N}.$$

We have $\operatorname{dist}(x_n, \mathcal{R}(A^{n+1})) \ge \frac{1}{2}$, because for all $y \in \mathcal{R}(A^{n+1})$ is

$$\|x_n - y\| = \frac{\|A^n x - (a_{n+1} + \|A^n x - a_{n+1}\| y)\|}{\|A^n x - a_{n+1}\|} \ge \frac{\operatorname{dist}\big(A^n x, \mathcal{R}(A^{n+1})\big)}{\|A^n x - a_{n+1}\|} \ge \frac{1}{2}.$$

Thus for $m > n$

$$\|Tx_n - Tx_m\| = \|x_n - (Ax_n + x_m - Ax_m)\| \ge \frac{1}{2},$$

because $Ax_n + x_m - Ax_m \in \mathcal{R}(A^{n+1})$. Therefore, $(Tx_n)_{n\in\mathbb{N}}$ has no convergent subsequence although $(x_n)_{n\in\mathbb{N}}$ is bounded. This is in contradiction to the compactness of T.

Step 4: codim$\mathcal{R}(A) \le$ dim$\mathcal{N}(A)$: By Step 1 $n :=$dim$\mathcal{N}(A) \in \mathbb{N}_0$. Let $x_1, \dots, x_n \in X$ be a basis of $\mathcal{N}(A)$.

Assume the statement is not true. Then there exist linear independent vectors $y_1, \dots, y_n \in X$ such that $\operatorname{span}\{y_1, \dots, y_n\} \oplus \mathcal{R}(A)$ is a strict subset of X. As a corollary of the Hahn–Banach theorem (see Corollary 2.5(iii) below) there exist $x_1', \dots, x_n' \in X'$ such that

$$x_k'(x_l) = \delta_{k,l}, \quad k, l = 1, \dots n.$$

Then

$$\tilde{T}x := Tx + \sum_{k=1}^{n} x_k'(x) y_k, \quad x \in X,$$

defines an operator $\tilde{T} \in K(X)$, because T is compact and $\tilde{T} - T$ is finite rank. Set

$$\tilde{A} := Id - \tilde{T}.$$

Then $x \in \mathcal{N}(\tilde{A})$ is equivalent to

$$0 = \tilde{A}x = Ax - \sum_{k=1}^{n} x_k'(x) y_k.$$

Hence $Ax = 0$ and $x_k'(x) = 0$, $k = 1, \dots, n$ (due to the choice of $y_1, \dots, y_n$). Since $x \in \mathcal{N}(A)$ we have

$$x = \sum_{k=1}^{n} \alpha_k x_k \quad \text{for some} \quad \alpha_1, \dots, \alpha_n \in \mathbb{K}.$$

Therefore

$$0 = x_l'(x) = \sum_{k=1}^{n} \alpha_k x_l'(x_k) = \alpha_l, \quad \text{for all } l = 1, \dots, n.$$

Thus $x = 0$, i.e. $\tilde{A}$ is injective. Now, applying Step 3 to $\tilde{A}$, we obtain $\mathcal{R}(\tilde{A}) = X$. Since $\tilde{A}x_l = -y_l$, $l = 1, \dots, n$, and

$$\tilde{A}\Big(x - \sum_{l=1}^{n} x_l'(x) x_l\Big) = Ax \quad \text{for all } x \in X$$

we have

$$X = \mathcal{R}(\tilde{A}) \subset \ \text{span}(y_1, \dots, y_n) \oplus \mathcal{R}(A).$$

Contradiction!

Step 5: codim$\mathcal{R}(A) \geq$ dim$\mathcal{N}(A)$: From Step 4 we know that

$$\mathbb{N} \ni m := \text{codim}\mathcal{R}(A) \leq n := \text{dim}\mathcal{N}(A).$$

First we reduce the problem to the case $m = 0$. Choose $x_1, \dots, x_n \in X$ and $x'_1, \dots, x'_n \in X'$ as in Step 4 and $y_1, \dots, y_m \in X$ such that

$$X = \operatorname{span}\{y_1, \dots, y_m\} \oplus \mathcal{R}(A).$$

As in Step 4 the mapping

$$\tilde{T}x := Tx + \sum_{k=1}^{m} x'_k(x) y_k, \quad x \in X,$$

is compact and $\tilde{A} := Id - \tilde{T}$ is surjective with

$$\mathcal{N}(\tilde{A}) = \operatorname{span}\{x_i \,|\, m < i \leq n\} \cup \{0\}.$$

Hence it remains to show that $\mathcal{N}(\tilde{A}) = \{0\}$ for surjective $\tilde{A}$. I.e., the problem is reduced to the case $m = 0$.

In the case $m = 0$ is $\mathcal{R}(A) = X$. We assume there exists $x_1 \in \mathcal{N}(A) \setminus \{0\}$. By surjectivity of A, inductively we can choose $x_k \in X$ such that

$$Ax_k = x_{k-1}, \quad k \geq 2.$$

Then

$$x_k \in \mathcal{N}(A^k) \setminus \mathcal{N}(A^{k-1}).$$

By Proposition E3.5, for $k \geq 2$ we can choose

$$z_k \in \mathcal{N}(A^k) \text{ with } \|z_k\| = 1 \text{ and } \operatorname{dist}(z_k, \mathcal{N}(A^{k-1})) \geq \frac{1}{2}.$$

Then we have for $l < k$

$$\|Tz_k - Tz_l\| = \|z_k - (Az_k + z_l - Az_l)\| \geq \frac{1}{2},$$

because $(Az_k + z_l - Az_l) \in \mathcal{N}(A^{k-1})$. I.e., $(Tz_k)_{k \in \mathbb{N}}$ has no convergent subsequence. Since $(z_k)_{k \in \mathbb{N}}$ is bounded, this is in contradiction to the compactness of T. ■

1.5 Spectral theorem

Theorem 1.17 (Riesz–Schauder) *For each operator $T \in K(X)$ holds:*
(i) $\sigma(T)\setminus\{0\}$ *consists of countable many (finite or infinite) eigenvalues with* 0 *as the only possible accumulation point. If $\sigma(T)$ has infinite many elements, then*

$$\sigma(T) = \sigma_p(T) \cup \{0\}.$$

(ii) *For $\lambda \in \sigma(T) \setminus \{0\}$ is*

$$1 \leq n_\lambda := \max\{n \in \mathbb{N} \,|\, \mathcal{N}((\lambda Id - T)^{n-1}) \neq \mathcal{N}((\lambda Id - T)^n)\} < \infty.$$

n_λ *is called* **order** *of λ and* $\dim\mathcal{N}(\lambda Id - T)$ **multiplicity** *of λ.*
(iii) **(Riesz decomposition)** *For $\lambda \in \sigma(T) \setminus \{0\}$ we have:*

$$X = \mathcal{N}((\lambda Id - T)^{n_\lambda}) \oplus \mathcal{R}((\lambda Id - T)^{n_\lambda}).$$

Both spaces are closed and T-invariant. $\mathcal{N}((\lambda Id - T)^{n_\lambda})$ is finite dimensional.
(iv) *For $\lambda \in \sigma(T) \setminus \{0\}$ is*

$$\sigma(T|_{\mathcal{R}((\lambda Id - T)^{n_\lambda})}) = \sigma(T) \setminus \{\lambda\}.$$

Proof: (i): Let $0 \neq \lambda \notin \sigma_p(T)$. Then

$$\mathcal{N}\Big(Id - \frac{T}{\lambda}\Big) = \{0\}, \text{ hence } \quad \mathcal{R}\Big(Id - \frac{T}{\lambda}\Big) = X$$

by Proposition 1.16. Hence $\lambda \in \rho(T)$. This shows

$$\sigma(T) \setminus \{0\} \subset \sigma_p(T).$$

If $\sigma(T)\setminus\{0\}$ has infinite many elements, then we choose $\lambda_n \in \sigma(T)\setminus\{0\}$, $n \in \mathbb{N}$, pairwise different with corresponding eigenvectors e_n, $n \in \mathbb{N}$. Set

$$X_n := \text{span}\{e_1, \dots, e_n\}, \quad n \in \mathbb{N}.$$

The eigenvectors are linear independent, because if we would have

$$e_n = \sum_{k=1}^{n-1} \alpha_k e_k, \quad \alpha_1, \dots, \alpha_{n-1} \in \mathbb{K},$$

with $e_1, \dots, e_{n-1}$ linear independent, then

$$0 = Te_n - \lambda_n e_n = \sum_{k=1}^{n-1} \alpha_k (Te_k - \lambda_n e_k) = \sum_{k=1}^{n-1} \alpha_k (\lambda_k - \lambda_n) e_k,$$

hence $\alpha_k = 0$ for $k = 1, \dots n-1$, i.e. $e_n = 0$. Contradiction!

Therefore X_{n-1} is a proper subspace of X_n. Hence by Proposition E3.5 there exits $x_n \in X_n$ with

$$\|x_n\| = 1 \quad \text{and} \quad \operatorname{dist}\bigl(x_n, X_{n-1}\bigr) \geq \frac{1}{2}.$$

Since $x_n = \alpha_n e_n + \tilde{x}_n$ for some $\alpha_n \in \mathbb{K}$ and $\tilde{x}_n \in X_{n-1}$, T-invariance of X_{n-1} implies

$$Tx_n - \lambda_n x_n = T\tilde{x}_n - \lambda_n \tilde{x}_n \in X_{n-1}.$$

Thus we have for $m < n$

$$\Bigl\| T\Bigl(\frac{x_n}{\lambda_n}\Bigr) - T\Bigl(\frac{x_m}{\lambda_m}\Bigr) \Bigr\| = \Bigl\| x_n + \frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - \frac{1}{\lambda_m} Tx_m \Bigr\| \geq \frac{1}{2},$$

because

$$\frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - \frac{1}{\lambda_m} Tx_m \in X_{n-1}.$$

Therefore

$$\Bigl(T\Bigl(\frac{x_n}{\lambda_n}\Bigr)\Bigr)_{n\in\mathbb{N}}$$

has no convergent subsequence. Since T is compact,

$$\Bigl(\frac{x_n}{\lambda_n}\Bigr)_{n\in\mathbb{N}}$$

can not have a bounded subsequence. Thus

$$\lim_{n\to\infty} \frac{1}{|\lambda_n|} = \lim_{n\to\infty} \Bigl\| \frac{x_n}{\lambda_n} \Bigr\| = \infty,$$

i.e.,

$$\lim_{n\to\infty} \lambda_n = 0.$$

Hence 0 is the only accumulation point of $\sigma(T) \setminus \{0\}$. In particular,

$$\#(\sigma(T) \setminus U_r(0)) < \infty \quad \text{for all } r > 0.$$

Therefore, $\sigma(T)$ is countable.

(ii): Set $A := \lambda Id - T$, $\lambda \in \sigma(T) \setminus \{0\}$. Then clearly

$$\mathcal{N}(A^{n-1}) \subset \mathcal{N}(A^n) \quad \text{for all } n \in \mathbb{N}.$$

Assume that $\mathcal{N}(A^{n-1})$ is a proper subset of $\mathcal{N}(A^n)$ for all $n \in \mathbb{N}$. Similarly as in (i) we can choose $x_n \in \mathcal{N}(A^n)$ with

$$\|x_n\| = 1 \quad \text{and} \quad \operatorname{dist}\big(x_n, \mathcal{N}(A^{n-1})\big) \geq \frac{1}{2},$$

due to Proposition E3.5. Thus we have for $m < n$

$$\|Tx_n - Tx_m\| = \Big\|\lambda x_n - (Ax_n + \lambda x_m - Ax_m)\Big\| \geq \frac{|\lambda|}{2} > 0,$$

because

$$Ax_n + \lambda x_m - Ax_m \in \mathcal{N}(A^{n-1}).$$

But $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence. This is in contradiction to the compactness of T. Consequently we find an $n \in \mathbb{N}$ such that $\mathcal{N}(A^{n-1}) = \mathcal{N}(A^n)$. Then we have for $m > n$

$$\begin{aligned} x \in \mathcal{N}(A^m) \text{ implies } & A^{m-n}x \in \mathcal{N}(A^n) = \mathcal{N}(A^{n-1}) \\ & \text{implies } A^{n-1+m-n}x = 0 \text{ implies } x \in \mathcal{N}(A^{m-1}). \end{aligned}$$

Thus $\mathcal{N}(A^m) = \mathcal{N}(A^{m-1})$. Now inductively we obtain $\mathcal{N}(A^m) = \mathcal{N}(A^n)$ for all $m \geq n$. Therefore $n_\lambda < \infty$. Since $\mathcal{N}(A) \neq \{0\}$, we have $n_\lambda \geq 1$.

(iii): Again set $A := \lambda Id - T$, $\lambda \in \sigma(T) \setminus \{0\}$. We have

$$\mathcal{N}(A^{n_\lambda}) \oplus \mathcal{R}(A^{n_\lambda}) \subset X,$$

because if

$$x \in \mathcal{N}(A^{n_\lambda}) \cap \mathcal{R}(A^{n_\lambda}),$$

then $A^{n_\lambda}x = 0$ and $x = A^{n_\lambda}y$ for some $y \in X$. Therefore, $A^{2n_\lambda}y = 0$, i.e.,

$$y \in \mathcal{N}(A^{2n_\lambda}) = \mathcal{N}(A^{n_\lambda}).$$

Hence

$$x = A^{n_\lambda} y = 0.$$

Now we can write

$$A^{n_\lambda} = \lambda^{n_\lambda} Id + \sum_{k=1}^{n_\lambda} \binom{n_\lambda}{k} \lambda^{n_\lambda - k} (-T)^k$$

and

$$\sum_{k=1}^{n_\lambda} \binom{n_\lambda}{k} \lambda^{n_\lambda - k} (-T)^k$$

is compact by Lemma 1.4. Hence $\mathcal{R}(A^{n_\lambda})$ is closed and

$$\mathcal{N}(A^{n_\lambda}) \oplus \mathcal{R}(A^{n_\lambda}) = X$$

by Proposition 1.16 together with Corollary 5.5 below.

Notice that T commutes with A, i.e., $AT = TA$, and therefore T also commutes with A^{n_λ}. Hence T leaves $\mathcal{N}(A^{n_\lambda})$ and $\mathcal{R}(A^{n_\lambda})$ invariant.

(iv): Denote by T_λ the restriction of T to $\mathcal{R}(A^{n_\lambda})$, $\lambda \in \sigma(T) \setminus \{0\}$. Then $T_\lambda \in K(\mathcal{R}(A^{n_\lambda}))$. Note that $\mathcal{R}(A^{n_\lambda})$ is a closed subspace of X by (iii), hence a Banach space. Furthermore

$$\mathcal{N}(\lambda Id - T_\lambda) = \mathcal{N}(A) \cap \mathcal{R}(A^{n_\lambda}) = \{0\},$$

because $\mathcal{N}(A) \subset \mathcal{N}(A^{n_\lambda})$. Thus

$$\mathcal{R}(\lambda Id - T_\lambda) = \mathcal{R}(A^{n_\lambda}),$$

by Proposition 1.16 applied to T_λ. Hence $\lambda \in \rho(T_\lambda)$. It remains to show that

$$\sigma(T_\lambda) = \sigma(T) \setminus \{\lambda\}.$$

Let $\mu \in \mathbb{K} \setminus \{\lambda\}$. As above we obtain that $(\mu Id - T)$ leaves $\mathcal{N}(A^{n_\lambda})$ invariant. Furthermore, $(\mu Id - T)$ is on this subspace injective, because

$$x \in \mathcal{N}(\mu Id - T) \text{ means } (\lambda - \mu)x = Ax.$$

If additionally $A^m x = 0$ for some $m \in \mathbb{N}$, then

$$(\lambda - \mu)A^{m-1}x = A^m x = 0,$$

i.e. $A^{m-1}x = 0$, because $\lambda \neq \mu$. Hence, inductively we obtain $x = 0$. This shows

$$\mathcal{N}(\mu Id - T) \cap \mathcal{N}(A^m) = \{0\}$$

for all $m \in \mathbb{N}$. For $m = n_\lambda$ this yields injectivity of $\mu Id - T$ on $\mathcal{N}(A^{n_\lambda})$. Since this space is finite dimensional, $\mu Id - T$ is also bijective on $\mathcal{N}(A^{n_\lambda})$. Hence

$$\mu \in \rho(T) \quad \text{iff} \quad \mu \in \rho(T_\lambda).$$

Therefore, by separating a finite dimensional characteristic subspace corresponding to the eigenvalue λ, we obtain a remaining operator T_λ with

$$\sigma(T_\lambda) = \sigma(T) \setminus \{\lambda\}.$$

■

1.6 Fredholm alternative and an application

Theorem 1.18 (Fredholm alternative) *If $T \in K(X)$ and $\lambda \neq 0$, then: Either the equation*

$$Tx - \lambda x = y$$

is for each $y \in X$ uniquely solvable or the equation

$$Tx - \lambda x = 0$$

has a non-trivial solution.

Proof: Follows immediately from Proposition 1.16. ■

Example 1.19 Consider the following Volterra type integral operator $T : C([0,1]) \to C([0,1])$:

$$(Tf)(x) := \int_0^x k(x,y) f(y)\, dy, \quad f \in C([0,1]),\, x \in [0,1],$$

where $k \in C([0,1]^2)$. We know that $T \in K(C([0,1]))$, see Exercise 3.1(iii). We are interested in solutions to

$$Tf - \lambda f = 0, \quad f \in C([0,1]),$$

where $\lambda \neq 0$. Such an equation is called **integral equation of second type**. (Integral equations of first type are given by $Tf = 0$ or $Tf = g$, respectively, and much more complicated to analyze.) Our aim is to show that for $\lambda \neq 0$ the operator $\lambda Id - T$ is injective. W.l.o.g. we may assume $\lambda = 1$ (otherwise consider $\frac{T}{\lambda}$). $Tf = f$, $f \in C([0,1])$, implies

$$|f(x)| = |(Tf)(x)| \leq \int_0^x |k(x,y)||f(y)|\, dy \leq x\|k\|_{\sup}\|f\|_{\sup}, \; x \in [0,1].$$

Hence

$$|f(x)| \leq \int_0^x |k(x,y)|\, y\, \|k\|_{\sup}\|f\|_{\sup}\, dy \leq \frac{x^2}{2}\|k\|_{\sup}^2\|f\|_{\sup}, \; x \in [0,1].$$

Then, inductively,

$$|f(x)| \leq \frac{x^n}{n!}\|k\|_{\sup}^n\|f\|_{\sup}, \; x \in [0,1].$$

Hence, in the limit $n \to \infty$ we obtain $f = 0$, i.e. $\lambda Id - T$ is injective. Now, by the Fredholm alternative, uniqueness implies the existence of a unique solution $f \in C([0,1])$ to the inhomogeneous equation

$$Tf - \lambda f = g$$

for all $g \in C([0,1])$.

1.7 Normal operators

In this subsection X is a Hilbert space over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with scalar product $(\cdot,\cdot)$.

Definition 1.20 *Let $T \in L(X)$. T is called* **normal**, *iff T commutes with T^*, i.e.*

$$T^*T - TT^* = 0.$$

Lemma 1.21 Let $T \in L(X)$.
(i) If T is self-adjoint, then T is also normal.
(ii) T is normal, iff

$$\|Tx\| = \|T^*x\| \quad \text{for all } x \in X.$$

(iii) If T is normal, then also $\lambda Id - T$ is normal for all $\lambda \in \mathbb{K}$.
(iv) If T is normal and $\lambda \in \mathbb{K}$, then

$$\mathcal{N}(\lambda Id - T) = \mathcal{N}(\overline{\lambda} Id - T^*).$$

Proof: (i): Obvious!
(ii): Let T be normal. Then

$$(Tx, Tx) = (x, T^*Tx) = (x, TT^*x) = (T^*x, T^*x) \quad \text{for all } x \in X.$$

Vice versa: By the **polarization identity**

$$\frac{1}{4}(\|u+v\|^2 - \|u-v\|^2) = \Re(u,v), \quad u, v \in X,$$

follows

$$\Re(Tx, Ty) = \Re(T^*x, T^*y) \quad \text{for all } x, y \in X.$$

In the case $\mathbb{K} = \mathbb{C}$ we replace y by iy and obtain

$$0 = (Tx, Ty) - (T^*x, T^*y) = (T^*Tx - TT^*x, y) \quad \text{for all } x, y \in X.$$

Thus $T^*T - TT^*$.
(iii): Obvious!
(iv): Follows immediately from (ii) together with (iii). ■

Lemma 1.22 Let $T \in L(X)$ be normal and $\mathbb{K} = \mathbb{C}$. If $X \neq \{0\}$, then

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \|T\|_{L(X)}.$$

Proof: We already now from Proposition 1.11 that

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{m \to \infty} \|T^m\|_{L(X)}^{\frac{1}{m}} \leq \|T\|_{L(X)}.$$

Hence it is sufficient to show that

$$\|T^m\| \geq \|T\|^m \quad \text{for all } m \in \mathbb{N}_0. \tag{1.4}$$

Let $T \neq 0$ (for $T = 0$ the statement is obvious). For $m = 0, 1$ the inequality in (1.4) is trivial. Let $m \geq 1$ and $x \in X$, then by Lemma 1.21(ii)

$$\begin{aligned}\|T^m x\|^2 = (T^*T^m x, T^{m-1}x) &\leq \|T^*T^m x\| \, \|T^{m-1}x\| \\ &\leq \|T^{m+1}x\| \, \|T^{m-1}x\| \leq \|T^{m+1}\| \, \|T\|^{m-1} \, \|x\|^2.\end{aligned}$$

Thus

$$\|T^m\|^2 \leq \|T^{m+1}\| \, \|T\|^{m-1}.$$

Therefore, if $\|T^m\| \geq \|T\|^m$, then

$$\|T^{m+1}\| \geq \frac{\|T^m\|^2}{\|T\|^{m-1}} \geq \|T\|^{2m-(m-1)} = \|T\|^{m+1}.$$

■

Example 1.23 Let $(e_k)_{k \in N}$, $N \subset \mathbb{N}$, be an orthonormal system in X and $\lambda_k \in \mathbb{K}$ such that $|\lambda_k| \leq r < \infty$, $k \in N$. Then

$$Tx := \sum_{k \in N} \lambda_k (x, e_k) e_k, \quad x \in X,$$

defines an operator $T \in L(X)$, see Exercise 1.3. Since

$$(Tx, y) = \sum_{k \in N} \lambda_k (x, e_k)(e_k, y) = \Big(x, \sum_{k \in N} \overline{\lambda_k (e_k, y)} e_k \Big),$$

is

$$T^* x = \sum_{k \in N} \overline{\lambda_k} (x, e_k) e_k, \quad x \in X.$$

Therefore

$$T^*Tx = TT^*x = \sum_{k \in N} |\lambda_k|^2 (x, e_k) e_k, \quad x \in X,$$

i.e., T is normal. If $\#N < \infty$, then is T finite rank and, in particular, $T \in K(X)$. If $N = \mathbb{N}$, then

$$T \in K(X) \quad \text{iff} \quad \lim_{n \to \infty} \lambda_n = 0,$$

see Exercise 1.3(ii).

1.8 Spectral theorem for normal operators

In this subsection X is a Hilbert space over the field $\mathbb{C}$.

Theorem 1.24 *Let $T \in K(X)$ be normal, $T \neq 0$. Then there exists an orthonormal system $\{e_k \,|\, k \in M\}$, $M \subset \mathbb{N}$, and $0 \neq \mu_k \in \mathbb{C}$, $k \in M$, such that:*
(i)

$$Te_k = \mu_k e_k, \quad k \in M, \quad \sigma(T) \setminus \{0\} = \{\mu_k \,|\, k \in M\},$$

i.e. the numbers μ_k are the eigenvalues of T different from zero with eigenvectors e_k, $k \in M$. (In this notation the eigenvalues μ_k for different k might be the same.) If $M = \mathbb{N}$, then $\lim_{k\to\infty} \mu_k = 0$.
(ii) *For the orders we have: $n_{\mu_k} = 1$ for all $k \in M$.*
(iii)

$$X = \mathcal{N}(T) \perp \overline{\mathrm{span}\{e_k \,|\, k \in M\}}.$$

(iv)

$$Tx = \sum_{k\in M} \mu_k (x, e_k) e_k \quad \textit{for all } x \in X.$$

Remark 1.25 If we write

$$X = Y \perp Z$$

for $Y, Z \subset X$ closed, subspaces, then this means

$$X = Y \oplus Z \quad \text{and} \quad (y, z) = 0 \quad \text{for all} \quad y \in Y,\, z \in Z.$$

In the proof we will also use the notation

$$X \supset (\perp_{n\in N} X_n) := \mathrm{span}\{x_1 \in X_1, x_2 \in X_2, x_3 \in X_3, \ldots\}$$

for $X_n \subset X$, $n \in N \subset \mathbb{N}$, closed, subspaces, pairwise orthogonal.

Proof: From Theorem 1.17 we know that $\sigma(T) \setminus \{0\}$ consists of eigenvalues λ_k, $k \in N \subset \mathbb{N}$, only. Furthermore, if N has infinitely many elements, then $\lim_{k\to\infty} \lambda_k = 0$. Here we choose the λ_k pairwise different for different $k \in N$. We also know from Theorem 1.17, that

$$N_k := \mathcal{N}(\lambda_k Id - T)$$

is finite dimensional for all $k \in N$. Set $N_0 := \mathcal{N}(T)$ and $\lambda_0 := 0$. Lemma 1.21(iv) implies

$$N_k = \mathcal{N}(\overline{\lambda_k} Id - T^*), \quad k \in N \cup \{0\}. \tag{1.5}$$

Observe that

$$N_k \perp N_l \quad \text{for } k, l \in N \cup \{0\},\ k \neq l,$$

because if $x_k \in N_k$ and $x_l \in N_l$, then

$$\lambda_k(x_k, x_l) = (Tx_k, x_l) = (x_k, T^* x_l) = (x_k, \overline{\lambda_l} x_l) = \lambda_l(x_k, x_l).$$

Since $\lambda_k \neq \lambda_l$ it follows that $(x_k, x_l) = 0$.

We claim that

$$X = \overline{\perp_{k\in N\cup\{0\}} N_k}. \tag{1.6}$$

In order to show this we choose

$$y \in Y = \big(\perp_{k\in N\cup\{0\}} N_k\big)^{\perp}.$$

Using (1.5), we can conclude for $x \in N_k$, $k \in N \cup \{0\}$,

$$(Ty, x) = (y, T^* x) = (y, \overline{\lambda_k} x) = \lambda_k(y, x) = 0.$$

Thus $Ty \in Y$, i.e. Y is T-invariant. Now consider

$$T_0 := T|_Y.$$

Then $T \in K(Y)$ and normal. If $Y \neq \{0\}$, then by Lemma 1.22 there exists $\lambda \in \sigma(T_0)$ with $|\lambda| = \|T_0\|$. If $T_0 \neq 0$, then λ would be an eigenvalue of T_0 (by Theorem 1.17) and therefore also of T. I.e. $N_k \cap Y \neq \{0\}$ for some $k \in N$. That is in contradiction with the definition of Y. Hence $T_0 = 0$, i.e. $Y \subset N_0$. But that is also in contradiction with the definition of Y. Thus $Y = \{0\}$, i.e. (1.6) is true.

Denote by P_0 the orthogonal projection on $\mathcal{N}(T)$. Since for all $x \in X$ we have $x = (Id - P_0)x + P_0 x$, from (1.6) we can infer

$$X = \mathcal{N}(T) \perp \overline{\perp_{k\in N} N_k}. \tag{1.7}$$

Now choose for each $k \in N$ an orthonormal basis $\{b_{k1}, \ldots, b_{kd_k}\}$ of N_k. Then by Proposition E5.8

$$\{b_{ki_k} \,|\, 1 \leq i_k \leq d_k, k \in N\} \tag{1.8}$$

is an orthonormal basis of $\overline{\perp_{k\in N} N_k}$ and together with (1.7) we obtain

$$X = \mathcal{N}(T) \perp \overline{\operatorname{span}\{b_{ki_k} \,|\, 1 \leq i_k \leq d_k, k \in N\}}. \tag{1.9}$$

Furthermore we can conclude that

$$x = \sum_{k\in N} \sum_{i=1}^{d_k} (x, b_{ki}) b_{ki} + P_0(x) \quad \text{for all } x \in X. \tag{1.10}$$

Applying T to (1.10) we obtain

$$Tx = \sum_{k\in N} \sum_{i=1}^{d_k} (x, b_{ki}) T b_{ki} + T P_0(x) = \sum_{k\in N} \sum_{i=1}^{d_k} \lambda_k (x, b_{ki}) b_{ki}, \quad x \in X. \tag{1.11}$$

Changing the notation of the orthonormal system in (1.8) into $\{e_k \,|\, k \in M\}$, $M \subset \mathbb{N}$, and adapting appropriately the notation for the eigenvalues, (i) follows by the above considerations. Furthermore, (iii) then is a equivalent to (1.9) and (iv) to (1.11).

(ii): Let $x \in \mathcal{N}((\mu_j Id - T)^2)$, $j \in M$. Then

$$(\mu_j Id - T)x \in \mathcal{N}(\mu_j Id - T) = \mathcal{N}(\overline{\mu_j} Id - T^*).$$

Therefore

$$\begin{aligned} 0 = (x, (\overline{\mu_j} Id - T^*)(\mu_j Id - T)x) \\ = ((\mu_j Id - T)x, (\mu_j Id - T)x) = \|(\mu_j Id - T)x\|^2. \end{aligned}$$

Thus $x \in \mathcal{N}(\mu_j Id - T)$, i.e. $n_{\mu_j} = 1$ for all $j \in M$. ■

Corollary 1.26 *Let $T \in L(X)$ be self-adjoint, i.e. $T^* = T$.*
(i) $\sigma_p(T) \subset [-\|T\|, \|T\|] \subset \mathbb{R}$. If additionally $T \in K(X)$, then $\|T\|$ or $-\|T\|$ is an eigenvalue.
(ii) If T is **positive semi-definite**, *i.e. $(Tx, x) \geq 0$ for all $x \in X$, then $\sigma_p(T) \subset [0, \|T\|]$. If additionally $T \in K(X)$, then $\|T\|$ is an eigenvalue.*

Proof: (i): Let x be an eigenvector with corresponding eigenvalue λ. Then

$$\lambda\|x\|^2 = (\lambda x, x) = (Tx, x) = (x, Tx) = (x, \lambda x) = \overline{\lambda}\|x\|^2.$$

Hence $\lambda = \overline{\lambda}$, because $x \neq 0$. Since

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \|T\| \tag{1.12}$$

(see Lemma 1.22) the first statement is shown. Then for $T \in K(X)$ (1.12) together with Theorem 1.17 implies, that $\|T\|$ or $-\|T\|$ is an eigenvalue.

(ii): Let x be an eigenvector with corresponding eigenvalue λ. Then

$$\lambda\|x\|^2 = (Tx, x) \geq 0.$$

Hence $\lambda \geq 0$, because $x \neq 0$. If $T \in K(X)$ from (i) we already know that $\|T\|$ or $-\|T\|$ is an eigenvalue. Thus $\|T\|$ is an eigenvalue. ■

2 Hahn–Banach theorem

2.1 Extension of linear functionals on spaces with sub-linear mappings

Theorem 2.1 *Let X be an $\mathbb{R}$-vector space and:*

(i) *$p : X \to \mathbb{R}$ is sub-linear, i.e., for all $x, y \in X$ and $\alpha \geq 0$ we have:*

$$p(x+y) \leq p(x) + p(y) \quad \text{and} \quad p(\alpha x) = \alpha p(x).$$

(ii) *$f : Y \to \mathbb{R}$ is linear, Y a subspace of X.*

(iii) *$f(x) \leq p(x)$ for all $x \in Y$.*

Then there exists a linear mapping $F : X \to \mathbb{R}$ such that

$$F(x) = f(x) \quad \text{for} \quad x \in Y \quad \text{and} \quad F(x) \leq p(x) \quad \text{for} \quad x \in X.$$

Proof: We consider the class of all extensions of f:

$$\mathcal{M} := \{(Z,g) \,|\, Z \text{ subspace}, \, Y \subset Z \subset X, \\ g: Z \to \mathbb{R} \text{ linear}, \, g = f \text{ on } Y, \, g(x) \le p(x) \text{ on } Z\}.$$

$\mathcal{M} \neq \emptyset$, because $(Y,f) \in \mathcal{M}$. Now consider an arbitrary $(Z,g) \in \mathcal{M}$ with $Z \neq X$ and $z_0 \in X \setminus Z$. At least, we want to extend g to

$$Z_0 := \operatorname{span}\{Z \cup \{z_0\}\} = Z \oplus \operatorname{span}\{z_0\}.$$

We try the ansatz

$$g_0(z + \alpha z_0) := g(z) + c\alpha, \quad z \in Z, \alpha \in \mathbb{R},$$

where we have to chose $c \in \mathbb{R}$ appropriately. Clearly, g_0 is linear on Z_0. Furthermore, $g_0 = g = f$ on Y. It remains to show that

$$g(z) + c\alpha \le p(z + \alpha z_0), \quad z \in Z, \alpha \in \mathbb{R}.$$

Since $g \le p$ on Z, it is fulfilled for $\alpha = 0$. For $\alpha > 0$ the inequality implies

$$c \le \frac{p(z + \alpha z_0) - g(z)}{\alpha} = p\left(\frac{z}{\alpha} + z_0\right) - g\left(\frac{z}{\alpha}\right),$$

and for $\alpha < 0$

$$c \ge \frac{p(z + \alpha z_0) - g(z)}{\alpha} = g\left(-\frac{z}{\alpha}\right) - p\left(-\frac{z}{\alpha} - z_0\right).$$

Hence, c has to fulfill

$$\sup_{z \in Z}(g(z) - p(z - z_0)) \le c \le \inf_{z \in Z}(p(z + z_0) - g(z)).$$

This is possible, because for $z, z' \in Z$ we have:

$$g(z') + g(z) = g(z' + z) \le p(z' + z) \\ = p(z' - z_0 + z + z_0) \le p(z' - z_0) + p(z + z_0),$$

and therefore

$$g(z') - p(z' - z_0) \le p(z + z_0) - g(z).$$

Our aim is to find via this extension procedure $(X, F) \in \mathcal{M}$. For this we use:

Lemma 2.2 (Zorn's lemma) *Let $(\mathcal{M}, \leq)$ be a non-empty partially ordered set such that each totally ordered subset $\mathcal{N}$ (i.e., $n_1, n_2 \in \mathcal{N}$ implies $n_1 \leq n_2$ or $n_2 \leq n_1$) possesses an upper bound (i.e., there exists $m \in \mathcal{M}$ such that $n \leq m$ for all $n \in \mathcal{N}$). Then $\mathcal{M}$ possesses a maximal element (i.e., there exists $m_0 \in \mathcal{M}$ such that for all $m \in \mathcal{M}$: $m_0 \leq m$ implies $m \leq m_0$).*

In our situation an order is defined by

$$(Z_1, g_1) \leq (Z_2, g_2) \quad \text{iff} \quad Z_1 \subset Z_2 \text{ and } g_2 = g_1 \text{ on } Z_1.$$

We have to verify the assumptions of Zorn's lemma. Let $\mathcal{N} \subset \mathcal{M}$ be totally ordered and define

$$Z_* := \bigcup_{(Z,g)\in\mathcal{N}} Z,$$
$$g_*(x) := g(x), \quad \text{if } x \in Z \quad \text{and} \quad (Z, g) \in \mathcal{N}.$$

It is to show that $(Z_*, g_*) \in \mathcal{M}$. We have $Y \subset Z_* \subset X$ and g_* is well defined. Indeed, if

$$x \in Z_1 \cap Z_2, \quad (Z_1, g_1), (Z_2, g_2) \in \mathcal{N},$$

then

$$(Z_1, g_1) \leq (Z_2, g_2) \quad \text{or} \quad (Z_2, g_2) \leq (Z_1, g_1) \quad (\mathcal{N} \text{ is totally ordered}).$$

W.l.o.g. we assume the first case (the second case we can treat analogously). Then

$$Z_1 \subset Z_2 \text{ and } g_2 = g_1 \text{ on } Z_1$$

and therefore

$$g_2(x) = g_1(x) \quad (\text{since } x \in Z_1).$$

The properties $g_* = f$ on Y and $g_* \leq p$ are inhered by construction.

Linearity of Z_* and g_*: Let $x, y \in Z_*$, then there exist $(Z_x, g_x), (Z_y, g_y) \in \mathcal{N}$ such that $x \in Z_x$ and $y \in Z_y$. Again we have

$$(Z_x, g_x) \leq (Z_y, g_y) \quad \text{or} \quad (Z_y, g_y) \leq (Z_x, g_x).$$

W.l.o.g. we assume $x, y \in Z_y$. Then

$$\alpha x + \beta y \in Z_y \subset Z_*, \quad \alpha, \beta \in \mathbb{R}.$$

Furthermore,

$$g_*(\alpha x + \beta y) = g_y(\alpha x + \beta y) = \alpha g_y(x) + \beta g_y(y) = \alpha g_*(x) + \beta g_*(y).$$

Now by Zorn's lemma $\mathcal{M}$ has an maximal element (Z, g). Suppose that $Z \neq X$, then the extension procedure from the beginning of the proof gives $(Z_0, g_0) \in \mathcal{M}$ such that

$$(Z, g) \leq (Z_0, g_0) \quad \text{and} \quad Z_0 \neq Z.$$

But (Z, g) is maximal. That's a contradiction! ■

2.2 Extension of continuous linear functionals

Theorem 2.3 *Let Y be a subspace of a normed $\mathbb{K}$-vector space X (where Y is equipped with the norm of X!). Then for each $y' \in Y'$ there exists an $x' \in X'$ such that*

$$x' = y' \text{ on } Y \quad \text{and} \quad \|x'\|_{X'} = \|y'\|_{Y'}.$$

Proof: First let $\mathbb{K} = \mathbb{R}$. Set

$$p(x) := \|y'\|_{Y'} \|x\|_X, \quad x \in X.$$

Then for $y \in Y$

$$y'(y) \leq \|y'\|_{Y'} \|y\|_Y = \|y'\|_{Y'} \|y\|_X = p(y).$$

Thus, the assumptions of Theorem 2.1 are fulfilled and we get a linear mapping $x' : X \to \mathbb{R}$ such that

$$x' = y' \quad \text{on } Y \quad \text{and} \quad x' \leq p \quad \text{on } X. \tag{2.1}$$

The second property in (2.1) implies

$$\pm x'(x) = x'(\pm x) \leq p(\pm x) = \|y'\|_{Y'} \|x\|_X,$$

i.e., $x' \in X'$ and $\|x'\|_{X'} \le \|y'\|_{Y'}$. The first property in (2.1) implies

$$\|y'\|_{Y'} = \sup_{y \in Y, \|y\|_X \le 1} |y'(y)| = \sup_{y \in Y, \|y\|_X \le 1} |x'(y)| \le \|x'\|_{X'}.$$

Next we consider the case $\mathbb{K} = \mathbb{C}$. View X and Y as $\mathbb{R}$-vector spaces $X_\mathbb{R}$ and $Y_\mathbb{R}$. Then

$$y'_{re} := \Re y' \in Y'_\mathbb{R} \quad \text{with} \quad \|y'_{re}\|_{Y'_\mathbb{R}} \le \|y'\|_{Y'},$$

and

$$y'(x) = \Re y'(x) + i\Im y'(x) = y'_{re}(x) - iy'_{re}(ix), \quad x \in Y.$$

Let x'_{re} be a extension of y'_{re} to $X_\mathbb{R}$ with $\|x'_{re}\|_{X'_\mathbb{R}} = \|y'_{re}\|_{Y'_\mathbb{R}}$ constructed as in the real case. Then define

$$x'(x) := x'_{re}(x) - ix'_{re}(ix), \quad x \in X.$$

Then $x' = y'$ on Y and $x' : X \to \mathbb{C}$ is $\mathbb{C}$-linear, because x' is $\mathbb{R}$-linear and

$$x'(ix) = x'_{re}(ix) - ix'_{re}(-x) = i(-ix'_{re}(ix) - x'_{re}(-x)) = ix'(x), \quad x \in X.$$

Now let $x \in X$ and $x'(x) = re^{i\theta}, r \ge 0, \theta \in [0, 2\pi)$. Then

$$|x'(x)| = r = \Re(e^{-i\theta}x'(x)) = \Re(x'(e^{-i\theta}x)) = x'_{re}(e^{-i\theta}x) \le \|x'_{re}\|_{X'_\mathbb{R}}\|x\|$$

and

$$\|x'_{re}\|_{X'_\mathbb{R}} = \|y'_{re}\|_{Y'_\mathbb{R}} \le \|y'\|_{Y'}.$$

Hence, $x' \in X'$ and $\|x'\|_{X'} \le \|y'\|_{Y'}$. On the other hand we have $\|x'\|_{X'} \ge \|y'\|_{Y'}$, because x' is a extension of y'. ■

2.3 Applications

Proposition 2.4 *Let Y be a closed, subspace of a normed vector space X and $x_0 \in X \setminus Y$. Then there exists an $x' \in X'$ such that*

$$x' = 0 \text{ on } Y, \ \|x'\| = 1 \quad \text{and} \quad x'(x_0) = \operatorname{dist}(x_0, Y).$$

Proof: On

$$Y_0 := \mathrm{span}(Y \cup \{x_0\}) = Y \oplus \mathrm{span}\{x_0\}$$

define

$$y_0'(y + \alpha x_0) := \alpha \mathrm{dist}(x_0, Y), \quad y \in Y,\ \alpha \in \mathbb{K}.$$

Then

$$y_0' : Y_0 \to \mathbb{K}$$

is linear and $y_0' = 0$ on Y. Now by Theorem 2.3 it is sufficient to show that $y_0' \in Y_0'$ and $\|y_0'\| = 1$. Since for $y \in Y$ and $\alpha \neq 0$

$$\mathrm{dist}(x_0, Y) \leq \left\| x_0 - \frac{-y}{\alpha} \right\|,$$

we have

$$|y_0'(y + \alpha x_0)| \leq |\alpha| \left\| x_0 - \frac{-y}{\alpha} \right\| = \|\alpha x_0 + y\|,$$

i.e. $y_0' \in Y_0'$ and $\|y_0'\| \leq 1$. Because Y is closed, we have $\mathrm{dist}(x_0, Y) > 0$. Hence for each $\varepsilon > 0$ the exist $y_\varepsilon \in Y$ such that

$$\|x_0 - y_\varepsilon\| \leq (1 + \varepsilon)\mathrm{dist}(x_0, Y).$$

Hence

$$y_0'(x_0 - y_\varepsilon) = \mathrm{dist}(x_0, Y) \geq \frac{1}{1+\varepsilon}\|x_0 - y_\varepsilon\|.$$

Because $x_0 - y_\varepsilon \neq 0$ this yields

$$\|y_0'\| \geq \frac{1}{1+\varepsilon} \quad \text{for all } \varepsilon > 0.$$

Thus $\|y_0'\| = 1$. ■

Corollary 2.5 Let X be a normed space and $x_0 \in X$.

(i) If $x_0 \neq 0$, then there exist $x_0' \in X'$ such that

$$\|x_0'\| = 1 \quad \text{and} \quad x_0'(x_0) = \|x_0\|.$$

(ii) If $x'(x_0) = 0$ for all $x' \in X'$, then $x_0 = 0$.

(iii) Let $x_1, \ldots, x_n \in X$ be linear independent. Then there exist $x'_1, \ldots, x'_n \in X'$ such that $x'_k(x_l) = \delta_{k,l}$, $1 \le k, l \le n$.

(iv) By $Tx' := x'(x_0), x' \in X'$, a linear functional $T \in L(X'; \mathbb{K}) = (X')'$ is defined with $\|T\|_{(X')'} = \|x_0\|$.

Proof: (i): Follows from Proposition 2.4 when setting $Y = \{0\}$.

(ii): Follows from (i).

(iii): To construct $x'_k \in X'$ apply Proposition 2.4 to

$$Y_k = \operatorname{span}\{x_l \,|\, l \neq k, 1 \le l \le n\}, \quad 1 \le k \le n,$$

and then normalize the obtained linear functional.

(iv): We have $|T(x')| \le \|x'\|_{X'} \|x_0\|$. If $x_0 \neq 0$, then $|Tx'_0| = \|x_0\|$ where x'_0 as in (i). Thus, $\|T\|_{(X')'} = \|x_0\|$. ■

Remark 2.6 Proposition 2.4 can be considered as a generalization of the projection theorem for Hilbert spaces, see Corollary E5.14. Because, if X is a Hilbert space we can define

$$x'(x) := \left(x, \frac{x_0 - Px_0}{\|x_0 - Px_0\|} \right), \quad x \in X,$$

where P is the orthogonal projection on Y. By construction $x' = 0$ on Y and therefore

$$x'(x_0) = x'(x_0 - Px_0) = \|x_0 - Px_0\| = \operatorname{dist}(x_0, Y).$$

Additionally, by Cauchy–Schwartz

$$|x'(x)| \le \|x\|$$

and

$$x'(x_0 - Px_0) = \left(x_0 - Px_0, \frac{x_0 - Px_0}{\|x_0 - Px_0\|} \right) = \|x_0 - Px_0\| \neq 0.$$

Thus, x' has the properties as in Proposition 2.4.

3 Uniform boundedness principle

3.1 Baire category theorem

Theorem 3.1 *Let (X, d) be a non-empty complete metric space and*

$$X = \bigcup_{k\in\mathbb{N}} A_k, \quad A_k \text{ closed}, \; k \in \mathbb{N}.$$

Then there exists a $k_0 \in \mathbb{N}$ such that $\mathring{A}_{k_0} \neq \emptyset$.

Proof: Assume that $\mathring{A}_k = \emptyset$ for all $k \in \mathbb{N}$. Then we have:

$$U \subset X \text{ open, not empty}, k \in \mathbb{N} \;\text{ implies }\; U \setminus A_k \text{ open, not empty}$$
$$\text{implies } \text{ there exists a ball } \overline{U_\epsilon(x)} \subset U \setminus A_k \text{ with } \epsilon \leq \frac{1}{k}.$$

Hence, we can choose inductively a sequence of balls $U_\epsilon(x_k)$ such that

$$\overline{U_{\epsilon_k}(x_k)} \subset U_{\epsilon_{k-1}}(x_{k-1}) \setminus A_k \text{ and } \epsilon_k \leq \frac{1}{k}, \quad k \geq 2,$$

with $U_{\epsilon_1}(x_1) \subset X \setminus A_1$. Then $x_l \in U_{\epsilon_k}(x_k)$ for all $l \geq k$ and $\lim_{k\to\infty} \epsilon_k = 0$. Thus, $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and there exists

$$x := \lim_{k\to\infty} x_k \in X.$$

Note that $x \in \overline{U_{\epsilon_k}(x_k)}$ for all $k \in \mathbb{N}$. Since $\overline{U_{\epsilon_k}(x_k)} \cap A_k = \emptyset$ we have

$$x \notin \bigcup_{k\in\mathbb{N}} A_k = X.$$

That is a contradiction. ■

3.2 Uniform boundedness principle

Theorem 3.2 *Let (X, d) be a non-empty complete metric space and Y a normed space. Consider a set of functions $\mathcal{F} \subset C^0(X; Y)$ such that*

$$\sup_{f\in\mathcal{F}} \|f(x)\| < \infty \quad \textit{for all } x \in X.$$

Then there exists $x_0 \in X$ and $\epsilon_0 > 0$ such that

$$\sup_{x\in\overline{U_{\epsilon_0}(x_0)}} \sup_{f\in\mathcal{F}} \|f(x)\| < \infty.$$

Proof: Set

$$A_k := \bigcap_{f\in\mathcal{F}} \{x \in X \mid \|f(x)\| \le k\}.$$

Then the A_k fulfill the assumptions of Theorem 3.1. Thus, there exists a k_0 such that $\mathring{A}_{k_0} \neq \emptyset$. In particular,

$$\sup_{x\in A_{k_0}} \sup_{f\in\mathcal{F}} \|f(x)\| \le k_0.$$

Now choose a ball $\overline{U_{\epsilon_0}(x_0)} \subset A_{k_0}$. ■

3.3 Banach–Steinhaus theorem

Theorem 3.3 (Banach–Steinhaus theorem) *Let X be a Banach space, Y a normed space and $\mathcal{T} \subset L(X;Y)$ such that*

$$\sup_{T\in\mathcal{T}} \|Tx\| < \infty \quad \textit{for all } x \in X.$$

Then

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

i.e., $\mathcal{T}$ is bounded in $L(X;Y)$.

Proof: Since $\mathcal{T} \subset L(X;Y) \subset C^0(X;Y)$ and $\mathcal{T}$ has the properties as in Theorem 3.2, there exists $x_0 \in X$, $\epsilon_0 > 0$ and a constant $C < \infty$ such that

$$\|Tx\| \le C \quad \text{for all } T \in \mathcal{T}, \|x - x_0\| \le \epsilon_0.$$

Then for all $T \in \mathcal{T}$ and $x \neq 0$

$$\|Tx\| = \frac{\|x\|}{\epsilon_0} \left\| T\left(x_0 + \epsilon_0 \frac{x}{\|x\|}\right) - T(x_0) \right\| \le \frac{\|x\|}{\epsilon_0} 2C,$$

i.e., $\|T\| \le \frac{2C}{\epsilon_0}$. ■

Notation 3.4 For $x \in X$ and $x' \in X'$ we write

$$\langle x, x' \rangle := x'(x)$$

and call it **duality product**. Because, if X is a Hilbert space, then the Riesz isomorphism J yields

$$\langle x, Jy \rangle = (x, y)_X.$$

Theorem 3.5 *Let X be a Banach space, Y a normed space and $\mathcal{T} \subset L(X;Y)$ such that for all $x \in X$ and $y' \in Y'$*

$$\sup_{T \in \mathcal{T}} |\langle Tx, y' \rangle| < \infty.$$

Then $\mathcal{T}$ is bounded in $L(X;Y)$.

Proof: For $x \in X$ and $T \in \mathcal{T}$

$$S_{x,T}(y') := \langle Tx, y' \rangle$$

defines an element of $(Y')'$ with $\|S_{x,T}\| = \|Tx\|_Y$, see Corollary 2.5(iii). Since for all $x \in X$

$$\sup_{T \in \mathcal{T}} |S_{x,T}(y')| < \infty, \quad \text{for all } y' \in Y',$$

and Y' is a Banach space, see Proposition E4.3(ii), Theorem 3.3 yields

$$\sup_{T \in \mathcal{T}} \|Tx\|_Y = \sup_{T \in \mathcal{T}} \|S_{x,T}\| < \infty, \quad \text{for all } x \in X.$$

Now the statement follows from the Banach–Steinhaus theorem. ∎

3.4 Open mapping theorem

Definition 3.6 Let X and Y be topological spaces. Then $f : X \to Y$ is open, iff

$$U \text{ open in } X \text{ implies } f(U) \text{ open in } Y.$$

Remark 3.7 (i) If f is bijective, then f is open if and only if f^{-1} is continuous.
(ii) If X, Y are normed spaces and $T : X \to Y$ is linear, then:

$$T \text{ is open iff there exists } \delta > 0 \text{ such that } U_\delta(0) \subset T(U_1(0)).$$

Proof: (i): clear!

(ii) Sufficiency: also clear!

Necessity: Let U be open and $x \in U$. Choose $\epsilon > 0$ such that $U_\epsilon(x) \subset U$. Since $U_\delta(0) \subset T(U_1(0))$ for some $\delta > 0$, we find $U_{\epsilon\delta}(Tx) \subset T(U_\epsilon(x)) \subset T(U)$. Thus, $T(U)$ is open. ∎

Theorem 3.8 *Let X and Y be a Banach spaces and $T \in L(X;Y)$. Then:*

$$T \text{ is surjective iff } T \text{ is open.}$$

Proof: Necessity: Since $U_\delta(0) \subset T(U_1(0))$ for some $\delta > 0$, see Remark 3.7(ii), we find $U_r(0) \subset T(U_{\frac{r}{\delta}}(0))$ for all $r > 0$.

Sufficiency: Since T is surjective we have

$$Y = \bigcup_{k \in \mathbb{N}} \overline{T(U_k(0))}.$$

By Baire category theorem there exists k_0 and a ball $U_{\epsilon_0}(y_0)$ in Y such that

$$U_{\epsilon_0}(y_0) \subset \overline{T(U_{k_0}(0))}.$$

This implies that for each $y \in U_{\epsilon_0}(0)$ there exists a sequence $(x_i)_{i\in\mathbb{N}}$ in $U_{k_0}(0)$ such that $\lim_{i\to\infty} Tx_i = y_0 + y$. If we choose $x_0 \in X$ with $Tx_0 = y_0$ this gives

$$\lim_{i\to\infty} T\left(\frac{x_i - x_0}{k_0 + \|x_0\|}\right) = \frac{y}{k_0 + \|x_0\|} \quad \text{and} \quad \left\|\frac{x_i - x_0}{k_0 + \|x_0\|}\right\| < 1 \quad \text{for all } i \in \mathbb{N}.$$

This yields

$$U_\delta(0) \subset \overline{T(U_1(0))} \tag{3.1}$$

for $\delta := \frac{\epsilon_0}{k_0 + \|x_0\|} > 0$.

We would like, however, to have such an inclusion without taking the closure. Note that (3.1) implies

$$\begin{aligned} y \in U_\delta(0) \text{ implies } & \text{ there exists } x \in U_1(0) \text{ such that } y - Tx \in U_{\frac{\delta}{2}}(0) \\ \text{implies } & \ 2(y - Tx) \in U_\delta(0). \end{aligned}$$

Hence we can choose inductively points $y_k \in U_\delta(0)$ and $x_k \in U_1(0)$ such that

$$y_0 = y \quad \text{and} \quad y_{k+1} = 2(y_k - Tx_k).$$

Then

$$2^{-(k+1)} y_{k+1} = 2^{-k} y_k - T(2^{-k} x_k),$$

and therefore

$$\lim_{m\to\infty} T\left(\sum_{k=0}^{m} 2^{-k} x_k\right) = y - \lim_{m\to\infty} 2^{-(m+1)} y_{m+1} = y.$$

Since

$$\sum_{k=0}^{m} \|2^{-k} x_k\| < \sum_{k=0}^{m} 2^{-k} \leq 2 < \infty \quad \text{is} \quad \left(\sum_{k=0}^{m} 2^{-k} x_k\right)_{m\in\mathbb{N}}$$

a Cauchy sequence in X. Because X is complete, there exists

$$x := \sum_{k=0}^{\infty} 2^{-k} x_k \in X \quad \text{with} \quad \|x\| < 2.$$

Then continuity of T implies

$$Tx = \lim_{m\to\infty} T\left(\sum_{k=0}^{m} 2^{-k} x_k\right) = y.$$

Hence we have shown that $U_\delta(0) \subset T(U_2(0))$, or $U_{\frac{\delta}{2}}(0) \subset T(U_1(0))$. Now by Remark 3.7(ii) we can conclude that T is open. ■

3.5 Inverse mapping theorem

Theorem 3.9 *Let X and Y be a Banach spaces and $T \in L(X;Y)$. Then*

$$T \text{ is bijective implies } T^{-1} \in L(Y;X).$$

Proof: T^{-1} is linear. By Theorem 3.8 T is open, hence T^{-1} is continuous, see Remark 3.7(i). ■

3.6 Closed graph theorem

Theorem 3.10 *Let X and Y be a Banach spaces and $T : X \to Y$ linear. Then*

$$\text{graph}(T) := \{(x, Tx) \in X \times Y \mid x \in X\}$$

is closed in $X \times Y$ iff $T \in L(X;Y)$.

Proof: Sufficiency: In the formulation of the theorem we view $X \times Y$ as a Banach space, e.g., equipped with the norm $\|(x,y)\| := \|x\|_X + \|y\|_Y$. As a closed subspace $Z := \text{graph}(T)$ is a Banach space. Set

$$P_X(x,y) := x, \ P_Y(x,y) := y, \ \text{for } (x,y) \in Z.$$

P_X and P_Y are linear and continuous and $P_X : Z \to X$ is bijective. By the inverse mapping theorem $P_X^{-1} \in L(X;Z)$, therefore $T = P_Y P_X^{-1} \in L(X;Y)$.

Necessity: Follows directly from continuity of T. ■

4 Weak convergence

In this section we assume X to be a Banach space and use the notation $\langle x, x' \rangle := x'(x)$, $x \in X$, $x' \in X'$, as fixed in Notation 3.4.

4.1 Definition, elementary properties and examples

Definition 4.1 *(i) A sequence $(x_k)_{k\in\mathbb{N}}$ in X* **converges weakly** *to $x \in X$ ($x_k \to x$ weakly in X as $k \to \infty$, or $x_k \rightharpoonup x$ as $k \to \infty$), iff*

$$\lim_{k\to\infty} \langle x_k, x' \rangle = \langle x, x' \rangle \quad \textit{for all} \quad x' \in X'.$$

(ii) A sequence $(x'_k)_{k\in\mathbb{N}}$ in X' **converges weakly*** *to $x' \in X'$ ($x'_k \to x'$ weakly* in X' as $k \to \infty$, or $x'_k \rightharpoonup^* x'$ as $k \to \infty$), if*

$$\lim_{k\to\infty} \langle x, x'_k \rangle = \langle x, x' \rangle \quad \textit{for all} \quad x \in X.$$

(iii) Weak and weak* Cauchy sequences are defined correspondingly.
(iv) A subset $M \subset X$ (X', resp.) is called **weak** *(***weak****, resp.)* **sequentially compact***, if each sequence in M possess a weak (weak*, resp.) convergent subsequence, whose weak (weak*, resp.) limit is also in M.*
(v) To distinguish norm convergence from weak convergence, in corresponding situations we call convergence w.r.t. the norm **strong convergence***.*

Proposition 4.2 (i) Via

$$\langle x', J_X x \rangle := \langle x, x' \rangle$$

an isometric mapping $J_X \in L(X; X'')$ is defined. Here $X'' := (X')'$ is the **bidual space** of X.

(ii) Let $x_k, x \in X$ for all $k \in \mathbb{N}$, then:

$$x_k \rightharpoonup x \text{ in } X \text{ as } k \to \infty \quad \text{iff} \quad J_X x_k \rightharpoonup^* J_X x \text{ in } X'' \text{ as } k \to \infty.$$

Proof: (i): See Corollary 2.5(iii).

(ii): For $x' \in X'$ is $\langle x_k, x' \rangle = \langle x', J_X x_k \rangle$ and $\langle x, x' \rangle = \langle x', J_X x \rangle$. ■

Proposition 4.3 (i) Corollary 2.5(ii) implies that the weak limit is uniquely determined. For the weak* limit this is trivially true.

(ii) Strong convergence implies weak (weak*) convergence.

(iii) From $x'_k \rightharpoonup^* x'$ in X' as $k \to \infty$ it follows that

$$\|x'\| \leq \liminf_{k\to\infty} \|x'_k\|.$$

(iv) From $x_k \rightharpoonup x$ in X as $k \to \infty$ it follows that

$$\|x\| \leq \liminf_{k\to\infty} \|x_k\|,$$

i.e., the norm is lower continuous w.r.t. weak convergence.

(v) Weakly (weakly*) convergent sequences are norm bounded.

(vi) If $x_k \to x$ strongly in X and $x'_k \rightharpoonup^* x'$ in X' as $k \to \infty$, then

$$\lim_{k\to\infty} \langle x_k, x'_k \rangle = \langle x, x' \rangle.$$

The same statement is true, if $x_k \rightharpoonup x$ in X and $x'_k \to x'$ strongly in X' as $k \to \infty$.

Proof: (iii): For all $x \in X$ we have

$$|\langle x, x' \rangle| = \lim_{k\to\infty} |\langle x, x'_k \rangle| \leq \liminf_{k\to\infty} \|x'_k\| \|x\|,$$

hence

$$\|x'\| \leq \liminf_{k\to\infty} \|x'_k\|.$$

(iv): As in the proof of (iii) we find

$$|\langle x, x' \rangle| \leq \|x'\| \cdot \liminf_{k\to\infty} \|x_k\|.$$

Now we choose x' with $\|x'\| = 1$ and $\langle x, x'\rangle = \|x\|$, see Corollary 2.5(i), and the statement is proven.

(v): If $x'_k \rightharpoonup^* x'$ in X', then

$$\sup_{k\in\mathbb{N}} |\langle x, x'_k\rangle| < \infty \quad \text{for all } x \in X.$$

Thus, by Banach–Steinhaus theorem

$$\sup_{k\in\mathbb{N}} \|x'_k\| < \infty.$$

If $x_k \rightharpoonup x$ in X, then $J_X x_k \rightharpoonup^* J_X x$ in X'', see Proposition 4.2(ii). Therefore, as above we find that $(J_X x_k)_{k\in\mathbb{N}}$ is bounded in X'' and then isometry of J_X yields that $(x_k)_{k\in\mathbb{N}}$ is bounded in X.

(vi): Under the first assumption we have:

$$|\langle x, x'\rangle - \langle x_k, x'_k\rangle| \le |\langle x, x' - x'_k\rangle| + \|x - x_k\|\|x'_k\|.$$

Since $(x'_k)_{k\in\mathbb{N}}$ is bounded in X', see (v), the statement is shown. Under the second assumption the statement can be shown analogously. ■

Example 4.4 Let $1 \le p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, where for $p = 1$ the complete measure space $(\mu, \mathcal{B}, S)$ is assumed to be σ-finite. Then

$$J(g)(f) := \int_S f\overline{g}\,d\mu, \qquad f \in L^p(\mu),\ g \in L^q(\mu),$$

defines an isometric conjugate linear isomorphism $J : L^q(\mu) \to L^p(\mu)'$ (proof will be given later). In the case $p = q = 2$ (Hilbert space) J coincides with the Riesz isomorphism. Hence

$$f_k \rightharpoonup f \quad \text{in} \quad L^p(\mu) \quad \text{as} \quad k \to \infty$$
$$\text{iff} \quad \lim_{k\to\infty} \int_S f_k\overline{g}\,d\mu = \int_S f\overline{g}\,d\mu \quad \text{for all} \quad g \in L^q(\mu).$$

4.2 Banach–Alaoglu theorem

Theorem 4.5 *Let X be separable. Then the closed unit ball $\overline{U_1(0)}$ in X' is weak* sequentially compact.*

Proof: Let $\{x_n \,|\, n \in \mathbb{N}\}$ be dense in X and $(x'_k)_{k\in\mathbb{N}}$ a sequence in X' with $\|x'_k\| \leq 1$. Then the $(\langle x_n, x'_k\rangle)_{k\in\mathbb{N}}$ are bounded sequences in $\mathbb{K}$. Hence, by dropping to subsequences and then to the diagonal sequence we obtain for all $n \in \mathbb{N}$ the existence of

$$\lim_{k\to\infty} \langle x_n, x'_k\rangle \;\in\; \mathbb{K}.$$

Then also for all $y \in Y := \operatorname{span}\{x_n \,|\, n \in \mathbb{N}\}$ there exists

$$\langle y, x'\rangle := \lim_{k\to\infty} \langle y, x'_k\rangle,$$

and $x' : Y \to \mathbb{K}$ is linear. Since

$$|\langle y, x'\rangle| = \lim_{k\to\infty} |\langle y, x'_k\rangle| \leq \|y\|,$$

the mapping $x' \in L(Y;\mathbb{K})$ and therefore can be extended to a continuous, linear mapping on $\overline{Y} = X$, see Exercise 1.1. Consequently, $x' \in X'$ with $\|x'\| \leq 1$. Additionally, for all $x \in X$ and $y \in Y$ we have

$$|\langle x, x' - x'_k\rangle| \leq |\langle x - y, x' - x'_k\rangle| + |\langle y, x' - x'_k\rangle| \leq 2\|x - y\| + |\langle y, x' - x'_k\rangle|.$$

The second term for each $y \in Y$ tends to zero as $k \to \infty$ and the first can be made arbitrarily small, because Y is dense in X. ■

Example 4.6 Theorem 4.5 in general does not hold, if X is not separable. E.g., take $X = L^\infty((0,1))$ and for $1 \geq \epsilon > 0$ define

$$T_\epsilon f := \frac{1}{\epsilon}\int_0^\epsilon f\,dx, \quad f \in L^\infty((0,1)).$$

Then $T_\epsilon \in L^\infty((0,1))'$ with $\|T_\epsilon\| \leq 1$. But there does not exist any zero sequence $(\epsilon_k)_{k\in\mathbb{N}}$ such that the sequence $(T_{\epsilon_k})_{k\in\mathbb{N}}$ converges weakly*.

Proof: Assume there exists such a zero sequence $(\epsilon_k)_{k\in\mathbb{N}}$. W.l.o.g. we can assume (by dropping to a subsequence) that

$$1 > \frac{\epsilon_{k+1}}{\epsilon_k} \text{ for all } k \in \mathbb{N} \text{ and } \lim_{k\to\infty} \frac{\epsilon_{k+1}}{\epsilon_k} = 0.$$

Observe the function

$$f(x) := (-1)^j, \quad \epsilon_{j+1} < x < \epsilon_j,\; j \in \mathbb{N}.$$

Then $f \in L^\infty((0,1))$. We have

$$T_{\epsilon_k} f = \frac{1}{\epsilon_k}\left((\epsilon_k - \epsilon_{k+1})(-1)^k + \int_0^{\epsilon_{k+1}} f\, dx\right),$$

and therefore

$$|T_{\epsilon_k} f - (-1)^k| \le \frac{1}{\epsilon_k}\left(\epsilon_{k+1} + \int_0^{\epsilon_{k+1}} |f|\, dx\right) \le \frac{2\epsilon_{k+1}}{\epsilon_k} \quad \text{for all} \quad k \in \mathbb{N}.$$

This shows that $(T_{\epsilon_k} f)_{k\in\mathbb{N}}$ has the two accumulation points $\{-1, 1\}$. Thus, $(T_{\epsilon_k})_{k\in\mathbb{N}}$ can not be weakly* convergent. ■

4.3 Reflexive spaces

Definition 4.7 Let J_X be the isometry as in Proposition 4.2. The space X is called reflexive, iff J_X is surjective.

Lemma 4.8 (i) If X is reflexive, then weak and weak* convergence in X' coincide.
(ii) If X is reflexive, then each closed subspace of X is reflexive.
(iii) Let $T : X \to Y$ be a continuous isomorphism (Y a Banach space). Then X is reflexive, iff Y is reflexive.
(iv) X is reflexive, iff X' is reflexive.
(v) X' separable, implies X separable.

Proof: (ii): Let $Y \subset X$ be a closed subspace. For $y'' \in Y''$ set

$$\langle x', x''\rangle := \langle x'\big|_Y, y''\rangle, \quad x' \in X'.$$

Then $x'' \in X''$. Define $x := J_X^{-1} x''$. Then we have for all $x' \in X'$ with $x' = 0$ on Y

$$\langle x, x'\rangle = \langle x', x''\rangle = \langle x'\big|_Y, y''\rangle = 0.$$

Therefore, $x \in Y$ by Proposition 2.4. If $x' \in X'$ is an extension of y', provided by Hahn–Banach theorem, we conclude for all $y' \in Y'$:

$$\langle x, y'\rangle = \langle x, x'\rangle = \langle x'\big|_Y, y''\rangle = \langle y', y''\rangle,$$

i.e., $y'' = J_Y(x)$. This yields surjectivity of J_Y.

(iii): Let X be reflexive and $y'' \in Y''$. Then

$$\langle x', x'' \rangle := \langle x' \circ T^{-1}, y'' \rangle, \quad x' \in X',$$

defines an element $x'' \in X''$ and we have for all $y' \in Y'$:

$$\langle y', y'' \rangle = \langle y' \circ T, x'' \rangle = \langle J_X^{-1} x'', y' \circ T \rangle = \langle T J_X^{-1} x'', y' \rangle.$$

Thus, $y'' = J_Y T J_X^{-1} x''$.

(iv) Let X be reflexive: If $x''' \in X'''$, then $x''' \circ J_X \in X'$ and we have for all $x'' \in X''$:

$$\langle x'', x''' \rangle = \langle J_X^{-1} x'', x''' \circ J_X \rangle = \langle x''' \circ J_X, x'' \rangle,$$

i.e., $x''' = J_{X'}(x''' \circ J_X)$.

Let X' be reflexive: Using the arguments as above we obtain that X'' is reflexive. Since J_X is isometric, $J_X(X)$ is a closed subspace of X''. Hence, by (ii) also reflexive. Now (iii) yields reflexivity of X.

(v): Let $\{x'_n \,|\, n \in \mathbb{N}\}$ be dense in X'. Choose $x_n \in X$ such that

$$\langle x_n, x'_n \rangle \geq \frac{1}{2} \|x'_n\| \quad \text{and} \quad \|x_n\| = 1$$

and define $Y := \overline{\operatorname{span}\{x_n \,|\, n \in \mathbb{N}\}}$. If now $x' \in X'$ with $x' = 0$ on Y, then for all $n \in \mathbb{N}$:

$$\|x' - x'_n\| \geq |\langle x_n, x' - x'_n \rangle| = |\langle x_n, x'_n \rangle| \geq \frac{1}{2}\|x'_n\| \geq \frac{1}{2}(\|x'\| - \|x'_n - x'\|).$$

Thus

$$\|x'\| \leq 3 \inf_{n \in \mathbb{N}} \|x' - x'_n\| = 0,$$

because $\{x'_n \,|\, n \in \mathbb{N}\}$ is assumed to be dense in X'. Now Proposition 2.4 yields $Y = X$. ■

Theorem 4.9 *Let X be reflexive. Then the closed unit ball $\overline{U_1(0)} \subset X$ is weak sequentially compact.*

Proof: Let $(x_k)_{k \in \mathbb{N}}$ be a sequence in X with $\|x_k\| \leq 1$ and

$$Y := \overline{\operatorname{span}\{x_n \,|\, n \in \mathbb{N}\}}.$$

Then also Y is reflexive, see Lemma 4.8(ii), and additionally separable. Consequently, also $Y'' = J_Y(Y)$ and Y', see Lemma 4.8(v), are separable. Therefore, we can apply Theorem 4.5 to $(J_Y x_k)_{k\in\mathbb{N}}$. Thus, there exists a $y'' \in Y''$ such that for a subsequence $(k_l)_{l\in\mathbb{N}}$

$$\lim_{l\to\infty} \langle y', J_Y x_{k_l} \rangle = \langle y', y'' \rangle \quad \text{for all} \quad y' \in Y'.$$

Set $x := J_Y^{-1} y'' \in Y$. Then

$$\lim_{l\to\infty} \langle x_{k_l}, y' \rangle = \lim_{l\to\infty} \langle y', J_Y x_{k_l} \rangle = \langle y', y'' \rangle = \langle x, y' \rangle$$

for all $y' \in Y'$. Since for $x' \in X'$ the mapping $x'|_Y$ lies in Y' we also have

$$\lim_{l\to\infty} \langle x_{k_l}, x' \rangle = \langle x, x' \rangle,$$

i.e., $x_{k_l} \rightharpoonup x$ in X as $l \to \infty$. ■

Example 4.10 (i) Each Hilbert space X is reflexive. Hence, together with Riesz representation we have: Let $(x_k)_{k\in\mathbb{N}}$ be a bounded sequence in X, then there exists a subsequence $(x_{k_l})_{l\in\mathbb{N}}$ and $x \in X$ such that

$$\lim_{l\to\infty} (x_{k_l}, y)_X = (x, y)_X \quad \text{for all} \quad y \in X.$$

(ii): $L^p(\mu)$ for $1 < p < \infty$ is reflexive. Therefore, together with Example 4.4: Let $(f_k)_{k\in\mathbb{N}}$ be a bounded sequence in $L^p(\mu)$, then there exists a subsequence $(f_{k_l})_{l\in\mathbb{N}}$ and $f \in L^p(\mu)$ such that

$$\lim_{l\to\infty} \int_S f_{k_l} g \, d\mu = \int_S f g \, d\mu \quad \text{for all} \quad g \in L^q(\mu).$$

(iii) Let μ be σ-finite. Then $L^1(\mu)$ and $L^\infty(\mu)$ are not reflexive, if the underlying σ-algebra has infinite many disjoint sets with positive finite measure (i.e., if and only if $L^1(\mu)$ and $L^\infty(\mu)$, resp., are infinite dimensional).

Proof: (i): Let $J : X \to X'$ be the (conjugate linear) isomorphism provided in the Riesz representation theorem. For $x'' \in X''$ define

$$\langle y, x' \rangle := \overline{\langle Jy, x'' \rangle}, \qquad y \in X.$$

Then $x' \in X'$. Set $x := J^{-1}x'$, then we have for all $y \in X$:

$$\langle Jy, x''\rangle = \overline{\langle y, Jx\rangle} = \overline{(y,x)_X} = \langle x, Jy\rangle,$$

i.e., $x'' = J_X x$. Thus, surjectivity of J_X is shown. Notice, that in the real case $J_X^{-1} = J^{-1}J'$, where $J' : X'' \to X'$ is the adjoint mapping to J.
(ii): The isometries

$$J_p : L^p(\mu) \to L^q(\mu)' \quad \text{and} \quad J_q : L^q(\mu) \to L^p(\mu)'$$

provided in Example 4.4 have the property:

$$\overline{\langle f, J_q g\rangle} = \langle g, J_p f\rangle, \quad f \in L^p(\mu),\, g \in L^q(\mu).$$

For $f'' \in L^p(\mu)''$ we define

$$\langle g, g'\rangle := \overline{\langle J_q g, f''\rangle}, \qquad g \in L^q(\mu).$$

We find $g' \in L^q(\mu)'$. Set $f := J_p^{-1}g'$, then we have for $g \in L^q(\mu)$:

$$\langle g, g'\rangle = \langle g, J_p f\rangle = \overline{\langle f, J_q g\rangle} = \overline{\langle J_q g, J_{L^p(\mu)} f\rangle}.$$

Therefore,

$$\langle J_q g, f''\rangle = \langle J_q g, J_{L^p(\mu)} f\rangle, \quad \text{for all} \quad g \in L^q(\mu).$$

Since J_q is surjective, we can conclude that $f'' = J_{L^p(\mu)} f$. Thus, $L^p(\mu)$ is reflexive. Notice, that in the real case $J_{L^p(\mu)}^{-1} = J_p^{-1}J_q'$, where $J_q' : L^p(\mu)'' \to L^q(\mu)'$ is the adjoint mapping to J_q.
(iii): Because of Lemma 4.8(iv), Example 4.4 for $p = 1$ and Lemma 4.8(iii), it suffices to prove this for $L^1(\mu)$. Let $F \in L^\infty(\mu)'$ and $J_\infty : L^\infty(\mu) \to L^1(\mu)'$ the isomorphism provided in Example 4.4. Then via

$$\langle f', G\rangle := \overline{\langle J_\infty^{-1} f', F\rangle}, \qquad f' \in L^1(\mu)',$$

an element $G \in L^1(\mu)''$ is defined.

Assume that $G = J_{L^1(\mu)} f$ for some $f \in L^1(\mu)$. Then we have for all $g \in L^\infty(\mu)$:

$$\overline{\langle g, F\rangle} = \langle J_\infty g, G\rangle = \langle J_\infty g, J_{L^1(\mu)} f\rangle = \langle f, J_\infty g\rangle = \int_S f\overline{g}\, d\mu,$$

i.e.,

$$\langle g, F\rangle = \int_S g\overline{f}\, d\mu, \quad \text{for all} \quad g \in L^\infty(\mu). \tag{4.1}$$

Now, under the assumptions on μ as in (iv), we construct an F which does not fulfill (4.1). Let $E_k \in \mathcal{B}$ such that

$$E_k \subset E_{k+1},\ \mu(E_k) < \mu(E_{k+1}) \text{ and } E := \bigcup_{k\in\mathbb{N}} E_k.$$

Consider the subspace

$$Y := \overline{\{g \in L^\infty(\mu) \,|\, g = 0 \text{ on } S \setminus E_k \text{ for some } k\}} \subset L^\infty(\mu).$$

Then $\chi_E \notin Y$. Thus, Proposition 2.4 yields the existence of an $F \in L^\infty(\mu)'$ such that $F = 0$ on Y and $F(\chi_E) = 1$. Hence, we have for all k:

$$F(\chi_{E_k}) = 0 \quad \text{and} \quad F(\chi_E) = 1.$$

But for all $f \in L^1(\mu)$ we have

$$\lim_{k\to\infty} \int_S \chi_{E_k}\overline{f}\, d\mu = \int_S \chi_E \overline{f}\, d\mu.$$

That stands in contradiction to (4.1). Therefore, $J_{L^1(\mu)}$ can not be surjective. ■

4.4 Separation theorem

Theorem 4.11 *Let X be a normed space, $M \subset X$ closed and convex, and $x_0 \in X \setminus M$. Then there exists $x' \in X'$ and $\alpha \in \mathbb{R}$ such that*

$$\Re\langle x, x'\rangle \le \alpha \quad \textit{for all} \quad x \in M \quad \textit{and} \quad \Re\langle x_0, x'\rangle > \alpha.$$

Proof: First we consider the case $\mathbb{K} = \mathbb{R}$. Without lost of generality we assume $0 \in \overset{\circ}{M}$ (translate M and x_0 by a point from M and substitute M by $\overline{U_r(M)}$ with $r < \operatorname{dist}(x_0, M)$). Let us consider the **Minkowski functional**

$$p(x) := \inf\left\{ r > 0 \,\middle|\, \frac{x}{r} \in M \right\}, \quad x \in X.$$

Since $0 \in \overset{\circ}{M}$, we have $0 \le p(x) < \infty$ for all $x \in M$. Additionally,

$$p \le 1 \quad \text{on} \quad M \quad \text{and} \quad p(x_0) > 1.$$

Furthermore,

$$\begin{aligned} &p(\alpha x) = \alpha p(x), \quad \alpha \ge 0, \\ &p(x+y) \le p(x) + p(y), \end{aligned}$$

i.e., p is sublinear. Indeed, because for $\alpha > 0$ we have

$$\frac{x}{r} \in M \quad \text{iff} \quad \frac{\alpha x}{\alpha r} \in M,$$

and convexity of M yields:

$$\frac{x}{r}, \frac{y}{s} \in M \quad \text{implies} \quad \frac{x+y}{r+s} = \frac{r}{r+s}\frac{x}{r} + \frac{s}{r+s}\frac{y}{s} \in M.$$

Define

$$f(\alpha x_0) := \alpha p(x_0), \quad \alpha \in \mathbb{R}.$$

Then

$$\begin{aligned} &f(\alpha x_0) = p(\alpha x_0), \quad \alpha \ge 0, \\ &f(\alpha x_0) \le 0 \le p(\alpha x_0), \quad \alpha \le 0. \end{aligned}$$

Now by Hahn–Banach (applied to $\operatorname{span}\{x_0\}$) there exists a linear extension F of f to X such that $F \le p$. Therefore

$$F \le p \le 1 \quad \text{on} \quad M \quad \text{and} \quad F(x_0) = f(x_0) = p(x_0) > 1.$$

Since $\overline{U_r(0)} \subset M$ for some $r > 0$, we have

$$x \in X \text{ implies } \frac{rx}{\|x\|} \in M \text{ implies } p(x) \le \frac{\|x\|}{r} \text{ implies } F(x) \le \frac{1}{r}\|x\|,$$

i.e., $F \in X'$.

In the case $\mathbb{K} = \mathbb{C}$ we consider X as a $\mathbb{R}$ vector space $X_{\mathbb{R}}$ and obtain an $F_{\mathbb{R}} \in X'_{\mathbb{R}}$ with the desired properties. Then as in the proof of Theorem 2.3 we define $F := F_{\mathbb{R}} - iF_{\mathbb{R}}(i\cdot) \in X'$. Since $\Re F = F_{\mathbb{R}}$, the proof is finished. ■

Proposition 4.12 *Let X be a normed space and $M \subset X$ closed and convex. Then M is* **weak sequentially** *closed, i.e., if $x_k \in M$ for all $k \in \mathbb{N}$ and $x_k \rightharpoonup x$ in X as $k \to \infty$, then also $x \in M$.*

Proof: Assume that $x \notin M$. Then by Theorem 4.11 there exists $x' \in X'$ and $\alpha \in \mathbb{R}$ such that

$$\Re\langle y, x'\rangle \leq \alpha \quad \text{for all} \quad y \in M \quad \text{and} \quad \Re\langle x, x'\rangle > \alpha.$$

Hence, $\Re\langle x_k, x'\rangle \leq \alpha$ and because of weak convergence also $\Re\langle x, x'\rangle \leq \alpha$. That is a contradiction. ∎

Proposition 4.13 (Lemma of Mazur) *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in a normed space X converging weakly to x. Then $x \in \overline{\operatorname{conv}\{x_k \,|\, k \in \mathbb{N}\}}$.*

Proof: $\operatorname{conv}\{x_k \,|\, k \in \mathbb{N}\}$ is convex, hence also its closure. Now apply Proposition 4.12. ∎

Theorem 4.14 *Let X be reflexive, $M \subset X$ non-empty, closed and convex. Then for $x_0 \in X$ there exists $x \in M$ such that*

$$\|x_0 - x\| = \operatorname{dist}(x_0, M).$$

Proof: Let $(x_k)_{k\in\mathbb{N}}$ be a minimizing sequence, i.e.,

$$x_k \in M \text{ for all } k \in \mathbb{N} \quad \text{and} \quad \lim_{k\to\infty} \|x_0 - x_k\| = \operatorname{dist}(x_0, M).$$

Then $(x_k)_{k\in\mathbb{N}}$ is a bounded sequence and Theorem 4.9 yields the existence of a subsequences $(k_l)_{l\in\mathbb{N}}$ and an $x \in X$ such that $x_{k_l} \rightharpoonup x$ as $l \to \infty$. By Proposition 4.12 $x \in M$. Since also $x_{k_l} - x_0 \rightharpoonup x - x_0$ as $l \to \infty$, lower continuity of the norm, see Proposition 4.3(iv), implies that $\|x_0 - x\| = \operatorname{dist}(x_0, M)$. ∎

5 Projections

In this section we assume X to be a $\mathbb{K}$ vector space.

5.1 Linear projections

Definition 5.1 *Let Y be a subspace of X. A linear mapping $P : X \to X$ is called* **(linear) projection** *on Y, iff*

$$P^2 = P \quad \text{and} \quad \mathcal{R}(P) = Y.$$

Proposition 5.2 *(i) P is a projection on a subspace $Y \subset X$, iff*

$$P : X \to Y \quad \text{and} \quad P = Id \text{ on } Y.$$

(ii) If $P : X \to X$ is a projection, then

$$X = \mathcal{N}(P) \oplus \mathcal{R}(P).$$

(iii) If $P : X \to X$ is a projection, then also $Id - P$ and

$$\mathcal{N}(Id - P) = \mathcal{R}(P) \quad \text{and} \quad \mathcal{R}(Id - P) = \mathcal{N}(P).$$

(iv) For each subspace $Y \subset X$ there exist a linear projection on Y.

Proof: (i): Obvious!

(ii): We have for all $x \in X$:

$$x = x - Px + Px.$$

Here $(x - Px) \in \mathcal{N}(P)$ and $Px \in \mathcal{R}(P)$. If $x \in \mathcal{N}(P) \cap \mathcal{R}(P)$, then $Px = 0$ and $P(x) = x$, thus $x = 0$.

(iii): We have

$$(Id - P)^2 = Id - 2P + P^2 = Id - 2P + P = Id - P.$$

Furthermore

$$x \in \mathcal{N}(Id - P) \text{ iff } x - Px = 0 \text{ iff } x \in \mathcal{R}(P),$$

hence $\mathcal{N}(Id-P) = \mathcal{R}(P)$. Then also $\mathcal{N}(P) = \mathcal{N}(Id-(Id-P)) = \mathcal{R}(Id-P)$.

(iv): As in the proof of Theorem 2.1 (Hahn–Banach) set

$$\mathcal{M} := \{(Z, P) \mid Z \text{ subspace}, Y \subset Z \subset X, \\ P : Z \to Y \text{ linear}, P = Id \text{ on } Y\}.$$

with the same order relation. Analogously as in the proof of Theorem 2.1 one can prove that $\mathcal{M}$ possesses a maximal element (Z, P). Suppose there exists $z_0 \in X \setminus Z$. Then

$$Z_0 := Z \oplus \text{span}\{z_0\}, \quad P_0(z + \alpha z_0) := P(z), \quad z \in Z, \alpha \in \mathbb{K},$$

defines an element $(Z_0, P_0) \in \mathcal{M}$ with $(Z, P) \leq (Z_0, P_0)$ and $Z_0 \neq Z$. But (Z, P) is maximal. That's a contradiction. ■

5.2 Continuous projections

Proposition 5.3 *Let X be a normed space and $P \in P(X)$ (linear continuous projection).*
(i) $\mathcal{N}(P)$ and $\mathcal{R}(P)$ are closed.
(ii) $\|P\| \geq 1$ or $P = 0$.

Proof: (i): Since the pre-image of a closed set under a continuous mapping is closed $\mathcal{N}(P) = P^{-1}(\{0\})$ is closed. By Proposition 5.2(iii) then also $\mathcal{R}(P)$ is closed.

(ii): Since $L(X)$ is a Banach algebra, we have $\|P\| = \|P^2\| \leq \|P\|^2$. Thus $\|P\| = 0$ or $\|P\| \geq 1$. ∎

5.3 Closed complement theorem

Theorem 5.4 *Let Y be a closed subspace of a Banach space X and Z a subspace such that $X = Y \oplus Z$. Then the following are equivalent:*
(i) There exists a continuous projection P on Y with $Z = \mathcal{N}(P)$.
(ii) Z is closed.

Proof: (i) implies (ii): $\mathcal{N}(P)$ is closed.

(ii) implies (i): Consider the Banach space

$$\widetilde{X} := Y \times Z, \quad \|(y,z)\|_{\widetilde{X}} := \|y\|_X + \|z\|_X,$$

and define $T(y,z) := y + z$. Since $X = Y \oplus Z$, $T : \widetilde{X} \to X$ is linear and bijective. Define $P_Y : X \to Y$ and $P_Z : X \to Z$ via

$$T^{-1}x = (P_Y x, P_Z x), \quad x \in X.$$

Then P_Y and P_Z are linear. Since $T^{-1}(y) = (y, 0)$ for $y \in Y$, $P_Y = Id$ on Y, i.e., P_Y is a projection on Y. Because $\|P_Y x\|_X \leq \|T^{-1}x\|_{\widetilde{X}}$, P_Y is continuous if T^{-1} is continuous. Since $\|T(y,z)\|_X \leq \|(y,z)\|_{\widetilde{X}}$, T is continuous and therefore also T^{-1} by the inverse mapping theorem. ∎

Corollary 5.5 Let Y be a finite dimensional subspace of a Banach space X and Z a closed subspace such that $X = Y \oplus Z$. If $W \cap Z = \{0\}$, then W is finite dimensional with $\dim(W) \leq \dim(Y)$ and $\dim(W) = \dim(Y)$, iff $X = W \oplus Z$.

Proof: Since Y is finite dimensional, it is closed. Let $P \in P(X)$ be the projection on Y with $Z = \mathcal{N}(P)$ provided in Theorem 5.4. Then

$$S := P|_W : W \to Y$$

is linear and injective. Indeed, if $Py = 0$, then $y \in Z \cap W = \{0\}$. Since Y is finite dimensional, this implies that also W is finite dimensional with $\dim(W) \leq \dim(Y)$.

If $X = W \oplus Z$, then as above (exchange Y and W) $\dim(Y) \leq \dim(W)$, i.e., $\dim(W) = \dim(Y)$.

If $\dim(W) = \dim(Y)$, then is S bijective. Thus for $x \in X$ is

$$y := S^{-1}Px \in W$$

with

$$Py = PS^{-1}Px = SS^{-1}Px = Px,$$

i.e., $x - y \in \mathcal{N}(P) = Z$. This proves $X = W \oplus Z$. ■

5.4 Orthogonal projections

Lemma 5.6 Let Y be a closed subspace of a Hilbert space X and P the orthogonal projection on Y provided in Corollary E5.14. Then:
(i): $P \in P(X)$.
(ii): $\mathcal{R}(P) = Y$ and $\mathcal{N}(P) = Y^\perp$.
(iii): $X = Y \perp Y^\perp$.
(iv): Let $Z \subset X$ a subspace such that $X = Y \perp Z$, then $Z = Y^\perp$. That is why $Y^\perp$ is called the **orthogonal complement** of Y.

Proof: (i), (ii): P as in Corollary E5.14 is characterized by

$$(x - Px, y) = 0 \qquad \forall y \in Y, \tag{5.1}$$

and from this we already concluded linearity of P. Additionally, P is continuous because when setting $y = Px$, (5.1) implies

$$\|Px\|^2 = (Px, Px) = (x, Px) \leq \|x\|\|Px\|,$$

thus $\|Px\| \leq \|x\|$. Furthermore, (5.1) immediately yields that $P \in P(X)$. Indeed, if $x \in Y$, then set $y = x - Px \in Y$ in (5.1) and obtain $x - Px = 0$, i.e., $P = Id$ on Y. Furthermore, (5.1) implies

$$x \in \mathcal{N}(P) \quad \text{iff} \quad Px = 0 \quad \text{iff} \quad (x, y) = 0 \ \forall y \in Y \quad \text{iff} \quad x \in Y^{\perp}.$$

(iii): Follows from Proposition 5.2(ii).

(iv): First observe that $Z \subset Y^{\perp}$. But, if $x \in Y^{\perp}$ with the representation $x = z + y$, $z \in Z$, $y \in Y$, then also $x - z \in Y^{\perp}$. Thus, $0 = (x - z, y) = \|y\|^2$, i.e., $x = z \in Z$. ■

Proposition 5.7 Let X be a Hilbert space and $P : X \to X$ linear. Then the following statements are equivalent:
(i) P is an orthogonal projection on $\mathcal{R}(P)$, i.e.,

$$\|x - Px\| \leq \|x - Py\| \quad \forall x, y \in X.$$

(ii) $(x - Px, Py) = 0$ for all $x, y \in X$.
(iii) $P^2 = P$ and $(x, Py) = (Px, y)$ for all $x, y \in X$ (i.e., P is self-adjoint).
(iv) $P \in P(X)$ and $\|P\| \leq 1$ (then $\|P\| = 1$ or $\|P\| = 0$ by Proposition 5.3(iv)).

Proof: (i) is equivalent to (ii): See the proofs of Proposition E5.13 and Corollary E5.14.

(ii) implies (iii): For $x, y \in X$ we have:

$$\begin{aligned} 0 &= (x - Px, Py) - \overline{(y - Py, Px)} \\ &= (x, Py) - (Px, Py) - \overline{(y, Px)} + \overline{(Py, Px)} \\ &= (x, Py) - (Px, y). \end{aligned}$$

Using this identity we get for $x \in X$:

$$(P^2x - Px, y) = (P(Px - x), y) = (Px - x, Py) = 0$$

for all $y \in X$. Thus, $P^2x = Px$.

(iii) implies (iv): Set $y = Px$ in (iii) and obtain

$$\|Px\|^2 = (x, P^2x) = (x, Px) \leq \|x\|\|Px\|.$$

Hence $\|Px\| \leq \|x\|$ and therefore $\|P\| \leq 1$. Now $P^2 = P$ yields $P \in P(X)$.

(iv) implies (ii): Let $x \in X$, $y \in \mathcal{R}(P)$ and set $z = x - Px$. Since $Py = y$ and $Pz = 0$ we have for $\varepsilon > 0$ and $|\alpha| = 1$:

$$\|y\|^2 = \|P(\varepsilon z + \alpha y)\|^2 \leq \varepsilon^2 \|z\|^2 + 2\varepsilon \Re(z, \alpha y) + \|y\|^2.$$

Thus

$$0 \leq \lim_{\varepsilon \to 0} \varepsilon \|z\|^2 + 2\Re(z, \alpha y) = 2\Re\overline{\alpha}(z, y).$$

Since this holds for all $|\alpha| = 1$ we have

$$0 = (z, y) = (x - Px, y).$$

■

6 Bounded operators

In this section we assume X and Y to be normed $\mathbb{K}$ vector spaces.

6.1 Adjoint operators

Let us recall the definition of the adjoint operator given in Definition E.4.4.

Definition 6.1 *For $T \in L(X; Y)$*

$$\langle x, T'y' \rangle := \langle Tx, y' \rangle, \quad x \in X,\, y' \in Y',$$

defines a linear mapping $T' : Y' \to X'$. T' is called the **adjoint operator** *to T. Since*

$$|\langle x, T'y' \rangle| \leq \|y'\|_{Y'} \|T\| \|x\|_X,$$

we have $T' \in L(Y'; X')$ with $\|T'\| \leq \|T\|$.

Example 6.2 Let $X = Y = l^1(\mathbb{K})$ and T the shift operator

$$T(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots), \quad (x_1, x_2, \ldots) \in l^1(\mathbb{K}).$$

Then $J_\infty^{-1} T' J_\infty : l^\infty(\mathbb{K}) \to l^\infty(\mathbb{K})$ is the operator

$$J_\infty^{-1} T' J_\infty (y_1, y_2, \ldots) = (y_2, y_3, \ldots), \quad (y_1, y_2, \ldots) \in l^\infty(\mathbb{K}).$$

Furthermore, $\|T\| = 1 = \|T'\|$.

Theorem 6.3 *Let X and Y be Banach spaces. The map $T \to T'$ is an isometric embedding of $L(X;Y)$ into $L(Y';X')$.*

Proof: The map $T \to T'$ is linear. Furthermore

$$\begin{aligned}\|T\| = \sup_{\|x\|_X \le 1} \|Tx\|_Y &= \sup_{\|x\|_X \le 1} \big(\sup_{\|y'\|_{Y'} \le 1} |\langle Tx, y' \rangle| \big) \\ &= \sup_{\|y'\|_{Y'} \le 1} \big(\sup_{\|x\|_X \le 1} |\langle x, T'y' \rangle| \big) = \sup_{\|y'\|_{Y'} \le 1} \|T'y'\| = \|T'\|.\end{aligned}$$

The second equality is a consequence of Corollary 2.5(i). ■

Let us recall the definition of the Hilbert space adjoint given in Definition E5.16.

Definition 6.4 *Let T be a bounded linear operator mapping a Hilbert space X into itself. The Banach space adjoint is then in $L(X')$. Recall the conjugate linear Riesz isomorphism*

$$J : X \to X'.$$

The **Hilbert space adjoint** *then is defined by*

$$T^* = J^{-1} T' J \in L(X).$$

$T \in L(X)$ *is called* **self-adjoint**, *iff* $T^* = T$.

The Hilbert space adjoint satisfies

$$(Tx, y) = \langle Tx, Jy \rangle = \langle x, T'Jy \rangle = (x, J^{-1}T'Jy) = (x, T^*y), \quad x, y \in X.$$

Proposition 6.5 *Let X be a Hilbert space and $T, S \in L(X)$.*
(i) $(T^*)^* = T$.
(ii) $(TS)^* = S^*T^*$.
(iii) $T \to T^*$ *is a conjugate linear isometric isomorphism of* $L(X)$ *onto* $L(X)$.
(iv) If T *has a bounded inverse, then* T^* *has a bounded inverse and* $(T^*)^{-1} = (T^{-1})^*$.
(v) $\|T^*T\| = \|T\|^2$.

Proof: (i), (ii): Easily checked.

(iii): Follows from Theorem 6.3, the fact that J is a conjugate linear isometry and (i).

(iv): Since $T^{-1}T = Id = TT^{-1}$ we have from (ii)

$$T^*(T^{-1})^* = Id^* = Id = Id^* = (T^{-1})^*T^*$$

which proves (iv).

(v) Note that by (iii)

$$\|T^*T\| \le \|T^*\|\|T\| = \|T\|^2$$

and

$$\|T^*T\| \ge \sup_{\|x\|\le 1} (T^*Tx, x) = \sup_{\|x\|\le 1} (Tx, Tx) = \sup_{\|x\|\le 1} \|Tx\|^2 = \|T\|^2.$$

■

Lemma 6.6 Let X be a Hilbert space and $T \in L(X)$ self-adjoint. Then

$$\|T\| = \sup_{\|x\|\le 1} |(Tx, x)|.$$

Proof: See Exercise 11.2. ■

6.2 Spectrum and resolvent

Proposition 6.7 *Let X be a Banach space and suppose $T \in L(X)$. Then for any two points $\lambda, \mu \in \rho(T)$, $R(\lambda; T)$ and $R(\mu; T)$ commute and*

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

Proof: The expression

$$R(\lambda; T) - R(\mu; T) = R(\lambda; T)(\mu Id - T)R(\mu; T) - R(\lambda; T)(\lambda Id - T)R(\mu; T)$$

proves the first resolvent equation. Interchanging λ and μ shows that $R(\lambda; T)$ and $R(\mu; T)$ commute. ■

The statement of the following lemma was already shown in Proposition 1.11. But here we give a different proof, which exemplary shows how to generalize results from Complex Analysis for mappings with values in $\mathbb{C}$ to mappings with values in a $\mathbb{C}$ Banach space.

Lemma 6.8 Let $X \neq \{0\}$ be a $\mathbb{C}$ Banach space, $T \in L(X)$. Then the spectrum of T is not empty.

Proof: If $|\lambda| > \|T\|$, then we have

$$R(\lambda; T) = \frac{1}{\lambda}\left(Id - \frac{T}{\lambda}\right)^{-1} = \frac{1}{\lambda}\left(Id + \sum_{n=1}^{\infty}\left(\frac{T}{\lambda}\right)^n\right)$$

(Neumann series). Thus

$$\lim_{|\lambda|\to\infty} \|R(\lambda; T)\| = 0. \tag{6.1}$$

Assume that $\sigma(T) = \emptyset$. Then by Proposition 1.9

$$R(\cdot; T) : \mathbb{C} \to L(X)$$

is a holomorphic mapping. Hence there exists a sequence $(T_n)_{n\in\mathbb{N}}$ in $L(X)$ such that

$$R(\lambda; T) = \sum_{n=0}^{\infty} T_n \lambda^n, \quad \lambda \in \mathbb{C}.$$

In particular, $R(\cdot; T)$ is a continuous mapping and therefore bounded on compact subsets of $\mathbb{C}$. This together with (6.1) yields the existence of a constant $0 < C < \infty$ (independent of $\lambda \in \mathbb{C}$) such that

$$\|R(\lambda; T)\| \leq C \quad \text{for all} \quad \lambda \in \mathbb{C}.$$

Hence for all $y' \in L(X)$

$$\langle R(\lambda; T), y'\rangle = \sum_{n=0}^{\infty} \langle T_n, y'\rangle \lambda^n, \quad \lambda \in \mathbb{C},$$

and

$$|\langle R(\lambda; T), y'\rangle| \leq \|R(\lambda; T)\|\|y'\| \leq C\|y'\| \quad \text{for all} \quad \lambda \in \mathbb{C}.$$

Therefore

$$\langle R(\cdot; T), y'\rangle : \mathbb{C} \to \mathbb{C}$$

is a bounded holomorphic function. By Liouville's theorem together with (6.1)

$$\langle R(\lambda; T), y' \rangle = 0 \quad \text{for all} \quad \lambda \in \mathbb{C}, y' \in L(X).$$

Then Corollary 2.5(i) implies

$$R(\lambda; T) = 0 \quad \text{for all} \quad \lambda \in \mathbb{C}.$$

This is impossible if $X \neq \{0\}$. Contradiction! Thus, $\sigma(T)$ is not empty. ■

Theorem 6.9 (Phillips) *Let X be a Banach space, $T \in L(X)$. Then $\sigma(T) = \sigma(T')$ and $R(\lambda; T') = R(\lambda; T)'$, $\lambda \in \rho(T) = \rho(T')$. If X is a Hilbert space, then $\sigma(T^*) = \{\lambda \mid \overline{\lambda} \in \sigma(T)\}$ and $R(\overline{\lambda}; T^*) = R(\lambda; T)^*$.*

Proof: Let $\lambda \in \rho(T)$, then

$$\begin{aligned} \langle x, x' \rangle &= \langle (\lambda Id - T) R(\lambda; T)x, x' \rangle \\ &= \langle R(\lambda; T)x, (\lambda Id - T')x' \rangle = \langle x, R(\lambda; T)'(\lambda Id - T')x' \rangle \end{aligned}$$

for all $x \in X$ and $x' \in X'$. The same holds when interchanging $(\lambda Id - T)$ and $R(\lambda; T)$. Therefore,

$$R(\lambda; T)'(\lambda Id - T') = Id = (\lambda Id - T')R(\lambda; T)',$$

i.e., $R(\lambda; T') = R(\lambda; T)'$ and, in particular, $\rho(T) \subset \rho(T')$. Starting with $\lambda \in \rho(T')$, $(\lambda Id - T')$ and $R(\lambda; T')$ in an analogous way we obtain $\rho(T') \subset \rho(T)$. Thus, $\rho(T) = \rho(T')$ and therefore also $\sigma(T) = \sigma(T')$.

The Hilbert space case follows from Proposition 6.5 or by an analogous consideration as above, but with the scalar product instead of the dual paring. ■

Example 6.10 Let T be the shift operator on $l^1(\mathbb{K})$ acting as

$$T(x_1, x_2, \ldots) = (x_2, x_3, \ldots), \quad (x_1, x_2, \ldots) \in l^1(\mathbb{K}).$$

Its adjoint $T' : l^\infty(\mathbb{K}) \to l^\infty(\mathbb{K})$ is the operator

$$T'(y_1, y_2, \ldots) = (0, y_1, y_2, \ldots), \quad (y_1, y_2, \ldots) \in l^\infty(\mathbb{K}).$$

(here we identify $J_\infty^{-1}T'J_\infty$ and T'). It is easy to check that $\|T\| = 1 = \|T'\|$. Thus all λ with $|\lambda| > 1$ are in $\rho(T)$ and $\rho(T')$.

Suppose $|\lambda| < 1$. Then the vector

$$x_\lambda := (1, \lambda, \lambda^2, \ldots)$$

is in $l^1(\mathbb{K})$ and satisfies

$$(\lambda Id - T)x_\lambda = 0.$$

Thus, all such λ are in the point spectrum of T. Since the spectrum is closed $\sigma(T) = \{\lambda \,|\, |\lambda| \leq 1\}$. By Theorem 6.9 this set is also the spectrum of T'. We want to show that T' has no point spectrum. Suppose that $y = (y_n)_{n\in\mathbb{N}} \in l^\infty(\mathbb{K})$ such that $(\lambda Id - T')y = 0$. Then

$$\lambda y_1 = 0, \quad \lambda y_2 - y_1 = 0, \quad \ldots.$$

These equations together imply that $y = 0$. So $(\lambda Id - T')$ is injective and T' has no point spectrum. Next suppose $|\lambda| < 1$. Then for all $y \in l^\infty(\mathbb{K})$

$$\langle x_\lambda, (\lambda Id - T')y\rangle = \langle (\lambda Id - T)x_\lambda, y\rangle = 0,$$

where $x_\lambda \in l^1(\mathbb{K})$ is the eigenvector with eigenvalue λ. By Corollary 2.5(i) we now that there exists an element in $l^\infty(\mathbb{K})$ which does not vanish on x_λ, so the range of $(\lambda Id - T')$ is not dense. Thus $\{\lambda \,|\, |\lambda| < 1\}$ is in the residual spectrum of T'.

It remains to consider the boundary $|\lambda| = 1$. Suppose that $|\lambda| = 1$ and $(\lambda Id - T)x = 0$ for some $x = (x_n)_{n\in\mathbb{N}} \in l^1(\mathbb{K})$. Then

$$x_2 = \lambda x_1, \quad x_3 = \lambda x_2, \quad \ldots.$$

So, $x = x_1(1, \lambda, \lambda^2, \ldots)$ which is not in $l^1(\mathbb{K})$. Thus, λ is not in the point spectrum. If the range of $(\lambda Id - T)$ were not dense, there would be a nonzero $y \in l^\infty(\mathbb{K})$ such that

$$\langle (\lambda Id - T)x, y\rangle = 0 \quad \forall x \in l^1(\mathbb{K}).$$

But then

$$\langle x, (\lambda Id - T')y\rangle = 0 \quad \forall x \in l^1(\mathbb{K})$$

which would imply that λ is in the point spectrum of T' which we have proven cannot occur. Thus, $\{\lambda \,|\, |\lambda| = 1\}$ is neither in the point spectrum of T nor in the residual spectrum of T, hence in the continuous spectrum of T.

Finally, we prove that $\{\lambda \,|\, |\lambda| = 1\}$ is in the residual spectrum of T' by explicitly finding an open ball disjoint from $\mathcal{R}(\lambda Id - T')$. If $x = (x_n)_{n\in\mathbb{N}}, y = (y_n)_{n\in\mathbb{N}} \in l^\infty(\mathbb{K})$ and obey $y = (\lambda Id - T')x$, then

$$y_1 = \lambda x_1, \ldots, y_n = \lambda x_n - x_{n-1}, \ldots.$$

Therefore,

$$x_n = \overline{\lambda}^{n+1} \sum_{m=1}^{n} \lambda^m y_m.$$

Let $z = (z_n)_{n\in\mathbb{N}} \in l^\infty(\mathbb{K})$ with $z_n = \overline{\lambda}^n$ and suppose that $w \in l^\infty(\mathbb{K})$ with $\|w - z\|_\infty \leq \frac{1}{2}$. Then

$$\Re(\lambda^n w_n) \geq \Re(\lambda^n z_n) - \|w - z\|_\infty \geq \frac{1}{2}.$$

Thus, if $(\lambda Id - T')v = w$ for some $v \in l^\infty(\mathbb{K})$, then since

$$v_n = \overline{\lambda}^{n+1} \sum_{m=1}^{n} \lambda^m w_m$$

$|v_n| \geq n/2$ which is impossible. Therefore, $\mathcal{R}(\lambda Id - T')$ does not intersect with the ball of radius $\frac{1}{2}$ about z. Thus, λ is in the residual spectrum.

Proposition 6.11 Let X be a Banach space, $T \in L(X)$. Then
(i) If λ is in the residual spectrum of T, then λ is in the point spectrum of T'.
(ii) If λ is in the point spectrum of T, then λ is either in the point spectrum or the residual spectrum of T'.

Proof: (i): Since $(\lambda Id - T)$ is not dense, by Proposition 2.4 there exists an $0 \neq x' \in X'$ such that

$$0 = \langle (\lambda Id - T)x, x' \rangle = \langle x, (\lambda Id - T')x' \rangle \quad \forall x \in X.$$

So x' is an eigenvector of T' corresponding to the eigenvalue λ.

(ii): Let x be an eigenvector of T corresponding to the eigenvalue λ, then

$$0 = \langle (\lambda Id - T)x, x' \rangle = \langle x, (\lambda Id - T')x' \rangle \quad \forall x' \in X'.$$

Furthermore, by Corollary 2.5(i) the exists an $x_0' \in X'$ such that $\langle x, x_0' \rangle \neq 0$. Therefore $\mathcal{R}(\lambda Id - T')$ cannot be dense in X'. If now $(\lambda Id - T')$ is not injective, then λ is in the point spectrum of T'. Otherwise λ is in the residual spectrum of T'. ■

Theorem 6.12 *Let $A \in L(X)$ be a self-adjoint operator on a Hilbert space X. Then,*
(i) A has no residual spectrum.
(ii) $\sigma(A) \subset \mathbb{R}$.
(iii) Eigenvectors corresponding to distinct eigenvalues of A are orthogonal.

Proof: (i): First note that the point spectrum is a subset of $\mathbb{R}$. Then (i) follows from Proposition 6.11 and the fact that the point and residual spectrum are disjoint by definition.

(ii): If λ and μ are real, we compute

$$\begin{aligned}\|((\lambda + i\mu)Id - A)x\|^2 &= (x, ((\lambda - i\mu)Id - A)((\lambda + i\mu)Id - A)x) \\ &= \|(\lambda Id - A)x\|^2 + \mu^2\|x\|^2, \quad x \in X.\end{aligned}$$

Thus

$$\|((\lambda + i\mu)Id - A)x\| \geq |\mu|\|x\|. \tag{6.2}$$

Now let $\mu \neq 0$. Then (6.2) implies that $((\lambda + i\mu)Id - A)$ is an injection and has bounded inverse on its range which is closed. Since A has no residual spectrum, $\mathcal{R}((\lambda + i\mu)Id - A) = X$. Therefore $(\lambda + i\mu) \in \rho(A)$ if $\mu \neq 0$. Thus $\sigma(A) \subset \mathbb{R}$.

(iii): Let $x_\mu, x_\lambda \in X$ be eigenvectors corresponding to $\mu \neq \lambda$, respectively. Then by (ii)

$$\lambda(x_\lambda, x_\mu) = (Ax_\lambda, x_\mu) = (x_\lambda, Ax_\mu) = \mu(x_\lambda, x_\mu).$$

Hence

$$(\lambda - \mu)(x_\lambda, x_\mu) = 0.$$

Since $\lambda \neq \mu$ this implies $(x_\lambda, x_\mu) = 0$. ■

6.3 Spectral theorem (continuous functional calculus)

In this subsection X is always assumed to be a $\mathbb{C}$ vector space.

Theorem 6.13 *Let A be a self-adjoint bounded operator on a Hilbert space X. Then, there exists a unique map $\phi : C(\sigma(A)) \to L(X)$ with the following properties:*
(i) ϕ is linear and an algebraic $*$-homomorphism, that is,

$$\phi(fg) = \phi(f)\phi(g) \qquad \phi(\lambda f) = \lambda\phi(f)$$
$$\phi(1) = Id \qquad \phi(\overline{f}) = \phi(f)^*$$

for all $f, g \in C(\sigma(A))$, $\lambda \in \mathbb{C}$.
(ii) ϕ is continuous, that is, $\|\phi(f)\| \leq C\|f\|_{C(\sigma(A))}$ for some $C < \infty$.
(iii) Let f be the function $f(x) = x$, $x \in \sigma(A)$, then $\phi(f) = A$.
Moreover, ϕ has the additional properties:
(iv) If $A\psi = \lambda\psi$, $\psi \in X$, then $\phi(f)\psi = f(\lambda)\psi$.
(v) $\sigma(\phi(f)) = \{f(\lambda) \mid \lambda \in \sigma(A)\}$ (spectral mapping theorem).
(vi) If $f \geq 0$, then $\phi(f) \geq 0$.
(vii) $\|\phi(f)\| = \|f\|_{C(\sigma(A))}$ (this strengthens (ii)).
We sometimes write $\phi_A(f)$ or $f(A)$ for $\phi(f)$ to emphasize the dependence on A.

Lemma 6.14 Let A be a bounded operator on a Banach space X and $P(x) = \sum_{n=0}^{N} a_n x^n$, $x, a_n \in \mathbb{C}, 0 \leq n \leq N$. Then

$$\sigma(P(A)) = \{P(\lambda) \mid \lambda \in \sigma(A)\}.$$

Proof: Let $\lambda \in \sigma(A)$. Since $x = \lambda$ is a root of $P(x) - P(\lambda)$, we have

$$P(x) - P(\lambda) = (x - \lambda)Q(x) = Q(x)(x - \lambda)$$

so

$$P(A) - P(\lambda)Id = (A - \lambda Id)Q(A) = Q(A)(A - \lambda Id).$$

Since $(A - \lambda Id)$ has no inverse neither does $P(A) - P(\lambda)Id$, that is, $P(\lambda) \in \sigma(P(A))$.

Conversely, let $\mu \in \sigma(P(A))$ and let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the roots of $P(x) - \mu$, that is,

$$P(x) - \mu = a(x - \lambda_1) \cdot \ldots \cdot (x - \lambda_n), \quad a \in \mathbb{C}.$$

The case $a = 0$ is trivial. Hence let $a \neq 0$. If $\lambda_1, \ldots, \lambda_n \notin \sigma(A)$, then

$$(P(A) - \mu Id)^{-1} = a^{-1}(A - \lambda_n Id)^{-1} \cdot \ldots \cdot (A - \lambda_1 Id)^{-1}.$$

So we conclude that some $\lambda_i \in \sigma(A)$, that is, $\mu = P(\lambda)$ for some $\lambda \in \sigma(A)$. ■

Lemma 6.15 Let A be a bounded self-adjoint operator on a Hilbert space X and $P(x) = \sum_{n=0}^{N} a_n x^n$, $x, a_n \in \mathbb{C}$, $0 \leq n \leq N$. Then

$$\|P(A)\| = \sup_{\lambda \in \sigma(A)} |P(\lambda)|.$$

Proof:

$$\begin{aligned}
\|P(A)\|^2 = \Big(\sup_{\|\varphi\| \leq 1} \sqrt{(P(A)\varphi, P(A)\varphi)}\Big)^2 &= \sup_{\|\varphi\| \leq 1} (P(A)\varphi, P(A)\varphi) \\
= \sup_{\|\varphi\| \leq 1} (\varphi, P(A)^* P(A)\varphi) &= \|P(A)^* P(A)\| = \|\overline{P}P(A)\| \\
= \sup_{\lambda \in \sigma(\overline{P}P(A))} |\lambda| = \sup_{\lambda \in \sigma(A)} |\overline{P}P(\lambda)| &= \sup_{\lambda \in \sigma(A)} |P(\lambda)|^2 = \Big(\sup_{\lambda \in \sigma(A)} |P(\lambda)|\Big)^2,
\end{aligned}$$

where we used Lemma 6.6, Lemma 1.22 and Lemma 6.14. ■

Proof of Theorem 6.13: Properties (i), (iii) imply that

$$\phi(P) = P(A)$$

for each polynomial. Then by Lemma 6.15

$$\|\phi(P)\| = \|P\|_{C(\sigma(A))}.$$

Therefore, ϕ has a unique continuous linear extension to the closure of the polynomials in $C(\sigma(A))$, i.e, to all of $C(\sigma(A))$ by Weierstraß approximation theorem. Hence, properties (i)-(iii) determine ϕ uniquely. Obviously, properties (i)-(iii), (vii) also hold for the closure.

(iv): Note that

$$\phi(P)\psi = P(\lambda)\psi$$

for all polynomials. Thus

$$\phi(f)\psi = f(\lambda)\psi$$

for all $f \in C(\sigma(A))$ by continuity.

(vi): Notice if $f \geq 0$ then $f = g^2$ with g real and $g \in C(\sigma(A))$. Thus, $\phi(f) = \phi(g)^2$ with $\phi(g)$ self-adjoint, so $\phi(f) \geq 0$.

(v): See Exercise 12.1. ∎

Example 6.16 (i) Theorem 6.13 gives the existence of the **square root** of positive semi-definite $A \in L(X)$ (see Corollary 1.26(ii) for the definiton of positive semi-definite).

First note that on a complex Hilbert space positive semi-definite operators are always self-adjoint (in the real case this is not true). Indeed, since

$$\mathbb{R} \ni (Ax, x) = \overline{(Ax, x)} = (x, Ax) \quad \text{for all} \quad x \in X,$$

we get

$$(Ax, y) = (x, Ay) \quad \text{for all} \quad x, y \in X,$$

by the polarization identities:

$$\begin{aligned}(Ax, y) = \frac{1}{4}\Big(&(A(x+y), x+y) - (A(x-y), x-y) \\ &+ i\big((A(x+iy), x+iy) - (A(x-iy), x-iy)\big)\Big)\end{aligned}$$

and

$$\begin{aligned}(x, Ay) = \frac{1}{4}\Big(&(x+y, A(x+y)) - (x-y, A(x-y)) \\ &+ i\big((x+iy, A(x+iy)) - (x-iy, A(x-iy))\big)\Big), \quad x, y \in X.\end{aligned}$$

Then $\sigma(A) \subset \mathbb{R}$ by Theorem 6.12(ii). Now let $\lambda < 0$, then

$$\begin{aligned}\|(\lambda Id - A)x\|^2 &= (\lambda x - Ax, x - Ax) \\ &= \lambda^2(x, x) - \lambda(x, Ax) - \lambda(Ax, x) + (Ax, Ax) \\ &\geq \lambda^2(x, x) = |\lambda|^2\|x\|^2 \quad \text{for all} \quad x \in X.\end{aligned}$$

Hence, as in the proof of Theorem 6.12(ii) it follow that $\lambda \in \rho(A)$. Thus $\sigma(A) \subset [0, \infty)$.

If $f = \sqrt{\cdot}$, then $f \in C(\sigma(A))$ and real valued. Thus, $\sqrt{A}$ is well-defined, self-adjoint and

$$\sqrt{A}\sqrt{A} = \phi(\sqrt{\cdot})\phi(\sqrt{\cdot}) = \phi((\sqrt{\cdot})^2) = A,$$

by Theorem 6.13.
(ii) If $A \in L(X)$, then obviously $A^*A \geq 0$. Hence we can define the **modulus** of A by

$$L(X) \ni |A| := \sqrt{A^*A} \geq 0.$$

(iii) From Theorem 6.13(vii) we see that

$$\|(\lambda Id - A)^{-1}\| = (\operatorname{dist}(\lambda, \sigma(A)))^{-1}$$

if A is bounded, self-adjoint, and $\lambda \notin \sigma(A)$.

7 Unbounded operators

In this section X is a Hilbert space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

7.1 Domains, graphs, adjoints, and spectrum

Example 7.1 (i) Consider the linear mapping $(L, D(L))$ in $X = L^2([0, \pi])$ from Example 1.3(v) given by

$$D(L) := \{f \in C^2([0, \pi]) | f(0) = f(\pi) = 0\} \subset L^2([0, \pi])$$

and

$$L^2([0, \pi]) \ni Lf := f'', \quad f \in D(L).$$

Since the functions

$$f_n = \sin(n \cdot) \in D(L)$$

are eigenfunctions to the eigenvalues $-n^2$, $n \in \mathbb{N}$, $(L, D(L))$ is not bounded.
(ii) Let $X = L^2(\mathbb{R})$ and

$$D(T) := \Big\{ f \in L^2(\mathbb{R}) \Big| \int_{\mathbb{R}} x^2 \, |f(x)|^2 \, dx < \infty \Big\}.$$

We define for $f \in D(T)$

$$Tf(x) := xf(x), \quad x \in \mathbb{R}, \quad (\textbf{position operator}).$$

Obviously, $Tf \in L^2(\mathbb{R})$. By choosing indicator functions of intervals with measure 1 having a large distance to the origin, one easily shows that the operator $(T, D(T))$ is unbounded.

From now on we consider linear mappings

$$T : D(T) \to X$$

which might be well-defined only on a linear subset $D(T) \subset X$. To stress this we write $(T, D(T))$ and call $(T, D(T))$ an **operator** in X. If $(T, D(T))$ is not bounded, i.e., there does not exist $0 < C < \infty$ such that

$$\|Tx\| \leq C\|x\| \quad \text{for all} \quad x \in D(T),$$

then we call $(T, D(T))$ an **unbounded operator**.

Definition 7.2 *Let $(T, D(T))$ be an operator in X. Then we define the* **graph** *of $(T, D(T))$ by*

$$\Gamma_T := \{[x, y] \in X \times X \mid y = Tx,\ x \in D(T)\}.$$

The **graph norm** *corresponding to $(T, D(T))$ is defined by*

$$\|x\|_{\Gamma_T} := \sqrt{\|x\|^2 + \|Tx\|^2}, \quad x \in D(T).$$

$(T, D(T))$ is called a **closed** *operator, iff Γ_T is a closed subset of $X \times X$. Here $X \times X$ is equipped with the scalar product*

$$([x_1, y_1], [x_2, y_2])_{X\times X} := (x_1, x_2)_X + (y_1, y_2)_X, \quad [x_1, y_1], [x_2, y_2] \in X \times X. \tag{7.1}$$

Lemma 7.3 An operator $(T, D(T))$ in X is closed, iff $(D(T), \|\cdot\|_{\Gamma_T})$ is complete.

Proof: Let $(T, D(T))$ be closed and let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $D(T)$ w.r.t $\|\cdot\|_{\Gamma_T}$. Then $(x_n)_{n\in\mathbb{N}}$ and $(Tx_n)_{n\in\mathbb{N}}$ are Cauchy sequences in X. Hence there exists

$$x = \lim_{n\to\infty} x_n, \ y = \lim_{n\to\infty} Tx_n \in X.$$

Set $y_n := Tx_n$, $n \in \mathbb{N}$. Then $([x_n, y_n])_{n\in\mathbb{N}}$ is a sequence in Γ_T which converges to $[x, y]$ in $X \times X$. Since Γ_T is closed, we have $x \in D(T)$ and $y = Tx$. Thus $(x_n)_{n\in\mathbb{N}}$ converges to x in $D(T)$ w.r.t $\|\cdot\|_{\Gamma_T}$.

Let $(D(T), \|\cdot\|_{\Gamma_T})$ be complete and $([x_n, y_n])_{n\in\mathbb{N}}$ a sequence in Γ_T which converges to $[x, y]$ in $X \times X$. Then $y_n = Tx_n$, $n \in \mathbb{N}$. Hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $D(T)$ w.r.t $\|\cdot\|_{\Gamma_T}$. Because $(D(T), \|\cdot\|_{\Gamma_T})$ is complete

$$x = \lim_{n\to\infty} x_n \in D(T) \text{ and } Tx = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} y_n = y$$

in X. Thus Γ_T is closed. ■

Definition 7.4 *Let* $(T_1, D(T_1))$ *and* $(T_2, D(T_2))$ *be operators in* X*. The operator* $(T_2, D(T_2))$ *is called an* **extension** *of* $(T_1, D(T_1))$*, iff*

$$\Gamma_{T_1} \subset \Gamma_{T_2}.$$

Or, equivalently,

$$D(T_1) \subset D(T_2) \quad \textit{and} \quad T_2|_{D(T_1)} = T_1.$$

Definition 7.5 *An operator* $(T, D(T))$ *in* X *we call* **closable***, iff it has a closed extension. Every closable operator has a smallest closed extension (see the proof of Proposition 7.6 below), called its* **closure***, which we denote by* $(\overline{T}, D(\overline{T}))$*.*

Proposition 7.6 *If* $(T, D(T))$ *is a closable operator in* X*, then* $\Gamma_{\overline{T}} = \overline{\Gamma_T}$*.*

Proof: Suppose $(S, D(S))$ is a closed extension of $(T, D(T))$. Then $\overline{\Gamma_T} \subset \Gamma_S$. Hence, if

$$[0, y] \in \overline{\Gamma_T}, \quad \text{then} \quad y = 0. \tag{7.2}$$

Furthermore, since $\Gamma_T \subset X \times X$ is a linear subset, also $\overline{\Gamma_T} \subset X \times X$ is a linear subset. Hence on

$$D(R) := \{x \in X \mid [x, y] \in \overline{\Gamma_T} \text{ for some } y \in X\}$$

we can define the linear mapping

$$Rx := y \quad \text{where } [x, y] \in \overline{\Gamma_T},$$

which due to (7.2) together with the linearity of $\overline{\Gamma_T}$ is well-defined on $D(R)$. Then $\Gamma_R = \overline{\Gamma_T}$. Thus $(R, D(R))$ is a closed extension of $(T, D(T))$. But $\Gamma_R \subset \Gamma_S$, which is an arbitrary closed extension of $(T, D(T))$. Thus $\Gamma_R = \Gamma_{\overline{T}}$. ■

Definition 7.7 *Let* $(T, D(T))$ *be a* **densely defined** *operator in* X *(i.e.,* $D(T) \subset X$ *is dense). Let* $D(T^*)$ *be the set of all elements* y *from* X *for which there exists* $z \in X$ *such that*

$$(Tx, y) = (x, z) \quad \textit{for all} \quad x \in D(T).$$

Since $D(T) \subset X$ is dense, this z is unique. Hence for each such $y \in D(T^*)$ we can define

$$T^*y := z.$$

$(T^*, D(T^*))$ is called the **(Hilbert space) adjoint** *of $(T, D(T))$. Obviously, $T^* : D(T^*) \to X$ is linear.*

Lemma 7.8 Let $(T, D(T))$ be a densely defined operator in X. Then $y \in D(T^*)$, iff there exists $0 \le C < \infty$ such that

$$|(Tx, y)| \le C\|x\| \quad \text{for all} \quad x \in D(T).$$

Proof: Let $y \in D(T^*)$, then by the Cauchy–Schwartz inequality

$$|(Tx, y)| = |(x, T^*y)| \le \|x\|\|T^*y\| \quad \text{for all} \quad x \in D(T).$$

Vice versa. Suppose there exist $0 \le C < \infty$ such that

$$|(Tx, y)| \le C\|x\| \quad \text{for all} \quad x \in D(T).$$

Then the mapping

$$D(T) \ni x \mapsto (Tx, y) \in \mathbb{K}$$

is linear and continuous. Since $D(T) \subset X$ is dense, it can be extended uniquely to a linear continuous mapping $F : X \to \mathbb{K}$. Hence by the Riesz representation theorem there exists a unique $z \in X$ such that

$$F(x) = (x, z) \quad \text{for all} \quad x \in X.$$

In particular

$$(Tx, y) = F(x) = (x, z) \quad \text{for all} \quad x \in D(T).$$

■

Theorem 7.9 *Let $(T, D(T))$ be a densely defined operator on X. Then:*
(i): $(T^*, D(T^*))$ is closed.
(ii): $(T, D(T))$ is closable, iff $D(T^*) \subset X$ is dense in which case $\Gamma_{\overline{T}} = \Gamma_{T^{**}}$.
(iii): If $(T, D(T)$ is closable, then $\Gamma_{(\overline{T})^*} = \Gamma_{T^*}$.

Proof: (i): We define the operator V on $X \times X$ by

$$V[x, y] = [-y, x], \quad [x, y] \in X \times X.$$

First note that

$$V(E^\perp) = V(E)^\perp \quad \text{for all subspaces} \quad E \subset X \times X.$$

Furthermore $[x, y] \in V(\Gamma_T)^\perp$, iff

$$([x, y], [-Tz, z])_{X \times X} = 0 \quad \text{for all} \quad z \in D(T).$$

That is equivalent to

$$(y, z)_X = (x, Tz)_X \quad \text{for all} \quad z \in D(T).$$

This in turn holds, iff $[x, y] \in \Gamma_{T^*}$. Thus

$$\Gamma_{T^*} = V(\Gamma_T)^\perp. \tag{7.3}$$

Since $V(\Gamma_T)^\perp \subset X \times X$ is closed, this proves (i).

(ii): Since $\Gamma_T \subset X \times X$ is a linear subset we have by using (7.3)

$$\overline{\Gamma_T} = ((\Gamma_T)^\perp)^\perp = ((V^2(\Gamma_T))^\perp)^\perp = (V((V(\Gamma_T))^\perp))^\perp = (V(\Gamma_{T^*}))^\perp.$$

Thus, by (7.3), if $D(T^*) \subset X$ is dense, then $\overline{\Gamma_T}$ is the graph of $(T^{**}, D(T^{**}))$. Hence, in this case $(T, D(T))$ is closable and $\Gamma_{\overline{T}} = \Gamma_{T^{**}}$.

Conversely, suppose that $D(T^*) \subset X$ is not dense and that $0 \neq z \in D(T^*)^\perp$. Then

$$[z, 0] \in (\Gamma_{T^*})^\perp$$

and therefore

$$[0, z] = V[z, 0] \in V((\Gamma_{T^*})^\perp) = (V(\Gamma_{T^*}))^\perp.$$

Hence

$$\overline{\Gamma_T} = (V(\Gamma_{T^*}))^\perp$$

can not be the graph of a linear mapping. Thus, by Proposition 7.6, $(T, D(T))$ is not closable.

(iii): Notice that if T is closable, then by (i) and (ii)

$$\Gamma_{T^*} = \Gamma_{\overline{T^*}} = \Gamma_{T^{***}} = \Gamma_{(\overline{T})^*}.$$

■

Definition 7.10 *Let $(T, D(T))$ be an operator in X. A $\lambda \in \mathbb{C}$ is in the* **resolvent set** *of $(T, D(T))$, $\rho(T)$, iff:*
(i) $\lambda Id - T : D(T) \to X$ is injective,
(ii) $\lambda Id - T : D(T) \to X$ is surjective,
(iii) $R(\lambda; T) := (\lambda Id - T)^{-1} \in L(X)$.

If $\lambda \in \rho(T)$, then $R(\lambda; T)$ is called the **resolvent** *of $(T, D(T))$ at λ. The* **spectrum**, **point spectrum**, *and* **residual spectrum** *are the same for unbounded operators as they are for bounded operators, see Definition 1.7.*

Theorem 7.11 *Let $(T, D(T))$ be an operator in X. Then $\rho(T) \subset \mathbb{K}$ is open and the resolvent function $R(\cdot; T)$ is a $\mathbb{K}$-analytic mapping from $\rho(T)$ to $L(X)$. Furthermore, for any two points $\lambda, \mu \in \rho(T)$, $R(\lambda; T)$ and $R(\mu; T)$ commute and*

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

Proof: The same as in the case of $T \in L(X)$, see Proposition 1.9 and Proposition 6.7. ■

7.2 Symmetric and self-adjoint operators

Definition 7.12 *A densely defined operator $(T, D(T))$ in X is called* **symmetric** *(or* **Hermitian***), iff $\Gamma_T \subset \Gamma_{T^*}$. Or, equivalently,*

$$(Tx, y) = (x, Ty) \quad \textit{for all} \quad x, y \in D(T).$$

Example 7.13 (i) In Example 1.3(v) we have already shown that the operator $(L, D(L))$ in $X = L^2([0, \pi])$ given by

$$D(L) = \{f \in C^2([0, \pi]) | f(0) = f(\pi) = 0\} \subset L^2([0, \pi])$$

and

$$L^2([0, \pi]) \ni Lf = f'', \quad f \in D(L),$$

is symmetric.
(ii) Consider the position operator $(T, D(T))$ in $L^2(\mathbb{R})$ from Example 7.1(ii). I.e.,

$$D(T) = \left\{ f \in L^2(\mathbb{R}) \,\middle|\, \int_{\mathbb{R}} x^2 \, |f(x)|^2 \, dx < \infty \right\}$$

and

$$Tf(x) = xf(x), \quad x \in \mathbb{R}, \quad f \in D(T).$$

$(T, D(T))$ is densely defined, since the indicator functions of bounded measurable sets, which are dense in $L^2(\mathbb{R})$, are contained in $D(T)$. Furthermore, $(T, D(T))$ is symmetric, because

$$\begin{aligned}(Tf, f) &= \int_{\mathbb{R}} Tf(x)\overline{f(x)}\,dx = \int_{\mathbb{R}} xf(x)\overline{f(x)}\,dx \\ &= \int_{\mathbb{R}} f(x)\overline{xf(x)}\,dx = (f, Tf) \quad \text{for all} \quad f \in D(T).\end{aligned}$$

From Theorem 7.9(i) we can conclude that $(T^*, D(T^*))$ is closed and therefore a closed extension of $(T, D(T))$. Hence, $(T, D(T))$ is closable. But we have even that $\Gamma_T = \Gamma_{T^*}$. Indeed, let $f \in D(T^*)$, then

$$\begin{aligned}\int_{\mathbb{R}} g(x)\overline{(T^*f)(x)}\,dx &= (g, T^*f) = (Tg, f) \\ &= \int_{\mathbb{R}} xg(x)\overline{f(x)}\,dx = \int_{\mathbb{R}} g(x)\overline{xf(x)}\,dx \quad \text{for all} \quad g \in D(T).\end{aligned}$$

Thus $T^*f(x) = xf(x)$ for dx-almost all $x \in \mathbb{R}$. Since $T^*f \in L^2(\mathbb{R})$, we have

$$\int_{\mathbb{R}} x^2|f(x)|^2\,dx < \infty,$$

i.e., $f \in D(T)$. Hence $(T, D(T))$ is self-adjoint in the sense of the following definition.

Definition 7.14 *A densely defined operator* $(T, D(T))$ *in* X *is called* **self-adjoint***, iff* $\Gamma_T = \Gamma_{T^*}$.

A symmetric operator is always closable, since $D(T^*) \supset D(T)$ is dense in X, see Theorem 7.9(ii).

If $(T, D(T))$ is symmetric, $(T^*, D(T^*))$ is a closed extension of $(T, D(T))$, see Theorem 7.9(i), so the smallest closed extension $(T^{**}, D(T^{**}))$, see Theorem 7.9(ii), must be contained in $(T^*, D(T^*))$.Thus for symmetric operators we have

$$\Gamma_T \subset \Gamma_{T^{**}} \subset \Gamma_{T^*}.$$

For closed symmetric operators

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

And, for self-adjoint operators

$$\Gamma_T = \Gamma_{T^{**}} = \Gamma_{T^*}.$$

Hence a closed symmetric operator $(T, D(T))$ is self-adjoint, iff $(T^*, D(T^*))$ is symmetric.

Definition 7.15 *A symmetric operator* $(T, D(T))$ *in* X *is called* **essentially self-adjoint**, *iff its closure* $(\overline{T}, D(\overline{T}))$ *is self-adjoint. If* $(T, D(T))$ *is self-adjoint, a subset* $D \subset D(T)$ *is called a* **core** *for* $(T, D(T))$ *iff* $\Gamma_{\overline{T|_D}} = \Gamma_T$.

Theorem 7.16 (the basic criterion for self-adjointness) *Let* $(T, D(T))$ *be a symmetric operator in a complex Hilbert space* X*. Then the following statements are equivalent:*
(i) $(T, D(T))$ *is self-adjoint.*
(ii) $(T, D(T))$ *is closed and* $\mathcal{N}(T^* \pm iId) = \{0\}$.
(iii) $\mathcal{R}(T \pm iId) = X$.

Proof: (i) implies (ii): A self-adjoint operator $(T, D(T))$ is always closed, because $(T^*, D(T^*))$ is closed by Theorem 7.9(i) and $\Gamma_T = \Gamma_{T^*}$.

Suppose $x \in D(T^*) = D(T)$ fulfills $T^*x = ix$. Then $Tx = ix$ and

$$i(x, x) = (ix, x) = (Tx, x) = (x, T^*x) = (x, Tx) = (x, ix) = -i(x, x).$$

Thus $x = 0$. A similar argument shows that $T^*x = -ix$ can hold only for $x = 0$.

(ii) implies (iii): Since $T^*x = -ix$ implies $x = 0$, $\mathcal{R}(T - iId)$ must be dense in X. Indeed, if

$$x \in \mathcal{R}(T - iId)^\perp,$$

then we have

$$((T - iId)y, x) = 0 \quad \text{for all} \quad y \in D(T).$$

Hence $x \in D(T^*)$ and

$$0 = (T - iId)^*x = T^*x + ix.$$

Thus $x = 0$ by (ii). Now we only have to show that $\mathcal{R}(T - iId)$ is closed to conclude that $\mathcal{R}(T - iId) = X$. But this follows from

$$\|(T - iId)x\|^2 = (Tx - ix, Tx - ix) = \|Tx\|^2 + \|x\|^2, \quad x \in D(T).$$

Indeed, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in $D(T)$ such that

$$\lim_{n\to\infty} (T - iId)x_n = z.$$

Then there exist $x, y \in X$ such that

$$\lim_{n\to\infty} x_n = x \quad \text{and} \quad \lim_{n\to\infty} Tx_n = y.$$

Since $(T, D(T))$ is closed, $x \in D(T)$ and $Tx = y$. Hence

$$z = \lim_{n\to\infty} (T - iId)x_n = (T - iId)x \in \mathcal{R}(T - iId).$$

Similarly one shows that $\mathcal{R}(T + iId) = X$.

(iii) implies (i): Let $x \in D(T^*)$. Since $\mathcal{R}(T - iId) = X$, there exists $y \in D(T)$ such that

$$(T - iId)y = (T^* - iId)x.$$

Since $\Gamma_T \subset \Gamma_{T^*}$, we have $x - y \in D(T^*)$ and

$$((T^* - iId)(x - y) = 0.$$

Since $\mathcal{R}(T + iId) = X$, we have $\mathcal{N}(T^* - iId) = \{0\}$. Thus $x = y \in D(T)$. This proves that $D(T^*) = D(T)$. Hence $(T, D(T))$ is self-adjoint. ■

Corollary 7.17 *Let $(T, D(T))$ be a symmetric operator in a complex Hilbert space X. Then the following statements are equivalent:*
(i) $(T, D(T))$ is essentially self-adjoint.
(ii) $\mathcal{N}(T^* \pm iId) = \{0\}$.
(iii) $\mathcal{R}(T \pm iId)$ are dense in X.

Proof: Follows from a careful analysis of the proof of Theorem 7.16. ■

References

[Alt02] H.W. Alt. *Lineare Funktionalanalysis*. Springer-Verlag, Berlin, 2002.

[RS75] M. Reed and B. Simon. *Methods of modern mathematical physics*, volume I. Academic Press, New York, London, 1975.