Lecture Notes
Functional Analysis
WS 2012/2013

February 13, 2013

Contents

Chapter I

Normed vector spaces, Banach spaces and metric spaces

1 Normed vector spaces and Banach spaces

In the following let X be a linear space (vector space) over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 1.1. *A* ***seminorm*** *on X is a map $p : X \to \mathbb{R}_+ = [0, \infty)$ s.t.*

(a) $p(\alpha x) = |\alpha| p(x) \quad \forall \alpha \in \mathbb{F}, \forall x \in X$ (homogeneity).

(b) $p(x + y) \leq p(x) + p(y) \quad \forall x, y \in X$ (triangle inequality).

If, in addition, one has

(c) $p(x) = 0 \Rightarrow x = 0$

then p is called a ***norm****. Usually one writes $p(x) = \|x\|$, $p = \|\cdot\|$. The pair $(X, \|\cdot\|)$ is called a* ***normed (vector) space****.*

Remark 1.2. • *If $\|\cdot\|$ is a seminorm on X then*

$$\big|\|x\| - \|y\|\big| \leq \|x - y\| \quad \forall x, y \in X \quad \text{(reverse triangle inequality).}$$

Proof.

$$\begin{aligned} \|x\| = \|x - y + y\| &\leq \|x - y\| + \|y\| \\ \Rightarrow \|x\| - \|y\| &\leq \|x - y\| \end{aligned}$$

Now swap $x\&y$: $\|x\| - \|y\| \leq \|y - x\| = \|(-1)(x - y)\| = \|x - y\|$.
Hence

$$\big|\|x\| - \|y\|\big| = max(\|x\| - \|y\|, \|y\| - \|x\|) \leq \|x - y\|.$$

□

- $\|\vec{0}\| = \|0.\vec{0}\| = |0|.\|\vec{0}\| = 0.$

Interpret $\|x - y\|$ as distance between x and y.

Definition 1.3. *Let $(x_n)_{n\in\mathbb{N}} = (x_n)_n$ be a sequence in a normed vector space X ($(X, \|\cdot\|)$). Then $(x_n)_n$* ***converges*** *to a limit $x \in X$ if $\forall \varepsilon > 0 \exists N_\varepsilon \in \mathbb{N}$ s.t. $\forall n \geq N_\varepsilon$ it holds $\|x_n - x\| < \varepsilon$ (or $\|x_n - x\| \leq \varepsilon$). One writes $x_n \to x$ or $\lim_{n\to\infty} x_n = x$.*
$(x_n)_n$ is a ***Cauchy sequence*** *if $\forall \varepsilon > 0 \exists N_\varepsilon \in \mathbb{N}$ s.t. $\forall n, m \geq N_\varepsilon$ it holds $\|x_n - x_m\| < \varepsilon$ (or $\|x_n - x_m\| \leq \varepsilon$).*
$(X, \|\cdot\|)$ is ***complete*** *if every Cauchy sequence converges.*
A complete normed space $(X, \|\cdot\|)$ is called a ***Banach space****.*

Remark 1.4. *Let X be a normed vector space, $(x_n)_n$ a sequence in X.*

(a) If $x_n \to x$ in X, then $(x_n)_n$ is a Cauchy sequence.

Proof. Given $\varepsilon > 0 \exists N_\varepsilon : \forall n \geq N_\varepsilon : \|x - x_n\| < \frac{\varepsilon}{2}$. Hence for $n, m \geq N_\varepsilon$ we have

$$\|x_n - x_m\| = \|x_n - x + x - x_m\| \leq \|x_n - x\| + \|x - x_m\| < \varepsilon.$$

□

(b) Limits are unique!
If $x_n \to x$ in X and $x_n \to y$ in X, then $x = y$.

Proof.

$$\begin{aligned}\|x - y\| &= \|x - x_n + x_n - y\| \\ &\leq \|x_n - x\| + \|x_n - y\| \to 0 + 0 = 0 \quad \text{as } n \to \infty.\end{aligned}$$

□

(c) If $(x_n)_n$ converges or is Cauchy, then it is bounded, i.e.

$$\sup_{n\in\mathbb{N}} \|x_n\| < \infty.$$

Proof. Take $\varepsilon = 1$. Then there exists $N \in \mathbb{N}$ s.t. $\forall n, m \geq N : \|x_n - x_m\| < 1$. In particular, $\forall n \geq N : \|x_n - x_N\| < 1$.

$$\Rightarrow \|x_n\| = \|x_n - x_N + x_N\| \leq \|x_n - x_N\| + \|x_N\| < 1 + \|x_N\| \quad (\forall n \in \mathbb{N})$$

$$\Rightarrow \forall n \in N : \|x_n\| \leq max(\|x_1\|, \|x_2\|, \dots, \|x_N\|, 1 + \|x_N\|) < \infty.$$

□

Let X be a normed vector space, $S \neq \emptyset$ a set. For functions $f, g : S \to X$, $\alpha, \beta \in \mathbb{F}$ define

$$f+g : \begin{cases} S \to X \\ s \mapsto (f+g)(s) = f(s) + g(s) \end{cases}$$

$$\alpha f : \begin{cases} S \to X \\ s \mapsto (\alpha f)(s) = \alpha f(s) \end{cases}$$

So the set of functions from S to X is a normed space itself!

In case $X = \mathbb{R}$, we write $f \geq \alpha$ (or $f > \alpha$) if $f(s) \geq \alpha$ for all $s \in S$ ($f(s) > \alpha$ for all $s \in S$). Similarly one defines $\alpha \leq f \leq \beta$, $f \leq g$, etc.

Example 1.5. *(a) $X = \mathbb{F}^d, d \in \mathbb{N}$ is a Banach space (or short B-space) with respect to (w.r.t) the norms*

$$|x|_p := \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{\frac{1}{p}}, \quad 1 \leq p < \infty,$$

$$|x|_\infty := \max_{j=1,\ldots,d} |x_j|.$$

Here $x = (x_1, x_2, \ldots, x_d) \in \mathbb{F}^d$.

(b) Let $\Omega \neq \emptyset$, $L^\infty(\Omega) = L^\infty(\Omega, \mathbb{R}) =$ the set of all real-valued functions on Ω which are bounded, i.e.

$$f \in L^\infty(\Omega) \text{ then } \exists M_f < \infty : |f(\omega)| \leq M_f \text{ for all } \omega \in \Omega.$$

Norm on $L^\infty(\Omega)$: for $f \in L^\infty(\Omega) : \|f\|_\infty = \sup_{\omega\in\Omega} |f(\omega)|$ (check that this is a norm!).

Claim: $(L^\infty(\Omega), \|\cdot\|_\infty)$ is a Banach space.

Proof. Normed vector space is clear.

Take (f_n) a Cauchy sequence in $L^\infty(\Omega)$ w.r.t. $\|\cdot\|_\infty$. We have: $\forall \varepsilon > 0 \exists N : \forall n, m \geq N : \|f_n - f_m\|_\infty < \varepsilon$.

Fix $\omega \in \Omega$, then $\big(f_n(\omega)\big)$ is a Cauchy sequence in $\mathbb{R}$ since

$$|f_n(\omega) - f_m(\omega)| \leq \sup_{\omega\in\Omega} |f_n(\omega) - f_m(\omega)| = \|f_n - f_m\|_\infty < \varepsilon \quad \forall n, m \geq N.$$

Since $\mathbb{R}$ is complete, $f(\omega) := \lim_{n\to\infty} f_n(\omega)$ exists (this f is the candidate for the limit). We have

$$\begin{aligned} |f(\omega)| &\leq |f(\omega) - f_n(\omega)| + |f_n(\omega)| \\ &= \lim_{m\to\infty} |f_m(\omega) - f_n(\omega)| + |f_n(\omega)| \leq \varepsilon + \underbrace{|f_n(\omega)|}_{<\infty} \end{aligned}$$

$$\Rightarrow sup_{\omega\in\Omega} |f(\omega)| \leq \infty,$$

i.e., $f \in L^\infty(\Omega)$.

Take $\varepsilon > 0$. Then

$$|f_n(\omega) - f(\omega)| = \lim_{m\to\infty} \underbrace{|f_n(\omega) - f_m(\omega)|}_{\leq\varepsilon \text{ if } n,m\geq N} \leq \varepsilon \quad \text{if } n \geq N$$

$$\Rightarrow \forall n \geq N : \|f_n - f\|_\infty \leq \varepsilon, \text{ i.e., } f_n \to f \text{ w.r.t. } \|\cdot\|_\infty.$$

□

(c) *$X = C([0,1]), \|f\|_1 := \int_0^1 |f(t)|dt$ is a norm, $\Big(C([0,1]), \|\cdot\|_1\Big)$ is not complete.*

Proof.

$$\|f\|_1 \geq 0 \quad \forall f \in C([0,1]).$$

$$\|f+g\|_1 = \int_0^1 \underbrace{|f(t)+g(t)|}_{\leq |f(t)|+|g(t)|} dt \leq \|f\|_1 + \|g\|_1$$

$$\|\alpha f\|_1 = \int_0^1 |\alpha f(t)|dt = |\alpha|\|f\|_1$$

So $\|\cdot\|$ is a seminorm.
If $f \not\equiv 0$ and f is continuous, we see that there exist an interval $I \subset [0,1]$, $\delta > 0$ such that $|f(t)| \geq \delta \ \forall t \in I$.

$$\Rightarrow \|f\|_1 = \int_0^1 |f(t)|dt \geq \int_I \underbrace{|f(t)|}_{\geq \delta} dt \geq \delta.\text{lehgth of } I > 0.$$

So $\|\cdot\|_1$ is a seminorm. Now take a special sequence

$$f_n(t) := \begin{cases} 0, & \text{if } 0 \leq t \leq \frac{1}{2} - \frac{1}{n} \\ nt - \frac{n}{2} + 1, & \text{if } \frac{1}{2} - \frac{1}{n} < t < \frac{1}{2} \\ 1, \text{if } t \geq \frac{1}{2} \end{cases} \qquad (n \geq 3)$$

For $m \geq n \geq 3$:

$$\|f_n - f_m\|_1 = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{n}} |f_n(t) - f_m(t)|dt \leq \frac{1}{n} \to 0 \quad \text{as } n \to \infty,$$

so (f_n) is a Cauchy sequence.
Assume that $f_n \to f \in C([0,1])$. Fix $\alpha \in \big[0, \frac{1}{2}\big), n : \frac{1}{2} - \frac{1}{n} \geq \alpha$

$$0 \leq \int_0^\alpha |f(t)|dt = \int_0^\alpha |f_n(t) - f(t)|dt$$

$$\leq \int_0^1 |f_n(t) - f(t)|dt = \|f_n - f\|_1 \to 0.$$

Hence $f(t) = 0$ for all $0 \leq t \leq \alpha$, all $0 \leq \alpha < \frac{1}{2}$

$$f(t) = 0 \quad \text{for all} 0 \leq t < \frac{1}{2}.$$

On the other hand

$$0 \leq \int_{\frac{1}{2}}^{1} |f(t) - 1| dt = \int_{\frac{1}{2}}^{1} |f(t) - f_n(t)| dt \leq \|f - f_n\|_1 \to 0 \quad \text{as } n \to \infty$$

Since f is continuous on $[0, 1]$, it follows that $f(t) = 1$ for $\frac{1}{2} \leq t \leq 1$.
So f cannot be continuous at $t = \frac{1}{2}$. A contradiction. □

2 Basics of metric spaces

Definition 2.1. *Given a set $M \neq \emptyset$, a metric (or distance) d on M is a function $d : M \times M \to \mathbb{R}$ such that*

(a) $d(x, y) \geq 0 \; \forall x, y \in M$ and $d(x, y) = 0 \iff x = y$.

(b) $d(x, y) = d(y, x) \; \forall x, y \in M$ (symmetry).

(c) $d(x, y) \leq d(x, z) + d(z, y) \; \forall x, y, z \in M$ (triangle inequality).

The pair (M, d) is called a ***metric space****. We often simply write M if it is clear what d is.*
A sequence $(x_n)_n$ in a metric space (M, d) ***converges*** *to $x \in M$ if $\forall \varepsilon > 0 \exists N_\varepsilon \in \mathbb{N} \forall n \geq N_\varepsilon : d(x, x_n) < \varepsilon$ (or $\leq \varepsilon$). One writes $\lim x_n = x$ or $x_n \to x$.*

One always has

$$\big|d(x, z) - d(z, y)\big| \leq d(x, y)$$

Hint for the proof:

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

and think and use symmetry.

Example 2.2.

- *$\mathbb{R}$ with $d(x, y) = |x - y|$;*
- *Any normed vector space $(X, \|\cdot\|)$ with $d(x, y) = \|x - y\|$;*
- *Eucledian space $\mathbb{R}^d$ (or $\mathbb{C}^d$) with $d_2(x, y) = \big(\sum_{j=1}^{d} |x_j - y_j|^2\big)^{\frac{1}{2}}$ or $d_p(x, y) = \big(\sum_{j=1}^{d} |x_j - y_j|^p\big)^{\frac{1}{p}}$, or $d_\infty(x, y) = \max_{j=1,\dots d} |x_j - y_j|$.*
- *$M \neq \emptyset$, define $d : M \times M \to \mathbb{R}$ by*

$$d(x, y) := \begin{cases} 0, & \text{if } x = y \\ 1, & \text{else} \end{cases}$$

  is discrete metric. (M, d) is called discrete metric space.
- *$M = (0, \infty)$, $d(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|$ is a metric.*

- *Paris metric*

$$d(x,y) := \begin{cases} |x-y|, & \text{if } x = \lambda y \text{ for some } \lambda > 0, \\ |x| + |y|, & \text{else.} \end{cases}$$

- *If (M,d) is a metric space, $N \subset M$, then (N,d) is a metric space. Example: $M = \mathbb{R}^2, N = \{x : |x| = 1\}$.*

- *$M = \mathbb{F}^{\mathbb{N}}$ = set of all sequences $(a_n)_n, a_n \in \mathbb{F}$ = set of all functions $a : \mathbb{N} \to \mathbb{F}$ is a metric space with metric*

$$d(a,b) := \sum_{j=1}^{\infty} 2^{-j} \frac{|a(j)-b(j)|}{1+|a(j)-b(j)|}.$$

Proof. $d(a,b) \geq 0, d(a,b) = 0 \Rightarrow a = b, d(a,b) = d(b,a)$ are clear.
Need $d(a,b) \leq d(a,c) + d(c,b)$ for all sequences a,b,c.
Note: $0 \leq t \mapsto \frac{t}{1+t}$ is increasing!

$$\begin{aligned} d(a,b) &= \sum_{j=1}^{\infty} 2^{-j} \frac{|a(j)-b(j)|}{1+|a(j)-b(j)|} \\ &\leq \sum_{j=1}^{\infty} 2^{-j} \frac{|a(j)-c(j)|+|c(j)-b(j)|}{1+|a(j)-c(j)|+|c(j)-b(j)|} \\ &\leq \sum_{j=1}^{\infty} 2^{-j} \left(\frac{|a(j)-c(j)|}{1+|a(j)-c(j)|} + \frac{|c(j)-b(j)|}{1+|c(j)-b(j)|} \right), \end{aligned}$$

since, by the triangle inequality, $|a(j)-b(j)| \leq |a(j)-c(j)| + |c(j)-b(j)|$.
Note: $(a_n)_n \subset \mathbb{F}^{\mathbb{N}}, a_n \to a$ in $\mathbb{F}^{\mathbb{N}} \iff \forall j \in \mathbb{N} : a_n(j) \to a(j)$ and this space is complete!

$$\begin{aligned} 2^{-j} \frac{|a_n(j)-a(j)|}{1+|a_n(j)-a(j)|} &\leq d(a_n,a) \quad \text{for fixed } j \\ \Rightarrow |a_n(j)-a(j)| \leq & \underbrace{2^j d(a_n,a)}_{\leq \frac{1}{2} \text{ for } n \text{ large enough}} (1+|a_n(j)-a(j)|) \\ &\leq 2^j d(a_n,a) + \frac{1}{2}|a_n(j)-a(j)| \text{ for } n \text{ large enough} \\ \Rightarrow \text{ for } n \text{ large enough:} &|a_n(j)-a(j)| \leq 2^{j+1} d(a_n,a) \to 0 \text{ as } n \to \infty, \end{aligned}$$

so $a_n \to a$ in $\mathbb{F}^{\mathbb{N}} \Rightarrow \forall j \in \mathbb{N} : a_n(j) \to a(j)$.
Need $\Leftarrow$:

$$\begin{aligned} d(a_n,a) &= \sum_{j=1}^{\infty} 2^{-j} \frac{|a_n(j)-a(j)|}{1+|a_n(j)-a(j)|} \\ &\leq \underbrace{\sum_{j=1}^{L} 2^{-j}|a_n(j)-a(j)|}_{\leq L \max_{j=1,\dots,L} |a_n(j)-a(j)|} + \underbrace{\sum_{L+1}^{\infty} 2^{-j}}_{< \frac{\varepsilon}{2} \text{ by choosing} L \text{ large enough}} \end{aligned}$$

$\square$

Definition 2.3. *Let (M, d) be a metric space.*

- *The* ***open ball*** *at x with radius $r > 0$: $B_r(x) := \{y \in M : d(x, y) < r\}$.*
- *$A \subset M$ is* ***open*** *if $\forall x \in A\ \exists r > 0$ with $B_r(x) \subset A$.*

Note: Every open ball is itself an open set! Indeed, $y \in B_r(x), r_1 := r - d(x, y) \Rightarrow B_{r_1}(x) \subset B_r(x)$ since, if $z \in B_{r_1}(y)$ then

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + r_1 = r$$

so $z \in B_r(x)$.

Theorem 2.4. *(a) M and $\emptyset$ are open.*

(b) An arbitrary union of open sets is open.

(c) Finite intersections of open sets are open.

Proof. (a) Clear.

(b) Take $(A_j)_{j \in J}, A_j \subset M$ open.

$$x \in \bigcup_{j \in J} A_j = \{y \in M : \exists j \in J \text{ with } y \in A_j\} \Rightarrow \exists j \in J : x \in A_j.$$

Since A_j is open, there exists $r > 0$ with $B_r(x) \subset A_j \subset \bigcup_{j \in J} A_j$. Hence $\bigcup_{j \in J} A_j$ is open.

(c) Take $\{A_1, \ldots A_n\}$ open sets in M

$$x \in A := \bigcap_{j=1}^{n} A_j = \{y \in M : y \in A_j \text{ for all } j = 1, \ldots n\}$$

A_j open $\Rightarrow \exists r_j > 0 : B_{r_j}(x) \subset A_j, j = 1, \ldots n$. Let $r := min(r_1, r_2, \ldots, r_n) > 0$. Then

$$B_r(x) \subset B_{r_j}(x) \subset A_j \text{ for all } j = 1, \ldots n$$

$$\Rightarrow B_r(x) \subset \bigcap_{j=1}^{n} A_j$$

□

Definition 2.5. *(a) $x \in A$ is called an* ***interior point*** *of A if $\exists r > 0 : B_r(x) \subset A$. The set of all interior points is denoted by A^o.*
Note:

- *A^o is the largest open subset of M contained in A.*
- *A is open $\iff A = A^o$.*

(b) $A \subset M$ is ***closed*** *if its complement $A^c := M \setminus A = \{x \in M : x \notin A\}$ is open;*

Theorem 2.6. *(a) M and $\emptyset$ are closed.*

(b) Arbitrary intersections of closed sets are closed.

(c) Finite unions of closed sets are closed.

Proof. (a) $M^c = \emptyset$, $\emptyset^c = M$ are open.

(b) $(A_j)_{j\in J}$ family of closed sets. By Theorem 2.4 and de Morgan's law

$$\Big(\bigcap_{j\in J} A_j\Big)^c = \bigcup_{j\in J} A_j^c \text{ is open,}$$

so $\bigcap_{j\in J} A_j$ is closed;

(c) Combine $\left(\bigcup_{j=1}^n A_j\right)^c = \bigcap_{j=1}^n A_j^c$ with (c) of Theorem 2.4.

□

Definition 2.7. *A point $x \in M$ is called* ***closure point*** *of $A \subset M$ if $\forall r > 0 :$ $B_r(x) \cap A \neq \emptyset$. The set of all closure points of A is denoted by $\overline{A}$ and it is called the* ***closure*** *of A.*
Clearly $A \subset \overline{A}$.

Theorem 2.8. *Let (M, d) be a metric space, $A \subset M$. Then $\overline{A}$ is the smallest closed set that contains A.*

Remark 2.9. *Let $\mathcal{F}_A := \{B \subset M : B$ is closed and $A \subset B\}$. Then the smallest closed subset of M that contains A is, of course, given by $\bigcap_{B\in\mathcal{F}_A} B$. (think about this!)*

Proof of Theorem 2.8. Let $A \subset M$.
Step 1: $\overline{A}$ is closed. Indeed, if $x \in (\overline{A})^c$, then $\exists r > 0$ with $B_r(x) \cap A = \emptyset$. We want to show that $B_r(x) \subset (\overline{A})^c$, because then $(\overline{A})^c$ is open, hence $\overline{A}$ is closed. Let $y \in B_r(x)$. Since $B_r(x)$ is open, there exists $\delta > 0$ with $B_\delta(y) \subset B_r(x)$

$$\Rightarrow B_\delta(y) \cap A \subset B_r(x) \cap A = \emptyset$$

$\Rightarrow y \notin \overline{A}$ and since $y \in B_r(x)$ was arbitrary, this shows

$$B_r(x) \cap \overline{A} = \emptyset$$

so $B_r(x) \subset (\overline{A})^c$, hence $(\overline{A})^c$ is open.
Step 2: Let $B \subset M$ be closed with $A \subset B$. We show $\overline{A} \subset B$. Indeed, take $x \in B^c$. Since B^c is open, there exists $r > 0$ with $B_r(x) \subset B^c$, i.e., $B_r(x) \cap B = \emptyset$. In particular, $B_r(x) \cap A \subset B_r(x) \cap B = \emptyset$. So no point in B^c is a closure point of $A \Rightarrow \overline{A} \subset (B^c)^c = B$. □

Corollary 2.10. *$A \subset M$ is closed $\Rightarrow A = \overline{A}$.*

Proof. Have a close look at Theorem 2.8. □

Remark 2.11.

- *For $a \in M$ and $r > 0$ call*

$$B_{\overline{r}}(a) := \{x \in M : d(x,a) \leq r\}$$

the closed ball at a with radius r. This is always a closed set. Indeed, assume $x \notin B_{\overline{r}}(a)$, i.e., $d(x,a) > r$ and set $r_1 := d(x,a) - r > 0$. If $y \in B_{r_1}(x)$, then

$$d(a,x) \leq d(a,y) + d(y,x)$$

$$\Leftrightarrow d(a,y) \geq d(a,x) - d(y,x) > d(a,x) - r_1 = r,$$

i.e., $y \notin B_{\overline{r}}(a)$, hence $B_{r_1}(x) \subset (B_{\overline{r}}(a))^c$ so $(B_{\overline{r}}(a))^c$ is open $\Leftrightarrow B_{\overline{r}}(a)$ is closed.

- *One always has $\overline{B_r(a)} \subset B_{\overline{r}}(a)$. In a discrete metric space the above inclusion can be strict! But, e.g., in $\mathbb{R}^d$ with the distance $d_p, 1 \leq p \leq \infty$, one always has $\overline{B_r(a)} = B_{\overline{r}}(a)$. (think about this!)*

Lemma 2.12. *If (M,d) is a metric space, then $A^o = (\overline{A^c})^c$.*

Proof.

$$\begin{aligned} x \in A^o &\Leftrightarrow \exists r > 0 : B_r(x) \subset A \\ &\Leftrightarrow B_r(x) \cap A^c = \emptyset \\ &\Leftrightarrow x \notin \overline{A^c} \\ &\Leftrightarrow x \in (\overline{A^c})^c. \end{aligned}$$

□

Definition 2.13. *Let (M,d) be a metric space, $A \subset M$. A point $x \in M$ is an* ***accumulation point*** *of A if*

$$\forall r > 0 \quad B_r(x) \cap (A \setminus \{x\}) \neq \emptyset,$$

i.e., every open ball around x contains an element of A different from x.

Note:

- It can be that $x \notin A$!
- Every accumulation point is a closure point of A.
- If one denotes the set of all accumulation points of A by A', then one has $\overline{A} = A \cup A'$ (why?).

Theorem 2.14. *Let $A \subset M$, (M,d) a metric space. Then $x \in M$ belongs to $\overline{A}$ if and only if (iff) there is a sequence $(x_n)_n \subset A$ with $\lim x_n = x$. Moreover, if x is an accumulation point of A, then there exists a sequence $(x_n)_n \subset A$ with $x \neq x_n \neq x_m, n \neq m$, i.e., all terms are distinct.*

Proof. Let $x \in \overline{A}$. Given $n \in \mathbb{N}$ pick x_n with $x_n \in B_r(x) \cap A (\neq \emptyset$ since $x \in \overline{A}$!). Then $x_n \in A$ and $\lim x_n = x$.
Conversely, if $x_n \in A$ and $\lim x_n = x$, then given $r > 0$ there exists $k \in \mathbb{N}$ such that $d(x, x_n) < r$ for all $n \geq k$. Therefore $B_r(x) \cap A \neq \emptyset$ for all $r > 0 \Rightarrow x \in \overline{A}$.
If x is an accumulation point of A, choose $x_1 \in A, x_1 \neq x$ and $d(x, x_1) < 1$. Then, inductively, if $x_1, \ldots x_n \in A \setminus \{x\}$ pich $x_{n+1} \in A \setminus \{x\}$ with

$$d(x, x_{n+1}) < \min\Big(\frac{1}{n+1}, d(x, x_n)\Big).$$

Thus $(x_n)_n$ is a sequence in $A \setminus \{x\}$, $x_n \neq x_m$ if $n \neq m$ and $\lim x_n = x$. □

Definition 2.15. *$A \subset M$ is **dense** in M if $\overline{A} = M$.*

Remark 2.16.
- *By Theorem 2.14, A is dense in M iff $\forall x \in M, \exists$ sequence $(x_n)_n \subset A$ with $\lim x_n = x$.*
- *A is dense in $M \Leftrightarrow V \cap A \neq \emptyset$ for every nonempty open set V.*

Definition 2.17. *Let $A \subset M$. $x \in M$ is a **boundary point** of A if $\forall r > 0$: $B_r(x) \cap A \neq \emptyset \neq B_r(x) \cap A^c$. The set of all boundary points of A is denoted by ∂A and it is called **boundary** of A.*

Note:

- By symmetry, $\partial A = \partial(A^c)$.
- $\partial A = \overline{A} \cap \overline{A^c}$ (Why?)

Definition 2.18 (Continuity)**.** *Let $(M, d), (N, \rho)$ be two metric spaces. A function $f : M \to N$ is*

- *continuous at a point $a \in M$ if $\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0$ with $\rho\big(f(x), f(a)\big) < \varepsilon$ for all $d(x, a) < \delta$.*
- *continuous on M (or simply continuous) if f is continuous at every point of M.*
- *sequentially continuous at a point $a \in M$ if for every sequence $(x_n)_n \subset M, x_n \to a$ one has $f(x_n) \to f(a)$.*
- *sequenctially continuous on M (or simply sequentially continuous) if it is sequentially continuous at every point of M.*
- *topologically continuous if for every open set $\mathcal{O}$ the set $f^{-1}(\mathcal{O}) \subset M$ is open.*

Theorem 2.19. *For a function $f : (M, d) \to (N, \rho)$ between two metric spaces, the following are equivalent:*

(a) f is continuous on M.

(b) f is topologically continuous on M.

(c) f is sequentially continuous on M.

(d) $f(\overline{A}) \subset \overline{f(A)}$ for every $A \subset M$.

(e) $f^{-1}(\mathcal{C}) \subset M$ *is closed for every closed subset* $\mathcal{C} \subset N$.

Remark 2.20. *For a fixed* $a \in M$*, the following are also equaivalent:*

(a') f is continuous at a.

(c') f is sequentially continuous at a.

(Prove this!)

Proof of Theorem 2.19. $\underline{(a) \Rightarrow (b)}$: Let $\mathcal{O} \subset N$ be open and $a \in f^{-1}(\mathcal{O})$. Since $f(a) \in \mathcal{O}$ and $\mathcal{O}$ is open, there exists $r > 0$ such that $B_r(f(a)) \subset \mathcal{O} \subset N$. f continuous implies that there exists $\delta > 0$ such that

$$d(x,a) < \delta \Rightarrow \rho(f(x), f(a)) < r,$$

i.e., $B_\delta(a) \subset f^{-1}(\mathcal{O})$ and so $f^{-1}(\mathcal{O})$ is open.
$\underline{(b) \Rightarrow (c)}$: Let $x_n \to x$ in M and $\varepsilon > 0$. Let $V := B_\varepsilon(f(x)) \subset N$, which is open. Then $f^{-1}(V)$ is open in M and since $x \in f^{-1}(V)$ there exists $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(V)$. Let $N \in \mathbb{N}$ be such that $n \geq N \Rightarrow x_n \in B_\delta(x)$ (i.e., $d(x_n, x) < \delta$ for all $n \geq N$), Then also $x_n \in f^{-1}(V)$, so $f(x_n) \in V$, i.e., $\rho(f(x_n), f(x)) < \varepsilon$ for all $n \geq N$. Thus $f(x_n) \to f(x)$.
$\underline{(c) \Rightarrow (d)}$: Let $A \subset M$. Assume $y \in f(\bar{A})$. Then there exists $x \in \bar{A}$ with $f(x) = y$. Since $x \in \overline{A}$, by Theorem 2.14, it follows that there exists a sequence $(x_n)_n \subset A$ with $x_n \to x$, but then by (c): $f(x_n) \to f(x)$ in N, i.e., $y \in \overline{f(A)}$. So $f(\bar{A}) \subset \overline{f(A)}$.
$\underline{(d) \Rightarrow (e)}$: Let $\mathcal{C} \subset N$ be closed, so $\bar{\mathcal{C}} = \mathcal{C}$. Let $A := f^{-1}(\mathcal{C})$. Then by (d) we have

$$f(\overline{A}) \subset \overline{f(A)} = \bar{\mathcal{C}} = \mathcal{C},$$

so $\bar{A} \subset f^{-1}(\mathcal{C}) = A$. Since $A \subset \bar{A}$ is always true, we must have $f^{-1}(\mathcal{C}) = A = \bar{A}$, i.e., $f^{-1}(\mathcal{C})$ is closed.
$\underline{(e) \Rightarrow (a)}$: Let $a \in M$ and $\varepsilon > 0$. Consider

$$\mathcal{C} := B_\varepsilon\big(f(a)\big)^c = \{y \in N : \rho(f(a), y) \geq \varepsilon\}$$

which is closed. By (e) $f^{-1}(\mathcal{C}) \subset M$ is closed, i.e., $\Big(f^{-1}(\mathcal{C})\Big)^c$ is open. Thus, since $a \notin f^{-1}(\mathcal{C})$, i.e., $a \in \Big(f^{-1}(\mathcal{C})\Big)^c$, there exists $\delta > 0$ such that $B_\delta(a) \subset \Big(f^{-1}(\mathcal{C})\Big)^c$. But then $d(x,a) < \delta \Rightarrow \rho(f(x), f(a)) < \varepsilon$, i.e., f is continuous. □

Remark 2.21. *It should be clear that compositions of continuous functions are continuous.*

Definition 2.22.
- *Two metric spaces* $(M,d), (N,\rho)$ *are* ***homeomorphic*** *if* $\exists$ *a one-to-one onto function (i.e., bijection)* $f : (M,d) \to (N,\rho)$ *such that both* f *and* f^{-1} *are continuous;*
- *Two metrics* d *and* ρ *on* M *are* ***equaivalent*** *if a sequence* $(x_n)_n \subset M$ *satisfies*

$$\lim d(x_n, x) = 0 \iff \lim \rho(x_n, x) = 0,$$

or equivalently, if any open set w.r.t. d *is open w.r.t.* ρ *and conversely.*

- *A metric space M is **bounded**, if $\exists 0 < M < \infty$ s.t. $d(x,y) \leq M \forall x,y \in M$. The **diameter** of $A \subset M$ is*

$$d(A) := \sup(d(x,y) : x,y \in A).$$

Note: If d is a metric on M

$$\rho(x,y) := \frac{d(x,y)}{1+d(x,y)}$$

is an equaivalent metric on M under which M is bounded!

- *A sequence $(x_n)_n$ in a metric space (M,d) is a **Cauchy sequence** if $\forall \varepsilon > 0 \exists N_\varepsilon \in \mathbb{N} : d(x_n, x_m) < \varepsilon$ for all $n,m \geq N_\varepsilon$.*

  Note: Every convergent sequence $(x_n)_n$ is a Cauchy sequence (Why?). The converse is not true, e.g. take $M = (0,\infty), d(x,y) = |x-y|$. Then $x_n = \frac{1}{n}$ is Cauchy but not convergent in M.

- *A metric space (M,d) is **complete** (or complete metric space) if every Cauchy sequence converges (in M).*

Example 2.23.
- $\mathbb{R}^d$ *with Eucledean metric or with $d_p, 1 \leq p \leq \infty$.*
- $L^\infty(S), S \neq \emptyset$, $D(f,g) := \sup_{x \in S} |f(s) - g(s)|$.

Theorem 2.24. *Let (M,d) be a complete metric space. Then $A \subset M$ is closed if and only if (A,d) is a complete metric space (in its own right).*

Proof. "$\Rightarrow$": Let $A \subset M$ be closed, $(x_n)_n \subset A$ be Cauchy $\Rightarrow$ $(x_n)_n$ is Cauchy in M. Since M is complete, it follows that $x = \lim_{n\to\infty} x_n$ exists in M. Since A is closed, we conclude that $x \in A$. So $(x_n)_n$ converges in A and thus (A,d) is complete.

"$\Leftarrow$": Let (A,d) be complete. Let $(x_n)_n \subset A$ converge to some $x \in M$. So $(x_n)_n$ is Cauchy in A, A is complete $\Rightarrow$ $(x_n)_n$ converges to some point in $A \subset M$. The limit is unique so $x = \lim_{n\to\infty} x_n \in A$. So A is closed. ☐

Lemma 2.25. *Let (M,d) be a metric space and $(x_n)_n, (y_n)_n \subset M$ s.t. $x_n \to x, y_n \to y$. Then*

$$\lim_{n\to\infty} d(x_n, y_n) = d(x,y).$$

Proof. By the triangle inequality one has

$$|d(x,z) - d(z,y)| \leq d(x,y)$$

$$\begin{aligned} \Rightarrow |d(x_n,y_n) - d(x,y)| &\leq |d(x_n,y_n) - d(x,y_n)| + |d(x,y_n) - d(x,y)| \\ &\leq d(x_n,x) + d(y_n,y) \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

☐

Definition 2.26. *A function $f : (M,d) \to (N,\rho)$ is called **uniformly continuous** if $\forall \varepsilon > 0 \exists \delta > 0 : x,y \in M, d(x,y) < \delta (or \leq \delta) \Rightarrow \rho(f(x),f(y)) < \varepsilon (or \leq \varepsilon)$.*

Remark 2.27.
- *Every uniformly continuous function is continuous.*
- *$M = (0,1], N = \mathbb{R}, d(x,y) = |x-y|, d : (0,1] \to \mathbb{R}, x \mapsto f(x) = x^2$ is uniformly continuous, $g : (0,1] \to \mathbb{R}, x \mapsto g(x) = \frac{1}{x}$ is continuous but not uniformly continuous.*

Theorem 2.28. *Let A be a subset of a metric space (M,d), (N,ρ) be a complete metric space. If $f : A \to N$ is uniformly continuous, then f has a unique uniformly continuous extension to the closure $\overline{A}$ of A.*

Remark 2.29. *This does not hold if f is only continuous!*
Example: 1) $f : (0,1] \to \mathbb{R}, x \to \frac{1}{x}$.
2) $f : \mathbb{Q} \to \mathbb{R}, x \mapsto \begin{cases} 1, & \text{if } x^2 \geq 2 \\ -1, & \text{if } x^2 < 2 \end{cases}$ is a continuous function in $\mathbb{Q}$!
Note that f also is differentiable on $\mathbb{Q}$ with zero derivative!
Look at $g(x) = x + 4f(x), x \in \mathbb{Q} \Rightarrow g'(x) = 1$. So g ¨must¨ be increasing! (?). But

$$g(-2) = 2,$$
$$g(0) = -4.$$

So g is not increasing!

Proof of Theorem 2.28. Step 1: Uniqueness should be clear (why?).
Step 2: Let $x \in \overline{A}$. By Theorem 2.14, there exists a sequence $(x_n)_n \subset A$ with $x_n \to x$.
Claim: $\lim_{n\to\infty} f(x_n)$ exists in (N,ρ)!
(N,ρ) is complete $\Rightarrow$ we only need to show that $(f(x_n))$ is Cauchy in (N,ρ). Let $\varepsilon > 0$. Since f is uniformly continuous, $\exists \delta > 0 : d(x,y) < \delta \Rightarrow \rho(f(x), f(y)) < \varepsilon$. So let $N_\varepsilon \in \mathbb{N}$ be such that $d(x_n, x_m) < \delta$ for all $n, m \geq N_\varepsilon \Rightarrow \rho(f(x_n), f(x_m)) < \varepsilon$ for all $n, m \geq N_\varepsilon$.
Step 3: The limit $\lim_{n\to\infty} f(x_n)$ in Step 2 is independent of the sequence as long as $x_n \to x$. Indeed, let $(x_n)_n, (y_n)_n \subset A, x_n \to x, y_n \to x$ in M. By Step 2 we know that $u = \lim f(x_n), v = \lim f(y_n)$ exist in N. We want to show $u = v$. For $n \in \mathbb{N}$, let $z_{2n} = x_n, z_{2n-1} = y_n \Rightarrow z_n \to x$ also, and, by Step 2: $\lim f(z_n)$ exists. We have

$$v = \lim f(y_n) = \lim f(z_{2n-1}) = \lim f(z_n) = \lim f(z_{2n}) = \lim f(x_n) = u.$$

Step 4: Define $f^* = \lim f(x_n), x_n \in A, x_n \to x$ (well defined by Steps 2&3). Of course $f^*(x) = f(x), x \in A$ is an extension of A to $\overline{A}$.
Step 5: $f^* : \overline{A} \to N$ is uniformly continuous. Indeed, given $\varepsilon > 0$, let $\delta > 0$ such that $x, y \in A, d(x,y) < \delta \Rightarrow \rho(f(x), f(y)) < \varepsilon$. Now if $x, y \in \overline{A}$ satisfy $d(x,y) < \delta$, let $(x_n)_n, (y_n)_n \subset A, x_n \to x, y_n \to y$. By Lemma 2.25

$$\lim d(x_n, y_n) = d(x,y) < \delta \Rightarrow \exists N_0 \in \mathbb{N} : d(x_n, y_n) < \delta \quad \text{or all } n \geq N_0.$$

Since f is uniformly continuous

$$\rho(f(x_n), f(y_n)) < \varepsilon$$

By Lemma 2.25

$$\rho(f(x), f(y)) = \lim \rho(f(x), f(y)) \leq \varepsilon$$

so f^* is uniformly continuous. □

Definition 2.30.

- *A function $f : (M,d) \to (N,\rho)$ is an* ***isometry*** *if $\rho(f(x), f(y)) = d(x,y)$ for all $x, y \in M$. If f is also onto, then (M,d) and (N,ρ) are isometric.*

  Note: any isometry is uniformly continuous!

- *A complete metric space (N,ρ) is called a* ***completion*** *of a metric space (M,d) if there exists an isometry $f : (M,d) \to (N,\rho)$ such that $f(M) = \{y \in N : \exists x \in M : y = f(x)\}$ is dense in N (w.r.t ρ).*
  If we think of M and $f(M)$ as identical, then M can be considered to be a subset of N.

Remark 2.31. *Any two completions of a metric space (M,d) must be isometric.*

Proof. Indeed, if N_1, N_2 are completions of M:

$$N_2 \supset \text{dense } g(M) \xleftarrow{g} M \xrightarrow{f} f(M) \text{ dense} \subset N_1$$

f, g are isometries. Define $h := g \circ f^{-1} : f(M) \to f(M)$. h is also an isometry (so, it is uniformly continuous). $f(M)$ is dense in N_1, N_2 is complete, so by Theorem 2.23 h has a unique uniformly continuous extension $\tilde{h} : N_1 \to N_2$.
Note: $\tilde{h}$ is an isometry from N_1 onto N_2! (Why?) (use that $g(M)$ is dense in N_2). ☐

Our approach to completeness: Given a metric space (M,d), find a complete metric space (N,ρ) and an isometry $f : (M,d) \to (N,\rho)$. $f(M)$ is then isometric to M. Take the closure $\overline{f(M)}$ in N. Then $(\overline{f(M)}, \rho) \subset (N,\rho)$ is a completion of (M,d)!

Theorem 2.32. *Every metric space (M,d) has a unique (up to isometries) completion.*

Proof. Goal: Embedd M in a complete metric space and take the closure!
We will use $(L^\infty(M), D)$, the bounded real-valued functions on M with

$$D(f,g) := \sup_{x \in M} |f(x) - g(x)|.$$

Fix $a \in M$. For $x \in M$ let

$$f_x : \begin{cases} M \to \mathbb{R}, \\ y \mapsto f_x(y) := d(x,y) - d(y,a). \end{cases}$$

By the reverse triangle inequality:

$$|f_x(y)| = |d(x,y) - d(y,a)| \leq d(x,a)$$

So $f_x \in L^\infty(M)$. Hence there exists a unique

$$f : \begin{cases} M \to L^\infty(M), \\ x \mapsto f_x. \end{cases}$$

Claim: f is an isometry!
Indeed, for $x, y, z \in M$:

$$\begin{aligned} |f_x(y) - f_z(y)| &= |d(x,y) - d(y,a) - (d(z,y) - d(y,a))| \\ &= |d(x,y) - d(z,y)| \leq d(x,z). \end{aligned}$$

$$\Rightarrow D(f_x, f_y) = \sup_{y \in M} |f_x(y) - f_z(y)| \le d(x, z).$$

Choose $y = z$:

$$|f_x(z) - f_z(z)| = |d(x, z) - d(z, z)| = d(x, z),$$

so

$$D(f_x, f_z) = d(x, z).$$

Since $(L^\infty(M), D)$ is a complete metric space $\Rightarrow (\overline{f(M)}, D)$ is a completion of (M, d). □

3 Compactness in metric space

In the following let (M, d) be a metric space.

Definition 3.1 (Totally bounded set). *A subset $A \subset M$ is totally bounded if $\forall \varepsilon > 0 \ \exists n \in \mathbb{N} : x_1, \dots, x_n \in M$ with $A \subset \bigcup_{i=1}^n B_\varepsilon(x_i)$ (so each $x \in A$ is within ε-distance from some x_i).*

Remark 3.2. *(a) Every $x \in A$ can be approximated up to error ε by one of the x_i.*

(b) In a finite dimensional (vector) space totally bounded is equivalent to bounded. In general totally bounded $\Rightarrow$ bounded, but the converse in wrong!

(c) In Definition 3.1 we could easily insist that each ε-ball is centered at some point in A. Indeed, let $\varepsilon > 0$, choose $x_1, \dots, x_n \in M$.

$$A \subset \bigcup_{i=1}^n B_{\frac{\varepsilon}{2}}(x_i).$$

W.l.o.g., we may assume that $B_{\frac{\varepsilon}{2}}(x_i) \cap A \neq \emptyset$. Then choose any $y_i \in A \cap B_{\frac{\varepsilon}{2}}(x_i)$. By the triangle inequality: $B_{\frac{\varepsilon}{2}}(x_i) \subset B_\varepsilon(y_i) \Rightarrow A \subset \bigcup_{i=1}^n B_\varepsilon(y_i)$.

Lemma 3.3. *$A \subset M$ is totally bounded $\Leftrightarrow \forall \varepsilon > 0$ there exist finitely many sets $A_1, \dots A_n$ with $diam(A_i) < \varepsilon$ for all $i = 1, \dots n$ and $A \subset \bigcup_{i=1}^n A_i$.*

Proof. "$\Rightarrow$": Let A be totally bounded. Given $\varepsilon > 0$ choose $x_1, \dots x_n \in M$ with $A \subset \bigcup_{i=1}^n B_\varepsilon(x_i)$. Let $A_i := A \cap B_\varepsilon(x_i)$ to see that $\bigcup_{i=1}^n A_i = \bigcup_{i=1}^n A \cap B_\varepsilon(x_i) = A \cap (\bigcup_{i=1}^n B_\varepsilon(x_i) = A$ and note that $diam(A_i) < 2\varepsilon$.
"$\Leftarrow$": Given $\varepsilon > 0$ assume that there are finitely many $A_i \subset A, i = 1, \dots, n$, $diam(A_i) < \varepsilon, A \subset \bigcup_{i=1}^n A_i$. Then choose any $x_i \in A_i \Rightarrow A_i \subset B_{2\varepsilon}(x_i)(\forall i = 1 \dots n) \Rightarrow A \subset \bigcup_{i=1}^n B_{2\varepsilon}(x_i)$. □

Remark 3.4. *In Lemma 3.3 we insisted on $A_i \subset A (\forall i = 1 \dots n)$. This is not a real constraint. If A is covered by $B_1, \dots, B_n \subset M$, $diam(B_i) < \varepsilon$. Then A is also covered by $A_i = A \cap B_i \subset A$ and $diam(A_i) \le diam(B_i) < \varepsilon$.*

There is also a sequential criterion for total boundedness. The Key observation is

Lemma 3.5. *Let $(x_n)_n \subset M, A = \{x_n : n \in \mathbb{N}\}$. Then*

(a) If $(x_n)_n$ is Cauchy, then A is totally bounded.

(b) If A is totally bounded, then $(x_n)_n$ has a Cauchy subsequence.

Proof. (a) Let $\varepsilon > 0$. Since $(x_n)_n$ is Cauchy, there exists $N \in \mathbb{N}$ with

$$d(x_n, x_m) < \frac{\varepsilon}{2} \quad \text{for all } n, m \geq N$$

$$\Rightarrow \sup_{n,m \geq N} d(x_n, x_m) \leq \frac{\varepsilon}{2} < \varepsilon$$

$$\Rightarrow diam\{x_n : n \geq N\} = \sup_{n,m \geq N} d(x_n, x_m) < \varepsilon$$

$$\Rightarrow \{x_n : n \geq N\} \subset B_\varepsilon(x_N).$$

(b) If A is finite, we are done because by pidgeonholing, there must be a point in A which the sequence $(x_n)_n$ hits infinitely often. Thus $(x_n)_n$ even has a constant subsequence in this case.
So assume that A is an infinite totally bounded set. Then A can be covered by finitely many sets of diameter < 1. At least one of them must contain infinitely many points of A. Call this set A_1. Note that A_1 is totally bounded, so it can itself be covered by finitely many sets of diameter $< \frac{1}{2}$. One of these, call it A_2, contains infinitely many points of A_1. Continuing inductively we find a decreasing sequence of sets $A \supset A_1 \supset A_2 \supset \cdots \supset A_n \supset A_{n+1} \supset \ldots$ where each A_k contains infinitely many x_n and where $diam(A_k) < \frac{1}{k}$.
Now choose a subsequence $(x_{n_k})_k, x_{n_k} \in A_k, k \in \mathbb{N}$. This subsequence is Cauchy, since

$$\sup(d(x_{n_l}, x_{n_m})l, m \geq k) \leq diam(A_k) < \frac{1}{k}.$$

□

Theorem 3.6 (Sequential characterization of total boundedness)**.** *A set $A \subset M$ is totally bounded $\iff$ every sequence in A has a Cauchy subsequence.*

Proof. "⇒": Clear by Lemma 3.5.
"⇐": Assume A is not totally bounded. So for some $\varepsilon > 0$, A cannot be covered by finitely many ε-balls. By induction, there is a sequence $(x_n)_n \subset A$ with $d(x_n, x_m) \geq \varepsilon$ for all $n \neq m$ (Why?). But this sequence has no Cauchy subsequence! □

Corollary 3.7 (Bolzano-Weierstraß)**.** *Every bounded infinite subset of $\mathbb{R}^d$ has an accumulation point.*

Proof. Let $A \subset \mathbb{R}^d$ be bounded and infinite. Then there is a sequence $(x_n)_n$ of distinct points in A. Since A is totally bounded ($\mathbb{R}^d$ has dimension $d < \infty$) there is a Cauchy subsequence of $(x_n)_n$, but $\mathbb{R}^d$ is complete, so $(x_n)_n$ converges to some $x \in \mathbb{R}^d$. This x is an accumulation point of A. □

Now we come to compactness.

Definition 3.8. • *A metric space (M, d) is compact if it is complete and totally bounded.*

- *A subset $A \subset M$ is compact, if (A, d) is a compact metric space.*

Example 3.9. *(a) $K \subset \mathbb{R}^d$ is compact $\iff$ K is closed and bounded.*

(b) Let l^∞ = set of all bounded sequences and let

$$e_n := \delta_n, \quad \delta_n(j) := \begin{cases} 1, & \text{if } j = n, \\ 0, & \text{else.} \end{cases}$$

Then the set $A := \{e_n | n \in \mathbb{N}\}$ is closed and bounded, but not totally bounded, since

$$d(e_n, e_m) = \sup_{j \in \mathbb{N}} |e_n(j) - e_m(j)| = 1, \quad \text{if } n \neq m,$$

hence, A cannot be covered by finitely many $\varepsilon = \frac{1}{2}$-balls! (Why?)

(c) A subset of a discrete metric space is compact $\iff$ A is finite. (Why?)

The sequential characterization of compactness is given by

Theorem 3.10. *(M, d) is compact $\iff$ every sequence in M has a convergent subsequence in M.*

Proof. By Lemma 3.5 and the definition of completenes:

$$\left\{ \begin{array}{c} \text{totally bounded} \\ + \\ \text{complete} \end{array} \right\} \iff \left\{ \begin{array}{c} \text{every sequence in } M \\ \text{has a Cauchy subsequence} \\ + \\ \text{Cauchy sequences converge} \end{array} \right\}$$

□

Compactness is an extremely useful property to have: if you happen to have a sequence in a compact space which does not converge, simply extract a convergent subsequence and use this one instead!

Corollary 3.11. *Let A be a subset of a metric space M. If A is compact, then A is closed in M (and totally bounded). If M is compact and A is closed, then A is compact.*

Proof. Assume that A is compact and let $x \in M$ and $(x_n)_n \subset A$ with $x_n \to x$. By Theorem 3.10, $(x_n)_n$ has a convergent subsequence whose limit is also in $A \Rightarrow x \in A$ so A is closed.
Assume M is compact, $A \subset M$ is closed. Given $(x_n)_n \subset A$, Theorem 3.10 supplies a convergent subsequence of $(x_n)_n$ which converges to a point $x \in M$. Since A is closed, we must have $x \in A$, so by Theorem 3.10 again, A is compact. □

Corollary 3.12. *Let (M,d) be compact and $f : M \to \mathbb{R}$ continuous. Then f attains its maximum and minimum, i.e., there are $x_{\min}, x_{\max} \in M$ such that*

$$f(x_{\min}) = \inf(f(x)|x \in M),$$
$$f(x_{\max}) = \sup(f(x)|x \in M),$$

In particular, inf *and* sup *are finite!*

Proof. Only for mininmum (otherwise look at $-f$).
Let $a := \inf(f(x)|x \in M)$. Note that there is always a minimizing sequence, i.e., a sequence $(x_n)_n \subset M$ such that

$$f(x_n) \to a \quad \text{as } n \to \infty.$$

Now if $(x_n)_n$ converges to some point $x \in M$, then we are done, since by continuity of f,

$$f(x) = \lim_{n\to\infty} f(x_n) = a = \inf(f(x)|x \in M).$$

If $(x_n)_n$ does not converge, use the fact that M is compact, so by Theorem 3.10 $(x_n)_n$ has a convergent subsequence and then use this subsequence instead! □

Corollary 3.13. *Let (N,ρ) be a metric space. If (M,d) is compact and $f : (M,d) \to (N,\rho)$ is continuous, then f is uniformly continuous.*

Proof. Recall the definition of uniform continuity:

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

So assume that f is not uniformly continuous. Then by negating the above one sees

$$\exists \varepsilon > 0 : \forall \delta > 0 \exists x, y \in M, d(x,y) < \delta \text{ and } \rho(f(x), f(y)) \geq \varepsilon.$$

Now fix this $\varepsilon > 0$ and let $\delta = \frac{1}{n}$. Then there must exist $x_n, y_n \in M, d(x_n, y_n) < \frac{1}{n}$ and $\rho(f(x_n), f(y_n)) \geq \varepsilon$. Since $(y_n)_n \subset M$ and M is compact, there exists a subsequence $(y_{n_l})_l$ of $(y_n)_n$ which converges to some point y. Look at $(x_{n_l})_l$. Again by compactness, there exists a subsequence $(x_{n_{l_k}})_k$ which converges to some point x. Since $x_{n_{l_k}} \to x$ and $y_{n_{l_k}} \to y$ we have

$$d(x,y) = \lim_{k\to\infty} d(x_{n_{l_k}}, y_{n_{l_k}}) = 0,$$

i,e, $x = y$.
But since $\rho(f(x_n), f(y_n)) \geq \varepsilon > 0$, we have

$$\lim_{k\to\infty} f(x_{n_{l_k}}) \neq \lim_{k\to\infty} f(y_{n_{l_k}})$$

so f is not continuous at x.
Thus f not uniformly continuous $\Rightarrow f$ not continuous $\iff f$ continuous $\Rightarrow f$ uniformly continuous. □

4 The sequence spaces $l^p(\mathbb{N}), 1 \leq p \leq \infty$

Definition 4.1.

- *$l^\infty(\mathbb{N})$ is the space of all bounded sequences $x : \mathbb{N} \to \mathbb{F}$ equipped with the norm*

$$\|x\|_\infty := \sup_{n\in\mathbb{N}} |x_n|.$$

- *Let $1 \leq p < \infty$. $l^p(\mathbb{N})$ is the space of all sequences $x : \mathbb{N} \to \mathbb{F}$ for which $\sum\limits_{n\in\mathbb{N}} |x_n|^p < \infty$. With*

$$\|x\|_p := \Big(\sum_{n\in\mathbb{N}} |x_n|^p\Big)^{\frac{1}{p}}$$

it becomes a normed vector space.

Lemma 4.2. *Let $1 \leq p \leq \infty$. Then $(l^p(\mathbb{N}), \|\cdot\|_p)$ is a normed vector space.*

Proof. Case 1: $p = \infty$ should be immediate.
Case 2: $1 \leq p < \infty$ is more complicated. It is not even obvious why $(l^p, \|\cdot\|_p)$ is a vector space. If $x \in l^p$ and $\alpha \in \mathbb{F}$, then $\alpha x \in l^p$ is clear, but if $x, y \in l^p(\mathbb{N})$ why is $x + y \in l^p(\mathbb{N})$?
Let $x, y \in l^p(\mathbb{N})$, i.e., $\|x\|_p, \|y\|_p < \infty$. Then

$$\begin{aligned}\sum_{n=1}^{\infty} |x_n + y_n|^p &\leq \sum_{n=1}^{\infty} (2\max(|x_n|, |y_n|))^p \\ &= 2^p \sum_{n=1}^{\infty} \max(|x_n|, |y_n|) \leq 2^p (\sum_{n=1}^{\infty} |x_n|^p + \sum_{n=1}^{\infty} |y_n|^p) < \infty\end{aligned}$$

so $x + y \in l^p(\mathbb{N})$.
To show that $\|\cdot\|_p$ is a norm, we only have to check the triangle-inequality and for this we need some more help.

Lemma 4.3 (Hölder inequality)**.** *Let $1 \leq p \leq \infty$ and define the **dual exponent** $q \in [1, \infty]$ by*

$$q = \begin{cases} \infty, & \text{if } p = 1, \\ 1, & \text{if } p = \infty, \\ \frac{p}{p-1} (\text{ i.e., } q \text{ is such that } \frac{1}{p} + \frac{1}{q} = 1), & \text{if } 1 < p < \infty. \end{cases}$$

Then if $x \in l^p(\mathbb{N}), y \in l^q(\mathbb{N})$ and if $x - y$ is defined by $(x \cdot y)_n := x_n \cdot y_n, n \in \mathbb{N}$, then $x \cdot y \in l^1(\mathbb{N})$ and

$$\|x \cdot y\|_1 \leq \|x\|_p \|y\|_q.$$

Armed with this, we can show that $\|\cdot\|_p$ is a norm for $1 \leq p \leq \infty$.

Let $x, y \in l^p(\mathbb{N})$, then

$$\begin{aligned}\|x+y\|_p^p &= \sum_{n=1}^{\infty} |x_n + y_n|^p \le \sum_{n=1}^{\infty} (|x_n| + |y_n|)^p \\ &= \sum_{n=1}^{\infty} (|x_n| + |y_n|)(|x_n| + |y_n|)^{p-1} \\ &= \sum_{n=1}^{\infty} |x_n|(|x_n| + |y_n|)^{p-1} + \sum_{n=1}^{\infty} |y_n|(|x_n| + |y_n|)^{p-1} =: (\star)\end{aligned}$$

We know already that $x + y \in l^p$. Let q be the dual exponent to p. Then $\frac{p}{q} = p - 1$ (with the convention that $\frac{1}{\infty} = 0$). So, since $(|x_n| + |y_n|)_n \in l^p$, one has

$$(|x_n| + |y_n|)^{p-1} = (|x_n| + |y_n|)^{\frac{p}{q}} \in l^q.$$

So the Hölder inequality applies to $(\star)$ and

$$\begin{aligned}\sum_{n=1}^{\infty} (|x_n| + |y_n|)^p &= \sum_{n=1}^{\infty} |x_n|(|x_n| + |y_n|)^{\frac{p}{q}} + \sum_{n=1}^{\infty} |y_n|(|x_n| + |y_n|)^{\frac{p}{q}} \\ &\le \Big(\sum_{n=1}^{\infty} |x_n|^p\Big)^{\frac{1}{p}} \Big(\sum_{n=1}^{\infty} (|x_n| + |y_n|)^{\frac{p}{q}\cdot q}\Big)^{\frac{1}{q}} \\ &\qquad + \Big(\sum_{n=1}^{\infty} |y_n|^p\Big)^{\frac{1}{p}} \Big(\sum_{n=1}^{\infty} (|x_n| + |y_n|)^{\frac{p}{q}\cdot q}\Big)^{\frac{1}{q}} \\ &= (\|x\|_p + \|y\|_p)\Big(\sum_{n=1}^{\infty} (|x_n| + |y_n|)^p\Big)^{\frac{1}{q}}\end{aligned}$$

$$\Rightarrow \underbrace{\Big(\sum_{n=1}^{\infty} (|x_n| + |y_n|)^p\Big)^{1-\frac{1}{q}}}_{\big(\sum_{n=1}^{\infty} (|x_n|+|y_n|)^p\big)^{\frac{1}{p}} = \|x+y\|_p} \le \|x\|_p + \|y\|_p.$$

So $\|x + y\|_p \le \|x\|_p + \|y\|_p$. □

It remains to prove Hölder inequality. For this we need

Lemma 4.4 (Young's inequality)**.** *Let $1 < p < \infty$. Then for all $a, b \ge 0$*

$$ab \le \frac{1}{p}a^p + \frac{1}{q}b^q,$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. For some suitable function G, we want to have an inequality of the form

$$a \cdot b \le G(a) + F(b) \quad \forall a, b \ge 0$$

for a suitable function F. How to guess F?
Certainly F given by

$$F(b) := \sup_{a>0}(ab - G(a)) \tag{$\star\star$}$$

works, since then

$$G(a) + F(b) \geq G(a) + ab - G(a) = ab.$$

So we need to find the supremum in ($\star\star$). If $G(a) = \frac{1}{p}a^p$ and $1 < p < \infty$, then $G(0) = 0$ and $\lim_{a\to\infty}(ab - G(a)) = -\infty$ so there will be a point a (depending on b) for which $a \mapsto ab - G(a)$ is maximal.
At this point the derivative

$$\frac{d}{da}(ab - G(a)) = b - G'(a) = b - a^{p-1}$$

must be zero $\Rightarrow a = b^{1/(p-1)}$.

$$\begin{aligned}\Rightarrow F(b) &= ab - \frac{1}{p}a^p = a(b - \frac{1}{p}a^{p-1})\\ &= b^{\frac{1}{p-1}}(b - \frac{1}{p}b) = b^{\frac{p}{p-1}}\frac{p}{p-1} = \frac{1}{q}b^q\end{aligned}$$

with $q = \frac{p}{p-1}$. □

Proof of Lemma 4.3. Let $(x_n)_n \in l^p$ and $(y_n)_n \in l^q, 1 \leq p \leq \infty$, q dual exponent of p.
The cases $p = 1$ or $p = \infty$ are easy (do them!).
So let $1 < p < \infty$.
Step 1: Assume $\|x\|_p = 1 = \|y\|_q$. Then

$$\|x \cdot y\|_1 \leq 1.$$

Indeed,

$$\|x \cdot y\|_1 \leq \sum_{n=1}^{\infty}|x_n y_n| = \sum_{n=1}^{\infty}|x_n||y_n|$$

and by Lemma 4.4

$$\begin{aligned}&\leq \sum_{n=1}^{\infty}\big(\frac{1}{p}|x_n|^p + \frac{1}{q}|y_n|^q\big)\\ &= \frac{1}{p}\|x\|_p^p + \frac{1}{q}\|y\|_q^q\\ &= \frac{1}{p} + \frac{1}{q} = 1.\end{aligned}$$

Step 2: Assume $x \neq 0, y \neq 0$. Then

$$\frac{\|x \cdot y\|_1}{\|x\|_p\|y\|_q} = \|\frac{x}{\|x\|_p} \cdot \frac{y}{\|y\|_q}\|_1 = \|\tilde{x} \cdot \tilde{y}\|_1$$

with $\tilde{x} = \left(\frac{x_n}{\|x\|_p}\right)_n, \tilde{y} = \left(\frac{y_n}{\|y\|_q}\right)_n$.
Note $\|\tilde{x}\|_p = 1 = \|\tilde{y}\|_q$. So by Step 1

$$\|\tilde{x} \cdot \tilde{y}\|_1 \leq 1,$$

hence

$$\|x \cdot y\|_1 = \|x\|_p \|y\|_q \|\tilde{x} \cdot \tilde{y}\|_1 \leq \|x\|_p \|y\|_q.$$

□

Theorem 4.5. *The spaces $(l^p(\mathbb{N}), \|\cdot\|_p)$ are Banach spaces, i.e., they are complete.*

Proof. Only completeness remains: we do only $1 \leq p < \infty$.
We write $x = (x(j))_{j\in\mathbb{N}} \in l^p(\mathbb{N})$.
So let $(x_n)_n \subset l^p(\mathbb{N})$ be Cauchy.
Step 1: A candidate for the limit: Fix $j \in \mathbb{N}$ and consider

$$|x_n(j) - x_m(j)| \leq \Big(\sum_{l=1}^{\infty} |x_n(l) - x_m(l)|^p\Big)^{\frac{1}{p}} = \|x_n - x_m\|_p$$

$\Rightarrow (x_n(j))_n \subset \mathbb{F}$ is Cauchy. By completeness of $\mathbb{F}$ $x(j) := \lim_{n\to\infty} x_n(j)$ exists.
Step 2: $x \in l^p(\mathbb{N})$!
Idea:

$$\|x\|_p^p = \sum_{j=1}^{\infty} |x(j)|^p = \sum_{j=1}^{\infty} \lim_{n\to\infty} |x_n(j)|^p \text{“} = \text{“} \lim_{n\to\infty} \underbrace{\sum_{j=1}^{\infty} |x_n(j)|^p}_{\|x_n\|_p^p} < \infty.$$

Let $L \in \mathbb{N}$. Note that

$$\begin{aligned}\sum_{j=1}^{L} |x(j)|^p &= \sum_{j=1}^{L} \lim_{n\to\infty} |x_n(j)|^p \\ &= \liminf_{n\to\infty} \sum_{j=1}^{L} |x_n(j)|^p \\ &\leq \liminf_{n\to\infty} \|x_n\|_p^p < \infty\end{aligned}$$

Thus, using the monotone convergence theorem, we conclude that

$$\sum_{j=1}^{\infty} |x(j)|^p = \lim_{L\to\infty} \sum_{j=1}^{L} |x(j)|^p \leq (\liminf \|x_n\|_p)^p$$

$$\Rightarrow \|x\|_p \leq \liminf \|x_n\|_p,$$

so $x \in l^p$!
Step 3: $x_n \to x$ in l^p: Given $\varepsilon > 0$, there exists

$$N \in \mathbb{N} : \|x_n - x_m\|_p < \varepsilon \quad \forall n, m \geq N.$$

Let $L \in \mathbb{N}$.

$$\begin{aligned}\sum_{j=1}^{L} |x(j) - x_n(j)|^p &= \lim_{m\to\infty} \sum_{j=1}^{L} |x_m(j) - x_n(j)|^p \\ &\leq \limsup_{m\to\infty} \|x_m - x_n\|_p^p \\ &\leq \varepsilon^p \text{ for } n \text{ large enough}\end{aligned}$$

$\Rightarrow$ for $n \geq N$:

$$\|x - x_n\|_p^p = \lim_{L\to\infty} \sum_{j=1}^{L} |x(j) - x_n(j)|^p \leq \varepsilon^p$$

or $\|x - x_n\|_p \leq \varepsilon$ for all n large enough!

□

5 Hahn-Banach type theorems

5.1 Some preparations

Definition 5.1. *Let $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ be normed vector spaces. A continuous linear map $T : X \to Y$ is called* ***operator****. If $Y = \mathbb{R}$ or $\mathbb{C}$ we call them functionals.*

Lemma 5.2. *Let X, Y be normed vector spaces and $T : X \to Y$ linear. Then the following are equivalent (t.f.a.e.):*

(a) T is continuous.

(b) T is continuous at 0.

(c) $\exists M \geq 0 : \|Tx\|_Y \leq M\|x\|_X \forall x \in X$.

(d) T is uniformly continuous.

Proof. $(c) \Rightarrow (d) \Rightarrow (a) \Rightarrow (b)$ is easy.
E.g.: $(c) \Rightarrow T$ is Lipschitz continuous, since

$$\|Tx - Tx_0\|_Y = \|T(x - x_0)\|_Y \leq M\|x - x_0\|_X.$$

So we only need to show $(b) \Rightarrow (c)$. Assume that (c) is wrong $\Rightarrow \forall n \in \mathbb{N} \exists x_n \in X : \|Tx_n\|_Y > n\|x_n\|_X \Rightarrow x_n \neq 0$. Then

$$y_n := \frac{x_n}{n\|x_n\|_X} \to 0 \text{ in } X.$$

But

$$\|Ty_n\|_Y = \frac{\|Tx_n\|_X}{n\|x_n\|_X} > 1$$

so $Ty_n \nrightarrow 0$, so T is not continuous at 0, a contradiction. □

Definition 5.3 (Operator-norm)**.** *Given $T : X \to Y$ linear*

$$\|T\| := \|T\|_{X\to Y} := \inf(M \geq 0 | \|Tx\|_Y \leq M\|x\|_X \text{ for all } x \in X)$$

*defines the **operator-norm** of T.*

Note:

$$\|T\| = \sup_{x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X=1} \|Tx\|_Y = \sup_{\|x\|_X\leq 1} \|Tx\|_Y.$$

and

$$\|Tx\|_Y \leq \|T\|\|x\| \quad \forall x \in X. \tag{I.1}$$

Indeed, let $M_0 := \sup_{x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X}$

$$\|Tx\|_Y = \frac{\|Tx\|_Y}{\|x\|_X}\|x\|_X \leq M_0\|x\|_X$$

$$\Rightarrow \|T\| \leq M_0.$$

On the other hand: given $\varepsilon > 0 \exists x_\varepsilon \neq 0 :$

$$\|Tx_\varepsilon\|_Y \geq M_0(1-\varepsilon)\|x_\varepsilon\|_X$$

$$\Rightarrow \|T\| \geq M_0(1-\varepsilon) \quad \forall \varepsilon < 0$$

$$\Rightarrow \|T\| \geq M_0$$

and thus $\|T\| = M_0$, so (I.1) holds.

Definition 5.4. *Let X, Y be normed spaces.*

$$L(X,Y) := \{T : X \to Y | T \text{ is linear and continuous}\}$$

is again a vector space.

$$(S+T)(x) := Sx + Tx,$$
$$(\lambda T)(x) := \lambda Tx.$$

Proposition 5.5. *(a) $\|T\| = \sup_{\|x\|_X\leq 1} \|Tx\|_Y$ defines a norm on $L(X,Y)$.*

(b) If Y is complete, then $L(X,Y)$ is also complete.

Proof. (a) Looking closely reveals

$$\|\lambda T\| = |\lambda|\|T\|$$

$$\|T\| = 0 \Rightarrow T = 0 \text{ (the zero linear map).}$$

Triangle-inequality:

$$\begin{aligned}\|S+T\| &= \sup_{\|x\|_X\leq 1} \underbrace{\|(S+T)x\|_Y}_{=\|Sx+Tx\|_Y} \\ &\leq \sup_{\|x\|_X\leq 1} \|Sx\|_Y + \sup_{\|x\|_X\leq 1} \|Tx\|_Y \\ &= \|S\| + \|T\|.\end{aligned}$$

(b) Let $(T_n)_n \subset L(X,Y)$ be Cauchy $\Rightarrow$ for fixed $x \in X$ $(T_n x)_n$ is Cauchy in Y!

$$\|T_n x - T_m x\|_Y = \|(T_n - T_m)x\|_Y \leq \|T_n - T_m\| \|x\|_X.$$

By completeness of $Y \Rightarrow Tx := \lim_{n\to\infty} T_n x$ exists.
Step 1: T is linear. Indeed

$$\begin{aligned} T(\lambda x_1 + \mu x_2) &= \lim_{n\to\infty} \underbrace{T_n(\lambda x_1 + \mu x_2)}_{\lambda T_n x_1 + \mu T_n x_2} \\ &= \lambda \lim_{n\to\infty} T_n x_1 + \mu \lim_{n\to\infty} T_n x_2 \\ &= \lambda T x_1 + \mu T x_2. \end{aligned}$$

Step 2: $T \in L(X,Y)$, i.e., $\|T\| < \infty$ and $\|T - T_n\| \to 0$.
Indeed, let $\varepsilon > 0$ and choose $N_1 \in \mathbb{N}$ so that

$$\|T_n - T_m\| < \varepsilon \quad \forall n, m \geq N_1.$$

Let $x \in X$, $\|x\|_X \leq 1$. Choose $N_\varepsilon := N_\varepsilon(\varepsilon, x) \geq N_1$ so that

$$\|T_{N_\varepsilon} x - Tx\|_Y \leq \varepsilon$$

Thus, for every $x \in X$ with $\|x\|_X \leq 1$:

$$\begin{aligned} \|T_n x - Tx\|_Y \leq & \underbrace{\|T_n x - T_{N_\varepsilon} x\|_Y}_{=\|(T_n - T_{N_\varepsilon})x\|_Y \leq \|T_n - T_{N_\varepsilon}\| \|x\|_X \leq \|T_n - T_{N_\varepsilon}\|} + \underbrace{\|T_{N_\varepsilon} x - Tx\|_Y}_{\leq \varepsilon} \\ &\leq \underbrace{\|T_n - T_{N_\varepsilon}\|}_{\leq \varepsilon, n \geq N_1} + \varepsilon \\ &\leq 2\varepsilon \text{ for all } n \geq N_1 \end{aligned}$$

$$\begin{aligned} &\|Tx\|_Y \leq \|T_n x - Tx\|_Y + \|T_n x\|_Y \leq 2\varepsilon + \|T_n\| < \infty \\ &\|T - T_n\| = \sup_{\|x\|_X \leq 1} \|Tx - T_n x\| \leq 2\varepsilon \text{ for all } n \geq N_1, \end{aligned}$$

so $T_n \to T$ in operator norm.

□

Definition 5.6. *Given a normed vector space X, its **dual** space is the space $X' = X^* := L(X, \mathbb{F})$ of continuous linear functionals.*

Corollary 5.7. *For any normed vector space X, its dual X' equipped with the norm*

$$\|x'\|_{X'} := \sup_{\|x\|_X \leq 1} |x'(x)| = \sup_{\|x\|_X = 1} |x'(x)|$$

is a Banach space.

5.2 The analytic form of Hahn-Banach: extension of linear functionals

Definition 5.8. *Let E be a vector space. A map $p : E \to \mathbb{R}$ is **sub-linear** if*

(a) $p(\lambda x) = \lambda p(x), \forall \lambda \geq 0, \forall x \in E$.

(b) $p(x + y) \leq p(x) + p(y), \forall x, y \in E$.

Example 5.9. *(i) Every semi-norm is sub-linear.*

(ii) Every linear functional on a real vector space is sub-linear.

(iii) On $l^\infty(\mathbb{N}, \mathbb{R})$ = bounded real-valued sequences, $t = (t_n)_n \mapsto \limsup_{n\to\infty} t_n$ is sub-linear.
On $l^\infty(\mathbb{N}, \mathbb{C})$, $t = (t_n)_n \mapsto \limsup_{n\to\infty} Re(t_n)$ is sub-linear.

(iv) A sub-linear map is often called Minkowski functional.

Theorem 5.10 (Hahn-Banach, analytic form)**.** *Let E be a real vector space, $p : E \to \mathbb{R}$ sub-linear, $G \subset E$ a subspace, and $g : G \to \mathbb{R}$ a linear functional with*

$$g(x) \leq p(x) \qquad \forall x \in G.$$

Then there exists a linear functional $f : E \to \mathbb{R}$ which extends g, i.e., $g(x) = f(x)\ \forall x \in G$, such that

$$f(x) \leq p(x) \qquad \forall x \in E.$$

For the proof we need Zorn's lemma, which is an important property of ordered sets.

Some notations:

- Let P be a set with a partial order relation $\leq$. A subset $Q \subset P$ is **totally ordered** if for any $a, b \in Q$ either $a \leq b$ or $b \leq a$ (or both!) holds.
- Let $Q \subset P$, then $c \in P$ is an **upper bound** for Q if $a \leq c$ for all $a \in Q$.
- We say that $m \in P$ is a **maximal element** of P if there is no element $x \in P$ such that $m \leq x$ except for $x = m$.
  Note that a maximal element of P need not be an upper bound for P!
- We say that P is **inductive** if every totally ordered subset $Q \subset P$ has an upper bound.

Lemma 5.11 (Zorn)**.** *Every non-empty ordered set which is inductive has a maximal element.*

Proof of Theorem 5.10. We say that h extends g if

$$D(h) \supset D(g) \quad \text{and} \quad h(x) = g(x)\ \forall x \in D(g).$$

Consider the set

$$P = \left\{ \begin{array}{rr} h : E \supset D(h) \to \mathbb{R}| & D(h) \text{ is a linear subspace of } E, \\ & h \text{ is linear, } G \subset D(h), h \text{ extends } g, \\ & h(x) \leq p(x) \quad \forall x \in D(h) \end{array} \right\}.$$

Note: $P \neq \emptyset$ since $g \in P$!
On P we define the order $h_1 \leq h_2 \iff h_2$ extends h_1.
Step 1: P is inductive.
Indeed, let $Q \subset P$ be totally ordered. Write $Q = (h_i)_{i \in I}$ and set

$$D(h) := \bigcup_{i \in I} D(h_i), \quad h(x) := h_i(x) \text{ if } x \in D(h_i) \text{ for some } i \in I.$$

It is easy to see that this definition is consistent and that h is an upper bound for Q.
Step 2: By Step 1 and Zorn's lemma, P has a maximal element $f \in P$.
Claim: $D(f) = E$ (which finishes the proof).
Assume that $D(f) \neq E$. Let $x_0 \notin D(f)$ and set $D(h) := D(f) + \mathbb{R}x_0$ and for $x \in D(f)$ set

$$h(x + tx_0) := f(x) + t\alpha, \quad t \in \mathbb{R},$$

where we will choose α so that $h \in P$. For this we need

$$f(x) + t\alpha \leq p(x + tx_0). \tag{I.2}$$

Let $t > 0$. Then

$$\begin{aligned}
\text{(I.2)} \iff t\alpha &\leq p(x + tx_0) - f(x) \\
\iff \alpha &\leq \frac{1}{t} p(x + tx_0) - \frac{1}{t} f(x) \\
&= p(\frac{x}{t} + x_0) - f(\frac{x}{t}) \\
&= p(u + x_0) - f(u),
\end{aligned}$$

where $u := \frac{x}{t}$.
If $t < 0$, then

$$\begin{aligned}
\text{(I.2)} \iff t\alpha &\leq p(x + tx_0) - f(x) \\
\iff -\alpha &\leq \frac{1}{-t} p(x + tx_0) - \frac{1}{-t} f(x) \\
&= p(\frac{x}{-t} + x_0) - f(\frac{x}{-t}) \\
&= p(w + x_0) - f(w) \\
\iff \alpha &\geq f(w) - p(w - x_0),
\end{aligned}$$

where $w := \frac{x}{-t}$.
Thus (I.2) holds if

$$f(w) - p(w - x_0) \leq \alpha \leq p(u + x_0) - f(u) \quad \forall u, w \in D(f). \tag{I.3}$$

Since $f \in P$, we have

$$f(x) \leq p(x) \quad \forall x \in D(f).$$

Hence $\forall u, w \in D(f)$ it holds

$$\begin{aligned}
f(u) + f(w) &= f(u + w) \\
&\leq p(u + w) \\
&= p(u + x_0 + w - x_0) \\
&\leq p(u + x_0) + p(w - x_0)
\end{aligned}$$

so

$$f(w) - p(w - x_0) \le p(u + x_0) - f(u)$$

and hence (I.3) holds with the choice

$$\alpha = \sup_{w \in D(f)} (f(w) - p(w - x_0)).$$

We have shown: If $D(f) \neq E$, then we can extend f by h, i.e., $f \le h$ but this contradicts that f is a maximal element of P! □

To extend Hahn-Banach to complex vector spaces, we need

Lemma 5.12. *Let X be a $\mathbb{C}$-vector space.*

(a) If $l : X \to \mathbb{R}$ is $\mathbb{R}$-linear, i.e.,

$$l(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 l(x_1) + \lambda_2 l(x_2) \quad \forall x_1, x_2 \in X, \lambda_1, \lambda_2 \in \mathbb{R}$$

then

$$\tilde{l}(x) := l(x) - il(ix), \quad x \in X$$

defines a $\mathbb{C}$-linear functional $\tilde{l} : X \to \mathbb{C}$ and $l = Re(\tilde{l})$.

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

(c) If $p : X \to \mathbb{R}$ is a semi-norm and $l : X \to \mathbb{C}$ is $\mathbb{C}$-linear, then

$$|l(x)| \le p(x) \;\; x \in X \iff |Re(l(x))| \le p(x) \;\; x \in X.$$

(d) If $\|\cdot\|$ is a norm on X and $l : X \to \mathbb{C}$ is $\mathbb{C}$-linear and continuous, then

$$\|l\| = \|Re(l)\|.$$

Remark 5.13. *Thus the map $l \mapsto Re(l)$ is a bijective $\mathbb{R}$-linear map between the space of $\mathbb{C}$-linear and $\mathbb{R}$-linear functionals and if X is a normed vector space, it is an isometry.*

Proof of Lemma 5.12. (a) Let l be $\mathbb{R}$-linear and $\tilde{l}(x) := l(x) - il(ix)$. Since $x \mapsto ix$ is $\mathbb{R}$-linear, we have that $\tilde{l}$ is $\mathbb{R}$-linear and, by construction, $Re(\tilde{l}) = l$. So we only need to check $\tilde{l}(ix) = i\tilde{l}(x)$.

$$\begin{aligned} \tilde{l}(ix) &= l(ix) - il(iix) \\ &= l(ix) - il(-x) = l(ix) + il(x) \\ &= i(l(x) - il(ix)) = i\tilde{l}(x). \end{aligned}$$

(b) If h is $\mathbb{C}$-linear, then $l = Re(h)$ is $\mathbb{R}$-linear.
Since $Im z = -Re(iz) \; \forall z \in \mathbb{C}$, we have

$$\begin{aligned} h(x) &= Re\, h(x) + iIm h(x) \\ &= Re\, h(x) - iRe(ih(x)) \\ &= Re\, h(x) - iRe(h(ix)) \\ &= l(x) - il(ix) = \tilde{l}(x). \end{aligned}$$

(c) $\forall z \in \mathbb{C} : |Re(z)| \leq |z|$ so "$\Rightarrow$" holds.
"$\Leftarrow$": Let $x \in X$ and write $l(x) = \lambda |l(x)|$ for some $\lambda = \lambda(x)$ with $|\lambda| = 1$. Then

$$|l(x)| = \lambda^{-1} l(x) = l(\lambda^{-1} x)$$

so

$$\begin{aligned} |l(x)| = Re|l(x)| = Re\, l(\lambda^{-1}x) &\leq |Re\, l(\lambda^{-1}x)| \\ &\leq p(\lambda^{-1}x) = p(x) \quad \forall x \in X. \end{aligned}$$

(d) Follows immediately from (c).

□

Theorem 5.14 (Hahn-Banach, complex form)**.** *Let E be a complex vector space, $p : E \to \mathbb{R}$ sub-linear, $G \subset E$ a subspace, and $l : G \to \mathbb{C}$ a linear functional with*

$$Re\, l(x) \leq p(x) \quad \forall x \in G.$$

Then there exists a linear functional $h : E \to \mathbb{C}$ which extends l with

$$Re\, h(x) \leq p(x) \quad \forall x \in E.$$

Proof. Consider $G \subset E$ as a real vector space and $Re\, l : G \to \mathbb{R}$ as a real functional. By Theorem 5.10 there exists and extension $f : E \to \mathbb{R}$ of $Re\, l$ with $f(x) \leq p(x)\ \forall x \in E$. By Lemma 5.12, the functional

$$h(x) := f(x) - i f(ix)$$

is $\mathbb{C}$-linear and

$$Re\, h(x) = f(x) \leq p(x) \quad \forall x \in E$$

and $h(x) = l(x)\ \forall x \in G$. □

Corollary 5.15. *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let X be a normed vector space and $U \subset X$ a linear subspace. Then for every continuous linear functional $g : U \to \mathbb{F}$ there exists $f \in X'$ such that*

$$f|_U = g \quad \textit{and} \quad \sup_{x \in B_X} |f(x)| = \|f\| = \|g\| = \sup_{x \in B_U} |g(x)|,$$

where

$$B_U = \{x \in U | \|x\| \leq 1\} \subset B_X = \{x \in X | \|x\| \leq 1\}.$$

Proof. Step 1: Let X be a real normed vector space. Put $\|g\| := \sup_{x \in B_U} |g(x)|$ and

$$p(x) := \|g\| \|x\|.$$

By Theorem 5.10 there exists $f : X \to \mathbb{R}$ which extends g and

$$f(x) \leq p(x) \quad \forall x \in X.$$

Since

$$-f(x) = f(-x) \leq p(-x) = p(x)$$

one has $|f(x)| \leq p(x)\ \forall x \in X$ and so $\|f\| \leq \|g\|$.
On the other hand

$$\|g\| = \sup_{x \in B_U} |g(x)| = \sup_{x \in B_U} |f(x)| \leq \sup_{x \in B_X} |f(x)| = \|f\|.$$

Step 2: If X is a complex vector space, apply Step 1, now combined with Theorem 5.14, to get a linear functional $f : X \to \mathbb{C}$ with

$$f|_U = g \quad \text{and} \quad \|Ref\| = \|g\|.$$

By Lemma 5.12 (d) we get $\|f\| = \|Ref\|$. □

Corollary 5.16. *In any normed vector space X to every $x_0 \in X$ there exists $f_0 \in X', \|f_0\| = 1, f_0(x_0) = \|x_0\|$. In particular, X' separates the points of X: $\forall x_1, x_2 \in X, x_1 \neq x_2 \Rightarrow \exists f \in X' : f(x_1) \neq f(x_2)$.*

Proof. By Corollary 5.15 we can extend linear functional $g : \mathbb{F}x_0 \to \mathbb{F}, g(\lambda) = \lambda\|x_0\|$ to X and preserve its norm.
If $x_1 \neq x_2$, consider $x = x_1 - x_2 \neq 0$. □

Corollary 5.17. *Let X be a normed vector space. Then it holds*

$$\|x\|_X = \underbrace{\sup\{|f(x)|\ |f \in X', \|f\|_{X'} \leq 1\}}_{=:a} \quad \forall x \in X. \tag{I.4}$$

Proof. By definition of $\|f\|_{X'}$

$$|f(x)| \leq \|f\|_{X'}\|x\|_X \leq \|x\|_X \quad \Rightarrow a \leq \|x\|_X.$$

By Corollary 5.16, there exists $f_0 \in X'$ such that $\|f_0\|_{X'} = 1$ and $f_0(x) = \|x\|$ $\Rightarrow a \leq \|x\|$. □

Remark 5.18. *Note the symmetry of (I.4) with*

$$\|f\|_{X'} = \sup\{|f(x)|\ |\ \|x\|_X \leq 1\}.$$

5.3 Geometric form of Hahn-Banach

In the following, let E be a normed vector space.

Definition 5.19. *An **affine hyperplane** is a subset $H \subset E$ of the form*

$$H = \{x \in E | Ref(x) = \alpha\}$$

where f is a linear functional on E which is not identically zero, $\alpha \in \mathbb{R}$.

$$H = [Ref = \alpha], \quad \text{or} \quad Ref = \alpha$$

is the equation of H.

Proposition 5.20. *The hyperplane* $H = [Ref = \alpha]$ *is closed* $\iff$ f *is continuous.*

Proof. "$\Leftarrow$": clear.
"$\Rightarrow$": H is closed $\Rightarrow$ H^c is open. Pick $x_0 \in H^c$ with $Ref(x_0) < \alpha$. Thus $\exists r > 0 : B_r(x_0) \subset H^c$.
Claim: $Ref(x) < \alpha \; \forall x \in B_r(x_0)$.
Assume $\exists x_1 \in B_r(x_0) : Ref(x_1) > \alpha$

$$\Rightarrow [x_0, x_1] = \{x_t = (1-t)x_0 + tx_1 | t \in [0,1]\} \subset B_r(x_0)$$

so, since $B_r(x_0) \subset H^c \Rightarrow Ref(x_t) \neq \alpha$ for all $t \in [0,1]$. But if

$$t = \frac{Ref(x_1) - \alpha}{Ref(x_1) - Ref(x_0)} \quad \Rightarrow \quad Ref(x_0) = \alpha,$$

a contradiction. So $Ref(x) < \alpha \; \forall x \in B_r(x_0)$.
Let $x \in E, \|x\| \leq 1$. Then

$$\tilde{x} = x_0 + rx \in B_r(x_0) \quad \Rightarrow \quad \underbrace{Ref(\tilde{x})}_{= rRef(x) + Ref(x_0)} < \alpha$$

$$\Rightarrow Ref(x) < \frac{1}{r}(\alpha - Ref(x_0))$$

and

$$Ref(-x) = -Ref(x) > \frac{-1}{r}(\alpha - Ref(x_0)).$$

So

$$|Ref(x)| \leq \frac{1}{r}(\alpha - Ref(x_0))$$

i.e.,

$$\|f\| = \|Ref\| = \sup_{\|x\| < 1} |Ref(x)| \leq \frac{1}{r}(\alpha - Ref(x_0)) < \infty$$

so f is continuous. □

Definition 5.21. *Let* $A, B \subset E$. $H = [Ref = \alpha]$ ***separates*** A *and* B *if*

$$Ref(x) \leq \alpha \quad \forall x \in A \quad \text{and} \quad Ref(x) \geq \alpha \quad \forall x \in B.$$

H ***strictly separates*** A *and* B *if there exists* $\varepsilon > 0$ *such that*

$$Ref(x) \leq \alpha - \varepsilon \quad \forall x \in A \quad \text{and} \quad Ref(x) \geq \alpha + \varepsilon \quad \forall x \in B.$$

$A \subset E$ *is* ***convex*** *if* $\forall x_1, x_2 \in A$

$$[x_1, x_2] = \{(1-t)x_1 + tx_2 | t \in [0,1]\} \subset A.$$

Theorem 5.22 (Hahn-Banach, first geometric form)**.** *Let $A, B \subset E, A, B \neq \emptyset$, convex, $A \cap B = \emptyset$ and one of them be open. Then there exists a closed hyperplane H which separates them.*

We need two Lemmata:

Lemma 5.23. *Let $C \subset E$ be open and convex with $0 \in C$. For $x \in E$ set*

$$p(x) = \inf\{\alpha > 0 | \alpha^{-1}x \in C\} \quad \text{(gauge, Minkowski functional)}. \tag{I.5}$$

Then p is sub-linear and

$$\exists M < \infty : 0 \leq p(x) \leq M\|x\| \quad \forall x \in E \tag{I.6}$$

$$C = \{x \in E | p(x) < 1\}. \tag{I.7}$$

Proof. Step 1: Clearly $p(\lambda x) = \lambda p(x)\ \forall x \in E, \lambda \geq 0$.
Step 2: $\exists r > 0 : B_r(0) \subset C$ (C is open). Thus $\forall \tilde{x} \in B_r(0)$

$$p(\tilde{x}) \leq 1$$

so if $x \in E \setminus \{0\}$ and $0 < \delta < r$ we have

$$\tilde{x} = (r - \delta)\frac{x}{\|x\|} \in B_r(0)$$

$$\Rightarrow 1 \geq p(\tilde{x}) = p((r - \delta)\frac{x}{\|x\|}) = (r - \delta)\frac{1}{\|x\|}p(x)$$

$$p(x) \leq \frac{\|x\|}{r - \delta}p(\tilde{x}) \quad \forall x \in E$$

$$\delta \to 0 \Rightarrow p(x) \leq \frac{1}{r}\|x\| \quad \forall x \in E$$

so $M = \frac{1}{r}$ works.
Step 3: $C = \{x \in E | p(x) < 1\}$.
Indeed, let $x \in C$. C is open $\Rightarrow (1 + \varepsilon)x \in C$ for small enough ε

$$\Rightarrow p(x) \leq \frac{1}{1 + \varepsilon} < 1 \quad \forall x \in C$$

$$\Rightarrow C \subset \{p < 1\}.$$

Conversely, if $p(x) < 1$

$$\exists \alpha \in (0, 1) : \frac{x}{\alpha} \in C.$$

$0 \in C$, C convex $\Rightarrow x = \alpha\frac{x}{\alpha} + (1 - \alpha)0 \in C \Rightarrow x \in C$.
Step 4: $\forall x, y \in E : p(x + y) \leq p(x) + p(y)$.
Indeed, let $\varepsilon > 0, \lambda = p(x) - \varepsilon, \mu = p(y) + \varepsilon$ and note

$$\frac{x}{\lambda} \in C, \quad \frac{y}{\mu} \in C$$

since

$$p(\frac{x}{\lambda}) = \frac{1}{\lambda}p(x) = \frac{p(x)}{p(x)+\varepsilon} < 1.$$

$$C(\text{convex}) \ni \frac{\lambda}{\lambda+\mu}\underbrace{\frac{x}{\lambda}}_{\in C} + \frac{\mu}{\lambda+\mu}\underbrace{\frac{y}{\mu}}_{\in C} = \frac{x+y}{\lambda+\mu} \in C$$

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

$$\Rightarrow p(x+y) < \lambda+\mu = p(x)+p(y)+2\varepsilon \quad \forall \varepsilon > 0$$

$$\Rightarrow p(x+y) \leq p(x)+p(y).$$

□

Lemma 5.24. *Let $C \subset E, C \neq \emptyset$, convex, $x_0 \in E \setminus C$. Then there exists $f \in E' : Ref(x) < Ref(x_0)\ \forall x \in C$.*

Proof. Step 1: $\mathbb{F} = \mathbb{R}$.
By translation: we may assume $0 \in C$. Let p be the Minkowski functional for C, $G = \mathbb{R}x_0, g : G \to \mathbb{R}$

$$g(tx_0) := tp(x_0) \quad t \in \mathbb{R}$$

$$\Rightarrow g(x) \leq p(x) \quad \forall x \in G.$$

By Theorem 5.10 $\exists$ linear functional f on E which extends g and

$$f(x) \leq p(x) \quad \forall x \in E.$$

In particular, $f(x_0) = p(x_0) \geq 1$, and by (I.6)

$$f(x) \leq p(x) \leq M\|x\| \quad \forall x \in E$$

$$|f(x)| \leq M\|x\| \quad \forall x \in E$$

so f is continuous.

$$\text{(I.7)} \Rightarrow f(x) \leq p(x) < 1 \quad \forall x \in C$$

$$\Rightarrow f(x) \neq f(x_0) \quad \forall x \in C.$$

Step 2: $\mathbb{F} = \mathbb{C}$. Use Step 1 and Lemma 5.12. □

Proof of Theorem 5.22. $A, B \neq \emptyset$ convex, $A \cap B = \emptyset$, A open.

$$C := A - B = \{x - y | x \in A, y \in B\}$$

is open ($C = \cup_{y \in B} \underbrace{(A - y)}_{\text{open}}$) and convex (check this!). $0 \notin C$ ($A \cap B = \emptyset$). By Lemma 5.24 there exists $f \in E'$ with

$$\begin{aligned} Ref(z) &< 0 \quad \forall z \in C \\ Ref(x - y) &< 0 \quad \forall x \in A, \forall y \in B \\ Ref(x) - Ref(y) &< 0 \quad \forall x \in A, \forall y \in B \end{aligned}$$

$$\Rightarrow Ref(x) < Ref(y) \quad \forall x \in A, \forall y \in B.$$

Take $\alpha = \sup_{x \in A} Ref(x)$

$$\Rightarrow Ref(x) \leq \alpha \leq Ref(y) \quad \forall x \in A, \forall y \in B.$$

□

Theorem 5.25 (Hahn-Banach, second geometric form)**.** *Let $A, B \subset E, A, B \neq \emptyset$, convex, $A \cap B = \emptyset, A$ closed, B compact. Then there exists a closed hyperplane that strictly separates A and B.*

Proof. $C := A - B = \{x - y | x \in A, y \in B\}$ is convex and closed (why?), and $0 \notin C \Rightarrow \exists r > 0 : B_r(0) \cap C = \emptyset$. By Theorem 5.22, there exists a closed hyperplane H which separates C and $B_r(0)$. So there exists $f \in E^*, f \not\equiv 0$, such that

$$Ref(x - y) \leq Ref(rz) = rRef(z) \quad \forall x \in A, y \in B, z \in B_1(0).$$

Since

$$\inf_{z \in B_1(0)} Ref(z) = - \sup_{z \in B_1(0)} |Ref(z)| = - \sup_{z \in B_1(0)} |f(z)| = -\|f\| < 0$$

$$\Rightarrow Ref(x) - Ref(y) \leq -r\|f\| \quad \forall x \in A, y \in B$$

$$\Leftrightarrow Ref(x) + \underbrace{\frac{r}{2}\|f\|}_{=:\varepsilon} \leq Ref(y) - \frac{r}{2}\|f\| \quad \forall x \in A, y \in B.$$

Choose $\alpha = \sup_{x \in A} Ref(x) + \varepsilon$

$$\Rightarrow Ref(x) + \varepsilon \leq \alpha \leq f(y) - \varepsilon \quad \forall x \in A, y \in B.$$

So $H = [Ref = \alpha]$ strictly separates A and B. □

Remark 5.26. *Assuming only that A, B are convex, $\neq \emptyset, A \cap B = \emptyset$ it is in general impossible to separate A and B by a closed hyperplane (except when E is finite-dimensional).*

Corollary 5.27. *Let $F \subset E$ linear subspace such that $\bar{F} \neq E$. Then there exists $f \in E^*, f \not\equiv 0$ such that $f(x) = 0 \; \forall x \in F$.*

Proof. Let $x_0 \in E \setminus \bar{F}$, so $x_0 \notin \bar{F}$. Let $A = \bar{F}, B = x_0$. By Theorem 5.25 there exists a closed hyperplane $H = [Ref = \alpha]$ which strictly separates $\bar{F}$ and $\{x_0\}$

$$\Rightarrow Ref(x) < \alpha < Ref(x_0) \quad \forall x \in \bar{F}.$$

Since F is linear: $x \in F \Rightarrow \lambda x \in F \; \forall \lambda \in \mathbb{R}$

$$\Rightarrow \lambda Ref(x) < \alpha < Ref(x_0) \quad \forall x \in F, \forall \lambda \in \mathbb{R}$$

$$\Rightarrow Ref(x) = 0 \quad \forall x \in F \quad \text{(why?)}$$

$$\Rightarrow 0 = Ref(-ix) = Re(-if(x)) = -Re(if(x)) = Imf(x) \quad \forall x \in F$$

so

$$f(x) = 0 \quad \forall x \in F \quad \text{and} \quad Ref(x_0) > 0.$$

□

Remark 5.28. *The main use of Corollary 5.27 is in the reverse direction: If every continuous linear functional f which vanishes in the subspace $F \subset E$ also vanishes on E, then F is dense in E!*

6 The Baire Category theorem and its applications

6.1 The Baire Category theorem

Recall that $E \subset M$ is dense in M if $\bar{E} = M$.

Definition 6.1. *Let (M, d) be a metric space.*

*(a) $E \subset M$ is **nowhere dense** if $\bar{E}$ has empty interior, i.e., $(\bar{E})^o = int(\bar{E}) = \emptyset$.*

*(b) $F \subset M$ is **meager** (or of 1st category) if it is the countable union of nowhwere dense sets, i.e., there exists $(F_n)_n$ of nowhere dense sets such that $F = \bigcup_{n \in \mathbb{N}} F_n$.*

*(c) F is **fat** (or of 2nd category) if F is not meager.*

*(d) $E \subset M$ is **generic** if E^c is meager.*

Note:

- A nowhere dense set is in no open set dense.
- In (b) we can always assume that $F_n = \bar{F}_n$.
- $F = \{x\}, x \in M$ is nowhere dense. So $\mathbb{Q}$ is meager in $\mathbb{R}$.

Theorem 6.2 (Baire)**.** *Let (M, d) be a complete metric space.*

(a) If for $n \in \mathbb{N}, F_n \subset M$ is closed and nowhere dense, then

$$(\bigcup_{n\in\mathbb{N}} F_n)^o = int(\bigcup_{n\in\mathbb{N}} F_n) = \emptyset.$$

(b) If for $n \in \mathbb{N}, \mathcal{O}_n \subset M$ is open and dense in M, then $\bigcap_{n\in\mathbb{N}} \mathcal{O}_n$ is dense in M.

Remark 6.3. *(a) says that in a complete metric space, meager sets are nowhere dense.*

(b) says that in a complete metric space, a generic set is dense.

Proof of Theorem 6.2. Step 1: $(a) \Leftrightarrow (b)$.
Indeed, note that for any $E \subset M$ one has

$$int(E^c) = \bar{E}^c \quad \text{(why?)} \tag{I.8}$$

Thus if (a) holds, $\mathcal{O}_n \subset M$ open and dense $\Rightarrow F_n := \mathcal{O}_n^c$ is closed and nowhere dense.

$$(\bigcap_{n\in\mathbb{N}} \mathcal{O}_n)^c = \bigcup_{n\in\mathbb{N}} \mathcal{O}_n^c = \bigcup_{n\in\mathbb{N}} F_n$$

has empty interior, so

$$\Rightarrow \emptyset = int((\bigcap_{n\in\mathbb{N}} \mathcal{O}_n)^c) \underbrace{=}_{(I.8)} \overline{(\bigcap_{n\in\mathbb{N}} \mathcal{O}_n)}^c$$

so $\overline{\bigcap_{n\in\mathbb{N}} \mathcal{O}_n} = M$.
Conversely, if (b) holds and F is meager, $F = \bigcup_{n\in\mathbb{N}} F_n$, F_n closed and nowhere dense, $\mathcal{O}_n = F_n^c$ is open and dense. By (I.8) $(\bar{\mathcal{O}}_n)^c = int(\mathcal{O}_n^c) = int F_n = \emptyset$, so $\mathcal{O}_n$ is dense in M

$$\Rightarrow \overline{\bigcap_{n\in\mathbb{N}} \mathcal{O}_n} = M \quad \text{so} \quad \emptyset = \overline{(\bigcap_{n\in\mathbb{N}} \mathcal{O}_n)}^c$$
$$= int((\bigcap_{n\in\mathbb{N}} \mathcal{O}_n)^c)$$
$$= int(\bigcup_{n\in\mathbb{N}} \mathcal{O}_n^c)$$
$$= int(\bigcup_{n\in\mathbb{N}} F_n)$$

so (a) holds.
Step 2: (b) is true: Let $\mathcal{O}_n \subset M$ open and dense, and set $D := \bigcap_{n\in\mathbb{N}} \mathcal{O}_n$.
So it is enough to show that every open ε-ball in M contains an element in D.
$\mathcal{O}_1$ is open and dense $\Rightarrow \mathcal{O}_1 \cap B_\varepsilon(x_0) \neq \emptyset \Rightarrow \exists x_1 \in \mathcal{O}_1 \cap B_\varepsilon(x_0)$ and $\exists 0 < \varepsilon_1 \leq \frac{1}{2}\varepsilon$ such that

$$B_{\overline{\varepsilon_1}}(x_0) \subset B_{2\varepsilon_1}(x_1) \subset \mathcal{O}_1 \cap B_\varepsilon(x_0).$$

Now consider $\mathcal{O}_2$ which is open and dense. As above $\Rightarrow \mathcal{O}_2 \cap B_{\varepsilon_1}(x_1) \neq \emptyset$, so $\exists x_2 \in \mathcal{O}_2, \varepsilon_2 < \frac{1}{2}\varepsilon_1$ such that

$$B_{\overline{\varepsilon_2}}(x_2) \subset B_{2\varepsilon_2}(x_2) \subset \mathcal{O}_2 \cap \mathcal{O}_1 \cap B_\varepsilon(x_0).$$

Continuing inductively, there exist sequences $(\varepsilon_n)_n, (x_n)_n$ such that

(1) $0 < \varepsilon_n < \frac{1}{2}\varepsilon_{n-1}$, in particular $\varepsilon_n < 2^{-n}\varepsilon$

(2) $B_{\overline{\varepsilon_n}}(x_n) \subset B_{2\varepsilon_n}(x_n) \subset \mathcal{O}_n \cap B_{\varepsilon_{n-1}}(x_{n-1}) \subset \cdots \subset \mathcal{O}_1 \cap \mathcal{O}_2 \ldots \mathcal{O}_n \cap B_\varepsilon(x_0)$ for all $n \in \mathbb{N}$. In particular, $x_n \subset B_{\varepsilon_N}(x_N) \subset B_{2^{-N}\varepsilon}(x_0)$ for all $n \geq N$.

$\Rightarrow (x_n)_n$ is Cauchy, M is complete, so $x = \lim x_n \in M$ exists

$$\Rightarrow x \in B_{\overline{\varepsilon_N}}(x_N) \quad \forall N \in \mathbb{N}.$$

So $D \cap B_\varepsilon(x_0) \neq \emptyset \; \forall \varepsilon > 0, x_0 \in M$, so D is dense in M. □

Corollary 6.4 (Baire)**.** *Let M be a complete metric space, $(F_n)_n \subset M$ closed such that $M = \bigcup_{n\in\mathbb{N}} F_n$. Then there exists $n_0 \in \mathbb{N} : intF_{n_0} \neq \emptyset$. So a complete metric space is not meager.*

Proof. F_n is closed. If $intF_n = \emptyset$, by Theorem 6.2(a), $F = \bigcup_{n\in\mathbb{N}} F_n$ has empty interior, so $M = F$ has empty interior, but $M^o = int(M) = M$, a contradiction. □

6.2 Application I: The set of discontinuities of a limit of continuous functions

Theorem 6.5. *Assume that $(f_n)_n : M \to \mathbb{C}$ are continuous, M is complete metric space and*

$$f(x) := \lim_{n\to\infty} f_n(x)$$

exists for all $x \in M$. Then the set of points where f is discontinuous is (at most) meager. In other words, the set of points where f is continuous is the complement of a meager set, in particular it is dense.

Proof. Let $D =$ set of discontinuities of f. The oscillations of the function f at a point x are

$$osc(f)(x) := \lim_{r\searrow 0} w(f)(r,x) = \inf_{r>0} w(f)(r,x)$$

with $w(f)(r,x) := \sup_{y,z\in B_r(x)} |f(y) - f(z)|$ (which is decreasing in r).
So $osc(f)(x) < \varepsilon \iff \exists$ ball B centered at x with $|f(y) - f(z)| < \varepsilon \; \forall z, y \in B$.
Note also

$$osc(f)(x) = 0 \quad \Leftrightarrow \quad f \text{ is continuous at } x \tag{I.9}$$

$$\forall \varepsilon > 0 \quad E_\varepsilon := \{x \in M | osc(f)(x) < \varepsilon\} \text{ is open} \tag{I.10}$$

(I.9) is immediate and for (I.10) note that if $x \in E_\varepsilon$ there exists $r > 0$ with

$$\sup_{y,z\in B_r(x)} |f(y) - f(z)| < \varepsilon.$$

So if $\tilde{x} \in B_{\frac{\varepsilon}{2}}(x)$ then $\tilde{x} \in E_\varepsilon$ since $B_{\frac{\varepsilon}{2}}(\tilde{x}) \subset B_\varepsilon(x)$ and hence

$$\sup_{y,z\in B_{\frac{\varepsilon}{2}}(\tilde{x})} |f(y) - f(z)| \leq \sup_{y,z\in B_\varepsilon(x)} |f(y) - f(z)| < \varepsilon.$$

Thus $B_{\frac{\varepsilon}{2}}(\tilde{x}) \subset E_\varepsilon$, so E_ε is open.

We need one more

Lemma 6.6. *Let $(f_n)_n$ be a sequence of continuous functions on a complete metric space M and $f_n(x) \to f(x)\ \forall x \in M$. Then given any open ball $B \subset M$ and $\varepsilon > 0$, there exist an open ball $B_0 \subset B$ and $m \in \mathbb{N}$ such that $|f_m(x) - f(x)| \leq \varepsilon\ \forall x \in B_0$.*

Proof. Let Y be a closed ball in M and recall that Y is itself a complete metric space. Define

$$\begin{aligned} E_l &:= \{x \in Y | \sup_{j,k \geq l} |f_j(x) - f_k(x)| \leq \varepsilon\} \\ &= \bigcap_{j,k \geq l} \underbrace{\{x \in Y | \ |f_j(x) - f_k(x)| \leq \varepsilon\}}_{\text{closed since } f_n \text{ is continuous}}. \end{aligned}$$

So E_l is closed and since $f_n(x) \to f(x)\ \forall x \in M$ we have $Y = \bigcup_{l=1}^{\infty} E_l$.
By Corollary 6.4 applied to Y, some set, say E_m, must contain an open ball B_0. But then

$$\sup_{j,k \geq l} |f_j(x) - f_k(x)| \leq \varepsilon \quad \forall x \in B_0$$

and letting $k \to \infty$ one sees

$$|f_m(x) - f_k(x)| \leq \varepsilon \quad \forall x \in B_0.$$

□

To finish the proof of Theorem 6.5 define

$$F_n := \{x \in M | osc(f)(x) \geq \frac{1}{n}\}.$$

So $F_n = E^c_{\frac{1}{n}}$ (from (I.10)) so F_n is closed and $D = \bigcup_{n \in \mathbb{N}} F_n$ is the set of discontinuities of f.
Final claim: Each F_n is nowhere dense!
Indeed, if not, let B be open ball with $B \subset F_n$. Then setting $\varepsilon = \frac{1}{4n}$ in Lemma 6.6, we get an open ball $B_0 \subset B$ and $m \in \mathbb{N}$ such that

$$|f_m(x) - f(x)| \leq \frac{1}{4n} \quad \forall x \in B_0.$$

f_m is continuous $\Rightarrow \exists$ ball $B' \subset B_0$ such that

$$|f_m(y) - f_m(z)| \leq \frac{1}{4n} \quad \forall y, z \in B'.$$

Then

$$\begin{aligned} |f(y) - f(z)| &\leq |f(y) - f_m(y)| + |f_m(y) - f_m(z)| + |f_m(z) - f(z)| \\ &\leq \frac{1}{4n} + \frac{1}{4n} + \frac{1}{4n} = \frac{3}{4n} < \frac{1}{n} \quad \forall y, z \in B' \subset B \subset F_n. \end{aligned}$$

So if x' is the center of B' then

$$osc(f)(x') < \frac{1}{n},$$

which contradicts $x' \in F_n$! □

6.3 Application II: Continuous but nowhere differentiable functions

Consider the complete metric space $C([0,1])$ with norm $\|f\|_\infty := \sup_{x\in[0,1]} |f(x)|$ and metric $d(f,g) = \|f-g\|_\infty$.

Theorem 6.7. *The set of functions in $C([0,1])$ which are nowhere differentiable is generic (in particular, it is dense!).*

Proof. Let $D =$ set of functions $f \in C([0,1])$ which are differentiable at at least one point. We have

$$D \subset \bigcup_{N\in\mathbb{N}} \underbrace{\{f \in C([0,1]) | \exists x^* \in [0,1] : \forall x \in [0,1] |f(x) - f(x^*)| \leq N|x - x^*|\}}_{=:E_N} \tag{I.11}$$

Claim:

(a) E_N is closed.

(b) E_N is nowhere dense, i.e., it has empty interior.

Then Theorem 6.2 yields the claim.

Proof of (a). Let $\{f_n\} \subset E_N$ with $f_n \to f$. Let x_n^* be the point for which (I.11) holds with f replaced by f_n. $[0,1]$ is compact $\Rightarrow \exists (x_{n_k}^*)$ which converges to a limit $x^* \in [0,1]$. Then

$$|f(x) - f(x^*)| \leq |f(x) - f_{n_k}(x)| + |f_{n_k}(x) - f_{n_k}(x^*)| + |f_{n_k}(x^*) - f(x^*)|. \tag{I.12}$$

Since $\|f_n - f\|_\infty \to 0$, for $\varepsilon > 0$ $\exists K$ such that

$$\forall k > K \quad |f(x) - f_{n_k}(x)| < \frac{\varepsilon}{2} \quad \text{and} \quad |f_{n_k}(x^*) - f(x^*)| < \frac{\varepsilon}{2}.$$

For the middle term in (I.12) note that $f_{n_k} \in E_N$ so

$$\begin{aligned} |f_{n_k}(x) - f_{n_k}(x^*)| &\leq |f_{n_k}(x) - f_{n_k}(x_{n_k}^*)| + |f_{n_k}(x_{n_k}^*) - f_{n_k}(x^*)| \\ &\leq N|x - x_{n_k}^*| + N|x_{n_k}^* - x^*| \end{aligned}$$

and so

$$\begin{aligned} |f(x) - f(x^*)| &\leq \varepsilon + N|x - x_{n_k}^*| + N|x_{n_k}^* - x^*| \\ &\to \varepsilon + |x - x^*| + N \cdot 0 \quad \text{as } k \to \infty. \end{aligned}$$

□

Proof of (b). Let $P \subset C([0,1])$ be the subspace of all continuous piecewise linear functions. For $0 < M$ let $P_M \subset P$ be the set of continuous piecewise linear functions with slope $\geq M$ or $\leq -M$. Think of P_M as the set of "zig-zag" functions!
Key fact: $P_M \cap E_N = \emptyset$ if $M > N$!

Lemma 6.8. *$\forall M > 0$ P_M is dense in $C([0,1])$.*

Now we finish the proof that E_N has no interior points: Let $f \in E_N$ and $\varepsilon > 0$. Fix $M > N$, then $\exists h \in P_M$ with $\|f - h\| < \varepsilon$ and $h \notin E_N$ since $P_M \cap E_N = \emptyset$ when $M > N$. So no open ball around f is entirely contained in E_n, i.e., E_N has no interior. □

Proof of Lemma 6.8. Step 1: P is dense in $C([0,1])$: Let $f \in C([0,1])$. Then f is uniformly continuous, since $[0,1]$ is compact. So there exists $g \in P$ with $\|f - g\| < \varepsilon$. Indeed, since f is uniformly continuous $\exists \delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon \quad \forall |x - y| < \delta.$$

Choose $n \in \mathbb{N}$ such that $\frac{1}{n} < \delta$ and let g be the piecewise linear function on each interval $[\frac{k}{n}, \frac{k+1}{n}], k = 0, \dots, n-1$ with $g(\frac{k}{n}) := f(\frac{k}{n})$, $g(\frac{k+1}{n}) := f(\frac{k+1}{n})$ and linearly interpolated in between. Then $\|f - g\| < \varepsilon$!

Step 2: P_M is dense in P: Let $g(x) = ax + b$ for $0 \le x \le \frac{1}{n}$ and

$$\varphi_\varepsilon(x) = g(x) + \varepsilon, \quad \psi_\varepsilon(x) = g(x) - \varepsilon.$$

Begin at $g(0)$, travel a slope $+M$ until you intersect φ_ε. Reverse direction and travel on a line segment of slope $-M$ until you intersect ψ_ε. This yields a function $h \in P_M$ with

$$\psi_\varepsilon(x) \le h(x) \le \varphi_\varepsilon(x) \quad \forall 0 \le x \le \frac{1}{n}$$

so

$$|g(x) - h(x)| \le \varepsilon \quad \text{in } [0, \frac{1}{n}].$$

Then begin at $h(\frac{1}{n})$ and repeat the argument on the interval $[\frac{1}{n}, \frac{2}{n}]$ and continue in this fashion.

$\Rightarrow$ get a function $h \in P_M$ with $\|g - h\| \le \varepsilon$. So $\|f - h\| \le 2\varepsilon$ □

□

6.4 Application III: The uniform boundedness principle

Recall: Let E, F be normed vector spaces. $L(E, F)$ = vector space of all bounded linear operators $T : E \to F$ with the norm $\|T\| := \sup_{x \in E, \|x\|_E \le 1} \|Tx\|$.

Theorem 6.9 (Banach-Steinhaus uniform boundedness principle)**.** *Let E be a Banach space and F be a normed vector space, $(T_i)_{i \in I}$ be a family (not necessarily countable) of continuous linear operators, $T_i \in L(E, F)$ $\forall i \in I$. Assume that*

$$\sup_{i \in I} \|T_i x\| < \infty \; \forall x \in E. \tag{I.13}$$

Then

$$\sup_{i \in I} \|T_i\| < \infty, \tag{I.14}$$

i.e., $\exists C < \infty : \|T_i x\| \le C \|x\| \; \forall x \in E, \forall i \in I$.

Remark 6.10. *The conclusion of Theorem 6.9 is quite remarkable and surprising. Just having the pointwise estimate* $\sup_{i\in I}\|T_i x\|$ *we get* $\sup_{i\in I}\sup_{\|x\|\le 1}\|T_i x\| < \infty$.

Proof. $\forall n \in \mathbb{N}$ let

$$F_n := \{x \in E | \forall i \in I, \|T_i x\| \le n\}.$$

F_n is closed and $\bigcup_{n\in\mathbb{N}} F_n = E$. By Corollary 6.4 $\Rightarrow \exists m \in \mathbb{N} : intF_m \ne \emptyset$. Then $\exists x_0 \in F_m, r > 0, B_r(x_0) \subset F_m$. Then

$$\|T_i(x_0 + r \cdot z)\| \le m, \quad \|z\| \le 1$$

$$\begin{aligned}\Rightarrow \|T_i(z)\| &= \frac{1}{r}\|T_i(rz)\| = \frac{1}{r}\|T_i(x_0+rz) - T_i(x_0)\| \\ &\le \frac{1}{r}\underbrace{\|T_i(x_0+rz)\|}_{\le m} + \frac{1}{r}\underbrace{\|T_i(x_0)\|}_{\le m} \le \frac{2m}{r} \quad \forall z \in E, \|z\| \le 1.\end{aligned}$$

□

Corollary 6.11. *Let* E, F *be Banach spaces,* $(T_n)_n \subset L(E,F)$ *such that for* $\forall x \in E, T_n x$ *converges and let* $Tx : \lim_{n\to\infty} T_n x$. *Then*

(a) $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$.

(b) $T \in L(E,F)$.

(c) $\|T\| \le \liminf_{n\to\infty} \|T_n\|$.

Proof. (a) Follows from Theorem 6.9 immediately.

(b) Also follows from Theorem 6.9 immediately.

(c)

$$\|Tx\| \leftarrow \|T_n x\| \le C\|x\| \; \forall x \in E$$

$$\|Tx\| \leftarrow \|T_n x\| \le \|T_n\|\|x\|$$

$$\|T\| \le \liminf \|T_n\|$$

□

Corollary 6.12. *Let* $B \subset G$ *and* G *be a normed vector space (not necessarily complete). Then the following are equivalent*

(a) B *is bounded.*

(b) $f(B)$ *is bounded for* $\forall f \in G^*$.

Proof. $(a) \Rightarrow (b)$ is obvious.
$(b) \Rightarrow (a)$: Recall that G^* is a Banach space.
For $x \in B$ and $f \in G^*$ let $T_x(f) := f(x)$, $T_x(f)$ is linear and bounded, because

$$\sup_{f \in G*, \|f\|=1} |T_x(f)| \leq \|x\| \|f\| \leq \|x\|.$$

By Theorem 6.9 and (b) with $E = G^*, F = \mathbb{F}$ and $I = B$, we conclude

$$\|T_x(f)\| \leq C\|f\| \quad \forall x \in B, \forall f \in G^*.$$

Then for $\forall x \in B$

$$\|x\| = \sup_{f \in G^*, \|f\| \leq 1} |f(x)| = \sup_{f \in G^*, \|f\| \leq 1} |T_x(f)| \leq C\|x\|.$$

□

Notation: $(x_n)_n \subset E$ converges weakly to $x \in E$ ($x_n \rightharpoonup x$) if $\forall f \in E^*$ it holds $f(x_n) \to f(x)$.

Corollary 6.13. *Weakly convergent sequences are bounded.*

Proof. If $(x_n)_n$ converges weakly, then for any $f \in E^*, (f(x_n))_n$ is bounded. The result follows from Corollary 6.12. □

Corollary 6.14 (Statement dual to 6.12)**.** *Let G be a Banach space and $B^* \subset G^*$. Then the following are equivalent*

(a) $\forall x \in G$ the set $B^*(x) := \{f(x) | f \in B^*\}$ is bounded.

(b) B^* is bounded.

Proof. $(b) \Rightarrow (a)$ is obvious ($\exists M, \|f\| \leq M \ \forall f \in B^*$).
$(a) \Rightarrow (b)$: We apply Theorem 6.9 with $E = G, F = \mathbb{F}, I = B^*$. For every $f \in B^*$ we set $T_f(x) := f(x)$, $x \in G$. Due to (a) and Theorem 6.9 $\exists C < \infty$

$$|f(x)| = |T_f(x)| \leq C\|x\| \quad \forall f \in B^*, x \in G.$$

By definition

$$\|f\| = \sup_{x \in F, \|x\| \leq 1} |f(x)| \leq C \quad \forall f \in B^*,$$

i.e., B^* is bounded. □

6.5 Application IV: The Open Mapping and the Closed Graph theorems

Theorem 6.15 (Open Mapping)**.** *Let E, F be Banach spaces and $T \in L(E, F)$ be surjective. Then there exists $C > 0$ such that*

$$T(B_1^E(0)) \supset B_c^F(0). \tag{I.15}$$

Remark 6.16. *Property (I.15) ensures that the image under T of any open set in E is open in F.*
Indeed, let U be open in E. Fix $y_0 \in T(U)$ so $y_0 = Tx_0, x_0 \in U$. Let $r_0 > 0$ be such that $B_{r_0}(x_0) \subset U$. Due to Theorem 6.15 it holds $T(B_{r_0}(0)) \supset B_{cr_0}(0)$ (use linearity).

$$\begin{aligned} B_{cr_0}(y_0) = y_0 + B_{cr_0}(0) &\subset T(x_0) + T(B_{r_0}(0)) \\ &= T(x_0 + B_{r_0}(0)) = T(\underbrace{B_{r_0}(x_0)}_{\subset U}) \end{aligned}$$

$\Rightarrow T(U)$ is open.

Corollary 6.17. *Let E, F be Banach spaces, $T \in L(E,F)$ bijective (i.e., injective and surjective). Then $T^{-1} \in L(F,E)$.*

Proof. Obviously, T^{-1} exists and it is linear. By (I.15)

$$T^{-1}T(B_1^E(0)) \supset T^{-1}B_C^F(0)$$

$$\Rightarrow B_1^E(0) \supset T^{-1}B_c(0).$$

So if $y \in F, \|y\| < C \Rightarrow \|T^{-1}(y)\| < 1$

$$\Rightarrow \|T^{-1}y\| < \frac{1}{C}, \quad \|y\| \leq 1$$

$$\|T^{-1}\| \leq 1.$$

□

Corollary 6.18. *Let E be a vector space with two norms $\|\cdot\|_1, \|\cdot\|_2$ and assume that E is complete w.r.t. either norm and there exists $C > 0$ such that $\|x\|_2 \leq C\|x\|_1 \; \forall x \in E$. Then the two norms are equivalent, i.e., there exists $C_1 > 0$ such that $\|x\|_1 \leq C_1\|x\|_2 \; \forall x \in E$.*

Proof. Apply Corollary 6.17 with $E = (E, \|\cdot\|_1), F = (E, \|\cdot\|_2), T = Id$. □

Proof of Theorem 6.15. Step 1: Assume T is linear surjective operator from E onto F. Then there exists $c > 0$ such that

$$\overline{T(B_1(0))} \supset B_{2c}(0). \tag{I.16}$$

Indeed, set

$$F_n := n\overline{T(B_1(0))}.$$

T is surjective $\Rightarrow F = \bigcup_{n=1}^{\infty} F_n$. So by Baire there exists $m \in \mathbb{N} : int(F_m) \neq \emptyset$.
By linearity $int(\overline{T(B_1(0))}) \neq \emptyset$!
Pick $c > 0$ and $y_0 \in F$ such that

$$B_{4c}(y_0) \subset \overline{T(B_1(0))},$$

in particular

$$y_0 \in \overline{T(B_1(0))}. \tag{I.17}$$

By symmetry

$$-y_0 \in \overline{T(B_1(0))}. \tag{I.18}$$

Adding (I.17) and (I.18) we get

$$B_{4c}(0) \subset \overline{T(B_1(0))} + \overline{T(B_1(0))}$$

and since $\overline{T(B_1(0))}$ is convex,

$$\overline{T(B_1(0))} + \overline{T(B_1(0))} = 2\overline{T(B_1(0))}$$

so (I.16) holds.
Step 2: Assume $T \in L(E, F)$ and (I.16) holds. Then (I.15) holds, i.e. $\overline{T(B_1(0))} \supset B_c(0)$. Indeed, choose any $y \in F, \|y\| < c$.
Aim: Find some $x \in E$ such that $\|x\| < 1$ and $Tx = y$ (because then (I.15) holds! (why?)).
By (I.16) we know that

$$\begin{aligned} \forall \alpha > 0 \quad \text{and} \quad \tilde{y} \in F \text{ with } \|\tilde{y}\| < \alpha C \\ \exists z \in E \text{ with } \|z\| < \frac{\alpha}{2} \text{ and } \|\tilde{y} - Tz\| < \varepsilon. \end{aligned} \tag{I.19}$$

(Hint: Use (I.16) and linearity to see this)
Choosing $\varepsilon = \frac{c}{2}$ we find $z_1 \in E$ such that

$$\|z_1\| < \frac{1}{2} \quad \text{and} \quad \|y - Tz_1\| < \frac{1}{2}C.$$

Now apply (I.19) to $\tilde{y} = y - Tz_1$. Since $\|\tilde{y}\| < \frac{1}{2}C, \alpha = \frac{1}{2}$ and by (I.19) with $\varepsilon = \varepsilon_2 = \frac{C}{2^2}$, $\exists z_2 \in E$ with

$$\|z_2\| < \frac{1}{4} \quad \text{and} \quad \|\tilde{y} - Tz_2\| = \|y - Tz_1 - Tz_2\| < \frac{\alpha}{2}c = \frac{c}{4}.$$

Proceeding inductively, using (I.19) repeatedly with $\varepsilon = \varepsilon_n = \frac{c}{2^n}, \alpha = \alpha_n = \frac{1}{2^n}$ we obtain a sequence $(z_n)_n$ such that

$$\|z_n\| < \frac{1}{2^n} \quad \text{and} \quad \|y - T(z_1 + z_2 + \ldots z_n)\| < \frac{C}{2^n} \quad \forall n \in \mathbb{N}.$$

So $x_n := z_1 + \ldots z_n$ is Cauchy and hence $x_n \to x$ for some $x \in E$.
Clearly

$$\|x\| \le \sum_{n=1}^{\infty} \|z_n\| < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1$$

and since T is continuous we have $y = Tx$. □

Theorem 6.19 (Closed Graph). *Let E, F be Banach spaces and T a linear operator from E to F. Then*

$$T \text{ is continuous} \quad \iff \quad \text{The graph of } T \text{ is closed.}$$

Remark 6.20.

- *Assume that $T : E \to F$. The graph of T is the set $G(T) := \{(x, T(x)) | x \in E\} \subset E \times F$.*
- *The set $G(T) \subset E \times F$ is closed if for every sequence $(x_n)_n \subset E$ for which $x_n \to x$ and $y_n := Tx_n \to y$ we have $y = Tx$.*

Proof. "⇒": Clear by continuity of T!
"⇐": Consider the two norms on E:

$$\|x\|_1 := \|x\|_E + \|Tx\|_F \quad \text{and} \quad \|x\|_2 := \|x\|_E.$$

The norm $\|\cdot\|_1$ is called the graph norm.
E is a Banach space w.r.t. $\|\cdot\|_2$ by assumption and certainly

$$\|x\|_2 \leq \|x\|_1 \quad \forall x \in E.$$

Let $(x_n)_n \subset E$ be Cauchy w.r.t. $\|\cdot\|_2$, i.e., $\forall \varepsilon > 0\ \exists N : \|x_n - x_m\|_2 < \varepsilon\ \forall n, m \geq N$. Then $y_n := Tx_n$ is Cauchy in F and x_n is Cauchy in E. Therefore $x = \lim x_n$ and $y = \lim Tx_n$ exist. Since $G(T)$ is closed, it follows that $y = Tx$. Thus

$$\|x - x_n\|_1 = \|x - x_n\|_E + \|y - Tx_n\|_F \to 0 \quad \text{as } n \to \infty$$

so x_n converges to x also in $\|\cdot\|_1$ norm, i.e., $(E, \|\cdot\|_1)$ is complete!
By Corollary 6.18 the two norms are equivalent, i.e., there exists $c > 0$ such that

$$\|x\|_1 \leq c\|x\|_2 = c\|x\|_E$$

so

$$\|Tx\|_F \leq \|x\|_E + \|Tx\|_F = \|x\|_1 \leq c\|x\|_E.$$

□

7 Weak Topologies. Reflexive Spaces. Separable Spaces. Uniform Convexity

7.1 The coarsest topology for which a collection of maps becomes continuous

Recall: Given a set X a topology τ on X is a collection of subsets of X, called the open sets, such that

(1) $\emptyset \in \tau$, $X \in \tau$,

(2) arbitrary unions of sets in τ are in τ,

(3) finite intersections of sets in τ are in τ.

(2) is called τ is stable under arbitrary unions (or $\bigcup_{\text{arbitrary}}$), (3) is called τ is stable under finite intersections (or $\bigcap_{\text{finite}}$).
A set X with a topology τ is called topological space.

Suppose X is a set (no structure yet) and $(Y_i)_{i\in I}$ a collection of topological spaces, and $(\varphi_i)_{i\in I}$ a collection of maps $\varphi_i : X \to Y_i$.

Problem 1: Construct a topology on X that makes all the maps $(\varphi_i)_{i\in I}$ continuous. Can one find a topology on X which is most economical in the sense that it contains the fewest open sets?

Note: If X is equipped with the discrete topology, then all subsets of X are open and hence every map $\varphi_i : X \to Y_i$ is continuous. But this topology is huge!

Want: The cheapest topology! It is called the coarsest or weakest topology associated with $(\varphi_i)_{i\in I}$,

If $\omega_i \subset Y_i$ is open then $\varphi_i^{-1}(\omega_i)$ is necessarily open in τ and as ω_i varies in the open subsets of Y_i and i runs through I, one gets a family of open subsets which is necessarily open in X! Call this $(U_\lambda)_{\lambda\in\Lambda}$.
More precisely: (Y_i, τ_i) topological spaces, $\varphi_i : X \to Y_i$,

$$\Lambda := \{\lambda = (i, \omega_i) | i \in I, \omega_i \in \tau_i\},$$

$$U_\lambda = \varphi_i^{-1}(\omega_i).$$

Catch: $(U_\lambda)_{\lambda\in\Lambda}$ does not need to be a topology!

$\Rightarrow$ Problem 2: Given a set X and a family $(U_\lambda)_{\lambda\in\Lambda}$ of subsets of X, construct the cheapest topology τ on X which contains $(U_\lambda)_{\lambda\in\Lambda}$.

So τ must be stable under $\bigcap_{\text{finite}}$ and $\bigcup_{\text{arbitrary}}$ and $U_\lambda \subset \tau\ \forall\lambda \in \Lambda$.

Step 1: Consider the enlarged family of all finite intersections of sets in $(U_\lambda)_{\lambda\in\Lambda}$: $\bigcap_{\lambda\in\Gamma} U_\lambda$, where $\Gamma \subset \Lambda$ is finite. Call this family Φ. It is stable under $\bigcap_{\text{finite}}$.

Step 2: Φ need not be stable under $\bigcup_{\text{arbitrary}}$ $\Rightarrow$ consider families $\mathcal{F}$ obtained from Φ by taking arbitrary unions of sets in Φ. So $\mathcal{F}$ is stable under $\bigcup_{\text{arbitrary}}$.

Lemma 7.1. *$\tau := \bigcup_{arbitrary}\bigcap_{finite} U_\lambda$ is stable under $\bigcap_{finite}$. Hence τ is a topology!*

Proof. See any book on point set topology. $\square$

A **basis of a neighborhood** of a point $x \in X$ is a family $(U_i)_{i\in I}$ of open sets containing x, such that any open set containing x contains an open set from the basis (i.e., from $(U_i)_{i\in I}$).

Example: In a metric space X, take the open balls centered at $x \in X$.

In our situation: Given $x \in X$, V_i a neighborhood of $\varphi_i(x)$ in Y_i

$$\bigcap_{\text{finite}} \varphi_i^{-1}(V_i)$$

yields a basis of neighborhoods of x in X.

In the following, we equip X with the topology τ which is the weakest (smallest, coarsest) topology for which all the $\varphi_i : X \to Y_i$ are continuous for all $i \in I$.

Proposition 7.2. *Let $(x_n)_n \subset X$. Then $x_n \to x$ in τ (i.e., for any $U \in \tau, x \in U, x_n \in U$ for almost all n) $\iff \varphi_i(x_n) \to \varphi_i(x)$ as $n \to \infty\ \forall i \in I$.*

Proof. $\Rightarrow$: Simple, since by definition φ_i is continuous $\forall i \in I$.
$\Leftarrow$: Let U be a neighborhood of x. From the discussion above we may assume U is of the form

$$U = \bigcap_{i \in J} \varphi_i^{-1}(V_i),$$

$J \subset I$ finite, $\varphi_i(x) \in V_i \in \tau_i$. Since $\varphi_i(x_n) \to \varphi_i(x)\ \forall i \in I \Rightarrow$ for $i \in J\ \exists N_i \in \mathbb{N} : \varphi_i(x_n) \in V_i\ \forall n \geq N_i$. Choose $N := \max_{i \in J}(N_i) < \infty$

$$\Rightarrow \varphi_i(x_n) \in V_i \quad \forall i \in J, \forall n \geq N$$

$$\Rightarrow x_n \in U \quad \forall n \geq N.$$

□

Proposition 7.3. *Let Z be a topological space, $\psi : Z \to X$. Then*

$$\psi \text{ is continuous} \iff \varphi_i \circ \psi : Z \to Y_i \text{ is continuous } \forall i \in I.$$

Proof. "$\Rightarrow$": Simple: use that compositions of continuous functions are continuous.
"$\Leftarrow$": Need to show: $\psi^{-1}(U)$ is open (in Z) $\forall$ open set U in X. U has the form

$$U = \bigcup_{\text{arbitrary}} \bigcap_{\text{finite}} \varphi_i^{-1}(V_i), \quad V_i \in \tau_i$$

$$\begin{aligned} \psi^{-1}(U) &= \bigcup_{\text{arbitrary}} \bigcap_{\text{finite}} \psi^{-1}(\varphi_i^{-1}(V_i)) \\ &= \underbrace{\bigcup_{\text{arbitrary}} \underbrace{\bigcap_{\text{finite}} \underbrace{(\varphi_i \circ \psi)^{-1}(V_i)}_{\text{open in } Z}}_{\text{is open}}}_{\text{is open!}} \end{aligned}$$

so $\psi^{-1}(U)$ is open in Z, so ψ is continuous. □

7.2 The weakest topology $\sigma(E, E^*)$

Let E be a Banach space, E^* the dual, so E has a norm $\|\cdot\| = \|\cdot\|_E$, $f \in E^*$ are continuous linear functionals on E. For $f \in E^*$ let

$$\varphi_f : \begin{cases} E \to \mathbb{F} \\ x \mapsto \varphi_f(x) := f(x) \end{cases}$$

Take $I = E^*, Y_i = \mathbb{F}, X = E$ with the usual topology on $\mathbb{R}$, resp. $\mathbb{C}$.

Definition 7.4. *The weak topology $\sigma(E, E^*)$ on E is the coarsest (smallest) topology associated with the collection $(\varphi_f)_{f \in E^*}$ in the sense of Section 7.1.*

Note: Since every map $\varphi_f = f$ is continuous linear functional the weak topology is weaker (it contains fewer open sets) than the usual topology on E induced by the norm on E!

Proposition 7.5. *The weak topology $\sigma(E,E^*)$ is Hausdorff (i.e., it separates points).*

Proof. Let $x_1, x_2 \in E, x_1 \neq x_2$. We need to construct open sets $\mathcal{O}_1, \mathcal{O}_2 \in \sigma(E,E^*)$ with $x_1 \in \mathcal{O}_1, x_2 \in \mathcal{O}_2, \mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$.
By Hahn-Banach (2nd geometric form) we can strictly separate $\{x_1\}, \{x_2\}$ by some $f \in E^*$, i.e., $\exists \alpha \in \mathbb{R}$

$$Ref(x_1) < \alpha < Ref(x_2).$$

Set

$$\begin{aligned}\mathcal{O}_1 &:= \{x \in E | Ref(x) < \alpha\} = \varphi_f^{-1}((-\infty, \alpha) + i\mathbb{R}) \in \sigma(E,E^*)\\ \mathcal{O}_2 &:= \{x \in E | Ref(x) > \alpha\} = \varphi_f^{-1}((\alpha, +\infty) + i\mathbb{R}) \in \sigma(E,E^*).\end{aligned}$$

Clearly $x_1 \in \mathcal{O}_1, x_2 \in \mathcal{O}_2, \mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$. □

Proposition 7.6. *Let $x_0 \in E$. Given $\varepsilon > 0$ and finitely many $f_1, \dots, f_k \in E^*$, and let*

$$V := V(f_1, \dots f_k, \varepsilon) := \{x \in E | \, |f_i(x - x_0)| < \varepsilon, \forall i = 1, \dots, k\}.$$

Then V is a neighborhood of x_0 in $\sigma(E,E^*)$ and we get a basis of neighborhoods of x_0 in $\sigma(E,E^*)$ by varying $\varepsilon > 0, k \in \mathbb{N}$, and $f_1, \dots f_k \in E^*$.

Proof.

$$x_0 \in V = \bigcap_{i=1}^{k} \varphi_{f_i}^{-1}(\{z \in \mathbb{C} | \, |z - \alpha_i| < \varepsilon\}) \in \sigma(E,E^*), \quad \alpha_i := f_i(x_0), \text{ is open!}$$

Conversely, let $x_0 \in U \in \sigma(E,E^*)$. By definition of $\sigma(E,E^*)$, U contains an open set $W \ni x_0$ of the form

$$W = \bigcap_{\text{finite}} \varphi_{f_i}^{-1}(V_i),$$

V_i neighborhood of $f_i(x_0) = \alpha_i$ in $\mathbb{F}$.

$$\Rightarrow \exists \varepsilon > 0 : \quad \{z \in \mathbb{C} | \, |z - \alpha_i| < \varepsilon\} \subset V_i \quad \forall i = 1, \dots k$$

so $x_0 \in V \subset W \subset U$. □

Notation: If $(x_n)_n \subset X$ converges to x in the weak topology $\sigma(E,E^*)$, we write $x_n \rightharpoonup x$ (or $x_n \rightharpoonup x$ in $\sigma(E,E^*)$, or $x_n \rightharpoonup x$ weakly in $\sigma(E,E^*)$, or $x_n \rightharpoonup x$ weakly). We say that $x_n \to x$ strongly if $\|x_n - x\| \to 0$ (usual convergence in E).

Proposition 7.7. *Let $(x_n)_n \subset E$ be a sequence. Then*

(a) $x_n \rightharpoonup x$ weakly $\iff$ $f(x_n) \to f(x) \, \forall f \in E^*$.

(b) If $x_n \to x$ strongly then $x_n \rightharpoonup x$ weakly.

(c) If $x_n \rightharpoonup x$ weakly, then $(\|x_n\|)_n$ is bounded in $\mathbb{R}$ and $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.

(d) If $x_n \rightharpoonup x$ weakly and $f_n \to f$ strongly in E^* (i.e., $\|f_n - f\|_{E^*} \to 0$) then $f_n(x_n) \to f(x)$.

Proof. (a) Follows from the definition of $\sigma(E, E^*)$ and Proposition 7.2.

(b) By (a)

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\|_{E^*}\|x_n - x\|_E \to 0.$$

(c) Note that $\forall f \in E^*, (f(x_n))_n \subset \mathbb{F}$ is bounded. Therefore, by the uniform boundedness principle

$$\infty > \sup_{n\in N} \underbrace{\sup_{f\in E^*, \|f\|_{E^*}\leq 1} |f(x_n)|}_{=\|x_n\|_E} = \sup_{n\in N} \|x_n\|_E.$$

$$|f(x)| \leftarrow |f(x_n)| \leq \|f\|_{E^*}\|x_n\|_E \leq \|x_n\|_E \quad \text{if } \|f\|_{E^*} \leq 1.$$

$$\Rightarrow |f(x)| \leq \liminf_{n\to\infty} \|x_n\|_E \quad \forall f \in E^*, \|f\|_{E^*} \leq 1$$

$$\Rightarrow \|x\|_E = \sup_{f\in E^*, \|f\|_{E^*}\leq 1} |f(x)| \leq \liminf_{n\to\infty} \|x_n\|_E.$$

(d) Note that by (a) and (c)

$$\begin{aligned} |f_n(x_n) - f(x)| &\leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)| \\ &\leq \|f_n - f\|_{E^*}\|x_n\| + |f(x_n - x)| \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

□

Proposition 7.8. *If E is finite-dimensional then $\sigma(E, E^*)$ and the usual topology are the same, so a sequence $(x_n)_n$ converges weakly $\Leftrightarrow$ $(x_n)_n$ converges strongly.*

Proof. Since $\sigma(E, E^*)$ contains fewer open sets than the strong topology it is enough to show that every (strongly) open set is weakly open.
Let $x_0 \in E$ and U strongly open with $x_0 \in U$. Need to find $f_1, \ldots, f_k \in E^*, \varepsilon > 0$ with

$$V := V(f_1, \ldots, f_k, \varepsilon) = \{x \in E |\, |f_i(x - x_0)| < \varepsilon \text{ for all } i = 1, \ldots, k\} \subset U.$$

Let $r > 0$ such that $B_r(x_0) \subset U$. Pick a basis $e_1, e_2, \ldots, e_k$ in E such that $\|e_i\| = 1$ for all $i = 1, \ldots k$. Note that

$$x = \sum_{j=1}^{k} x_j e_j \quad \text{and} \quad x \mapsto x_j =: f_j(x)$$

are continuous linear functionals on E. Also

$$\begin{aligned}\|x - x_0\| &= \|\sum_{j=1}^{k} f_j(x - x_0)e_j\| \\ &\leq \sum_{j=1}^{k} |f_j(x - x_0)|\|e_j\| \\ &= \sum_{j=1}^{k} |f_j(x - x_0)| \leq k \cdot \varepsilon \quad \forall x \in V.\end{aligned}$$

Choose $r = \frac{\varepsilon}{k}$ to get $V \subset U$. □

Remark 7.9. *Weakly open (resp. closed) sets are always open (resp. closed) in the strong topology! If E is infinite-dimensional, the weak topology $\sigma(E, E^*)$ is strictly coarser (smaller) than the strong topology.*

Example. *Let E be infinite-dimensional. The unit sphere*

$$S := \{x \in E | \, \|x\| = 1\} \quad \Rightarrow \quad \overline{S}^{\sigma(E,E^*)} = B_E = \{x \in E | \, \|x\| \leq 1\}!$$

Proof. Step 1: $\{x \in E | \, \|x\| \leq 1\} \subset \overline{S}^{\sigma(E,E^*)}$.
Indeed, let $x_0 \in V \subset \sigma(E, E^*)$. Need to show that $V \cap S \neq \emptyset$!.
By Proposition 7.6, we may assume

$$V = \{x \in E | \, |f_i(x - x_0)| < \varepsilon \quad \forall i = 1, \ldots, k\}$$

for some $\varepsilon > 0, f_1, \ldots, f_k \in E^*$.
Claim: $\exists y_0 \in E \setminus \{0\}$ with $f_i(y_0) = 0 \; \forall i = 1, \ldots, k$.
If not, the map

$$\varphi : \begin{cases} E \to \mathbb{F}^k \\ x \mapsto \varphi(x) := (f_1(x), f_2(x), \ldots, f_k(x)) \end{cases}$$

is injective (why?) and hence φ would be injective and surjective from E onto $\varphi(E) \subset \mathbb{F}^k$. Since $\varphi(E) \subset \mathbb{F}^k$ is a Banach space, the inverse mapping theorem would give that φ and φ^{-1} are continuous so E is homeomorphic to a finite-dimensional space, hence E would be finite-dimensional. So the claim is true.
Note that $x_0 + ty_0 \in V$ for all $t \in \mathbb{R}$. Since $g(t) := \|x_0 + ty_0\|$ is continuous on $[0, \infty)$, $g(0) = \|x_0\| < 1, \lim_{t\to\infty} g(t) = \infty$, there exists $t_0 > 0 : \|x_0 + ty_0\| = 1 \Rightarrow x_0 + t_0 y_0 \subset S \cap V$. By Step 1 we see

$$S \subset B_E \subset \overline{S}^{\sigma(E,E^*)}. \tag{$*$}$$

Step 2: B_E is closed in the weak topology.
Indeed,

$$B_E = \bigcap_{f \in E^*, \|f\|_{E^*} \leq 1} \underbrace{\{x \in E | \, |f(x)| \leq 1\}}_{\text{weakly closed}} \quad \text{is weakly closed.}$$

By ($*$) and Step 2: $B_E = \overline{S}^{\sigma(E,E^*)}$ since $\overline{S}^{\sigma(E,E^*)}$ is the smallest weakly closed set containing B_E and B_E is weakly closed. □

Example. *The unit ball $U = \{x \in E | \, \|x\| < 1\}$, E infinite-dimensional, is* <u>*not*</u> *weakly open.*
Indeed, if U were weakly open then $U^c = \{x \in E | \, \|x\| \geq 1\}$ is weakly closed and hence

$$S = B_E \cap U^c$$

is weakly closed which by the previous Example it is not!

7.3 Weak topology and convex sets

Recall that every weakly closed set is strongly closed, but the converse is false if E is infinite-dimensional.
But: convex + strongly closed $\Rightarrow$ weakly closed.

Theorem 7.10. *Let $C \subset E$ be convex. Then C is closed if and only if C is weakly closed.*

Proof. "$\Leftarrow$": Clear since C^c is weakly open, hence open.
"$\Rightarrow$": Need to check that C^c is weakly open. Let $x_0 \notin C$. By Hahn-Banach, there exists a closed hyperplane which strictly separates $\{x_0\}$ and C, i.e., there exists $f \in E^*, \alpha \in \mathbb{R}$ such that

$$Ref(x_0) < \alpha < Ref(y) \quad \forall y \in C.$$

Set

$$V := \{x \in E | Ref(x) < \alpha\} \in \sigma(E, E^*).$$

Then $x_0 \in V, V \cap C = \emptyset$ so $V \subset C^c$. ☐

Remark 7.11. *The above proof shows that $C = \bigcap H_c$ where the intersection is over all closed half-spaces H_C which contain C.*

Corollary 7.12 (Mazur)**.** *Assume that $x_n \rightharpoonup x$ weakly. Then there exists a sequence $(y_n)_n$ of convex combinations of x_n which converges strongly to x.*

Proof. Let $C := conv(\bigcup_{l=1}^{\infty} \{x_l\})$ be the convex hull of x_n. Since x belongs to the weak closure of $\bigcup_{l=1}^{\infty} \{x_l\}$, it also belongs to the weak closure of C! By Theorem 7.10 we get $x \in \bar{C}$, the strong closure of C! ☐

Corollary 7.13. *Assume $\varphi : E \to (-\infty, +\infty]$ is convex and lower semi-continuous (l.s.c) in the strong topology. Then φ is l.s.c. in the weak topology.*

Proof. φ is (strongly) l.s.c. if for every sequence $(x_n)_n \subset E, x_n \to x$ one has

$$\liminf_{n \to \infty} \varphi(x_n) \geq \varphi(x)$$

and similarly for weakly l.s.c. (replace $x_n \to x$ by $x_n \rightharpoonup x$).
In terms of the level sets:

Lemma. *$\varphi : E \to (-\infty, +\infty]$ is strongly (resp. weakly) l.s.c. if for all $\lambda \in \mathbb{R}$ the sets*

$$A_\lambda := \{x \in E | \varphi(x) \leq \lambda\}$$

are strongly (resp. weakly) closed.

Proof. If φ is strongly l.s.c and $x_n \in A_\lambda$ with $x_n \to x$, then

$$\varphi(x) \leq \liminf_{n\to\infty} \underbrace{\varphi(x_n)}_{\leq\lambda} \leq \lambda$$

so $x \in A_\lambda$, i.e., A_λ is closed.

For the converse, assume that φ is not l.s.c. at some point x but A_λ is closed $\forall \lambda \in \mathbb{R}$. So there exists a sequence $(x_n)_n \subset E, x_n \to x$ and

$$\liminf_{n\to\infty} \varphi(x_n) < \varphi(x).$$

Thus there exists a subsequence, also called $(x_n)_n$, and $\lambda \in \mathbb{R}$ such that

$$\varphi(x_n) < \lambda < \varphi(x) \quad \forall n \in \mathbb{N}.$$

But then $x_n \in A_\lambda$ $\forall n \in \mathbb{N}$ and since $x_n \to x$ and A_λ is closed, also $x \in A_\lambda$, i.e., $\varphi(x) \leq \lambda$, a contradiction.

For the statement with strongly replaced by weakly, just replace $x_n \to x$ by $x_n \rightharpoonup x$ in the proof. $\square$

Continuing the proof of the Corollary, we have

$$A_\lambda = \{x \in E | \varphi(x) \leq \lambda\}$$

is closed, since φ is strongly l.s.c. Since φ is convex, we also have that A_λ is convex (why?)! So A_λ is convex and strongly closed and by Theorem 7.10 it is weakly closed! $\square$

Example. *$\varphi(x) = \|x\|$ is convex and strongly continuous so it is weakly l.s.c. Hence if $x_n \rightharpoonup x$ weakly, then $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$ (compare with Proposition 7.7).*

Theorem 7.14. *Let E, B be Banach spaces and $T : E \to B$ linear. Then T is continuous in the strong topologies on E and B if and only if T is continuous in the weak topologies on E and B.*

Proof. "$\Rightarrow$": By Proposition 7.3, we need to show that for any $f \in B^*$ the composition $f \circ T$, i.e., the map $x \mapsto f(Tx)$ is continuous from $(E, \sigma(E, E^*))$ to $\mathbb{F}$.
Since $x \mapsto f(Tx) \in E^*$ it is automatically also continuous w.r.t. $\sigma(E, E^*)$!
"$\Leftarrow$": Assume $T : (E, \sigma(E, E^*)) \to (B, \sigma(B, B^*))$ is continuous. Then

$$G(T) = \{(x, Tx) | x \in E\} \subset E \times B$$

is closed in $E \times B$ equipped with the product topology $\sigma(E, E^*) \times \sigma(B, B^*) = \sigma(E \times B, (E \times B)^*)$. So $G(T)$ is weakly closed, but then also strongly closed in $E \times B$. By the Closed graph theorem it follows that $T : E \to B$ is continuous in the strong topology. $\square$

7.4 The weak$*$ topology $\sigma(E^*, E)$

Consider the dual space E^* of a normed vector space E. So far, we have two topologies on E^*:

(a) The usual (strong) topology associated to the norm on E^*, $\|f\|_{E^*} := \sup_{\|x\|_E \leq 1} |f(x)|$.

(b) The weak topology $\sigma(E^*, E^{**})$, where $E^{**} = (E^*)^*$ is the dual of E^*, from the last sections.

Note that we can always consider E as a subset of $E^{**} = \{$ continuous linear functionals on $E^*\}$ by the following device: Given $x \in E$ let $\varphi_X : E^* \to \mathbb{F}$ be defined by

$$\varphi_x(f) := f(x).$$

Then $\varphi_x \in E^{**}$ corresponds to $x \in E$ and $x \mapsto \varphi_x$ is injective since if $\varphi_{x_1} = \varphi_{x_2}$ then for all $f \in E^*$ one has

$$f(x_1) = \varphi_{x_1} = \varphi_{x_2} = f(x_2)$$

and since E^* separates the points in E this means $x_1 = x_2$! So the map $x \mapsto \varphi_x$ yields an injection of E into E^{**}.

Definition 7.15. *The weak$*$ topology $\sigma(E^*, E)$ is the smallest topology on E^* associated with the family $(\varphi_x)_{x\in E}$, i.e., it is the smallest topology on E^* which makes all the maps $\varphi_x : E^* \to \mathbb{F}, x \in E$, continuous.*

Remark 7.16.
- *Since $E \subset E^{**}$ it is clear that $\sigma(E^*, E)$ contains fewer open sets than the weak topology $\sigma(E^*, E^{**})$ which in turn has fewer open sets that the strong topology on E^*.*
- *The reason why one wants to study these different notions of weak topologies is that the fewer open sets a topology has, the more sets are compact! Since compact sets are fundamentally important – e.g., in the proof of existence of minimizers – it is easy to understand the importance of the weak$*$ topology.*

Proposition 7.17. *The weak$*$ topology is Hausdorff.*

Proof. Given $f_1, f_2 \in E^*$ with $f_1 \neq f_2$, there exists $x \in E$ such that $f_1(x) \neq f_2(x)$ (this DOES NOT use Hahn-Banach, but just the fact that $f_1 \neq f_2$!).
W.l.o.g., we can assume that $Re\ f_1(x) \neq Re\ f_2(x)$. If not, then $Im\ f_1(x) \neq Im\ f_2(x)$ and hence

$$Re(-if_1(x)) = Im\ f_1(x) \neq Im\ f_2(x) = Re(-if_2(x))$$

so consider $-if_1, -if_2$ instead of f_1 and f_2.
W.l.o.g., $Re\ f_1(x) < Re\ f_2(x)$ and choose $\alpha \in \mathbb{R} : Re\ f_1(x) < \alpha < Re\ f_2(x)$. Set

$$\mathcal{O}_1 := \{f \in E^* | Ref(x) < \alpha\} = \varphi_x^{-1}([-\infty, \alpha) + i\mathbb{R})$$
$$\mathcal{O}_2 := \{f \in E^* | Ref(x) > \alpha\} = \varphi_x^{-1}((\alpha, \infty) + i\mathbb{R})$$

Then $\mathcal{O}_1, \mathcal{O}_2 \in \sigma(E^*, E), f_1 \in \mathcal{O}_1, f_2 \in \mathcal{O}_2$ and $\mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$. □

Proposition 7.18. *Let $f_0 \in E^*, n \in \mathbb{N}, \{x_1, x_2, \dots, x_n\} \subset E$ and $\varepsilon > 0$. Consider*

$$V = V(x_1, \dots, x_n, \varepsilon) := \{f \in E^* | \ |(f - f_0)(x_j)| < \varepsilon \text{ for all } j = 1, \dots, n\}.$$

Then V is a neighborhood of f_0 in $\sigma(E^*, E)$. Moreover, we obtain a basis of neighborhoods of f_0 in $\sigma(E, E^*)$ by varying $\varepsilon > 0, n \in \mathbb{N}$, and $x_1, \dots, x_n \in E$.

Proof. A literal transcription of the proof of Proposition 7.6. □

Notation: If a sequence $(f_n)_n \subset E^*$ converges to f in the weak$*$ topology, we write $f_n \overset{*}{\rightharpoonup} f$.

To avoid confusion, we sometimes emphasize "$f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E)$", "$f_n \rightharpoonup f$ in $\sigma(E^*, E^{**})$" and "$f_n \to f$ strongly".

Proposition 7.19. *Let $(f_n)_n \subset E$. Then*

(a) $f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E) \iff f_n(x) \to f(x), \forall x \in E$ (i.e., convergence of functionals in $\sigma(E^*, E)$ is the same as pointwise convergence of f_n to f!).

(b) If $f_n \to f$ strongly, then $f_n \rightharpoonup f$ in $\sigma(E^*, E^{**})$.
If $f_n \rightharpoonup f$ in $\sigma(E^*, E^{**})$, then $f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E)$.

(c) If $f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E)$, then $(\|f_n\|)_n$ is bounded and $\|f\| \leq \liminf \|f_n\|$.

(d) If $f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E)$ and if $x_n \to x$ strongly in E, then $f_n(x) \to f(x)$.

Proof. Copy the proof of Proposition 7.7. □

Remark 7.20. *When E is finite-dimensional, then the three topologies (strong, weak, weak$*$) on E^* coincide! Indeed, then the canonical injection $J : E \to E^{**}$ given by $x \mapsto \varphi_x, \varphi_x(f) := f(x), f \in E^*$ is surjective (since $\dim E = \dim E^{**}$) and therefore $\sigma(E^*, E) = \sigma(E^*, E^{**})$.*

The main result about compactness in the weak$*$ topology is the famous

Theorem 7.21 (Banach-Alaoglu-Bourbaki)**.** *The closed unit ball*

$$B_{E^*} := \{f \in E^* | \ \|f\|_{E^*} \leq 1\}$$

is compact in the weak$*$ topology $\sigma(E^*, E)$.

Note: This compactness property is the most essential property of the weak$*$ topology!

Proof. We will reformulate the problem slightly: Consider the cartesian product

$$Y := \mathbb{F}^E = \{\text{maps } \omega : E \to \mathbb{F}\} = (\omega(x))_{x \in E} \text{ with } \omega(x) \in \mathbb{F}.$$

We equip Y with the standard product topology, i.e., the smallest topology on Y such that the collection of maps

$$\mathbb{F}^E = Y \ni \omega \mapsto \omega(x) \in \mathbb{F}, x \in E$$

is continuous for all $x \in E$. This is the same as the topology of pointwise convergence, i.e., $(\omega_n)_n \subset Y$ converges to ω if $\forall x \in E, \omega_n(x) \to \omega(x)$ (see Munkres: Topology, A First Course, Prentice Hall, 1975 or Dixmier: General Topology, Springer 1984, or Knapp: Basic Real Analysis, Birkhäuser, 2005).

A very useful fact on product topology:

Theorem (Tychonov's theorem)**.** *An arbitrary product of compact spaces is compact in the product topology.*

Proof. See the above books. □

Note: E^* consists of very special maps from E to $\mathbb{F}$, namely the continuous linear maps. So we may consider E^* as a subset of Y!

More precisely, let

$$\Phi : E^* \to Y$$

be the canonical injection from E^* to Y given by

$$\Phi(f) := (\Phi(f)_x)_{x \in E} = (f(x))_{x \in E}.$$

Clearly, Φ is continuous from E^* into Y. To see this, simply use Proposition 7.3 and note that for each fixed $x \in E$, the map

$$E^* \ni f \mapsto (\Phi(f))_x = f(x)$$

is continuous!

Since the inverse $\Phi^{-1} : \Phi(E^*) \to E^*$ is given by

$$\omega \mapsto (E \ni x \mapsto \Phi^{-1}(\omega)(x) := \omega(x)),$$

one sees that $\Phi^{-1} : Y \supset \Phi(E^*) \to E^*$ is also continuous when Y is given the product topology. Indeed, using Proposition 7.3 again, it is enough to check, for each fixed $x \in E$, that the map $\omega \mapsto \Phi^{-1}(\omega)(x) := \omega(x))$ is continuous on $\Phi(E^*) \subset Y$. But this is obvious, since Y is given the product topology, so if $\omega_n \to \omega$ in Y then $\omega_n(x) \to \omega(x)$ for all $x \in E$, so

$$\Phi^{-1}(\omega_n)(x) = \omega_n(x) \to \omega(x) = \Phi^{-1}(\omega)(x) \quad \text{as } n \to \infty.$$

Upshot: Φ is a homeomorhism from E^* onto $\Phi(E^*) \subset Y$ where E^* is given the weak$*$ topology $\sigma(E^*, E)$ and Y is given the product topology.

Note: $\Phi(B_{E^*}) = K$, where the set $K \subset Y$ is given by

$$\begin{aligned} K = \{\omega \in Y |\, |\omega(x)| \leq \|x\|_E, \omega &\text{ is linear, i.e.,} \\ \omega(x+y) &= \omega(x) + \omega(y) \text{ and} \\ \omega(\lambda x) &= \lambda \omega(x)\ \forall \lambda \in \mathbb{F}, x, y \in E\}. \end{aligned}$$

Now we only have to check that K is a compact subset of Y!
We can write $K = K_1 \cap K_2$ where

$$K_1 = \{\omega \in Y |\, |\omega(x)| \leq \|x\|_E\ \forall x \in E\}$$

and

$$K_2 := \Phi(E^*) = \{\omega \in Y | \omega \text{ is linear}\}.$$

Note that K_1 can be written as

$$K_1 = \prod_{x\in E} [-\|x\|, \|x\|] \subset \mathbb{R}^E \quad \text{if } \mathbb{F} = \mathbb{R}$$

or

$$K_1 = \prod_{x\in E} \{z \in \mathbb{C} | \, |z| \leq \|x\|\} \subset \mathbb{C}^E \quad \text{if } \mathbb{F} = \mathbb{C}$$

and by Tychonov's theorem K_1 is a compact subset of Y!

So we only have to show that K_2 is closed (since the intersection of a closed set and a compact set is compact!).
Let

$$B_{x,y,\lambda_1,\lambda_2} := \{\omega \in Y | \omega(\lambda_1 x + \lambda_2 y) - \lambda_1\omega(x) - \lambda_2\omega(y) = 0\}$$

which are closed subsets of Y, since if $\omega_n \in B_{x,y,\lambda_1,\lambda_2}$ then, if $\omega_n \to \omega$ in Y, then

$$\begin{aligned} 0 =& \omega_n(\lambda_1 x + \lambda_2 y) - \lambda_1\omega_n(x) - \lambda_2\omega_n(y) \\ &\to \omega(\lambda_1 x + \lambda_2 y) - \lambda_1\omega(x) - \lambda_2\omega(y) \quad \text{as } n \to \infty \end{aligned}$$

so $\omega \in B_{x,y,\lambda_1,\lambda_2}$.

So

$$K_2 := \bigcap_{x,y\in E,\lambda_1,\lambda_2\in\mathbb{F}} B_{x,y,\lambda_1,\lambda_2}$$

is closed in Y!

Hence $K = K_1 \cap K_2$ is compact and so $B_{E^*} = \Phi^{-1}(K)$ is compact in E^* w.r.t $\sigma(E^*, E)$. □

7.5 Reflexive spaces

Definition 7.22. *Let E be a Banach space and $J : E \to E^{**}$ the canonical injection from E into E^{**} given by*

$$(J(x))(f) := \varphi_x(f) := f(x) \quad \forall x \in E, f \in E^*.$$

The space E is ***reflexive*** *if J is surjective, i.e., $J(E) = E^{**}$.*

Note: When E is reflexive, E^{**} is usually identified with E!

Remark 7.23. *(a) Finite-dimensional spaces are reflexive (since $dimE = dimE^* = dimE^{**}$).*
Later we will see that L^p and l^p are reflexive if $1 < p < \infty$.

(b) Every Hilbert space is reflexive.

(c) L^1, L^∞, l^1 and l^∞ are not reflexive.

$C(K)$ = space of continuous functions on an infinite compact metric space K

is not reflexive.

*(d) It is essential to use the canonical injection J in the definition of reflexive spaces. See R.C. James: A non-reflexive Banach space isometric with its second conjugate, Proc. Nat. Acad. Sci USA **37** (1951), pp 174-177, for a non-reflexive Banach space for which E is isometric to E^{**}.*

Theorem 7.24 (Kakutani)**.** *Let E be a Banach space. Then E is reflexive if and only if $B_E = \{x \in E| \, \|x\| \leq 1\}$ is compact in the weak topology $\sigma(E, E^*)$.*

Proof. "$\Rightarrow$": Here $J(B) = B_{E^{**}}$ by assumption. By Theorem 7.21 we know that $B_{E^{**}}$ is compact in the weak$*$ topology $\sigma(E^{**}, E^*)$. So it is enough to check that $J^{-1} : E^{**} \to E$ is continuous when E^{**} is equipped with the weak$*$ topology $\sigma(E^{**}, E^*)$ and E is equipped with the weak topology $\sigma(E, E^*)$.
But a map $J^{-1} : E^{**} \to E$ is continuous when E is given the weak topology if and only if $\forall f \in E^*$ the map $\xi \mapsto f(J^{-1}(\xi))$ is continuous.
Note that $f(J^{-1}(\xi)) = \xi(f), \xi \in E^{**}$ but for fixed f the map $E^{**} \ni \xi \mapsto \xi(f)$ is continuous on E^{**} with the weak$*$ topology $\sigma(E^{**}, E^*)$! So J^{-1} is continuous and $B_E = J^{-1}(B_{E^{**}})$ is compact.
"$\Leftarrow$:" We need the following two lemmata

Lemma 7.25. *Let E be a Banach space, $f_1, \dots f_k \in E^*$ and $\gamma_1, \dots, \gamma_k \in \mathbb{F}$. Then*

(a) $\forall \varepsilon \; \exists x_\varepsilon \in E$ with $\|x_\varepsilon\| \leq 1$ and $|f_l(x_\varepsilon) - \gamma| < \varepsilon \; \forall l = 1, \dots, k$

is equivalent to

(b) $|\sum_{l=1}^{k} \beta_l \gamma_l| \leq \|\sum_{l=1}^{k} \beta_l f_l\| \; \forall \beta_1, \dots, \beta_k \in \mathbb{F}$.

Proof. Only for $\mathbb{F} = \mathbb{C}$.
"$(a) \Rightarrow (b)$": Fix $\beta_1, \dots, \beta_k \in \mathbb{C}$, $S := \sum_{l=1}^{k} |\beta_l|$. By (a) we have

$$|\sum_{l=1}^{k} \beta_l f_l(x_\varepsilon) - \sum_{l=1}^{k} \beta_l \gamma_l| \leq \varepsilon S$$

and hence

$$\begin{aligned} |\sum_{l=1}^{k} \beta_l \gamma_l| &\leq |\sum_{l=1}^{k} \beta_l f_l(x_\varepsilon)| + \varepsilon S \\ &\leq \|\sum_{l=1}^{k} \beta_l f_l\|_{E^*} \|x_\varepsilon\|_E + \varepsilon S \quad \forall \varepsilon > 0. \end{aligned}$$

"$(b) \Rightarrow (a)$": We will show that not $(b) \Rightarrow$ not (a):
Let $\gamma = (\gamma_1, \dots, \gamma_k) \in \mathbb{C}^k$ and let $\varphi : E \to \mathbb{C}^k$ be given by

$$\varphi(x) := (f_1(x), f_2(x), \dots, f_k(x)).$$

Then (a) can be rephrased as follows

$$\gamma \in \overline{\varphi(B_E)} \quad \text{(closure in } \mathbb{C}^k)$$

and not (a) means $\gamma \notin \overline{\varphi(B_E)}$, i.e., $\{\gamma\}$ and $\overline{\varphi(B_E)}$ can be striclty separated in $\mathbb{C}^k$ by a hyperplane, i.e., there exist $\beta = (\beta_1, \ldots, \beta_k) \in \mathbb{C}^k = (\mathbb{C}^k)^*$ and $\alpha \in \mathbb{R}$ such that for all $x \in B_E$

$$\begin{aligned} Re(\beta(\varphi(x))) = Re(\beta \cdot \varphi(x)) &= Re \sum_{l=1}^{k} \beta_l f_l(x) \\ < \alpha < Re(\beta \cdot \gamma) &= Re \sum_{l=1}^{k} \beta_l \gamma_l. \end{aligned}$$

Therefore (take sup over $\|x\| \leq 1$)

$$\| \sum_{l=1}^{k} \beta_l f_l \| \leq \alpha < Re \sum_{l=1}^{k} \beta_l \gamma_l \leq | \sum_{l=1}^{k} \beta_l \gamma_l |,$$

i.e., not (b) is true! □

Lemma 7.26. *Let E be a Banach space. Then $J(B_E)$ is dense in $B_{E^{**}}$ w.r.t. the weak∗ topology $\sigma(E^{**}, E^*)$ on E^{**}. Consequently, $J(E)$ is dense in E^{**} w.r.t. the weak∗ topology $\sigma(E^{**}, E^*)$ on E^{**}.*

Proof. Let $\xi \in B_{E^{**}}$ and V be a neighborhood of ξ in $\sigma(E^{**}, E^*)$. Need to show $V \cap J(B_E) \neq \emptyset$. As usual, we may assume that V is of the form

$$V = \{\eta \in E^{**} | \ |(\eta - \xi)(f_j)| < \varepsilon, \ \forall j = 1, \ldots, k\}$$

for some $f_1, \ldots f_k \in E^*, \varepsilon > 0$.
We have to find $x \in B_E$ with $J(x) \in V$, i.e.,

$$|f_l(x) - \xi(f_l)| < \varepsilon \quad \forall l = 1, \ldots k.$$

Set $\gamma_l := \xi(f_l)$. By Lemma 7.25 we need to check

$$| \sum_{l=1}^{k} \beta_l \gamma_l | \leq \| \sum_{l=1}^{k} \beta_l f_l \|$$

but this is clear since

$$\sum_{l=1}^{k} \beta_l \gamma_l = \sum_{l=1}^{k} \beta_l \xi(f_l) = \xi(\sum_{l=1}^{k} \beta_l f_l) \quad (\xi \in E^{**})$$

so

$$| \sum_{l=1}^{k} \beta_l \gamma_l | = |\xi(\sum_{l=1}^{k} \beta_l f_l) \leq \| \sum_{l=1}^{k} \beta_l f_l \|_{E^*} \underbrace{\|\xi\|_{E^{**}}}_{\leq 1}.$$

□

Remark 7.27. *$J(B_E)$ is always closed in $B_{E^{**}}$ in the strong topology on E^{**}! Indeed, if $\xi_n = J(x_n) \to \xi$ then, since J is an isometry, x_n must be Cauchy in B_E, so $x_n \to x$ and $\xi = J(x)$. Thus $J(B_E)$ is not dense in $B_{E^{**}}$ in the strong topology unless $J(B_E) = B_{E^{**}}$, i.e., E is reflexive!*

Continuing the proof of Theorem 7.24 "$\Leftarrow$":
The canonical injection $J : E \to E^{**}$ is always continuous from $\sigma(E, E^*)$ into $\sigma(E^{**}, E^*)$ since for fixed $f \in E^*, x \mapsto (Jx)(f) = f(Jx)$ is continuous w.r.t. $\sigma(E, E^*)$. Assuming that B_E is weakly compact (i.e., in $\sigma(E, E^*)$ topology) we see that $J(B_E)$ is compact and thus closed in E^{**} w.r.t. $\sigma(E^{**}, E^*)$.
But by Lemma 7.26, $J(B_E)$ is dense in $B_{E^{**}}$ for the same topology! Therefore $J(B_E) = B_{E^{**}}$, hence $J(E) = E^{**}$, i.e., E is reflexive. □

Theorem 7.28. *Assume that E is a reflexive Banach space and $(x_n)_n \subset E$ a bounded sequence. Then there exists a subsequence (x_{n_l}) that converges weakly.*

Remark 7.29. *A result of Eberlein-Šmulian says that if E is a Banach space such that every bounded sequence has a weakly convergent subsequence then E is reflexive! (See Holmes: Geometric Functional Analysis and its Applications, Springer, 1975).*

Proposition 7.30. *Let E be a reflexive Banach space and $M \subset E$ a closed linear subspace of E. Then M is reflexive.*

Proof. M, equipped with the norm from E has a-priori two distinct weak topologies:

(a) the topology induced by $\sigma(E, E^*)$

(b) its own weak topology $\sigma(M, M^*)$.

Fact: these two topologies are the same since by Hahn-Banach, every continuous linear functional on M is the restriction of a continuous linear functional on E!
By Theorem 7.24 we need to check that B_M is compact in the weak topology $\sigma(M, M^*)$, or equaivalently, in the topology $\sigma(E, E^*)$! We know that B_E is compact in the weak topology and since M is (strongly) closed and convex it is also weakly closed by Theorem 7.10. So $B_M = M \cap B_E$ is weakly compact! □

Corollary 7.31. *A Banach space E is reflexive if and only if E^* is reflexive.*

Proof. "$\Rightarrow$": Roughly: $E^{**} = E \Rightarrow E^{***} = E^*$.
More precisely, let $J : E \to E^{**}$ be the canonical isometry. Let $\varphi \in E^{***}$. The map

$$x \mapsto f_\varphi(x) := \varphi(J(x))$$

is a continuous linear functional on E, so $f \in E^*$.
Note:

$$\varphi(J(x)) = f(x) = (J(x))(f) \quad \forall x \in E, J(x) \in E^{**}. \tag{$*$}$$

By assumption $J : E \to E^{**}$ is surjective so for every $\xi \in E^{**}$ $\exists x \in E, \xi = J(x)$. So ($*$) yields

$$\varphi(\xi) = \xi(f) \quad \forall \xi \in E^{**},$$

i.e., the canonical injection $E^* \to E^{***}$ is surjective.
"$\Leftarrow$": Let E^* be reflexive. By "$\Rightarrow$" above we know that E^{**} is reflexive. Since $J(E) \subset E^{**}$ is a closed subspace in the strong topology, Theorem 7.30 yields that $J(E)$ is reflexive. Thus E is reflexive! □

Corollary 7.32. *Let E be a reflexive Banach space, $K \subset E$ a bounded, closed and convex subset. Then K is compact in the weak topology $\sigma(E, E^*)$.*

Proof. By Theorem 7.10 K is closed in the weak topology. Since K is bounded there exists $m \in \mathbb{N}$ with $K \subset mB_E$ and mB_E is weakly compact by Theorem 7.24. So K is a weakly closed subset of a weakly compact set and thus K is weakly compact. □

Corollary 7.33. *Let E be a reflexive Banach space and let $A \subset E$ be non-empty, closed and convex. Let $\varphi : A \to (-\infty, \infty]$ be a convex lower semi-continuous (l.s.c.) function such that $\varphi \not\equiv +\infty$ and*

$$\lim_{x \in A, \|x\| \to \infty} \varphi(x) = \infty \quad (\text{no assumption if } A \text{ is bounded}). \tag{$**$}$$

Then φ achieves its minimum on A, i.e., there exists some $x_0 \in A$ such that

$$\varphi(x_0) = \inf_{x \in A} \varphi(x).$$

Proof. Fix any $a \in A$ such that $\varphi(a) < \infty$ and define

$$\tilde{A} := \{x \in A | \varphi(x) \leq \varphi(a)\}.$$

Then $\tilde{A}$ is closed, convex and bounded (by $(**)$) and thus compact in the weak topology $\sigma(E, E^*)$ by Corollary 7.32! By Corollary 7.13, φ is also l.s.c. in the weak topology $\sigma(E, E^*)$ (since φ is convex and strongly l.s.c).

Let $(x_n)_n \subset \tilde{A}$ be a minimizing sequence in $\tilde{A}$ (i.e., $x_n \in \tilde{A}, \varphi(x_n) \to \inf_{x \in \tilde{A}} \varphi(x)$). Since $\tilde{A}$ is weakly compact, $(x_n)_n$ has a weakly convergent subsequence, i.e.

$$x_0 := \text{weak} - \lim_{j \to \infty} x_{n_j} \text{ exists}$$

for some subsequence $(x_{n_j})_j$ of (x_n). Since $\tilde{A}$ is weakly closed it follows that $x_0 \in \tilde{A}$ and by the weak l.s.c. property of φ we get

$$\inf_{x \in \tilde{A}} \varphi(x) \leq \varphi(x_0) \leq \liminf_{l \to \infty} \varphi(x_{n_l}) = \inf_{x \in \tilde{A}} \varphi(x)$$

so $\varphi(x_0) = \inf_{x \in \tilde{A}} \varphi(x)$.
If $x \in A \setminus \tilde{A}$, then

$$\varphi(x_0) \leq \varphi(a) < \varphi(x),$$

thus $\varphi(x_0) < \varphi(x)\ \forall x \in A$. □

Remark 7.34. *Corollary 7.33 is the main reason why reflexive spaces and convex functions are so important in many problems in the calculus of variations.*

7.6 Separable spaces

Definition 7.35. *A metric space E is separable if there exists a countable dense subset $D \subset E$.*

Note: Many important spaces are separable. Finite-dimensional spaces are separable, also L^p and $l^p, 1 \leq p < \infty$ are separable. $C(K), K$ compact, is separable, but L^∞ and l^∞ are not separable.

Proposition 7.36. *Let E be a separable metric space and $F \subset E$ any subset. Then F is separable.*

Proof. Let $(u_n)_n \subset E$ be a countable dense subset of E and $r_m > 0, r_m \to \infty$ as $m \to \infty$. Choose any point $a_{m,n} \in B_{r_m}(u_n) \cap F$ whenever this is non-empty. Then $(a_{m,n})_{m,n}$ is countable and dense in F. □

Theorem 7.37. *Let E be a Banach space such that E^* is separable. Then E is separable.*

Remark 7.38. *The converse is not true! E.g., $E = L^1$ is separable, but $E^* = L^\infty$ is not.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be countable and dense in E^*. Since $\|f_n\| := \|f_n\|_{E^*} := \sup_{x\in E, \|x\|_E=1} |f_n(x)|$, there is some $x_n \in E$ such that

$$\|x_n\| = 1 \quad \text{and} \quad |f_n(x_n)| \geq \frac{1}{2}\|f_n\|. \tag{$*$}$$

Let L be the vector space over $\mathbb{F}$ generated by the $(x_n)_{n\in\mathbb{N}}$ (i.e., the set of finite linear combinations of the x_n).
Claim 1: L is dense in E.
Indeed, according to Remark 5.28 we have to check that any $f \in E^*$ which vanishes on L must be identically zero.
Given $\varepsilon > 0$ $\exists N \in \mathbb{N}$ such that $\|f - f_N\| < \varepsilon$. Then

$$\|f\| \leq \|f - f_N\| + \|f_N\|.$$

Note that since $f(x_N) = 0$ (f vanishes on L) and ($*$) we have

$$\frac{1}{2}\|f_N\| \leq \|f_N(x_N)\| = \|(f - f_N)(x_N)\| \leq \|f - f_N\|\|x_N\| = \|f - f_N\|.$$

So

$$\|f\| \leq \|f - f_N\| + 2\|f - f_N\| < 3\varepsilon$$

and since this holds for all $\varepsilon > 0, \|f\| = 0$, i.e., $f \equiv 0$.

If $\mathbb{F} = \mathbb{R}$, let L_0 be the vector space over $\mathbb{Q}$ generated by the $(x_n)_n$. If $\mathbb{F} = \mathbb{C}$ let L_0 be the vector space over $\mathbb{Q} + i\mathbb{Q}$ generated by the $(x_n)_n$. I.e., the set of all finite linear combinations with coefficients in $\mathbb{Q}$, resp. in $\mathbb{Q} + i\mathbb{Q}$.

Then L_0 is dense in L and hence dense in E (since L is dense in E by Claim 1).
Claim 2: L_0 is countable!
Indeed, for $n \in \mathbb{N}$ let Λ_n be the vector space over $\mathbb{Q}$, resp. over $\mathbb{Q}+i\mathbb{Q}$, generated by $(x_k)_{1\leq k\leq n}$. Λ_n is countable and

$$L_0 = \bigcup_{n\in\mathbb{N}} \Lambda_n$$

is countable, as a countable union of countable sets. □

Corollary 7.39. *Let E be a Banach space. Then E is reflexive and separable if and only if E^* is reflexive and separable.*

Proof. We already know by Theorem 7.37 and Corollary 7.31 that

$$E^*\text{reflexive and separable} \Rightarrow E\text{reflexive and separable.}$$

Conversely, if E is reflexive and separable, then $E^{**} = J(E)$ is reflexive and separable. Since $E^{**} = (E^*)^*$, the "$\Rightarrow$" direction applied to E^* yields E reflexive and separable. □

There is also a connection between separability and metrizability of the weak topologies.

Theorem 7.40. *Let E be a separable Banach space. Then B_{E^*} is metrizable in the weak∗ topology $\sigma(E^*, E)$. Conversely, if B_{E^*} is metrizable in $\sigma(E^*, E)$, then E^* is separable.*

There is a dual statement.

Theorem 7.41. *Let E be a Banach space such that E^* is separable. Then B_E is metrizable in the weak topology $\sigma(E, E^*)$. Conversely, if B_E is metrizable in $\sigma(E, E^*)$, then E^* is separable.*

Proof of Theorem 7.40. Let $(x_n)_n \subset B_E$ be a dense countable subset of B_E. For $f \in E^*$ set

$$[f] := \sum_{n=1}^{\infty} \frac{1}{2^n} |f(x_n)|.$$

Then $[\cdot]$ is a norm on E^* and $[f] \leq \|f\|_{E^*}$ (Why?). Put $d(f, g) := [f - g]$. We have to show that the topology induced bu d on B_{E^*} is the same as the weak∗ topology $\sigma(E^*, E)$ restricted to B_E.
<u>Step 1:</u> Let $f_0 \in B_{E^*}$ and V a neighborhood of f_0 in $\sigma(E^*, E)$. Have to find some $r > 0$ such that

$$U_r = \{f \in B_{E^*} | d(f, f_0) < r\} \subset V.$$

As before, we can assume that V is of the form

$$V = \{f \in B_{E^*} |\ |(f - f_0)(y_i)| < \varepsilon,\ \forall i = 1, \ldots, k\}$$

for some $\varepsilon > 0, y_1, \ldots, y_k \in E$.
W.l.o.g., $\|y_i\| \leq 1,\ i = 1, \ldots, k$.
Since $(x_n)_n$ is dense in B_E, we know that $\forall i = 1, \ldots, k, \exists n_i \in \mathbb{N}$ such that

$$\|y_i - x_{n_i}\| < \frac{\varepsilon}{4}.$$

Choose $r > 0$ small enough such that

$$2^{n_i} r < \frac{\varepsilon}{2}, \quad i = 1, \ldots, k.$$

Claim: $U_r \subset V$!
Indeed, if

$$r > d(f, f_0) = \sum_{n=1}^{\infty} \frac{1}{2^{n_i}} |(f - f_0)(x_{n_i})|$$

then

$$\frac{1}{2^{n_i}} |(f - f_0)(x_{n_i})| < r, \quad \forall i = 1, \dots, k.$$

Hence, for $i = 1, \dots, k$

$$\begin{aligned} |(f - f_0)(y_i)| &= |(f - f_0)(y_i - x_{n_i}) + (f - f_0)(x_{n_i})| \\ &\leq \underbrace{\|f - f_0\|}_{\leq 2} \underbrace{\|y_i - x_{n_i}\|}_{< \frac{\varepsilon}{4}} + \underbrace{|(f - f_0)(x_{n_i})|}_{< \frac{\varepsilon}{2}} < \varepsilon \end{aligned}$$

so $f \in V$.
Step 2: Let $f_0 \in B_{E^*}$. Given $r > 0$, we have to find some neighborhood V in $\sigma(E^*, E)$ such that

$$V \subset U = \{f \in B_{E^*} | \, d(f, f_0) < r\}.$$

We choose V to be of the form

$$V := \{f \in B_{E^*} | \; |(f - f_0)(x_i)| < \varepsilon\} \quad \forall i = 1, \dots, k$$

with ε and k to be determined so that $V \subset U$.
If $f \in V$, then

$$\begin{aligned} d(f, f_0) &= \sum_{n=1}^{\infty} \frac{1}{2^n} |(f - f_0)(x_n)| \\ &= \sum_{n=1}^{k} \frac{1}{2^n} \underbrace{|(f - f_0)(x_n)|}_{< \varepsilon} + \sum_{n=k+1}^{\infty} \frac{1}{2^n} \underbrace{|(f - f_0)(x_n)|}_{\leq 2} \\ &< \varepsilon + 2 \sum_{n=k+1}^{\infty} \frac{1}{2^n} = \varepsilon + \frac{1}{2^{k-1}} \end{aligned}$$

so it is enough to take $\varepsilon = \frac{r}{2}$ and $k \in \mathbb{N}$ such that $\frac{1}{2^{k-1}} < \frac{r}{2}$.

Conversely, suppose that B_{E^*} is metrizable in $\sigma(E^*, E)$ and let us prove that E is separable.
Set

$$U_n := \{f \in B_{E^*} | \, d(f, 0) < \frac{1}{n}\}$$

and let V_n be a neighborhood of 0 in $\sigma(E^*, E)$ such that $V_n \subset U_n$. Again, we may assume that V_n has the form

$$V_n := \{f \in B_{E^*} | \, |f(x)| < \varepsilon_n \; \forall x \in \Phi_n\}$$

with $\varepsilon_n > 0$ and Φ_n some finite subset of E. Set

$$D := \bigcup_{n\in\mathbb{N}} \Phi_n$$

so that D is countable.
Claim: The vector space generated by D is dense in E (this implies E is separable!).
Suppose $f \in E^*$ is such that $f(x) = 0\ \forall x \in D$. Then $f \in V_n \subset U_n\ \forall n \in \mathbb{N}$. Thus $f \equiv 0$ (i.e., *span*(D) is dense in E). □

"Proof of Theorem 7.41": The implication

$$E^* \text{ separable} \quad\Rightarrow\quad B_E \text{ is metrizable in } \sigma(E, E^*)$$

is exactly as above.
The proof of the converse is trickier (where does the above argument break down?). See Dunford-Schwartz: Linear Operators, Interscience, 1972. □

Corollary 7.42. *Let E be a Banach space and $(f_n)_n$ a bounded sequence in E^*. Then there exists a subsequence $(f_{n_l})_l$ that converges in the weak$*$ topology $\sigma(E^*, E)$.*

Proof. W.l.o.g. $\|f_n\| \le 1\ \forall n \in \mathbb{N}$. The set B_{E^*} is compact (by Banach-Alaoglu) and metrizable (by Theorem 7.40) in the weak$*$ topology $\sigma(E^*, E)$. So every sequence in B_{E^*} has a convergent subsequence! □

Proof of Theorem 7.28. Let $M_0 = span(x_n, n \in \mathbb{N})$ and $M = \overline{M_0}$. Clearly M is separable and $M \subset E$ is also reflexive (by Theorem 7.30). Thus $B_M =$ unit ball in M is compact and metrizable in the weak topology $\sigma(M, M^*)$, since M^* is separable (see Corollary 7.39 and Theorem 7.40). Hence there exists a subsequence $(x_{n_l})_l$ which converges weakly w.r.t. $\sigma(M, M^*)$ and hence $(x_{n_l})_l$ converges weakly w.r.t. $\sigma(E, E^*)$ also (see Proof of Theorem 7.30). □

7.7 Uniformly convex spaces

Definition 7.43. *A Banach space E is* ***uniformly convex*** *if $\forall\varepsilon > 0 \exists\delta > 0$ such that*

$$x, y \in E, \|x\| \le 1, \|y\| \le 1, \text{ and } \|x - y\| > \varepsilon \quad\Rightarrow\quad \left\|\frac{x+y}{2}\right\| < 1 - \delta.$$

This is a geometric property of the unit ball. If one slides a ruler of length $\varepsilon > 0$ in the unit ball, then its midpoint must stay within a ball of radius $1 - \delta$ for some $\delta > 0$, i.e., it measures how round the unit sphere is.

Example. *(1) $E = \mathbb{R}^2, \|x\|_2 = (x_1^2 + x_2^2)^{\frac{1}{2}}$ is uniformly convex. Here the curvature of the unit sphere is positive.*
But

$$\|x\|_1 = |x_1| + |x_2| \quad (\textit{Manhattan norm})$$
$$\|x\|_\infty = \max(|x_1|, |x_2|)$$

are not uniformly convex. They both have a flat surface!

(2) L^p spaces are uniformly convex for $1 < p < \infty$. Any Hilbert space is uniformly convex.

Theorem 7.44. *[Milman-Pettis) Every uniformly convex Banach space is reflexive.*

Note:

- Uniform convexity is a **geometric property of the norm**, an equivalent norm need not be uniformly convex.
  Reflexivity is a **topological property**: a reflexive space remains reflexive for an equivalent norm.
  Thus Theorem 7.44 is somewhat surprising: a geometric property implies a topological property.

- Uniform convexity is often used to prove reflexivity, but this is only sufficient. There are (weird) reflexive Banach spaces that do not have any uniformly convex equaivalent norm!

Proof. Assume E is a real Banach space. Let $\xi \in E^{**}$, $\|\xi\| = 1$ and $J : E \to E^{**}$ be the canonical injection given by

$$J(x)(f) := f(x) \quad \forall f \in E^*, x \in E.$$

Have to show: $\xi \in J(B_E)$.
Since J is an isometry, $J(B_E) \subset E^{**}$ is closed in the strong topology on E^{**}. So it is enough to show

$$\forall \varepsilon > 0 \; \exists x \in B_E \text{ such that } \|\xi - J(x)\| \leq \varepsilon. \tag{$*$}$$

Fix $\varepsilon > 0$ and let $\delta = \delta_\varepsilon > 0$ be the modulus of uniform convexity. Choose some $f \in E^*$ with $\|f\| = 1$ and

$$\xi(f) > 1 - \frac{\delta}{2} \quad \text{(if } E \text{ is real, otherwise work with } Re\xi(f)).$$

This is possible since $\|\xi\| = 1$.
Set

$$V := \{\eta \in E^{**} | \; |(\eta - \xi)(f)| < \frac{\delta}{2}$$

so $\xi \in V \in \sigma(E^{**}, E^*)$.
Since $J(B_E)$ is dense in $B_{E^{**}}$ w.r.t. weak$*$ topology $\sigma(E^{**}, E^*)$ thanks to Lemma 7.26 we have $V \cap J(B_E) \neq \emptyset$. Thus there is $x \in B_E$ such that $J(x) \in V$!
Claim: x satisfies ($*$).
If not, then $\|\xi - J(x)\| > \varepsilon$, i.e.

$$\xi \in (J(x) + \varepsilon B_{E^{**}})^c := W \in \sigma(E^{**}, E^*) \quad (\text{since } B_{E^{**}} \text{ is closed in } \sigma(E^{**}, E^*)).$$

Then, again by Lemma 7.26, it follows $V \cap W \cap J(B_E) \neq \emptyset$, i.e.

$$\exists y \in B_E \text{ such that } J(y) \in V \cap W \subset V.$$

Note: Since $J(y) \in W$, we have $\|J(x) - J(y)\| \geq \varepsilon$, and since J is isometric, we must have

$$\|x - y\| > \varepsilon. \tag{$**$}$$

Since $J(x), J(y) \in V$ we have the inequalities

$$\frac{\delta}{2} > |(J(x) - \xi)(f)| = |f(x) - \xi(f)| \geq \xi(f) - f(x)$$
$$\frac{\delta}{2} > |(J(y) - \xi)(f)| = |f(y) - \xi(f)| \geq \xi(f) - f(y)$$

$$\Rightarrow 2\xi(f) < f(x+y) + 2\delta \leq \|x+y\| + \delta$$

or

$$\left\|\frac{x+y}{2}\right\| > \xi(f) - \frac{\delta}{2} > 1 - \frac{\delta}{2} - \frac{\delta}{2} = 1 - \delta.$$

But by uniform convexity, this means

$$\|x - y\| < \varepsilon$$

contradicting ($**$). □

Proposition 7.45. *Let E be a uniformly convex Banach space and $(x_n)_n \subset E$ with $x_n \rightharpoonup x$ weakly in $\sigma(E, E^*)$ and*

$$\limsup \|x_n\| \leq \|x\|. \tag{I.20}$$

Then $x_n \to x$ strongly.

Remark. *We always have $x_n \rightharpoonup x \Rightarrow \|x\| \leq \liminf \|x_n\|$ (by Proposition 7.7), so (I.20) says that the sequence $\|x_n\|$ does not loose "mass" as $n \to \infty$.*

Proof. Assume $x \neq 0$ (otherwise trivial).
Idea: renormalize!
Set

$$\lambda_n := \max(\|x_n\|, \|x\|), \quad y_n := \frac{1}{\lambda_n} x_n, \quad y := \frac{x}{\|x\|}, \quad \text{so } \|y_n\| \leq 1, \|y\| = 1.$$

Note: $y_n \to y$ strongly implies $x_n \to x$ strongly (check this!).
Further note $\lambda_n \to \lambda$ and hence (since $x_n \rightharpoonup x$ weakly), $y_n \rightharpoonup y$ weakly (check this!). Thus

$$\frac{y_n + y}{2} \rightharpoonup y$$

and by Proposition 7.7

$$1 = \|y\| = \left\|\frac{y+y}{2}\right\| \leq \liminf \underbrace{\left\|\frac{y_n+y}{2}\right\|}_{\leq \frac{1}{2}(\|y_n\| + \|y\|) \leq 1}$$

$$\Rightarrow \left\|\frac{y_n+y}{2}\right\| \to 1 \quad \text{as } n \to \infty.$$

By the uniform convexity we get

$$\|y_n - y\| \to 0 \quad \text{as } n \to \infty,$$

i.e., $y_n \to y$ strongly. □

8 L^p spaces

Some notation: $(\Omega, \mathcal{A}, \mu)$ **measure space**, i.e., Ω is a set and

(i) $\mathcal{A}$ is a σ-algebra in Ω: a collection of subsets of Ω (so $\mathcal{A} \subset \mathcal{P}(\Omega)$) such that

  (a) $\emptyset \in \mathcal{A}$

  (b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

  (c) $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ whenever $A_n \in \mathcal{A}\ \forall n \in \mathbb{N}$

(ii) μ is a measure, i.e., $\mu : \mathcal{A} \to [0, \infty]$ with

  (a) $\mu(\emptyset) = 0$

  (b) $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum\limits_{n=1}^{\infty} \mu(A_n)$ whenever $(A_n)_n \subset \mathcal{A}$ are disjoint

We will also assume that

  (iii) Ω is σ-finite, i.e., there exist $\Omega_n \in \mathcal{A}, n \in \mathbb{N}$ which exhaust Ω, i.e., $\bigcup_{n\in\mathbb{N}} \Omega_n = \Omega$, and $\mu(\Omega_n) < \infty\ \forall n \in \mathbb{N}$.

The sets $N \in \mathcal{A}$ such that $\mu(N) = 0$ are called null sets.

A property holds almost everywhere (a.e.) or for almost all $x \in \Omega$, if it holds everywhere on $\Omega \setminus N$, where N is a null set.

See Bauer: Measure theory, 4th edition, and the handout for details on **measurable functions** $f : \Omega \to \mathbb{R}$ (or $\Omega \to \mathbb{C}$).

We denote by $L^1(\Omega, \mu)$ (or simply $L^1(\Omega)$, or just L^1) the space of integrable function from Ω to $\mathbb{R}/\mathbb{C}$.

We often write $\int f = \int f d\mu = \int\limits_{\Omega} f d\mu$,

$$\|f\|_1 = \|f\|_{L^1} = \int\limits_{\Omega} |f| d\mu = \int |f|.$$

As usual, we identify functions which coincide a.e.!

8.1 Some results from integration everyone must know

Theorem (Monotone convergence, Beppo-Levi)**.** *Let $(f_n)_n$ be a sequence of non-negative functions in L^1 which is increasing,*

$$f_1 \le f_2 \le \cdots \le f_n \le f_{n+1} \le \ldots \text{ a.e. on } \Omega,$$

and bounded, $\sup_{n\in\mathbb{N}} \int f_n d\mu < \infty$*. Then*

$$f(x) := \lim_{n\to\infty} f_n(x)$$

exists a.e., $f \in L^1$, and $\|f - f_n\|_{L^1} \to 0$.

Theorem (Dominated convergence, Lebesgue)**.** *Let $(f_n)_n \subset L^1$ be such that*

(a) $f_n(x) \to f(x)$ a.e. on Ω

(b) there exists $g \in L^1$ such that for all $n \in \mathbb{N}$

$$|f_n(x)| \le g(x) \quad a.e.$$

Then $f \in L^1$ and $\|f_n - f\|_1 \to 0$.

Lemma (Fatou). *Let $(f_n)_n \subset L^1$ with*

(a) $\forall n \in \mathbb{N} : f_n(x) \ge 0$ a.e.

(b) $\sup_{n\in\mathbb{N}} \int f_n d\mu < \infty$

Set $f(x) := \liminf_{n\to\infty} f_n(x) \le \infty$. Then $f \in L^1$ and

$$\int f d\mu \le \liminf_{n\in\mathbb{N}} \int f_n d\mu.$$

Basic example: $\Omega = \mathbb{R}^d, \mathcal{A} =$ Borel-measurable sets (or Lebesgue-measurable sets) and $\mu =$ Lebesgue measure on $\mathbb{R}^d$.

Notation: $C_c(\mathbb{R}^d) =$ space of continuous functions on $\mathbb{R}^d$ with compact support, i.e.,

$$C_c(\mathbb{R}^d) = \{f \in C(\mathbb{R}^d) | \exists K \subset \mathbb{R}^d \text{ compact such that } f(x) = 0 \; \forall x \in K^c\}.$$

Theorem (Density). *$C_c(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$, i.e., $\forall f \in L^1(\mathbb{R}^d) \forall \varepsilon > 0 \; \exists g \in C_c(\mathbb{R}^d)$ with $\|f - g\|_1 < \varepsilon$.*

The case of product measures (and spaces): $(\Omega, \mathcal{A}_1, \mu_1), (\Omega, \mathcal{A}_2, \mu_2)$ two σ-finite measure spaces

$$\Omega := \Omega_1 \times \Omega_2$$
$$\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$$
$$\mu = \mu_1 \otimes \mu_2 \quad \text{by } \mu(A_1 \times A_2) := \mu_1(A_1) \cdot \mu_2(A_2) \; \forall A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2.$$

Theorem (Tonelli). *Let $F(= F(x, y)) : \Omega_1 \times \Omega_2 \to [0, \infty]$ be measurable and*

(a) $\int_{\Omega_2} F(x, y) d\mu_2 < \infty$ a.e. in Ω_1,

(b) $\int_{\Omega_1} \Big(\int_{\Omega_2} F(x, y) d\mu_2 \Big) d\mu_1 < \infty$.

Then $F \in L^1(\Omega_1 \times \Omega_2, \mu_1 \otimes \mu_2)$ and

$$\int_{\Omega_1} \Big(\int_{\Omega_2} F(x, y) d\mu_2 \Big) d\mu_1 = \int_{\Omega_2} \Big(\int_{\Omega_1} F(x, y) d\mu_1 \Big) d\mu_2 = \int \int_{\Omega_1 \times \Omega_2} F(x, y) d(\mu_1 \otimes \mu_2).$$

Theorem (Fubini). *If $F \in L^1(\Omega_1 \times \Omega_2)$, i.e.,*

$$\int_{\Omega_1 \times \Omega_2} |F(x, y)| d(\mu_1 \otimes \mu_2) < \infty,$$

then

(a) for a.e. $x \in \Omega_1 : F(x,\cdot) \in L^1(\Omega_2)$ *and* $\int\limits_{\Omega_2} F(x,y)d\mu_2 \in L^1_x(\Omega_1)$

(b) for a.e. $y \in \Omega_2 : F(\cdot,y) \in L^1(\Omega_1)$ *and* $\int\limits_{\Omega_1} F(x,y)d\mu_1 \in L^1_y(\Omega_2)$

Moreover

$$\int\limits_{\Omega_1}\Big(\int\limits_{\Omega_2} F(x,y)d\mu_2\Big)d\mu_1 = \int\limits_{\Omega_2}\Big(\int\limits_{\Omega_1} F(x,y)d\mu_1\Big)d\mu_2$$
$$= \int\int\limits_{\Omega_1\times\Omega_2} F(x,y)d\mu_1 d\mu_2.$$

8.2 Definition and some properties of L^p spaces

Definition 8.1. • $1 \le p < \infty$:

$$L^p = L^p(\Omega,\mathbb{F}) := \{f : \Omega \to \mathbb{F} | f \text{ is measurable and } |f|^p \in L^1\},$$

$$\|f\|_p := \|f\|_{L^p} := \Big(\int\limits_\Omega |f(x)|^p d\mu\Big)^{\frac{1}{p}}.$$

• $p = \infty$:

$$L^\infty = L^\infty(\Omega,\mathbb{F}) := \{f : \Omega \to \mathbb{F} | f \text{ is measurable and there exists a constant } C < \infty \text{ such that } |f(x)| \le C a.e. \text{ on } \Omega\},$$

$$\|f\|_\infty := \|f\|_{L^\infty} := \inf(C|\ |f(x)| \le C a.e. \text{ on } \Omega\} =: esssup_{x\in\Omega}|f(x)|.$$

Remark. *If* $f \in L^\infty$ *then*

$$|f(x)| \le \|f\|_\infty \quad a.e. \text{ on } \Omega.$$

Indeed, by definition of $\|f\|_\infty$*, there exists* $C_n \searrow \|f\|_\infty$ *(e.g.* $C_n = \|f\|_\infty + \frac{1}{n}$*)*
such that

$$|f(x)| \le C_n \quad a.e. \text{ on } \Omega,$$

i.e., $\exists N_n$ *such that* $|f(x)| \le C_n\ \forall x \in \Omega \setminus N_n$ *and* $\mu(N_n) = 0$.
Set $N := \bigcup_n N_n$ *and note*

$$\mu(N) \le \sum_{n\in\mathbb{N}} \mu(N_n) = 0$$

and for all $n \in \mathbb{N}$:

$$|f(x)| \le C_n \quad \forall x \in \Omega \setminus N$$

$$\Rightarrow |f(x)| \le \|f\|_\infty \quad \forall x \in \Omega \setminus N.$$

Notation: If $1 \le p \le \infty$, then p' given by $\frac{1}{p} + \frac{1}{p'} = 1$ is the **dual exponent** of p.

Theorem 8.2 (Hölder)**.** *Let $f \in L^p$ and $g \in L^{p'}$ with $1 \le p \le \infty$. Then $fg \in L^1$ and*

$$\|fg\|_1 \le \|f\|_p \|g\|_{p'}.$$

Proof. Obvious for $p = 1$ or $p = \infty$.
So assume $1 < p < \infty$, and note that for all $a, b \ge 0$

$$\begin{aligned} ab &= \frac{1}{p}a^p + ab - \frac{1}{p}a^p \\ &\le \frac{1}{p}a^p + \sup_{b \ge 0}(ab - \frac{1}{p}a^p) \\ &= \frac{1}{p}a^p + \frac{1}{p'}b^{p'} \quad (\text{ also called Young's inequality}) \end{aligned}$$

Thus

$$\begin{aligned} |f(x)g(x)| &\le \frac{1}{p}|f(x)|^p + \frac{1}{p'}|g(x)|^{p'} \quad \text{a.e.} \\ &\in L^1 \text{ since } f \in L^p, g \in L^{p'}. \end{aligned}$$

Moreover,

$$\int |fg| d\mu \le \frac{1}{p}\|f\|_p^p + \frac{1}{p'}\|g\|_{p'}^{p'}.$$

So for $\lambda > 0$,

$$\begin{aligned} \int |fg| d\mu = \int |\lambda f \frac{1}{\lambda} g| d\mu &\le \frac{1}{p}\|\lambda f\|_p^p + \frac{1}{p'}\|\lambda^{-1} g\|_{p'}^{p'} \\ &= \frac{\lambda^p}{p}\|f\|_p^p + \frac{\lambda^{-p'}}{p'}\|g\|_{p'}^{p'} = h(\lambda). \end{aligned}$$

Minimizing over $\lambda > 0$ yields the claim, since

$$\inf_{\lambda > 0} h(\lambda) = \|f\|_p \|g\|_{p'} \quad (\text{check this!})$$

□

Remark. *There is a very useful extension of Hölder in the form: If $f_1, f_2, \ldots, f_k$ are such that $f_j \in L^{p_j}$ for $1 \le j \le k$ and $\frac{1}{p} = \sum_{j=1}^k \frac{1}{p_j}$, then $f = f_1 \cdot f_2 \cdot \ldots f_k \in L^p$ and*

$$\|f\|_p \le \prod_{j=1}^k \|f_j\|_{p_j}.$$

In particular, if $f \in L^p \cap L^q$ for some $1 \le p \le q \le \infty$, then $f \in L^r$ for all $p \le r \le q$ and

$$\|f\|_r \le \|f\|_p^\theta \|g\|_q^{1-\theta} \quad \textit{with } \frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q},\ 0 \le \theta \le 1.$$

Theorem 8.3. *L^p is a vector space and $\|\cdot\|_p$ is a norm for any $1 \le p \le \infty$.*

Proof. The cases $p = 1$ and $p = \infty$ are easy, so assume $1 < p < \infty$.
If $f, f \in L^p$, then

$$\begin{aligned}|f+g|^p &\le (|f|+|g|)^p \le (2\max(|f|,|g|))^p \\ &= 2^p \max(|f|^p,|g|^p) \le 2^p(|f|^p+|g|^p) \in L^p.\end{aligned}$$

Moreover,

$$\begin{aligned}\|f+g\|_p^p &= \int |f+g|^{p-1}|f+g|d\mu \\ &\le \int |f+g|^{p-1}|f|d\mu + \int |f+g|^{p-1}|g|d\mu. \qquad (*)\end{aligned}$$

Note that $p' = \frac{p}{p-1}$, so

$$\left(|f+g|^{p-1}\right)^{p'} = |f+g|^p \in L^1$$

so $|f+g|^{p-1} \in L^{p'}$ and by Hölder, $(*)$ yields

$$\|f+g\|_p^p \le \||f+g|^{p-1}\|_{p'}(\|f\|_p + \|g\|_p) = \|f+g\|_p^{p-1}(\|f\|_p + \|g\|_p).$$

Since $\|f+g\|_p \le \infty$, this yields

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

□

Theorem 8.4 (Fischer-Riesz)**.** *L^p is a Banach space for $1 \le p \le \infty$.*

Proof. We distinguish the cases $p = \infty$ and $1 \le p < \infty$.
Case 1: $p = \infty$: Let $(f_n)_n \subset L^\infty$ be Cauchy. Given $k \in \mathbb{N}\ \exists N_k \in \mathbb{N}$ such that $\|f_m - f_n\|_\infty \le \frac{1}{k}$ for $m, n \ge N_k$. Hence there exists a set $E_k \subset \Omega, \mu(E_k) = 0$, such that

$$|f_m(x) - f_n(x)| \le \frac{1}{k} \quad \forall x \in \Omega \setminus E_k \text{ and all } m, n \ge N_k.$$

Put $E := \bigcup_{k\in\mathbb{N}} E_k$, note $\mu(E) = 0$ and

$$\forall x \in \Omega \setminus E: \quad |f_m(x) - f_n(x)| \le \frac{1}{k} \quad \text{for all } m, n \ge N_k, \qquad (*)$$

that is, the sequence $(f_n(x))_n$ is Cauchy (in $\mathbb{R}$). So

$$f(x) := \lim_{n\to\infty} f_n(x)$$

exists for all $x \in \Omega \setminus E$ and we simply set $f(x) := 0$ for $x \in E$.
Letting $m \to \infty$ in $(*)$, we also see

$$|f(x) - f_n(x)| \le \frac{1}{k} \quad \forall x \in \Omega \setminus E \text{ and all } n \ge N_k.$$

So

$$|f(x)| \leq \underbrace{|f(x) - f_n(x)|}_{\leq \frac{1}{k}} + \underbrace{f_n(x)}_{\leq \|f_n\|_\infty} \quad \text{for a.a. } x \in \Omega.$$

Hence $f \in L^\infty$ and $\|f - f_n\|_\infty \leq \frac{1}{k}$ for all $n \geq N_k$. Thus $f_n \to f$ in L^∞!
Case 2: $1 \leq p < \infty$:
Step 1: Let $(f_n)_n \subset L^p$ be Cauchy. It is enough to show that there is a subsequence $(f_{n_l})_l$ that converges to some $f \in L^p$. Indeed, assume that $f_{n_l} \to f$ in L^p. Then

$$\|f - f_m\|_p \leq \|f - f_{n_l}\|_p + \|f_{n_l} - f_m\|_p,$$

so if $\varepsilon > 0$ there exists $N_1 \in \mathbb{N}$ such that

$$\|f - f_{n_l}\|_p < \frac{\varepsilon}{2} \quad \forall l \geq N_1$$

and there exists $N_2 \in \mathbb{N}$ such that

$$\|f_n - f_m\|_p < \frac{\varepsilon}{2} \quad \forall m, n \geq N_2.$$

Note that $n_l \geq n$ (because of subsequence) so

$$\|f_{n_l} - f_m\|_p < \frac{\varepsilon}{2} \quad \forall l, m \geq N_2.$$

Hence for $l \geq \max(N_1, N_2)$ one has

$$\|f - f_m\|_p \leq \|f - f_{n_l}\|_p + \|f_{n_l} - f_m\|_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall m \geq N_1,$$

i.e., $f_, \to f$ in L^p.
Step 2: There exists a subsequence (f_{n_l}) which converges in L^p.
Extract a subsequence (f_{n_l}) such that

$$\|f_{n_{l+1}} - f_{n_l}\|_p \leq \frac{1}{2^l} \quad \forall l \in \mathbb{N}.$$

(To see that this exists proceed inductively: Choose $n_1 \in \mathbb{N}$ such that $\|f_m - f_n\|_p < \frac{1}{2}$ $\forall m, n \geq n_1$. Then choose $n_2 \geq n_1$ such that $\|f_m - f_n\|_p < \frac{1}{2^2}$ $\forall m, n \geq n_2$, etc.).
Claim: f_{n_l} converges to some f in L^p. Indeed, writing f_l instead of f_{n_l}, we have

$$\|f_{l+1} - f_l\|_p < \frac{1}{2^l} \quad \forall l \in \mathbb{N}.$$

Set

$$g_n(x) := \sum_{l=1}^{n} |f_{l+1}(x) - f_l(x)|$$

and note that the sequence $(g_n)_n$ is increasing. Also note that

$$\|g_n\|_p \leq \sum_{l=1}^{n} \|f_{l+1} - f_l\|_p < \sum_{l=1}^{n} \frac{1}{2^l} < \sum_{l=1}^{\infty} \frac{1}{2^l} = 1.$$

So

$$\sup_n \|g_n\|_p \leq 1$$

and hence, by monotone convergence, $g_n(x)$ converges to a finite limit, say

$$g(x) = \lim_{n\to\infty} g_n(x) = \sup_n g_n(x) \quad \text{for a.a. } x.$$

If $m, n \geq 2$, then

$$\begin{aligned} |f_m(x) - f_n(x)| &\leq |f_m(x) - f_{m-1}(x)| + \cdots + |f_{n+1}(x) - f_n(x)| \\ &\leq g(x) - g_{n-1}(x) \to 0 \quad \text{a.e.} \end{aligned}$$

So for a.e. x, $(f_n(x))_n$ is Cauchy and converges to some finite limit, denoted by $f(x)$, say. Letting $m \to \infty$, we also see, for a.e. x,

$$|f(x) - f_n(x)| \leq g(x) - g_{n-1}(x) \leq g(x) \quad \text{for } n \geq 2.$$

In particular, $f \in L^p$ and, since $g^p \in L^1$ and $f(x) - f_n(x) \to 0$ a.e. as $n \to \infty$, we can also apply dominated convergence to see

$$\|f - f_n\|_p \to 0 \quad \text{as } n \to \infty.$$

□

8.3 Reflexivity, Separability. The Dual of L^p

We will consider the three cases

(A) $1 < p < \infty$

(B) $p = 1$

(C) $p = \infty$

(A) Study of L^p for $1 < p < \infty$.

This is the most favorable case: L^p is reflexive, separable, and the dual of L^p is $L^{p'}$.

Theorem 8.5. *L^p is reflexive for $1 < p < \infty$.*

Proof. Step 1: (Clarkson's first inequality) Let $2 \leq p < \infty$. Then

$$\left\|\frac{f+g}{2}\right\|_p^p + \left\|\frac{f-g}{2}\right\|_p^p \leq \frac{1}{2}(\|f\|_p^p + \|g\|_p^p) \quad \forall f, g \in L^p. \tag{1}$$

Proof of (1). Enough to show

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leq \frac{1}{2}(|a|^p + |b|^p) \quad \forall a, b \in \mathbb{R}.$$

Note that

$$\alpha^p + \beta^p \leq (\alpha^2 + \beta^2)^{\frac{p}{2}} \quad \forall \alpha, \beta \geq 0. \tag{2}$$

Indeed, if $\beta > 0$, then (2) is equivalent to

$$\left(\frac{\alpha}{\beta}\right)^p + 1 \leq \left(\left(\frac{\alpha}{\beta}\right)^2 + 1\right)^{\frac{p}{2}} \tag{3}$$

and the function $(x^2+1)^{\frac{p}{2}} - x^p - 1$ increases on $[0, \infty)$ and equals 0 at $x = 0$, so

$$(x^2+1)^{\frac{p}{2}} - x^p - 1 \geq 0 \quad \forall x \geq 0.$$

Hence (3) and thus (2) hold.

Now choose $\alpha = \left|\frac{a+b}{2}\right|, \beta = \left|\frac{a-b}{2}\right|$ in (2) to see

$$\begin{aligned}\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p &\leq \left(\left|\frac{a+b}{2}\right|^2 + \left|\frac{a-b}{2}\right|^2\right)^{\frac{p}{2}} \\ &= \left(\frac{a^2+b^2}{2}\right)^{\frac{p}{2}} \leq \frac{1}{2}(a^p + b^p),\end{aligned}$$

where in the last inequality we used the convexity of the function $x \mapsto x^{\frac{p}{2}}$ for $p \geq 2$.

□

Step 2: L^p is uniformly convex, and thus reflexive, for $2 \leq p < \infty$.
Indeed, let $f, g \in L^p, \|f\|_p \leq 1, \|g\|_p \leq 1$ and $\|f - g\| \geq \varepsilon$. Then from (1) we get

$$\left|\left|\frac{f+g}{2}\right|\right|_p^p \leq \frac{1}{2}(\|f\|_p^p + \|g\|_p^p) - \left|\left|\frac{f-g}{2}\right|\right|_p^p \leq 1 - \left(\frac{\varepsilon}{2}\right)^p$$

$$\Rightarrow \left|\left|\frac{f+g}{2}\right|\right|_p \leq \left(1 - \left(\frac{\varepsilon}{2}\right)^p\right)^{\frac{1}{p}} = 1 - \underbrace{\left(1 - \left(1 - \left(\frac{\varepsilon}{2}\right)^p\right)^{\frac{1}{p}}\right)}_{=\delta_\varepsilon > 0}.$$

So $L^p, 2 \leq p < \infty$, is uniformly convex and hence reflexive by Theorem 7.44.
Step 3: L^p is reflexive for $1 < p \leq 2$.
Indeed, let $1 < p < \infty$ and consider $T : L^p \to (L^{p'})^*, \frac{1}{p} + \frac{1}{p'} = 1$, defined as follows: given $u \in L^p$, the mapping

$$L^{p'} \ni f \mapsto \int u f d\mu$$

is a continuous linear functional on $L^{p'}$ (by Hölder) and thus defines an element $Tu \in (L^{p'})^*$ such that

$$(Tu)(f) = \int u f d\mu \quad \forall f \in L^{p'}.$$

Claim:

$$\|Tu\|_{(L^{p'})^*} = \|u\|_{L^p} \quad \forall u \in L^p.$$

Proof. By Hölder

$$|Tu(f)| = |\int ufd\mu| \leq \int |u||f|d\mu \leq \|u\|_p\|f\|_{p'} \quad \forall f \in L^{p'}$$

so

$$\|Tu\|_{(L^{p'})^*} = \sup_{\|f\|_p=1} |\int ufd\mu| \leq \|u\|_p.$$

On the other hand, given $u \in L^p$, we set

$$f_0(x) := \begin{cases} \lambda|u(x)|^{p-2}\overline{u(x)}, & \text{if } u(x) \neq 0 \\ 0, & \text{else} \end{cases}$$

and note that, since $p' = \frac{p}{p-1}$,

$$\int |f_0(x)|^{p'} d\mu = \lambda^{p'} \int (|u|^{p-1})^{p'} d\mu = \lambda^{p'} \int |u|^p d\mu = \lambda^{p'} \|u\|_p^p$$

so

$$\|f_0\|_{p'} = \lambda\|u\|_p^{p-1} = 1 \quad \text{if } \lambda = \frac{1}{\|u\|_p^{p-1}}.$$

With this choice of f, we have

$$\|Tu\|_{(L^{p'})^*} \geq |Tu(f_0)| = \|u\|_p$$

so the claim follows and $T : L^p \to (L^{p'})^*$ is an isometry!. Since L^p is a Banach space, we see that $T(L^p)$ is a closed subspace of $(L^{p'})^*$.

Now assume $1 < p \leq 2$. Since $2 < p' < \infty$, we know from Step 2, that $L^{p'}$ is reflexive. Since a Banach space E is reflexive if and only if its dual E^* is reflexive, we see that $(L^{p'})^*$ is also reflexive and since every closed subspace of a reflexive space is also reflexive, we see that $T(L^p)$ is reflexive and thus L^p too. □

□

Remark. *L^p is also uniformly convex for $1 < p \leq 2$ due to Clarkson's second inequality*

$$\Big\|\frac{f+g}{2}\Big\|_p^{p'} + \Big\|\frac{f-g}{2}\Big\|_p^{p'} \leq \Big(\frac{1}{2}(\|f\|_p^p + \|g\|_p^p)\Big)^{\frac{1}{p-1}}$$

which is trickier to prove than his first inequality.

Theorem 8.6 (Riesz representation theorem)**.** *Let $1 < p < \infty$ and $\phi \in (L^p)^*$. Then there exists a unique $u \in L^{p'}$ such that*

$$\phi(f) = \int ufd\mu.$$

Moreover,

$$\|u\|_{p'} = \|\phi\|_{(L^p)^*}.$$

Remark. *Theorem 8.6 is extremely important! It says that every continuous linear functional on L^p with $1 < p < \infty$ can be represented in a "concrete way" as an integral. The mapping $\phi \mapsto u$ is linear and surjective and allows us to identify the abstract space $(L^p)^*$ with $L^{p'}$! It is the sole reason why one always makes identification $(L^p)^* = L^{p'}$ for $1 < p < \infty$.*

Proof. Consider $T : L^{p'} \to (L^p)^*$ defined by

$$Tu(f) := \int uf d\mu \quad \forall u \in L^{p'}, f \in L^p$$

and note that by Step 3 in the proof of Theorem 8.5 one has

$$\|Tu\|_{(L^p)^*} = \|u\|_{p'} \quad \forall u \in L^{p'}.$$

So we only have to check that T is surjective. Indeed, let $E = T(L^{p'})$ which is a closed subspace of $(L^p)^*$. So it is enough to show that E is dense in $(L^p)^*$. For this, let $h \in (L^p)^{**}$ satisfy

$$h(\phi) = 0 \quad \forall \phi \in E,$$

i.e., $h(Tu) = 0 \; \forall u \in L^{p'}$. Since L^p is reflexive, $h \in L^p$ and

$$h(Tu) = Tu(h) = \int uh d\mu.$$

So we have

$$\int uh d\mu = 0 \quad \forall u \in L^{p'}.$$

Choosing

$$u = |h|^{p-2}\bar{h} \in L^{p'}$$

one sees

$$0 = \int uh d\mu = \int |h|^p d\mu$$

so $h = 0$. Hence every continuous linear functional on $E \subset (L^p)^*$ vanishes on $(L^p)^*$, so E is dense in $(L^p)^*$. □

Theorem 8.7. *The space $C_c(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$ for every $1 \leq p < \infty$.*

Some notations:

- Truncation operator $T_n : \mathbb{R} \to \mathbb{R}$,

$$T_n(r) := \begin{cases} r, & \text{if } |r| \leq n, \\ \frac{nr}{|r|}, & \text{if } |r| > n. \end{cases}$$

- Characteristic function: for $E \subset \Omega$ let

$$\mathbb{1}_E(x) = \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{else.} \end{cases}$$

Proof of Theorem 8.7. Step 1: $L^p \cap L^\infty_c$ is dense in L^p. (L^∞_c = bounded functions with compact support).
Indeed, let $f \in L^p$. Put

$$g_n := \mathbb{1}_{B_n} T_n(f) \in L^\infty_c,$$

where $B_n = B_n(0) = \{x \in \mathbb{R}^d \,|\, |x| < n\}$. Since $|g_n| \leq |f| \in L^p \; \forall n$ and $g_n \to f$ a.e., Dominated convergence yields

$$\|g_n - f\|_p \to 0 \quad \text{as } n \to \infty.$$

Step 2: $C_c(\mathbb{R}^d)$ is dense in $L^p \cap L^\infty_c$ w.r.t. $\|\cdot\|_p$.
Indeed, let $f \in L^p \cap L^\infty_c$. Since f is bounded and has compact support, we have $f \in L^1$ also. Let $\varepsilon > 0$. By density of $C_c(\mathbb{R}^d)$ in L^1, for any $\delta > 0$ there exists $g \in C_c(\mathbb{R}^d)$ such that

$$\|f - g\|_1 < \delta.$$

W.l.o.g., we may assume that $\|g\|_\infty \leq \|f\|_\infty$, otherwise simply replace g by $T_n(g)$ with $n = \|f\|_\infty$.
Now note

$$\|f - g\|_p \leq \|f - g\|_1^{\frac{1}{p}} \|f - g\|_\infty^{1-\frac{1}{p}} \leq \delta^{\frac{1}{p}} (2\|f\|_\infty)^{1-\frac{1}{p}}.$$

Choosing δ so small that $\delta^{\frac{1}{p}} (2\|f\|_\infty)^{1-\frac{1}{p}} < \varepsilon$ we see

$$\|f - g\|_p < \varepsilon.$$

□

Theorem 8.8. *$L^p(\mathbb{R}^d)$ is separable for any $1 \leq p < \infty$.*

Remark. *As a consequence, $L^p(\Omega)$ is separable for any measurable subset $\Omega \subset \mathbb{R}^d$. Indeed, let I be the canonical isometry from $L^p(\Omega)$ into $L^p(\mathbb{R}^d)$ by extending a function $f : \Omega \to \mathbb{F}$ to $\mathbb{R}^d$ by setting it zero outside Ω. Then $L^p(\Omega)$ may be identified with a subspace of $L^p(\mathbb{R}^d)$, hence $L^p(\Omega)$ is also separable, whenever $L^p(\mathbb{R}^d)$ is! (see Theorem 7.36).*

Proof of Theorem 8.8. Let $\mathcal{R}$ be the countable family of sets of the form

$$R = \prod_{k=1}^{d} (a_k, b_k), \quad a_k, b_k \in \mathbb{Q}$$

and $\mathcal{E}$ = vector space over $\mathbb{Q}$ (or $\mathbb{Q} + i\mathbb{Q}$) generated by the functions $(\mathbb{1}_R)_{R \in \mathcal{R}}$. So $\mathcal{E}$ is countable, since $\mathcal{E}$ consists of finite linear combinations with rational coefficients of functions $\mathbb{1}_R$.
Claim: $\mathcal{E}$ is dense in $L^p(\mathbb{R}^d)$.
Indeed, given $f \in L^p(\mathbb{R}^d), \varepsilon > 0 \; \exists f_1 \in C_c(\mathbb{R}^d)$ such that $\|f - f_1\|_p < \frac{\varepsilon}{2}$.
Let $R \in \mathcal{R}$ be any cube such that $supp(f) \subset R$.
Subclaim: Given any $\delta > 0$, there exists a function $f_2 \in \mathcal{E}$ such that $\|f_1 - f_2\|_p < \delta$ and $supp(f_2) \subset R$.

Indeed, simply split R into small cubes in $\mathcal{R}$ where the oscillation (sup − inf) of f_1 is less than δ. Then

$$\|f_1 - f_2\|_p \leq \|f_1 - f_2\|_\infty |R|^{\frac{1}{p}} < \delta |R|^{\frac{1}{p}},$$

where $|R|$ = volume of R. By choosing $\delta > 0$ such that $\delta|R|^{\frac{1}{p}} < \frac{\varepsilon}{2}$ we have

$$\|f - f_2\|_p \leq \|f - f_1\|_p + \|f_1 - f_2\|_p < \varepsilon$$

and $f_2 \in \mathcal{E}$. □

(B) Study of L^1.
The dual space to L^1 is described in

Theorem 8.9 (Riesz representation theorem)**.** *Let $\phi \in (L^1)^*$. Then there exists a unique function $u \in L^\infty$ such that*

$$\phi(f) = \int u f d\mu \quad \forall f \in L^1.$$

Moreover

$$\|u\|_\infty = \|\phi\|_{(L^1)^*}.$$

Remark. *Again, Theorem 8.9 allows us to identify every abstract continuous linear functional $\phi \in (L^1)^*$ with a concrete integral. The mapping $\phi \mapsto u$, which is a linear surjective isometry allows to identify the abstract space $(L^1)^*$ with L^∞. Therefore, one usually makes the identification $(L^1)^* = L^\infty$.*

Proof. Recall that we assume that Ω is σ-finite, i.e., there exists a sequence $\Omega_n \subset \Omega$ of measurable sets such that $\Omega = \bigcup_n \Omega_n$ and $\mu(\Omega_n) < \infty$ $\forall n$. Set $\chi_n := \mathbb{1}_{\Omega_n}$.
Uniqueness of u: Suppose $u_1, u_2 \in L^\infty$ satisfy

$$\phi(f) = \int u_1 f d\mu = \int u_2 f d\mu \quad \forall f \in L^1.$$

Then $u = u_1 - u_2$ satisfies

$$\int u f d\mu = 0 \quad \forall f \in L^1. \tag{$*$}$$

Let

$$sign\ u = \begin{cases} \frac{\bar{u}}{|u|^2}, & \text{if } u \neq 0, \\ 0, & \text{if } u = 0, \end{cases}$$

and choose $f = \mathbb{1}_n sign\ u$ in ($*$). Then

$$\int\limits_{\Omega_n} |u| d\mu = 0 \quad \forall n$$

so $u = 0$ on Ω_n, hence $u = 0$.
Existence of u: Step 1: There is a function $\theta \in L^2$ such that

$$\theta(x) \geq \varepsilon_n > 0 \quad \forall x \in \Omega_n \ \forall n.$$

Indeed, let $\theta = \alpha_1$ on Ω_1, $\theta = \alpha_2$ on $\Omega_2 \setminus \Omega_1$, $\dots, \theta = \alpha_n$ on $\Omega_n \setminus \Omega_{n-1}$, etc. and adjust the constants $\alpha_n > 0$ so that $\theta \in L^2$.
Step 2: Given $\theta \in (L^1)^*$, the mapping

$$L^2 \ni f \mapsto \phi(\theta f)$$

defines a continuous linear functional on L^2! So by the Riesz representation theorem for L^2, there exists a function $v \in L^2$ such that

$$\phi(\theta f) = \int v f d\mu \quad \forall f \in L^2. \tag{$**$}$$

Set $u(x) := \frac{v(x)}{\theta(x)}$ (well-defined since $\theta > 0$ on Ω). Note that u is measurable and, with $\chi_n := \mathbb{1}_{\Omega_n}$, we have $u\chi_n \in L^2 \ \forall n$.
Claim: u has all the required properties.
Choosing $f = \chi_n \frac{g}{\theta} \in L^2$ for $g \in L^\infty$ in ($**$) we have

$$\phi(\chi_n g) = \int u \chi_n g d\mu \quad \forall g \in L^\infty. \tag{$***$}$$

Claim: $u \in L^\infty$ and $\|u\|_\infty \leq \|\phi\|_{(L^1)^*}$.

Proof. Fix $C > \|\phi\|_{(L^1)^*}$ and set

$$A := \{x \in \Omega | \ |u(x)| > C\}.$$

We need to show that $\mu(A) = 0$.
Choosing $g = \chi_A sign\ u$ in ($***$), we see

$$\begin{aligned} \int\limits_{A \cap \Omega_n} |u| d\mu = \int u \chi_n g d\mu &= \phi(\chi_n g) \\ &\leq \|\phi\|_{(L^1)^*} \|\chi_n g\|_1 \\ &= \|\phi\|_{(L^1)^*} \mu(A \cap \Omega_n). \end{aligned}$$

Note that $|u| > C$ on A, so

$$\int\limits_{A \cap \Omega_n} |u| d\mu \geq C \int\limits_{A \cap \Omega_n} d\mu = C\mu(A \cap \Omega_n)$$

and thus

$$C\mu(A \cap \Omega_n) \leq \|\phi\|_{(L^1)^*} \mu(A \cap \Omega_n),$$

so, since $C > \|\phi\|_{(L^1)^*}$, we must have

$$\mu(A \cap \Omega_n) = 0 \quad \forall n$$

and since $A = A \cap \left(\bigcup_n \Omega_n\right) = \bigcup_n A \cap \Omega_n$

$$\mu(A) = \mu(\bigcup_n A \cap \Omega_n) \leq \sum_n \mu(A \cap \Omega_n) = 0.$$

So A is a null set and $\|u\|_\infty \leq \|\phi\|_{(L^1)^*}$. □

Claim:

$$\phi(h) = \int uhd\mu \quad \forall h \in L^1. \qquad (****)$$

Indeed, choose $g = T_n h$ in $(***)$ and note that $\chi_n T_n h \to h$ in L^1.
Claim:

$$\|\phi\|_{(L^1)^*} = \|u\|_\infty.$$

Indeed, by $(****)$ one sees

$$|\phi(h)| \leq \|u\|_\infty \|h\|_1 \quad \forall h \in L^1$$

so $\|\phi\|_{(L^1)^*} \leq \|u\|_\infty$. □

Remark 8.10. *The space L^1 is never reflexive, except in the trivial case where Ω consists of a finite number of atoms. Then L^1 is finite-dimensional! Indeed, assume that L^1 is reflexive and consider two cases*

(i) $\forall \varepsilon > 0\ \exists A_\varepsilon \subset \Omega$ *measurable with* $0 < \mu(A_\varepsilon) < \varepsilon$.

(ii) $\exists \varepsilon > 0$ *such that* $\mu(A) \geq \varepsilon$ *for every measurable set* $A \subset \Omega$ *with* $\mu(A) > 0$.

In case (i) there exists a decreasing sequence A_n of measurable sets such that

$$0 < \mu(A_n) \to 0 \quad \text{as } n \to \infty.$$

(Choose first any sequence B_n such that

$$0 < \mu(B_n) < 2^{-n}$$

and set $A_n := \bigcup_{k=n}^\infty B_k$.)
Let $\chi_n := \mathbb{1}_{A_n}$ and set

$$u = \frac{\chi_n}{\|\chi_n\|_1}.$$

Since $\|u\|_1 = 1$ and since we assume that L^1 is reflexive, Theorem 7.28 applies and gives us a subsection (which we still denote by $(u_n)_n$) and $u \in L^1$ such that $u_n \rightharpoonup u$ weakly in L^1, i.e.,

$$\int u_n \phi d\mu \to \int u \phi d\mu \quad \forall \phi \in L^\infty.$$

Moreover, for fixed j and $n > j$ we have

$$\int_{A_j} u_n d\mu = \int u_n \chi_j d\mu = 1$$

so letting $n \to \infty$, we have

$$\int_{A_j} u d\mu = \int u\chi_j d\mu = \lim_{n\to\infty} \int u_n \chi_j d\mu = 1 \quad \forall j \in \mathbb{N}.$$

But, by dominated convergence, we have

$$\int u\chi_j d\mu \to 0 \quad \textit{as } j \to \infty$$

which is a contradiction. So L^1 is not reflexive.
In case (ii) the space Ω is purely atomic and consists of a countable number of distinct atoms (a_n), unless there are only finitely many atoms. In this case, L^1 is isomorphic to $l^1(\mathbb{N})$ and we need only to show that l^1 is not reflexive. Consider the canonical basis

$$e_n = (0, \ldots, 0, \underbrace{1}_{n-th\ slot}, 0, \ldots).$$

Assuming that l^1 is reflexive, there exists a subsequence (e_{n_k}) and some $x \in l^1$ such that $e_{n_k} \rightharpoonup x$ in the weak topology $\sigma(l^1, l^\infty)$, i.e.

$$\underbrace{(\varphi, e_{n_k})}_{=\sum \varphi(j)e_{n_k}(j)} \to (\varphi, x) \quad \forall \varphi \in l^\infty.$$

Choosing $\varphi = \varphi_j = (0, 0, \ldots, 0, \underbrace{1}_{j-th\ slot}, 1, \ldots)$ we get

$$(\varphi_j, x) = \lim_{k\to\infty} \underbrace{(\varphi_j, e_{n_k})}_{=1\ \forall k \geq j} = 1$$

but

$$(\varphi_j, x) = \sum_{n\geq j} x(j) \to 0 \quad \textit{as } j \to \infty,$$

since $x \in l^1$, a contradiction.

(C) Study of L^∞.
This is more complicated and we will not give a full answer. We already know $L^\infty = (L^1)^*$ by Theorem 8.9. Being a dual space, L^∞ has some nice properties, in particular

- The closed unit ball B_{L^∞} is compact in the weak∗ topology $\sigma(L^\infty, L^1)$ by Theorem 7.2.

- If $\Omega \subset \mathbb{R}^d$ is measurable and $(f_n)_n$ is a bounded sequence in $L^\infty(\Omega)$, there exists a subsequence $(f_{n_k})_k$ and some $f \in L^\infty$ such that $f_{n_k} \rightharpoonup f$ in the weak$*$ topology $\sigma(L^\infty, L^1)$. This is a consequence of Corollary 7.42 which applies, since L^∞ is the dual space of the separable space L^1.

However, L^∞ is not reflexive, except in the case where Ω consists of a finite number of points, since otherwise $L^1(\Omega)$ were reflexive (since a Banach space E is reflexive if and only if E^* is reflexive), and we know by the previous discussion that L^1 is not reflexive (Remark 8.10)! Thus, the dual space $(L^\infty)^*$ contains L^1, since $L^\infty = (L^1)^*$, and $(L^\infty)^*$ is stricIty bigger than L^1. Thus there are continuous linear functionals ϕ on L^∞ which cannot be represented as

$$\phi(f) = \int u f d\mu \quad \forall f \in L^\infty \text{ and some } u \in L^1.$$

Example. *Let $\phi_0 : C_c(\mathbb{R}^d) \to \mathbb{R}$ (or $\mathbb{C}$) be defined by*

$$\phi_0(f) := f(0) \quad \forall f \in C_c(\mathbb{R}^d).$$

This is a continuous linear functional on $C_c(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d)$ and by Hahn-Banach, we may extend ϕ_0 to a continuous linear functional ϕ on $L^\infty(\mathbb{R}^d)$ and

$$\phi(f) = f(0) \quad \forall f \in C_c(\mathbb{R}^d).$$

BUT there is no $u \in L^1$ such that

$$\phi(f) = \int u f d\mu \quad \forall f \in L^\infty. \tag{$*$}$$

Assuming that such a function $u \in L^1$ exists, we get from ($*$) that

$$\int u f dx = 0 \quad \forall f \in C_c(\mathbb{R}^d), f(0) = 0.$$

By some result from measure theory, this implies that $u = 0$ a.e. on $\mathbb{R}^d \setminus \{0\}$, hence $u = 0$ a.e. on $\mathbb{R}^d$, but then

$$\phi(f) = \int u f d\mu = 0 \quad \forall f \in L^\infty,$$

a contradicion.

Remark. *In fact, the dual space of L^∞ is the space of (complex valued) Radon measures.*

Theorem 8.11. *$L^\infty(\mathbb{R}^d)$ is not separable. (In fact, $L^\infty(\Omega)$ is not separable, except if Ω consists of a finite number of atoms).*

Lemma 8.12. *Let E be a Banach space. Assume that there exists a family $(\mathcal{O}_i)_{i\in I} \subset E$ such that*

(a) $\forall i \in I, \mathcal{O}_i \neq \emptyset$ is open

(b) $\mathcal{O}_i \cap \mathcal{O}_j = \emptyset$ *if* $i \neq j$

(c) I *is uncountable*

Then E *is not separable!*

Proof. Suppose that E is separable and let $(u_n)_{n\in\mathbb{N}}$ be a dense countable set in E. For each $i \in I$ the set $\mathcal{O}_i \cap (u_n)_{n\in\mathbb{N}} \neq \emptyset$ so we can choose $n(i)$ such that $u_{n(i)} \in \mathcal{O}_i$.
Note that the map $I \ni i \mapsto n(i) \in \mathbb{N}$ is injective, since, if $n(i) = n(j)$, then

$$u_{n(i)} = u_{n(j)} \in \mathcal{O}_i \cap \mathcal{O}_j$$

so by (b) we must have $i = j$!
Therefore, I is countable, a contradicion! □

Proof of Theorem 8.11. Let $I = \mathbb{R}^d$ and $\omega_i := B_1(i)$ (ball of radius one in $\mathbb{R}^d$ centered at $i \in \mathbb{R}^d$).
Note:

$$\omega_i \triangle \omega_j = (\omega_i \setminus \omega_j) \vee (\omega_j \setminus \omega_i) \neq 0 \quad \text{if } i \neq j.$$

Let

$$\mathcal{O}_i := \{f \in L^\infty(\mathbb{R}^d) |\ \|f - \mathbb{1}_{\omega_i}\|_\infty < \frac{1}{2}\}$$

and check that $(\mathcal{O}_i)_{i\in I}$ obeys the assumptions of Lemma 8.12 (for this note that $\|\mathbb{1}_{\omega_i} - \mathbb{1}_{\omega_j}\|_\infty = 1$ if $i \neq j$!) so by Lemma 8.12, L^∞ is not separable! □

	Reflexive	Separable	Dual space
$L^p, 1 < p < \infty$	YES	YES	$L^{p'}$
L^1	NO	YES	L^∞
L^∞	NO	NO	strictly bigger than L^1!

9 Hilbert spaces

9.1 Some elementary properties

Definition 9.1. *(a) Let* H *be a real vector space. A (real)* ***scalar product*** $< u, v >$ *on* H *is a bilinear form* $< \cdot, \cdot >: H \times H \to \mathbb{R}$ *that is linear in both variables such that* $\forall u, v \in H$

$$< u, v > = < v, u > \quad (symmetry)$$
$$< u, u > \geq 0 \quad (positivity)$$
$$< u, u > = 0 \Rightarrow u = 0$$

(b) If H *is a complex vector space, a (complex) scalar product on* H *is a map* $< \cdot, \cdot >: H \times H \to \mathbb{C}$ *such that* $\forall u, w, x \in H, \alpha, \beta \in \mathbb{C}$*:*

$$< x, \alpha u + \beta w > = \alpha < x, u > + \beta < x, w >$$
$$< u, w > = \overline{< w, u >}$$
$$< u, u > \geq 0 \quad \text{and} < u, u >= 0 \Rightarrow u = 0$$

So $<\cdot,\cdot>$ is linear in the second argument and

$$\begin{aligned} <\alpha u+\beta w, x> &= \overline{<x,\alpha u+\beta w>} \\ &= \bar{\alpha}\overline{<x,u>}+\bar{\beta}\overline{<x,w>} \\ &= \bar{\alpha}<u,x>+\bar{\beta}<w,x> \end{aligned}$$

so it is "anti"-linear in the first component.

One always has the Cauchy-Schwarz inequality

$$|<u,v>| \leq <u,u>^{\frac{1}{2}}<v,v>^{\frac{1}{2}} .$$

Proof. W.l.o.g., $u, v \neq 0$.

$$\begin{aligned} 0 \leq &<tu-sv, tu-sv> = \bar{t}<u, tu-sv> - \bar{s}<v, tu-sv> \\ &= |t|^2<u,u> - \bar{t}s<u,v> - \bar{s}t\underbrace{<v,u>}_{=\overline{<u,v>}} + |s|^2<v,v> \\ &= |t|^2<u,u> + |s|^2<v,v> - 2Re(\bar{t}s<u,v>) \\ &= |t|^2<u,u> + |s|^2<v,v> - 2Re(\bar{t}se^{i\theta}|<u,v>|) \end{aligned}$$

where θ is such that $<u,v>=|<u,v>|e^{i\theta}$.
Choose $s=re^{-i\theta}, r,t>0$ to get

$$0 \leq t^2<u,u>+r^2<v,v>-2\underbrace{Re(tr|<u,v>|)}_{=tr|<u,v>|}$$

$$\Rightarrow |<u,v>| \leq \frac{1}{2}\Big(\frac{t}{r}<u,u>+\frac{r}{t}<v,v>-<tu-re^{-i\theta}v, tu-re^{-i\theta}v>\Big).$$

Now choose t, r such that $\lambda=\frac{t}{r}=\frac{<v,v>^{\frac{1}{2}}}{<u,u>^{\frac{1}{2}}}$

$$\Rightarrow |<u,v>| \leq <u,u>^{\frac{1}{2}}<v,v>^{\frac{1}{2}} - \frac{1}{2}\underbrace{<\ldots,\cdots>}_{\geq 0}$$

so we have the inequality, and if

$$|<u,v>| = <u,u>^{\frac{1}{2}}<v,v>^{\frac{1}{2}}$$

then we must have

$$<tu-re^{-i\theta}v, tu-re^{-i\theta}v>=0$$

for some choice of $t, r>0$. So $tu-re^{-i\theta}v=0$, hence u and v are linearly dependant! ☐

Because of the Cauchy-Schwarz,

$$|u| := \sqrt{<u,u>} \quad \text{(the norm induced by } <\cdot,\cdot>\text{)}$$

is a norm (we write $|u|$ instead of $\|u\|$ if the norm comes from a scalar product). Indeed,

$$\begin{aligned}|u+v|^2 &= < u+v, u+v > = < u,u > + 2Re < u,v > + < v,v > \\ &\le |u|^2 + 2| < u,v > | + |v|^2 \\ &\le |u|^2 + 2|u||v| + |v|^2 \\ &= (|u|+|v|)^2\end{aligned}$$

so

$$|u+v| \le |u| + |v|.$$

Recall the parallelogram law

$$\begin{aligned}\left|\frac{a+b}{2}\right|^2 + \left|\frac{a-b}{2}\right|^2 &= \frac{1}{4}(< a+b, a+b > + < a-b, a-b >) \\ &= \frac{1}{4}(|a|^2 + < a,b > + < b,a > + |b|^2 \\ &\qquad + |a|^2 - < a,b > - < b,a > + |b|^2) \\ &= \frac{1}{2}(|a|^2 + |b|^2).\end{aligned}$$

Definition 9.2. *A Hilbert space is a (real or complex) vector space equipped with a scalar product $< \cdot, \cdot >$ such that H is complete w.r.t. the norm induced by $< \cdot, \cdot >$.*

Example.

- $L^2(\Omega)$ *with*

$$< u,v > := \int_\Omega \bar{u} v d\mu$$

is a Hilbert space.

- $l^2(\mathbb{N})$ *with*

$$< x,y > := \sum_{n\in\mathbb{N}} \bar{x_n} y_n$$

is a Hilbert space.

Proposition 9.3. *Any Hilbert space H is uniformly convex and thus reflexive.*

Proof. Let $\varepsilon > 0, u,v \in H, |u| \le 1, |v| \le 1$ and $|u-v| > \varepsilon$. Then, by the parallelogram law

$$\left|\frac{u+v}{2}\right|^2 \le 1 - \left|\frac{u-v}{2}\right|^2 < 1 - \frac{\varepsilon^2}{4}$$

so

$$\left|\frac{u+v}{2}\right| \le 1-\delta$$

with $\delta = 1 - (1 - \frac{\varepsilon^2}{4})^{\frac{1}{2}} > 0$. □

Theorem 9.4 (Projection theorem)**.** *Let H be a Hilbert space and $K \subset H, K \neq \emptyset$, a closed convex set. Then for every $f \in H$ there exists a unique $u \in K$ such that*

$$|f-u| = \inf_{v \in K} |f-v| =: dist(f,K). \tag{1}$$

Moreover, u is characterized by the property

$$u \in K \quad and \quad Re < f-u, v-u > \leq 0 \; \forall v \in K. \tag{2}$$

Notation: The above element u is called **projection** of f onto K and is denoted by

$$u =: P_K f.$$

Proof. Existence: 1st proof: The function

$$K \ni v \mapsto \varphi(v) := |f-v|$$

is convex, continuous and

$$\lim_{v \in K, |v| \to \infty} \varphi(v) = \infty.$$

So by Corollary 7.33 we know that φ attains its minimum on K since H is reflexive.
2nd proof: Now a direct argument: Let $(v_n)_n \subset K$ be a minimizing sequence for (1), i.e., $v_n \in K$ and

$$d_n := |f-v_n| \to d := \inf_{v \in K} |f-v|.$$

Claim 1: $v := \lim_{n\to\infty} v_n$ exists and $v \in K$.
Indeed, apply the parallelogram identity to $a = f - v_n$ and $b = f - v_m$ to see

$$\left|f - \frac{v_n+v_m}{2}\right|^2 + \left|\frac{v_n-v_m}{2}\right|^2 = \frac{1}{2}(|f-v_n|^2 + |f-v_m|^2) = \frac{1}{2}(d_n^2 + d_m^2).$$

Since K is convex, $\frac{v_n+v_m}{2} \in K$, so

$$\left|f - \frac{v_n+v_m}{2}\right|^2 \geq d^2$$

and hence

$$\left|\frac{v_n-v_m}{2}\right|^2 \leq \frac{1}{2}(d_n^2 + d_m^2) - d^2 \to 0 \quad \text{as } n, m \to \infty$$

so

$$\lim_{n.m\to\infty} |v_n - v_m| = 0,$$

and $(v_n)_n$ is Cauchy! Thus $v = \lim_{n\to\infty} v_n$ exists and since K is closed, $v \in K$.
Equivalence of (1) and (2): Assume $u \in K$ satisfies (1) and let $w \in K$. Then

$$v := (1-t)u + tw \in K \quad \forall t \in [0,1]$$

so

$$|f-u| \le |f-v| = |(f-u)-t(w-u)|$$

$$\Rightarrow |f-u|^2 \le |f-u|^2 - 2tRe < f-u, w-u > + t^2|w-u|^2$$

so

$$\begin{aligned} 2Re < f-u, w-u > &\le t|w-u|^2 \quad \forall t \in (0,1] \\ &\to 0 \quad \text{as } t \to 0 \end{aligned}$$

so (2) holds.
On the other hand, if (2) holds, then for $v \in K$,

$$\begin{aligned} |u-f|^2 - |v-f|^2 &= < u-f, u-f > - < v-f, v-f > \\ &= |u|^2 - 2Re < f, u > + |f|^2 - |v|^2 + 2Re < f, v > - |f|^2 \\ &= |u|^2 - |v|^2 + 2Re < f, v-u > \\ &= |u|^2 - |v|^2 + 2Re < f-u, v-u > + 2Re < u, v-u > \\ &= -|u|^2 - |v|^2 + 2Re < u, v > + 2Re < f-u, v-u > \\ &= -|u-v|^2 + \underbrace{2Re < f-u, v-u >}_{\le 0 \text{ by } (2)} \le 0, \end{aligned}$$

so (1) holds.
Uniqueness: Assume that $u_1, u_2 \in K$ satisfy (1). Then

$$Re < f-u_1, v-u_1 > \le 0 \quad \forall v \in K \tag{3}$$

$$Re < f-u_2, v-u_2 > \le 0 \quad \forall v \in K \tag{4}$$

Choose $v = u_2$ in (3) and $v = u_1$ in (4). Then

$$\begin{aligned} Re < f-u_1, u_2-u_1 > &\le 0, \\ Re < f-u_2, u_2-u_1 > &\ge 0. \end{aligned}$$

$$\begin{aligned} \Rightarrow 0 &\ge Re < f-u_1, u_2-u_1 > - Re < f-u_2, u_2-u_1 > \\ &= Re < -u_1, u_2-u_1 > + Re < u_2, u_2-u_1 > \\ &= Re < u_2-u_1, u_2-u_1 > \\ &= |u_2-u_1|^2 \ge 0 \end{aligned}$$

so $|u_2-u_1| = 0$, i.e., $u_2 = u_1$. □

Remark. *(1) It is not at all surprising to have a minimization problem related to a system of inequalities. Let $F : [0,1] \to \mathbb{R}$ be differentiable (with left and right derivatives at 1 and 0, resp.) and let $u \in [0,1]$ be a point at which F achieves its minimum. Then we have three cases:*

$$\begin{aligned} \text{either } u \in (0,1) \quad &\text{and} \quad F'(u) = 0 \\ \text{or } u = 0 \quad &\text{and} \quad F'(0) \ge 0 \\ \text{or } u = 1 \quad &\text{and} \quad F'(1) \le 1 \end{aligned}$$

All three cases can be summarized as

$$u \in [0,1] \quad \textit{and} \quad F'(u)(v-u) \geq 0 \quad \forall v \in [0,1].$$

(2) Let E be a uniformly convex Banach space, $K \subset E, K \neq \emptyset$, closed and convex. Then $\forall f \in E$ there exists a unique $u \in K$ such that

$$\|f-u\| = \inf_{v \in K} \|f-v\| =: dist(f,K).$$

Proposition 9.5. *Let $K \subset H, K \neq \emptyset$, closed and convex. Then P_K does not increase distance, i.e.,*

$$|P_K f_1 - P_K f_2| \leq |f_1 - f_2| \quad \forall f_1, f_2 \in H.$$

Proof. Let $u_j := P_K f_j$. Then as in the uniqueness proof of Theorem 9.4, we have by (2)

$$\begin{aligned} Re < f_1 - u_1, v - u_1 > &\leq 0 \quad \forall v \in K, \\ Re < f_2 - u_2, v - u_2 > &\leq 0 \quad \forall v \in K. \end{aligned}$$

Choose $v = u_2$ in the first inequality and $v = u_1$ in the second to see

$$\begin{aligned} Re < f_1 - u_1, u_2 - u_1 > &\leq 0, \\ Re < f_2 - u_2, u_2 - u_1 > &\geq 0. \end{aligned}$$

Therefore

$$\begin{aligned} 0 &\geq Re < f_1 - u_1, u_2 - u_1 > + Re < f_2 - u_2, u_1 - u_2 > \\ &= Re < f_1 - u_1 - f_2 + u_2, u_2 - u_1 > \\ &= Re < f_1 - f_2, u_2 - u_1 > - |u_2 - u_1|^2. \end{aligned}$$

So

$$\begin{aligned} |u_2 - u_1|^2 &\leq Re < f_1 - f_2, u_2 - u_1 > \\ &\leq | < f_1 - f_2, u_2 - u_1 > | \\ &\underbrace{\leq}_{\text{CSI}} |f_1 - f_2||u_2 - u_1|. \end{aligned}$$

Thus

$$|u_2 - u_1| \leq |f_1 - f_2|.$$

□

Corollary 9.6. *Assume that $M \subset H$ is a linear subspace. Let $f \in H$. Then $u = P_M f$ is characterized by*

$$u \in M \quad \textit{and} \quad < f-u, v > = 0 \quad \forall v \in M, \tag{6}$$

i.e., $f - u$ is perpendicular to all $v \in M$. Moreover, P_M is a linear operator called the ***orthogonal projection****.*

Proof. Step 1: By (2) we have

$$Re < f - u, v - u >= 0 \quad \forall v \in M.$$

Since M is a subspace, $tv \in M$ $\forall t \in \mathbb{R}, v \in M$. Hence

$$\underbrace{Re < f - u, tv - u >}_{=tRe<f-u,v>-Re<f-u,v>} \leq 0 \quad \forall t \in \mathbb{R}$$

and thus for $t > 0$:

$$Re < f - u, v > \leq \frac{1}{t} Re < f - u, v > \to 0 \quad \text{as } t \to \infty$$

so

$$Re < f - u, v > \leq 0$$

and for $t < 0$:

$$Re < f - u, v > \geq \frac{1}{t} Re < f - u, v > \to 0 \quad \text{as } t \to \infty$$

so

$$Re < f - u, v > \leq 0 \quad \text{and} \quad Re < f - u, v > \geq 0,$$

i.e.,

$$Re < f - u, v >= 0 \quad \forall v \in M.$$

Replace v by $-iv$. Then

$$0 = Re < f - u, -iv >= Re(-i < f - u, v >) = Im < f - u, v >$$

so (6) holds.
Step 2:

$$|P_M f| \leq |f| \quad \forall f \in H.$$

Indeed, since M is linear, $0 \in M$ and $P_M 0 = 0$, so by Proposition 9.5

$$|P_M f| = |P_M f - P_M 0| \leq |f - 0| = |f|.$$

Step 3: If u satisfies (6), then $u = P_M f$.
Indeed, if

$$< f - u, v >= 0 \quad \forall v \in M,$$

then, since $u \in M$, and M is linear, $v - u \in M$, so

$$< f - u, v - u >= 0.$$

Hence (2) holds which characterizes $u = P_M f$!
Step 4: $P_M : H \to M$ is linear.
Indeed, if $f_1, f_2 \in H, u_j = P_M f_j, \alpha_1, \alpha_2 \in \mathbb{F}$, then

$$< f_1 - u_1, v >= 0 \quad \forall v \in M,$$
$$< f_2 - u_2, v >= 0 \quad \forall v \in M.$$

Thus

$$\begin{aligned}0 &=< \alpha_1 f_1 - \alpha_1 u_1, v > + < \alpha_2 f_2 - \alpha_2 u_2, v > \\ &=< \alpha_1 f_1 + \alpha_2 f_2 - (\alpha_1 u_1 + \alpha_2 u_2), v >,\end{aligned}$$

i.e.,

$$\alpha_1 u_1 + \alpha_2 u_2 = P_M(\alpha_1 f_1 + \alpha_2 f_2).$$

□

9.2 The dual space of a Hilbert space

There are plenty of continuous linear functionals on a Hilbert space H. Simply pick $f \in H$ and consider

$$u \mapsto < f, u > .$$

The remarkable fact is that all continuous linear functionals on H are of this form!

Theorem 9.7 (Riesz-Fréchet representation theorem)**.** *Given any $\varphi \in H^*$ there exists a unique $f = f_\varphi \in H$ such that*

$$\varphi(u) =< f, u > \quad \forall u \in H.$$

Moreover,

$$|f| = \|\varphi\|_{H^*}.$$

Proof. 1st: Consider the map $T : H \to H^*$,

$$Tf :=< f, \cdot > \in H^*,$$

i.e.,

$$Tf(u) :=< f, u > .$$

It is clear that $\|Tf\|_{H^*} = |f|$ (why?), so T is an isometry from H onto $T(H)$, i.e., $T(H)$ is a closed subspace of H^*. Assume $h \in (H^*)^*$ which vanishes on $T(H)$. Since H is reflexive, $h \in H$ and

$$Tf(h) =< f, h >= 0 \quad \forall f \in H.$$

Take $f = h$. Then

$$|h|^2 =< h, h >= 0 \quad \Rightarrow \quad h = 0,$$

i.e., $T(H)$ is dense in H^* and thus $T(H) = H$.
2nd: Given $\varphi \in H^*$, let

$$M := \varphi^{-1}(\{0\}) \subset H$$

and note that M is closed since φ is continuous.
Assume $M \neq H$ (otherwise $\varphi \equiv 0$ and we take $f = 0$). Pick any $g_0 \in H$ such that $\varphi(g_0) \neq 0$ and set $g_1 := P_M g_0 \in M$. Note

$$\varphi(g_0 - g_1) = \varphi(g_0) - \underbrace{\varphi(g_1)}_{=0} = \varphi(g_0) \neq 0$$

so

$$g_0 - g_1 \neq 0.$$

Put

$$g := \frac{g_0 - g_1}{|g_0 - g_1|}.$$

Then $|g| = 1$,

$$\varphi(g) = \frac{\varphi(g_0)}{|g_0 - g_1|} \neq 0$$

and

$$< g, v >= \frac{1}{|g_0 - g_1|} < g_0 - g_1, v >= \frac{1}{|g_0 - g_1|} < g_0 - P_M g_0, v >= 0$$

by Corollary 9.6.

Given $u \in H$ let

$$v = u - \lambda g$$

and choose λ such that $v \in M$, i.e.,

$$\lambda = \frac{\varphi(u)}{\varphi(g)}.$$

But then

$$\begin{aligned} 0 &=< g, v >=< g, u - \lambda g > \\ &=< g, v > -\lambda \underbrace{< g, g >}_{=1} \\ &=< g, v > -\frac{\varphi(u)}{\varphi(g)}. \end{aligned}$$

Thus

$$\varphi(u) = \varphi(g) < g, u >=< \overline{\varphi(g)} g, u >$$

so $f := \overline{\varphi(g)}$ works. □

9.3 The Theorems of Stampacchia and Lax-Milgram

In the following, let H be a real Hilbert space.

Definition. *A bilinear form $a : H \times H \to \mathbb{R}$ is said to be*

(i) continuous, if there exists $C > 0$ such that

$$|a(u, v)| \leq C|u||v| \quad \forall u, v \in H;$$

(ii) coercive, if there exists $\alpha > 0$ such that

$$a(v, v) \geq \alpha|v|^2 \quad \forall v \in H.$$

Theorem 9.8 (Stampacchia)**.** *Assume that a is a continuous coercive bilinear form on a real Hilbert space H. Let $K \subset H, K \neq \emptyset$ closed and convex. Then given $\varphi \in H^*$ there exists a unique $u \in K$ such that*

$$a(u, v - u) \geq \varphi(v - u) \quad \forall v \in K. \tag{1}$$

Moreover, if a is symmetric, then u is characterized by

$$u \in K \quad \textit{and} \quad \frac{1}{2}a(u, u) - \varphi(u) = \inf_{v \in K} \Big(\frac{1}{2}a(v, v) - \varphi(v)\Big). \tag{2}$$

We need

Theorem 9.9 (Banach fixed point theorem)**.** *Let $X \neq \emptyset$ be a complete metric space and $S : X \to X$ be a strict contraction, i.e.,*

$$d(S(x_1), S(x_2)) \leq kd(x_1, x_2) \quad \forall x_1, x_2 \in X \textit{ with } k < 1.$$

Then S has a unique fixed point u, i.e., $u = S(u)$.

Proof of Theorem 9.8. By Riesz representation theorem there exists $f \in H$ such that

$$\varphi(v) = < f, v > \quad \forall v \in H.$$

Note that also the maps $v \mapsto a(u, v) \in H^*$, so again there exists a unique element in H, denoted by Au such that

$$a(u, v) = < Au, v > \quad \forall v \in H.$$

Note: A is a linear operator from H to H and

$$|Au| \leq C|u| \quad \forall u \in H,$$
$$< Au, u > \geq \alpha|u|^2 \quad \forall u \in H.$$

So problem (1) says we should find $u \in K$ such that

$$< Au, v - u > \geq < f, v - u > \quad \forall v \in K. \tag{3}$$

Let $\rho > 0$ and note that (3) is equivalent to

$$< \rho f - \rho Au + u - u, v - u > \leq 0 \quad \forall v \in K, \tag{4}$$

i.e.,

$$u = P_K(\rho f - \rho A u + u).$$

For $v \in K$ set

$$S(v) := P_K(\rho f - \rho A v + v).$$

Claim: Choosing $\rho > 0$ cleverly, S is a strict contraction, so it has a unique fixed point!
Indeed,

$$\begin{aligned} |Sv_1 - Sv_2| &= |P_K(\rho f - \rho A v_1 + v_1) - P_K(\rho f - \rho A v_2 + v_2)| \\ &\leq |\rho f - \rho A v_1 + v_1 - \rho f + \rho A v_2 - v_2| \\ &= |(v_1 - v_2) + \rho(A v_1 - A v_2)| \end{aligned}$$

$$\begin{aligned} \Rightarrow |Sv_1 - Sv_2|^2 &= |v_1 - v_2|^2 - 2\rho \underbrace{< Av_1 - Av_2, v_1 - v_2 >}_{\geq \alpha |v_1 - v_2|^2} + \rho^2 |Av_1 - Av_2|^2 \\ &\leq |v_1 - v_2|^2 (1 - 2\rho\alpha + \rho^2 C^2). \end{aligned}$$

Choose ρ so that

$$K^2 = 1 - 2\rho\alpha + \rho^2 C^2 < 1,$$

i.e., $0 < \rho < \frac{2\alpha}{C^2}$. Then S has a unique fixed point.

Assume now that a is symmetric. Then $a(u, v)$ defines a new scalar product on H with norm $\sqrt{a(u,u)}$ which is equivalent to the old norm $|u|$. Thus H is a Hilber space for this new scalar product. By Riesz-Fréchet for $a(u, v)$, it follows that given $\varphi \in H^*$ there exists a unique $g \in H$ such that

$$\varphi(u) = a(g, u) \quad \forall u \in H.$$

Note that problem (1) amounts to finding some $u \in K$ such that

$$a(g - u, v - u) \leq 0 \quad \forall v \in K \tag{5}$$

but the solution to (5) is the projection onto K of g for the new scalar product a! By Theorem 9.4 $u \in K$ is the unique element which achieves

$$\inf \sqrt{a(g - v, g - v)},$$

i.e., one minimizes on K the function

$$\begin{aligned} v \mapsto a(g - v, g - v) &= a(v, v) - 2a(g, v) + a(g, g) \\ &= a(v, v) - 2\varphi(u) + a(g, g) \end{aligned}$$

or equivalently, the function

$$v \mapsto \frac{1}{2} a(v, v) - \varphi(u).$$

□

Corollary 9.10 (Lax-Milgram)**.** *Assume that $a(u,v)$ is a continuous coercive bilinear form on H. Then given any $\varphi \in H^*$ there exists a unique $u \in H$ such that*

$$a(u,v) = \varphi(u) \quad \forall v \in H. \tag{6}$$

Moreover, if a is symmetric, then u is characterized by

$$u \in H \quad \textit{and} \quad \frac{1}{2}a(u,u) - \varphi(u) = \inf_{\frac{1}{2}v \in H}(a(v,v) - \varphi(v)). \tag{7}$$

Proof. Apply Theorem 9.8 with $K = H$ and use linearity of H as in the proof of Corollary 9.6. □