Functional Analysis - The irreducible Minimum

Contents

1 Structures

1.1 Topological spaces

We will deal with various structures, in a sense, they build a certain hierachy. Topological spaces as a framework, then metric spaces, normed spaces, Banach spaces and conclusively Hilbert Spaces.

Definition 1.1. *A topological space is a pair (X, τ), consisting of a set X and a family $\tau \subset 2^X$ that we call a topology on X, such that*

1. $\emptyset, X \in \tau$.
2. *Stable under countable intersection of open sets.*
3. *Stable under arbitrary union of open sets.*

Remark 1.1. Thanks to the laws of De Morgan we easily conclude from the definition that arbitrary intersections and finite unions of closed sets are again closed.

Definition 1.2. *Let (X, τ) be a topological space and $A \subset X$ a subset. The closure $\overline{A}$ of A is defined by*

$$\overline{A} = \bigcap \{B \subset X : B \text{ is closed and } A \subset B\}$$

by definition, it is the smallest closed set that contains A. The set A is called dense in X if $\overline{A} = X$. The space X is called separable, if it contains a countable dense set. Moreover, the interior of A is defined through

$$A^\circ = \bigcup \{B \subset A : B \text{ is open}\}$$

in other words, A° is the largest open set contained in A. Finally, the boundary of A is defined as $\partial A = \overline{A} \setminus A^\circ$.

Remark 1.2. Evidently, the closure and interior of a set depend on the choice of the topology.

Definition 1.3. *Let (X, τ) be a topological space, $x \in X$. A set $U \subset X$ is called an open neighbourhood of x if $U \in \tau$ (i.e. U is open) and $x \in U$.*

Definition 1.4. *Let (X, τ) be a topological space and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. We say that x_n converges to $x \in X$ as $n \to \infty$, written $x_n \to X$, if*

For every open neighbourhood U of x, there exists $n_0 \in \mathbb{N} : x_n \in U \ \forall n \geq n_0$.

Definition 1.5. *Let $(X, \tau), (Y, \mathcal{S})$ be two topological spaces. A function $f : X \to Y$ is called continuous if*

$$f^{-1}(V) \in \tau, \ \forall V \in \mathcal{S}.$$

I.e., if the pre-image of every open set in Y is an open set in X.

Remark 1.3. In metric spaces, the notion of convergence completely characterizes the topology. This is however not necessarily true in topological spaces.

Definition 1.6. *A topological space (X, τ) is called Hausdorff if*

$$\forall x, y \in X, x \neq y \implies \exists U_x, U_y \in \tau \text{ with } x \in U_x, y \in U_y \text{ and } U_x \cap U_y = \emptyset$$

Definition 1.7. *A topological space (X, τ) is called compact if it is Hausdorff and if for every*

$$(U_\lambda)_{\lambda\in\Lambda} \text{ family in } \tau \text{ with } \bigcup_{\lambda\in\Lambda} U_\lambda = X$$
$$\implies \exists n \in \mathbb{N} \text{ and } \lambda_1, \ldots, \lambda_n \in \Lambda : \bigcup_{j=1}^{n} U_{\lambda_j} = X$$

i.e. if for every open covering of X there exists a finite sub-covering.

Theorem 1.1. *Let (X, τ) be a compact space and $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Then there exists at least one accumulation point of x_n on X.*

Theorem 1.2 (Lemma of Urysohn)**.** *Let (X, τ) be a compact space and $A, B \subset X$ disjoint, non-empty, closed subsets of X. Then there exists a continuous function $g : X \to [0, 1]$ with $g(A) = 0$ and $g(B) = 1$.*

Remark 1.4. The Lemma of Urysohn is important because it helps us avoid some of the pathological situations (like only constant functions being continuous). In particular, the theorem implies that on compact spaces the topology is large enough for us to do some meaningful analysis on.

Another way to think about is that the Hausdorff property in a topological space (X, τ) guarantees that there are 'enough' open disjoint sets for us to work with, coupled with the Lemma of Urysohn we know that in compact spaces there are enough closed subsets which are disjoint and we can define 'meaningful' functions between them (i.e. non constant).

Definition 1.8. *Let K be a compact space. We define*

$$C_{\mathbb{K}}(K) := \{f : K \to \mathbb{K} \text{ continuous}\}$$

As a consequence of Urysohn's Lemma we can show that $C_{\mathbb{K}}(K)$ separates the points of K.

Corollary 1.1. *Let K be a compact space. Then $C_{\mathbb{K}}(K)$ separates the points of K. In other words, for every $x, y \in K$ with $x \neq y$, there exists $f \in C_{\mathbb{K}}(K)$ such that $f(x) \neq f(y)$.*

Proof. Since K is Hausdorff, for every $x \neq y$ we can find two open neighbourhoods U_x, U_y of x and y respectively with $U_x \cap U_y = \emptyset$. We can further find closed sets A, B with $A \subset U_x, B \subset U_y$ with $x \in A$ and $y \in B$. In particular we have $A \cap B = \emptyset$. Urysohn's Lemma then gives that there exists $f \in C_{\mathbb{K}}(K)$ with $f(x) = 0$ and $f(y) = 1$. □

1.2 Metric Spaces

Definition 1.9. *A metric space is a pair (X, d) with X being an arbitrary set and a map $d : X \times X \to [0, \infty)$ called a metric on X, with the following properties*

1. *$d(x, y) = 0$ if and only if $x = y$.*
2. *$d(x, y) = d(y, x)$.*
3. *(Δ-inequality) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.*

Remark 1.5.

1. Every metric space (X, d) is a topological space (X, τ_d) with topology τ_d inducedby the metric. The topology τ_d is defined by the condition that $A \in \tau_d$ if and only if for all $x \in A$, there exists $\epsilon > 0$ with $B_\epsilon(x) = \{y \in X : d(x, y) < \epsilon\} \subset A$.

2. In contrast with general topological spaces, on metric spaces the notion of convergence characterizes the topology. That is we have the convenient characterizations

   (a) A set $A \subset X$ is closed if and only if for every sequence x_n in A with $x_n \to x$ in X, we have that $x \in A$.

   (b) $\overline{A} = \{x \in X : \exists (x_n)_{n \in \mathbb{N}} \text{ sequence in } A \text{ with } x_n \to x\}$.

   (c) $A \subset X$ is dense if and only if $\forall x \in X, \exists (x_n)_{n \in \mathbb{N}}$ sequence in A with $x_n \to x$.

   (d) The function $f : X \to Y$ between the two metric spaces $(X, d_1), (Y, d_2)$ is continuous at the point $x \in X$ if and only if for every sequence $(x_n)_{n \in \mathbb{N}}$ in X with $x_n \to x$ we have $f(x_n) \to f(x)$.

3. It is clear that all topologies taht are induced by a metric are Hausdorff. Hence, non-Hausdorff topologies are never metrizable. Example: Zariski Topology.

Definition 1.10. *Let (X, d) be a metric space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is called a Cauchy sequence (or is said to have the Cauchy property) if $d(x_n, x_m) \to 0$ as $n, m \to \infty$. Every convergent sequence is indeed Cauchy. The metric space (X, d) is called complete, if every Cauchy sequence is convergent.*

1.3 Normed spaces

Definition 1.11. *A normed space is a pair $(X, \|\cdot\|)$, consisting of a vector space V over a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) and a map $\|\cdot\| : V \to [0, \infty)$ called a norm on V with the following properties*

1. *$\|x\| = 0$ if and only if $x = 0$.*
2. *(Homogenity) $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in \mathbb{K}, x \in V$.*
3. *(Δ-inequality) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$.*

Remark 1.6. Every norm induces the metric $d(x, y) = \|x-y\|$, hence every normed space is also a metric space and therefore also a topological space.

Definition 1.12. *A normed space $(X, \|\cdot\|)$ is called complete, if X, equipped with the induced metric $d(x, y) = \|x - y\|$, is a complete metric space. A complete normed space is called a Banach space.*

Completeness is very important for analysis. It is not a coincidence that we always do analysis on $\mathbb{R}$ instead of $\mathbb{Q}$. For this reason, it is useful to have a general recipe to complete normed spaces.

Definition 1.13. *Let $(X, \|\cdot\|)$ be a normed space. A completion of $(X, \|\cdot\|)$ is a 3-tuple $(Y, \|\cdot\|_Y, \phi)$ consisting of a Banach space $(Y, \|\cdot\|_Y)$ and an isometric linear map $\phi : X \to Y$, with $\overline{\phi(X)} = Y$.*

Theorem 1.3. *Every normed space $(X, \|\cdot\|)$ has a completion, which is unique, up to linear isometric isomorphisms.*

Proof. The proof is constructive, we give a sketch. Let $\mathcal{C}_X$ denote the set of all Cauchy sequence on X. We can easily give this space the structure of a vector space over $\mathbb{K}$. Next we define the linear subspace $\mathcal{N}_X \subset \mathcal{C}_X$ consisting of all null-sequence on X, i.e.

$$\mathcal{N}_X := \{x = (x_n)_{n\in\mathbb{N}} \in \mathcal{C}_X : x_n \to 0\}$$

Moreover we define $Y := \mathcal{C}_X/\mathcal{N}_X$ as the quotient space of $\mathcal{C}_X$ w.r.t. the equivalence relation defined by $x \sim y :\iff x - y \in \mathcal{N}_X$. In other words, in Y, we identify Cauchy sequences whose difference converges to zero. Y is also a vector space over $\mathbb{K}$.

Next we want to introduce a norm on Y. To this end, we define the function $p : \mathcal{C}_X \to [0, \infty)$ through

$$p(x) = \lim_{n\to\infty} \|x_n\|$$

Thanks to the reverse triangle inequality we've got

$$|\|x_k\| - \|x_l\|| \leq \|x_k - x_l\| \to 0 \text{ as } k, l \to \infty$$

which shows that $(\|x_n\|)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$ whenever $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{C}_X$. Thus the limit above is well-defined and finite. We now set

$$\|[x]\|_Y := p(x) = \lim_{n\to\infty} \|x_n\|$$

We then can verify that $\|\cdot\|_Y$ is indeed a norm on Y.

Next we define the map $\phi : X \to Y$ by $\phi(z) = [(z, z, \dots)]$, i.e. $\phi(z)$ denotes the equivalence class of all sequences on X that converge to z in the limit. This map is clearly linear and since

$$\|\phi(z)\|_Y = \|z\|_X$$

it defines an isometry. We now claim that $(Y, \|\cdot\|_Y, \phi)$ is a completion of $(X, \|\cdot\|_X)$. In order to show that we can show that $(Y, \|\cdot\|_Y)$ is always complete and that $\phi(X)$ is dense in Y. That is we want to show that for all $[x] \in Y$, we can find $\tilde{x} \in X$ with $\|\phi(\tilde{x}) - [x]\|_Y < \epsilon$. We start with the density and then use this to show that the space is complete.

Finally we can show that the uniqueness of the completion is up to isometric isomorphisms. □

Remark 1.7. The completion of $\mathbb{Q}$ is $\mathbb{R}$. For $\Omega \subset \mathbb{R}^n$, the completion of the space $C(\overline{\Omega})$ of continuous functions on $\overline{\Omega}$ is $L^p(\Omega)$.

1.4 Hilbert Spaces

Definition 1.14. *Let H be a vector space over the field $\mathbb{K}$. A scalar product (or an inner product) on H is a map $(\cdot,\cdot) : H \times H \to \mathbb{K}$ with the properties*

1. $(z, x + \lambda y) = (z, x) + \lambda(z, y)$ *for all* $x, y, z \in H, \lambda \in \mathbb{K}$.
2. $(x, y) = \overline{(y, x)}$.
3. $(x, x) > 0$ *for all* $x \neq 0$.

A pair $(H, (\cdot,\cdot))$ consisting of a vector space H over $\mathbb{K}$ and a scalar product $(\cdot,\cdot)$ is called a pre-Hilbert space.

Remark 1.8. We defiend the scalar product to be linear in its second argument and anti-linear in its first argument.

Lemma 1.1 (Cauchy Schwarz Inequality)**.** *Let $(H, (\cdot, \cdot))$ be a pre-Hilbert space. Then*

$$|(x, y)|^2 \leq (x, x)(y, y).$$

Remark 1.9. The Cauchy-Schwarz inequality allows us to use the scalar product to define a norm on every pre-Hilbert space.

Corollary 1.2. *Let $(H, (\cdot, \cdot))$ be a pre-Hilbert space. Then*

$$\|x\| := \sqrt{(x, x)}$$

defines a norm on H.

Remark 1.10. The triangle inequality follows from the Cauchy-Schwarz inequality. The remaining properties follow from the properties of scalar products.

Definition 1.15. *A pre-Hilbert space is called a Hilbert space if H, equipped with the norm $\|x\| = \sqrt{(x, x)}$ induced by the scalar product, is a Banach space (i.e. if H is complete).*

Remark 1.11.

1. Every pre-Hilbert space can be completed into a Hilbert space.

2. Every Hilbert space is a metric space and therefore a topological space. However, clearly, not every Banach space is a Hilbert space, simply because not all norms on a vector space can be induced by a scalar product.

Theorem 1.4. *Let $(H, (\cdot, \cdot))$ be a Hilbert space, $K \subset H$ a closed convex set in H and $x_0 \in H$. Then there exists a unique $y \in K$ such that*

$$\|x_0 - y\| = \mathit{dist}(x_0, K) := \inf_{x \in K} \|x_0 - x\|$$

As an application of this theorem, we show that every Hilbert space H can be decomposed in the direct sum of an arbitrary closed subspace and of its orthogonal complement.

Theorem 1.5. *Let $(H, (\cdot, \cdot))$ be a Hilbert space and $M \subset H$ a closed linear subspace. Then the orthogonal complement $M^\perp$ of M, defined through*

$$M^\perp := \{x \in H : (x, m) = 0, \text{ for all } m \in M\}$$

is also a linear closed subspace of H and $H = M \oplus M^\perp$, meaning that $H = M + M^\perp$ and $M \cap M^\perp = \{0\}$.

Proof. Clearly $M^\perp$ is linear. In order to see that it is closed take x_n to be a sequence in $M^\perp$ such that $x_n \to x$ in H. Then we have

$$(x, m) = \lim_{n\to\infty} (x_n, m) = 0$$

because thanks to the Cauchy-Schwarz inequality we have

$$|(x - x_n, m)| \leq \|x - x_n\| \|m\| \to 0 \text{ as } n \to \infty$$

in particular $|(x - x_n, m)| = |(x, m) - (x_n, m)| \to 0$ and it follows that $x \in M^\perp$, i.e. $M^\perp$ is closed.

Moreover, the fact that $M \cap M^\perp = \{0\}$ follows, because $(x, x) = 0$ implies that $x = 0$. Thus it only rmeains to show that $M + M^\perp = H$. To this end, we fix $x \in H$. Since $M \subset H$ is a closed linear subspace (and therefore in particular convex) we can apply the previous theorem to find $z \in M$ such that $\operatorname{dist}(x, M) = \|x - z\|$.

We now claim that $x - z \in M^\perp$, which gives that $x = z + (x - z)$ is the desired decomposition of H into M and $M^\perp$. Lets assume for contradiction that $(x - z) \notin M^\perp$. Then there exists $\alpha \in M$ with $(x - z, \alpha) > 0$. For $t \in [-1, 1]$ let $z_t = z + t\alpha$. Then we have $z_t \in M$ for all t and

$$\|x - z_t\|^2 = \|x - z\|^2 + t^2\|\alpha\|^2 - 2t(x - z, \alpha) < \|x - z\|^2 = \operatorname{dist}(x, M)$$

for $t > 0$ small enough. But this condradicts the definition of $\operatorname{dist}(x, M)$. $\square$

Definition 1.16. *An orthonormal system in $(H, (\cdot, \cdot))$ is a family $(x_\alpha)_{\alpha\in A} \subset H$ for an arbitrary index-set A with $(x_\alpha, x_\beta) = \delta_{\alpha,\beta}$. In the case when $A = \mathbb{N}$, we also call the orthonormal system an orthonormal sequence.*

Lemma 1.2. *Let H be a Hilbert space, $(x_n)_{n\in\mathbb{N}}$ an orthonormal system (orthonormal sequence) and $(\alpha_n)_{n\in\mathbb{N}}$ a sequence in $\mathbb{K}$. Then we have*

1. *$\sum_{k=1}^{\infty} \alpha_k x_k$ converges if and only if $\sum_{k=1}^{\infty} |\alpha_k|^2$ converges.*
2. *$\| \sum_{k=1}^{n} \alpha_k x_k\|^2 = \sum_{k=1}^{n} |\alpha_k|^2$.*
3. *If $\sum_{k=1}^{\infty} \alpha_k x_k$ converges, then the limit is independent of the order of the terms.*

Lemma 1.3. *Let $(H, (\cdot, \cdot))$ be a Hilbert space, A an arbitrary set an $(x_\alpha)_{\alpha\in A}$ an orthonormal system in H. Then $\sum_{\alpha\in A}(x_\alpha, x)x_\alpha$ converges for every $x \in H$. Moreover, the linear map $\phi : H \to H$ defined through $\phi(x) = \sum_{\alpha\in A}(x_\alpha, x)x_\alpha$ is the continuous projection onto*

$$M := \overline{\mathit{span}\{x_\alpha : \alpha \in A\}} \text{ along its orthogonal complement } M^\perp$$

In particular for $x \in M$, we find that $x = \sum_{\alpha\in A}(x_\alpha, x)x_\alpha$.

Remark 1.12. In particular, if $M = \text{span}\{x_\alpha : \alpha \in A\}$ is dense in H, i.e. if $\overline{M} = H$, then gives us the previous Lemma a representation for every vector $x \in H$. In this case, we say that $(x_\alpha)_{\alpha\in A}$ is a Hilbert space basis.

Definition 1.17. *Let H be a Hilbert space. A Hilbert space basis is an orthonormal system $(x_\alpha)_{\alpha\in A}$ with*

$$\overline{\mathit{span}\{x_\alpha : \alpha \in A\}} = H$$

Theorem 1.6 (Characterizations of Hilbert space bases)**.** *Let H be a Hilbert space, and $(x_\alpha)_{\alpha\in A}$ an orthonormal system. Then the following statements are equivalent:*

1. *$(x_\alpha)_{\alpha\in A}$ is a Hilbert space basis.*
2. *$x = \sum_{\alpha\in A}(x_\alpha, x)x_\alpha$, for all $x \in H$.*
3. *$\|x\|^2 = \sum_\alpha |(x_\alpha, x)|^2$ for all $x \in H$.*
4. *$(x_\alpha, x) = 0$ for all $\alpha \in A$ implies that $x = 0$.*
5. *$(x_\alpha)_{\alpha\in A}$ is a maximal orthonormal system in the sense of inclusions.*

Remark 1.13. Using the maximality property, i.e. point 5 in the theorem above, it follows easily from the Lemma of Zorn that every pre-Hilbert space has a Hilbert space basis. In particular, it follows that every separable Hilbert space admits a countable orthonormal basis. The observation that separable Hilbert spaces have countable orthonormal bases can be used to identify separable Hilbert spaces with the sequence space $\ell^2(\mathbb{K})$.

Theorem 1.7. *Let H be an infinite dimensional separable Hilbert space over $\mathbb{K}$. Then there exists a linear Isomorphism $\phi : H \to \ell^2(\mathbb{K})$ with*

$$(\phi(x), \phi(y))_{\ell^2} = (x, y)_H$$

for all $x, y \in H$. In particular the isomorphism is isometric.

2 Function Spaces

2.1 Continuous Functions on Compact Spaces

Lemma 2.1. *Let $f \in C_{\mathbb{K}}(K)$ where K is compact. Then f is bounded and its supremum and infimum are attained.*

This allows us to define the sup/max norm of f on the space $C_{\mathbb{K}}(K)$.

Definition 2.1. *For $f \in C_{\mathbb{K}}(K)$ let*

$$\|f\| := \sup_{x\in K} |f(x)| = \max_{x\in K} |f(x)|$$

It is then simple to check that $\|\cdot\|$ defines a norm on $C_{\mathbb{K}}(K)$ (well-defined because of the above Lemma). Hence, the pair $(C_{\mathbb{K}}(K), \|\cdot\|)$ is a normed space.

Theorem 2.1. *$(C_{\mathbb{K}}(K), \|\cdot\|)$ is a Banach space.*

Proof. Let f_n be a Cauchy sequence in $C_{\mathbb{K}}(K)$. Then, trivially we have for any $x \in K$ that

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| \to 0 \text{ as } n, m \to \infty$$

It follows that $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy Sequence in $\mathbb{K}$, but since $\mathbb{K}$ is complete we conclude that $f_n(x)$ converges. Let $f(x) := \lim f_n(x)$ denote its limit. We easily note that

$$|f(x)| \leq \limsup_{n\to\infty} |f_n(x)| \leq \limsup_{n\to\infty} \|f_n\|.$$

Moreover, since for every $x \in K$, we have

$$|f(x) - f_n(x)| = \lim_{m\to\infty} |f_m(x) - f_n(x)| \leq \limsup_{m\to\infty} \|f_n - f_m\|$$

we conclude that

$$\sup_{x\in K} |f(x) - f_n(x)| \leq \limsup_{m\to\infty} \|f_m - f_n\| \to 0 \text{ as } n \to \infty.$$

since this convergences in uniform over all $x \in K$, we conclude that f is indeed continuous, thus our Cauchy sequence f_n converges to $f \in C_{\mathbb{K}}(K)$ and our space is complete. $\square$

Definition 2.2. *$\mathbb{A} \subset C_{\mathbb{K}}(K)$ is called a subalgebra if $\mathbb{A}$ is a linear subspace of $C_{\mathbb{K}}(K)$ and if, for every $f, g \in \mathbb{A}$, we also have $f \cdot g \in \mathbb{A}$. We say that the subalgebra $\mathbb{A}$ separates the points of K, if*

$$\forall x, y \in K \text{ with } x \neq y, \exists f \in \mathbb{A} : f(x) \neq f(y).$$

Remark 2.1. We already know that if $\mathbb{A} = C_{\mathbb{K}}(K)$ then indeed $C_{\mathbb{K}}(K)$ is a subalgebra separating the points of K.

Theorem 2.2 (Stone-Weierstrass Theorem, $\mathbb{K} = \mathbb{R}$)**.** *Let $\mathbb{A}$ be a subalgebra of $C_{\mathbb{R}}(K)$ separating the points of K. Then we have either $\overline{\mathbb{A}} = C_{\mathbb{R}}(K)$ or there exists a unique point $x_0 \in K$ such that $\overline{\mathbb{A}} = \{f \in C_{\mathbb{R}}(K) : f(x_0) = 0\}$.*

Proof. We first assume that for all $x \in K$ there exists $f \in \mathbb{A}$ with $f(x) \neq 0$, under this assumption we then want to show that $\overline{\mathbb{A}} = C_{\mathbb{R}}(K)$. To establish this result we require 6 steps.

- Step 1: Let $x_1, x_2 \in \mathbb{K}$ with $x_1 \neq x_2$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ then we have that that there exists $f \in \mathbb{A}$ such that $f(x_1) = \alpha_1$ and $f(x_2) = \alpha_2$.
  - This can be shown by using that $\mathbb{A}$ separates the points of K.
- Step 2: $\overline{\mathbb{A}}$ is a subalgebra. (Was an Exercise)
- Step 3: $f \in \mathbb{A} \implies |f| \in \overline{\mathbb{A}}$
  - W.l.o.g. we can assume that f is uniformly bounded by 1. We then write $|f| = \sqrt{f^2}$ and approximate $\sqrt{s}$ uniformly on $[0, 1]$ through polynomials. With such a sequence p_n the claim follows because $f \in \mathbb{A} \implies p_n(f^2) \in \mathbb{A}$ for all n and because $p_n(f^2) \to |f|$ and since $\overline{\mathbb{A}}$ is closed we conclude that $|f| \in \overline{\mathbb{A}}$.
- Step 4: For $f, g \in \overline{\mathbb{A}}$ we have $\min\{f, g\}, \max\{f, g\} \in \overline{\mathbb{A}}$.
  - Trivial because $\min\{f, g\} = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|$ and identity for max is just $\max\{f, g\} = \frac{1}{2}(f + g) + \frac{1}{2}|f + g|$.
- Step 5: Let $g \in C_{\mathbb{R}}(K)$ be arbitrary and $\epsilon > 0$. Then $\forall x \in K$ there exists $f_x \in \overline{\mathbb{A}}$ with $f_x(x) = g(x)$ and $f_x(y) \leq g(y) + \epsilon$ for all $y \in K$.
  - By step 1 we find $f_{x,y} \in \mathbb{A}$ for $x, y \in K$ such that $f_{x,y}(x) = g(x)$ and $f_{x,y} = g(y)$. Then $U_{x,y} := \{z \in K : f_{x,y}(z) < g(z) + \epsilon\} \ni y$ is an open covering of the compact K. Hence we can set $f_x := \min\{f_{x,y_1}, f_{x,y_2}, \ldots, f_{x,y_n}\}$, by step 4 $f_x \in \overline{\mathbb{A}}$ and satisfies the claim.

- Step 6: Let $g \in C_{\mathbb{R}}(K)$ be arbitrary and $\epsilon > 0$, then there exists $f \in \overline{\mathbb{A}}$ with $\|f - g\| < \epsilon$, which entails that $\overline{\mathbb{A}} = C_{\mathbb{R}}(K)$.
  - We use step 5 and have for $x \in K$ that $V_x := \{z \in K : f_x(z) > g(z) - \epsilon\}$ is an open covering of the compact K and hence $f := \max\{f_{x_1}, \ldots, f_{x_m}\}$ is in $\overline{\mathbb{A}}$ by step 4 and satisfies $|f(x) - g(x)| < \epsilon$ for all $x \in K$.

Assume now that $\exists x_0 \in K$ with $f(x_0) = 0$ for all $f \in \mathbb{A}$ and we want to show that $\overline{\mathbb{A}} = \{f \in C_{\mathbb{K}}(K) : f(x_0) = 0\}$. The uniqueness of x_0 follows because $\mathbb{A}$ separates the points of K. For the existence notice that $f \in \overline{\mathbb{A}}$ gives $f(x_0) = 0$ and thus one inclusion is trivial, then $\mathbb{B} = \{f + \lambda : f \in \mathbb{A}, \lambda \in \mathbb{R}\}$ is a subalgebra of $C_{\mathbb{R}}(K)$ that separates the points of K gives "$\supset$". □

Remark 2.2. The Stone-Weierstrass Theorem allows us for example to approximate continuous functions on compact subsets of $\mathbb{R}^n$ through series of polynomials, see next example.

Example 2.1. *Let $K \subset \mathbb{R}^n$ be compact and $\mathbb{A}$ be the set of all polynomials in the variables $x_1, \ldots, x_n$, i.e.*

$$\mathbb{A} = \left\{ p(x) = \sum_{\alpha : |\alpha| \leq n} b_\alpha x^\alpha, \ m \in \mathbb{N}, \ b_\alpha \in \mathbb{R} \right\}$$

It is then easy to check that $\mathbb{A}$ is a subalgebra of $C_{\mathbb{K}}(K)$, separating the points of K, moreover that $\{f \in C_{\mathbb{K}}(K) : f(x_0) = 0\}$ cannot hold true, because evidently we have $1 \in \mathbb{A}$ and if we were in the second case, then there would exists $x_0 \in K$ such that $0 = 1(x_0) = 1$ which is a contradiction. Thus by Stone-Weierstrass we concludethat $\overline{\mathbb{A}} = C_{\mathbb{K}}(K)$.

Theorem 2.3 (Stone-Weierstrass, $\mathbb{K} = \mathbb{C}$)**.** *Let $\mathbb{A}$ be a subalgebra of $C_{\mathbb{C}}(K)$, separating the points of K and such that for all $f \in \mathbb{B}$ we have $\overline{f} \in \mathbb{A}$. Then we either have $\overline{\mathbb{A}} = C_{\mathbb{C}}(K)$ or*

$$\exists! x_0 \in K \text{ with } \overline{\mathbb{A}} = \{f \in C_{\mathbb{C}}(K) : f(x_0) = 0\}.$$

Proof. We consider $\mathbb{A}_{\mathbb{R}} := C_{\mathbb{R}}(K) \cap \mathbb{A}$ and notice that $\mathbb{A}_{\mathbb{R}}$ is a subalgebra of $C_{\mathbb{R}}(K)$ that also separates the points of K. Moreover $f \in \mathbb{A}$ implies both that Ref and Im f are in $\mathbb{A}_{\mathbb{R}}$. The claim then follows by noticing that

$$\overline{\mathbb{A}} = \overline{\mathbb{A}_{\mathbb{R}} + i\mathbb{A}_{\mathbb{R}}} = \overline{\mathbb{A}_{\mathbb{R}}} + i\overline{\mathbb{A}_{\mathbb{R}}}$$

□

2.2 Lebesgue Spaces

This should already be known, we will only repeat the most relevant results without much explanation in this section. We work on some measure space (Ω, Σ, μ) and consider the space

$$\tilde{L}^p(\Omega, d\mu) = \{f : \Omega \to \mathbb{R} : f \text{ measurable}, \int_\Omega |f|^p d\mu < \infty\}$$

Theorem 2.4 (Hölder's inequality)**.** *Let $1 \leq p, q \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Let (Ω, Σ, μ) be a measure space and $f \in L^p(\Omega), g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$ and*

$$\left|\int_\Omega fg d\mu\right| \leq \int_\Omega |f||g| d\mu \leq \|f\|_p \cdot \|g|_q$$

Theorem 2.5 (Fischer-Riesz)**.** *Let (Ω, Σ, μ) be a measure space and $1 \leq p \leq \infty$. Then $L^p(\Omega, \Sigma, \mu)$ is complete.*

Remark 2.3. For $p = 2$ the Lebesgue space L^2 is even a Hilbert space with scalar product on $L^2(\Omega, \Sigma, \mu)$ given by

$$(f, g)_{L^2} = \int \overline{f(x)} g(x) d\mu(x)$$

an exercise from class entails that this is on the only Lebesgue Space that is a Hilbert Space, i.e. only for the case $p = 2$.

Definition 2.3. *Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{K}$. We define the dual space X^* of X as the space of all continuous linear functionals on X, i.e.*

$$X^* := \{f : X \to \mathbb{K} \mid f \text{ is linear and continuous}\}$$

We will discuss this space in much more detail later on, for now we note that X^* is always a Banach space w.r.t. the norm

$$\|f\|_{X^*} := \sup_{x \in X : \|x\| = 1} |f(x)|$$

Theorem 2.6 (Duality of L^p spaces)**.** *Let $1 \leq p < \infty, 1 < q \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ and let $(\Omega, \mathcal{A}, \mu)$ be a measure space. If $p = 1$, we assume additionally that the measure space $(\Omega, \mathcal{A}, \mu)$ is σ-finite. Then the map*

$$\phi : \begin{cases} L^q(\Omega, \mathcal{A}, \mu) & \longrightarrow (L^p(\Omega, \mathcal{A}, \mu))^* \\ f & \longmapsto \phi_f \end{cases}$$

with

$$\phi_f(g) = \int_\Omega \overline{f(x)} g(x) d\mu(x)$$

for all $g \in L^p(\Omega, \mathcal{A}, \mu)$ is an anti-linear isometric isomorphism. In other words, we have

$$L^p(\Omega, \mathcal{A}, \mu)^* \cong L^q(\Omega, \mathcal{A}, \mu)$$

Remark 2.4. The theorem determines the dual space of $L^p(\Omega, \mathcal{A}, \mu)$, it shows that every continuous linear functional L on $L^p(\Omega, \mathcal{A}, \mu)$ for $1 \leq p < \infty$ has the form

$$L(f) = \int_\Omega \overline{g(x)} f(x) d\mu(x)$$

for a $g \in L^q(\Omega, \mathcal{A}, \mu)$ with $\frac{1}{p} + \frac{1}{q} = 1$.

Our last goal in this section is to show that smooth functions are dense in L^p, for all $1 \leq p < \infty$. For simplicity we work on $\mathbb{R}^n$, then the next theorem shows how to approximate L^p-functions with sequences of C^∞-functions.

Theorem 2.7. *Let $j \in L^1(\mathbb{R}^n)$ with $\int j d\lambda_n(x) = 1$. For $\epsilon > 0$ we let $j_\epsilon(x) = \epsilon^{-1} j(x/\epsilon)$ so that $\int j_\epsilon d\lambda_n = 1$ for all $\epsilon > 0$. Let now $f \in L^p(\mathbb{R}^n)$ for a $1 \leq p < \infty$ and set $f_\epsilon = f * j_\epsilon$. Then $f_\epsilon \in L^p(\mathbb{R}^n)$ with $\|f_\epsilon\|_p \leq \|j\|_1 \|f\|_p$ and $f_\epsilon \to f$ in $L^p(\mathbb{R}^n)$ as $\epsilon \to 0$. If $j \in C_c^\infty(\mathbb{R}^n)$, then we have $f_\epsilon \in C^\infty(\mathbb{R}^n)$ and $D^\alpha f_\epsilon = (D^\alpha j_\epsilon) * f$ for all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$.*

Lemma 2.2. *Let $\Omega \subset \mathbb{R}^n$ open, $K \subset \Omega$ compact. Then there exists $J_K \in C_c^\infty(\Omega)$ with $0 \leq J_K(x) \leq 1$ for all $x \in \Omega$, $J_K(x) = 1$ for all $x \in K$. Therefore, there exists a sequence $(g_j)_{j \in \mathbb{N}}$ in $C_C^\infty(\Omega)$ with $0 \leq g_j(x) \leq 1$ for all $j \in \mathbb{N}$ and $\lim g_j(x) = 1$ for all $x \in \Omega$. Hence if $(f_j)_{j \in \mathbb{N}}$ is a sequence in $C^\infty(\Omega)$ with $f_j \to f$ in $L^p(\Omega)$ for $1 \leq p < \infty$, then $g_j f_j \in C_c^\infty(\Omega)$ and $g_j f_j \to f$ in $L^p(\Omega)$. The Theorem above implies therefore that $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for all $1 \leq p < \infty$.*

As an application of the density of $C_c^\infty(\mathbb{R}^n)$ in $L^p(\mathbb{R}^n)$ we have the following theorem.

Theorem 2.8. *For every measurable $\Omega \subset \mathbb{R}^n$ and $1 \leq p < \infty$, the Banach space $L^p(\Omega)$ is seperable.*

2.3 Sobolev Spaces

Functions in L^p-spaces are characterized by their integrability properties. In analysis, however, it is often important to consider derivatives. The notion of derivative does not combine very well with L^p spaces, functions in L^p spaces are typically not differentiable. For this reason, we are going to introduce a weaker notion of derivative, Sobolev spaces consists then of L^p-functions, whose weak derivative is again an L^p-function.

We work with $\Omega \subset \mathbb{R}^n$ open and let $\Sigma \subset 2^\Omega$ be the Borel σ-algebra over Ω, moreover let dx denotes Lebesgue mass. We define the normed space

$$X = \{f \in C^\infty(\Omega) : \|f\|_X < \infty\}, \text{ with } \|f\|_X = \sum_{|\alpha| \leq m} \|D^\alpha f\|_{L^p(\Omega)}$$

where the sum above runs over all $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ with $|\alpha| = \sum_{i=1}^n \alpha_i$ and $D^\alpha f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$. It is easy to check that $(X, \|\cdot\|_X)$ is indeed a normed space but that it is not complete. We define $\tilde{X}$ as the completion of this normed space and we want to find a good cahracterization of $\tilde{X}$.

To this end, let $[(f_j)_{j\in\mathbb{N}}] \in \tilde{X}$, i.e. let $(f_j)_{j\in\mathbb{N}}$ be a Cauchy sequence in X. Then by the definition of the norm on X, $(D^\alpha f_j)_{j\in\mathbb{N}}$ defines a Cauchy sequence on $L^p(\Omega)$ because we have for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$

$$\|D^\alpha f_j - D^\alpha f_l\|_{L^p} \le \|D^\alpha f_j - D^\alpha f_l\|_X \to 0 \text{ as } j, l \to \infty.$$

Since $L^p(\Omega)$ is a complete space, there exiss $f^{(\alpha)} \in L^p(\Omega)$ such that $D^\alpha f_j \to f^{(\alpha)}$ in L^p as $j \to \infty$. For any $\xi \in C_c^\infty(\Omega)$ we obtain by the above convergence, integration by parts and the fact that ξ is compactly supported that

$$\begin{aligned}\int_\Omega D^\alpha \xi f^{(0)} dx &= \lim_{j\to\infty} \int_\Omega D^\alpha \xi f_j dx \\ &= (-1)^{|\alpha|} \lim_{j\to\infty} \int_\Omega \xi D^\alpha f_j dx = (-1)^{|\alpha|} \int_\Omega \xi f^{(\alpha)} dx\end{aligned}$$

This identity gives a relationship between $f^{(\alpha)}$ and $f^{(0)} = f$, for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$.

Motivated by this remark, we define the Sobolev space of order $m \in \mathbb{N}$ with exponent $1 \le p \le \infty$ as

$$\begin{aligned}H^{m,p}(\Omega) = \{f \in L^p(\Omega) : \forall \alpha \in \mathbb{N}^n \text{ with } 1 \le |\alpha| \le m, \exists f^{(\alpha)} \in L^p(\Omega) \text{ such that} \\ \int_\Omega D^\alpha \xi f dx = (-1)^{|\alpha|} \int_\Omega \xi f^{(\alpha)} dx \text{ for all } \xi \in C_c^\infty(\Omega)\}\end{aligned}$$

We equip $H^{m,p}(\Omega)$ with the norm

$$\|f\|_{H^{m,p}} := \sum_{|\alpha|\le m} \|f^{(\alpha)}\|_{L^p}.$$

It is then easy to check that $(H^{m,p}(\Omega), \|\cdot\|_{H^{m,p}})$ is a normed space. For $f \in H^{m,p}(\Omega)$ we call the functions $f^{(\alpha)}, 1 \le |\alpha| \le m$, the weak derivatives of f. We often use the notation $f^{(\alpha)} = \partial^\alpha f$. The following Lemma shows that the weak derivatives of a function are unique.

Lemma 2.3. *Let $\Omega \subset \mathbb{R}^n$ open and $f \in L^p(\Omega, dx)$, $1 \le p \le \infty$, with*

$$\int_\Omega \xi f dx = 0, \text{ for all } \xi \in C_c^\infty(\Omega)$$

then we have $f = 0$.

Remark 2.5. If $f \in L^p(\Omega) \cap C^m(\Omega)$, then the classical derivates $D^\alpha f$ coincide with the weak derivatives, this can be shown by using integration by parts. The advantage of the weak derivatives is of course the fact that they exist for a much larger class of functions.

Next we show that the completeness of the L^p spaces implies also the completeness of the Sobolev spaces.

Theorem 2.9. *For all $m \in \mathbb{N},\ 1 \leq p \leq \infty,\ H^{m,p}(\Omega)$ is a Banach space.*

Proof. Let $(f_j)_{j\in\mathbb{N}}$ be a Cauchy sequence in $H^{m,p}(\Omega)$, then clearly f_j is also a Cauchy sequence in $L^p(\Omega)$, hence there exists $f \in L^p(\Omega)$ with $f_j \to j$ in $L^p(\Omega)$. Moreover, for every $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m$, we also have that $\partial^\alpha f$ is a Cauchy sequence in $L^p(\Omega)$. Hence there exists $g_\alpha \in L^p(\Omega)$ with $\partial^\alpha f_j \to g_\alpha$ in $L^p(\Omega)$. We claim that that in fact $f \in H^{m,p}(\Omega)$ and that $\partial^\alpha f = g_\alpha$. In fact, we observe that for arbitray $\xi \in C_c^\infty(\Omega)$ we have

$$\int_\Omega D^\alpha \xi f dx = \lim_{j\to\infty} \int_\Omega D^\alpha \xi f_j dx = (-1)^{|\alpha|} \lim_{j\to\infty} \int_\Omega \xi \partial^\alpha f_j dx = (-1)^{|\alpha|} \int_\Omega \xi g_\alpha dx$$

□

Theorem 2.10 (Approximation Theorem)**.** *Let $f \in H^{m,p}(\Omega)$, $1 \leq p < \infty$, $m \in \mathbb{N}$. Then there exists a sequence in $H^{m,p}(\Omega) \cap C^\infty(\Omega)$ with $\|f_j - f\|_{H^{m,p}}$ as $j \to \infty$. Moreover, if $\Omega = \mathbb{R}^n$, then we can choose $f_k \in C_c^\infty(\mathbb{R}^n)$ with $\|f_k - f\|_{H^{m,p}} \to 0$ as $k \to \infty$.*

Proof. We consider a Dirac sequence in order to show the existence of cutoff functions. Then we show the existence of the partition of the identity. To approximate functions in $H^{m,p}(\Omega)$ we first approximate them locally, then we use an appropriate partition of the identity to reduce the global problem to a local one.

Finally we find a locally finite covering of Ω and then take $(\eta_k)_{k\in\mathbb{N}}$ to be the corresponding partition of the identity, we than have $f_{k,\epsilon} \in C^\infty(U_k)$ and ultimately

$$f_\epsilon := \sum_{k\in\mathbb{N}} \eta_k f_{k,\epsilon} \to f \text{ in } H^{m,p}(\Omega)$$

□

Remark 2.6. The theorem shows that, for all $1 \leq p < \infty$, $H^{m,p}(\Omega)$ is the completion of $C^\infty(\Omega)$ with respect to the norm of $H^{m,p}$.

3 Compactness

3.1 Compact Sets in Metric Spaces

Definition 3.1. *Let (X, d) be a metric space. A subset $A \subset X$ is called pre-compact if, for all $\epsilon > 0$, A has a finite covering with ϵ-balls, i.e. if there are $x_1, \ldots, x_n \in X$ with $\cup_{i=1}^{n} B_\epsilon(x_i) \supset A$.*

Remark 3.1. We have the following

1. Subsets of pre-compact sets are pre-compact.
2. $A \subset X$ pre-compact implies that A is bounded.
3. $A \subset X$ pre-compact implies that $\overline{A}$ is closed and pre-compact.

Theorem 3.1. *Let (X, d) be a metric space. For $A \subset X$, the following statements are equivalent:*

1. *A is compact.*
2. *A is sequentially compact, i.e. every sequence in A has a convergent subsequence.*
3. *(A, d) is complete and A is pre-compact.*

Remark 3.2.

1. We see that what's missing for a pre-compact set to be compact is completeness.
2. $A \subset X$ compact implies that A is closed (since A complete implies that A is closed)
3. If X is complete, then $A \subset X$ is pre-compact if and only if $\overline{A}$ is compact (since closed subsets of complete spaces are complete).

On finite dimensional vector spaces (for example normed spaces with finite dimensional V vector space) all norms are equivalent.

Lemma 3.1. *Let X be a finite dimensional vector space over $\mathbb{K}$. Then all norms on X are equivalent. In other words, if $\|\cdot\|_1, \|\cdot\|_2$ are two norms on X, then there exists $C > 0$ such that*

$$\frac{1}{C}\|x\|_2 \leq \|x\|_1 \leq C\|x\|_2, \text{ for all } x \in X.$$

Corollary 3.1. *Every finite dimensional subspace of a normed space is complete and therefore closed.*

Theorem 3.2 (Heine-Borel)**.** *Let $(X, \|\cdot\|)$ be a finite dimensional vector space. Then $A \subset X$ is compact if and only if A is bounded and closed.*

Proof. " $\Longrightarrow$ " Readily follows from the Theorem above, because A is compact if and only if (A, d) is complete and A is pre-compact, which implies that A is bounded and closed.

"$\Longleftarrow$" Assume that $A \subset X$ is bounded and closed. Since $X \subset X$ is a finite dimensional subspace, X is complete and therefore closed. Thus it is enough to show that $A \subset X$ is pre-compact and we know that this is the case if and only if $\overline{A} = A$ is complete.

Let $\{e_1, \ldots, e_n\}$ be a basis of X. Then we define the norm

$$\|x\|_\infty = \max_{j=1,\ldots,n} |x_j|, \text{ for } x = \sum_{k=1}^{n} x_j e_j \in X$$

on X. Since all norms are equivalent, it is enough to prove the claim w.r.t. to this norm $\|\cdot\|_\infty$. To this end, it is enough to show that for $R > 0, \epsilon > 0$ there exists $n(R, \epsilon) \in \mathbb{N}$ and $x^{(1)}, \ldots, x^{(n(R,\epsilon))} \in X$ such that

$$B_R(0) \subset \bigcup_{j=1}^{n(R,\epsilon)} B_\epsilon(x^{(j)})$$

$\square$

Proposition 3.1. *Let X be a normed space. Then $\overline{B_1(0)}$ is compact if and only if* $\dim X < \infty$

Proof. "$\Longleftarrow$" readily follows from the Theorem of Heine Borel, indeed since our normed space is of finite dimension and evidently $\overline{B_1(0)}$ is both closed and bounded, we the claim follows.

" $\Longrightarrow$ " Let $y_1, \ldots, y_n \in X$ be such that

$$\overline{B_1(0)} \subset \bigcup_{j=1}^{n} B_{1/2}(y_j) \tag{*}$$

and let $Y = \text{span}\{y_1, \ldots, y_n\} \subset X$, then since Y is a finite dimensional subspace of the normed space X it is complete and in particular closed (see previous corollary). Assume then for contradiction that $Y \neq X$, then for all $\theta \in (0, 1)$ we can find $x_\theta \in X \setminus Y$ with $\|x_\theta\| = 1$ (i.e. $x_\theta \in \overline{B_1(0)}$) and $\text{dist}(x_\theta, Y) \geq \theta$. Choosing $\theta > 1/2$ then yields a contradiction to (*) above. $\square$

3.2 Compact Subsets of $C(K)$

For finite dimensional normed spaces we found a simple characterization of compactness, namely if $\dim X < \infty$ then $A \subset X$ is compact if and only if A is closed and bounded. We try to find a similar simple characterization of compact subsets for the case of infinite dimensional spaces. We denote $C(K) = C_{\mathbb{K}}(K)$ where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Definition 3.2. *$S \subset C(K)$ is called equicontinuous at $x \in K$ if*

$\forall \epsilon > 0, \exists U_x$ *open neighbourhood of x in K* : $|f(x) - f(y)| \leq \epsilon, \forall y \in U_x, \forall f \in S.$

$S \subset C(K)$ is called equicontinuous, if S is equicontinuous at x, for all $x \in K$.

Remark 3.3. The open neighbourhood of x U_x may solely depend on ϵ and the point x, it mustn't depend on $f \in S$.

Theorem 3.3 (Arzelá-Ascoli)**.** *A subset $S \subset C(K)$ is pre-compact if and only if S is bounded and equicontinuous.*

Remark 3.4. We know that $S \subset C(K)$ is pre-compact if and only if $\overline{S}$ is compact. Hence from Arzelá-Ascoli we obtain that $S \subset C(K)$ is compact if and only S is closed, bounded and equicontinuous. In particular, in comparison with Heine-Borel Theorem we require to satisfy 1 additional conditional, namely the condition of equicontinuity.

Proof. " $\Longrightarrow$ " Let $\epsilon > 0$. Since S is pre-compact, it has a finite covering with ϵ-balls. In other words, there exists $f_1, \ldots, f_n \in S$ with $S \subset \cup_{j=1}^{n} B_\epsilon(f_j)$. Hence, for all $f \in S$, there exists $j \in \{1, \ldots, n\}$ such that $f \in B_\epsilon(f_j)$ and therefore for all $f \in S$

$$\|f\| \ \leq \|f_j\| + \epsilon$$

This proves that

$$\sup_{j \in S} \|f\| \leq \max_{j=1,\ldots,n} \|f_j\| + \epsilon$$

and also that S is bounded. Now, fix $x \in K$. Since the f_i are continuous, we have for $i = 1, \ldots, n$ there exists an open neighbourhood U_i of x in K such that

$$|f_i(x) - f_i(y)| \leq \epsilon, \ \forall y \in U_i$$

Let now $U := \cap_{i=1}^{n} U_i$. Then U is an open neighbourhood of $x \in K$ for all $f \in S$ and for all $y \in U$ we have

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| \\ &\leq 2\|f - f_i\| + |f_i(x) - f_i(y)| \leq 3\epsilon \end{aligned}$$

for an appropriate choice of $i \in \{1, \ldots, n\}$. This shows the equicontinuity of S.

For "$\Longleftarrow$", since S is equicontinuous we can for $\epsilon > 0$ and $x \in K$ always find U_x open neighbourhood of x such that

$$|f(x) - f(y)| \leq \epsilon \; \forall f \in S, \forall y \in U_x.$$

Then $\{U_x\}_{x \in K}$ is an open covering of K. I.e. there exists $x_1, \ldots, x_n \in K$ with $K = \cup_{j=1}^n U_{x_j}$. Since moreover S is bounded, we have that

$$R = \{(f(x_1), \ldots, f(x_n)) : f \in S\}$$

is a bounded subset of $\mathbb{C}^n$, equipped with the max-norm. Therefore, R is pre-compact in $\mathbb{C}^n$ (since $\mathbb{C}^n$ is of finite dimension), i.e. there exists $f_1, \ldots, f_m \in S$, with

$$R \subset \bigcup_{i=1}^m B_\epsilon((f_i(x_1), \ldots, f_i(x_n))$$

we can than show that

$$S \subset \bigcup_{i=1}^m B_{3\epsilon}(f_i).$$

which shows that S is pre-compact. □

3.3 Compact Subsets of L^p-Spaces

We can ask the same question as in the previous section but now for subsets of L^p-spaces. In this case the description of compact subsets is given by the Theorem of Riesz.

Theorem 3.4 (Riesz's Theorem)**.** *Let $1 \leq p < \infty$, $A \subset L^p(\mathbb{R}^n)$. Then A is pre-compact if and only if the following conditions are satisfied:*

1. *A is bounded*
2. $\sup\limits_{f \in A} \|f(\cdot + h) - f\|_p \to 0$ *as* $h \to 0$
3. $\sup\limits_{f \in A} \|f\|_{L^p(\mathbb{R}^n \setminus B_r(0))} \to 0$ *as* $R \to \infty$.

Remark 3.5.

1. The conditions 2 and 3 are nontrivial in the theorem of Riesz because they have to hold uniformly. If we drop the uniformity they are trivial (3 for instance follows by dominated convergence).
2. The theorem holds true for subsets of $L^p(\mathbb{R}^n)$. There are also extensions to subsets of $L^p(\Omega)$, for measurable $\Omega \subset \mathbb{R}^n$.

Proof. " $\Longrightarrow$ " Since $A \subset L^p(\mathbb{R}^n$ is pre-compact, we have for arbitrary $\epsilon > 0$ the existence of $g_1, \ldots, g_n \in L^p(\mathbb{R}^n)$ with $A \subset \cup_{j=1}^n B_\epsilon(g_j)$. Now for $f \in A$ we find $f \in B_\epsilon(g_j)$ for an appropriate $j = 1, \ldots, n$ and thus

$$\sup_{f \in A} \|f\|_p \leq \epsilon + \max_{j=1,\ldots,n} \|g_j\|$$

and therefore

$$\sup_{f \in A} \|f(. + h) - f\|_p \leq 2\epsilon + \max_{j=1,\ldots,n} \|g_j(. + h) - g_j\|_p$$
$$\sup_{f \in A} \|f\|_{L^p(\mathbb{R}^n \setminus B_R(0))} \leq \epsilon + \max_{j=1,\ldots,n} \|g_j\|_{L^p(\mathbb{R}^n \setminus B_R(0))}$$

The RHS of the two inequalities above are certainly smaller than 3ϵ, respecitively, 2ϵ if $|h|$ is small enough, respectively, if R is large enough. To summarize, in a sketch of the proof we could say that " $\Longrightarrow$ " is trivial.

" $\Longleftarrow$ " Let $(\varphi_\epsilon)_{\epsilon > 0}$ be a standard Dirac sequence and R_ϵ be such that $R_\epsilon \to \infty$ as $\epsilon \to 0$. For $f \in A$ let

$$T_\epsilon f(x) = \int_{B_{R_\epsilon}(0)} \varphi_\epsilon(x - y) f(y) dy = (\varphi_\epsilon * 1_{B_{R_\epsilon}(0)} f)(x)$$

We then know that $T_\epsilon f \in L^p(\mathbb{R}^n)$ and with Hölder's inequality we can estimate

$$\|T_\epsilon f - f\|_p^p \leq \left(\sup_{|h| \leq \epsilon} \sup_{f \in A} \|f(. + h) - f\|^p + \sup_{f \in A} \|f\|^p_{L^p(\mathbb{R}^n \setminus B_{R_\epsilon - \epsilon}(0)}\right)^p =: \kappa_\epsilon^p$$

From the assumption we have $\kappa_\epsilon \to 0$ as $\epsilon \to 0$. Moreover we have $T_\epsilon f \in C_c^\infty(B_{R_\epsilon + \epsilon}(0))$. The set

$$B = \{T_\epsilon f : f \in A\} \subset C(\overline{B_{R_\epsilon + \epsilon}(0)})$$

is a bounded and equicontinuous subset of $C(\overline{B_{R_\epsilon + \epsilon}(0)})$ for every $\epsilon > 0$. The equicontinuity follows from

$$|T_\epsilon f(x) - T_\epsilon f(y)| \leq \|\Delta T_\epsilon f\|_\infty |x - y| \leq C(\epsilon)|x - y|$$

where $C(\epsilon)$ depends on $\epsilon > 0$ but not on $f \in A$. So from Arzela-Ascoli Theorem we can conclude that B is pre-compact and thus cover it with epsilon balls. We then use that the L^p norm on $\overline{B_{R_\epsilon + \epsilon}(0)}$ can be estimated by the C-norm and we obtain that we can find $g_1, \ldots, g_n \in L^p(\mathbb{R}^n$ such that

$$A \subset \bigcup_{j=1}^n B_{\epsilon'}(g_j)$$

which entails that $A \subset L^p(\mathbb{R}^n)$ is pre-compact. $\square$

4 Linear Operators & Functionals on Normed Spaces

4.1 Continuous Operators

Definition 4.1. *Let X, Y be two normed spaces. A (linear) operator $T : X \to Y$ is a linear map. A continuous operator $T : X \to Y$ is a continuous linear map. An operator $T : X \to Y$ is called bounded, if there exists a constant $C > 0$ with*

$$\|Tx\|_Y \leq C\|x\|_X, \text{ for all } x \in X$$

Proposition 4.1 (Characterization of continuous opetaros)**.** *Let X, Y be normed spaces and $T : X \to Y$ a linear operator. The following statements are equivalent:*

1. *T is continuous.*
2. *T is bounded.*
3. *T is continuous at $x = 0$.*

Definition 4.2. *Let X, Y be normed spaces and $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$. Then we denote by $\mathcal{L}(X, Y)$ the space of all continuous linear operators from X to Y. $\mathcal{L}(X, Y)$ has the structure of a vector space over $\mathbb{K}$ when we consider the canonical operators for additions and multiplication with scalars. Moreover, for $T \in \mathcal{L}(X, Y)$ we define*

$$\|T\| := \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\| \leq 1} \|Tx\| = \sup_{\|x\|=1} \|Tx\|$$

in particular $(\mathcal{L}(X, Y), \|\cdot\|)$ is a normed space. We also denote $\mathcal{L}(X) = \mathcal{L}(X, X)$.

Proposition 4.2. *Let X be a normed space and Y be a Banach space. Then $\mathcal{L}(X, Y)$ is a Banach space.*

Proof. Let $(T_l)_{l \in \mathbb{N}}$ be a cauchy sequence on $\mathcal{L}(X, Y)$ and $x \in X$. Then we always have the inequality (!)

$$\|T_l x - T_k x\| \leq \|T_l - T_k\| \|x\| \to 0 \text{ as } l, k \to \infty.$$

Thus $(T_l x)_{l \in \mathbb{N}}$ is a Cauchy sequence in Y. Since Y is by assumption complete, there exists

$$Tx := \lim_{l \to \infty} T_l x$$

The map $T : X \to Y$ given as above is clearly linear. Moreover, $T \in \mathcal{L}(X, Y)$ since

$$\|Tx\| = \lim_{l \to \infty} \|T_l x\| \leq \limsup_{l \to \infty} \|T_l\| \|x\| \leq C\|x\|,$$

since $(T_l)_{l\in\mathbb{N}}$ is a Cauchy sequence and therefore bounded. We finally get

$$\|Tx - T_l x\| = \lim_{k\to\infty} \|T_k x - T_l x\| \leq \limsup_{k\to\infty} \|T_k - T_l\|\|x\| \to 0 \text{ as } l \to \infty$$

because $(T_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, this completes the proof. □

Definition 4.3. *Let X be a normed space over the field $\mathbb{K}$. We define the dual space X^* of X through $X^* := \mathcal{L}(X, \mathbb{K})$. That is elements of X^* are continuous linear functionals (i.e. continuous linear operators on X with values on $\mathbb{K}$). Since $\mathbb{K}$ is always complete, it follows that X^* is always a banach space, independentely of the completeness of X.*

4.2 The Hahn-Banach Theorem and its Applications

Let us recall again the statement of the Lemma of Zorn. A partial order on a set P is a relation $\preceq$ on P with the following properties:

1. $a \preceq a$ (Reflexivity)
2. $a \preceq b$ and $b \preceq a$ implies that $a = b$ (Antisymmetry)
3. If $a \preceq b$ and $b \preceq c$ then $a \preceq c$ (Transitivity)

A subset $M \subset P$ of a partially ordered set P is called totally ordered (or a chain) if for all $a, b \in M$ with $a \neq b$ we either have $a \preceq b$ or $b \preceq a$. Moreover, an element $b \in P$ is called an upper bound for $M \subset P$ if $a \preceq b$ for all $a \in M$. Naturally, we call $b \in P$ a maximal element in P, if there is no other upper bound $a \in P$ with $a \neq b$ and $b \leq a$ for said upper bound.

Lemma 4.1 (Lemma of Zorn)**.** *Let P be a partially ordered set, so that every totally ordered subset of P has an upper bound. Then P contains at least one maximal element.*

Definition 4.4. *Let X be a vector space over $\mathbb{R}$. A map $p : X \to \mathbb{R}$ is called a sublinear functional if*

1. *$p(\lambda x) = \lambda p(x)$ for all $x \in X$ and $\lambda \geq 0$.*
2. *$p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.*

Lemma 4.2. *Let X be a vector space over $\mathbb{R}$. $M \subset X$ a linear subspace, $p : X \to \mathbb{R}$ a sublinear functional, $f : M \to \mathbb{R}$ a linear functional, $x_0 \in X \setminus M$. Assume $f(x) \leq p(x)$ for all $x \in M$. Then there exists $F : \widetilde{M} := M + \mathbb{R}x_0 \to \mathbb{R}$ linear, with $F(x) \leq p(x)$ for all $x \in \widetilde{M}$ and $F_{|M} = f$.*

Theorem 4.1 (Hahn-Banach, vector spaces over $\mathbb{R}$)**.** *Let X be a vector space over $\mathbb{R}$, $M \subset X$ a linear subspace and $f : M \to \mathbb{R}$ linear. Let $p : X \to \mathbb{R}$ be a sublinear functional with $f(x) \le p(x)$ for all $x \in M$. Then there exists $F : X \to \mathbb{R}$ linear with $F(x) \le p(x)$ for all $x \in X$ and $F_{|M} = f$.*

Proof. We consider the set

$$\begin{aligned}\mathcal{F} := \{(\widetilde{M}, \tilde{f}) : \widetilde{M} \subset X \text{ linear with } \widetilde{M} \supset M, \tilde{f} \text{ a linear function on } \widetilde{M},\\ \text{with } \tilde{f}_{|M} = f \text{ and } \tilde{f}(x) \le p(x) \text{ for all } x \in \widetilde{M}\}\end{aligned}$$

We have $\mathcal{F} \neq \emptyset$ since by assumption $(M, f) \in \mathcal{F}$. On $\mathcal{F}$ we define a partial order by setting $(\widetilde{M}, \tilde{f}) \preceq (\widetilde{N}, \tilde{g})$ if $\widetilde{M} \subset \widetilde{N}$ and $\tilde{g}_{|\widetilde{M}} = \tilde{f}$. We claim that every totally ordered subset of $\mathcal{F}$ has an upper bound.

To this end let $\mathcal{G} = \{(M_i, g_i) : i \in I\}$ be a totally ordered subset of $\mathcal{F}$. Since $\mathcal{G}$ is totally ordered we can define the linear space

$$\widetilde{M} = \bigcup_{i \in I} M_i$$

and $\tilde{g} : \widetilde{M} \to \mathbb{R}$, so that for $x \in M_i, \tilde{g}(x) = g_i(x)$. We claim that $\tilde{g}$ is linear. Indeed, since $\mathcal{G}$ is totally ordered we have for any given $x, y \in \widetilde{M}$ and $\lambda \in \mathbb{R}$ there exists $i \in I$ (since $\mathcal{G}$ is totally ordered) such that $x, y\lambda y \in M_i$, hence we have

$$\tilde{g}(x + \lambda y) = \tilde{g}(x) + \lambda \tilde{g}(y)$$

since the g_i are linear. Moreover we clearly have $\tilde{g}_{|M} = f$ and $\tilde{g}(x) \le p(x)$ for all $x \in \widetilde{M}$.

Thus $(\widetilde{M}, \tilde{g}) \in \mathcal{F}$ is an upper bound for $\mathcal{G}$. The Lemma of Zorn implies that there exists a maximal element $(N, F) \in \mathcal{F}$. We want to establish that $N = X$. Assume that $N \neq X$ then $X \setminus N \neq \emptyset$ and we can choose $x_0 \in X \setminus N$, if we apply now the previous Lemma to said x_0 then there exists G with $(N + \mathbb{R}x_0, G) \in \mathcal{F}, G_{|N} = F$. But we clearly have $(N, F) \preceq (N + \mathbb{R}x_0, G)$ and $(N, F) \neq (N + \mathbb{R}x_0, G)$. This is a contradiction to the maximilty of (N, F) and thus $N = X$. □

We want to extend the theorem to vector spaces over $\mathbb{C}$, we thus need to intrtoduce the notion of a seminorm.

Definition 4.5. *Let X be a vector space over $\mathbb{K}$ and $q : X \to \mathbb{R}$ a map with the properties $q(\alpha x) = |\alpha| q(x)$ for all $\alpha \in \mathbb{K}, x \in X$ and with $q(x + y) \le q(x) + q(y)$ for all $x, y \in X$. Then we call q a seminorm.*

Theorem 4.2 (Hahn-Banach, general version, for real and complex vector spaces)**.** *Let X be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, $q : X \to \mathbb{R}$ a seminorm, $M \subset X$ a linear subspace and $f : M \to \mathbb{K}$ a linear functional with $|f(x)| \leq q(x)$ for all $x \in M$. Then there exists a linear functional $F : X \to \mathbb{K}$ with $F_{|M} = f$ and $|F(x)| \leq q(x)$ for all $x \in X$.*

Proof. If $\mathbb{K} = \mathbb{R}$ we can apply Hahn-Banach Theorem (real version) to find $F : X \to \mathbb{R}$ linear with $F_{|M} = f$ and with $F(x) \leq q(x)$ for all $x \in X$. Then we have

$$-F(x) = F(-x) \leq q(-x) = q(x) \implies F(x) \geq -q(x), \text{ thus } |F(x)| \leq q(x).$$

So, let us now consider the case when $\mathbb{K} = \mathbb{C}$. We can consider X and M as vector spaces over $\mathbb{R}$. Notice that Re $f : M \to \mathbb{R}$ is $\mathbb{R}$-linear with $|\text{Re } f(x)| \leq |f(x)| \leq q(x)$. We apply Hahn-Banach for vector spaces over $\mathbb{R}$ and find $G : X \to \mathbb{R}$, $\mathbb{R}$-linear with $G_{|M} = \text{Re } f$ and with $|G(x)| \leq q(x)$ for all $x \in X$.

Let us now define $F(x) = G(x) - iG(ix)$. Then F is clearly $\mathbb{R}$ liniear, however since we have

$$\begin{aligned} F(ix) &= G(ix) - iG(i^2x) = G(ix) - iG((-1)x) = G(ix) + iG(x) \\ &= iG(x) + G(ix) = i(G(x) - iG(ix)) = iF(x) \end{aligned}$$

we can also conclude that F is $\mathbb{C}$-linear. Moreover we have for $x \in M$

$$\begin{aligned} &\text{Re } F(x) = G(x) = \text{Re } f(x) \\ &\text{Im } F(x) = -G(ix) = -\text{Re } f(ix) = -\text{Re } if(x) = \text{Im } f(x) \end{aligned}$$

This shows that $F_{|M} = f$. We still need to verify that $|F(x)| \leq q(x)$ for all $x \in X$. Let $x \in X$ be arbitraryand we write $F(x) = re^{i\theta}$. Then $e^{-i\theta}F(x)$ is real and therefore

$$|F(x)| = |e^{-i\theta}F(x)| = |F(e^{-i\theta}x)| \overset{\text{real}}{=} |G(e^{-i\theta}x)| \leq q(e^{-i\theta x}) = q(x).$$

$\square$

The following corollaries are immediate consequences of the Hahn-Banach theorem. Of course, Hahn-Banach's Theorem is used to construct linear functionals.

Corollary 4.1. *Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{K}$, $M \subset X$, a linear subspace and $f \in M^*$. Then there exists $F \in X^*$ with $F_{|M} = f$ and $\|F\|_{X^*} = \|f\|_{M^*}$.*

Proof. We define $q(x) = \|f\|_{M^*}\|x\|$. Then q is clearly a seminorm and we have $|f(x)| \leq q(x)$ for all $x \in M$. By Hahn-Banach, there exists a linear functional $F : X \to \mathbb{K}$ with $F_{|M} = f$ and $|F(x)| \leq q(x) = \|f\|_{M^*}\|x\|$ for all $x \in X$. This readily implies that $F \in X^*$. Moreover it shows that $\|F\|_{X^*} \leq \|f\|_{M^*}$. On the other hand we have

$$\|f\|_{M^*} = \sup_{x\in M, \|x\|\leq 1} |f(x)| = \sup_{x\in M, \|x\|\leq 1} |F(x)| \leq \sup_{x\in X, \|x\|\leq 1} |F(x)| = \|F\|_{X^*}$$

Hence we established $\|F\|_{X^*} = \|f\|_{M^*}$. □

Corollary 4.2. *Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{K}$, $y \in X \setminus \{0\}$. Then there exists $f \in X^*$ with $\|f\| = 1$ and $f(y) = \|y\|$.*

Proof. We define $g : \mathbb{K} \cdot y =: M \to \mathbb{K}$ through $g(ty) = t\|y\|$. Then clearly $g \in M^*$ with $\|g\|_{M^*} = 1$ and $g(y) = \|y\|$. From the above Corollarly, there exists $f \in X^*$ with $f_{|M} = g$ and hence in particular $f(y) = \|y\|$ and moreover $\|f\|_{X^*} = \|g\|_{M^*} = 1$. □

Corollary 4.3. *Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{K}, Z \subset X$ a linear subspace and $y \in X \setminus Z$. Let $d = dist(y, Z) = \inf_{z\in Z} \|z - y\| > 0$. Then there exists $F \in X^*$ with $\|F\| = 1$, $F_{|Z} = 0$ and $F(y) = d$.*

Proof. Let $M = Z + \mathbb{K}y$ and define $f : M \to \mathbb{K}$ through $f(z + \alpha y) = \alpha d$, for all $\alpha \in \mathbb{K}$. Then f is clearly linear. We claim that $\|f\| = 1$. In fact, since for $\alpha \in \mathbb{K}$ and $z \in Z$ we have that $-z/\alpha \in Z$ we can conclude that

$$|f(z + \alpha y)| = |\alpha|d \leq |\alpha| \cdot \|y - (-z/\alpha)\| = \|\alpha y + z\|$$

In other words, we have $\|f\|_{M^*} \leq 1$. On the other hand, let $(z_n) \in Z$ be a sequence with $\|y - z_n\| \to d$. Then

$$d = f(y - z_n) \leq \|f\|_{M^*}\|y - z_n\|$$

As $n \to \infty$ we find that $\|f\|_{M^*} \geq 1$. We summarize, $f \in M^*$ with $f_{|Z} = 0$ and $\|f\|_{M^*} = 1$. From Corollary 4.1. there exists $F \in X^*$ with $\|F\|_{X^*} = \|f\|_{M^*} = 1$ and $F_{|M} = f$ which gives in particular that $F(y) = f(y) = d$ and $F_{|Z} = 0$. □

Remark 4.1. Yet another consequence of Hahn-Banach Theorem is that if X is a normed space, then X^* seperates the points of X, i.e. for every $x_1, x_2 \in X$ with $x_1 \neq x_2$, there exists $f \in X^*$ with $f(x_1) \neq f(x_2)$. In fact, given $x_1, x_2 \in X$ with $x_1 \neq x_2$ we set $y = x_1 - x_1 \neq 0$. Then from Corollary 4.2. there exists $f \in X^*$ with $f(y) = \|y\| \neq 0$ which implies that $f(x_1) \neq f(x_2)$.

Definition 4.6. *Let X, Y be two normed spaces and $T : X \to Y$ a continuous linear operator. The adjoint operator to T, denoted by T^*, is a linear map $T^* : Y^* \to X^*$, defined by $T^*(f) = f \circ T$, for all $f \in Y^*$.*

Remark 4.2. Notice that if $T \in \mathcal{L}(X, Y)$ and $S \in \mathcal{L}(X, Z)$, then we have $(ST)^* = T^* S^*$.

Proposition 4.3. *Let $T \in \mathcal{L}(X, Y)$. Then $T^* \in \mathcal{L}(Y^*, X^*)$ with $\|T^*\| = \|T\|$.*

4.3 Reflexive normed spaces

Let X be a normed space. Apart from the dual space X^* also the bidual space X^{**} plays an important role. We construct the canonical inclusion of X in X^{**} as follows: For $x \in X$ we define the linear function $\tilde{x} : X^* \to \mathbb{K}$ through

$$\tilde{x}(f) := f(x)$$

Then indeed $\tilde{x}$ is continuous since $f \in X^*$ is continuous, hence we have $\tilde{x} \in X^{**}$. We define $J_X : X \to X^{**}$ as $J_X(x) = \tilde{x}$ and call it the canonical inclusion of X in X^{**}.

Theorem 4.3. *Let X be a normed vector space over $\mathbb{K}$. Then the canonical inclusion $J_X : X \to X^{**}$ is a linear isometry.*

Definition 4.7. *A normed space $(X, \|\cdot\|)$ is called reflexive, if the map J_X is surjective (and therefore an isometric isomorphism, in particular $X \cong X^{**}$).*

Remark 4.3. The space $X^{**} = \mathcal{L}(X^*, \mathbb{K})$ is always a Banach space, hence X reflexive implies that X is also a Banach space, i.e. X is complete.

Theorem 4.4. *Let X be a reflexive Banach space and $M \subset X$ a closed linear subspace. Then M is reflexive as well.*

Theorem 4.5. *Let X be a Banach space. Then X is reflexive if and only if X^* is reflexive.*

4.4 Hilbert Space Methods

Hilbert spaces can be identified with their dual spaces, this is the subject of Riesz's Representation Theorem.

Theorem 4.6 (Riesz's Representation Theorem)**.** *Let $(H, (\cdot, \cdot))$ be an Hilbert space. Let $v \in H$ be arbitrary. The map $R_H^v : H \to H^*$ defined through $R_H^v(u) = (R_H(u))(v) = (u, v) = \langle u, v\rangle_H$ is an anti-linear isometric isomorphism.*

Proof. • R_H is well defined because $R_H(u)$ is clearly linear and by $|(R_H(u))(v)| \leq |\langle u,v\rangle| \leq \|u\|\|v\|$ it follows that $R_H(u)$ is continuous.
• R_H is anti-linear, because of the anti-linearity of the inner product in its first argument.
• R_H is isometric (and thus injective) because from our first point we immediatly get $\|R_H(u)\| \leq \|u\|$ and since $R_H(u) = \langle u,u\rangle = \|u\|^2$ it follows that $\|R_H(u)\| \geq \|u\|$.

It remains to show that R_H is surjective. Let $L \in H^*$ and set $M := \ker L = \{u \in H : L(u) = 0\}$. Then clearly M is a libear subspace of H. Moreoever, since L is continuous, the set M is also (quite trivially) closed. Thus we know that we can decompose our Hilbert space as $H = M \oplus M^\perp$.

W.l.o.g. we can assume that $L \neq 0$. Then it follows that $\dim M^\perp = 1$, because if $u, v \in M^\perp$ then by the linearty of L we get that

$$\begin{aligned} L(L(v)u - L(u)v) = 0 &\implies L(v)u - L(u)v \in M \cap M^\perp = \{0\} \\ &\implies L(v)u = L(u)v \implies u = \frac{L(u)}{L(v)}v \end{aligned}$$

i.e. $u, v \in M^\perp$ are related in the sense of $u = \lambda v$, thus $M^\perp = \mathbb{K}v$ for an appropriate $v \in H \setminus \{0\}$ which gives us dimension 1.

We claim now that

$$L(u) = \left(\frac{\overline{L(v)}}{\|v\|^2}v, u\right), \text{ for all } u \in H,$$

which entails that $L = R_H(\overline{L(v)}/\|v\|^2 v)$, i.e. R_H is surjective. To verify our claim we use the decomposition of H and write for $H \ni u = \tilde{u} + \lambda v$ for $\tilde{u} \in M = \ker L$ and $\lambda \in \mathbb{K}$. Then

$$\left(\frac{\overline{L(v)}}{\|v\|^2}v, u\right) = \left(\frac{\overline{L(v)}}{\|v\|^2}v, \lambda v\right) = \lambda L(v) = L(u)$$

□

Remark 4.4.

1. It follows from Riesz's Representation Theorem that H^* also has the structure of a Hilbert space (clearly H^* is a Banach space). For $f, g \in H^*$ we define

$$\langle f, g\rangle_{H^*} := \langle R_H^{-1} g, R_H^{-1} f\rangle_H$$

then it is simple to check that $\langle ., , \rangle$ is a scalar product and that

$$\|f\|_{H^*}^2 = \langle f, f \rangle_{H^*} = \langle R_H^{-1} f, R_H^{-1} f \rangle_H = \|R_H^{-1} f\|_H^2$$

since R_H and thus also R_H^{-1} are isometric.

2. The theorem states that every element of H^* can be written uniquely in this form. We also use the notation $(R_H(u))(v) = (u, v)$.

Corollary 4.4. *Every Hilbert space is reflexive.*

Proof. We claim that $J_H = R_{H^*} \circ R_H$. Then J_H is surjective, as the composition of two surjective maps. Let $u \in H$ and $f \in H^*$ be arbitrary. Then

$$\begin{aligned}((R_{H^*} \circ R_H)(u))(f) &= (R_{H^*}(R_H(u))(f) = \langle R_H(u), f \rangle_{H^*} \\ &= \langle R_H^{-1} f, u \rangle_H = (R_H(R_H^{-1} f))(u) = f(u) = (J_H u)(f)\end{aligned}$$

□

4.4.1 Partial differential equations

A consequence of Riesz's representation theorem which is very useful in applications to the study of partial differential equations is the following result of Lax-Milgram.

Theorem 4.7 (Lax-Milgram). *Let H be a Hilbert space over $\mathbb{K}$ and let $a : H \times H \to \mathbb{K}$ be a sesquilinear form (i.e. a is linear in the second argument and antilinear in the first argument). Suppose that there are constants $0 < c_0 \leq C_0$ with*

$$\begin{aligned}|a(x, y)| &\leq C_0 \|x\| \|y\|, \textit{ for all } x, y \in H \textit{ (continuity)} \\ Re\ a(x, x) &\geq c_0 \|x\|^2, \textit{ for all } x \in H \textit{ (coercivity)}\end{aligned}$$

Then there exists a linear operator $A : H \to H$ with

$$a(x, y) = (Ax, y), \textit{ for all } x, y \in H$$

Moreover $A \in \mathcal{L}(H)$ is invertible with

$$\|A\| \leq C_0 \textit{ and } \|A^{-1}\| \leq \frac{1}{c_0}$$

Proof. For all fixed $x \in H$, the map $y \mapsto a(x, y)$ is in H^* and we have

$$\|a(x, \cdot)\|_{H^*} = \sup_{\|y\| \leq 1} |a(x, y)| \leq C_0 \|x\|$$

From Riesz's representation Theorem there exists for arbitrary $x \in H$, exactly one $A(x) \in H$ such that $a(x, y) = (A(x), y) = (R_H(A(x)))(y)$ for all

$y \in H$ with $\|A(x)\| = \|a(x,\cdot)\|_{H^*} \leq C_0\|x\|$.

Moreover, we have that A is linear because

$$\begin{aligned}(A(x_1 + \lambda x_2), y) &= a(x_1 + \lambda x_2, y) = a(x_1, y) + \overline{\lambda} a(x_2, y) \\ &= (A(x_1), y) + \overline{\lambda}(A(x_2), y) \\ &= (A(x_1) + \lambda A(x_2), y), \text{ for all } y \in H,\end{aligned}$$

hence we have $A(x_1 + \lambda x_2) = A(x_1) + \lambda A(x_2)$ and thus $A \in \mathcal{L}(H)$ with $\|A\| \leq C_0$. We now show that A is invertible, as follows:

To show that A is injective we use the coercivity assumption, i.e.

$$\begin{aligned}c_0\|x\|^2 \leq \text{Re } a(x,x) = \text{Re } (A(x), x) \leq |(Ax, x)| \leq \|Ax\|\|x\| \\ \implies \|Ax\| \geq c_0\|x\| > 0, \text{ for all } x \in H \setminus \{0\}\end{aligned}$$

this shows that $\ker A = \{0\}$, and thus A is injective.

To show that A is surjective, we first establish that Ran$A$= $\{Ax : x \in H\}$ is closed, we can use the inequality established in the injectivity part to entail this. Let Ax_n be a sequence in Ran$A$ with $Ax_n \to y \in H$, then indeed Ax_n is Cauchy in H (since it convergent) but also (from the above inequality)

$$c_0\|x_n - x_l\| \leq \|Ax_n - Ax_l\| \to 0 \text{ as } n, l \to \infty$$

Thus x_n is a Cauchy sequence in H and therefore there exists $x \in H$ such that $x_n \to x$. Because of continuity of A we conclude that $Ax_n \to Ax = y$ because the limit in unique and thus $y \in \text{Ran}A$. Thus Ran$A$ is indeed a closed (linear) subspace of H and we decompose our Hilbert space as $H = \text{Ran}A \oplus \text{Ran}A^{\perp}$.

Assume now for contradiction that Ran$A \neq H$. Then there exists $x_0 \in \text{Ran}A^{\perp}$ with $x_0 \neq 0$ with $(y, x_0) = 0$ for all $y \in \text{Ran}A$ and in particular $(Ax_0, x_0) = 0$ which implies that

$$\underbrace{\text{Re } (Ax_0, x_0)}_{=0} \geq c_0\|x_0\| > 0, \text{ because } x_0 \neq 0.$$

Hence Ran$A = H$ and A is bijective. Thus A^{-1} is well defined and linear. If we use again the inequality established in the injectivity part we get for all $x \in H$

$$c_0\|A^{-1}x\| \leq \|A(A^{-1}x)\| = \|x\| \implies \|A^{-1}\| \leq \frac{1}{c_0}$$

□

We next show how to use the Theorem of Lax-Milgram to prove the existence and uniqueness of solutions of certain partial differential equations.

Abstract setting: We work with $a : H \times H \to \mathbb{K}$ as in the Theorem of Lax-Milgram and fix $x \in H^*$. The unique solution of the problem

$$a(x, y) = x^*(y), \text{ for all } y \in H$$

is given by $x = A^{-1} R_H^{-1} x^*$. Indeed we have, thanks to Lax Milgram, for all $y \in H$

$$a(x, y) = (Ax, y) = (AA^{-1} R_H^{-1} x^*, y) = (R_H^{-1} x^*, y) = R_H(R_H^{-1} x^*)(y) = x^*(y)$$

Moreover we notice that the problem $a(x, y) = x^*(y)$ is stable in the following sense, given $x_1^*, x_2^* \in H^*$ the corresponding solutions $x_1 = A^{-1} R_H^{-1} x_1^*, x_2 = A^{-1} R_H^{-1} x_2^*$ have distance

$$\|x_1 - x_2\|_H = \|A^{-1} R_H^{-1}(x_1^* - x_2^*)\| \leq \frac{1}{c_0} \|x_1^* - x_2^*\|_{H^*}$$

where in the last inequality we used Lax-Milgram Thm and the fact that R_H^{-1} is an isometry. We apply now this general theory to solve so called elliptic boundary value problems.

Elliptic boundary value problem: Let $\Omega \subset \mathbb{R}^n$ be open and bounded, we consider functions on Ω with values in $\mathbb{K} = \mathbb{R}$. We look for a function $u \in C^2(\Omega)$, solving hte partial differential equation

$$-\sum_{i=1}^{n} \partial_i \sum_{j=1}^{n} a_{ij} \partial_j u + bu + f = 0 \tag{*}$$

Here $b, f \in C(\Omega)$ have values in $\mathbb{R}$ and the coefficients $a_{ij} = a_{if}(x) \in C^1(\Omega)$ are real valued and satisfy the positivity condition

$$\sum_{i,j=1}^{n} a_{ij}(x) z_i z_j \geq c_0 |z|^2, \text{ for all } z \in \mathbb{R}^n, x \in \Omega$$

We say that the matrix $(a_{ij}(x))$ satisfying the above condition is uniformly elliptic. We can also notice that (*) is a generalization of the Poisson equation $(-\Delta + b)u = -f$.

Boundary conditions: In order to find a unique solution of (*) we have to specify boundary conditions. We investigate two kind of boundary conditions.

- **Dirichlet boundary conditions:** In this case, we lok for $u \in C^2(\Omega) \cap C(\overline{\Omega})$ solving (*) in Ω and so that $u = g$ on $\partial\Omega$ for a given $g \in C(\partial\Omega)$.
- **Neumann boundary conditions:** In this case, we look for $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ solving (*) in Ω and such that

  $$-\sum_{i=1}^{n} \nu_i \sum_{j=1}^{n} a_{ij} \partial_j u = g, \text{ on } \partial\Omega \text{ for a given } g \in C(\partial\Omega)$$

  Here $\nu = (\nu_1, \dots, \nu_)$ denotes the outwards pointing normal to $\partial\Omega$.

Remark 4.5. In Dirichlet boundary value problems, we specify the values of u on the boundary $\partial\Omega$. In the Neumann boundary value problems, we specify the values of the normal derivates of u on the boundary $\partial\Omega$.

Reduction to a homogeneous problem: We can show that it is enough to solve the boundary value problems with homogenoeous boundary conditions, i.e. $g = 0$ on $\partial\Omega$.

Integral formulation of boundary value problems: We can show that the elliptic problem (*) is equivalent to solving (w.l.o.g. boundary $g = 0$ by the above)

$$\int \left(\sum_{i=1}^{n} \partial_i \xi \sum_{j=1}^{n} a_{ij} \partial_j u + \xi(bu + f) \right) = 0, \text{ for all } \xi \in C_0^\infty(\Omega)$$

Weak solutions of boundary value problems: Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $b \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$ and $a_{ij} \in L^\infty(\Omega)$ such that the uniform elliptic condition holds for almost all $x \in \Omega$.

- We call u a **weak solution of the homogeneous Dirichlet problem**, if $u \in H_0^1(\Omega) = H_0^{1,2}(\Omega)$ and

  $$\int \left(\sum_{i=1}^{n} \partial_i \xi \sum_{j=1}^{n} a_{ij} \partial_j u + (bu + f)\xi \right) = 0, \text{ for all } \xi \in H_0^1(\Omega)$$

  (since $u \in H^1(\Omega)$ enough to verify the above for all $\xi \in C_0^\infty(\Omega)$).
- We call u a **weak solution of the homogeneous Neumann problem**, if $u \in H^1(\Omega)$ and

  $$\int \left(\sum_{i=1}^{n} \partial_i \xi \sum_{j=1}^{n} a_{ij} \partial_j u + (bu + f)\xi \right) = 0, \text{ for all } \xi \in H^1(\Omega)$$

  (enough to verify for all $\xi \in C^\infty(\overline{\Omega})$)

Relationship between weak and classical solutions of boundary value problems: So far, we argued that every classical solution of the original boundary value problem is also a weak solution. It is natural to ask whether weak solutions, satisfying the Dirichlet/Neumann boundary value problems in a weak sense, are also classical solutions. To prove that this is the case, one needs to know that weak solutions are regular enough (in order to be able to integrate by parts). One should keep in mind, that once the existence and uniqueness of weak solution is know, one can typically conclude existence and uniqueness of classical solutions of boundary value problems, proving that weak solutions are regular and therefore also classical solutions.

Existence and uniqueness of weak solutions of boundary value problems: Making use of the Lax-Milgram theorem we can prove, the existence and uniqueness of the weak solutions of Neumann boundary value problems.

Theorem 4.8. *We assume that all conditions for the NBVP are satisfied. Furthermore, we assume that there is $b_0 > 0$ with $b(x) \geq b_0$ for almost every $x \in \Omega$. Then there exists exactly one weak solution $u \in H^1(\Omega)$ of the Neumann boundary value problem. Moreoever, there exists $C > 0$ with $\|u\|_{H^1(\Omega)} \leq C\|f\|_2$.*

For weak solutions of Dirichlet problems, the same result holds true. In fact, for Dirichlet problems, we do not need to assume that b is strictly positive, $b \geq 0$ is enough. This is a consequence of the Poincaré inequality.

Theorem 4.9. *We assume that all conditions in the definition of weak solutions for DBVP are satisfied. We assume furthermore that $b(x) \geq 0$ for almost every $x \in \Omega$. Then there exists exactly one weak solution $u \in H_0^1(\Omega)$ of the Dirichlet problem. Moreover, there is a constant $C > 0$ with $\|u\|_{H^1(\Omega)} \leq C\|f\|_2$.*

5 Baire Category Theorem and Consequences

5.1 Theorems of Baire and of Banach-Steinhaus

Lemma 5.1. *Let (X, τ) be a topological space The following statements are equivalent.*

1. *Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of closed sets in X. If the interior of each A_i is empty, then also the interior of $\cup_{j=1}^{\infty} A_i$ is empty.*

2. *Let $(B_i)_{i\in\mathbb{N}}$ be a sequence of open sets in R. If each B_i is dense in X, then also $\cap_{j=1}^{\infty} B_j$ is dense in X.*

Proof. The equivalence follows from the remark that a set is dense if and only if its complement has an empty interior, then apply De Morgan. $\square$

Definition 5.1. *A topological space (X, τ) is called a Baire's space, if condition 1) or condition 2) (and thus both) are satisfied.*

Theorem 5.1 (Baire)**.** *Every complete metric space is a Baire space.*

Proof. Let (X, d) be a complete metric space with $X \neq \emptyset$. Let $(B_i)_{i\in\mathbb{N}}$ be a sequence of open dense sets in X. We have to show that $L := \cap_{j=1}^{\infty} B_j$ is again dense in X. To this end, it is enough to show that for every non-empty open set $G \subset X$, we have $G \cap L \neq \emptyset$.

Since B_1 is open and dense, $G \cap B_1 \neq \emptyset$ and open. Hence, we can find $\epsilon_1 \in (0, 1]$ and $x_1 \in X$ such that

$$\overline{B_{\epsilon_1}(x_1)} \subset B_1 \cap G.$$

Since B_2 is open and dense, $B_{\epsilon_1}(x_1) \cap B_2$ is also non-empty and open. Hence we can find $\epsilon_2 \in (0, 1/2]$ and $x_2 \in X$ with

$$\overline{B_{\epsilon_2}(x_2)} \subset B_{\epsilon_1}(x_1) \cap B_2.$$

Iteratively, we find $\epsilon_n \in (0, 1/n]$ and $x_n \in X$ with

$$\overline{B_{\epsilon_n}(x_n)} \subset B_{\epsilon_{n-1}}(x_{n-1}) \cap B_n.$$

The sequence $(x_n)_{n\in\mathbb{N}}$ is then a Cauchy sequence in X. In fact, by construction we have for every $m \geq n$ that $B_{\epsilon_m}(x_m) \subset B_{\epsilon_n}(x_n)$ and therefore $d(x_n, x_m) \leq \epsilon_n \leq 1/n \to 0$ as $n, m \to \infty$. Since X is complete, the limit $x = \lim x_n$ exists.

From $x_k \in \overline{B_{\epsilon_n}(x_n)}$ for all $k \geq n$, we deduce that $x \in \overline{B_{\epsilon_n}(x_n)}$ for all $n \in \mathbb{N}$. Hence

$$x \in G \cap \bigcap_{j=1}^{\infty} B_j = G \cap L \implies G \cap L \neq \emptyset.$$

$\square$

As a first application of Baire's Theorem, we prove the Theorem of Banach-Steinhaus.

Theorem 5.2 (Banach-Steinhaus)**.** *Let X be a Banach space, Y be a normed space and $\mathcal{F} \subset \mathcal{L}(X,Y)$. Assume that for all $x \in X$, there exists $c_x \geq 0$ with*

$$\sup_{T \in \mathcal{F}} \|Tx\| \leq c_x.$$

Then there exists $c \geq 0$ with

$$\sup_{T \in \mathcal{F}} \|T\| \leq c.$$

Proof. For $k \in \mathbb{N}$ we consider the set

$$A_k := \{x \in X : \|Tx\| \leq k \text{ for all } T \in \mathcal{F}\}.$$

Then A_k is closed for all $k \in \mathbb{N}$. In fact, if $(x_j)_{j \in \mathbb{N}}$ denotes a sequence in A_k with $x_j \to x$ in X, then we have for every $T \in \mathcal{F}$,

$$\begin{aligned} \|Tx\| &\overset{\Delta}{\leq} \|Tx_j\| + \|T(x - x_j)\| \leq k + \|T\|\|x - x_j\| \\ &\implies \|Tx\| \leq k \text{ for } j \in \mathbb{N} \text{ sufficiently large} \end{aligned}$$

Since this holds for all $T \in \mathcal{F}$, we conclude that

$$\sup_{T \in \mathcal{F}} \|Tx\| \leq k, \text{ i.e. } x \in A_k \text{ and thus } A_k \text{ is closed.}$$

We have that

$$\bigcup_{k=1}^{\infty} A_k = X.$$

From Baire's Category Theorem (see condition 1, contraposition of that) we obtain that there exists $k_0 \in \mathbb{N}$ with $A_{k_0}^\circ \neq \emptyset$.

We find therefore $x_0 \in X$ and $\epsilon_0 >$ with $\overline{B_{\epsilon_0}(x_0)} \subset A_{k_0}^\circ$. Hence for any $y \in X$ with $\|y\| \leq \epsilon_0$ we get

$$\|Ty\| \leq \|T(y + x_0)\| + \|Tx_0\| \leq k_0 + c_{x_0}, \text{ for all } T \in \mathcal{F}.$$

Consequently we have for all $y \in X$ with $y \neq 0$ and for all $T \in \mathcal{F}$,

$$\|Ty\| = \left\| T\left(\frac{y}{\|y\|}\epsilon_0 \frac{\|y\|}{\epsilon_0}\right)\right\| = \frac{\|y\|}{\epsilon_0}\left\| T\left(\frac{y}{\|y\|}\epsilon_0\right)\right\| \leq \frac{k_0 + c_{x_0}}{\epsilon_0}\|y\|.$$

We conclude that

$$\|T\| = \sup_{\|y\| \leq 1} \|Ty\| \leq \frac{k_0 + c_{x_0}}{\epsilon_0} =: c$$

$\square$

5.2 The open map and the closed graph theorems

Theorem 5.3 (Open map theorem)**.** *Let X, Y be Banach spaces and let $T \in \mathcal{L}(X, Y)$ be surjective (i.e. $T(X) = Y$). Then $T(U)$ is open in Y for all open sets $U \subset X$.*

An equivalent formulation of the open map theorem is given by the following inverse map theorem.

Theorem 5.4 (Inverse Map Theorem)**.** *Let X, Y be Banach spaces, $T : X \to Y$ a continuous linear bijection. Then $T^{-1} \in \mathcal{L}(X, Y)$ i.e. is also linear and continuous, which implies that there exists $c > 0$ with*

$$\frac{1}{c}\|x\| \leq \|Tx\| \leq c\|x\|, \text{ for all } x \in X.$$

Proof. Clearly, T^{-1} is well-defined and linear. From the inverse mapping theorem we know that there exists $r > 0$ such that $T(B_1^X(0)) \supset B_r^Y(0)$, i.e. $T^{-1}(B_r^Y(0)) \subset B_1^X(0)$. This implies that for arbitrary $y \in Y$ we have

$$\|T^{-1}y\| = \frac{2\|y\|}{r}\left\|T^{-1}\frac{ry}{2\|y\|}\right\| \leq \frac{2}{r}\|y\|.$$

Hence T^{-1} is bounded and therefore continuous. □

A consequence of the inverse map theorem is the following closed graph theorem.

Theorem 5.5 (Closed Graph Theorem)**.** *Let X, Y be Banach spaces, $T : X \to Y$ linear. Then are the following statements equivalent:*

1. *graph$(T) = \{(x, TX) \in X \times Y : x \in X\}$ is closed in $X \times Y$ w.r.t. the norm $\|(x, y)\|_{X \times Y} := \|x\|_X + \|y\|_Y$.*
2. *T is continuous.*

Proof. "2) $\implies$ 1)". Let $(x, y) \in \overline{\text{graph}(T)}$. Then there exists a sequence $(x_k)_{k\in\mathbb{N}}$ in X with $(x_k, T_{x_k}) \to (x, y)$. Hence $x_k \to x$ in X and $T_{x_k} \to y$ in Y. Since T is continuous, we have that $y = Tx$ and thus $(x, y) \in \text{graph}(T)$. Thus graph$(T)$ is closed.

"1) $\implies$ 2)" Define $\phi : X \to X \times Y$ through $\phi(x) = (x, Tx)$. The image of ϕ is then given by

$$R(\phi) := \{\phi(x) : x \in X\} = \text{graph}(T)$$

and therefore, by assumption, is a closed subset of $X \times Y$. Hence $R(\phi)$ is a Banach space w.r.t. the norm $\|(x, y)\| = \|x\|_X + \|y\|_Y$ and the map $\phi : X \to R(\phi)$ is a linear bijection between Banach spaces. We want to show that ϕ is continuous, to this end we discuss it's inverse. The inverse map $\phi^{-1} : R(\phi) \to X$ is given by $\phi^{-1}(x, Tx) = x$ is also a linear bijection. Moreover, since

$$\|\phi^{-1}(x, Tx)\|_X = \|x\|_X \leq \|(x, Tx)\|_{X \times Y}$$

we have that ϕ^{-1} is a continuous linear bijection. The inverse map theorem implies that also $(\phi^{-1})^{-1} = \phi$ is continuous. Since the projection $p : X \times Y \to X$, defined as $p(x, y) = y$ is clearly continuous, we conclude that $T = p \circ \phi$ is continuous as well. $\square$

Definition 5.2. *Let X be a Banch space. A continuous map $P : X \to X$ with $P^2 = P \circ P = P$ is called a projection on X. It is easy to see that if P is a projection, then also $1 - P$ is a projection.*

Theorem 5.6. *Let X be a Banach space and $P \in \mathcal{L}(X)$ a projection. Then $X = P(X) \oplus (1 - P)(X)$.*

Proof. Let us define $A = \text{Ran } P = \{Px : x \in X\}$ and $B = \text{Ran } (1 - P) = \{(1 - P)x : x \in X\}$. Then we have:

• A is closed. In fact, if $(y_k)_{k \in \mathbb{N}}$ is a sequence in A with $y_k \to y$ in X, then by definition $y_k = Px_k$ for appropriate $x_k \in X$. Hence

$$Py_k = P(Px_k) = P^2 x_k = Px_k = y_k \text{ for all } k \in \mathbb{N}.$$

if we now let $k \to \infty$ we find, since $P \in \mathcal{L}(X)$ is continuous, that $Py = y$ and thus $y \in A$ i.e. A is closed.

• B is also closed, using the same argument as above with P replaced by $1 - P$.

• $X = A + B$ is clear, since $x = Px + x - Px = Px + (1 - P)x$.

• $A \cap B = \{0\}$. If $x \in A$ then we have $x = Px$ (because if $x = Py$ then $Px = P^2 y = Py = x$) and if $x \in B$ we have $x = (1 - P)x$, thus $x = Px = P(1 - P)x = 0$. $\square$

It follows from the last theorem that finite dimensional subspaces always have a topological complement.

Theorem 5.7. *Let X be a Banach space, $A \subset X$ a finite dimensional linear subspace. Then there exists a closed subspace $B \subset X$ with $X = A \oplus B$.*

6 Weak Topologies on normed spaces

6.1 The weak and the weak-* topologies

Let X be a set and let $\mathcal{F}$ denote a family of maps $f : X \to Y_f$, where Y_f is a topological space. On X, we introduce the topology $\tau_{\mathcal{F}}$, defined as the smallest topology on X so that all functions in $\mathcal{F}$ are continuous. We know that the arbitrary union of a family of topologies on X is again a topology on X, thus for

$$\mathcal{S} := \{f^{-1}(V) : V \subset Y_f \text{ open}, f \in \mathcal{F}\}$$

we define

$$\tau_{\mathcal{F}} = \bigcap \{\tau : \tau \text{ is a topology on } X \text{ and } \mathcal{S} \subset \tau\}$$

Then $\tau_{\mathcal{F}}$ is the smallest topology on X, with the property that all maps in $\mathcal{F}$ are continuous. As always, we care about the Hausdorff property.

Proposition 6.1. *Let X be a set and $\mathcal{F}$ a family of maps $f : X \to Y_f$, with Y_f Hausdorff for every $f \in \mathcal{F}$. Then $\mathcal{F}$ separates the points of X and $(X, \tau_{\mathcal{F}})$ is again a Hausdorff space.*

We now specialize our construction to the case of normed spaces.

Definition 6.1. *Let X be a normed space, X^* its dual space and X^{**} the bidual space. The smallest topology on X, with the property that all $f \in X^*$ are continuous is called the weak topology on X, it is denoted by τ_W.*
Since X^* is again a normed space, we can introduce also on X^* the weak topology, defined as the smallest topology with the property that all $f \in X^{**}$ are continuous. Moreover we define on X^* the weak- topology, denoted by T_W^*, it is defined as the smallest topology such that all $f \in J_X(X) \subset X^{**}$ are continuous where $J_X : X \to X^{**}$ is the canonical inclusion.*

Remark 6.1.

1. On X we have $\tau_W \subset T_{\|\cdot\|}$, i.e. convergence (w.r.t. the norm on X) always implies weak convergence. This follow because by definition $f \in X^*$ is always continuous w.r.t. $\tau_{\|\cdot\|}$ and because T_W is the smallest topology on X such that $f \in X^*$ is continuous.

2. On X^* we have $\tau_W^* \subset \tau_W \subset \tau_{\|\cdot\|}$, because $J_X(X) \subset X^{**}$.

3. If X is reflexive, then $J_X(X) = X^{**}$ and $\tau_W = \tau_W^*$ as topologies on X^*.

On infinite dimensional vector spaces, norm- and weak topologies are never the same (i.e. not equivalent), as the next Lemma entails

Lemma 6.1. *Let X be a normed space. Then $\tau_W = \tau_{\|\cdot\|}$ if and only if $\dim X < \infty$. If $\dim X = \infty$, then every τ_W-open neighbourhood of 0 contains an infinitely dimensional linear subspace of X.*

6.2 The notion of weak convergence

Let X be a normed space. As every topology, the weak topology τ_W induces a notion of convergence of sequences in X. In this context, we say that a sequence x_k in X converges weakly towards $x \in X$ and we write $x_k \rightharpoonup x$., if the sequence converges w.r.t. τ_W.

Remark 6.2.

1. Since (X, τ_W) and (X^*, τ_W^*) are both Hausdorff, weak and weak-* limits are always unique if they exist.

2. Since $\tau_W \subset \tau_{\|\cdot\|}$, $x_k \to x$ w.r.t. $\tau_{\|\cdot\|}$ always implies $x_k \rightharpoonup x$. Analogously, since $\tau_W^* \subset \tau_W \subset \tau_{\|\cdot\|}$ both, norm and weak convergence, imply weak-* convergence on X^*.

Lemma 6.2 (Characterisation of weak convergence)**.** *Let X be a normed space. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in X, then we have $x_k \rightharpoonup x$ if and only if $f(x_k) \to f(x)$ for all $f \in X^*$.*

The Lemma above says that we can identify weak convergence with convergence w.r.t. the norm on $\mathbb{K}$.

Lemma 6.3 (Characterisation of weak-* convergence)**.** *Let $(f_i)_{i\in\mathbb{N}}$ be a sequence in X^*. Then f_i convergence towards $f \in X^*$ w.r.t. to τ_W^* if and only if $f_i(x) \to f(x)$ for all $x \in X$.*

Despite the fact, that in general weak convergent sequences do not converge w.r.t. the norm topology, they are always bounded. This is a consequence of the Banach-Steinhaus theorem and subject of the next Proposition.

Proposition 6.2. *Let X be a normed space, $(x_k)_{k\in\mathbb{N}}$ be a sequence in X with $x_k \rightharpoonup x$. then $(x_k)_{k\in\mathbb{N}}$ is bounded and moreover we have*

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

Proof. By assumption $f(x_k) \to f(x)$ for all $f \in X^*$. Hence, for every $f \in X^*$ there exists $c_f > 0$ such that $|J(x_k)(f)| = |f(x_k)| \leq c_f$ and thanks to Banach Steinhaus there exists $c > 0$ such that $\|J(x_k)\| \leq c$ which implies that $\|x_k\| \leq c$. From Hahn-Banach there exists $f \in X^*$ with $\|f\| = 1$ and $f(x) = \|x\|$ and thus we have

$$\|x\| = f(x) = \lim f(x_k) \leq \liminf \|f\|\|x_k\| = \liminf \|x_k\|$$

□

6.3 Weak-* topology and compactness

For finite dimensional normed spaces, it is very easy to characterize compact sets thanks to the Heine Borel Theorem, a set M is compact if and only if M is bounded and closed. For infinite dimensional normed spaces, this is no longer true. We have also seen in Proposition 3.1 that the closed unit ball $\overline{B_1(0)}$ is compact if and only if $\dim X < \infty$, in particular in the infinite dimensional case $\overline{B_1(0)}$ is neither covering nor sequentially compact.

Weak topologies are important because often they allow us to "restore" the compactness of the closed ball. The statement that the closed unit ball in X^* is compact w.r.t. τ_W^* is known as the Banach-Alaoglu Theorem.

Theorem 6.1 (Tychonov Theorem)**.** *Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of compact sets. Then the product $X = \prod_{\lambda\in\Lambda} X_\lambda$ is compact, w.r.t. the product topology defined on X.*

Remark 6.3. The product topology on X is generated by product sets having the form

$$\prod_{\lambda\in\Lambda} U_\lambda = \{(x_\lambda)_{\lambda\in\Lambda} : x_\lambda \in U_\lambda\}, \text{ where } U_\lambda \subset X_\lambda \text{ are open for all } \lambda \in \Lambda.$$

Equivalently, the product topology is the smallest topology on X, so that all projections $p_\lambda : X \to X_\lambda$, defined through $p((x_\lambda)_{\lambda\in\Lambda}) = x_\lambda$, are continuous.

Theorem 6.2 (Banach-Alaoglu)**.** *Let X be a normed space and X^* its dual space. Then the closed unit ball $k_{X^*} = \{f \in X^* : \|f\| \le 1\}$ in X^* is compact w.r.t the weak-* topology τ_W^*.*

Proof. For $x \in X$ let $D_x := \{\alpha \in \mathbb{K} : |\alpha| \le \|x\|\}.$, then D_x is a compact space (in $\mathbb{K}$). If we set $P = \prod_{x\in X} D_x$ and equip it with the product topology τ_P then thanks to Tychonov's Theorem P is also compact. We have $k_{X^*} = X^* \cap P$ because elements in P can be thought of as maps $\varphi(x) = \alpha_x$ i.e. $|\varphi(x)| \le \|x\|$.

We show that k_{X^*} is a closed subset of P w.r.t. τ_P which implies that it's compact w.r.t to τ_P. This follows because for $\sigma_{x,y,\lambda} : P \to \mathbb{K}$ given by

$$\sigma_{x,y,\lambda}(\varphi) = \varphi(x + \lambda y) - \varphi(x) - \lambda\varphi(y)$$

is continuous w.r.t to τ_P and since $\{0\} \subset \mathbb{K}$ is closed it also follows that $\sigma^{-1}_{x,y,\lambda}(\{0\})$ is closed in P but then

$$k_{X^*} = \bigcap_{x,y\in X,\lambda\in\mathbb{K}} \sigma^{-1}_{x,y,\lambda}(\{0\}), \text{ is closed in P}$$

We then show that $\tau_{W^*} |_{k_{X^*}} \subset \tau_P |_{k_{X^*}}$ and conclude the statement. $\square$

Since in general the weak-* topology is not metrizable, the notion of sequentially compactness is not equivalent to the notion of (covering) compactness. In fact, in general the unit ball in X^* is not sequentially compact w.r.t τ_W^*. The next theorem shows, however, that the unit ball in X^* is sequentially compact, under the assumption that X is also seperable.

Theorem 6.3. *Let X be a separable normed space. Then the closed unit ball k_{X^*} in X^* is sequentially compact, w.r.t. the τ_W^* topology.*

6.4 Reflexivity and compactness

In this section, we consider reflexive Banach spaces and we show that, under appropriate assumptions, the closed unit ball in X is compact and sequential compact w.r.t the τ_W topology.

Proposition 6.3. *Let X be a Banach space. Then X is reflexive if and only if $k_X = \{x \in X : \|x\| \leq 1\}$ is compact w.r.t the τ_W topology.*

For reflexive Banach spaces, k_X is not only compact, it is also sequentially compact w.r.t. the topology τ_W.

Theorem 6.4. *Let X be a reflexive Banach space. Then the closed unit ball k_X is sequentially compact w.r.t. the topology τ_W. In other words, every bounded sequence in X has a τ_W-convergent subsequence.*

6.5 Weak topology and convexity

A weakly convergent sequence $(x_n)_{n\in\mathbb{N}}$ in a Banach space X has one additional neat property, that is often useful. If x_n converges weakly, there exists namely another sequence, consisting of convex combinations of the elements of x_n, converging strongly (i.e. w.r.t. the norm topology). This results is known as the Lemma of Mazur.

Theorem 6.5 (Separation theorem)**.** *Let X be a normed space over $\mathbb{K}$ and $M \subset X$ closed and convex and $x_0 \in X \setminus M$. Then there exists $F \in X^*$ and $\alpha \in \mathbb{R}$ with*

$$\begin{aligned} Re\ F(x) &\leq \alpha, \text{ for all } x \in M \text{ and} \\ Re\ F(x_0) &> \alpha \end{aligned}$$

i.e. M and x_0 are divided/separated by the hyperplane $Re\ F(x) = \alpha$.

Proof. Consider only $\mathbb{K} = \mathbb{R}$. W.l.o.g. assume that 0 is contained in the interior of M and thus $\exists \rho > 0$ such that $B_\rho(0) \subset M$. We define then the Minkowski function $p(x) := \inf\{r > 0 : x/r \in M\}$ for all $x \in X$. We can then show that p is a sublinear function and satisfies $p(x) \leq 1$ for all $x \in M$ and $p(x_0) > 1$. Moreover we have $p(x) \leq 1/\rho\|x\|$

Let now $f : \mathbb{R} \cdot x_0 \to \mathbb{R}$ be given by $f(\lambda x_0) = \lambda p(x_0)$ then we can show that $f(x) \leq p(x)$ for all $x \in \mathbb{R} \cdot x_0$ and thus from Hahn-Banach theorem there exists $F : X \to \mathbb{R}$ linear with $F_{|\mathbb{R}\cdot x_0} = f$ and $F(X) \leq p(x)$ for all $x \in X$. Therefore $F(X) \leq 1$ for all $x \in M$ and $F(x_0) = p(x_0) > 1$. Moreover we can easily show that $|F(x)| \leq (1/\rho)\|x\|$ and thus $F \in X^*$. $\square$

Theorem 6.6. *Let X be a normed space over $\mathbb{K}$ and $A \subset X$ convex. Then $\overline{A}^{\tau_{\|\cdot\|}} = \overline{A}^{\tau_W}$, i.e. the closure of A w.r.t. the (strong) norm topology is the same as the closure of A w.r.t. the weak topology.*

Proof. Consider only $\mathbb{K} = \mathbb{R}$ Since $\tau_W \subset \tau_{\|\cdot\|}$ we have $\overline{A}^{\tau_W} \supset \overline{A}^{\tau_{\|\cdot\|}}$. Let now $x_0 \in \overline{A}^{\tau_W}$ and assume for contradiction that $x_0 \notin \overline{A}^{\tau_{\|\cdot\|}}$. Now $\overline{A}^{\tau_{\|\cdot\|}}$ is closed and convex w.r.t τ_W we get from the separation theorem that there are $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f(x) \leq \alpha$ for all $x \in \overline{A}^{\tau_{\|\cdot\|}}$ and $f(x_0) > \alpha$. Since

$$(f^{-1}((-\infty, \alpha]))^c = \{x \in X : f(x) > \alpha\} = f^{-1}((\alpha, \infty)) \in \tau_W$$

we get that $f^{-1}((-\infty, \alpha])$ is τ_W closed and thus $\overline{A}^{\tau_W} \subset f^{-1}((-\infty, \alpha])$ since the closure of A w.r.t τ_W is the smallest closed set containing A. But we assumed that $x_0 \in \overline{A}^{\tau_W} \subset f^{-1}((-\infty, \alpha])$ and thus $f(x_0) < \alpha$ which is a contradiction to $f(x_0) > \alpha$. $\square$

Corollary 6.1 (Mazur's Lemma). *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a Banach space $(X, \|\cdot\|)$, converging weakly towards x. Then there exists a sequence $(y_k)_{k\in\mathbb{N}}$ so that every y_k is a finite convex combination of the elements x_n, i.e. for every $k \in \mathbb{N}$ we have*

$$y_k = \sum_{n\in\mathbb{N}} t_n^{(k)} x_n, \text{ s.t. for every fixed } k \in \mathbb{N},\ t_n^{(k)} \neq 0 \text{ for only finitely many } n$$

$$\text{and } \sum_{n\in\mathbb{N}} t_n^{(k)} = 1$$

and $y_k \to x$ wr.t. the norm topology (i.e. strongly).

Proof. Let $A = \{$finite convex combinations of $(x_n)_{n\in\mathbb{N}}\}$. Then A is convex and thanks to the above theorem we have $\overline{A}^{\tau_{\|\cdot\|}} = \overline{A}^{\tau_W}$. Since $x_n \rightharpoonup x$ we have $x \in \overline{A}^{\tau_W} = \overline{A}^{\tau_{\|\cdot\|}}$. Hence, there exists a sequence $(y_k)_{k\in\mathbb{N}}$ in A such that $y_k \to x$, because the norm closure of A is exactly the set of all norm limits of sequences in A. $\square$

7 Spectral Theorem for Compact Operators

In this entire section we consider bounded linear operators defined on a Banach space over $\mathbb{C}$.

7.1 The spectrum of bounded operators

Let X be a Banach space over $\mathbb{C}$.

Definition 7.1. *Let $T \in \mathcal{L}(X)$. We define the resolvent set of T as*

$$\rho(T) = \{\lambda \in \mathbb{C} : \ker(\lambda Id - T) = \{0\} \text{ and } ran(\lambda Id - T) = X\}$$

i.e. $\rho(T)$ is the set of all $\lambda \in \mathbb{C}$ such that $\lambda Id - T$ is invertible. The spectrum of T is then defined as the complement of the resolvent set, i.e. $\sigma(T) = \mathbb{C} \setminus \rho(T)$. The spectrum $\sigma(T)$ can be further decomposed into the point spectrum

$$\sigma_p(T) = \{\lambda \in \mathbb{C} : \ker(\lambda Id - T) \neq \{0\}\},$$

the continuous spectrum

$$\sigma_c(T) = \{\lambda \in C : \ker(\lambda Id - T) = \{0\}, ran(\lambda Id - T) \neq X, \text{ but } \overline{ran(\lambda Id - T)} = X\}$$

and finally the residual spectrum

$$\sigma_r(T) = \{\lambda \in \mathbb{C} : \ker(\lambda Id - T) = \{0\}, \ \overline{ran(\lambda Id - T)} \neq X\}$$

Remark 7.1.

1. We have $\lambda \in \sigma_p(T)$ if and only if there exists $x \neq 0$ with $Tx = \lambda x$. Such an x is called an eigenvector of T associated with the eigenvalue λ. We call $\ker(\lambda - T)$ the eigenspace of T with eigenvalue λ.

2. For $\lambda \in \rho(T)$, the operator $(\lambda - T)$ is bijetive and invertible. It follows from the inverse function theorem that the inverse $(\lambda - T)^{-1}$ is automatically continuous and thus contained in the space $\mathcal{L}(X)$.

Definition 7.2. *For $\lambda \in P(T)$ we define $R_T(\lambda) = (\lambda - T)^{-1} \in \mathcal{L}(X)$ (defined on $\rho(T)$), the function R_T is called the resolvent of T.*

Theorem 7.1. *The resolvent set $\rho(T)$ of T is open and the resolvent $R_T(\lambda)$ is an analytic function on $\rho(T)$ with values in $\mathcal{L}(X)$, such that $\|R_T(\lambda)\| \geq 1/dist(\lambda, \sigma(T))$.*

Remark 7.2. Recall that a map $F : D \to Y$, where $D \subset \mathbb{C}$ is open and Y is a Banach space, is called analytic if for all $\lambda \in D$, there exists a ball $B_{r_0}(\lambda) \subset D$ such that

$$F(\mu) = \sum_{n=0}^{\infty} A_n (\mu - \lambda)^n$$

where $A_n \in Y$ for all $n \in \mathbb{N}$ and the series above is absolutely convergent.

In order to prove the previous Theorem we will use the following lemma.

Lemma 7.1 (Von Neumann Series)**.** *Let $T \in \mathcal{L}(X)$, with $\|T\| < 1$. Then $1 - T$ is bijective, with $(1-T)^{-1} \in \mathcal{L}(X)$ and*

$$(1-T)^{-1} = \sum_{j=0}^{\infty} T^n$$

Proof. Let $S_l = \sum_{n=0}^{l} T^n$. Then for $k < l$ we have,

$$\|S_l - S_k\| \leq \sum_{n=k+1}^{l} \|T^n\| \leq \sum_{n=k+1}^{l} \|T\|^n \to 0$$

as $k, l \to \infty$, because $\|T\| < 1$ and the geometric series converges, hence the tail has to vanish. In other words, S_l is a Cauchy sequence in $\mathcal{L}(X)$ and since X is a Banach space, we know that $\mathcal{L}(X)$ is complete. Hence S_l convergeces, let $S := \lim S_l$. Then we obtain

$$(1-T)S = \lim_{l\to\infty} (1-T)S_l \overset{\text{Tele.}}{=} \lim_{l\to\infty} 1 - T^{l+1} = 1$$

because $\|T^{l+1}\| \leq \|T\|^{l+1} \to 0$ as $l \to \infty$ thanks to $\|T\| < 1$. $\square$

Theorem 7.2 (Spectral radius)**.** *The spectrum $\sigma(T)$ of T is a compact, non-emptyset (if $X \neq \{0\}$) and*

$$\sup_{\lambda\in\sigma(T)} |\lambda| = \lim_{m\to\infty} \|T^m\|^{1/m} \leq \|T\|$$

To conclude this section, we study the relation between the spectrum of $T \in \mathcal{L}(X)$ and the spectrum of the adjoint operator $T^* \in \mathcal{L}(X^*)$, defined through $(T^*f)(x) = f(Tx)$ for all $f \in X^*$ and $x \in X$.

Proposition 7.1. *Let $T \in \mathcal{L}(X)$. Then $\sigma(T) = \sigma(T^*)$.*

Proof. The proof follows from the fact that $T \in \mathcal{L}(X)$ is invertible if and only if T^* is invertible (covered in exercises). Then

$$\begin{aligned} \lambda \in \rho(T) &\iff \lambda - T \text{ invertible} \iff (\lambda - T)^* \text{ invertible} \\ &\iff \lambda - T^* \text{ invertible} \iff \lambda \in \rho(T^*) \end{aligned}$$

thus, since the resolvent sets are equal we have in particular $\sigma(T) = \sigma(T^*)$. $\square$

7.2 Compact Operators on Banach Spaces

If $\dim X < \infty$ the spectrum of $T \in \mathcal{L}(X)$ consists only of finitely many eigenvalues. If instead $\dim X = \infty$, the spectrum can be more complicated. In this section, we consider a class of bounded operators, known as compact operators, whose spectrum is similar to the case $\dim X < \infty$.

Definition 7.3. *Let X, Y be Banach spaces and $B_X = \{x \in X : \|x\| < 1\}$ denote the open unit ball in X. An operator $T \in \mathcal{L}(X,Y)$ is called compact if TB_X is pre-compact, i.e. if for all $\epsilon > 0$, TB_X has a finite covering with ϵ-balls. In other words, the operator T is comapct if $\overline{TB_X}$ is compact. Equivalently, T is compact, if for every bounded sequence $(x_n)_{n\in\mathbb{N}}$ in X, the sequence Tx_n in Y has a convergent subsequence. We denote by $\mathcal{K}(X,y)$ the set of all compact operators from X to Y. We also denote $\mathcal{K}(X) = \mathcal{K}(X,X)$.*

Example 7.1. *We say that the operator $T \in \mathcal{L}(X,Y)$ has finite rank, if* $\dim TX < \infty$*. Every finite rank operator is compact. Indeed, if x_n is a bounded sequence in X, then Tx_n is a bounded sequence in RanT. Since* $\dim$ *RanT* $< \infty$ *it follows that Tx_n has a convergent subsequence since RanT is closed (because dim Ran T $< \infty$). In particular, every continuous linear function $f \in X^*$ is compact, because* $\dim f(X) \leq 1$.

The next theorem gives some important properties of compact operators.

Theorem 7.3. *Let X, Y, Z be Banach spaces.*

1. *$\mathcal{K}(X,Y)$ is a closed subspace of $\mathcal{L}(X,Y)$.*
2. *$T \in \mathcal{K}(X,Y), S \in \mathcal{L}(Y,Z), R \in \mathcal{L}(Z,X)$. Then $ST \in \mathcal{K}(X,Z), TR \in \mathcal{K}(Z,Y)$.*
3. *Let $T \in \mathcal{L}(X,Y)$. Then $T \in \mathcal{K}(X,Y)$ if and only if $T^* \in \mathcal{K}(Y^*,X^*)$.*

We consider now the spectrum of compact operators. We will also use the notion of Fredholm operator.

Definition 7.4. *Let X, Y be Banach spaces. $A \in \mathcal{L}(X,Y)$ is called a Fredholm operator if* $\dim\ker(A) < \infty$, *Ran(A) is closed and codim Ran$(A) < \infty$. The index of A is then defined as ind$(A) = \dim\ker(A) -$ codim Ran(A).*

Remark 7.3. The codimension codim Z of a closed subspace $Z \subset X$ is the dimension of the quotient space $X/\sim$ where $\sim$ is the equivalence relation defined on X by $x \sim y :\iff x - y \in Z$. If Z has codimension $m < \infty$, then there exist linear independent $x_1, \dots, x_m \in X$ with $X = Z \oplus \operatorname{span}(x_1, \dots, x_m)$.

The next Theorem shows how we can get a Fredholm operator out of a compact operator.

Theorem 7.4. *Let $T \in \mathcal{K}(X)$. Then $A = 1 - T$ is a Fredholm operator with index* 0.

Proof. The proof is divided into 5 steps.
Step 1: $\dim\ker A < \infty$.
Let $Ax = 0$, then it follows that $Tx = x$. Hence $B_1(0) \cap \ker A \subset T(B_1(0))$. Since T is a compact operator, $T(B_1(0))$ is pre-compact, hence $B_1(0) \cap \ker A$ is also pre-compact. But we know that this can only be the case if $\dim\ker A < \infty$ from Prop. 3.1.

Step 2: Ran$A$ is closed.
Let $x \in \overline{\text{Ran}A}$ and $Ax_n \to x$ as $n \to \infty$. We can assume that

$$\|x_n\| \leq 2d_n, \text{ where } d_n = \text{dist}(x_n \ker A)$$

and we can show that d_n is bounded, thus x_n is bounded. Thus there is a subsequence x_{n_j} of x_n and $z \in X$ with $Tx_{n_j} \to z$ as $j \to \infty$. Hence

$$x = \lim_{j\to\infty} Ax_{n_j} = \lim_{j\to\infty} A(Ax_{n_j}Tx_{n_j}) = A(x+z) \implies x \in \text{Ran}A.$$

Step 3: $\ker A = \{0\}$ implies that $\text{Ran}A = X$.
Assume that $\ker A = \{0\}$ and for condratiction that there exists $x \in X \setminus \text{Ran}A$. Then we can shows that $A^n x \in \text{Ran}A^n \setminus \text{Ran}A^{n+1}$. We can then show that $\text{Ran}A^{n+1}$ is closed. We can then construct a sequence x_n in $\text{Ran}A^n$ which is bounded but $\|Tx_n - Tx_m\| \geq 1/2$ i.e. the sequence Tx_n has no convergent subsequence, although x_n is a bounded sequence in X, which is contradiction to the compactness of T.

Step 4: codim Ran $A \leq \dim\ker A$.
From Step 1 we know that $\dim\ker A < \infty$. Let $x_1, \ldots, x_n$ be a basis of $\ker A$. Then there are $f_1, \ldots, f_n \in X^*$ with $f_i(x_l) = \delta_{il}$. If the claim was wrong, there would be linearly independent $y_1, \ldots, y_n \in X$ such that $\text{span}\{y_1, \ldots, y_n\} \oplus \text{Ran}A \subset X$ would be a proper subspace of X. Then for

$$\tilde{T}x := Tx + \sum_{k=1}^{n} f_k(x)y_k$$

$\tilde{T} \in \mathcal{K}(X)$ and $\ker\tilde{A} = \{0\}$ where $\tilde{A} = 1 - \tilde{T}$. Step 3 then implies that $\text{Ran}\tilde{A} = X \subset \text{Ran}A \oplus \text{span}\{y_1, \ldots, y_n\}$ which is a condtradiction.

Step 5: $\dim\ker A \leq$ codim Ran$A$.
Let $A^* = 1 - T^* \in \mathcal{L}(X^*)$. Then by Step 4 we have $\text{Ran}A^* \leq \dim\ker A^*$. The claim follows if we can provfe that 1) $\dim\ker A \leq \dim\ker A^{**}$ and 2) $\dim\ker A^* =$ codim Ran$A$ because then we have

$$\dim\ker A \leq \dim\ker A^{**} = \text{codim Ran } A^* \leq \dim\ker A^* = \text{codim Ran } A.$$

□

We use the last theorem to describe the spectrum of compact operators.

Theorem 7.5 (Spectral theorem for compact operators; Riesz-Schauder)**.** *Let $T \in \mathcal{K}(X)$, then we have*

1. *The set $\sigma(T)\setminus\{0\}$ consists of countably (finite or infinite) many eigenvalues with 0 as the only possible accumulation point.*
2. *For $\lambda \in \sigma(T) \setminus \{0\}$ we have $\dim\ker(\lambda - T) < \infty$ and*
$$1 \le n_\lambda := \max\{n \in \mathbb{N} : \ker(\lambda - T)^{n-1} \neq \ker(\lambda - T)^n\} < \infty$$
n_λ is called the order of λ, $\dim\ker(\lambda - T)$ the multiplicity of λ.
3. *For $\lambda \in \sigma(T) \setminus \{0\}$ we have the decomposition*
$$X = Ran(\lambda - T)^{n_\lambda} \oplus \ker(\lambda - T)^{n_\lambda}.$$
Both subspaces are closed and invariant w.r.t. T and we have that $\dim\ker(\lambda - T)^{n_\lambda} < \infty$.
4. *For $\lambda \in \sigma(T) \setminus \{0\}$ we have $\sigma(T_{|Ran(\lambda-T)^{n_\lambda}}) = \sigma(T) \setminus \{\lambda\}$.*
5. *For $\lambda \in \sigma(T) \setminus \{0\}$ let E_λ denote the projection on $\ker(\lambda - T)^{n_\lambda}$ along $Ran(\lambda - T)^{n_\lambda}$. Then $E_\lambda E_\mu = \delta_{\lambda,\mu} E_\lambda$ for all $\lambda, \mu \in \sigma(T) \setminus \{0\}$.*

7.3 Normal compact operators

Let X be a Banach space and $T \in \mathcal{L}(X)$. We recall again that the adjoint operator $T^* \in \mathcal{L}(X^*)$ is defined through $T^*f(x) = f(Tx)$ for all $f \in X^*, x \in X$. We already know that $\|T^*\| = \|T\|$ and $(T+\lambda S)^* = T^*+\lambda S^*$, $(TS)^* = S^*T^*$ and also $\sigma(T^*) = \sigma(T)$, finally we have already seen that T is compact if and only if T^* is compact.

If H is a Hilbert space, then by the Riesz's Representation Theorem we can identify H^* with H through the map $R_H : H \to H^*$ which is an anti-linear isometric bijection between H and H^*. (Recall that $(R_H(x))(y) = \langle x, y\rangle_H$). Thus we can interprete the adjoint operator in this case as a map in $\mathcal{L}(H)$. In other words, we can define the adjoint to $T \in \mathcal{L}(H)$ as

$$T^\dagger = R_H^{-1} T^* R_H$$

Then $T^\dagger \in \mathcal{L}(H)$ is characterized by the equation for all $x \in H$:

$$\begin{aligned} R_H T^\dagger(x) = T^* R_H(x) &\iff (R_H T^\dagger x)(y) = (R_H x)(Ty), \text{ for all } y \in H \\ &\iff \langle T^\dagger x, y\rangle = \langle x, Ty\rangle, \text{ for all } y \in H. \end{aligned}$$

From the above we deduce that $T^\dagger \in \mathcal{L}(H)$ is the adjoint to $T \in \mathcal{L}(H)$ if and only if $\langle x, Ty\rangle = \langle T^* x, y\rangle$ for all $x, y \in H$.

It is easy to derive from the properties of $T^* \in \mathcal{L}(H^*)$, the corresponding properties of $T^\dagger \in \mathcal{L}(H)$. For $T, S \in \mathcal{L}(H), \lambda \in \mathbb{C}$ we have $(T + \lambda S)^\dagger = T^\dagger + \overline{\lambda} S^\dagger$ (because R_H is anti-linear). Moreover $(TS)^\dagger = S^\dagger T^\dagger$, $\|T^\dagger\| = \|T\|$ and T is compact if and only if $T^\dagger$ is compact and $\sigma(T^\dagger) = \overline{\sigma(T)}$.

Lemma 7.2. *Let H be a Hilbert space, $T \in \mathcal{L}(H)$. Then* $\ker T^\dagger = (RanT)^\perp$ *and* $RanT^\dagger = (\ker T)^\perp$

Proof. Follows easily from

$$\begin{aligned} x \in \ker T^\dagger &\iff T^\dagger x = 0 \iff \langle T^\dagger x, y \rangle = 0 \text{ for all } y \in H \\ &\iff \langle x, Ty \rangle = 0 \text{ for all } y \in H \iff x \perp \mathrm{Ran} T. \end{aligned}$$

□

Definition 7.5. *An operator $T \in \mathcal{L}(H)$ is called self-adjoint if $T^\dagger = T$ i.e. if $\langle x, Ty \rangle = \langle Tx, y \rangle$ for all $x, y \in H$. $T \in \mathcal{L}(H)$ is called normal if T and $T^\dagger$ commute, i.e. if $TT^\dagger = T^\dagger T$. Every self-adjoint operator is also normal.*

A couple of important properties of normal operators are collected in the next theorem.

Theorem 7.6. *Let $T \in \mathcal{L}(H)$. Then*

1. *T normal implies that $T - \lambda$ is normal for all $\lambda \in \mathbb{C}$.*
2. *T is normal if and only if $\|Tx\| = \|T^\dagger x\|$ for all $x \in H$.*
3. *T normal implies that* $\ker(\lambda - T) = \ker(\overline{\lambda} - T^\dagger)$.

Another important property of normal operators is proved in the following lemma.

Lemma 7.3. *Let $T \in \mathcal{L}(H)$ be normal. Then $\|T^m\| = \|T\|^m$ for all $m \in \mathbb{N}$.*

With the help of the above Lemma we immediately arrive that

Corollary 7.1 (Spectral radius for normal operators)**.** *Let $T \in \mathcal{L}(H)$ be normal. Then*

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

Theorem 7.7 (Spectral theorem for normal compact operators)**.** *Let H be a Hilbert space over $\mathbb{C}, T \in \mathcal{K}(H)$ be also normal, $T \neq 0$. Then we have*

1. *There exists an orthonormal system $(e_k)_{k\in N}$ in H and a sequence $(\lambda_k)_{k\in N}$ in $\mathbb{C}$ with $N \subset \mathbb{N}$, such that $\lambda_k \neq 0$ for all $k \in N$ and*

   $$\begin{aligned} Te_k &= \lambda_k e_k, \ \textit{for all } k \in N \\ \sigma(T) \setminus \{0\} &= \{\lambda_k : k \in N\} \end{aligned}$$

   Moreover, if $|N| = \infty$, then $\lambda_k \to 0$ (notice that in this statement, the λ_k do not need to be all different).

2. *For all $k \in N$, we have*

   $$n_{\lambda_k} = \max\{n \in \mathbb{N} : \ker(\lambda_k - T)^{n-1} \neq \ker(\lambda_k - T)^n\} = 1$$

3. *We can decompose*

   $$H = \ker T \oplus \overline{\mathit{span}\{e_k : k \in N\}}$$

   and the two spaces are orthogonal to each other.

4. *For all $h \in H$, we have*

   $$Th = \sum_{k\in N} \lambda_k \langle e_k, h\rangle e_k$$

   i.e. $T = \sum_{k\in N} \lambda_k P_k$, where P_k is the orthogonal projection onto e_k.