Functional Analysis–Math 920 (Spring 2003)

Casim Abbas

April 25, 2003

Contents

Chapter 1

Preliminary remarks, Notation

1.1 Notation

Functional Analysis is a fundamental part of Mathematics developed in the first half of the 20th century. It has become a very important tool in modern mathematics, in particular for partial differential equations. If I had to say in a few lines what Functional Analysis is about I would say this: Functional Analysis is about solving equations $F(x) = y$, where F is a linear map between vector spaces X and Y. If X and Y were finite dimensional then this would just be Linear Algebra. The vector spaces we are concerned with will be infinite dimensional. In fact, they will mostly be function spaces. For example, if $\Omega \subset \mathbf{R}^n$ is a domain in $\mathbf{R}^n$, i.e. an open connected subset of $\mathbf{R}^n$, and if $C^k(\Omega)$ denotes the set of all functions $f : \Omega \to \mathbf{R}$ which are k times differentiable with continuous derivatives then the Laplace operator

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial {x_i}^2}$$

is a linear map from $C^2(\Omega)$ into $C^0(\Omega)$. Finding a solution $u \in C^2(\Omega)$ to Poisson's equation $\Delta u = f$ with given $f \in C^0(\Omega)$ can then be viewed as solving an inhomogeneous linear equation between suitable vector spaces. It is this abstract point of view that makes Functional Analysis so powerful: There is a large number of partial differential equations (elliptic partial differential equations) which can all be treated in the same way because they have the same abstract functional analytic origin. The vector spaces considered in Functional Analysis will carry particular topological structures, and the linear maps F will mostly be continuous with respect to the given topologies on X and Y.
Some remarks on notation: Vector spaces will always be over the real or over

the complex numbers. We denote the sets of natural numbers, integers, rational, real and complex numbers by $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ respectively. The set $\mathbf{N}$ contains 0, otherwise we write $\mathbf{N}^*$. The letter Ω will denote a domain in $\mathbf{R}^n$ (not necessarily bounded). I will also write $'x := y'$ if I want to *define* x to be y. The term $'x = y'$ means that x, y are both defined and I am claiming that they are equal. Sometimes we will write $\partial_k u$, $\partial_{x_k} u$ or $D_k u$ for the partial derivative $\frac{\partial}{\partial x_k}$. We will frequently use the following notation: If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an n–tuple of integers $\alpha_k \geq 0$ then we write $|\alpha| := \alpha_1 + \cdots + \alpha_n$ and

$$D^\alpha u := \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

We will write $C^k(\overline{\Omega})$ for the set of all functions $f : \Omega \to \mathbf{R}$ so that all derivatives $D^\alpha f$ exist whenever $|\alpha| \leq k$, they satisfy $\sup_{x\in\Omega} |D^\alpha f(x)| < +\infty$, and they can be extended continuously up to the closure $\overline{\Omega}$ of Ω. We also write

$$C_0^k(\Omega) := \{f \in C^k(\Omega) \,|\, \mathrm{supp}(f) \subset \Omega \text{ is compact}\}$$

with

$$\mathrm{supp}(f) := \overline{\{x \in \Omega \,|\, f(x) \neq 0\}}$$

("functions with compact support"). We use the notation $C^\infty(\Omega)$ for the set of all infinitely differentiable functions ("smooth functions").

1.2 Baire's lemma

In this section we will prove Baire's lemma which we will need later on. Let (X, d) be a metric space.

Definition 1.2.1 *A subset $A \subset X$ is called nowhere dense if $\overset{\circ}{\overline{A}} = \emptyset$.*

In particular, a nowhere dense set does not contain any open ball.

Theorem 1.2.2 *Let (X, d) be a complete metric space, and let $(U_i)_{i\in\mathbf{N}}$ be a sequence of open dense sets. Then the countable intersection*

$$\bigcap_{i\in\mathbf{N}} U_i$$

is also dense in X.

Proof:
Let $x \in X$. We have to show that for all numbers $\varepsilon > 0$

$$B_\varepsilon(x) \cap \left(\bigcap_{i\in\mathbf{N}} U_i\right) \neq 0$$

with $B_\varepsilon(x) := \{y \in X \,|\, d(x,y) < \varepsilon\}$. We know that $B_\varepsilon(x) \cap U_1$ is not empty and open. Pick $x_1 \in B_\varepsilon(x) \cap U_1$ and $\varepsilon_1 < \frac{\varepsilon}{2}$ so that

$$\overline{B_{\varepsilon_1}(x_1)} \subset B_\varepsilon(x) \cap U_1.$$

Now $B_{\varepsilon_1}(x_1) \cap U_2$ is not empty and open as well. Pick $x_2 \in B_{\varepsilon_1}(x_1) \cap U_2$ and $\varepsilon_2 < \frac{\varepsilon_1}{2}$ so that

$$\overline{B_{\varepsilon_2}(x_2)} \subset B_{\varepsilon_1}(x_1) \cap U_2.$$

Continuing this iteration, we obtain sequences of points $(x_n)_{n\in\mathbf{N}}$ and positive numbers $(\varepsilon_n)_{n\in\mathbf{N}}$ so that

$$\overline{B_{\varepsilon_{n+1}}(x_{n+1})} \subset B_{\varepsilon_n}(x_n) \cap U_{n+1}$$

and

$$\varepsilon_{n+1} < \frac{\varepsilon_n}{2}.$$

Because X is complete the intersection $\bigcap_{n\in\mathbf{N}} \overline{B_{\varepsilon_n}(x_n)}$ is not empty. Let $\tilde{x}$ be an element in this intersection. By construction

$$\tilde{x} \in \overline{B_{\varepsilon_n}(x_n)} \subset B_\varepsilon(x) \cap U_n$$

for all $n \in \mathbf{N}$, hence

$$\tilde{x} \in B_\varepsilon(x) \cap \left(\bigcap_{i\in\mathbf{N}} U_i \right).$$

❑

Exercise 1.2.3 *Let (X,d) be a metric space. Show that the following two statements are equivalent:*

- *The space (X,d) is complete, i.e. every Cauchy sequence is convergent.*
- *Let $B_{\varepsilon_k}(x_k)$ be any sequence of open balls with*

  $$B_{\varepsilon_k}(x_k) \subset \overline{B_{\varepsilon_k}(x_k)} \subset B_{\varepsilon_{k-1}}(x_{k-1}) \text{ and } \varepsilon_k \searrow 0.$$

  Then $\bigcap_k \overline{B_{\varepsilon_k}(x_k)} \neq \emptyset$.

We continue with some equivalent formulations of Baire's lemma. The set $U_i \subset X$ is open and dense if and only if the complement $X \backslash U_i$ is closed and nowhere dense. Indeed, if U_i is open and dense then the complement is closed and any open ball around a point $x \in X \backslash U_i$ has to intersect U_i. This means that $X \backslash U_i = \overline{X \backslash U_i}$ does not contain any open ball, i.e. it is nowhere dense. Conversely, if $X \backslash U_i$ is closed and nowhere dense then U_i is open and $X \backslash U_i$ does not contain any open ball which implies that U_i is dense.

Theorem 1.2.4 *Let (X, d) be a complete metric space, and let $(A_i)_{i\in\mathbf{N}}$ be a sequence of closed nowhere dense subsets of X. Then the union $\bigcup_{i\in\mathbf{N}} A_i$ has no interior points.*

❑

Theorem 1.2.5 *Let (X, d) be a complete metric space, and let $(A_i)_{i\in\mathbf{N}}$ be a sequence of closed subsets of X. Assume that the union $\bigcup_{i\in\mathbf{N}} A_i$ contains an open ball. Then there is some $k \in \mathbf{N}$ so that the set A_k also contains an open ball.*

❑

Exercise 1.2.6 *Prove the above two versions of Baire's lemma.*

Remark:
Theorem 1.2.5 is the most commonly used version of Baire's lemma.

1.3 Brief review of Integration

In this section we briefly review some basic facts about integration without proofs, and I assume that you are familiar with them. You may find the proofs in books about measure theory, for example in the book by Wheeden and Zygmund. Here are the 'big three', the convergence theorems of Lebesgue integration (theorem of monotone convergence by Beppo Levi, Lebesgue's convergence theorem and Fatou's lemma):

Theorem 1.3.1 (B. Levi, monotone convergence)
Let $f_n : \Omega \to \mathbf{R}$ be a sequence of functions in $L^1(\Omega)$ such that $f_n(x) \leq f_{n+1}(x)$ almost everywhere and $\sup_n \int_\Omega f_n(x)dx < \infty$. Then the sequence f_n converges pointwise almost everywhere to some limit f which is also in $L^1(\Omega)$ and $\|f_n - f\|_{L^1(\Omega)} \to 0$ as $n \to \infty$.

❑

Theorem 1.3.2 (H. Lebesgue)
Let $f_n : \Omega \to \mathbf{R}$ be a sequence of functions in $L^1(\Omega)$. Suppose that

- *$f_n(x) \to f(x)$ for almost every $x \in \Omega$,*
- *There is an integrable function $g : \Omega \to \mathbf{R}$ such that $|f_n(x)| \leq |g(x)|$ for all n and for almost every $x \in \Omega$.*

Then $\|f_n - f\|_{L^1(\Omega)} \to 0$ as $n \to \infty$.

□

Lemma 1.3.3 (Fatou's lemma)
Let $f_n : \Omega \to \mathbf{R}$ be a sequence of functions in $L^1(\Omega)$ such that

- *for almost every $x \in \Omega$ and every n we have $f_n(x) \geq 0$,*
- $\sup_n \int_\Omega f_n(x)dx < \infty$.

For each $x \in \Omega$ we define $f(x) := \liminf_n f_n(x)$. Then f is integrable and

$$\int_\Omega f(x)dx \leq \liminf_n \int_\Omega f_n(x)dx.$$

□

The following result is very important:

Theorem 1.3.4 *The space $C_0^\infty(\Omega)$ is dense in $L^1(\Omega)$, i.e. for every $\varepsilon > 0$ and $f \in L^1(\Omega)$ there is some $\phi \in C_0^\infty(\Omega)$ such that*

$$\|\phi - f\|_{L^1(\Omega)} < \varepsilon.$$

□

We conclude our summary of integration theory with the theorems of Tonelli and Fubini. Assume that $\Omega_1 \subset \mathbf{R}^{n_1}$ and $\Omega_2 \subset \mathbf{R}^{n_2}$ are open domains. Moreover, let $F : \Omega_1 \times \Omega_2 \to \mathbf{R}$ be a measureable function. Here is Fubini's theorem:

Theorem 1.3.5 (Fubini)
Suppose that $F \in L^1(\Omega_1 \times \Omega_2)$. Then for almost all $x \in \Omega_1$

$$F(x, *) \in L^1(\Omega_2) \text{ and } \int_{\Omega_2} F(*, y)dy \in L^1(\Omega_1).$$

Also for almost all $y \in \Omega_2$

$$F(*, y) \in L^1(\Omega_1) \text{ and } \int_{\Omega_1} F(x, *)dx \in L^1(\Omega_2).$$

Moreover,

$$\int_{\Omega_1} \left(\int_{\Omega_2} F(x, y)dy \right) dx = \int_{\Omega_2} \left(\int_{\Omega_1} F(x, y)dx \right) dy = \int_{\Omega_1 \times \Omega_2} F(x, y)dxdy.$$

□

Hence finiteness of the integral $\int_{\Omega_1\times\Omega_2} F(x,y)dxdy$ implies finiteness of the iterated integrals $\int_{\Omega_1}\left(\int_{\Omega_2} F(x,y)dy\right)dx$ and $\int_{\Omega_2}\left(\int_{\Omega_1} F(x,y)dx\right)dy$. The converse is not true, even if the iterated integrals both exist and are equal, the function F need not be integrable over $\Omega_1 \times \Omega_2$ (see Wheeden-Zygmund p. 91 for a counterexample). However, the converse is true if F is not negative, which is Tonelli's theorem.

Theorem 1.3.6 (Tonelli)
Assume that F is not negative. Then for almost every $x \in \Omega_1$ the function $F(x,*)$ is a measureable function on Ω_2. Moreover, as a function of x, $\int_{\Omega_2} F(x,y)dy$ is measureable on Ω_1 and

$$\int_{\Omega_1}\left(\int_{\Omega_2} F(x,y)dy\right)dx = \int_{\Omega_1\times\Omega_2} F(x,y)dxdy.$$

❑

Chapter 2

Normed Linear Spaces

2.1 Norms

Let X be a vector space over the real or over the complex numbers. A norm on X is a real valued function $X \to \mathbf{R}$, which we denote by $|x|$ satisfying the following conditions:

- $|x| \geq 0$ with equality if and only if $x = 0$,
- $|x + y| \leq |x| + |y|$, 'subadditivity'
- For all $\lambda \in \mathbf{R}$ we have $|\lambda x| = |\lambda| \cdot |x|$.

A norm on a vector space X induces a metric on X by

$$d(x, y) := |x - y|.$$

This metric is invariant under translations and homogeneous, i.e.

$$d(x + z, y + z) = d(x, y)\ ,\ d(\lambda x, \lambda y) = |\lambda|\, d(x, y).$$

Definition 2.1.1 *If a vector space X equipped with a norm is complete, i.e. every Cauchy sequence converges, then $(X, |\,.\,|)$ is called a Banach space.*

Definition 2.1.2 *Let X be a vector space. Two different norms $|.|_1$ and $|.|_2$ are called equivalent if there is a constant $c > 0$ such that*

$$c\,|x|_1 \leq |x|_2 \leq \frac{1}{c}|x|_1$$

for all $x \in X$.

Equivalent norms on X induce the same topology on X. We observe the following:

1. A subspace Y of a normed linear space is again a normed linear space.
2. If X, Y are two normed linear spaces, then we denote the set of all ordered pairs (x, y) with $x \in X$, $y \in Y$ by $X \oplus Y$. The space $X \oplus Y$ can be equipped with a norm by defining

   $$|(x, y)|_1 := |x| + |y| \ , \ |(x, y)|_2 := \max\{|x|, |y|\} \text{ or } |(x, y)|_3 := \sqrt{|x|^2 + |y|^2}.$$

Exercise 2.1.3 *Show that $|.|_k$, $k = 1, 2, 3$ above are indeed norms and show that they are equivalent norms.*

Let X be a normed linear space and let Y be a subspace. If $Y \subset X$ is closed then there is a natural norm on the quotient space X/Y as follows:

Proposition 2.1.4 *Let X, Y be as above with Y closed. If $[x] \in X/Y$ is an equivalence class of elements of X modulo Y then the following defines a norm on X/Y:*

$$|[x]| := \inf_{x \in [x]} |x| = \inf_{y \in Y} |x + y|.$$

If moreover X is a Banach space then X/Y is also a Banach space with the above norm.

Proof:
We first check that $|[x]|$ is indeed a norm. If $\lambda \in \mathbf{R}$ we trivially have $|\lambda[x]| = |\lambda|\,|[x]|$. In order to check the triangle inequality, let $\varepsilon > 0$ and pick representatives $x_\varepsilon \in [x]$, $y_\varepsilon \in [y]$ so that

$$|x_\varepsilon| < |[x]| + \varepsilon \text{ and } |y_\varepsilon| < |[y]| + \varepsilon$$

which is possible by definition of the norm $|[x]|$. Since $x_\varepsilon + y_\varepsilon$ is a representative of the class $[x] + [y]$ we estimate

$$\begin{aligned} |[x] + [y]| &= \inf_{z \in [x]+[y]} |z| \\ &\leq |x_\varepsilon + y_\varepsilon| \\ &\leq |x_\varepsilon| + |y_\varepsilon| \\ &\leq |[x]| + |[y]| + 2\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we conclude $|[x] + [y]| \leq |[x]| + |[y]|$. We clearly have $|[x]| \geq 0$ for all $[x] \in X/Y$. Assume now that $|[x]| = 0$. We would like to show that $[x] = 0$, i.e. $x \in Y$ for any representative $x \in [x]$. Since

$$0 = |[x]| = \inf_{x \in [x]} |x|,$$

there is a sequence $x_k \in [x]$ with $|x_k| \to 0$ as $k \to \infty$. Since the elements x_k are all equivalent modulo Y, we can find a sequence $(y_k)_{k\in\mathbf{N}} \subset Y$ such that

$$x_k = x_1 - y_k \ , \ k \geq 2.$$

Now $|x_k| = d(x_1, y_k) \to 0$, i.e. viewing X as a metric space the sequence y_k converges to x_1. Because Y is closed by assumption the element x_1 then also belongs to Y which implies $[x] = 0$.

Assume now that X is a Banach space. We know now that the quotient X/Y is a normed space with the norm $|[x]|$ as above. Let $|[x_n]|$ be a Cauchy sequence in X/Y, i.e. $|[x_n] - [x_m]|$ converges to zero as n, m tend to infinity. We have to show that the sequence $([x_n])_{n\in\mathbf{N}}$ converges in X/Y. It is sufficient to show that the sequence $([x_n])_{n\in\mathbf{N}}$ has a convergent subsequence.

Since $([x_n])_{n\in\mathbf{N}}$ is a Cauchy sequence we may find a subsequence $([x_{n_k}])_{k\in\mathbf{N}}$ such that

$$|[x_{n_{k+1}}] - [x_{n_k}]| < \frac{1}{2^k} \ \forall \ k \in \mathbf{N}.$$

We claim now that every class $[x] \in X/Y$ has a representative $x \in [x]$ such that

$$|x| < 2\,|[x]|.$$

If this were not true then there would be some $[x] \in X/Y$ such that for all representatives $x \in [x]$

$$|x| \geq 2\,|[x]| = 2 \inf_{x\in[x]} |x|$$

which is clearly absurd proving the claim. Pick now representatives $x_{n_k} \in [x_{n_k}]$ so that

$$|x_{n_{k+1}} - x_{n_k}| < 2\,|[x_{n_{k+1}}] - [x_{n_k}]| < \frac{1}{2^{k-1}}.$$

We then get for $l \in \mathbf{N}$

$$\begin{aligned}
|x_{n_{k+l}} - x_{n_k}| &\leq \sum_{m=1}^{l} |x_{n_{k+m}} - x_{n_{k+m-1}}| \\
&< \sum_{m=1}^{l} \frac{1}{2^{k+m-2}} \\
&= \frac{1}{2^{k-1}} \sum_{m=0}^{l-1} \frac{1}{2^m} \\
&= \frac{1}{2^k}(1 - \frac{1}{2^l}),
\end{aligned}$$

hence the sequence $(x_{n_k})_{k\in\mathbf{N}}$ is a Cauchy sequence in X. Since X is complete it has a limit $x \in X$. Now

$$|[x_{n_k}] - [x]| = \inf_{y\in Y} |x_{n_k} - x + y| \leq |x_{n_k} - x| \to 0$$

as $k \to \infty$ completing the proof.

$\square$

Proposition 2.1.5 *If X is a finite dimensional vector space then any two norms on X are equivalent.*

Proof:
Let $e_1, \ldots, e_n$ be a basis for X. Then every element $x \in X$ has a unique representation

$$x = \sum_{i=1}^{n} x_i \, e_i$$

and

$$\|x\|_{max} := \max_{1 \leq i \leq n} |x_i|$$

is a norm on X. It suffices to show that any norm on X is equivalent to the norm $\|.\|_{max}$. Let $\|.\|$ be a norm on X. We estimate

$$\|x\| \leq \sum_{i=1}^{n} |x_i| \, \|e_i\| \leq \Big(\max_{1 \leq i \leq n} |x_i| \Big) \sum_{i=1}^{n} \|e_i\| = c \, \|x\|_{max}$$

where $c = \sum_{i=1}^{n} \|e_i\|$. We have to show the reverse inequality, i.e. we have to show that there is some positive constant c' such that $\|x\|_{max} \leq c' \|x\|$ for all $x \in X$. If this were not true then for every $\varepsilon > 0$ there would be some $x^\varepsilon \in X$ such that

$$\varepsilon \|x^\varepsilon\|_{max} > \|x^\varepsilon\|.$$

Hence $x^\varepsilon \neq 0$ and we may assume without loss of generality that

$$\|x^\varepsilon\|_{max} = \max_{1 \leq i \leq n} |x_i^\varepsilon| = 1$$

(otherwise consider $x^\varepsilon / \|x^\varepsilon\|_{max}$ instead of x^ε). We can now find a sequence $\varepsilon_k \searrow 0$ and some $1 \leq i_0 \leq n$ so that

$$|x_{i_0}^{\varepsilon_k}| = 1 \;\forall\, k \in \mathbf{N}$$

and

$$x_i^{\varepsilon_k} \longrightarrow \xi_i \text{ as } k \to \infty.$$

Let

$$x := \sum_{i=1}^{n} \xi_i e_i$$

so that

$$x - x^{\varepsilon_k} = \sum_{i=1}^{n} (\xi_i - x_i^{\varepsilon_k}) \, e_i.$$

We estimate

$$\begin{aligned} \|x\| &\leq \|x - x^{\varepsilon_k}\| + \|x^{\varepsilon_k}\| \\ &\leq \Big(\max_{1 \leq i \leq n} |\xi_i - x_i^{\varepsilon_k}| \Big) \sum_{i=1}^{n} \|e_i\| + \varepsilon_k, \end{aligned}$$

which converges to zero as $k \to \infty$. This implies $x = 0$ and also $\xi_i = 0$ for all $1 \le i \le n$ in contradiction to $|\xi_{i_0}| = 1$.

❑

2.2 Examples of Banach spaces

2.2.1 $C_b^k(\Omega)$, $C^k(\overline{\Omega})$ and Hölder spaces

We denote by $C_b^k(\Omega)$ the space of k–times continuously differentiable functions such that all derivatives up to order k are bounded in the supremum–norm, i.e. we define for $f \in C^k(\Omega)$

$$\|f\|_{C^k(\Omega)} := \sum_{0 \le |\alpha| \le k} \sup_{x \in \Omega} |D^\alpha f(x)|$$

and

$$C_b^k(\Omega) := \{f \in C^k(\Omega) \,|\, \|f\|_{C^k(\Omega)} < \infty\}.$$

If $0 < \beta \le 1$ then we define for $f \in C^k(\Omega)$

$$\|f\|_{C^{k,\beta}(\Omega)} := \|f\|_{C^k(\Omega)} + \sum_{|\alpha| = k} \sup_{x,y \in \Omega\,,\, x \ne y} \frac{|D^\alpha f(y) - D^\alpha f(x)|}{|x-y|^\beta}$$

and

$$C^{k,\beta}(\Omega) := \{f \in C^k(\Omega) \,|\, \|f\|_{C^{k,\beta}(\Omega)} < \infty\}.$$

Functions in $C^{0,\beta}(\Omega)$ are called Hölder–continuous and Lipschitz–continuous in the case $\beta = 1$. We will refer to the spaces $C^{k,\beta}(\Omega)$ simply as Hölder spaces. If Ω is a bounded domain, we define

$$C^k(\overline{\Omega}) := \{f \in C^k(\Omega) \,|\, D^\alpha f \text{ extends continuously onto } \overline{\Omega} \text{ for all } 0 \le |\alpha| \le k\}$$

and

$$C^{k,\beta}(\overline{\Omega}) := \{f \in C^k(\overline{\Omega}) \,|\, \|f\|_{C^{k,\beta}(\overline{\Omega})} < \infty\}$$

where the norms $\|.\|_{C^k(\overline{\Omega})}$ and $\|.\|_{C^{k,\beta}(\overline{\Omega})}$ are defined in a similar way as above, just replace Ω by $\overline{\Omega}$ in the definition. Hölder spaces are extremely important in the theory of partial differential equations.

Exercise 2.2.1 *Show that the product of two Hölder continuous functions $f_1 \in C^{0,\beta_1}(\overline{\Omega})$ and $f_2 \in C^{0,\beta_2}(\overline{\Omega})$ is again Hölder continuous, i.e. there is $\gamma \in (0,1]$ such that $f_1 f_2 \in C^{0,\gamma}(\overline{\Omega})$. What is the correct Hölder exponent γ ?*

Theorem 2.2.2 *Let $\Omega \subset \mathbf{R}^n$ be a domain. Then the spaces $C_b^k(\Omega)$ and $C^{k,\beta}(\Omega)$ are Banach spaces with the norms $\|.\|_{C^k(\Omega)}$ and $\|.\|_{C^{k,\beta}(\Omega)}$ respectively.*

Proof:
We will only consider the spaces $C^1(\Omega)$ and $C^{0,\beta}(\Omega)$. The general case follows easily by iteration. First, the space $C_b^0(\Omega)$ is a Banach space with the supremum norm $\|.\|_{C^0(\Omega)}$ for the following reason: If $(f_n) \subset C_b^0(\Omega)$ is a Cauchy sequence then for each $\varepsilon > 0$ there is $N \in \mathbf{N}$ so that

$$|f_n(x) - f_m(x)| < \varepsilon \; \forall\, n, m \geq N \, , \, x \in \Omega.$$

The sequence of real numbers $(f_n(x))$ is then also a Cauchy sequence for every $x \in \Omega$, hence it has a limit $f(x)$ by the completeness of the real numbers. On the other hand, we also have $\|f_n - f\|_{C^0(\Omega)} \to 0$ since

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \liminf_{m\to\infty} \|f_n - f_m\|_{C^0(\Omega)}$$

and

$$||f_n - f\|_{C^0(\Omega)} \leq \liminf_{m\to\infty} \|f_n - f_m\|_{C^0(\Omega)} \to 0 \text{ for } n \to \infty.$$

Since $\|f_n - f\|_{C^0(\Omega)} \to 0$ the sequence (f_n) converges uniformly to f so that f is continuous and also bounded. Let now $(f_n)_{n\in\mathbf{N}}$ be a Cauchy sequence in $C_b^1(\Omega)$. Then the sequences $(\partial_i f_n)_{n\in\mathbf{N}}$, $(f_n)_{n\in\mathbf{N}}$, are Cauchy sequences with respect to the supremum norm and therefore have continuous limits which we denote by g_i and f respectively. It remains to show that the limit f is differentiable and that $\partial_i f = g_i$. We define $g := (g_1, \ldots, g_d)$ (with d being the dimension of the domain Ω), $\nabla f_n(x) := (\partial_1 f_n(x), \ldots, \partial_d f_n(x))$, and for given $x \in \Omega$ we pick $y \in \Omega$ such that $x_t := (1-t)x + ty \in \Omega$ for all $0 \leq t \leq 1$. Then

$$\begin{aligned} |f_n(y) - f_n(x) - \nabla f_n(x)\cdot(y-x)| &= \left| \int_0^1 (\nabla f_n(x_t) - \nabla f_n(x))\cdot(y-x)\,dt \right| \\ &\leq |y-x| \int_0^1 |\nabla f_n(x_t) - \nabla f_n(x)|\,dt \\ &\leq |y-x|(2\,\|\nabla f_n - g\|_{C^0(\Omega)} + \\ &\quad + \sup_{0\leq t\leq 1} |g(x_t) - g(x)|). \end{aligned}$$

For $n \to \infty$ we obtain

$$|f(y) - f(x) - g(x)\cdot(y-x)| \leq |y-x| \sup_{0\leq t\leq 1} |g(x_t) - g(x)|,$$

but $\sup_{0\leq t\leq 1} |g(x_t) - g(x)|$ converges to zero as $y \to x$. This means that f is differentiable in x with $\nabla f(x) = g(x)$. This shows that $C_b^1(\Omega)$ is a Banach space.
Let us now assume that $(f_n)_{n\in\mathbf{N}}$ is a Cauchy sequence in $C^{0,\beta}(\Omega)$. The sequence (f_n) is also Cauchy for the supremum norm, hence there is a bounded continuous function f so that $\|f_n - f\|_{C^0(\Omega)} \to 0$ as $n \to \infty$. We have to show that also

$$\sup_{x,y\in\Omega\,,\,x\neq y} \frac{|f(y) - f(x)|}{|x-y|^\beta} < \infty.$$

If $\varepsilon > 0$ then there is $N \in \mathbf{N}$ such that for all $n, m \geq N$

$$\sup_{x,y\in\Omega\,,\,x\neq 0} \frac{|f_n(x) - f_m(x) - (f_n(y) - f_m(y))|}{|x-y|^\beta} < \varepsilon$$

because (f_n) is a Cauchy sequence with respect to the norm $\|.\|_{C^{0,\beta}(\Omega)}$. For each pair $x, y \in \Omega$ with $x \neq y$ we may pass to the limit $m \to \infty$, and we obtain

$$\frac{|f_n(x) - f(x) - (f_n(y) - f(y))|}{|x-y|^\beta} \leq \varepsilon,$$

which implies

$$\frac{|f(x) - f(y)|}{|x-y|^\beta} \leq \varepsilon + \frac{|f_n(x) - f_n(y)|}{|x-y|^\beta},$$

hence $f \in C^{0,\beta}(\Omega)$, and $\|f - f_n\|_{C^{0,\beta}(\Omega)} \to 0$ as $n \to \infty$.

❑

In the same way we have

Theorem 2.2.3 *Let $\Omega \subset \mathbf{R}^n$ be a bounded domain. Then the spaces $C^k(\overline{\Omega})$ and $C^{k,\beta}(\overline{\Omega})$ are Banach spaces with the norms $\|.\|_{C^k(\overline{\Omega})}$ and $\|.\|_{C^{k,\beta}(\overline{\Omega})}$ respectively.*

❑

The following crucial theorem characterises precompact sets in $C^0(\overline{\Omega})$. It is called the Ascoli–Arzela theorem.

Theorem 2.2.4 (Ascoli–Arzela)
Let $\Omega \subset \mathbf{R}^d$ be a bounded domain. Then a subset $A \subset C^0(\overline{\Omega})$ is precompact if and only if the following two conditions are satisfied

1. $$c := \sup_{f\in A} \sup_{x\in\overline{\Omega}} |f(x)| < \infty \quad \textit{(”uniformly bounded”)},$$

2. $$\sup_{f\in A} |f(x) - f(y)| \to 0 \ \textit{ as } |x-y| \to 0.$$

   (”equicontinuous”)

❑

(here precompact means that every sequence in A has a subsequence which converges in $C^0(\overline{\Omega})$.)

Proof:
Let $(f_n)_{n\in\mathbf{N}} \subset A$ be a sequence. We have to show that it has a convergent subsequence. We will actually show that (f_n) has a subsequence which is a Cauchy sequence with respect to the norm $\|.\|_{C^0(\overline{\Omega})}$ which is sufficient since $(C^0(\overline{\Omega}), \|.\|_{C^0(\overline{\Omega})})$ is a Banach space. We first pick a sequence $(x_i)_{i\in\mathbf{N}}$ of points in $\overline{\Omega}$ which is dense in $\overline{\Omega}$, for example take $\overline{\Omega} \cap \mathbf{Q}^d$, i.e. all points in $\overline{\Omega}$ with rational coordinates. This is a countable set, we enumerate it and get $\{x_i \,|\, i \in \mathbf{N}\} = \overline{\Omega} \cap \mathbf{Q}^d$. Since

$$\sup_n \sup_{x\in\overline{\Omega}} |f_n(x)| = c < \infty$$

we have in particular

$$\sup_n |f_n(x_1)| \leq c.$$

Hence the sequence (f_n) has a subsequence $(f_{1n})_{n\in\mathbf{N}}$ so that $f_{1n}(x_1)$ converges as $n \to \infty$ by the completeness property of the real numbers. We still have

$$\sup_n \sup_{x\in\overline{\Omega}} |f_{1n}(x)| \leq c,$$

hence the sequence (f_{1n}) has a subsequence (f_{2n}) so that $f_{2n}(x_2)$ converges for $n \to \infty$. Continuing this iteration we obtain a subsequence $(f_{kn})_{n\in\mathbf{N}}$ of $(f_{k-1,n})_{n\in\mathbf{N}}$ so that $f_{kn}(x_k)$ converges as $n \to \infty$. We consider now the 'diagonal sequence' $(f_{nn})_{n\in\mathbf{N}}$. It has the property that $f_{nn}(x_i)$ converges for any $i \in \mathbf{N}$ as $n \to \infty$ ('Cantor's diagonal process').
We will now show that the sequence f_{nn} is Cauchy with respect to the norm $\|.\|_{C^0(\overline{\Omega})}$. Pick $\varepsilon > 0$. Then

1. Choose $\delta > 0$ so that

   $$|f_{kk}(y) - f_{kk}(x)| < \frac{\varepsilon}{3} \;\forall\; k \in \mathbf{N}$$

   whenever $|x - y| < \delta$. We have used here that the set A is equicontinuous.

2. Choose $M \in \mathbf{N}$ so that for each $x \in \overline{\Omega}$ there is an integer i between 1 and M so that $|x - x_i| < \delta$.
   We used here that $\overline{\Omega} \subset \mathbf{R}^d$ is compact, i.e. only a finite number of balls $B_\delta(x_i)$ is necessary to cover $\overline{\Omega}$.

3. Choose $N \in \mathbf{N}$ so that for all $1 \leq i \leq M$ and $n, m \geq N$ we have

   $$|f_{nn}(x_i) - f_{mm}(x_i)| < \frac{\varepsilon}{3}.$$

   Note that we are talking here about finitely many points x_i, hence there is no problem with finding such a number N which is good for all i.

Note that N depends on M, M depends on $\delta > 0$ which in turn depends on $\varepsilon > 0$. We now estimate for arbitrary $x \in \overline{\Omega}$, $n, m \geq N$ and for x_i as in (2)

$$\begin{aligned} |f_{nn}(x) - f_{mm}(x)| &\leq |f_{nn}(x) - f_{nn}(x_i)| + \\ &\quad + |f_{nn}(x_i) - f_{mm}(x_i)| + |f_{mm}(x_i) - f_{mm}(x)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon. \end{aligned}$$

Hence for given $\varepsilon > 0$ there is an integer $N > 0$ such that for all $n, m \geq N$ we have

$$\|f_{nn} - f_{mm}\|_{C^0(\overline{\Omega})} < \varepsilon.$$

This is what we wanted to show.

❑

Exercise 2.2.5 *Use the Ascoli–Arzela theorem to derive the following useful Corollary: Let $\Omega \subset \mathbf{R}^n$ be a bounded domain. Let $(f_n)_{n \in \mathbf{N}}$ be a sequence in $C^\infty(\overline{\Omega})$ with the following property: For every multi–index α there is a constant $c_\alpha > 0$ such that*

$$\sup_{x \in \overline{\Omega}} \sup_{n \in \mathbf{N}} |D^\alpha f_n(x)| \leq c_\alpha.$$

Then the sequence $(f_n)_{n \in \mathbf{N}}$ has a subsequence $(f_{n_l})_{l \in \mathbf{N}}$ so that f_{n_l} converges in $C^k(\overline{\Omega})$ for every $k \in \mathbf{N}$.

We will see later that $C^\infty(\overline{\Omega}) \subset C^k(\overline{\Omega})$ is dense for all $k \geq 0$ if Ω is a bounded domain with sufficiently 'nice' boundary $\partial\Omega$ (we will be more precise later if we prove this statement). This implies that $C^\infty(\overline{\Omega})$ is not a Banach space if equipped with any C^k–norm. This motivates the following two exercises:

Exercise 2.2.6 *Is it true that $C^\infty(\overline{\Omega})$ is dense in $C^{m,\alpha}(\overline{\Omega})$ if $\alpha > 0$?*

Hint: Let $\Omega = (-1, +1)$, $\alpha = 1/2$. Show that the function $f(x) := \sqrt{|x|}$ is in $C^{0,1/2}(\overline{\Omega})$ try to approximate it in the Hölder norm by smooth functions.

Exercise 2.2.7 Fréchet–metric on $C^\infty(\overline{\Omega})$
Let Ω be a bounded domain. We define for $f, g \in C^\infty(\overline{\Omega})$

$$d(f, g) := \sum_{k=0}^{\infty} 2^{-k} \frac{\|f - g\|_{C^k(\overline{\Omega})}}{1 + \|f - g\|_{C^k(\overline{\Omega})}}.$$

Show that the above expression is always finite, and verify that d defines a metric on $C^\infty(\overline{\Omega})$ such that $(C^\infty(\overline{\Omega}), d)$ is a complete metric space. Moreover, show that $d(f_n, f) \to 0$ is equivalent to the convergence of the sequence f_n to f with respect to any C^k–norm.

Here are some more problems:

Exercise 2.2.8 Prove Dini's theorem:
Let Ω be a bounded domain and let $f_n \in C^0(\overline{\Omega})$, $n \in \mathbf{N}$ so that $f_n(x) \to 0$ and $f_n(x) \geq f_{n+1}(x)$ for all $x \in \overline{\Omega}$. Then

$$\|f_n\|_{C^0(\overline{\Omega})} \longrightarrow 0 \text{ as } n \to \infty,$$

i.e. the sequence (f_n) converges uniformly to zero.

Exercise 2.2.9 Comparison of Hölder spaces:
Let Ω be a bounded domain and $0 < \alpha < \beta \leq 1$. Show that bounded subsets of $C^{0,\beta}(\overline{\Omega})$ are precompact in $C^{0,\alpha}(\overline{\Omega})$.

2.2.2 $L^p(\Omega)$ and l^p

If $\Omega \subset \mathbf{R}^n$ is a domain and $p \in \mathbf{R}$ with $1 \leq p < \infty$ we define for measurable $f : \Omega \to \mathbf{R}$

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f(x)|^p dx \right)^{\frac{1}{p}}$$

and

$$\mathcal{L}^p(\Omega) := \{ f : \Omega \to \mathbf{R} \,|\, f \text{ is measureable and } \|f\|_{L^p(\Omega)} < \infty \}.$$

The vector space $(\mathcal{L}^p(\Omega), \|.\|_{L^p(\Omega)})$ is not a normed space since $\|f\|_{L^p(\Omega)} = 0$ does not imply $f \equiv 0$. We rather introduce the following equivalence relation on the vector space $\mathcal{L}^p(\Omega)$. We say $f, g \in \mathcal{L}^p(\Omega)$ are equivalent if the set

$$\{ x \in \Omega \,|\, f(x) \neq g(x) \}$$

has measure zero. We denote the vector space of equivalence classes by $L^p(\Omega)$ which then becomes a normed space. We remark that the proof of the triangle inequality $\|f + g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}$ is not trivial (it is also called "Minkowski inequality"). We will discuss the proof in a moment. For measureable $f : \Omega \to \mathbf{R}$ we define

$$\begin{aligned} \text{ess sup}_{x \in \Omega} f(x) &:= \inf\{ c \in \mathbf{R} \cup \{\infty\} \,|\, f(x) \leq c \text{ for almost all } x \in \Omega \} \\ &= \inf\Big\{ \sup_{x \in \Omega \setminus N} |f(x)| \,:\, N \subset \Omega\,,\, |N| = 0 \Big\} \end{aligned}$$

and we denote by $L^\infty(\Omega)$ the equivalence classes of all measureable functions with

$$\|f\|_{L^\infty(\Omega)} := \text{ess sup}_{x \in \Omega} |f(x)| < \infty.$$

These are functions which are bounded except on a set of measure zero.

Exercise 2.2.10 *Show that $(L^\infty(\Omega), \|.\|_{L^\infty(\Omega)})$ is a Banach space.*

We will often be somewhat sloppy and talk about a measureable function being in the space $L^p(\Omega)$ instead of referring to its equivalence class.
We will show a fundamental inequality ("Hölder inequality") for L^p–spaces which will imply among other things Minkowski's inequality. After that we will show that the spaces $L^p(\Omega)$ are Banach spaces for $1 \leq p < \infty$.

Theorem 2.2.11 (Hölder's inequality)
Let $\Omega \in \mathbf{R}^n$ be a domain and $f \in L^p(\Omega)$, $g \in L^q(\Omega)$ with $1 \leq p, q \leq \infty$ such that

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

Then $fg \in L^1(\Omega)$ and

$$\|fg\|_{L^1(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}.$$

Proof:
The theorem is obvious if $p = 1$ and $q = \infty$ or vice versa. Hence we assume that $1 < p, q < \infty$. Recall Young's inequality which is

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q \quad \forall\, a, b \geq 0.$$

The proof is evident: Since the logarithm function is concave on $(0, \infty)$ we have

$$\log\left(\frac{1}{p} a^p + \frac{1}{q} b^q\right) \geq \frac{1}{p} \log a^p + \frac{1}{q} \log b^q = \log(ab).$$

Therefore,

$$|f(x)|\, |g(x)| \leq \frac{1}{p} |f(x)|^p + \frac{1}{q} |g(x)|^q$$

for almost all $x \in \Omega$. We conclude that $fg \in L^1(\Omega)$ and that

$$\int_\Omega |fg|(x) dx \leq \frac{1}{p} \|f\|^p_{L^p(\Omega)} + \frac{1}{q} \|g\|^q_{L^q(\Omega)}.$$

Replacing now f by λf, where $\lambda > 0$ we obtain

$$\int_\Omega |fg|(x) dx \leq \frac{\lambda^{p-1}}{p} \|f\|^p_{L^p(\Omega)} + \frac{1}{\lambda q} \|g\|^q_{L^q(\Omega)}.$$

Choosing now $\lambda = \|f\|^{-1}_{L^p(\Omega)} \|g\|^{q/p}_{L^q(\Omega)}$ we obtain

$$\begin{aligned} \int_\Omega |fg|(x) dx &\leq \frac{\|f\|_{L^p(\Omega)}}{p} \|g\|^{q\frac{p-1}{p}}_{L^q(\Omega)} + \frac{\|f\|_{L^p(\Omega)}}{q} \|g\|^{q-q/p}_{L^q(\Omega)} \\ &= \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)} \left(\frac{1}{p} + \frac{1}{q}\right). \end{aligned}$$

□

Before we proceed, let us note some useful consequences of Hölder's inequality. The first one is Minkowski's inequality:

Theorem 2.2.12 (Minkowski's inequality)
Let $1 \leq p \leq \infty$ and $f, g \in L^p(\Omega)$. Then $f + g \in L^p(\Omega)$ and

$$\|f + g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}.$$

Proof:
The cases $p = 1$ and $p = \infty$ are obvious, so let us assume that $1 < p < \infty$. We have

$$|f(x) + g(x)|^p \leq (|f(x)| + |g(x)|)^p \leq 2^{p-1}(|f(x)|^p + |g(x)|^p)$$

so that $f + g \in L^p(\Omega)$. Let now $q := \frac{p}{p-1}$ so that $\frac{1}{p} + \frac{1}{q} = 1$. We use the trivial inequality

$$|f(x) + g(x)|^p \leq |f(x)|\,|f(x) + g(x)|^{p-1} + |g(x)|\,|f(x) + g(x)|^{p-1}$$

and we note that the function $|f(x) + g(x)|^{p-1}$ is in $L^q(\Omega)$. Then we conclude from Hölder's inequality

$$\begin{aligned}\int_\Omega |f(x) + g(x)|^p dx &\leq \|f\|_{L^p(\Omega)} \||f + g|^{p-1}\|_{L^q(\Omega)} + \|g\|_{L^p(\Omega)} \||f + g|^{p-1}\|_{L^q(\Omega)} \\ &= (\|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}) \left(\int_\Omega |f(x) + g(x)|^p dx\right)^{1-\frac{1}{p}}.\end{aligned}$$

If $\int_\Omega |f + g|^p = 0$ then Minkowski's inequality is trivially true. Otherwise we divide the above inequality by $\left(\int_\Omega |f(x) + g(x)|^p dx\right)^{1-\frac{1}{p}}$.

□

Exercise 2.2.13 *Prove the following generalization of Hölder's inequality: Let $p_1, \ldots, p_n \geq 1$ so that*

$$\frac{1}{p_1} + \cdots + \frac{1}{p_n} = 1$$

and $f_k \in L^{p_k}(\Omega)$, $k = 1, \ldots, n$. Then

$$\int_\Omega |f_1(x) \cdots f_n(x)| dx \leq \|f_1\|_{L^{p_1}(\Omega)} \cdots \|f_n\|_{L^{p_n}(\Omega)}.$$

Here are some simple consequences of Hölder's inequality

Corollary 2.2.14 *Let $\Omega \subset \mathbf{R}^n$ be a bounded domain and $1 \leq p \leq q \leq \infty$. Then $L^q(\Omega) \subset L^p(\Omega)$ and*

$$|\Omega|^{-\frac{1}{p}} \|f\|_{L^p(\Omega)} \leq |\Omega|^{-\frac{1}{q}} \|f\|_{L^q(\Omega)} \;\forall f \in L^q(\Omega).$$

Proof:
The case $q = \infty$ is obvious, hence assume that $q < \infty$. Using Hölder's inequality we obtain

$$\begin{aligned} \|f\|^p_{L^p(\Omega)} &= \int_\Omega 1 \cdot |f(x)|^p dx \\ &\leq \|1\|_{L^{\frac{q}{q-p}}(\Omega)} \||f|^p\|_{L^{\frac{q}{p}}(\Omega)} \\ &= |\Omega|^{1-\frac{p}{q}} \|f\|^p_{L^q(\Omega)} \end{aligned}$$

since $\frac{p}{q} + \frac{q-p}{q} = 1$.

❑

Corollary 2.2.15 (interpolation inequality)
Assume that $1 \leq p \leq q \leq r$ *and* $0 \leq \lambda \leq 1$ *with*

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r}.$$

If $f \in L^p(\Omega) \cap L^r(\Omega)$ *then also* $f \in L^q(\Omega)$ *and*

$$\|f\|_{L^q(\Omega)} \leq \|f\|^\lambda_{L^p(\Omega)} \cdot \|f\|^{1-\lambda}_{L^r(\Omega)}.$$

Proof:
We have $\frac{1}{p_1} + \frac{1}{p_2} = 1$ if we choose $p_1 = \frac{p}{\lambda q}$ and $p_2 = \frac{r}{(1-\lambda)q}$. Then we obtain from Hölder's inequality

$$\begin{aligned} \int_\Omega |f(x)|^q dx &= \int_\Omega |f(x)|^{\lambda q} \cdot |f(x)|^{(1-\lambda)q} dx \\ &\leq \||f|^{\lambda q}\|_{L^{p_1}(\Omega)} \||f|^{(1-\lambda)q}\|_{L^{p_2}(\Omega)} \\ &= \|f\|^{\lambda q}_{L^p(\Omega)} \|f\|^{(1-\lambda)q}_{L^r(\Omega)}. \end{aligned}$$

❑

The following interesting result explains why the space $L^\infty(\Omega)$ is called like this (if Ω is a bounded domain).

Proposition 2.2.16 *Let* $\Omega \subset \mathbf{R}^n$ *be a bounded domain. For* $f \in L^p(\Omega)$, $1 \leq p < \infty$ *we define*

$$\Phi_p(f) := \left(\frac{1}{|\Omega|} \int_\Omega |f(x)|^p dx \right)^{1/p}.$$

If $|f|^p$ *is merely measureable, but not integrable we set* $\Phi_p(f) := +\infty$. *Then for every measureable function* $f : \Omega \to \mathbf{R} \cup \{\pm\infty\}$

$$\lim_{p \to \infty} \Phi_p(f) = \|f\|_{L^\infty(\Omega)}.$$

Proof:
We have
$$\Phi_p(f) = |\Omega|^{-\frac{1}{p}} \|f\|_{L^p(\Omega)}.$$
By corollary 2.2.14, $\Phi_p(f)$ viewed as a function of p is increasing with
$$\Phi_p(f) \le \|f\|_{L^\infty(\Omega)}.$$
Therefore, the limit $\lim_{p\to\infty} \Phi_p(f) \in \mathbf{R} \cup \{\infty\}$ exists, and it remains to show that
$$\|f\|_{L^\infty(\Omega)} \le \lim_{p\to\infty} \Phi_p(f).$$
For $K \in \mathbf{R}$ let
$$A_K := \{x \in \Omega \,|\, |f(x)| \ge K\}.$$
The set A_k is measureable since f is and $|A_K| > 0$ if $K < \|f\|_{L^\infty(\Omega)}$. Moreover,
$$\Phi_p(f) \ge |\Omega|^{-\frac{1}{p}} \left(\int_{A_K} |f(x)|^p dx \right)^{1/p} \ge |\Omega|^{-\frac{1}{p}} |A_K|^{\frac{1}{p}} K.$$
Passing to the limit $p \to \infty$ we obtain
$$\lim_{p\to\infty} \Phi_p(f) \ge K.$$
Because this holds for all $K < \|f\|_{L^\infty(\Omega)}$ we conclude
$$\lim_{p\to\infty} \Phi_p(f) \ge \|f\|_{L^\infty(\Omega)}.$$

❑

Theorem 2.2.17 (Fischer–Riesz)
The space $(L^p(\Omega), \|.\|_{L^p(\Omega)})$ is a Banach space.

Proof:
Let $(f_k)_{k\in\mathbf{N}} \subset L^p(\Omega)$ be a Cauchy sequence. It suffices to show that (f_k) has a convergent subsequence. For every $i \in \mathbf{N}$ there is an integer N_i so that
$$\|f_n - f_m\|_{L^p(\Omega)} \le 2^{-i} \text{ whenever } n, m \ge N_i.$$
We construct a subsequence $(f_{k_i}) \subset (f_k)$ so that
$$\|f_{k_{i+1}} - f_{k_i}\|_{L^p(\Omega)} \le 2^{-i}$$
by setting $k_i := \max\{i, N_i\}$. In order to simplify notation we will from now on assume that
$$\|f_{k+1} - f_k\|_{L^p(\Omega)} \le 2^{-k}$$
so that
$$M := \sum_{k\in\mathbf{N}} \|f_{k+1} - f_k\|_{L^p(\Omega)} < \infty.$$

We define

$$g_l(x) := \sum_{k=1}^{l} |f_{k+1}(x) - f_k(x)|.$$

The sequence $(g_l^p(x))_{l\in\mathbf{N}}$ is monotone increasing and consists of nonnegative integrable functions since we have

$$\int_\Omega g_l^p(x)dx = \|g_l\|_{L^p(\Omega)}^p \leq \left(\sum_{k=1}^{l} \|f_{k+1} - f_k\|_{L^p(\Omega)}\right)^p \leq M^p.$$

By the theorem on monotone convergence the sequence $(g_l^p)_{l\in\mathbf{N}}$ converges pointwise almost everywhere to some integrable function h. This implies by definition of g_l that the sequence $(f_k(x))_{k\in\mathbf{N}}$ is a Cauchy sequence in $\mathbf{R}$ for almost every $x \in \Omega$ so that the pointwise limit

$$f(x) := \lim_{k\to\infty} f_k(x)$$

exists almost everywhere. We now apply Fatou's lemma to the sequence of integrable functions $(|f_k - f_l|^p)_{k\in\mathbf{N}}$ and conclude

$$\begin{aligned}
\int_\Omega |f(x) - f_l(x)|^p dx &\leq \liminf_{k\to\infty} \int_\Omega |f_k(x) - f_l(x)|^p dx \\
&= (\liminf_{k\to\infty} \|f_k - f_l\|_{L^p(\Omega)})^p \\
&\leq \left(\sum_{k\geq l} \|f_{k+1} - f_k\|_{L^p(\Omega)}\right)^p
\end{aligned}$$

which tends to zero as $l \to \infty$. This shows that the sequence (f_k) converges to f in the L^p–norm and it also shows that $f - f_l \in L^p(\Omega)$ and therefore $f \in L^p(\Omega)$.

❑

During the proof of theorem 2.2.17 we have also proved the following:

Corollary 2.2.18 *Let $1 \leq p \leq \infty$ and let $(f_k)_{k\in\mathbf{N}} \subset L^p(\Omega)$ be a sequence which converges in $L^p(\Omega)$ to some $f \in L^p(\Omega)$. Then there is a subsequence which converges pointwise almost everywhere to f.*

❑

Exercise 2.2.19 *Find a sequence $(f_k) \subset L^p(\Omega)$ which converges in $L^p(\Omega)$ but which does not converge pointwise almost everywhere. This means that the above corollary only holds for a suitable subsequence not for the whole sequence (f_k).*

We have mentioned earlier that integrable functions can be approximated in L^1 by continuous functions with compact support. We will show that $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$. We will use the concept of 'mollifiers', a convenient method to obtain approximations by smooth functions.

Theorem 2.2.20 *$C^0(\Omega)$ is dense in $L^p(\Omega)$ if $1 \le p < \infty$.*

Proof:
We use the result from measure theory that every integrable function can be approximated in the L^1–norm by continuous functions with compact support. First, we view $f \in L^p(\Omega)$ as a function on $L^p(\mathbf{R}^n)$ simply by continuing it trivially outside the domain Ω. It is also sufficient to consider the case $f \ge 0$, otherwise consider $\max\{f, 0\}$ and $-\min\{0, f\}$ separately. We then define for $k \in \mathbf{N}$

$$f_k(x) := \begin{cases} \min\left(f(x), k\right) & \text{for} \quad |x| \le k \\ 0 & \text{for} \quad |x| > k. \end{cases}$$

so that each function f_k is integrable. In view of $|f_k - f|^p \le |f|^p$ the convergence theorem of H. Lebesgue implies that $f_k \to f$ with respect to the $L^p(\mathbf{R}^n)$–norm. Hence for every $\varepsilon > 0$ there is some integer k so that

$$\|f_k - f\|_{L^p(\mathbf{R}^n)} < \frac{\varepsilon}{2}.$$

We can also find a continuous function $\phi \in C_0^0(\mathbf{R}^n)$ so that

$$\|f_k - \phi\|_{L^1(\mathbf{R}^n)} < \frac{\varepsilon^p}{2^{2p-1} k^{p-1}}.$$

Since $0 \le f_k \le k$ we may assume that $0 \le \phi \le k$ as well, otherwise replace ϕ by $\min\{\max\{\phi, 0\}, k\}$. We then have $|f_k - \phi| \le 2k$ and

$$|f_k - \phi|^p \le (2k)^{p-1} |f_k - \phi|$$

and therefore

$$\begin{aligned} \|f_k - \phi\|^p_{L^p(\Omega)} &\le \|f_k - \phi\|^p_{L^p(\mathbf{R}^n)} \\ &\le 2^{p-1} k^{p-1} \frac{\varepsilon^p}{2^{2p-1} k^{p-1}} \\ &< \left(\frac{\varepsilon}{2}\right)^p. \end{aligned}$$

We obtain

$$\|f - \phi\|_{L^p(\Omega)} < \varepsilon.$$

❑

We introduce the concept of convolution:

Proposition 2.2.21 *Let $1 \leq p \leq \infty$, $f \in L^p(\mathbf{R}^n)$ and $\phi \in L^1(\mathbf{R}^n)$. The following integral exists*

$$(f * \phi)(x) := \int_{\mathbf{R}^n} \phi(x-y)f(y)dy = \int_{\mathbf{R}^n} \phi(y)f(x-y)dy$$

and

$$\|f * \phi\|_{L^p(\mathbf{R}^n)} \leq \|f\|_{L^p(\mathbf{R}^n)} \|\phi\|_{L^1(\mathbf{R}^n)}.$$

If $\phi \in C_0^\infty(\mathbf{R}^n)$ then $f * \phi \in C^\infty(\mathbf{R}^n)$ and

$$D^\alpha (f * \phi)(x) = \int_{\mathbf{R}^n} D_x^\alpha \phi(x-y) f(y)dy,$$

where D_x^α denotes differentiation with respect to the variable x.

Definition 2.2.22 *We call $f * \phi$ the convolution of f with ϕ. In the case where ϕ is smooth with compact support we call $f * \phi$ a mollifier of f.*

Proof:
The case $p = \infty$ is trivial, it follows from the translation invariance of the Lebesgue measure. Let us consider first the case $p = 1$. We remark that the function $\Phi(x, y) := \phi(x - y)$ is measureable on $\mathbf{R}^n \times \mathbf{R}^n$ if ϕ is (show this as an exercise). Assume for the moment that both f and ϕ are not negative. Then the product $\phi(x - y)f(y)$ is a non negative measureable function on $\mathbf{R}^n \times \mathbf{R}^n$. We may apply Tonelli's theorem and obtain

$$\begin{aligned}
\int_{\mathbf{R}^n \times \mathbf{R}^n} \phi(x-y)f(y)dxdy &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} \phi(x-y)f(y)dy \right) dx \\
&= \int_{\mathbf{R}^n} f(y) \left(\int_{\mathbf{R}^n} \phi(x-y)dx \right) dy \\
&= \left(\int_{\mathbf{R}^n} f(y)dy \right) \left(\int_{\mathbf{R}^n} \phi(x)dx \right)
\end{aligned}$$

which can be written as

$$\int_{\mathbf{R}^n} (f * \phi)(x)dx = \left(\int_{\mathbf{R}^n} f(y)dy \right) \left(\int_{\mathbf{R}^n} \phi(x)dx \right).$$

This proves the case $p = 1$ for $\phi, f \geq 0$. The general case then follows from the estimate $|f * \phi| \leq |f| * |\phi|$.
We are left with the case where $1 < p < \infty$. Choosing $1 < q < \infty$ so that $1/p + 1/q = 1$ we use Hölder's inequality

$$\begin{aligned}
|(f * \phi)(x)| &\leq \int_{\mathbf{R}^n} \left| \left[f(y)\phi^{\frac{1}{p}}(x-y) \right] \phi^{\frac{1}{q}}(x-y) \right| dy \\
&\leq \left(\int_{\mathbf{R}^n} |f(y)|^p \, |\phi(x-y)| \, dy \right)^{1/p} \left(\int_{\mathbf{R}^n} |\phi(x-y)| \, dy \right)^{1/q} \\
&= (|f|^p * |\phi|)^{1/p}(x) \|\phi\|_{L^1(\mathbf{R}^n)}^{1/q}
\end{aligned}$$

Raise to the power p and integrate so that

$$\begin{aligned} \|f * \phi\|_{L^p(\mathbf{R}^n)}^p &\leq \|\phi\|_{L^1(\mathbf{R}^n)}^{p/q} \||f|^p * |\phi|\|_{L^1(\mathbf{R}^n)} \\ &\leq \|\phi\|_{L^1(\mathbf{R}^n)}^{1+\frac{p}{q}} \||f|^p\|_{L^1(\mathbf{R}^n)} \\ &= \|\phi\|_{L^1(\mathbf{R}^n)}^p \|f\|_{L^p(\mathbf{R}^n)}^p \end{aligned}$$

which is the desired inequality. We leave the situation where ϕ is smooth with compact support as an exercise for the reader. One has to take care of the following points: Using that ϕ is smooth with compact support show first that $f * \phi$ is continuous, i.e.

$$|(f * \phi)(x + h) - (f * \phi)(x)| \to 0 \text{ as } |h| \to 0.$$

Then justify differentiation under the integral sign.

❑

Let Ω be a domain in $\mathbf{R}^n$. If $\Omega' \subset \Omega$ is a bounded subdomain so that $\overline{\Omega'} \subset \Omega$ as well, then we use the shorthand notation $\Omega' \subset\subset \Omega$. Let ρ be a nonnegative smooth function with support in the unit ball in $\mathbf{R}^n$ so that

$$\int_{\mathbf{R}^n} \rho(x)dx = 1.$$

An example for such a function is

$$\rho(x) := \begin{cases} c \exp\left(\frac{1}{|x|^2-1}\right) & \text{for} \quad |x| < 1 \\ 0 & \text{for} \quad |x| \geq 1, \end{cases}$$

where the constant $c > 0$ is chosen so that the integral of ρ equals 1. Let $f \in L^1_{loc}(\Omega)$, i.e. every point in Ω has a neighborhood over which f is integrable. Let $\Omega' \subset\subset \Omega$ and $\varepsilon < \operatorname{dist}(\Omega', \partial\Omega)$. We then define

$$f_\varepsilon(x) := (f * \rho_\varepsilon)(x) \ , x \in \Omega',$$

where

$$\rho_\varepsilon(z) := \frac{1}{\varepsilon^n}\rho\left(\frac{z}{\varepsilon}\right).$$

Remarks: The function

$$y \longmapsto f(y)\,\rho\left(\frac{x-y}{\varepsilon}\right)$$

has support in the ball $B_\varepsilon(x)$. Therefore the function f_ε is only defined on the smaller domain Ω' unless we extend f trivially onto all of $\mathbf{R}^n$. The function f_ε is smooth.

We use now the smooth functions f_ε for approximating L^p–functions. This procedure will also be used later for Sobolev spaces.

Proposition 2.2.23 *If $f \in C^0(\Omega)$ then the functions f_ε converge on every subdomain $\Omega' \subset\subset \Omega$ uniformly to f as $\varepsilon \to 0$.*

Proof:
We have

$$\begin{aligned} f_\varepsilon(x) &= \frac{1}{\varepsilon^n}\int_{\mathbf{R}^n} f(y)\rho\left(\frac{x-y}{\varepsilon}\right)dy \\ &= \frac{1}{\varepsilon^n}\int_{B_\varepsilon(x)} f(y)\rho\left(\frac{x-y}{\varepsilon}\right)dy \\ &= \int_{|z|\le 1}\rho(z)f(x-\varepsilon z)dz. \end{aligned}$$

If $\Omega' \subset\subset \Omega$, $\varepsilon < \frac{1}{2}\mathrm{dist}(\Omega', \partial\Omega)$ and recalling that the integral of the function ρ over the unit ball equals 1 we obtain

$$\begin{aligned} \sup_{x\in\Omega'}|f(x)-f_\varepsilon(x)| &= \sup_{x\in\Omega'}\left|\int_{|z|\le 1}\rho(z)[f(x)-f(x-\varepsilon z)]dz\right| \\ &\le \sup_{x\in\Omega'}\int_{|z|\le 1}\rho(z)\,|f(x)-f(x-\varepsilon z)|dz \\ &\le \sup_{x\in\Omega'}\sup_{|z|\le 1}|f(x)-f(x-\varepsilon z)|\int_{|z|\le 1}\rho(z)dz \\ &= \sup_{x\in\Omega'}\sup_{|z|\le 1}|f(x)-f(x-\varepsilon z)|. \end{aligned}$$

The function f is uniformly continuous on the compact set $\{x \in \Omega \,|\, \mathrm{dist}(x, \Omega') \le \varepsilon\}$, therefore the right hand side tends to zero as $\varepsilon \to 0$.

❑

Theorem 2.2.24 *Let $1 \le p < \infty$ and $f \in L^p(\Omega)$. Viewing f as an element in $L^p(\mathbf{R}^n)$ by trivial extension we have*

$$f_\varepsilon \longrightarrow f \text{ in } L^p(\mathbf{R}^n)$$

as $\varepsilon \to 0$.

Proof:
We write as before

$$\begin{aligned} |f_\varepsilon(x)| &= \left|\int_{|z|\le 1}\rho(z)f(x-\varepsilon z)dz\right| \\ &\le \int_{|z|\le 1}\rho(z)^{1-\frac{1}{p}}\,\rho(z)^{\frac{1}{p}}|f(x-\varepsilon z)|dz \end{aligned}$$

$$
\begin{aligned}
&\le \left(\int_{|z|\le 1}\rho(z)dz\right)^{1-\frac{1}{p}}\left(\int_{|z|\le 1}\rho(z)|f(x-\varepsilon z)|^p dz\right)^{\frac{1}{p}} \\
&\quad \text{(with Hölder's inequality)} \\
&= \left(\int_{|z|\le 1}\rho(z)|f(x-\varepsilon z)|^p dz\right)^{\frac{1}{p}}.
\end{aligned}
$$

Let $\tilde{\Omega} \subset \Omega$ be a subdomain and define a slightly larger domain by

$$\tilde{\Omega}_\varepsilon := \{x \in \Omega \,|\, \mathrm{dist}(x, \tilde{\Omega}) \le \varepsilon\}.$$

We conclude with $\int_{|z|\le 1}\rho(z)dz = 1$

$$
\begin{aligned}
\int_{\tilde{\Omega}} |f_\varepsilon(x)|^p dx &\le \int_{\tilde{\Omega}}\int_{|z|\le 1}\rho(z)|f(x-\varepsilon z)|^p dz\, dx \\
&= \int_{|z|\le 1}\rho(z)\left(\int_{\tilde{\Omega}}|f(x-\varepsilon z)|^p dx\right)dz \\
&\quad \text{(Fubini's theorem)} \\
&\le \left(\int_{|z|\le 1}\rho(z)dz\right)\left(\int_{\tilde{\Omega}_\varepsilon}|f(y)|^p dy\right) \\
&= \int_{\tilde{\Omega}_\varepsilon}|f(y)|^p dy.
\end{aligned}
$$

If $\tilde{\Omega} = \Omega$ then we obtain $\|f_\varepsilon\|_{L^p(\Omega)} \le \|f\|_{L^p(\Omega)}$. Let now $\varepsilon' > 0$. We claim that we can choose $R > 0$ so large that

$$\|f - f_\varepsilon\|_{L^p(\Omega\setminus\overline{B_R(0)})} < \frac{\varepsilon'}{4} \tag{2.1}$$

for all sufficiently small $\varepsilon > 0$. It is important here that R does not depend on ε. This follows from the inequality

$$\|f_\varepsilon\|_{L^p(\Omega\setminus\overline{B_R(0)})} \le \|f\|_{L^p(\Omega\setminus\overline{B_{R-\varepsilon}(0)})}$$

which we have just proved. We can then choose R so that

$$\|f\|_{L^p(\Omega\setminus\overline{B_{R-1}(0)})} < \frac{\varepsilon'}{8}$$

which implies (2.1) for all $\varepsilon < 1$. By theorem 2.2.20 we can find $\phi \in C_0^0(\mathbf{R}^n)$ with

$$\|f - \phi\|_{L^p(\mathbf{R}^n)} < \frac{\varepsilon'}{4}.$$

By proposition 2.2.23 we have for sufficiently small $\varepsilon > 0$

$$\|\phi - \phi_\varepsilon\|_{L^p(\Omega\cap B_R(0))} \le \sup_{x\in\Omega\,,\,|x|\le R}|\phi(x) - \phi_\varepsilon(x)| < \frac{\varepsilon'}{4}$$

(we apply proposition 2.2.23 as follows: Take $\Omega' = \Omega \cap B_R(0)$ and take for Ω a slightly larger domain). Using the fact that the L^p–norm of $(f-\phi)_\varepsilon = f_\varepsilon - \phi_\varepsilon$ is bounded by the L^p–norm of $f-\phi$ we get

$$\begin{aligned}
\|f-f_\varepsilon\|_{L^p(\Omega)} &\leq \|f-f_\varepsilon\|_{L^p(\Omega\setminus\overline{B_R(0)})} + \|f-f_\varepsilon\|_{L^p(\Omega\cap B_R(0))} \\
&\leq \frac{\varepsilon'}{4} + \|f-\phi\|_{L^p(\mathbf{R}^n)} + \\
&\quad + \|\phi-\phi_\varepsilon\|_{L^p(\Omega\cap B_R(0))} + \|f_\varepsilon-\phi_\varepsilon\|_{L^p(\mathbf{R}^n)} \\
&\leq \frac{\varepsilon'}{4} + \frac{\varepsilon'}{4} + \frac{\varepsilon'}{4} + \|f-\phi\|_{L^p(\mathbf{R}^n)} \\
&\leq \varepsilon'.
\end{aligned}$$

❑

Theorem 2.2.25 *The set $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$ if $1 \leq p < \infty$.*

Proof:
Given $f \in L^p(\Omega)$ and $\varepsilon' > 0$ we have to find a smooth function ϕ with compact support in Ω such that

$$\|f-\phi\|_{L^p(\Omega)} < \varepsilon'.$$

First we choose a domain $\Omega' \subset\subset \Omega$ such that

$$\|f\|_{L^p(\Omega\setminus\overline{\Omega'})} < \frac{\varepsilon'}{3}.$$

Then we define

$$\tilde{f}(x) := \begin{cases} f(x) & \text{if} \quad x \in \overline{\Omega'} \\ 0 & \text{if} \quad x \in \mathbf{R}^n\setminus\overline{\Omega'} \end{cases}.$$

By theorem 2.2.24 there is some $0 < \varepsilon < \frac{1}{2}\mathrm{dist}(\Omega', \partial\Omega)$ so that

$$\|\tilde{f}-\tilde{f}_\varepsilon\|_{L^p(\Omega)} < \frac{\varepsilon'}{3}.$$

Since $\tilde{f} \equiv 0$ outside the domain Ω' we also have

$$\|\tilde{f}_\varepsilon\|_{L^p(\Omega\setminus\overline{\Omega'})} = \|\tilde{f}-\tilde{f}_\varepsilon\|_{L^p(\Omega\setminus\overline{\Omega'})} \leq \|\tilde{f}-\tilde{f}_\varepsilon\|_{L^p(\Omega)} < \frac{\varepsilon'}{3}$$

so that

$$\|f-\tilde{f}_\varepsilon\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega\setminus\overline{\Omega'})} + \|\tilde{f}_\varepsilon\|_{L^p(\Omega\setminus\overline{\Omega'})} + \|\tilde{f}-\tilde{f}_\varepsilon\|_{L^p(\Omega')} < \varepsilon'.$$

The function $\tilde{f}_\varepsilon$ is smooth and its support is compact and contained in Ω by our choice $0 < \varepsilon < \frac{1}{2}\mathrm{dist}(\Omega', \partial\Omega)$

❑

Proposition 2.2.26 *Let $f \in L^1(\Omega)$ so that*

$$\int_{\Omega} f(x)\phi(x)dx = 0$$

for all $\phi \in C_0^\infty(\Omega)$. Then $f \equiv 0$ almost everywhere.

Proof:
Let E be a bounded measureable set with $\overline{E} \subset \Omega$ and $\operatorname{dist}(\overline{E}, \partial\Omega) > 0$. Denote the characteristic function of E by χ. i.e. $\chi|_E \equiv 1$ and zero otherwise. Define now

$$\zeta_\varepsilon(x) := (\chi * \rho_\varepsilon)(x)$$

which equals

$$\frac{1}{\varepsilon^n}\int_E \rho\left(\frac{x-y}{\varepsilon}\right) dy = \int_{\mathbf{R}^n} \chi(x-\varepsilon z)\rho(z)\, dz.$$

For sufficiently small $\varepsilon > 0$ the functions ζ_ε are smooth with compact support in Ω and $0 \le \zeta_\varepsilon(x) \le 1$. By theorem 2.2.24 we have $\zeta_\varepsilon \to \chi$ in $L^p(\mathbf{R}^n)$ as $\varepsilon \to 0$ for all $1 \le p < \infty$. By corollary 2.2.18 we can extract a subsequence which converges pointwise almost everywhere. Without loss of generality we will therefore assume that $\zeta_\varepsilon \to \chi$ pointwise almost everywhere. The convergence theorem of H. Lebesgue now implies that $f\zeta_\varepsilon \to f\chi$ in $L^1(\Omega)$ and

$$0 = \int_\Omega f(x)\zeta_\varepsilon(x)dx \to \int_\Omega f(x)\chi(x)dx = \int_E f(x)dx,$$

i.e.

$$\int_E f(x)dx = 0 \tag{2.2}$$

which holds for arbitrary measureable sets E as specified above. Let now $\Omega' \subset\subset \Omega$ be a bounded subdomain. Define

$$\Omega'_\pm := \{x \in \Omega' \mid \pm f(x) > 0\}.$$

Apply now (2.2) to the measureable sets $\Omega'_\pm$ so that

$$\int_{\Omega'} |f(x)|dx = \int_{\Omega'_+} f(x)dx - \int_{\Omega'_-} f(x)dx = 0.$$

We conclude $f|_{\Omega'} \equiv 0$ almost everywhere and also $f \equiv 0$ almost everywhere since Ω' was arbitrary.

❑

Using convolutions we will now prove the following theorem which provides criteria when a subset $A \subset L^p(\mathbf{R}^n)$ is precompact. It is the L^p–version of the theorem of Ascoli–Arzela. Let us first insert the following topological definition and lemma:

Definition 2.2.27 *A subset A of a metric space (X,d) is called totally bounded if for every $\varepsilon > 0$ there is an integer $N = N_\varepsilon > 0$ and finitely many balls $B_1, \ldots, B_N$ of radius ε such that*

$$\bigcup_{1\le k\le N} B_k \supset A.$$

Lemma 2.2.28 *Let (X,d) be a complete metric space, and let $A \subset X$ be a subset. Then A is totally bounded if and only if it is precompact.*

Proof:
Assume that A is precompact. If A was not totally bounded then we could find some $\varepsilon > 0$ so that A can not be covered by finitely many balls of radius ε. We can now define a sequence $(x_k)_{k\in\mathbf{N}} \subset A$ inductively by

$$x_{k+1} \in A \setminus \bigcup_{1\le i\le k} B_\varepsilon(x_i).$$

This sequence has no convergent subsequence contradicting precompactness.
Assume now that A is totally bounded and let $(x_k)_{k\in\mathbf{N}}$ be a sequence in A. We have to show that it has a convergent subsequence. Let ε_k be a sequence of positive numbers converging to zero. We can cover A with finitely many balls of radius ε_1. At least one of these balls, say B_{ε_1}, contains infinitely many members of the sequence (x_k). We may then cover the ball B_{ε_1} with finitely many balls of radius ε_2, and at least one of those again contains infinitely many of the points $\{x_k\}_{k\in\mathbf{N}} \cap B_{\varepsilon_1}$. Let this ball be B_{ε_2}. Now cover $B_{\varepsilon_1} \cap B_{\varepsilon_2}$ with finitely many balls of radius ε_3. Then one of them, B_{ε_3} will contain infinitely many of the points $\{x_k\}_{k\in\mathbf{N}} \cap B_{\varepsilon_1} \cap B_{\varepsilon_2}$. Continuing this process we obtain an infinite sequence of nested sets

$$C_l := \bigcap_{1\le i\le l} B_{\varepsilon_i}\ ,\ C_{l+1} \subset C_l$$

Each of the C_l contains infinitely many elements of the sequence (x_k) and the diameters of the sets C_l tend to zero as $l \to \infty$. For each $l \in \mathbf{N}$ we pick

$$y_l \in \{x_k\}_{k\in\mathbf{N}} \cap C_l$$

so that $|y_l - y_{l'}| \le \max\{\operatorname{diam}(C_l), \operatorname{diam}(C_{l'})\} \to 0$ as $l, l' \to \infty$. Hence (y_l) is a Cauchy sequence which converges since X is complete.

❑

Theorem 2.2.29 (Fréchet–Kolmogorov)
Let $1 \le p < \infty$, $\Omega \subset \mathbf{R}^n$ a domain and $\Omega' \subset\subset \Omega$ a bounded subdomain. Then a subset $A \subset L^p(\Omega)$ is precompact in $L^p(\Omega')$ if the following conditions are satisfied:

1. $\sup_{f\in A}\|f\|_{L^p(\Omega)} < \infty$,

2. For all $\varepsilon' > 0$ *there is a number* $0 < \delta < dist(\Omega', \partial\Omega)$ *so that*

$$\sup_{f\in A}\|f(\,\cdot\,+h) - f\|_{L^p(\Omega')} \le \varepsilon'$$

for all $h \in \mathbf{R}^n$ *with* $|h| < \delta$. *Here* $f(\,\cdot\,+h)$ *denotes the function* $x \mapsto f(x+h)$.

Proof:
The idea of the proof is the following: We 'mollify' the whole family A. We then get a family of smooth functions which will satisfy the assumptions of the Ascoli–Arzela theorem so that every sequence will have a uniformly convergent subsequence. Uniform convergence on a bounded domain implies L^p–convergence.
We may assume first that the domain Ω is bounded since we are only interested in L^p–convergence on the bounded domain Ω'. Otherwise replace Ω with a bounded one which contains $\overline{\Omega'}$. Then we assume that all the elements $f \in A$ are actually defined on all of $\mathbf{R}^n$ by trivially extending them outside Ω. Denote the extended version of f by $\tilde{f}$. Then the set

$$\tilde{A} := \{\tilde{f} \in L^p(\mathbf{R}^n)\,|\, f \in A\}$$

is bounded in $L^p(\mathbf{R}^n)$ and also in $L^1(\mathbf{R}^n)$ (remember that Ω is now bounded !). We claim that

$$\sup_{\tilde{f}\in\tilde{A}}\|\tilde{f} * \rho_\varepsilon - \tilde{f}\|_{L^p(\Omega')} \le \varepsilon'$$

for all $\varepsilon < \delta$. Using the same method as in the proof of theorem 2.2.24, we estimate

$$\begin{aligned}
|(\tilde{f} * \rho_\varepsilon)(x) - \tilde{f}(x)| &= \left|\int_{|z|\le 1}\rho(z)[\tilde{f}(x) - \tilde{f}(x-\varepsilon z)]dz\right| \\
&\le \left(\int_{|z|\le 1}\rho(z)|\tilde{f}(x) - \tilde{f}(x-\varepsilon z)|^p dz\right)^{\frac{1}{p}}.
\end{aligned}$$

Again, as in the proof of theorem 2.2.24 we use Fubini's theorem and get

$$\begin{aligned}
\|\tilde{f} * \rho_\varepsilon - \tilde{f}\|^p_{L^p(\Omega')} &= \int_{|z|\le 1}\rho(z)\left(\int_{\Omega'}|\tilde{f}(x) - \tilde{f}(x-\varepsilon z)|^p dx\right)dz \\
&\le \int_{|z|\le 1}\rho(z)\sup_{|h|\le\varepsilon}\|\tilde{f} - \tilde{f}(.+h)\|^p_{L^p(\Omega')}dz \;\le\; (\varepsilon')^p
\end{aligned}$$

if $\varepsilon < \delta$. We consider now the mollified families

$$\mathcal{M}_\varepsilon := \{(\tilde{f} * \rho_\varepsilon)|_{\overline{\Omega'}}\,|\, f \in A\},$$

where $\varepsilon < \delta$. We claim that each family $\mathcal{M}_\varepsilon$ satisfies the assumptions of the Ascoli–Arzela theorem. First, we have

$$\|\tilde{f} * \rho_\varepsilon\|_{L^\infty(\mathbf{R}^n)} \leq \|\rho_\varepsilon\|_{L^\infty(\mathbf{R}^n)}\|\tilde{f}\|_{L^1(\mathbf{R}^n)} \leq C_\varepsilon$$

for all $\tilde{f} \in \tilde{A}$ with a constant C_ε depending on ε only. Recall that the set $\tilde{A}$ is bounded in $L^1(\mathbf{R}^n)$. Now if $x_1, x_2 \in \mathbf{R}^n$ and $\tilde{f} \in \tilde{A}$ then we get using the mean value theorem

$$\begin{aligned} |(\tilde{f} * \rho_\varepsilon)(x_1) - (\tilde{f} * \rho_\varepsilon)(x_2)| &= \left| \int_{\mathbf{R}^n} \tilde{f}(y)[\rho_\varepsilon(x_1 - y) - \rho_\varepsilon(x_2 - y)]dy \right| \\ &\leq |x_1 - x_2| \, \|\rho_\varepsilon\|_{C^1(\mathbf{R}^n)} \, \|\tilde{f}\|_{L^1(\mathbf{R}^n)}. \end{aligned}$$

This means that each set $\mathcal{M}_\varepsilon$ is precompact in $C^0(\overline{\Omega'})$ and therefore also in $L^p(\Omega')$. We conclude the proof as follows: Given $\varepsilon' > 0$ we now fix $\varepsilon < \delta$ such that

$$\sup_{\tilde{f} \in \tilde{A}} \|\tilde{f} * \rho_\varepsilon - \tilde{f}\|_{L^p(\Omega')} \leq \varepsilon'.$$

Because the set $\mathcal{M}_\varepsilon$ is precompact in $L^p(\Omega')$ it is also totally bounded. Hence we can cover it with finitely many balls of radius ε' with respect to the $L^p(\Omega')$–norm. Because of the above inequality we can now cover the set A with finitely many balls of radius $2\varepsilon'$. Hence A is also totally bounded in $L^p(\Omega')$ and therefore precompact.

❑

Exercise 2.2.30 *Prove the following version of the Fréchet–Kolmogorov theorem:*

Theorem 2.2.31 *Let $1 \leq p < \infty$. Then a subset $A \subset L^p(\mathbf{R}^n)$ is precompact if the following conditions are satisfied:*

1. $\sup_{f \in A} \|f\|_{L^p(\mathbf{R}^n)} < \infty$,
2. $\sup_{f \in A} \|f(\,\cdot\, + h) - f\|_{L^p(\mathbf{R}^n)} \to 0$ *as* $h \to 0$ *(here $f(\,\cdot\, + h)$ denotes the function $x \mapsto f(x + h)$),*
3. $\sup_{f \in A} \|f\|_{L^p(\mathbf{R}^n \setminus B_R(0))} \longrightarrow 0$ *as* $R \nearrow \infty$.

We close this section with the following definition: Denote by $\mathbf{R}^\mathbf{N}$ the set of all sequences of real numbers. We write $x = (x_k)_{k \in \mathbf{N}} \in \mathbf{R}^\mathbf{N}$ and

$$\|x\|_{l^p} := \left(\sum_k |x_k|^p \right)^{1/p},$$

$$\|x\|_{l^\infty} := \sup_{k \in \mathbf{N}} |x_k|.$$

Definition 2.2.32

$$l^p := \{x \in \mathbf{R^N} \mid \|x\|_{l^p} < \infty\},$$

$$l^\infty := \{x \in \mathbf{R^N} \mid \|x\|_{l^\infty} < \infty\}.$$

The spaces $(l^p, \|.\|_{l^p})$ and $(l^\infty, \|.\|_{l^\infty})$ are Banach spaces. In fact, the spaces L^p can be defined on an arbitrary measure space instead of $\Omega \subset \mathbf{R}^n$ equipped with the Lebesgue measure. In this context the spaces l^p, l^∞ are then the spaces L^p, L^∞, where the underlying measure space are the natural numbers with the discrete measure.

2.2.3 Sobolev spaces

In this section we will introduce a very important class of Banach spaces, the so–called Sobolev spaces. These consist of L^p–functions which have 'weak derivatives'. They are extensively used in the theory of partial differential equations (we will see some of their applications later on).

Definition 2.2.33 *1. Let $\Omega \subset \mathbf{R}^n$ be a domain and $f \in L^1_{loc}(\Omega)$. We say that $v^\alpha \in L^1_{loc}(\Omega)$ is a weak derivative of f of order α if*

$$\int_\Omega f(x)\, D^\alpha \phi(x)\, dx = (-1)^{|\alpha|} \int_\Omega v^\alpha(x)\, \phi(x)\, dx$$

for all $\phi \in C_0^\infty(\Omega)$.

2. *Assume that $1 \le p \le \infty$ and $k \in \mathbf{N}$. We then define the Sobolev–space $W^{k,p}(\Omega)$ to be the set of all $f \in L^p(\Omega)$ which have weak derivatives up to order k so that all weak derivatives are contained in $L^p(\Omega)$.*

We note that weak derivatives are unique. In fact if v^α and w^α were weak derivatives of order α of the same function f then

$$\int_\Omega (v^\alpha(x) - w^\alpha(x))\phi(x)dx = 0$$

for all $\phi \in C_0^\infty(\Omega)$, but then $v^\alpha \equiv w^\alpha$ almost everywhere by proposition 2.2.26. We equip the Sobolev space $W^{k,p}(\Omega)$ with the following norm:

$$\|f\|_{k,p,\Omega} := \sum_{0 \le |\alpha| \le k} \|D^\alpha f\|_{L^p(\Omega)},$$

where $D^\alpha f$ denotes the weak derivative of f of order α. If $f \in W^{k,p}(\Omega) \cap C^k(\Omega)$ then we have by partial integration

$$\int_\Omega f(x)\, D^\alpha \phi(x)\, dx = (-1)^{|\alpha|} \int_\Omega D^\alpha f(x)\, \phi(x)\, dx$$

for all $\phi \in C_0^\infty(\Omega)$, hence the 'strong' derivative in the usual sense coincides almost everywhere with the weak derivative. This justifies the name 'weak derivative'.
Example:
Let $\Omega = (-1, +1)$. The function $f(x) := |x|$ has a weak derivative, namely

$$g(x) := \begin{cases} -1 & \text{if} \quad x \leq 0 \\ +1 & \text{if} \quad x > 0 \end{cases} .$$

Indeed, if ϕ is smooth with compact support in $(-1, 1)$ then

$$\begin{aligned} \int_{-1}^{1} \phi'(x)|x|dx &= -\int_{-1}^{0} x\phi'(x)dx + \int_{0}^{1} x\phi'(x)dx \\ &= -x\phi(x)|_{-1}^{0} + \int_{-1}^{0} \phi(x)dx + x\phi(x)|_{0}^{1} - \int_{0}^{1} \phi(x)dx \\ &= -\int_{-1}^{1} g(x)\phi(x)dx. \end{aligned}$$

On the other hand, the function g has no weak derivative. If it had then its weak derivative h would have to satisfy

$$\int_{-1}^{1} g(x)\phi'(x)dx = -2\phi(0) = -\int_{-1}^{1} h(x)\phi(x)dx$$

for all $\phi \in C_0^\infty((-1, 1))$. In particular, we have for all $\psi \in C_0^\infty((-1, 0))$

$$0 = \int_{-1}^{0} h(x)\psi(x)dx,$$

which implies $h \equiv 0$ almost everywhere on $(-1, 0)$. Similarly we conclude that $h \equiv 0$ almost everywhere on $(0, 1)$. We then obtain

$$\phi(0) = 0 \ \forall \ \phi \in C_0^\infty((-1, 1)),$$

a contradiction.

Proposition 2.2.34 *Let $1 \leq p \leq \infty$ and $k \in \mathbf{N}$. Then $(W^{k,p}(\Omega), \|.\|_{k,p,\Omega})$ is a Banach space*

Proof:
If $(f_k)_{k\in\mathbf{N}}$ is a Cauchy sequence with respect to the $W^{k,p}$–norm then all the sequences $(D^\alpha f_k)_{k\in\mathbf{N}}$ are Cauchy sequences in $L^p(\Omega)$ for all $0 \leq |\alpha| \leq k$. Therefore there are $g^\alpha \in L^p(\Omega)$ so that

$$D^\alpha f_k \stackrel{L^p(\Omega)}{\longrightarrow} g^\alpha$$

as $k \to \infty$. In particular, the sequence f_k itself converges in L^p to some function $g := g^0$. It remains to show that g has derivatives up to order k and that they are given by the functions g^α, i.e. $D^\alpha g = g^\alpha$. We carry out the argument for one derivative of first order. Higher derivatives are then done by iteration. We know that

$$\int_\Omega f_k(x)\frac{\partial \phi}{\partial x_1}(x)dx = -\int_\Omega \frac{\partial f_k}{\partial x_1}(x)\phi(x)dx$$

for all k and for all $\phi \in C_0^\infty(\Omega)$ by the definition of weak derivative. Note that $\partial f_k/\partial x_1$ denotes here a weak derivative of f. Let Ω_ϕ be the support of ϕ. Then

$$\int_\Omega \frac{\partial f_k}{\partial x_1}(x)\phi(x)dx = \int_{\Omega_\phi} \frac{\partial f_k}{\partial x_1}(x)\phi(x)dx$$

and we can justify

$$\lim_{k\to\infty}\int_{\Omega_\phi} \frac{\partial f_k}{\partial x_1}(x)\phi(x)dx = \int_{\Omega_\phi} \lim_{k\to\infty}\frac{\partial f_k}{\partial x_1}(x)\phi(x)dx$$

as follows: The sequence $(\partial f_k/\partial x_1)_{k\in\mathbf{N}}$ converges in $L^p(\Omega)$ to some $g^1 \in L^p(\Omega)$. In particular, we have also convergence in $L^p(\Omega_\phi)$. Since Ω_ϕ is bounded the convergence is also in $L^1(\Omega_\phi)$. Since ϕ is smooth with compact support we also have

$$\frac{\partial f_k}{\partial x_1}\phi \longrightarrow g^1\phi$$

in $L^1(\Omega_\phi)$, which implies

$$\lim_{k\to\infty}\int_{\Omega_\phi} \frac{\partial f_k}{\partial x_1}(x)\phi(x)dx = \int_\Omega \lim_{k\to\infty}\frac{\partial f_k}{\partial x_1}(x)\phi(x)dx = \int_\Omega g^1(x)\phi(x)dx.$$

The same argument yields

$$\lim_{k\to\infty}\int_\Omega f_k(x)\frac{\partial \phi}{\partial x_1}(x)dx = \int_\Omega g(x)\frac{\partial \phi}{\partial x_1}(x)dx$$

so that

$$\int_\Omega g(x)\frac{\partial \phi}{\partial x_1}(x)dx = -\int_\Omega g^1(x)\phi(x)dx$$

for all $C_0^\infty(\Omega)$. Hence g^1 is the weak derivative of g and both g, g^1 are in $L^p(\Omega)$.

❑

We consider now the normed space

$$(C^\infty(\Omega) \cap W^{k,p}(\Omega), \|.\|_{k,p,\Omega})$$

which is a linear subspace of $W^{k,p}(\Omega)$. We define now

Definition 2.2.35

$$H^{k,p}(\Omega) := \overline{C^\infty(\Omega) \cap W^{k,p}(\Omega)}$$

$$H_0^{k,p}(\Omega) := \overline{C_0^\infty(\Omega) \cap W^{k,p}(\Omega)},$$

where $\overline{X}$ *denotes the closure of* X *with respect to the norm* $\|.\|_{k,p,\Omega}$.

The following theorem states that we can identify $W^{k,p}(\Omega)$ with $H^{k,p}(\Omega)$.

Theorem 2.2.36 (Meyers–Serrin)
Let $\Omega \subset \mathbf{R}^n$ *be a domain. Then* $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ *is dense in* $W^{k,p}(\Omega)$*, i.e. for every* $u \in W^{k,p}(\Omega)$ *there is a sequence* $(u_j)_{j\in\mathbf{N}} \subset C^\infty(\Omega) \cap W^{k,p}(\Omega)$ *so that* $\|u_j - u\|_{k,p,\Omega} \to 0$ *as* $j \to \infty$.

We first prove a weaker version of the above theorem:

Lemma 2.2.37 *Let* $\Omega \subset \mathbf{R}^n$ *be a domain and* $u \in W^{k,p}(\Omega)$*. Then for every subdomain* $\Omega' \subset\subset \Omega$ *there is a sequence* $(u_j)_{j\in\mathbf{N}} \subset C^\infty(\Omega) \cap W^{k,p}(\Omega)$ *so that*

$$\|u_j - u\|_{k,p,\Omega'} \longrightarrow 0$$

as $j \to \infty$.

Proof:
Let $\varepsilon < \mathrm{dist}(\Omega', \partial\Omega)$. If we extend the function u trivially onto all of $\mathbf{R}^n$ then the extended function may not have weak derivatives, i.e. the extension of u onto in $\mathbf{R}^n$ is of class L^p, but not of class $W^{k,p}$. Whenever we use weak derivatives of u we have to be careful that u is evaluated on Ω only. Let $u_\varepsilon = u * \rho_\varepsilon$ be the mollifier of u, and denote the weak derivative of u of order α by $D^\alpha u$. We have for $x \in \Omega'$

$$\begin{aligned}
D^\alpha(u * \rho_\varepsilon)(x) &= \frac{1}{\varepsilon^n}\int_\Omega D_x^\alpha \rho\left(\frac{x-y}{\varepsilon}\right) u(y)dy \\
&= \frac{(-1)^{|\alpha|}}{\varepsilon^n}\int_\Omega \left[D_y^\alpha \rho\left(\frac{x-y}{\varepsilon}\right)\right] u(y)dy \\
&= \frac{1}{\varepsilon^n}\int_\Omega \rho\left(\frac{x-y}{\varepsilon}\right) D^\alpha u(y)dy \\
&= (D^\alpha u * \rho_\varepsilon)(x),
\end{aligned}$$

i.e. the operations 'mollifying' and 'differentiating' commute, or shortly

$$D^\alpha u_\varepsilon = (D^\alpha u)_\varepsilon.$$

(note that $D^\alpha u_\varepsilon$ makes sense on all of $\mathbf{R}^n$ while $(D^\alpha u)_\varepsilon$ is only defined on Ω'). We have shown earlier that $u_\varepsilon \to u$ in $L^p(\Omega')$ as $\varepsilon \to 0$. By the above equation we also have $D^\alpha u_\varepsilon \to D^\alpha u$ in $L^p(\Omega')$. This just means that

$$\|u_\varepsilon - u\|_{k,p,\Omega'} \longrightarrow 0.$$

If $\varepsilon_j \searrow 0$, then we take $u_j = u_{\varepsilon_j}$. Smoothness of u_j is clear. We have $D^\alpha u_j \in L^p(\Omega)$ because of

$$\|D^\alpha u_j\|_{L^p(\Omega)} \leq \|D^\alpha (u * \rho_{\varepsilon_j})\|_{L^p(\Omega)} \leq \|D^\alpha \rho_{\varepsilon_j}\|_{L^1(\mathbf{R}^n)} \|u\|_{L^p(\Omega)} < \infty.$$

❑

In order to prove the Meyers–Serrin theorem we have to decompose the open set Ω and the function u in 'smaller' pieces and apply the local lemma 2.2.37 to each of them. There are less fancy versions of the Meyers–Serrin theorem which require a domain with sufficiently regular boundary $\partial\Omega$. They are based on extending u as a $W^{k,p}$–function (!) onto all of $\mathbf{R}^n$ where finally lemma 2.2.37 can be applied. The tool for cutting a function into smaller pieces is called a 'partition of unity'. We just give the definition here.

Definition 2.2.38 *Let $A \subset \mathbf{R}^n$ be a subset and let $(U_j)_{j\in\mathbf{N}}$ be an open covering, i.e. each set $U_j \subset \mathbf{R}^n$ is open and their union contains A. An open covering is called locally finite if every point $x \in A$ has a neighborhood $B_\varepsilon(x)$ so that the set*

$$\{j \in \mathbf{N} \,|\, U_j \cap \overline{B_\varepsilon(x)} \neq \emptyset\}$$

is finite.

Simple example: The intervals $(-1/n, 1/n)_{n\in\mathbf{N}}$ are an open covering of $A = [-1/2, 1/2]$, but any neighborhood of 0 hits infinitely many of these intervals, so the covering is not locally finite. If we rather put $A = [1/4, 1/2]$ then the same intervals are a locally finite open covering.

Definition 2.2.39 *Let $(U_j)_{j\in\mathbf{N}}$ be a locally finite open covering of a set $A \subset \mathbf{R}^n$ so that all the U_j are bounded sets. A partition of unity associated to the locally finite open covering $(U_j)_{j\in\mathbf{N}}$ is a family of smooth functions $(\eta_j)_{j\in\mathbf{N}}$ with the following properties:*

1. *The function η_j has compact support in U_j,*
2. *$\eta_j \geq 0$,*
3. *$\sum_j \eta_j(x) = 1$ for all $x \in A$.*

(Note that the sum in 3. is finite !)

We give the following statement without proof. The result is usually proved in topology books for continuous partitions of unity (see J. Munkres, Topology-A first course, chapter 4-5). Their proofs can easily be modified to the smooth version. Another option are books about differentiable manifolds (M. Spivak, Calculus on manifolds, p.63).

Proposition 2.2.40 *Let $(U_j)_{j\in\mathbf{N}}$ be a locally finite open covering of a set $A \subset \mathbf{R}^n$ so that all the U_j are bounded sets. Then there is a partition of unity.*

We can now prove the Meyers–Serrin theorem:

Proof:
Let $(U_j)_{j\in\mathbf{N}}$ be a locally finite covering of Ω so that each set U_j is bounded and $\overline{U_j} \subset \Omega$ so that $h_j := \operatorname{dist}(\partial\Omega, \overline{U_j}) > 0$. An example for such a covering is the following:

$$\tilde{U}_j := \{x \in \Omega \,|\, \frac{1}{2} \cdot 2^{-j} < \operatorname{dist}(x, \partial\Omega) < 2 \cdot 2^{-j}\Big\} \text{ with } j \in \mathbf{Z}.$$

The above sets are a locally finite covering of Ω. If we take

$$U_j := [\tilde{U}_j \cap B_{2\cdot 2^j}(0)] \cup \left(\bigcup_{i<j} \tilde{U}_i \cap B_{2\cdot 2^j}(0) \backslash B_{\frac{1}{2}\cdot 2^j}(0) \right)$$

then the sets U_j are also bounded. Let $(\eta_j)_{j\in\mathbf{N}}$ be a partition of unity and let

$$0 < c_j \leq \frac{1}{2^{j+1}\|\eta_j\|_{C^k(\overline{\Omega})}}.$$

If $\varepsilon > 0$ then we can find by lemma 2.2.37 functions $u_{j,\varepsilon} \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that

$$\|u - u_{j,\varepsilon}\|_{k,p,U_j} \leq \varepsilon c_j.$$

We define

$$u_\varepsilon := \sum_{j\in\mathbf{N}} \eta_j u_{j,\varepsilon}$$

so that

$$u_\varepsilon - u = \sum_{j\in\mathbf{N}} \eta_j (u_{j,\varepsilon} - u)$$

and all the sums above are finite sums. If $\phi \in C_0^\infty(\Omega)$ then

$$\int_\Omega \eta_j u \partial_i \phi = \int_\Omega u(\partial_i(\eta_j\phi) - \phi\partial_i\eta_j) = -\int_\Omega \phi(\eta_j\partial_i u + u\partial_i\eta_j),$$

therefore $\eta_j u \in W^{1,p}(\Omega)$ with weak derivative given by the product rule. We can deal with higher derivatives by induction, and we conclude that $\eta_j u \in W^{k,p}(\Omega)$ with

$$D^\alpha(\eta_j u) = \sum_{0\leq|\gamma|\leq|\alpha|} c_{\alpha,\gamma}[D^{\alpha-\gamma}\eta_j]\, D^\gamma u, \quad |\alpha| \leq k,$$

where $c_{\alpha,\gamma} > 0$ are suitable constants, and

$$D^\alpha u_\varepsilon - D^\alpha u = \sum_{0\leq|\gamma|\leq|\alpha|} c_{\alpha,\gamma} \sum_{j\in\mathbf{N}} [D^{\alpha-\gamma}\eta_j]\,(D^\gamma u_{j,\varepsilon} - D^\gamma u).$$

Hence there is a constant C depending only on k and the dimension n of the domain so that

$$\begin{aligned}\|D^\alpha u_\varepsilon - D^\alpha u\|_{L^p(\Omega)} &\leq C \sum_{j\in\mathbf{N}} \|\eta_j\|_{C^k(\Omega)} \|u_{j,\varepsilon} - u\|_{k,p,U_j} \\ &\leq C\varepsilon \sum_{j\in\mathbf{N}} c_j \|\eta_j\|_{C^k(\Omega)} \\ &\leq C\varepsilon.\end{aligned}$$

❑

The Approximation property by smooth functions is very useful. Instead of proving statements about Sobolev functions directly one first does it for smooth Sobolev functions since strong derivatives are easier to handle than weak derivatives. Then the Meyers–Serrin theorem often implies that the desired properties are also true for non–smooth Sobolev functions. Use approximation to do the following exercises:

Exercise 2.2.41 (Product and Chain rule for Sobolev–functions)

1. *Let $\Omega \subset \mathbf{R}^n$ be a domain and $1 \leq p, q \leq \infty$ so that $1/p + 1/q = 1$. If $f \in W^{k,p}(\Omega)$ and $g \in W^{k,q}(\Omega)$ then the product fg is in $W^{k,1}(\Omega)$ and the weak derivatives of fg are given by the product rule.*
2. *Let $\Omega, \tilde{\Omega} \subset \mathbf{R}^n$ be domains and let $\tau : \tilde{\Omega} \to \Omega$ be a C^1–diffeomorphism, i.e. τ is bijective with τ, τ^{-1} continuously differentiable so that the derivatives of τ and τ^{-1} are bounded. If $f \in W^{k,p}(\Omega)$ then also $f \circ \tau \in W^{k,p}(\tilde{\Omega})$ and the weak derivatives of $f \circ \tau$ are given by the Chain rule.*

In contrast to the Meyers–Serrin theorem we only have $H_0^{k,p}(\Omega) \subset W^{k,p}(\Omega)$ with strict inclusion unless $\Omega = \mathbf{R}^n$. Let $I = (0,1)$. The following example demonstrates that the spaces $H^{1,1}(I)$ and $H_0^{1,1}(I)$ are indeed different. Consider a smooth function ϕ with compact support in I. We have

$$\phi(x) = \int_0^x \phi'(t)dt$$

and therefore

$$|\phi(x)| \leq \int_0^1 |\phi'(t)|dt$$

so that

$$\int_I |\phi| \leq \int_I |\phi'|.$$

Denote by 1 the constant function 1. We have

$$1 \leq |\phi(x)| + |1 - \phi(x)|.$$

Integrating over the interval I we obtain

$$\begin{aligned} 1 &\leq \int_I |\phi(x)|dx + \int_I |1-\phi(x)|dx \\ &\leq \int_I |\phi'(x)|dx + \int_I |1-\phi(x)|dx \\ &\leq \int_I |(1-\phi(x))'|dx + \int_I |1-\phi(x)|dx \\ &= \|1-\phi\|_{1,1,I} \end{aligned}$$

i.e. smooth functions with compact support cannot approximate the constant function 1 with respect to the $H^{1,1}(I)$–norm.

Exercise 2.2.42

$$H_0^{k,p}(\mathbf{R}^n) = W^{k,p}(\mathbf{R}^n)$$

Sketch of proof: Because of Meyer–Serrin it is sufficient to approximate $u \in C^\infty(\mathbf{R}^n) \cap W^{k,p}(\mathbf{R}^n)$ by a smooth function with compact support. Take $\eta \in C_0^\infty(\mathbf{R}^n)$ with the following property $\eta(x) \equiv 1$ if $|x| \leq 1$ and $\eta(x) \equiv 0$ if $|x| \geq 2$. Consider now the sequence of smooth functions

$$u_j(x) := \eta\left(\frac{x}{j}\right) u(x) \in C_0^\infty(\mathbf{R}^n)$$

and show that $\|u - u_j\|_{k,p,\mathbf{R}^n} \to 0$ as $j \to \infty$.

An element $u \in W^{1,p}(\Omega)$ is an equivalence class of L^p–functions with some other properties. We would like to study the question whether we can do better than that. For example, is it possible to choose a representative in the class of u which is continuous, bounded or even differentiable. We are lead to the Sobolev–embedding theorem, the most important result about Sobolev spaces. We start with the following lemma:

Lemma 2.2.43 (Morrey)
Let $I_l \subset \mathbf{R}^n$ be a cube whose edges are parallel to the coordinate axes and have length l. Assume that $p > n$. Then

$$\left| \frac{1}{|I_l|} \int_{I_l} u(x)dx \; - \; u(x) \right| \leq \frac{l^{1-\frac{n}{p}}}{1-\frac{n}{p}} \sum_{j=1}^n \|\partial_j u\|_{L^p(I_l)}$$

for all $l > 0$, $x \in I_l$ and $u \in C^\infty(\mathbf{R}^n)$.

Proof:
Because of translation invariance it suffices to consider the case $x = 0$. We have by definition

$$\frac{1}{|I_l|} \int_{I_l} u(x)dx - u(0) = \frac{1}{|I_l|} \int_{I_l} (u(x) - u(0))dx.$$

If $x \in I_l$ then

$$\begin{aligned}
|u(x) - u(0)| &= \left| \int_0^1 \frac{d}{dt} u(tx) dt \right| \\
&= \left| \int_0^1 \sum_{j=1}^n \partial_j u(tx) x_j dt \right| \\
&\leq l \int_0^1 \sum_{j=1}^n |\partial_j u(tx)| dt,
\end{aligned}$$

where we have used that $|x_j| \leq l$. Combining the two formulas yields

$$\begin{aligned}
\left| \frac{1}{|I_l|} \int_{I_l} u(x) dx - u(0) \right| &\leq \frac{l}{|I_l|} \int_0^1 \left(\sum_{j=1}^n \int_{I_l} |\partial_j u(tx)| dx \right) dt \\
&= \frac{l}{|I_l|} \int_0^1 \frac{1}{t^n} \left(\sum_{j=1}^n \int_{t\, I_l} |\partial_j u(y)| dy \right) dt.
\end{aligned}$$

We now estimate the above integral over y with Hölder's inequality. Let χ_{tI_l} be the characteristic function of the cube tI_l and q such that $1/p + 1/q = 1$.

$$\begin{aligned}
\int_{t\, I_l} |\partial_j u(y)| dy &= \int_{t\, I_l} \chi_{tI_l}(y) |\partial_j u(y)| dy \\
&\leq \left(\int_{tI_l} \chi_{tI_l}(y) dy \right)^{1/q} \left(\int_{tI_l} |\partial_j u(y)|^p dy \right)^{1/p} \\
&\leq |t\, I_l|^{1/q} \left(\int_{I_l} |\partial_j u(y)|^p dy \right)^{1/p}
\end{aligned}$$

because $0 \leq t \leq 1$. Recalling that $|tI_l| = t^n\, l^n$ we obtain

$$\begin{aligned}
\left| \frac{1}{|I_l|} \int_{I_l} u(x) dx - u(0) \right| &\leq \frac{l}{|I_l|} \int_0^1 \frac{1}{t^n} t^{n/q} |I_l|^{1/q} \sum_{j=1}^n \|\partial_j u\|_{L^p(I_l)} \\
&\leq l^{1-\frac{n}{p}} \left(\int_0^1 t^{-n(1-1/q)} dt \right) \sum_{j=1}^n \|\partial_j u\|_{L^p(I_l)}
\end{aligned}$$

The integral over t is finite because of $p > n$, and it equals $(1 - n/p)^{-1}$. We get finally

$$\left| \frac{1}{|I_l|} \int_{I_l} u(x) dx - u(0) \right| \leq \frac{l^{1-\frac{n}{p}}}{1 - \frac{n}{p}} \sum_{j=1}^n \|\partial_j u\|_{L^p(I_l)}.$$

❑

We will prove now a special case of the Sobolev–embedding theorem:

Theorem 2.2.44 *Let $p > n$ and $\beta = 1 - \frac{n}{p}$. Then there is a continuous embedding*

$$W^{1,p}(\mathbf{R}^n) \hookrightarrow C^{0,\beta}(\mathbf{R}^n),$$

i.e. there is a constant $M > 0$ so that every element $u \in W^{1,p}(\mathbf{R}^n)$ has a representative in $C^{0,\beta}(\mathbf{R}^n)$ such that

$$\|u\|_{C^{0,\beta}(\mathbf{R}^n)} \leq M \|u\|_{1,p,\mathbf{R}^n}.$$

Proof:
We assume first that $u \in C^\infty(\mathbf{R}^n) \cap W^{1,p}(\mathbf{R}^n)$. Let $x, y \in \mathbf{R}^n$ and let $I_l \subset \mathbf{R}^n$ ba a cube as in Morrey's lemma which contains both x and y. We may choose $l = 2|x - y|$. Morrey's lemma implies

$$\begin{aligned} |u(x) - u(y)| &\leq \left| u(x) - \frac{1}{|I_l|} \int_{I_l} u(x)dx \right| + \left| \frac{1}{|I_l|} \int_{I_l} u(x)dx - u(y) \right| \\ &\leq \frac{2l^{1-\frac{n}{p}}}{1 - \frac{n}{p}} \sum_{j=1}^{n} \|\partial_j u\|_{L^p(I_l)} \\ &= \frac{2^{\beta+1}|x-y|^\beta}{\beta} \sum_{j=1}^{n} \|\partial_j u\|_{L^p(I_l)}, \end{aligned}$$

so that

$$\frac{|u(x) - u(y)|}{|x - y|^\beta} \leq c \|u\|_{1,p,\mathbf{R}^n}, \tag{2.3}$$

where $c = 2^{\beta+1}/\beta$. Now let $l = 1$ and $x \in I_1$ so that $|I_1| = 1$. Then

$$|u(x)| \leq \left| \int_{I_l} u(x)dx - u(x) \right| + \left| \int_{I_l} u(x)dx \right|.$$

If χ is the characteristic function of the cube I_1 then we estimate with Hölder's inequality

$$\left| \int_{I_l} u(x)dx \right| \leq \int_{I_1} \chi(x)|u(x)|dx \leq \|u\|_{L^p(\mathbf{R}^n)}.$$

Morrey's lemma yields

$$\left| \int_{I_l} u(x)dx - u(x) \right| \leq \frac{1}{1 - \frac{n}{p}} \sum_{j=1}^{n} \|\partial_j u\|_{L^p(I_l)},$$

so that

$$|u(x)| \leq c \|u\|_{1,p,\mathbf{R}^n} \tag{2.4}$$

with a positive constant c depending on β only. Consider now the general case where $u \in W^{1,p}(\mathbf{R}^n)$. Using the Meyers–Serrin theorem we can find a sequence u_j in $C^\infty(\mathbf{R}^n) \cap W^{1,p}(\mathbf{R}^n)$ so that

$$\|u - u_j\|_{1,p,\mathbf{R}^n} \longrightarrow 0 \text{ as } j \to \infty.$$

Inequalities (2.3) and (2.4) imply that the sequence (u_j) is a Cauchy sequence in the Hölder space $C^{0,\beta}(\mathbf{R}^n)$, hence $u_j \to \tilde{u}$ in $C^{0,\beta}(\mathbf{R}^n)$, in particular, the convergence is uniform and also pointwise. We therefore have

$$\frac{|\tilde{u}(x) - \tilde{u}(y)|}{|x-y|^\beta} \leq c\,\|u\|_{1,p,\mathbf{R}^n}$$

and

$$|\tilde{u}(x)| \leq c\,\|u\|_{1,p,\mathbf{R}^n}.$$

Because the sequence u_j converges to u in $L^p(\mathbf{R}^n)$, there is a subsequence (u_{j_k}) which converges pointwise almost everywhere, i.e.

$$\lim_{k\to\infty} u_{j_k}(x) = u(x)$$

for almost all x. On the other hand,

$$\lim_{k\to\infty} u_{j_k}(x) = \tilde{u}(x)$$

for all x, so that $\tilde{u} \equiv u$ almost everywhere. Then $\tilde{u}$ is the desired representative.

❑

If $\Omega \subset \mathbf{R}^n$ is an open set, $m \geq 1$, $p > n$ and $\beta = 1 - \frac{n}{p}$ then we can prove exactly in the same way that $H_0^{m,p}(\Omega)$ is continuously embedded into $C^{m-1,\beta}(\overline{\Omega})$. We simply apply Morrey's lemma to $u \in C_0^\infty(\Omega)$ and all its derivatives up to order $m-1$. We then use that

$$\overline{C_0^\infty(\Omega)} = H_0^{m,p}(\Omega).$$

We conclude

$$\bigcap_{m\geq 1} H_0^{m,p}(\Omega) = C^\infty(\overline{\Omega}).$$

Hence we may replace in theorem 2.2.44 the domain $\mathbf{R}^n$ by Ω if (!) we substitute $W^{1,p}(\mathbf{R}^n)$ with $H_0^{1,p}(\Omega)$.

Remark: The limit case $p = n$
Theorem 2.2.44 is false if $p \leq n$. As an example, consider the function

$$u(x) := \begin{cases} \left(\log \frac{1}{|x|}\right)^\alpha & \text{if} \quad |x| \leq 1 \\ 0 & \text{if} \quad |x| > 1 \end{cases},$$

with $0 < \alpha < 1 - 1/n$. The function u is obviously not continuous in the origin, but it is in $W^{1,n}(\mathbf{R}^n)$.

Exercise 2.2.45 *Show that $u \in W^{1,n}(\mathbf{R}^n)$.*

The generalized version of theorem 2.2.44 is the following:

Theorem 2.2.46 (Sobolev-embedding-theorem I) *Let $\Omega \subset \mathbf{R}^n$ be a domain, $1 \leq m \in \mathbf{N}$ and $p \geq 1$. We assume that*

$$m - \frac{n}{p} = k + \beta$$

for some integer $k \geq 0$ and $0 < \beta < 1$. Then

$$H_0^{m,p}(\Omega) \hookrightarrow C^{k,\beta}(\overline{\Omega})$$

is embedded continuously, i.e. there is a constant $M > 0$ such that for all $u \in H_0^{m,p}(\Omega)$ there is a representative in $C^{k,\beta}(\overline{\Omega})$ which satisfies

$$\|u\|_{C^{k,\beta}(\overline{\Omega})} \leq M \, \|u\|_{m,p,\Omega}.$$

Note that this is not a straightforward consequence of theorem 2.2.44 because we may have $m - n/p = k + \beta$ also if $p \leq n$ where theorem 2.2.44 does not apply. We need a version of the embedding theorem which also works for $p \leq n$. Of course, we cannot expect an embedding into Hölder space but there is an embedding $H_0^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for a suitable $q > p$. So we trade one derivative for a better 'p'. The hope is then that $q > n$ so that theorem 2.2.44 is applicable again. The following is another special case of the Sobolev embedding theorem for $W^{1,p}(\mathbf{R}^n)$ and $p < n$.

Theorem 2.2.47 *Let $1 \leq p, q < \infty$ with*

$$\frac{n}{q} = \frac{n}{p} - 1.$$

Then we have for all $u \in W^{1,p}(\mathbf{R}^n)$

$$\|u\|_{L^q(\mathbf{R}^n)} \leq q \cdot \frac{n-1}{n} \|\nabla u\|_{L^p(\mathbf{R}^n)},$$

where $\|\nabla u\|_{L^p(\mathbf{R}^n)} := max_{1 \leq k \leq n} \|\partial_k u\|_{L^p(\mathbf{R}^n)}$.

Remark: The assumptions in the theorem imply that $p < n$ and that $n \geq 2$. In the one–dimensional case we have

$$\|u\|_{L^\infty(\mathbf{R})} \leq \|u'\|_{L^1(\mathbf{R})} \ \forall \ u \in W^{1,1}(\mathbf{R}).$$

Show this as an exercise.

Proof:
Because of $W^{1,p}(\mathbf{R}^n) = H_0^{1,p}(\mathbf{R}^n)$ it suffices to show the estimate for smooth functions on $\mathbf{R}^n$ with compact support. Indeed, if $u \in W^{1,p}(\mathbf{R}^n)$ is arbitrary we pick a sequence of smooth functions with compact support such that $u_j \to u$ in the $W^{1,p}(\mathbf{R}^n)$–norm. The estimate for smooth functions with compact support then implies that (u_j) is a Cauchy sequence in $L^q(\mathbf{R}^n)$ as well, hence (u_j) also

converges in $L^q(\mathbf{R}^n)$ to some limit which agrees almost everywhere with u. The inequality follows then easily for u.
So let u be a smooth function on $\mathbf{R}^n$ with compact support. Consider first the case where $p = 1$, i.e. $q = \frac{n}{n-1}$ (note that $n \geq 2$ by assumption). Then

$$\begin{aligned} |u(x)| &= \left| \int_{x_i}^{\infty} \partial_i u(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n) d\xi \right| \\ &\leq \int_{\mathbf{R}} |\partial_i u(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n)| d\xi, \end{aligned}$$

which we write shortly as

$$|u(x)| \leq \int_{\mathbf{R}} |\partial_i u| d\xi_i.$$

This actually proves our exercise above. For $i = 1, \dots, n$ these are n inequalities which we all multiply with each other. Hence

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^{n} \left(\int_{\mathbf{R}} |\partial_i u| d\xi_i \right)^{\frac{1}{n-1}}.$$

We integrate over the variable x_1 and obtain

$$\int_{\mathbf{R}} |u|^{\frac{n}{n-1}} d\xi_1 \leq \left(\int_{\mathbf{R}} |\partial_1 u| d\xi_1 \right)^{\frac{1}{n-1}} \cdot \int_{\mathbf{R}} \prod_{i=2}^{n} \left(\int_{\mathbf{R}} |\partial_i u| d\xi_i \right)^{\frac{1}{n-1}} d\xi_1.$$

We use Hölder's inequality in the form

$$\int |f_2 \cdots f_n| \leq \prod_{i=2}^{n} \|f_i\|_{L^{p_i}} \text{ with } \sum_{i=2}^{n} 1/p_i = 1$$

where

$$f_i := \left(\int_{\mathbf{R}} |\partial_i u| d\xi_i \right)^{\frac{1}{n-1}} \text{ and } p_i = n - 1.$$

This implies

$$\int_{\mathbf{R}} \prod_{i=2}^{n} \left(\int_{\mathbf{R}} |\partial_i u| d\xi_i \right)^{\frac{1}{n-1}} d\xi_1 \leq \prod_{i=2}^{n} \left(\int_{\mathbf{R}^2} |\partial_i u| d\xi_1 \, d\xi_i \right)^{\frac{1}{n-1}}$$

so that

$$\int_{\mathbf{R}} |u|^{\frac{n}{n-1}} d\xi_1 \leq \left(\int_{\mathbf{R}} |\partial_1 u| d\xi_1 \right)^{\frac{1}{n-1}} \prod_{i=2}^{n} \left(\int_{\mathbf{R}^2} |\partial_i u| d\xi_1 \, d\xi_i \right)^{\frac{1}{n-1}}.$$

We integrate this inequality now with respect to x_2. If $n = 2$ then we obtain the desired estimate. If not then we have to use Hölder's inequality again:

$$\begin{aligned}
\int_{\mathbf{R}^2} |u|^{\frac{n}{n-1}} d\xi_1 d\xi_2 &\leq \int_{\mathbf{R}} \left(\int_{\mathbf{R}} |\partial_1 u| d\xi_1 \right)^{\frac{1}{n-1}} \cdot \prod_{i=3}^{n} \left(\int_{\mathbf{R}^2} |\partial_i u| d\xi_1\, d\xi_i \right)^{\frac{1}{n-1}} d\xi_2 \cdot \\
&\quad \cdot \left(\int_{\mathbf{R}} |\partial_2 u| d\xi_1\, d\xi_2 \right)^{\frac{1}{n-1}} \\
&\leq \left(\int_{\mathbf{R}} |\partial_1 u| d\xi_1\, d\xi_2 \right)^{\frac{1}{n-1}} \left(\int_{\mathbf{R}} |\partial_2 u| d\xi_1\, d\xi_2 \right)^{\frac{1}{n-1}} \cdot \\
&\quad \cdot \prod_{i=3}^{n} \left(\int_{\mathbf{R}^3} |\partial_i u| d\xi_1\, d\xi_2\, d\xi_i \right)^{\frac{1}{n-1}}.
\end{aligned}$$

We continue by iteration and obtain

$$\int_{\mathbf{R}^j} |u|^{\frac{n}{n-1}} d\xi_1 \cdots d\xi_j \leq \prod_{i=1}^{j} \left(\int_{\mathbf{R}^j} |\partial_i u| d\xi_1 \cdots d\xi_j \right)^{\frac{1}{n-1}} \prod_{i=j+1}^{n} \left(\int_{\mathbf{R}^{j+1}} |\partial_i u| d\xi_1 \cdots d\xi_j\, d\xi_i \right)^{\frac{1}{n-1}},$$

hence

$$\int_{\mathbf{R}^n} |u|^{\frac{n}{n-1}} \leq \prod_{i=1}^{n} \left(\int_{\mathbf{R}^n} |\partial_i u| \right)^{\frac{1}{n-1}} \leq \left(\int_{\mathbf{R}^n} |\nabla u| \right)^{\frac{n}{n-1}}$$

and therefore

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbf{R}^n)} = \|u\|_{L^q(\mathbf{R}^n)} \leq \|\nabla u\|_{L^1(\mathbf{R}^n)}. \tag{2.5}$$

If $p > 1$ then we would like to apply the above estimate to the function

$$v := |u|^{\frac{q(n-1)}{n}}.$$

The function u is smooth with compact support and

$$\frac{q(n-1)}{n} > p > 1.$$

Then $v \in C_0^1(\mathbf{R}^n)$ with

$$|\nabla v| = \frac{q(n-1)}{n} |u|^{\frac{q(n-1)}{n}-1} |\nabla u|.$$

We conclude now

$$\begin{aligned}
\left(\int_{\mathbf{R}^n} |u|^q \right)^{\frac{n-1}{n}} &= \left(\int_{\mathbf{R}^n} v^{\frac{n}{n-1}} \right)^{\frac{n-1}{n}} \\
&\leq \int_{\mathbf{R}^n} |\nabla v| \text{ used (2.5) here} \\
&= \frac{q(n-1)}{n} \int_{\mathbf{R}^n} |u|^{\frac{q(n-1)}{n}-1} \cdot |\nabla u| \\
&\leq \frac{q(n-1)}{n} \left(\int_{\mathbf{R}^n} |u|^{(\frac{q(n-1)}{n}-1)r} \right)^{1/r} \|\nabla u\|_{L^p(\mathbf{R}^n)}
\end{aligned}$$

using Hölder's inequality with $1/p + 1/r = 1$, i.e. $r = p/(p-1)$. We have by assumption

$$\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$$

and

$$\frac{n-1}{n} - \frac{1}{r} = \frac{n-1}{n} - \frac{p-1}{p} = \frac{1}{q}$$

so that

$$(\frac{q(n-1)}{n} - 1)r = q.$$

Then

$$\left(\int_{\mathbf{R}^n} |u|^q\right)^{\frac{1}{q}} \leq \frac{q(n-1)}{n} \|\nabla u\|_{L^p(\mathbf{R}^n)}.$$

❑

We note the following simple generalization.

Theorem 2.2.48 (Sobolev-embedding-theorem II) *Let $\Omega \subset \mathbf{R}^n$ be a domain. Moreover, let $1 \leq p, q < \infty$ and $m_1 > m_2 \geq 0$ so that*

$$m_1 - \frac{n}{p} = m_2 - \frac{n}{q}$$

(this implies in particular that $q > p$). Then there is a continuous embedding

$$H_0^{m_1,p}(\Omega) \hookrightarrow H_0^{m_2,q}(\Omega),$$

i.e. for every $u \in H_0^{m_1,p}(\Omega)$ we have

$$\|u\|_{m_2,q,\Omega} \leq C \|u\|_{m_1,p,\Omega}$$

with a suitable positive constant C only depending on n, m_1 and p.

Proof:
Sketch only. Fill out the details as an exercise.
Convince yourself that it suffices to consider the case $m_2 = m_1 - 1$ since we can obtain the result for smaller m_2 by iterated application. If $u \in H_0^{m_1,p}(\Omega)$ then it is also in $H^{m_1,p}(\mathbf{R}^n)$ by trivial extension (the subscript '0' is of course the reason why it works). Then all weak derivatives of order up to $m_1 - 1 = m_2$ are in $H^{1,p}(\mathbf{R}^n)$ and then also in $L^q(\mathbf{R}^n)$.

❑

Proof:
(Theorem 2.2.46, Sobolev-embedding-I)

If $m = 1$ then $p > n$ and $k = 0$, but we have already covered this case. So assume that $m > 1$. We note that

$$\frac{n}{p} - (m - k - 1) = 1 - \beta \in (0, 1)$$

hence there is a number $n < q < \infty$ so that

$$\frac{n}{q} = \frac{n}{p} - (m - k - 1) \text{ i.e. } 1 - \frac{n}{q} = m - k - \frac{n}{p}$$

We now invoke theorem 2.2.48 (Sobolev-embedding-II) and we conclude that $H_0^{m-k,p}(\Omega) \subset H_0^{1,q}(\Omega)$. On the other hand, the space $H_0^{1,q}(\Omega)$ embeds into $C^{0,\gamma}(\overline{\Omega})$ with $\gamma = 1 - \frac{n}{q} = m - k - \frac{n}{p} = \beta$ since $q > n$. Now all the weak derivatives $D^\alpha u$ with $|\alpha| \leq k$ are in the space $H_0^{m-k,p}(\Omega)$ and therefore also in $C^{0,\beta}(\overline{\Omega})$. Moreover, we have the estimate

$$\|D^\alpha u\|_{C^{0,\beta}(\overline{\Omega})} \leq C \, \|D^\alpha u\|_{m-k,p,\Omega} \; \forall \; u \in H_0^{m,p}(\Omega) \, , \; |\alpha| \leq k.$$

At this point we do not know whether u is differentiable in the classical sense. We have just shown that all weak derivatives up to order k have Hölder continuous representatives. Assuming now that u is smooth with compact support we may write the above estimate as

$$\|u\|_{C^{k,\beta}(\overline{\Omega})} \leq C \, \|u\|_{m,p,\Omega},$$

If a sequence of smooth functions with compact support u_j now approximates an arbitrary $u \in H_0^{m,p}(\Omega)$, i.e. $\|u - u_j\|_{m,p,\Omega} \to 0$ as $j \to \infty$. Then the sequence u_j is also a Cauchy sequence in $C^{k,\beta}(\overline{\Omega})$ by the estimate above, hence it converges and the limit coincides almost everywhere with u.

❑

Remarks:

1. The Sobolev–embedding theorems 2.2.46 and 2.2.48 remain correct with $H_0^{m,p}(\Omega)$ replaced by the larger spaces $H^{m,p}(\Omega)$ if the domain Ω has Lipschitz–boundary, i.e. if the boundary is locally the graph of a Lipschitz continuous function. If this is the case then any $u \in H^{m,p}(\Omega)$ can be extended to a slightly larger domain $\tilde{\Omega}$ **as a Sobolev–function in** $H_0^{m,p}(\tilde{\Omega})$. Then one uses the theorems 2.2.46 and 2.2.48 that we proved in the lecture. We are not going to prove this extension result here.

2. In theorem 2.2.47 we proved an estimate of the form

$$\|u\|_{L^q(\mathbf{R}^n)} \leq C \, \|\nabla u\|_{L^p(\mathbf{R}^n)},$$

where $1 \leq p, q < \infty$ with
$$\frac{n}{q} = \frac{n}{p} - 1.$$

While the proof was rather tedious it is pretty easy to see why the exponent q has to satisfy the above equation: Assume we want to prove the inequality $\|u\|_{L^q(\mathbf{R}^n)} \leq C\,\|\nabla u\|_{L^p(\mathbf{R}^n)}$ for all $u \in W^{1,p}(\mathbf{R}^n)$ but we are not sure what is the right number q. If the inequality holds for $u \in W^{1,p}(\mathbf{R}^n)$ then it must also hold for $u_\lambda \in W^{1,p}(\mathbf{R}^n)$ where $u_\lambda(x) = u(\lambda x)$, $\lambda > 0$. Now

$$\|u_\lambda\|_{L^q} = \left(\int_{\mathbf{R}^n} |u(\lambda x)|^q\,dx\right)^{1/q} = \lambda^{-n/q}\left(\int_{\mathbf{R}^n} |u(y)|^q\,dy\right)^{1/q}$$

and

$$\|\nabla u_\lambda\|_{L^p} = \left(\int_{\mathbf{R}^n} |\lambda\,\nabla u(\lambda x)|^p\,dx\right)^{1/p} = \lambda^{1-n/p}\left(\int_{\mathbf{R}^n} |\nabla u(y)|^p\,dy\right)^{1/p}$$

so that

$$\|u\|_{L^q} \leq C\,\lambda^{1-\frac{n}{p}+\frac{n}{q}}\,\|\nabla u\|_{L^p}.$$

The expression $\lambda^{1-\frac{n}{p}+\frac{n}{q}}$ can only equal 1 for all $\lambda > 0$ if the exponent $1 - \frac{n}{p} + \frac{n}{q}$ equals 0.

3. In the Sobolev embedding theorems the conclusion holds also if we have

$$m - \frac{n}{p} \geq k + \beta \text{ and } m_1 - \frac{n}{p} \geq m_2 - \frac{n}{q}$$

respectively instead of the corresponding equalities. The second one is not so obvious, but the first one is clear since we have the trivial embeddings $H^{m,p} \hookrightarrow H^{m-1,p}$ and $C^{k,\beta} \hookrightarrow C^{k-1,\beta}$.

The following theorem is very useful for proving compactness.

Theorem 2.2.49 (Rellich–Kondrachov)

1. *Let $\Omega \subset \mathbf{R}^n$ be a bounded domain. Moreover, let $1 \leq p, q < \infty$ and $m_1 > m_2 \geq 0$ integers so that*

$$m_1 - \frac{n}{p} > m_2 - \frac{n}{q}$$

Then the embedding

$$H_0^{m_1,p}(\Omega) \hookrightarrow H_0^{m_2,q}(\Omega)$$

from theorem 2.2.48 is compact, i.e. bounded sets in $H_0^{m_1,p}(\Omega)$ are precompact in $H_0^{m_2,q}(\Omega)$.

2. *Let $\Omega \subset \mathbf{R}^n$ be a bounded domain, $1 \leq m \in \mathbf{N}$ and $p \geq 1$. We assume that*

$$m - \frac{n}{p} > k + \beta$$

for some integer $k \geq 0$ and $0 < \beta < 1$. Then the embedding from theorem 2.2.46

$$H_0^{m,p}(\Omega) \hookrightarrow C^{k,\beta}(\overline{\Omega})$$

is compact.

Proof:
(Second part only)
We may choose $\tilde{p} < \infty$, $\tilde{k} \geq 0$ and $0 < \tilde{\beta} < 1$ so that

$$m - \frac{n}{p} \geq m - \frac{n}{\tilde{p}} = \tilde{k} + \tilde{\beta} > k + \beta.$$

Choose now $r > 0$ so large that $\Omega \subset B_r(0)$. Functions in $H_0^{m,p}(\Omega)$ can be extended trivially to functions in $H_0^{m,p}(B_r(0))$, and the embedding

$$H_0^{m,p}(B_r(0)) \hookrightarrow H_0^{m,\tilde{p}}(B_r(0))$$

is continuous by Hölder's inequality. On the other hand, the embedding

$$H_0^{m,\tilde{p}}(B_r(0)) \hookrightarrow C^{\tilde{k},\tilde{\beta}}(\overline{B_r(0)})$$

is continuous by the Sobolev–embedding theorem 2.2.46. The embedding

$$C^{\tilde{k},\tilde{\beta}}(\overline{B_r(0)}) \hookrightarrow C^{k,\beta}(\overline{B_r(0)})$$

is compact because of the Ascoli–Arzela theorem.

❑

Remark:
We placed the domain Ω into a ball for the following reason: There is only an embedding $C^1(\overline{\Omega}) \subset C^{0,\beta}(\overline{\Omega})$, or $C^{\tilde{k},\tilde{\beta}}(\overline{\Omega}) \subset C^{k,\beta}(\overline{\Omega})$ with $\tilde{\beta} < \beta$ if Ω has Lipschitz boundary, i.e. if the boundary is locally the graph of a Lipschitz continuous function. Consider the following domain

$$\Omega := \{(x,y) \in \mathbf{R}^2 \,|\, y < \sqrt{|x|}\ ,\ x^2 + y^2 < 1\},$$

which has no Lipschitz boundary since there is a cusp at the origin. The function

$$u(x,y) := \begin{cases} \operatorname{sign}(x)\, y^\alpha & \text{if} \quad y > 0 \\ 0 & \text{if} \quad y \leq 0 \end{cases}$$

is in $C^1(\overline{\Omega})$ if $1 < \alpha < 2$. Let $\alpha/2 < \beta < 1$. Then

$$\frac{|u(x,\sqrt{|x|}) - u(-x,\sqrt{|-x|})|}{2|x|^\beta} = |x|^{\alpha/2-\beta}$$

which tends to ∞ as $|x| \to 0$. Therefore $u \notin C^{0,\beta}(\overline{\Omega})$. If we change the domain Ω to

$$\{(x,y) \in \mathbf{R}^2 \,|\, y < |x|\ ,\ x^2 + y^2 < 1\}$$

which has Lipschitz boundary, then $u \in C^1(\overline{\Omega})$ as well, but $u \in C^{0,\beta}(\overline{\Omega})$ for all $0 < \beta \leq 1$.

2.3 Noncompactness of the unit ball, Uniform Convexity

We will prove the following theorem which is the main tool for showing whether a normed space is finite or infinite dimensional.

Theorem 2.3.1 *The normed vector space X is infinite dimensional if and only if the closed unit ball $B := \{x \in X \mid |x| \leq 1\}$ is not compact.*

We start with a lemma

Lemma 2.3.2 *Let Y be a closed proper subspace of the normed vector space X. Then there is $z \in X$ with $|z| = 1$ and*

$$|z - y| > \frac{1}{2} \text{ for all } y \in Y. \tag{2.6}$$

Proof:
Pick a point $x \in X \backslash Y$. Because Y is closed we have

$$\inf_{y \in Y} |x - y| = d > 0.$$

We can then find a point $y_0 \in Y$ such that

$$|x - y_0| < 2d.$$

We define $z' := x - y_0$ and note that for any $y \in Y$

$$|z' - y| = |x - y_0 - y| \geq \inf_{y \in Y} |x - y| = d.$$

We now define

$$z := \frac{z'}{|z'|}$$

so that for any $y \in Y$

$$|z - y| = \frac{1}{|z'|} \left| z' - |z'| y \right| \geq \frac{1}{2d} \cdot d = \frac{1}{2}.$$

❑

Proof:
(theorem 2.3.1)
If X is finite dimensional then all norms are equivalent to the standard Euclidean norm. In this case we have the Heine–Borel property, i.e. all closed bounded sets are compact. We will show now that the closed unit ball is not compact if X is infinite dimensional. We construct a sequence $(x_n)_{n \in \mathbf{N}}$ recursively as

follows: Let $0 \neq x_1 \in X$ be arbitrary with $|x_1| = 1$. If $x_1, \ldots, x_{n-1}$ are given, then we construct x_n as follows: The space

$$Y_n := \text{Span}\{x_1, \ldots, x_{n-1}\}$$

is finite–dimensional, hence it is closed, and it is a proper subspace of X (recall that X is infinite dimensional). Using the above lemma we can find an element $x_n \in X$ with $|x_n| = 1$ and $|x_n - y| > 1/2$ for all $y \in Y_n$, in particular

$$|x_n - x_j| > \frac{1}{2} \ \forall\, j < n.$$

The sequence constructed like this has the property that any two elements x_k and x_j with $j \neq k$ satisfy $|x_j - x_k| > 1/2$. Hence $(x_n)_{n \in \mathbf{N}} \subset B$ has no convergent subsequence. Therefore B is not compact.

❑

Definition 2.3.3 1. *Let $(X, |.|)$ be a normed vector space. The norm $|.|$ is called strictly subadditive if $|x + y| = |x| + |y|$ and $x, y \neq 0$ implies that x is a nonnegative multiple of y.*

2. *Let $(X, |.|)$ be a normed vector space and denote by B the closed unit ball. The space $(X, |.|)$ is called uniformly convex if for every $\varepsilon > 0$ there is some $\delta > 0$ such that*

$$x, y \in B \text{ and } |x - y| \geq \varepsilon$$

imply that

$$\left| \frac{x + y}{2} \right| \leq 1 - \delta.$$

Example: The property of being uniformly convex is a geometric property, not a topological property. A normed vector space may be uniformly convex with respect to one norm, but not with respect to another **equivalent** norm. If $X = \mathbf{R}^2$ and

$$|x|_1 := \sqrt{x_1^2 + x_2^2}\ , \ |x|_2 := |x_1| + |x_2|$$

then $(X, |.|_1)$ is uniformly convex, while $(X, |.|_2)$ is not.

If X is a normed vector space with strictly subadditive norm and $x \neq y \in \partial B$ then they are not nonnegative multiples of each other and therefore $|x + y| < |x| + |y| = 2$, i.e.

$$\left| \frac{x + y}{2} \right| < 1.$$

Uniformly convex means that this condition holds uniformly.

Proposition 2.3.4 *If $(X, |.|)$ is uniformly convex then the norm $|.|$ is strictly subadditive.*

Proof:
Assume that $|x+y| = |x| + |y|$ and $x, y \neq 0$. Dividing by $|x|$ we may assume that $|x| = 1$ and $|x+y| = 1 + |y|$. Then

$$\begin{aligned} \left| x + y - \frac{x+y}{|x+y|} \right| &= \left| \left(1 - \frac{1}{|x+y|}\right)(x+y) \right| \\ &= |x+y| \left(1 - \frac{1}{|x+y|}\right) \\ &= |x+y| - 1 \\ &= |y| \end{aligned}$$

and also

$$|(x+y) - x| = |y|.$$

We have

$$|y| \geq \left| \frac{1}{2}\left(x + y - \frac{x+y}{|x+y|}\right) + \frac{1}{2}((x+y) - x) \right| = \left| x + y - \frac{1}{2}\left(x + \frac{x+y}{|x+y|}\right) \right|.$$

Then

$$\begin{aligned} |x+y| &= \left| x + y - \frac{1}{2}\left(x + \frac{x+y}{|x+y|}\right) + \frac{1}{2}\left(x + \frac{x+y}{|x+y|}\right) \right| \\ &\leq |y| + \frac{1}{2}\left| x + \frac{x+y}{|x+y|} \right| \end{aligned}$$

and consequently

$$1 \leq \frac{1}{2}\left| x + \frac{x+y}{|x+y|} \right|.$$

The two vectors x and $(x+y)/|x+y|$ have norm 1. They cannot be different because of uniform convexity, hence

$$x = \frac{x+y}{|x+y|}$$

and $x = \lambda y$ with $\lambda = (|x+y| - 1)^{-1}$. This proves that the norm is strictly subadditive.

❑

The following theorem illustrates the importance of uniformly convex spaces.

Theorem 2.3.5 *Let X be a uniformly convex Banach space. Moreover, let $C \subset X$ be a closed convex subset of X and let $z \in X$. Then there is a unique point $y_0 \in C$ so that*

$$|y_0 - z| = \inf_{y \in C} |y - z|.$$

If the set C was compact then the conclusion of the theorem would be trivial since the continuous function $\phi : C \to [0,\infty)$, $\phi(y) := |y - z|$ would attain a minimum at some point $y_0 \in C$. The theorem is useful because the minimum is attained on a closed convex subset which is not necessarily compact. In an infinite dimensional Banach space there are much more closed, convex sets than compact sets.

Proof:
If $z \in C$ then $y_0 = z$, so assume that $z \notin C$. By translation we may assume that $z = 0$. We have

$$\inf_{y\in C} |y| > 0$$

since $0 \notin C$ and C is closed. Let $(y_n')_{n\in\mathbf{N}} \subset C$ be a minimizing sequence, i.e.

$$|y_n'| \longrightarrow \inf_{y\in C} |y|.$$

We write now $y_n := y_n'/|y_n'|$ and

$$\frac{y_n + y_m}{2} = \frac{1}{2|y_n'|} y_n' + \frac{1}{2|y_m'|} y_m' = \left(\frac{1}{2|y_n'|} + \frac{1}{2|y_m'|}\right)(c_n y_n' + c_m y_m')$$

where $c_n, c_m > 0$ with $c_n + c_m = 1$. Because C is convex by assumption we have

$$c_n y_n' + c_m y_m' \in C$$

as well, so that

$$|c_n y_n' + c_m y_m'| \geq \inf_{y\in C} |y|.$$

We conclude that

$$\left|\frac{y_n + y_m}{2}\right| \geq \left(\frac{1}{2|y_n'|} + \frac{1}{2|y_m'|}\right) \inf_{y\in C} |y|.$$

The right hand side of the above inequality converges to 1 as $n, m \to \infty$. Uniform convexity implies that the sequence $(y_n)_{n\in\mathbf{N}}$ must be a Cauchy sequence. Otherwise, the expression

$$\left|\frac{y_n + y_m}{2}\right|$$

would have to be smaller than $1 - \delta$ for some $\delta > 0$. By completeness the sequence (y_n) converges. The sequence $y_n' = |y_n'| \, y_n$ also converges to some limit $y_0 \in C$ since C is closed. By continuity we have $|y_0| = \inf_{y\in C} |y|$.

❑

We will show later that the L^p–spaces and Sobolev $W^{m,p}$–spaces are uniformly convex if $1 < p < \infty$. In the case of $p = 1$, $p = \infty$ or C^k–spaces the corresponding norms are not even strictly subadditive as the following trivial example illustrates: Let $X = C^0([-1,1])$ with the supremum norm. Take $f(x) = |x|$ and g so that $g(x) \equiv 0$ for $x \le 0$ and $g(x) = x$ for $x > 0$. Then

$$\|f + g\|_{C^0([-1,1])} = 2 = \|f\|_{C^0([-1,1])} + \|g\|_{C^0([-1,1])},$$

but f, g are not multiples of each other. Hence the supremum norm on $X = C^0([-1,1])$ is not strictly subadditive and $(X, \|.\|_{C^0([-1,1])})$ is not uniformly convex either. The next example shows that the above minimizing result may fail if the Banach space is not uniformly convex.

Example:
Consider $X = C^0([-1,1])$ with the supremum norm. Define

$$C := \left\{ f \in X \,\Big|\, \int_{-1}^{0} f(t)dt = 0\,,\ \int_{0}^{1} f(t)dt = 0 \right\}$$

which is a closed linear subspace of X and therefore convex. Let now z be a continuous function so that

$$\int_{-1}^{0} z(t)dt = 1\,,\ \int_{0}^{1} z(t)dt = -1.$$

Then we have for all $f \in C$

$$\int_{-1}^{0} (z - f)(t)dt = 1\,,\ \int_{0}^{1} (z - f)(t)dt = -1$$

hence

$$\max_{-1 \le t \le 0} (z(t) - f(t)) \ge 1$$

and similarly

$$\min_{0 \le t \le 1} (z(t) - f(t)) \le -1.$$

Equality holds if and only if

$$z(t) - f(t) \equiv \begin{cases} 1 & \text{if} \quad -1 \le t \le 0 \\ -1 & \text{if} \quad 0 \le t \le 1 \end{cases}$$

which is certainly impossible because $z - f$ is continuous. Hence we have

$$\max_{-1 \le t \le 0} (z(t) - f(t)) \ge 1 \text{ or } \min_{0 \le t \le 1} (z(t) - f(t)) \le -1.$$

Therefore

$$\|z - f\|_{C^0([-1,1])} > 1.$$

On the other hand we can choose $f \in C$ so that

$$\max_{-1 \le t \le 0} (z(t) - f(t)) \text{ and } \min_{0 \le t \le 1} (z(t) - f(t)).$$

are as close to $+1$ and -1 as we wish, hence

$$\inf_{f\in C} \|z-f\|_{C^0([-1,1])} = 1.$$

The distance of z to the set C equals 1, but there is no element in C that has distance 1 from the point z. There would have to be one if X were uniformly convex. We will now show that the space $L^p(\Omega)$ is uniformly convex for $2 \leq p < \infty$. The proof for $1 < p < 2$ is more difficult, we will not do it here.

Lemma 2.3.6 (first inequality of Clarkson)
Let $2 \leq p < \infty$ and let $\Omega \subset \mathbf{R}^n$ be a domain. Then

$$\left\|\frac{f+g}{2}\right\|^p_{L^p(\Omega)} + \left\|\frac{f-g}{2}\right\|^p_{L^p(\Omega)} \leq \frac{1}{2}(\|f\|^p_{L^p(\Omega)} + \|g\|^p_{L^p(\Omega)}) \ \forall\ f,g \in L^p(\Omega).$$

Proof:
It is enough to show the inequality

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leq \frac{1}{2}(|a|^p + |b|^p) \tag{2.7}$$

for all $a, b \in \mathbf{R}$ and $2 \leq p$. We claim that

$$\alpha^p + \beta^p \leq (\alpha^2+\beta^2)^{p/2} \ \forall\ \alpha, \beta \geq 0. \tag{2.8}$$

The inequality (2.8) is trivially true for $\beta = 0$. Otherwise, dividing by β^p on both sides, it suffices to show (2.8) for the case $\beta = 1$. Then there is nothing more to do because the function

$$F : [0,\infty) \longrightarrow \mathbf{R}$$

$$F(x) := (x^2+1)^{p/2} - x^p - 1$$

is an increasing function with $F(0) = 0$. Hence we have proved (2.8). Take now

$$\alpha = \left|\frac{a+b}{2}\right| \ , \ \beta = \left|\frac{a-b}{2}\right|$$

and recall that the function $x \mapsto |x|^{p/2}$ is convex for $p \geq 2$. Then

$$\begin{aligned}
\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p &\leq \left(\left|\frac{a+b}{2}\right|^2 + \left|\frac{a-b}{2}\right|^2\right)^{p/2} \\
&= \left(\frac{a^2+b^2}{2}\right)^{p/2} \\
&\leq \frac{1}{2}(|a|^p + |b|^p).
\end{aligned}$$

□

Theorem 2.3.7 *Let $\Omega \subset \mathbf{R}^n$ be a domain and $2 \leq p < \infty$. Then the space $L^p(\Omega)$ is uniformly convex.*

Proof:
Let $\varepsilon > 0$ and suppose that $f, g \in L^p(\Omega)$ with

$$\|f\|_{L^p(\Omega)}, \|g\|_{L^p(\Omega)} \leq 1 \text{ and } \|f - g\|_{L^p(\Omega)} \geq \varepsilon.$$

Then Clarkson's first inequality implies that

$$\left\| \frac{f+g}{2} \right\|_{L^p(\Omega)}^p \leq 1 - \frac{\varepsilon^p}{2^p},$$

hence

$$\left\| \frac{f+g}{2} \right\|_{L^p(\Omega)} \leq 1 - \delta$$

with

$$\delta = 1 - \left(1 - \frac{\varepsilon^p}{2^p}\right)^{1/p} > 0.$$

□

The proof of the following result is identical to the above theorem:

Theorem 2.3.8 *Let $\Omega \subset \mathbf{R}^n$ be a domain, $k \geq 1$ an integer and $2 \leq p < \infty$. Then the spaces $W^{k,p}(\Omega)$ and $H_0^{k,p}(\Omega)$ are uniformly convex.*

Chapter 3

Linear Operators

3.1 Definitions and elementary properties

In this section X and Y will be normed vector spaces over the real numbers or over the complex numbers. We denote the norms by $\|.\|_X$ and $\|.\|_Y$, but we will sometimes drop the subscripts X and Y if there is no danger of confusion. We will investigate linear maps $T : X \to Y$, in particular continuous linear maps. If X is finite dimensional then a linear map T as above is always continuous. This is not the case if X is infinite dimensional. In a later chapter we will prove that X is finite dimensional if and only if every linear map $T : X \to \mathbf{R}$ is continuous. We start with the following lemma

Lemma 3.1.1 *If $T : X \to Y$ is linear then the following statements are equivalent:*

1. *T is continuous*
2. *There is a point $x_0 \in X$ so that T is continuous in x_0*
3. $$\sup_{\|x\|_X \leq 1} \|Tx\|_Y < \infty$$
4. *There is a constant $C > 0$ so that $\|Tx\|_Y \leq C \|x\|_X$ for all $x \in X$.*

Proof:
We show that (2) implies (3): There is some $\delta > 0$ so that

$$T(\overline{B_\delta(x_0)}) \subset B_1(T(x_0))$$

(here $B_r(x)$ denotes the ball with radius r centered at x). If $\|x\|_X \le 1$, i.e. $x \in \overline{B_1(0)}$, then $x_0 + \delta x \in \overline{B_\delta(x_0)}$, hence

$$T(x_0) + \delta T(x) = T(x_0 + \delta x) \in B_1(T(x_0))$$

and therefore

$$T(x) \in B_{1/\delta}(0), \text{i.e. } \|Tx\|_Y \le \frac{1}{\delta}.$$

We now show that (3) implies (4): Let C be the supremum in (3). Then if $x \ne 0$

$$\|Tx\|_Y = \|x\|_X \cdot \left\| T\left(\frac{x}{\|x\|_X}\right)\right\|_Y \le C\|x\|_X.$$

We show that (4) implies (1): Let $x, x_0 \in X$. Then

$$\|Tx - Tx_0\|_Y = \|T(x - x_0)\|_Y \le C\|x - x_0\|_X$$

which converges to zero if $x \to x_0$, hence T is continuous. The implication (1)$\Longrightarrow$(2) is trivial.

❑

Definition 3.1.2 *We define*

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

We call elements $T \in L(X,Y)$ 'continuous operators' or 'bounded operators'. The following expression is called the operator norm of $T \in L(X,Y)$

$$\|T\|_{L(X,Y)} := \sup_{\|x\|_X \le 1} \|Tx\|_Y.$$

We wll also use the notation $\|T\|$ instead.

By the previous lemma $\|T\| < \infty$ if T is continuous. It is the smallest number so that for all $x \in X$

$$\|Tx\|_Y \le \|T\| \; \|x\|_X.$$

It is clear that $L(X,Y)$ is a vector space. The operator norm is also a norm on $L(X,Y)$.

Exercise 3.1.3 *Verify that the operator norm is a norm.*

Proposition 3.1.4 *Assume that X, Y, Z are normed vector spaces. Let $T \in L(X,Y)$ and $S \in L(Y,Z)$. Then $ST \in L(X,Z)$ and*

$$\|ST\| \le \|S\| \, \|T\|.$$

If Y is a Banach space then $L(X,Y)$ equipped with the operator norm is also a Banach space.

Proof:
The first statement is immediate since the composition of continuous maps is again continuous. The second statement follows from

$$\begin{aligned}
\|ST\| &= \sup_{0\neq\|x\|\leq 1} \|STx\| \\
&= \sup_{0\neq\|x\|\leq 1} \|Tx\| \cdot \left\| S\left(\frac{Tx}{\|Tx\|}\right)\right\| \\
&\leq \sup_{0\neq\|x\|\leq 1} \|Tx\| \cdot \sup_{0\neq\|y\|\leq 1} \|Sy\| \\
&= \|S\|\,\|T\|.
\end{aligned}$$

Let T_k be a Cauchy sequence in $L(X,Y)$. Since $\|T_k x - T_l x\|_Y \to 0$ as $k,l\to\infty$ for any $x\in X$, the sequence $(T_k x)_{k\in\mathbf{N}}$ is a Cauchy sequence in Y. Hence for every $x\in X$ the pointwise limits

$$Tx := \lim_{k\to\infty} Tx_k \in Y$$

exist and T is linear. Then

$$\begin{aligned}
\|(T-T_k)x\|_Y &= \lim_{l\to\infty} \|(T_l - T_k)x\|_Y \\
&\leq \liminf_{l\to\infty} \|T_l - T_k\| \cdot \|x\|_X
\end{aligned}$$

so that $T - T_k \in L(X,Y)$ and

$$\|T-T_k\|_{L(X,Y)} \leq \liminf_{l\to\infty} \|T_k - T_l\| \to 0$$

as $k\to\infty$.

❑

We continue with a few definitions:

Definition 3.1.5

- $$L(X) := L(X,X),$$
- *We denote by 'Id' ('identity') the operator in $L(X)$ which maps x onto x.*
- *The space $X' := L(X,\mathbf{R})$ is called the dual space of X. Elements in X' are also called linear functionals. We will discuss dual spaces in detail when we prove the Hahn–Banach theorem and when we consider weak and weak*–topologies.*
- *The set of compact operators from X to Y is defined by*

  $$K(X,Y) := \{T\in L(X,Y)\,|\,\overline{T(B_1(0))} \text{ is compact}\},$$

  where $B_1(0) := \{x\in X\,|\,|x|<1\}$. If Y is complete we may replace the above definition by "$T(B_1(0))$ is precompact".

- *A linear map $P \in L(X)$ is called a projection if $P^2 = P$.*
- *If $T \in L(X, Y)$ then we define the kernel of T*

$$\ker T := \{x \in X \,|\, Tx = 0\}$$

and the range of T

$$R(T) := \{y \in Y \,|\, \exists x \in X \,:\, Tx = y\}.$$

Since T is continuous the kernel of T is a closed subspace of X. The range of T is in general not closed (see example below).

Example:
Let $I = (0,1) \subset \mathbf{R}$. We define a bounded linear operator $T \in L(C^0(\bar{I}))$ as follows:

$$(Tf)(x) := \int_0^x f(t)dt \ , \ f \in C^0(\bar{I}).$$

This operator is indeed bounded, $\|T\| = 1$, and

$$R(T) = \{f \in C^1(\bar{I}) \,|\, f(0) = 0\} \subset C^0(\bar{I}),$$

which is not closed.

Exercise 3.1.6 *Let $P, Q : X \to X$ be linear operators with $PQ - QP = Id$. Then at least one of the operators P and Q has to be unbounded. This relation comes up in Quantum Mechanics. It is called the Heisenberg uncertainty principle.*

Exercise 3.1.7 *Consider the Dirac–sequence ρ_ε from the previous chapter. We have shown that*

$$(T_\varepsilon f)(x) := (f * \rho_\varepsilon)(x)$$

defines an operator $T_\varepsilon \in L(L^p(\mathbf{R}^n))$ with $\|T_\varepsilon\| \leq 1$ because of $\|f * \rho_\varepsilon\|_{L^p(\mathbf{R}^n)} \leq \|\rho_\varepsilon\|_{L^1(\mathbf{R}^n)}\|f\|_{L^p(\mathbf{R}^n)}$. We have also shown that for $p < \infty$

$$(T_\varepsilon - Id)f \stackrel{\varepsilon \to 0}{\longrightarrow} 0 \ \forall f \in L^p(\mathbf{R}^n).$$

Answer the following question: Is it true that $T_\varepsilon \longrightarrow Id$ in $L(L^p(\mathbf{R}^n))$?

We conclude this section with the following useful result

Proposition 3.1.8 *Let X be a Banach space and $T \in L(X)$ with $\limsup_{m\to\infty} \|T^m\|^{\frac{1}{m}} < 1$ (this is for example satisfied if $\|T\| < 1$). Then $Id - T$ has a continuous inverse and*

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

where the infinite series converges in $L(X)$.

Proof:
Let $S_k := \sum_{n=0}^k T^n$ and choose m and $\theta < 1$ such that $\|T^n\| \le \theta^n$ for all $n \ge m$. Then we conclude for $m \le k < l$

$$\|S_l - S_k\| = \left\| \sum_{k<n\le l} T^n \right\| \le \sum_{k<n\le l} \|T^n\| \le \sum_{k<n\le\infty} \theta^n \to 0$$

as $k \to \infty$. Because $L(X)$ is complete the limit

$$S := \lim_{k\to\infty} S_k$$

exists in $L(X)$. Moreover,

$$(\mathrm{Id} - T)S_k x = \sum_{n=0}^{k} (T^n - T^{n+1})x = x - T^{k+1}x$$

which converges to x as $k \to \infty$ because for large k we have $\|T^{k+1}x\| \le \theta^{k+1}\|x\| \to 0$. On the other hand, the left hand side of the above equation converges to $(\mathrm{Id} - T)Sx$, so that

$$(\mathrm{Id} - T)S = \mathrm{Id}.$$

The proof of $S(\mathrm{Id} - T) = \mathrm{Id}$ is similar and we omit it.

❑

3.2 The Banach–Steinhaus theorem

We start with the following fundamental result which is a consequence of Baire's lemma

Proposition 3.2.1 *Let (X, d) be a complete metric space. Moreover, let Y be a normed vector space and let $\mathcal{X} \subset C^0(X, Y)$ be a subset of the set of continuous maps from X to Y. Assume that for every $x \in X$*

$$\sup_{f\in\mathcal{X}} \|f(x)\| < \infty.$$

Then there is a ball $\overline{B_\varepsilon(x_0)} \subset X$ so that

$$\sup_{|x-x_0|\le\varepsilon} \sup_{f\in\mathcal{X}} \|f(x)\| < \infty.$$

Proof:
For $k \in \mathbf{N}$ we define the sets

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

Every $x \in X$ is contained in one of the sets A_k by assumption, hence

$$X = \bigcup_{k \in \mathbf{N}} A_k.$$

On the other hand, each set A_k is closed because the maps f are continuous. By Baire's lemma there is an integer k_0 so that the set A_{k_0} does contain an open ball $B_\varepsilon(x_0)$, i.e.

$$\sup_{x \in B_\varepsilon(x_0)} \sup_{f \in \mathcal{X}} \|f(x)\| \leq k_0.$$

❑

The Banach–Steinhaus theorem (also known as the principle of uniform boundedness) is a special case.

Theorem 3.2.2 *Let X be a Banach space, and let Y be a normed vector space. Assume that $\mathcal{T} \subset L(X, Y)$ with*

$$\sup_{T \in \mathcal{T}} \|Tx\| < \infty \ \forall \ x \in X.$$

Then

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

Proof:
We define

$$\mathcal{X} := \{f \in C^0(X, \mathbf{R}) \,|\, f(x) = \|Tx\|\,,\, T \in \mathcal{T}\}.$$

By the above proposition there are a constant $C < \infty$ and a ball $B_\varepsilon(x_0)$ so that

$$\|Tx\| \leq C \ \forall \ T \in \mathcal{T}\,,\, |x - x_0| \leq \varepsilon.$$

If $x \neq 0$ is arbitrary then we get for any $T \in \mathcal{T}$

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

i.e. $\|T\| \leq 2C/\varepsilon$.

❑

3.3 The open mapping principle and corollaries

Definition 3.3.1 *If X, Y are metric spaces then a map $f : X \to Y$ is called open if*

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

Remarks:
If f is bijective then f is open if and only if f^{-1} is continuous. Moreover, if X, Y are normed vector spaces and $T : X \to Y$ is linear then it is open if and only if there is some $\delta > 0$ such that $B_\delta(0) \subset T(B_1(0)) \subset Y$.

Theorem 3.3.2 (open mapping principle) *Let X, Y be Banach spaces, and let $T \in L(X, Y)$. Then T is open if and only if T is surjective.*

Proof:
Let us start with the easy direction first. Assume that T is open. Then $B_\delta(0) \subset T(B_1(0))$ for some positive δ implies that $B_r(0) \subset T(B_{r/\delta}(0))$ for any $r > 0$, i.e. any ball in Y centered at the origin is contained in the range of T.
Assume now that T is surjective. We have to show that there is a positive number δ such that $B_\delta(0) \subset T(B_1(0))$. Since T is surjective we have

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

Baire's lemma implies that there is an integer k_0 so that the set $\overline{T(B_{k_0}(0))}$ contains an open ball $B_\varepsilon(y_0)$. This implies the following: For any $y \in B_\varepsilon(0)$ there are points $x_j \in B_{k_0}(0)$ so that

$$Tx_j \stackrel{j\to\infty}{\longrightarrow} y_0 + y \in B_\varepsilon(y_0).$$

Recalling that T is surjective we pick $x_0 \in X$ with $Tx_0 = y_0$. We then obtain

$$T\left(\frac{x_j - x_0}{2(k_0 + |x_0|)}\right) = \frac{1}{2(k_0 + |x_0|)}(Tx_j - y_0) \longrightarrow \frac{y}{2(k_0 + |x_0|)}$$

and

$$\left\|\frac{x_j - x_0}{2(k_0 + |x_0|)}\right\| \le \left\|\frac{x_j}{2(k_0 + |x_0|)}\right\| + \left\|\frac{x_0}{2(k_0 + |x_0|)}\right\| < 1.$$

We have shown that any element in the ball $B_\delta(0)$ with

$$\delta = \frac{\varepsilon}{2(k_0 + |x_0|)}$$

is the limit of a sequence of the form Tx'_j where $|x'_j| < 1$, i.e.

$$B_\delta(0) \subset \overline{T(B_1(0))}, \tag{3.1}$$

The second step of the proof consists of showing that there is a possibly smaller radius δ so that $B_\delta(0) \subset T(B_1(0))$. Note that (3.1) implies the following: If $y \in B_\delta(0)$ then we can find $x \in B_1(0)$ so that $y - Tx \in B_{\delta/2}(0)$ which means that $2(y - Tx) \in B_\delta(0)$. We use this procedure to construct sequences $y_k \in B_\delta(0)$ and $x_k \in B_1(0)$ by demanding for

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

We conclude
$$\frac{y_{k+1}}{2^{k+1}} = \frac{y_k}{2^k} - T(2^{-k}x_k).$$
In the following sum, all the terms except two cancel each other
$$T\left(\sum_{k=0}^{m} 2^{-k}x_k\right) = y - \frac{y_{m+1}}{2^{m+1}} \longrightarrow y$$
as $m \to \infty$. We estimate
$$\sum_{k=0}^{m} |2^{-k}x_k| \le \sum_{k=0}^{m} 2^{-k} \le 2 < \infty,$$
hence the sequence
$$\left(\sum_{k=0}^{m} 2^{-k}x_k\right)_{m\in\mathbf{N}}$$
is a Cauchy sequence in X which converges to some
$$x := \sum_{k=0}^{\infty} 2^{-k}x_k \text{ and } |x| \le 2.$$
Since T is continuous we get $Tx = y$. Since $y \in B_\delta(0)$ was arbitrary we have shown that
$$B_\delta(0) \subset T(\overline{B_2(0)}) \subset T(B_3(0))$$
and
$$B_{\delta/3}(0) \subset T(B_1(0)).$$

❑

We note the following corollary

Corollary 3.3.3 (inverse map theorem) *Let X, Y be Banach spaces, and let $T \in L(X,Y)$ be bijective. Then the inverse T^{-1} is continuous.*

Proof:
The inverse is linear and by the open mapping principle T is open. Hence T^{-1} is continuous.

❑

Remark:
Let X be a vector space equipped with two norms $|\,.\,|_1$ and $|\,.\,|_2$ so that X is a Banach space with respect to each of these norms. Assume moreover, that there is a constant $C > 0$ such that
$$|x|_2 \le C\,|x|_1 \;\forall\; x \in X.$$

Then the two norms are equivalent. This can be seen as follows: Consider the Banach spaces $X_1 = (X, |\,.\,|_1)$ and $X_2 = (X, |\,.\,|_2)$. The above inequality is the same as saying that the identity map

$$\mathrm{Id} : X_1 \longrightarrow X_2$$

is continuous. By the inverse map theorem its inverse $\mathrm{Id} : X_2 \to X_1$ is also continuous, i.e. there is another constant $c > 0$ such that

$$|x|_1 \le c\,|x|_2 \;\forall\, x \in X.$$

So the two norms are equivalent.

Theorem 3.3.4 (Closed graph theorem) *Let X, Y be Banach spaces. Let $T : X \to Y$ be a linear map so that its graph $G(T) := \{(x, y) \in X \times Y \mid y = Tx\}$ is closed in $X \times Y$. Then T is continuous.*

Proof:
We take advantage of the remark above. We consider the following norms on X:

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

We claim that X equipped with the norm $|x|_1$ is a Banach space. Indeed, assume that $(x_n)_{n\in\mathbf{N}}$ is a Cauchy sequence with respect to the norm $|\,.\,|_1$. Then

$$|x_n - x_m|\,,\; |Tx_n - Tx_m| \to 0 \text{ as } n, m \to \infty.$$

Hence $x_n \to x$ with respect to the usual norm $|\,.\,|_2$ on X and $Tx_n \to y \in Y$, i.e.

$$G(T) \ni (x_n, Tx_n) \longrightarrow (x, y).$$

Because the graph of T is closed we have $x, y) \in G(T)$, i.e. $y = Tx$ and $|x_n - x|_1 \to 0$. We have trivially $|x|_2 \le |x|_1$. By our remark above the two norms are equivalent so that

$$|Tx| \le |x|_1 = |x| + |Tx| \le C\,|x| \;\forall\, x \in X$$

which means that T is continuous.

❑

3.4 Topological complements, right and left inverses of operators

In this section we are going to investigate closed subspaces of Banach spaces. We will establish some properties which follow from the open mapping principle.

Definition 3.4.1 *Let $Y \subset X$ be a closed subspace of a Banach space X. Another subspace Z is called a topological complement of Y if*

- *Z is closed*
- *$X = Y \oplus Z$*

Proposition 3.4.2 *Let $Y \subset X$ be a closed subspace of a Banach space X, and let Z be a topological complement of Y. Because every $x \in X$ has a unique decomposition $x = y + z$ with $y \in Y$, $z \in Z$ we define the projections $\pi_Y : X \to Y$, $\pi_Z : X \to Z$ by $\pi_Y(x) := y$, $\pi_Z(x) := z$. Then the projections π_Y, π_Z are continuous linear operators.*

Proof:
We equip the product space $Y \times Z$ with the norm $\|(y, z)\| := |y| + |z|$ so that it becomes a Banach space (note that both Y, Z are closed). Then the linear operator

$$T : Y \times Z \longrightarrow X$$

$$T(y, z) := y + z$$

is continuous and surjective. By the open mapping principle there is $\delta > 0$ so that

$$B_\delta(0) \subset T(B_1(0)).$$

This means that every $x \in X$ with $|x| < \delta$ can be written as a sum $x = y + z$, $y \in Y$, $z \in Z$, so that $|y| + |z| < 1$. If x is now an arbitrary nonzero element in X then

$$x = \frac{2|x|}{\delta}\, y' + \frac{2|x|}{\delta}\, z' =: y + z$$

with $|y'| + |z'| < 1$ and

$$|y| + |z| < \frac{2}{\delta}\, |x|.$$

This implies

$$|\pi_Y(x)| = |y| \leq \frac{2}{\delta}|x| \text{ and } |\pi_Z(x)| = |z| \leq \frac{2}{\delta}|x|$$

which concludes the proof.

❑

The reverse of the above proposition is true in normed vector spaces (easy!)

Exercise 3.4.3 *Assume that $Y, Z \subset X$ are subspaces of a normed vector space X so that $X = Y \oplus Z$. If the projection operators π_Y, π_Z are continuous then Y, Z must be closed.*

Closed subspaces $Y \subset X$ of a Banach space which are finite dimensional or finite co–dimensional have topological complements. The Hahn–Banach theorem is needed for the part where $\dim(Y) < \infty$. We will show in the next chapter that every closed subspace of a Hilbert space has a topological complement. In contrast to the Hilbert space situation, Lindenstrauss and Tzafiri (On the complemented subspaces problem, Israel J. Math., 9, (1971)) have shown that every Banach space which is not isomorphic to a Hilbert space has a closed subspace that does not admit any topological complement. Having a topological complement or not is important for the construction of continuous right– or left inverses of continuous linear operators.

Proposition 3.4.4 *Assume X, Y are Banach spaces and $T \in L(X,Y)$ surjective. Then the following properties are equivalent:*

1. *T admits a (continuous) right inverse, which is an operator $S \in L(Y,X)$ so that $T \circ S = Id_Y$.*
2. *The kernel of T admits a topological complement.*

Proof:
Let S be a continuous right inverse. Then $R(S) = S(Y)$ is a topological complement of $\ker(T)$. Indeed, every $x \in X$ can be written as the sum of an element in $\ker T$ and one in $R(S)$:

$$x = (x - STx) + STx.$$

If $x_0 \in \ker T \cap R(S)$ then $Tx_0 = 0$ and there is $y_0 \in Y$ so that $x_0 = Sy_0$. This implies

$$0 = Tx_0 = TSy_0 = y_0 \text{ and } x_0 = 0,$$

hence $X = \ker T \oplus R(S)$. In order to check whether $R(S)$ is closed we pick a sequence $x_k \in R(S)$ which converges to some $x \in X$, and we have to show that $x \in R(S)$ as well. We have $x_k = Sy_k$ for suitable $y_k \in Y$. Applying T we get $y_k = Tx_k \to Tx$ since T is continuous. Then $x_k = Sy_k \to STx = x$.
Assume now that Z is a topological complement of $\ker T$. Then the projection $\pi_Z : X \to Z$ is continuous. If $y \in Y$ then let $x \in X$ be any point so that $Tx = y$. We then define $Sy := \pi_Z x$. Note that this definition does not depend on the particular choice of x.

❑

We leave the following proposition as an exercise.

Proposition 3.4.5 *Assume X, Y are Banach spaces and $T \in L(X,Y)$ injective. Then the following properties are equivalent:*

1. *T admits a (continuous) left inverse, which is an operator $S \in L(Y,X)$ so that $S \circ T = Id_X$.*
2. *The range of T is closed and admits a topological complement.*

Chapter 4

Hilbert spaces

4.1 Definitions, orthogonal complement, Uniform convexity

Definition 4.1.1 *Let X be a (real) vector space. A scalar product is a map $X \times X \to \mathbf{R}$, denoted by (x, y), which satisfies the following conditions for all $x, x_1, x_2, y \in X$:*

- *Symmetry:* $(x, y) = (y, x)$,
- $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$, $(\alpha x, y) = \alpha(x, y)\ \forall\ \alpha \in \mathbf{R}$
- *Positive definite:* $(x, x) > 0$ *if* $x \neq 0$

Two vectors $x, y \in X$ are called orthogonal if $(x, y) = 0$.

A scalar product induces a norm on X by $|x| := \sqrt{(x, x)}$ (we will prove the triangle inequality $|x + y| \leq |x| + |y|$ below).

Definition 4.1.2 *A vector space with a scalar product is called a Hilbert space if it is complete with respect to the induced norm.*

Example: The space $L^2(\Omega)$ becomes a Hilbert space if we define the scalar product by

$$(f, g)_{L^2(\Omega)} := \int_\Omega f(x)\overline{g(x)}dx \text{ or } (f, g)_{L^2(\Omega)} := \int_\Omega f(x)g(x)dx,$$

where we adopt the first definition if f, g are complex–valued. In the real valued case the two are the same. In the same way $W^{k,2}(\Omega)$ and $H_0^{k,2}(\Omega)$ also become Hilbert spaces with

$$(f, g)_{W^{k,2}(\Omega)} := \sum_{0 \leq |\alpha| \leq k} (D^\alpha f, D^\alpha g)_{L^2(\Omega)}.$$

If the vector space under consideration is a complex vector space then we have to modify the definition of scalar product as follows:

- $(\, . \, , \, . \,) : X \times X \to \mathbf{C}$
- $(y, x) = \overline{(x, y)}$ "skew–symmetry"
- $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$, $(\alpha x, y) = \alpha(x, y) \; \forall \; \alpha \in \mathbf{C}$
- $(x, x) > 0$ if $x \neq 0$

Note that

$$(\lambda x, y) = \lambda \, (x, y) \; , \; (x, \lambda y) = \overline{\lambda} \, (x, y)$$

where $\lambda \in \mathbf{C}$. We derive now some simple inequalities and identities. Let X be a complex vector space with a scalar product. Denote by $| \, . \, |$ the norm induced by the scalar product. If $t \in \mathbf{R}$ and $0 \neq y \in X$ then

$$|x + ty|^2 = |x|^2 + 2t \operatorname{Re}(x, y) + t^2 |y|^2 \geq 0. \tag{4.1}$$

Choosing now $t = -\operatorname{Re}(x, y)/|y|^2$ and multiplying (4.1) by $|y|^2$ we obtain

$$|x|^2 |y|^2 - 2(\operatorname{Re}(x, y))^2 + (\operatorname{Re}(x, y))^2 \geq 0$$

and

$$(\operatorname{Re}(x, y))^2 \leq |x|^2 |y|^2.$$

Replace now x by λx where λ is a complex number with $|\lambda| = 1$ so that $\lambda(x, y) \in \mathbf{R}$. Then we obtain the *Cauchy–Schwarz inequality*

$$|(x, y)| \leq \; |x| \, |y|.$$

Choosing $t = 1$ in (4.1) and estimating the term in the middle with the Cauchy–Schwarz inequality we obtain

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

This is the triangle inequality for the induced norm. Setting $t = \pm 1$ in (4.1) we obtain the *parallelogram identity*

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

Hilbert spaces have many nice properties in comparison to general Banach spaces. We will mention some of them in this section. Given a Banach space $(X, |.|)$ it is a natural question to ask whether X is actually a Hilbert space in the following sense: Is there a scalar product on X so that the induced norm equals the given norm $| \, . \, |$? There are some results in this direction. Here are some of them

Theorem 4.1.3 (Fréchet–von Neumann–Jordan) *Let $(X, |.|)$ be a Banach space and assume that the norm satisfies the parallelogram identity. Then X is a Hilbert space.*

See the book by Yosida for a proof.

Theorem 4.1.4 (Kakutani) *Let X be a normed vector space with dimension equal or greater than three. Assume that every two–dimensional subspace Y of X admits a projection P of norm ≤ 1, i.e. there is $P \in L(X, Y)$ with $Py = y$ for all $y \in Y$ and $\|P\| \leq 1$. Then the norm on X is induced by a scalar product.*

For a proof see S. Kakutani, Some characterizations of Euclidean spaces, Jap. J. Math., 16, (1940), pp. 93–97. It follows from the Hahn–Banach theorem (which we will prove in the following chapter) that every one–dimensional subspace has a projection P as in the theorem.

Theorem 4.1.5 *Let X be a Hilbert space. Denote the norm induced by the scalar product by $|.|$. Then $(X, |.|)$ is a uniformly convex Banach space.*

Proof:
Let $\varepsilon > 0$, $x, y \in X$ with $|x|, |y| \leq 1$ and $|x - y| \geq \varepsilon$. We obtain from the parallelogram identity

$$\left|\frac{x+y}{2}\right|^2 = 2\left|\frac{x}{2}\right|^2 + 2\left|\frac{y}{2}\right|^2 - \left|\frac{x-y}{2}\right|^2 \leq 1 - \frac{\varepsilon^2}{4}.$$

Hence

$$\left|\frac{x+y}{2}\right| \leq 1 - \delta$$

with

$$\delta = 1 - \sqrt{1 - \frac{\varepsilon^2}{4}}.$$

❑

We have the following result

Theorem 4.1.6 *Let X be a Hilbert space. Assume that $C \subset X$ is a nonempty closed convex subset of X. Then for every point $x_0 \in X$ there is a unique point $y_0 \in C$ such that*

$$|x_0 - y_0| = \inf_{y \in C} |x_0 - y|,$$

i.e. the point y_0 is closer to x_0 than any other point in C. Moreover, the point y_0 is characterized by the property

$$y_0 \in C\,,\ (x_0 - y_0, y - y_0) \leq 0\ \forall\ y \in C.$$

Proof:
Existence of y_0 follows from our earlier result about uniformly convex Banach spaces. We actually forgot in the earlier proof to show uniqueness of y_0. So we

will do it here. Apart from this, the only new feature here is the alternative characterisation of the minimizer in terms of the scalar product. Let us first show uniqueness. Assume that there are two distinct minimizers y_0 and y_1, i.e. $|y_1 - y_0| \geq \varepsilon$ for some $\varepsilon > 0$. Let $M := \inf_{y \in C} |x_0 - y|$. Then $(y_0 + y_1)/2 \in C$ and

$$M \leq \left| \frac{y_0 + y_1}{2} - x_0 \right| \leq \left| \frac{y_0 - x_0}{2} + \frac{y_1 - x_0}{2} \right| \leq M/2 + M/2$$

hence

$$1 = \left| \frac{y_0 - x_0}{2M} + \frac{y_1 - x_0}{2M} \right|$$

violating uniform convexity. Assume now that $y_0 \in C$ satisfies

$$|y_0 - x_0| = \inf_{y \in C} |y - x_0|.$$

Let $y \in C$ so that $z = (1-t)y_0 + ty \in C$ for all $t \in (0, 1]$. We obtain

$$|y_0 - x_0| \leq |x_0 - [(1-t)y_0 + ty]| = |(x_0 - y_0) - t(y - y_0)|$$

and

$$|y_0 - x_0|^2 \leq |x_0 - y_0|^2 - 2t(x_0 - y_0, y - y_0) + t^2|y - y_0|^2.$$

Dividing by t we get

$$2(x_0 - y_0, y - y_0) \leq t|y - y_0|^2 \;\forall\; 0 < t < 1$$

which implies

$$(x_0 - y_0, y - y_0) \leq 0.$$

Assume now that $(x_0 - y_0, y - y_0) \leq 0$ for all $y \in C$. Then

$$|y_0 - x_0|^2 - |y - x_0|^2 = |y_0|^2 - 2(x_0, y_0) - |y|^2 + 2(x_0, y)$$

which equals

$$2(x_0 - y_0, y - y_0) - |y_0 - y|^2 \leq 0.$$

Therefore,

$$|y_0 - x_0| \leq |y - x_0|$$

for all $y \in C$.

❑

Definition 4.1.7 *Let X be a Hilbert space and let Y be a closed subspace. We define another subspace of X, the orthogonal complement of Y, by*

$$Y^\perp := \{x \in X \,|\, (x, y) = 0 \;\forall\; y \in Y\}.$$

The following proposition demonstrates that topological complements of closed subspaces always exist in the Hilbert space setting.

Proposition 4.1.8 *Let X be a Hilbert space and let Y be a closed subspace of X. Then $Y^\perp$ is a topological complement of Y and $(Y^\perp)^\perp = Y$.*

Proof:
By the properties of the scalar product $Y^\perp$ is a vector space. Assume that x_k is a sequence in $Y^\perp$ which converges to some $x \in X$. If we can show that $x \in Y^\perp$ then $Y^\perp$ is closed. We compute for arbitrary $y \in Y$

$$(x, y) = (x - x_k, y) + (x_k, y)$$

and

$$|(x, y)| \leq |x - x_k|\,|y|$$

by the Cauchy Schwarz inequality. Then $(x, y) = 0$ for all $y \in Y$ and $Y^\perp$ is closed. We now show that $X = Y \oplus Y^\perp$. Given $x \in X$ there is an element $y \in Y$ which is closest to x, i.e. for any $z \in Y$ and any $t \in \mathbf{R}$

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

i.e. $2\,\mathrm{Re}\,(x - y, z) \geq -t|z|^2$ for all positive t and $2\,\mathrm{Re}\,(x - y, z) \leq -t|z|^2$ for all negative t. Therefore, $\mathrm{Re}\,(x - y, z) = 0$ and every $x \in X$ can be decomposed as a sum $x = y + x - y$ where $y \in Y$ and $x - y \in Y^\perp$. If $y \in Y \cap Y^\perp$ then we have in particular $(y, y) = |y|^2 = 0$. i.e. $y = 0$. The last statement is an immediate consequence.

❑

4.2 Riesz–Fischer representation theorem and Lax–Milgram lemma

In this section we assume that X is a vector space over the complex numbers. The theorems we are going to prove have obvious counterparts for the real case.

Theorem 4.2.1 (Riesz–Fischer representation theorem)
Let X be a Hilbert space. The following map is a conjugate linear isometric isomorphism between X and its dual space $X' = L(X, \mathbf{C})$:

$$J : X \longrightarrow X'$$

$$J(x)y := (y, x)\ ,\ \ x, y \in X$$

('conjugate' linear refers to $J(\alpha x) = \overline{\alpha}\,J(x)$ for $\alpha \in \mathbf{C}$).

Proof:
It is clear that $J(x) : X \to \mathbf{C}$ is linear. The same applies to $J(\alpha x) = \overline{\alpha}\,J(x)$ for $\alpha \in \mathbf{C}$ and $J(x_1 + x_2) = J(x_1) + J(x_2)$. The Cauchy–Schwarz inequality implies

that $J(x)$ is continuous, hence J is well–defined. Again, by Cauchy–Schwarz' inequality

$$\|J(x)\|_{X'} = \sup_{|y|\leq 1} |J(x)y| \leq |x|.$$

On the other hand,

$$\left| J(x)\frac{x}{|x|} \right| = |x|$$

so that $\|J(x)\|_{X'} = |x|$. Hence J is an isometry and injective. The real issue is to show that J is surjective, i.e. we have to show that for every element $\ell \in X'$ there is some $x \in X$ such that $\ell(y) = J(x)y$ for all $y \in Y$. Assume that $0 \neq \ell \in X'$. Then $\ker \ell$ is a closed proper subspace of X. Applying theorem 4.1.6, we can find for any element $x \in X$ a unique point $Px \in \ker \ell$ which is characterized by

$$|x - Px| = \inf_{y \in \ker \ell} |x - y|$$

and

$$\mathrm{Re}(x - Px, y - Px) \leq 0 \ \forall\ y \in \ker \ell.$$

Pick now $e \in X$ so that $\ell(e) = 1$ and define

$$x_0 := e - Pe$$

so that $\ell(x_0) = \ell(e) - \ell(Pe) = 1$, in particular, $x_0 \neq 0$. We have now for all $y \in \ker \ell$

$$\mathrm{Re}(e - Pe, y - Pe) = \mathrm{Re}(x_0, y - Pe) \leq 0.$$

But this implies

$$(x_0, y) = 0 \ \forall\ y \in \ker \ell.$$

We compute for arbitrary $x \in X$

$$\begin{aligned} (x, x_0) &= (x - \ell(x)x_0, x_0) + (\ell(x)x_0, x_0) \\ &= \ell(x)|x_0|^2 \end{aligned}$$

and

$$\ell(x) = \left(x, \frac{x_0}{|x_0|^2} \right) = J\left(\frac{x_0}{|x_0|^2} \right) x.$$

❑

The following theorem is a very effective tool for proving existence of solutions to linear elliptic partial differential equations.

Theorem 4.2.2 (Lax–Milgram lemma)
Let X be a Hilbert space, and let $a : X \times X \to \mathbf{C}$ ba a map which is linear with respect to the first variable and conjugate linear with respect to the second ($a(x, \alpha y) = \overline{\alpha}\, a(x, y)$). Moreover, assume that

- *a is bounded, i.e.* $|a(x, y)| \le C\,|x|\,|y|$ *for some constant* $C > 0$,
- $a(x, x) \ge c\,|x|^2$ *for some constant* $c > 0$,
- $a(x, y) = \overline{a(y, x)}$,

where $0 < c \le C < \infty$. *Then there is a unique bijective linear operator* $A \in L(X)$ *so that*

$$a(y, x) = (y, Ax)\ \forall\ x, y \in X.$$

In addition, we have $\|A\| \le C$ *and* $\|A^{-1}\| \le \frac{1}{c}$.

Proof:
For every $x \in X$ the map $y \mapsto a(y, x)$ is linear functional (this means an element in the dual space X') since

$$\|a(\,\cdot\,, x)\|_{X'} = \sup_{|y|\le 1} |a(y, x)| \le\ C|x|.$$

By the Riesz–Fischer representation theorem there is a unique point in X which we denote by Ax, so that

$$a(y, x) = (y, Ax)\ \forall\ y \in X.$$

Clearly, Ax depends linearly on x and

$$|Ax|^2 = (Ax, Ax) = |a(Ax, x)| \le\ C\,|Ax|\,|x|$$

so that

$$|Ax| \le\ C|x|,$$

hence $A \in L(X)$ and $\|A\| \le C$. We estimate

$$c|x|^2 \le a(x, x) = (x, Ax) \le |x|\,|Ax|$$

so that $c|x| \le |Ax|$ and therefore $\ker A = \{0\}$. The same estimate implies that $R(A)$ is closed. Indeed, let Ax_k be a sequence in $R(A)$ which converges to some point $y \in X$. Then

$$|x_k - x_l| \le \frac{1}{c}|Ax_k - Ax_l| \to 0 \text{ as } k, l \to \infty,$$

so (x_k) is a Cauchy sequence in X which must converge to some point $x \in X$. Since A is continuous we obtain $Ax_k \to Ax$ which must equal y. This shows that $R(A)$ is closed. It remains to show that $R(A) = X$. Assume this is not the case. We apply the orthogonal projection result, theorem 4.1.6: For every $x_0 \in X\backslash R(A)$ there is a unique element $Px_0 \in R(A)$ so that

$$\mathrm{Re}(x_0 - Px_0, y - Px_0) \le 0\ \forall\ y \in R(A)$$

or

$$\mathrm{Re}(x_0 - Px_0, y) \le\ \mathrm{Re}(x_0 - Px_0, Px_0) \le 0\ \forall\ y \in R(A)$$

which is only possible if $\mathrm{Re}(x_0 - Px_0, y) = 0 \ \forall\, y \in R(A)$. This is the same as

$$(x_0 - Px_0, y) = 0 \ \forall\, y \in R(A).$$

If we take $y := A(x_0 - Px_0)$ we obtain the following contradiction

$$0 = (x_0 - Px_0, A(x_0 - Px_0)) = a(x_0 - Px_0, x_0 - Px_0) \geq c|x_0 - Px_0|^2$$

which implies that $x_0 = Px_0 \in R(A)$ contradicting our initial assumption $x_0 \in X \backslash R(A)$.

❑

4.3 Some Applications

4.3.1 Dirichlet Problem

A. Classical Formulation

Let $\Omega \subset \mathbf{R}^n$ be a bounded domain, $f \in C^0(\overline{\Omega})$, $a_{ij} \in C^1(\overline{\Omega})$, $i, j = 1, \ldots, n$ are given real valued functions so that $a_{ij} = a_{ji}$. We also assume that there is a constant $c_0 > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq c_0|\xi|^2 \ \forall\, x \in \Omega\,,\ \xi \in \mathbf{R}^n.$$

We then say that $(a_{ij})_{1\leq i,j\leq n}$ is elliptic (Note that for fixed $x \in \Omega$ and $c > 0$ the set $\{\xi \in \mathbf{R}^n \mid \sum_{i,j} a_{ij}(x)\xi_i\xi_j = c$ is an ellipse). We are looking for a function $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ that solves the following boundary value problem ('Dirichlet–problem')

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u}{\partial x_j}(x)\right) = f(x) \text{ for } x \in \Omega \tag{4.2}$$

$$u(x) = g(x) \text{ for } x \in \partial\Omega,$$

where $g \in C^2(\Omega) \cap C^0(\overline{\Omega})$ is a given function. In this section we will use the notation $\partial_i := \partial/\partial x_i$. We can reduce this to the case $u|_{\partial\Omega} \equiv 0$ by replacing u with $u - g$. The Dirichlet–problem then becomes

$$\sum_{i,j} \partial_i(a_{ij}\partial_j u) = f - \sum_{i,j} \partial_i(a_{ij}\partial_j g) \text{ in } \Omega \tag{4.3}$$

$$u \equiv 0 \text{ on } \partial\Omega.$$

In order to shorten notation we write also

$$e_i := \sum_j a_{ij}\partial_j g.$$

We multiply the PDE ('partial differential equation') (4.3) with $\phi \in C_0^\infty(\Omega)$ and we perform partial integration. We obtain

$$\int_\Omega \left(\sum_{i,j} \partial_i\phi \, a_{ij} \partial_j u + \sum_i \partial_i \phi \, e_i + \phi \, f \right) \; = \; 0 \; \forall \; \phi \in C_0^\infty(\Omega). \tag{4.4}$$

Conversely, if the above equation is satisfied for all $\phi \in C_0^\infty(\Omega)$ then we can reverse the partial integration, and we conclude that u satisfies the differential equation (4.3). The functions ϕ are also called test functions.

B. Weak formulation in the Hilbert space $H_0^{1,2}(\Omega)$

Instead of searching directly for classical solutions of the partial differential equation (4.3) we define a more general notion of solution. Inspired by the integral identity (4.4) we make the following definition

Definition 4.3.1 (Weak solution)
We call u a weak solution of the Dirichlet problem (4.3) if $u \in H_0^{1,2}(\Omega)$ and

$$\int_\Omega \left(\sum_{i,j} \partial_i\phi \, a_{ij} \, \partial_j u + \sum_i \partial_i \phi \, e_i + \phi \, f \right) = 0 \; \textit{for all} \; \phi \in H_0^{1,2}(\Omega). \tag{4.5}$$

In comparison to (4.4) the space of solutions and the space of test functions has been chosen larger. Of course, every classical solution is also a weak solution.

C. Existence of a weak solution with Lax–Milgram lemma

We will prove the existence of weak solutions as in (4.5) using the Lax–Milgram lemma. In the classical formulation (4.4) the function spaces involved are not well–behaved while we are dealing with a Hilbert space in the case of weak solutions. We only need to assume here that $a_{ij} \in L^\infty$ and $e_i, f \in L^2(\Omega)$. Define for $v, w \in H_0^{1,2}(\Omega)$ the following bilinear map

$$a(v,w) := \int_\Omega \sum_{i,j} \partial_i v \, a_{ij} \, \partial_j w$$

which is also symmetric. We claim that the Lax Milgram lemma can be applied to a. We have

$$|a(v,w)| \leq \sum_{i,j} \|a_{ij}\|_{L^\infty(\Omega)} \|\partial_i v\|_{L^2(\Omega)} \|\partial_j v\|_{L^2(\Omega)} \leq C\|v\|_{1,2,\Omega}\|w\|_{1,2,\Omega}.$$

Ellipticity implies that there is a constant $c_0 > 0$ such that

$$\sum_{i,j=1}^n a_{ij}(x)\partial_i v(x) \partial_j v(x) \geq c_0 |\nabla v(x)|^2 \; \forall \; x \in \Omega \, , \; \nabla v = (\partial_1 v, \ldots, \partial_n v)$$

hence

$$|a(v,v)| \geq c_0 \int_\Omega |\nabla v(x)|^2 dx.$$

If we can show the inequality

$$\int_\Omega |v(x)|^2 dx \leq C_0 \int_\Omega |\nabla v(x)|^2 dx \ \forall\ v \in H_0^{1,2}(\Omega), \tag{4.6}$$

which is called *Poincaré inequality*, then the assumptions of the Lax Milgram lemma are satisfied and there is a linear operator $A \in L(H_0^{1,2}(\Omega))$ such that

$$a(w,v) = (w, Av)_{H^{1,2}(\Omega)} \ \forall\ v, w \in H_0^{1,2}(\Omega). \tag{4.7}$$

Now $v \in H_0^{1,2}(\Omega)$ is a weak solution if and only if

$$a(v,w) = F(w) \ \forall\ w \in H_0^{1,2}(\Omega)$$

where

$$F(w) := -\int_\Omega \sum_i \partial_i w\, e_i - \int_\Omega f\, w.$$

We note that $F \in (H_0^{1,2}(\Omega))'$ since

$$\begin{aligned} |F(w)| &\leq \|f\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)} + \sum_i \|e_i\|_{L^2(\Omega)} \|\partial_i w\|_{L^2(\Omega)} \\ &\leq C\,(\max_i \|e_i\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}) \|w\|_{1,2,\Omega}. \end{aligned}$$

The Riesz representation theorem guarantees the existence of some $u \in H_0^{1,2}(\Omega)$ such that

$$F(w) = (w,u)_{H^{1,2}(\Omega)}.$$

Putting all the pieces together we obtain for any $w \in H_0^{1,2}(\Omega)$

$$a(A^{-1}u, w) = a(w, A^{-1}u) \stackrel{(4.7)}{=} (w,u)_{H^{1,2}(\Omega)} \stackrel{Riesz}{=} F(w).$$

This means that $A^{-1}u$ is the desired weak solution. We are left with the proof of the Poincaré inequality (4.6). If suffices to show the inequality for a smooth function with compact support in Ω by an approximation argument. Let $u \in C_0^\infty(\Omega)$. We view u as a smooth function on all of $\mathbf{R}^n$ by trivial extension. Let $Q = (a_1, b_1) \times \cdots \times (a_n, b_n)$ be a cube containing the closure of Ω. We write $x = (x_1, \ldots, x_n)$ and $a = (a_1, \ldots, a_n)$. We estimate for $a_k \leq x_k \leq b_k$

$$\begin{aligned} |u(x)|^2 &= |u(x) - u(a)|^2 \\ &= \left| \int_{a_k}^{x_k} \partial_k u(x_1, \ldots, x_{k-1}, \xi, x_{k+1}, \ldots, x_n) dx_k \right|^2 \\ &\leq (x_k - a_k) \int_{a_k}^{x_k} |\partial_k u(x_1, \ldots, x_{k-1}, \xi, x_{k+1}, \ldots, x_n)|^2 dx_k \\ &\quad \text{(with Hölder's inequality)} \\ &\leq (b_k - a_k) \int_{a_k}^{b_k} |\partial_k u(x_1, \ldots, x_{k-1}, \xi, x_{k+1}, \ldots, x_n)|^2 dx_k. \end{aligned}$$

Integrating over $\bar{Q}$ yields

$$\int_{\bar{Q}} |u(x)|^2 dx \leq (b_k - a_k)^2 \int_{\bar{Q}} |\partial_k u(x)|^2 dx$$

which is the same as integrating over the domain Ω. We get an estimate like this for each k. Adding all of them yields the Poincaré inequality.

D. Regularity of weak solutions (some remarks only)

Having found a weak solution we pose the question whether a weak solution is actually a classical solution in the sense of (4.3) or (4.4). This is a complicated issue. Some references are the books by D. Gilbarg and N. Trudinger (Elliptic partial differential equations of second order), A. Friedman (Partial Differential Equations), J. Jost (Postmodern Analysis) or L. Evans (Partial Differential Equations). The coefficients a_{ij} will have to be more regular as previously specified in the existence proof. Without going into details, the typical regularity result for a weak solution u of (4.5) is an estimate of the form

$$\|u\|_{m+2,2,\Omega} \leq C\,(\|f\|_{m,2,\Omega} + \|g\|_{m+2,2,\Omega} + \|u\|_{m,2,\Omega}), \tag{4.8}$$

this means the regularity of the weak solution is always two notches better than the regularity of the data f. In particular, the weak solution will be smooth if f and g are (by the Sobolev embedding theorem).

E. Remarks

We could prove existence of weak solutions of linear elliptic partial differential equations with Dirichlet boundary conditions as in (4.2) in the Hilbert space $H_0^{1,2}(\Omega)$ using the Lax Milgram lemma and the Riesz Fischer representation theorem. So we did not need any Sobolev spaces $W^{k,p}$ with $p \neq 2$. So why bother with them ? There is also a notion of ellipticity for nonlinear differential equations and the picture changes drastically in this case. In the nonlinear case Sobolev spaces with $p \neq 2$ or Hölder spaces are usually used. We will discuss variational methods later on which are a valuable tool to prove existence of weak solutions for certain nonlinear PDE's. We note that there is no universal existence theorem of weak solutions for nonlinear elliptic partial differential equations. What about the regularity issue ? Regularity estimates like (4.8) for linear elliptic PDE's also exist in the general case $p > 1$ but there are much harder to prove than the Hilbert space case. There is also a similar estimate for Hölder spaces (equally hard to prove). See the book by D. Gilbarg and N. Trudinger or the paper by S. Agmon, A. Douglis and L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I, Communications on Pure and Applied Math., 12, (1959). We remark that there is no regularity estimate relating $\|u\|_{C^{k+2}(\Omega)}$ to the C^k–norms of f, g. In the chapter about Fredholm operators we will see why (another reason why the classical spaces $C^k(\Omega)$ are not suitable for partial differential equations). Solving the regularity question for nonlinear elliptic PDE's is very difficult. There is no complete answer known to the question for which nonlinear elliptic PDE's there is regularity of weak solutions and for which there is not.

4.3.2 Radon–Nikodym theorem

Let $(M, \mathcal{M}, \nu)$ and $(M, \mathcal{M}, \mu)$ be two σ–finite measure spaces with the same underlying set M and the same σ–algebra $\mathcal{M}$. We assume that $\nu(M) < \infty$. The measure ν is called absolutely continuous with respect to μ if

$$A \in \mathcal{M}\ ,\ \mu(A) = 0 \implies \nu(A) = 0.$$

The Radon–Nikodym theorem then promises us a nonnegative integrable function h (with respect to the measure μ) such that

$$\nu(A) = \int_A h \, d\mu \ \forall\ A \in \mathcal{M}.$$

This is a classical result in measure theory, and we will see shortly that it can be proved with the Riesz–Fischer representation theorem (the proof is due to von Neumann). We confine ourselves to the case $\mu(M) < \infty$, i.e. the measure of the total space is finite with respect to both measures. Let X be the real Hilbert space $L^2(\mu+\nu)$ with the norm

$$\|f\| := \left(\int_M f^2 \, d(\mu+\nu) \right)^{\frac{1}{2}}.$$

It follows from the Cauchy–Schwarz inequality that

$$L^2(\mu+\nu) \subset L^1(\mu+\nu) \subset L^1(\nu)$$

Then for any $f \in L^2(\mu+\nu)$ we can define

$$\ell(f) := \int_M f \, d\nu.$$

Again, by Cauchy–Schwarz we have

$$|\ell(f)| = |(1, f)_{L^2(\nu)}| \le |1|_{L^2(\nu)} |f|_{L^2(\nu)} \le |1|_{L^2(\nu)} |f|_{L^2(\mu+\nu)},$$

hence $\ell \in (L^2(\mu+\nu))'$. By the Riesz representation theorem we can find $g \in L^2(\mu+\nu)$ such that

$$\ell(f) = \int_M f \, d\nu = \int_M f g \, d(\mu+\nu) \ \forall\ f \in L^2(\mu+\nu).$$

Write this as

$$\int_M f(1-g)\,d\nu = \int_M fg\,d\mu \;\forall\; f \in L^2(\mu+\nu). \tag{4.9}$$

We define the set $Z := \{x \in M \,|\, g(x) \leq 0\}$, and we claim that $\mu(Z) = 0$. Take now $f \equiv 1$ on Z and $f \equiv 0$ on the complement of Z. Then (4.9) becomes

$$\int_Z (1-g)\,d\nu = \int_Z g\,d\mu. \tag{4.10}$$

By definition of Z we have for $\varepsilon > 0$

$$\begin{aligned} 0 &\leq \int_Z (1-g)\,d\nu = \int_Z g\,d\mu \\ &= \int_{\{g<-\varepsilon\}} g\,d\nu + \int_{\{0\geq g\geq -\varepsilon\}} g\,d\nu \\ &\leq -\varepsilon \cdot \mu(\{g < -\varepsilon\}), \end{aligned}$$

so that $\mu(\{g < -\varepsilon\}) = 0$ for all ε and also $\mu(Z) = 0$ proving the claim. We now define $Z := \{x \in M \,|\, g(x) \geq 1\}$, and we claim again that $\mu(Z) = 0$. We argue indirectly and assume that $\mu(Z) > 0$. Take $f \equiv 1$ on Z and $f \equiv 0$ on the complement of Z. Then we obtain again (4.10). This time the right hand side is positive, but the left hand side is zero or negative since $g \geq 1$ on Z. By absolute continuity of ν with respect to μ we also get $\nu(Z) = 0$ in both cases. Summarizing, we have shown that the function g satisfies

$$0 < g(x) < 1 \quad (\mu+\nu)\text{-almost everywhere.}$$

We then modify the function g on a set of μ-measure zero so that $0 < g(x) < 1$ for all $x \in M$. Because ν is absolutely continuous with respect to μ, the equation (4.9) still holds for the modified function (which we denote again by g). We claim that the desired function h in the theorem is given by

$$h(x) := \frac{g(x)}{1-g(x)}.$$

Let $E \in \mathcal{M}$ be a measureable set and denote its characteristic function by χ_E. Then we have for $k \in \mathbf{N}$

$$f := \frac{1-g^k}{1-g}\chi_E = \left(\sum_{j=0}^{k-1} g^j\right)\chi_E \in L^\infty(\mu+\nu) \subset L^2(\mu+\nu)$$

so that we can insert it into equation (4.9). We obtain

$$\int_M (1-g^k)\chi_E\,d\nu = \int_M (1-g^k)\chi_E \frac{g}{1-g}\,d\mu \;\forall\; k$$

We have $(\mu+\nu)$-almost everywhere $0 \leq (1-g^k)\chi_E \nearrow \chi_E$ as $k \to \infty$. The monotone convergence theorem then implies that

$$\begin{aligned}
\nu(E) &= \int_M \lim_{k\to\infty} (1-g^k)\chi_E d\nu \\
&= \lim_{k\to\infty} \int_M (1-g^k)\chi_E d\nu \\
&= \lim_{k\to\infty} \int_M (1-g^k)\chi_E \frac{g}{1-g}\, d\mu.
\end{aligned}$$

Fatou's lemma then implies that

$$\int_M \frac{g}{1-g}\chi_E \, d\mu = \int_M \liminf_{k\to\infty} (1-g^k)\chi_E \frac{g}{1-g}\, d\mu \leq \nu(E),$$

in particular, $\frac{g}{1-g}\chi_E \in L^1(\mu)$. Applying the monotone convergence theorem again (or the dominated convergence theorem) we finally get

$$\nu(E) = \int_E \frac{g}{1-g} d\mu$$

which is the assertion of the theorem.

Chapter 5

The Hahn–Banach theorem, Dual Spaces, Reflexivity

5.1 The different versions of the Hahn–Banach theorem

Before we discuss the Hahn–Banach theorem, let us recall Zorn's lemma. The lemma deals with partially ordered sets. Let P be a set so that a relation is defined on some pairs of points in P. We denote this relation by $\leq$. (The 'official' definition is this: A relation R on the set P is a subset of $P \times P$. We write $x \leq y$ if $(x, y) \in R$). A partial ordering on P is then a relation $\leq$ that satisfies the following conditions:

- If $x \leq y$ and $y \leq z$ then $x \leq z$,
- We have $x \leq x$ for all $x \in P$,
- If $x \leq y$ and $y \leq x$ then $x = y$.

The set P is called totally ordered if for any pair $(x, y) \in P \times P$ either $x \leq y$ or $y \leq x$. For example the real numbers are totally ordered with the relation $\leq=$ 'less or equal'. The set P of all real valued function on $[0, 1]$ becomes partially ordered if we define $f \leq g$ as $f(x) \leq g(x)$ for all $x \in [0, 1]$. Given a subset $Q \subset P$, an element $z \in P$ is called an upper bound for the set Q if $x \leq z$ for all $x \in Q$. An element z of a partially ordered set P is called maximal if $z \leq x$ for any $x \in P$ implies that $x \leq z$ as well.

Lemma 5.1.1 (Zorn's lemma)
If every totally ordered subset of a nonempty partially ordered set has an upper bound then the partially ordered set has a maximal element.

Theorem 5.1.2 (Hahn-Banach theorem, analytic version)
Let X be a vector space over the real numbers. Moreover, let $p : X \to \mathbf{R}$ be a map satisfying the following conditions:

- $p(\lambda x) = \lambda\, p(x)$ *for all* $x \in X$ *and* $\lambda > 0$,
- $p(x + y) \le p(x) + p(y)$ *for all* $x, y \in X$.

Assume also that $Y \subset X$ is a linear subspace and let $g : Y \to \mathbf{R}$ be a linear map satisfying $g(x) \le p(x)$ for all $x \in Y$. Then there is a linear map $G : X \to \mathbf{R}$ such that

- $G(x) = g(x)$ *if* $x \in Y$,
- $G(x) \le p(x)$ *for all* $x \in X$.

Proof:
We consider the following set:

$$\begin{aligned} P \;:=\; & \{(Z, h) \,|\, Z \text{ is a linear subspace with } Y \subset Z \subset X \text{ and} \\ & h : Z \to \mathbf{R} \text{ is a linear map so that } h(x) = g(x)\ \forall x \in Y \text{ and} \\ & h(x) \le p(x)\ \forall x \in Z\} \end{aligned}$$

We define a partial ordering on P as follows: We define

$$(Z_1, h_1) \le (Z_2, h_2) \;:\Longleftrightarrow\; Z_1 \subset Z_2 \text{ and } h_2|_{Z_1} \equiv h_1,$$

i.e. h_2 is an extension of h_1. This is just a partial ordering, not a total ordering, because the map g in general has many different possible extensions onto larger subspaces. The set P is not empty since it contains the element (Y, g). Assume now that $Q \subset P$ is a totally ordered subset, i.e.

$$Q = \bigcup_{i \in I} \{(Z_i, h_i)\},$$

where I is some index set. Then for $i, j \in I$ either h_i is an extension of h_j or vice versa (or both if $Z_i = Z_j$ and $h_1 = h_j$). The set Q has an upper bound $(Z, h) \in P$ as follows:

$$Z = \bigcup_{i \in I} Z_i \;,\; h(x) := h_i(x) \text{ if } x \in Z_i.$$

Note that h is well–defined because Q is totally ordered and that $(Z, h) \in P$. By Zorn's lemma the partially ordered set P then has a maximal element which we denote by (X', G). Our job will be to show that X' must be the whole space X. Then G is the desired map and the proof is complete. We argue indirectly and assume that $X' \ne X$. We pick $x_0 \in X \backslash X'$ and we define the following linear subspace of X:

$$Z := X' \oplus \mathbf{R}x_0.$$

We claim that we can extend G onto the space Z so that $(Z, G_{extended}) \in P$. This would violate the maximality property of (X', G), i.e. we have $(X', G) \leq (Z, G_{extended})$, but we do *not* have $(Z, G_{extended}) \leq (X', G)$, and it would complete the proof. We define

$$G_{extended}(x + tx_0) := G(x) + t\alpha$$

for a suitable constant α which we will determine now such that $(Z, G_{extended}) \in P$, i.e. we would like to have

$$G(x) + t\alpha \leq p(x + tx_0) \ \forall x \in X' \, , \, t \in \mathbf{R}. \tag{5.1}$$

This inequality is satisfied for $t = 0$ since $(X', G) \in P$. For $t > 0$ it leads to

$$\alpha \leq \frac{p(x + tx_0) - G(x)}{t} = p\left(\frac{x}{t} + x_0\right) - G\left(\frac{x}{t}\right),$$

and for $t < 0$ it leads to

$$\alpha \geq \frac{p(x + tx_0) - G(x)}{t} = -p\left(-\frac{x}{t} - x_0\right) + G\left(-\frac{x}{t}\right).$$

In particular, (5.1) is satisfied if

$$\sup_{x \in X'} (G(x) - p(x - x_0)) \leq \alpha \leq \inf_{x \in X'} (p(x + x_0) - G(x)). \tag{5.2}$$

If $x, x' \in X'$ are arbitrary points then we estimate

$$\begin{aligned} G(x) + G(x') &= G(x + x') \\ &\leq p(x + x') \\ &= p(x - x_0 + x' + x_0) \\ &\leq p(x - x_0) + p(x + x_0), \end{aligned}$$

so that

$$G(x) - p(x - x_0) \leq p(x' + x_0) - G(x') \ \forall \ x, x' \in X'$$

which means that (5.2) can be satisfied for some α.

❑

We continue with a few easy corollaries.

Corollary 5.1.3 *Let $(X, |\,.\,|)$ be a normed vector space. Let Y be a linear subspace and let $g : Y \to \mathbf{R}$ be a linear continuous map with norm*

$$\|g\|_{Y'} := \sup_{x \in Y \, , \, |y| \leq 1} |g(x)|.$$

Then there is a linear continuous map $G \in X'$ so that $G|_Y \equiv g$ and

$$\|G\|_{X'} = \|g\|_{Y'}.$$

Proof:
Apply the Hahn–Banach theorem with $p(x) := \|g\|_{Y'}|x|$.

❑

Corollary 5.1.4 *Let $(X, |\,.\,|)$ be a normed vector space. Then for every $x_0 \in X$ there is some $f \in X'$ such that*

$$\|f\|_{X'} = |x_0| \text{ and } f(x_0) = |x_0|^2.$$

Proof:
Apply the previous corollary with $Y = \mathbf{R}\, x_0$ and $g(tx_0) = t|x_0|^2$ so that $\|g\|_{Y'} = |x_0|$.

❑

The following corollary characterizes the norm on X by the norm on the dual space:

Corollary 5.1.5 *Let $(X, |\,.\,|)$ be a normed vector space. For every $x \in X$*

$$|x| = \sup_{f \in X',\, \|f\| \leq 1} |f(x)| = \max_{f \in X',\, \|f\| \leq 1} |f(x)|.$$

Proof:
The assertion is clear if $x = 0$. Suppose that $x \neq 0$. Then

$$|f(x)| \leq \|f\|_{X'}\, |x| \leq |x|.$$

On the other hand there is $f \in X'$ such that $\|f\| = |x|$ and $f(x) = |x|^2$. Define now $g := |x|^{-1} f$ so that

$$\|g\|_{X'} = 1 \text{ and } |g(x)| = |x|.$$

❑

We mentioned earlier the following statement which we will now prove using the Hahn–Banach theorem:

Corollary 5.1.6 *Every finite dimensional subspace Y of a normed vector space X has a topological complement.*

Proof:
Assume that $\{y_1, \ldots, y_n\}$ is a basis for Y. Then by the Hahn–Banach theorem there are linear functionals $\ell_1, \ldots, \ell_n \in X'$ such that

$$\ell_j(y_k) = \begin{cases} 1 & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases}$$

(define ℓ_j on Y by the above formula, Hahn–Banach guarantees a continuous linear extension onto all of X). Because all the ℓ_j are continuous the spaces $\ker \ell_j \subset X$ are closed for all j. Then

$$Z := \bigcap_{1 \leq j \leq n} \ker \ell_j \subset X$$

is also a closed subspace, and it satisfies $X = Y \oplus Z$.

❑

The following theorem generalizes corollary 5.1.5

Theorem 5.1.7 *Let X be a (real) normed vector space, and let Y be a linear subspace. For any $z \in X$ we write*

$$d(z) := \inf_{y \in Y} |z - y|$$

for the distance between z and the set Y. We also define

$$D(z) := \max\{ |\ell(z)| \ : \ \ell \in X' \, , \, \|\ell\|_{X'} \leq 1 \, , \, \ell|_Y \equiv 0\}.$$

Then for every $z \in X$

$$d(z) = D(z).$$

Proof:
The assertion of the theorem is trivially true if $z \in Y$. We have for all $y \in Y$

$$|\ell(z)| = |\ell(z) - \ell(y)| = |\ell(z - y)| \leq |z - y|$$

and therefore

$$D(z) \leq d(z). \tag{5.3}$$

In order to show equality, we consider the linear space Z consisting of all elements of the form $y + \lambda z$, $y \in Y$, $\lambda \in \mathbf{R}$. We define a linear functional ℓ_Z on Z by $\ell_Z(y + \lambda z) := \lambda d(z)$. We have

$$|\ell_Z(y + \lambda z)| \leq |\lambda| \inf_{y \in Y} |z - y| = \inf_{y \in Y} |\lambda z - \lambda y| \leq |\lambda z + y|$$

so that

$$\|\ell_Z\|_{Z'} \leq 1.$$

We extend ℓ_Z to a linear map $L \in X'$ so that $\|L\|_{X'} \leq 1$ as well. If we set now $y = 0$ and $\lambda = 1$ then

$$L(z) = d(z) \geq D(z)$$

because of (5.3). On the other hand we also have $L(z) \leq D(z)$ by definition of $D(z)$. This shows that $D(z) = d(z)$.

□

Definition 5.1.8 *Let $(X, |\,.\,|)$ be a normed vector space and let $A \subset X$ be a subset. The closed linear span of A is the smallest closed linear subspace of X which contains A, i.e. it is the intersection of all closed linear subspaces containing A.*

The following theorem is a very valuable tool for deciding whether a given point $z \in X$ is contained in the closed linear span of a set A.

Theorem 5.1.9 *A point z in a normed vector space X belongs to the closed linear span of a subset A if and only if every $\ell \in X'$ which satisfies $\ell|_A \equiv 0$ also satisfies $\ell(z) = 0$.*

Proof:
If z is contained in the closed linear span of A, then there are elements $x_k \in X$ of the form

$$x_k = \sum_{j=1}^{n(k)} \alpha_{kj}\, a_{kj}\ ,\ a_{kj} \in A,\ \alpha_{kj} \in \mathbf{R}$$

so that $x_k \to z$. Now $\ell(x_k) = 0$ for all k and ℓ is continuous, hence $\ell(z) = 0$. In order to prove the converse statement, assume that z does not belong to the closed linear span Y of A. We have to find $\ell \in X'$ such that $\ell|_Y \equiv 0$ but $\ell(z) \neq 0$. We define $\tilde{Y}$ to be the linear subspace of X consisting of all elements of the form $y + \lambda\, z$ where $y \in Y$ and $\lambda \in \mathbf{R}$. We define $\tilde{\ell} \in \tilde{Y}'$ by

$$\tilde{\ell}(y + \lambda z) := \lambda.$$

We have

$$|y + \lambda z| = |\lambda| \left| \frac{y}{\lambda} + z \right| \geq c\,|\lambda|,$$

where

$$c = \inf_{y \in Y} |z - y| > 0.$$

This shows that indeed $\tilde{\ell} \in \tilde{Y}'$ with norm bounded by c^{-1}. We may extend $\tilde{\ell}$ to a linear functional $\ell \in X'$ by the Hahn–Banach theorem. Then $\ell(y) = 0$ for all $y \in Y$ and $\ell(z) = 1$.

□

We note the following corollary. This is how we prove that a subspace of a normed vector space is dense.

Corollary 5.1.10 *Assume that X is a normed vector space and that Y is a linear subspace such that $\overline{Y} \neq X$. Then there is $f \in X'$ such that $f \not\equiv 0$ but $f|_Y \equiv 0$.*

There is a geometric version of the Hahn–Banach theorem which we will discuss next. In the following let $(X, |\,.\,|)$ be a normed vector space over the real numbers.

Definition 5.1.11 *An (affine) hyperplane in X is a set of the form*

$$H := \{x \in X \mid f(x) = \lambda\},$$

where $f : X \to \mathbf{R}$ is a nontrivial linear map and $\lambda \in \mathbf{R}$. We will sometimes write shortly $\{f = \lambda\}$.

Lemma 5.1.12 *A hyperplane $H = \{f = \lambda\}$ is closed if and only f is continuous.*

Proof:
The hyperplane H is clearly closed if f is continuous. Conversely, if H is closed then its complement $X \backslash H$ is open and not empty (note that $f \not\equiv 0$). Pick a point x_0 in the complement and assume that $f(x_0) < \lambda$ (the case $f(x_0) > \lambda$ is handled similarly, we leave it as an exercise to the reader). Then pick an open ball $B_\varepsilon(x_0) \subset X \backslash H$ centered at x_0. We then have $f(x) < \lambda$ for all $x \in B_\varepsilon(x_0)$. Indeed, we can never have $f(x) = \lambda$ on $B_\varepsilon(x_0)$ because the ball is contained in the complement of H. On the other hand, $f(x_1) > \lambda$ for some $x_1 \in B_\varepsilon(x_0)$ is also impossible because $tx_1 + (1-t)x_0 \in B_\varepsilon(x_0)$ for all $0 \le t \le 1$ and $f(tx_1 + (1-t)x_0) = \lambda$ if

$$t = \frac{\lambda - f(x_0)}{f(x_1) - f(x_0)},$$

a contradiction. It follows that

$$f(x_0 + \varepsilon \frac{z}{|z|}) < \lambda \text{ for all } z \in X,$$

i.e. $f(z) < \frac{|z|}{\varepsilon}(\lambda - f(x_0))\ \forall\ z \in X$ and also

$$|f(z)| < |z|\, \frac{\lambda - f(x_0)}{\varepsilon}\ \forall\ z \in X$$

so that f is continuous

❑

Definition 5.1.13 *Let $A, B \subset X$ be subsets. We say that a hyperplane $H = \{f = \lambda\}$ separates the sets A and B if*

$$f(x) \le \lambda\ \forall\ x \in A \textit{ and } f(x) \ge \lambda\ \forall\ x \in B.$$

We say that H separates A and B in the strict sense if there is $\varepsilon > 0$ such that

$$f(x) \le \lambda - \varepsilon\ \forall\ x \in A \textit{ and } f(x) \ge \lambda + \varepsilon\ \forall\ x \in B.$$

Theorem 5.1.14 (Hahn–Banach theorem: First geometric version) *Assume that $A, B \subset X$ are not empty, convex and disjoint subsets of X. Assume moreover that A is open. Then there is a closed hyperplane which separates A and B.*

We start with two lemmas.

Lemma 5.1.15 *Let $C \subset X$ be an open convex set containing the origin. We define for $x \in X$*

$$\mu_C(x) := \inf\{\alpha > 0 \,|\, \frac{x}{\alpha} \in C\}.$$

The map $\mu_C : X \to [0, \infty)$ is called the gauge of C and it has the following properties:

1. $\mu_C(\lambda x) = \lambda\, \mu_C(x) \;\forall\; x \in X, \lambda > 0,$
2. $\mu_C(x + y) \leq \mu_C(x) + \mu_C(y) \;\forall\; x, y \in X,$
3. *there is $K > 0$ so that $0 \leq \mu_C(x) \leq K\,|x|$,*
4. $C = \{x \in X \,|\, \mu_C(x) < 1\}.$

Proof:
The first property is obvious. Let us start with property 4. Assume that $x \in C$. Since C is open we also have $(1 + \varepsilon)x \in C$ if ε is sufficiently small. Hence

$$\mu_C(x) \leq \frac{1}{1 + \varepsilon} < 1.$$

Now assume that $\mu_C(x) < 1$. Then there is a number $0 < \alpha < 1$ such that $x/\alpha \in C$. Then

$$x = \alpha(\frac{x}{\alpha}) + (1 - \alpha) \cdot 0$$

is also in C because C is convex and contains the origin. Let us now prove property 2. Pick $x, y \in X$ and $\varepsilon > 0$. By properties 1. we have

$$\mu_C\left(\frac{x}{\mu_C(x) + \varepsilon}\right) = \frac{\mu_C(x)}{\mu_C(x) + \varepsilon} < 1,$$

and property 4. then implies

$$\frac{x}{\mu_C(x) + \varepsilon} \in C.$$

By convexity of C we have

$$\tau \frac{x}{\mu_C(x) + \varepsilon} + (1 - \tau) \frac{y}{\mu_C(y) + \varepsilon} \in C \;\forall\; \tau \in [0, 1].$$

If we choose $\tau = (\mu_C(x) + \varepsilon)/(\mu_C(x) + \mu_C(y) + 2\varepsilon)$ we obtain in particular

$$\frac{x + y}{\mu_C(x) + \mu_C(y) + 2\varepsilon} \in C.$$

Properties 1. and 4. yield

$$\frac{\mu_C(x + y)}{\mu_C(x) + \mu_C(y) + 2\varepsilon} < 1$$

and $\mu_C(x+y) < \mu_C(x)+\mu_C(y)+2\varepsilon$ for all $\varepsilon > 0$ which implies property 2. We are left with property 3. Let $\varepsilon > 0$ so that $B_\varepsilon(0) \subset C$ (note that $0 \in C$ and C is open). Then

$$\mu_C(x) \leq \frac{1}{\varepsilon}|x|$$

by definition of μ_C, hence $K = 1/\varepsilon$.

❑

Lemma 5.1.16 *Let $\emptyset \neq C \subset X$ be an open convex set and let $x_0 \in X \backslash C$. Then there is a linear functional $f \in X'$ such that $f(x) < f(x_0)$ for all $x \in C$. This means that the hyperplane $\{f = f(x_0)\}$ separates the point $\{x_0\}$ and the convex set C.*

Proof:
We may assume by translation that the set C contains the origin. Consider the linear subspace $Y = \mathbf{R} \cdot x_0$ and the linear map $g : Y \to \mathbf{R}$ defined by

$$g(\lambda x_0) := \lambda.$$

We claim that for all $x \in Y$

$$g(x) \leq \mu_C(x).$$

Indeed, if $x = \lambda x_0$ with $\lambda > 0$ then by the previous lemma

$$\mu_C(x) = \lambda \mu_C(x_0) \geq \lambda = g(\lambda x_0) = g(x)$$

since $\mu_C(x_0) \geq 1$. If $\lambda \leq 0$ the inequality is trivially true because $\mu_C(x) \geq 0$ and $g(x) \leq 0$. By the Hahn–Banach theorem (analytic version) we can extend g to $f \in X'$ such that $f(x) \leq \mu_C(x)$ for all $x \in X$ and $f(x_0) = 1$. Property 4. of the previous lemma then implies that $f(x) < 1$ whenever $x \in C$, completing the proof.

❑

We can now proceed with the proof of the geometric version of the Hahn Banach theorem (theorem 5.1.14).

Proof:
Define $C := A - B := \{x \in X \,|\, x = a - b\,,\, a \in A\,,\, b \in B\}$. Then C is open since

$$C = \bigcup_{b \in B} A - b.$$

We leave it as an exercise to show that C is also convex. Moreover, $0 \notin C$ since $A \cap B = \emptyset$ by assumption. We translate the set C so that the translated set $C + x_0$ contains the origin. The set $C + x_0$ is of course still convex and open and it does not contain x_0. Using the lemma above we can find $f \in X'$ such that $f(\tilde{x}) < f(x_0)$ for all $\tilde{x} \in C + x_0$. We conclude by linearity of f that $f(x) < 0$ for all $x \in C$. Writing $x = a - b$ with $a \in A$ and $b \in B$ we get

$$f(a) < f(b) \;\forall\; a \in A\,,\, b \in B.$$

Choose now α so that $\sup_A f \leq \alpha \leq \inf_B f$, hence the hyperplane $\{f = \alpha\}$ separates the sets A and B.

❑

Theorem 5.1.17 (Hahn–Banach theorem: Second geometric version)
Let $A, B \subset X$ be non empty, convex disjoint subsets so that A is closed and B is compact. Then there is a closed hyperplane which separates A and B in the strict sense.

Proof:
If $\varepsilon > 0$ and $A_\varepsilon := A + B_\varepsilon(0)$, $B_\varepsilon := B + B_\varepsilon(0)$ then A_ε and B_ε are both open, not empty and convex. The notation $A + B_\varepsilon(0)$ refers to $\{x \in X \,|\, x = a + z\,,\, a \in A\,,\, |z| < \varepsilon\}$. If $\varepsilon > 0$ is sufficiently small then $A_\varepsilon \cap B_\varepsilon = \emptyset$. Indeed, if this were not true then we could find sequences $\varepsilon_k \searrow 0$, $x_k \in A$, $y_k \in B$ such that $|x_k - y_k| \leq 2\varepsilon_k$. Since B is compact by assumption, the sequence (y_k) has a convergent subsequence, hence we may assume that $y_k \to y \in B$. Then also $x_k \to y$ which has to be in A because A is closed by assumption. Hence $y \in A \cap B$, a contradiction since A and B are disjoint. Using the previous version of Hahn–Banach we can now separate the sets A_ε and B_ε by a closed hyperplane $\{f = \alpha\}$, i.e.

$$f(x + \varepsilon z) \leq \alpha \leq f(y + \varepsilon z) \;\forall\; x \in A\,,\, y \in B\,,\, |z| < 1.$$

This implies that

$$f(x) + \varepsilon \|f\|_{X'} \leq \alpha \leq f(y) - \varepsilon \|f\|_{X'} \;\forall\; x \in A\,,\, y \in B,$$

which is the assertion of the theorem.

❑

We note (without proof) the following complex version of the Hahn–Banach theorem

Theorem 5.1.18 (Hahn–Banach: complex version)
Let X be a vector space over the complex numbers and let $p : X \to \mathbf{R}$ be a map satisfying

- $p(\lambda x) = |\lambda|\, p(x)\ \forall\ x \in X, \lambda \in \mathbf{C}$,
- $p(x + y) \le p(x) + p(y)\ \forall\ x, y \in X$.

Let Y be a linear subspace of X, and let $g : Y \to \mathbf{C}$ be a linear map satisfying

$$|g(y)| \le p(y)\ \forall\ y \in Y.$$

Then there is a linear map $G : X \to \mathbf{C}$ so that $G|_Y \equiv g$ and $|G(x)| \le p(x)$ for all $x \in X$.

5.2 Reflexivity, the dual space of $L^p(\Omega)$

We introduce the important concept of reflexivity. We will use it here to characterize the dual space of $L^p(\Omega)$. We will explore it further in the following chapter about weak convergence. Let $(X, |\,.\,|)$ be a normed vector space over the real numbers. We denote by X'' the so–called bidual space which is the dual space of X'. There is a natural map from X into its bidual space

$$J_X : X \longrightarrow X''$$

$$J_X(x)\ell := \ell(x)\ ,\ \ \ell \in X'\,,\ x \in X.$$

We have

$$\|J_X(x)\|_{X''} = \sup_{\ell \in X'\,,\, \|\ell\|_{X'} \le 1} |\ell(x)| = |x|$$

by corollary 5.1.5, the dual characterization of the norm. Hence the map J is an isometry, in particular it is injective.

Exercise 5.2.1 *If $(X, |\,.\,|)$ is a Banach space then the range of J_X is closed in X''.*

Definition 5.2.2 *A Banach space $(X, |\,.\,|)$ is called reflexive if the natural map J_X above is surjective, i.e. J_X identifies the Banach space X with its bidual space X''.*

Remark:
It is important in the definition of reflexivity that the natural map $J_X : X \to X''$

is used for the identification. It is possible to construct an example of a non–reflexive Banach space so that there exists a surjective isometry from X into X'' (see R.C. James, A non reflexive Banach space isometric with its second conjugate space, Proc. Nat. Acad. Sci. USA, 37, (1951)).

So which Banach spaces are reflexive ? We state the following theorem which we will prove in the next chapter:

Theorem 5.2.3 (Milman)
Every uniformly convex Banach space is reflexive.

In particular, Hilbert spaces are reflexive. We have shown earlier that the spaces $L^p(\Omega)$ are uniformly convex if $2 \leq p < \infty$. Then they are also reflexive for $2 \leq p < \infty$ by Milman's theorem. We will show that $L^p(\Omega)$ is actually reflexive for $1 < p < \infty$. The Hahn–Banach theorem permits us to establish some properties of reflexivity:

Theorem 5.2.4 *A Banach space $(X, |\,.\,|)$ is reflexive if and only if its dual space X' is reflexive.*

Proof:
Assume that X is reflexive. We have to show that the map

$$J_{X'} : X' \longrightarrow X'''$$

is surjective. Pick $x''' \in X'''$. Define $\ell := x''' \circ J_X \in X'$. Take now $x'' \in X''$ which is of the form $J_X(x)$ for some $x \in X$ since X is reflexive. We compute

$$J_{X'}(\ell)x'' = x''(\ell) = x''(x''' \circ J_X) = J_X(x)(x''' \circ J_X) = x'''(J_X(x)) = x'''(x''),$$

hence $x''' = J_{X'}(\ell)$ and X' is also reflexive. Assume now that X' is reflexive, i.e. $J_{X'} : X' \to X'''$ is surjective. Arguing indirectly, we assume that X is not reflexive, hence there is $x'' \in X'' \backslash J_X(X)$. By the Hahn–Banach theorem (or one of its corollaries) there is $x''' \in X'''$ such that $x'''(x'') \neq 0$ and $x'''|_{J_X(X)} \equiv 0$. By reflexivity of X' we can find $\ell \in X'$ such that $J_{X'}(\ell) = x'''$. Then ℓ is non trivial and for all $y'' \in X''$ we have $x'''(y'') = J_{X'}(\ell)y'' = y''(\ell)$. On the other hand, $0 = x'''(J_X(x))$ for all $x \in X$, which implies $J_X(x)\ell = \ell(x) = 0$ for all $x \in X$, i.e. $\ell \equiv 0$, a contradiction.

❑

Theorem 5.2.5 *Every closed linear subspace of a reflexive Banach space is again a reflexive Banach space.*

Proof:
Let Y be a closed subspace of X. Consider the map

$$J_Y : Y \longrightarrow Y''.$$

Let $y'' \in Y''$ and let $i : Y \hookrightarrow X$ be the inclusion. We define an element $x'' \in X''$ by $x''(\ell) := y''(\ell \circ i)$ where $\ell \in X'$. Since X is reflexive we can find $x \in X$ such that $J_X(x) = x''$, i.e.

$$\ell(x) = J_X(x)\ell = y''(\ell \circ i) \ \forall \ \ell \in X'.$$

We claim now that $x \in Y$. If we had $x \notin Y$ then we could find by the Hahn Banach theorem some $x' \in X'$ so that $x'(x) \neq 0$ but $x'|_Y \equiv 0$. This is a contradiction since

$$x'(x) \neq 0 \text{ but } x' \circ i \equiv 0.$$

Hence $x \in Y$. It remains to show that $J_Y(x) = y''$. Every bounded linear functional $\ell_Y \in Y'$ can be extended to a bounded linear functional on X, which we denote by ℓ. We have for all $y \in Y$, using $x'' = J_X(x)$ and $x''(\ell) = y''(\ell \circ i)$

$$J_Y(x)\ell_Y = \ell_Y(x) = \ell(x) = x''(\ell) = y''(\ell \circ i) = y''(\ell_Y),$$

which holds for all $\ell_Y \in Y'$.

❑

Using Milman's theorem we will characterize the dual space of $L^p(\Omega)$.

Theorem 5.2.6 *The dual space of $L^p(\Omega)$, $1 < p < \infty$, is $L^q(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$.*

Proof:
Let us assume first that $2 \leq p < \infty$. In this case we have shown that $L^p(\Omega)$ is uniformly convex (uniform convexity also holds for $1 < p < 2$, we only mentioned it without proof). We define an operator

$$T : L^q(\Omega) \longrightarrow (L^p(\Omega))'$$

by

$$(Tu)v := \int_\Omega u(x)v(x)\,dx \ , u \in L^q(\Omega)\, , \, v \in L^p(\Omega),$$

which is well–defined by Hölder's inequality. Also by Hölder's inequality

$$\|Tu\|_{(L^p(\Omega))'} = \sup_{\|v\|_{L^p(\Omega)} \leq 1} |(Tu)v| \leq \|u\|_{L^q(\Omega)}.$$

On the other hand, if we define

$$f(x) := \begin{cases} 0 & \text{if} \quad u(x) = 0 \\ |u(x)|^{q-2}u(x) & \text{if} \quad u(x) \neq 0 \end{cases}$$

then $f \in L^p(\Omega)$ since

$$\int_\Omega |f(x)|^p dx = \int_\Omega |u(x)|^{p(q-1)} dx = \int_\Omega |u(x)|^q dx,$$

i.e.

$$\|f\|_{L^p(\Omega)} = \|u\|_{L^q(\Omega)}^{q-1}.$$

Moreover,

$$(Tu)f = \int_\Omega f(x)u(x)dx = \|u\|_{L^q(\Omega)}^q$$

so that

$$\|u\|_{L^q(\Omega)} = \frac{(Tu)f}{\|f\|_{L^p(\Omega)}} \leq \|Tu\|_{(L^p(\Omega))'}.$$

This is true for all $u \in L^q(\Omega)$, hence

$$\|Tu\|_{(L^p(\Omega))'} = \|u\|_{L^q(\Omega)},$$

and the map T is an isometry. We claim that it is surjective as well. Because T is an isometry, the space $T(L^q(\Omega)) \subset (L^p(\Omega))'$ is closed. We have to show that it is also dense. This is the same (as a consequence of the Hahn Banach theorem) as showing that every $\phi \in (L^p(\Omega))''$ which satisfies $\phi|_{T(L^q(\Omega))} \equiv 0$ must be trivial. By uniform convexity and Milman's theorem, the space $L^p(\Omega)$ is reflexive, hence every $\phi \in (L^p(\Omega))''$ can be written as $J_{L^p(\Omega)}(h)$ for a suitable $h \in L^p(\Omega)$. If $u \in L^q(\Omega)$ we conclude

$$0 = \phi(Tu) = \int_\Omega u(x)h(x)\,dx \ \forall\ u \in L^q(\Omega).$$

This implies that $h \equiv 0$ almost everywhere, for example by choosing $u = |h|^{p-2}h$, which implies $\phi = 0$ so that $L^q(\Omega)$ and $(L^p(\Omega))'$ are isometrically isomorphic via the map T if $2 \leq p < \infty$. Recall that a Banach space is reflexive if and only if its dual space is. Because T is an isometry, the spaces $L^p(\Omega)$ are also reflexive for $1 < p < 2$. Then the above proof also works for $1 < p < 2$.

❑

Exercise 5.2.7 *Let $m \geq 1$ and $1 < p < \infty$. Show that the Sobolev space $W^{m,p}(\Omega)$ is reflexive. Hint: Identify $W^{m,p}(\Omega)$ with a closed subspace of the $(m+1)$–fold product of $L^p(\Omega)$. You should then also show that $\times_{m+1} L^p(\Omega)$ is reflexive.*

Recall that the space $C^0([-1,1])$ furnished with the maximum–norm is not uniformly convex as we have shown earlier. It even fails to be reflexive.

Theorem 5.2.8 *The space $C^0([-1,1])$ furnished with the maximum–norm is not reflexive.*

Proof:
If $C^0([-1,1])$ were reflexive then we could identify it with its bidual space via the isometry

$$J_{C^0([-1,1])} : C^0([-1,1]) \longrightarrow (C^0([-1,1]))''$$

$$J_{C^0([-1,1])}(f)\ell = \ell(f).$$

Let $\ell \in (C^0([-1,1]))'$. By the dual characterization of the norm we have

$$\|\ell\|_{(C^0([-1,1]))'} = \max_{\phi \in (C^0([-1,1]))'',\, \|\phi\|_{(C^0([-1,1]))''}=1} |\phi(\ell)|.$$

Then there is $f \in C^0([-1,1])$ such that

$$\|\ell\|_{(C^0([-1,1]))'} = \ell(f) \text{ and } \|f\|_{C^0([-1,1])} = 1.$$

Define now $\ell \in (C^0([-1,1]))'$ by

$$\ell(g) := \int_{-1}^0 g(x)dx - \int_0^1 g(x)dx$$

so that for every $g \in C^0([-1,1])$

$$|\ell(g)| \leq 2\,|g|_{C^0([-1,1])}.$$

On the other hand, for any $\varepsilon > 0$ we can find a bounded continuous function g on $[-1,1]$ so that

$$|\ell(g)| > (2-\varepsilon)\,|g|_{C^0([-1,1])}.$$

This shows that $\|\ell\|_{C^0([-1,1]))'} = 2$. For $g = f$ we now obtain a contradiction since there is no continuous (!) function with $\sup|f| = 1$ and

$$\ell(f) = \int_{-1}^0 f(x)dx - \int_0^1 f(x)dx = 2.$$

❑

We conclude this section with some remarks: Theorem 5.2.6 is also called the Riesz representation theorem. In a similar spirit one can show the following theorem

Theorem 5.2.9 *Let $\phi \in (L^1(\Omega))'$. Then there is $u \in L^\infty(\Omega)$ such that*

$$\phi(v) = \int_\Omega u(x)v(x)dx \;\forall\; v \in L^1(\Omega).$$

Moreover, $\|u\|_{L^\infty(\Omega)} = \|\phi\|_{(L^1(\Omega))'}$. Hence we may identify the dual space of $L^1(\Omega)$ with $L^\infty(\Omega)$.

For a proof, see the book by H. Brezis. The space $L^1(\Omega)$ is not reflexive (we can prove this in the next chapter). Then by the above theorem $L^\infty(\Omega)$ is also not reflexive. In fact, the dual space of $L^\infty(\Omega)$ contains $L^1(\Omega)$, but it is strictly larger than $L^1(\Omega)$. For a description of $(L^\infty(\Omega))'$ see the book by Yosida p. 118. The following is a nice exercise related to the Hahn Banach theorem

Exercise 5.2.10 *Show that there is $\phi \in (L^\infty(\Omega))'$ so that there is no $u \in L^1(\Omega)$ satisfying*

$$\phi(f) = \int_\Omega u(x)f(x)dx \ \forall f \in L^\infty(\Omega).$$

Here is another way to see that $C^0([-1,1])$ is not reflexive.

Definition 5.2.11 *A normed vector space $(X, |\,.\,|)$ is called separable if it contains a dense countable set.*

Theorem 5.2.12 *Let $(X, |\,.\,|)$ be a normed vector space. If its dual space $(X', |\,.\,|_{X'})$ is separable then so is $(X, |\,.\,|)$.*

Proof:
There is a sequence $(\ell_k)_{k\in\mathbf{N}}$ which is dense in X'. We can find a sequence $(x_k)_{k\in\mathbf{N}} \subset X$ such that

$$|x_k| = 1 \text{ and } \ell_k(x_k) > \frac{1}{2}|\ell_k|_{X'}$$

by the definition of the norm on X'. We claim that the closed linear span of the set (x_k) is all of X. Suppose this is not true. Then we can find $\ell \in X'$ such that $\ell(x_k) = 0$ for all k but $|\ell|_{X'} = 1$. The sequence (ℓ_k) is dense in X', hence there is some ℓ_k such that

$$|\ell - \ell_k|_{X'} < \frac{1}{3}.$$

The norm of ℓ equals 1, therefore

$$|\ell_k|_{X'} > \frac{2}{3}.$$

We arrive at the following contradiction:

$$\frac{1}{3} > |\ell(x_k) - \ell_k(x_k)| = |\ell_k(x_k)| > \frac{1}{2}|\ell_k|_{X'} > \frac{1}{3}.$$

Hence there is no such ℓ and the closed linear span is all of X. This means that the set of all finite linear combinations of elements in $\{x_k\}$ is dense. Then the set of all finite linear combinations of elements in $\{x_k\}$ with rational coefficients is also dense in X, but this is a countable set. Hence X is separable.

❑

The Banach space $C^0([-1,1])$ (with maximum–norm) is clearly separable: Every continuous function can be approximated by piecewise linear functions with rational data (rational nodes, slope). The dual space $(C^0([-1,1]))'$ however, is not separable: Define linear functionals $(\ell_t)_{-1\le t\le 1}$ by

$$\ell_t(f) := f(t).$$

We have

$$|\ell_t|_{(C^0([-1,1]))'} \le 1 \text{ and } |\ell_{t'} - \ell_t|_{(C^0([-1,1]))'} = 2 \text{ if } t \neq t'.$$

So we have found a non–countable set in $(C^0([-1,1]))'$ where two distinct elements have distance 2 from each other. Consider the set $U = \bigcup_{-1\le t\le 1} B_{1/2}(\ell_t)$ which is a non–countable union of pairwise disjoint balls. If $(C^0([-1,1]))'$ were separable with dense set $D = (x'_k)_{k\in\mathbf{N}} \subset (C^0([-1,1]))'$ then $D \cap B_{1/2}(\ell_t) \neq \emptyset$ for all t. We define a map

$$\Phi : D \cap U \longrightarrow [-1,1]$$

so that $\Phi(x'_k)$ is the number t such that $x'_k \in B_{1/2}(\ell_t)$. This map is surjective, but this is not possible since $D \cap U$ is countable but $[-1,1]$ is not. This shows that $(C^0([-1,1]))'$ is not separable.
If $C^0([-1,1])$ were reflexive then $C^0([-1,1])$ and $(C^0([-1,1]))''$ are isometrically isomorphic, in particular, $(C^0([-1,1]))''$ would be separable. Then $(C^0([-1,1]))'$ would also be separable by the above theorem which is a contradiction.
This type of argument also works for L^p–spaces after the following exercise:

Exercise 5.2.13 *Show that $L^p(\Omega)$ is separable if $1 \le p < \infty$, but not for $p = \infty$.*

5.3 Application: Existence of a Green's function for the Laplace operator

Let $\Omega \subset \mathbf{R}^2$ be a bounded domain with C^1–boundary. Let $f \in C^0(\overline{\Omega})$ and $\phi \in C^2(\partial\Omega)$. Assume that $u \in C^2(\overline{\Omega})$ is a solution of the boundary value problem

$$\begin{aligned} \Delta u(p) &= f(p) \text{ for } p \in \Omega, \\ u(p) &= \phi(p) \text{ for } p \in \partial\Omega. \end{aligned}$$

We would like to find a function $G(p,q)$, defined for $p \in \overline{\Omega}$, $q \in \Omega$, $p \neq q$, so that

$$G(p,q) = 0 \text{ for } p \in \partial\Omega$$

and for $q \in \Omega$

$$u(q) = \int_{\partial\Omega} \phi(p) \frac{\partial G(p,q)}{\partial \nu}\, ds + \int_\Omega G(p,q)\, f(p)\, dp, \tag{5.4}$$

where $\frac{\partial}{\partial\nu}$ denotes the outer normal derivative with respect to the p–variable, and all integration is with respect to the p–variable. We try the following approach: We write

$$G(p,q) := -\frac{1}{2\pi} \log|p-q| + g_0(p,q),$$

where g_0 is defined on $\overline{\Omega} \times \Omega$. For $q \in \Omega$ we want g_0 to satisfy the following boundary value problem

$$\Delta g_0(p,q) \;=\; 0 \text{ for } p \in \Omega, \tag{5.5}$$
$$g_0(p,q) = \frac{1}{2\pi} \log |p-q| \text{ for } p \in \partial\Omega.$$

Here, Δ is the Laplace operator with respect to the p–variable. If we can succeed to solve the boundary value problem (5.5) then trivially $G(p,q) = 0$ if $p \in \partial\Omega$ and the representation formula (5.4) also holds.

Exercise 5.3.1 *Prove that (5.4) holds. Fixing* $q \in \Omega$ *and defining* $v(p) := -\frac{1}{2\pi} \log |p-q|$ *use Green's second identity*

$$\int_{\Omega_\varepsilon} (v(p)\Delta u(p) - u(p)\Delta v(p))dp = \int_{\partial\Omega_\varepsilon} \left(v(p)\frac{\partial u}{\partial \nu}(p) - u(p)\frac{\partial v}{\partial \nu}(p) \right) ds$$

on the domain $\Omega_\varepsilon := \Omega \backslash \overline{B_\varepsilon(q)}$ *(or read the book by D. Gilbarg and N. Trudinger, pp. 17–19).*

We will use the Hahn–Banach theorem in order to show that we can solve the boundary value problem (5.5). Denote by $\mathcal{C}$ the space of (real–valued) continuous functions on $\partial\Omega$ endowed with the maximum–norm. Define now the following linear subspace of $\mathcal{C}$:

$$\mathcal{H} := \{h \in \mathcal{C} \,|\, \exists\, H \in C^2(\Omega) \cap C^0(\overline{\Omega}) \;:\; \Delta H = 0\,,\; H|_{\partial\Omega} \equiv h\}.$$

Fix now a point $q \in \Omega$. Define now a linear functional $\ell_q : \mathcal{H} \to \mathbf{R}$ as follows:

$$\ell_q(h) := H(q).$$

Harmonic functions v on a bounded domain Ω satisfy the maximum and minimum principle

$$\inf_{\partial\Omega} v \le v(x) \le \sup_{\partial\Omega} v \;\forall\; x \in \Omega$$

(see D. Gilbarg, N. Trudinger, p. 15 for a proof). This implies in particular that H is uniquely determined by its boundary condition h, so ℓ_q is well–defined. The inequality

$$\ell_q(h) = H(q) \le \max_{p \in \partial\Omega} |h(p)| = |h|$$

can be read as follows: The functional ℓ_q is in $\mathcal{H}'$ and its norm is bounded by 1. The Hahn Banach theorem implies that ℓ_q can be extended from $\mathcal{H}$ to $\mathcal{C}$ so that its norm is still bounded by 1. If $w \in \mathbf{R}^2 \backslash \partial\Omega$ then we define $k(w) \in \mathcal{C}$ by

$$k(p,w) := \frac{1}{2\pi} \log |p-w| \text{ where } p \in \partial\Omega.$$

We observe that k depends differentiably on the parameter w, and k viewed as a function of w is harmonic in $\mathbf{R}^2 \backslash \partial\Omega$. Moreover, if $w \notin \overline{\Omega}$ then $k(w) \in \mathcal{H}$ (just permit $p \in \Omega$ in this case). We now define a function $g(w,q)$ by

$$g(w,q) := \ell_q(k(w)).$$

Lemma 5.3.2 *1. The function $w \mapsto g(w,q)$ is harmonic on $\mathbf{R}^2 \backslash \partial\Omega$.*

2. If $w \notin \overline{\Omega}$ then

$$g(w,q) = \frac{1}{2\pi} \log |q-w|.$$

3. The function $w \mapsto g(w,q)$ is continuous as w crosses the boundary of Ω, i.e. $w \mapsto g(w,q)$ can be extended continuously onto all of $\mathbf{R}^2$.

Before we prove the lemma we remark that this implies the existence of the Green's function G, take $g_0 = g$.

Proof:
We compute using the linearity of ℓ_q

$$\frac{1}{\varepsilon}(g(w+\varepsilon u,q) - g(w,q)) = \ell_q\left(\frac{1}{\varepsilon}(k(w+\varepsilon u) - k(w))\right)$$

We pass to the limit $\varepsilon \to 0$ and we use the fact that ℓ_q is continuous so that

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} g(w+\varepsilon u,q) = \ell_q\left(\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} k(w+\varepsilon u)\right).$$

Denoting the Laplace operator in the w–variable by Δ and using the above property for second derivatives we obtain

$$\Delta g(w,q) = \ell_q(\Delta k(w)) = 0$$

because $k(w)$ is harmonic in w. This proves the first statement. If $w \notin \overline{\Omega}$ then $k(w) \in \mathcal{H}$ and the original definition of ℓ_q can be used, i.e. $\ell_q(h) = H(q)$. Then we get

$$g(w,q) = \ell_q(k(w)) = k(q,w) = \frac{1}{2\pi}\log|q-w|.$$

If $w \in \Omega$ is a point close to $\partial\Omega$ then let $w_0 \in \partial\Omega$ be the point on the boundary closest to w. Then choose $w' \in \mathbf{R}^2 \backslash \overline{\Omega}$ such that $(w+w')/2 = w_0$. The point w' is the reflection of w at the boundary. By definition of the function g and by linearity of ℓ_q we compute

$$g(w,q) - g(w',q) = \ell_q(k(w) - k(w')) = \ell_q\left(\frac{1}{2\pi}\log\frac{|*-w|}{|*-w'|}\right),$$

where $*$ stands for a point on $\partial\Omega$. Remember that $k(w), k(w')$ are functions defined on $\partial\Omega$. We assumed that $\partial\Omega$ has a C^1–boundary. Then the tangents to $\partial\Omega$ at $p \in \partial\Omega$ depend continuously on p and we get

$$\frac{|p-w|}{|p-w'|} \longrightarrow 1$$

uniformly in $p \in \partial\Omega$ as $\operatorname{dist}(w, \partial\Omega) \to 0$. Then

$$\sup_{p \in \partial\Omega} \log \frac{|p-w|}{|p-w'|} \longrightarrow 0$$

as w approaches the boundary. We have shown in the previous step that $g(w', q) = \frac{1}{2\pi} \log|w' - q|$, hence we conclude that

$$\lim_{w \to p \in \partial\Omega} g(w, q) = \frac{1}{2\pi} \log|q - p|,$$

which completes the proof of the lemma.

❑

Chapter 6

Weak and weak* convergence

6.1 Weak and Weak* convergence

In this chapter we will introduce weaker notions of convergence on a normed vector space $(X, |\,.\,|)$ and its dual space X' which will have nicer compactness properties than the norm–convergence.

Definition 6.1.1 *Let $(X, |\,.\,|)$ be a normed vector space. Let $x \in X$ and let $(x_n)_{n\in\mathbf{N}} \subset X$ be a sequence. We say that (x_n) converges weakly to x (and we write $x_n \rightharpoonup x$) if*

$$\phi(x_n) \to \phi(x) \;\forall\; \phi \in X'.$$

On the dual space X' of a normed vector space $(X, |\,.\,|)$ we consider the following notions of convergence:

1. 'Strong convergence', i.e. $\ell_n \to \ell$ if $\|\ell_n - \ell\|_{X'} \to 0$,
2. 'Weak convergence', i.e. $\ell_n \rightharpoonup \ell$ if $\phi(\ell_n) \to \phi(\ell)$ for all $\phi \in X''$,
3. 'Weak* convergence', i.e. $\ell_n \overset{*}{\rightharpoonup} \ell$ if $\phi(\ell_n) \to \phi(\ell)$ for all $\phi \in J_X(X) \subset X''$, where $J_X : X \to X''$ is the natural isometry.

By definition of J_X, weak* convergence $\ell_n \overset{*}{\rightharpoonup} \ell$ just means that

$$\ell_n(x) \longrightarrow \ell(x) \;\forall\; x \in X.$$

The notion of weak* convergence on X' is weaker than the notion of weak convergence on X'. On reflexive Banach spaces, however, these two notions coincide.

Remarks:

- The weak limit of a sequence is unique. Indeed, if we had $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$ with $x \neq y$ then we could separate the sets $\{x\}$ and $\{y\}$ in the strict sense (geometric Hahn Banach theorem), and we would obtain a contradiction,
- If $x_n \to x$, i.e. $|x_n - x| \to 0$ ('strong convergence') then also $x_n \rightharpoonup x$. The converse is true if X is finite dimensional. In infinite dimensions weak convergence usually does not imply strong convergence, but there are exceptions: In the space l^1 every weakly convergent sequence also converges strongly. Such examples however, should be regarded as pathological.

The following proposition shows that weak convergence still has some properties of norm–convergence: A weakly convergent sequence is bounded, and the norm is lower–semi–continuous with respect to weak convergence. In the case of strong convergence we have continuity of the norm: $x_n \to x$ implies $|x_n| \to |x|$.

Proposition 6.1.2 *Let $(x_n)_{n\in\mathbf{N}} \subset X$ be a sequence with $x_n \rightharpoonup x$ for some $x \in X$. Then $(|x_n|)_{n\in\mathbf{N}}$ is a bounded sequence and*

$$|x| \leq \liminf_{n\to\infty} |x_n|.$$

Proof:
Since $\phi(x_n) \to \phi(x)$ for all $\phi \in X'$ there are constants $c_\phi > 0$ such that $|\phi(x_n)| \leq c_\phi$ for all $n \in \mathbf{N}$. Consider now the natural isometry

$$J_X : X \longrightarrow X'' \ , \ J_X(x)\phi := \phi(x)$$

so that

$$|J_X(x_n)\phi| \leq c_\phi \ \forall \ n \in \mathbf{N} \, , \ \phi \in X'.$$

We can apply the Banach–Steinhaus theorem to the family

$$(J_X(x_n) : X' \to \mathbf{R})_{n\in\mathbf{N}},$$

and we obtain existence of a positive constant c such that

$$|x_n| = \|J_X(x_n)\|_{X''} \leq c \ \forall \ n \in \mathbf{N}.$$

This proves the first assertion of the proposition. By the Hahn Banach theorem there is $\phi \in X'$ such that $\phi(x) = |x|$ and $\|\phi\|_{X'} = 1$. We conclude

$$|x| = \phi(x) = \lim_{n\to\infty} \phi(x_n) = \liminf_{n\to\infty} \phi(x_n) \leq \liminf_{n\to\infty} \|\phi\|_{X'}|x_n| = \liminf_{n\to\infty} |x_n|.$$

❑

Here is the weak* version of the above proposition. Since the proof is very similar, we leave it as an exercise.

Proposition 6.1.3 *Let $(\ell_n)_{n\in\mathbf{N}} \subset X'$ be a sequence with $\ell_n \overset{*}{\rightharpoonup} \ell$ for some $\ell \in X'$. Then $(\|\ell_n\|_{X'})_{n\in\mathbf{N}}$ is a bounded sequence and*

$$\|\ell\|_{X'} \leq \liminf_{n\to\infty} \|\ell_n\|_{X'}.$$

Definition 6.1.4 *Let $(X, |\,.\,|)$ be a normed vector space. We say that a subset $A \subset X$ is weakly closed if for every sequence $(x_n)_{n\in\mathbf{N}} \subset A$ with $x_n \rightharpoonup x$ we have $x \in A$. The weak closure of a set A, which we denote by $\overline{A}^w$, is defined by*

$$\overline{A}^w := \{x \in X \,|\, \exists (x_n)_{n\in\mathbf{N}} \subset A \,:\, x_n \rightharpoonup x\}.$$

The following proposition characterizes the weak closure of a convex subset $A \subset X$.

Proposition 6.1.5 *Let $(X, |\,.\,|)$ be a normed vector space and let $A \subset X$ be convex. Then*

$$\overline{A} = \overline{A}^w.$$

Proof:
We note that $\overline{A} \subset \overline{A}^w$ without the convexity assumption on A. Indeed, let $x \in \overline{A}$. Then there is a sequence x_n in A which converges to x strongly. Since strong convergence also implies weak convergence, we also have $x \in \overline{A}^w$.
As for the reverse direction, assume that $x_0 \notin \overline{A}$ and show that also $x_0 \notin \overline{A}^w$. By the Hahn Banach theorem (second geometric version) we can separate the sets $\{x_0\}$ and $\overline{A}$ in the strict sense, i.e. we can find $\phi \in X'$, $\gamma \in \mathbf{R}$ and $\varepsilon > 0$ such that

$$\phi(x_0) \leq \gamma - \varepsilon \text{ and } \phi(x) \geq \gamma + \varepsilon \;\forall\; x \in \overline{A}$$

(if X is a complex normed vector space we have instead $\mathrm{Re}(\phi(x_0)) \leq \gamma - \varepsilon$ and $\mathrm{Re}(\phi(x)) \geq \gamma + \varepsilon \;\forall\; x \in \overline{A}$). But then there can not be any sequence $(x_n)_{n\in\mathbf{N}} \subset A$ with $x_n \rightharpoonup x_0$, hence $x_0 \notin \overline{A}^w$.

❑

6.2 Weak sequential compactness

We know that the closed unit ball in an infinite dimensional normed vector space is never compact, i.e. bounded sequences usually do not have convergent subsequences. If we relax the notion of convergence to weak or weak* convergence then the situation looks much better.

Definition 6.2.1 *Let $(X, |\,.\,|)$ be a normed vector space and let $(X', \|\,.\,\|_{X'})$ be its dual space. A subset $M \subset X$ (or $M \subset X'$) is called weakly (or weakly*) sequential compact if every sequence in M has a subsequence which converges weakly (or weakly*) to some limit in M.*

Theorem 6.2.2 *Let X be a separable Banach space. Then the closed unit ball in the dual space*

$$\overline{B_1(0)} := \{\ell \in X' \mid \|\ell\|_{X'} \le 1\}$$

is weakly* sequential compact.

Proof:
Let $\{x_n\}_{n\in\mathbf{N}} \subset X$ be a dense countable set and let $(\ell_k)_{k\in\mathbf{N}} \subset X'$ be a sequence with $\|\ell_k\|_{X'} \le 1$. We have to show that there is a subsequence ℓ_{k_l} which is weakly* convergent. We have for every $n \in \mathbf{N}$

$$|\ell_k(x_n)| \le \|\ell_k\|_{X'}|x_n| \le |x_n| < \infty.$$

Then the sequence (ℓ_k) has a subsequence, which we denote by $(\ell_k^1)_{k\in\mathbf{N}}$, so that $(\ell_k^1(x_1))_{k\in\mathbf{N}} \subset \mathbf{R}$ converges. We can then extract another subsequence from $(\ell_k^1)_{k\in\mathbf{N}}$, call it $(\ell_k^2)_{k\in\mathbf{N}}$, so that $(\ell_k^2(x_2))_{k\in\mathbf{N}} \subset \mathbf{R}$ converges. We iterate this procedure and take the diagonal sequence $(\ell_k^k)_{k\in\mathbf{N}}$ which has the property that

$$\lim_{k\to\infty} \ell_k^k(x_n) =: \ell(x_n)$$

exists for all $n \in \mathbf{N}$. We extend ℓ as a linear map onto the linear span of the set $Z = \{x_n\}_{n\in\mathbf{N}}$. We have for $z \in Z$

$$|\ell(z)| \le \liminf_{k\to\infty} \|\ell_k\|_{X'}|z| \le |z|,$$

so that ℓ extends continuously to the closed linear span of $\{x_n\}_{n\in\mathbf{N}}$ which is all of X. Let now $x \in X$ and $\varepsilon > 0$. Then we can find a sequence $(z_l) \subset Z$ such that $z_l \to x$ as $l \to \infty$. We estimate

$$\begin{aligned}
|\ell_k^k(x) - \ell(x)| &= |(\ell_k^k - \ell)(x)| \\
&\le |(\ell_k^k - \ell)(x - z_l)| + |(\ell_k^k - \ell)(z_l)| \\
&\le \|\ell_k^k - \ell\|_{X'}|x - z_l| + |(\ell_k^k - \ell)(z_l)| \\
&\le 2\,|x - z_l| + |(\ell_k^k - \ell)(z_l)| \\
&\le 3\,\varepsilon
\end{aligned}$$

where $l \ge l(\varepsilon)$ such that $|x - z_l| \le \varepsilon$ and then $k \ge k(\varepsilon, l)$ such that $|(\ell_k^k - \ell)(z_l)| \le \varepsilon$. This shows weak* convergence.

❑

As an example consider $X = L^1(\Omega)$ which is separable and $X' = L^\infty(\Omega)$. In this concrete example the above theorem implies the following: Let $(u_k)_{k\in\mathbf{N}} \subset L^\infty(\Omega)$ be a bounded sequence. Then there is a subsequence $(u_{k_l})_{l\in\mathbf{N}}$ and $u \in L^\infty(\Omega)$ such that

$$\int_\Omega u_{k_l}(x)g(x)dx \overset{l\to\infty}{\longrightarrow} \int_\Omega u(x)g(x)dx \;\forall\; g \in L^1(\Omega).$$

What is the corresponding statement for a bounded sequence in $L^p(\Omega)$ with $1 < p < \infty$? Next we will see that the closed unit ball in $(L^\infty([0,1]))'$ is not weakly* sequential compact, in particular $L^\infty([0,1])$ is not separable. Consider $\ell_t \in (L^\infty([0,1]))'$, $0 < t \le 1$, with

$$\ell_t(f) := \frac{1}{t}\int_0^t f(x)dx \;,\; f \in L^\infty([0,1]).$$

We have

$$\|\ell_t\|_{(L^\infty([0,1]))'} = \sup_{\|f\|_{L^\infty([0,1])}\le 1} |\ell_t(f)| \le 1.$$

Assume now that there is a sequence $t_k \searrow 0$ such that $\ell_{t_k} \overset{*}{\rightharpoonup} \ell$ as $k \to \infty$. By passing to a suitable subsequence of t_k we may assume that the ratios t_{k+1}/t_k converge to zero. Define now

$$f := \sum_k (-1)^k \chi_{[t_{k+1},t_k)} \in L^\infty([0,1]),$$

where $\chi_{[t_{k+1},t_k)}$ denotes the characteristic function of the interval $[t_{k+1}, t_k)$. We compute

$$\begin{aligned}
\ell_{t_k}(f) &= \frac{1}{t_k}\sum_{l=k}^{\infty}(-1)^l(t_l - t_{l+1}) \\
&= (-1)^k\frac{t_k - t_{k+1}}{t_k} + \frac{1}{t_k}\sum_{l=k+1}^{\infty}(-1)^l(t_l - t_{l+1}) \\
&= (-1)^k\frac{t_k - t_{k+1}}{t_k} + \frac{t_{k+1}}{t_k}\ell_{t_{k+1}}(f).
\end{aligned}$$

We conclude that

$$|\ell_{t_k}(f) - (-1)^k| \le \frac{t_{k+1}}{t_k}|(-1)^k - \ell_{t_{k+1}}(f)| \le \frac{2\,t_{k+1}}{t_k} \to 0,$$

so that the sequence $(\ell_{t_k}(f))_{k\in\mathbf{N}}$ has the two accumulation points $+1$ and -1, hence $(\ell_{t_k})_{k\in\mathbf{N}}$ can not be weak* convergent.
In the case of a reflexive Banach space bounded sequences have weakly convergent subsequences.

Theorem 6.2.3 *Let X be a reflexive Banach space. Then the closed unit ball in X is weakly sequential compact.*

Proof:
Let us start with a sequence $(x_k)_{k\in\mathbf{N}} \subset X$ so that $|x_k| \leq 1$. Let Z be the closed linear span of the set $\{x_k\}_{k\in\mathbf{N}}$. The space Z is a closed linear subspace of X which is also separable and reflexive (remember that closed linear subspaces of reflexive Banach spaces are reflexive). By reflexivity Z is isometrically isomorphic to Z'' which is then also separable. This implies that Z' is separable. Then the closed unit ball in Z'' is weakly* sequential compact. This means that after passing to a suitable subsequence, the bounded sequence $J_Z(x_k) \subset Z''$ converges in weak* to some $J_Z(x) \in Z''$, i.e.

$$J_Z(x_k)\ell = \ell(x_k) \longrightarrow J_Z(x)\ell = \ell(x) \;\forall\; \ell \in Z'.$$

Since every $\ell \in X'$ is also contained in Z' by restriction to Z, we obtain

$$\ell(x_k) \longrightarrow \ell(x) \;\forall\; \ell \in X',$$

i.e. $x_k \rightharpoonup x$.

❑

6.3 Lower semi–continuity and convexity

Definition 6.3.1 *A function $F : X \supset A \to \mathbf{R}$ defined on a subset A of a normed vector space is called weakly sequential lower semi–continuous in the point $x \in A$ if for every sequence $(x_n)_{n\in\mathbf{N}} \subset A$ with $x_n \rightharpoonup x$ we have*

$$F(x) \leq \liminf_{n\to\infty} F(x_n).$$

If F is weakly sequential lower semi–continuous for every $x \in A$ then we say it has 'property (W)'.

Example (for lower semi–continuous):
In this example $X = \mathbf{R}$ so that weak convergence and strong convergence are the same. We define functions F_1, F_2 by

$$F_1(x) := \begin{cases} x & \text{if } \; x < 1 \\ x-1 & \text{if } \; x \geq 1 \end{cases}$$

$$F_2(x) := \begin{cases} x & \text{if } \; x \leq 1 \\ x-1 & \text{if } \; x > 1 \end{cases}$$

The function F_1 is lower semi–continuous in the point $x = 1$, but F_2 is not: If (x_n) is a sequence converging to 1 so that the sign of $x_n - 1$ alternates then the sequences $F_1(x_n)$ and $F_2(x_n)$ do not converge, they have accumulation points at 1 and 0. The limit inferior is the smallest accumulation point, hence $\liminf_{n\to\infty} F_1(x_n) = \liminf_{n\to\infty} F_2(x_n) = 0$, but $F_1(1) = 0$ while $F_2(1) = 1$.

Property (W) together with the compactness result, theorem 6.2.3, implies the solvability of a variety of minimizing problems. The following theorem is familiar to us in thr framework of uniformly convex Banach spaces and Hilbert spaces.

Theorem 6.3.2 *Let X be a reflexive Banach space. Assume $A \subset X$ is not empty, convex and closed, $x_0 \in X \backslash A$. Then there exists a point $a \in A$ such that*

$$|a - x_0| = \inf_{x \in A} |x - x_0|.$$

We will prove first the following generalization:

Theorem 6.3.3 (Variational Principle)
Let X be a reflexive Banach space and let $A \neq \emptyset$ be a weakly closed subset, i.e. $A = \overline{A}^w$. Suppose also that $F : A \to \mathbf{R}$ is coercive on A and has property (W), i.e.

- *$F(x_n) \to +\infty$ for all sequences $(x_n)_{n \in \mathbf{N}} \subset A$ with $|x_n| \to \infty$,*
- *For every $x \in A$ and every sequence $(x_n)_{n \in \mathbf{N}} \subset A$ with $x_n \rightharpoonup x$*

$$F(x) \leq \liminf_{n \to \infty} F(x_n).$$

Then F is bounded from below on A, and there is $x_0 \in A$ such that

$$F(x_0) = \inf_{x \in A} F(x),$$

i.e. F attains its infimum on A.

Remarks:

1. If A is bounded then the assumption of coerciveness is always satisfied,
2. An important class of examples for weakly closed sets are closed convex sets A because $\overline{A}^w = \overline{A} = A$ in this case. In particular, closed linear subspaces are weakly closed.
3. Theorem 6.3.2 is a consequence of the above theorem if we set $F(x) = |x - x_0|$.

Proof:
Let us show first that F is bounded from below. Arguing indirectly we assume that $\inf_A F = -\infty$. Then there is a sequence $(x_n) \subset A$ such that $F(x_n) < -n$. If the sequence (x_n) is not bounded then we obtain immediately a contradiction with the assumption that F is coercive. On the other hand, if the sequence (x_n) is bounded then it has a weakly convergent subsequence (use theorem 6.2.3 and the fact that X is reflexive), hence we may assume that $x_n \rightharpoonup x$ for some $x \in X$.

Actually, $x \in A$ because A was assumed to be weakly closed. Therefore, using property (W),

$$-\infty < F(x) \leq \liminf_{n\to\infty} F(x_n) = -\infty,$$

which is a contradiction. Hence,

$$\inf_A F = \alpha \in \mathbf{R}.$$

Take now a minimizing sequence, i.e. a sequence $(x_n)_{n\in\mathbf{N}} \subset A$ such that

$$F(x_n) \longrightarrow \alpha$$

(such a sequence exists because α is the infimum of F over the set A). By coerciveness, the sequence (x_n) must be bounded. Arguing as before, we obtain

$$x_n \rightharpoonup x \in A$$

after passing to a suitable subsequence. We then obtain

$$F(x) \leq \liminf_{n\to\infty} F(x_n) = \alpha,$$

hence $F(x) = \alpha$

❑

Remark:
The Variational Principle only provides the existence of a minimum of F. In order to show uniqueness of a minimum we need additional assumptions. For example, if we assume that A is conves and F is strictly convex in the sense that

$$F(tx + (1-t)y) < t\,F(x) + (1-t)\,F(y)\ \forall\ x \neq y\,,\ 0 < t < 1$$

then there is only one minimum. Indeed, if we had $x_1 \neq x_2$ with $F(x_1) = F(x_2) < F(x)\ \forall\ x \in A$ then $0 < t < 1$ yields the contradiction

$$F(x_1) \leq F(t\,x_1 + (1-t)\,x_2) < tF(x_1) + (1-t)F(x_2) = F(x_2) = F(x_1).$$

Also in theorem 6.3.2 the minimum is in general not unique. There is only one minimum if the norm is strictly subadditive, i.e.

$$x, y \neq 0 \text{ and } |x+y| = |x| + |y| \text{ implies } x = ty$$

for some $t > 0$. In the case of a uniformly convex Banach space the norm is always strictly aubadditive, as we have shown, therefore the minimum is unique (the same applies of course to the Hilbert space setting).

The following example shows that minimizing functions on infinite dimensional normed vector spaces is much different than in the finite dimensional situation.

If X is a finite dimensional vector space, $F : X \supset M \to \mathbf{R}$ is continuous, the set $M \subset X$ is closed and F is coercive on M, then F attains its infimum on M. Let us compare with the following infinite dimensional example:
Define $X = C^1([0,1])$ with the obvious norm and

$$M := \{u \in X \,|\, u(0) = 0\,,\, u'(1) = 1\},$$

which is a closed linear subspace of X. Define

$$F(u) := \|u'\|_{C^0([0,1])} + \int_0^1 (u'(x))^2 dx,$$

which is a continuous function on X (Indeed, $u_n \to u$ in the C^1–norm implies that $F(u_n) \to F(u)$). Moreover, the map F is coercive because

$$F(u_n) \geq \|u_n'\|_{C^0([0,1])} \to \infty \text{ if } \|u_n\|_{C^1([0,1])} \to \infty \textbf{ and } u_n \in M.$$

We claim that F does not attain its infimum on M. We note that

$$F(u) \geq \|u_n'\|_{C^0([0,1])} \geq |u'(1)| = 1 \;\forall\; u \in M.$$

On the other hand, the functions $u_\lambda(x) := \lambda^{-1} x^\lambda$ are in M if $\lambda > 1$ and

$$\|u_\lambda'\|_{C^0([0,1])} = 1 \;\forall\; \lambda > 1$$

$$\int_0^1 (u_\lambda'(x))^2 dx = \frac{1}{2\lambda - 1} \to 0 \text{ as } \lambda \to \infty.$$

Therefore,

$$\inf_M F = 1.$$

The existence of a minimum $u \in M$ would imply in view of $u'(1) = 1$

$$\|u'\|_{C^0([0,1])} = 1 \text{ and } \int_0^1 (u'(x))^2 dx = 0,$$

which is a contradiction. The space $C^1([0,1])$ is not reflexive (use a similar argument as we did in the case of C^0). Although F is continuous (with respect to the C^1–norm) it it may not have property (W).

6.4 An application to a partial differential equation

Let $(X, |\,.\,|)$ be a Banach space. A continuous map $F : X \to \mathbf{R}$ is called Fréchet differentiable at the point $x \in X$ if there exists a linear functional $DF(x) \in X'$ such that

$$\lim_{h \to 0} \frac{1}{|h|} |F(x+h) - F(x) - DF(x)h| = 0$$

If F is everywhere Fréchet differentiable then the map $DF : X \to X'$ is called the Fréchet derivative of F. The directional derivative of F in the direction of $h \in X$ is given by

$$\frac{d}{d\varepsilon} F(x + \varepsilon h)\Big|_{\varepsilon=0} = DF(x)h.$$

We call a point $x \in X$ critical if $DF(x) \equiv 0$. The equation

$$DF(x) = 0$$

is also called the Euler–Lagrange equation for the function F. Relative Maxima and Minima are examples for critical points, but there are also saddle type critical points x, i.e. every neighborhood of x contains points x_1, x_2 such that $F(x_1) < F(x) < F(x_2)$. The Calculus of Variations deals with finding critical points of maps $F : X \to \mathbf{R}$ as above. The Variational Principle which we proved in the previous section guarantees the existence of a minimum under suitable assumtions. The Calculus of Variations is a very old and vast part of mathematics. The purpose of this section is only to demonstrate that our rather abstract Variational principle has very concrete applications. Weak solutions for many nonlinear partial differential equations can be identified with critical points of suitable functions F. In such cases the Calculus of Variations yields existence theorems for weak solutions. Our example is from the first section of the nice textbook by Michael Struwe (Variational Methods and their Applications to nonlinear partial differential equations and Hamiltonian systems). Let us introduce some notation: We denote the standard Euclidean scalar product on $\mathbf{R}^n$ by $\langle \, . \, , \, . \, \rangle$. If $\Omega \subset \mathbf{R}^n$ is a domain and if $u : \Omega \to \mathbf{R}$, $\xi = (\xi_1, \ldots, \xi_n) : \Omega \to \mathbf{R}^n$ are twice differentiable maps then we write

$$\nabla u := (\partial_1 u, \ldots, \partial_n u)$$

for the gradient and

$$\nabla \cdot \xi := \sum_{k=1}^{n} \partial_k \xi_k$$

for the divergence of ξ.

Theorem 6.4.1 *Let $\Omega \subset \mathbf{R}^n$ be a bounded domain, $2 \le p < \infty$ and q such that $1/p + 1/q = 1$. Moreover, let $f \in L^q(\Omega)$ be given. Then there exists a weak solution $u \in H_0^{1,p}(\Omega)$ to the boundary value problem*

$$-\nabla \cdot (\, |\nabla u|^{p-2} \, \nabla u \,) \; = \; f \; \text{in } \Omega,$$

$$u \; \equiv \; 0 \; \text{on } \partial\Omega$$

in the sense that

$$\int_\Omega \Big(|\nabla u(x)|^{p-2} \langle \nabla u(x), \nabla \phi(x) \rangle - f(x)\phi(x) \Big) \, dx \; = \; 0 \;\; \forall \; \phi \in C_0^\infty(\Omega). \qquad (6.1)$$

Remark:
An elementary computation using partial integration shows that the existence of a classical solution u to the equation $-\nabla \cdot (|\nabla u|^{p-2}\nabla u) = f$ would imply formula (6.1). The operator $-\nabla \cdot (|\nabla u|^{p-2}\nabla u)$ is called the p-Laplacian. In the case $p = 2$ it reduces to $-\Delta$.

Proof:
Recall the Poincaré inequality which we have proved earlier in (4.6),

$$\int_\Omega |v(x)|^2 dx \le C_0 \int_\Omega |\nabla v(x)|^2 dx \ \forall\ v \in H_0^{1,2}(\Omega).$$

Our original proof can be slightly modified so that for $p \ge 2$

$$\int_\Omega |v(x)|^p dx \le C_0 \int_\Omega |\nabla v(x)|^p dx \ \forall\ v \in H_0^{1,p}(\Omega)$$

(we leave this fact as an easy exercise). This implies that there is a constant $c > 0$ depending on p and on Ω so that

$$\|\nabla v\|_{L^p(\Omega)} \le \|v\|_{1,p,\Omega} \le c\,\|\nabla v\|_{L^p(\Omega)} \ \forall\ v \in H_0^{1,p}(\Omega). \tag{6.2}$$

Hence

$$\|v\| := \|\nabla v\|_{L^p(\Omega)}$$

is a norm on the Sobolev space $H_0^{1,p}(\Omega)$ which is equivalent to the usual $W^{1,p}$–norm. Note that this is only true for bounded domains and only for $H_0^{1,p}(\Omega)$, not for $W^{1,p}(\Omega)$. Consider the following map from the Banach space $X = (H_0^{1,p}(\Omega), \|\,.\,\|)$ into the real numbers.

$$F(u) := \frac{1}{p}\int_\Omega |\nabla u(x)|^p \, dx - \int_\Omega f(x)u(x)dx.$$

The map F is well–defined and continuous by Hölder's inequality (and the above version of the Poincaré inequality). We compute for $h \in X$

$$\begin{aligned}
\frac{d}{d\varepsilon}F(u+\varepsilon h)\Big|_{\varepsilon=0} &= \frac{1}{p}\,\frac{d}{d\varepsilon}\int_\Omega |\nabla u(x) + \varepsilon\nabla h(x)|^p \, dx\Big|_{\varepsilon=0} - \\
&\quad - \int_\Omega f(x)h(x)dx \\
&= \frac{1}{p}\int_\Omega p|\nabla u(x)|^{p-1}\,\frac{d}{d\varepsilon}|\nabla u(x) + \varepsilon\nabla h(x)|\Big|_{\varepsilon=0} dx - \\
&\quad - \int_\Omega f(x)h(x)dx \\
&= \int_\Omega \Big(|\nabla u(x)|^{p-2}\langle\nabla u(x), \nabla h(x)\rangle - f(x)h(x)\Big)\, dx,
\end{aligned}$$

which is exactly the left hand side of equation (6.1). On the other hand, the linear map

$$DF(u) : X \longrightarrow \mathbf{R}$$

$$DF(u)h := \int_\Omega \Big(|\nabla u(x)|^{p-2} \langle \nabla u(x), \nabla h(x) \rangle - f(x)h(x) \Big)\, dx$$

is continuous. Using Hölder's inequality and $q = \frac{p}{p-1}$ we obtain

$$\begin{aligned} \left| \int_\Omega |\nabla u(x)|^{p-2} \langle \nabla u(x), \nabla h(x) \rangle \, dx \right| &\leq \int_\Omega |\nabla u(x)|^{p-1} |\nabla h(x)|\, dx \\ &\leq \left(\int_\Omega |\nabla u(x)|^p dx \right)^{\frac{p-1}{p}} \left(\int_\Omega |\nabla h(x)|^p dx \right)^{1/p} \\ &=: C\, \|h\|. \end{aligned}$$

The linear functional $DF(u)$ also satisfies the condition of Fréchet derivative of F at the point $u \in X$. Because $C_0^\infty(\Omega)$ is dense in $H_0^{1,p}(\Omega)$, the condition that $u \in X$ is a weak solution (i.e. solves (6.1)), is then equivalent to the condition of u being a critical point of the function F. In particular, we have found a weak solution if we can show that F has a minimum. Hence we will complete the proof by checking the assumptions of the Variational principle of the previous section. The Banach space X is reflexive because $L^p(\Omega)$ is reflexive for $1 < p < \infty$ (see next section, Milman's theorem) and X is isometrically isomorphic to a closed linear subspace in $\Pi := L^p(\Omega) \times \ldots \times L^p(\Omega)$ (product (n+1) times) via the isometry

$$H_0^{1,p}(\Omega) \longrightarrow \Pi$$

$$u \mapsto (u, \partial_1 u, \ldots, \partial_n u).$$

Moreover, the map F above is coercive since with (6.2)

$$\begin{aligned} F(u) &\geq \frac{1}{p}\|u\|^p - \|f\|_{L^q(\Omega)} \|u\|_{L^p(\Omega)} \\ &\geq \frac{1}{p}(\|u\|^p - c'\, \|u\|) \end{aligned}$$

for a suitable constants $c', C > 0$. Since $p \geq 2$ we obtain $F(u) \to +\infty$ if $\|u\| \to \infty$. Assume now that $u_n \rightharpoonup u$ in $H_0^{1,p}(\Omega)$. We have to show that

$$F(u) \leq \liminf_{n\to\infty} F(u_n).$$

First we note that the map

$$H_0^{1,p}(\Omega) \ni u \longmapsto \int_\Omega f(x)u(x)dx$$

is a continuous linear functional if $f \in L^q(\Omega)$. By definition of weak convergence we then get

$$\int_\Omega u_n(x) f(x) dx \longrightarrow \int_\Omega u(x) f(x) dx.$$

On the other hand, we have shown that $u_n \rightharpoonup u$ implies

$$\|u\| \leq \liminf_{n\to\infty} \|u_n\|$$

which takes care of the first term in F. We now apply the Variational Principle with $A = X$ which implies that F is bounded from below, and there is $u_0 \in X$ such that

$$F(u_0) = \inf_X F,$$

i.e. u_0 is an absolute minimum of F and therefore a critical point of F completing the proof.

❑

6.5 Weak topologies

We will put the concepts of weak and weak* convergence into a more general framework. We start with some general remarks from point set topology. A good reference is the book by James Munkres (Topology, a first course). If $(X, |\,.\,|)$ is a normed vector space then a subset $U \subset X$ is called open if for every $x \in X$ there is $\varepsilon > 0$ and an open ball $B_\varepsilon(x) = \{y \in X \mid |x - y| < \varepsilon\} \subset U$. We note that open sets in a normed vector space have the following properties:

- Unions of open sets are open,
- Finite intersections of open sets are open.

This can be formalized as follows:

Definition 6.5.1 *Let X be a set and let $\mathcal{T}$ be a set consisting of subsets of X so that*

- $\emptyset, X \in \mathcal{T}$,
- *unions of sets $U_i \in \mathcal{T}$ are again in $\mathcal{T}$,*
- *if $U_1, \ldots, U_k \in \mathcal{T}$ then*
$$\bigcap_{1\leq i\leq k} U_i \in \mathcal{T}.$$

Then the pair $(X, \mathcal{T})$ is called a topological space, the system $\mathcal{T}$ is called a topology on X. If $(X, \mathcal{T})$ is a topological space then every set $U \in \mathcal{T}$ is called an open set.

We just pointed out that open sets in a normed vector space (in the usual sense) satisfy the conditions in the above definition. We say that the norm on X induces a topology on X (also called the norm–topology). In general, every metric space comes with a natural topology. On the other hand, there are many topological spaces which do not carry any metric inducing the given topology. We will see that the concepts of weak convergence and weak* convergence come from a certain topology on the Banach space $(X, |\,.\,|)$ which is always different from the norm topology if X is infinite dimensional and which is often not coming from any metric on X.
If $(X_1, \mathcal{T}_1)$ and $(X_2, \mathcal{T}_2)$ are topological spaces then we can define the notions of continuous maps, compactness and convergence of sequences.

Definition 6.5.2

- *A map $f : X_1 \to X_2$ is called continuous if for every open set $U \subset X_2$ the preimage $f^{-1}(U) \subset X_1$ is also open.*
- *Let $x \in X_1$ and $(x_n)_{n \in \mathbf{N}}$. We say the sequence x_n converges to x with respect to the topology $\mathcal{T}_1$ if for every set $U \in \mathcal{T}_1$ containing x there is a positive integer N such that $x_n \in U$ for all $n \geq N$.*
- *A subset $A \subset X$ is called compact if the following is true: If $(U_i)_{i \in I} \subset \mathcal{T}_1$ is any system of open maps such that $\bigcup_{i \in I} U_i \supset A$ ('an open covering of A') then finitely many of the sets U_i already cover the set A.*

If the topological space is a normed vector space with the norm topology then the above notions are equivalent to the usual '$\varepsilon - \delta$–definitions'.

Definition 6.5.3 *Let X be a set with two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$ defined on it. We say that $\mathcal{T}_1$ is finer than $\mathcal{T}_2$ if $\mathcal{T}_2 \subset \mathcal{T}_1$, i.e. the topology $\mathcal{T}_1$ has more open sets than the topology $\mathcal{T}_2$.*

The finest topology on a set X is

$$\mathcal{T}_{discrete} := \{A \,|\, A \subset X\},$$

the set of all subsets of X. The least fine topology on a set X is the one which just consists of the two sets X and the empty set, $\mathcal{T}_{coarse} := \{X, \emptyset\}$. We will always equip $\mathbf{R}^n$ with the topology induced by the Euclidean norm. Since all norms on $\mathbf{R}^n$ are equivalent they all induce the same topology on $\mathbf{R}^n$ (you may verify this as an exercise). The topology on a set X determines which maps

$$f : X \longmapsto \mathbf{R}$$

are continuous and which are not. For example, if we equip X with the finest topology possible $\mathcal{T}_{discrete}$ then any map f is continuous. On the other hand, only finite subsets of X are compact. There are so many open sets and so many possibilities to assemble open coverings of a set such that the chance to extract a finite subcovering are very slim. If we equip X with the topology $\mathcal{T}_{coarse}$ then only constant maps $f : X \to \mathbf{R}$ are continuous. On the other hand every subset of X is compact because every open covering of any set just consists of X alone. Summarizing, the finer the topology on a set X, the bigger the chance that a given map $f : X \to \mathbf{R}$ is continuous and the smaller the chance that a given subset $A \subset X$ is compact.

Let now $(X, |\,.\,|)$ be a normed vector space. The dual space X' consists of all linear maps $X \to \mathbf{R}$ which are continuous (with respect to the norm topology). Denote the norm topology on X by $\mathcal{T}$. But the norm topology may not be the most effective topology to make all $\ell \in X'$ continuous: Is it possible to find a less fine topology $\mathcal{T}_w$ on X such that all $\ell \in X'$ are still continuous with respect to the new topology $\mathcal{T}_w$ on X ? We make the following definitions:

Definition 6.5.4 *Let $(X, |\,.\,|)$ be a normed vector space. We define $\mathcal{T}_w$ to be the least fine topology on X such that all maps $\ell \in X'$ are still continuous with respect to $\mathcal{T}_w$ on X. This topology is called the weak topology on X.*

Definition 6.5.5 *Let $(X, |\,.\,|)$ be a normed vector space. We define $\mathcal{T}_w^*$ to be the least fine topology on X' such that all maps $\phi \in J_X(X) \subset X''$ are still continuous with respect to $\mathcal{T}_w^*$ on X'. This topology is called the weak* topology on X'.*

At this moment it is not clear whether $\mathcal{T}$ and $\mathcal{T}_w$ are really different. It will turn out that they are if X is infinite dimensional. If $(X, \mathcal{T})$ is a topological space then the system $\mathcal{T}$ is not very convenient to handle because it is usually very large.

Definition 6.5.6 *If X is a set then a basis on X is a nonempty collection $\mathcal{B}$ of subsets of X which satisfy the following conditions:*

- *For every $x \in X$ there is at least one set $B \in \mathcal{B}$ such that $x \in B$,*
- *If $x \in B_1 \cap B_2$ for some $B_1, B_2 \in \mathcal{B}$ then there is a set $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.*

Then the topology $\mathcal{T}_{\mathcal{B}}$ induced by $\mathcal{B}$ is defined as follows: A set $U \subset X$ is said to be open if for each $x \in U$ there is some $B \in \mathcal{B}$ such that $x \in B$ and $B \subset U$.

Exercise 6.5.7 *Verify that $\mathcal{T}_{\mathcal{B}}$ as defined above satisfies the conditions of a topology.*

Every set $B \in \mathcal{B}$ is trivially open, i.e. $\mathcal{B} \subset \mathcal{T}_{\mathcal{B}}$. If $\mathcal{T}'$ is a topology so that $\mathcal{B} \subset \mathcal{T}' \subset \mathcal{T}_{\mathcal{B}}$ then we must have $\mathcal{T}' = \mathcal{T}_{\mathcal{B}}$. This means that the topology on X induced by the basis $\mathcal{B}$ is the least fine topology containing all the sets in

$\mathcal{B}$. Indeed, if $U \in \mathcal{T}_{\mathcal{B}}$ then for every $x \in U$ we can find a set $B_x \in \mathcal{B}$ such that $x \in B_x \subset U$. On the other hand, $B_x \in \mathcal{T}'$, and since $\mathcal{T}'$ is a topology

$$U = \bigcup_{x \in U} B_x \in \mathcal{T}',$$

hence $\mathcal{T}_{\mathcal{B}} \subset \mathcal{T}'$.

Exercise 6.5.8 *Let $\mathcal{B}_1$ be the set consisting of all open balls in the plane $\mathbf{R}^2$. Show that it is a basis for the standard topology on $\mathbf{R}^2$. Show the same for the set of all open rectangles in the plane.*

We wish to characterize bases for the weak and the weak* topologies, and we want to show that convergence with respect to these topologies coincides with the notions of weak and weak* convergence that we have introduced earlier.

Proposition 6.5.9 *Let $(X, |\,.\,|)$ be a normed vector space. We define a system $\mathcal{B}$ of subsets of X as follows: We say that $U \in \mathcal{B}$ for $U \subset X$ if there are a point $x_0 \in X$, a number $\varepsilon > 0$ and finitely many linear functionals $\ell_1, \ldots, \ell_k \in X'$ such that*

$$U = \{x \in X \mid |\ell_i(x - x_0)| < \varepsilon \ \forall\ i = 1, \ldots, k\}.$$

Then $\mathcal{B}$ is a basis which induces the weak topology $\mathcal{T}_w$ on X.

Proof:
First, we have to check that $\mathcal{B}$ is a basis. Pick $x_0 \in X$ and any $\ell \in X'$. Then

$$x_0 \in B := \{x \in X \mid |\ell(x - x_0)| < \varepsilon\} \text{ and } B \in \mathcal{B}.$$

Now pick $B_1, B_2 \in \mathcal{B}$, i.e.

$$B_1 := \{x \in X \mid |\ell_j^{(1)}(x - x_1)| < \varepsilon_1 \ \forall\ j = 1, \ldots, k_1\}$$

and

$$B_2 := \{x \in X \mid |\ell_i^{(2)}(x - x_2)| < \varepsilon_2 \ \forall\ i = 1, \ldots, k_2\}$$

for suitable $\ell_j^{(1)}, \ell_i^{(2)} \in X'$, $\varepsilon_1, \varepsilon_2 > 0$, $x_1, x_2 \in X$ and $k_1, k_2 \in \mathbf{N}$. Assume that $x_3 \in B_1 \cap B_2$. Hence

$$\delta_1 := \varepsilon_1 - |\ell_j^{(1)}(x_3 - x_1)| > 0 \text{ and } \delta_2 := \varepsilon_2 - |\ell_i^{(2)}(x_3 - x_2)| > 0 \ \forall\ i, j.$$

Define now

$$B_3 := \{x \in X \mid |\ell_j^{(1)}(x - x_3)|\,, |\ell_i^{(2)}(x - x_3)| < \min\{\delta_1, \delta_2\} \ \forall\ i, j\}$$

which is a set from the collection $\mathcal{B}$. Then trivially $x_3 \in B_3$. By linearity of $\ell_j^{(1)}, \ell_i^{(2)}$ and the triangle inequality we conclude also that $B_3 \subset B_1 \cap B_2$. This shows that $\mathcal{B}$ is a basis.

Now we have to make sure that the weak topology $\mathcal{T}_w$ contains all the sets in $\mathcal{B}$. The weak topology must contain each of the sets

$$\{x \in X \mid \ell_i(x_0) - \varepsilon < \ell_i(x) < \ell_i(x_0) + \varepsilon\} \ , \ i = 1, \ldots, k$$

and therefore also U, which is a finite intersection of them. The topology $\mathcal{T}_{\mathcal{B}}$ induced by the basis $\mathcal{B}$ leaves all $\ell \in X'$ continuous, hence $\mathcal{T}_w \subset \mathcal{T}_{\mathcal{B}}$, and by our remarks before $\mathcal{T}_w = \mathcal{T}_{\mathcal{B}}$ which is the assertion of the proposition.

❑

There is a similar statement for the weak* topology:

Proposition 6.5.10 *Let $(X, |\,.\,|)$ be a normed vector space. We define a system $\mathcal{B}^*$ of subsets of X' as follows: We say that $U \in \mathcal{B}^*$ for $U \subset X'$ if there are a point $\ell_0 \in X'$, a number $\varepsilon > 0$ and finitely many linear functionals $\phi_1, \ldots, \phi_k \in J_X(X) \subset X''$ such that*

$$U = \{\ell \in X' \mid |\phi_i(\ell - \ell_0)| < \varepsilon \ \forall \ i = 1, \ldots, k\}.$$

Then $\mathcal{B}^*$ is a basis which induces the weak* topology $\mathcal{T}_w^*$ on X'.

Theorem 6.5.11 *Let $(X, |\,.\,|)$ be a normed vector space. The norm–topology $\mathcal{T}$ and the weak topology on X coincide if and only if X is finite dimensional.*

Proof:
Assume that X is finite dimensional. We have to show that $\mathcal{T} \subset \mathcal{T}_w$ since the reverse inclusion holds by definition. Hence let $x_0 \in X$ and let U be a neighborhood of x_0 with respect to the norm–topology. Let $R > 0$ such that $B_R(x_0) \subset U$. We have to construct a weak neighborhood V of x_0 such that $V \subset U$ (then U would be weakly open). Choose a basis $e_1, \ldots, e_n$ of X so that $|e_i| = 1$ for all $i = 1, \ldots, n$. Every point $x \in X$ has a unique decomposition $x = \sum_{i=1}^n x_i e_i$. The maps

$$x \overset{\ell_i}{\longmapsto} x_i$$

define elements in X'. We have

$$|x - x_0| \leq \sum_{i=1}^n |\ell_i(x - x_0)| < n \max_{1 \leq i \leq n} |\ell_i(x - x_0)|.$$

Define now

$$V := \{x \in X \mid |\ell_i(x - x_0)| < \frac{R}{n} \ \forall \ i = 1, \ldots, n\}.$$

This is a weak neighborhood of x_0 and it satisfies $V \subset B_R(x_0) \subset U$.
As for the converse direction, assume that X is infinite dimensional. Any set V as above with $x_0 = 0$ contains an infinite dimensional dimensional linear subspace of X. In particular open balls $B_R(0)$ are not open with respect to the weak topology.

□

Exercise 6.5.12 *Let $(X, |\,.\,|)$ be an infinite dimensional normed vector space. Show that the sphere $S = \{x \in X \mid |x| = 1\}$ is not closed with respect to the weak topology (weakly closed=complement is weakly open).*

We note that the notion of convergence with respect to the weak topology and weak* topology coincides with the definitions that we made earlier. Let us carry this out for the weak topology (weak* topology works out in the same way). A sequence (x_n) converges to x with respect to the weak topology if for every $U \in \mathcal{T}_w$ containing x there is an integer N such that $x_n \in U$ for all $n \geq N$. Let us show that this implies $x_n \rightharpoonup x$. Pick $\ell \in X'$, $\varepsilon > 0$ and define

$$U := \{y \in X \mid |\ell(y - x)| \leq \varepsilon\}.$$

For every $\varepsilon > 0$ we can find N such that $|\ell(x_n - x)| \leq \varepsilon$ if $n \geq N$. But this means that $\ell(x_n) \to \ell(x)$, hence $x_n \rightharpoonup x$ because this consideration applies to any $\ell \in X'$. On the other hand, if U is any weakly open neighborhood of x then it contains a set of the form

$$V = \{y \in X \mid |\ell_i(y - x)| < \varepsilon \;\forall\; i = 1, \ldots, k\}$$

and $\ell_i(x_n) \to \ell_i(x)$ for all $i = 1, \ldots, k$ implies that $x_n \in V \subset U$ for sufficiently large n.

6.6 Proof of Milman's theorem

We have seen in the previous sections that reflexive Banach spaces are very important because of the compactness properties with respect to weak convergence. Milman's theorem states that uniformly convex Banach spaces are reflexive. We then know that the spaces $L^p(\Omega)$ (and $W^{k,p}(\Omega)$, $H_0^{k,p}(\Omega)$) are reflexive for $1 < p < \infty$ since we have already shown that they are uniformly convex (although reflexivity can be established in these particular cases more directly using the Riesz representation theorem). Milman's theorem is nevertheless very useful because uniform convexity is often easier to verify than the definition of reflexivity. This does not work all the time because there are Banach spaces which are reflexive but not uniformly convex.

Theorem 6.6.1 (D.P. Milman)
Every uniformly convex Banach space is reflexive.

Before we can prove the theorem we need two lemmas:

Lemma 6.6.2 *Let X be a Banach space. Moreover let $\ell_1, \ldots, \ell_n \in X'$ and $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$. Then the following two properties are equivalent:*

1. For any $\varepsilon > 0$ there is some $x_\varepsilon \in X$ with $|x_\varepsilon| \leq 1$ such that

$$|\ell_i(x_\varepsilon) - \alpha_i| < \varepsilon \;\forall\; i = 1, \ldots, n.$$

2.

$$\left|\sum_{i=1}^{n} \beta_i \alpha_i\right| \leq \left\|\sum_{i=1}^{n} \beta_i \ell_i\right\|_{X'} \quad \forall\; \beta_1, \ldots, \beta_n \in \mathbf{R}.$$

Proof:
Show that 1. implies 2.: For this purpose pick $\beta_1, \ldots, \beta_n \in \mathbf{R}$ and define $M := \sum_{i=1}^{n} |\beta_i|$ so that

$$\left|\sum_{i=1}^{n} \beta_i \ell_i(x_\varepsilon) - \sum_{i=1}^{n} \beta_i \alpha_i\right| = \left|\sum_{i=1}^{n} \beta_i(\ell_i(x_\varepsilon) - \alpha_i)\right| < M\varepsilon$$

and

$$\left|\sum_{i=1}^{n} \beta_i \alpha_i\right| \leq \left\|\sum_{i=1}^{n} \beta_i \ell_i\right\|_{X'} |x_\varepsilon| + M\varepsilon \leq \left\|\sum_{i=1}^{n} \beta_i \ell_i\right\|_{X'} + M\varepsilon,$$

which is true for any $\varepsilon > 0$ so that 2. follows. Let us now show that 2. also implies 1.: Consider $\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbf{R}^n$ and define a linear map

$$\Phi : X \longrightarrow \mathbf{R}^n$$

by

$$\Phi(x) := (\ell_1(x), \ldots, \ell_n(x)).$$

Property 1. above is the same as saying

$$\alpha \in \overline{\Phi(B)},$$

where B is the closed unit ball in X. Arguing indirectly, we assume that $\alpha \notin \overline{\Phi(B)}$. In $\mathbf{R}^n$ we may separate strictly the sets $\{\alpha\}$ and $\overline{\Phi(B)}$, i.e. there is $\beta = (\beta_1, \ldots, \beta_n) \in \mathbf{R}^n$ and $c \in \mathbf{R}$ such that

$$\langle \Phi(x), \beta \rangle < c < \langle \alpha, \beta \rangle \;\forall\; x \in B.$$

We obtain by definition of Φ

$$\left|\sum_{i=1}^{n} \beta_i \ell_i(x)\right| < c < \sum_{i=1}^{n} \alpha_i \beta_i \text{ whenever } x \in B.$$

Taking the supremum over all $x \in B$ we get

$$\left\|\sum_{i=1}^{n} \beta_i \ell_i\right\|_{X'} \leq c < \sum_{i=1}^{n} \alpha_i \beta_i$$

contradicting condition 2.

❑

If X is a Banach space we consider now the canonical embedding into its bidual space

$$J_X : X \longrightarrow X''.$$

Again, we denote the closed unit ball in X by B. Since J_X is an isometry the set $J_X(B)$ will be closed in $B_{X''} \subset X''$, where $B_{X''}$ denotes the closed unit ball in the bidual space. If the Banach space is reflexive then $B_{X''} = J_X(B)$. If we equip X'' with the weak* topology then $J_X(B)$ is dense in $B_{X''}$.

Lemma 6.6.3 *Let X be a Banach space. Then $J_X(B)$ is dense in $B_{X''}$ with respect to the weak* topology on X''.*

Proof:
Pick $\phi_0 \in B_{X''} \subset X''$ and a neighborhood U of ϕ_0 with respect to the weak* topology. We have to show that $U \cap J_X(B) \neq \emptyset$. We characterized earlier a basis for the weak* topology, hence we may assume without loss of generality that

$$U = \{\phi \in X'' \mid |f_i(\phi - \phi_0)| < \varepsilon \ \forall\ i = 1, \ldots, k\},$$

where $f_i = J_{X'}(\ell_i) \subset X'''$ are suitable elements in X''' which are actually in the image of $J_{X'} : X' \to X'''$. This means that

$$U = \{\phi \in X'' \mid |(\phi - \phi_0)(\ell_i)| < \varepsilon \ \forall\ i = 1, \ldots, k\}.$$

Hence we have to find a point $x \in B$ such that

$$|\ell_i(x) - \phi_0(\ell_i)| < \varepsilon \ \forall\ i = 1, \ldots, k.$$

Writing now $\alpha_i := \phi_0(\ell_i)$ and using that $\|\phi_0\|_{X''} \leq 1$ we estimate for any $\beta_1, \ldots, \beta_n \in \mathbf{R}$

$$\left|\sum_{i=1}^{n} \beta_i \alpha_i\right| = \left|\phi_0\left(\sum_{i=1}^{n} \beta_i \ell_i\right)\right| \leq \left\|\sum_{i=1}^{n} \beta_i \ell_i\right\|_{X'}$$

since $\|\phi_0\|_{X''} \leq 1$, but this is property 2. of lemma 6.6.2, hence there is $x_\varepsilon \in X$ with $|x_\varepsilon| \leq 1$ so that

$$|\ell_i(x_\varepsilon) - \alpha_i| < \varepsilon \ \forall\ i = 1, \ldots, n.$$

❑

Proof:
(Milman's theorem, proof due to Kakutani)
Let $(X, |\,.\,|)$ be a Banach space and $\phi \in X''$ with $\|\phi\|_{X''} = 1$. We have to find some $x \in X$ with $|x| = 1$ such that $J_X(x) = \phi$, where J_X denotes the canonical

isometry from X into its bidual space. Denote the closed unit ball in X by B. Because the norm of ϕ equals 1 we may find a sequence $(\ell_n)_{n\in\mathbf{N}} \subset X'$ such that

$$1 \geq \phi(\ell_n) \geq 1 - \frac{1}{n}.$$

By the previous lemma $J_X(B)$ is dense in X'' with respect to the weak* topology. A typical weak* neighborhood of $\phi \in X''$ in X'' is

$$U_n := \{\psi \in X'' \,|\, |\alpha_i(\psi - \phi)| < \frac{1}{n} \ \forall i = 1, \ldots, n\},$$

where $\alpha_i = J_{X'}(\ell_i)$, so that

$$U_n := \{\psi \in X'' \,|\, |\psi(\ell_i) - \phi(\ell_i)| < \frac{1}{n} \ \forall i = 1, \ldots, n\}.$$

Since $J_X(B) \cap U_n \neq \emptyset$ there are points $x_n \in X$ with $|x_n| \leq 1$ such that

$$|J_X(x_n)(\ell_i) - \phi(\ell_i)| = |\ell_i(x_n) - \phi(\ell_i)| < \frac{1}{n} \ \forall\ i = 1, \ldots, n. \tag{6.3}$$

We choose now $m \geq n$ and obtain with the triangle inequality

$$\begin{aligned} 2 &\geq |x_n + x_m| \\ &\geq \|\ell_n\|_{X'} |x_n + x_m| \\ &\geq \ell_n(x_n + x_m) \\ &\geq \phi(\ell_n) - \frac{1}{n} + \phi(\ell_n) - \frac{1}{m} \\ &\geq 1 - \frac{3}{n} + 1 - \frac{1}{m} \\ &\geq 2 - \frac{4}{n}. \end{aligned}$$

which implies that

$$\lim_{n,m\to\infty} \left| \frac{x_n + x_m}{2} \right| = 1.$$

By uniform convexity of X the sequence $(x_n)_{n\in\mathbf{N}}$ must then be a Cauchy sequence, i.e. $x_n \to x$ for some $x \in B$. Inequality (6.3) then implies that

$$\phi(\ell_i) = \ell_i(x) \ \forall\ i \in \mathbf{N}.$$

We want to show that $J_X(x)\ell = \ell(x) = \phi(\ell)$ for all $\ell \in X'$ which would conclude the proof of the theorem. We claim that x is unique. Assuming that there is another $\bar{x} \in X$ with $\phi(\ell_i) = \ell_i(\bar{x})$ for all integers i, we define a sequence (x'_n) in X by $(x, \bar{x}, x, \bar{x}, \ldots)$. The sequence (x'_n) trivially satisfies inequality (6.3) implying that (x'_n) is a Cauchy sequence and $x = \bar{x}$. Let now $\ell \in X'$ be an arbitrary element with norm equals 1. We replace now the original sequence (ℓ_i)

by the sequence $(\ell, \ell_1, \ell_2, \ldots)$. Running the same argument as in the beginning once again we get $\bar{x} \in X$ with

$$\phi(\ell) = \ell(\bar{x}) \text{ and } \phi(\ell_i) = \ell_i(\bar{x}) \; \forall \; i.$$

By the uniqueness which we have just proved we have $x = \bar{x}$, hence $\phi(\ell) = \ell(x)$, and we are done.

❑

Chapter 7

Spectrum of compact operators

7.1 Spectrum

In this chapter we assume that $(X, |\,.\,|)$ is a complex Banach space.

Definition 7.1.1 *Let $T \in L(X)$.*

1. *We define the resolvent set of T by*

$$\rho(T) := \{\lambda \in \mathbf{C} \,|\, \ker(\lambda Id - T) = \{0\}\,,\, R(\lambda Id - T) = X\}.$$

2. *We define the spectrum of T by*

$$\sigma(T) := \mathbf{C} \backslash \rho(T).$$

3. *We decompose the spectrum into the following sets:*

   (a) *The point spectrum*

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

   (b) *The continuous spectrum*

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

   (c) *The residual spectrum*

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

We have $\lambda \in \rho(T)$ if and only if $\lambda \mathrm{Id} - T$ is bijective. By the inverse mapping theorem the inverse $(\lambda \mathrm{Id} - T)^{-1}$ is again in $L(X)$. We call $(\lambda \mathrm{Id} - T)^{-1}$ the resolvent of T in λ, and we denote it by $R(\lambda, T)$.
We have $\lambda \in \sigma_p(T)$ if and only if there is $x \neq 0$ such that $Tx = \lambda x$. We then call x an eigenvector associated to the eigenvalue λ. If the Banach space X is finite dimensional then the spectrum of a linear operator consists of the point spectrum only, i.e. every point in the spectrum is an eigenvalue. In general, this is not true as the following example shows:

Example:
Consider the following operator $T \in L(C^0([0,1]))$

$$(Tf)(x) := \int_0^x f(t)dt.$$

We have $R(T) = \{f \in C^1([0,1]) \,|\, f(0) = 0\}$ which is not closed in $C^0([0,1])$. On the other hand, $\ker(T) = \{0\}$. Therefore $0 \in \sigma_r(T)$. We remark for later reference that the operator T is compact by the Ascoli–Arzela theorem (recall that T compact means that $T(B_1(0))$ is precompact).

Theorem 7.1.2 *Let $T \in L(X)$. The resolvent set $\rho(T) \subset \mathbf{C}$ is an open set, and the resolvent function*

$$\mathbf{C} \supset \rho(T) \longrightarrow L(X)$$

$$\lambda \longmapsto R(\lambda, T)$$

is an analytic function which satisfies

$$\|R(\lambda, T)\|^{-1} \leq \mathit{dist}(\lambda, \sigma(T)).$$

Proof:
Let $\lambda \in \rho(T)$. We have for any $\mu \in \mathbf{C}$

$$(\lambda - \mu)\mathrm{Id} - T = (\lambda \mathrm{Id} - T)(\mathrm{Id} - \mu\, R(\lambda, T)).$$

The operator $S(\mu) := \mathrm{Id} - \mu\, R(\lambda, T)$ is invertible if

$$|\mu| \cdot \|R(\lambda, T)\| < 1$$

(this follows from the Neumann series, proposition 3.1.8). Under this condition we have $\lambda - \mu \in \rho(T)$ showing that the resolvent set is open. Using again proposition 3.1.8 we obtain

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

This shows that the resolvent function is analytic. If $r = \|R(\lambda, T)\|^{-1}$ then $B_r(\lambda) \subset \rho(T)$, as we have just shown. But this implies that $\mathrm{dist}(\lambda, \sigma(T)) \geq r$.

□

Theorem 7.1.3 *Let $T \in L(X)$ with $X \neq \{0\}$. Then $\sigma(T) \subset \mathbf{C}$ is compact and not empty. Moreover,*

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{m \to \infty} \sqrt[m]{\|T^m\|} \leq \|T\|.$$

The number $\sup_{\lambda \in \sigma(T)} |\lambda|$ is called the spectral radius of T.

Proof:
Let $\lambda \neq 0$. By proposition 3.1.8 (Neumann series) the operator $\mathrm{Id} - \frac{T}{\lambda}$ is invertible if $\|T\| < \lambda$, and in this case

$$R(\lambda, T) = \frac{1}{\lambda}\left(\mathrm{Id} - \frac{T}{\lambda}\right)^{-1} = \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}. \tag{7.1}$$

This shows that the spectral radius r satisfies $r \leq \|T\|$. Defining

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

we obtain

$$\lambda^m \mathrm{Id} - T^m = (\lambda \mathrm{Id} - T) P_m(T) = P_m(T)(\lambda \mathrm{Id} - T)$$

This shows that $\lambda \in \sigma(T)$ implies that $\lambda^m \in \sigma(T^m)$. Because the spectral radius of T^m is bounded by the operator norm of T^m we conclude that $|\lambda^m| \leq \|T^m\|$ and therefore $|\lambda| \leq \sqrt[m]{\|T^m\|}$, i.e.

$$r \leq \liminf_{m \to \infty} \sqrt[m]{\|T^m\|}.$$

Our aim is now to show that also

$$r \geq \limsup_{m \to \infty} \sqrt[m]{\|T^m\|}.$$

We know that the function $\lambda \mapsto R(\lambda, T)$ is analytic, where $\lambda \in \mathbf{C} \backslash \overline{B_r(0)}$ (analytic on the whole plane if $\sigma(T) = \emptyset$). By Cauchy's integral theorem the integral

$$\frac{1}{2\pi i} \int_{\partial B_s(0)} \lambda^j R(\lambda, T) d\lambda$$

does not depend on s as long as $j \geq 0$ and $s > r$. Choosing $s > \|T\|$ we may use the formula (7.1). Then

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial B_s(0)} \lambda^j R(\lambda, T) d\lambda &= \frac{1}{2\pi i} \int_{\partial B_s(0)} \sum_{k=0}^{\infty} \lambda^{j-k-1} T^k d\lambda \\ &\overset{\lambda = s e^{i\theta}}{=} \frac{1}{2\pi} \sum_{k=0}^{\infty} s^{j-k} \left(\int_0^{2\pi} e^{i(j-k)\theta} d\theta \right) T^k \\ &= T^j, \end{aligned}$$

since $\int_0^{2\pi} e^{i(j-k)\theta} d\theta \neq 0$ if and only if $j = k$. Hence we have for $j \geq 0$ and $s > r$

$$\|T^j\| = \frac{1}{2\pi} \left\| \int_{\partial B_s(0)} \lambda^j R(\lambda, T) d\lambda \right\| \leq s^{j+1} \sup_{|\lambda|=s} \|R(\lambda, T)\|.$$

If $s > r$ and $j(k) \to \infty$ as $k \to \infty$ we obtain

$$\sqrt[j(k)]{\|T^{j(k)}\|} \leq s \sqrt[j(k)]{s \sup_{|\lambda|=s} \|R(\lambda, T)\|},$$

where the right hand side converges to s. Then

$$\limsup_{j\to\infty} \sqrt[j]{\|T^j\|} \leq s$$

for all $s > r$. This implies the assertion about the spectral radius. If the spectrum was empty we choose $j = 0$ and $s \searrow 0$ so that

$$\|\mathrm{Id}\| \leq s \sup_{|\lambda|\leq 1} \|R(\lambda, T)\| \longrightarrow 0,$$

hence $\mathrm{Id} = 0$ and therefore $X = \{0\}$.

❑

7.2 The spectral theorem for compact operators

Before we state and partly prove the spectral theorem for compact operators $T \in K(X)$ we make some simple remarks: If X is infinite dimensional and if T is a compact operator then $0 \in \sigma(T)$. Indeed, if we had $0 \in \rho(T)$ then $T^{-1} \in L(X)$. But then $\mathrm{Id} = T^{-1}T$ would also be a compact operator since T is compact. On the other hand, the identity operator can only be compact in finite dimensional normed vector spaces since these are the only spaces where the open unit ball is precompact. Although 0 is in the spectrum of any compact operator, it may not be an eigenvalue (see our previous example). One of the statements of the spectral theorem is that all nonzero points in the spectrum must be eigenvalues.

Theorem 7.2.1 (spectral theorem for compact operators, Riesz–Schauder) *Let $T \in K(X)$. Then*

1. $\sigma(T)\backslash\{0\}$ *consists of at most countably many points which are all eigenvalues, and which may only accumulate at* 0.
2. *For* $\lambda \in \sigma(T)\backslash\{0\}$ *we have*

$$1 \leq n_\lambda := \max\{n \in \mathbf{N} \mid \ker(\lambda Id - T)^{n-1} \neq \ker(\lambda Id - T)^n\} < \infty.$$

The number n_λ is called the order or the index of λ while the dimension of $\ker(\lambda Id - T)$ *is called the multiplicity of* λ.

3. *For $\lambda \in \sigma(T)\backslash\{0\}$ we have*

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

Both subspaces are closed and T–invariant. The space $\ker(\lambda Id - T)^{n_\lambda}$ is finite dimensional. Moreover,

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

4. *Let E_λ be the projection onto the subspace $\ker((\lambda Id - T)^{n_\lambda})$ with respect to the direct sum decomposition in 3. Then*

$$E_\lambda \circ E_\mu = 0 \text{ if } \lambda \neq \mu.$$

Before we embark on the proof of parts of the theorem we quickly insert a lemma due to M. Riesz:

Lemma 7.2.2 ('Almost approximation lemma', M. Riesz)
Let $(X, |\,.\,|)$ be a normed vector space and let Y be a closed proper subspace. Then for every $\varepsilon > 0$ there exists $x \in X$ such that $|x| = 1$ and

$$dist(x, Y) \geq 1 - \varepsilon.$$

Proof:
Let $z \in X$ with $z \notin Y$. Since Y is closed, it must have positive distance d from the point z. We may now choose $y \in Y$ such that

$$d \leq |z - y| \leq \frac{d}{1 - \varepsilon}.$$

Then

$$x := \frac{z - y}{|z - y|}$$

does the job. Indeed, if $y' \in Y$ is any point then

$$|y' - x| = \left| \frac{z - y}{|z - y|} - y' \right| = \frac{|z - y - y'|z - y|\,|}{|z - y|} \geq \frac{1 - \varepsilon}{d} \cdot d = 1 - \varepsilon.$$

❑

We need another result first.

Lemma 7.2.3 *Assume $(X, |\,.\,|)$ is a Banach space and $T \in L(X)$ a compact operator. Then the operator $Id - T$ has finite dimensional kernel and closed range. Moreover, if $Id - T$ is injective then it is also surjective.*

The above lemma makes up the first half of the proof of the following theorem.

Theorem 7.2.4 *Assume $(X, |\,.\,|)$ is a Banach space and $T \in L(X)$ a compact operator. Then the operator $Id - T$ is a Fredholm operator of index zero, i.e. its kernel and its cokernel $X/R(Id - T)$ are finite dimensional, its range is closed and*

$$\mathit{index}(Id - T) = \mathit{dim}(\ker(Id - T)) - \mathit{dim}(\mathit{coker}(Id - T)) = 0.$$

Remark: Lemma 7.2.3 implies the Fredholm alternative: The equation

$$(\mathrm{Id} - T)x = y$$

has either a unique solution or there are finitely many linearly independent solutions to the homogeneous equation $(\mathrm{Id} - T)x = 0$.

Proof:
(Lemma 7.2.3)
We organise the proof in several steps.
First step: Show that $\mathbf{Id} - T$ has finite dimensional kernel

If $x \in \ker(\mathrm{Id} - T)$ then $x = Tx$ and

$$B_1(0) \cap \ker(\mathrm{Id} - T) \subset T(B_1(0)).$$

Since T is a compact operator the set $T(B_1(0))$ is precompact, hence the unit ball in $\ker(\mathrm{Id} - T)$ is precompact. But this is only possible if $\ker(\mathrm{Id} - T)$ is finite dimensional.

Second step: Show that the range of $\mathbf{Id} - T$ is closed

Assume that $x \in \overline{R(\mathrm{Id} - T)}$ and that $(\mathrm{Id} - T)x_n \to x$ for a suitable sequence $(x_n)_{n \in \mathbf{N}} \subset X$. We define

$$d_n := \mathrm{dist}(x_n, \ker(\mathrm{Id} - T))$$

and we may assume without loss of generality that $|x_n| \leq 2\, d_n$. Otherwise, if this is not true, then we may pick $y_n \in \ker(\mathrm{Id} - T)$ such that $|x_n - y_n| \leq 2\, d_n$ and we consider $\tilde{x}_n := x_n - y_n$ instead of x_n. We first assume that the sequence (d_n) is not bounded. Then we may assume that $d_n \to \infty$ after passing to a suitable subsequence. Defining $z_n := x_n/d_n$ we obtain

$$(\mathrm{Id} - T)z_n = \frac{(\mathrm{Id} - T)x_n}{d_n} \longrightarrow 0$$

since $((\mathrm{Id} - T)x_n)$ is a bounded sequence. We have $|z_n| \leq 2$ and T is a compact operator, hence the sequence Tz_n has a convergent subsequence, i.e. assume that $Tz_n \to z$ after passing to some subsequence. We conclude

$$z_n = (\mathrm{Id} - T)z_n + Tz_n \longrightarrow z$$

and $(\mathrm{Id}-T)z = 0$ since T is continuous. But then

$$|z_n - z| \geq \operatorname{dist}(z_n, \ker(\mathrm{Id}-T)) = \operatorname{dist}(\frac{x_n}{d_n}, \ker(\mathrm{Id}-T)) = \frac{\operatorname{dist}(x_n, \ker(\mathrm{Id}-T))}{d_n} = 1,$$

a contradiction. So we have shown that the sequence (d_n) must be bounded, which also implies that the sequence (x_n) is bounded. After passing to a subsequence the sequence (Tx_n) then converges to some $z \in X$. But then

$$(\mathrm{Id}-T)x_n = (\mathrm{Id}-T)\left[(\mathrm{Id}-T)x_n + Tx_n\right] \longrightarrow (\mathrm{Id}-T)(x+z)$$

and also

$$(\mathrm{Id}-T)x_n \longrightarrow x$$

hence x is in the range of $\mathrm{Id}-T$.

Third Step: Show that injectivity of Id$-T$ implies surjectivity of Id$-T$

We argue indirectly and assume that there is some $x \in X \backslash R(\mathrm{Id}-T)$. We claim that $(\mathrm{Id}-T)^n x \in R((\mathrm{Id}-T)^n) \backslash R((\mathrm{Id}-T)^{n+1})$. Indeed if we had $(\mathrm{Id}-T)^n x = (\mathrm{Id}-T)^{n+1} y$ for some n and some y then

$$(\mathrm{Id}-T)^n[x - (\mathrm{Id}-T)y] = 0$$

and by injectivity of $\mathrm{Id}-T$ we conclude that $x - (\mathrm{Id}-T)y = 0$, i.e. x has to be in the range of $\mathrm{Id}-T$ which is a contradiction to our assumptions. This proves the claim that

$$(\mathrm{Id}-T)^n x \in R((\mathrm{Id}-T)^n) \backslash R((\mathrm{Id}-T)^{n+1}).$$

We also claim that the range of $(\mathrm{Id}-T)^{n+1}$ is closed as well. Indeed, the operator $(\mathrm{Id}-T)^{n+1}$ can be written as identity plus some compact operator, and we have just shown in the second step that such operators have closed range. We have

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

and we recall that a composition of a compact operator with a linear continuous operator has to be compact as well. We may now pick points $a_{n+1} \in R((\mathrm{Id}-T)^{n+1})$ such that

$$|(\mathrm{Id}-T)^n x - a_{n+1}| \leq 2 \operatorname{dist}((\mathrm{Id}-T)^n x, R((\mathrm{Id}-T)^{n+1}))) \neq 0.$$

Considering

$$x_n := \frac{(\mathrm{Id}-T)^n x - a_{n+1}}{|(\mathrm{Id}-T)^n x - a_{n+1}|}$$

we estimate for $y \in R((\mathrm{Id}-T)^{n+1})$

$$\begin{aligned} |x_n - y| &= \frac{\Big|(\mathrm{Id}-T)^n x - [\, a_{n+1} + |(\mathrm{Id}-T)^n x - a_{n+1}| \cdot y \,]\Big|}{|(\mathrm{Id}-T)^n x - a_{n+1}|} \\ &\geq \frac{\mathrm{dist}\Big((\mathrm{Id}-T)^n x, R((\mathrm{Id}-T)^{n+1})\Big)}{|(\mathrm{Id}-T)^n x - a_{n+1}|} \\ &\geq \frac{1}{2}. \end{aligned}$$

For $m > n$ we get

$$|Tx_n - Tx_m| = |x_n - (\,(\mathrm{Id}-T)x_n + x_m - (\mathrm{Id}-T)x_m)\,| \geq \frac{1}{2}$$

because $(\mathrm{Id}-T)x_n + x_m - (\mathrm{Id}-T)x_m)$ is in the range of $(\mathrm{Id}-T)^{n+1}$ if $m > n$. This means that the sequence (Tx_n) does not have any convergent subsequence. On the other hand, the sequence (x_n) is bounded and T is a compact operator, a contradiction.

❑

We are now able to prove the first part of theorem 7.2.1. Assume that $0 \neq \lambda \notin \sigma_p(T)$. Then the kernel of $\mathrm{Id} - \frac{T}{\lambda}$ is trivial, and by lemma 7.2.3, the operator $\mathrm{Id} - \frac{T}{\lambda}$ is also surjective. Hence $\lambda \in \rho(T)$, and we have shown that

$$\sigma(T)\backslash\{0\} \subset \sigma_p(T),$$

i.e. every nonzero point in the spectrum must be an eigenvalue. Assume now that the set $\sigma(T)\backslash\{0\}$ is not finite. Then we pick pairwise distinct eigenvalues $\lambda_n \in \sigma(T)\backslash\{0\}$ and corresponding eigenvectors $e_n \neq 0$. Define

$$X_n := \mathrm{Span}\{e_1, \dots, e_n\}.$$

We leave it as an exercise to the reader to show that the eigenvectors e_n are linear independent because they all correspond to different eigenvalues. Hence X_{n-1} is a proper subspace of X_n. Using the 'almost approximation lemma' we find $x_n \in X_n$ with

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

We may write $x_n = a_n e_n + \tilde{x}_n$ for suitable a suitable vector $\tilde{x}_n \in X_{n-1}$ and some scalar a_n. The subspace X_{n-1} is invariant under the operator T by definition, hence

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

We then estimate for $m < n$

$$\left| T(\frac{x_n}{\lambda_n}) - T(\frac{x_m}{\lambda_m}) \right| = \left| x_n + \frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - \frac{1}{\lambda_m} T x_m \right| \geq \frac{1}{2}.$$

Hence the sequence (Tx_n/λ_n) has no convergent subsequence. Because T is compact the sequence x_n/λ_n can not contain any bounded subsequence. This implies that

$$\frac{1}{|\lambda_n|} = \left|\frac{x_n}{\lambda_n}\right| \longrightarrow \infty,$$

i.e. $\lambda_n \longrightarrow 0$ which implies that 0 is the only accumulation point of the set $\sigma(T)\backslash\{0\}$. In particular, the set $\sigma(T)\backslash B_r(0)$ must be finite for any $r > 0$, therefore $\sigma(T)\backslash\{0\}$ is countable.