Functional Analysis
Winter 2018/2019

Peer Christian Kunstmann
Karlsruher Institut für Technologie (KIT)
Institut für Analysis
Englerstr. 2, 76131 Karlsruhe
e-mail: peer.kunstmann@kit.edu

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

1 Metric und normed spaces

1.1. Definition: Let $X \neq \emptyset$ be a set. A *metric on* X is a map $d : X \times X \to [0, \infty)$ satisfying

(M1) $\forall x, y \in X$: $d(x, y) = 0 \Leftrightarrow x = y$ (*definiteness*);

(M2) $\forall x, y \in X$: $d(x, y) = d(y, x)$ (*symmetry*);

(M3) $\forall x, y, z \in X$: $d(x, y) \leq d(x, z) + d(z, y)$ (*triangle inequality*).

If d is a metric on X then the pair (X, d) is called a *metric space*.

If d only satisfies (N2) and (N3) and $d(x, x) = 0$ for all $x \in X$ then d is called a *semimetric* on X.

Remarks: If d is a semimetric on X then

$$\begin{aligned} \forall x, y, \in X \quad &: \quad |d(x, z) - d(x, y)|| \leq d(z, y) \quad (\textit{reverse triangle inequality}); \\ \forall x, y, u, v \in X \quad &: \quad |d(x, y) - d(u, v)| \leq d(x, u) + d(y, v). \end{aligned}$$

The first inequality follows from $d(x, z) \leq d(x, y) + d(y, z)$ and $d(x, y) \leq d(x, z) + d(z, y)$. The proof of the second inequality is similar.

Examples: (1) $d(x, y) := |x - y|$ defines a metric on $\mathbb{R}$.

(2) $d((x_1, x_2), (y_1, y_2)) := |x_1 - y_1|$ define a semimetric on $\mathbb{R}^2$.

1.2. Norms: We always consider vecctor spaces over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, i.e. real vector spaces for $\mathbb{K} = \mathbb{R}$ or complex vector spaces for $\mathbb{K} = \mathbb{C}$.

Definition: Let X be a $\mathbb{K}$-vector space. A *norm on* X is a map $\| \cdot \| : X \to \mathbb{R}$ satisfying

(N1) $\forall x \in X$: $\|x\| = 0 \Longrightarrow x = 0$ (*definiteness*);

(N2) $\forall x \in X, \alpha \in \mathbb{K}$: $\|\alpha x\| = |\alpha| \|x\|$ (*homogeneity*);

(N3) $\forall x, y \in X$: $\|x + y\| \leq \|x\| + \|y\|$ (*triangle inequality*).

Then $(X, \| \cdot \|)$ is called a *normed space*.

If $\| \cdot \|$ only satisfies (N2) and (N3) then $\| \cdot \|$ is called a *seminorm* on X.

Remarks: (1) If $\| \cdot \|$ is a seminorm on X then $\|x\| \geq 0$ for all $x \in X$. [By (N2) we obtain $\|0\| = 0$; further $0 = \|x - x\| \leq \|x\| + \| - x\| = 2\|x\|$, hence $\|x\| \geq 0$.]

Again, the triangle inequality implies $|\|x\| - \|y\|| \leq \|x - y\|$ for all $x, y \in X$.

(2) If $\| \cdot \|$ is a norm [resp. seminorm] on X then $d(x, y) := \|x - y\|$ defines a metric [resp. semimetric] on X.

Examples: (1) Let $n \in \mathbb{N}$. Then $\|(x_j)_{j=1}^n\|_1 := \sum_{j=1}^n |x_j|$ defines a norm on $\mathbb{K}^n$.

(2) Let $M \neq \emptyset$ be a set and $B(M) := \{f : M \to \mathbb{K} : \ f \text{ is bounded}\}$. Then $B(M)$ is a $\mathbb{K}$-vector space (as a linear subspace of the $\mathbb{K}$-vector space $\mathbb{K}^M$ of all functions $M \to \mathbb{K}$) and

$$\|f\|_\infty := \sup\{|f(\omega)| : \omega \in M\}$$

defines a norm on $B(M)$: (N1) and (N2) are clear. For the proof of (N3) take $f, g \in B(M)$. For any $\omega \in M$ we then have

$$|(f+g)(\omega)| \leq |f(\omega)| + |g(\omega)| \leq \|f\|_\infty + \|g\|_\infty,$$

and taking the sup over $\omega \in M$ yields (N3).

(3) Let $p \in (1, \infty)$ and $n \in \mathbb{N}$. Then

$$\|(x_j)_{j=1}^n\|_p := \Big(\sum_{j=1}^n |x_j|^p\Big)^{1/p}$$

defines a norm $\mathbb{K}^n$. Again, (N1) and (N2) are clear, but (N3) needs some work.

Young's inequality: For all $a, b \geq 0$ we have $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ where $q \in (1, \infty)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. This is clear for $b = 0$. So fix $b > 0$ and consider $f : [0, \infty) \to \mathbb{R}$, $f(a) := ab - a^p/p$. Then $f(0) = 0$, $f(a) \to -\infty$ for $a \to \infty$, f is differentiable on $[0, \infty)$ with $f'(a) = b - a^{p-1}$. Since f' is strictly decreasing, f attains its maximum at the only zero $a_0 = b^{1/(p-1)}$ of f', and for any $a \geq 0$:

$$f(a) \leq f(a_0) = b^{1/(p-1)}b - \frac{b^{p/(p-1)}}{p} = \frac{b^q}{q},$$

since $\frac{p}{p-1} = q$. □

Hölder's inequality: For all $n \in \mathbb{N}$ and $a_1, \ldots, a_n, b_1 \ldots, b_n \geq 0$ we have

$$\sum_{j=1}^n a_j b_j \leq \Big(\sum_{j=1}^n a_j^p\Big)^{1/p} \Big(\sum_{j=1}^n b_j^q\Big)^{1/q}.$$

Proof. This is clear if $\sum_j a_j^p = 0$ or $\sum_j b_j^q = 0$. By homogeneity it thus suffices to consider the case $\sum_j a_j^p = 1$ and $\sum_j b_j^q = 1$ (otherwise let $\tilde{a}_j := a_j/(\sum_k a_k^p)^{1/p}$ and $\tilde{b}_j := b_j/(\sum_k b_k^p)^{1/p}$). We then have by Young

$$\sum_j a_j b_j \leq \frac{1}{p} \sum_j a_j^p + \frac{1}{q} \sum_j b_j^q = \frac{1}{p} + \frac{1}{q} = 1.$$

□

Minkowski's inequality: Let $n \in \mathbb{N}$ and $x = (x_j)_{j=1}^n, y = (y_j)_{j=1}^n \in \mathbb{K}^n$. Then

$$\Big(\sum_{j=1}^n |x_j + y_j|^p\Big)^{1/p} \le \Big(\sum_{j=1}^n |x_j|^p\Big)^{1/p} + \Big(\sum_{j=1}^n |y_j|^p\Big)^{1/p}.$$

This finally is (N3) for $\|\cdot\|_p$ on $\mathbb{K}^n$.

Proof. By $|x_j + y_j| \le |x_j| + |y_j|$ it suffices to take $x_j, y_j \ge 0$. The assertion is clear for $\sum_j (x_j + y_j)^p = 0$, so we may assume $\sum_j (x_j + y_j)^p = 1$. Then, by Hölder,

$$\begin{aligned} 1 &= \sum_j (x_j + y_j)^p = \sum_j x_j \underbrace{(x_j + y_j)^{p-1}}_{=:z_j} + \sum_j y_j \underbrace{(x_j + y_j)^{p-1}}_{=z_j} \\ &\le \Big(\sum_{j=1}^n x_j^p\Big)^{1/p}\Big(\sum_{j=1}^n z_j^q\Big)^{1/q} + \Big(\sum_{j=1}^n y_j^p\Big)^{1/p}\Big(\sum_{j=1}^n z_j^q\Big)^{1/q}, \end{aligned}$$

and $z_j^q = (x_j + y_j)^{qp-q} = (x_j + y_j)^p$ (since $pq - q = p$). Hence $\sum_j z_j^q = 1$ and the inequality is proved. □

1.3. Open and closed sets: Let (X, d) be a metric space. For $r > 0$, $x_0 \in X$ let

$$\begin{aligned} B(x_0, r) &:= \{x \in X : d(x, x_0) < r\} \quad \text{open ball with radius } r \text{ around } x_0 \\ \overline{B}(x_0, r) &:= \{x \in X : d(x, x_0) < r\} \quad \text{closed ball with radius } r \text{ around } x_0. \end{aligned}$$

Further we define

$$\begin{aligned} Q \subseteq X \text{ is called } \textit{open} &:\Longleftrightarrow \forall x_0 \in Q\, \exists r > 0 : B(x_0, r) \subseteq Q. \\ A \subseteq X \text{ is called } \textit{closed} &:\Longleftrightarrow X \setminus A \text{ is open.} \end{aligned}$$

Examples: X, $\emptyset$, $B(x_0, r)$ are open; X, $\emptyset$, $\overline{B}(x_0, r)$ are closed: The assertions on X and $\emptyset$ are clear. The triangle inequality shows for $x \in B(x_0, r)$ that $B(x, r - d(x, x_0)) \subseteq B(x_0, r)$ and for $x \notin \overline{B}(x_0, r)$ that $B(x, d(x, x_0) - r) \subseteq X \setminus \overline{B}(x_0, r)$.

Remarks: (1) Arbitrary unions of open sets are open, finite intersections of open sets are open.

(2) Arbitrary intersections of closed sets are closed, finite unions of closed sets are closed.

(3) If $\emptyset \ne M \subseteq X$ then $(M, d|_{M\times M})$ is a metric space. A subset $Q \subseteq M$ is called *relatively open* [resp. *relatively closed*] if M is open in the metric space $(M, d|_{M\times M})$.

Example: $[0, 1)$ is relatively open in $[0, 1]$, but not open in $\mathbb{R}$.

Proposition: Let $Q \subseteq M$. Then Q is relatively open in M if and only if there exists $\widetilde{Q} \subseteq X$ open in X such that $Q = \widetilde{Q} \cap M$.

Proof. "⇐" is clear. For "⇒" put

$$\widetilde{Q} := \bigcup\{B_X(x_0, r) : x_0 \in Q, r > 0, B_M(x_0, r) \subseteq Q\},$$

where $B_X(x_0, r) = \{x \in X : d(x, x_0) < r\}$ and $B_M(x_0, r) = \{x \in M : d(x, x_0) < r\}$. Observe that $B_M(x_0, r) = B_X(x_0, r) \cap M$ and thus

$$\begin{aligned} \widetilde{Q} \cap M &= \bigcup\{B_X(x_0, r) \cap M : x_0 \in Q, r > 0, B_M(x_0, r) \subseteq Q\} \\ &= \bigcup\{B_M(x_0, r) : x_0 \in Q, r > 0, B_M(x_0, r) \subseteq Q\} = Q, \end{aligned}$$

where for the last inclusion we use that Q is relatively open in M. □

end Mon 15.10.18

Definition: A subset $U \subseteq X$ is called a *neighborhood* of $x_0 \in X$ if there exists $r > 0$ such that $B(x_0, r) \subseteq U$.

Let $M \subseteq X$. We define the *interior of M*

$$\operatorname{int}(M) := \{x \in X : M \text{ is a neighborhood of } x\},$$

the *closure of M*

$$\overline{M} := \{x \in X : U \cap M \neq \emptyset \text{ for any neighborhood } U \text{ of } x\},$$

and the *boundary of M*

$$\partial M := \{x \in X : U \cap M \neq \emptyset \text{ and } U \cap (X \setminus M) \neq \emptyset \text{ for any neighborhood } U \text{ of } x\}.$$

An $x \in X$ is called *interior point of M*, if $x \in \operatorname{int}(M)$, and *boundary point of M* if $x \in \partial M$.

Remarks: (1) $\operatorname{int}(M)$ is the largest open subset of M, $\overline{M}$ is the smallest closed superset of M, and $\operatorname{int}(X \setminus M) = X \setminus \overline{M}$, $\overline{X \setminus M} = X \setminus \operatorname{int}(M)$.

(2) M is open ⇔ $\operatorname{int}(M) = M$; M is closed ⇔ $\overline{M} = M$.

(3) $\partial M = \overline{M} \cap \overline{X \setminus M}$, in particular ∂M is closed.

(4) $\overline{M}$ is the disjoint union of $\operatorname{int}(M)$ and ∂M, and X is the disjoint union of $\operatorname{int}(M)$, ∂M and $\operatorname{int}(X \setminus M)$.

1.4. Convergence: Let (X, d) be a metric space and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X and $x \in X$. Then (x_n) is said to be *convergent to x* ($x_n \to x$) with respect to d if $d(x_n, x) \to 0$, and (x_n) is said to be a *Cauchy sequence* with respect to d if

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall n, m \geq n_0 : d(x_n, x_m) < \varepsilon.$$

By (M1) the limit of a convergent sequence is uniquely determined.

Proposition: Let $M \subseteq X$, $x_0 \in M$. Then $x_0 \in \overline{M}$ if and only if there exists a sequence (x_n) in M such that $x_n \to x_0$.

Proof. "$\Rightarrow$": For any $n \in \mathbb{N}$ we find $x_n \in B(x_0, 1/n) \cap M$. Then (x_n) is a sequence in M with $d(x_n, x_0) \leq 1/n \to 0$.

"$\Leftarrow$": Let $r > 0$. We find $n \in \mathbb{N}$ with $d(x_n, x_0) < r$. Then $x_n \in B(x_0, r) \cap M$. We have shown $x_0 \in \overline{M}$. □

Examples: (1) If (X, d) is a metric space and $x_n \to x$, $y_n \to y$ then $d(x_n, y_n) \to d(x, y)$. This follows from $|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y)$.

(2) If $(X, \|\cdot\|)$ is a normed space and $x_n \to x$, $y_n \to y$ in X and $\alpha_n \to \alpha \in \mathbb{K}$ then $x_n + y_n \to x + y$, $\alpha_n x_n \to \alpha x$ and $\|x_n\| \to \|x\|$. This follows from

$$\begin{aligned} \|x_n + y_n - (x + y)\| &\leq \|x_n - x\| + \|y_n - y\|, \\ \|\alpha_n x_n - \alpha x\| &\leq |\alpha_n| \|x_n - x\| + |\alpha_n - \alpha| \|x\|, \\ |\|x_n\| - \|x\|| &\leq \|x_n - x\|. \end{aligned}$$

In particular, a convergent sequence (x_n) in X is bounded, i.e. satisfies $\sup_n \|x_n\| < \infty$.

(X, d) is said to be *complete* if any Cauchy sequence in X is convergent to some $x \in X$.

Examples: For the metric induced by $|\cdot|$, $\mathbb{R}$ and $\mathbb{C}$ are complete, but $\mathbb{Q}$ is not complete.

A *Banach space* is a normed space which is complete for the metric induced by the norm.

Examples: (1) Let M be a set. Then $(B(M), \|\cdot\|_\infty)$ is a Banach space: Let (f_n) be a $\|\cdot\|_\infty$-Cauchy sequence in $B(M)$.

Step 1 Find a candidate: Fix $\omega \in M$. We have $|f_n(\omega) - f_m(\omega)| \leq \|f_n - f_m\|_\infty$, so $(f_n(\omega))$ is Cauchy in $\mathbb{K}$, hence convergent. Thus we can define

$$f : M \to \mathbb{K}, \quad \omega \mapsto f(\omega) := \lim_{n\to\infty} f_n(\omega).$$

Step 2 Show that f is bounded, i.e. $f \in B(M)$, and $\|f_n - f\|_\infty \to 0$: Let $\varepsilon > 0$. By assumption we find n_0 such that $\|f_n - f_m\|_\infty \leq \varepsilon$ for all $n, m \geq n_0$. Fix $n \geq n_0$. For any $\omega \in M$ we have

$$|f_n(\omega) - f(\omega)| = \lim_{m\to\infty} \underbrace{|f_n(\omega) - f_m(\omega)|}_{\leq \|f_n - f_m\|_\infty} \leq \varepsilon.$$

Taking the sup over $\omega \in M$ shows $\|f_n - f\|_\infty \leq \varepsilon$. This implies $f = f_{n_0} + (f - f_{n_0}) \in B(M)$ and $\|f_n - f\|_\infty \to 0$.

Special case $M = \mathbb{N}$: $l^\infty := B(\mathbb{N})$.

(2) Let $p \in [1, \infty)$. Then

$$l^p := \left\{ x = (x_j)_{j\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} : \sum_{j=1}^{\infty} |x_j|^p < \infty \right\}$$

is a vector space and $\|x\|_p := (\sum_{j=1}^\infty |x_j|^p)^{1/p}$ defines a norm on l^p. Moreover, $(l^p, \|\cdot\|_p)$ is a Banach space.

It is clear that $0 \in l^p$ and that, for $x \in l^p$ and $\alpha \in \mathbb{K}$, $\alpha x \in l^p$ and $\|\alpha x\|_p = |\alpha|\|x\|_p$. So let $x, y \in l^p$. We have to show $x + y \in l^p$ and the triangle inequality. For any $N \in \mathbb{N}$ we have by Minkowski

$$\Big(\sum_{j=1}^N |x_j + y_j|^p\Big)^{1/p} \le \Big(\sum_{j=1}^N |x_j|^p\Big)^{1/p} + \Big(\sum_{j=1}^N |y_j|^p\Big)^{1/p} \le \|x\|_p + \|y\|_p.$$

Letting $N \to \infty$ we obtain $\Big(\sum_{j=1}^\infty |x_j + y_j|^p\Big)^{1/p} \le \|x\|_p + \|y\|_p < \infty$, hence $x + y \in l^p$ and $\|x + y\|_p \le \|x\|_p + \|y\|_p$.

Completeness: Let $(x^{(n)})$ be a $\|\cdot\|_p$-Cauchy sequence in l^p. For any $j \in \mathbb{N}$, $(x_j^{(n)})_n$ is Cauchy in $\mathbb{K}$ and $x_j^{(0)} := \lim_{n\to\infty} x_j^{(n)}$ exists. Setting $x^{(0)} := (x_j^{(0)})_j$ we have a candidate for the limit.

We show $x^{(0)} \in l^p$ and $\|x^{(n)} - x^{(0)}\|_p \to 0$: Let $\varepsilon > 0$. By assumption we find n_0 such that $\|x^{(n)} - x^{(m)}\|_p \le \varepsilon$ for all $n, m \ge n_0$. For $n \ge n_0$ and any $N \in \mathbb{N}$ we then have

$$\Big(\sum_{j=1}^N |x_j^{(n)} - x_j^{(0)}|^p\Big)^{1/p} = \lim_{m\to\infty} \underbrace{\Big(\sum_{j=1}^N |x_j^{(n)} - x_j^{(m)}|^p\Big)^{1/p}}_{\le \|x^{(n)} - x^{(m)}\|_p} \le \varepsilon.$$

This implies for $N \to \infty$: $\|x^{(n)} - x^{(0)}\|_p < \infty$ for all $n \ge n_0$. Hence $x^{(0)} = x^{(n_0)} - (x^{(n_0)} - x^{(0)}) \in l^p$ and $\|x^{(n)} - x^{(0)}\|_p \to 0$.

1.5. Lemma: Let (X, d) be a complete metric space and $\emptyset \ne M \subseteq X$. Then $(M, d|_{M\times M})$ is complete if and only if M is closed in X.

Proof. If M is complete and (x_n) is a sequence in M converging to $x \in X$ then (x_n) is Cauchy and therefore has a limit $\widetilde{x} \in M$. But then $x = \widetilde{x} \in M$, and we have shown that M is closed. If M is closed in X and (x_n) is a Cauchy sequence in M then we find, by completeness of X, an $x \in X$ such that $x_n \to x$. By closedness of M we obtain $x \in M$ and (x_n) has a limit in M. □

Example: $c_0 := \{x = (x_j) \in \mathbb{K}^{\mathbb{N}} : \lim_{j\to\infty} x_j = 0\}$ is a Banach space with respect to $\|\cdot\|_\infty$, i.e. c_0 is a closed subspace of l^∞: Let $(x^{(n)})$ be a sequence in c_0 converging to $x^{(0)} \in l^\infty$. We have to show $x^{(0)} \in c_0$. So let $\varepsilon > 0$. By assumption we find $n \in \mathbb{N}$ with $\|x^{(n)} - x^{(0)}\|_\infty \le \varepsilon/2$ and by $x^{(n)} \in c_0$ we find $j_0 \in \mathbb{N}$ such that $|x_j^{(n)}| \le \varepsilon/2$ for all $j \ge j_0$. For any $j \ge j_0$ we then have

$$|x_j^{(0)}| \le |x_j^{(0)} - x_j^{(n)}| + |x_j^{(n)}| \le \|x^{(n)} - x^{(0)}\|_\infty + \varepsilon/2 \le \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which proves $x_j^{(0)} \to 0$ as $j \to \infty$, i.e. $x^{(0)} \in c_0$.

Also $c := \{x = (x_j) \in \mathbb{K}^{\mathbb{N}} : (x_j) \text{ converges in } \mathbb{K}\}$ is a closed subspace of l^∞. end Thu 18.10.18

1.6. Denseness and separability: Let (X, d) be a metric space and $D, B \subseteq X$. Then D is called *dense in* B if $B \subseteq \overline{D}$, i.e. if an $x \in B$ can be approximated by a sequence in D. If D is dense in X it is simply called *dense*.

The space (X, d) is called *separable* if it contains a countable dense subset.

Examples: $\mathbb{R}^n$ is separable since $\mathbb{Q}^n$ is a countable dense subset. $\mathbb{C}^n$ is separable since $\mathbb{Q}^n + i\mathbb{Q}^n$ is a countable dense subset.

Proposition: A normed space $(X, \|\cdot\|)$ is separable if and only if it contains a countable subset A such that $\text{lin}(A)$ is dense in X.

Here $\text{lin}(A)$ denotes the linear span of A, i.e. the smallest linear subspace containing A or, in other words, the set of all linear combinations of elements of A.

Proof. If D is dense then also $\text{lin}(D)$ is dense. So assume that A is countable and $\text{lin}(A)$ is dense. Then the set D of all linear combinations of elements of A with rational coefficients is countable and dense in $\text{lin}(A)$, and thus also dense in X. $\square$

Examples: The spaces c_0 and l^p, $p \in [1, \infty)$, are separable: The space of finite sequences

$$\varphi := \{x = (x_j) \in \mathbb{K}^{\mathbb{N}} : \exists j_0 \in \mathbb{N}\ \forall j > j_0 : x_j = 0\}$$

is dense and we have $\varphi = \text{lin}\{e_n : n \in \mathbb{N}\}$ where $e_n := (\delta_{jn})_{j \in \mathbb{N}}$ for any $n \in \mathbb{N}$ denotes the n-th unit vector.

The space l^∞ is not separable: The set $M := \{0, 1\}^{\mathbb{N}} \subseteq l^\infty$ is uncountable and for any two vectors $x, y \in M$ with $x \neq y$ we have $\|x - y\|_\infty = 1$. If D is dense in l^∞ then D is dense in M and for every $x \in M$ we have $D_x := B(x, 1/4) \cap D \neq \emptyset$, but $D_x \cap D_y = \emptyset$ for $x \neq y$, i.e. D contains uncountably many disjoint non-empty subsets and thus cannot be countable.

Lemma: Any subset of a separable space is separable.

Proof. left as an exercise. $\square$

1.7. Continuity: Let X, Y be metric spaces. A map $f : X \to Y$ is called *continuous* if, for any open set $Q \subseteq Y$, the pre-image $f^{-1}(Q)$ is open in X.

f is called *continuous at* $x_0 \in X$ if for any neighborhood V of $f(x_0)$ there exists a neighborhood U of x_0 such that $f(U) \subseteq V$.

We clearly have:

$$\begin{aligned} f \text{ continuous at } x_0 &\iff \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X : d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \varepsilon \\ &\iff x_n \to x_0 \text{ in } X \ \Rightarrow\ f(x_n) \to f(x_0) \text{ in } X \end{aligned}$$

The first condition may be rephrased as $\forall \varepsilon > 0\ \exists \delta > 0 : B(x_0, \delta) \subseteq f^{-1}(B(f(x_0), \varepsilon))$.

Proposition: A map $f : X \to Y$ is continuous if and only if it is continuous at every point.

Proof. "$\Rightarrow$": Let f be continuous, $x_0 \in X$, and $\varepsilon > 0$. Then $B(f(x_0), \varepsilon))$ is open and thus also $M := f^{-1}(B(f(x_0), \varepsilon))$ is open by assumption. Since $x_0 \in M$ we find $\delta > 0$ such that $B(x_0, \delta) \subseteq M$. We have shown that f is continuous at x_0.

"$\Leftarrow$": Let f be continuous at every point and $Q \subseteq Y$ be open. Let $x_0 \in f^{-1}(Q)$. Then $f(x_0) \in Q$ and we find $\varepsilon > 0$ with $B(f(x_0), \varepsilon) \subseteq Q$ and then $\delta > 0$ with $B(x_0, \delta) \subseteq f^{-1}(B(f(x_0), \varepsilon) \subseteq f^{-1}(Q)$. Hence $f^{-1}(Q)$ is open. □

Examples: (1) If (X_j, d_j), $j = 1, 2$, are metric spaces then $X_1 \times X_2$ is metric space for the metric d given by $d((x_1, x_2), (y_1, y_2)) := d_1(x_1, y_1) + d_2(x_2, y_2)$. For any metric space (X, d) the metric $d : X \times X \to [0, \infty)$ is continuous.

(2) If $(X_j, \|\cdot\|_j)$, $j = 1, 2$, are normed spaces then $X_1 \times X_2$ is a normed space for the norm $\|\cdot\|$ given by $\|(x_1, x_2)\| = \|x_1\| + \|x_2\|$. For any normed space $(X, \|\cdot\|)$ the maps $+ : X \times X \to X$ and $\cdot : \mathbb{K} \times X \to X$ are continuous. Moreover, the norm $\|\cdot\| : X \to \mathbb{R}$ is continuous.

1.8. Uniform continuity: Let X, Y be metric spaces. A map $f : X \to Y$ is called *uniformly continuous* if

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, y \in X : d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon.$$

Remarks: (1) A uniformly continuous map is clearly continuous.

(2) A uniformly continuous map maps Cauchy sequences to Cauchy sequences: If (x_n) is Cauchy in X and $\varepsilon > 0$ we find $\delta > 0$ according to uniform continuity of f and then n_0 such that $d(x_n, x_m) < \delta$ for all $n, m \geq n_0$. By choice of δ this implies $d(f(x_n), f(x_m)) < \varepsilon$ for all $n, m \geq n_0$. Hence $(f(x_n))$ is Cauchy in Y.

(3) A map $f : X \to Y$ is called *Lipschitz continuous* if there exists an $L > 0$ such that $d(f(x), f(y)) \leq L d(x, y)$ for all $x, y \in X$. Any Lipschitz continuous map is unifomrly continuous (one can take $\delta = \varepsilon/L$).

Examples: (1) If (X, d) is a metric space then $d : X \times X \to [0, \infty)$ is Lipschitz continuous (with $L = 1$).

(2) If $(X, \|\cdot\|)$ is a normed space then $\|\cdot\| : X \to \mathbb{R}$ is Lipschitz continuous (with $L = 1$) and $+ : X \times X \to X$ is Lipschitz continuous (with $L = 1$). If $X \neq \{0\}$ then $\cdot : \mathbb{K} \times X \to X$ is not uniformly continuous: Choose $x_0 \in X$ with $\|x_0\| = 1$. For $\delta > 0$ and $k \in \mathbb{N}$ the pairs $(k + \delta, (k + \delta)x_0)$ and (k, kx_0) have distance 2δ but

$$\|(k + \delta) \cdot (k + \delta)x_0 - k \cdot kx_0\| = (k + \delta)^2 - k^2 = 2k\delta + \delta^2 \to \infty \quad (k \to \infty).$$

The following principle has countless applications in praxis.

1.9. Proposition: Let X, Y be metric spaces, $D \subseteq X$, Y complete, and $f : D \to Y$ uniformly continuous. Then there exists a unique continuous extension $\widetilde{f} : \overline{D} \to Y$ of f. This extension $\widetilde{f}$ is uniformly continuous.

Proof. The extension necessarily has to satisfy $\widetilde{f}(x) = \lim_{n\to\infty} f(x_n)$ if (x_n) is a sequence in D with $x_n \to x$. We take this as the definition of $\widetilde{f}$ and have to show that $\widetilde{f}(x)$ is well-defined, i.e. the limit exists and does not depend on the chosen sequence. If (x_n) is a sequence in D with $x_n \to x$ then (x_n) is Cauchy in D, so $(f(x_n))$ is Cauchy in Y and the limit exists by completeness of Y. If (y_n) is another sequence in D with $y_n \to x$ then also $\lim_n f(y_n)$ exists. Letting $z_n := x_{(n+1)/2}$ for odd n and $z_n := y_{n/2}$ for even n we have defined a sequence (z_n) in D with $z_n \to x$ and conclude $\lim_n f(x_n) = \lim_n f(z_n) = \lim_n f(y_n)$.

Hence $\widetilde{f} : \overline{D} \to Y$ is well-defined and it rests to show uniform continuity. To this end let $\varepsilon > 0$. We find $\delta > 0$ according to uniform continuity of f. Now let $x, y \in \overline{D}$ such that $d(x, y) < \delta/3$. we find sequences (x_n) and (y_n) in D with $x_n \to x$ and $y_n \to y$ and we find $n_0 \in \mathbb{N}$ such that $d(x_n, x) < \delta/3$ and $d(y_n, y) < \delta/3$ for all $n \geq n_0$. For $n \geq n_0$ we then obtain

$$d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n) < \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta,$$

hence $d(f(x_n), f(y_n)) < \varepsilon$ by choice of δ, and we conclude

$$d(\widetilde{f}(x), \widetilde{f}(y)) = \lim_{n\to\infty} d(f(x_n), f(y_n)) \leq \varepsilon,$$

which finishes the proof of uniform continuity. □

end Mon 22.10.18

1.10. Linear operators: Let X, Y be normed spaces (over $\mathbb{K}$) and $T : X \to Y$ be $\mathbb{K}$-linear. The following are equivalent:

(i) T is continuous,

(ii) T is continuous at 0,

(iii) there exists $C > 0$ such that, for all $x \in X$, $\|Tx\| \leq C\|x\|$,

(iv) T is Lipschitz-continuous,

(v) T is uniformly continuous.

Proof. (iv) ⇒ (v) ⇒ (i) ⇒ (ii) is clear. (ii) ⇒ (iii): If (iii) does not hold we find a sequence (x_n) in X satisfying $\|x_n\| = 1/n$ and $\|Tx_n\| > 1$. Then $x_n \to 0$ but $Tx_n \not\to 0$, so (ii) does not hold.

(iii) ⇒ (iv): By (iii) and linearity of T we obtain, for all $x, y \in X$,

$$\|Tx - Ty\| = \|T(x - y)\| \leq C\|x - y\|.$$

Hence T is Lipschitz continuous. □

Linear operators satisfying (iii) are called *bounded*. We let

$$\mathcal{L}(X,Y) := \{T : X \to Y : \ X \text{ is linear and bounded}\}.$$

Then $\mathcal{L}(X,Y)$ is a vector space and

$$\|T\| := \sup\Big\{\frac{\|Tx\|}{\|x\|} : x \in X \setminus \{0\}\Big\}$$

defines a norm on $\mathcal{L}(X,Y)$, the *operator norm.* (For $X = \{0\}$ we use the convention $\sup\emptyset := 0$ here).

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

$$\begin{aligned} \|T\| &= \inf\{C \geq 0 : \ \text{(iii) holds}\} \\ &= \sup\{\|Tx\| : x \in X, \|x\| = 1\} \\ &= \sup\{\|Tx\| : x \in X, \|x\| \leq 1\}. \end{aligned}$$

Proof. (N1) is clear. (N2) follows from (N2) for the norm of Y. For the proof of (N3) let $S, T \in \mathcal{L}(X,Y)$ and $x \in X$ with $\|x\| \leq 1$. Then

$$\|(S+T)x\| \leq \|Sx\| + \|Tx\| \leq \|S\| + \|T\|,$$

hence $\|S+T\| \leq \|S\| + \|T\|$. □

Remark: If Z is another normed space and $T \in \mathcal{L}(X,Y)$, $S \in \mathcal{L}(Y,Z)$ then clearly $ST \in \mathcal{L}(X,Z)$ and $\|ST\| \leq \|S\|\|T\|$.

Definition: In case $Y = \mathbb{K}$ the space $X' := \mathcal{L}(X,\mathbb{K})$ is called *the dual space* of X. The elements of X' are called (linear) *functionals on* X.

Remark: If $T \in \mathcal{L}(X,Y)$ then the kernel $N(T) = T^{-1}(\{0\})$ of T is a closed linear subspace of X.

1.11. Proposition: Let X be a normed space and Y a Banach space. Let $D \subseteq X$ be a dense linear subspace and $T \in \mathcal{L}(D,Y)$. Then there exists a unique continuous extension $\widetilde{T} : X \to Y$ of T, and for this extension we have $\widetilde{T} \in \mathcal{L}(X,Y)$ and $\|\widetilde{T}\| = \|T\|$.

Proof. Use 1.9, linearity and boundedness of T. □

1.12. Proposition: Fix $p \in (1,\infty)$. Let $q \in [1,\infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. We show that $J : l^q \to (l^p)'$ defined by $(Jy)(x) = \sum_{j=1}^{\infty} x_j y_j$ for $x = (x_j) \in l^p$, $y = (y_j) \in l^q$, defines an *isometry*, i.e. $J : l^q \to (l^p)'$ is linear and bijective and satisfies $\|Jy\|_{(l^p)'} = \|y\|_{l^q}$ for all $y \in l^q$.

Proof. If $y \in l^q$ then we have by Hölder

$$|(Jy)(x)| = |\sum_{j=1}^{\infty} x_j y_j| \leq \|x\|_{l^p} \|y\|_{l^q} \quad \text{for all } x \in l^p.$$

Hence $Jy \in (l^p)'$ and $\|Jy\|_{(l^p)'} \leq \|y\|_{l^q}$. Conversely, for a given $\phi \in (l^p)'$ we let $y_j := \phi(e_j)$. For any $x = (x_j) \in l^p$ we have $x = \sum_j x_j e_j$ where the series converges in l^p. Hence $\phi(x) = \sum_j x_j \phi(e_j) = \sum_j x_j y_j$ where the series converges in $\mathbb{K}$. In particular, we see that ϕ is uniquely determined by the sequence $y := (y_j)$. We also have $|\phi(x)| \leq \|x\|_{l^p} \|y\|_{l^q}$ which implies $\|\phi\| \leq \|y\|_{l^q}$. We only have to show $y \in l^q$ and $\|y\|_{l^q} \leq \|\phi\|$ to finish the proof. To this end we fix $N \in \mathbb{N}$ and define $x^{(N)} = \sum_{j=1}^N \overline{y_j} |y_j|^{q-2} e_j$. Then we have

$$\sum_{j=1}^{N} |y_j|^q = |J(y)(x^{(N)})| = |\phi(x^{(N)})| \leq \|\phi\| \|x^{(N)}\|_{l^p}$$

and

$$\|x^{(N)}\|_{l^p} = \Big(\sum_{j=1}^{N} |y_j|^{pq-p} \Big)^{1/p} = \Big(\sum_{j=1}^{N} |y_j|^q \Big)^{1-1/q}.$$

We thus obtain

$$\Big(\sum_{j=1}^{N} |y_j|^q \Big)^{1/q} \leq \|\phi\|,$$

which implies for $N \to \infty$ that $y \in l^q$ and $\|y\|_{l^q} \leq \|\phi\|$. □

Remark: So for $p \in (1, \infty)$, the dual space of l^p is l^q where the action of $y = (y_j)$ on $x = (x_j)$ is given by $J(y)(x) = \sum_{j=1}^{\infty} x_j y_j$. This is written as *duality bracket*

$$\langle x, y \rangle := \sum_{j=1}^{\infty} x_j y_j.$$

Further examples: In the same way, using also $\langle x, y \rangle = \sum_{j=1}^{\infty} x_j y_j$, one can show $(c_0)' = l^1$ and $(l^1)' = l^\infty$. However, as we shall see later, not every functional on l^∞ is given by an l^1-sequence.

1.13. Proposition: If X is a normed space and Y is a Banach space then $\mathcal{L}(X, Y)$ is a Banach space. In particular, the *dual space* X' of a normed space X is always a Banach space.

Proof. Let (T_n) be a Cauchy sequence in $\mathcal{L}(X, Y)$. By $\|T_n x - T_m\| \leq \|T_n - T_m\| \|x\|$, $(T_n x)$ is Cauchy in Y for any $x \in X$, hence convergent. Thus we can define $T : X \to Y$ by $x \mapsto Tx := \lim_{n\to\infty} T_n x$. It is easy to see that T is linear (use 1.4). Boundedness follows from

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq \lim_{n\to\infty} \|T_n\| \|x\|.$$

Let $\varepsilon > 0$. We find n_0 such that $\|T_n - T_m\| < \varepsilon$ for all $n, m \geq n_0$. For $n \geq n_0$ and $x \in X$ with $\|x\| \leq 1$ we then have

$$\|T_n x - Tx\| = \lim_{m\to\infty} \underbrace{\|T_n x - T_m x\|}_{\leq \|T_n - T_m\|} \leq \varepsilon,$$

hence $\|T_n - T\| \leq \varepsilon$. We have shown $\|T_n - T\| \to 0$. □

1.14. Proposition (completion): Let X be normed space. There exists a Banach space $\widetilde{X}$ and a linear map $J : X \to \widetilde{X}$ such that $J(X)$ is dense in $\widetilde{X}$ and $\|Jx\| = \|x\|$ for all $x \in X$.

Any such pair $(\widetilde{X}, J)$ is called a *completion of* X.

Proof. We set $\widetilde{X} := \{x = (x_j) \in \mathbb{K}^{\mathbb{N}} : (x_j) \text{ is Cauchy in } X\}/\{x = (x_j) \in \mathbb{K}^{\mathbb{N}} : x_j \to 0\}$ with norm $\|[(x_j)]\| := \lim_{j\to\infty} \|x_j\|$ and define $J : X \to \widetilde{X}$ by $Jx := [(x)_{j\in\mathbb{N}}]$. Then J is linear and $\|Jx\| = \|x\|$ for all $x \in X$.

We show that $\widetilde{X}$ is complete. To this end let $(z_n) = ([x^{(n)}])$ be Cauchy in $\widetilde{X}$. For any fixed n we find $j_0^{(n)}$ such that $\|x_j^{(n)} - x_k^{(n)}\| < 2^{-n}$ for all $j, k \geq j_0^{(n)}$. We let $\widetilde{x}_n := x^{(n)}_{j_0^{(n)}}$ and $\widetilde{z}_n := J\widetilde{x}_n$. Then we have

$$\|z_n - \widetilde{z}_n\| = \lim_{j\to\infty} \|x_j^{(n)} - \widetilde{x}_n\| = \lim_{j\to\infty} \|x_j^{(n)} - x^{(n)}_{j_0^{(n)}}\| \leq 2^{-n}.$$

In particular, $(\widetilde{z}_n)$ is Cauchy in $J(X) \subseteq \widetilde{X}$ and thus $(\widetilde{x}_n)$ is Cauchy in X. Now we let $z_0 := [(\widetilde{x}_j)_j] \in \widetilde{X}$. □

end Thu 25.10.18

Remark: If $(\widetilde{X}_j, J_j)$, $j = 1, 2$, are completions of X then there exists a linear bijection $J : X_1 \to X_2$ such that $\|Jx_1\| = \|x_1\|$ for all $x_1 \in X_1$. This follows from 1.11, applied to $J_2 J_1^{-1}$, and an approximation argument for the isometric property.

Remark: One can also construct completions of metric spaces.

1.15. Proposition: Let X be a normed space and $Y \subseteq X$ be a closed linear subspace. Then the quotient space $X/Y = \{x + Y : x \in X\}$ is a vector space and

$$||x + Y|| := \inf\{\|z\| : z \in x + Y\}$$

defines a norm on X/Y. If X is a Banach space then $(X/Y, ||\cdot||)$ is a Banach space.

As we shall see in the proof, $||\cdot||$ is just a seminorm if Y is not closed in X. We will use the following lemma.

1.16. Lemma: A normed space X is a Banach space if and only if each absolutely convergent series is convergent, i.e. if $\sum_{j=1}^{\infty} \|x_j\| < \infty$ implies convergence of $\sum_{j=1}^{\infty} x_j$ in X.

Proof. "$\Rightarrow$": By $\| \sum_{j=n}^{m} x_j \| \leq \sum_{j=n}^{m} \|x_j\|$ the partial sums of an absolutely convergent sequence are a Cauchy sequence, hence convergent.

"$\Leftarrow$": Let (x_j) be Cauchy sequence in X. We shall find its limit as the sum of an absolutely convergent series. By Exercise Sheet 1 we find a subsequence $(x_{k(j)})_j$ of (x_j) such that, for $y_j := x_{k(j+1)} - x_{k(j)}$ we have $\|y_j\| \leq 2^{-j}$ for all $j \in \mathbb{N}$. Then $\sum_{j=1}^{\infty} y_j$ is absolutely convergent, hence convergent with sum $y \in X$, say. By $\sum_{j=1}^{N} y_j = x_{k(N+1)} - x_{k(1)}$ we see that $x_{k(j)} \to y + x_{k(1)} =: x$ as $j \to \infty$. This implies that $x_j \to x$, and we have shown completeness of X. □

Proof of 1.15. As usual we write $[x] := x + Y$. (N1): If $\|[x]\| = 0$ there exists a sequence (z_n) in $[x]$ with $z_n \to 0$. By closedness of Y we obtain $0 \in [x]$, i.e. $[x] = [0]$. Clearly $[0] = 0$. For $\alpha \neq 0$ we have $z \in x + Y \Leftrightarrow \alpha z \in \alpha x + Y$. This implies (N2). For the proof of (N3) let $x_j \in X$, $j = 1, 2$, and $\varepsilon > 0$. We find $z_j \in [x_j]$, $j = 1, 2$, with $\|z_j\| \leq ||[x_j]|| + \varepsilon$. Then $z_1 + z_2 \in [x_1] + [x_2]$ and

$$\|z_1 + z_2\| \leq \|z_1\| + \|z_2\| \leq ||[x_1]|| + ||[x_2]|| + 2\varepsilon,$$

hence $||[x_1] + [x_2]|| \leq ||[x_1]|| + ||[x_2]|| + 2\varepsilon$, and we obtain (N3) since $\varepsilon > 0$ was arbitrary.

Now let X be a Banach space. We use Lemma 1.16 and let $\sum_j ||[x_j]|| < \infty$. For any $j \in \mathbb{N}$ we find $z_j \in [x_j]$ such that $\|z_j\| \leq ||[x_j]|| + 2^{-j}$. Hence $\sum_j \|z_j\| < \infty$, and thus the series converges with sum $z \in X$, say. But then $\sum_j [x_j] = \sum_j [z_j]$ converges with sum $[z]$, since

$$||[z] - \sum_{j=1}^{N} [z_j]|| \leq \|z - \sum_{j=1}^{N} z_j\| \to 0 \quad (N \to \infty).$$

□

Remark: If X is a metric space, $x \in X$ and $\emptyset \neq A \subseteq X$ then the *distance* $d(x, A)$ *of* x *to* A is given by

$$d(x, A) := \inf\{d(x, y) : y \in A\}.$$

It is easy to see that $d(x, A) = 0 \Leftrightarrow x \in \overline{A}$ and that $|d(x, A) - d(y, A)| \leq d(x, y)$, so $d(\cdot, A)$ is Lipschitz-continuous.

For the norm $|| \cdot ||$ we have

$$||[x]|| = \inf\{\|x - y\| : y \in Y\} = d(x, Y),$$

i.e. the norm of $[x] = x + Y$ in the quotient space X/Y is the distance of x to the closed linear subspace Y.

Remark: Let X, Z be normed spaces and $T \in \mathcal{L}(X,Z)$. Then $N(T)$ is a closed linear subspace of X and we can factorize T as $T = \widehat{T} \circ q$ where $q : X \to X/N(T)$, $x \mapsto [x]$ is the quotient map and $\widehat{T} \in \mathcal{L}(X/N(T), Z)$ is given by $\widehat{T}([x]) := Tx$. The operator $\widehat{T}$ is injective and satisfies $\|\widehat{T}\| = \|T\|$.

1.17. Lebesgue spaces: Let (Ω, Σ, μ) be a measure space, i.e. Σ is a σ-algebra in Ω and $\mu : \Sigma \to [0,\infty]$ is σ-additive. Usually we omit Σ in notation as it is implicitly given as the domain of definition of μ. We denote by 1_A the *characteristic function of* $A \subseteq \Omega$ given by $1_A(\omega) = 1$ if $\omega \in A$ and $1_A(\omega) = 0$ otherwise. We denote by

$$S(\mu) := \text{lin}\{1_A : A \in \Sigma, \mu(A) < \infty\}$$

the vector space of μ-*simple functions*.

1) For $\varphi = \sum_{j=1}^n \xi_j 1_{A_j} \in S(\mu)$ let $\int \varphi \, d\mu := \sum_{j=1}^n \xi_j \mu(A_j)$. This defines a linear map

$$\int \cdot \, d\mu : S(\mu) \to \mathbb{K}, \quad \varphi \mapsto \int \varphi \, d\mu.$$

2) If $\varphi \in S(\Sigma, \mu)$ and $p \in [1,\infty)$ then $|\varphi|^p \in S(\mu)$. We recall that any $\varphi \in S(\mu)$ has representations

$$\varphi = \sum_{j=1}^n \xi_j 1_{A_j}$$

where the $A_j \in \Sigma$ with $\mu(A_j) < \infty$ are pairwise disjoint. Then

$$|\varphi|^p = \sum_{j=1}^n |\xi_j|^p 1_{A_j}.$$

3) Seminorms: If $p \in [1,\infty)$ then

$$\|\varphi\|_p := \Big(\int |\varphi|^p \, \mu \Big)^{1/p}$$

defines a seminorm on $S(\mu)$ and $\|\varphi\|_p = 0 \Leftrightarrow \varphi(\omega) = 0$ for μ-a.e. $\omega \in \Omega$. We clearly have, for all $\varphi \in S(\mu)$,

$$\Big| \int \varphi \, d\mu \Big| \le \int |\varphi| \, d\mu = \|\varphi\|_1.$$

Moreover, if $p \in (1,\infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$, then we have, for all $\varphi, \psi \in S(\mu)$,

$$\Big| \int \varphi\psi \, d\mu \Big| \le \|\varphi\psi\|_1 \le \|\varphi\|_p \|\psi\|_q \qquad \text{(Hölder's inequality)}.$$

For the proof we find pairwise disjoint A_j and $\xi_j, \eta_j \in \mathbb{K}$ such that $\varphi = \sum_{j=1}^n \xi_j 1_{A_j}$ and $\psi = \sum_{j=1}^n \eta_j 1_{A_j}$. Then $\varphi\psi = \sum_{j=1}^n \xi_j \eta_j 1_{A_j}$ and

$$\|\varphi\psi\|_1 = \sum_j |\xi_j| \mu(A_j)^{1/p} |\eta_j| \mu(A_j)^{1/q} \le \underbrace{\Big(\sum_j |\xi_j|^p \mu(A_j) \Big)^{1/p}}_{=\|\varphi\|_p} \underbrace{\Big(\sum_j |\eta_j|^q \mu(A_j) \Big)^{1/q}}_{=\|\psi\|_q}.$$

Remark: As in 1.14 we could construct a "completion" of the seminormed space $(S(\mu), \|\cdot\|_p)$ by considering equivalence classes of $\|\cdot\|_p$-Cauchy sequences. But one can find functions that represent this rather abstract concept.

4) Preparation: If (φ_n) is a $\|\cdot\|_p$-Cauchy sequence in $S(\mu)$ then it contains a subsequence $(\varphi_{k(n)})_n$ that converges μ-a.e. to a function f. If $(\varphi_{l(n)})_n$ is another subsequence converging μ-a.e. to a function g then $f = g$ μ-a.e.. In order to see this we prove a lemma.

Lemma: If (φ_n) is a sequence in $S(\mu)$ such that $C := \sum_{n=1}^\infty \|\varphi_n\|_p < \infty$ then $\sum_{n=1}^\infty \varphi_n(\omega)$ converges for μ-a.e. $\omega \in \Omega$.

Proof. For any $N \in \mathbb{N}$ we have

$$\int \underbrace{\Big(\sum_{n=1}^N |\varphi_n|\Big)^p}_{=:\rho_N} d\mu = \|\sum_{n=1}^N |\varphi_n|\|_p^p \le \Big(\sum_{n=1}^\infty \|\varphi_n\|_p\Big)^p \le C^p.$$

Now monotone convergence yields $\int \sup_N \rho_N \, d\mu \le C^p < \infty$, hence $\sum_{n=1}^\infty |\varphi_n| < \infty$ μ-a.e., moreover, $(\sum_{n=1}^\infty |\varphi_n|)^p$ is integrable.[1] □

Now let (φ_n) be $\|\cdot\|_p$-Cauchy. Then we find a subsequence $(\varphi_{k(n)})_n$ satisfying

$$\|\varphi_{k(n+1)} - \varphi_{k(n)}\|_p \le 2^{-n} \quad \text{for all } n \in \mathbb{N}.$$

By the lemma

$$(\varphi_{k(n+1)})_n := \Big(\varphi_{k(1)} + \sum_{j=1}^n (\varphi_{k(j+1)} - \varphi_{k(j)})\Big)_n$$

converges μ-a.e. to a function f. If $(\varphi_{l(n)})_n$ is another subsequence that converges μ-a.e. to a function g then we can arrange for $\|\varphi_{k(n)} - \varphi_{l(n)}\|_p \le 2^{-n}$ (by taking subsequences if necessary). By the lemma $\sum_{n=1}^\infty (\varphi_{k(n)} - \varphi_{l(n)})$ converges μ-a.e., so $\varphi_{k(n)} - \varphi_{l(n)} \to 0$ μ-a.e.. On the other hand $\varphi_{k(n)} - \varphi_{l(n)} \to f - g$ μ-a.e., so $f = g$ μ-a.e.. end Mon 29.10.18

5) Definition: For $p \in [1, \infty)$ we define $\mathscr{L}^p(\Omega, \mu)$ to be the set of all functions $f : \Omega \to \mathbb{K}$ such that there exists a $\|\cdot\|_p$-Cauchy sequence $(\varphi_n)_{n\in\mathbb{N}}$ such that $\varphi_n(\omega) \to f(\omega)$ for μ-a.e. $\omega \in \Omega$. In this case we set $\|f\|_p := \lim_{n\to\infty} \|\varphi_n\|_p$ (recall that the limit exists).

---

[1] One can avoid the use of monotone convergence: We let

$$M := \{\omega \in \Omega : \sup_N \rho_N(\omega) = \infty\} = \bigcap_{k\in\mathbb{N}} \underbrace{\{\omega \in \Omega : \sup_N \rho_N(\omega) > k\}}_{=:B_k} = \bigcap_{k\in\mathbb{N}} \bigcup_{N\in\mathbb{N}} \underbrace{\{\omega \in \Omega : \rho_N(\omega) > k\}}_{=:A_k^N} \in \Sigma.$$

For any $k \in \mathbb{N}$ we have $k\mu(A_k^N) \le \int \rho_N \, d\mu \le C$ (this is Tchebyshev's inequality), hence $\mu(A_k^N) \le C/k$. By $A_k^N \subseteq A_k^{N+1}$ this implies $\mu(B_k) \le C/k$ and we conclude $\mu(M) \le \inf_k (C/k) = 0$, i.e. $\sum_{n=1}^\infty |\varphi_n|$ converges μ-a.e..

This is well-defined: It suffices to show that $\|\varphi_n\|_p \to 0$ holds if (φ_n) is $\|\cdot\|_p$-Cauchy with $\varphi_n \to 0$ μ-a.e.. Resorting to a subsequence, we may assume that $\|\varphi_{j+1} - \varphi_j\|_p \leq 4^{-j}$ for all $j \in \mathbb{N}$. This follows from 4) (integrability of $(\sum_{n=1}^{\infty} |\varphi_n|)^p$) and dominated convergence.[2]

We remark that, for fixed k, $(\varphi_n - \varphi_k)_n$ is $\|\cdot\|_p$-Cauchy and converges μ-a.e. to $f - \varphi_k$, hence $\|f - \varphi_k\|_p = \lim_{n\to\infty} \|\varphi_n - \varphi_k\|_p$. By the Cauchy property of (φ_n) this implies $\|f - \varphi_k\|_p \to 0$ as $k \to \infty$, so f represents the equivalence class of the Cauchy sequence (φ_n).

It is clear that $\mathscr{L}^p(\Omega, \mu)$ is a $\mathbb{K}$-vector space. Moreover, $\|\cdot\|_p$ is a seminorm on $\mathscr{L}^p(\Omega, \mu)$.

6) Completeness: The seminormed space $(\mathscr{L}^p(\Omega, \mu), \|\cdot\|_p)$ is complete in the sense that, for any $\|\cdot\|_p$-Cauchy sequence (f_n), there exists $f \in \mathscr{L}^p(\Omega, \mu)$ such that $\|f - f_n\|_p \to 0$.

We find a sequence (φ_n) in $S(\mu)$ such that $\|f_n - \varphi_n\|_p < 2^{-n}$ for all n, hence also (φ_n) is $\|\cdot\|_p$-Cauchy. According to 4) and 5) we find f such that a subsequence of (φ_n) converges to f w.r.t. $\|\cdot\|_p$. This implies $\|f - f_n\|_p \to 0$.

7) Hölder's inequality: If $f \in \mathscr{L}^p(\Omega, \mu)$ and $g \in \mathscr{L}^q(\Omega, \mu)$ then $fg \in \mathscr{L}^1(\Omega, \mu)$ and $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

We find sequences (φ_n) for f and (ψ_n) for g in $S(\mu)$ as in 5). By Hölder's inequality in 3) we obtain that $(\varphi_n \psi_n)$ is $\|\cdot\|_1$-Cauchy. Since $\varphi_n \psi_n \to fg$ μ-a.e. we obtain $fg \in \mathscr{L}^1(\Omega, \mu)$ and the inequality follows by 3) and a limit argument.

---

[2] One can avoid the use of dominated convergence here. We proceed in two steps.
Step 1: By $0 = \varphi_1 + \sum_{j=1}^{\infty} (\varphi_{j+1} - \varphi_j)$ (see 4)) outside $N \in \Sigma$ with $\mu(N) = 0$, we have $\varphi_{n+1} = \sum_{j=n+1}^{\infty} (\varphi_j - \varphi_{j+1})$ on $\Omega \setminus N$, and thus, for any $\alpha > 0$:

$$\{|\varphi_{n+1}| > \alpha\} \subseteq N \cup \bigcup_{j=n+1}^{\infty} \{|\varphi_{j+1} - \varphi_j| > \alpha 2^{n-j}\}.$$

Now we use Tchebyshev's inequality again:

$$\mu\{|\varphi_{n+1}| > \alpha\} \leq \sum_{j=n+1}^{\infty} \mu\{|\varphi_{j+1} - \varphi_j| > \alpha 2^{n-j}\} \leq \alpha^{-p} \sum_{j=n+1}^{\infty} 2^{p(j-n)} \|\varphi_{j+1} - \varphi_j\|_p^p \leq \alpha^{-p} 2^{-np} \underbrace{\sum_{j>n} 2^{-jp}}_{\leq 1},$$

which shows that $\mu\{|\varphi_n| > \alpha\} \to 0$ as $n \to \infty$ for any $\alpha > 0$.
Step 2: Let $\varepsilon > 0$. We find n such that $\|\varphi_m - \varphi_n\|_p < \varepsilon$. We let $U := \{\varphi_n \neq 0\}$ and $M := \max\{\mu(U)^{1/p}, \|\varphi_n\|_\infty\}$. According to Step 1 we find $m_0 \geq n$ such that $\mu\{|\varphi_m| > \varepsilon/M\} < (\varepsilon/M)^p$ for all $m \geq m_0$. For $m \geq m_0$ and $P := \{|\varphi_m| > \varepsilon/M\}$ we then have

$$\|\varphi_m\|_p^p = \int 1_{(\Omega\setminus P)\cap U} |\varphi_m|^p \, d\mu + \int 1_{P\cap U} |\varphi_m|^p \, d\mu + \int 1_{\Omega\setminus U} |\varphi_m|^p \, d\mu = I_1 + I_2 + I_3.$$

Outside P we have $|\varphi_m| \leq \varepsilon/M$, hence $I_1 \leq (\varepsilon/M)^p \mu(U) \leq \varepsilon^p$. Moreover,

$$I_2^{1/p} \leq \|1_P \varphi_n\|_p + \|\varphi_n - \varphi_m\|_p \leq \mu(P)^{1/p} \|\varphi_n\|_\infty + \varepsilon \leq 2\varepsilon,$$

and finally $I_3 = \|1_{\Omega\setminus U}(\varphi_m - \varphi_n)\|_p^p \leq \|\varphi_n - \varphi_m\|_p^p \leq \varepsilon^p$. We have shown $\|\varphi_m\|_p \leq 4\varepsilon$ for any $m \geq m_0$, i.e. $\|\varphi_m\|_p \to 0$.

8) Banach spaces: For $p \in [1, \infty)$ we have that

$$\mathscr{N}(\mu) := \{f \in \mathscr{L}^p(\Omega, \mu) : \|f\|_p = 0\} = \{f : \Omega \to \mathbb{K} : \; f = 0 \; \mu\text{-a.e.}\},$$

is a linear subspace of $\mathscr{L}^p(\Omega, \mu)$ and for $f, g \in \mathscr{L}^p(\Omega, \mu)$ with $f - g \in \mathscr{N}(\mu)$ we have $\|f\|_p = \|g\|_p$. Then the quotient space

$$L^p(\Omega, \mu) := \mathscr{L}^p(\Omega, \mu)/\mathscr{N}(\mu) = \{f + \mathscr{N}(\mu) : f \in \mathscr{L}^p(\Omega, \mu)\}$$

is a Banach space with respect to $\|f + \mathscr{N}(\mu)\|_p := \|f\|_p$. Usually this is phrased as "identifying functions that are equal μ-a.e.".

Remark: In the same way one can define spaces $\mathscr{L}^p(\Omega, \mu; X)$ and $L^p(\Omega, \mu; X)$ for $p \in [1, \infty)$, where $\mathscr{L}^p(\Omega, \mu; X)$ consists of functions $f : \Omega \to X$ and X is a Banach space, and the *Bochner integral* $\int f \, d\mu$ for functions $f \in \mathscr{L}^1(\Omega, \mu; X)$.

Any $f \in \mathscr{L}^p(\Omega, \mu; X)$ is μ-a.e. limt of a sequence of μ-simple functions (φ_n) of the form $\varphi_n = \sum_{j=1}^{m_n} x_j^{(n)} 1_{A_j^{(n)}}$ with $x_j^{(n)} \in X$ and $A_j^{(n)} \in \Sigma$, $\mu(A_j^{(n)}) < \infty$. Hence there exist $N \in \Sigma$ with $\mu(N) = 0$ and a separable set $A \subseteq X$ such that $f(\omega) \in A$ for $\omega \in \Omega \setminus N$ (just take $A = \overline{\{x_j^{(n)} : n \in \mathbb{N}, j = 1, \ldots, m_n\}}$).

1.18. Measurable functions: The *Borel σ-algebra* $\mathcal{B} = \mathcal{B}(\mathbb{K})$ of $\mathbb{K}$ is the smallest σ-algebra in $\mathbb{K}$ containing all open subsets of $\mathbb{K}$, i.e. $\mathcal{B}(\mathbb{K}) = \sigma(\{Q \subseteq \mathbb{K} : \; Q \text{ is open}\})$. For $\mathbb{K} = \mathbb{R}$ we have

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(a, b] : a, b \in \mathbb{R}, a < b\}) = \sigma(\{(-\infty, b] : b \in \mathbb{R}\}) = \sigma(\{(a, \infty) : a \in \mathbb{R}\}).$$

1) Σ-measurable functions: Let (Ω, Σ) be a measurable space, i.e. Σ is a σ-algebra in the set Ω. A function $f : \Omega \to \mathbb{K}$ is called Σ-*measurable* if $f^{-1}(B) \in \Sigma$ for any $B \in \mathcal{B}(\mathbb{K})$. For $\mathbb{K} = \mathbb{R}$ this is the case if and only if $\{\omega \in \Omega : f(\omega) \leq b\} \in \Sigma$ for any $b \in \mathbb{R}$. For $\mathbb{K} = \mathbb{C}$ a function f is Σ-measurable if and only if its real and imaginary parts are Σ-measurable.

Lemma: If (f_n) is a sequence of Σ-measurable functions $\Omega \to \mathbb{K}$ that converges pointwise to a function $f : \Omega \to \mathbb{K}$ then f is Σ-measurable. If we let $S(\Sigma) := \text{lin}\{1_A : A \in \Sigma\}$ then any Σ-measurable function is the pointwise limit of a sequence in $S(\Sigma)$. end Mon 05.11.18

2) Bounded Σ-measurable functions: We let $B(\Omega, \Sigma)$ denote the set of all bounded Σ-measurable functions $f : \Omega \to \mathbb{K}$. This is a closed linear subspace of $(B(\Omega), \|\cdot\|_\infty)$, hence a Banach space for the sup-norm $\|\cdot\|_\infty$. Any $f \in B(\Omega, \Sigma)$ is the uniform limit of a sequence in $S(\Sigma)$.

3) Remark: Let (Ω, Σ, μ) be a measure space and $p \in [1, \infty)$. A function $f \in \mathscr{L}^p(\Omega, \mu)$ is μ-a.e. limit of a sequence in $S(\mu) \subseteq S(\Sigma)$, but might not be Σ-measurable: For $M \subseteq \Omega$ the function 1_M is Σ-measurable if and only if $M \in \Sigma$, but $1_M = 0$ μ-a.e. if and only if there exists $N \in \Sigma$ with $\mu(N) = 0$ and $M \subseteq N$. This is as bad as it gets: Denoting

$$\widetilde{\Sigma} := \sigma\big(\Sigma \cup \{M \subseteq \Omega : \exists N \in \Sigma : \mu(N) = 0, M \subseteq N\}\big),$$

the measure μ has a unique extension to a measure $\widetilde{\mu} : \widetilde{\Sigma} \to [0, \infty]$ and we have

4) Proposition: Let $p \in [1, \infty)$ and $f \in \mathscr{L}^p(\Omega, \mu)$. Then f is $\widetilde{\Sigma}$-measurable and there exists a Σ-measurable function f_0 such that $f = f_0$ μ-a.e.. In particular, $L^p(\Omega, \mu)$ can also be represented as

$$\{f \in \mathscr{L}^p(\Omega, \mu) : \ f \text{ is } \Sigma\text{-measurable}\}/\{f \in \mathscr{N}(\mu) : \ f \text{ is } \Sigma\text{-measurable}\}.$$

Moreover, $\mathscr{L}^p(\Omega, \widetilde{\mu}) = \mathscr{L}^p(\Omega, \mu)$, $\mathscr{N}(\widetilde{\mu}) = \mathscr{N}(\mu)$, and $L^p(\Omega, \widetilde{\mu}) = L^p(\Omega, \mu)$. Finally, a function $g : \Omega \to \mathbb{K}$ is $\widetilde{\Sigma}$-measurable if and only if g is *μ-measurable*, i.e. g is the μ-a.e. limit of a sequence in $S(\Sigma)$.

Remark: For functions $f : \Omega \to X$ where X is a Banach space this last property is used and thus there exists a separable subset A of X and a μ-null set N such that $f(\omega) \in A$ for all $\omega \in \Omega \setminus N$.

1.19. Essentially bounded functions: A μ-measurable function $f : \Omega \to \mathbb{K}$ is called *essentially bounded* if there exists $N \in \Sigma$ with $\mu(N) = 0$ such that $1_{\Omega \setminus N} f$ is bounded. We denote by $\mathscr{L}^\infty(\Omega, \mu)$ the space of all μ-measurable and essentially bounded functions $f : \Omega \to \mathbb{K}$. On this space,

$$\|f\|_\infty := \inf\{ \sup_{\omega \in \Omega \setminus N} |f(\omega)| \ : \ N \in \Sigma, \mu(N) = 0\} = \inf\{C > 0 : \{|f| > C\} \text{ is a } \mu\text{-null set}\}$$

defines a seminorm with $\|f\|_\infty = 0 \Leftrightarrow f \in \mathscr{N}(\mu)$. The quotient space

$$L^\infty(\Omega, \mu) := \mathscr{L}^\infty(\Omega, \mu)/\mathscr{N}(\mu)$$

is a Banach space for the induced norm $\|\cdot\|_\infty$.[3] The assertion of 1.18 4) also holds for $p = \infty$: In every co-class we find essentially bounded Σ-measurable functions, and $L^\infty(\Omega, \mu)$ may also be represented as

$$\{f \in \mathscr{L}^\infty(\Omega, \mu) : \ f \text{ is } \Sigma\text{-measurable}\}/\{f \in \mathscr{N}(\mu) : \ f \text{ is } \Sigma\text{-measurable}\}.$$

As for $\mathscr{L}^p$ with $p \in [1, \infty)$ the choice whether to work with Σ-measurable or μ-measurable functions is thus rather a matter of taste or context.

Lemma: Let $g \in L^\infty(\Omega, \mu)$. For all $p \in [1, \infty]$, the operator $L^p(\Omega, \mu) \to L^p(\Omega, \mu)$, $f \mapsto gf$, is bounded with norm $\le \|g\|_\infty$.

1.20. Proposition: Let $p \in [1, \infty)$ and $f : \Omega \to \mathbb{K}$. Then $f \in \mathscr{L}^p(\Omega, \mu)$ if and only if f is μ-measurable and $|f|^p \in \mathscr{L}^1(\Omega, \mu)$.

[3] Observe that for each $f \in \mathscr{L}^\infty(\Omega, \mu)$ there exists $N \in \Sigma$ with $\mu(N) = 0$ such that $\|f\|_\infty = \sup_{\omega \in \Omega \setminus N} |f(\omega)|$.

Proof. If $f \in \mathscr{L}^p(\Omega, \mu)$ and (φ_n) is a $\|\cdot\|_p$-Cauchy sequence with $\varphi_n \to f$ μ-a.e., then f is clearly μ-measurable. Moreover, $|\varphi_n|^p \to |f|^p$ μ-a.e., and we have, for $p = 1$,

$$\||\varphi_n| - |\varphi_m|\|_1 \leq \|\varphi_n - \varphi_m\|_1$$

by the reverse triangle inequality. For $p > 1$ we have

$$\big||\varphi_n|^p - |\varphi_m|^p\big| \leq |\varphi_n - \varphi_m| p\big(|\varphi_n|^{p-1} + |\varphi_m|^{p-1}\big)$$

by the mean value theorem. Integration and Hölder yield (with $\frac{1}{p} + \frac{1}{q} = 1$)

$$\||\varphi_n|^p - |\varphi_m|^p\|_1 \leq p\|\varphi_n\|_p\big(\|\varphi_n\|_p^{p/q} + \|\varphi_m\|_p^{p/q}\big).$$

Since every Cauchy sequence is bounded we obtain that $(|\varphi_n|^p)$ is $\|\cdot\|_1$-Cauchy, hence $|f|^p \in \mathscr{L}^1(\Omega, \mu)$.

For the reverse implication let f be μ-measurable and $|f|^p \in \mathscr{L}^1(\Omega, \mu)$. Then $\operatorname{sgn} f$[4]is μ-measurable, hence $\operatorname{sgn} f \in \mathscr{L}^\infty(\Omega, \mu)$. By the lemma in 1.19 we only have to show $|f| \in \mathscr{L}^p(\Omega, \mu)$. We find a $\|\cdot\|_1$-Cauchy sequence (φ_n) such that $\varphi_n \to |f|^p$ μ-a.e., where we may assume that $\varphi_n \geq 0$ (considering $|\varphi_n|$ otherwise). Then $\varphi_n^{1/p} \to |f|$ μ-a.e. and by

$$\big|\varphi_n^{1/p} - \varphi_m^{1/p}\big| \leq \big|\varphi_n - \varphi_m\big|^{1/p}$$

and integration we obtain that $(\varphi_n^{1/p})$ is $\|\cdot\|_p$-Cauchy. Hence $|f| \in \mathscr{L}^p(\Omega, \mu)$. □

Corollary: Let $p, q, r \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Then $(f, g) \mapsto fg$ is continuous $L^p(\Omega, \mu) \times L^q(\Omega, \mu) \to L^r(\Omega, \mu)$ and $\|fg\|_r \leq \|f\|_p\|g\|_q$.

1.21. Proposition: Let (Ω, Σ, μ) be a measure space and $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then $J : L^q(\Omega, \mu) \to (L^p(\Omega, \mu))'$ given by $Jg(f) := \int fg\, d\mu$ is well-defined linear and isometric. If Ω is σ-finite then the assertion holds for $p = 1$ and $q = \infty$.

Here we use the

Definition: A subset $A \subseteq \Omega$ is called *σ-finite* if there exists a sequence (A_j) in Σ such that $\mu(A_j) < \infty$ for each $j \in \mathbb{N}$ and $A \subseteq \bigcup_{j\in\mathbb{N}} A_j$.

Notice that, if $p \in [1, \infty)$ and $f \in \mathscr{L}^p(\Omega, \mu)$, then $\{f \neq 0\}$ is σ-finite.

Proof. Assume $p \in [1, \infty)$. By Hölder's inequality we have $|(Jg)(f)| \leq \|fg\|_1 \leq \|g\|_q\|f\|_p$, so J is well-defined and linear and $\|Jg\|_{(L^p)'} \leq \|g\|_q$.

Now let $p \in (1, \infty)$ and $g \in L^q(\Omega, \mu)$. We use 1.20. Set $f := \overline{\operatorname{sgn} g}|g|^{q-1}$. Then f is μ-measurable and $|f|^p = |g|^{pq-p} = |g|^q \in L^1(\Omega, \mu)$, hence $f \in L^p(\Omega, \mu)$ and $\|f\|_p^p = \|g\|_q^q$. Since $\frac{q}{p} = q - 1$, we obtain

$$Jg(f) = \int fg\, d\mu = \int |g|^q\, d\mu = \|g\|_q^q = \|g\|_q\|f\|_p,$$

which implies $\|Jg\|_{(L^p)'} \geq \|g\|_q$. end Thu 08.11.18

Finally let $p = 1$, let Ω be σ-finite, and let $g \in L^\infty(\Omega, \mu)$. We may assume $\|g\|_\infty > 0$ and fix $\varepsilon \in (0, \|g\|_\infty)$. We also assume that g is σ-measurable. Then $M_0 := \{\omega \in \Omega : |g(\omega)| > \|g\|_\infty - \varepsilon\} \in \Sigma$ satisfies $\mu(M_0) > 0$. Since Ω is σ-finite we find a subset $M \subseteq M_0$ in Σ such that $\mu(M) \in (0, \infty)$. Now we let $f := (\operatorname{sgn} g)1_M$. Then $f \in L^1(\Omega, \mu)$ and $\|f\|_1 = \mu(M)$. Moreover

$$Jg(f) = \int fg \, d\mu = \int 1_M |g| \, d\mu \geq (\|g\|_\infty - \varepsilon)\mu(M) = (\|g\|_\infty - \varepsilon)\|f\|_1.$$

We conclude $\|Jg\|_{(L^1)'} \geq \|g\|_\infty$. □

We collect without proof some useful results in

1.22. Proposition: (a) Let $p \in [1, \infty)$. If (f_n) is a sequence in $\mathscr{L}^p(\Omega, \mu)$ and $f \in \mathscr{L}^p(\Omega, \mu)$ such that $\|f - f_n\|_p \to 0$ then there exists a subsequence $(f_{k(n)})$ and a function $g \geq 0$ in $\mathscr{L}^p(\Omega, \mu)$ such that $f_{k(n)} \to f$ μ-a.e. and $\sup_n |f_{k(n)}| \leq g$ μ-a.e..

(b) **Monotone Convergence Theorem:** Suppose (f_n) is a sequence in $\mathscr{L}^1(\Omega, \mu)$ satisfying $f_n \leq f_{n+1}$ and $\sup_n \int f_n \, d\mu < \infty$. Then $f := \sup_n f_n < \infty$ μ-a.e., $f 1_{\{f<\infty\}} \in \mathscr{L}^1(\Omega, \mu)$ and $\int f \, d\mu = \sup_n \int f_n \, d\mu$.

(c) **Dominated Convergence Theorem:** Suppose that (f_n) is a sequence in $\mathscr{L}^1(\Omega, \mu)$, f is function such that $f_n \to f$ μ-a.e., and there exists a function $g \in \mathscr{L}^1(\Omega, \mu)$ such that $\sup_n |f_n| \leq g$ μ-a.e.. Then $f \in \mathscr{L}^1(\Omega, \mu)$ and $\int f \, d\mu = \lim_n \int f_n \, d\mu$.

1.23. Definition: Let (X, d) be a metric space and $K \subseteq X$. A family $\mathcal{Q}$ of subsets is called an *open cover* of K if all $Q \in \mathcal{Q}$ are open and $K \subseteq \bigcup \mathcal{Q} = \bigcup_{Q \in \mathcal{Q}} Q$.

K is called *compact* if any open cover $\mathcal{Q}$ of K contains a finite subcover, i.e. a finite subset $\mathcal{E}$ which is still an open cover of K.

K is called *relatively compact* if its closure $\overline{K}$ is compact.

K is called *totally bounded* if, for every $\varepsilon > 0$, there exists a finite set $F \subseteq X$ such that $K \subseteq \bigcup_{x \in F} B(x, \varepsilon)$.

K is called *sequentially compact* if every sequence in K contains a convergent subsequence with limit in K.

Remark: If K is totally bounded and $\varepsilon > 0$ we can also find a finite set F in the smaller set K such that $K \subseteq \bigcup_{x \in F} B(x, \varepsilon)$.

By the Proposition in 1.3 it makes no difference whether the $Q \in \mathcal{Q}$ are open in X or relatively open in K, i.e. compactness of K does not depend on the ambient space X. Similarly, total boudedness of K does not depend on he ambient space X.

[4] Here $\operatorname{sgn} f(\omega) = \operatorname{sgn}(f(\omega))$ where $\operatorname{sgn} 0 := 0$ and $\operatorname{sgn} z := z/|z|$ for $z \in \mathbb{C} \setminus \{0\}$.

1.24. Proposition: Let (X, d) be a metric space and $K \subseteq X$. Then the following are equivalent:

(i) K is compact.

(ii) K is *sequentially compact.*

(iii) K is totally bounded and complete.

Proof. We show that (i) implies (ii): If (x_n) is a sequence in K without convergent subsequence then we find, for every $x \in K$ a radius $\varepsilon_x > 0$ such that $B(x, \varepsilon_x)$ contains x_n only for finitely many n. But then the open cover $\mathcal{Q} = \{B(x, \varepsilon_x) : x \in K\}$ contains a finite subcover $\bigcup_{x \in F} B(x, \varepsilon_x)$ which can only contain x_n for finitely many n, a contradiction.

Before we proceed we show a lemma.

Lemma: K is totally bounded if and only if every sequence in K contains a Cauchy subsequence.

Proof. If K is totally bounded and (x_n) is sequence in K then it contains a Cauchy subsequence: We prove this by a diagonal argument. For $\varepsilon = 1$ we find an open ball B_1 of radius 1 that contains a subsequence $(x_{k_1(n)})$ of (x_n). For $\varepsilon = 1/2$ we find an open ball B_2 of radius $1/2$ that contains a subsequence $(x_{k_2(n)})$ of $(x_{k_1(n)})$ etc. In other words, subsequently we find, for each $j \in \mathbb{N}$, an open ball B_j of radius $1/j$ that contains a subsequence $(x_{k_j(n)})$ of $(x_{k_{j-1}(n)})$ (where we put $k_0(n) := n$). Finally we let $l(n) := k_n(n)$. Then $(x_{l(n)})$ is a subsequence of (x_n) and for all $m \geq n$ we have $x_{l(n)}, x_{l(m)} \in B_n$, hence $d(x_{l(n)}, x_{l(m)}) < 2/n$, i.e. $(x_{l(n)})$ is Cauchy.

Conversely, suppose that K is not totally bounded. then we find a radius $\varepsilon > 0$ such that K cannot be covered by finitely many open balls of radius ε. Choose $x_1 \in K$. By assumption we find $x_2 \in K \setminus B(x_1, \varepsilon)$, then $x_3 \in K \setminus (B(x_1, \varepsilon) \cup B(x_2, \varepsilon))$ etc. We thus construct a sequence (x_n) in K with $x_n \notin \bigcup_{j=1}^{n-1} B(x_j, \varepsilon)$ for every n. In particular, $d(x_n, x_m) \geq \varepsilon$ for $n \neq m$, hence (x_n) does not contain a Cauchy subsequence. $\square$

By the lemma, (iii) implies (ii). Clearly, (ii) implies completeness of K, so by the lemma, (ii) implies (iii). We finally assume $K = X$ and that (ii) and (iii) hold and show (i). To this end we assume that $\mathcal{Q}$ is an open cover of K which contains no finite subcover. We find subsequently finite subsets F_k of K such that the open balls $B(x, 2^{-k})$, $x \in F_k$, cover K. By assumption we find $x_1 \in F_1$ such that $\mathcal{Q}$ contains no finite subcover of $U_1 := B(x_1, 2^{-1})$. Then $\{U_1 \cap B(x, 2^{-2}) : x \in F_2\}$ is a finite cover of U_1 and we find $x_2 \in F_2$ such that $\mathcal{Q}$ contains no finite subcover of $U_2 := U_1 \cap B(x_2, 2^{-2})$ etc. In this way we construct a sequence (x_n) such that $\mathcal{Q}$ contains no finite subcover of $U_n = \bigcap_{j=1}^{n} B(x_j, 2^{-j})$. In particular, $B(x_n, 2^{-n}) \cap B(x_{n+1}, 2^{-(n+1)}) \neq \emptyset$, so $d(x_n, x_{n+1}) < 2^{-n} + 2^{-(n+1)} \leq 2^{1-n}$ for every n, which implies that (x_n) is Cauchy. By (ii), (x_n) has a convergent subsequence with limit $x_0 \in K$, so $x_n \to x_0$. But then we find $Q_0 \in \mathcal{Q}$ that contains x_0. Since Q_0 is open we find $r > 0$ such that $B(x_0, r) \subseteq Q_0$. Then we find n_0 such that $d(x_n, x_0) < r/2$ and $2^{-n} < r/2$ for $n \geq n_0$. We conclude that, for $n \geq n_0$,

$$U_n \subseteq B(x_n, 2^{-n}) \subseteq B(x, r) \subseteq Q_0,$$

a contradiction as $\mathcal{Q}$ contains no finite subcover of U_n. □

Remark: (a) If K is totally bounded then K is *bounded*, i.e. $\sup\{d(u, v) : u, v \in K\} < \infty$.[5]

end Mon 12.11.18

Proof. We find a finite subset $E \subseteq K$ such that $K \subseteq \bigcup_{x \in E} B(x, 1)$ and set $M := \max\{d(x, y) : x, y \in E\}$. Let $u, v \in K$. We find $x, y \in E$ such that $d(x, u) <$ and $d(y, v) < 1$. Then

$$d(u, v) \leq d(u, x) + d(x, y) + d(y, v) < 1 + M + 1 = M + 2.$$

□

(b) If $K \subseteq X$ is complete then K is closed in X: If (x_n) is a sequence in K with $x_n \to X \in X$ then (x_n) is Cauchy in K and hence has a limit in K. Since the limit is unique we conclude $x \in K$.

(c) By (a) and (b) we have that a compact subset K of a metric space is always closed and bounded. It should be known that the converse holds in $(\mathbb{K}^n, \|\cdot\|_\infty)$, and that any non-empty compact subset of $\mathbb{R}$ contains a maximal and a minimal element.

(d) Any totally bounded subset of a metric space is separable.

(e) Any closed subset of a compact set is compact. Any subset of a compact set is relatively compact.

1.25. Proposition: Let X, Y be metric spaces where X is compact and let $f : X \to Y$ be continuous. Then f is uniformly continuous and $f(X)$ is compact.

Proof. Let $\varepsilon > 0$. Then

$$\mathcal{Q} := \{B(x, \delta/2) : x \in X, \delta > 0, f(B(x, \delta)) \subseteq B(f(x), \varepsilon/2)\}$$

is an open cover of X (since f is continuous). Hence we find a finite subcover $\mathcal{E} = \{B(x_j, \delta_j/2) : j = 1, \ldots, n\} \subseteq \mathcal{Q}$. We set $\delta_0 := \min_j \delta_j/2$, then $\delta_0 > 0$. Let $x, y \in X$ with $d(x, y) < \delta_0$. Since $\mathcal{E}$ covers X we find $j \in \{1, \ldots, n\}$ such that $d(x, x_j) < \delta_j/2$. Then $d(y, x_j) \leq d(x, y) + d(x, x_j) < \delta_0 + \delta_j/2 \leq \delta_j$. Hence

$$d(f(x), f(y)) \leq d(f(x), f(x_j)) + d(f(x_j), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

[5] If $(X, \|\cdot\|)$ is a normed space the $K \subseteq X$ is bounded if and only if $\sup\{\|x\| : x \in K\} < \infty$.

We have proved uniform continuity of f. Now we prove compactness of $f(X)$. Let $\mathcal{Q}$ be an open cover of $f(X)$. Then $\{f^{-1}(Q) : Q \in \mathcal{Q}\}$ is an open cover of X (since f is continuous). Since X is compact we find a finite subset $\mathcal{E} \subseteq \mathcal{Q}$ such that $\{f^{-1}(Q) : Q \in \mathcal{E}\}$ covers X. But then $\mathcal{E}$ covers $f(X)$. □

Corollary: If X is a compact metric space then

$$C(X) = C_b(X) = BUC(X)$$

and for any $f \in C(X)$ we find $x_0 \in X$ such that $\|f\|_\infty = |f(x_0)|$, i.e. $\|f\|_\infty = \max_{x\in X} |f(x)|$.

1.26. Equivalence of norms: Let X be a $\mathbb{K}$-vector space. Two norms $\|\cdot\|_0$ and $\|\cdot\|_1$ on X are equivalent if there exist constants $c, C > 0$ such that

$$c\|x\|_1 \le \|x\|_2 \le C\|x\|_1 \quad \text{for all } x \in X,$$

or, in other words, if the identity map $I : (X, \|\cdot\|_j \to (X, \|\cdot\|_{1-j})$ is continuous for $j = 0, 1$. This clearly defines an equivalence relation on the set of norms on X. Moreover, it is clear that open and closed sets, convergent sequences and Cauchy sequences, compact subsets and continuous, uniformly continuous and Lipschitz continuous functions for equivalent norms coincide.

Example: For $x \in \mathbb{K}^n$ and $p \in (1, \infty)$ we have

$$\|x\|_\infty \le \|x\|_p \le \|x\|_1 \le n\|x\|_\infty,$$

hence all norms $\|\cdot\|_p$, $p \in [1, \infty]$, are equivalent on $\mathbb{K}^n$.

Proposition: Let X be a finite-dimensional $\mathbb{K}$-vector space. Then all norms on X are equivalent.

Proof. Let $\|\cdot\|$ be a norm on X and $n := \dim X$. Choosing a basis $e_1, \dots, e_n \in X$ we define a norm on $\mathbb{K}^n$ by letting $\|(x_1, \dots, x_n)\| := \|\sum_j x_j e_j\|$. Hence we may assume that $X = \mathbb{K}^n$, and it suffices to show that $\|\cdot\|$ is equivalent to the Euclidean norm $\|\cdot\|_2$. For $x = (x_j) = \sum_j x_j e_j$ we have by Cauchy-Schwarz

$$\|x\| = \|\sum_j x_j e_j\| \le \sum_j |x_j| \|e_j\| \le \underbrace{\Big(\sum_j \|e_j\|^2\Big)^{1/2}}_{=:C} \|x\|_2.$$

This implies $|\|x\| - \|y\|| \le \|x - y\| \le C\|x - y\|_2$, which means that $\|\cdot\|$ is (even Lipschitz) continuous on $(X, \|\cdot\|_2)$. Since $S := \{y \in \mathbb{K}^n : \|y\|_2 = 1\}$ is closed and bounded it is a compact subset of $(X, \|\cdot\|_2)$ and the continuous function $\|\cdot\|$ attains its minimum on S, i.e. we find $y_0 \in S$ such that $c := \|y_0\| \le \|y\|$ for all $y \in S$. If $x \in X \setminus \{0\}$ then $x/\|x\|_2\|S\|$ hence

$$c \le \|x/\|x\|_2\|, \quad \text{i.e. } c\|x\|_2 \le \|x\|.$$

□

Corollary: If X, Y are normed spaces with $\dim X < \infty$ and $T : X \to Y$ is linear then T is continuous, i.e. $T \in \mathcal{L}(X, Y)$.

Proof. Let $n := \dim X$. We find a basis $e_1, \ldots, e_n$ of X. Then $\|\sum_j x_j e_j\| := \sum_j |x_j|$ defines a new norm on X which is equivalent to the original norm. We then have, for any $x = \sum_j x_j e_j \in X$,

$$\|Tx\|_Y = \|\sum_j x_j Te_j\|_Y \leq \sum_j |x_j| \|Te_j\|_Y \leq \underbrace{\max_j \|Te_j\|_Y}_{=:C} \sum_j |x_j| = C\|x\|,$$

which proves $T \in \mathcal{L}(X, Y)$. □

1.27. Lemma (Riesz): Let X be normed space, $Y \neq X$ a closed subspace, and $\delta \in (0, 1)$. Then there exists $x_\delta \in X$ with $\|x_\delta\| = 1$ and $\|x_\delta - y\| \geq 1 - \delta$ for all $y \in Y$.

Proof. By 1.15 the quotient space X/Y is a normed space, which is nontrivial by $Y \neq X$. Hence we find $x \in X$ such that $\|[x]\|_{X/Y} = 1 - \delta$ and then $\widetilde{x} \in [x]$ with $\|\widetilde{x}\| \leq 1$. Letting $x_\delta := \widetilde{x}/\|\widetilde{x}\|$ we have $\|x_\delta\| = 1$ and, for any $y \in Y$,

$$\|x_\delta - y\| \geq \|[x_\delta]\|_{X/Y} = \frac{\|[x]\|_{X/Y}}{\|\widetilde{x}\|} = \frac{1-\delta}{\|x\|} \geq 1 - \delta.$$

□

1.28. Proposition: Let X be a normed space. The following are equivalent:

(i) $\dim X < \infty$,

(ii) $B_X := \{x \in X : \|x\| \leq 1\}$ is compact,

(iii) any closed and bounded subset is compact.

Proof. We know that (i) implies (iii) and it is clear that (iii) implies (ii). We assume that $\dim X = \infty$ and show that B_X is not compact. We find $x_1 \in X$ with $\|x_1\| = 1$ and let $Y_1 := \text{lin}\{y_1\}$ and choose $\delta = 1/2$. Then Y_1 is a closed subspace of X (since it is complete) and $Y_1 \neq X$. By 1.28 we find $x_2 \in X$ with $\|x_2\| = 1$ and $\|x_1 - x_2\| \geq 1/2$. We set $Y_2 := \text{lin}\{y_1, y_2\}$ and find, by 1.28, $x_3 \in X$ with $\|x_3\| = 1$ and $d(x_3, Y_2) \geq 1/2$, in particular we have $\|x_3 - x_j\| \geq 1/2$ for $j = 1, 2$. Continuing like this we find, for any $n \in \mathbb{N}$ with $n \geq 2$, an $x_n \in X$ with $\|x_n\| = 1$ and $\|x_n - x_j\| \geq 1/2$ for all $j = 1, \ldots, n-1$. Then (x_n) is a sequence in B_X that has no Cauchy subsequence. Hence B_X is not compact. □

end Thu 15.11.18

The following compactness criterion in $C(K)$ is fundamental for various other compactness results.

1.29. Theorem (Arzelà-Ascoli): Let K be a compact metric space and $M \subseteq C(K)$. Then M is relatively compact if and only if M is bounded and *equicontinuous*, i.e. for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in K$ with $d(x,y) < \delta$ and all $f \in M$ we have $|f(x) - f(y)| \leq \varepsilon$.

Proof. If M is relatively compact then $\overline{M}$ is compact, hence bounded. For a given $\varepsilon > 0$ we find a finite subset $E \subseteq M$ such that $M \subseteq \bigcup_{g \in E} B(g, \varepsilon/3)$. Then we find a common $\delta > 0$ for $\varepsilon/3$ according to uniform continuity of $g \in E$. Now let $x, y \in K$ with $d(x,y) < \delta$ and $f \in M$. Then we find $g \in E$ such that $\|f - g\|_\infty \leq \varepsilon/3$, and we obtain

$$|f(x) - f(y)| \leq |f(x) - g(x)| + |g(x) - g(y)| + |g(y) - f(y)| \leq 2\|f - g\|_\infty + \varepsilon/3 \leq \varepsilon,$$

so M is equicontinuous.

Now let M be bounded and equicontinuous and let (f_n) be a sequence in M. Since K is compact we find a countable dense subset $D = \{x_1, x_2, \ldots\}$ in K. Since $(f_n(x_1))$ is bounded we find a convergent subsequence $(f_{k_1(n)}(x_1))$. Since $(f_{k_1(n)}(x_2))$ is bounded we find a convergent subsequence $(f_{k_2(n)}(x_2))$ etc. Now we set $l(n) := k_n(n)$ for all $n \in \mathbb{N}$. Then $(f_{l(n)})$ is a subsequence of (f_n) and $(f_{l(n)}(y))_n$ converges for each $y \in D$. We shall show that $(f_{l(n)})$ is $\|\cdot\|_\infty$-Cauchy. So let $\varepsilon > 0$. For $\varepsilon/3$ we find $\delta > 0$ according to equicontinuity of M, and then a finite $E \subseteq K$ such that $K \subseteq \bigcup_{z \in E} B(z, \delta/2)$. For each $z \in E$ we find $y_z \in D$ with $d(y_z, z) < \delta/2$. We denote the set of all those y_z, $z \in E$, by D_E. Note that D_E is a finite subset of D and that $K \subseteq \bigcup_{y \in D_E} B(y, \delta)$. Now we find n_0 such that, for all $n, m \geq n_0$ and $y \in D_E$, we have $|f_{l(m)}(y) - f_{l(n)}(y)| \leq \varepsilon/3$. Now let $n, m \geq n_0$ and $x \in K$. We find $y \in D_E$ with $d(x,y) < \delta$ and thus obtain

$$|f_{l(n)}(x) - f_{l(m)}(x)| \leq \underbrace{|f_{l(m)}(x) - f_{l(m)}(y)|}_{\leq \varepsilon/3} + \underbrace{|f_{l(m)}(y) - f_{l(n)}(y)|}_{\leq \varepsilon/3} + \underbrace{|f_{l(n)}(y) - f_{l(n)}(x)|}_{\leq \varepsilon/3} \leq \varepsilon.$$

Hence we have shown $\|f_{l(n)} - f_{l(m)}\|_\infty \leq \varepsilon$ for all $n, m \geq n_0$, and $(f_{l(n)})$ is $\|\cdot\|_\infty$-Cauchy. Since $(C(K), \|\cdot\|_\infty)$ is complete, we have shown that M is relatively compact. □

Remark: The proof shows that it is sufficient that M is bounded pointwise, i.e. $\{f(x) : f \in M\}$ is bounded for every $x \in K$. There are versions for functions $f : K \to Y$ where Y is a metric space. Then one has to require that $\{f(x) : f \in M\}$ is a relatively compact subset of Y for every $x \in K$.

1.30. Definition: Let X, Y be normed spaces. A linear operator $T : X \to Y$ is called *compact* if $T(B_X)$ is relatively compact in Y. The set of all compact operators $X \to Y$ is denoted by $\mathcal{K}(X,Y)$.

An operator $T \in \mathcal{L}(X,Y)$ is called of *finite rank* if $\dim T(X) < \infty$[6]. The set of all finite rank operators is denoted $\mathcal{F}(X,Y)$.

[6] As for matrices, the *rank of* T is $\dim T(X)$.

Remark: (1) Any compact operator is bounded, i.e. $\mathcal{K}(X,Y) \subseteq \mathcal{L}(X,Y)$.

(2) Any $T \in \mathcal{F}(X,Y)$ is compact, i.e. $\mathcal{F}(X,Y) \subseteq \mathcal{K}(X,Y)$.

(3) $T \in \mathcal{L}(X,Y)$ is compact if and only if there exists a compact superset $K \subseteq Y$ of $T(B_X)$.

(4) If Y is a Banach space then $T \in \mathcal{K}(X,Y)$ if and only if $T(B_X)$ is totally bounded.

1.31. Proposition: Let X, Y, Z be normed spaces.

(a) $\mathcal{K}(X,Y)$ and $\mathcal{F}(X,Y)$ are linear subspaces of $\mathcal{L}(X,Y)$.

(b) If Y is a Banach space then $\mathcal{K}(X,Y)$ is closed in $\mathcal{L}(X,Y)$, in particular, $\mathcal{K}(X,Y)$ is a Banach space and any limit of a sequence in $\mathcal{F}(X,Y)$ is compact.

(c) Let $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$. If T or S is compact then ST is compact (*ideal property*).

Proof. (a) Clearly, $T \in \mathcal{K}(X,Y)$ $[\in \mathcal{F}(X,Y)]$ implies $\lambda T \in \mathcal{K}(X,Y)$ $[\in \mathcal{F}(X,Y)]$. If $S, T \in \mathcal{K}(X,Y)$ then $(S+T)(B_X) \subseteq \overline{S(B_X)} + \overline{T(B_X)}$ which is compact. If $S, T \in \mathcal{F}(X,Y)$ then $(S+T)(X) \subseteq S(X) + T(X)$ which has finite dimension.

(b) We show that $\mathcal{L}(X,Y) \setminus \mathcal{K}(X,Y)$ is open. So let $T \in \mathcal{L}(X,Y)$ be not compact. Then $T(B_X)$ is not totally bounded and we find $\varepsilon > 0$ such that $T(B_X)$ cannot be covered by finitely many balls of radius ε. By the argument we used in the proof of the lemma in 1.24 we find a sequence (x_n) in B_X such that $\|Tx_k - Tx_l\| \geq \varepsilon$ for all $k \neq l$. Let $S \in \mathcal{L}(X,Y)$ with $\|S - T\| \leq \varepsilon/3$. Then, for all $k \neq l$ we have

$$\|Sx_k - Sx_l\| \geq \|Tx_k - Tx_l\| - \|(S-T)(x_k - x_l)\| \geq \varepsilon - 2\|S - T\| \geq \varepsilon - 2\varepsilon/3 = \varepsilon/3,$$

hence (Sx_n) has no Cauchy subsequence and $S(B_X)$ is not totally bounded.

(c) If T is compact then $K := \overline{T(B_X)}$ is compact, hence $S(K)$ is compact. Since $(ST)(B_X) \subseteq S(K)$, $(ST)(B_X)$ is relatively compact. If S is compact, then $\overline{S(B_Y)}$ is compact, hence also $\|T\|\overline{S(B_Y)}$ is compact. But by $T(B_X) \subseteq \|T\| B_Y$ we obtain $(ST)(B_X) \subseteq \|T\| S(B_Y)$, hence $(ST)(B_X)$ is relatively compact. □

1.32. Examples: (1) Let $p \in [1, \infty]$ and $(a_n) \in l^\infty$. Define T by $Tx = (a_n x_n)$ for $x = (x_n) \in l^p$. Then $T \in \mathcal{L}(l^p)$ with $\|T\| = \sup_n |a_n|$. We have $T \in \mathcal{K}(l^p)$ if and only if $(a_n) \in c_0$: If $(a_n) \in c_0$ define $(a_n^{(m)})_n$ for $m \in \mathbb{N}$ by $a_n^{(m)} := \begin{cases} a_n & , n \leq m \\ 0 & , n > m \end{cases}$ and T_m by $T_m x = (a_n^{(m)} x_n)_n$. Then $T_m \in \mathcal{F}(l^p)$ and $\|T - T_m\| = \sup_{n>m} |a_n|$, which tends to 0 for $m \to \infty$. We conclude that $T \in \overline{\mathcal{F}(l^p)} \subseteq \mathcal{K}(l^p)$. If $(a_n) \notin c_0$ we find $\varepsilon > 0$ such that $|a_{k(n)}| \geq \varepsilon$ for a subsequence $(k(n))$ of (n). Bu then $(e_{k(n)})_n$ is a sequence in B_{l^p} such that $(Te_{k(n)}) = (a_{k(n)} e_{k(n)})$ satisfies $\|Te_{k(n)} - Te_{k(m)}\|_p \geq \varepsilon$ for all $n \neq m$, thus $(Te_{k(n)})$ has no convergent subsequence.

(2) Let X be a Banach space. Then $I_X \in \mathcal{K}(X)$ if and only if $\dim X < \infty$ (by 1.28). end Mon 19.11.18

(3) Let $K = [a,b]$ and $k \in C(K \times K)$, and define $T : C(K) \to C(K)$ by $Tf(x) = \int_K k(x,y)f(y)\,dy$, $x \in K$. Then T is bounded:

$$|Tf(x)| \leq \int_K |k(x,y)||f(y)|\,dy \leq (b-a)\|k\|_\infty \|f\|_\infty.$$

Moreover, T is compact: Let $\varepsilon > 0$. Since $K \times K$ is compact, f is uniformly continuous by 1.25. For $\varepsilon > 0$ we thus find $\delta > 0$ such that $|k(x,y) - k(\widetilde{x},\widetilde{y})| \leq \varepsilon$ for $|x - \widetilde{x}| + |y - \widetilde{y}| \leq \delta$. Then, for $x, \widetilde{x} \in K$ with $|x - \widetilde{x}| \leq \delta$ and $f \in C(K)$ with $\|f\|_\infty \leq 1$, we have

$$|Tf(x) - Tf(\widetilde{x})| \leq \int_K |k(x,y) - k(\widetilde{x},y)||f(y)|\,d\mu \leq \varepsilon(b-a)\|f\|_\infty \leq (b-a)\varepsilon.$$

We have shown that $T(B_{C(K)})$ is equicontinuous.

(4) Let (Ω, Σ, μ) be a measure space and $k \in L^2(\Omega \times \Omega, \mu \otimes \mu)$. Here we recall that $\mu \otimes \mu(A \times B) := \mu(A)\mu(B)$ for $A, B \in \Sigma$ induces a measure on the σ-algebra $\Sigma \otimes \Sigma = \sigma\{A \times B : A, B \in \Sigma\}$ and that $(\Omega \times \Omega, \Sigma \otimes \Sigma, \mu \otimes \mu)$ is a measure space. We also quote the following.

Theorem (Fubini): Let $p \in [1, \infty)$ and $f \in L^1(\Omega \times \Omega, \mu \otimes \mu)$. Then $y \mapsto f(x,y) \in L^1(\Omega, \mu)$ for μ-a.e. $x \in \Omega$, $x \mapsto \int_\Omega f(x,y)\,d\mu(y) \in L^1(\Omega, \mu)$ and

$$\int_{\Omega \times \Omega} f(x,y)\,d(\mu \otimes \mu)(x,y) = \int_\Omega \int_\Omega f(x,y)\,d\mu(y)\,d\mu(x).$$

Theorem (Tonelli) Let $f : \Omega \times \Omega \to \mathbb{K}$ be $\mu \otimes \mu$-measurable and suppose that $\int_\Omega \int_\Omega |f(x,y)|\,d\mu(y)\,d\mu(x) < \infty$. Then $f \in L^1(\Omega \times \Omega, \mu \otimes \mu)$ and exists

$$\int_\Omega \int_\Omega f(x,y)\,d\mu(y)\,d\mu(x) = \int_{\Omega \times \Omega} f(x,y)\,d(\mu \otimes \mu)(x,y) = \int_\Omega \int_\Omega f(x,y)\,d\mu(x)\,d\mu(y).$$

For a μ-measurable function $g \geq 0$ we use here the convention that $\int_\Omega g\,d\mu = \infty$ if $g \notin L^1(\Omega, \mu)$.

Remark: Similar results hold for the products of two different measure spaces. Via 1.20 one also has L^p-versions. We shall use $p = 2$.

We claim that $T_k f(x) := \int_\Omega k(x,y)f(y)\,d\mu(y)$ defines a compact operator $T_k \in \mathcal{K}(L^2(\Omega, \mu))$. By $k \in L^2(\Omega \times \Omega, \mu \otimes \mu)$ we have $k(x,\cdot) \in L^2(\Omega, \mu)$ for μ-a.e. $x \in \Omega$. For any $f \in L^2(\Omega, \mu)$ we hence have that

$$k(x,\cdot)f \in L^1(\Omega,\mu) \quad \text{and} \quad \left|\int_\Omega k(x,y)f(y)\,d\mu\right| \leq \|k(x,\cdot)\|_2 \|f\|_2.$$

μ-a.e. $x \in \Omega$. By Fubini we have $x \mapsto \|k(x,\cdot)\|_2 \in L^2(\Omega, \mu)$ and obtain $Tf \in L^2(\Omega, \mu)$ and $\|Tf\|_2 \leq \|k\|_2 \|f\|_2$. Hence $T \in \mathcal{L}(L^2(\Omega, \mu))$. For the proof of compactness we approximate $k \in L^2(\Omega \times \Omega, \mu \otimes \mu)$ in $\|\cdot\|_2$ by a sequence of simple functions (k_n) of the form $\sum_j \xi_j 1_{A_j \times B_j}$.

By the estimate we proved the T_{k_n} tend to T_k in operator norm. So it rests to show $T_{k_n} \in \mathcal{F}(L^2(\Omega, \mu))$. But by

$$T_{1_{A \times B}} f(x) = \int_\Omega 1_A(x) 1_B(y) f(y)\, d\mu(y) = 1_A(x) \int_\Omega 1_B f\, d\mu$$

the operator $T_{1_{A \times B}}$ has rank ≤ 1, thus $T_{k_n} \in \mathcal{F}(L^2(\Omega, \mu))$ as a linear combination of such operators.

2 Baire Categories, Uniform Boundedness Principle, and Open Mapping Theorem

2.1. Theorem (Baire): Let (X, d) be a complete metric space and (Q_n) be a sequence of open and dense subsets. Then $\bigcap_n Q_n$ is dense.

Proof. It suffices to show that any open ball contains an element of $D := \bigcap_n Q_n$, so let $B(x_0, r_0)$ be an open ball in X. Then $B(x_0, r_0) \cap Q_1 \neq \emptyset$ is open and thus contains a closed ball $\overline{B}(x_1, r_1)$ where we may arrange for $r_1 \in (0, r_0/2)$. Then $B(x_1, r_1) \cap Q_2 \neq \emptyset$ is open and contains a closed ball $\overline{B}(x_2, r_2)$ with $r_2 \in (0, r_1/2)$. We thus have

$$\overline{B}(x_2, r_2) \subseteq B(x_1, r_1) \cap Q_2 \subseteq B(x_0, r_0) \cap Q_1 \cap Q_2.$$

Proceeding in this way we obtain a sequence $(\overline{B}(x_n, r_n))$ of closed balls such that $r_n \in (0, r_{n-1}/2)$, in particular $r_n < 2^{-n} r_0$, and

$$\overline{B}(x_n, r_n) \subseteq B(x_{n-1}, r_{n-1}) \cap Q_n \subseteq B(x_0, r_0) \cap \bigcap_{j=1}^{n} Q_j.$$

By the following lemma we find $x \in \bigcap_{n \in \mathbb{N}} \overline{B}(x_n, r_n) \subseteq B(x_0, r_0) \cap D$. □

Lemma: Let (A_n) be a sequence of closed subsets such that $\emptyset \neq A_{n+1} \subseteq A_n$ for all $n \in \mathbb{N}$ and $\alpha_n := \operatorname{diam} A_n := \sup\{d(x, y) : x, y \in A_n\} \to 0$ as $n \to \infty$. Then $\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$.

Proof. For each $n \in \mathbb{N}$ we find $x_n \in A_n$. Then, for $m \geq n$, we have $x_m \in A_n$, hence $d(x_m, x_n) \leq \alpha_n$. We conclude that (x_n) is Cauchy, hence has a limit $x \in X$, which satisfies $x \in \overline{A_n} = A_n$ for any n. □

2.2. Variants of Baire's Theorem: Let (X, d) be a complete metric space and (A_n) be a sequence of closed subsets such that $\operatorname{int}(A_n) = \emptyset$ for any $n \in \mathbb{N}$. Then $\operatorname{int}(\bigcup n \in \mathbb{N} A_n) = \emptyset$.

Proof. Apply 2.1 for $Q_n = X \setminus A_n$. □

Definition: Let (X, d) be a metric space and $M \subseteq X$.

(a) M is called *nowhere dense* if $\operatorname{int}(\overline{M}) = \emptyset$.

(b) M is called *of first category* or *meager* if M is the union of a sequence of nowhere dense subsets.

(c) M is called *of second category* if M is not of first category.

Any subset of a meager set is meager, and the countable union of meager sets is meager.

Corollary: In a complete metric space the complement of a set of first category is dense.

Corollary: A complete metric space is of second category in itself.

Baire's Theorem can be used for existence proofs. These are non-constructive but usually show that existence is not rare in the sense of category.

2.3. Proposition: There exist functions $f \in C([0,1])$ that are nowhere differentiable.

Proof. We show that the set of functions that are differentiable at one point $t_0 \in [0,1]$ is of first category. For any $n \in \mathbb{N}$ let

$$E_n := \{f \in C([0,1]) : \exists t_0 \in [0,1] : A(f,t_0) \le n\},$$

where $A(f,t_0) := \sup_{0<|h|\le 1} \left|\frac{f(t_0+h)-f(t_0)}{h}\right|$ (we extend f by constants to a continuous function on $[-1,2]$). If f is differentiable at t_0 then $A(f,t_0) < \infty$ and thus $f \in \bigcup_n E_n$. By Baire it suffices to show that each E_n is closed and has empty interior. So let $n \in \mathbb{N}$.

E_n is closed: Let (f_k) be a sequence in E_n and $f \in C([0,1])$ such that $\|f_k - f\|_\infty \to 0$ as $k \to \infty$. Then we find a corresponding sequence (t_k) in $[0,1]$ satisfying $A(f_k,t_k) \le n$. Since $[0,1]$ is compact we find a convergent subsequence, which we again denote (t_k), with limit $t_0 \in [0,1]$. We shall use that then $\|g_k - g\|_\infty \to 0$, where g is continuous, implies $g_k(t_k) \to g(t_0)$[7]. For fixed $h \in [-1,1] \setminus \{0\}$ we apply this to $g_k = f_k(\cdot + h) - f_k$ and $g = f(\cdot + h) - f$ and obtain

$$|f(t_0+h) - f(t_0)| = \lim_{k\to\infty} \underbrace{|f_k(t_k+h) - f_k(t_k)|}_{\le n|h|} \le n|h|.$$

Hence $A(f,t_0) \le n$ and thus $f \in E_n$.

$\operatorname{int}(E_n) = \emptyset$: Let $f \in E_n$ and $\varepsilon > 0$. We have to find $g \in B(f,\varepsilon) \setminus E_n$. We first approximate f by a C^1-functionf_δ given by $f_\delta(t) := \delta^{-1} \int_t^{t+\delta} f(s)\,ds$ for $t \in [0,1]$. Since f is uniformly continuous we have $\|f_\delta - f\|_\infty \to 0$ as $\delta \to 0+$[8]. Hence we find $\delta > 0$ such that $\|f_\delta - f\|_\infty < \varepsilon/2$. Now it suffices to find $g \in B(f_\delta, \varepsilon/2) \setminus E_n$. To this end we show a uniform bound on $A(f_\delta, t)$: For $0 < |h| \le 1$ and $t \in [0,1]$ we have

$$|f_\delta(t+h) - f_\delta(t)| = \delta^{-1}\Big|\int_{t+h}^{t+h+\delta} - \int_t^{t+\delta}\Big| = \delta^{-1}\Big|\int_{t+\delta}^{t+h+\delta} - \int_t^{t+h}\Big| \le 2|h|\delta^{-1}\|f\|_\infty,$$

hence $A(f,t) \le 2\delta^{-1}\|f\|_\infty =: M$ for each $t \in [0,1]$. Now we choose $m > M+n$ and let ϕ be a zig-zag-function with slopes $-m$ and m but $\|\phi\|_\infty < \varepsilon/2$. Then $g := f_\delta + \phi \in B(f_\delta, \varepsilon/2)$ and, for any $0 < |h| \le 1$ and $t \in [0,1]$, we have

$$|g(t+h) - g(t)| \ge |\phi(t+h) - \phi(t)| - |f_\delta(t+h) - f_\delta(t)| \ge |\phi(t+h) - \phi(t)| - M|h|,$$

[7] Just observe $|g_k(t_k) - g(t_0)| \le |g_k(t_k) - g(t_k)| + |g(t_k) - g(t_0)| \le \|g_k - g\|_\infty + |g(t_k) - g(t_0)|$.

[8] $|f_\delta(t) - f(t)| \le \delta^{-1} \int_t^{t+\delta} |f(s) - f(t)|\,ds \le \sup_{|s-t|\le\delta} |f(s) - f(t)| \to 0$ as $\delta \to 0+$.

hence $A(g,t) \geq m - M > n$ for each $t \in [0,1]$, which means $g \notin E_n$. □

2.4. Theorem (Uniform Boundedness Principle): Let X be a Banach space, Y be normed space, and $\mathfrak{T} \subseteq \mathcal{L}(X,Y)$. If $\mathfrak{T}$ is bounded pointwise, i.e. $\sup\{\|Tx\| : T \in \mathfrak{T}\} < \infty$ for any $x \in X$, then $\mathfrak{T}$ is uniformly bounded, i.e. $\sup\{\|T\| : T \in \mathfrak{T}\} < \infty$.

Proof. For any $n \in \mathbb{N}$, let $A_n := \{x \in X : \forall T \in \mathfrak{T} : \|Tx\| \leq n\}$. Then each A_n is closed and $X = \bigcup_{n\in\mathbb{N}} A_n$ by assumption. By Baire we find $n_0 \in \mathbb{N}$ such that $\operatorname{int}(A_{n_0}) \neq \emptyset$, hence also a ball $B(x_0, r) \subseteq A_{n_0}$. For $x \in X$ with $\|x\| < 1$ and $T \in \mathfrak{T}$ we then have $y := x_0 + rx \in B(x_0, r)$ and

$$\|Tx\| = \|T(y - x_0)/r)\| \leq \|Ty\| + r^{-1}\|Tx_0\| \leq n + r^{-1}\sup\{\|Tx_0\| : T \in \mathfrak{T}\} =: C < \infty,$$

which implies $\sup\{\|T\| : T \in \mathfrak{T}\} \leq C$. □

2.5. Proposition: Let X be a Banach space, Y be normed space, and (T_n) be a sequence in $\mathcal{L}(X,Y)$ such that $Tx := \lim_{n\to\infty} T_n x$ exists for every $x \in X$. Then $T \in \mathcal{L}(X,Y)$ and

$$\|T\| \leq \liminf_{n\to\infty} \|T_n\| < \infty.$$

Proof. By linearity of limits we obtain that T is linear. By $\lim_{n\to\infty} \|T_n x\| = \|Tx\|$, the sequence (T_n) is bounded pointwise, hence uniformly bounded, i.e. $\sup_n \|T_n\| =: M < \infty$. We obtain

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq \liminf_{n\to\infty} \|T_n\| \cdot \|x\| \leq M \|x\|$$

for any $x \in X$. □

2.6. Theorem (Banach-Steinhaus): Let X, Y be Banach spaces and (T_n) be a sequence in $\mathcal{L}(X,Y)$. The following are equivalent:

(i) There exists $T \in \mathcal{L}(X,Y)$ such that $Tx = \lim_{n\to\infty} T_n x$ for all $x \in X$.

(ii) $\sup_n \|T_n\| < \infty$ and there exists a dense subset $D \subseteq X$ such $(T_n x)$ is a Cauchy sequence in Y for any $x \in X$.

Proof. By 2.5, (i) implies (ii). For the converse assume that (ii) holds. We set $M := \sup_n \|T_n\|$ and let $x \in X$, $\varepsilon > 0$. We find $z \in D$ with $\|x - z\| < \varepsilon$ and by the Cauchy property of $(T_n z)$ we find n_0 such that $\|T_n z - T_m z\| \leq \varepsilon$ for all $n, m \geq n_0$. For $n, m \geq n_0$ we thus have

$$\|T_n x - T_m x\| \leq \|T_n x - T_n z\| + \|T_n z - T_m z\| + \|T_m z - T_m x\| \leq M\varepsilon + \varepsilon + M\varepsilon = (2M+1)\varepsilon.$$

We have shown that, for any $x \in X$, $(T_n x)$ is Cauchy in Y, hence convergent. Defining T by $Tx := \lim_{n\to\infty} T_n x$ we have that $T : X \to Y$ is linear and by $\|Tx\|_Y = \lim_{n\to\infty} \|T_n x\|_Y \leq M\|x\|_X$ for any $x \in X$ we have $T \in \mathcal{L}(X,Y)$. □

Remark: Observe that only the proof of "(i) $\Rightarrow$ (ii)" relies on Baire's Theorem. The reverse implication is shown directly.

In the following theorem we denote by p_j the function given by $p_j(t) = t^j$ and we consider real-valued functions.

2.7. Theorem (Korovkin): Let (T_n) be a sequence on $\mathcal{L}(C([0,1],\mathbb{R}))$ such that each T_n is positive, i.e. $f \geq 0$ implies $T_n f \geq 0$. Suppose that $T_n p_j \to p_j$ for $j = 0, 1, 2$. Then $T_n f \to f$ for any $f \in C([0,1],\mathbb{R})$.

Proof. Let $f \in C[0,1]$ and $\varepsilon > 0$. Then f is uniformly continuous and we find $\delta > 0$ such that $|s - t| \leq \sqrt{\delta}$ implies $|f(s) - f(t)| \leq \varepsilon$. Letting $\alpha := 2\|f\|_\infty/\delta$ we then have

$$|f(s) - f(t)| \leq \varepsilon + \alpha(t-s)^2.$$

Indeed, the left hand side is $\leq \varepsilon$ if $|s-t| \leq \sqrt{\delta}$, and for $|s - t| > \sqrt{\delta}$ we have

$$|f(s) - f(t)| \leq |f(s)| + f(t)| \leq 2\|f\|_\infty \leq \alpha(t-s)^2.$$

For fixed $t \in [0,1]$ we let $g_t(s) := (t-s)^2$. Then we have

$$-\varepsilon - \alpha g_t \leq f - f(t) \leq \varepsilon + \alpha g_t, \qquad \text{for all } t \in [0,1].$$

By positivity of T_n this yields

$$-\varepsilon T_n p_0 - \alpha T_n g_t \leq T_n f - f(t) T_n p_0 \leq \varepsilon T_n p_0 + \alpha T_n g_t \quad \text{for all } t \in [0,1].$$

This means, for any $n \in \mathbb{N}$,

$$|T_n f - f(t) T_n p_0| \leq \varepsilon T_n p_0 + \alpha T_n g_t \quad \text{for all } t \in [0,1],$$

and in particular

$$|T_n f(t) - f(t) T_n p_0(t)| \leq \varepsilon T_n p_0(t) + \alpha T_n g_t(t) \quad \text{for all } t \in [0,1].$$

By $g_t = t^2 p_0 - 2t p_1 + p_2$ we have end Mon 26.11.18

$$|T_n g_t(t)| \leq \|T_n g_t - g_t\|_\infty \leq \|T_n p_0 - p_0\|_\infty + 2\|T_n p_1 - p_1\|_\infty + \|T_n p_2 - p_2\|_\infty.$$

Hence $\sup_t |T_n g_t(t)| \to 0$ as $n \to \infty$, and we obtain

$$\begin{aligned}\limsup_n \|T_n f - f\|_\infty &\leq \limsup_n \|T_n f - f T_n p_0\|_\infty + \limsup_n \|f(T_n p_0 - p_0)\|_\infty \\ &\leq \varepsilon \lim_n \|T_n p_0\|_\infty + \alpha \limsup_n \sup_t |T_n g_t(t)| + \|f\|_\infty \lim_n \|T_n p_0 - p_0\|_\infty = \varepsilon,\end{aligned}$$

which proves the claim. $\square$

2.8. Theorem (Weierstraß): The set of all polynomial functions is dense in $C([a,b],\mathbb{R})$.

Proof. We may restrict to $[a,b]=[0,1]$. We shall apply 2.7 to T_n given by

$$T_nf(t)=\sum_{j=0}^{n}\binom{n}{j}t^j(1-t)^{n-j}f(j/n).$$

Clearly, T_nf is a polynomial of degree $\leq n$ (the n-th *Bernstein polynomial* of f) and $|T_nf(t)|\leq\sum_{j=0}^{n}\binom{n}{j}t^j(1-t)^{n-j}\|f\|_\infty$, hence $\|T_n\|\leq 1$[9]. Moreover, each T_n is clearly positive and $T_np_0=p_0$,

$$\begin{aligned}T_np_1(t) &= \sum_{j=0}^{n}\binom{n}{j}t^j(1-t)^{n-j}\frac{j}{n}=\sum_{j=1}^{n}\binom{n-1}{j-1}t^j(1-t)^{n-j}\\ &= \sum_{j=0}^{n-1}\binom{n-1}{j}t^{j+1}(1-t)^{n-1-j}=t,\end{aligned}$$

i.e. $T_np_1=p_1$, as well as

$$\begin{aligned}T_np_2(t) &= \sum_{j=0}^{n}\binom{n}{j}t^j(1-t)^{n-j}\frac{j^2}{n^2}=\sum_{j=0}^{n-1}\binom{n-1}{j}t^{j+1}(1-t)^{n-1-j}\frac{j+1}{n}\\ &= \frac{t}{n}+t\frac{n-1}{n}\sum_{j=0}^{n-1}\binom{n-1}{j}t^j(1-t)^{n-1-j}\frac{j}{n-1}\\ &= \frac{t}{n}+\frac{t^2(n-1)}{n}=t^2+\frac{t(1-t)}{n},\end{aligned}$$

which implies $T_np_2\to p_2$ in $\|\cdot\|_\infty$. □

We give an application to numerical integration and consider quadrature formulae of the form

$$\int_0^1 f(t)\,dt\approx Q_nf=\sum_{j=0}^{n}\alpha_j^{(n)}f(t_j^{(n)}),$$

where, for fixed n, all $t_j^{(n)}$, $j=0,1,\ldots,n$, are assumed to be different.

2.9. Theorem (Szegö): If the Q_n are as above then the following are equivalent:

(i) $Q_nf\to\int_0^1 f(t)\,dt$ for all $f\in\mathbb{C}[0,1]$.

(ii) $Q_np\to\int_0^1 p(t)\,dt$ for all polynomials p and $\sup_n\sum_{j=0}^{n}|\alpha_j^{(n)}|<\infty$.

[9]Observe that $\sum_{j=0}^{n}t^j(1-t)^{n-j}=(t+1-t)^n=1$.

Proof. Since by 2.8 the polynomials are dense in $C[0,1]$, this follows from Banach-Steinhaus (2.6), when we have proved $\|Q_n\| = \sum_{j=0}^n |\alpha_j^{(n)}|$ for each fixed n. In order to see this we take a piecewise linear interpolating function f with $f(t_j^{(n)}) = \operatorname{sgn}(\alpha_j^{(n)})$. This has $\|f\|_\infty \le 1$ and $Q_n f = \sum_{j=0}^n |\alpha_j^{(n)}|$. □

Remark: Clearly the second condition in (ii) follows from the first in (ii) if all $\alpha_j^{(n)} \ge 0$, since then

$$\sum_{j=0}^n |\alpha_j^{(n)}| = \sum_{j=0}^n \alpha_j^{(n)} = Q_n p_0 \to 1 \quad (n \to \infty).$$

For $t_j^{(n)} = \frac{j}{n}$, $j = 0, 1, \dots, n$, the requirement $Q_n p = \int_0^1 p(t)\,dt$ determines $\alpha_j^{(n)}$, $j = 0, 1, \dots, n$, uniquely (*Newton-Cotes formulae*). Unfortunately, for large n negative $\alpha_j^{(n)}$ appear, and it is known that (i) does not hold.

2.10. Open mappings: Let X, Y be metric spaces. A mapping $f : X \to Y$ is called *open* if $f(Q)$ is open for any open $Q \subseteq X$.

Remark: If X and Y are normed spaces then a linear operator $T : X \to Y$ is open if and only if there exists $\varepsilon > 0$ such that $B(0, \varepsilon) \subseteq T(B(0,1))$. "Only if" is clear. For "if" let $Q \subseteq X$ be open and let $x \in Q$. We find $r > 0$ such that $B(x, r) \subset Q$. Then $T(B(x,r)) \subseteq T(Q)$ and, by linearity, $T(B(x,r)) = Tx + rT(B(0,1)) \supseteq Tx + rB(0,\varepsilon) = B(Tx, r\varepsilon)$, hence $B(Tx, r\varepsilon) \subseteq T(Q)$. Hence $T(Q)$ is open.

Example: If Y is a closed linear subspace of the normed space X and $q : X \to X/Y$ is the quotient map then $q(B(0,1)) = B(0,1)$, in particular, q is open.

2.11. The Open Mapping Theorem: Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$ be surjective. Then T is open.

Proof. We have $Y = \bigcup_{n \in \mathbb{N}} T(B(0,n))$. by Baire we find n_0 and an open ball $B(y_0, r) \subseteq \overline{T(B(0,n_0))}$. Then also $B(-y_0, r) = -B(y_0, r) \subseteq \overline{-T(B(0,n_0))} = \overline{T(B(0,n_0))}$. Moreover, $T(B(0,n_0))$ is convex, hence also its closure. For $y \in B(0,r)$ we thus have

$$y = \frac{1}{2}(y_0 + y) + \frac{1}{2}(-y_0 + y) \in \overline{T(B(0,n_0))},$$

hence $B(0,r) \subseteq \overline{T(B(0,n_0))}$ and by linearity $B(0, r/n_0) \subseteq \overline{T(B(0,1))}$.

Now let $\varepsilon_0 := r/n_0$. We claim that $B(0,\varepsilon_0) \subseteq T(B(0,1))$. So let $\|y\| < \varepsilon_0$. Choose $\varepsilon \in (0, \varepsilon_0)$ such that $\|y\| < \varepsilon$ and put $\overline{y} := \frac{\varepsilon_0}{\varepsilon} y$. Then $\|\overline{y}\| < \varepsilon_0$. Hence we find $x_0 \in B(0,1)$ such that $y_0 := Tx_0$ satisfies $\|\overline{y} - y_0\| < \alpha\varepsilon_0$ where we choose $\alpha \in (0,1)$ so small that $\frac{\varepsilon}{\varepsilon_0}\frac{1}{1-\alpha} < 1$. Then $\|(\overline{y} - y_0)/\alpha\| < \varepsilon_0$ and we find $x_1 \in B(0,1)$ such that $y_1 := Tx_1$ satisfies $\|\overline{y} - y_0)/\alpha - y_1\| < \alpha\varepsilon_0$. Then $\|\overline{y} - (y_0 + \alpha y_1)\| < \alpha^2\varepsilon_0$ and $\|(\overline{y} - y_0)/\alpha^2 - y_1/\alpha\| < \varepsilon_0$.

Hence we find $x_2 \in B(0,1)$ such that $y_2 := Tx_2$ satisfies $\|(\overline{y} - y_0)/\alpha^2 - y_1/\alpha - y_2\| < \alpha\varepsilon_0$. Then $\|\overline{y} - (y_0 + \alpha y_1 + \alpha^2 y_2)\| < \alpha^3\varepsilon_0$ etc.

In ths way we find a sequence (x_n) in $B(0,1)$ such that $\|\overline{y} - T(\sum_{j=0}^n \alpha^j x_j)\| < \alpha^{n+1}\varepsilon_0$ for each n. By $\alpha \in (0,1)$ the series $\sum_{j=0}^\infty \alpha^j x_j$ is absolutely convergent, hence convergent with value $\overline{x}$, we have $\|\overline{x}\| \le \sum_{j=0}^\infty \alpha^j = (1-\alpha)^{-1}$ and $T\overline{x} = \overline{y}$. Letting $x := \frac{\varepsilon}{\varepsilon_0}\overline{x}$ we thus have $\|x\| < 1$ and $Tx = y$. □

end Thu 29.11.18

2.12. Corollary: Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$.

(a) If T is bijective then the linear operator T^{-1} is bounded and T is an isomorphism.

(b) If T is injective and $Z = T(X)$, then $T^{-1} \in \mathcal{L}(Z,X)$ if and only if Z is closed in Y.

Proof. (a) By 2.10, T is open. Thus, if $Q \subseteq X$ is open, then $(T^{-1})^{-1}(Q) = T(Q)$ is open. Hence T^{-1} is continuous.

(b) follows from (a). This is part of an exercise. □

2.13. Corollary: Let $\|\cdot\|_0$ and $\|\cdot\|_1$ be norms on a $\mathbb{K}$-vector space X such that X s a Banach space for both of them. If there exists $M > 0$ such that

$$\|x\|_0 \le M\|x\|_1 \quad \text{for all } x \in X,$$

then $\|\cdot\|_0$ and $\|\cdot\|_1$ are equivalent.

Proof. Apply 2.11(a) to the bounded operator $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_0)$. □

For the Closed Graph Theorem we need a bit of notation.

2.14. Closed operators: Let X and Y be normed spaces, $D \subseteq X$ a linear subspace and let $T : D \to Y$ be a linear operator. Then the *graph of* T.

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

is a linear subspace of $X \times Y$. We call $D(T) := D$ the *domain of* T and write $T : X \supseteq D(T) \to Y$. Such an operator T is called *closed* if its graph $\mathrm{gr}(T) = \{(x, Tx) : x \in D(T)\}$ is closed in $X \times Y$.

Recall that $X \times Y$ is a normed space space for the norm given by $\|(x,y)\|_{X\times Y} = \|x\|X + \|y\|Y$, and that $X \times Y$ is a Banach space if X and Y are Banach spaces.

Obviously, an operator $T : X \supseteq D(T) \to Y$ is closed if and only if for each sequence (x_n) in $D(T)$ with $x_n \to x \in X$ and $Tx_n \to y \in Y$ we have $x \in D(T)$ and $Tx = y$.

Example: Let $X = Y = C[0,1]$ and $D = D(T) = C^1[0,1]$, $Tf := f'$. Then $T : X \supseteq D(T) \to Y$ is linear and closed: If (f_n) is a sequence in $C^1[0,1]$ such that $f_n \to f \in C[0,1]$ and $f_n' \to g \in C[0,1]$ (convergence in $\|\cdot\|_\infty$) then $f \in C^1[0,1]$ and $f' = g$.

For the Closed Graph Theorem we are only interested in the case $D = D(T) = X$. Then it is clear that closedness is a weaker property than continuity: if T is continuous then $x_n \to x$ already implies convergence of (Tx_n) with limit Tx; if T is closed we are allowed to assume that (Tx_n) converges in order to obtain convergence to Tx.

2.15. Closed Graph Theorem: Let X and Y be Banach spaces and $T : X \to Y$ be linear and closed. Then T is continuous, i.e. $T \in \mathcal{L}(X,Y)$.

Proof. As a closed linear subspace of the Banach space $X \times Y$ the graph $\mathrm{gr}(T)$ is itself a Banach space. The projections $\pi_1(x,y) := x$ and $\pi_2(x,y) := y$ are linear and continuous $X \times Y \to X$ and $X \times Y \to Y$, respectively. We consider the restriction $S := \pi_1|_{\mathrm{gr}(T)}$ of π_1 to $\mathrm{gr}(T)$. Then $S : \mathrm{gr}(T) \to X$, $(x, Tx) \mapsto x$, is linear, bijective and continuous. By 2.10, the inverse $S^{-1} : X \to \mathrm{gr}(T)$ is continuous. Hence also $T = \pi_2 S^{-1} : X \to Y$ is continuous. □

We give a short application.

2.16. Lemma: Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. Let $Z \subseteq Y$ be a linear subspace which is a Banach space for a norm $\|\cdot\|_Z$ such there exists $C > 0$ with $\|z\|_Y \le C\|z\|_Z$ for all $z \in Z$. If $T(X) \subseteq Z$ then $T \in \mathcal{L}(X,Z)$.

Proof. By 2.15 we only have to show that $T : X \to Z$ is closed. So let (x_n) be a sequence in X such that $\|x_n - x\|_X \to 0$ and $\|Tx_n - z\|_Z \to 0$ where $x \in X$ and $z \in Z$. By $T \in \mathcal{L}(X,Y)$ we obtain $\|Tx_n - Tx\|_Y \to 0$ and by $\|\cdot\|_Y \le C\|\cdot\|_Z$ we obtain $\|Tx_n - z\|_Y \to 0$. We conclude $z = Tx$ and have shown closedness of $T : X \to Z$. □

3 Hahn-Banach type theorems

3.1. Theorem (Hahn-Banach): Let X be a real vecor space and $q : X \to \mathbb{R}$ be *sublinear*, i.e. $q(\lambda x) = \lambda q(x)$ for all $\alpha \in \mathbb{R}$, $x \in X$ (q is positive homogeneous) and $q(x+y) \leq q(x) + q(y)$ for all $x, y \in X$ (q is sub-additive). Let $U \subseteq X$ be a linear subspace and $l : U \to \mathbb{R}$ be a linear functional that satisfies

$$l(u) \leq q(u) \quad \text{for all } u \in U.$$

Then there exists a linear extension $L : X \to \mathbb{R}$ of l that satisfies

$$L(x) \leq q(x) \quad \text{for all } x \in X.$$

Proof. Step 1: Assume $U \neq X$, $x_0 \in X \setminus U$, and $V = \operatorname{lin}\{x_0\} + U$. We extend l on V by letting $\widetilde{l}(\alpha x_0 + u) := \alpha r + l(u)$ for all $\alpha \in \mathbb{R}$ and $u \in U$ where $r = \widetilde{l}(x_0)$ ha to chosen suitably. We want to have $\alpha r + l(u) = \widetilde{l}(\alpha x_0 + u) \leq q(\alpha x_0 + u)$ for all $\alpha \in \mathbb{R}$, $u \in U$. For $\alpha = 0$ this is clear. For $\alpha > 0$ this holds for all $u \in U$ if and only if

$$r \leq \alpha^{-1}\big(q(\alpha x_0 + u) - l(u)\big) = q(x_0 + \alpha^{-1}u) - l(\alpha^{-1}u) \quad \text{for all } u \in U,$$

which in turn is equivalent to

$$r \leq q(x_0 + u) - l(u) \quad \text{for all } u \in U,$$

since U is a linear subspace. Similarly, we have, for any $\alpha > 0$,

$$-\alpha r + l(u) \leq q(-\alpha x_0 + u) \quad \text{for all } u \in U,$$

if and only if

$$-r \leq \alpha^{-1}\big(q(-\alpha x_0 + u) - l(u)\big) = q(-x_0 + \alpha^{-1}u) - l(\alpha^{-1}u) \quad \text{for all } u \in U,$$

which in turn is equivalent to

$$r \geq l(u) - q(-x_0 + u) \quad \text{for all } u \in U.$$

Hence, a suitable choice of r is possible if

$$l(w) - q(-x_0 + w) \leq q(x_0 + u) - l(u) \quad \text{for all } u, w \in U,$$

which is equivalent to

$$l(u + w) \leq q(x_0 + u) + q(-x_0 + w) \quad \text{for all } u, w \in U.$$

But for fixed $u, w \in U$ we have by assumption

$$l(u + w) \leq q(u + w) = q(x_0 + u + (-x_0) + w) \leq q(x_0 + u) + q(-x_0 + w).$$

Hence, by a suitable choice of r we can extend l to a linear functional $\widetilde{l} : V \to \mathbb{R}$ with the desired properties.

Step 2: Apply Zorn's Lemma: Let $\mathcal{A}$ be the set of all pairs (V, L_V) where $V \supseteq U$ is a linear subspace of X and $L_V : V \to \mathbb{R}$ is linear with $L_V|_U = l$ and $L_V(v) \leq q(v)$ for all $v \in V$. We define a partial order on $\mathcal{A}$ by

$$(V_1, L_{V_1}) \leq (V_2, L_{V_2}) \;:\Longleftrightarrow\; V_1 \subseteq V_2 \text{ and } L_{V_2}|_{V_1} = L_{V_1}.$$

Then each totally ordered subset $\mathcal{K} \subseteq \mathcal{A}$ has an upper bound in $\mathcal{A}$: If $\mathcal{K} = \{V_\lambda : \lambda \in \Lambda\}$ is a totally ordered subset of $\mathcal{A}$ then $V := \bigcup_{\lambda\in\Lambda} V_\lambda$, $L_V v := L_{V_\lambda} v$ if $v \in V_\lambda$ defines an element $(V, L_V) \in \mathcal{A}$ which is an upper bound of $\mathcal{K}$. end Mon 03.12.18

Hence, by Zorn's Lemma, $\mathcal{A}$ contains a maximal element (Y, L_Y). We have $Y = X$ since otherwise we could, by Step 1, construct a stricly larger element in $\mathcal{A}$. We conclude that $L := L_Y$ has the desired properties. $\square$

3.2. Reminder: Let M be a set. A relation $R \subseteq M \times M$ (we write $x \leq y$ instead of $(x, y) \in R$) is a *partial order* if, for all $x, y, z \in M$, the following hold

$$x \leq x; \qquad x \leq y \text{ and } y \leq x \implies x = y; \qquad x \leq y \text{ and } y \leq z \implies x \leq z.$$

A subset $K \subseteq M$ of a partially ordered set M is called *totally ordered* if, for any $x, y \in K$, we have $x \leq y$ or $y \leq x$.

An *upper bound* of a subset $K \subseteq M$ is an element $m \in M$ such that $x \leq m$ for all $x \in K$.

A *maximal element of* M is an element $m \in M$ such that, for any $x \in M$, $m \leq x$ implies $x = m$.

Zorn's Lemma: Let M be a partially ordered set such that each totally ordered subset $K \subseteq M$ has an upper bound in M. Then M contains a maximal element.

Remark: A maximal element is not necessarily unique.

We prepare the version for complex linear functionals by a lemma on the relation of $\mathbb{R}$-linear and $\mathbb{C}$-linear functionals.

3.3. Lemma: Let X be a $\mathbb{C}$-vector space.

(a) If $l : X \to \mathbb{R}$ is $\mathbb{R}$-linear and $\widetilde{l}(x) := l(x) - il(ix)$ for $x \in X$ then $\widetilde{l} : X \to \mathbb{C}$ is $\mathbb{C}$-linear and $l = \operatorname{Re} \widetilde{l}$.

(b) If $h : X \to \mathbb{C}$ is $\mathbb{C}$-linear then $l := \operatorname{Re} h$ is $\mathbb{R}$-linear and, if we contruct $\widetilde{l}$ as in (a), then $\widetilde{l} = h$.

(c) If X is a normed space (for $\mathbb{K} = \mathbb{C}$) and $h : X \to \mathbb{C}$ is $\mathbb{C}$-linear and continuous then $\|\operatorname{Re} h\| = \|h\|$.

Proof. (a) Obviously, $\widetilde{l}$ is $\mathbb{R}$-linear. For any $x \in X$ we have

$$\widetilde{l}(ix) = l(ix) - il(-x) = i\big(l(x) - il(ix)\big) = i\widetilde{l}(x).$$

Now $\mathbb{C}$-linearity follows.

(b) Obviously, l is $\mathbb{R}$-linear. Let $x \in X$. By $h(ix) = ih(x)$ we obtain $\operatorname{Re} h(ix) = -\operatorname{Im} h(x)$ which yields

$$h(x) = \operatorname{Re} h(x) + i\operatorname{Im} h(x) = \operatorname{Re} h(x) - i\operatorname{Re} h(ix) = l(x) - il(ix) = \widetilde{l}(x).$$

(c) Let $x \in X$ with $\|x\| \le 1$. By $|\operatorname{Re} h(x)| \le |h(x)|$ we obtain "$\le$". Now let $\alpha = \overline{\operatorname{sgn} h(x)}$. Then, by $\|\alpha x\| = |\alpha|\|x\| \le \|x\| \le 1$, we obtain

$$|h(x)| = \alpha h(x) = h(\alpha x) = |\operatorname{Re} h(\alpha x)| \le \|\operatorname{Re} h\|,$$

which proves "$\ge$". □

3.4. Theorem (Hahn-Banach, complex version): Let X be a $\mathbb{C}$-vector space and $q : X \to \mathbb{R}$ be sublinear. Let U be a linear subspace of X and $h : U \to \mathbb{C}$ be linear such that $\operatorname{Re} h(u) \le q(u)$ for all $u \in U$. Then there exists a linear extension $L : X \to \mathbb{C}$ of h such that $\operatorname{Re} L(x) \le q(x)$ for all $x \in X$.

Proof. Let $l := \operatorname{Re} h$, which is $\mathbb{R}$-linear. By 3.1 we obtain an $\mathbb{R}$-linear extension $F : X \to \mathbb{R}$ of l such that $F(x) \le q(x)$ for all $x \in X$. Now we apply 3.3 to obtain a $\mathbb{C}$-linear functional $L := \widetilde{F}$ such that $\operatorname{Re} L = F$. Clearly, L is an extension of $\widetilde{l} = h$ (use 3.3(b) here). □

3.5. Theorem (Hahn-Banach, norm-preserving extensions of linear functionals): Let X be normed $\mathbb{K}$-vector space and U a linear subspace. If $\phi : U \to \mathbb{K}$ is a continuous linear functional then there exists a continuous linear functional $\psi : X \to \mathbb{K}$ such that $\psi|_U = \phi$ and $\|\psi\| = \|\phi\|$.

Remark: The assertion also holds in semi-normed spaces, as the proof will show.

Proof. Case $\mathbb{K} = \mathbb{R}$: Set $q(x) := \|\phi\|\|x\|$ for $x \in X$. Then $q : X \to \mathbb{R}$ is sublinear and $\phi(u) \le q(u)$ for all $u \in U$. By 3.1 we obtain a linear extension $\psi : X \to \mathbb{R}$ of ϕ such that $\psi(x) \le q(x)$ for all $x \in X$. But then also $-\psi(x) = \psi(-x) \le q(-x) = q(x)$ for all $x \in X$, which implies $|\psi(x)| \le q(x)$ for all $x \in X$. We conclude $\|\psi\| \le \|\phi\|$. The reverse inequality is clear, since ψ is an extension of ϕ.

Case $\mathbb{K} = \mathbb{C}$: For q as above we apply 3.4 and obtain a $\mathbb{C}$-linear extension $\psi : X \to \mathbb{C}$ of ϕ satisfying $\operatorname{Re} \psi(x) \le q(x)$ for all $x \in X$. As before we obtain $|\operatorname{Re} \psi(x)| \le q(x)$ for all $x \in X$. Hence $\|\psi\| = \|\operatorname{Re} \psi\| \le \|\phi\|$ (we use 3.3(d) here). Again, $\|\psi\| \ge \|\phi\|$ is clear. □

As a consequence the dual space of a normed space is sufficiently “rich”.

3.6. Corollaries: Let X be a normed space and $x \in X$.

(a) If $x \neq 0$ then there exists $x' \in X'$ with $\|x'\| = 1$ and $x'(x) = \|x\|$. In particular, X' *separates the points of* X, i.e. for $x_1, x_2 \in X$, $x_1 \neq x_2$ there exists $x' \in X'$ with $x'(x_1) \neq x'(x_2)$.

(b) We have $\|x\| = \sup\{|x'(x)| : x' \in B_{X'}\} = \sup\{|x'(x)| : x' \in X', \|x'\| \leq 1\}$.

(c) If U is a closed linear subspace of X and $x \notin U$ then there exists $x' \in X'$ such that

$$x'|_U = 0 \quad \text{and} \quad x'(x) \neq 0.$$

(d) If U is a linear subspace of X then the following are equivalent
  (i) U is dense in X.
  (ii) Each $x' \in X'$ that vanishes on U vanishes on X.

Proof. (a) Apply 3.5 to the linear functional $u' : \operatorname{lin}\{x_0\} \to \mathbb{K}$, $u'(\alpha x_0) := \alpha\|x\|$, observe $\|u'\| = 1$. We obtain a separating functional if we apply this to $x = x_1 - x_2 \neq 0$.

(b) Follows from (a) since $|x'(x)| \leq \|x'\|\|x\| \leq \|x\|$ for $x' \in B_{X'}$ is clear.

(c) Consider the quotient map $q : X \to X/U$. We have $q(x) \neq 0$, and by (a) we find $\phi \in (X/U)'$ such that $\phi(q(x)) \neq 0$. Now set $x' := \phi \circ q \in X'$.

(d) By 1.11 any $x' \in X'$ vanishing on U vanishes on $\overline{U}$. This proves (i) $\Rightarrow$ (ii), and shows that (d) follows from (c). □

3.7. Dual spaces for closed subspaces and quotient spaces: Let X be a normed space and $U \subseteq X$, $V \subseteq X'$ be linear subspaces. We let

$$\begin{aligned} U^\perp &:= \{x' \in X' : \forall\, x \in U : x'(x) = 0\}, \\ V_\perp &:= \{x \in X : \forall\, x' \in V : x'(x) = 0\}, \end{aligned}$$

where $U^\perp$ is called the *annihilator of* U *in* X' and $V_\perp$ is called the *annihilator of* V *in* X. Clearly, $U^\perp$ and $V_\perp$ are losed linear subspaces of X' and X, respectively.

Proposition: If U is a closed linear subspace of a normed space X then there exist canonical isometric isomorphims

$$\begin{aligned} (X/U)' &\cong U^\perp, \\ U' &\cong X'/U^\perp. \end{aligned}$$

Proof. Will be an exercise. □

3.8. Proposition: A normed space with a separable dual space is separable.

Proof. Let X be normed and X' be separable. Then $S_{X'} = \{x' \in X' : \|x'\| = 1\}$ is separable, hene contains a dense sequence (x'_n). For each $n \in \mathbb{N}$ we find an $x_n \in S_X$ such that $|x'_n(x_n)| \geq 1/2$. We set $U := \text{lin}\{x_n : n \in \mathbb{N}\}$ and claim that U is dense in X. We use 3.6(d) and let $x' \in X'$ vanish on U. If $x' \neq 0$ we may assume $\|x'\| = 1$ and find $n \in \mathbb{N}$ such that $\|x'_n - x'\| \leq 1/4$. Then we obtain

$$\frac{1}{2} \leq |x'_n(x_n)| = |x'_n(x_n) - x'(x_n)| \leq \|x'_n - x'\|\|x_n\| \leq \frac{1}{4},$$

a contradiction. We conclude $x' = 0$ as desired. □

end Thu 06.12.18

Remark: In particular, we see that the dual space of l^∞ (not separable) is not isomorphic to l^1 (separable).

We give another application to a representation of the dual space of $C[a, b]$.

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

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

Clearly, $BV[a, b]$ is a $\mathbb{K}$-vector space and $\|\cdot\|_{BV}$ is seminorm on $BV[a, b]$.

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

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

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

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

Remarks:[10] (a) Any monotone function $g : [a, b] \to \mathbb{R}$ is of bounded variation and $\|g\|_{BV} = |g(b) - g(a)|$.

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

[10]without proof

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

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

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

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

3.10. Theorem (Dual space of $C[a,b]$): Any $\phi \in (C[a,b])'$ has a representation as a Stieltjes integral

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

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

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

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

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

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

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

converges to f as $n \to \infty$ in $(B[a,b], \|\cdot\|_\infty)$. Hence

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

which proves the desired representation. Moreover, by the estimate in 3.9 this implies

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

i.e. $\|\phi\| \le \|g\|_{BV}$. □

Remark: The $g \in BV_0[a,b]$ in 3.10 is not unique. We can obtain uniqueness if we require in addition that g is continuous from the right at avery point $t \in (a,b)$. This does not affect the integrals $\int_a^b f(t)\, dg(t)$ for continuous f, and in the situation of 3.10 it also does not affect the norm $\|g\|_{BV}$.

We shall use 3.1 to obtain results on separation of convex sets by hyperplanes. This needs some preparation.

3.11. Convex sets: Let X be a $\mathbb{K}$-vector space. A subset $K \subseteq X$ is called *convex* if, with any $x, y \in K$, K contains also the line segment $[x,y] := \{(1-\lambda)x + \lambda y : \lambda \in [0,1]\}$ between x and y.

Remark: If $K \subseteq X$ is convex and $T : X \to Y$ is linear, then also $T(K)$ is convex. If X is a normed space and $K \subseteq X$ convex then also $\overline{K}$ is convex. If $K_1, K_2 \subseteq X$ are convex then also $K_1 \pm K_2$ is convex.

A subset $K \subseteq X$ is called *absolutely convex* if, for all $x, y \in K$ and $\lambda, \mu \in \mathbb{K}$ with $|\lambda| + |\mu| \le 1$ one has $\lambda x + \mu y \in K$.

Remark: If X is a normed space then the closed unit ball $B_X := \{x \in X : \|x\| \le 1\}$ is absolutely convex.

Minkowski functional: Let X be a $\mathbb{K}$-vector space and $A \subseteq X$. The *Minkowski functional* p_A *of* A is given by

$$p_A(x) := \inf\{\lambda > 0 : x \in \lambda A\}, \quad x \in X.$$

If $\bigcup_{\lambda>0} \lambda A = X$ then $p_A < \infty$ on X.

If A is convex and $\bigcup_{\lambda>0}\lambda A = X$ then $0 \in A$ and $p_A : X \to \mathbb{R}$ is sublinear: $p_A(\alpha x) = \alpha p_A(x)$ for $\alpha > 0$ is clear. If $x \in \lambda A$ and $y \in \mu A$ where $\lambda, \mu > 0$ then

$$\frac{x+y}{\lambda+\mu} = \frac{x}{\lambda+\mu} + \frac{y}{\lambda+\mu} \in \frac{\lambda}{\lambda+\mu}A + \frac{\mu}{\lambda+\mu}A \subseteq A,$$

where we used convexity of A for the inclusion. Taking the inf over λ and μ we obtain $p_A(x+y) \leq p_A(x) + p_A(y)$.

If Q is convex and open with $0 \in Q$ then $\bigcup_{\lambda>0}\lambda Q = X$ then $p_Q(x) < 1$ if and only if $x \in Q$. end Mon 10.12.18

3.12. Lemma: Let X be a normed space, $V \subseteq X$ be open and convex and $0 \notin V$. Then there exists $x' \in X'$ such that $\operatorname{Re} x'(x) < 0$ for all $x \in V$.

Proof. Case $\mathbb{K} = \mathbb{R}$: We may assume $V \neq \emptyset$ and thus find $x_0 \in V$. We let $y_0 := x_0$ and $U := y_0 + V$. Then U is open and convex and $0 \in U$. Let $q := p_U$ denote the Minkowski functional of U. We define the linear functional $y' : \operatorname{lin}\{y_0\} \to \mathbb{R}$ by $y'(\alpha y_0) := \alpha p_U(y_0)$. Then $y'(\alpha y_0) \leq p_U(\alpha y_0)$ for all $\alpha \in \mathbb{R}$.

By 3.1 we find a linear extension $x' : X \to \mathbb{R}$ of y' such that $x'(x) \leq p_U(x)$ for all $x \in X$. As before we have $-x'(x) = x'(-x) \leq p_U(-x)$, hence $|x'(x)| \leq \max\{p_U(x), p_U(-x)\}$ for all $x \in X$. As U is open we find $\varepsilon > 0$ such that $B(0, \varepsilon) \subseteq U$. But then $p_U(\pm x) \leq \varepsilon^{-1}\|\pm x\| = \varepsilon^{-1}\|x\|$ for all $x \in X$, i.e. $x' \in X'$.

Now let $x \in V$. Then $u := x + y_0 \in U$. By (the last assertion in) 3.11 we have $p_U(u) < 1 \leq p_U(y_0)$, hence

$$x'(x) = x'(u) - x'(y_0) \leq p_U(u) - p_U(y_0) < 0.$$

The case $\mathbb{K} = \mathbb{C}$ then follows by 3.3. □

3.13. Theorem (Hahn-Banach, geometric version I): Let X be a normed space, $V_1, V_2 \subseteq X$ be convex with $V_1 \cap V_2 = \emptyset$ and V_1 be open. Then there exists $x' \in X'$ such that, for all $v_1 \in V_1$ and $v_2 \in V_2$, we have $\operatorname{Re} x'(v_1) < \operatorname{Re} x'(v_2)$, i.e. V_1 and V_2 are separated by the ($\mathbb{R}$-linear) functional $\operatorname{Re} x'$.

Proof. Let $V := V_1 - V_2$. Then V is convex (by 3.11) and open (by $V = \bigcup_{v_1 \in V_1} v_1 - V_2$). by $V_1 \cap V2 = \emptyset$ we have $0 \notin V$. By 3.12 we find $x' \in X'$ such that $\operatorname{Re}(v_1 - v_2) < 0$ for all $v_1 \in V_1$, $v_2 \in V_2$. This proves the assertion. □

3.14. Theorem (Hahn-Banach, geometric version II): Let X be a normed space, let $V \subseteq X$ be closed and convex and let $x_0 \notin V$. Then there exists $x' \in X'$ such that

$$\operatorname{Re} x'(x_0) < \inf\{\operatorname{Re} x'(v) : v \in V\}.$$

Proof. We find $r > 0$ with $B(x_0, r) \cap V = \emptyset$, since V is closed. By 3.13 we find $x' \in X'$ such that $\operatorname{Re} x'(x_0 + ru) < \operatorname{Re} x'(v)$ for all $u \in B(0, 1)$, $v \in V$. We observe $\operatorname{Re} x'(x_0 + u) = \operatorname{Re} x'(x_0) + r \operatorname{Re} x'(u)$, take the sup over $u \in B(0, 1)$ and obtain $\operatorname{Re} x'(x_0) + r\|x'\| \leq \operatorname{Re} x'(v)$ for all $v \in V$, which proves the assertion. $\square$

4 Hilbert spaces

4.1. Salar products: Let X be a $\mathbb{K}$-vector space. A *scalar product* (or *inner product*) on X is a map $(\cdot|\cdot) : X \times X \to \mathbb{K}$ satisfying

(S1) $\forall x, y \in X$: $(x|y) = \overline{(y|x)}$ (*symmetry*),

(S2) $\forall x, y, z \in X$, $\alpha \in \mathbb{K}$: $(\alpha x + y|z) = \alpha(x|z) + (y|z)$ (*linearity in the first component*),

(S3) $\forall x \in X \setminus \{0\}$: $(x|x) > 0$ (*definiteness*).

The pair $(X, (\cdot|\cdot))$ is called a *pre-Hilbert space* or an *inner product space.*

Remark: A scalar product on X has the following properties

$$\forall x, y \in X, \alpha \in \mathbb{K}\colon (x|\alpha y + z) = \overline{\alpha}(x|y) + (x|z),$$
$$\forall x \in V\colon (x|0) = (0|x) = 0,$$
$$\forall x, y \in V : |(x|y)| \leq \sqrt{(x|x)}\sqrt{(y|y)} \text{ (\textit{Cauchy-Schwarz inequality})}.$$

For a proof of the Cauchy-Schwarz inequality for $y \neq 0$ we write

$$0 \leq (x - \alpha y|x - \alpha y) = (x|x) - \alpha(y|x) - \overline{\alpha}(x|y) + |\alpha|^2(y|y) = (x|x) - \frac{|(x|y)|^2}{(y|y)},$$

where $\alpha = \frac{(x|y)}{(y|y)}$. We also see that "=" holds if and only if x, y are linearly dependent.

Remark: A scalar product on X induces a norm by $\|x\| := \sqrt{(x|x)}$, $x \in X$. Definiteness is clear and homogeneity is easy. The triangle inequality is obtained by

$$\|x+y\|^2 = (x+y|x+y) = \|x\|^2 + 2\mathrm{Re}\,(x|y) + \|y\|^2 \leq \|x\|^2 + 2|(x|y)| + \|y\|^2 \leq (\|x\| + \|y\|)^2,$$

where we used (S1) and Cauchy-Schwarz.

Remark: By Cauchy-Schwarz, the scalar product is continuous for the induced norm.

Definition: An inner product space X is called a *Hilbert space* if X is complete for the norm induced by the scalar product.

Examples: (1) $\mathbb{K}^n$ is Hilbert space w.r.t. $(x|y) = \sum_{j=1}^{n} x_j \overline{y_j}$.

(2) l^2 is a Hilbert space w.r.t. $(x|y) = \sum_{j=1}^{\infty} x_j \overline{y_j}$.

(3) If (Ω, Σ, μ) is a measure space then $L^2(\Omega, \mu)$ is a Hilbert space w.r.t.

$$(f|g) = \int f\overline{g}\, d\mu.$$

4.2. Proposition: If X is an inner product space with induced norm $\|\cdot\|$ then the *parallelogramm identity*

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big), \quad x, y \in X,$$

holds. Conversely, if $(X, \|\cdot\|)$ is a normed space such that the parallelogramm identity holds, then $\|\cdot\|$ is induced by a scalar product which can be obtained by *polarization*

$$\begin{aligned}(x|y) &= \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big), \quad \mathbb{K} = \mathbb{R}, \\ (x|y) &= \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big), \quad \mathbb{K} = \mathbb{C}.\end{aligned}$$

Proof. (S3) is clear since $(x|x) = \|x\|^2$. (S1) is clear for $\mathbb{K} = \mathbb{R}$. For $\mathbb{K} = \mathbb{C}$ we observe

$$\|x \pm iy\| = \|ix \pm i^2 y\| = \|ix \mp y\| = \|y \mp ix\|,$$

which implies $(x|y) = \overline{(y|x)}$. We now show (S2) for $\mathbb{K} = \mathbb{R}$. For $x, y, z \in X$ we have

$$\begin{aligned}(x|z) + (y|z) &= \frac{1}{4}\Big(\big(\|x+z\|^2 + \|y+z\|^2\big) - \big(\|x-z\|^2 + \|y-z\|^2\big)\Big) \\ &= \frac{1}{8}\Big(\big(\|x+y+2\|^2 + \|x-y\|^2\big) - \big(\|x+y-2z\|^2 + \|x-y\|^2\big)\Big) \\ &= \frac{1}{2}\Big(\Big\|\frac{x+y}{2} + z\Big\| - \Big\|\frac{x+y}{2} - z\Big\|\Big) \\ &= 2\Big(\frac{x+y}{2}\Big|z\Big).\end{aligned}$$

For $y = 0$ we obtain $(x|z) = 2(x/2|z)$ for all $x, z \in X$, and combining these two relations yields

$$(x|z) + (y|z) = (x+y|z) \qquad \text{for all } x, y, z \in X.$$

Moreover, we conclude that $(\alpha x|z) = \alpha(x|z)$ for all $\alpha = \frac{m}{2^n}$ where $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. From the definition we see that $(\cdot|z)$ is continuous, and we thus obtain

$$(\alpha x|z) = \alpha(x|z) \qquad \text{for all } x, z \in X,\ \alpha \in \mathbb{R}.$$

So (S2) is proved for $\mathbb{K} = \mathbb{R}$. For $\mathbb{K} = \mathbb{C}$ we let $(x|z)_{\mathbb{R}} := \operatorname{Re}(x|z)$, $x, z \in X$. Then we have by definition

$$(x|z) = (x|z)_{\mathbb{R}} - i(ix|z)_{\mathbb{R}} \qquad \text{for all } x, z \in X.$$

For fixed $z \in X$ we have shown that $(\cdot|z)_{\mathbb{R}}$ is $\mathbb{R}$-linear. By 3.3 we obtain that $(\cdot|z)$ is $\mathbb{C}$-linear, i.e. (S2) holds for $\mathbb{K} = \mathbb{C}$. □

end Thu 13.12.18

Corollary: The completion of an inner product space is a Hilbert space.

4.3. Proposition: Let H be a Hilbert space and $\emptyset \neq K \subseteq H$ be a closed and convex subset. For each $x \in H$ there exists a unique $y_0 \in K$ such that

$$\|x - y_0\| = \inf\{\|x - y\| : y \in K\},$$

i.e. a unique *best approximation* in K.

Proof. For $x \in K$ there is nothing to show, so let $x \notin K$ and $\delta := \inf\{\|x-y\| : y \in K\}$. Let (y_n) be a sequence in K such that $\|x - y_n\| \to \delta$. For $n, m \in \mathbb{N}$ we then have $\frac{1}{2}(y_n + y_m) \in K$ by convexity and

$$\delta \leq \left\| \frac{y_n + y_m}{2} - x \right\| \leq \frac{1}{2}\big(\|y_n - x\| + \|y_m - x\|\big) \to \delta \quad (n, m \to \infty),$$

hence $\|(y_n + y_m)/2 - x\| \to \delta$, $(n, m \to \infty)$. By 4.2 we have

$$4\|(y_n + y_m)/2 - x\|^2 + \|y_n - y_m\|^2 = 2\big(\|y_n - x\|^2 + \|y_m - x\|^2\big),$$

and we thus obtain $\|y_n - y_m\| \to 0$ as $n, m \to \infty$, i.e. (y_n) is a Cauchy sequence. Since H is complete we find $y_0 = \lim_{n\to\infty} y_n$, and by closedness of K we have $y_0 \in K$. Clearly,

$$\|x - y_0\| = \lim_{n\to\infty} \|x - y_n\| = \delta.$$

The argument also shows uniqueness: if $y_0^* \in K$ is a best approximation of x in K then $(y_0, y_0^*, y_0, y_0^*, \ldots)$ is a sequence as (y_n) above, hence convergent. We conclude $y_0^* = y_0$. $\square$

Remark: If in 4.3 we denote $P_K x := y_0$, then $P_K : H \to H$ is a projection, i.e. $P_K^2 = P_K$.

4.4. Proposition: Let H be a Hilbert space and $\emptyset \neq K \subseteq H$ be a closed and convex subset, $x \in H$ and $y_0 \in K$. The following are equivalent:

(i) $\|x - y_0\| \leq \|x - y\|$ for all $y \in K$ (i.e. y_0 is best approximation of y_0 in K),

(ii) $\mathrm{Re}\,(x - y_0 | y - y_0) \leq 0$ for all $y \in K$.

Proof. Let (ii) hold. For any $y \in K$ we have

$$\|x - y\|^2 = \|x - y_0 + y_0 - y\|^2 = \|x - y_0\|^2 + 2\mathrm{Re}\,(x - y_0 | y_0 - y) + \|y_0 - y\|^2 \geq \|x - y_0\|^2,$$

hence (i) holds. Conversely, let (i) hold and, for fixed $y \in K$, set $y_t := (1-t)y_0 + ty = y_0 + t(y - y_0)$ where $t \in (0,1)$. Then $y_t \in K$ and

$$\|x - y_0\|^2 \leq \|x - y_t\|^2 = \|x - y_0\|^2 - 2t\mathrm{Re}\,(x - y_0 | y - y_0) + t^2\|y - y_0\|^2,$$

and hence

$$\mathrm{Re}\,(x - y_0 | y - y_0) \leq \frac{t}{2}\|y - y_0\|^2.$$

Letting $t \to 0+$ we obtain (ii). $\square$

4.5. Corollary: If $Y \subseteq H$ is a closed linear subspace of a Hilbert space H, then $(x - P_Y x|y) = 0$ for all $x \in H$, $y \in Y$,[11] the projection operator $P_Y : H \to H$ is linear and bounded and $\|P_Y\| \leq 1$. We have

$$N(P_Y) = \{x \in H : \forall y \in Y : (x|y) = 0\}$$

and, if $Y \neq \{0\}$, then $\|P_Y\| = 1$.

Proof. Let $x \in H$ and $y_0 = P_Y x$. Since Y is a linear subspace, condition (ii) in Proposition 4.4 reads $\operatorname{Re}(x - P_Y x|y) \leq 0$ for all $y \in Y$. Since $(x - P_Y x|\cdot)$ is linear this implies $(x - P_Y x|\cdot) = 0$ on Y. This condition then implies linearity of $x \mapsto P_Y x$ and the assertion on $N(P_Y)$. Finally we have, for $x \in H$,

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

which implies $\|P_Y\| \leq 1$. Since $P_Y y = y$ for $y \in Y$, we have $\|P_Y\| = 1$ if $Y \neq \{0\}$. □

In a Hilbert space, all continuous linear functionals can be represented via the scalar product.

4.6. Theorem (Fréchet-Riesz): Let H be a Hilbert space. The map $J : H \to H'$, $y \mapsto (\cdot|y)$, is an antilinear and isometric bijection. In particular, for each $\phi \in H'$ there exists a unique element $y \in H$ such that $\phi(x) = (x|y)$ for all $x \in H$ and we have $\|y\| = \|\phi\|$.

Proof. For each $y \in H$ the map $Jy : H \to \mathbb{K}$, $x \mapsto (x|y)$, is linear and, by Cauchy-Schwarz, its norm is $\leq \|y\|$. By $Jy(y) = (y|y) = \|y\|^2$ we have $\|Jy\|_{H'} = \|y\|$. Hence J is isometric. Antilinearity of J is clear, so it rests to prove surjectivity.

Let $\phi \in H' \setminus \{0\}$ where we assume $\|\phi\| = 1$. Then $Z := N(\phi)$ is a closed linear subspace of H of codimension 1. Consider the projection P_Z and let $Y := N(P_Z)$. We find $y \in Y$ such that $\phi(y) = 1$. Then, clearly, $1 = \phi(y) \leq \|y\|$ by $\|\phi\| = 1$.

On the other hand, each $x \in H$ has a unique representation $x = \alpha y + z$ where $\alpha \in \mathbb{R}$ and $z \in Z$. Then $\phi(x) = \alpha$, and by $(y|z) = 0$ we have $\|x\|^2 = |\alpha|^2\|y\|^2 + \|z\|^2$. Since $\|\phi\| = 1$ we find sequences (α_n) in $\mathbb{K}$ and (z_n) in Z such that $|\alpha_n| \to 1$ and $|\alpha_n|^2\|y\|^2 + \|z_n\|^2 = 1$, which implies $\|y\| = 1$. But then $\phi(\alpha y + z) = \alpha = \alpha(y|y) = (\alpha y + z|y)$ for all $\alpha \in \mathbb{K}$ and $z \in Z$, i.e. $\phi = Jy$. □

Remark: Let H^* denote the vector space of continuous *antilinear functionals* $H \to \mathbb{K}$. Then $\overline{J} : H \to H^*$, $x \mapsto (x|\cdot)$ is a linear and isometric bijection. Moreover, we see that H^* is a Hilbert space with scalar product $(\phi|\psi)_{H^*} = (\overline{J}^{-1}\phi|\overline{J}^{-1}\psi)_H$.

[11] More precisely, for fixed $x \in H$, $P_Y x$ is the unique element $z \in Y$ such $(x - z|y) = 0$ for all $y \in Y$.

4.7. Orthogonality: Let X be an inner product space. For $x, y \in X$ we define

$$x \perp y \;:\Longleftrightarrow\; (x|y) = 0 \qquad (x \text{ and } y \text{ are } \textit{orthogonal}).$$

Pythagoras: If $x \perp y$ then

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

For $M, N \subseteq X$ we define $M \perp N :\Longleftrightarrow \forall x \in M, y \in N$: $x \perp y$.

For $A \subseteq X$ we define $A^\perp := \{x \in X : \forall y \in A : x \perp y\}$. Clearly, $A \subseteq B$ implies $A^\perp \supseteq B^\perp$.

Remark: If $A \subseteq X$ then $A^\perp$ is a closed linear subspace of X and $A^\perp = \big(\overline{\operatorname{lin} A}\big)^\perp$.

Proof. "$\supseteq$" is clear. If $x \in A^\perp$ then x is orthogonal to linear combinations of elements of A, i.e. $x \in (\operatorname{lin} A)^\perp$. Continuity of the scalar product then proves the assertion. □

Remark: If Y is a closed linear subspace of a Hilbert space H then $(Y^\perp)^\perp = Y$, $Y^\perp = N(P_Y)$, and $I - P_Y = P_{Y^\perp}$.

end Mon 17.12.18

Proof. By 4.5 we have $P_Y x = 0$ if and only if $(x|y) = 0$ for all $y \in Y$, i.e. $N(P_Y) = Y^\perp$. Hence H is the direct sum of Y and $Y^\perp$ (see below). Applied to $Y^\perp$ in place of Y we have that H is the direct sum of $Y^\perp$ and $(Y^\perp)^\perp$. By $Y \subseteq (Y^\perp)^\perp$ we then obtain $Y = (Y^\perp)^\perp$. Finally, $I - P_Y$ is a projection with range $Y^\perp$ and $N(I - P_Y) = Y = (Y^\perp)^\perp$, and we conclude that $P_{Y^\perp} = I - P_Y$. □

Remark: We call a normed space X the (topological) *direct sum* of linear subspaces Y and Z if $Y \cap Z = \{0\}$ and $Y + Z = X$ and if the bounded linear operator $S : Y \times Z \to X$, $y + z \mapsto x$, has a continuous inverse $S^{-1} : X \to Y \times Z$. We write this as $X = Y \oplus Z$. If X is a Banach space then it is sufficient that Y and Z are closed linear subspaces with $Y \cap Z = \{0\}$ and $Y + Z = X$ (use the Open Mapping Theorem).

4.8. Orthonormal systems and Bessel's inequality: A family $(e_\lambda)_{\lambda\in\Lambda}$ in an inner product space X is called an *orthonormal system* (*ONS*) if $(e_\lambda|e_\mu) = \delta_{\lambda\mu}$ for all $\lambda, \mu \in \Lambda$.

Bessel's inequality I: If $(e_\lambda)_{\lambda\in\Lambda}$ is an ONS, $x \in X$, and $F \subseteq \Lambda$ is finite then

$$\Big\| \sum_{\lambda\in F} (x|e_\lambda)e_\lambda \Big\|^2 = \sum_{\lambda\in F} |(x|e_\lambda)|^2 \le \|x\|^2.$$

Proof. First we observe

$$\Big\| \sum_{\lambda\in F} (x|e_\lambda)e_\lambda \Big\|^2 = \Big(\sum_{\lambda\in F} (x|e_\lambda)e_\lambda \Big| \sum_{\mu\in F} (x|e_\mu)e_\mu \Big) = \sum_{\lambda,\mu\in F} (x|e_\lambda)\overline{(x|e_\mu)}(e_\lambda|e_\mu) = \sum_{\lambda\in F} |(x|e_\lambda)|^2.$$

Then we have

$$\begin{aligned} 0 &\le \Big|x - \sum_{\lambda\in F}(x|e_\lambda)e_\lambda\Big|^2 = \|x\|^2 - 2\mathrm{Re}\Big(x\Big|\sum_{\lambda\in F}(x|e_\lambda)e_\lambda\Big) + \sum_{\lambda\in F}|(x|e_\lambda)|^2 \\ &= \|x\|^2 - 2\mathrm{Re}\sum_{\lambda\in F}\overline{(x|e_\lambda)}(x|e_\lambda) + \sum_{\lambda\in F}|(x|e_\lambda)|^2 \\ &= \|x\|^2 - 2\mathrm{Re}\Big(x\Big|\sum_{\lambda\in F}(x|e_\lambda)e_\lambda\Big) + \sum_{\lambda\in F}|(x|e_\lambda)|^2 = \|x\|^2 - \sum_{\lambda\in F}|(x|e_\lambda)|^2. \end{aligned}$$

□

Bessel's inequality II: If $(e_\lambda)_{\lambda\in\Lambda}$ is an ONS in a Hilbert space H and $x \in H$ then $((x|e_\lambda)e_\lambda)_{\lambda\in\Lambda}$ is summable in H, $\big(|(x|e_\lambda)|^2\big)_{\lambda\in\Lambda}$ is summable in $\mathbb{R}$ and

$$\Big\|\sum_{\lambda\in\Lambda}(x|e_\lambda)e_\lambda\Big\|^2 = \sum_{\lambda\in\Lambda}|(x|e_\lambda)|^2 \le \|x\|^2.$$

Proof. By the first part we have

$$\alpha := \sup_{F\subseteq\Lambda\,\text{finite}}\Big\|\sum_{\lambda\in F}(x|e_\lambda)e_\lambda\Big\|^2 = \sup_{F\subseteq\Lambda\,\text{finite}}\sum_{\lambda\in F}|(x|e_\lambda)|^2 \le \|x\|^2 < \infty.$$

For a given $\varepsilon \in (0, \alpha)$ we can find a finite subset $F \subseteq \Lambda$ such that

$$\Big\|\sum_{\lambda\in F}(x|e_\lambda)e_\lambda\Big\|^2 = \sum_{\lambda\in F}|(x|e_\lambda)|^2 > \alpha - \varepsilon.$$

If $G \subseteq \Lambda$ is finite and $G \cap F = \emptyset$ then

$$\Big\|\sum_{\lambda\in G}(x|e_\lambda)e_\lambda\Big\|^2 = \sum_{\lambda\in G}|(x|e_\lambda)|^2 = \sum_{\lambda\in F\cup G}|(x|e_\lambda)|^2 - \sum_{\lambda\in F}|(x|e_\lambda)|^2 < \alpha - (\alpha - \varepsilon) = \varepsilon,$$

which proves the claim, since H is complete. □

Remark: We recall that summability of a family $(x_\lambda)_{\lambda\in\Lambda}$ with value $x = \sum_{\lambda\in\Lambda} x_\lambda$ implies that $x_\lambda \ne 0$ at most for a countable subset $\{\lambda_j\}_{j=1}^N$ of Λ (where $N \in \mathbb{N} \cup \{\infty\}$) and that

$$x = \sum_{j=1}^N x_{\lambda_j}$$

is invariant under permutation of the λ_j (unconditional convergence).

4.9. Orthonormal bases: Let H be a Hilbert space and $(e_\lambda)_{\lambda\in\Lambda}$ be an ONS in H. Then the following are equivalent:

(i) $\operatorname{lin}\{e_\lambda : \lambda \in \Lambda\}$ is dense in H.

(ii) If $x \perp e_\lambda$ for all $\lambda \in \Lambda$ then $x = 0$.

(iii) For all $x \in H$: $\sum_{\lambda\in\Lambda}(x|e_\lambda)e_\lambda = x$.

(iv) For all $x \in H$: $\sum_{\lambda\in\Lambda}|(x|e_\lambda)|^2 = \|x\|^2$.

(v) For all $x, y \in H$: $(x|y) = \sum_{\lambda\in\Lambda}(x|e_\lambda)\overline{(y|e_\lambda)}$.

If any of these properties hold the ONS is said to be *complete* or an *orthonormal basis* (*ONB*) of H.

Proof. By Hahn-Banach and 4.6, (i) and (ii) are equivalent. By 4.8 (and its proof), (iii) and (iv) are equivalent. Clearly, (v) implies (iv), and by polarization (iv) implies (v) (we also obtain summability of $((x|e_\lambda)\overline{(y|e_\lambda)})_{\lambda\in\Lambda}$ by polarization). Clearly, (iii) implies (ii). Conversely, if (ii) does not hold there exists $x \in H$ with $y := x - \sum_{\lambda\in\Lambda}(x|e_\lambda)e_\lambda \neq 0$. But then $y \perp e_\lambda$ for all $\lambda \in \Lambda$, and by $y \neq 0$ we see that (ii) does not hold. □

Proposition: Any Hilbert space has an ONB. If $(e_\lambda)_{\lambda\in\Lambda}$ is an ONB of H and

$$l^2(\Lambda, \mathbb{K}) := \{(\xi_\lambda)_{\lambda\in\Lambda} \in \mathbb{K}^\Lambda : \sum_{\lambda\in\Lambda}|\xi_\lambda|^2 < \infty\} \quad \text{with norm } \|(\xi_\lambda)\|_{l^2} := \Big(\sum_{\lambda\in\Lambda}|\xi_\lambda|^2\Big)^{\frac{1}{2}},$$

then $e_\lambda \mapsto (\delta_{\lambda\mu})_{\mu\in\Lambda}$ induces a linear and isometric bijection $H \to l^2(\Lambda, \mathbb{K})$. In particular, an infinite-dimensional Hilbert space H is separable if and only if one (and then all) ONBs are countable if and only if H is isometrically isomorphic to $l^2(\mathbb{N}, \mathbb{K})$.

Proof. The first part is an application of Zorn's Lemma. Define $\mathcal{A} := \{M \subseteq X : M \text{ is an ONS}\}$, ordered by inclusion. If $\mathcal{K}$ is a chain in $\mathcal{A}$, then $\bigcup\mathcal{K}$ is an upper bound of $\mathcal{K}$ in $\mathcal{A}$. Hence $\mathcal{A}$ contains maximal elements. Any maximal element in an ONB (cp. (ii) above). The second part is easy. □

Example: We consider the complex Hilbert space $H = L^2(0, 2\pi)$ with scalar product $(f|g) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$ and let $e_n(t) := e^{int}$ for $n \in \mathbb{Z}$, $t \in [0, 2\pi]$. Then $(e_n)_{n\in\mathbb{Z}}$ is an ONS in H. This ONS is complete, i.e. it is an ONB.

To see this we denote by $\hat{f}(n) := (f|e_n) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt$, $n \in \mathbb{Z}$, the *Fourier coefficients* of $f \in H$. We have $(\hat{f}(n))_{n\in\mathbb{Z}} \in l^2(\mathbb{Z}, \mathbb{C})$ for any $f \in H$.

1) If $(\alpha_n)_{n\in\mathbb{Z}} \in l^1(\mathbb{Z})$ the series $\sum_{n\in\mathbb{Z}}\alpha_n e_n$ converges uniformly to a continuous function $g : [0, 2\pi] \to \mathbb{C}$ with $g(0) = g(2\pi)$, and we obtain $\hat{g}(n) = \alpha_n$ for all $n \in \mathbb{Z}$.

2) If $f \in C^1[0, 2\pi]$ with $f(0) = f(2\pi)$ then $f, f' \in H$ and

$$\widehat{f'}(n) = \frac{1}{2\pi}\int_0^{2\pi} f'(t)e^{-int}\,dt = \frac{in}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt = in\hat{f}(n), \quad n \in \mathbb{Z},$$

which implies $(\hat{f}(n)) \in l^1(\mathbb{Z})$. The set D of these f is dense in H.

3) It suffices to show (iii) for all $f \in D$. Starting with an $f \in D$ as in 2) we have to show that the function g in 1) corresponding to $(\alpha_n) = (\hat{f}(n))$ equals f. This follows if we show that a continuous function h with vanishing Fourier coefficients vanishes.

Assume that $\hat{h}(n) = 0$ for all $n \in \mathbb{Z}$ and that h is real (consider real and imaginary part separately). For any $t_0 \in [0, 2\pi]$ and $\delta \in (0, \pi/2)$ the function Q given by $Q(t) = \cos(t - t_0) - \cos\delta + 1$ and all its powers Q^m, $m \in \mathbb{N}$, are in $\operatorname{lin}\{e_n : n \in \mathbb{Z}\}$. Q satisfies $Q(t) >$ for $|t - t_0| < \delta$ and $|Q(t)| < 1$ if $|t - t_0| > \delta$ (taken modulo 2π). Denoting the first interval I_1 and the second I_2 we have $(h 1_{I_2} | Q^m) \to 0$ by dominated convergence. Since $(h|Q^m) = 0$ for all $m \in \mathbb{N}$ by assumption we obtain $(h 1_{I_1} | Q^m) \to 0$. Since $Q^m \to \infty$ on I_1, h can neither be strictly positive nor strictly negative on I_1. We conclude that $h = 0$. end Thu 20.12.18

We close this chapter by stating the Lax-Milgram Lemma which is proved in the exercises.

4.10. Lemma (Lax-Milgram): Let H be a Hilbert space and $\beta : H \times H \to \mathbb{K}$ be a sesquilinear form which is *continuous*, i.e. there exists $M > 0$ such that

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

and *corecive*, i.e. there exists $\eta > 0$ such that

$$\operatorname{Re}\beta(x, x)| \geq \eta\|x\|^2 \qquad \text{for all } x \in H.$$

Then there exists a unique linear operator $B : H \to H$ with $(Bx|y) = \beta(x, y)$ for all $x, y \in H$. B is an isomorphism of H with $\|B\| \leq M$ and $\|B^{-1}\| \leq \eta^{-1}$.

5 Sobolev spaces

Let $\Omega \subseteq \mathbb{R}^n$ be open. We denote by $L^1_{\mathrm{loc}}(\Omega)$ the set of all (equivalence classes of) measurable functions $f : \Omega \to \mathbb{K}$ such that $f 1_K \in L^1(\Omega)$ for any compact subset $K \subseteq \Omega$. We denote by $C_c^\infty(\Omega)$ the set of all C^∞-functions $\varphi : \Omega \to \mathbb{K}$ whose *support*

$$\operatorname{supp} \varphi := \overline{\{x \in \Omega : \varphi(x) \neq 0\}}$$

is a compact subset of Ω. We recall that a *domain* is an open and connected subset of $\mathbb{R}^n$.

We use the usual multi-index notation for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$:

$$\partial^\alpha := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} \qquad \text{where} \quad \partial_j := \frac{\partial}{\partial x_j}, \quad j = 1, \ldots, n,$$

and $|\alpha| := \alpha_1 + \ldots + \alpha_n$.

5.1. Weak derivatives: Let $\Omega \subseteq \mathbb{R}^n$ be a domain, $f \in L^1_{\mathrm{loc}}(\Omega)$ and $\alpha \in \mathbb{N}_0^n$. Then f is said to have a *weak α-th derivative in* Ω if here exists $g \in L^1_{\mathrm{loc}}(\Omega)$ such that, for all $\varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega f \, \partial^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_\Omega g \, \varphi \, dx.$$

In this case $\partial^\alpha f := g$ (we show in 5.2 below that g is unique up to "=" a.e.).

Motivation: If $f \in C^1(\Omega)$ and $\alpha = e_j$ then integration by parts yields

$$\int_\Omega f \, \partial_j \varphi \, dx = -\int_\Omega \partial_j f \, \varphi \, dx.$$

Example: Let $\Omega = (-1, 1)$, consider f and g given by $f(x) = |x|$, $g(x) = \operatorname{sgn} x$. For $\varphi \in C_c^\infty(-1, 1)$ we have

$$\begin{aligned} \int_{-1}^1 f \varphi' \, dx &= -\int_{-1}^0 x \varphi'(x) \, dx + \int_0^1 x \varphi'(x) \, dx \\ &= -x\varphi(x)\Big|_{-1}^0 + \int_{-1}^0 \varphi \, dx + x\varphi(x)\Big|_0^1 - \int_0^1 \varphi \, dx \\ &= \int_{-1}^1 g \, \varphi \, dx. \end{aligned}$$

Hence f has the weak derivative g on $(-1, 1)$.

5.2. Mollifiers: Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$.

(a) Define $f * \varphi : \mathbb{R}^n \to \mathbb{K}$ by $f * \varphi(x) := \int f(y) \varphi(x - y) \, dy$. Then $f * \varphi \in C^\infty(\mathbb{R}^n)$ and $\partial^\alpha(f * \varphi) = f * (\partial^\alpha \varphi)$ for any $\alpha \in \mathbb{N}_0^n$.

(b) Let $p \in [1, \infty)$ and $f \in L^p(\mathbb{R}^n)$. Assume $\operatorname{supp} \varphi \subseteq B(0, 1)$ and $\int \varphi \, dx = 1$, and let $\varphi_k := k^n \varphi(k \, \cdot)$ for $k \in \mathbb{N}$. Then $f * \varphi \in L^p(\mathbb{R}^n)$ and $\|f * \varphi_k - f\|_{L^p} \to 0$ for $k \to \infty$.

The function $f * \varphi$ is called the *convolution* of f with φ.

Proof. (a) Since $\operatorname{supp}\varphi(x-\cdot) = x - \operatorname{supp}\varphi$ is compact for any $x \in \mathbb{R}^n$ the function $f*\varphi$ is well-defined. It suffices to show the formula for $\alpha = e_j$. So let $h \neq 0$. Then we have

$$\begin{aligned} &\frac{1}{h}\big(f*\varphi(x+he_j) - f*\varphi(x)\big) - f*(\partial_j\varphi)(x) \\ = \; &\int f(y)\Big(\frac{\varphi(x+he_j-y)-\varphi(x-y)}{h} - \partial_j\varphi(x-y)\Big)\,dy \end{aligned}$$

By the Fundamental Theorem of Calculus we thus obtain

$$\begin{aligned} &\Big|\frac{1}{h}\big(f*\varphi(x+he_j) - f*\varphi(x)\big) - f*(\partial_j\varphi)(x)\Big| \\ \leq \; &\int |f(y)|\frac{1}{|h|}\int_{\min\{0,h\}}^{\max\{0,h\}} |\partial_j\varphi(x+te_j-y) - \varphi(x-y)|\,dt\,dy. \end{aligned}$$

For $|h| < 1$ the integral w.r.t. y is only over $K := x - \operatorname{supp}\varphi + \overline{B}(0,1)$, so this is

$$\leq \int_K |f(y)|\,dy \sup_{|\xi-\eta|\leq|h|} |\partial_j\varphi(\xi) - \partial_j\varphi(\eta)|,$$

where the integral is a finite constant and the sup tends to 0 as $h \to 0$ by uniform continuity of $\partial_j\varphi$.

(b) Observe that $\int \varphi_k\,dx = 1$ and $\|\varphi_k\|_{L^1} = \|\varphi\|_{L^1}$ for all $k \in \mathbb{N}$. A simple substitution shows $f*\varphi_k(x) = \int f(x-y)\varphi_k(y)\,dy$. We write this as

$$f*\varphi_k = \int \tau_y f\varphi_k(y)\,dy,$$

where $\tau_y f(x) := f(x-y)$. For $f \in L^p(\mathbb{R}^n)$ the map $y \mapsto \tau_y f$ is continuous $\mathbb{R}^n \to L^p(\mathbb{R}^n)$ (since $\|\tau_y\| = 1$ for all y it suffices to show this on characteristic functions of rectangles where it is clear; we use $p < \infty$ here!). Hence the integral above is a Riemann integral (or a Bochner-integral) with values in $L^p(\mathbb{R}^n)$. Then we have, by the continuous Minkowski inequality

$$\Big\|f*\varphi_k - f\Big\|_{L^p} = \Big\|\int \varphi_k(y)(\tau_y f - f)\,dy\Big\|_{L^p} \leq \int |\varphi_k(y)|\|\tau_y f - f\|_{L^p}\,dy.$$

We observe $\operatorname{supp}\varphi_k \subseteq B(0,1/k)$. Hence we have

$$\Big\|f*\varphi_k - f\Big\|_{L^p} \leq \|\varphi_k\|_{L^1} \sup_{|y|\leq 1/k} \|\tau_y f - f\|_{L^p} = \|\varphi\|_{L^1} \sup_{|y|\leq 1/k} \|\tau_y f - f\|_{L^p},$$

which proves the claim since $\lim_{y\to 0}\|\tau_y f - f\|_{L^p} = 0$. □

Corollary: If $g \in L^1_{\text{loc}}(\Omega)$ such that $\int_\Omega g\varphi\,dx = 0$ for all $\varphi \in C_c^\infty(\Omega)$ then $g = 0$ a.e..

Proof. We can resort to the case where $g \in L^1(\Omega)$ where $\Omega = B(x_0, r_0)$ is an open ball. Now let $r \in (0, r_0)$ and $g_r := g1_{B(x_0,r)}$. Take a mollifier φ as in 5.2(b). Then we have, for $x \in B(x_0, r - 1/k)$:

$$0 = g * \varphi_k(x) = g_r * \varphi_k(x),$$

and, by 5.2, $g_r * \varphi_k \to g_r$ in $\|\cdot\|_{L^1}$. We conclude that $g_r = 0$ a.e. on $B(x_0, r)$, hence also on $\Omega = B(x_0, r_0)$. □

5.3. Sobolev spaces: Let $\Omega \subseteq \mathbb{R}^n$ be a domain, $p \in [1, \infty]$, and $m \in \mathbb{N}$. Then $W^{m,p}(\Omega)$ is defined as the space of all (equivalence classes) of functions $f \in L^p(\Omega)$ such that, for all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq m$, f has a weak α-th derivative $\partial^\alpha f \in L^p(\Omega)$. For $p = 2$ we define $H^m(\Omega) := W^{m,2}(\Omega)$.

Proposition: $W^{m,p}(\Omega)$ is a Banach space for the norm given by

$$\|f\|_{W^{m,p}(\Omega)} := \Big(\sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{L^p(\Omega)} \Big)^{1/p}$$

(for $p = \infty$ we use $\|f\|_{W^{m,\infty}(\Omega)} := \max_{|\alpha| \leq m} \|\partial^\alpha f\|_{L^\infty(\Omega)}$), and $H^m(\Omega)$ is a Hilbert space with respect to the scalar product given by

$$(f|g)_{H^m(\Omega)} := \sum_{|\alpha| \leq m} (\partial^\alpha f | \partial^\alpha g)_{L^2(\Omega)}.$$

end Mon 08.01.19

Proof. Let (f_k) be a $\|\cdot\|_{W^{m,p}}$-Cauchy sequence. Then, for every $|\alpha| \leq m$, $(\partial^\alpha f_k)$ is a $\|\cdot\|_{L^p}$-Cauchy sequence, converging w.r.t. $\|\cdot\|_{L^p}$ to some $g_\alpha \in L^p(\Omega)$. Let $f := g_0$. For any $|\alpha| \leq m$ and $\varphi \in C_c^\infty(\Omega)$ we then have

$$\int_\Omega f_k\, \partial^\alpha \varphi\, dx = (-1)^{|\alpha|} \int_\Omega \partial^\alpha f_k\, \varphi\, dx \quad \text{for each } k.$$

Since $\operatorname{supp} \varphi$ is a compact subset of Ω we obtain be Hölder that, for $k \to \infty$, the left hand side tends to $\int f\, \partial^\alpha \varphi\, dx$ whereas the right hand side tends to $\int g_\alpha\, \varphi\, dx$. Hence

$$\int_\Omega f\, \partial^\alpha \varphi\, dx = (-1)^{|\alpha|} \int_\Omega g_\alpha\, \varphi\, dx \quad \text{for all } \varphi \in C_c^\infty(\Omega),$$

i.e. $g_\alpha = \partial^\alpha f$. We conclude $f \in W^{m,p}(\Omega)$ and $\|f_k - f\|_{W^{m,p}} \to 0$ as $k \to \infty$. □

5.4. Proposition: For $p \in [1, \infty)$ and $m \in \mathbb{N}$, $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{m,p}(\mathbb{R}^n)$.

In the proof we shall use the following “product rule”.

Lemma 1: If $j \in \{1, \dots, n\}$ and $f \in L^1_{\mathrm{loc}}(\Omega)$ has a weak j-th partial derivative in Ω and $g \in C^\infty(\Omega)$ then fg has a weak j-th partial derivative in Ω and

$$\partial_j(fg) = (\partial_j f)g + f(\partial_j g).$$

Proof. Observe that $fg, (\partial_j f)g, f(\partial_j g) \in L^1_{\mathrm{loc}}(\Omega)$. For $\varphi \in C_c^\infty(\Omega)$ we then have

$$\int_\Omega fg\,\partial_j\varphi\,dx = \int_\Omega f\big(\partial_j(g\varphi) - (\partial_j g)\varphi\big)\,dx = -\int_\Omega (\partial_j f)g\,\varphi\,dx - \int_\Omega f(\partial_j g)\,\varphi\,dx,$$

which proves the lemma. □

We also need that weak derivatives commute with the convolution in 5.2. The basis is that classical derivatives commute with translation.

Lemma 2: Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ such that f has a weak α-th derivative on $\mathbb{R}^n$. Let $\varphi \in C_c^\infty(\mathbb{R}^n)$. Then $\partial^\alpha(f * \varphi) = (\partial^\alpha f) * \varphi$.

Proof. For $\rho \in C_c^\infty(\mathbb{R}^n)$ we have by Fubini

$$\begin{aligned}
\int (f*\varphi)(x)(\partial^\alpha\rho)(x)\,dx &= \int\int f(x-y)\varphi(y)\,dy\,(\partial^\alpha\rho)(x)\,dx \\
&= \int\int f(x-y)(\partial^\alpha\rho)(x)\,dx\,\varphi(y)\,dy \\
&= \int\int f(x)(\partial^\alpha\rho)(x+y)\,dx\,\varphi(y)\,dy \\
&= -\int\int (\partial^\alpha f)(x)\rho(x+y)\,dx\,\varphi(y)\,dy \\
&= -\int \big((\partial^\alpha f)*\varphi\big)(x)\rho(x)\,dx,
\end{aligned}$$

which proves the claim. □

Finally we define the *support* of L^1_{loc}-functions: If $f \in L^1_{\mathrm{loc}}(\Omega)$ then $\operatorname{supp} f$ is the complement (in Ω) of the largest open subset M of Ω for which we have $f = 0$ a.e. on M.

Proof of Proposition 5.4. We first choose a function $\psi \in C_c^\infty(\mathbb{R}^n)$ such that $\psi = 1$ on $B(0,1)$ and $\operatorname{supp}\psi \subseteq B(0,2)$ and let $\psi_k := \psi(\cdot/k)$ for $k \in \mathbb{N}$. By Lemma 1 and the Leibniz rule we obtain, for $|\alpha| \le m$,

$$\partial^\alpha(f\psi_k) = \sum_{\beta\le\alpha}\binom{\alpha}{\beta}\partial^\beta f\,\partial^{\alpha-\beta}\psi_k,$$

and

$$\begin{aligned}\|\partial^\alpha f - \partial^\alpha(f\psi_k)\|_{L^p} &\le \|(1-\psi_k)\partial^\alpha f\|_{L^p} + \sum_{\beta<\alpha}\binom{\alpha}{\beta}\|\partial^\beta f\|_{L^p}\|\partial^{\alpha-\beta}\psi_k\|_\infty \\ &\le \big(1+\|\psi\|_\infty\big)\|1_{|x|\ge k}\partial^\alpha f\|_{L^p} + \sum_{\beta<\alpha}\binom{\alpha}{\beta}\|\partial^\beta f\|_{L^p}k^{-|\alpha-\beta|}\|\partial^{\alpha-\beta}\psi\|_\infty.\end{aligned}$$

As $\partial^\alpha f \in L^p$ and $|\alpha-\beta|\ge 1$ for each term in the last sum this tends to 0 as $k\to\infty$. We conclude $f\psi_k \to f$ in $\|\cdot\|_{W^{m,p}}$ and recall that $f\psi_k = 0$ a.e. outside $B(0,2k)$. Hence it suffices to approximate functions with compact support.

So let $f\in W^{m,p}(\mathbb{R}^n)$ have compact support. Taking a mollifier φ as in 5.2(b) we have that $f*\varphi_k \in C_c^\infty(\mathbb{R}^n)$ (if $\operatorname{supp} f \subseteq B(0,r)$ then $\operatorname{supp} f*\varphi_k \subseteq B(0,r+1/k)$). By Lemma 2 and 5.2(b) we have, for each $|\alpha|\le m$,

$$\|\partial^\alpha f - \partial^\alpha(f*\varphi_k)\|_{L^p} = \|\partial^\alpha f - (\partial^\alpha f)*\varphi_k\|_{L^p} \to 0 \quad (k\to\infty).$$

We conclude that $\|f - f*\varphi_k\|_{W^{m,p}} \to 0$ as $k\to\infty$. $\square$

5.5. A Dirichlet problem for the Poisson equation: Let $\Omega\subseteq\mathbb{R}^n$ be a bounded domain and $f\in L^2(\Omega)$. Consider the problem of finding a function u such that

$$\left\{\begin{aligned} -\Delta u &= f \quad \text{in } \Omega,\\ u &= 0 \quad \text{on } \partial\Omega.\end{aligned}\right. \tag{D}$$

Multiply the first line with $\varphi\in C_c^\infty(\Omega)$ and integrate over Ω to obtain

$$-\int_\Omega (\Delta u)\varphi\,dx = \int_\Omega f\varphi\,dx,$$

and observe

$$-\int_\Omega(\Delta u)\varphi\,dx = -\sum_{j=1}^n\int(\partial_j^2 u)\varphi\,dx = \sum_{j=1}^n\int_\Omega(\partial_j u)(\partial_j\varphi)\,dx = \int_\Omega \nabla u\cdot\nabla\varphi\,dx.$$

This opens the way to the *weak formulation* of (D): Instead of $u\in H^2(\Omega)$ we only require $u\in H^1(\Omega)$ and incorporate the boundary condition into the space by defining

$$W_0^{m,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{W^{m,p}}}, \quad H_0^m(\Omega) := W_0^{m,2}(\Omega), \qquad m\in\mathbb{N}, p\in[1,\infty).$$

Writing $\overline{\varphi}$ in place of φ we arrive at:

$$\text{Find } u\in H_0^1(\Omega) \text{ such that } \quad \int_\Omega \nabla u\cdot\overline{\nabla\varphi}\,dx = \int_\Omega f\overline{\varphi}\,dx \quad \text{for all } \varphi\in C_c^\infty(\Omega).$$

We define the *Dirichlet form* $\mathfrak{a} : H^1(\Omega) \times H^1(\Omega) \to \mathbb{C}$ by $\mathfrak{a}(u, v) := \int_\Omega \nabla u \cdot \overline{\nabla v}\, dx$ which is sesquilinear and continuous (since $|\mathfrak{a}(u, v)| \leq \|u\|_{H^1}\|v\|_{H^1}$). Then the weak formulation of (D) reads

$$\text{Find } u \in H_0^1(\Omega) \text{ such that } \quad \mathfrak{a}(u, v) = (f|v)_{L^2} \quad \text{for all } v \in H_0^1(\Omega), \qquad (D_w)$$

where we have used denseness of $C_c^\infty(\Omega)$ in $H_0^1(\Omega)$ w.r.t. $\|\cdot\|_{H^1}$. The antilinear functional $v \mapsto (f|v)_{L^2}$ is continuous on $H_0^1(\Omega)$ (by Cauchy-Schwarz and $\|\cdot\|_{L^2} \leq \|\cdot\|_{H^1}$).

We now need the following.

5.6. The Poincaré inequality: Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain. For $p \in [1, \infty)$ and all $u \in W_0^{1,p}(\Omega)$ we have

$$\|u\|_{L^p} \leq \operatorname{diam}(\Omega)\|\nabla u\|_{L^p},$$

where $\operatorname{diam}(\Omega) := \sup\{|x - y| : x, y \in \Omega\}$.

Proof. We let $d := \operatorname{diam}(\Omega)$. We may assume $\Omega \subseteq \{x = (x', x_n) \in \mathbb{R}^n : 0 < x_n < d\}$ and $u \in C_c^\infty(\Omega)$. We extend u to a function on $\mathbb{R}^n$ by letting $u(x) = 0$ for $x \notin \Omega$. For $x = (x', x_n) \in \mathbb{R}^n$ with $x_n \in (0, d)$ we have by Hölder

$$|u(x', x_n)|^p = |u(x', x_n) - u(x', 0)|^p = \Big| \int_0^{x_n} \partial_n u(x', \xi)\, d\xi \Big|^p \leq d^{p-1} \int_0^d |\partial_n u(x', \xi)|^p\, d\xi.$$

Now we get by Fubini

$$\begin{aligned}
\int_\Omega |u(x)|^p\, dx &= \int_{\mathbb{R}^{n-1}} \int_0^d |u(x', x_n)|^p\, dx_n\, dx' \\
&\leq d^{p-1} \int_{\mathbb{R}^{n-1}} \int_0^d \int_0^d |\partial_n u(x', \xi)|^p\, d\xi\, dx_n\, dx' \\
&= d^p \int_{\mathbb{R}^{n-1}} \int_0^d |\partial_n u(x', \xi)|^p\, d\xi\, dx' \\
&\leq d^p \|\nabla u\|_{L^p}^p,
\end{aligned}$$

which ends the proof. □

end Thu 10.01.19

Continuation of 5.5: On $H_0^1(\Omega)$ the Dirichlet form $\mathfrak{a}$ is coercive. Indeed, by Poincaré's inequality we have for $u \in H_0^1(\Omega)$:

$$\|u\|_{H^1}^2 = \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2 \leq (d^2 + 1)\|\nabla u\|_{L^2}^2,$$

which yields

$$\operatorname{Re} \mathfrak{a}(u, u) = \int_\Omega |\nabla u|^2\, dx = \|\nabla u\|_{L^2}^2 \geq (d^2 + 1)^{-1}\|u\|_{H^1}^2.$$

By Lax-Milgram (see 4.10), applied to $H := H_0^1(\Omega)$, the operator $H \to H^*$, $u \mapsto \mathfrak{a}(u, \cdot)$, is an isomorphism. Since $(f|\cdot)_{L^2} \in H^*$ there is a exactly one $u \in H_0^1(\Omega)$ such that $\mathfrak{a}(u,v) = (f|v)_{L^2}$, i.e. we have proved

Proposition: For any $f \in L^2(\Omega)$ the weak formulation (D_w) of the Dirichlet problem has a unique solution $u \in H_0^1(\Omega)$.

Remark: We studied here a special case of more general *elliptic equations*. In this case we may also have used Fréchet-Riesz (Theorem 4.6) instead of Lax-Milgram: By Poincaré's inequality $\mathfrak{a}$ is a scalar product on $H_0^1(\Omega)$ which is *equivalent* to the scalar product $(\cdot|\cdot)_{H^1}$ (i.e., the corresponding norms are equivalent). Hence we can apply 4.6 directly to $H := (H_0^1(\Omega), \mathfrak{a})$, since $(f|\cdot)_{L^2} \in H^*$ also for this H.

We denote the solution operator $R : L^2(\Omega) \to H_0^1(\Omega)$, $f \mapsto u$, by $R =: (-\Delta_D)^{-1}$. R is continuous $L^2(\Omega) \to H_0^1(\Omega)$ and $L^2(\Omega) \to L^2(\Omega)$. Moreover, for $f, g \in L^2(\Omega)$ and $u := Rf$, $v := Rg$ we have $u, v \in H_0^1(\Omega)$ and

$$(Rf|g)_{L^2} = \overline{(g|u)} = \overline{\mathfrak{a}(v,u)} = \mathfrak{a}(u,v) = (f|Rg)_{L^2}.$$

The question whether $u = Rf \in H^2(\Omega) \cap H_0^1(\Omega)$ for all $f \in L^2(\Omega)$ is one of *elliptic regularity* and cannot be settled by abstract functional analytic arguments. The answer is "no" in general but "yes" under assumptions on Ω, e.g., $\partial\Omega \in C^2$.

Remark: More general than right hand sides $f \in L^2(\Omega)$ in (D) or $(f|\cdot)_{L^2}$ in (D_w) the given arguments allow to obtain unique solutions $u \in H_0^1(\Omega)$ for right hand sides $\phi \in (H_0^1(\Omega))^* = H^*$.

Example: For $F \in L^2(\Omega)^n$ define $\operatorname{div} F \in (H_0^1(\Omega))^*$ by

$$(\operatorname{div} F)(v) := \int_\Omega F \cdot \overline{\nabla v}\, dx \qquad \text{for all } v \in H_0^1(\Omega).$$

Motivation: For $F \in C^1(\Omega)^n$ and $\varphi \in C_c^\infty(\Omega)$ integration by parts yields

$$\int_\Omega \operatorname{div} F\, \overline{\varphi}\, dx = \sum_{j=1}^n \int_\Omega \partial_j F_j\, \overline{\varphi}\, dx = -\sum_{j=1}^n \int_\Omega F_j\, \partial_j \overline{\varphi}\, dx = -\int_\Omega F \cdot \overline{\nabla \varphi}\, dx.$$

If, in addition, $F \in L^2(\Omega)^n$ and $\operatorname{div} F \in L^2(\Omega)$ then we get by approximation

$$\int_\Omega \operatorname{div} F\, \overline{v}\, dx = \int_\Omega F \cdot \overline{\nabla v}\, dx \qquad \text{for all } v \in H_0^1(\Omega).$$

Notation: $H^{-1}(\Omega) := (H_0^1(\Omega))^*$. For $\phi \in H^{-1}(\Omega)$ and $v \in H_0^1(\Omega)$ write $\phi(v) =: \langle \phi, v \rangle_{H^{-1} \times H_0^1} = \langle \phi, v \rangle$.

Result: For each $F \in L^2(\Omega)^n$ the problem

$$\text{Find } u \in H_0^1(\Omega) \text{ such that } \mathfrak{a}(u, \cdot) = \operatorname{div} F \text{ in } H^{-1}(\Omega)$$

has a unique solution $u \in H_0^1(\Omega)$.

5.7. Theorem (Meyers-Serrin): Let $\Omega \subseteq \mathbb{R}^n$ be a domain, $p \in [1,\infty)$, and $m \in \mathbb{N}$. Then $C^\infty(\Omega) \cap W^{m,p}(\Omega)$ is dense in $W^{m,p}(\Omega)$.

Idea of proof. Find a sequence (ψ_j) in $C_c^\infty(\Omega)$ such that $u = \sum_j \psi_j u$. Take a mollifier φ as in 5.2(b) and consider $v_j := (\psi_j u) * \varphi_{k_j}$ for k_j large enough. Then put $v := \sum_j v_j$ (arrange for a bounded overlap of the supports of the v_j). □

5.8. Definition: Let $\Omega \subseteq \mathbb{R}^n$ be a domain. Ω has the *segment property* if, for each $x_0 \in \partial\Omega$, there exists an open neighborhood U and a vector $y \in \mathbb{R}^n \setminus \{0\}$ such that $x + ty \in \Omega$ for all $x \in U \cap \overline{\Omega}$ and all $t \in (0,1)$.

5.9. Theorem: Let Ω have the segment property, $p \in [1,\infty)$, and $m \in \mathbb{N}$. Then the set $\{\varphi|_\Omega : \varphi \in C_c^\infty(\mathbb{R}^n)\}$ is dense in $W^{m,p}(\Omega)$.

Sketch of proof. By localization we can work near a boundary point x_0 (taking y and U as in the definition above), and it suffices to approximate $u \in W^{m,p}(\Omega)$ with $u = 0$ a.e. outside V where V is an open and bounded set with $x_0 \in V \subseteq \overline{V} \subseteq U$. Define $u(x) := 0$ for $x \notin \Omega$ and, for $t \in (0,1)$ the function u_t on $\mathbb{R}^n$ by $u_t(x) := u(x + ty)$. This means that we translate u in direction $-y$. For $\Gamma := \partial\Omega \cap \overline{V}$ we have $\Gamma - ty \cap \overline{\Omega} = \emptyset$: if z was in the intersection then by the segment property $z + ty \in \Gamma \cap \Omega = \emptyset$.

By continuity of translation we have $u_t \to u$ in $\|\cdot\|_{W^{m,p}(\Omega)}$ as $t \to 0+$, so it suffices to approximate u_t for fixed $t \in (0,1)$. Since $\Gamma - ty$ has a positive distance from $\overline{\Omega}$ we can multiply u_t with a smooth cutoff ψ such that $u_t\psi \in W^{m,p}(\mathbb{R}^n)$. Then we approximate $u_t\psi$ in $W^{m,p}(\mathbb{R}^n)$ with a $C_c^\infty(\mathbb{R}^n)$-function. □

5.10. Theorem (Rellich): Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain and $p \in [1,\infty)$. Then the embedding $W_0^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is compact.

end Mon 14.01.19

Proof. The basic idea is to apply Arzelà-Ascoli for suitable regularizations and then use the following.

Lemma: Let $u \in W^{1,p}(\mathbb{R}^n)$. Then

$$\|\tau_y u - u\|_{L^p} \le |y| \|\nabla u\|_{L^p} \qquad \text{for all } y \in \mathbb{R}^n.$$

If φ is a mollifier as in 5.2(b) then

$$\|u * \varphi_k - u\|_{L^p} \le \frac{1}{k} \|\varphi\|_{L^1} \|\nabla u\|_{L^p} \qquad \text{for all } k \in \mathbb{N}.$$

Proof. It suffices to show the estimates for $u \in C_c^\infty(\mathbb{R})$. For such u and $x, y \in \mathbb{R}^n$ we have

$$|u(x+y) - u(x)| \leq \int_0^1 |(\nabla u)(x+ty) \cdot y| \, dt,$$

hence by Hölder and Fubini

$$\|\tau_y u - u\|_{L^p}^p \leq |y|^p \int_0^1 \int |(\nabla u)(x+ty)|^p \, dy \, dt = |y|^p \int_0^1 \|\tau_{ty}(\nabla u)\|_{L^p}^p \, dt = |y|^p \, \|\nabla u\|_{L^p}^p.$$

Taking a mollifier as in 5.2(b) we have

$$\begin{aligned} \|u * \varphi_k - u\|_{L^p} &\leq \int |\varphi_k(y)| \, \|\tau_y u - u\|_{L^p} \, dy \leq \int_{|y| \leq 1/k} |y||\varphi_k(y)| \, dy \, \|\nabla u\|_{L^p} \\ &\leq \frac{1}{k} \underbrace{\|\varphi_k\|_{L^1}}_{=\|\varphi\|_{L^1}} \|\nabla u\|_{L^p}, \end{aligned}$$

which ends the proof. □

First we consider $u \in C_c^\infty(\Omega)$ with $\|u\|_{W^{1,p}} \leq 1$. We extend such u by 0 to all of $\mathbb{R}^n$ so that $u \in C_c^\infty(\mathbb{R}^n)$. We take a mollifier φ as in 5.2(b) and have, for $x, y \in \mathbb{R}^n$,

$$\begin{aligned} |u * \varphi_k(x) - u * \varphi_k(y)| &= \Big| \int_0^1 \nabla(u * \varphi_k)(y + t(x-y)) \cdot (x-y) \, dt \Big| \\ &= \Big| \int_0^1 \big((\nabla u) * \varphi_k\big)(y + t(x-y)) \cdot (x-y) \, dt \Big| \\ &\leq \|(\nabla u) * \varphi_k\|_\infty \, |x-y|. \end{aligned}$$

Moreover,

$$|(\partial_j u) * \varphi_k(x)| \leq \int |\partial_j u(y)||\varphi_k(x-y)| \, dy \leq \|\partial_j u\|_{L^p} \|\varphi_k\|_{L^q}$$

where $p^{-1} + q^{-1} = 1$. Hence

$$\|u * \varphi_k\|_\infty \leq \|u\|_{L^p} \|\varphi_k\|_{L^q},$$

and

$$|u * \varphi_k(x) - u * \varphi_k(y)| \leq \Big(\sum_{j=1}^n \|\partial_j u\|_{L^p}^2 \Big)^{1/2} \|\varphi_k\|_{L^q} \, |x-y|.$$

All functions have support contained in some large compact $K \subseteq \mathbb{R}^n$.

If we now have a sequence (u_m) in $C_c^\infty(\Omega)$ such that $\|u_m\|_{W^{1,p}} \leq 1$ we find subsequently subsequences $(l_1(m))$ of (m), $(l_2(m))$ of $(l_1(m))$ etc such that $(u_{l_k(m)} * \varphi_k)_m$ converges uniformly on K to some continuous function $f_k \in C(K)$. The diagonal sequence $(v_m) := (u_{l_m(m)})$ satisfies $v_m * \varphi_k \to f_k$ uniformly for each k, hence also

$$\|v_m * \varphi_k - f_k\|_{L^p} \to 0 \qquad (m \to \infty)$$

for each k. We now show that $(f_k)_k$ is $\|\cdot\|_{L^p}$-Cauchy. Using the lemma we have

$$\begin{aligned}
\|f_k - f_l\|_{L^p} &= \lim_{m\to\infty} \|v_m * \varphi_k - v_m * \varphi_l\|_{L^p} \\
&\leq \limsup_{m\to\infty} \big(\|v_m * \varphi_k - v_m\|_{L^p} + \|v_m - v_m * \varphi_l\|_{L^p}\big) \\
&\leq \|\varphi\|_{L^1}(k^{-1} + l^{-1}),
\end{aligned}$$

i.e. $(f_k)_k$ is $\|\cdot\|_{L^p}$-Cauchy and thus has a limit $f \in L^p(\mathbb{R}^n)$ with $\operatorname{supp} f \subset K$. We now have

$$\|f - v_m\|_{L^p} \leq \|f - f_k\|_{L^p} + \|f_k - v_m * \varphi_k\|_{L^p} + \|v_m * \varphi_k - v_m\|_{L^p}.$$

Let $\varepsilon > 0$. We find k such that $\|f_k - f\|_{\leq}\varepsilon/3$ and $\|v_m * \varphi_k - v_m\|_{\leq}\varepsilon$ for all m. Then we find m_0 such that $\|f_k - v_m * \varphi_k\|_{\leq}\varepsilon$ for all $m \geq m_0$. We conclude that $\|v_m - f\|_{\leq}\varepsilon$ for $m \geq m_0$, i.e. we have proved $v_m \to f$ in $\|\cdot\|_{L^p}$.

For a general sequence (u_m) in $W_0^{1,p}(\Omega)$ with $\|u_m\|_{W^{1,p}}$ we first find functions $\widetilde{u}_m \in C_c^\infty(\Omega)$ with $\|\widetilde{u}_m\|_{W^{1,p}} \leq 1$ and $\|u_m - \widetilde{u}_m\|_{W^{1,p}} \leq 1/m$. An $\|\cdot\|_{L^p}$-limit of a subsequence of $(\widetilde{u}_m)$ is then also a limit for the corresponding subsequence of (u_m). □

Corollary: In the situation of the Problem in 5.5 we have that $R = (-\Delta_D)^{-1} : L^2(\Omega) \to H_0^1(\Omega)$ is a bounded linear operator. By 5.10 the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact. The ideal property of compact operators then implies that $R = (-\Delta_D)^{-1}$ a compact linear operator in $L^2(\Omega)$.

In the following we only consider $\mathbb{K} = \mathbb{R}$, i.e. real-valued functions.

5.11. Proposition (Composition rule): (a) Let $F \in C^1(\mathbb{R})$ such that $F(0) = 0$ and there exists $M > 0$ with $|F'(t)| \leq M$ for all t. Then, for $p \in [1, \infty)$ and $u \in W^{1,p}(\Omega)$ we have $F \circ u \in W^{1,p}(\Omega)$ and

$$\nabla(F \circ u) = (F' \circ u)\nabla u.$$

(b) For $p \in [1, \infty)$ and $u \in W_0^{1,p}(\Omega)$ we have $u_+ := \max\{u, 0\}, u \wedge 1 := \min\{u, 1\} \in W_0^{1,p}(\Omega)$ with $\nabla(u_+) = 1_{\{u>0\}}\nabla u$ and $\nabla(u \wedge 1) = 1_{\{u<1\}}\nabla u$.

Proof. (a) If $u \in C_c^\infty(\Omega)$ we have $F \circ u \in C_c^1(\Omega) \subseteq W_0^{1,p}(\Omega)$ (here we use an exercise on approximation of C_c^1-functions by C_c^∞-functions) and $\partial_j(F \circ u) = (F' \circ u)\partial_j u$ for the classical derivative, hence also for the weak derivative. For a general $u \in W_0^{1,p}(\Omega)$ we find a sequence (u_m) in $C_c^\infty(\Omega)$ converging to u in $\|\cdot\|_{W^{1,p}}$. Resorting to a subsequence (if necessary) we may assume that $u_m \to u$ pointwise a.e., $\nabla u_m \to \nabla u$ pointwise a.e. and there exists $g \in L^p(\Omega)$ such that $|u_m|, |\partial_j u_m| \leq g$ pointwise a.e. for all m and j. Then $F \circ u_m \to F \circ u$ pointwise a.e. and $(F' \circ u_m)\partial_j u_m \to (F' \circ u)\partial_j u$ pointwise a.e.. Moreover,

$$|F \circ u_m - F \circ u| \leq M|u_m - u|,$$

which implies $F \circ u_m \to F \circ u$ in $\|\cdot\|_{L^p}$. We also have

$$|(F' \circ u_m)\partial_j u_m| \leq Mg \quad \text{a.e.},$$

hence dominated convergence implies $(F' \circ u_m)\nabla u_m \to (F' \circ u)\nabla u$ in $\|\cdot\|_{L^p}$. If $\varphi \in C_c^\infty(\Omega)$ we thus can pass to the limit in the equation

$$\int_\Omega F \circ u \, \partial_j \varphi \, dx = -\int_\Omega (F' \circ u)\partial_j u \, \varphi \, dx,$$

which proves the formula in (a). We also have $\|F \circ u_m - F \circ u\|_{W^{1,p}(\Omega)} \to 0$ which shows $F \circ u \in W_0^{1,p}(\Omega)$.

(b) Let $F(t) := \max\{t, 0\}$ and $F_\varepsilon(t) := \frac{1}{\varepsilon}\int_{t-\varepsilon}^t \max\{s, 0\}\, ds$ for $t \in \mathbb{R}$ and $\varepsilon > 0$. Then $F_\varepsilon(0) = 0$, $F_\varepsilon \in C^1(\mathbb{R})$,

$$F_\varepsilon'(t) = \frac{1}{\varepsilon}(\max\{t, 0\} - \max\{t - \varepsilon, 0\}) = \begin{cases} 0 & , \ t \leq 0 \\ t/\varepsilon & , \ t \in (0, \varepsilon) \\ 1 & , \ t \geq \varepsilon \end{cases},$$

$0 \leq F_\varepsilon'(t) \leq 1$, and $F_\varepsilon \to F$, $F_\varepsilon' \to 1_{(0,\infty)}$ pointwise as $\varepsilon \to 0+$. For $u \in W_0^{1,p}(\Omega)$ we have $F_\varepsilon \circ u \to F \circ u = u_+$ pointwise and $|F_\varepsilon \circ u| \leq |u|$ hence $F_\varepsilon \circ u \to u_+$ in $\|\cdot\|_{L^p}$ as $\varepsilon \to 0+$. Similarly, we have $(F_\varepsilon' \circ u)\nabla u \to (1_{(0,\infty)} \circ u)\nabla u$ pointwise and $|F_\varepsilon' \circ u| \leq 1_{\{u>0\}}$, which implies $(F_\varepsilon' \circ u)\nabla u \to 1_{\{u>0\}}\nabla u$ in $\|\cdot\|_{L^p}$. As in (a) we now obtain that u_+ has a weak derivative and $\nabla(u_+) = 1_{\{u>0\}}\nabla u$, and then $F_\varepsilon \circ u \to u_+$ in $\|\cdot\|_{W^{1,p}}$. By (a) we have $F_\varepsilon \circ u \in W_0^{1,p}(\Omega)$, hence also $u_+ \in W_0^{1,p}(\Omega)$.

The proof for $u \wedge 1$ is similar, where now $F(t) := \min\{t, 1\}$. We let $F_\varepsilon(t) := \frac{1}{\varepsilon}\int_t^{t+\varepsilon} \min\{s, 1\}\, ds$. Then $F_\varepsilon \in C^1(\mathbb{R})$, $F_\varepsilon(0) = 0$, $0 \leq F_\varepsilon'(t) \leq 1$, and $F\varepsilon' \to 1_{(-\infty,1)}$ pointwise. We can now argue as before. □

5.12. Applications: Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain.

(a) As in 5.5 we consider the weak formulation (D_w) of the Dirichlet problem (D) for real-valued functions and the solution operator $R = (-\Delta_D)^{-1} : L^2(\Omega) \to H_0^1(\Omega)$. Let $f \geq 0$ a.e.. We claim that $u = Rf \geq 0$ a.e., i.e. we claim that $u_- = 0$ a.e. where $u_- := -\min\{u, 0\} = \max\{-u, 0\}$. Observe that $u_- \geq 0$ a.e., that $u = u_+ - u_-$, and that $u_- \in H_0^1(\Omega)$ by 5.11(b) with $\nabla(u_-) = 1_{\{u<0\}}\nabla u$[12]. Hence we have

$$0 \leq (f|u_-)_{L^2} = \mathfrak{a}(u, u_-) = \mathfrak{a}(u_+, u_-) - \mathfrak{a}(u_-, u_-) = -\mathfrak{a}(u_-, u_-),$$

where $\mathfrak{a}(u_+, u_-) = 0$ since $\nabla(u_+) \cdot \nabla(u_-) = 0$ a.e.. By Poincaré's inequality we obtain $\|u_-\|_{L^2} = 0$, i.e. $u_- = 0$ a.e. (otherwise $\mathfrak{a}(u_-, u_-) > 0$, a contradiction). end Thu 17.01.19

(b)[13] We consider the slightly different problem

$$\begin{cases} u - \mu\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{D,μ}$$

where $\mu > 0$ in its weak formulation

$$\text{Find } u \in H_0^1(\Omega) \text{ such that } \quad \mathfrak{a}_\mu(u, v) = (f|v)_{L^2} \quad \text{for all } v \in H_0^1(\Omega), \tag{D_w,μ}$$

where $\mathfrak{a}_\mu : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is given by

$$\mathfrak{a}_\mu(u, v) := \int_\Omega \mu\nabla u \cdot \overline{\nabla v} + u\overline{v}\, dx.$$

Then $\mathfrak{a}$ is a symmetric and continuous sesquilinear form on $H_0^1(\Omega)$ and clearly

$$\min\{\mu, 1\}\|u\|_{H^1} \leq \mathfrak{a}(u, u) \leq \max\{\mu, 1\}\|u\|_{H^1} \quad \text{for all } u \in H_0^1(\Omega).$$

By Lax-Milgram (or Fréchet-Riesz) we obtain that, for any $f \in L^2(\Omega)$, the problem (D_w, μ) has a unique solution $u \in H_0^1(\Omega)$, which we denote by $R_\mu f := u$.

Now we consider $\mathbb{K} = \mathbb{R}$, i.e. real-valued functions. As in (a) we can show that $R_\mu f \geq 0$ a.e. if $f \geq 0$ a.e.. Now let $f \leq 1$ a.e.. We claim that $u = R_\mu f \leq 1$ a.e.. Since $u = u \wedge 1 + (u-1)_+$ this means that we claim $(u-1)_+ = 0$ a.e.. By 5.11(b) we have $u \wedge 1 \in H_0^1(\Omega)$ and thus also $(u-1)_+ \in H_0^1(\Omega)$ and $\nabla\big((u-1)_+\big) = 1_{\{u \geq 1\}}\nabla u$. Now we have

$$\begin{aligned} \int_\Omega (u-1)_+\, dx &\geq (f|(u-1)_+)_{L^2} = \mathfrak{a}_\mu(u, (u-1)_+) = \int_{\{u\geq 1\}} \mu|\nabla u|^2\, dx + \int_\Omega u(u-1)_+\, dx \\ &\geq \mu\|\nabla(u-1)_+\|_{L^2} + \int_\Omega (u-1)_+\, dx, \end{aligned}$$

which implies $\big(\nabla(u-1)_+\big) = 0$ a.e. and thus $(u-1)_+ = 0$ a.e. by Poincaré's inequality.

As a consequence we obtain that $R_\mu : L^\infty(\Omega) \to L^\infty(\Omega)$ with $\|R_\mu\| \leq 1$ (since Ω is bounded we ahve $L^\infty(\Omega) \subseteq L^2(\Omega)$): Let $f \in L^\infty(\Omega)$ with $\|f\|_{L^\infty} \leq 1$. Then $f = f_+ - f_-$ where $0 \leq f_\pm \leq 1$ a.e., hence $0 \leq R_\mu f_\pm \leq 1$ a.e., and we obtain

$$-1 \leq R_\mu f = R_\mu f_+ - R_\mu f_- \leq 1 \quad \text{a.e.},$$

i.e. $\|R_\mu f\|_{L^\infty} \leq 1$, and the claim is proved.

[12] Notice also that this means we have $\nabla u = 0$ a.e. on $\{u = 0\}$.

[13] Should have been part of the lecture on 17.01.19, but wasn't.

6 Dual operators, reflexive spaces, weak convergence

6.1. Dual operators: Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. The dual operator $T' \in \mathcal{L}(Y', X')$ of T is given by

$$T'\phi := \phi \circ T : X \to \mathbb{K} \qquad \text{for all } \phi \in Y'.$$

Using duality brackets this means

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

We have $\|T'\| = \|T\|$. Indeed, by Hahn-Banach,

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

Properties: The map $\mathcal{L}(X, Y) \to \mathcal{L}(Y', X')$, $T \mapsto T'$ is linear and isometric but in general not surjective. If Z is another Banach space then

$$(ST)' = T'S' \qquad \text{for all } T \in \mathcal{L}(X, Y),\ S \in \mathcal{L}(Y, Z).$$

Examples: (1) Let $p \in [1, \infty)$ and $L : l^p \to l^p$ be the *left shift* given by $L(x_n) = (x_2, x_3, x_4, \ldots)$. The dual space of l^p is l^q (where $\frac{1}{p} + \frac{1}{q} = 1$) for the duality bracket $\langle x, y\rangle = \sum_{n=1}^{\infty} x_n y_n$ for $x = (x_n) \in l^p$ and $y = (y_n) \in l^q$. We then have

$$\langle Lx, y\rangle = \sum_{n=1}^{\infty} x_{n+1} y_n = \sum_{n=2}^{\infty} x_n y_{n-1} = \langle x, Ry\rangle,$$

where $Ry = (0, y_1, y_2, y_3, \ldots)$ is the *right shift*. Hence $L' = R$.
(2) Define $T : l^1 \to l^1$ by $Tx := (\sum_n x_n, 0, 0, 0, \ldots)$. Then, for $y \in l^\infty = (l^1)'$,

$$\langle Tx, y\rangle = \Big(\sum_{n=1}^{\infty} x_n\Big) y_1 = \sum_{n=1}^{\infty} x_n y_1 = \langle x, T'y\rangle$$

hence $T'y = (y_1, y_1, y_1, \ldots)$.
(3) If $T = \phi : X \to \mathbb{K}$ then $T' : \mathbb{K} \to X'$, $\alpha \mapsto \alpha\phi$.

It is often easier to determine the kernel $N(T)$ of an operator T than its *range* $R(T)$. In the following we use the notation from 3.7.

6.2. Proposition: Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Then

$$R(T)^\perp = N(T') \qquad \text{and} \qquad \overline{R(T)} = N(T')|_\perp.$$

Proof. For $y' \in Y'$ we have

$$y' \in N(T') \;\Leftrightarrow\; T'y' = 0 \;\Leftrightarrow\; \forall x \in X : \underbrace{\langle x, T'y' \rangle}_{=\langle Tx, y' \rangle} = 0 \;\Leftrightarrow\; y' \in R(T)^{\perp}.$$

The second equality then follows from Hahn-Banach: By 3.6(d) we have $y \in \overline{R(T)}$ if and only if every $y' \in Y'$ vanishing on $R(T)$ vanishes on y, i.e. (by the first equality) $y \in \overline{R(T)}$ if and only if $y' \in N(T)$ vanishes on y. □

6.3. The bidual of a normed space: Let X be normed space. The dual $X'' := (X')'$ of the dual space X' of X is called the *bidual of* X. For any $x \in X$ the functional $\delta_x : X' \to \mathbb{K}$, $x' \mapsto \langle x, x' \rangle$, is an element of X''. The *canonical map* $j_X : X \to X''$, $x \mapsto \delta_x$, is linear and isometric.

Proof. By Hahn-Banach we have, for any $x \in X$,

$$\|x\|_X = \sup_{\|x'\|_{X'} \leq 1} |\langle x, x' \rangle| = \sup_{\|x'\|_{X'} \leq 1} |\delta_x(x')| = \|\delta_x\|_{X''}.$$

□

Corollary: If X is a normed space then $\left(\overline{j_X(X)}, j_X\right)$ is a completion of X.

Lemma: Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Then $T'' := (T')' : X'' \to Y''$ with norm $\|T''\| = \|T\|$ and $T'' \circ j_X = j_Y \circ T$.

Proof. By 6.1 we have $\|T''\| = \|T'\| = \|T\|$. Moreover, for $x \in X$ and $y' \in Y'$,

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

□

One thus can regard a Banach space X as a closed linear subspace of X'' (by identifying $j_X(X)$ and X) and T'' as an extension of T.

Examples: (1) We have $c_0'' = l^\infty$ with $j_{c_0}(x) = x$ for all $x \in c_0$[14]. In particular, $j_{c_0} : c_0 \to l^\infty$ is not surjective. For the left shift $L : c_0 \to c_0$ the operator L'' is the left shift on l^∞.

(2) The operator $T : l^1 \to l^1$ from Example 6.2(2) has $T' : l^\infty \to l^\infty$ and T' has rank 1, i.e. $\dim R(T') = 1$, with $R(T') \cap c_0 = \{0\}$. It follows that there is no operator $S : c_0 \to c_0$ such that $S' = T$: otherwise we had $T' = S'' : l^\infty \to l^\infty$ as an extension of S which imply $S = 0$ and then $T' = 0$, a contradiction.

In particular, we see that $\mathcal{L}(c_0) \to \mathcal{L}(l^1)$, $S \mapsto S'$ is not surjective.

[14] This is easy, but has to be checked!

6.4. Theorem (Schauder): Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. Then T is compact if and only if T' is compact.

Proof. "$\Rightarrow$": Let T be compact and (ψ_n) be a sequence in Y' with $\|\psi_n\|_{Y'} \leq 1$. Let $K := \overline{T(B_X)}$, which is a compact subset of Y by assumption. For all $n, m \in \mathbb{N}$ we have

$$\begin{aligned}\|T'\psi_n - T'\psi_m\|_{X'} &= \sup_{x \in B_X} |\langle x, T'(\psi_n - \psi_m)\rangle| = \sup_{x \in B_X} |\langle Tx, \psi_n - \psi_m\rangle| \\ &= \sup_{y \in K} |(\psi_n - \psi_m)(y)| = \|\psi_n|_K - \psi_m|_K\|_{C(K)}.\end{aligned}$$

Hence it suffices to find a subsequence $(k(n))$ such that $(\psi_{k(n)}|_K)$ is Cauchy in $C(K)$. The subset $M := \{\psi_n|_K : n \in \mathbb{N}\}$ is by Arzelà-Ascoli relatively compact in $C(K)$ since, for each $n \in \mathbb{N}$, $\|\psi_n|_K\|_{C(K)} = \|T'\psi_n\|_{X'} \leq \|T'\| = \|T\|$ and ψ_n is Lipschitz-continuous with constant ≤ 1, i.e. M is bounded and equicontinuous.

"$\Leftarrow$": Let T' be compact. By what we have shown $T'' : X'' \to Y''$ is compact. By $T''j_X = j_Y T$ the operator $j_Y T$ is compact. Since j_Y is isometric, also T is compact. □

6.5. Proposition: Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. Then $T : X \to Y$ is an isomorphism if and only if $T : Y' \to X'$ is an isomorphism.

Proof. "$\Rightarrow$": We have $T^{-1}T = I_X$, $TT^{-1} = I_Y$, thus by 6.1 $T'(T^{-1})' = I_{X'}$ and $(T^{-1})'T' = I_{Y'}$. Hence T' is an isomorphism and $(T')^{-1} = (T^{-1})'$.

"$\Leftarrow$": $T'' : X'' \to Y''$ is an isomorphism, and by $T''j_X = j_Y T$, T is injective. Since T' is injective, $R(T)$ is dense in Y by 6.2. Since T is a restriction of T'', $(T'')^{-1}$ is an extension of $T^{-1} : R(T) \to X$. Hence $T^{-1} : R(T) \to X$ is continuous, and this implies that $R(T)$ is closed in Y. We conclude that $T : X \to Y$ is bijective, and that $T^{-1} : Y \to X$ is continuous. □

end Mon 21.01.19

6.6. Closed Range Theorem: Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. The following are equivalent:

(i) $R(T)$ is closed.

(ii) $R(T) = N(T')_\perp$.

(iii) $R(T')$ is closed.

(iv) $R(T') = N(T)^\perp$.

Proof. (i) $\Rightarrow$ (ii) holds by 6.2. (ii) $\Rightarrow$ (i) and (iv) $\Rightarrow$ (iii) are clear.

(i) $\Rightarrow$ (iv): We factorize T as $\widehat{T} \circ q$ where $q : X \to X/N(T)$ is the quotient map and $\widehat{T} : X/N(T) \to Y$ is injective. Then $T' = q'(\widehat{T})'$, where $q' : (X/N(T))' \to X'$ has $R(q') = N(T)^\perp$ by 3.7. It thus suffices to show $R((\widehat{T})') = (X/N(T))'$, in other words: we may assume that T is injective and have to show that then $R(T') = X'$. So let $\phi \in X'$ and consider $\widetilde{\psi} := \phi \circ T^{-1} : R(T) \to \mathbb{K}$ (T^{-1} is continuous on $R(T)$ by the Open Mapping Theorem). We extend $\widetilde{\psi}$ by Hahn-Banach to an element $\psi \in Y'$. Then

$$T'\psi = \psi \circ T = \widetilde{\psi} \circ T = \phi \circ T^{-1} \circ T = \phi,$$

so $\phi \in R(T')$, and $R(T') = X'$ is proved.

(iii) $\Rightarrow$ (i): Let $R(T')$ be closed, set $Z := \overline{R(T)}$, and consider $S : X \to Z$, $x \mapsto Tx$. Then $S' : Z' \to X'$ is injective by 6.2. We claim $R(S') = R(T')$. Indeed, for $\psi \in Z'$ we have

$$S'\psi = \psi \circ S = \widetilde{\psi} \circ T = T'(\widetilde{\psi}) \in R(T'),$$

where $\widetilde{\psi} \in Y'$ is an Hahn-Banach extension of ψ, and for $\phi \in Y'$ we have

$$T'\phi = \phi \circ T = \phi|_Z \circ S = S'(\phi|_Z) \in R(S'),$$

where $\phi|_Z \in Z'$. Hence $R(S') = R(T')$ is closed, and consequently there exists $c > 0$ such that

$$c\|\psi\|_{Z'} \leq \|S'\psi\|_{X'} \quad \text{for all } \psi \in Z'. \tag{$*$}$$

We claim that $B_Z(0, c) \subseteq \overline{S(B_X(0, 1))}$. As in the second part of the Open Mapping Theorem this implies $B_Z(0, c) \subseteq S(B_X(0, 1))$, and S is surjective, i.e. $Z = R(S) = R(T)$, which finishes the proof.

To prove the claim let $z_0 \in Z$ with $\|z_0\| < c$ and assume $z_0 \notin \overline{S(B_X(0, 1))} =: V$. Then V is closed and convex and by Hahn-Banach we find $\psi \in Z'$, $\alpha \in \mathbb{R}$ such that

$$\operatorname{Re} \psi(z) \leq \alpha < \operatorname{Re} \psi(z_0) = |\psi(z_0)| \quad \text{for all } z \in V.$$

By $0 \in D$ we have $\alpha \geq 0$. Hence we may assume $\alpha > 0$ and even $\alpha = 1$ (replace ψ by ψ/α otherwise). Since $\lambda z \in V$ for $z \in V$ and $\lambda \in \mathbb{K}$ with $|\lambda| \leq 1$ we thus have

$$|\psi(z)| \leq 1 < |\psi(z_0)| \quad \text{for all } z \in V.$$

For $x \in B_X(0, 1)$ we now have $Sx \in V$ hence $|S'\psi(x)| = |\psi(Sx)| \leq 1$, i.e. $\|S'\psi\|_{X'} \leq 1$. On the other hand, by $(*)$ we have

$$1 < |\psi(z_0)| \leq \|\psi\|_{Z'}\|z_0\|_Z < c\|\psi\|_{Z'} \leq \|S'\psi\|_{X'},$$

a contradiction. □

6.7. Reflexive spaces: A Banach space X is called *reflexive* if the isometric canonical map $j_X : X \to X''$, $x \mapsto \delta_x$ is surjective.

Remark: There exists a non-reflexive space X that is isometrically isomorphic to its bidual X'' (James 1951).

Examples: Finite dimenisonal spaces are reflexive. Hilbert spaces are reflexive. l^p is reflexive for $p \in (1, \infty)$. c_0 and l^1 are not reflexive.

6.8. Proposition: (a) A Banach space is reflexive if and only if its dual space is reflexive.

(b) If X is reflexive and Y is a closed subspace then Y and X/Y are reflexive.

Proof. (a) Let X be reflexive and $\rho : X'' \to \mathbb{K}$ be linear and continuous (so $\rho \in (X'')'$). Set $\psi := \rho \circ j_X \in X'$. We claim $\rho = j_{X'}\psi$. To check this let $\phi \in X''$. We find $x \in X$ with $\phi = \delta_x$ and have

$$j_{X'}\psi(\phi) = \phi(\psi) = \delta_x(\psi) = \psi(x) = \rho(\delta_x) = \rho(\phi).$$

Hence $j_{X'}$ is surjective, i.e. X' is reflexive.

Now let X' be reflexive. It suffices to show that $R(j_X)$ is dense in X''. So let $\rho \in (X'')' = (X')''$ vanish on $R(j_X)$. We claim that $\rho = 0$ (then $R(j_X)$ is dense in X'' by 3.6(d)). We find $\psi \in X'$ with $\rho = j_{X'}\psi$ and have to show $\psi = 0$. For any $x \in X$ we have

$$\psi(x) = \delta_x(\psi) = (j_{X'}\psi)(\delta_x) = \rho(\delta_x) = 0,$$

and the proof is finished.

(b) Y is reflexive: By Hahn-Banach $q : X' \to Y'$, $\psi \to \psi|_Y$ is continuous, linear, and surjective with $N(q) = Y^\perp$ (cp. 3.7). Now let $\phi \in Y''$. Then $\phi \circ q \in X''$ and we find $x \in X$ such that $\delta_x = \phi \circ q$, i.e. such that

$$\psi(x) = \delta_x(\psi) = \phi \circ q(\psi) = \phi(\psi|_Y) \quad \text{for all } \psi \in X'.$$

For all $\psi \in N(q) = Y^\perp$ we then have $\psi(x) = 0$ which by Hahn-Banach (see 3.6(d)) implies $x \in \overline{Y} = Y$.

X/Y is reflexive: By 3.7 we have $(X/Y)' \cong Y^\perp$. Since $Y^\perp$ is a closed subspace of X', which is reflexive by (a), $Y^\perp$ is reflexive by the first part of (b). This implies that $(X/Y)'$ is reflexive, and then X/Y is reflexive, again by (a). □

6.9. Weak convergence: Let X be Banach space, (x_n) be a sequence in X, and $x_0 \in X$. Then (x_n) is said to be *weakly convergent* with *weak limit* $x_0 \in X$, in notation $x_n \rightharpoonup x_0$, if

$$\langle x_n, \phi \rangle \to \langle x_0, \phi \rangle \quad \text{for all } \phi \in X'.$$

By Hahn-Banach, a weak limit is unique. As in 2.5 we see that $x_n \rightharpoonup x_0$ implies $\|x_0\| \le \liminf_{n\to\infty} \|x_n\|$[15]. Convergence in norm $\|x_n - x_0\| \to 0$ is sometimes called *strong convergence.*

[15] $x_n \rightharpoonup x_0$ means pointwise convergence of (δ_{x_n}) to δ_{x_0} in $(X')' = X''$. By the Uniform Boundedness Principle we thus have $\sup_n \|x_n\|_X = \sup_n \|\delta_{x_n}\|_{X''} < \infty$ and 2.5 yields $\|x_0\| = \|\delta_{x_0}\| \le \liminf_{n\to\infty} \|\delta_{x_n}\| = \liminf_{n\to\infty} \|x_n\|$.

Examples: (1) If $(e_n)_{n\in\mathbb{N}}$ is an ONS in a Hilbert space then $e_n \rightharpoonup 0$: For any $x \in H$ we have by 4.8: $\sum_n |(x|e_n)|^2 \le \|x\|^2 < \infty$, hence $(x|e_n) \to 0$ as $n \to \infty$.

(2) Similarly, we have $e_n \rightharpoonup 0$ in l^p if $p \in (1,\infty)$, and $e_n \rightharpoonup 0$ in c_0. However, (e_n) is not weakly convergent in l^1. end Thu 24.01.19

6.10. Theorem (Banach-Alaoglu): Let X be a separable Banach space and $(\phi_n)_n$ be a bounded sequence in X'. Then there is $\phi_0 \in X'$ and a subsequence $(\phi_{k(n)})_n$ such that

$$\langle x, \phi_{k(n)} \rangle \to \langle x, \phi_0 \rangle \quad \text{for all } x \in X.$$

Proof. We find a dense subset $(x_l)_{l\in\mathbb{N}}$ in X. For each $l \in \mathbb{N}$ the sequence $(\langle x_l, \phi_n \rangle)_n$ is bounded in $\mathbb{K}$. Hence we find subsequences $(k_1(n))_n$ of (n), $(k_2(n))_n$ of $(k_1(n))_n$, $(k_3(n))_n$ of $(k_2(n))_n$ etc. such that, for each $l \in \mathbb{N}$, the sequence $(\langle x_l, \phi_{k_l(n)} \rangle)_n$ has a limit $\alpha_l \in \mathbb{K}$ as $n \to \infty$. Letting $(k(n)) := (k_n(n))$ we have $\langle x_l, \phi_{k(n)} \rangle \to \alpha_l$ as $n \to \infty$ for every $l \in \mathbb{N}$ and finish the proof by 2.6 (Banach-Steinhaus). □

Remark: Pointwise convergence in X' (as we have it in Theorem 6.10) is called *w^*-convergence* (“*weak-star*”). Theorem 6.10 is but a special case, the general theorem states that $B_{X'}$ is compact for the *w^*-topology*, which is induced by the functions $\phi \mapsto \phi(x)$, $x \in X$. If X is separable then the restriction of the w^*-topology to $B_{X'}$ is induced by a metric, so compactness is equivalent to sequential compactness.

6.11. Corollary: (a) If X is separable and reflexive then each bounded sequence in X has a weakly convergent subsequence.

(b) If X is reflexive then each bounded sequence in X has a weakly convergent subsequence.

Proof. (a) By reflexivity, weak convergence in X is the same as pointwise convergence in $(X')'$. Since $(X')'$ is separable, X' is separable by 3.8. Hence we can apply 6.10 in $(X')'$.

(b) is an exercise. □

Also the following is an exercise.

6.12. Proposition: Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. If T is compact then T maps weakly convergent sequences to strongly convergent sequences. The converse holds if X is reflexive.

6.13. Proposition: Let X be a Banach space and $V \subseteq X$ convex and closed. If (x_n) is a sequence in V such that $x_n \rightharpoonup x_0 \in X$ then $x_0 \in V$.

Proof. Follows from 3.14 (geometric version of Hahn-Banach). □

6.14. Uniformly convex spaces: A Banach space X is called *uniformly convex* if, for every $\varepsilon > 0$, there exists a $\delta < 0$ such that, for all $x, y \in X$ with $\|x\| = \|y\| = 0$, one has

$$\left\|\frac{x+y}{2}\right\| > 1 - \delta \implies \|x - y\| < \varepsilon.$$

This means: if the midpoint between x and y is close to the unit sphere S_X then x and y have to be close. The contraposition is

$$\|x - y\| \geq \varepsilon \implies \left\|\frac{x+y}{2}\right\| \leq 1 - \delta,$$

i.e. if x and y have a certain distance then the midpoint is a certain amount away from S_X. In particular, S_X cannot contain straight lines: if $x, y, \frac{x+y}{2} \in S_X$ then $x = y$ (this property is called *strict convexity*). The above is a quantitavie version of this, which holds uniformly on S_X.

Example: Hilbert spaces H are uniformly convex since, for $x, y \in H$ with $\|x\| = \|y\| = 1$, we have

$$\left\|\frac{x+y}{2}\right\|^2 + \left\|\frac{x-y}{2}\right\|^2 = 1.$$

by 4.2. c_0, l^1, and l^∞ are not uniformly convex since these spaces are not strictly convex.

6.15. Proposition: If X is uniformly convex then $x_n \rightharpoonup x_0$ and $\|x_n\| \to \|x_0\|$ together imply $\|x_n - x_0\| \to 0$.

Proof. We may assume $x_0 \neq 0$ and $x_n \neq 0$ for all n. Let $\widetilde{x}_n := x_n/\|x_n\|$ for $n \in \mathbb{N}_0$. It suffices to show $\widetilde{x}_n \to \widetilde{x}_0$. So we may assume $x_n \in S_X$ for $n \in \mathbb{N}_0$. It suffices to show that (x_n) is $\|\cdot\|$-Cauchy.

So we let $\varepsilon > 0$ and find $\delta > 0$ according to uniform convexity. By Hahn-Banach we find $\phi \in X'$ such that $\|\phi\| = 1$ and $\langle x_0, \phi\rangle = \|x_0\| = 1$. By weak convergence we thus have $\operatorname{Re}\langle x_n, \phi\rangle \to 1$ and find n_0 such that $\operatorname{Re}\langle x_n, \phi\rangle > 1 - \delta$ for all $n \geq n_0$. For $n, m \geq n_0$ we then have

$$\left\|\frac{x_n + x_m}{2}\right\| \geq \frac{1}{2}\big|\langle x_n, \phi\rangle + \langle x_m, \phi\rangle\big| \geq \frac{1}{2}\big(\operatorname{Re}\langle x_n, \phi\rangle + \operatorname{Re}\langle x_m, \phi\rangle\big) > 1 - \delta,$$

hence $\|x_n - x_m\| < \varepsilon$. □

Next we show that L^p-spaces are uniformly convex for $p \in (1, \infty)$.

6.16. Clarkson Inequalities: Let (Ω, Σ, μ) be a measure space and $p \in (1, \infty)$. Then we have, for all $f, g \in L^p(\Omega, \mu)$,

$$\begin{aligned}
\left\|\frac{f+g}{2}\right\|_{L^p}^p + \left\|\frac{f-g}{2}\right\|_{L^p}^p &\leq \frac{1}{2}\big(\|f\|_{L^p}^p + \|g\|_{L^p}^p\big) && \text{if } p \geq 2,\\
\left\|\frac{f+g}{2}\right\|_{L^p}^q + \left\|\frac{f-g}{2}\right\|_{L^p}^q &\leq \Big(\frac{1}{2}\|f\|_{L^p}^p + \frac{1}{2}\|g\|_{L^p}^p\Big)^{q/p} && \text{if } p \in (1, 2),
\end{aligned}$$

where $\frac{1}{p} + \frac{1}{q} = 1$, as usual.

Remark: Given $\varepsilon \in (0,1)$ and $f, g \in L^p$ with $\|f\|_{L^p} = \|g\|_{L^p} = 1$ and $\|f-g\|_{L^p} \geq 2\varepsilon$ we have

$$\Big\|\frac{f+g}{2}\Big\|_{L^p} \leq 1-\delta \quad \text{where } \delta = \begin{cases} 1-\big(1-\varepsilon^p\big)^{1/p} & , \ p \geq 2 \\ 1-\big(1-\varepsilon^q\big)^{1/q} & , \ p \in (1,2), \end{cases}$$

which shows uniform convexity of L^p for $1 < p < \infty$.

6.17. Theorem (Dual space of L^p): Let (Ω, Σ, μ) be a measure space and $p, q \in (1,\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then $J : L^q(\Omega,\mu) \to (L^p(\Omega,\mu))'$ given by $Jg(f) := \int fg\,d\mu$ is linear, isometric, and surjective. If Ω is σ-finite then the assertion holds for $p = 1$ and $q = \infty$.

Corollary: $L^p(\Omega,\mu)$ is reflexive for $p \in (1,\infty)$.

Remark: One can show that every uniformly convex Banach space is reflexive (see, e.g., Brézis' book).

Proof. By 1.21 we only have to prove surjectivity of J.

Case $p \in (1,\infty)$: We take $\phi \in (L^p(\Omega,\mu))'$ and may assume $\|\phi\| = 1$. We denote by $S := S_{L^p}$ the unit sphere of L^p. First we show that ϕ attains its maximum on S, i.e. that there exists $f_0 \in S$ such that $\phi(f_0) = 1 = \sup_{f\in S} |\phi(f)| = \sup_{f \in S} \operatorname{Re}\phi(f)$.

Let (f_n) be a sequence in S such that $\operatorname{Re}\phi(f_n) \to 1$. We show that (f_n) is Cauchy. For this we use uniform convexity of L^p. We let $\varepsilon > 0$ and find $\delta > 0$ accordingly. Then we find n_0 such that $\operatorname{Re}\phi(f_n) > 1-\delta$ for $n \geq n_0$. For $n, m \geq 0$ we then have

$$1-\delta < \frac{1}{2}\big(\operatorname{Re}\phi(f_n) + \operatorname{Re}\phi(f_m)\big) \leq \Big\|\frac{f_n+f_m}{2}\Big\|_{L^p},$$

hence $\|f_n - f_m\|_{L^p} < \varepsilon$. We thus find $f_0 \in L^p$ such that $\|f_n - f_0\|_{L^p} \to 0$ and conclude $f_0 \in S$ and $\operatorname{Re}\phi(f_0) = 1$, $\phi(f_0) = 1$.

We now find $g \in L^q$ with $\|g\|_{L^q} = 1$ such that $Jg(f_0) = 1$ (in the same way we did before). For any $f \in L^p$ we denote let $D_p(f) := |f|^{p-1}\overline{\operatorname{sgn} f}$. Then

$$|D_p(f)|^q = |f|^{qp-q} = |f|^p, \qquad D_p(f) \in L^q, \qquad \|D_p(f)\|_{L^q}^q = \|f\|_{L^p}^p = J(D_p(f))f.$$

In particular, $g_0 := D_p(f_0) \in L^q$ with $\|g_0\|_{L^q} = 1$ and $Jg_0(f_0) = 1$. end Mon 28.01.19

We claim that $\phi = Jg_0 = JD_p(f_0)$, i.e. $\phi(f) = (JD_p(f_0))(f)$ for all $f \in L^p$. To prove this we fix $f \in S$ and consider

$$h(t) := \frac{|\phi(f_0+tf)|^p}{\|f_0+tf\|_{L^p}^p} \qquad \text{for } |t| < 1.$$

Clearly, h is maximal at $t = 0$, hence $h'(0) = 0$ as h is differentiable, as we shall show. We have

$$\begin{aligned}\frac{d}{dt}\big|\phi(f_0+tf)\big|^p &= \frac{d}{dt}\big(\phi(f_0+tf)\overline{\phi(f_0+tf)}\big)^{p/2}\\ &= \frac{p}{2}|\phi(f_0+tf)|^{p-2}\cdot 2\mathrm{Re}\,\Big(\overline{\phi(f_0+tf)}\frac{d}{dt}\phi(f_0+tf)\Big)\\ &= p|\phi(f_0+tf)|^{p-1}\,\mathrm{Re}\,\big(\overline{\mathrm{sgn}\,\phi(f_0+tf)}\phi(f)\big).\end{aligned}$$

Similarly, we have

$$\partial_t\big|f_0+tf\big|^p = p|f_0+tf|^{p-1}\,\mathrm{Re}\,\big(\overline{\mathrm{sgn}\,(f_0+tf)}\,f\big).$$

We thus see $\partial_t\big|f_0+tf\big|^p\in L^1$ and

$$\Big|\partial_t\big|f_0+tf\big|^p\Big|\le p|f_0+tf|^{p-1}|f|\le 2^p p\big(|f_0|^p+|f|^p\big).$$

By dominated convergence we obtain

$$\begin{aligned}\frac{d}{dt}\|f_0+tf\|_{L^p}^p &= \int_\Omega \partial_t|f_0+tf|^p\,d\mu = p\int_\Omega |f_0+tf|^{p-1}\mathrm{Re}\,\big(\overline{\mathrm{sgn}\,(f_0+tf)}\,f\big)\,d\mu\\ &= p\,\mathrm{Re}\,\Big(J\big(D_p(f_0+tf)\big)f\Big).\end{aligned}$$

Now we can finish the proof

$$\begin{aligned}0 &= h'(0)=\frac{d}{dt}\left(\frac{|\phi(f_0+tf)|^p}{\|f_0+tf\|_{L^p}^p}\right)\Bigg|_{t=0}\\ &= \|f_0\|_{L^p}\frac{d}{dt}\,\big(\big|\phi(f_0+tf)\big|^p\big)\big|_{t=0}-\big|\phi(f_0)\big|^p\|f_0\|_{L^p}^{-2p}\frac{d}{dt}\big(\|f_0+tf\|_{L^p}^p\big)\big|_{t=0}\\ &= p\,\mathrm{Re}\,\big(\overline{\mathrm{sgn}\,\phi(f_0)}\phi(f)\big)-p\,\mathrm{Re}\,\big(J\big(D_p(f_0)\big)f\big)\\ &= p\,\big(\mathrm{Re}\,\big(\phi(f)\big)-\mathrm{Re}\,\big(J\big(D_p(f_0)\big)f\big)\big)\,,\end{aligned}$$

which first implies $\mathrm{Re}\,\phi=\mathrm{Re}\,J(D_p(f_0))$ and then $\phi=J(D_p(f_0))$ by Lemma 3.3.

Case $p=1$: Let Ω be σ-finite and $\phi\in(L^1(\Omega,\mu))'$ with $\|\phi\|_{(L^1)'}=1$. For fixed $A\in\Sigma$ with $\mu(A)<\infty$ we define ϕ_A by $\phi_A(f):=\phi(1_A f)$. For any $p\in(1,\infty)$ and we have by Hölder

$$|\phi_A(f)|\le\|\phi\|_{(L^1)'}\int_A|f|\,d\mu\le\mu(A)^{1/q}\|f\|_{L^p},$$

where $\frac{1}{p}+\frac{1}{q}=1$. Hence $\phi_A\in(L^p(\Omega,\mu))'$. By the first part we find $g_A\in L^q$ such that $\phi_A(f)=\int g_A f\,d\mu$ for all $f\in L^p$ and we have $\|g_A\|_{L^q}\le\mu(A)^{1/q}$. Since $\phi_A(f)$ if $f=0$ a.e. on A we find that $g_A=0$ a.e. on $\Omega\setminus A$. It is easy to see that g_A does not depend on $p\in(1,\infty)$. Considering the probability space $(A,\Sigma\cap A,\mu/\mu(A))$ and letting $q\to\infty$ we obtain $g_A \in L^\infty(A)$ and $g_A \le 1$ a.e. on A[16]. Since Ω is σ-finite we find a sequence (A_j) in Σ such that $A_j \subseteq A_{j+1}$ and $\mu(A_j) < \infty$. Clearly $g_{A_{j+1}} = g_{A_j}$ a.e. on A_j and $g := \lim_j g_{A_j}$ is well-defined a.e. with $|g| \le 1$ a.e. in Ω, i.e. $g \in L^\infty(\Omega, \mu)$ and $\|g\|_{L^\infty} \le 1$. For $f \in L^p(\Omega, \mu)$ for some $p > 1$ and $f = 0$ a.e. outside some A_j we have $\phi(f) = Jg(f)$. Since the set of all such f is dense in $L^1(\Omega, \mu)$ we conclude that $\phi = Jg$ as desired. □

We still have to prove 6.16. The next is a matrix version of a "complex interpolation" result due to E.M. Stein. We shall use it in the proof of 6.16. Such results are a very powerful tool in functional analysis, harmonic analysis, and in various areas of the large field of partial differential equations. We shall need some complex analysis for formulation and proof.

6.18. Stein interpolation (matrix version): Let $S := \{z \in \mathbb{C} : \operatorname{Re} z \in (0,1)\}$ and $n \in \mathbb{N}$. Let $T : \overline{S} \to \mathbb{C}^{n\times n}$ be continuous and bounded on $\overline{S}$ and analytic in S. Let $p_j, q_j \in [1,\infty]$ and $M_j > 0$ for $j = 0, 1$ and assume that

$$\|T(z)x\|_{q_j} \le M_j \|x\|_{p_j} \qquad \text{for all } x \in \mathbb{C}^n,\ \operatorname{Re} z = j, \text{ and } j = 0, 1.$$

Then, for $\theta \in (0,1)$,

$$\|T(\theta)x\|_q \le M_0^{1-\theta} M_1^{\theta} \|x\|_p \qquad \text{for all } x \in \mathbb{C}^n,$$

where $p, q \in [1, \infty]$ are given by $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, $\frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$.

Notation: For $r \in [1,\infty]$ we shall use the *dual exponent* $r' \in [1,\infty]$ given by $\frac{1}{r} + \frac{1}{r'} = 1$.

Proof. We let $T(z) = \big(T_{kl}(z)\big)_{k,l=1}^n$ for $z \in \overline{S}$ and fix $\theta \in (0,1)$. It suffices to show

$$\Big|\sum_{k,l=1}^n T_{kl}(z) x_l y_k\Big| \le M_0^{1-\theta} M_1^{\theta}$$

for all $x = (x_l), y = (y_k) \in \mathbb{C}^n$ with $\|x\|_p \le 1$ and $\|y\|_{q'} \le 1$. For such vectors $x, y \in \mathbb{C}^n$ and $z \in \overline{S}$, $k, l \in \{1, \dots, n\}$ we define

$$\begin{aligned} x_l(z) &:= |x_l|^{\alpha(z)} \operatorname{sgn} x_l, & \alpha(z) &= p\Big(\frac{1-z}{p_0} + \frac{z}{p_1}\Big),\\ y_k(z) &:= |y_k|^{\beta(z)} \operatorname{sgn} y_k, & \beta(z) &= q'\Big(\frac{1-z}{q_0'} + \frac{z}{q_1'}\Big), \end{aligned}$$

[16]Otherwise, we find $\delta > 0$ such that $B := \{|g_A| \ge 1 + \delta\}$ has $\mu(B) > 0$, and then

$$(1+\delta)\mu(B)^{1/q} \le \|g_A\|_{L^q} \le \mu(A)^{1/q}$$

leads for $q \to \infty$ to $1 + \delta \le 1$, a contadicton.

where we let $al(z) := 1$ if $p_0 = p_1$, $\beta(z) := 1$ if $q_0 = q_1$ [17], and understand $0^w := 0$ (in any case, for $x = 0$ we have $\operatorname{sgn} x = 0$). Observe that $\alpha(j) = p/p_j$, $\beta(j) = q'/q'_j$ for $j = 0, 1$ (where $\infty/\infty := 1$). We set

$$I(z) := M_0^{1-z} M_1^{-z} \sum_{k,l=1}^{n} T_{kl}(z) x_l(z) y_k(z) \quad \text{for } z \in \overline{S}.$$

Then $z \mapsto I(z)$ is continuous and bounded on $\overline{S}$, analytic in S, and satisfies $|I(z)| \leq 1$ for $\operatorname{Re} z = 0$ and for $\operatorname{Re} z = 1$: By

$$|x_l(z)| = \big| |x_l|^{\alpha(z)} \big| = |x_l|^{\operatorname{Re}(\alpha(z))} = |x_l|^{\alpha(\operatorname{Re} z)}$$

we have, for $j = 0, 1$ and $\operatorname{Re} z = j$,

$$\|(x_l(z))\|_{p_j}^{p_j} = \sum_l |x_l(z)|^{p_j} = \sum_l |x_l(z)|^{\alpha(j) p_j} = \sum_l |x_l|^p = \|x\|_p^p \leq 1,$$

and, similarly, $\|(y_k(z))\|_{q'_j}^{q'_j} = \|(y_k(z))\|_{q'}^{q'} \leq 1$ [18]. Hence we get $|I(z)| \leq 1$ from the assumptions.

The three lines lemma from complex analysis then yields $|I(z)| \leq 1$ for all $z \in \overline{S}$. In particular, we have $|I(\theta)| \leq 1$, which ends the proof. □

Proof of 6.16. $p \geq 2$: We have for all $\alpha, \beta \geq 0$:

$$\alpha^p + \beta^p \leq \big(\alpha^2 + \beta^2\big)^{p/2}.$$

The derivative of $h : [0, \infty) \to \mathbb{R}$, $t \mapsto t^{p/2}$ is increasing and $h(0) = 0$, hence, for $s, t \geq 0$,

$$h(t) + h(s) = \Big(\int_0^t + \int_0^s \Big) h'(\tau)\, d\tau \leq \Big(\int_0^t + \int_t^{s+t} \Big) h'(\tau)\, d\tau = h(s + t),$$

and we obtain the claim for $t = \alpha^2$, $s = \beta^2$.

Using this for $\alpha = |x + y|/2$, $\beta = |x - y|/2$ we obtain

$$\Big|\frac{x+y}{2}\Big|^p + \Big|\frac{x-y}{2}\Big|^p \leq \Big(\frac{|x+y|^2}{4} + \frac{|x-y|^2}{4} \Big)^{p/2} = \Big(\frac{|x|^2 + |y|^2}{2} \Big)^{p/2} \leq \frac{1}{2}\big(|x|^p + |y|^p\big),$$

where the last inequality is due to convexity of h.[19] The assertion follows by integration.

[17] also for $p_0 = p_1 = \infty$, $q_0 = q_1 = 1$!

[18] Observe that, for $p < \infty$, $p_0 = \infty$, say, we have $p_1 < \infty$, $\alpha(z) = pz/p_1$, and $\big| |x_l|^{\alpha(z)} \big| = 1$ for $\operatorname{Re} z = 0$ if $x_l \neq 0$. Hence $\|(x_l(z))\|_{p_0} \leq 1$. If $p_0 = p_1 = \infty$ then also $p = \infty$ and $x_l(z) = x_l$ for all z. So also in this case $\|(x_l(z))\|_{p_j} \leq 1$.

[19] In the lecture we argued via 6.8 for this inequality.

$p \in (1,2)$: The key is the inequality

$$\Big|\frac{x+y}{2}\Big|^q + \Big|\frac{x-y}{2}\Big|^q \le \Big(\frac{|x|^p + |y|^p}{2}\Big)^{q/p}$$

for all $x, y \in \mathbb{C}$, which is equivalent to

$$\Big\|\Big(\frac{x+y}{2}, \frac{x-y}{2}\Big)\Big\|_q \le 2^{-1/p}\|(x,y)\|_p \quad \text{for all } (x,y) \in \mathbb{C}^2.$$

We apply 6.18 for $n = 2$ and $T = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. The inequality is clear for $p_0 = q_0 = 2$ with $M_0 = 2^{-1/2}$, but it is also clear for $p_1 = 1$, $q_1 = \infty$ with $M_1 = 2^{-1}$, since

$$\max\Big\{\frac{|x+y|}{2}, \frac{|x-y|}{2}\Big\} \le \frac{1}{2}\big(|x| + |y|\big) \quad \text{for all } (x,y) \in \mathbb{C}^2.$$

By 6.18 we obtain it for $p \in (1,2)$: If $\theta \in (0,1)$ and $\frac{1}{p} = \frac{1-\theta}{2} + \frac{\theta}{1} = \frac{1+\theta}{2}$ then

$$\frac{1}{q} = 1 - \frac{1}{p} = 1 - \frac{1+\theta}{2} = \frac{1-\theta}{2} + \frac{\theta}{\infty},$$

and the constant is

$$M_0^{1-\theta} M_1^{\theta} = 2^{-(1-\theta)/2} 2^{-\theta} = 2^{-(1+\theta)/2} = 2^{-1/p}$$

as claimed. Next we recall Minkowski's inequality for $r > 1$:

$$\Big(\sum_j \Big(\sum_k a_{jk}\Big)^r\Big)^{1/r} \le \sum_k \Big(\sum_j a_{jk}^r\Big)^{1/r} \quad \text{for all } a_{jk} \ge 0.$$

Taking this to the power of r and replacing $a_{jk} = b_{jk}^{1/r}$ we obtain

$$\sum_j \Big(\sum_k b_{jk}^{1/r}\Big)^r \le \Big(\sum_k \Big(\sum_j b_{jk}\Big)^{1/r}\Big)^r \quad \text{for all } b_{jk} \ge 0.$$

Replacing $c_{jk} = b_{jk} w_k^r$ we obtain the weighted version

$$\sum_j \Big(\sum_k c_{jk}^{1/r} w_k\Big)^r \le \Big(\sum_k \Big(\sum_j c_{jk}\Big)^{1/r} w_k\Big)^r \quad \text{for all } c_{jk} \ge 0,\ w_k \ge 0.$$

We now take simple functions f, g and may assume $f = \sum_k x_k 1_{A_k}$, $g = \sum_k y_k 1_{A_k}$ where the A_k are pairwise disjoint. We apply the weighted inequality for $r = q/p$, $j = 1, 2$, $c_{1k} = |(x_k + y_k)/2|^q$, $c_{2k} = |(x_k - y_k)/2|^q$, $w_k = \mu(A_k)$. Then

end Thu 31.01.19

$$\left(\left\|\frac{f+g}{2}\right\|_{L^p}^q+\left\|\frac{f-g}{2}\right\|_{L^p}^q\right)^{1/q} = \left(\left(\sum_k\left|\frac{x_k+y_k}{2}\right|^p w_k\right)^{q/p}+\left(\sum_k\left|\frac{x_k-y_k}{2}\right|^p w_k\right)^{q/p}\right)^{1/q}$$
$$\leq \left(\sum_k\left(\left|\frac{x_k+y_k}{2}\right|^q+\left|\frac{x_k-y_k}{2}\right|^q\right)^{p/q} w_k\right)^{1/p}$$

and we continue by the key inequality

$$\leq \left(\sum_k \frac{|x_k|^p+|y_k|^p}{2} w_k\right)^{1/p} = 2^{-1/p}\Big(\|f\|_{L^p}^p+\|g\|_{L^p}^p\Big)^{1/p},$$

which finishes the proof. □

6.19. Remark: In the same way as 6.18 one can prove a version with weighted norms on $\mathbb{C}^n$: Let $\mu_l, \nu_k > 0$ for $k, l = 1, \dots, n$ and replace $\|x\|_p = \big(\sum_l |x_l|^p\big)^{1/p}$, $\|x\|_{p_j}$ by $\|x\|_{p,\mu} = \big(\sum_l |x_l|^p \mu_l\big)^{1/p}$, $\|x\|_{p_j,\mu}$ and $\|y\|_q$, $\|y\|_{q_j}$ by $\|y\|_{q,\nu}$, $\|y\|_{q_j,\nu}$. The duality bracket between $(\mathbb{C}^n, \|\cdot\|_{q,\nu})$ and $(\mathbb{C}^n, \|\cdot\|_{q',\nu})$ is now given by $\langle x, y\rangle = \sum_k x_k y_k \nu_k$.

Moreover, one has versions for matrices $T(z) \in \mathbb{C}^{n\times m}$ with $m \neq n$, which follow from those for quadratic matrices.

We shall use the weighted version to prove the following result.

6.20. Theorem (Riesz-Thorin): Let (Ω, Σ, μ) and (Ω', Σ', ν) be σ-finite measure spaces and $p_j, q_j \in [1, \infty]$ for $j = 0, 1$. Let T be a linear operator and $M_0, M_1 > 0$ such that

$$\|Tf\|_{L^{q_j}(\nu)} \leq M_j \|f\|_{L^{p_j}(\mu)} \quad \text{for all } f \in L^{p_j}(\Omega, \mu) \text{ and } j = 0, 1.$$

If $\theta \in (0, 1)$ and $p, q \in [1, \infty]$ are defined by $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ and $\frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$ then

$$\|Tf\|_{L^q(\nu)} \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p(\mu)} \quad \text{for all } f \in L^{p_0}(\Omega, \mu) \cap L^{p_1}(\Omega, \nu).$$

Proof. For simple functions $f = \sum_l x_l 1_{A_l}$ with the $A_l \in \Sigma$ pairwise disjoint of positive μ-measure and $g = \sum_k y_k 1_{B_k}$ with $B_k \in \Sigma'$ pairwise disjoint of positive ν-measure we have

$$\langle Tf, g\rangle = \sum_{k,l} \langle T1_{A_l}, 1_{B_k}\rangle x_l y_k = \sum_{k,l} \nu_k^{-1} \langle T1_{A_l}, 1_{B_k}\rangle x_l y_k \nu_k$$

and it suffices to show

$$|\langle Tf, g\rangle| \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p(\mu)} \|g\|_{L^{q'}(\nu)}$$

for all such functions. Here $\|f\|_{L^p(\mu)}$ equals the weighted l^p-norm of the vector (x_l) with weights $\mu_l = \mu(A_l)$ and likewise for g. Letting $T_{kl} := \nu_k^{-1}\langle T1_{A_l}, 1_{B_k}\rangle$ and $\widetilde{T} := (T_{kl})_{k,l}$ we thus have to show, with $\widetilde{\mu} = (\mu_l)_l$ and $\widetilde{\nu} = (\nu_k)_k$,

$$\|\widetilde{T}x\|_{l^q(\widetilde{\nu})} \leq M_0^{1-\theta}M_1^{\theta}\|x\|_{l^p(\widetilde{\mu})} \quad \text{for all vectors } x = (x_l)_l.$$

Since we have by assumption, for $j = 0, 1$,

$$\|\widetilde{T}x\|_{l^{q_j}(\widetilde{\nu})} \leq M_j\|x\|_{l^{p_j}(\widetilde{\mu})} \quad \text{for all vectors } x = (x_l)_l.$$

this follows from the weighted version of 6.18. □

We state the following application of 6.11 and 6.13.

6.21. Proposition: Let X be a reflexive Banach space, let $K \subseteq X$ closed and convex. Assume that $f : K \to \mathbb{R}$ is continuous and convex[20]. if K is not bounded assume in addition that $f(x) \to \infty$ if $\|x\| \to \infty$ with $x \in K$. Then f attains its minimum on K.

Proof. is an exercise. Apply 6.13 with $V = \overline{\text{conv}\{x_n : n \in \mathbb{N}\}}$. □

[20]This means $f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$ for all $x, y \in K$ and $t \in [0, 1]$.

7 Self-adjoint compact operators

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

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

i.e. $T^* = J_{H_1}^{-1} T' J_{H_2}$ where $T' \in \mathcal{L}(H_2', H_1')$ denotes the dual operator of T and, for $j = 1, 2$, $J_{H_j} : H - j \to H_j'$, $y \mapsto (\cdot|y)_{H_j}$ (cp. 4.6). We clearly have

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

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

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

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

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

Examples: (1) Let Y be a closed linear subspace of H and P_Y the linear projection operator onto Y from 4.5. Then $R(P_Y) = Y$, $N(P_Y) = Y^\perp$ and we obtain, for $x, y \in H$,[21]

$$(P_Y x|y) = (P_Y x|P_Y y + P_{Y^\perp} y) = (P_Y x|P_Y y) = (P_Y x + P_{Y^\perp} x|P_Y y) = (x|P_Y y).$$

Hence P_Y is self-adjoint. The operator P_Y is called the *orthogonal projection onto* Y.

(2) If $\Omega \subseteq \mathbb{R}^n$ is bounded then the operator $R = (-\Delta_D)^{-1} \in \mathcal{L}(L^2(\Omega))$ is self-adjoint since we have $(Rf|g) = (f|Rg)$ for all $f, g \in L^2(\Omega)$ (see "Continuation of 5.5").

7.2. Lemma: If S is self-adjoint then $\|S\| = \sup\{|(Sx|x)| : \|x\| \le 1\}$.

Proof. Let M denote the sup in the claim and take x, y with $\|x\| = \|y\| = 1$. Then

$$\begin{aligned} 4\mathrm{Re}\,(Sx|y) &= 2(Sx|y) + 2(Sy|x) = (S(x+y)|S(x+y)) - (S(x-y)|x-y) \\ &\le M\big(\|x+y\|^2 + \|x-y\|^2\big) = 2M\big(\|x\|^2 + \|y\|^2\big) = 4M, \end{aligned}$$

and we conclude

$$\|S\| = \sup_{\|x\|,\|y\| \le 1} |(Sx|y)| = \sup_{\|x\|,\|y\| \le 1} \mathrm{Re}\,(Sx|y) \le M.$$

The converse inequality is clear. □

end Mon 04.02.19

7.3. Eigenvectors of self-adjoint operators: Let $T \in \mathcal{L}(H)$. A $\lambda \in \mathbb{C}$ is called *eigenvalue of* T if $N(\lambda I - T) \neq \{0\}$. In this case $E_\lambda := N(\lambda I - T)$ is called the corresponding *eigenspace* and any $x \in E_\lambda \setminus \{0\}$ is called a corresponding *eigenvector*. In this case we always have $|\lambda| \leq \|T\|$.

Proof. For $|\lambda| > \|T\|$ we have $\lambda I - T = \lambda(I - \frac{T}{\lambda})$ and $\|\frac{T}{\lambda}\| < 1$. Hence the *Neumann series* $R := \sum_{k=0}^{\infty} \lambda^{-k} T^k$ is absolutely convergent. Clearly, R commutes with T and we obtain $R(I - \frac{T}{\lambda}) = (I - \frac{T}{\lambda})R = I$, hence $\lambda I - T : H \to H$ is an isomorphism and $(\lambda I - T)^{-1} = \lambda - 1R$. In particular, λ cannot be an eigenvalue of T. $\square$

Now let $S \in \mathcal{L}(H)$ be self-adjoint and λ be an eigenvalue of S. Then $\lambda \in \mathbb{R}$, hence $-\|S\| \leq \lambda \leq \|S\|$. We also have $R(\lambda I - S)^{\perp} = N(\lambda I - S)$. If $\mu \neq \lambda$ is another eigenvalue then $E_\lambda \perp E_\mu$, in particular $E_\mu \subseteq E_\lambda^{\perp}$.

Proof. We find $x \in H$ with $\|x\| = 1$ and $Sx = \lambda x$. Hence

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

i.e. $\lambda \in \mathbb{R}$. The formulas on kernel and range hold by $S^* = S$ and 6.2 (formulated for adjoint operators). If $y \in E_\mu \setminus \{0\}$ then

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

hence $x \perp y$ by $\lambda \neq \mu$. $\square$

The following result states that self-adjoint compact operators can be diagonlized via an orthonormal basis of eigenvectors (just as Hermitian matrices on $\mathbb{K}^n$).

7.4. Theorem: Let H be Hilbert space with $\dim H = \infty$, let $S \in \mathcal{L}(H)$ be self-adjoint and compact. Then S has at most countably many eigenvalues. Denote pairwise different non-zero eigenvalues of S by λ_j, $j \in I$, where $I = \emptyset$ or $I = \{1, \ldots, N\}$ for some $N \in \mathbb{N}$ or $I = \mathbb{N}$. For $j \in I$ write $E_j := E_{\lambda_j}$ and $P_j := P_{E_j}$ for the orthogonal projection onto E_j. Set $\lambda_0 := 0$, $I_0 := I \subset \{0\}$, and write P_0 for the orthogonal projection onto $N(S)$. Then the following holds.

(a) For all $j \in I$ we have $\nu_j := \dim E_j < \infty$ and $P_j S = S P_j$.

(b) $\max_{j \in I_0} |\lambda_j| = \|S\|$ and $\lambda_j \to 0$ if $I = \mathbb{N}$.

(c) $S = \sum_{j \in I} \lambda_j P_j$ with convergence in operator norm.

[21] Recall 4.7: we have $I - P_Y = P_{Y^\perp}$.

Proof. Let λ be an eigenvalue, $x \in E_\lambda$ and $y \in E_\lambda^\perp$. Then $\lambda \in \mathbb{R}$, $Tx = \lambda x \in E_\lambda$, and

$$(Sy|x) = (y|Sx) = (y|\lambda x) = 0,$$

thus $Sy \in E_\lambda^\perp$. Hence S leaves E_λ and $E_\lambda^\perp$ invariant. We conclude that S commutes with every P_j, $j \in I_0$:

$$P_jS = P_jSP_j + P_jS(I - P_j) = P_jSP_j = P_jSP_j + (I - P_j)SP_j = SP_j.$$

Now let $j \in I$ and (x_n) be a bounded sequence in E_j. Since T is compact, $(x_n) = (\lambda^{-1}Sx_n)$ has a convergent subsequence. By 1.28 we have $\dim E_j < \infty$. We have proved (a).

(b) and (c) are clear if $S = 0$ with $J = \emptyset$. So let $S \neq 0$. We show that $-\|S\|$ or $\|S\|$ is an eigenvalue. By 7.2 we find a sequence (x_n) with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $|(Sx_n|x_n)| \to \|S\|$. Since $(Sx_n|x_n)$ is real we may assume $(Sx_n|x_n) \to \lambda \in \{-\|S\|, \|S\|\}$. By

$$\begin{aligned} \|Sx_n - \lambda x_n\|^2 &= \|Sx_n\|^2 - 2\mathrm{Re}\,(Sx_n|\lambda x_n) + \lambda^2\|x_n\|^2 \\ &\leq 2\lambda^2 - 2\lambda(Sx_n|x_n) \to 0 \quad (n \to \infty) \end{aligned}$$

we see that $Sx_n - \lambda x_n \to 0$. On the other hand, (Sx_n) has convergent subseqeunce with $Sx_{k(n)} \to y \in H$. Hence also $\lambda x_{k(n)} \to y$, $x_{k(n)} \to \lambda^{-1}y$, and we infer $\|y\| = |\lambda| \neq 0$ and

$$y = \lim_n Sx_{k(n)} = S(\lambda^{-1}y) = \lambda^1 Sy,$$

i.e. y is an eigenvector for the eigenvalue λ. We put $\lambda_1 = \lambda$, denote the eigenspace by E_1 and let $H_2 := E_1^\perp$, $S_2 := S|_{H_2}$. Then $S_2 \in \mathcal{L}(H_2)$ is self-adjoint and compact. If $S_2 = 0$ then $I = \{1\}$ and we stop. If $S_2 \neq 0$ then we can repeat the step, find λ_2 and E_2, and set $H_3 := E_2^\perp$, $S_3 := S_2|_{H_3} = S|_{H_3}$ etc.

If the procedure stops after N steps then $I = \{1, \dots, N\}$, otherwise $I = \mathbb{N}$. Observe that $(|\lambda|_j) = (\|S_j\|)$ is decreasing (if we set $H_1 := H$, $S_1 := S$). In case $I = \mathbb{N}$ we have $\lambda_j \to 0$: For each $j \in \mathbb{N}$ we find $x_j \in E_j$ with $\|x_j\| = 1$. Since S is compact, we find a convergent subsequence of $(Sx_j) = (\lambda_j x_j)$ such that $Sx_{k(j)} \to y$ as $j \to \infty$. Since (x_j) is an ONB of $\widetilde{H} := \overline{\mathrm{lin}\{x_j : j \in \mathbb{N}\}}$ we have $y \in H$. However $(y|x_n) = \lim_{j\to\infty}(x_{k(j)}|x_n) = 0$ for any n, hence $y = 0$ and thus $\lim_j \lambda_{k(j)} = 0$, i.e. $\inf_j |\lambda_j| = 0$. Since $(|\lambda_j|)$ is decreasing this proves $\lambda_j \to 0$.

For each $n \in I$ we have $H_{n+1}^\perp = E_1 \oplus E_2 \oplus \dots \oplus E_n$ where the sum is orthogonal. On this space we clearly have $S = \sum_{j=1}^n \lambda_j P_j + S_{n+1}Q_{n+1}$ where $Q_{n+1} = I - (P_1 + \dots + P_n)$. This proves (c) if $I = \{1, \dots, N\}$. If $I = \mathbb{N}$ we have for each n

$$\Big\|S - \sum_{j=1}^n \lambda_j P_j\Big\| = \|S_{n+1}Q_{n+1}\| \leq \|S_{n+1}\| = |\lambda_{n+1}| \to 0,$$

which proves (c) in this case. □

Remark: (1) In the situation of 7.4 we have $\dim N(S) = \infty$ if I is finite. (2) We can choose finite ONBs in each eigenspace E_j, $j \in I$, and a (possibly uncountable) ONB in $N(S)$. We thus see that H has ONB consisting of eigenvectors of S. Letting $M := \sum_{j\in I} \nu_j$ we also see that in case $S \neq 0$ we find a sequence $(\mu_m)_{m=1}^M$ of eigenvalues and an ONS $(e_m)_{m=1}^M$ of corresponding eigenvectors satisfying $|\mu_m| \geq |\mu_{m+1}| > 0$ and

$$Sx = \sum_{m=1}^{M} \mu_m (x|e_m) e_m \quad \text{for all } x \in H.$$

If S is injective then (e_m) is an ONB of H.

7.5. Application to the Dirichlet problem: Let $\Omega \subseteq \mathbb{R}^n$ be bounded. Then the operator $R = (-\Delta_D)^{-1} \in \mathcal{L}(L^2(\Omega))$ is self-adjoint (by 7.2) and compact (by the corollary to 5.10), and it is clearly injective. By 7.4 we find a sequence $(\mu_m)_{m\in\mathbb{N}}$ of eigenvalues and an ONB $(e_m)_{m\in\mathbb{N}}$ in $H = L^2(\Omega)$ such that

$$Rf = \sum_{m=1}^{\infty} \mu_m (f|e_m)_{L^2}\ e_m \quad \text{for all } f \in L^2(\Omega).$$

Observe that, for each $m \in \mathbb{N}$, we have $Re_m = \mu_m e_m$, $Re_m \in H_0^1(\Omega)$, and thus

$$\mu_m = (e_m|Re_m)_{L^2} = \mathfrak{a}(Re_m, Re_m) = \|\nabla Re_m\|_{L^2}^2 \geq \operatorname{diam}(\Omega)^{-2}\|Re_m\|_{L^2}^2 = \operatorname{diam}(\Omega)^{-2}\mu_m^2$$

by 5.6. Hence $0 < \mu_m \leq \operatorname{diam}(\Omega)^2$ and we can order $\mu_1 \geq \mu_2 \geq \ldots \to 0$.

Defining the linear operator $A = -\Delta_D$ by $A := R^{-1}$ on $D(A) = R(R)$ we then have $e_m \in D(A)$ and

$$Ae_m = \mu_m^{-1} e_m \quad \text{for all } m \in \mathbb{N}, \qquad 0 < \mu^{-1} \leq \mu_2^{-1} \leq \mu_3^{-1} \leq \ldots \to \infty.$$

Hence the Dirichlet Laplacian on Ω has an ONB of eigenfunctions and a corresponding sequence of positive eigenvalues tending to ∞.