Variational calculus

Winter 2022/23, Tübingen

Sascha Eichmann

April 6, 2023

These lecture notes are mainly based on a lecture given by Hans-Christoph Grunau in 2014 in Magdeburg, see [5], which itself is based on [11]. Another great source is the book by Evans, [1], which encompasses a big part of the content in this lecture.

Contents

Part I: Direct Method of variational calculus

1 Introduction

One of the most classical boundary value problems in the theory of partial differential equations is the Dirichlet problem for the Laplace operator: Given an $f \in C^0(\overline{\Omega})$ with $\Omega \subseteq \mathbb{R}^n$ open and bounded, we look for a $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ such that

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{1.1}$$

To show existence, we introduce the so called Dirichlet energy $E : C^1(\overline{\Omega}) \to \mathbb{R}$ by

$$E(v) := \frac{1}{2} \int_\Omega |\nabla v|^2 - f \cdot v \, d\mathcal{L}^n. \tag{1.2}$$

Let us assume, we have found a minimiser u of E in the following set

$$u \in M := \{ v \in C^2(\overline{\Omega}) | \ v|_{\partial\Omega} = 0 \},$$

i.e. for all $v \in M$ we have $E(u) \leq E(v)$. Now let $\varphi \in C_0^\infty(\Omega)$ be arbitrary. Then the one-dimensional function

$$t \mapsto E(u + t\varphi)$$

has a minimum at $t = 0$, hence

$$\frac{d}{dt} E(u + t\varphi)|_{t=0} = 0.$$

Let us calculate this derivative:

$$\begin{aligned}
&\frac{d}{dt} E(u + t\varphi)|_{t=0} = \frac{d}{dt} \frac{1}{2} \int_\Omega |\nabla u + t \nabla \varphi|^2 \, d\mathcal{L}^n - \int_\Omega f(u + t\varphi) \, d\mathcal{L}^n |_{t=0} \\
=& \frac{d}{dt} \left(\frac{1}{2} \int_\Omega |\nabla u|^2 \, d\mathcal{L}^n + \frac{1}{2} 2t \int_\Omega \langle \nabla u, \nabla \varphi \rangle \, d\mathcal{L}^n + \frac{1}{2} t^2 \int_\Omega |\nabla \varphi|^2 \, d\mathcal{L}^n \right) |_{t=0} \\
&- \int_\Omega f \varphi \, d\mathcal{L}^n \\
=& \int_\Omega \langle \nabla u, \nabla \varphi \rangle \, d\mathcal{L}^n - \int_\Omega f \varphi \, d\mathcal{L}^n \overset{p.I.}{=} - \int_\Omega \Delta u \varphi \, d\mathcal{L}^n - \int_\Omega f \varphi \, d\mathcal{L}^n \\
=& \int \varphi(-\Delta u - f) \, d\mathcal{L}^n.
\end{aligned}$$

Since $\varphi \in C_0^\infty(\Omega)$ is arbitrary, the fundamental lemma of the calculus of variations (exercise) yields

$$-\Delta u = f.$$

Since u is trivially zero on the boundary, the boundary data hold as well. Hence u solves (1.1). We say, that u is a critical point of E or that u satisfies the Euler-Lagrange equation of E.

The first part of the lecture is now concerned with showing the existence of such minimisers. The method to do this is called the direct method of the calculus of variations. The main idea is already encapsulated in the proof of the following Theorem, which is usually done in Analysis 1:

Theorem 1.1. *Let $a < b$ and $E : [a,b] \to \mathbb{R}$ be lower semicontinuous, i.e. for all converging sequences $x_k \in [a,b]$ with limit $x \in [a,b]$ we have*

$$E(x) \leq \liminf_{k\to\infty} E(x_k).$$

Then there exists an $x_{Min} \in [a,b]$ such that

$$E(x_{Min}) = \inf_{x\in[a,b]} E(x).$$

Proof. Let $x_k \in [a,b]$ be a minimising sequence for E, i.e.

$$E(x_k) \to \inf_{x\in[a,b]} E(x).$$

Since $[a,b] \subseteq \mathbb{R}$ is bounded and closed (hence it is sequentially compact), we find a converging subsequence, such that after relabeling we have

$$x_k \to x_{Min} \in [a,b].$$

By lower semicontinuity we get

$$\inf_{x\in[a,b]} E(x) \leq E(x_{Min}) \leq \liminf_{k\to\infty} E(x_k) = \inf_{x\in[a,b]} E(x).$$

Hence x_{Min} is a desired minimiser. □

The plan of the lecture is now as follows:
Since we like to adapt the proof of Thm. 1.1 to function spaces, which are usually infinite dimensional, the classical notion of sequential compactness and convergence will not suffice. In §2 we will therefore define another notion of convergence called weak convergence and introduce a theorem, which will guarantee good compactness properties. In §3 we will examine Sobolev spaces, which are suitable function spaces for our direct method. This will also lead to the notion of weak solutions.
Starting from section 4 we will make the notion of a critical point more precise and introduce a so called mountain-pass lemma. This lemma will guarantee the existence of such critical points of saddle point type and is usually employed to show nonuniqueness.
Due to time constraints, we will not be able to prove everything, but there will be appropriate references.

Remark 1.2.

1. The Dirichlet energy (1.2) only controls the first derivative, hence we cannot expect a-priori that a limit of a minimising sequence will have a second derivative. Therefore we will introduce so called weak solutions in section 3.
   This principle will also be applied to other equations. Depending on the specific equation, one may be able to do some regularity theory, see e.g. [4, Chapter 8] to obtain more derivatives. This is out of the scope of this lecture though.

2. These energies, which we will minimise, usually do have a physical background. In Physics these type of energies are usually called actions.

2 Functionalanalytic background

We repeat some of the most important facts for variational calculus about Banach spaces. These can be found in e.g. [8].

Definition 2.1. Let B be a real vector space. Furthermore let $\|\cdot\| : B \to [0, \infty)$ be a norm. We then call $(B, \|\cdot\|)$ a normed space. If $(B, \|\cdot\|)$ is complete w.r.t. to $\|\cdot\|$, we call it a Banach space.
If additionally $\|\cdot\|$ is induced by an inner product $\langle\cdot,\cdot\rangle$, we call $(B, \langle\cdot,\cdot\rangle)$ a Hilbert space.

Our goal is to find a suitable notion of convergence, for which we will have good compactness properties. To this end we need dual spaces:

Definition 2.2. Let $(B, \|\cdot\|)$ be a normed real vectorspace. Then we define the dualspace B^* of B to consist of all linear continuous functionals, i.e.

$$B^* := \{L : B \to \mathbb{R} |\ L \text{ continuous w.r.t. } \|\cdot\|\}.$$

We equip B^* with the following operatornorm:

$$\|L\|_{B^*} := \sup_{x \in B,\ \|x\| \leq 1} |Lx|.$$

Remark 2.3. Let $(B, \|\cdot\|)$ be a normed space. Then we have the following results (see also the exercises)

1. Let $L : B \to \mathbb{R}$ be linear. Then L is continuous if and only if $\|L\|_{B^*} < \infty$ or equivalently there exists a $C > 0$ such that for all $x \in B$

   $$|L(x)| \leq C\|x\|.$$

   If L satisfies this last condition, it is called bounded.

2. Since $\mathbb{R}$ is complete, the normed vectorspace $(B^*, \|\cdot\|_{B^*})$ is always complete, i.e. a Banach space, see e.g. [8, Thm. 4.1].

3. If $(B, \langle\cdot,\cdot\rangle)$ is a Hilbert space, the Riesz representation theorem states (see e.g. [4, Thm .5.7]) that for any $L \in B^*$ we find a unique $x_L \in B$ such that for all $y \in B$ we have

   $$L(y) = \langle x_L, y\rangle \text{ and } \|L\|_{B^*} = \|x_L\|.$$

   Hence B^* is isometric to B and we can define an inner product on B^* by

   $$\langle L, G\rangle_{B^*} := \langle x_L, x_G\rangle_B = L(x_G) = G(x_L).$$

Example 2.4. Let μ be a σ-finite measure on some set Ω and $1 \leq p < \infty$. Then by the Riesz-Fischer Theorem the space of p-integrable functions $L^p(\mu)$ is a Banach space. Furthermore by [9, Thm. 6.16] we can characterise the dual space:
Let $q \in (1, \infty]$ such that

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

Then the map

$$\Lambda_p : L^q(\mu) \to (L^p(\mu))^*$$

given by

$$\Lambda_p(g)(f) := \int_\Omega g \cdot f \, d\mu$$

is a bijective isometry.

The following theorem (or set of theorems) is a very important buildung block in functional analysis, see [8, §3.1] for proofs.

Theorem 2.5 (Hahn-Banach)**.** *Let $(B, \|\cdot\|)$ be a normed vectorspace. Then we have the following theorems:*

1. *Let $p : B \to \mathbb{R}$ be sublinear, i.e. satisfy*

$$\text{(i) } p(x+y) \le p(x) + p(y) \ \forall x, y \in B$$
$$\text{(ii) } p(\alpha x) = \alpha p(x) \ \forall \alpha \ge 0, \ \forall x \in B.$$

   Let $M \subseteq B$ be linear subspace, $\ell : M \to \mathbb{R}$ linear such that for all $x \in M$ we have

$$\ell(x) \le p(x).$$

   Then there exists an extension $L : B \to \mathbb{R}$ (i.e. $L|_M = \ell$) such that L is linear and for all $x \in B$ we have

$$L(x) \le p(x).$$

2. *Let $W \subseteq B$ be a subspace. Then every bounded linear functional $\ell \in W^*$ admits an extension $L \in B^*$ (i.e. $L|_W = \ell$) and such that*

$$\|\ell\|_{W^*} = \|L\|_{B^*}.$$

3. *(Seperation Theorem) Let $A, C \subseteq B$ be non-empty disjoint convex sets. If A is open, then there exists an $L \in B^*$ and $\gamma \in \mathbb{R}$, such that*

$$\forall a \in A \ \forall c \in C : \ L(a) < \gamma \le L(c).$$

   If A is compact and C is closed, there also exists an $L \in B^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\forall a \in A \ \forall c \in C : \ L(a) < \gamma_1 < \gamma_2 < L(c).$$

Before we proceed we apply the above theorem to show a useful characterisation of a norm:

Lemma 2.6. *Let $(B, \|\cdot\|)$ be a normed vectorspace. Then for every $x \in B$ we have*

$$\|x\| = \sup_{L \in B^*, \|L\|_{B^*} \le 1} |L(x)|. \tag{2.1}$$

Proof. W.l.o.g we assume $x \neq 0$. By the definition of the operatornorm we have

$$\sup_{L \in B^*, \|L\|_{B^*} \le 1} |L(x)| \le \sup_{L \in B^*, \|L\|_{B^*} \le 1} \|L\|_{B^*} \|x\| \le \|x\|.$$

The other estimates needs the Hahn-Banach theorem 2.5. We define a one-dimensional subspace $V \subseteq B$ by

$$V := \operatorname{span}(x)$$

and a linear continuous functional $L_V : V \to \mathbb{R}$ by

$$L_V(x) := \|x\|.$$

Then for any $y \in V$ we find a unique $r_y \in \mathbb{R}$, such that $y = r_y x$. Hence

$$L_V(y) = r_y L_V(x) = r_y \|x\|,$$

by which we get

$$\|L_V\|_{V^*} = \sup_{y \in V \setminus \{0\}} \frac{|L_V(y)|}{\|y\|} = \sup_{y \in V \setminus \{0\}} \frac{|r_y|}{\|y\|} |L_V(x)| = \sup_{y \in V \setminus \{0\}} \frac{\|y\|}{\|x\|\|y\|} \|x\| = 1.$$

By Hahn-Banach we can extend L_V to a continuous linear functional on the whole of B with the same operatornorm. Hence

$$\|x\| = |L_V(x)| \leq \sup_{L \in B^*, \|L\|_{B^*} \leq 1} |L(x)|,$$

which finishes the proof. □

From this point on $(B, \|\cdot\|)$ is always a Banach space (although some results and Definitions carry over to normed spaces).
We introduce the notion of convergence, which we will mainly use in our direct method of variations. As we will see in Theorem 2.14 this notion will enjoy good compactness properties.

Definition 2.7. Let $(B, \|\cdot\|)$ be a Banach space.

1. Let $x_k, x \in B$. We say $x_k \to x$ weakly, if and only if for all $L \in B^*$ we have
$$L(x_k) \to L(x).$$
2. Let $L_k, L \in B^*$. We say $L_k \to L$ weak* (or in a weak* sense), if and only if for all $x \in B$ we have
$$L_k(x) \to L(x).$$

Example 2.8. Let ν be the counting measure on $\mathbb{N}$. We define the sequence $e_i \in L^2(\nu)$ to be

$$e_i(j) := \delta_{ij}.$$

Then

$$e_i \to 0 \text{ weakly,}$$

but e_i and every subsequence does not converge w.r.t. $\|\cdot\|_{L^2(\nu)}$.

Proof. First we check, that the sequence does not converge in the norm:

$$\|e_i - e_k\|^2_{L^2(\nu)} = \int_{\mathbb{N}} |e_i - e_k|^2 \, d\nu = \sum_{j=1}^{\infty} |e_i(j) - e_k(j)|^2 = \sum_{j=1}^{\infty} |\delta_{ij} - \delta_{kj}|^2 = 2(1 - \delta_{ik}).$$

Hence this sequence and all its subsequences are not Cauchy and therefore cannot converge.
For the weak convergence we use Remark 2.3 3). Hence we have to check for every $f \in L^2(\nu)$, that

$$\langle f, e_i \rangle_{L^2(\nu)} = \int f \cdot e_i \, d\nu = \sum_{j=1}^{\infty} f(j) e_i(j) = f(i)$$

converges to zero. Since $f \in L^2(\nu)$ we have

$$\|f\|_{L^2(\nu)}^2 = \sum_{j=1}^{\infty} |f(j)|^2 < \infty,$$

hence

$$\lim_{j \to \infty} f(j) = 0$$

and the result follows. $\square$

Remark 2.9. Let $(B, \|\cdot\|)$ be a Banach space. Then we have the following properties concerning weak and weak* convergence (see also the exercises):

1. Limits of weak and weak* converging sequences are unique, see e.g. [8, § 3.1].
2. If $x_k \in B$ converges w.r.t. to $\|\cdot\|$ it also weakly converges to the same limit (Exercise).
3. If $x_k \in B$ weakly converges, then the sequence is bounded (see [8, Thm 3.18]), i.e. there exists a $C > 0$ such that for all $k \in \mathbb{N}$ we have

$$\|x_k\| \leq C.$$

Not all Banach spaces will enjoy good compactness properties. The ones we are examining here are called reflexive. To define this, we need the canonical embedding or also called evaluation map:

Definition 2.10. Let $(B, \|\cdot\|)$ be a Banach space. We call

$$B^{**} := (B^*)^*$$

the bidual of B. Furthermore we define the canonical embedding $i_B : B \to B^{**}$ to be

$$i_B(x)(L) = L(x).$$

Remark 2.11. By Hahn-Banach Theorem 2.5 rsp. Lemma 2.6 we see that i_B is always linear, injective and norm preserving (exercise).

Now we can define one of the central assumptions for the direct method to work.

Definition 2.12. Let $(B, \|\cdot\|)$ be a Banach space. We call it reflexive if the canonical embedding i_B is surjective.

Example 2.13. We have the following examples of reflexive spaces:

1. By Rieszs representation theorem (see Remark 2.3 3)), every Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$ is reflexive:
   Let $q^{**} \in H^{**}$ be arbitrary. Then by the Riesz representation theorem we find a unique $q^* \in H^*$ such that for all $z^* \in H^*$ we have
   $$\langle z^*, q^*\rangle_{H^*} = q^{**}(z^*).$$
   Repeating this step, yields a $q \in H$, such that for every $z \in H$ we have
   $$\langle z, q\rangle_H = q^*(z).$$
   For every $L \in H^*$ we find a $z_L \in H$ with
   $$L(q) = \langle z_L, q\rangle_H$$
   Hence applying the canonical embedding to q yields for every $L \in H^*$
   $$i_H(q)(L) = L(q) = \langle z_L, q\rangle_H = q^*(z_L) = \langle q^*, L\rangle_{H^*} = q^{**}(L)$$
   hence $i_H(q) = q^{**}$ and the result follows.

2. Let $1 < p < \infty$ and μ be a σ-finite measure on some set Ω. Then by Example 2.4 $L^p(\mu)$ is reflexive:
   Let $\Lambda_p : L^q(\mu) \to (L^p(\mu))^*$ (again with $\frac{1}{p} + \frac{1}{q} = 1$) be as in Example 2.4, i.e. an isometry given by
   $$\Lambda_p(g)(f) = \int gf\, d\mu.$$
   Now let $f^{**} \in (L^p(\mu))^{**}$ be arbitrary but fixated. Since Λ_p is bijective, we can define an $f^* := f^{**} \circ \Lambda_p \in (L^q(\mu))^*$, and for all $g^* \in (L^p(\mu))^*$ we find a unique $g \in L^q(\mu)$ with $g^* = \Lambda_p(g)$. Therefore
   $$f^{**}(g^*) = f^{**}(\Lambda_p(g)) = f^{**} \circ \Lambda_p(g) =: f^*(g).$$
   Let $f := \Lambda_q^{-1}(f^*) \in L^p(\mu)$. Then
   $$f^{**}(g^*) = f^*(g) = \Lambda_q(f)(g) = \int fg\, d\mu.$$
   Hence
   $$i_{L^p(\mu)}(f)(g^*) = g^*(f) = \Lambda_p(g)(f) = \int fg\, d\mu = f^{**}(g^*)$$
   and therefore $i_{L^p(\mu)}(f) = f^{**}$ and this map is surjective.

The following theorem encompasses the needed compactness properties. Since we do not have the time needed for a complete proof, we will just give a citation.

Theorem 2.14. *Let $(B, \|\cdot\|)$ be a Banach space. Then*

1. *(see [8, § 4, Ex. 1 c)]) Let B be reflexive. Let $x_k \in B$ be a bounded sequence w.r.t. to $\|\cdot\|$. Then there exists a weakly converging subsequence of x_k.*

2. *(Banach-Alaoglu, [8, Thm. 3.15 and Thm. 3.16]) Let B be seperable (i.e. there exists a countable dense set in B). Let $L_k \in B^*$ be a bounded sequence w.r.t. $\|\cdot\|_{B^*}$. Then there exists a subsequence, which converges in the weak* sense.*

Now we turn our focus to lower semicontinuity:

Definition 2.15. Let $(B, \|\cdot\|)$ be a Banach space. Furthermore let $X \subseteq B$ and $F : X \to \mathbb{R}$. We call F weakly lower semicontinuous, if for all sequences $x_k \in X$ weakly converging to an $x \in X$, we have

$$F(x) \leq \liminf_{k\to\infty} F(x_k)$$

The following gives a very important example for weakly lower semicontinuous functions:

Theorem 2.16. *Let $(B, \|\cdot\|)$ be a Banach space. Then the norm $\|\cdot\|$ is weakly lower semicontinuous on B.*

Proof. Let $x_k \in B$ with $x_k \to x \in B$ weakly. With the help of Hahn-Banachs theorem 2.5 rsp. Lemma 2.6 we can write

$$\|x\| = \sup_{L\in B^*, \|L\|\leq 1} |L(x)|.$$

Now for every $\varepsilon > 0$ we can find an $\tilde{L} \in B^*$ with $\|\tilde{L}\| \leq 1$ such that

$$\begin{aligned} \|x\| &\leq \tilde{L}(x) + \varepsilon = \lim_{k\to\infty} \tilde{L}(x_k) + \varepsilon \leq \liminf_{k\to\infty} \sup_{L\in B^*, \|L\|\leq 1} |L(x_k)| + \varepsilon \\ &= \liminf_{k\to\infty} \|x_k\| + \varepsilon. \end{aligned}$$

Letting $\varepsilon \downarrow 0$ finishes the proof of the lower semicontinuity. □

Now we are ready to prove a theorem encompassing the direct method. The proof itself is more important than the theorem itself, since one usually has to adapt the method to the situation at hand.

Theorem 2.17 (Direct method)**.** *Let $(B, \|\cdot\|)$ be a reflexive Banach space, $\emptyset \neq X \subseteq B$ weakly sequentially closed, i.e. for all $x_k \in X$ weakly converging to an $x \in B$ we have $x \in X$ and $F : X \to \mathbb{R}$. Furthermore F should satisfy*

1. *F is coercive, i.e. if $y_k \in X$ with $\|y_k\| \to \infty$, then $F(y_k) \to \infty$.*

2. *F is weakly lower semicontinuous.*

Then F is bounded below and admits a minimum in X, i.e. we find an $x_{min} \in X$, such that

$$F(x_{min}) = \inf_{x\in X} F(x).$$

Proof. We have $\inf_{x\in X} F(x) < \infty$ and let $x_k \in X$ be a minimising sequence for F, i.e.

$$F(x_k) \to \inf_{x\in X} F(x) < \infty.$$

Suppose $\|x_k\| \to \infty$. Since F is coercive we would have $F(x_k) \to \infty$, a contradiction. Hence we find a $C > 0$, such that for all $k \in \mathbb{N}$

$$\|x_k\| \leq C.$$

By Theorem 2.14 we find a weakly converging subsequence with limit $y \in B$, i.e. after relabeling

$$x_k \to y \text{ weakly.}$$

Since X is weakly sequentially closed, we have $y \in X$. Since F is weakly lower semicontinuous, we finally have

$$\inf_{x \in X} F(x) \leq F(y) \leq \liminf_{k \to \infty} F(x_k) = \inf_{x \in X} F(x).$$

Hence we have equality and y is a desired minimiser. $\square$

The next theorem identifies weakly sequentially closed sets and shows that convexity in the context of variational calculus is usually desired.

Theorem 2.18. *Let $(B, \|\cdot\|)$ be a Banach space. Let $X \subseteq B$ be closed and convex. Then X is sequentially weakly closed.*

Proof. We proceed by contradiction and assume we find a sequence $x_k \in X$ weakly converging to an $x \in B \setminus X$. By Hahn-Banach 2.5 3) we find an $L \in B^*$, and $\gamma_1, \gamma_2 \in \mathbb{R}$, such that for all $y \in X$ we have

$$L(x) < \gamma_1 < \gamma_2 < L(y).$$

Hence

$$\gamma_2 < L(x_k) \to L(x) < \gamma_1,$$

which yields

$$\gamma_2 \leq \gamma_1,$$

a contradiction. $\square$

3 Sobolev Spaces and weak solutions to PDEs

In this section we introduce a notion of weak derivatives and weak solutions to differential equations. As before due to time constraints we will not be able to prove everything. Details can be found in e.g. [4, Chapter 7] or [2, Chapter 4]. We start with the notion of weak derivatives, which are inspired by the fundamental lemma of variational calculus and partial integration:

Definition 3.1. Let $\Omega \subseteq \mathbb{R}^n$ be open, $u \in L^1_{loc}(\Omega)$ and α be a multiindex. We then call a locally integrable function $v \in L^1_{loc}(\Omega)$ the α^{th} weak derivative of u, if for all $\varphi \in C_0^{|\alpha|}(\Omega)$ we have

$$\int_\Omega \varphi v \, d\mathcal{L}^n = (-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi \, d\mathcal{L}^n.$$

In this case we write

$$D^\alpha u := v.$$

Remark 3.2.

1. Every function $u \in C^k(\Omega)$ possesses weak derivatives to the k-th order. These are given by the classical derivatives of u. To show this, let α be a multiindex with $|\alpha| \leq k$ and $\varphi \in C_0^{|\alpha|}(\Omega)$ be arbitrary. In the following calculation we denote with $D^\alpha u$ the classical derivative. By partial integration, the compact support of φ and Schwarzes theorem we obtain

$$\begin{aligned}
&\int D^\alpha u \varphi \, d\mathcal{L}^n = \int D^{\alpha_1} D^{\alpha_2} \ldots D^{\alpha_{|\alpha|}} u \varphi \, d\mathcal{L}^n \\
=&(-1) \int D^{\alpha_2} \ldots D^{\alpha_{|\alpha|}} u D^{\alpha_1} \varphi \, d\mathcal{L}^n = \ldots = (-1)^{|\alpha|} \int u D^{\alpha_{|\alpha|}} \ldots D^{\alpha_1} \varphi \, d\mathcal{L}^n \\
=&(-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi \, d\mathcal{L}^n.
\end{aligned}$$

   Hence the classical derivatives are weak derivatives as well.

2. The weak derivative is unique up to a set of Lebesgue measure zero (Exercise).

3. Weak derivatives always interchange, i.e. we always have up to a set of Lebesgue measure zero

$$D^\alpha D^\beta u = D^\beta D^\alpha u,$$

   if the weak derivatives exist (Exercise).

4. By ∇u we also denote the weak gradient and with Du the first weak derivative, i.e.

$$Du = (D^{(1)} u, \ldots, D^{(n)} u) = (\nabla u)^T.$$

5. With the help of some smoothing techniques, it is actually enough to require $\varphi \in C_0^\infty(\Omega)$ instead of $C_0^{|\alpha|}(\Omega)$.

The following example illustrates, that even if a function is smooth outside of a set of zero measure, it may not be weakly differentiable:

Example 3.3. Let $u : \mathbb{R} \to \mathbb{R}$ be given by $u(x) = |x|$. Then u does possess a first weak derivative, but is not twice weakly differentiable on $\mathbb{R}$:
By Remark 3.2 1), the weak derivative (if it exists) should coincide with the classical derivative, where it exists. Outside of $x = 0$ the function u is smooth. Hence we make an Ansatz for the first derivative $v : \mathbb{R} \to \mathbb{R}$ to be

$$v(x) = \operatorname{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

Now let $\varphi \in C_0^1(\mathbb{R})$. Then we have by partial integration

$$\begin{aligned}
&\int_{\mathbb{R}} v\varphi\, dx = \int_{-\infty}^{0} v\varphi\, dx + \int_{0}^{\infty} v\varphi\, dx \\
=&(-1)\int_{-\infty}^{0} d_x(x)\varphi\, dx + \int_{0}^{\infty} d_x(x)\varphi\, dx \\
=&(-1)\varphi(0)\cdot 0 - (-1)\int_{-\infty}^{0} xD^{(1)}\varphi\, dx - \varphi(0)\cdot 0 - \int_{0}^{\infty} xD^{(1)}\varphi\, dx \\
=&-\int_{-\infty}^{0} |x|D^{(1)}\varphi(x)\, dx - \int_{0}^{\infty} |x|D^{(1)}\varphi(x)\, dx = \int_{\mathbb{R}} u(x)D^{(1)}\varphi\, dx.
\end{aligned}$$

Hence v is the first weak derivative of u on $\mathbb{R}$, although u is not differentiable at zero.
Now we show, that v does not possess a weak derivative: We proceed by contradiction and assume $g \in L^1_{loc}(\mathbb{R})$ is such a weak derivative. Choosing a $\varphi \in C_0^1((0,\infty))$ yields

$$\int_{\mathbb{R}} g\varphi\, dx = -\int_{\mathbb{R}} vD^{(1)}\varphi\, dx = -\int_{0}^{\infty} D^{(1)}\varphi\, dx = 0$$

by the Fundamental theorem of calculus. Then the fundamental lemma of variational calculus yields

$$g|_{(0,\infty)} = 0,\ \mathcal{L}^1 - \text{a.e.}.$$

The same argument works on $(-\infty, 0)$, hence $g = 0$ almost everywhere. Now the definition of a weak derivative yields for a $\varphi \in C_0^1(\mathbb{R})$:

$$\begin{aligned}
0 = &\int_{\mathbb{R}} g\varphi\, dx = -\int_{\mathbb{R}} vD^{(1)}\varphi\, dx = -\left(\int_{-\infty}^{0} (-1)D^{(1)}\varphi\, dx + \int_{0}^{\infty} D^{(1)}\varphi\, dx\right) \\
=&-\left((-1)\varphi(0) + (-1)\varphi(0)\right) = 2\varphi(0).
\end{aligned}$$

Choosing a $\varphi \in C_0^1(\mathbb{R})$ with $\varphi(0) \neq 0$ results in a contradiction.

We summarize some rules for calculating weak derivatives. A proof of these rules would involve a mollification scheme, see e.g. [4, Ch. 7.2, Thm 7.4], hence it is out of scope of this lecture.

Theorem 3.4. *Let $\Omega \subseteq \mathbb{R}^n$ be open, furthermore let $u, v \in L^1_{loc}(\Omega)$. Then*

1. *If u, v are weakly differentiable w.r.t. the i-th coordinate, then we have a product rule*
$$D^{(i)}(u \cdot v) = D^{(i)}uv + uD^{(i)}v.$$
(see e.g. [4, Eq. 7.18])

2. *If Ω is connected and u is weakly differentiable, i.e. $D^{(1)}u, \dots D^{(n)}u$ exist, we have*
$$u = const \;\Leftrightarrow\; D^{(1)}u = \ldots = D^{(n)}u = 0.$$
(see e.g. [4, Lemma 7.6/7.7])

3. *Let $f \in C^1(\mathbb{R})$, such that $f' \in L^\infty(\mathbb{R})$. Furthermore $D^{(i)}u$ should exists weakly. Then we have a chain rule, i.e. $f \circ u$ is weakly differentiable w.r.t. to the i-th component and*
$$D^{(i)}(f \circ u) = f'(u)D^{(i)}u.$$
(see e.g. [4, Lemma 7.5])

Now we introduce Sobolev spaces, in which we will actually implement our direct method.

Definition 3.5. Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $k \in \mathbb{N}$ and $1 \leq p \leq \infty$. We define the Sobolev space $W^{k,p}(\Omega)$ by

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) |\, D^\alpha u \text{ exists weakly for all } \alpha \text{ with } |\alpha| \leq k,\ D^\alpha u \in L^p(\Omega)\}.$$

If $p < \infty$, we equip this space with the following norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\int_\Omega \sum_{|\alpha| \leq k} |D^\alpha u|^p \, dx \right)^{\frac{1}{p}}.$$

If $p = \infty$, we equip it with

$$\|u\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)}$$

instead.

Remark 3.6. By some standard estimates on finite dimensional spaces, we have some freedom in the definition of $\| \cdot \|_{W^{k,p}(\Omega)}$. For example

$$u \mapsto \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}$$

yields an equivalent norm for all $p \in [1, \infty)$.

Sobolev spaces have the needed functional analytic properties, cf. Theorem 2.14:

Theorem 3.7. *Let $\Omega \subseteq \mathbb{R}^n$ be open, $k \in \mathbb{N}$ and $1 \leq p \leq \infty$. Then we have the following:*

1. *$W^{k,p}(\Omega)$ is a complete normed space, i.e. a Banach space.*

2. *If furthermore $1 < p < \infty$, we have that $W^{k,p}(\Omega)$ is reflexive.*

3. *If $p = 2$, then $W^{k,2}(\Omega)$ is a Hilbert space.*

Proof. Sketch: Identify $W^{k,p}(\Omega)$ with a closed subspace of $(L^p(\Omega))^m$ with a suitable $m \in \mathbb{N}$. Since $(L^p(\Omega))^m$ is complete rsp. reflexive if $1 < p < \infty$ (by e.g. the same argument as in Example 2.13, 2)), the result follows. Working out the details is an exercise. $\square$

Since $\partial\Omega$ is a zero set if it is for example a C^1-submanifold, we cannot prescribe boundary values pointwise. Here we go another route via approximations:

Definition 3.8. Let $\Omega \subseteq \mathbb{R}^n$ be open, $k \in \mathbb{N}$, $1 \leq p \leq \infty$. We define the Sobolev space with zero boundary data $W_0^{k,p}(\Omega)$ as the closure of $C_0^\infty(\Omega) \subseteq W^{k,p}(\Omega)$ w.r.t. $\|\cdot\|_{W^{k,p}(\Omega)}$.

Remark 3.9.

1. By definition $W_0^{k,p}(\Omega) \subseteq W^{k,p}(\Omega)$ is a closed subspace, hence complete and reflexive if $1 < p < \infty$.

2. If $k \geq 1$, we interprete $u \in W_0^{k,p}(\Omega)$ as having zero boundary values, i.e. we think of it as

$$D^\alpha u|_{\partial\Omega} = 0 \text{ for all multiindicees } \alpha \text{ with } |\alpha| \leq k-1,$$

   with $D^0 u := u$.
   This notion can be made precise, if $\partial\Omega$ is of e.g. $C^{k-1,1}$ regularity, see e.g. [2, Ch. 4.3, Thm 1].

3. If we want to prescribe certain nontrivial boundary values g for a Sobolev-function $u \in W^{k,p}(\Omega)$, we assume $g \in W^{k,p}(\Omega)$ and require

$$u - g \in W_0^{k,p}(\Omega).$$

Theorem 3.10 (Poincaré inequality)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded, $1 \leq p < \infty$. Then there exists a $C = C(\mathrm{diam}(\Omega), n, p) > 0$, such that for all $u \in W_0^{1,p}(\Omega)$ we have*

$$\|u\|_{L^p(\Omega)} \leq C \|Du\|_{L^p(\Omega)}.$$

Proof. Sketch: By Definition 3.8 it suffices to show the estimate if $u \in C_0^\infty(\Omega)$. Then write $1 = \frac{\mathrm{div}(x - x_0)}{n}$, use partial integration and the Hölder inequality to obtain the result. Working out the details is an exercise. $\square$

Remark 3.11. Iterating Theorem 3.10 yields that

$$\|u\|_{W_0^{k,p}(\Omega)} := \left(\int_\Omega \sum_{|\alpha| = k} |D^\alpha u|^p \, dx \right)^{\frac{1}{p}}$$

is an equivalent norm to $\|\cdot\|_{W^{k,p}(\Omega)}$ on $W_0^{k,p}(\Omega)$, if $\Omega \subseteq \mathbb{R}^n$ is open and boundend.
A nontrivial constant function shows that $\|\cdot\|_{W_0^{k,p}(\Omega)}$ is not even a norm on $W^{k,p}(\Omega)$, although it is always a semi-norm.

Next we will introduce the so called Sobolev embedding theorems. They give a notion on how good a weakly differentiable function in a classical sense is. More importantly they yield certain compactness results, which will be very valuable for our variational calculus.
First we introduce the notion of compact operators:

Definition 3.12 (Compact operator)**.** Let $(B, \|\cdot\|_B)$, $(V, \|\cdot\|_V)$ be Banach spaces and $T : B \to V$ be linear. We call T compact, if for every bounded sequence $x_k \in B$ there exists a subsequence x_{k_j} and $y \in V$, such that

$$T(x_{k_j}) \to y \text{ w.r.t. } \|\cdot\|_V.$$

Before we can state the embedding, we have to introduce the spaces in which the embedding will happen:

Definition 3.13 (Hölder spaces)**.** Let $\Omega \subseteq \mathbb{R}^n$ be open, $0 < \alpha \leq 1$. We define the Hölder seminorm $|\cdot|_{\alpha,\Omega}$ for a function $u : \Omega \to \mathbb{R}^m$ by

$$|u|_{\alpha,\Omega} := \sup_{x,y\in\Omega, x\neq y} \frac{|u(x) - u(y)|}{|x-y|^\alpha}.$$

We define on $C^k(\Omega)$ for $k \in \mathbb{N}$ the Höldernorm

$$\|u\|_{C^{k,\alpha}(\Omega)} := \|u\|_{C^k(\Omega)} + \sum_{|\beta|\leq k} |D^\beta u|_{\alpha,\Omega}$$

For $k \in \mathbb{N}_0$ we can then define the Hölderspaces

$$C^{k,\alpha}(\Omega) := \{u \in C^k(\Omega) : \|u\|_{C^{k,\alpha}(\Omega)} < \infty\},$$

which is a complete subspace of $C^k(\Omega)$.

Now we can state our embedding result:

Theorem 3.14 (Sobolev embedding, see e.g. [4] p. 168, Thm. 7.26)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open, bounded and connected. $1 \leq p < \infty$, $k \in \mathbb{N}$ and*

1. *Let $\ell \in \mathbb{N}$, $1 \leq q < \infty$ with*

$$k - \frac{n}{p} \geq \ell - \frac{n}{q}.$$

Then the embedding $T : W_0^{k,p}(\Omega) \to W_0^{\ell,q}(\Omega)$ given by $T(u) = u$ is well defined and continuous. Moreover if $k > \ell$ and the the above inequality is strict, i.e.

$$k - \frac{n}{p} > \ell - \frac{n}{q}$$

we even have that T is compact.

2. *If $0 \leq m < k - \frac{n}{p} < m + 1$ the operator $T : W_0^{k,p}(\Omega) \to C^{m,\alpha}(\Omega)$ with $T(u) = u$ is well defined and continuous for all $0 < \alpha \leq k - \frac{n}{p} - m$. It is compact, if we have strict inequality, i.e. $0 < \alpha < k - \frac{n}{p} - m$.*

If Ω is additionally a $C^{0,1}$ domain, i.e. the boundary is Lipschitz, the results above are true for $W^{k,p}(\Omega)$ instead of $W_0^{k,p}(\Omega)$ as well.

Remark: The continuity of T is equivalent to the so called Sobolev inequality: For $\Omega \subset \mathbb{R}^n$ open, bounded and connected exists a $C > 0$, such that for all $u \in W_0^{k,p}(\Omega)$ we have

$$\|u\|_{W_0^{\ell,q}(\Omega)} \leq C\|u\|_{W_0^{k,p}(\Omega)} \text{ rsp.}$$

$$\|u\|_{C^{m,\alpha}(\Omega)} \leq C\|u\|_{W_0^{k,p}(\Omega)}.$$

If $\partial\Omega$ is a $C^{0,1}$-domain, we have the same kind of estimates in $W^{k,p}(\Omega)$.

Now our preparations concerning Sobolev spaces are complete and we turn our attention to weak solutions of PDEs. We start with the Dirichlet problem for the Laplace operator, see (1.1):

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{3.1}$$

Here $\Omega \subseteq \mathbb{R}^n$ open, bounded and $f \in L^2(\Omega)$. The goal is to find a formulation of (3.1), which satisfies the following requirements:

1. In the new formulation a solution does not have to be twice differentiable.
2. If a smooth solution satisfies the reformulation, it is also a classical solution in the sense of (3.1).

For this let $\varphi \in C_0^\infty(\Omega)$. We multiply the differential equation with φ and integrate:

$$\int_\Omega f\varphi\, dx = -\int_\Omega \Delta u\varphi\, dx = \int_\Omega \langle \nabla u, \nabla\varphi\rangle\, dx. \tag{3.2}$$

The last step is partial integration and using that φ has compact support in Ω. This now yields our weak formulation of the partial differential equation:

Definition 3.15 (weak solution)**.** We say a $u : \Omega \to \mathbb{R}$ solves (3.1) weakly, if $u \in W_0^{1,2}(\Omega)$ and such that for all $\varphi \in C_0^\infty(\Omega)$ we have

$$\int_\Omega \langle \nabla u, \nabla\varphi\rangle\, dx = \int_\Omega f\varphi\, dx.$$

Remark 3.16.

1. $u \in W_0^{1,2}(\Omega)$ represents the boundary condition $u|_{\partial\Omega} = 0$.
2. If u is smooth, the calculation in (3.2) can be made backwards and then the fundamental lemma of variational calculus yields u to be a classical solution of (3.1).
3. By density, one can also use $\varphi \in W_0^{1,2}(\Omega)$.
4. φ is usually called test function, i.e. one tests the equation with φ.
5. Since we can write $\Delta = \operatorname{div}\nabla$, we say, that (3.1) is in divergence form. If an equation is in such a form, we can employ partial integration and define a weak formulation of the differential equation.

Our direct method now yields a solution to our new weak formulation:

Theorem 3.17. *Let $\Omega \subseteq \mathbb{R}^n$ be bounded, open and connected. Let $f \in L^2(\Omega)$. Then there exists a weak solution $u \in W_0^{1,2}(\Omega)$ in the sense of Definition 3.15.*

Proof. We define the Dirichlet energy (cf. (1.2)) $E : W_0^{1,2}(\Omega) \to \mathbb{R}$ by

$$E(v) := \frac{1}{2}\int_\Omega |\nabla v|^2\,dx - \int_\Omega f v\,dx. \tag{3.3}$$

If $u \in W_0^{1,2}(\Omega)$ would be a minimiser in $W_0^{1,2}(\Omega)$, i.e.

$$\forall v \in W_0^{1,2}(\Omega)\ E(v) \geq E(u),$$

then the one dimensional function defined by $\varphi \in C_0^\infty(\Omega) \subseteq W_0^{1,2}(\Omega)$

$$t \mapsto E(u + t\varphi)$$

has a minimum in $t = 0$. Furthermore by linearity of the integral, it is smooth in t. Hence

$$\begin{aligned}
0 =& \frac{d}{dt}E(u+t\varphi)|_{t=0} \\
=& \frac{d}{dt}\left(\frac{1}{2}\int_\Omega |\nabla u|^2 + 2t\langle \nabla u, \nabla\varphi\rangle + t^2|\nabla\varphi|^2\,dx - \int_\Omega f(u+t\varphi)\,dx\right)|_{t=0} \\
=& \int \langle \nabla u, \nabla\varphi\rangle - f\varphi\,dx.
\end{aligned}$$

Since $\varphi \in C_0^\infty(\Omega)$ is arbitrary, u is then a weak solution to (3.1) in the sense of Definition 3.15.
Now we need to find a minimiser: Let $u_k \in W_0^{1,2}(\Omega)$ be a minimising sequence, i.e.

$$\inf_{v \in W_0^{1,2}(\Omega)} E(v) = \lim_{k\to\infty} E(u_k).$$

First we show a coercivity estimate for E:

$$\begin{aligned}
E(u_k) =& \frac{1}{2}\int_\Omega |\nabla u_k|^2\,dx - \int_\Omega f u_k\,dx \\
\overset{\text{C.S.}}{\geq}& \frac{1}{2}\int_\Omega |\nabla u_k|^2\,dx - \sqrt{\int_\Omega f^2\,dx}\sqrt{\int_\Omega u_k^2\,dx} \\
\overset{\text{Poincaré}}{\geq}& \frac{1}{2}\int_\Omega |\nabla u_k|^2\,dx - C\sqrt{\int_\Omega f^2\,dx}\sqrt{\int_\Omega |\nabla u_k|^2\,dx} \\
\overset{\text{Young}}{\geq}& \frac{1}{2}\int_\Omega |\nabla u_k|^2\,dx - C(\varepsilon)\int_\Omega f^2\,dx - C\varepsilon\int_\Omega |\nabla u_k|^2\,dx.
\end{aligned}$$

Choosing $\varepsilon < \frac{1}{4C}$ yields, that there exists another constant $C > 0$ (independent of k), such that (please note, that the energy has to be bounded above, since u_k is a minimising sequence)

$$C \geq \int_\Omega |\nabla u_k|^2\,dx.$$

By Remark 3.11 we have that $\|\cdot\|_{W_0^{1,2}(\Omega)}$ is a norm on $W_0^{1,2}(\Omega)$, such that that space is a Hilbert space. Hence by Theorem 2.14 we find $u \in W_0^{1,2}(\Omega)$ and a weakly converging subsequence such that

$$u_k \to u \text{ weakly in } W_0^{1,2}(\Omega).$$

The Sobolev embedding 3.14 yields another subsequence (please note that the limit is the same although we work in different spaces. This can bee seen in more general terms and is an exercise. This is also the only time the connectedness has to be employed. By seeing, that $L : W_0^{1,2}(\Omega) \to \mathbb{R}$, $L(v) = \int fv\,dx$ is linear and continuous, we would get the same final result without connectedness.), such that

$$u_k \to u \text{ in } L^2(\Omega).$$

Since norms are weakly lower semicontinuous (cf. Theorem 2.16), we then have

$$\begin{aligned}
&\inf_{v\in W_0^{1,2}(\Omega)} E(v) \leq E(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,dx - \int fu\,dx \\
\leq &\liminf_{k\to\infty} \int_\Omega |\nabla u_k|^2\,dx - \lim_{k\to\infty}\int_\Omega fu_k\,dx \\
= &\liminf_{k\to\infty} \frac{1}{2}\int_\Omega |\nabla u_k|^2\,dx - \int_\Omega fu_k\,dx = \inf_{v\in W_0^{1,2}(\Omega)} E(v)
\end{aligned}$$

and u is the desired minimiser. □

Remark 3.18.

1. The Riesz representation theorem for Hilbert spaces would also yield a weak solution, but that method is not as suited for nonlinear problems.

2. One can do regularity theory and show that u is as smooth as the data, i.e. f and $\partial\Omega$, allows, see e.g. [4, Ch. 8.3/8.4].

3. By testing Definition 3.15 with the solution itself, we see, that it is actually unique.

Part II: A MinMax Method

In the second chapter we like to find sufficient conditions for an energy to just have a critical point instead of a minimum. A critical point will be defined below, but is essentially a point in which the first derivative vanishes. This critical point, which we will construct, will be a so called saddle Point.
Moreover we will only work in Hilbert spaces. The whole theory can be expanded into Banach spaces but dealing with the Dualspace is by the Rieszs representation theorem (see Remark 2.3) more easily. The general theory can be found in e.g. [11, Ch. 2.2-2.3]. Another source which is more appropiate for beginners may be [1, §8.5].
Along the way we will also develop an example on how to apply this theory (see [5, §9-11] and [11, Thm 6.2]).

4 Palais-Smale Condition

Constructing these saddle points will require good compactness properties. The space itself will not be able to do that, hence the functional (or energy) has to shoulder some of that burden. In essence this is what the Palais-Smale condition embodies. We will make it precise in this section.
Throughout this exposition H will always be a real Hilbert space with scalarproduct $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. First we define what it means for an energy (or functional) $E : H \to \mathbb{R}$ to be differentiable:

Definition 4.1. Let H be a Hilbert space, $E : H \to \mathbb{R}$ is called (Fréchet-) differentiable in $u \in H$, iff there exists a bounded linear functional $DE(u) \in H^*$, $\delta > 0$ and a function $\phi : B_\delta(u) \to \mathbb{R}$, such that for all $v \in B_\delta(u)$ we have

$$E(v) = E(u) + DE(u)(v-u) + \phi(v) \text{ and } \lim_{v \to u} \frac{\phi(v)}{\|v-u\|} = 0.$$

Furthermore E is called continuously differentiable on H (written as $E \in C^1(H) = C^1(H, \mathbb{R})$), iff $DE(u) \in H^*$ exists for all $u \in H$ and the map

$$u \mapsto DE(u)$$

is continuous w.r.t. the topology of H and H^*.

Remark 4.2.

1. Rieszs representation theorem yields an $\nabla E(u) \in H$, such that for all $\varphi \in H$ we have
$$DE(u)(\varphi) = \langle \nabla E(u), \varphi \rangle.$$
Be careful, in $H = \mathbb{R}^n$ this identification is canonical, but in infinite dimensions, this can be quite tricky to do explicitly. For example $H = W_0^{1,2}(\Omega)$ with $\Omega \subseteq \mathbb{R}^n$ open and bounded is a Hilbert space. Given $f := DE(u)$ and $v = \nabla E(u)$ we have the following correlation for all $\varphi \in W_0^{1,2}(\Omega)$:
$$f(\varphi) = \langle v, \varphi \rangle_{W_0^{1,2}(\Omega)} = \int_\Omega \langle \nabla v, \nabla \varphi \rangle \, dx,$$
i.e. we would have to solve in a weak sense
$$-\Delta v = f \text{ in } \Omega \text{ and } v|_{\partial\Omega} = 0.$$

2. With $\langle \cdot, \cdot \rangle$ we will denote the H-scalarproduct as well as the dual $H - H^*$ pairing, i.e.

$$\langle DE(u), \varphi \rangle := \langle DE(u), \varphi \rangle_{H-H^*} := DE(u)(\varphi) = \langle \nabla E(u), \varphi \rangle_H =: \langle \nabla E(u), \varphi \rangle.$$

3. The condition $\lim_{v \to u} \frac{\phi(v)}{\|v-u\|} = 0$ is often also written as $o(\|v - u\|)$, i.e.

$$\phi(v) = o(\|v - u\|),$$

if the above limit holds.

4. If $E : H \to \mathbb{R}$ is Fréchet-differentiable, then one calculates $DE(u)$ with the help of the so called directional or Gâteaux -derivative: Let $\varphi \in H$ be arbitrary and $|t|$ close to zero. Then by the definition we would have in small o-notation

$$E(u + t\varphi) = E(u) + t\langle DE(u), \varphi \rangle + o(|t|).$$

Hence for $t \to 0$

$$\langle DE(u), \varphi \rangle = \lim_{t \to 0} \left(\frac{1}{t}(E(u + t\varphi) - E(u)) + \frac{1}{t} o(|t|) \right)$$
$$= \lim_{t \to 0} \frac{E(u + t\varphi) - E(u)}{t} = \frac{d}{dt} E(u + t\varphi)|_{t=0}.$$

This opens a way to show differentiability: First one calculates $\frac{d}{dt}E(u + t\varphi)|_{t=0}$ uses the result as a candidate for $DE(u)$ and verifies Definition 4.1

5. This last calculation shows, that satisfying an Euler-Lagrange equation as in section 1 means, that the Gâteaux derivative vanishes. Hence we say that we have found a critical point.

Now we introduce our running example, which will accompany us in the development of this theory:

Example 4.3. We consider the following Dirichlet boundary value problem

$$-\Delta u = g(\cdot, u) \text{ in } \Omega, \; u|_{\partial\Omega} = 0. \tag{4.1}$$

Here $g : \Omega \times \mathbb{R} \to \mathbb{R}$ is an x-dependent nonlinearity.
We assume the following:
$\Omega \subseteq \mathbb{R}^n$ $(n \geq 2)$ is open, bounded and connected, i.e. a bounded domain. The boundary $\partial\Omega$ is supposed to be sufficiently smooth. Furthermore

$$g : \Omega \times \mathbb{R} \to \mathbb{R}, \; (x, t) \mapsto g(x, t) \text{ is continuous.} \tag{4.2}$$

We introduce an integrability exponent $p > 0$, which satisfies

$$1 < p < \begin{cases} \infty, & n = 2 \\ \frac{n+2}{n-2}, & n > 2 \end{cases} \tag{4.3}$$

There exists also a constant $C > 0$, such that we have the following growths condition

$$|g(x, t)| \leq C(1 + |t|^p) \text{ for all } (x, t) \in \Omega \times \mathbb{R}. \tag{4.4}$$

We denote with G the primitive of g, i.e. we set

$$G(x,t) := \int_0^t g(x,\tau)\,d\tau, \tag{4.5}$$

which exists, since g is continuous. Finally for $u \in W_0^{1,2}(\Omega)$ we define the energy

$$E(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,dx - \int_\Omega G(x,u(x))\,dx. \tag{4.6}$$

Our goal is to show, that there exists a nontrivial weak solution $u \neq 0$ to (4.1).

Lemma 4.4. *The energy $E : W_0^{1,2}(\Omega) \to \mathbb{R}$ defined in Example 4.3 is well defined and satisfies $E \in C^1(W_0^{1,2}(\Omega))$.*

For the proof of this Lemma and other related theorems, we need the following facts, which are usually done in Analysis 1-4:

1. Youngs inequality:
   Let $x, y \geq 0$, $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $\varepsilon > 0$. Then

   $$xy \leq \varepsilon x^p + \frac{(p\varepsilon)^{1-p}}{q} y^q.$$

2. Addendum to Riesz-Fischer theorem:
   Let $f_k, f \in L^1(\mu)$ with μ being a measure, Furthermore let $f_k \to f$ in $L^1(\mu)$. Then there exists a subsequence $f_{k_j} \in L^1(\mu)$ such that

   $$f_{k_j}(x) \to f(x) \text{ for } \mu - \text{almost every } x.$$

3. Vitalis convergence theorem:
   Let $\Omega \subseteq \mathbb{R}^n$ be bounded, $f_k, f \in L^1(\Omega)$ with $f_k \to f$ pointwise $\mathcal{L}^n$-a.e.. Further let the sequence f_k have uniformely absolutely continuous integrals, i.e. for all $\varepsilon > 0$ exists a $\delta > 0$, such that for all measurable sets $A \subseteq \Omega$ with $\mathcal{L}^n(A) < \delta$ we have for all $k \in \mathbb{N}$

   $$\int_A |f_k|\,dx < \varepsilon.$$

   Then$f_k \to f$ in $L^1(\Omega)$, especially integral and pointwise limit interchanges, i.e.

   $$\int_\Omega f(x)\,dx = \lim_{k\to\infty} \int_\Omega f_k(x)\,dx.$$

   As a remark: If $f \in L^1(\Omega)$, then f has an absolutely continuous integral, i.e. for all $\varepsilon > 0$ there exists a $\delta > 0$, such that for all measurable $A \subseteq \Omega$ with $\mathcal{L}^n(A) < \delta$, we have

   $$\int_A |f|\,dx < \varepsilon.$$

Let us turn to the Proof of Lemma 4.4:

Proof. 1. We start with E being well defined:
By the Sobolev embedding theorem 3.14 there exists a constant $C = C(n,p,\Omega) > 0$, such that for all $v \in W_0^{1,2}(\Omega)$ we have

$$\|v\|_{L^{p+1}(\Omega)} \leq C\|v\|_{W_0^{1,2}(\Omega)}. \tag{4.7}$$

This yields for all $x \in \Omega$ and $u \in W_0^{1,2}(\Omega)$

$$\begin{aligned}|G(x,u(x))| \leq \max_{|\tau|\leq|u(x)|} |g(x,\tau)||u(x)| &\overset{(4.4)}{\leq} C(1+|u(x)|^p)|u(x)| \\ &\overset{\text{Young}}{\leq} C(1+|u(x)|^{p+1}).\end{aligned}$$

Since $\Omega \subseteq \mathbb{R}^n$ is bounded, (4.7) yields

$$E(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 - G(x,u(x))\,dx$$

to be well defined.

2. Finding a candidate for the derivative:
Let $u, \varphi \in W_0^{1,2}(\Omega)$. Then we calculate the Gâteaux derivative (cf. Remark 4.2 4)). For this we need to calculate the derivative of the integrand and afterwards verify, that we can interchange integral and derivative:

$$\frac{d}{dt}G(x,u(x)+t\varphi(x)) = \varphi(x)g(x,u(x)+t\varphi(x)).$$

For $|t| \leq 1$ we need to find a dominating integrable function independent of t:

$$\begin{aligned}&|\varphi(x)g(x,u(x)+t\varphi(x))| \leq C|\varphi|(1+|u+t\varphi|^p) \\ \leq& C|\varphi(x)|(1+|u(x)|^p+|\varphi(x)|^p) \overset{\text{Young}}{\leq} C(1+|u|^{p+1}+|\varphi|^{p+1}).\end{aligned}$$

In the second inequality we used (which is a result of Hölders inequality)

$$|x|+|y| \leq 2^{\frac{p-1}{p}}(|x|^p+|y|^p)^{\frac{1}{p}}. \tag{4.8}$$

By (4.7) this is an integrable function. Hence by the dominated convergence theorem

$$\begin{aligned}&\frac{d}{dt}E(u+t\varphi)|_{t=0} \\ =&\frac{d}{dt}\frac{1}{2}\int_\Omega |\nabla(u+t\varphi)|^2\,dx|_{t=0} - \frac{d}{dt}\int_\Omega G(x,u(x)+t\varphi(x))\,dx|_{t=0} \\ =&\int_\Omega \langle\nabla u,\nabla\varphi\rangle\,dx - \int_\Omega \frac{d}{dt}G(x,u(x)+t\varphi(x))|_{t=0}\,dx \\ =&\int_\Omega \langle\nabla u,\nabla\varphi\rangle\,dx - \int_\Omega \varphi(x)g(x,u(x))\,dx.\end{aligned}$$

Hence if u is a critical point of E, it satisfies (4.1) weakly.
Now let us define $DE(u)$ and check aftwerwards, that it is indeed the (Fréchet-) derivative of E:

$$\langle DE(u), \varphi\rangle = DE(u)(\varphi) := \int_\Omega \langle \nabla u, \nabla\varphi\rangle \, dx - \int_\Omega \varphi(x) g(x, u(x)) \, dx. \quad (4.9)$$

Trivially $DE(u) : W_0^{1,2}(\Omega) \to \mathbb{R}$ is linear. Let us also show it is bounded, i.e. continuous: By (4.4) and Hölders inequality we obtain

$$\begin{aligned}
&|\langle DE(u), \varphi\rangle| \leq \|u\|_{W_0^{1,2}(\Omega)} \|\varphi\|_{W_0^{1,2}(\Omega)} + C \int_\Omega (1 + |u|^p)|\varphi| \, dx \\
\leq &\|u\|_{W_0^{1,2}(\Omega)} \|\varphi\|_{W_0^{1,2}(\Omega)} + C(\|\varphi\|_{L^1(\Omega)} + \|u\|^p_{L^{p+1}(\Omega)} \|\varphi\|_{L^{p+1}(\Omega)}) \\
\leq &\|u\|_{W_0^{1,2}(\Omega)} \|\varphi\|_{W_0^{1,2}(\Omega)} + C(C\|\varphi\|_{L^{p+1}(\Omega)} + \|u\|^p_{L^{p+1}(\Omega)} \|\varphi\|_{L^{p+1}(\Omega)}) \\
\overset{(4.7)}{\leq} &\|\varphi\|_{W_0^{1,2}(\Omega)} C \left(\|u\|_{W_0^{1,2}(\Omega)} + 1 + \|u\|^p_{W_0^{1,2}(\Omega)} \right)
\end{aligned}$$

Hence by Youngs inequality

$$\|DE(u)\|_{(W_0^{1,2}(\Omega))^*} \leq C(1 + \|u\|^p_{W_0^{1,2}(\Omega)})$$

with $C = C(n, p, \Omega) > 0$. Therefore $DE(u) \in (W_0^{1,2}(\Omega))^*$.

3. Check that $DE(u)$ is indeed the Fréchet-derivative of E:
For $\varphi \in W_0^{1,2}(\Omega) \setminus \{0\}$ we define

$$\Psi(\varphi) := \frac{1}{\|\varphi\|_{W_0^{1,2}(\Omega)}} \left(E(u + \varphi) - E(u) - \langle DE(u), \varphi\rangle \right).$$

We need to show

$$\lim_{\|\varphi\|_{W_0^{1,2}(\Omega)} \to 0} \Psi(\varphi) = 0.$$

Since the limit is supposed to be unique, it is enough to show the following: For all sequences $\varphi_k \in W_0^{1,2}(\Omega) \setminus \{0\}$ with $\varphi_k \to 0$ w.r.t. $\|\cdot\|_{W_0^{1,2}(\Omega)}$ exists a subsequence k_ℓ, such that $\Psi(\varphi_{k_\ell}) \to 0$.
Let φ_k be such a sequence. By (4.7) and the addendum of the Riesz-Fischer theorem we have after extracting a subsequence

$$\varphi_k \to 0 \text{ in } L^{p+1}(\Omega), \ \varphi_k \to 0 \text{ a.e.}$$

Furthermore by Definition and Hölders inequality we have

$$\begin{aligned}
&|\Psi(\varphi_k)| \\
= &\frac{1}{\|\varphi_k\|_{W_0^{1,2}(\Omega)}} \Bigg| \frac{1}{2} \int_\Omega |\nabla u|^2 + 2\langle \nabla u, \nabla\varphi_k\rangle + |\nabla\varphi_k|^2 - 2G(x, u(x) + \varphi_k(x)) \, dx \\
&- \frac{1}{2} \int_\Omega |\nabla u|^2 - 2G(x, u(x)) \, dx - \int_\Omega \langle \nabla u, \nabla\varphi_k\rangle + \varphi_k g(x, u(x)) \, dx \Bigg| \\
\leq &\frac{1}{2} \|\varphi_k\|_{W_0^{1,2}(\Omega)} + \frac{1}{\|\varphi_k\|_{W_0^{1,2}(\Omega)}} \left| \int_\Omega \int_0^1 \frac{d}{dt} G(x, u(x) + t\varphi_k(x)) \, dt \, dx \right.
\end{aligned}$$

$$
\begin{aligned}
&- \int_\Omega \varphi_k g(x,u(x)))\,dx\Big| \\
\le& \frac{1}{2}\|\varphi_k\|_{W_0^{1,2}(\Omega)} \\
&+ \frac{1}{\|\varphi\|_{W_0^{1,2}(\Omega)}} \int_\Omega \int_0^1 |\varphi_k||g(x,u(x)+t\varphi_k(x)) - g(x,u(x))|\,dt\,dx\Big| \\
\le& \frac{1}{2}\|\varphi_k\|_{W_0^{1,2}(\Omega)} \\
&+ \frac{\|\varphi_k\|_{L^{p+1}(\Omega)}}{\|\varphi_k\|_{W_0^{1,2}(\Omega)}} \left(\int_\Omega \int_0^1 |g(x,u(x)+t\varphi_k(x)) - g(x,u(x))|^{\frac{p+1}{p}}\,dt\,dx \right)^{\frac{p}{p+1}} \\
\overset{(4.7)}{\le}& \frac{1}{2}\|\varphi_k\|_{W_0^{1,2}(\Omega)} + C\left(\int_\Omega \int_0^1 |g(x,u(x)+t\varphi_k(x)) - g(x,u(x))|^{\frac{p+1}{p}}\,dt\,dx \right)^{\frac{p}{p+1}}.
\end{aligned}
$$

Now we only need to show

$$
\int_\Omega \int_0^1 |g(x,u(x)+t\varphi_k(x)) - g(x,u(x))|^{\frac{p+1}{p}}\,dt\,dx \to 0 \text{ for } k \to \infty.
$$

We apply Vitalis convergence theorem:
Since $\varphi_k \to 0$ a.e. and since g is continuous, we have

$$
|g(x,u(x)+t\varphi_k(x)) - g(x,u(x))|^{\frac{p+1}{p}} \to 0 \text{ pointwise a.e.}
$$

Now we need to check that the integral is uniformely absolutely continuous. For that let $\varepsilon > 0$ $A \subseteq \Omega \times (0,1)$ be measurable with $\mathcal{L}^{n+1}(A) < \delta$ and $\delta > 0$ to be found. The growth condition (4.4), norms in $\mathbb{R}^n$ are all equivalent and (4.8) yield

$$
\begin{aligned}
&\int_A |g(x,u(x)+t\varphi_k(x)) - g(x,u(x))|^{\frac{p+1}{p}}\,d\mathcal{L}^{n+1}(x,t) \\
\le& C\int_A (1 + |u+t\varphi_k|^p + |u|^p)^{\frac{p+1}{p}}\,d\mathcal{L}^{n+1} \\
\le& C\int_A (1 + |u+t\varphi_k| + |u|)^{p+1}\,d\mathcal{L}^{n+1} \le C\int_A (1+|u|+|\varphi_k|)^{p+1}\,d\mathcal{L}^{n+1} \\
\le& C\int_A 1 + |u|^{p+1} + |\varphi_k|^{p+1}\,d\mathcal{L}^{n+1}
\end{aligned}
$$

Since $u \in L^{p+1}(\Omega)$ we can employ, that it is absolutely continuous w.r.t. to the integral, i.e. we find a $\delta_u > 0$, such that

$$
C\int_A |u|^{p+1}\,dx \le \frac{\varepsilon}{2}
$$

if $\mathcal{L}^{n+1}(A) < \delta_u$. By (4.7) we have

$$
\int_A |\varphi_k|^{p+1}\,d\mathcal{L}^{n+1} \le \int_\Omega |\varphi_k|^{p+1}\,d\mathcal{L}^n \le C\|\varphi_k\|_{W_0^{1,2}(\Omega)}^{p+1} < \frac{\varepsilon}{2},
$$

if $k \ge k_0(\varepsilon) \in \mathbb{N}$ by the convergence to zero in $W_0^{1,2}(\Omega)$. For $k = 1, \ldots, k_0 - 1$ we can employ the same argument as for u and obtain $\delta_1, \ldots, \delta_{k_0-1}$. Choosing

$$
\delta := \min\{\delta_u, \delta_1, \ldots, \delta_{k_0-1}\}
$$

yields

$$C \int_A 1 + |u|^{p+1} + |\varphi_k|^{p+1} \, d\mathcal{L}^{n+1} < \varepsilon.$$

Hence by Vitalis convergence theorem we have shown, that E is Fréchet-differentiable in $u \in W_0^{1,2}(\Omega)$ arbitrary with

$$DE(u)(\varphi) = \int_\Omega \langle \nabla u, \nabla \varphi \rangle - \varphi g(x, u(x)) \, dx.$$

4. To finish the proof we have to show, that the map

$$W_0^{1,2}(\Omega) \ni u \mapsto DE(u) \in W_0^{1,2}(\Omega)^*$$

is continuous. This then finally yields $E \in C^1(W_0^{1,2}(\Omega))$.
Again it is enough to show that for $u_k \to u$ in $W_0^{1,2}(\Omega)$ we find a subsequence k_ℓ, such that

$$\sup_{\varphi \in W_0^{1,2}(\Omega),\ \|\varphi\|_{W_0^{1,2}(\Omega)} \leq 1} |\langle DE(u_{k_\ell}) - DE(u), \varphi \rangle| \to 0 \text{ for } \ell \to \infty.$$

Let u_k, u be as above. Then by (4.7) we find a subsequence such that

$$u_k \to u \text{ in } L^{p+1}(\Omega) \text{ and } u_k \to u \text{ pointwise a.e..}$$

Cauchy-Schwartzes and Hölders inequality yield

$$\begin{aligned}
&\sup_{\|\varphi\|_{W_0^{1,2}(\Omega)} \leq 1} |\langle DE(u_{k_\ell}) - DE(u), \varphi \rangle| \\
= &\sup_{\|\varphi\|_{W_0^{1,2}(\Omega)} \leq 1} \left| \int_\Omega \langle \nabla u_k - \nabla u, \nabla \varphi \rangle \, dx - \int_\Omega (g(x, u_k(x)) - g(x, u(x)) \varphi(x) \, dx \right| \\
\leq &\|u_k - u\|_{W_0^{1,2}(\Omega)} + \left(\int_\Omega |g(x, u_k(x)) - g(x, u(x))|^{\frac{p+1}{p}} \, dx \right)^{\frac{p}{p+1}}.
\end{aligned}$$

The right term can now be handled with Vitalis theorem as above. Working out the details is an exercise. All in all this yields, that it all converges to 0. Therefore we finally have $E \in C^1(W_0^{1,2}(\Omega))$. □

Remark 4.5. The integrability p is choosen (see (4.3)), such that the Sobolev embedding Theorem 3.14 yields a compact embedding (see the proof of Lemma 4.5). This is not necessary for the differentiability of E, but will prove central for the later introduced Palais-Smale condition. The condition on p is also called subcritical.
If $p = \frac{n+2}{n-2}$, the case is called critical . Then the variational formulation (i.e. weak formulation) still makes sense, but the behaviour of the existence of solutions completely changes, see e.g. [3, Thm 7.31].
If $p > \frac{n+2}{n-2}$, the variational formulation looses its value completely (this case is called supercritical). Nevertheless other methods can be employed to examine such problems and again yield a different behaviour, see e.g. [3, §7.11].

Next we introduce the concept of a Palais-Smale condition, which let the energy shoulder the burden of sequential compactness:

Definition 4.6. Let H be a Hilbert space and $E \in C^1(H)$.

1. A sequence $(u_k)_{k\in\mathbb{N}} \subseteq H$ is called Palais-Smale sequence for E, iff
   (a) $\lim_{k\to\infty} E(u_k)$ exists in $\mathbb{R}$.
   (b) $\|DE(u_k)\|_{H^*} \to 0$ for $k \to \infty$.
2. We say E satisfies a Palais-Smale condition, iff every Palais-Smale sequence for E admits a (strongly, i.e. in norm) convergent subsequence.

Example 4.7. Every coercive and continuously differentiable $E : \mathbb{R}^n \to \mathbb{R}$ satisfies a Palais-Smale condition. This is due to the fact, that every Palais-Smale sequence of E is bounded by the coercivity and since $\mathbb{R}^n$ is finite dimensional, it admits a converging subsequence.

Now we examine our running example 4.3 in view of a Palais-Smale condition. Please note, that if the growth would be critical (see Remark 4.5), then the following Lemma 4.8 is not correct.
Furthermore in the Lemma we will have to add a 'superlinearity' condition at infinity, see (4.10).

Lemma 4.8. *E, g and G are the same as in example 4.3 and satisfy the same assumptions, i.e.* (4.2)-(4.6). *Furthermore we require that there exists an $R_0 > 0$ and $q > 2$, such that for all $x \in \Omega$ and all $|t| \geq R_0$ we have*

$$qG(x,t) \leq g(x,t)t. \tag{4.10}$$

Then E satisfies a Palais-Smale condition.

Proof. Let $u_k \in W_0^{1,2}(\Omega)$ be a Palais-Smale sequence for

$$E(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 - G(x,u(x))\,dx.$$

First we show this that this sequence is bounded:
The Palais-Smale properties of the sequence yield

$$\begin{aligned}&qE(u_k) - \langle DE(u_k)\, u_k\rangle \leq C + \|DE(u_k)\|_{W_0^{1,2}(\Omega)^*}\|u_k\|_{W_0^{1,2}(\Omega)}\\ \leq &C + o(1)\|u_k\|_{W_0^{1,2}(\Omega)}.\end{aligned}$$

On the other hand the definition of E and (4.10) yield

$$\begin{aligned}&qE(u_k) - \langle DE(u_k), u_k\rangle = \frac{q}{2}\int_\Omega |\nabla u_k|^2\,dx - \int_\Omega qG(x,u_k(x))\,dx\\ &-\int_\Omega |\nabla u_k|^2\,dx + \int_\Omega g(x,u_k(x))u_k(x)\,dx\\ \geq&\frac{q-2}{2}\int_\Omega |\nabla u_k|^2\,dx - \mathcal{L}^n(\Omega)\sup_{x\in\Omega,t\in\mathbb{R}}(qG(x,t)-g(x,t)t)\\ \overset{(4.10)}{\geq}&\frac{q-2}{2}\|u_k\|^2_{W_0^{1,2}(\Omega)} - \max\left\{0, \sup_{x\in\Omega,|t|\leq R_0}(qG(x,t)-g(x,t)t)\right\}\end{aligned}$$

By the growth condition G and g are bounded, if $|t| \leq R_0$, which yields that the last term is a finite constant. By combining both inequality we obtain

$$O(1) + o(1)\|u_k\|_{W_0^{1,2}(\Omega)} \geq \frac{q-2}{2}\|u_k\|^2_{W_0^{1,2}(\Omega)} - O(1).$$

By rearranging and using $q > 2$ we get

$$\|u_k\|_{W_0^{1,2}(\Omega)} \leq O(1).$$

Here we used $O(1)$ to denote some constant.
Theorem 2.14, the Sobolev embedding Theorem 3.14 and the addendum to the Riesz-Fischer Theorem yield a subsequence and a $u \in W_0^{1,2}(\Omega)$ such that

$$u_k \to u \text{ weakly in } W_0^{1,2}(\Omega), \text{ strongly in } L^{p+1}(\Omega) \text{ and pointwise a.e.}$$

We claim, that u satisfies the following boundary problem weakly:

$$\text{for all } \varphi \in W_0^{1,2}(\Omega) \text{ we have } \int_\Omega \langle \nabla u, \nabla \varphi \rangle \, dx = \int_\Omega g(x, u(x))\varphi(x) \, dx. \tag{4.11}$$

We show this claim now: Let $\varphi \in W_0^{1,2}(\Omega)$ be arbitrary but fixated. The idea of the proof is contained in the next calculation: Since u_k is a Palais-Smale sequence for E, we have by (4.9)

$$\begin{aligned} o(1) =& \langle DE(u_k), \varphi \rangle = \int_\Omega \langle \nabla u_k, \nabla \varphi \rangle \, dx - \int_\Omega g(x, u_k(x))\varphi(x) \, dx \\ =& \int_\Omega \langle \nabla u, \nabla \varphi \rangle - g(x, u(x))\varphi \, dx + o(1). \end{aligned}$$

The last equality still has to be shown. For this we prove

$$g(\cdot, u_k(\cdot)) \to g(\cdot, u(\cdot)) \text{ in } L^{\frac{p+1}{p}}(\Omega), \tag{4.12}$$

which will be a consequence of Vitalis Theorem. The Rest of the proof of (4.11) is then Hölders inequality. The argument will be similar to the one given in Lemma 4.5.
First we note, that by the pointwise a.e. convergence and g being continuous, we get

$$|g(x, u_k(x)) - g(x, u_k(x))|^{\frac{p+1}{p}} \to 0 \text{ for almost every } x \in \Omega.$$

Next we show the uniform absolute continuity w.r.t. the integral:
Let $\varepsilon > 0$ be arbitrary. Let $\delta > 0$ to be calulated later and let $A \subseteq \Omega$ be measurable with $\mathcal{L}^n(A) < \delta$. Then the growth condition (4.4) yields

$$\begin{aligned} &|g(x, u_k(x)) - g(x, u(x))|^{\frac{p+1}{p}} \leq C(1 + |u|^p + |u_k|^p)^{\frac{p+1}{p}} \\ \leq& C(1 + |u|^p + |(u_k - u) + u|^p)^{\frac{p+1}{p}} \overset{(4.8)}{\leq} C(1 + |u|^p + |u_k - u|^p)^{\frac{p+1}{p}} \\ \leq& C(1 + |u| + |u_k - u|)^{p+1} \overset{(4.8)}{\leq} C(1 + |u|^{p+1} + |u_k - u|^{p+1}) \end{aligned}$$

The second to last inequality comes from the fact, that $\|\cdot\|_1$ and $\|\cdot\|_p$ are equivalent norms on $\mathbb{R}^n$. Hence

$$\int_A |g(x,u_k(x)) - g(x,u(x))|^{\frac{p+1}{p}}\,dx \le C\int_A 1+|u|^{p+1}\,dx + C\int_A |u_k-u|^{p+1}\,dx$$

Since $1+|u|^{p+1} \in L^1(\Omega)$ we can find a $\delta_u > 0$ such that

$$C\int_A 1+|u|^{p+1} < \frac{\varepsilon}{2}.$$

Furthermore the other term converges and therefore there is a $k_0 \in \mathbb{N}$, such that for all $k > k_0$ we have

$$C\int_A |u_k-u|^{p+1}\,dx \le C\int_\Omega |u_k-u|^{p+1}\,dx < \frac{\varepsilon}{2}.$$

For $k = 1,\dots,k_0$ we argue in an analogue way to the u-term. Since these are only finitely many, we find an appropriate $\delta > 0$. This shows (4.12).
Hölders inequality now yields

$$\int_\Omega |g(x,u_k(x)) - g(x,u(x))||\varphi(x)|\,dx \le \|g(\cdot,u_k(\cdot)) - g(\cdot,u(\cdot))\|_{L^{\frac{p+1}{p}}(\Omega)}\|\varphi\|_{L^{p+1}(\Omega)} \to 0$$

which now shows (4.11). Next by Hölders inequality we have

$$\begin{aligned}
&\left|\int_\Omega g(x,u_k(x))u_k(x)\,dx - \int_\Omega g(x,u(x))u(x)\,dx\right| \\
&\le \int_\Omega |g(x,u_k(x)) - g(x,u(x))||u_k(x)|\,dx + \int_\Omega |g(x,u(x))||u_k(x)-u(x)|\,dx \\
&\le \|g(\cdot,u_k(\cdot)) - g(\cdot,u(\cdot))\|_{L^{\frac{p+1}{p}}(\Omega)}\|u_k\|_{L^{p+1}(\Omega)} + \|g(\cdot,u(\cdot))\|_{L^{\frac{p+1}{p}}(\Omega)}\|u_k-u\|_{L^{p+1}(\Omega)} \\
&\overset{(4.12)}{=} o(1)\cdot O(1) + O(1)\cdot o(1) \to 0 \text{ for } k\to\infty.
\end{aligned}$$

This in turn together with the Palais-Smale condition $\|DE(u_k)\|_{W_0^{1,2}(\Omega)^*} \to 0$ yields

$$\begin{aligned}
&o(1) = \|DE(u_k)\|_{W_0^{1,2}(\Omega)^*}\|u_k\|_{W_0^{1,2}(\Omega)} \ge \langle DE(u_k), u_k\rangle \\
&= \int_\Omega |\nabla u_k|^2\,dx - \int_\Omega g(x,u_k(x))u_k(x)\,dx = \int_\Omega |\nabla u_k|^2\,dx - \int_\Omega g(x,u(x))\,dx + o(1) \\
&\overset{(4.11)}{=} \int_\Omega |\nabla u_k|^2\,dx - \int_\Omega |\nabla u|^2\,dx + o(1).
\end{aligned}$$

Together with the weak lower semicontinuity of a norm (see Lemma 2.16)

$$\int_\Omega |\nabla u|^2\,dx \ge \limsup_{k\to\infty}\int_\Omega |\nabla u_k|^2\,dx \ge \liminf_{k\to\infty}\int_\Omega |\nabla u_k|^2\,dx \ge \int_\Omega |\nabla u|^2\,dx,$$

hence

$$\lim_{k\to\infty}\int_\Omega |\nabla u_k|^2\,dx = \int_\Omega |\nabla u|^2\,dx.$$

The weak convergence in conjunction with the norms converging to each other yield convergence w.r.t. the norm in a Hilbert space:

$$
\begin{aligned}
&\|u_k - u\|^2_{W_0^{1,2}(\Omega)} = \|u_k\|^2_{W_0^{1,2}(\Omega)} - 2\langle u_k, u\rangle_{W_0^{1,2}(\Omega)} + \|u\|^2_{W_0^{1,2}(\Omega)} \\
=&\|u_k\|^2_{W_0^{1,2}(\Omega)} - 2\|u\|^2_{W_0^{1,2}(\Omega)} + \|u\|^2_{W_0^{1,2}(\Omega)} + o(1) \\
=&\|u_k\|^2_{W_0^{1,2}(\Omega)} - \|u\|^2_{W_0^{1,2}(\Omega)} + o(1) = o(1) \to 0.
\end{aligned}
$$

All in all we have

$$u_k \to u \text{ in } W_0^{1,2}(\Omega),$$

which is the desired result. □

Next we define some important technical sets, in which we will look for our saddle points:

Definition 4.9. Let H be a Hilbert space and $E \in C^1(H)$. Furthermore let $\beta \in \mathbb{R}$ and $\delta, \rho > 0$. We then define

$$
\begin{aligned}
E_\beta :=&\{u \in H |\ E(u) < \beta\} \\
K_\beta :=&\{u \in H |\ E(u) = \beta \text{ and } DE(u) = 0\} \\
N_{\beta,\delta} :=&\{u \in H |\ |E(u) - \beta| < \delta \text{ and } \|DE(u)\|_{H^*} < \delta\} \\
U_{\beta,\rho} :=&\{v \in H |\ \exists u \in K_\beta :\ \|u - v\|_H < \rho\} = \bigcup_{u \in K_\beta} B_\rho(u) =: U_\rho(K_\beta).
\end{aligned}
$$

Lemma 4.10. *Let H be a Hilbert space and $E \in C^1(H)$ satisfy a Palais-Smale condition. Then we have*

1. *K_β is compact.*
2. *The system $(N_{\beta,\delta})_{\delta>0}$ is a neighbourhood basis of K_β, i.e. for all open sets $U \supseteq K_\beta$ exists a $\delta > 0$, such that*

$$N_{\beta,\delta} \subseteq U.$$

3. *The system $(U_{\beta,\rho})_{\rho>0}$ is a neighbourhood basis of K_β.*

Proof. 1. Let $u_k \in K_\beta$ be a sequence. Since $E(u_k) = \beta$ and $DE(u_k) = 0$, the sequence u_k is a Palais-Smale sequence of E. Since E satisfies a Palais-Smale condition, it has a convergent subsequence, i.e. we can assume

$$u_k \to u \in H.$$

Since E is continuous, we have $E(u) = \beta$. Since $E \in C^1(H)$ we furthermore have, that $DE(u) = 0$. Hence $u \in K_\beta$ and therefore the set is sequentially compact.

2. We proceed by contradiction and assume there exists an open set $U \supseteq K_\beta$, such that for all $\delta > 0$ we have $N_{\beta,\delta} \nsubseteq U$. Hence we find a sequence $u_k \in N_{\beta,\frac{1}{k}}$ with $u_k \notin U$. So

$$E(u_k) \to \beta,\ DE(u_k) \to 0 \text{ in } H^* \text{ and } u_k \notin U.$$

This yields u_k to be a Palais-Smale sequence of E. Therefore we find a converging subsequence $u_k \to u \in H$. Since $E \in C^1(H)$ we have

$$E(u) = \beta,\ DE(u) = 0 \ \Rightarrow\ u \in K_\beta.$$

Since $u \in K_\beta \subseteq U$ is open, this is a contradiction.

3. We again proceed by contradiction and assume there exists an open $U \supseteq K_\beta$, such that for all $\rho > 0$ we have $U_{\beta,\rho} \nsubseteq U$. As above we find sequences $v_k \in U_{\beta,\frac{1}{k}}$ with $v_k \notin U$. Hence there are $u_k \in K_\beta$ with $\|u_k - v_k\| < \frac{1}{k}$. Therefore $\|u_k - v_k\| \to 0$. Since K_β is compact, we find a converging subsequence for u_k, i.e we can assume $u_k \to u \in K_\beta$. By triangle inequality we then have

$$\|u - v_k\| \leq \|u - u_k\| + \|u_k - v_k\| \to 0,$$

i.e. $v_k \to u$, which is a contradiction to $u \in K_\beta \subseteq U$ being open.

$\square$

5 A deformation Lemma

This section is essentially taken from [11, § II.3], see also [5, § 10].
Here we develop a useful Lemma, which allows us to deform a Hilbert space H, such that the energy $E : H \to \mathbb{R}$ we are examining is decreased, i.e. we look for a one parameter homeomorphism $\Phi : \mathbb{R} \times H \to H$, such that $t \mapsto E(\Phi(t, x))$ is decreasing for all $x \in H$. This will allow us to approach critical points, if Φ is choosen well. Its construction will be similar to a gradient flow of E. Because of a lack of regularity for E, we will have to construct a pseudo-gradient, which will be Lipschitz. Then we can invoke a suitable version of Picard-Lindelöfs theorem to obtain a flow. This is not possible with just ∇E, because ∇E is just continuous. The Peano existence theorem (which works in finite dimensions) is not applicable here, because it essentially builds upon the Arzelá-Ascoli, wich only works in compact sets in H. Since we will work on open sets, this is not enough.
Before we start though, we have to collect some useful topological results, which we formulate in metric spaces (X, d), i.e. X a set and $d : X \times X \to \mathbb{R}$ a metric on X.

Definition 5.1. Let (X, d) be a metric space and $(U_i)_{i\in I} \subseteq X$ be a cover of X, i.e. $\bigcup_{i\in I} U_i = X$. We call $(V_j)_{j\in J}$ a refinement of the cover $(U_i)_{i\in I}$, if $(V_j)_{j\in J}$ covers X and for all $j \in J$ we find an $i \in I$, such that

$$V_j \subseteq U_i.$$

Remark: A subcover is always a refinement, but the inverse is in general false.

Definition 5.2. A cover $(U_i)_{i\in I}$ of a metric space (X, d) is called locally finite, if for all $x \in X$ exists an open neighbourhood $U \subseteq X$, such that only for finitely many $i \in I$ we have

$$U \cap U_i \neq \emptyset.$$

The following theorem cannot be proven here, due to time constraints, but it is a central tool in developing our pseudo-gradient:

Theorem 5.3 (see e.g. [6], Corollary 5.35, p. 160)**.** *Let (X, d) be a metric space. Then it is paracompact, i.e. it is*

1. *Hausdorff: $\forall x \neq y \in X$ exists $U, V \subseteq X$ open with $x \in U$, $y \in V$ and $U \cap V = \emptyset$.*
2. *For every open cover of X exists an open locally finite refinement.*

Now we define precisely what we mean with pseudo-gradient:

Definition 5.4. Let H be a Hilbert space and $E \in C^1(H)$. We define the set of regular points of E by

$$\tilde{H} := \{u \in H : \ \nabla E(u) \neq 0\}.$$

We call a locally Lipschitz-continuous map $G : \tilde{H} \to H$ a pseudo-gradient of E, if for all $u \in \tilde{H}$ we have

$$\|G(u)\|_H < 2\min\{1, \|\nabla E(u)\|\} \tag{5.1}$$

and
$$\langle \nabla E(u), G(u)\rangle_H > \frac{1}{2}\min\{1, \|\nabla E(u)\|\} \cdot \|\nabla E(u)\|. \tag{5.2}$$

Theorem 5.5. *Let H be a Hilbert space and $E \in C^1(H)$. Then E admits a (locally Lipschitz continuous) pseudo-gradient $G : \tilde{H} \to H$.*

Proof. As a preliminary we define for $u \in \tilde{H} = \{v \in H | \; \nabla E(v) \neq 0\}$

$$\tilde{G}(u) := \begin{cases} \nabla E(u), & \text{if } \|\nabla E(u)\| < 1 \\ \frac{\nabla E(u)}{\|\nabla E(u)\|}, & \text{if } \|\nabla E(u)\| \geq 1. \end{cases} \tag{5.3}$$

Let $u \in \tilde{H}$ be fixated. Since $v \mapsto \nabla E(v)$ is continuous, we find an open neighbourhood $W(u) \subseteq \tilde{H}$ of u, such that for all $v \in W(u)$ we have

1. $\|\tilde{G}(u)\| < 2\min\{\|\nabla E(v)\|, 1\}$
2. $\langle \nabla E(v), \tilde{G}(u)\rangle > \frac{1}{2}\min\{\|\nabla E(v)\|, 1\}\|\nabla E(v)\|$

and $W(u)$ is bounded. Since $\tilde{G}$ is not necessarily locally Lipschitz, we have to invest more work: Since
$$\tilde{H} = \bigcup_{v \in \tilde{H}} W(v)$$
is an open cover, by Thm. 5.3 we find an open locally finite refinement $(W_i)_{i\in I}$, $W_i \subseteq \tilde{H}$ open, hence e.g.

$$\tilde{H} = \bigcup_{i \in I} W_i, \;\forall i \in I \;\exists u_i \in \tilde{H} \text{ such that } W_i \subseteq W(u_i).$$

Now we define a locally Lipschitz continuous seperation of unity w.r.t. $(W_i)_{i\in I}$: We start with
$$\tilde{\varphi}_i(v) := \operatorname{dist}(v, H \setminus W_i).$$
Since W_i is open, we then have

$$\tilde{\varphi}_i(v) \neq 0 \;\Leftrightarrow\; v \in W_i.$$

The triangle inequality yields for $v, v' \in H$

$$\begin{aligned} \operatorname{dist}(v, H \setminus W_i) \leq& \|v - v'\| + \operatorname{dist}(v', H \setminus W_i) \\ \operatorname{dist}(v', H \setminus W_i) \leq& \|v - v'\| + \operatorname{dist}(v, H \setminus W_i), \end{aligned}$$

hence
$$|\tilde{\varphi}_i(v) - \tilde{\varphi}_i(v')| \leq \|v - v'\|$$
and $\tilde{\varphi}_i$ is Lipschitz. Since $W_i \subseteq W(u_i)$ with $W(u_i)$ bounded, $\tilde{\varphi}_i$ is bounded as well. Since $(W_i)_{i\in I}$ is a locally finite cover, the following is well defined, because the denominator only has finitely many nontrivial summands:

$$\varphi_i(v) := \frac{\tilde{\varphi}_i(v)}{\sum_{j \in I} \tilde{\varphi}_j(v)}.$$

This yields
$$0 \leq \varphi_i \leq 1, \; \sum_{i \in I} \varphi_i(v) = 1 \text{ for } v \in \tilde{H}$$

and

$$\varphi_i(v) \neq 0 \;\Leftrightarrow\; v \in W_i.$$

Furthermore φ_i is locally Lipschitz, because at least one $\tilde{\varphi}_j$ is bounded locally below by some constant. Here the definition of $\tilde{\varphi}_j$ is needed.
Finally we define our pseudo-gradient

$$\tilde{H} \ni v \mapsto G(v) := \sum_{i \in I} \varphi_i(v) \tilde{G}(u_i).$$

G is locally Lipschitz, because locally the sum above is finite and φ_i is locally Lipschitz. Let us now check the requirements of Definition 5.4: The properties of $\tilde{G}$ yield for all $v \in \tilde{H}$

$$\begin{aligned} \|G(v)\| \leq \sum_{i \in I} \varphi_i(v) \|\tilde{G}(u_i)\| &= \sum_{i \in I,\ v \in W_i \subseteq W(u_i)} \varphi_i(v) \|\tilde{G}(u_i)\| \\ &< 2 \left(\sum_{i \in I} \varphi_i(v) \right) \min\{\|\nabla E(v)\|, 1\} = 2 \min\{\|\nabla E(v)\|, 1\}, \end{aligned}$$

which is (5.1). Similarly we have

$$\begin{aligned} \langle \nabla E(v), G(v) \rangle &= \sum_{i \in I,\ v \in W(u_i)} \varphi_i(v) \langle \nabla E(v), \tilde{G}(u_i) \rangle \\ &> \frac{1}{2} \left(\sum_{i \in I} \varphi_i(v) \right) \min\{1, \|\nabla E(v)\|\} \|\nabla E(v)\| = \frac{1}{2} \min\{1, \|\nabla E(v)\|\} \|\nabla E(v)\|, \end{aligned}$$

which yields (5.2) and therefore finishes the proof.

□

Theorem 5.6 (Deformation lemma)**.** *Let H be a Hilbert space and $E \in C^1(H)$ satisfies a Palais-Smale condition. Let $\beta \in \mathbb{R}$, $\varepsilon_0 > 0$ and N a neighbourhood of $K_\beta = \{u \in H |\ E(u) = \beta,\ \nabla E(u) = 0\}$. Then there exists an $\varepsilon \in (0, \varepsilon_0)$ and continuous 1-parameter family of homeomorphisms $(\Phi(t, \cdot))_{t \in \mathbb{R}}$ of H, such that*

a) $\Phi(t, u) = u$, if $t = 0$ or $\nabla E(u) = 0$ or $|E(u) - \beta| \geq \varepsilon_0$;

b) $\forall u \in H$ the mapping $t \mapsto E(\Phi(t, u))$ is non-increasing.

c) $\Phi(1, E_{\beta+\varepsilon} \setminus N) \subseteq E_{\beta-\varepsilon}$ and $\Phi(1, E_{\beta+\varepsilon}) \subseteq E_{\beta-\varepsilon} \cup N$

Proof. As a reminder: $E_\beta := \{u \in H |\ E(u) < \beta\}$, $N_{\beta,\delta} = \{u \in H |\ |E(u) - \beta| < \delta,\ \|\nabla E(u)\| < \delta\}$ and $U_{\beta,\rho} := \bigcup_{u \in K_\beta} B_\rho(u)$.
We proceed in several steps:

1. First we construct suitable neighbourhoods:
   Since E satisfies a Palais-Smale condition, Lemma 4.10 is applicable. Hence $(N_{\beta,\delta})_{\delta > 0}$ and $(U_{\beta,\rho})_{\rho > 0}$ are neighbourhood bases for K_β. Therefore we find $\rho \in (0, 1)$ and $\delta \in (0, 1)$ such that

   $$K_\beta \subseteq N_{\beta,\delta} \subseteq U_{\beta,\rho} \subseteq U_{\beta,2\rho} \subseteq N.$$

   Furthermore there are $\rho^*, \delta^* \in (0, \frac{\delta}{2})$ with

   $$N_{\beta,\delta^*} \subseteq U_{\beta,\rho^*} \subseteq U_{\beta,2\rho^*} \subseteq N_{\beta,\delta}.$$

This allows us to find an $\eta \in C^{0,1}(H, \mathbb{R})$ satisfying

$$0 \leq \eta \leq 1,\ \eta(u) = \begin{cases} 1, & \text{if } u \in H \setminus N_{\beta,\delta} \\ 0, & \text{if } u \in N_{\beta,\delta^*}. \end{cases}$$

For example we can set

$$\eta(u) = \min\{1, \frac{1}{\rho^*}\operatorname{dist}(u, N_{\beta,\delta^*})\}.$$

Please note, that in case of $N = K_\beta = \emptyset$ the whole argument still makes sense by setting $\eta = 1$ and the other open sets to be $\emptyset$ as well.

2. We construct a suitable vectorfield:
Let $\varphi \in C^{0,1}(\mathbb{R}, \mathbb{R})$ be with

$$0 \leq \varphi \leq 1,\ \varphi(s) = \begin{cases} 1, \text{ if } |s - \beta| \leq \min\{\frac{\delta^*}{4}, \frac{\varepsilon_0}{2}\} \\ 0, \text{ if } |s - \beta| \geq \min\{\frac{\delta^*}{2}, \varepsilon_0\}. \end{cases}$$

This φ is needed to extend the pseudo-gradient from Lemma 5.5 to the whole of H.
Since ∇E is continuous, E is locally Lipschitz. Be careful, since H is usually infinite dimensional, the continuity of ∇E is really required:
Let $u \in H$ be fixated. Since ∇E is continuous, so is $v \mapsto \|\nabla E(v)\|$. Hence

$$W := (v \mapsto \|\nabla E(v)\|)^{-1}(]\|\nabla E(u)\| - 1, \|\nabla E(u)\| + 1[)$$

is open. Furthermore W is a neighbourhood of u, such that ∇E is bounded on it. Let $B_r(u) \subseteq W$ with $r > 0$ small enough. Integrating yields for $v \in B_r(u)$

$$\begin{aligned} &|E(u) - E(v)| = |\int_0^1 \frac{d}{dt}E(u + t(v-u))\,dt| \\ =&|\int_0^1 \langle \nabla E(u + t(v-u)), v - u\rangle\,dt| \leq \int_0^1 \sup_{w \in B_r(u)} \|\nabla E(w)\|\|v - u\|\,dt \\ \leq& \sup_{w \in B_r(u)} \|\nabla E(w)\|\|v - u\|. \end{aligned}$$

Since $\sup_{w \in B_r(u)} \|\nabla E(w)\| < \infty$, E is locally Lipschitz. Furthermore $\varphi \circ E$ is locally Lipschitz as well.
Now let $G : \tilde{H} \to H$ be a pseudo-gradient as in Def. 5.4 ($\tilde{H} = \{u \in H |\ \nabla E(u) \neq 0\}$) We define

$$e(u) := \begin{cases} -\eta(u)\varphi(E(u))G(u), & \text{if } u \in \tilde{H} \\ 0, & \text{if } u \notin \tilde{H}. \end{cases}$$

Now we show, that $e : H \to H$ is well defined and locally Lipschitz:
Let $u_0 \notin \tilde{H}$, i.e. $\nabla E(u_0) = 0$. Since ∇E is continuous, there exists an open neighbourhood V of u_0, such that $\|\nabla E(u)\| < \delta^*$ for all $u \in V$. Therefore for all $u \in V$ we have $e(u) = 0$, because

   (a) Case $|E(u) - \beta| < \delta^*$ yields $u \in N_{\beta,\delta^*}$. Hence $\eta(u) = 0$ which results in $e(u) = 0$.

(b) Case $|E(u) - \beta| \geq \delta^*$ yields $\varphi(E(u)) = 0$, which also results in $e(u) = 0$.

This yields e to be well defined and locally Lipschitz. Finally we have

$$\|G(u)\| \leq 2\min\{1, \|\nabla E(u)\|\} \leq 2,$$

which gives us

$$\forall u \in H: \ \|e(u)\| \leq 2. \tag{5.4}$$

3. Here we integrate the vectorfield from the last step to obtain a corresponding flow, which will yield the desired properties:
We consider the following initial value problem:

$$\frac{\partial}{dt}\Phi(t,u) = e(\Phi(t,u)), \ \Phi(0,u) = u. \tag{5.5}$$

The arguments for the Picard-Lindelöf theorem also apply here. Beware though that one would need to introduce integration with Hilbert space valued integrands to make it precise. This can be done similarly to the measure and integration theory in $\mathbb{R}^n$, but due to time constraints it is out of the scope of this lecture. Details can be found in [10], more precisely in [10, Thm 4.2.6] for the Hilbert space variant of Picard-Lindelöf. By (5.4) we have for a solution of (5.5)

$$\|\frac{\partial\Phi}{\partial t}(t,u)\| \leq 2 \ \Rightarrow \ \|\Phi(t,u) - \Phi(t',u)\| \leq 2|t - t'|$$

for all t, t' in the existence domain of $\Phi(\cdot, u)$. Hence there is no blow-up and we have together with [10, Thm 4.2.6] a unique, and continuous

$$\Phi : \mathbb{R} \times H \to H$$

satisfying (5.5). The continuity w.r.t. u can be shown via Gronwalls Lemma and the proof is mostly analogue to the finite dimensional case. Hence we skip that part as well.
The uniqueness also yields as in the finite dimensional case

$$\Phi(t, \Phi(s,u))) = \Phi(s+t, u),$$

and therefore

$$\Phi(t,\cdot)^{-1} = \Phi(-t,\cdot).$$

Hence $\Phi(t,\cdot) : H \to H$ is a homeomorphism for all $t \in \mathbb{R}$.

4. Here we check the desired properties:

(a) Since (5.5) has a unique solution, we have $\Phi(t,u) = u$, if e.g. $t = 0$ or if e.g. $\nabla E(u) = 0$ In the latter case we have $e(u) = 0$ and therefore $t \mapsto u$ is a solution to (5.5).
Furthermore if $|E(u) - \beta| \geq \varepsilon_0$, then $\varphi(E(u)) = 0$ and hence $e(u) = 0$. Again $t \mapsto u$ then solves (5.5). Uniqueness then yields the first point.

(b) We assume $(t, u) \in \mathbb{R} \times H$ to satisfy

$$\nabla E(\Phi(t, u)) \neq 0.$$

Then

$$\begin{aligned}&\frac{\partial}{\partial t} E(\Phi(t, u)) = \langle \nabla E(\Phi(t, u)), \frac{\partial}{\partial t}\Phi(t, u)\rangle \\ =& -\eta(\Phi(t, u))\varphi(E(\Phi(t, u)))\langle \nabla E(\Phi(t, u)), G(\Phi(t, u))\rangle \leq 0,\end{aligned}$$

because the scalar product is by Def. 5.4 nonnegative and the other terms are nonnegative as well.
Finally if $t_0 \in \mathbb{R}$ is with $\nabla E(\Phi(t_0, u)) = 0$, then from than point onwards the constant mapping $t \mapsto \Phi(t_0, u)$ solves (5.5) with initial value $\Phi(t_0, u)$. By the uniqueness property we therefore have for all t

$$\Phi(t, u) = \Phi(t_0, u)$$

and therefore the energy is non-increasing.

(c) We choose

$$0 < \varepsilon < \min\left\{\frac{\varepsilon_0}{2}, \frac{\delta^*}{4}, \frac{\rho\delta^2}{8}\right\}. \tag{5.6}$$

Let $u \in H$ satisfy

$$u \in E_{\beta+\varepsilon}, \text{ i.e. } E(u) < \beta + \varepsilon$$

and

$$u \notin N \text{ or } \Phi(1, u) \notin N.$$

Then we have to show, that

$$\Phi(1, u) \in E_{\beta-\varepsilon}, \text{ i.e. } E(\Phi(1, u)) < \beta - \varepsilon.$$

We proceed by contradiction and assume

$$E(\Phi(1, u)) \geq \beta - \varepsilon.$$

Since $t \mapsto E(\Phi(t, u))$ is nondecreasing and by the assumptions on u, we have

$$\forall t \in [0, 1] \ |E(\Phi(t, u)) - \beta| \leq \varepsilon < \min\left\{\frac{\varepsilon_0}{2}, \frac{\delta^*}{2}\right\}.$$

Hence we have $\varphi(\Phi(t, u)) = 1$ for these t. We define

$$S := \{t \in [0, 1] |\ \Phi(t, u) \notin N_{\beta,\delta}\}.$$

For all $t \in S$ we have

$$\|\nabla E(\Phi(t, u))\| \geq \delta.$$

By a distinction of cases (and showing in each case, that the integrands are nonnegativ), whether the gradient of E is zero or not, we get by chain rule

$$\begin{aligned}
&E(\Phi(1,u)) = E(\Phi(0,u)) + \int_0^1 \frac{\partial}{\partial t} E(\Phi(t,u))\, dt \\
=&E(u) + \int_0^1 \langle \nabla E(\Phi(t,u)), \partial_t \Phi(t,u)\rangle\, dt \\
\leq&E(u) + \int_S \langle \nabla E(\Phi(t,u)), \partial_t \Phi(t,u)\rangle\, dt \\
=&E(u) - \int_S \eta(\Phi(t,u)) \langle \nabla E(\Phi(t,u)), G(\Phi(t,u))\rangle\, dt \\
<&\beta + \varepsilon - \frac{1}{2}\int_S \eta(\Phi(t,u)) \min\{1, \|\nabla E(\Phi(t,u))\|\} \|\nabla E(\Phi(t,u))\|\, dt.
\end{aligned}$$

Since for all $t \in S$ we have

$$\Phi(t,u) \notin N_{\beta,\delta} \ \Rightarrow\ \eta(\Phi(t,u)) = 1.$$

Furthermore $S \neq \emptyset$: Since $\Phi(t_0,u) \notin N$ for $t_0 = 0$ or $t_0 = 1$ we have

$$\Phi(t_0,u) \notin N_{\beta,\delta} \subset N.$$

Therefore $t_0 \in S$. Hence we have $S \neq \emptyset$.
Since for all $t \in S$ we have $\|\nabla E(\Phi(t,u)\| \geq \delta$, we get

$$\begin{aligned}
&E(\Phi(1,u)) \\
<&\beta + \varepsilon - \frac{1}{2}\int_S \eta(\Phi(t,u)) \min\{1, \|\nabla E(\Phi(t,u))\|\} \|\nabla E(\Phi(t,u))\|\, dt \\
\leq&\beta + \varepsilon - \frac{\delta^2}{2}\mathcal{L}^1(S) = \beta + \varepsilon - \frac{\delta^2}{2}\mathcal{L}^1(\{t \in [0,1] |\ \Phi(t,u) \notin N_{\beta,\delta}\}).
\end{aligned}$$

Since $t_0 \in \{0,1\}$ with $\Phi(t_0,u) \notin N \supseteq N_{\beta,\delta}$ and $N_{\beta,\delta} \subseteq U_{\beta,\rho} \subseteq U_{\beta,2\rho} \subseteq N$ the map $t \mapsto \Phi(t,u)$ has to be outside of $N_{\beta,\delta}$ for at least a length of ρ. By $\|e(\cdot)\| \leq 2$ we have

$$\mathcal{L}^1(\{t \in [0,1] |\ \Phi(t,u) \notin N_{\beta,\varepsilon}\} \geq \frac{\rho}{2}.$$

Hence

$$E(\Phi(1,u)) < \beta + \varepsilon - \frac{1}{2}\delta^2 \frac{\rho}{2} = \beta + \varepsilon - \frac{\delta^2 \rho}{4} < \beta + \varepsilon - 2\varepsilon = \beta - \varepsilon,$$

which is a contradiction.

□

Remark 5.7.

1. $K_\beta = \emptyset$ is explicitly allowed. In this case $N = \emptyset$ can be choosen. The resulting homeomorphism then yields

$$\Phi(1, E_{\beta+\varepsilon}) \subseteq E_{\beta-\varepsilon},$$

which will be central in the proof of Mountain-Pass-Lemma 6.1 in the next section.

2. Φ is also called pseudo-gradient flow of E.

6 Mountain-Pass Lemma (by Ambrosetti-Rabinowitz)

Theorem 6.1 (Mountain-Pass Lemma)**.** *Let H be a Hilbert space and $E \in C^1(H)$ satisfies a Palais-Smale condition. Furthermore we assume*

1. $E(0) = 0$;
2. $\exists \alpha, \rho > 0$ *such that* $\|u\| = \rho \ \Rightarrow E(u) \geq \alpha$,
3. $\exists u_1 \in H$ *with* $\|u_1\| > \rho$ *satisfying* $E(u_1) < \alpha$.

We denote

$$\mathcal{P} := \{p \in C^0([0,1], H) |\ p(0) = 0,\ p(1) = u_1\}.$$

Then

$$\beta := \inf_{p \in \mathcal{P}} \sup_{u \in p([0,1])} E(u)$$

is a critical value of E. Furthemore E admits a critical point $u_{crit} \in H$ with $E(u_{crit}) = \beta \geq \alpha > 0$ and $\nabla E(u_{crit}) = 0$.

Proof. If we show

$$K_\beta = \{u \in H |\ E(u) = \beta,\ \nabla E(u) = 0\} \neq \emptyset$$

we are done. Hence we proceed by contradication and assume

$$K_\beta = \{u \in H |\ E(u) = \beta,\ \nabla E(u) = 0\} = \emptyset.$$

We set

$$\varepsilon_0 := \min\{\alpha, \alpha - E(u_1)\} > 0.$$

We are therefore allowed to choose $N = \emptyset$ in Theorem 5.6 and obtain a pseudo-gradient flow

$$\Phi : \mathbb{R} \times H \to H$$

with Φ being continuous and

$$\Phi(1, u) = u, \text{ if } |E(u) - \beta| \geq \varepsilon_0.$$

Furthermore there exists an $\varepsilon \in (0, \varepsilon_0)$ with

$$\Phi(1, E_{\beta+\varepsilon}) \subseteq E_{\beta-\varepsilon}.$$

By the definition of β we find a path $p_0 \in \mathcal{P}$, such that for all $s \in [0,1]$ we have

$$E(p_0(s)) < \beta + \varepsilon.$$

We define a new path $\tilde{p}$ by

$$\tilde{p}(s) := \Phi(1, p_0(s)).$$

$\tilde{p} \in \mathcal{P}$ because

$$|E(p_0(0)) - \beta| = |E(0) - \beta| = \beta \geq \alpha \geq \varepsilon_0$$

and

$$|E(p_0(1)) - \beta| = |E(u_1) - \beta| = \beta - E(u_1) \geq \alpha - E(u_1) \geq \varepsilon_0 > 0.$$

Hence the properties of Φ yield $\tilde{p}(0) = 0$ and $\tilde{p}(1) = u_1$, which in turn gives us $\tilde{p} \in \mathcal{P}$. Furthermore we have for all $s \in [0,1]$, that $p_0(s) \in E_{\beta+\varepsilon}$, hence the pseudo-gradient flow yields for all $s \in [0,1]$

$$\tilde{p}(s) = \Phi(1, p_0(s)) \in E_{\beta-\varepsilon}.$$

Hence

$$\beta = \inf_{p \in \mathcal{P}} \sup_{u \in p([0,1])} E(u) \leq \sup_{u \in \tilde{p}([0,1])} E(u) < \beta - \varepsilon,$$

which is a contradiction. So $K_\beta \neq \emptyset$ and the result follows. □

Now we return to example 4.3, i.e. we like to find a weak solution $u \neq 0$ of

$$\begin{cases} -\Delta u = g(\cdot, u), & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases} \tag{6.1}$$

The final theorem is as follows:

Theorem 6.2. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain, i.e. open, connected and bounded, $n \geq 2$. The nonlinearity $g : \Omega \times \mathbb{R} \to \mathbb{R}$ shall be continuous and satisfy*

1. *There exists a $p > 1$ with*

$$p < \begin{cases} \infty, & n = 2 \\ \frac{n+2}{n-2}, & n > 2 \end{cases}$$

   and a constant $C > 0$, such that for all $x \in \Omega, t \in \mathbb{R}$ we have subcritical growth, i.e.

$$|g(t,x)| \leq C(1 + |t|^p)$$

2. *We assume $g(x,0) = 0$. This yields $u = 0$ to be a weak solution to (6.1).*

3. *We define $G(x,t) := \int_0^t g(x,\tau)\, d\tau$. We assume there exists a $q > 2$ and $R_0 > 0$ such that for all $x \in \Omega$ and $t \in \mathbb{R}$ we have*

$$|t| \geq R_0 \;\Rightarrow\; 0 < qG(x,t) \leq tg(x,t),$$

   i.e. a superlinearity at ∞.

4. *$\forall \varepsilon > 0$ exists a $\delta > 0$ such that for all $t \in \mathbb{R}$ and all $x \in \Omega$ we have*

$$0 < |t| < \delta \;\Rightarrow\; \frac{g(x,t)}{t} \leq \varepsilon,$$

   i.e. a superlinearity at zero.

If we assume the above, there exists a weak nontrivial solution $u \in W_0^{1,2}(\Omega) \setminus \{0\}$ to (6.1).

Proof. The lemmata 4.4 and 4.8 yield

$$E : W_0^{1,2}(\Omega) \to \mathbb{R}, v \mapsto \frac{1}{2}\int_\Omega |\nabla u|^2\, dx - \int_\Omega G(x, u(x))\, dx$$

to be in $C^1(W_0^{1,2}(\Omega))$ and satisfy a Palais-Smale condition. Now we have to check the assumptions for the Mountain-Pass lemma 6.1:

1. $E(0) = 0$ is readily given, because $G(x, 0) = \int_0^0 g(x, \tau)\, d\tau = 0$.

2. We combine the subcritical growth condition with the superlinearity at zero and obtain that for all $\varepsilon > 0$ there exists a $C(\varepsilon) > 0$ such that for all $t \in \mathbb{R}$ and $x \in \Omega$ we have

   $$\operatorname{sgn}(t) g(x, t) \le \varepsilon |t| + C(\varepsilon)(p + 1)|t|^p.$$

   (The first term is for $|t| \le \delta$, δ from the superlinearity comndition, and the second term comes from the subcritical growth, i.e. writing $1 = \frac{\delta}{\delta} \le \frac{|t|}{\delta}$ to estimate $1 + |t|^p \le C(\delta)|t|^p$). This yields

   $$G(x, t) \le \frac{\varepsilon}{2} t^2 + C(\varepsilon)|t|^{p+1}.$$

   We define the frist eigenvalue of the Laplace operator

   $$\lambda_1 := \inf_{v \in W_0^{1,2}(\Omega)\setminus\{0\}} \frac{\|v\|_{W_0^{1,2}(\Omega)}}{\|v\|_{L^2(\Omega)}}.$$

   By the direct method one can show that this infimum is actually attained and therefore a minimum. Moreover $\lambda_1 > 0$. This is left as an exercise. All in all we estimate for $v \in W_0^{1,2}(\Omega)$ by Sobolev embedding 3.14

   $$\begin{aligned} E(v) \ge& \frac{1}{2}\int_\Omega |\nabla v|^2\, dx - \frac{\varepsilon}{2}\int_\Omega v^2\, dx - C(\varepsilon)\int_\Omega |v|^{p+1}\, dx \\ \ge& \frac{1}{2}(1 - \frac{\varepsilon}{\lambda_1^2})\int_\Omega |\nabla v|^2\, dx - C(\varepsilon, \Omega)\|v\|^{p+1}_{W_0^{1,2}(\Omega)} \\ =& \left(\frac{1}{2} - \frac{\varepsilon}{2\lambda_1^2} - C(\varepsilon, \Omega)\|v\|^{p-1}_{W_0^{1,2}(\Omega)}\right)\|v\|^2_{W_0^{1,2}(\Omega)}. \end{aligned}$$

   We choose $\varepsilon > 0$ in such a way, that $\frac{\varepsilon}{2\lambda_1^2} = \frac{1}{4}$. Then we choose $\rho = \rho(\lambda_1) > 0$ suitably, i.e. small enough, such that

   $$\inf_{\|v\|_{W_0^{1,2}(\Omega)} = \rho} E(v) \ge \left(\frac{1}{4} - C(\lambda_1, \Omega)\rho^{p-1}\right)\rho^2 =: \alpha > 0.$$

   This yields the second assumption of the Mountain-Pass lemma.

3. Now we need to show, that there exists a $u_1 \in W_0^{1,2}(\Omega)$ with

   $$\|u_1\|_{W_0^{1,2}(\Omega)} > \rho,\ E(u_1) < \alpha :$$

   The superlinearity at ∞ yields for $|t| \ge R_0$

   $$\begin{aligned} t|t|^q \frac{\partial}{\partial t}\left(|t|^{-q} G(x, t)\right) =& -qt|t|^q|t|^{-q-2} t G(x, t) + t|t|^q|t|^{-q} g(x, t) \\ =& -qG(x, t) + tg(x, t) \ge 0. \end{aligned}$$

Hence $t \geq R_0$ gives us

$$\frac{\partial}{\partial t}\left(|t|^{-q} G(x,t)\right) \geq 0,$$

which in turn yields monotinicity and therefore

$$|t|^{-q} G(x,t) \geq R_0^{-q} G(x,R_0).$$

If $t \leq -R_0$ we similarly have

$$\frac{\partial}{\partial t}\left(|t|^{-q} G(x,t)\right) \leq 0 \;\Rightarrow\; |t|^{-q} G(x,t) \geq R_0^{-q} G(x,-R_0).$$

All in all we have for $|t| \geq R_0$

$$G(x,t) \geq \gamma_0(x)|t|^q$$

with

$$\gamma_0(x) := R_0^{-q} \min\{G(x,R_0), G(x,-R_0)\}.$$

The superlinearity condition further yields

$$\forall x \in \Omega \; G(x,R_0) > 0 \text{ and } G(x,-R_0) > 0, \text{ i.e. } \gamma_0(x) > 0.$$

The subcritical growth condition implies that $\gamma_0(\cdot)$ is bounded.
Let $u_0 \in C_0^\infty(\Omega) \subseteq W_0^{1,2}(\Omega)$ with $u_0 \neq 0$. Further let $\sigma \geq \sigma_0$ with σ_0 to be suitably given later. We then have

$$\begin{aligned}
&\int_\Omega G(x,\sigma u_0(x))\,dx = \int_{\{x\in\Omega |\;\; \sigma|u_0(x)|\geq R_0\}} G(x,\sigma u_0(x))\,dx \\
&+ \int_{\{x\in\Omega |\;\; \sigma|u_0(x)|< R_0\}} G(x,\sigma u_0(x))\,dx \\
\geq &\int_{\{x\in\Omega |\;\; \sigma|u_0(x)|\geq R_0\}} \gamma_0(x)|\sigma u_0(x)|^q\,dx - \mathcal{L}^n(\Omega) \sup_{x\in\Omega, |t|\leq R_0} |G(x,t)| \\
\geq &\sigma^q \int_{\{x\in\Omega |\;\; \sigma_0|u_0(x)|\geq R_0\}} \gamma_0(x)|u_0(x)|^q\,dx - \mathcal{L}^n(\Omega) \sup_{x\in\Omega, |t|\leq R_0} |G(x,t)| \\
=: &\sigma^q C_0(u_0) - \mathcal{L}^n(\Omega) \sup_{x\in\Omega, |t|\leq R_0} |G(x,t)|
\end{aligned}$$

Since $\gamma_0(\cdot) > 0$ we have if $\sigma_0 > 0$ is big enough, that $C_0(u_0) > 0$. For $\sigma \geq \sigma_0$ we then have

$$\begin{aligned}
E(\sigma u_0) &= \frac{1}{2}\int_\Omega |\nabla(\sigma u_0)|^2\,dx - \int_\Omega G(x,\sigma u_0(x))\,dx \\
&\leq \frac{\sigma^2}{2}\int_\Omega |\nabla u_0|^2\,dx - \sigma^q C_0(u_0) + \mathcal{L}^n(\Omega) \sup_{x\in\Omega, |t|\leq R_0} |G(x,t)|.
\end{aligned}$$

Since $q > 2$ we therefore have

$$\lim_{\sigma\to\infty} E(\sigma u_0) = -\infty.$$

Hence we can choose $\sigma \geq \sigma_0$ to be so big, that

$$\|\sigma u_0\|_{W_0^{1,2}(\Omega)} > \rho \text{ and } E(\sigma u_0) < 0 < \alpha.$$

Setting $u_1 := \sigma u_0$ finishes this part of the proof.

All in all the Mountain-Pass lemma 6.1 is now applicable and it yields a critical point $u_{crit} \in W_0^{1,2}(\Omega)$ for E with $E(u_{crit}) > 0$, i.e. $DE(u_{crit}) = 0$. Since $E(0) = 0$ we therefore have $u_{crit} \neq 0$ and the result follows. $\square$

Remark: By a distinction of cases, i.e. $t > 0$ or $t < 0$, one can show, that assumption 4 implies assumption 2 in Thm. 6.2

Appendix

A BV-functions, minimising in a nonreflexive Banachspace

In case we have some time left at the end of the semester, we will talk about this chapter. We will introduce so called functions of bounded variation (in short BV-functions) and their corresponding Banachspace. In a sense the space of all BV-functions on an open set $\Omega \subseteq \mathbb{R}^n$ is the weak closure of $W^{1,1}(\Omega)$. This will allow us to define the so called Perimeter of a set, which is a generalisation of the area of the boundary of said set. Sets having such a perimeter are called Caccioppoli sets.
This is a wide field and a possible gateway to so called geometric measure theory. Good introductions are [2] and [7].
Afterwards we will give a geometric application, see [11, Thm. 1.4] (or [5, §2]).

Definition A.1. Let $\Omega \subseteq \mathbb{R}^n$ open and $u \in L^1_{loc}(\Omega)$. Then we define the total variation of u as

$$\int_\Omega |Du| := \sup \left\{ \int_\Omega u \operatorname{div} \varphi \, dx, \ \varphi \in C^1_0(\Omega, \mathbb{R}^n), |\varphi| \leq 1 \right\}.$$

We say $u \in L^1_{loc}(\Omega)$ is of bounded variation, if $\int_\Omega |Du| < \infty$.

Remark: Be careful, unlike the $W^{1,1}$-Norm, the total variation is in general not an integral w.r.t. Lebesgue measure. Hence there is no dx in the notation and the Symbol $\int_\Omega |Du|$ has to be interpreted as one Symbol.

Remark A.2. If $u \in L^1_{loc}(\Omega)$ is of bounded variation, i.e $\int_\Omega |Du| < \infty$, we find a vector valued Radon measure, which can be interpreted as the distributional derivative of u:
We define $L : C^1_0(\Omega, \mathbb{R}^n) \to \mathbb{R}$ as

$$L(\varphi) = \int_\Omega u \operatorname{div} \varphi \, dx.$$

By definition of the total variation, we get

$$|L(\varphi)| \leq \int_\Omega |Du| \|\varphi\|_{C^0(\Omega)}.$$

Since $C^1_0(\Omega) \subseteq C^0_0(\Omega)$ is dense w.r.t. to the supremum norm, we can extend L continuously to $C^0_0(\Omega)$, because L is linear.
Then the Riesz representation theorem for Radon measures (see e.g. [7, Thm 4.7, Ex. 4.19]) yields a Radon measure μ on Ω and a $\nu \in L^1(\mu, \mathbb{R}^n)$ with $|\nu| = 1$ μ-a.e. such that for all $\varphi \in C^0_0(\Omega, \mathbb{R}^n)$ we have

$$L(\varphi) = -\int_\Omega \langle \varphi, \nu \rangle \, d\mu.$$

The pair (ν, μ) is called a vector valued Radon measure. All in all this yields for $\varphi \in C^1_0(\Omega, \mathbb{R}^n)$

$$\int_\Omega u \operatorname{div} \varphi \, dx = L(\varphi) = -\int \langle \nu, \varphi \rangle \, d\mu.$$

Hence we interpret (ν, μ) as a measured valued derivative of u. Therefore we define the following notation

$$Du := (\nu, \mu), \;\; |Du| := \mu, \;\; \frac{Du}{|Du|} := \nu, \;\; \int \langle \varphi, Du \rangle := \int \langle \varphi, \nu \rangle \, d\mu.$$

Hence we have the following kind of partial integration

$$\int_\Omega u \operatorname{div} \varphi \, dx = -\int \langle \varphi, Du \rangle = -\int_\Omega \left\langle \varphi, \frac{Du}{|Du|} \right\rangle d|Du|.$$

Definition A.3. Let $\Omega \subseteq \mathbb{R}^n$ be open. We define the space of all functions of bounded variation or short BV-space as

$$BV(\Omega) := \left\{ u \in L^1(\Omega) | \int_\Omega |Du| < \infty \right\}$$

and equip it with the BV-norm

$$\|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + \int_\Omega |Du|.$$

Theorem A.4. *Let $\Omega \subseteq \mathbb{R}^n$ be open. Then $BV(\Omega)$ is a Banachspace, i.e. $\|\cdot\|_{BV(\Omega)}$ is a norm and $BV(\Omega)$ is complete.*

Proof. The triangle inequality of $u \mapsto \int |Du|$ follows by the subadditivity of the supremum, hence $BV(\Omega)$ is a normed space.
Further it is complete: Let $u_m \in BV(\Omega)$ be a Cauchy sequence. Since $\|\cdot\|_{L^1(\Omega)} \leq \|\cdot\|_{BV(\Omega)}$ it is also a Cauchy sequence in $L^1(\Omega)$. Hence there exists a $u \in L^1(\Omega)$, such that

$$u_m \to u \text{ in } L^1(\Omega).$$

Let $\varphi \in C_0^1(\Omega, \mathbb{R}^n)$, $\varepsilon > 0$. Then we find an $m_0 \in \mathbb{N}$, such that for all $\ell, m \geq m_0$ we have

$$\varepsilon \geq \|u_\ell - u_m\|_{BV(\Omega)}.$$

Therefore

$$\varepsilon \geq \int_\Omega (u_m - u_\ell) \operatorname{div}(\varphi) \, dx \to \int_\Omega (u_m - u) \operatorname{div}(\varphi) \, dx \text{ for } \ell \to \infty,$$

by Hölders inequality. Since φ is an arbitrary admissible function, we therefore have

$$\int_\Omega |D(u_m - u)| \leq \varepsilon$$

for $m \in \mathbb{N}$ big enough. Therefore

$$u_m \to u \text{ in } BV(\Omega)$$

and $u \in BV(\Omega)$. $\square$

The following Example A.5 shows, that $W^{1,1}(\Omega)$ is a closed subspace of $BV(\Omega)$.

Example A.5. Let $v \in C^1(\Omega) \cap W^{1,1}(\Omega)$. Then

$$\int_\Omega |\nabla v|\, dx = \int_\Omega |Dv|. \tag{A.1}$$

By the definition of $\int |Dv|$ and partial integration we get

$$\begin{aligned}
\int_\Omega |Dv| &= \sup\left\{\int_\Omega v \operatorname{div} \varphi\, dx,\ \varphi \in C_0^1(\Omega, \mathbb{R}^n), |\varphi| \le 1\right\} \\
&= \sup\left\{-\int_\Omega \langle \nabla v, \varphi\rangle\, dx,\ \varphi \in C_0^1(\Omega, \mathbb{R}^n), |\varphi| \le 1\right\} \\
&\le \sup\left\{\int_\Omega |\nabla v|\, dx \sup_\Omega |\varphi|\ \varphi \in C_0^1(\Omega, \mathbb{R}^n), |\varphi| \le 1\right\} \\
&\le \int_\Omega |\nabla v|\, dx.
\end{aligned}$$

Hence v is of bounded variation. The other estimate is as follows: Let $\nu_\varepsilon \in C_0^\infty(\Omega, [0,1])$ with $\nu_\varepsilon \uparrow 1$ locally uniformely. We set

$$\varphi_\varepsilon(x) := -\frac{\nabla v}{\sqrt{\varepsilon^2 + |\nabla v|^2}} \nu_\varepsilon(x).$$

Then $\varphi_\varepsilon \in C_0^0(\Omega)$ and $|\varphi_\varepsilon| \le 1$. By the above calculation and Remark A.2 it is an admissible function for the total variation. Hence

$$\begin{aligned}
\int_\Omega |Dv| \ge -\int_\Omega \nabla v \varphi_\varepsilon\, dx &= \int_\Omega \frac{|\nabla v|^2}{\sqrt{\varepsilon^2 + |\nabla v|^2}} \nu_\varepsilon \\
&\to \int_\Omega |\nabla v|\, dx \text{ for } \varepsilon \downarrow 0,
\end{aligned}$$

by Beppo-Levis monotone convergence theorem and

$$\frac{|\nabla v|^2}{\sqrt{\varepsilon^2 + |\nabla v|^2}} \nu_\varepsilon \uparrow |\nabla v| \text{ pointwise and monoton.}$$

Therefore

$$\int_\Omega |Dv| \ge \int_\Omega |\nabla v|\, dx$$

and the result follows.
Now for $u \in W^{1,1}(\Omega)$:
If we assume $\Omega \subset\subset \mathbb{R}^n$, we obtain by [4, Thm. 7.9] that $C^\infty(\Omega) \cap W^{1,1}(\Omega) \subseteq W^{1,1}(\Omega)$ is dense w.r.t. the $W^{1,1}$-Norm. Hence u can be approximated by $u_m \in C^\infty(\Omega) \cap W^{1,1}(\Omega)$ in the $W^{1,1}$-Norm. (A.1) yields

$$\|u\|_{BV(\Omega)} \leftarrow \|u_m\|_{BV(\Omega)} = \|u_m\|_{W^{1,1}(\Omega)} \to \|u\|_{W^{1,1}(\Omega)},$$

since the BV-Norm is continuous w.r.t. the $W^{1,1}$-Norm:

$$\begin{aligned}
&\int_\Omega |D(u - u_m)| = \sup\left\{\int_\Omega (u - u_m) \operatorname{div}(\varphi)\, dx|\ \varphi \in C_0^1(\Omega, \mathbb{R}^n),\ |\varphi| \le 1\right\} \\
&= \sup\left\{-\int_\Omega \nabla(u - u_m) \cdot \varphi\, dx|\ \varphi \in C_0^1(\Omega, \mathbb{R}^n),\ |\varphi| \le 1\right\} \\
&\le \int_\Omega |\nabla(u_m - u)|\, dx \to 0.
\end{aligned}$$

Hence all in all we have

$$\|u\|_{BV(\Omega)} = \|u\|_{W^{1,1}(\Omega)}. \tag{A.2}$$

Next we give an example, that shows, that $W^{1,1}(\Omega)$ is in general not the whole $BV(\Omega)$. The example is in essence an indicator function, i.e. a function with a 'jump' (cf. Example 3.3)

Example A.6. Let $\emptyset \neq E \subseteq \mathbb{R}^n$ be bounded, open, connected (i.e. a bounded domain) and with C^2-boundary. Then

$$\operatorname{area}(\partial E) = \int_{\mathbb{R}^n} |D\chi_E|.$$

with

$$\chi_E(x) = \begin{cases} 1, & x \in E \\ 0, & x \notin E. \end{cases}$$

Furthermore $\chi_E \notin W^{1,1}(\mathbb{R}^n)$.

Proof. First we show the equality for the area: Let $\varphi \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^n)$ with $|\varphi| \leq 1$. Let $\nu_E : \partial E \to \partial B_1(0)$ be the outer unit normal of E. Then partial integration yields

$$\begin{aligned} &\int_{\mathbb{R}^n} \chi_E \operatorname{div} \varphi \, dx = \int_E \operatorname{div} \varphi \, dx = \int_{\partial E} \langle \varphi, \nu_E \rangle \, darea_{\partial E} \\ \leq &\int_{\partial E} |\varphi||\nu_E| \, darea_{\partial E} = \operatorname{area}(\partial E). \end{aligned}$$

By the definition of the supremum we therefore have

$$\int_{\mathbb{R}^n} |D\chi_E| \leq \operatorname{area}(\partial E).$$

Let us turn our attention to the other inequality. Since the boundary of E is in C^2 and E is bounded, we find a function $\psi \in C_0^1(\mathbb{R}^n, \mathbb{R}^n)$ with $|\psi| \leq 1$ and

$$\psi|_{\partial E} = \nu_E.$$

The idea to construct such a ψ is to work locally around ∂E and extend the outer normal ν_E constantly in a small neigbourhood after straightening the boundary. Then one multiplies it with a cut off function and uses a partition of unity to define a global ψ. For the sake of time we will not work out the details here.
Then

$$\int_{\mathbb{R}^n} \chi_E \operatorname{div} \psi \, dx = \int_{\partial E} \langle \nu_E, \psi \rangle \, darea_{\partial E} = \int_{\partial E} darea_{\partial_E} = \operatorname{area}(\partial E).$$

Now we show, that $\chi_E \notin W^{1,1}(\mathbb{R}^n)$: We proceed by contradiction and assume $\chi_E \in W^{1,1}(\mathbb{R}^n)$. Therefore $\nabla \chi_E \in L^1(\mathbb{R}^n)$ should exist weakly. Let $\varphi \in C_0^1(\mathbb{R}^n \setminus \partial E, \mathbb{R}^n)$. Hence we have by Gauss's Theorem

$$-\int \nabla \chi_E \varphi \, dx = \int \chi_E \operatorname{div} \varphi \, dx = \int_E \operatorname{div} \varphi \, dx = \int_{\partial E} \nu_E \cdot \varphi \, darea_{\partial E} = 0.$$

Therefore $\nabla\chi_E = 0$ almost everywhere, since $\mathcal{L}^n(\partial E) = 0$, because it is in C^2. If we choose $\varphi \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^n)$, we therefore get

$$0 = -\int_{\mathbb{R}^n} \nabla\chi_E \cdot \varphi \, dx = \int_E \operatorname{div}(\varphi) \, dx = \int_{\partial E} \langle \varphi, \nu_E \rangle \, darea_{\partial E}.$$

In general the last term is not zero (see e.g. construction for ψ above), which results in a contradication. $\square$

Remark: With a small modification of the proof, one can also show

$$\operatorname{area}(\partial E \cap \Omega) = \int_\Omega |D\chi_E|$$

for $\Omega \subseteq \mathbb{R}^n$ open and bounded (Homework).

This example motivates the following definition, which generalises the notion of area measure of a boundary:

Definition A.7. Let $\Omega \subseteq \mathbb{R}$ be open and $E \subseteq \mathbb{R}^n$ measurable. We define the (relative) Perimeter of E in Ω by

$$P(E, \Omega) := \int_\Omega |D\chi_E| \in [0, \infty].$$

If $\Omega = \mathbb{R}^n$, we set

$$P(E) := \int |D\chi_E|.$$

If $P(E) < \infty$, E is called a Caccioppoli set.

We will apply the direct method later. Hence we need some kind of weak lower semicontinuity for the BV-seminorm. The following will play that role:

Theorem A.8. *Let $\Omega \subset \mathbb{R}^n$ be open, $v_k \in BV(\Omega)$ with*

$$v_k \to v \text{ in } L^1(\Omega).$$

Then

$$\int_\Omega |Dv| \leq \liminf_{k\to\infty} \int_\Omega |Dv_k|.$$

This also yields if $v_k \to v$ in $L^1(\Omega)$ and $\|v_k\|_{BV(\Omega)} \leq C$, then $v \in BV(\Omega)$.

Proof. W.l.o.g. we assume $M := \liminf_{k\to\infty} \int_\Omega |Dv_k| < \infty$. Let $\varepsilon > 0$ be arbitrary. The definition of $\liminf$ yields a subsequence k_ℓ, such that

$$\int_\Omega |Dv_{k_\ell}| \leq M + \varepsilon.$$

Let $\varphi \in C_0^1(\Omega, \mathbb{R}^n)$ with $|\varphi| \leq 1$ arbitrary. Then by Hölders inequality

$$M + \varepsilon \geq \int_\Omega v_{k_\ell} \operatorname{div}\varphi \, dx \to \int_\Omega v \operatorname{div}\varphi \, dx \text{ for } \ell \to \infty.$$

Hence

$$\int_\Omega |Dv| \leq M + \varepsilon = \liminf_{k\to\infty} \int_\Omega |Dv_k| + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the result follows. $\square$

Now we present a compactness result. The proof would involve a smoothing procedure, for which we do not have time. Hence we will not do a proof here.

Theorem A.9. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain with $C^{0,1}$-regularity, i.e. a Lipschitz boundary. Furthermore let $v_k \in BV(\Omega)$ and $C > 0$ such that*

$$\|v_k\|_{BV(\Omega)} = \int_\Omega |Dv_k| + \|v_k\|_{L^1(\Omega)} \leq C,$$

i.e. a bounded sequence. Then there exists a subsequence v_{k_ℓ} and a $v \in BV(\Omega)$, such that

$$v_{k_\ell} \to v \text{ in } L^1(\Omega).$$

Proof. See [2, §5.2.3] for a proof. ☐

Now we apply our theory for BV-functions to a geometric problem. This has been taken out of [11, Chapter 1, Thm 1.4].

Theorem A.10 (Minimal bisecting hypersurfaces)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain with Lipschitz boundary. We define the set of all bisecting sets of Ω*

$$M := \left\{ E \subseteq \Omega \text{ measurable} \middle| \, \mathcal{L}^n(E) = \mathcal{L}^n(\Omega \setminus E) = \frac{1}{2}\mathcal{L}^n(\Omega) \right\}.$$

Then $M \neq \emptyset$ and there exists a $G \in M$ minimising the perimeter, i.e. for all $E \in M$ we have

$$P(G, \Omega) \leq P(E, \Omega).$$

Proof. We start by showing $M \neq \emptyset$:
Since Ω is bounded, we find an $R > 0$, such that

$$\Omega \subseteq \,]-R, R[^n =: \tilde{E}.$$

Then we define

$$\tilde{E}_t := \tilde{E} + te_1$$

and

$$f(t) := \mathcal{L}^n(\tilde{E}_t \cap \Omega).$$

f is continuous: Let $t_k \to t$. Since $\tilde{E}_t \cap \Omega$ has a Lipschitz boundary, we have

$$\mathcal{L}^n(\partial(\tilde{E}_t \cap \Omega)) = 0.$$

If $x \notin \partial(\tilde{E}_t \cap \Omega)$, then there is a small open neighbourhood U of x, such that $U \cap \partial(\tilde{E}_t \cap \Omega) = \emptyset$. Hence there exists a $k_0 \in \mathbb{N}$, such that we either have

$$x \in \tilde{E}_{t_k} \cap \Omega \text{ for all } k \geq k_0 \text{ or } x \notin \tilde{E}_{t_k} \cap \Omega \text{ for all } k \geq k_0.$$

Therefore

$$\chi_{\tilde{E}_{t_k} \cap \Omega} \to \chi_{\tilde{E}_t \cap \Omega} \text{ pointwise a.e.}$$

Since Ω is bounded, the dominated convergence theorem yields the continuity of f.
Now for $t = 0$ we have $f(0) = \mathcal{L}^n(\Omega) > 0$. Since Ω is bounded, there is a $t_0 \in \mathbb{R}$,

such that $\tilde{E}_{t_0} \cap \Omega = \emptyset$. Hence $f(t_0) = 0$. The intermediate value theorem yields a $\tilde{t}$ such that

$$\mathcal{L}^n(\tilde{E}_{\tilde{t}} \cap \Omega) = \frac{1}{2}\mathcal{L}^n(\Omega).$$

The additivity of the measure now yields

$$\tilde{E}_{\tilde{t}} \cap \Omega \in M.$$

Now we turn to the minimisation of the perimeter. We follow the same spirit of proof as in Theorem 2.17:
Let $E_k \in M$ be a minimising sequence, i.e.

$$\lim_{k\to\infty} P(E_k, \Omega) = \inf_{E\in M} P(E, \Omega).$$

If the infimum is ∞, $\tilde{E}_{\tilde{t}} \cap \Omega$ will be a minimiser. If it is finite, there exists a constant $C > 0$, such that (for k large, i.e. choosing a subsequence)

$$\|\chi_{E_k}\|_{BV(\Omega)} = \mathcal{L}^n(E_k) + P(E_k, \Omega) \leq C.$$

By Theorem A.9 we find a $u \in BV(\Omega)$ and a subsequence, such that after relabeling we have

$$\chi_{E_k} \to u \text{ in } L^1(\Omega).$$

Then Theorem A.8 yields

$$\int_\Omega |Du| \leq \liminf_{k\to\infty} \int_\Omega |D\chi_{E_k}| = \liminf_{k\to\infty} P(E_k, \Omega).$$

By the addendum to the Riesz-Fischer theorem we can extract another subsequence and obtain

$$\chi_{E_k} \to u \text{ pointwise a.e.}$$

Since $\chi_{E_k}(x) \in \{0, 1\}$ we have a.e.

$$u(x) = \chi_G(x)$$

for some measurable set $G \subseteq \Omega$. The L^1-convergence yields

$$\frac{1}{2}\mathcal{L}^n(\Omega) = \mathcal{L}^n(E_k) = \int_\Omega \chi_{E_k}\, dx \to \int_\Omega u\, dx = \int_\Omega \chi_G\, dx = \mathcal{L}^n(G).$$

Therefore $G \in M$. Then the lower semicontinuity property yields

$$\inf_{E\in M} P(E, \Omega) \leq P(G, \Omega) \leq \liminf_{k\to\infty} P(E_k, \Omega) = \inf_{E\in M} P(E, \Omega).$$

Hence G is a desired minimiser. □

References

[1] L. C. Evans. *Partial Differential Equations.* AMS, Graduate Studies in Mathematics, second edition, 1998.

[2] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions.* CRC Press, 1992.

[3] F. Gazzola, H.-Chr. Grunau, and G. Sweers. *Polyharmonic Boundary Value Problems.* Springer, 2010. Lecture Notes in Mathematics 1991.

[4] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer-Verlag Berlin, 3rd edition, 1998.

[5] H.-Chr. Grunau. Variationsmethoden. Lecture given in Magdeburg, Germany, 2013-14.

[6] J. L. Kelley. *General topology.* van Nostrand, 1955.

[7] F. Maggi. *Sets of Finite Perimeter and Geometric Variational Problems.* Cambridge University Press, 1st edition, 2012.

[8] W. Rudin. *Functional Analysis.* Mc Graw Hill, 1973.

[9] W. Rudin. *Real and Complex Analysis.* McGraw Hill, third edition, 1987.

[10] C. Seifert, S. Trostorff, and M. Waurick. *Evolutionary Equations Picard's Theorem for Partial Differential Equations, and Applications.* Birkhäuser, 2022. `https://doi.org/10.1007/978-3-030-89397-2_4`.

[11] Michael Struwe. *Variational Methods.* Springer, 2008.

Index