PDEs in Science and Engineering

Lecture Notes

FERNANDO QUIRÓS GRACIÁN

8th January 2020

Chapter 1

Sobolev spaces

1.1 Some results about integration that everyone must know

We assume that the student is familiar with the notion of measurable function and integrable function (we will work always with Lebesgue's measure). If this is not the case, you should review this material in a good book on measure theory, for example [5]. You should at least understand and know how to use the results below.

Definition 1.1.1. Let $\Omega \subset \mathbb{R}^d$ be an open set. We set

$$L^1(\Omega) = \{f : \Omega \to \mathbb{R}; f \text{ measurable }, \int_\Omega |f| < \infty\}, \qquad \|f\|_{L^1(\Omega)} = \|f\|_1 = \int_\Omega |f|.$$

Theorem 1.1.1. *$L^1(\Omega)$ is a vector space and $\|\cdot\|_1$ is a norm. The space $L^1(\Omega)$ with this norm is a Banach space.*

Theorem 1.1.2 (Monotone Convergence Theorem, Beppo Levi). *Let (f_n) be a sequence of functions in $L^1(\Omega)$ that satisfy*

$$\text{(a) } f_n \leq f_{n+1} \text{ a.e. on } \Omega \text{ for all } n \in \mathbb{N}; \quad \text{(b) } \sup_n \int_\Omega f_n < \infty.$$

Then f_n converges a.e. on Ω to a finite limit, which we denote by f; the function f belongs to $L^1(\Omega)$, and $\|f_n - f\|_1 \to 0$.

Theorem 1.1.3 (Dominated Convergence Theorem, Lebesgue). *Let (f_n) be a sequence of functions in $L^1(\Omega)$ that satisfy*
(a) *$f_n \to f$ a.e. on Ω;*
(b) *there is a function $g \in L^1(\Omega)$ such that $|f_n| \leq g$ a.e. in Ω for all n.*
Then $f \in L^1(\Omega)$, and $\|f_n - f\|_1 \to 0$.

Remark. If $f_n \to f$ in $L^1(\Omega)$, then $\int_\Omega f_n \to \int_\Omega f$ (we can pass to the limit inside the integral). Indeed,

$$\left|\int_\Omega f_n - \int_\Omega f\right| \leq \int_\Omega |f_n - f| = \|f_n - f\|_{L^1(\Omega)} \to 0.$$

Lemma 1.1.4 (Fatou's lemma). *Let (f_n) be a sequence of functions in $L^1(\Omega)$ that satisfy*

$$\text{(a) } \textit{for all } n,\ f_n \geq 0 \textit{ a.e. in } \Omega \textit{ for all } n \in \mathbb{N}; \quad \text{(b) } \sup_n \int_\Omega f_n < \infty.$$

For almost every $x \in \Omega$ we set $f(x) = \liminf_{n\to\infty} f_n(x) \leq \infty$. Then $f \in L^1(\Omega)$, and

$$\int_\Omega f \leq \liminf_{n\to\infty} \int_\Omega f_n.$$

Theorem 1.1.5 (density). *The space $C_c(\Omega)$ is dense in $L^1(\Omega)$.*

Theorem 1.1.6 (Tonelli). *Let $F : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a measurable function satisfying*

$$\text{(a) } \int_{\Omega_2} |F(x,y)|\,dy < \infty \textit{ for a.e. } x \in \Omega_1; \quad \text{(b) } \int_{\Omega_1} \left(\int_{\Omega_2} |F(x,y)|\,dy \right) dx < \infty.$$

Then $F \in L^1(\Omega_1 \times \Omega_2)$.

Theorem 1.1.7 (Fubini). *Assume that $F \in L^1(\Omega_1 \times \Omega_2)$. Then, for a.e. $x \in \Omega_1$, $F(x,\cdot) \in L^1(\Omega_2)$ and $\int_{\Omega_2} F(\cdot,y)\,dy \in L^1(\Omega_1)$. Similarly, for a.e. $y \in \Omega_2$, $F(\cdot,y) \in L^1(\Omega_1)$ and $\int_{\Omega_1} F(x,\cdot)\,dx \in L^1(\Omega_2)$. Moreover,*

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

Definition 1.1.2. Let U and V be open subsets of $\mathbb{R}^d$. We say that U is *strongly included* in V, and we write $U \subset\subset V$, if $\overline{U}$ is compact and $\overline{U} \subset V$.

Notation. We wil denote the indicator function of a set A by $\mathbb{1}_A$, that is,

$$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

Definition 1.1.3. We say that $f \in L^1_{\text{loc}}(\Omega)$ if $f\mathbb{1}_{\Omega'} \in L^1(\Omega')$ for all $\Omega' \subset\subset \Omega$.

Theorem 1.1.8 (Lebesgue's differentiation theorem). *If $f \in L^1_{\text{loc}}(\Omega)$ then*

$$\lim_{r\to 0^+} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\,dy = 0 \quad \textit{for almost every } x \in \Omega.$$

As a consequence, $\displaystyle\lim_{r\to 0^+} \frac{1}{|B_r(x)|} \int_{B_r(x)} f = f(x)$ *for almost every $x \in \Omega$.*

Remark. A point $x \in \Omega$ for which the result holds is called a Lebesgue point for f.

1.2 A review of L^p spaces

You are expected to be familiar with some basic results on L^p spaces listed below. If this is not the case, it may be a good idea to read chapters 4 and 5 of the book [2].

Definition 1.2.1. Let $\Omega \subset \mathbb{R}^d$ be an open set and $1 \le p < \infty$. We set

$$L^p(\Omega) = \{f : \Omega \to \mathbb{R};\ f \text{ measurable}\ , \int_\Omega |f|^p < \infty\}, \qquad \|f\|_{L^p(\Omega)} = \|f\|_p = \Big(\int_\Omega |f|^p\Big)^{1/p}.$$

Definition 1.2.2. Let $\Omega \subset \mathbb{R}^d$ be an open set. A measurable function $f : \Omega \to \mathbb{R}$ is *essentially bounded* if there is a constant C such that $|f| \le C$ a.e. on Ω

Definition 1.2.3. Let $\Omega \subset \mathbb{R}^d$ be an open set. We set

$$L^\infty(\Omega) = \{f : \Omega \to \mathbb{R};\ f \text{ measurable and essentially bounded}\},$$
$$\|f\|_{L^\infty(\Omega)} = \|f\|_\infty = \inf\{C > 0 : |f| \le C \ \text{ a.e. on } \Omega\}.$$

Notation. Let $1 \le p \le \infty$; we denote by p' the *conjugate exponent*,

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

Theorem 1.2.1 (Hölder's inequality). *Let $f \in L^p$, $g \in L^{p'}$, $1 \le p \le \infty$. Then $fg \in L^1$ and*

$$\|fg\|_1 \le \|f\|_p \|g\|_{p'}.$$

Theorem 1.2.2. *$L^p(\Omega)$ is a vector space and $\|\cdot\|_p$ is a norm for any p, $1 \le p \le \infty$.*

Theorem 1.2.3. *The bilinear form $(f, g) = \int_\Omega fg$ is a scalar product in $L^2(\Omega)$.*

Theorem 1.2.4 (Fischer-Riesz). *$L^p(\Omega)$ with the norm $\|\cdot\|_p$ is a Banach space for all p, $1 \le p \le \infty$.*

Remark. $\|f\|_2 = \sqrt{(f, f)}$. Hence L^2 with the scalar product defined above is a Hilbert space.

Theorem 1.2.5. *$C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for all $p \in [1, \infty)$.*

Remark. The result is not true for $p = \infty$.

Theorem 1.2.6. *Let (f_n) be a sequence in L^p, and let $f \in L^p$ be such that $\|f_n - f\|_p \to 0$. Then, there exists a subsequence (f_{n_k}) and a function $h \in L^p$ such that*

$$f_{n_k} \to f \quad \text{a.e. in } \Omega, \qquad |f_{n_k}| \le h \quad \text{a.e. in } \Omega \text{ for all } k.$$

Definition 1.2.4. We say that $f \in L^p_{\text{loc}}(\Omega)$, $1 \le p \le \infty$, if $f\mathbb{1}_{\Omega'} \in L^p(\Omega')$ for all $\Omega' \subset\subset \Omega$.

Definition 1.2.5. Let p, $1 \le p \le \infty$. A sequence (f_n) in $L^p_{\text{loc}}(\Omega)$ converges in $L^p_{\text{loc}}(\Omega)$ to a function $f \in L^p_{\text{loc}}(\Omega)$ if $f_n\mathbb{1}_{\Omega'} \to f\mathbb{1}_{\Omega'}$ for all $\Omega' \subset\subset \Omega$.

1.3 Weak derivatives, Sobolev Spaces

Definition 1.3.1. A function $f \in L^1_{\text{loc}}(\Omega)$ is weakly differentiable with respect to x_i if there exists a function $g_i \in L^1_{\text{loc}}(\Omega)$ such that

$$\int_\Omega f \partial_i \phi = -\int_\Omega g_i \phi \quad \text{for all } \phi \in C_{\text{c}}^\infty(\Omega).$$

The function g_i is called the weak ith partial derivative of f and is denoted by $\partial_i f$.

Thus, for weak derivatives the integration by parts formula

$$\int_\Omega f \partial_i \phi = -\int_\Omega \partial_i f \phi$$

holds by definition for all $\phi \in C_{\text{c}}^\infty(\Omega)$.

Weak derivatives are unique. To show this we will need the following standard lemma.

Lemma 1.3.1. *Let $u \in L^1_{\text{loc}}(\Omega)$ be such that*

$$\int_\Omega u\phi = 0 \quad \text{for all } \phi \in C_{\text{c}}^\infty(\Omega).$$

Then $u = 0$ a.e. in Ω.

Proof. This is problem 3 in worksheet 2. □

Lemma 1.3.2 (uniqueness of weak derivatives). *Let $g, h \in L^1_{\text{loc}}(\Omega)$ be weak derivatives with respect to x_i of $f \in L^1_{\text{loc}}(\Omega)$. Then $g = h$ a.e. in Ω.*

Proof. Because of the definition of weak derivative,

$$\int_\Omega g\phi = -\int_\Omega f \partial_i \phi = \int_\Omega h\phi$$

for all $\phi \in C_{\text{c}}^\infty(\Omega)$. Hence,

$$\int_\Omega (g - h)\phi = 0 \quad \text{for all } \phi \in C_{\text{c}}^\infty(\Omega)$$

and the result follows by Lemma 1.3.1. □

Remark. The weak derivative of a continuously differentiable function agrees with the pointwise derivative. The existence of a weak derivative is, however, not equivalent to the existence of a pointwise derivative almost everywhere.

Example. Let $f \in C(\mathbb{R})$ given by

$$f(x) = \{x\}_+ = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

Then f is weakly differentiable with weak derivative given by the so-called Heaviside function,

$$f'(x) = H(x) := \mathbb{1}_{[0,\infty)}(x). \tag{1.1}$$

The choice of the value of $f'(x)$ at $x = 0$ is irrelevant, since the weak derivative is only defined up to pointwise almost everywhere equivalence. To prove (1.1), note that for any $\phi \in C_c^\infty(\mathbb{R})$ an integration by parts gives

$$\int_{\mathbb{R}} f\phi' = \int_0^\infty x\phi' = -\int_0^\infty \phi = -\int_{\mathbb{R}} H\phi.$$

Example. The Heaviside function is not weakly differentiable. Indeed, if it had a weak derivative $H' \in L^1_{\text{loc}}(\mathbb{R})$, then

$$\int_{-\infty}^\infty H'\phi = -\int_{-\infty}^\infty H\phi' = -\int_0^\infty \phi' = \phi(0) \quad \text{for all } \phi \in C_c^\infty(\mathbb{R}).$$

Choose a sequence (ϕ_k) in $C_c^\infty(\mathbb{R})$ such that $0 \le \phi_k \le 1$, $\phi_k(x) = 0$ if $x \notin [-1,1]$, $\phi_k(0) = 1$, $\phi_k(x) \to 0$ for all $x \neq 0$. Notice that $|H'\phi_k| \le |H'|\mathbb{1}_{(-1,1)} \in L^1(\mathbb{R})$, since, by assumption, $H' \in L^1_{\text{loc}}(\mathbb{R})$. Then, applying Lebesgue's dominated convergence theorem we would get $0 = 1$, a contradiction.

Let us show now see for the one-dimensional case that if the weak derivative is 0 in an interval, the function is a constant there. For the result in higher dimensions see problem 1 in worksheet number 3.

Lemma 1.3.3. *Let $I = (a,b)$, with $-\infty \le a < b \le \infty$. Let $f \in L^1_{\text{loc}}(I)$ be such that*

$$\int_I f\phi' = 0 \quad \text{for all } \phi \in C_c^\infty(I). \tag{1.2}$$

Then there exists a constant C such that $f = C$ almost everywhere in I.

Proof. Fix a function $\psi \in C_c^\infty(I)$ such that $\int_I \psi = 1$. For any function $w \in C_c^\infty(I)$ there exists $\phi \in C_c^\infty(I)$ such that

$$\phi' = h, \quad \text{where } h := w - \psi \int_I w.$$

Indeed, the function $h \in C_c^\infty(I)$, and $\int_I h = 0$. Therefore h has a (unique) primitive in $C_c^\infty(I)$. We deduce from (1.2) that

$$\int_I f\Big(w - \psi \int_I w\Big) = 0 \quad \text{for all } w \in C_c^\infty(I),$$

that is,

$$\int_I \Big(f - \int_I (f\psi)\Big) w = 0 \quad \text{for all } w \in C_c^\infty(I),$$

and therefore, by Lemma 1.3.1, $f - \int_I (f\psi) = 0$ a.e. on I, i.e., $f = C$ a.e. on I with $C = \int_I (f\psi)$. □

Example. Let us recall the construction of the Cantor set. Let K_0 be the interval $[0,1]$. For each positive integer n we construct a compact set K_n by removing from K_{n-1} the open middle third of each of the intervals making up K_{n-1}. The *Cantor set* K is given by $K = \cap_{n=0}^{\infty} K_n$. It is well known that K is a compact set which has the cardinality of the continuum, but has Lebesgue measure zero.

The *Cantor singular function* is the function $f : [0,1] \to \mathbb{R}$ defined as follows. For each x in the interval $(1/3, 2/3)$ let $f(x) = 1/2$. Thus f is now defined in each point removed from $[0,1]$ in the construction of K_1. Next define f at each point removed from K_1 in the construction of K_2 by letting $f(x) = 1/4$ if $x \in (1/9, 2/9)$ and letting $f(x) = 3/4$ if $x \in (7/9, 8/9)$. Continue in this way, letting $f(x) = 1/2^n, 3/2^n, 5/2^n, \dots$ on the various intervals removed from K_{n-1} in the construction of K_n. Now f is defined in the open set $[0,1] \setminus K$, is non-decreasing, and has values in $[0,1]$. Extend it to all of $[0,1]$ by letting $f(0) = 0$ and letting

$$f(x) = \sup\{f(t) : t \in [0,1] \setminus K, t < x\} \quad \text{if } x \in K, x \neq 0.$$

It is easy to check that f is non-decreasing and continuous, with $f(0) = 0$ and $f(1) = 1$.

Let us see that the Cantor singular function is not weakly differentiable in $(0,1)$. Indeed, let I be any of the open middle third intervals removed in the construction of the Cantor set. Taking test functions ϕ whose supports are compactly contained in I, and using the fact that f is constant on I, we find that $\int_I f\phi' = c\int_I \phi' = 0$. Hence, if f is weakly differentiable in $(0,1)$, then $f' = 0$ almost everywhere in $(0,1)$. But the only weakly differentiable functions with zero weak derivative in a connected set are the ones that are equivalent to a constant function. This is a contradiction, so the Cantor function is not weakly differentiable.

Remark. In defining *strong* solutions of a differential equation, that is, functions that satisfy the equation pointwise a.e., but which are not necessarily continuously differentiable classical solutions, it is important to include the condition that the solutions are weakly differentiable. For example, up to pointwise a.e. equivalence, the only weakly differentiable functions $u : \mathbb{R} \to \mathbb{R}$ ODE $u' = 0$ pointwise a.e. are the constant functions. There are, however, many non-constant functions that are differentiable pointwise a.e. and satisfy the ODE pointwise a.e., but these solutions are not weakly differentiable; the step function and the Cantor function are examples.

Next lemma is a "weak version" of the Fundamental Theorem of Calculus.

Lemma 1.3.4. *Let $I = (a,b)$, with $-\infty \le a < b \le \infty$. Let $g \in L^1_{\rm loc}(I)$; for y_0 fixed in I, set*

$$v(x) = \int_{y_0}^{x} g, \quad x \in I.$$

Then $v \in C(I)$ and

$$\int_I v\phi' = -\int_I g\phi \quad \text{for all } \phi \in C_{\rm c}^{\infty}(I).$$

Proof. We have

$$\int_I v\phi' = \int_I \Big(\int_{y_0}^{x} g(t)\,dt\Big)\phi(x)\,dx = -\int_{y_0}^{a}\int_{x}^{y_0} g(t)\phi(x)\,dtdx + \int_{y_0}^{b}\int_{y_0}^{x} g(t)\phi(x)\,dtdx.$$

By Fubini's theorem,

$$\int_I v\phi' = -\int_a^{y_0}\int_a^t \phi'(x)\,dx\,g(t)\,dt + \int_{y_0}^b\int_t^b \phi'(x)\,dx\,g(t)\,dt = -\int_I g(t)\phi(t)\,dt.$$

□

Proposition 1.3.5. *If $f \in L^1_{\mathrm{loc}}(\Omega)$ has a weak partial derivative $\partial_i f \in L^1_{\mathrm{loc}}(\Omega)$ and $\psi \in C^\infty(\Omega)$, then ψf is weakly differentiable with respect to x_i and*

$$\partial_i(\psi f) = f\partial_i\psi + \psi\partial_i f.$$

Proof. Let $\phi \in C^\infty_{\mathrm{c}}(\Omega)$ be any test function. Then $\psi\phi \in C^\infty_{\mathrm{c}}(\Omega)$ and the weak differentiability of f implies that

$$\int_\Omega f\partial_i(\psi\phi) = -\int_\Omega \psi\phi\partial_i f.$$

Expanding $\partial_i(\psi\phi) = \psi\partial_i\phi + \phi\partial_i\psi$, and rearranging the result we get

$$\int_\Omega \psi f\partial_i\phi = -\int_\Omega (f\partial_i\psi + \psi\partial_i f)\phi.$$

□

Notation. Given a multi-index $\alpha \in \mathbb{N}_0^d$ we denote $D^\alpha u := \partial_1^{\alpha_1}\cdots\partial_d^{\alpha_d}u$. The order of α is defined as $|\alpha| := \alpha_1 + \cdots + \alpha_d$.

Definition 1.3.2. Let Ω be an open subset of $\mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$ a multi-index. We say that $v \in L^1_{\mathrm{loc}}(\Omega)$ is the α-weak derivative of $u \in L^1_{\mathrm{loc}}(\Omega)$ if

$$\int_\Omega uD^\alpha\phi = (-1)^{|\alpha|}\int_\Omega v\phi \quad \text{for all } \phi \in C^\infty_{\mathrm{c}}(\Omega).$$

The function v is denoted by $D^\alpha u$.

The order of weak differentiation does not matter.

Proposition 1.3.6. *Suppose that $f \in L^1_{\mathrm{loc}}(\Omega)$ and that the weak derivatives $D^\alpha f$, $D^\beta f$ exist for multi-indices $\alpha, \beta \in \mathbb{N}_0^d$. Then if any one of the weak derivatives $D^{\alpha+\beta}f$, $D^\alpha D^\beta f$, $D^\beta D^\alpha f$ exists, all three derivatives exist and are equal.*

Proof. Using the existence of $D^\alpha f$, and the fact that $D^\beta\varphi \in C^\infty_{\mathrm{c}}(\Omega)$ for any $\varphi \in C^\infty_{\mathrm{c}}(\Omega)$, we have

$$\int_\Omega D^\alpha f D^\beta\varphi = (-1)^{|\alpha|}\int_\Omega fD^{\alpha+\beta}\varphi.$$

This equation shows that $D^{\alpha+\beta}f$ exists if and only if $D^\beta D^\alpha f$ exists, and in that case the weak derivatives are equal. Using the same argument with α and β exchanged, we get the result. □

Definition 1.3.3 (Sobolev space). Let $\Omega \subset \mathbb{R}^d$ be an open set and $1 \le p \le \infty$. The Sobolev space $W^{k,p}(\Omega)$ is defined by

$$W^{k,p}(\Omega) := \{u \in L^1_{\mathrm{loc}}(\Omega) : D^\alpha u \in L^p(\Omega) \text{ for all } \alpha \in \mathbb{N}_0^d, |\alpha| \le k\}.$$

Notation. $H^k(\Omega) = W^{k,2}(\Omega)$.

Proposition 1.3.7. *The Sobolev space $W^{k,p}(\Omega)$ is a Banach space when equipped with the norm*

$$\|f\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha|\le k} \|D^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p} \qquad \text{for } p\in[1,\infty),$$
$$\|f\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha|\le k} \|D^\alpha f\|_{L^\infty(\Omega)}.$$

The Sobolev space $H^k(\Omega)$ is a Hilbert space when equipped with the scalar product

$$(f,g)_{H^k(\Omega)} = \sum_{|\alpha|\le k} (D^\alpha f, D^\alpha g)_{L^2(\Omega)}.$$

Proof. Proving that $\|\cdot\|_{W^{k,p}(\Omega)}$ is a norm and that $(\cdot,\cdot)_{H^k(\Omega)}$ is a scalar product is an easy task that I leave as an exercise.

As for the fact that Sobolev spaces are Banach spaces, I will prove it for $k=1$, leaving the general case as an exercise.

Let (f_n) be a Cauchy sequence in $W^{1,p}(\Omega)$. Then the sequences (f_n), $(\partial_i f_n)$, $i = 1,\dots,d$, are Cauchy sequences in $L^p(\Omega)$, which is a Banach space. Let us denote the corresponding limits by f, g_i, $i=1,\dots,d$. In order to finish the proof it is enough to check that $g_i = \partial_i f$, $i=1,\dots,d$.

Using Hölder's inequality we get

$$\left|\int_\Omega f_n\partial_i\phi - \int_\Omega f\partial_i\phi\right| \le \int_\Omega |f_n - f||\partial_i\phi| \le \|f_n - f\|_p\|\partial_i\phi\|_{p'} \to 0,$$
$$\left|\int_\Omega \partial_i f_n\phi - \int_\Omega g_i\phi\right| \le \int_\Omega |\partial_i f_n - g_i||\phi| \le \|\partial_i f_n - g_i\|_p\|\phi\|_{p'} \to 0,$$

which means that $\int_\Omega f_n\partial_i\phi \to \int_\Omega f\partial_i\phi$, $\int_\Omega \partial_i f_n\phi \to \int_\Omega g_i\phi$ for all $\phi\in C_c^\infty(\Omega)$. Therefore, since

$$\int_\Omega f_n\partial_i\phi = -\int_\Omega \partial_i f_n\phi,$$

passing to the limit we get.

$$\int_\Omega f\partial_i\phi = -\int_\Omega g_i\phi,$$

that is, $\partial_i f = g_i$. □

We end with a result that is direct corollary of Proposition 1.3.5.

Proposition 1.3.8. *If $f\in W^{1,p}(\Omega)$ and $\psi\in C_c^\infty(\Omega)$, then $\psi f\in W^{1,p}(\Omega)$ and*

$$\partial_i(\psi f) = f\partial_i\psi + \psi\partial_i f.$$

1.4 Distributions

Although we will not make extensive use of the theory of distributions, it is useful to understand the interpretation of a weak derivative as a distributional derivative. In what follows $\Omega \subset \mathbb{R}^d$ is assumed to be an open set.

Definition 1.4.1. A sequence (φ_n) of functions $\varphi_n \in C_c^\infty(\Omega)$ converges to $\varphi \in C_c^\infty(\Omega)$ in the sense of test functions if:

(a) there exists $\Omega' \subset\subset \Omega$ such that $\operatorname{supp}\varphi_n \subset \Omega'$ for every $n \in \mathbb{N}$;

(b) $D^\alpha \varphi_n \to D^\alpha \varphi$ as $n \to \infty$ uniformly as $n \to \infty$ for every $\alpha \in \mathbb{N}_0^d$.

The topological vector space $\mathcal{D}(\Omega)$ consists of $C_c^\infty(\Omega)$ equipped with the topology that corresponds to convergence in the sense of test functions.

Note that since the supports of the functions φ_n are contained in the same compactly supported subset, the limit has compact support; and since the derivatives of all orders converge uniformly, the limit is smooth.

A linear functional on $\mathcal{D}(\Omega)$ is a linear map $T : \mathcal{D}(\Omega) \to \mathbb{R}$. We denote the value of T acting on a test function φ by $\langle T, \varphi\rangle$; thus, T is linear if $\langle T, \lambda\varphi + \mu\psi\rangle = \lambda\langle T, \varphi\rangle + \mu\langle T, \psi\rangle$ for all $\lambda, \mu \in \mathbb{R}$ and $\varphi, \psi \in \mathcal{D}(\Omega)$.

A functional T is continuous if $\varphi_n \to \varphi$ in the sense of test functions implies that $\langle T, \varphi_n\rangle \to \langle T, \varphi\rangle$ in $\mathbb{R}$.

Definition 1.4.2. A distribution on Ω is a continuous linear functional $T : \mathcal{D}(\Omega) \to \mathbb{R}$. A sequence (T_n) of distributions converges to a distribution T, and we will write $T_n \rightharpoonup T$, if $\langle T_n, \varphi\rangle \to \langle T, \varphi\rangle$ for every $\varphi \in \mathcal{D}(\Omega)$. The topological vector space $\mathcal{D}'(\Omega)$ consists of the distributions on Ω equipped with the topology corresponding to this notion of convergence.

Thus, the space of distributions is the topological dual of the space of test functions, hence the notation $\mathcal{D}'(\Omega)$.

Example. The delta-function supported at $a \in \Omega$ is the distribution $\delta_a : \mathcal{D}(\Omega) \to \mathbb{R}$ defined by evaluation of a test function at a: $\langle \delta_a, \varphi\rangle = \varphi(a)$. This functional is continuous since $\varphi_n \to \varphi$ in the sense of test functions implies, in particular, that $\varphi_n(a) \to \varphi(a)$.

Example. Any function $f \in L^1_{\mathrm{loc}}(\Omega)$ defines a distribution $T_f \in \mathcal{D}'(\Omega)$ by $\langle T_f, \varphi\rangle = \int_\Omega f\varphi$. The linear functional T_f is continuous, since if $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega)$, then $\sup_{\Omega'} |\varphi_n - \varphi| \to 0$ on a set $\Omega' \subset\subset \Omega$ that contains the supports of the φ_n, so

$$|\langle T, \varphi_n\rangle - \langle T, \varphi\rangle| = \Big| \int_{\Omega'} f(\varphi_n - \varphi)\Big| \le \sup_{\Omega'} |\varphi_n - \varphi| \int_{\Omega'} |f| \to 0.$$

Any distribution associated with a locally integrable function in this way is called a regular distribution. We typically regard the function f and the distribution T_f as equivalent.

One of the main advantages of distributions is that, in contrast to functions, every distribution is differentiable.

Definition 1.4.3. For $i \in \{1, \dots, d\}$, the partial derivative of a distribution $T \in \mathcal{D}'(\Omega)$ with respect to x_i is the distribution $\partial_i T \in \mathcal{D}'(\Omega)$ defined by $\langle \partial_i T, \varphi\rangle = -\langle T, \partial_i\varphi\rangle$ for all $\varphi \in \mathcal{D}(\Omega)$. For $\alpha \in \mathbb{N}_0^d$, the derivative $D^\alpha T \in \mathcal{D}'(\Omega)$ of order $|\alpha|$ is defined by $\langle D^\alpha T, \varphi\rangle = (-1)^{|\alpha|}\langle T, D^\alpha\varphi\rangle$ for all $\varphi \in \mathcal{D}(\Omega)$.

Note that if $T \in \mathcal{D}'(\Omega)$, then it follows from the linearity and continuity of the derivative $D^\alpha : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ on the space of test functions that $D^\alpha T$ is a continuous linear functional on $\mathcal{D}(\Omega)$. Thus, $D^\alpha T \in \mathcal{D}'(\Omega)$ for any $T \in \mathcal{D}'(\Omega)$. It also follows that the distributional derivative $D^\alpha : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ is linear and continuous on the space of distributions; in particular if $T_n \rightharpoonup T$, then $D^\alpha T_n \rightharpoonup D^\alpha T$.

Let $f \in L^1_{\text{loc}}(\Omega)$ be a locally integrable function and $T_f \in \mathcal{D}'(\Omega)$ the associated regular distribution. Suppose that the distributional derivative of T_f is a regular distribution $\partial_i T_f = T_{g_i}$, $g_i \in L^1_{\text{loc}}(\Omega)$. Then it follows from the definitions that $\int_\Omega f\partial_i\varphi = -\int_\Omega g_i\varphi$ for all $\varphi \in C_c^\infty(\Omega)$. Thus, Definition 1.3.1 may be restated as follows: A locally integrable function is weakly differentiable if its distributional derivative is regular, and its weak derivative is the locally integrable function corresponding to the distributional derivative.

Let us remark that the distributional derivative of a function exists even if it is not weakly differentiable.

Examples. (a) The distributional derivative of the step (Heaviside) function is the delta-function, which is not a regular distribution.

(b) The derivative of the delta-function δ_a supported at a, is the distribution $\partial_i\delta_a$ defined by $\langle\partial_i\delta_a, \varphi\rangle = -\partial_i\varphi(a)$. This distribution is not regular.

Differential equations are typically thought of as equations that relate functions. The use of weak derivatives and distribution theory leads to an alternative point of view of linear differential equations as linear functionals acting on test functions. Using this perspective, given suitable estimates, one can obtain simple and general existence results for weak solutions of linear PDEs by the use of the Hahn-Banach, Riesz representation, or other duality theorems for the existence of bounded linear functionals.

1.5 Approximation by smooth functions

Our goal is to approximate "bad" functions by good (smooth) ones. The idea is to average, giving increasing importance to nearby points and decreasing importance to the rest.

Definition 1.5.1. A *mollifying* family is any family of functions $\{\rho_\delta\}_{\delta>0}$ such that

$$\rho_\delta \in C_c^\infty(\mathbb{R}^d), \qquad \operatorname{supp}\rho_\delta \subset \overline{B(0,\delta)}, \qquad \int_{\mathbb{R}^d}\rho_\delta = 1, \qquad \rho_\delta \geq 0.$$

It is easy to generate a family of mollifiers starting from a function $\rho \in C_c^\infty(\mathbb{R}^d)$ such that $\operatorname{supp}\rho \subset \overline{B(0,1)}$, $\rho \geq 0$, $\rho \not\equiv 0$, for instance

$$\rho(x) = \begin{cases} \mathrm{e}^{1/(|x|^2-1)} & \text{si } |x| < 1, \\ 0 & \text{si } |x| > 1. \end{cases}$$

To this aim just define $\rho_\delta(x) = C\delta^{-d}\rho(x/\delta)$, with $C = 1/\int_{\mathbb{R}^d}\rho$.

Remark. The mollifying family with ρ as above is known as the *standard mollifier*.

Definition 1.5.2. Given $\delta > 0$, let $\Omega_\delta = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \delta\}$. The *mollification* of $f \in L^1_{\text{loc}}(\Omega)$ is the function $f^\delta : \Omega_\delta \to \mathbb{R}$ given by

$$f^\delta(x) = \underbrace{\int_\Omega \rho_\delta(x-y) f(y)\, dy}_{(\rho_\delta \star f)(x)} . \tag{1.3}$$

Notation. $f \star g$ is denoted as the *convolution* of f and g.

Remarks. (a) Since $B_\delta(x) \subset \Omega$ if $x \in \Omega_\delta$, we have room to average and f^δ is well defined in Ω_δ.

(b) $(\rho_\delta \star f)(x) = \displaystyle\int_{B_\delta(x)} \rho_\delta(x-y) f(y)\, dy = \int_{B_\delta(0)} \rho_\delta(y) f(x-y)\, dy = (f \star \rho_\delta)(x)$.

The next result shows that mollifications are smooth.

Theorem 1.5.1. *Let $f \in L^1_{\text{loc}}(\Omega)$, $\delta > 0$, and f^δ as in* (1.3)*, with $\{\rho_\delta\}_{\delta>0}$ a mollifying family. Then $f^\delta \in C^\infty(\Omega_\delta)$ and $D^\alpha(\rho_\delta \star f) = (D^\alpha \rho_\delta) \star f$ for all multi-index $\alpha \in \mathbb{N}_0^d$.*

Proof. Let $x \in \Omega_\delta$, $i \in \{1, ..., d\}$. Let $h \in \mathbb{R}$ be small enough, so that $x + he_i \in \Omega_\delta$. There exists $\Omega' \subset\subset \Omega$ such that

$$\frac{\rho_\delta(x + he_i - y) - \rho_\delta(x-y)}{h} - \partial_i \rho_\delta(x-y) = 0 \quad \text{if } y \notin \Omega'.$$

On the other hand, using Taylor's expansion we get

$$\left| \frac{\rho_\delta(x + he_i - y) - \rho_\delta(x-y)}{h} - \partial_i \rho_\delta(x-y) \right| \le C|h| \quad \text{if } y \in \Omega'.$$

We conclude that

$$\begin{aligned} &\left| \frac{f^\delta(x + he_i) - f^\delta(x)}{h} - (\partial_i \rho_\delta \star f)(x) \right| \\ &\qquad = \left| \int_\Omega \left(\frac{\rho_\delta(x + he_i - y) - \rho_\delta(x-y)}{h} - \partial_i \rho_\delta(x-y) \right) f(y)\, dy \right| \le C|h| \|f\|_{L^1(\Omega')}, \end{aligned}$$

which implies that $\partial_i f^\delta(x)$ exists and is equal to $(\partial_i \rho_\delta \star f)(x)$.

The proof for higher derivatives proceeds by induction. □

The mollification of a function resembles the function, as we show next.

Theorem 1.5.2 (local approximation by smooth functions). (a) *If $f \in C(\Omega)$ then f^δ converges to f uniformly in compact subsets of Ω as $\delta \to 0^+$.*

(b) *If $f \in L^p(\Omega)$, $1 \le p < \infty$, then $\|f^\delta\|_{L^p(\Omega_\delta)} \le \|f\|_{L^p(\Omega)}$ and $f^\delta \to f$ in $L^p_{\text{loc}}(\Omega)$ as $\delta \to 0^+$.*

(c) *If $f \in W^{1,p}(\Omega)$, $1 \le p < \infty$, then $\partial_i f^\delta = \rho_\delta \star \partial_i f$, and f^δ converges to f in $W^{1,p}_{\text{loc}}(\Omega)$ as $\delta \to 0^+$.*

Proof. (a) Let $K \subset \Omega$ be a compact set. Let $\delta_1 > 0$ be such that $K \subset \Omega_{\delta_1}$. Given $\varepsilon > 0$, due to the uniform continuity of f on compact sets, there is a value $\delta \in (0, \delta_1]$ such that for all $x \in K$ and $y \in B_\delta(0)$ we have that $|f(x-y) - f(x)| < \varepsilon$. Therefore,

$$|f^\delta(x) - f(x)| = \left| \int_{B_\delta(0)} \rho_\delta(y)(f(x-y) - f(x))\, dy \right| \leq \int_{B_\delta(0)} \rho_\delta(y)|f(x-y) - f(x)|\, dy < \varepsilon$$

for all $x \in K$.

(b) Let $x \in \Omega_\delta$. Then, applying Hölder's inequality,

$$\begin{aligned} |f^\delta(x)| &= \left| \int_{B_\delta(x)} \rho_\delta(x-y) f(y) dy \right| \leq \int_{B_\delta(x)} \rho_\delta(x-y)^{1/p'} \rho_\delta(x-y)^{1/p} |f(y)|\, dy \\ &\leq \Big(\int_{B_\delta(x)} \rho_\delta(x-y)\, dy \Big)^{1/p'} \Big(\int_{B_\delta(x)} \rho_\delta(x-y)|f(y)|^p\, dy \Big)^{1/p}. \end{aligned}$$

Therefore, using Fubini's theorem

$$\begin{aligned} \int_{\Omega_\delta} |f^\delta(x)|^p\, dx &\leq \int_{\Omega_\delta} \left(\int_{B_\delta(x)} \rho_\delta(x-y)|f(y)|^p\, dy \right) dx \\ &\leq \int_\Omega |f(y)|^p \left(\int_{\mathbb{R}^d} \rho_\delta(x-y)\, dx \right) dy = \int_\Omega |f(y)|^p\, dy = \|f\|^p_{L^p(\Omega)}. \end{aligned}$$

Let $\Omega' \subset\subset \Omega$ and $\delta > 0$ such that $\Omega' \subset \Omega_\delta$. There exists a function $f_1 \in C(\Omega)$ such that $\|f - f_1\|_{L^p(\Omega)} \leq \varepsilon$. Hence, using (a) and the estimate that we have just obtained,

$$\begin{aligned} \|f^\delta - f\|_{L^p(\Omega')} &\leq \|f^\delta - f_1^\delta\|_{L^p(\Omega')} + \|f_1^\delta - f_1\|_{L^p(\Omega')} + \|f_1 - f\|_{L^p(\Omega')} \\ &\leq \|(f - f_1)^\delta\|_{L^p(\Omega_\delta)} + \varepsilon + \|f_1 - f\|_{L^p(\Omega_\delta)} \leq 2\|f - f_1\|_{L^p(\Omega)} + \varepsilon \leq \varepsilon' \end{aligned}$$

if δ is small enough.

(c) Let $\Omega' \subset\subset \Omega$ and $\delta > 0$ such that $\Omega' \subset \Omega_\delta$. Let $x \in \Omega'$. Denoting $g^x(y) = \rho_\delta(x-y)$, we observe that $\partial_i g^x(y) = -\partial_i \rho_\delta(x-y)$. Therefore, integrating by parts,

$$\begin{aligned} \partial_i f^\delta(x) &= \int_\Omega \partial_i \rho_\delta(x-y) f(y)\, dy = -\int_\Omega \partial_i g^x(y) f(y)\, dy = \int_\Omega g^x(y) \partial_i f(y)\, dy \\ &= \int_\Omega \rho_\delta(x-y) \partial_i f(y)\, dy = \rho_\delta \star \partial_i f(x). \end{aligned}$$

Now, thanks to (b) we know that $f^\delta \to f$ in $L^p_{\text{loc}}(\Omega)$. On the other hand, applying this result to $\partial_i f$, which belongs to $L^p(\Omega)$, since $f \in W^{1,p}(\Omega)$, we get that

$$\partial_i f^\delta = (\partial_i f)^\delta \to \partial_i f \quad \text{in } L^p_{\text{loc}}(\Omega).$$

$\square$

Partitions of unity allow us to piece together global results from local results.

Theorem 1.5.3 (partitions of unity). *Let $\mathcal{O}$ be a family of open subsets of $\mathbb{R}^d$. There is a collection Φ of C^∞ functions φ defined in $\Omega = \cup_{\omega \in \mathcal{O}} \omega$ with the following properties:*

(a) $0 \le \varphi \le 1$ *in* Ω;

(b) *for each compact subset* K *of* Ω, *the set* $\{\varphi \in \Phi : \operatorname{supp} \varphi \cap K \neq \emptyset\}$ *is finite;*

(c) $\sum_{\varphi \in \Phi} \varphi = 1$ *on* Ω *(by (b), this sum is finite in every compact subset of* Ω*);*

(d) *for each* $\varphi \in \Phi$ *there is an open set* $U \in \mathcal{O}$ *such that* $\operatorname{supp} \varphi \subset U$.

Proof. You can find a proof in [8, Page 63]. □

Definition 1.5.3. A collection Φ satisfying (a)–(c) is called a C^∞ *partition of unity* for Ω. If it also satisfies (d), it is said to be *subordinate to the cover* $\mathcal{O}$.

Theorem 1.5.4 (global approximation by smooth functions, Meyers-Serrin). *Let* $\Omega \subset \mathbb{R}^d$ *be an open set and let* $1 \le p < \infty$. *Then* $C^\infty(\Omega) \cap W^{1,p}(\Omega)$ *is dense in* $W^{1,p}(\Omega)$.

Proof. I am taking the proof of this nice theorem from the book [6].

Take a family $\{\Omega_i\}_{i=1}^\infty$ of open subsets of Ω such that $\Omega_i \subset\subset \Omega_{i+1}$ and

$$\Omega = \cup_{i=1}^\infty \Omega_i.$$

For each $i \in \mathbb{N}$, let $V_i := \Omega_{i+1} \setminus \overline{\Omega_{i-1}}$, where $\Omega_0 = \emptyset$. We consider a smooth partition of unity $\mathcal{F}$ subordinated to the open cover $\{V_i\}_{i=1}^\infty$. For each $i \in \mathbb{N}$ let ψ_i be the sum of all the finitely many $\psi \in \mathcal{F}$ such that $\operatorname{supp} \psi \subset V_i$ and such that they have not already been selected at previous steps $j < i$. Then $\psi_i \in C_c^\infty(V_i)$ and $\sum_{i=1}^\infty \psi_i = 1$ in Ω. Hence, $\psi_i u \in W^{1,p}(\Omega)$ and $\operatorname{supp}(\psi_i u) \subset V_i$.

Let $\varepsilon > 0$. For all $i \in \mathbb{N}$ there is a value $\delta_i > 0$ such that the mollification $u_i = \rho_{\delta_i} \star (\psi_i u)$ satisfies that $\operatorname{supp} u_i \subset V_i$ and

$$\|u_i - \psi_i u\|_{W^{1,p}(V_i)} \le \frac{\varepsilon}{2^{i+1}}.$$

We define $v := \sum_{i=0}^\infty u_i \in C^\infty(\Omega)$ (in each $\Omega' \subset\subset \Omega$ the sum has only a finite number of terms different from 0). Let us also note that $u = \sum_{i=1}^\infty (\psi_i u)$.

For $x \in \Omega_\ell$,

$$u(x) = \sum_{i=1}^{\ell} (\psi_i u)(x), \qquad v(x) = \sum_{i=0}^{\ell} u_i(x)$$

Therefore,

$$\|v - u\|_{W^{1,p}(\Omega_\ell)} = \left\| \sum_{i=0}^{\ell} u_i - \sum_{i=0}^{\ell} \psi_i u \right\|_{W^{1,p}(\Omega_\ell)} \le \sum_{i=0}^{\ell} \|u_i - \psi_i u\|_{W^{1,p}(V_i)} \le \sum_{i=0}^{\infty} \frac{\varepsilon}{2^{i+1}} = \varepsilon.$$

Letting $\ell \to \infty$, it follows from the monotone convergence theorem that $\|v - u\|_{W^{1,p}(\Omega)} \le \varepsilon$. This also implies that $u - v$ (and, in turn, v) belongs to the space $W^{1,p}(\Omega)$. □

Corollary 1.5.5 (density of $C_c^\infty(\mathbb{R}^d)$ in $W^{1,p}(\mathbb{R}^d)$). *Let* $1 \le p < \infty$. *Then* $C_c^\infty(\mathbb{R}^d)$ *is dense in* $W^{1,p}(\mathbb{R}^d)$.

Proof. Given $u \in W^{1,p}(\mathbb{R}^d)$ and $\varepsilon > 0$, we consider a function $v \in C^\infty(\mathbb{R}^d) \cap W^{1,p}(\mathbb{R}^d)$ such that $\|u - v\|_{W^{1,p}(\mathbb{R}^d)} \le \varepsilon/2$. The existence of such function is guaranteed by Meyers-Serrin theorem.

Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \le \varphi \le 1$, $\varphi(x) = 1$ if $|x| \le 1$, $\varphi(x) = 0$ if $|x| \ge 2$. We define $\varphi_R(x) = \varphi(x/R)$ (which is known as a *cut-off* function). We claim that $\|\varphi_R v - v\|_{W^{1,p}(\mathbb{R}^d)} \le \varepsilon/2$ if R is large enough, from where the result follows immediately, since $\varphi_R v \in C_c^\infty(\mathbb{R}^d)$.

It is easily checked that $\lim_{R\to\infty} |\varphi_R v - v| = 0$ a.e. Hence, since

$$|\varphi_R v - v|^p = |1 - \varphi_R|^p |v|^p \le |v|^p \in L^1(\mathbb{R}^d),$$

we get that $\|\varphi_R v - v\|_p \to 0$ as $R \to \infty$. On the other hand,

$$|D(\varphi_R v) - Dv| = |(\varphi_R - 1)Dv + vD\varphi_R| \le |1 - \varphi_R||Dv| + \frac{\|D\varphi\|_\infty}{R}|v|.$$

Thus, $\lim_{R\to\infty} |D(\varphi_R v) - Dv| = 0$ a.e. Hence, since

$$|D(\varphi_R v) - Dv|^p \le C(|v| + |Dv|)^p \in L^1(\mathbb{R}^d),$$

we get that $\|D(\varphi_R v) - Dv\|_p \to 0$ as $R \to \infty$.

Summarizing, $\|\varphi_R v - v\|_{W^{1,p}(\mathbb{R}^d)} \to 0$ as $R \to \infty$, and the result follows. □

Let now $\Omega \subset \mathbb{R}^d$ be a bounded open set. Can we approximate a function in $W^{1,p}(\Omega)$ by a function in $C^\infty(\overline{\Omega})$, instead of only by functions in $C^\infty(\Omega) \cap W^{1,p}(\Omega)$? This is indeed the case if $\partial\Omega$ is smooth.

Definition 1.5.4. Let $\Omega \subset \mathbb{R}^d$ be a bounded open set. We say that its boundary, $\partial\Omega$, is C^1 if for all $x_0 \in \partial\Omega$ there is a radius $r > 0$ and a C^1 function $\gamma : \mathbb{R}^{d-1} \to \mathbb{R}$ such that (relabeling and reorienting the axes if needed),

$$\Omega \cap B_r(x_0) = \{x \in B_r(x_0) : x_d > \gamma(x_1, \cdots, x_{d-1})\}.$$

Theorem 1.5.6 (global approximation by functions smooth up to the boundary). *Let $\Omega \subset \mathbb{R}^d$ be a bounded open set with C^1 boundary $\partial\Omega$ and let $1 \le p < \infty$. Then $C^\infty(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$.*

Proof. Let $x_0 \in \partial\Omega$. Let $r > 0$ and a C^1 function $\gamma : \mathbb{R}^{d-1} \to \mathbb{R}$ such that (relabeling and reorienting the axes if needed),

$$\Omega \cap B_r(x_0) = \{x \in B_r(x_0) : x_d > \gamma(x_1, \cdots, x_{d-1})\}.$$

Set $V := \Omega \cap B_{\frac{r}{2}}(x_0)$. Let $\lambda > 1$. Given $x \in V$ and $\delta > 0$, we define the shifted point

$$x^\delta := x + \lambda\delta e_d.$$

Observe that the ball $B_\delta(x^\delta)$ lies in $\Omega \cap B_r(x_0)$ for all $x \in V$ and all small $\delta > 0$.

For $x \in V$ we define $u_\delta(x) := u(x^\delta)$. This is the function u translated a distance $\lambda\delta$ in the e_d direction. Next write $v_\delta = \rho_\delta \star u_\delta$. The idea is that we have moved up enough so that there is room to mollify within Ω. Clearly $v_\delta \in C^\infty(\overline{V})$.

We now claim that $v_\delta \to u$ in $W^{1,p}(V)$. To confirm this, observe that

$$\begin{aligned}
&\|v_\delta - u\|_{L^p(V)} \le \|v_\delta - u_\delta\|_{L^p(V)} + \|u_\delta - u\|_{L^p(V)},\\
&\|\partial_i v_\delta - \partial_i u\|_{L^p(V)} \le \|\partial_i v_\delta - \partial_i u_\delta\|_{L^p(V)} + \|\partial_i u_\delta - \partial_i u\|_{L^p(V)}, \quad i = 1, \dots, d.
\end{aligned}$$

Reasoning like in the proof of Theorem 1.5.2-(c) we can check that the first term on the right-hand side of each line goes to 0 with δ. The second term also vanishes in the limit, since translations are continuous in the L^p-norm (see problem 4 in worksheet 2), and the claim is proved.

Choose $\varepsilon > 0$. Since $\partial\Omega$ is compact, we can find finitely many points $x_j^0 \in \partial\Omega$, radii $r_j > 0$, corresponding sets $V_j = \Omega \cap \mathcal{O}_j$ where $\mathcal{O}_j := B_{\frac{r_j}{2}}(x_j^0)$, and functions $v_j \in C^\infty(\overline{V_j})$, $j = 1, \dots, M$, such that $\partial\Omega \subset \cup_{j=1}^M \mathcal{O}_j$, and

$$\|v_j - u\|_{W^{1,p}(V_j)} \le \varepsilon.$$

Take an open set $\mathcal{O}_0 \subset\subset \Omega$ such that $\Omega \subset \cup_{j=0}^M \mathcal{O}_j$ and select, using Theorem 1.5.2, a function $v_0 \in C^\infty(\overline{\mathcal{O}_0})$ satisfying

$$\|v_0 - u\|_{W^{1,p}(\mathcal{O}_0)} \le \varepsilon.$$

We consider a smooth partition of unity $\mathcal{F}$ subordinated to the open cover $\{\mathcal{O}_j\}_{j=0}^M$. Let ψ_j be the sum of all the finitely many $\psi \in \mathcal{F}$ such that $\operatorname{supp}\psi \in \mathcal{O}_j$, $\operatorname{supp}\psi \cap \overline{\Omega} \ne \emptyset$, and such that they have not already been selected at previous steps $k < j$. Then $\psi_j \in C_{\mathrm{c}}^\infty(V_j)$ and $\sum_{j=1}^M \psi_j = 1$ in $\overline{\Omega}$. We define $v := \sum_{j=0}^M \psi_j v_j$. Then clearly $v \in C^\infty(\overline{\Omega})$. In addition, since $u = \sum_{j=0}^M \psi_j u$ in Ω, we finally obtain

$$\|v - u\|_{W^{1,p}(\Omega)} \le \sum_{j=1}^M \|\psi_j(v_j - u)\|_{W^{1,p}(\mathcal{O}_j\cap\Omega)}.$$

But $\|\psi_j(v_j - u)\|_{L^p(\mathcal{O}_j\cap\Omega)} \le C\|v_j - u\|_{L^p(\mathcal{O}_j\cap\Omega)}$, and

$$\begin{aligned}
\|\partial_i(\psi_j(v_j - u))\|_{L^p(\mathcal{O}_j\cap\Omega)} &\le \|(v_j - u)\partial_i\psi_j\|_{L^p(\mathcal{O}_j\cap\Omega)} + \|\psi_j\partial_i(v_j - u)\|_{L^p(\mathcal{O}_j\cap\Omega)}\\
&\le C(\|v_j - u\|_{L^p(\mathcal{O}_j\cap\Omega)} + \|\partial_i(v_j - u)\|_{L^p(\mathcal{O}_j\cap\Omega)},
\end{aligned}$$

and we conclude that

$$\|v - u\|_{W^{1,p}(\Omega)} \le C\sum_{j=0}^M \|v_j - u\|_{W^{1,p}(\mathcal{O}_j\cap\Omega)} \le C(M+1)\varepsilon.$$

□

1.6 Extensions

Our next goal is to extend functions in the Sobolev space $W^{1,p}(\Omega)$ to become functions in the Sobolev space $W^{1,p}(\mathbb{R}^d)$.

Theorem 1.6.1 (Extension theorem). *Let $\Omega \subset \mathbb{R}^d$ be an open set with C^1 boundary and $1 \le p \le \infty$. Let V be a bounded open set such that $\Omega \subset\subset V$. Then there exists a bounded linear operator*

$$E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^d)$$

such that for all $u \in W^{1,p}(\Omega)$:

(a) *$Eu = u$ a.e. in Ω;*

(b) *Eu has support within V;*

(c) *$\|Eu\|_{W^{1,p}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\Omega)}$, the constant C depending only on p, Ω and V.*

We call Eu an extension of u to $\mathbb{R}^d$.

Proof. I will give a proof that uses global approximations by functions that are smooth up to the boundary, Theorem 1.5.6, which hence is only valid for $1 \le p < \infty$. I took it from [4]. You can find a (not so) different proof that is also valid for $p = \infty$ in [2, Theorem 9.7].

Fix $x^0 \in \partial\Omega$ and suppose first that $\partial\Omega$ is flat near x^0, lying in the plane $\{x_d = 0\}$. Then we may assume that there exists an open ball $B = B_r(x^0)$ for some radius $r > 0$, such that

$$B^+ := B \cap \{x_d \ge 0\} \subset \overline{\Omega}, \qquad B^- := B \cap \{x_d \le 0\} \subset \mathbb{R}^d \setminus \Omega.$$

Let us also assume temporarily that $u \in C^1(\overline{\Omega})$. Then we define

$$\bar{u}(x) := \begin{cases} u(x) & \text{if } x \in B^+, \\ -3u(x_1, \dots, x_{d-1}, -x_d) + 4u(x_1, \dots, x_{d-1}, -x_d/2) & \text{if } x \in B^-. \end{cases}$$

This is called a *higher-order reflection* of u from B^+ to B^-.

We claim that $\bar{u} \in C^1(B)$. We only have to check what happens at the plane $\{x_d = 0\}$, since smoothness in the rest of the domain under consideration is obvious.

Let us write $u^- := \overline{u}\big|_{B^-}$, $u^+ := \overline{u}\big|_{B^+}$. A straightforward computation shows that

$$\partial_d u^-(x) = 3\partial_d u(x_1, \dots, x_{d-1}, -x_d) - 2\partial_d u(x_1, \dots, x_{d-1}, -x_d/2),$$

and hence

$$\partial_d u^-\big|_{x_d=0} = \partial_d u^+\big|_{x_d=0}.$$

On the other hand, it is trivial to check that $u^- = u^+$ on $\{x_d = 0\}$. This implies moreover that

$$\partial_i u^-\big|_{x_d=0} = \partial_i u^+\big|_{x_d=0}, \quad i = 1, \dots, d-1,$$

and the claim is proved.

The above computations yield easily (check it!) that

$$\|\bar{u}\|_{W^{1,p}(B)} \le C\|u\|_{W^{1,p}(B^+)}$$

for some constant C that does not depend on u.

If $\partial\Omega$ is not flat near x^0, we *straighten the boundary* near x^0. To be more specific, we set

$$y_i = x_i =: \Phi_i(x), \quad i = 1, \dots, d-1, \qquad y_d = x_d - \gamma(x_1, \dots, x_{d-1}) =: \Phi_d(x),$$

where γ is the C^1 function describing the boundary in a neighbourhood of x^0, and write $y = \Phi(x)$. Similarly, we set

$$x_i = y_i =: \Psi_i(y), \quad i = 1, \dots, d-1, \qquad x_d = y_d + \gamma(y_1, \dots, y_{d-1}) =: \Psi_d(y),$$

and write $x = \Psi(y)$. Then $\Phi = \Psi^{-1}$, and the mapping $x \to \Phi(x) = y$ *straightens out* $\partial\Omega$ near x^0. Observe that $\det D\Phi = \det D\Psi = 1$.

Let us define $u'(y) := u(\Psi(y))$. Choose a small ball B (in the y variable) as before. As shown above, we can extend u' from B^+ to a function $\bar{u}'$ defined on all of B, such that $\bar{u}'$ is C^1 and we have the estimate

$$\|\bar{u}'\|_{W^{1,p}(B)} \le C\|u'\|_{W^{1,p}(B^+)}.$$

Let $W := \Psi(B)$. Going back to the x variables, we obtain an extension $\bar{u}$ of u to W with

$$\|\bar{u}\|_{W^{1,p}(W)} \le C\|u\|_{W^{1,p}(\Omega)}.$$

To conclude the proof in the case of functions that are smooth up to the boundary we use partitions of unity. Since $\partial\Omega$ is compact, there exist finitely many points $x_j^0 \in \partial\Omega$, open sets W_j, and extensions $\bar{u}_j$ of u to W_j, $j = 1, \dots M$, as above, such that $\partial\Omega \subset \cup_{j=1}^M W_j$. Take $W_0 \subset\subset \Omega$ such that $\Omega \subset \cup_{j=0}^M W_j$ and $\bar{u}_0 = 0$. Note that we may choose the open sets W_j so that $\cup_{j=0}^M W_j \subset\subset V$. Arguing as in the proof of Theorem 1.5.6, it is easy to check that there is a family of C^∞ functions (a partition of unity) $\{\psi_j\}_{j=0}^M$ such that $0 \le \psi_j \le 1$, $\operatorname{supp} \psi_j \subset W_j$, $\sum_{j=0}^M \psi_j = 1$ in $\overline{\Omega}$. We define $\bar{u} := \sum_{j=0}^M \psi_j \bar{u}_j$. Then,

$$\|\bar{u}\|_{W^{1,p}(\mathbb{R}^d)} \le \sum_{j=0}^M \|\psi_j \bar{u}_j\|_{W^{1,p}(W_j)} \le C\|u\|_{W^{1,p}(\Omega)} \tag{1.4}$$

for some constant C that does not depend on u. We write $Eu := \bar{u}$ and observe that the mapping $u \to Eu$ is linear.

We finally get rid of the assumption $u \in C^1(\Omega)$ using the global approximations smooth up to the boundary provided by Theorem 1.5.6. It is here that we are using that p is finite. Let $u \in W^{1,p}(\Omega)$. We choose $(u_k) \subset C^\infty(\overline{\Omega})$ converging to u in $W^{1,p}(\Omega)$. Using (1.4) and the linearity of E we get

$$\|Eu_k - Eu_j\|_{W^{1,p}(\mathbb{R}^d)} \le C\|u_k - u_l\|_{W^{1,p}(\Omega)}.$$

Therefore, (Eu_k) is a Cauchy sequence in $W^{1,p}(\mathbb{R}^d)$ and so converges to $\bar{u} =: Eu$. This extension, which does not depend on the particular choice of the approximating sequence (check it!), satisfies the conclusions of the theorem. □

1.7 Traces

We discuss now the possibility of assigning "boundary values" along $\partial\Omega$ to a function $u \in W^{1,p}(\Omega)$ assuming that $\partial\Omega$ is C^1. If $u \in C(\overline{\Omega})$ then u has values on $\partial\Omega$ in the usual sense. The problem is that a typical function $u \in W^{1,p}(\Omega)$ is not in general continuous and, even worse, is only defined a.e. in Ω. Since $\partial\Omega$ has d-dimensional Lebesgue measure zero, there is no direct meaning we can give to the expression "u restricted to $\partial\Omega$". The notion of a *trace operator* solves this problem.

Theorem 1.7.1 (Trace theorem). *Let $1 \le p < \infty$, and assume that $\Omega \subset \mathbb{R}^d$ is a bounded open set with smooth boundary. Then there exists a bounded linear operator*

$$T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that

(a) $Tu = u\big|_{\partial\Omega}$ *if* $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$,

(b) $\|Tu\|_{L^p(\partial\Omega)} \le C\|u\|_{W^{1,p}(\Omega)}$ *for all* $u \in W^{1,p}(\Omega)$,

where C is a constant depending only on p and Ω. We call Tu the trace of u on $\partial\Omega$.

Proof. We assume first that $u \in C^1(\overline{\Omega})$, the general case following by approximation.

We also assume that the boundary is flat near $x^0 \in \partial\Omega$, lying in the plane $\{x_d = 0\}$. We choose an open ball B as in the proof of the previous theorem, and we denote by $\hat{B}$ the concentric ball with radius $r/2$.

Select $\zeta \in C_c^\infty(B)$ with $\zeta \ge 0$ in B, $\zeta \equiv 1$ on $\hat{B}$ (such a function is known as a cut-off function). Denote by Γ the portion of $\partial\Omega$ within $\hat{B}$. Set $x' = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1} = \{x_d = 0\}$. Then, applying the Fundamental Theorem of Calculus, and also Young's inequality[1] if $p > 1$,

$$\begin{aligned} \int_\Gamma |u|^p\,dx' &\le \int_{\{x_d=0\}} \zeta|u|^p\,dx' = -\int_{B^+} \partial_d(\zeta|u|^p)\,dx \\ &= -\int_{B^+} \Big(|u|^p\partial_d\zeta + \zeta p|u|^{p-1}(\text{sign } u)\partial_d u\Big)\,dx \le C\int_{B^+} \Big(|u|^p + |\partial_d u|^p\Big)\,dx, \end{aligned}$$

which yields

$$\|u\|_{L^p(\Gamma)} \le C\|u\|_{W^{1,p}(\Omega)}. \tag{1.5}$$

In order to compute the weak derivative of $|u|^p$ we have used problems 3 and 4 in worksheet 3.

If $\partial\Omega$ is not flat around $x^0 \in \partial\Omega$, we straighten out the boundary, do the above computation, and then undo the change of variables, in the same fashion as in the proof of the previous theorem, to obtain again (1.5), where now Γ is some open subset of $\partial\Omega$ containing x^0.

Since $\partial\Omega$ is compact, there exist finitely many point $x_j^0 \in \partial\Omega$ and open subsets $\Gamma_j \subset \partial\Omega$ containing x_j^0, $j = 1, \dots, M$, such that $\partial\Omega = \cup_{j=1}^M \Gamma_j$ and

$$\|u\|_{L^p(\Gamma_j)} \le C\|u\|_{W^{1,p}(\Omega)}, \quad j = 1, \dots, M.$$

[1] $ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}$, $a, b > 0$, $1 < p, p' < \infty$, $\frac{1}{p} + \frac{1}{p'} = 1$.

Consequently, if we write $Tu := u\big|_{\partial\Omega}$, then

$$\|Tu\|_{L^p(\partial\Omega)} \le C\|u\|_{W^{1,p}(\Omega)} \tag{1.6}$$

for some appropiate constant C which does not depend on u.

Let us now consider $u \in W^{1,p}(\Omega)$ not necessarily in $C^1(\overline{\Omega})$. There exist functions $u_k \in C^\infty(\overline{\Omega})$ converging to u in $W^{1,p}(\Omega)$. According to (1.6) we have

$$\|Tu_k - Tu_j\|_{L^p(\partial\Omega)} \le C\|u_k - u_j\|_{W^{1,p}(\Omega)}, \tag{1.7}$$

so that $(Tu_k)_{k=1}^\infty$ is a Cauchy sequence in $L^p(\partial\Omega)$. We define

$$Tu := \lim_{k\to\infty} Tu_k,$$

the limit taken in $L^p(\partial\Omega)$. Acccording to (1.7) this definition does not depend on the particular choice of smooth functions approximating u.

Finally, if $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$, we note that the functions of the approximation of u constructed in Theorem 1.5.6 converge uniformly to u on $\overline{\Omega}$. Hence $Tu = u\big|_{\partial\Omega}$. □

Definition 1.7.1. The space $W_0^{1,p}(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $W^{1,p}(\Omega)$. It is customary to write $H_0^1(\Omega)$ to denote $W_0^{1,2}(\Omega)$.

Remark. Corollary 1.5.5 implies that $W_0^{1,p}(\mathbb{R}^d) = W^{1,p}(\mathbb{R}^d)$ if $1 \le p < \infty$.

Theorem 1.7.2 (Trace-zero functions in $W^{1,p}(\Omega)$)**.** *Let $1 \le p < \infty$. Assume that Ω is a bounded open subset of $\mathbb{R}^d$ with smooth boundary. Let $u \in W^{1,p}(\Omega)$. Then,*

$$u \in W_0^{1,p}(\Omega) \text{ if and only if } Tu = 0 \text{ on } \partial\Omega.$$

Proof. We prove that $u \in W_0^{1,p}(\Omega) \Rightarrow Tu = 0$ on $\partial\Omega$. A proof of the other implication can be found for example in [4, Section 5.5].

If $u \in W_0^{1,p}(\Omega)$, by definition, there exist functions $u_k \in C_c^\infty(\Omega)$ such that u_k converges to u in $W^{1,p}(\Omega)$. Since $Tu_k = 0$ on $\partial\Omega$ for all k, and $T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is a bounded linear operator, hence continuous, we deduce that $Tu = 0$ on $\partial\Omega$. □

Thus, asking a function to belong to some space $W_0^{1,p}(\Omega)$ is a reasonable (weak) way to ask it to be zero at $\partial\Omega$.

Definition 1.7.2. The space $W_0^{k,p}(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$. It is customary to write $H_0^k(\Omega)$ to denote $W_0^{k,2}(\Omega)$.

Thanks to Theorem 1.7.2, we know that $W_0^{k,p}(\Omega)$ consists of those functions in $W^{k,p}(\Omega)$ such that $D^\alpha u = 0$ on $\partial\Omega$ for all $|\alpha| \le k-1$ in the sense of traces.

1.8 Sobolev embeddings

Definition 1.8.1. We say that a Banach space X is *continuously* embedded, or embedded for short, in a Banach space Y if there is a one-to-one, bounded linear map $i : X \to Y$.

We often think of i as identifying elements of the smaller space X with elements of the larger space Y; if X is a subset of Y then i is the inclusion map. The boundedness of i means that ther is a constant C such that $\|ix\|_Y \le C\|x\|_X$ for all $x \in X$, so the weaker Y-norm of ix is controlled by the stronger X-norm of X.

We write an embedding as $X \hookrightarrow Y$, or as $X \subset Y$ when the boundedness is understood.

We wonder whether $W^{1,p}(\mathbb{R}^d)$ is continuously embedded in some other spaces. The answer will be "yes", but in wich other spaces depends on the relation between p and d.

1.8.1 Gagliardo-Nirenberg-Sobolev inequality.

Let $1 \le p < d$. We start by asking ourselves whether we can establish an estimate of the form

$$\|u\|_{L^q(\mathbb{R}^d)} \le C\|Du\|_{L^p(\mathbb{R}^d)} \text{ for all } u \in C_c^\infty(\mathbb{R}^d) \tag{1.8}$$

for some constants $C > 0$ and $q \in [1, \infty)$. The point is that the constants C and q should not depend on u. Here Du denotes the gradient of u and $\|Du\|_{L^p(\mathbb{R}^d)}$ stands for $\||Du|\|_{L^p(\mathbb{R}^d)}$.

Let us first show that such an inequality can only hold for a very specific value of q. For this, choose a function $u \in C_c^\infty(\mathbb{R}^d)$, $u \not\equiv 0$, and define for $\lambda > 0$ the rescaled function

$$u_\lambda(x) := u(\lambda x), \quad x \in \mathbb{R}^d.$$

Applying (1.8) to u_λ, we find

$$\|u_\lambda\|_{L^q(\mathbb{R}^d)} \le C\|Du_\lambda\|_{L^p(\mathbb{R}^d)} \tag{1.9}$$

Now,

$$\int_{\mathbb{R}^d} |u_\lambda|^q = \int_{\mathbb{R}^d} |u(\lambda x)|^q\, dx = \frac{1}{\lambda^d}\int_{\mathbb{R}^d} |u(y)|^q\, dy,$$
$$\int_{\mathbb{R}^d} |Du_\lambda|^p = \lambda^p\int_{\mathbb{R}^d} |Du(\lambda x)|^p\, dx = \frac{\lambda^p}{\lambda^d}\int_{\mathbb{R}^d} |Du(y)|^p\, dy.$$

Inserting these equations into (1.9), we discover that

$$\frac{1}{\lambda^{d/q}}\|u\|_{L^q(\mathbb{R}^d)} \le C\frac{\lambda}{\lambda^{d/p}}\|Du\|_{L^p(\mathbb{R}^d)},$$

and so

$$\|u\|_{L^q(\mathbb{R}^d)} \le C\lambda^{1-\frac{d}{p}+\frac{d}{q}}\|Du\|_{L^p(\mathbb{R}^d)}.$$

Therefore, if $1 - \frac{d}{p} + \frac{d}{q} \ne 0$ we obtain a contradiction upon sending λ to either 0 or ∞. Thus, the inequality (1.8) cannot hold unless $1 - \frac{d}{p} + \frac{d}{q} = 0$, that is, unless $q = \frac{dp}{d-p}$.

Definition 1.8.2. The *Sobolev conjugate exponent* of $p \in [1, d)$ is

$$p^* := \frac{dp}{d-p}.$$

Note that $\dfrac{1}{p^*} = \dfrac{1}{p} - \dfrac{1}{d}$, $p^* > p$.

Theorem 1.8.1 (Gagliardo-Nirenberg-Sobolev inequality). *Let $p \in [1, d)$. There exists a constant C, depending only on p and d, such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^d)} \leq C\|Du\|_{L^p(\mathbb{R}^d)} \text{ for all } u \in C^1_{\rm c}(\mathbb{R}^d) \tag{1.10}$$

Proof. Step 1. We first prove the estimate for the case $p = 1$. The general case will follow from it.

Since u has compact support, for each $i \in \{1, \dots, d\}$ and $x \in \mathbb{R}^d$ we have

$$u(x) = \int_{-\infty}^{x_i} \partial_i u(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_d)\, dy_i,$$

and so

$$|u(x)| \leq \int_{-\infty}^{\infty} |Du(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_d)|\, dy_i, \quad i \in \{1, \dots, d\}.$$

Consequently

$$|u(x)|^{\frac{d}{d-1}} \leq \prod_{i=1}^{d} \left(\int_{-\infty}^{\infty} |Du(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_d)|\, dy_i \right)^{\frac{1}{d-1}}.$$

Integrating this inequality with respect to x_1 and using the general Hölder's inequality[2] with $p_i = d-1$, $i = 1, \dots, d-1$, we obtain

$$\begin{aligned}
\int_{-\infty}^{\infty} |u(x)|^{\frac{d}{d-1}}\, dx_1 &\leq \int_{-\infty}^{\infty} \prod_{i=1}^{d} \left(\int_{-\infty}^{\infty} |Du|\, dy_i \right)^{\frac{1}{d-1}} dx_1 \\
&= \left(\int_{-\infty}^{\infty} |Du|\, dy_1 \right)^{\frac{1}{d-1}} \int_{-\infty}^{\infty} \prod_{i=2}^{d} \left(\int_{-\infty}^{\infty} |Du|\, dy_i \right)^{\frac{1}{d-1}} dx_1 \\
&\leq \left(\int_{-\infty}^{\infty} |Du|\, dy_1 \right)^{\frac{1}{d-1}} \left(\prod_{i=2}^{d} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du|\, dx_1 dy_i \right)^{\frac{1}{d-1}}.
\end{aligned}$$

We now integrate with respect to x_2,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u(x)|^{\frac{d}{d-1}}\, dx_1 dx_2 \leq \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du|\, dx_1 dy_2 \right)^{\frac{1}{d-1}} \int_{-\infty}^{\infty} \prod_{i=1, i\neq 2}^{d} I_i^{\frac{1}{d-1}}\, dx_2,$$

where

$$I_1 := \int_{-\infty}^{\infty} |Du|\, dy_1, \quad I_i := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du|\, dx_1 dy_i, \quad i = 3, \dots, d.$$

[2] Given $1 < p_1, p_2, \dots, p_n < \infty$ with $\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_n} = 1$ and $f_i \in L^{p_i}(\Omega)$, $i = 1, 2, \dots, n$,

$$\left| \int_{\Omega} f_1\, f_2 \, \dots f_n \right| \leq \|f_1\|_{p_1} \|f_2\|_{p_2} \dots \|f_n\|_{p_n}.$$

Applying once more the extended Hölder inequality, we find

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u(x)|^{\frac{d}{d-1}}\, dx_1 dx_2 \\
&\qquad \le \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\, dx_1 dy_2\right)^{\frac{1}{d-1}} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\, dy_1 dx_2\right)^{\frac{1}{d-1}} \\
&\qquad\qquad \prod_{i=3}^{d}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\, dx_1 dx_2 dy_i\right)^{\frac{1}{d-1}}.
\end{aligned}
$$

We continue by integrating with respect to $x_3, \dots, x_d$, eventually to find

$$
\begin{aligned}
\int_{\mathbb{R}^d} |u(x)|^{\frac{d}{d-1}}\, dx &\le \prod_{i=1}^{d}\left(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} |Du|\, dx_1 \dots dy_i \dots dx_d\right)^{\frac{1}{d-1}} \\
&= \left(\int_{\mathbb{R}^d} |Du|\, dx\right)^{\frac{d}{d-1}},
\end{aligned}
\tag{1.11}
$$

which is estimate (1.10) for $p = 1$.

STEP 2. We consider now the case $p \in (1, d)$. We apply estimate (1.11) to $v := |u|^\gamma$, where $\gamma > 1$ is to be selected. Then, applying Hölder's inequality

$$
\begin{aligned}
\left(\int_{\mathbb{R}^d} |u|^{\frac{\gamma d}{d-1}}\right)^{\frac{d-1}{d}} &\le \int_{\mathbb{R}^d} |D|u|^\gamma| = \gamma \int_{\mathbb{R}^d} |u|^{\gamma-1}|Du| \\
&\le \gamma \left(\int_{\mathbb{R}^d} |u|^{(\gamma-1)\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^d} |Du|^p\right)^{\frac{1}{p}}.
\end{aligned}
\tag{1.12}
$$

We now choose γ so that $\frac{\gamma d}{d-1} = (\gamma - 1)\frac{p}{p-1}$. That is, we set

$$
\gamma := \frac{p(d-1)}{d-p} > 1,
$$

in which case $\frac{\gamma d}{d-1} = (\gamma - 1)\frac{p}{p-1} = p^*$, and (1.12) becomes the desired inequality (1.10). □

Remark. We really do need u to have compact support for (1.8.2) to hold, as the example $u \equiv 1$ shows. But remarkably the constant C does not depend at all upon the size of the support of u.

Corollary 1.8.2. *Let $p \in [1, d)$ and p^* the Sobolev conjugate of p. The Gagliardo-Nirenberg-Sobolev inequality* (1.10) *holds for all $u \in W^{1,p}(\mathbb{R}^d)$. Moreover, for every $q \in [p, p^*]$ there is a constant $C = C(d, p, q)$ such that*

$$
\|u\|_{L^q(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)} \text{ for all } u \in W^{1,p}(\mathbb{R}^d), \tag{1.13}
$$

so that $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$.

Proof. Let $(u_k)_{k=1}^{\infty}$ be a sequence of functions in $C_c^\infty(\mathbb{R}^d)$ such that $u_k \to u$ in $W^{1,p}(\mathbb{R}^d)$. According to inequality (1.10),

$$
\|u_k - u_j\|_{L^{p^*}(\mathbb{R}^d)} \le C\|Du_k - Du_j\|_{L^p(\mathbb{R}^d)} \text{ for all } k, j \ge 1.
$$

Thus, $u_k \to \tilde{u}$ in $L^{p^*}(\mathbb{R}^d)$ for some $\tilde{u} \in L^{p^*}(\mathbb{R}^d)$. In particular, we know from Theorem 1.2.6 that a subsequence of $(u_k)_{k=1}^\infty$ converges pointwise almost everywhere to $\tilde{u}$. Since this subsequence converges in $L^p(\mathbb{R}^d)$ to u, we know, using again Theorem 1.2.6, that it has a subsequence that converges almost everywhere to u, and we conclude that $\tilde{u} = u$ almost everywhere. Thus $u \in L^{p^*}(\mathbb{R}^d)$. Moreover, passing to the limit in the inequality $\|u_k\|_{L^{p^*}(\mathbb{R}^d)} \le C\|Du_k\|_{L^p(\mathbb{R}^d)}$, we obtain the inequality in (1.10) for all functions in $W^{1,p}(\mathbb{R}^d)$.

On the other hand, using the interpolation inequality[3] and Young's inequality (with exponents $1/\theta$ and $1/(1-\theta)$) we obtain

$$\|u\|_q \le \|u\|_p^\theta \|u\|_{p^*}^{1-\theta} \le \theta\|u\|_p + (1-\theta)\|u\|_{p^*} \le \theta\|u\|_p + (1-\theta)\|Du\|_p \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}.$$

□

Notation. $W^{0,p}(\Omega) := L^p(\Omega)$.

Corollary 1.8.3. *Let $p \in [1,\infty)$, $kp < d$. Then,*

$$W^{k,p}(\mathbb{R}^d) \hookrightarrow W^{k-j,q_j}(\mathbb{R}^d) \quad \text{for all } j \in \{1,\dots,k\}, \text{ where } q_j = \frac{dp}{d-jp}.$$

Proof. It is a consequence of a repeated application of the embedding for the case $k = 1$. Indeed, if $u \in W^{k,p}(\mathbb{R}^d)$, then $D^\alpha u \in W^{1,p}(\mathbb{R}^d)$ for all multi-indexes α of order $k-1$. Hence, since $d > kp > p$, the embedding for $k = 1$ yields $W^{k,p}(\mathbb{R}^d) \hookrightarrow W^{k-1,q_1}(\mathbb{R}^d)$, where

$$\frac{1}{q_1} = \frac{1}{p} - \frac{1}{d}.$$

Iterating the argument we obtain the desired result, with

$$\frac{1}{q_j} = \frac{1}{p} - \frac{j}{d}, \quad j \in \{1,\dots,k\}.$$

Notice that $q_j < d$ if $j \in \{1,\dots,k-1\}$, as required, since $kp < d$. □

Corollary 1.8.4. *Let $p \in [1,d)$ and $q \in [1,p^*]$. Let Ω be a bounded, open subset of $\mathbb{R}^d$ with C^1 boundary. Then there is a constant $C = C(d,p,q,\Omega)$ such that*

$$\|u\|_{L^q(\Omega)} \le C\|u\|_{W^{1,p}(\Omega)} \text{ for all } u \in W^{1,p}(\Omega). \tag{1.14}$$

The result holds valid for arbitrary bounded open sets Ω if $W^{1,p}(\Omega)$ is replaced by $W_0^{1,p}(\Omega)$.

Proof. Under our assumptions, there exists an extension $Eu = \bar{u} \in W^{1,p}(\mathbb{R}^d)$ such that

$$\bar{u} = u \text{ in } \Omega,\ \bar{u} \text{ has compact support, and } \|\bar{u}\|_{W^{1,p}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\Omega)}. \tag{1.15}$$

Let $(u_k)_{k=1}^\infty$ be a sequence of functions in $C_c^\infty(\mathbb{R}^d)$ such that $u_k \to \bar{u}$ in $W^{1,p}(\mathbb{R}^d)$. According to inequality (1.10),

$$\|u_k - u_j\|_{L^{p^*}(\mathbb{R}^d)} \le C\|Du_k - Du_j\|_{L^p(\mathbb{R}^d)} \text{ for all } k,j \ge 1.$$

[3]If $f \in L^p(\Omega) \cap L^r(\Omega)$, $p < r$, then $f \in L^q(\Omega)$ for all $q \in [p,r]$, and $\|f\|_q \le \|f\|_p^\theta \|f\|_r^{1-\theta}$, where $\theta \in [0,1]$ is given by $\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{r}$.

Thus, $u_k \to \bar{u}$ in $L^{p^*}(\mathbb{R}^d)$ as well. Since we also have $\|u_k\|_{L^{p^*}(\mathbb{R}^d)} \leq C\|Du_k\|_{L^p(\mathbb{R}^d)}$, passing to the limit we obtain $\|\bar{u}\|_{L^{p^*}(\mathbb{R}^d)} \leq C\|D\bar{u}\|_{L^p(\mathbb{R}^d)}$, which combined with (1.15) yields

$$\|u\|_{L^{p^*}(\Omega)} = \|\bar{u}\|_{L^{p^*}(\Omega)} \leq \|\bar{u}\|_{L^{p^*}(\mathbb{R}^d)} \leq C\|D\bar{u}\|_{L^p(\mathbb{R}^d)} \leq \|\bar{u}\|_{W^{1,p}(\mathbb{R}^d)} \leq C\|u\|_{W^{1,p}(\Omega)},$$

which gives (1.14) for $q = p^*$.

Let now $q \in [1, p^*)$. By Hölder's inequality

$$\|u\|_q = \left(\int_\Omega \mathbb{1}_\Omega |u|^q\right)^{\frac{1}{q}} \leq \|u\|_{p^*} |\Omega|^{\frac{p^*-q}{p^*}},$$

from where the result follows using (1.14) with $q = p^*$. □

1.8.2 The limit case.

Theorem 1.8.5. *$W^{1,d}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$ for all $q \in [d, \infty)$.*

Proof. Combining (1.12) with Young's inequality we get the estimate

$$\|u\|_{\frac{d\gamma}{d-1}} \leq C(\|u\|_{(\gamma-1)\frac{d}{d-1}} + \|Du\|_d) \quad \text{valid for all } \gamma \geq 1.$$

Taking $\gamma = d$ we obtain

$$\|u\|_{\frac{d^2}{d-1}} \leq C\|u\|_{W^{1,d}(\mathbb{R}^d)},$$

which implies, by interpolation, the desired inclusion for all $q \in [d, d^2/(d-1)]$.

Take now $\gamma = d + 1$. We obtain, using also the previous step, the estimate

$$\|u\|_{\frac{d(d+1)}{d-1}} \leq C(\|u\|_{\frac{d^2}{d-1}} + \|Du\|_d) \leq C\|u\|_{W^{1,d}(\mathbb{R}^d)},$$

which implies, by interpolation, the desired inclusion for all $q \in [d, d(d+1)/(d-1)]$.

Taking $\gamma = d + 2$, $\gamma = d + 3$, ..., we eventually reach any finite exponent $q > d$. □

Remark. If $d > 1$, it is not true that $W^{1,d}(\Omega) \hookrightarrow L^\infty(\Omega)$. For a counterexample see problem 1 in worksheet 4.

Corollary 1.8.6. *$W^{k,d/k}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$ for all $q \in [d, \infty)$.*

Proof. Use $k - 1$ times the embedding (1.13) and one time the embedding from Theorem 1.8.5. □

1.8.3 Morrey's inequality.

Let now $p \in (d, \infty)$. We will show that if $u \in W^{1,p}$ then u is in fact *Hölder continuous*, after possibly being redefined on a set of measure zero.

Let us start by explaining what a Hölder space is. Assume $\Omega \subset \mathbb{R}^d$ is open. Let us recall that a function $u : \Omega \to \mathbb{R}$ is said to be *Lipschitz continuous* in Ω if there exists some constant C such that

$$|u(x) - u(y)| \leq C|x - y| \quad \text{for all } x, y \in \Omega.$$

This estimate of course implies that u is continuous. More importantly, it allows to quantify how close should be x and y in order for $u(y)$ to be close to $u(x)$. Such quantitative estimates are important in the analysis of PDEs. Hölder continuous functions satisfy a variant of the above estimate.

Definition 1.8.3. A function $u : \Omega \to \mathbb{R}$ is said to be *Hölder continuous* in Ω with exponent $\gamma \in (0, 1]$ if there exists some constant C such that

$$|u(x) - u(y)| \le C|x - y|^{\gamma} \quad \text{for all } x, y \in \Omega.$$

Definition 1.8.4. (i) If $u : \Omega \to \mathbb{R}$ is bounded and continuous we write

$$\|u\|_{C(\overline{\Omega})} := \sup_{x \in \Omega} |u(x)|.$$

(ii) The γ^{th}-*Hölder seminorm* of $u : \Omega \to \mathbb{R}$ is

$$\left[u\right]_{C^{0,\gamma}(\overline{\Omega})} := \sup_{x,y \in \Omega, x \neq y} \frac{|u(x) - u(y)|}{|x - y|^{\gamma}},$$

and the γ^{th}-*Hölder norm* is

$$\|u\|_{C^{0,\gamma}(\overline{\Omega})} := \|u\|_{C(\overline{\Omega})} + \left[u\right]_{C^{0,\gamma}(\overline{\Omega})}.$$

Definition 1.8.5. The *Hölder space* $C^{k,\gamma}(\overline{\Omega})$ consists of all functions $u \in C^k(\overline{\Omega})$ for which the norm

$$\|u\|_{C^{k,\gamma}(\overline{\Omega})} := \sum_{|\alpha| \le k} \|D^{\alpha} u\|_{C(\overline{\Omega})} + \sum_{|\alpha| = k} \left[D^{\alpha} u\right]_{C^{(0,\gamma)}(\overline{\Omega})}$$

is finite

So the space $C^{k,\gamma}(\overline{\Omega})$ consists of those functions u that are k times continuously differentiable and whose k^{th}-partial derivatives are bounded and Hölder continuous with exponent γ. Such functions are well behaved, and the space $C^{k,\gamma}(\overline{\Omega})$ itself posesses a good mathematical structure.

Theorem 1.8.7 (Hölder spaces as function spaces). *The space of functions* $C^{k,\gamma}(\overline{\Omega})$ *is a Banach space.*

Proof. I leave it as an exercise (it may be helpful to use Ascoli-Arzela's compactness criterion). □

Theorem 1.8.8 (Morrey's inequality). *Let* $p \in (d, \infty]$. *Then there exists a constant* C, *depending only on* p *and* d, *such that*

$$\|u\|_{C^{0,\gamma}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)} \text{ for all } u \in C^1(\mathbb{R}^d), \text{ where } \gamma = 1 - \frac{d}{p}.$$

Proof. STEP 1. We claim that there exists a constant C, depending only on d, such that

$$\fint_{B_r(x)} |u(y) - u(x)|\, dy \le C \int_{B_r(x)} \frac{|Du(y)|}{|y - x|^{d-1}}\, dy. \tag{1.16}$$

To prove the claim, fix any point $w \in \partial B_1(0)$. Then, if $0 < s < r$,

$$|u(x+sw)-u(x)| = \left|\int_0^s \frac{d}{dt}u(x+tw)\,dt\right| = \left|\int_0^s Du(x+tw)\cdot w\,dt\right| \le \int_0^s |Du(x+tw)|\,dt.$$

Hence, integrating in $\partial B_1(0)$,

$$\begin{aligned}\int_{\partial B_1(0)} |u(x+sw)-u(x)|\,dS(w) &\le \int_0^s \int_{\partial B_1(0)} |Du(x+tw)|\,dS(w)dt\\ &= \int_0^s \int_{\partial B_t(x)} \frac{|Du(y)|}{t^{d-1}}\,dS(y)dt\\ &= \int_{B_s(x)} \frac{|Du(y)|}{|x-y|^{d-1}}\,dy\\ &\le \int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{d-1}}\,dy,\end{aligned}$$

where we put $y = x+tw$, $t = |x-y|$. Therefore,

$$\int_{\partial B_s(x)} |u(z)-u(x)|\,dS(z) = s^{d-1}\int_{\partial B_1(0)} |u(x+sw)-u(x)|\,dS(w) \le s^{d-1}\int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{d-1}}\,dy.$$

Integrating with respect to s from 0 to r we finally arrive to

$$\int_{B_r(x)} |u(y)-u(x)|\,dy \le \frac{r^d}{d}\int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{d-1}}\,dy,$$

which proves the claim.

Step 2. Let us prove now that u is bounded. Given any $x \in \mathbb{R}^d$, estimate (1.16) yields, applying also Hölder's inequality,

$$\begin{aligned}|u(x)| &= ⨍_{B_1(x)} |u(x)|\,dy \le ⨍_{B_1(x)} |u(x)-u(y)|\,dy + ⨍_{B_1(x)} |u(y)|\,dy\\ &\le C\int_{B_1(x)} \frac{|Du(y)|}{|x-y|^{d-1}}\,dy + C\|u\|_{L^p(B_1(x))}\\ &\le C\|Du\|_{L^p(B_1(x))}\left(\int_{B_1(x)} \frac{1}{|x-y|^{(d-1)\frac{p}{p-1}}}\,dy\right)^{\frac{p-1}{p}} + C\|u\|_{L^p(\mathbb{R}^d)}.\end{aligned}$$

Since $p > d$, then $(d-1)\frac{p}{p-1} < d$. Hence

$$\int_{B_1(x)} \frac{1}{|x-y|^{(d-1)\frac{p}{p-1}}}\,dy < \infty,$$

and we conclude that $|u(x)| \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}$. Since x is arbitrary, it follows that

$$\sup_{\mathbb{R}^d} |u| \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}.$$

Step 3. We now obtain a bound for the Hölder seminorm.

Choose any two points $x, y \in \mathbb{R}^d$ and write $r = |x-y|$. Let $W := B_r(x) \cap B_r(y)$. Then

$$|u(x) - u(y)| \leq \fint_W |u(x) - u(z)|\, dz + \fint_W |u(y) - u(z)|\, dz. \tag{1.17}$$

Estimate (1.16) combined with Hölder's inequality allows to obtain

$$\begin{aligned}
\fint_W |u(x) - u(z)|\, dz &\leq \frac{|B_r|}{|W|} \fint_{B_r(x)} |u(x) - u(z)|\, dz \\
&\leq C\|Du\|_{L^p(B_r(x))} \left(\int_{B_r(x)} \frac{1}{|x-z|^{(d-1)\frac{p}{p-1}}}\, dz \right)^{\frac{p-1}{p}} \\
&\leq C\|Du\|_{L^p(\mathbb{R}^d)} \left(r^{d-(d-1)\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \\
&= C\|Du\|_{L^p(\mathbb{R}^d)} r^{1-\frac{d}{p}}.
\end{aligned}$$

Likewise,

$$\fint_W |u(y) - u(z)|\, dz \leq C\|Du\|_{L^p(\mathbb{R}^d)} r^{1-\frac{d}{p}}.$$

Plugging these two estimates in (1.17) we get

$$|u(x) - u(y)| \leq C\|Du\|_{L^p(\mathbb{R}^d)} |x-y|^{1-\frac{d}{p}},$$

which yields the bound

$$[u]_{C^{0,1-\frac{p}{d}}(\mathbb{R}^d)} \leq C\|Du\|_{L^p(\mathbb{R}^d)} \tag{1.18}$$

and Morrey's inequality follows. □

Definition 1.8.6. We say u^* is a *version* of a given function u provided $u = u^*$ a.e.

Theorem 1.8.9 (Estimates for $W^{1,p}$, $p \in (d, \infty]$). *Let $p \in (d, \infty]$ and $\gamma = 1 - \frac{d}{p}$. Let Ω be a bounded, open subset of $\mathbb{R}^d$ with C^1 boundary. Then there is a constant $C = C(d, p, q, \Omega)$ such that every $u \in W^{1,p}(\Omega)$ has a version $u^* \in C^{0,\gamma}(\overline{\Omega})$ satisfying*

$$\|u^*\|_{C^{0,\gamma}(\overline{\Omega})} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

Remark. In view of this theorem, we will henceforth always identify a function $u \in W^{1,p}(\Omega)$, $p > d$, with its continuous version.

Proof. We give the proof when $p \in (d, \infty)$. The case $p = \infty$ is left as an exercise; see problem 2 in worksheet 5.

Thanks to the properties of Ω, we know that there exists an extension $Eu = \bar{u} \in W^{1,p}(\mathbb{R}^d)$ such that

$$\bar{u} = u \text{ in } \Omega, \quad \bar{u} \text{ has compact support}, \quad \|\bar{u}\|_{W^{1,p}(\mathbb{R}^d)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

Meyer-Serrin's theorem guaranties the existence of a sequence of functions $u_k \in C_{\mathrm{c}}^\infty(\mathbb{R}^d)$ such that $u_k \to \bar{u}$ in $W^{1,p}(\mathbb{R}^d)$. Now, thanks to Morrey's inequality, we have

$$\|u_k - u_j\|_{C^{0,\gamma}(\mathbb{R}^d)} \leq C\|u_k - u_j\|_{W^{1,p}(\mathbb{R}^d)} \text{ for all } k, j \geq 1$$

whence there exists a function $u^* \in C^{0,\gamma}(\mathbb{R}^d)$ such that $u_k \to u^*$ in $C^{0,\gamma}(\mathbb{R})$. The convergence of u_k in $L^p(\mathbb{R}^d)$ implies that along some subsequence $u_{k_j} \to \bar{u}$ a.e. Hence, $u^* = \bar{u}$ a.e., and $u^* = u$ a.e. in Ω, so that u^* is a version of u in Ω. Morrey's inequality also implies $\|u_k\|_{C^{0,\gamma}(\mathbb{R}^d)} \le C\|u_k\|W^{1,p}(\mathbb{R}^d)$. Passing to the limit we obtain $\|\bar{u}\|_{C^{0,\gamma}(\mathbb{R}^d)} \le C\|\bar{u}\|W^{1,p}(\mathbb{R}^d)$, from where the desired inequality follows easily. □

Theorem 1.8.10. *Let $p \in (d,\infty)$ and $\gamma = 1 - \frac{d}{p}$. There is a constant $C = C(d,p)$ such that every $u \in W^{1,p}(\mathbb{R}^d)$ has a version $u^* \in C^{0,\gamma}(\overline{\Omega})$ satisfying*

$$\|u^*\|_{C^{0,\gamma}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}.$$

Proof. It is left as an exercise for the student. □

Remark. If $u \in W^{1,p}(\mathbb{R}^d)$, $p \in (d,\infty)$, then $\lim_{|x|\to\infty} u(x) = 0$. The proof is left as an exercise for the student (*Hint:* Approximate u by functions in $C_c^\infty(\mathbb{R}^d)$ and use Morrey's inequality to pass to the limit in $C^{0,\gamma}(\mathbb{R}^d)$). The result is not true for $p = \infty$.

1.9 Compact embeddings

We have seen that the Gagliardo-Nirenberg-Sobolev inequality implies the embedding of $W^{1,p}(\Omega)$ into $L^{p^*}(\Omega)$ for $1 \le p < d$, $p^* = \frac{dp}{d-p}$. We will prove now that in fact $W^{1,p}(\Omega)$ is in fact *compactly* embedded in $L^q(\Omega)$ for $q \in [1,p^*)$ (assuming that Ω is bounded and smooth). This compactness is fundamental for applications of linear and nonlinear functional analysis to PDEs.

Definition 1.9.1. Let X and Y be Banach spaces, $X \subset Y$. We say that X is *compactly embedded* in Y, written $X \subset\subset Y$, provided:

(i) $\|u\|_Y \le C\|u\|_X$ for all $u \in X$ for some fixed constant C;
(ii) each bounded sequence in X is precompact in Y.

More precisely, condition (ii) means that if $(u_k)_{k=1}^\infty$ is a sequence in X with $\sup_k \|u\|_X < \infty$, then some subsequence $(u_{k_j})_{j=1}^\infty \subset (u_k)_{k=1}^\infty$ converges in Y to some limit u:

$$\lim_{j\to\infty} \|u_{k_j} - u\|_Y = 0.$$

Theorem 1.9.1 (Rellich-Kondrachov compactness theorem). *Let Ω be a bounded open subset of $\mathbb{R}^d$ with C^1 boundary, and $p \in [1,d)$. Then*

$$W^{1,p}(\Omega) \subset\subset L^q(\Omega) \quad \textit{for each } q \in [1,p^*).$$

Proof. STEP 1. Let $q \in [1,p^*)$. We already now that $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$. Hence it only remains to show that if $(u_k)_{k=1}^\infty$ is a bounded sequence in $W^{1,p}(\Omega)$, there exists a subsequence $(u_{k_j})_{j=1}^\infty$ which converges in $L^q(\Omega)$.

STEP 2. In view of the Extension Theorem, we may with no loss of generality assume that $\Omega = \mathbb{R}^d$ and that the functions $(u_k)_{k=1}^\infty$ all have compact support contained in a compact subset of some bounded open set $V \subset \mathbb{R}^d$. We may also assume

$$\sup_k \|u_k\|_{W^{1,p}(V)} < \infty. \tag{1.19}$$

Step 3. We considered the smoothed functions $u_k^\delta := \rho_\delta \star u_k$, where ρ_δ denotes the standard mollifier. The supports of the functions $(u_k^\delta)_{k=1}^\infty$ are all contained in some set $K \subset\subset V$ if $\delta \in (0, \delta_0)$ for some δ_0 small.

We claim that

$$u_k^\delta \to u_k \text{ in } L^q(V) \text{ as } \delta \to 0 \textit{ uniformly} \text{ in } k.$$

To prove this, we first note that if u_k is smooth, then

$$\begin{aligned} u_k^\delta(x) - u_k(x) &= \frac{1}{\delta^d}\int_{B_\delta(x)} \rho\left(\frac{x-z}{\delta}\right)(u_k(z) - u_k(x))\,dz \\ &= \int_{B_1(0)} \rho(y)(u_k(x-\delta y) - u_k(x))\,dy \\ &= \int_{B_1(0)} \rho(y)\int_0^1 \frac{d}{dt}\big(u_k(x-\delta ty)\big)\,dtdy \\ &= -\delta\int_{B_1(0)} \rho(y)\int_0^1 Du_k(x-\delta ty)\cdot y\,dtdy. \end{aligned}$$

Thus, since $u_k(x-\delta ty) = 0$ for all $k \in \mathbb{N}$, $\delta \le \delta_0$ and $t \in [0,1]$ if $x \notin K$, we have

$$\begin{aligned} \int_V |u_k^\delta - u_k| &\le \delta\int_{B_1(0)} \rho(y)\int_0^1\int_V |Du_k(x-\delta ty)|\,dxdtdy \\ &= \delta\int_{B_1(0)} \rho(y)\int_0^1\int_K |Du_k(\xi)|\,d\xi dtdy = \delta\int_V |Du_k|. \end{aligned}$$

By approximation this estimate holds if $u_k \in W^{1,p}(V)$ (check it!). Hence

$$\|u_k^\delta - u_k\|_{L^1(V)} \le \delta\|Du_k\|_{L^1(V)} \le \delta C\|Du_k\|_{L^p(V)}, \tag{1.20}$$

the latter inequality holding since V is bounded. Owing to (1.19), we thereby discover

$$u_k^\delta \to u_k \quad \text{in } L^1(V) \text{ as } \delta \to 0^+, \text{ uniformly in } k.$$

On the other hand, since $1 \le q < p^*$, the interpolation inequality yields

$$\|u_k^\delta - u_k\|_{L^q(V)} \le \|u_k^\delta - u_k\|_{L^1(V)}^\theta \|u_k^\delta - u_k\|_{L^{p^*}(V)}^{1-\theta},$$

where $\frac{1}{q} = \theta + \frac{1-\theta}{p^*}$, $0 < \theta < 1$. Consequently, (1.19) and the Gagliardo-Nirenberg-Sobolev inequality imply (remember also that $\|f^\delta\|_p \le \|f\|_p$),

$$\|u_k^\delta - u_k\|_{L^q(V)} \le C\|u_k^\delta - u_k\|_{L^1(V)}^\theta,$$

whence the claim follows from (1.20).

Step 4. We now prove that for each fixed $\delta > 0$ the sequence $(u_k^\delta)_{k=1}^\infty$ is uniformly bounded and equicontinuous[4]. Indeed,

$$|u_k^\delta(x)| \le \int_{B_\delta(0)} \rho_\delta(x-y)|u_k(y)|\,dy \le \|\rho_\delta\|_{L^\infty(\mathbb{R}^d)}\|u_k\|_{L^1(V)} \le \frac{C}{\delta^d} = C_\delta < \infty,$$

[4]A sequence $(f_k)_{k=1}^\infty$ of real-valued functions defined on $K \subset \mathbb{R}^d$ is uniformly *equicontinuous* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $|x-y| < \delta$ implies $|f_k(x) - f_k(y)| < \varepsilon$ for $x, y \in K$, $k \in \mathbb{N}$.

where C_δ does not depend on k; that is, the family is uniformly bounded. Similarly,

$$|Du_k^\delta(x)| \leq \int_{B_\delta(x)} |D\rho_\delta(x-y)||u_k(y)|\,dy \leq \|D\rho_\delta\|_{L^\infty(\mathbb{R}^d)}\|u_k\|_{L^1(V)} \leq \frac{C}{\delta^{d+1}} = C_\delta < \infty,$$

where again C_δ does not depend on k. This yields immediately the equicontinuity.

STEP 5. Let $\varepsilon > 0$. We will show that there exists a subsequence $(u_{k_j})_{j=1}^\infty \subset (u_k)_{k=1}^\infty$ such that

$$\limsup_{j,l\to\infty} \|u_{k_j} - u_{k_l}\|_{L^q(V)} \leq \varepsilon.$$

To see this, we use Step 3 to select $\delta > 0$ so small that

$$\|u_k^\delta - u_k\|_{L^q(V)} \leq \varepsilon/2 \quad \text{for } k \in \mathbb{N}.$$

We now observe that since the functions $(u_k)_{k=1}^\infty$, and thus also the functions $(u_k^\delta)_{k=1}^\infty$ if δ is small enough, have support contained in some fixed bounded set V, we may use Step 4 and the Arzelà-Ascoli compactness criterion[5] to obtain a subsequence $(u_{k_j}^\delta)_{j=1}^\infty \subset (u_k^\delta)_{k=1}^\infty$ which converges *uniformly* on V. Therefore

$$\limsup_{j,l\to\infty} \|u_{k_j}^\delta - u_{k_l}^\delta\|_{L^q(V)} = 0.$$

Since

$$\|u_{k_j} - u_{k_l}\|_{L^q(V)} \leq \|u_{k_j} - u_{k_j}^\delta\|_{L^q(V)} + \|u_{k_j}^\delta - u_{k_l}^\delta\|_{L^q(V)} + \|u_{k_l}^\delta - u_{k_l}\|_{L^q(V)},$$

we get the claim.

STEP 6. The proof will now follow employing Step 5 with $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ plus a standard diagonal argument. Indeed, we start with $\varepsilon = 1$ and consider a subsequence of $(u_k)_{k=1}^\infty$, that we denote by $(u_k^{(1)})_{k=1}^\infty$, such that

$$\limsup_{k,j\to\infty} \|u_k^{(1)} - u_j^{(1)}\|_{L^q(V)} \leq 1.$$

Next we consider a subsequence of $(u_k^{(1)})_{k=1}^\infty$, that we denote by $(u_k^{(2)})_{k=1}^\infty$, such that

$$\limsup_{k,j\to\infty} \|u_k^{(2)} - u_j^{(2)}\|_{L^q(V)} \leq \frac{1}{2}.$$

We iterate this argument, considering a subsequence of $(u_k^{(n-1)})_{k=1}^\infty$, that we denote by $(u_k^{(n)})_{k=1}^\infty$, such that

$$\limsup_{k,j\to\infty} \|u_k^{(n-1)} - u_j^{(n-1)}\|_{L^q(V)} \leq \frac{1}{n}.$$

We finally consider the subsequence of the original sequence given by $(u_k^{(k)})_{k=1}^\infty$, which is a Cauchy sequence in $L^q(V)$, and hence convergent. □

[5] Arzelà-Ascoli compactness criterion asserts that any sequence of real valued functions defined in a compact subset K of $\mathbb{R}^d$ which is uniformly bounded and equicontinuous has a subsequence that converges uniformly on K. The limit is obviously a continuous function.

Corollary 1.9.2. *Let Ω be a bounded open subset of $\mathbb{R}^d$ with C^1 boundary. Then*

$$W^{1,d}(\Omega) \subset\subset L^q(\Omega) \quad \text{for each } q \in [1,\infty).$$

Proof. Given $q \in [1,\infty)$, take $p \in [1,d)$ such that $p* > q$ (any $p \in (\frac{q}{d+q}, d)$ will do the job). Since Ω is bounded, a bounded sequence in $W^{1,d}(\Omega)$ will also be bounded in $W^{1,p}(\Omega)$, and because of the compact embedding $W^{1,p}(\Omega) \subset\subset L^r(\Omega)$ for all $r \in [1,p^*)$, it will have a convergent subsequence in $L^q(\Omega)$, since $q \in [1,p^*)$. □

Lemma 1.9.3. *Let Ω be a bounded, open set in $\mathbb{R}^d$, and $\alpha \in (0,1]$. Then*

$$\begin{array}{lll} \text{(a)} & C^{0,\alpha}(\overline{\Omega}) \subset\subset C^{0,\beta}(\overline{\Omega}) & \text{for all } \beta \in (0,\alpha), \\ \text{(b)} & C^{0,\alpha}(\overline{\Omega}) \subset\subset L^q(\Omega) & \text{for all } q \in [1,\infty]. \end{array}$$

Proof. (a) By definition,

$$\begin{aligned} \big[u\big]_{C^{0,\beta}(\overline{\Omega})} &= \sup_{x,y\in\Omega, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\beta} = \sup_{x,y\in\Omega, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\alpha} |x-y|^{\alpha-\beta} \\ &\leq \sup_{x,y\in\Omega, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\alpha} \sup_{x,y\in\Omega} |x-y|^{\alpha-\beta} = \big[u\big]_{C^{0,\alpha}(\overline{\Omega})} \operatorname{diam}(\Omega)^{\alpha-\beta}, \end{aligned}$$

which implies that $C^{0,\alpha}(\overline{\Omega}) \hookrightarrow C^{0,\beta}(\overline{\Omega})$ if $\beta < \alpha$.

Let us see that the embedding is also compact. The proof is based in the following estimate,

$$\begin{aligned} \big[u\big]_{C^{0,\beta}(\overline{\Omega})} &= \sup_{x,y\in\Omega, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\beta} = \sup_{x,y\in\Omega, x\neq y} \frac{|u(x)-u(y)|^{\frac{\beta}{\alpha}}}{|x-y|^\beta} |u(x)-u(y)|^{1-\frac{\beta}{\alpha}} \\ &\leq \left(\sup_{x,y\in\Omega, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\alpha}\right)^{\frac{\beta}{\alpha}} \sup_{x,y\in\Omega} |u(x)-u(y)|^{1-\frac{\beta}{\alpha}} \\ &\leq \left(\big[u\big]_{C^{0,\alpha}(\overline{\Omega})}\right)^{\frac{\beta}{\alpha}} \left(2\|u\|_{L^\infty(\Omega)}\right)^{1-\frac{\beta}{\alpha}}. \end{aligned}$$

Let $(u_k)_{k=1}^\infty$ be a bounded sequence in $C^{0,\alpha}(\overline{\Omega})$. It is obviously uniformly bounded and equicontinuous in $\overline{\Omega}$. Hence, thanks to Arzelà-Ascoli compactness criterion we know that it has a subsequence $(u_{k_j})_{j=1}^\infty$ that converges uniformly in $\overline{\Omega}$. Using the above estimate and the boundedness of the sequence we get, for some constant C,

$$\begin{aligned} \|u_{k_j} - u_{k_l}\|_{C^{0,\beta}(\overline{\Omega})} &= \|u_{k_j} - u_{k_l}\|_{L^\infty(\Omega)} + \big[u_{k_j} - u_{k_l}\big]_{C^{0,\beta}(\overline{\Omega})} \\ &\leq \|u_{k_j} - u_{k_l}\|_{L^\infty(\Omega)} + \left(\big[u_{k_j} - u_{k_l}\big]_{C^{0,\alpha}(\overline{\Omega})}\right)^{\frac{\beta}{\alpha}} \left(2\|u_{k_j} - u_{k_l}\|_{L^\infty(\Omega)}\right)^{1-\frac{\beta}{\alpha}} \\ &\leq \|u_{k_j} - u_{k_l}\|_{L^\infty(\Omega)} + C\left(\|u_{k_j} - u_{k_l}\|_{L^\infty(\Omega)}\right)^{1-\frac{\beta}{\alpha}}. \end{aligned}$$

Therefore, since $(u_{k_j})_{j=1}^\infty$ is a Cauchy sequence in $L^\infty(\Omega)$, it is also a Cauchy sequence in $C^{0,\beta}(\overline{\Omega})$, $\beta \in (0,\alpha)$, hence convergent.

(b) As we have seen in the proof of (a), a bounded sequence in $C^{0,\alpha}(\overline{\Omega})$ has a bounded subsequence that converges uniformly in $\overline{\Omega}$, hence in $L^q(\Omega)$, $q \in [1,\infty]$, since Ω is bounded. □

Corollary 1.9.4. *Let Ω be a bounded open subset of $\mathbb{R}^d$ with C^1 boundary, and $p \in (d, \infty]$. Then*

$$W^{1,p}(\Omega) \subset\subset C^{0,\beta}(\Omega) \quad \text{for each } \beta \in \left(0, 1 - \frac{d}{p}\right).$$

As a consequence, $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$ for all $q \in [1, \infty]$.

Proof. Since in this range $W^{1,p}(\Omega) \hookrightarrow C^{0,1-\frac{d}{p}}(\overline{\Omega})$, a bounded sequence in $W^{1,p}(\Omega)$ is also bounded in $C^{0,1-\frac{d}{p}}(\overline{\Omega})$. But the latter space is compactly embedded in $C^{0,\beta}(\overline{\Omega})$ if $\beta \in \left(0, 1 - \frac{d}{p}\right)$, and the result follows. □

Remark. As a consequence of the previous results, we have in particular

$$W^{1,p}(\Omega) \subset\subset L^p(\Omega) \quad \text{for all } p \in [1, \infty]$$

if Ω is a bounded, open subset of $\mathbb{R}^d$ with C^1 boundary. Note also that

$$W^{1,p}_0(\Omega) \subset\subset L^p(\Omega) \quad \text{for all } p \in [1, \infty]$$

even if we do not assume $\partial\Omega$ to be C^1.

1.10 Poincaré inequalities

Poincaré-type inequalities are estimates of L^q norms in bounded sets in terms of L^p norms *of the gradient.* The first inequality of this type is a consequence of the results on compact embeddings.

Theorem 1.10.1 (Poincaré-Friedrichs' inequality). *Let Ω be a bounded, open subset of $\mathbb{R}^d$. There are constants $C = C(d, p, q, \Omega)$ such that*

$$\|u\|_{L^q(\Omega)} \le C\|Du\|_{L^p(\Omega)} \quad \text{for all } u \in W^{1,p}_0(\Omega) \tag{1.21}$$

for all q such that:

(a) $q \in [1, p^*]$ *if* $p \in [1, d)$;
(b) $q \in [1, \infty)$ *if* $p = d$;
(c) $q \in [1, \infty]$ *if* $p \in (d, \infty]$.

Proof. (a) Since $u \in W^{1,p}_0(\Omega)$, there exists a sequence $(u_k)_{k=1}^\infty$ of functions in $C^\infty_c(\Omega)$ converging to u in $W^{1,p}(\Omega)$. We extend each function u_k to be 0 on $\mathbb{R}^d \setminus \overline{\Omega}$. Let $p \in [1, d)$. Theorem 1.8.1 yields $\|u_k\|_{L^{p^*}(\Omega)} \le C\|Du_k\|_{L^p(\Omega)}$. Passing to the limit we get (1.21) with $q = p^*$. The result for $q \in [1, p^*)$ then follows using Hölder's inequality, since $|\Omega| < \infty$.
(b) If $p = d$, given any $q \in [1, \infty)$, we take $r \in [1, d)$ such that $r* > q$ (any $r \in (\frac{q}{d+q}, d)$ will do the job). The result now follows from (a) and Hölder's inequality, since

$$\|u\|_{L^q(\Omega)} \le C\|u\|_{L^{r^*}(\Omega)} \le C\|Du\|_{L^r(\Omega)} \le C\|Du\|_{L^d(\Omega)}.$$

(c) Let now $p \in (d, \infty]$. As in (a), we consider a sequence $(u_k)_{k=1}^\infty$ of functions in $C^\infty_c(\Omega)$ converging to u in $W^{1,p}(\Omega)$ and extend them to be 0 on $\mathbb{R}^d \setminus \overline{\Omega}$. Take $x_0 \in \partial\Omega$ and $x \in \Omega$.

Estimate (1.18), which was obtained in the course of the proof of Morrey's inequality, yields

$$|u_k(x)| = |u_k(x) - u_k(x_0)| \le \left[u_k\right]_{C^{0,1-\frac{d}{p}}(\overline{\Omega})} |x - x_0|^{1-\frac{d}{p}} \le C\|Du_k\|_{L^p(\Omega)}.$$

The result for $q = \infty$ follows after passing to the limit (recall that bounded sequences in $W^{1,p}(\Omega)$, $p \in (d, \infty]$, have a subsequence that converges uniformly). This then yields the result for finite q by means of Hölder's inequality. □

Remark. In view of Poincaré-Friedrichs' inequality, on $W_0^{1,p}(\Omega)$ the norm $\|Du\|_{L^p(\Omega)}$ is equivalent to $\|u\|_{W^{1,p}(\Omega)}$ if Ω is bounded and smooth.

The compactness results in the previous section can be used to generate new inequalities.

Theorem 1.10.2 (Poincaré-Wirtinger's inequality). *Let Ω be a bounded, connected, open subset with C^1 boundary. Assume $1 \le p \le \infty$. Then there exists a constant C, depending only on d, p and Ω, such that*

$$\left\| u - \fint_\Omega u \right\|_{L^p(\Omega)} \le C\|Du\|_{L^p(\Omega)} \quad \text{for all } u \in W^{1,p}(\Omega).$$

Proof. We argue by contradiction. Were the stated estimate false, there would exist for each integer $k = 1, \dots$ a function $u_k \in W^{1,p}(\Omega)$ satisfying

$$\left\| u_k - \fint_\Omega u_k \right\|_{L^p(\Omega)} > k\|Du_k\|_{L^p(\Omega)}. \tag{1.22}$$

We normalize by defining

$$v_k := \frac{u_k - \fint_\Omega u_k}{\left\|u_k - \fint_\Omega u_k\right\|_{L^p(\Omega)}}, \quad k = 1, \dots.$$

Then, $\fint_\Omega v_k = 0$, $\|v_k\|_{L^p(\Omega)} = 1$, and (1.22) implies $\|Dv_k\|_{L^p(\Omega)} < 1/k$, $k = 1, 2, \dots$. In particular, the sequence $(v_k)_{k=1}^\infty$ is bounded in $W^{1,p}(\Omega)$. Therefore, thanks to the compactness results of the previous section, we know that there exists a subsequence $(v_{k_j})_{j=1}^\infty \subset (v_k)_{k=1}^\infty$ and a function $v \in L^p(\Omega)$ such that $v_{k_j} \to v$ in $L^p(\Omega)$. It is then immediate to check (do it!) that

$$\fint_\Omega v = 0, \qquad \|v\|_{L^p(\Omega)} = 1.$$

On the other hand, for each $i \in \{1, \dots, d\}$ and $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega v\partial_i\phi = \lim_{k_j\to\infty} \int_\Omega v_{k_j}\partial_i\phi = -\lim_{k_j\to\infty} \int_\Omega \partial_i v_{k_j}\phi = 0,$$

since $\|Dv_{k_j}\|_{L^p(\Omega)} < 1/k_j$. Consequently $v \in W^{1,p}(\Omega)$ with $Dv = 0$ a.e. Thus v is a constant, since Ω is connected (see problem 1 in worksheet 3). Since $\fint_\Omega v = 0$, this constant must be 0. But this would imply that $\|v\|_{L^p(\Omega)} = 0$, which contradicts $\|v\|_{L^p(\Omega)} = 1$. □

Chapter 2

Linear elliptic equations

The aim of this chapter is to investigate the solvability of uniformly elliptic, second-order partial differential equations, subject to prescribed boundary conditions by means of Hilbert spaces tools. Some of the ideas that we will present are also valid for systems and for higher order equations.

2.1 Duals. Riesz-Fréchet representation theorem

In what follows X denotes a real Banach space.

Definition 2.1.1. (i) A bounded linear operator $F : X \to \mathbb{R}$ is called a bounded linear functional on X.

(ii) We write X^* to denote the collection of all bounded linear functionals on X; X^* is the *dual* space of X.

Remarks. (a) Linear operators are bounded if and only if they are continuous. Thus, X^* is the collection of all continuous linear functionals on X. It is sometimes denoted as the topological dual, to distinguish it from the algebraic dual.

(b) The dual of X is also denoted as X'.

Definition 2.1.2. (i) If $u \in X$, $F \in X^*$, we write $\langle F, u\rangle$ to denote the real number $F(u)$. The symbol $\langle \cdot, \cdot\rangle$ denotes the *pairing* of X^* and X.

(ii) We define a norm in X^* by

$$\|F\|_{X^*} := \sup\{\langle F, u\rangle : \|u\| \le 1\}.$$

(iii) A Banach space X is *reflexive* if $(X^*)^* = X$. More precisely, this means that for each $u^{**} \in (X^*)^*$ there exists $u \in X$ such that

$$\langle u^{**}, u^*\rangle = \langle u^*, u\rangle \quad \text{for all } u^* \in X^*.$$

In a Hilbert space H it is very easy to write down continuous linear functionals. Pick any $f \in H$. Then the map $u \mapsto (f, u)$ is a continous linear functional on H. It is a remarkable fact that *all* continuous linear functionals on H are obtained in this fashion.

Theorem 2.1.1 (Riesz-Fréchet representation theorem). *Given any $F \in H^*$ there exists a unique $f \in H$ such that*

$$\langle F, u\rangle = (f, u) \quad \text{for all } u \in H.$$

Morevorer, $\|f\|_H = \|F\|_{H^*}$.

You can find a proof of this important theorem for instance in [2, Chapter 5]. If you are not acquainted with the basic theory of Hilbert spaces, I strongly recommend you to read carefully that chapter.

The mapping $F \mapsto f$, which is a linear surjective isometry, allows us to identify H^* with H. In particular, any Hilbert space is reflexive.

Continuous linear functionals on L^p with p finite can be represented "concretely" as an integral.

Theorem 2.1.2 (Riesz representation theorem). *Let $p \in [1, \infty)$, p' its dual exponent, and $F \in (L^p)^*$. Then there exists a unique function $f \in L^{p'}$ such that*

$$\langle F, u\rangle = \int fu \quad \text{for all } u \in L^p.$$

Moreover, $\|f\|_{p'} = \|F\|_{(L^p)^*}$.

You can find a proof for instance in [2, Chapter 4].

The mapping $F \mapsto f$, which is a linear surjective isometry, allows us to identify the "abstract" space $(L^p)^*$, $p \in [1, \infty)$, with $L^{p'}$. In particular, if $p \in (1, \infty)$, then L^p is reflexive.

Remarks. (a) Though $(L^\infty)^* \supset L^1$, $(L^\infty)^* \neq L^1$. Hence, L^1 is not reflexive. Neither is L^∞.
(b) The spaces $W^{k,p}$ are reflexive if and only if $1 < p < \infty$.

Definition 2.1.3. Let Ω be an open subset of $\mathbb{R}^d$. We denote the dual of $W_0^{1,p}(\Omega)$ by $W^{-1,p'}(\Omega)$, and the dual of $H_0^1(\Omega)$ by $H^{-1}(\Omega)$.

Theorem 2.1.3 (Characterization of $H^{-1}(\Omega)$). *Let Ω be an open subset of $\mathbb{R}^d$ and $F \in H^{-1}(\Omega)$.*

(a) *There exist functions $f^0, f^1 \dots, f^d$ in $L^2(\Omega)$ such that*

$$\langle F, u\rangle = \int_\Omega \Big(f^0 u + \sum_{i=1}^d f^i \partial_i u\Big) \quad \text{for all } u \in H_0^1(\Omega). \tag{2.1}$$

(b) $$\|F\|_{H^{-1}(\Omega)} = \inf\left\{\left(\sum_{i=0}^d \|f^i\|^2_{L^2(\Omega)}\right)^{1/2} : f^0, f^1, \dots, f^d \text{ satisfy (2.1)}\right\}.$$

Proof. Given $F \in H^{-1}(\Omega)$, Riesz's Representation Theorem implies the existence of a unique function $f \in H_0^1(\Omega)$ such that

$$\langle F, u\rangle = (f, u)_{H_0^1(\Omega)} \quad \text{for all } v \in H_0^1(\Omega). \tag{2.2}$$

Defining $f^0 = f$, $f^i = \partial_i f$, $i = 1, \dots, d$, wet get the desired representation (2.1).

Assume that there are functions $f^0, f^1, \dots, f^d \in L^2(\Omega)$ such that

$$\langle F, v\rangle = \int_\Omega \Big(g^0 v + \sum_{i=1}^d g^i \partial_i v\Big) \quad \text{for all } v \in H_0^1(\Omega). \tag{2.3}$$

Setting $u = f$ in (2.2) and using (2.3), and Cauchy-Schwarz's inequality we get

$$\begin{aligned}\int_\Omega \sum_{i=0}^d |f^i|^2 &= \langle F, f\rangle = \int_\Omega \Big(g^0 f + \sum_{i=1}^d g^i \partial_i f\Big) \le \int_\Omega \Big(\sum_{i=0}^d |g^i|^2\Big)^{\frac12} \Big(\sum_{i=0}^d |f^i|^2\Big)^{\frac12} \\ &\le \Big(\int_\Omega \sum_{i=0}^d |g^i|^2\Big)^{\frac12} \Big(\int_\Omega \sum_{i=0}^d |f^i|^2\Big)^{\frac12};\end{aligned}$$

that is,

$$\Big(\sum_{i=0}^d \|f^i\|^2_{L^2(\Omega)}\Big)^{1/2} \le \Big(\sum_{i=0}^d \|g^i\|^2_{L^2(\Omega)}\Big)^{1/2}.$$

On the other hand, from (2.2) we have that

$$|\langle F, v\rangle| \le \Big(\sum_{i=0}^d \|f^i\|^2_{L^2(\Omega)}\Big)^{1/2} \quad \text{if } \|v\|_{H_0^1(\Omega)} \le 1.$$

Consequently, $\|F\|_{H^{-1}(\Omega)} \le \Big(\sum_{i=0}^d \|f^i\|^2_{L^2(\Omega)}\Big)^{1/2}$. But, setting $u = f/\|f\|_{H_0^1(\Omega)}$ in (2.2), we conclude that $\|F\|_{H^{-1}(\Omega)} \ge \Big(\sum_{i=0}^d \|f^i\|^2_{L^2(\Omega)}\Big)^{1/2}$, hence the result. □

Notation. We write $F = f^0 - \sum_{i=1}^d \partial_i f^i$ for general Ω and $F = -\sum_{i=1}^d \partial_i f^i$ if Ω is bounded.

Remarks. (a) If Ω is bounded we may take $f^0 = 0$.
(b) There is a similar characterizacion of $W^{-1,p'}(\Omega)$: given $F \in W^{-1,p'}(\Omega)$ there exist functions $f^0, f^1 \dots, f^d$ in $L^{p'}(\Omega)$ such that

$$\langle F, u\rangle = \int_\Omega \Big(f^0 u + \sum_{i=1}^d f^i \partial_i u\Big) \quad \text{for all } u \in W_0^{1,p}(\Omega), \tag{2.4}$$

and $\|F\|_{W^{-1,p'}(\Omega)} = \inf \left\{ \Big(\sum_{i=0}^d \|f^i\|^{p'}_{L^{p'}}\Big)^{1/p'} : f^0, f^1, \dots, f^d \text{ satisfy (2.4)} \right\}$.

2.2 Weak solutions for the Dirichlet problem

To fix ideas, we will start by considering *Dirichlet* boundary conditions, prescribing the value of the unknown u at the boundary of the domain. To make things simpler, we will prescribe the value 0, so that we have *homogeneous Dirichlet* boundary conditions.

Given $\Omega \subset \mathbb{R}^d$ open and bounded, and $f : \Omega \longrightarrow \mathbb{R}$, we consider the *homogeneous Dirichlet problem*

$$Lu = f \text{ in } \Omega, \quad u = 0 \text{ in } \partial\Omega, \tag{HDP}$$

where L denotes a second-order partial differential operator in *divergence form*,

$$Lu := -\sum_{i,j=1}^{d} \partial_j(a^{ij}(x)\partial_i u) + \sum_{i=1}^{d} b^i(x)\partial_i u + c(x)u. \tag{2.5}$$

Remark. If the coefficients $a^{ij} \in C^1(\Omega)$, then L can be written in the *non-divergence form*

$$Lu = -\sum_{i,j=1}^{d} a^{ij}(x)\partial_{ij} u + \sum_{i=1}^{d} \widetilde{b^i}(x)\partial_i u + c(x)u,$$

with $\widetilde{b^i} = b^i - \sum_j^d \partial_j a^{ij}$. Though this form may offer sometimes advantages, in this chapter we will always deal with operators L written in divergence form.

In what follows we will always assume the symmetry condition $a^{ij} = a^{ji}$, plus the boundedness of the coefficients, $a^{ij}, b^i, c \in L^\infty(\Omega)$.

Definition 2.2.1. We say the partial differential operator L is (uniformly) *elliptic* if there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2, \quad \text{for almost every } x \in \Omega \text{ and all } \xi \in \mathbb{R}^d.$$

Ellipticity thus means that for each point $x \in \Omega$ the symmetric $d \times d$ matrix $A(x) = (a^{ij}(x))_{i,j=1}^d$ is positive definite, with smallest eigenvalue greater or equal than θ.

Remark. We also have, using Cauchy-Schwarz's inequality, $\sum_{i,j} a^{ij}(x)\xi_i\xi_j \leq \Lambda|\xi|^2$, with $\Lambda = d \max_{i,j} \|a^{ij}\|_{L^\infty(\Omega)}$.

Example. If $a^{ij}(x) \equiv \delta_{ij}$, $b^i \equiv 0$ and $c \equiv 0$, then $L = -\Delta$. Solutions for the general second-order elliptic partial differential equation $Lu = 0$ are similar in many ways to harmonic functions. However, for these partial differential equations we do not have available the various explicit formulas developed for harmonic functions.

Physical interpretation. Second-order elliptic PDE generalize Laplace's and Poisson's equations. Thus, u may represent the density of some quantity, say a chemical concentration, at equilibrium whithin a region u. The second order term represents the *diffusion* of u within Ω, the coefficients a^{ij} describing the anisotropic, heterogeneous nature of the medium. In particular, $F := -ADu$ is the diffusive flux density, and the ellipticity condition implies $F \cdot Du \leq 0$, that is, the flow is from regions of higher to lower concentration. The first order term represents *transport* within Ω, and the zeroth-order term describes the local *increase* or *depletion* of the chemical (owing, say, to reactions).

Nonlinear second-order elliptic PDE also arise naturally in the calculus of variations, as the Euler-Lagrange equations of convex energy integrands (we will learn more about this in a later chapter), and in differential geometry, as expressions involving curvatures.

If both the coefficients a^{ij}, b^i, c and the right-hand side f are smooth, we can look for *classical solutions*.

Definition 2.2.2. A function $u : \overline{\Omega} \in \mathbb{R}$ is sayed to be a *classical solution* to (HDP) if $u \in C^2(\Omega) \cap C(\overline{\Omega})$ and

$$Lu(x) = f(x) \quad \text{for all } x \in \Omega, \qquad u(x) = 0 \quad \text{for all } x \in \partial\Omega.$$

Even when all the data of the problem (the domain, the coefficients and the right-hand side) are smooth, proving the existence of a classical solution may be a hard task. Moreover, if any of the data is not smooth, it may be the case that a classical solution does not exist. Hence, it may be convenient to follow a different approach. One tries to weaken the notion of solution so that is is easier to prove the existence of one. But we have to be careful not to weaken it too much, since we would like to have only one solution. In a second step, one tries to show that if the data are smooth, the solution is smooth, and hence a classical solution. Thus we consider separately two main issues: existence (and uniqueness) and regularity. The second step is usually the hardest.

Any reasonable notion of weak solution should satisfy the following two requirements:

- A classical solution should be a weak solution.
- A smooth weak solution should be a classical solution.

A third requirement, which is not always achievable, is the existence of a unique weak solution if the data are in a reasonable class.

When the operator L is in divergence form, a possibility to weaken the notion of solution is to multiply the equation satisfied by classical solutions by a test function $\varphi \in C_c^\infty(\Omega)$ and integrate by parts, so that the highest derivative in the formulation is now a first order one, instead of a second order one, and appears in an integrated form. Hence, in order for the weak notion of solution to make sense we will only need the solution to be in some first order Sobolev space. Let us apply this idea to (HDP).

Let u be a classical solution to (HDP) (all the data are assumed to be smooth). We multiply the equation by $v \in C_c^\infty(\Omega)$ and integrate by parts in the first term to obtain

$$\int_\Omega \Big(\sum_{i,j=1}^d a^{ij} \partial_i u \partial_j v + \sum_i b^i \partial_i u v + cuv \Big) = \int_\Omega fv. \tag{2.6}$$

There are no boundary terms since $v = 0$ on $\partial\Omega$. By density, the same identity holds with the smooth function v replaced by any $v \in H_0^1(\Omega)$. The resulting identity makes sense if only $u \in H_0^1(\Omega)$. We choose the space $H_0^1(\Omega)$ instead of $H^1(\Omega)$ to incorporate the boundary data. Notice that if $u, v \in H_0^1(\Omega)$ the identity makes sense requiring only a^{ij}, b and c to be in $L^\infty(\Omega)$, $f \in L^2(\Omega)$, and without any smoothness assumption on the domain.

Definition 2.2.3. Let Ω be an open, bounded subset of $\mathbb{R}^d$, $a^{i,j}, b^i, c \in L^\infty(\Omega)$, and $f \in L^2(\Omega)$.

(i) The *bilinear form* $B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ associated with the divergence form elliptic operator L given in (2.5) is defined by

$$B[u, v] := \int_\Omega \Big(\sum_{i,j=1}^d a^{ij} \partial_i u \partial_j v + \sum_i b^i \partial_i u v + cuv \Big).$$

(ii) We say that $u : \Omega \to \mathbb{R}$ is a *weak solution* to (HDP) if $u \in H_0^1(\Omega)$ and

$$B[u, v] = (f, v)_{L^2(\Omega)} \quad \text{for all } v \in H_0^1(\Omega), \tag{2.7}$$

where $(\cdot, \cdot)$ denotes the inner product in $L^2(\Omega)$.

Indentity (2.7) is sometimes called the *variational formulation* of (HDP). The origin of this terminology will become clear later in the course.

Remark. More generally, we may consider a right-hand side $f \in H^{-1}(\Omega) \equiv (H_0^1(\Omega))^*$. We say then that $u \in H_0^1(\Omega)$ is a weak solution of (HDP) if

$$B[u, v] = \langle f, v \rangle \quad \text{for all } v \in H_0^1(\Omega). \tag{2.8}$$

Remember that any $f \in H^{-1}(\Omega)$ can be characterized as $f = f^0 - \sum_{i=1}^d \partial_i f^i$ for some functions $f^0, f^1, \dots, f^d \in L^2(\Omega)$, meaning that

$$\langle f, v \rangle = \int_\Omega \left(f^0 v + \sum_{i=1}^d f^i \partial_i v \right) \quad \text{for all } v \in H_0^1(\Omega).$$

It should be clear from the way we reached the notion of weak solution that classical solutions are weak solutions. On the other hand, if a weak solution belongs to $C^2(\Omega) \cap C(\overline{\Omega})$, then, since $u \in H_0^1(\Omega)$, we have that $u(x) = 0$ for all $x \in \partial\Omega$. If the data of the problem are moreover smooth, integrating by parts in (2.6) we obtain

$$\int_\Omega (Lu - f)v = 0 \quad \text{for all } v \in H_0^1(\Omega).$$

Since $Lu - f \in C(\Omega)$, we conclude that $(Lu - f)(x) = 0$ for all $x \in \Omega$, and hence that u is a classical solution.

Example. Riesz's Representation Theorem yields existence of a weak solution for

$$-\Delta u = f \text{ in } \Omega, \qquad u = 0 \in \partial\Omega$$

where $\Omega \subset \mathbb{R}^d$ is a bounded open set.

2.3 Lax–Milgram's Theorem

We now introduce a fairly simple abstract principle from linear functional analysis which will provide under suitable hypotheses the existence and uniqueness of weak solutions to several boundary-value problems.

Along this section H is a real Hilbert space, with inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$. By $\langle \cdot, \cdot \rangle$ we will denote the pairing of H with its dual space H^*.

Theorem 2.3.1 (Lax-Milgram's Theorem). *Assume that $B : H \times H \longrightarrow \mathbb{R}$ is a bilinear mapping for which there exist constants $\alpha, \beta > 0$ such that*

- $|B[u, v]| \le \alpha \|u\| \|v\|$ *(continuity);*
- $B[u, u] \ge \beta \|u\|^2$ *(coercivity).*

Then, for all $f \in H^*$ there exists a unique $u \in H$ such that

$$B[u,v] = \langle f, v\rangle \quad \textit{for all } v \in H. \tag{2.9}$$

The application $\mathcal{T} : H^* \to H$ given by $\mathcal{T} f = u$ is an isomorphism[1] *between H^* and H.*

If B is moreover symmetric, that is, if $B[u,v] = B[v,u]$ for all $u, v \in H$, then the unique solution $u \in H$ of (2.9) *is the unique minimizer of the functional*

$$J : H \to \mathbb{R}, \qquad J(v) = \frac{1}{2} B[v,v] - \langle f, v\rangle.$$

Proof. For each fixed element $u \in H$, the mapping $v \mapsto B[u,v]$ is a bounded linear functional on H, whence Riesz's Representation Theorem asserts the existence of a unique element $w_u \in H$ satisfying $B[u,v] = (w_u, v)$ for all $v \in H$. We denote $Au = w_u$, so that

$$B[u,v] = (Au, v) \quad \text{for all } v \in H. \tag{2.10}$$

We claim that $A : H \to H$ is a bounded linear operator. Indeed, if $\lambda_1, \lambda_2 \in \mathbb{R}$ and $u_1, u_2 \in H$, we see for each $v \in H$ that

$$\begin{aligned}
(A(\lambda_1 u_1 + \lambda_2 u_2), v) &= B[\lambda u_1 + \lambda_2 u_2, v] \quad \text{by (2.10)},\\
&= \lambda_1 B[u_1, v] + \lambda_2 B[u_2, v]\\
&= \lambda_1 (Au_1, v) + \lambda_2 (Au_2, v) \quad \text{by (2.10) again},\\
&= (\lambda_1 Au_1 + \lambda_2 Au_2, v).
\end{aligned}$$

Since this equality holds for each $v \in H$, we conclude $A(\lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 Au_1 + \lambda_2 Au_2$, that is, A is linear. On the other hand, for all $u \in H$ we have

$$\|Au\|^2 = (Au, Au) = B[u, Au] \le \alpha \|u\| \|Au\|,$$

where we have used once more (2.10), and the continuity of the bilinear form. Consequently, $\|Au\| \le \alpha \|u\|$, and hence A is bounded ($\|A\| \le \alpha$).

We next observe that for all $u \in H$ we have, using the coercivity of B and Cauchy-Schwarz's inequality,

$$\beta \|u\|^2 \le B[u,u] = (Au, u) \le \|Au\| \|u\| \quad \text{for some } \beta > 0.$$

Hence $\beta \|u\| \le \|Au\|$, and since A is linear,

$$\beta \|u - v\| \le \|Au - Av\| \quad \text{for all } u, v \in H.$$

This immediately implies that A is one-to-one. On the other hand, the same estimate shows that if $(Au_k)_{k=1}^\infty$ is a Cauchy sequence, then $(u_k)_{k=1}^\infty$ is also a Cauchy sequence, hence convergent to some $u \in H$. Since A is continuous, we conclude that $Au_k \to Au$ in H. Hence $R(A)$ is closed in H.

[1]Two normed spaces X and Y are *isomorphic* if there exists a linear bijection $\mathcal{T} : X \to Y$ such that $\mathcal{T}$ and its inverse $\mathcal{T}^{-1}$ are continuous. If one of the two spaces X or Y is complete (or reflexive, separable, etc.) then so is the other space.

Let us prove now that $R(A) = H$. Otherwise, since $R(A)$ is closed in H, there would exist a nonzero element $w \in (R(A))^\perp$. But then

$$\beta\|w\|^2 \leq B[w, w] = (Aw, w) = 0,$$

a contradiction.

Given $f \in H^*$, Riesz's Representation Theorem implies the existence of a unique $w \in H$ such that $\langle f, v\rangle = (w, v)$ for all $v \in H$. Since we have already proved that $A : H \to H$ is a bijection, we have that there exists a unique $u \in H$ such that $Au = w$. For this u we have

$$B[u, v] = (Au, v) = (w, v) = \langle f, v\rangle \quad \text{for all } v \in H,$$

which is the desired existence result.

Let $u_1, u_2 \in H$ such that

$$B[u_1, v] = \langle f, v\rangle = B[u_2, v] \quad \text{for all } v \in H.$$

Then $B[u_1 - u_2, v] = 0$ for all $v \in H$. Taking $v = u_1 - u_2$ and using the coerciveness we get

$$0 = B[u_1 - u_2, u_1 - u_2] \geq \beta\|u_1 - u_2\|^2.$$

Hence $u_1 - u_2 = 0$, and we have uniqueness.

It should be clear from the previous steps that the application $\mathcal{T} : H^* \to H$ is a linear bijection. It is also bounded, hence continuous. Indeed, if $u = \mathcal{T}f$,

$$\beta\|u\|^2 \leq B[u, u] = \langle f, u\rangle \leq \|f\|_{H^*}\|u\|,$$

hence $\|\mathcal{T}f\| \leq \frac{1}{\beta}\|f\|_{H^*}$. By the Open Mapping Theorem[2] it is an isomorphism.

Assume now that B is symmetric. Let u be the unique solution to (2.9) and $v \in H$, $v \neq u$. Then

$$\begin{aligned} 0 < \tfrac{\beta}{2}\|u - v\|^2 &\leq \tfrac{1}{2}B[u - v, u - v] = \tfrac{1}{2}B[v, v] + \tfrac{1}{2}B[u, u] - B[u, v] \\ &= J(v) - J(u) + B[u, u - v] - \langle f, u - v\rangle = J(v) - J(u). \end{aligned}$$

Therefore, $J(u) < J(v)$ for all $v \in H$, $v \neq u$.

□

Remark. If the bilinear form B is symmetric we can fashion a much simpler proof by noting that$((u, v)) := B[u, v]$ is a new inner product on H, to which Riesz's Representation Theorem directly applies[3]. Consequently, Lax-Milgram's Theorem is primarily significant in that it does not require the symmetry of B.

[2]**Open Mapping Theorem.** If X and Y are Banach spaces and $A : X \to Y$ is a surjective continuous linear operator, then A is an open map (i.e. if U is an open set in X, then $A(U)$ is open in Y).

As a corollary, if $A : X \to Y$ is a bijective continuous linear operator between the Banach spaces X and Y, then the inverse operator $A^{-1} : Y \to X$ is continuous as well. This result, known as the *Bounded Inverse Theorem*, is the one we are using here.

[3]The story behind this result – the story might be completely true or completely false – is that Lax and Milgram attended a seminar where the speaker proved existence for a symmetric PDE by use of the Riesz representation theorem, and one of them asked the other if symmetry was required; in half an hour, they convinced themselves that is wasn't, giving birth to the Lax- Milgram "lemma."

2.4 Applications to the Dirichlet problem

We now go back to the Dirichlet problem (HDP), with $f \in H^{-1}(\Omega)$. Our aim is to give conditions guaranteing that the bilinear form associated to an elliptic operator L satisfies the hypotheses of Lax-Milgram's theorem. We remark that B is symmetric if and only if $b^i = 0$ for all $i = 1, \dots d$.

Proposition 2.4.1 (Energy estimates). *Let Ω be an open subset of $\mathbb{R}^d$, and let B be the bilinear form associated to the elliptic operator L in divergence form given in (2.5). We assume that the coefficients of the operator are bounded.*

(i) *There exists a constant $\alpha > 0$ such that*

$$|B(u,v)| \le \alpha \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \quad \text{for all } u, v \in H^1(\Omega).$$

(ii) *If Ω is moreover bounded, there exist constants $\beta > 0$ and $\gamma \ge 0$ such that*

$$\beta \|u\|^2_{H^1_0(\Omega)} \le B[u,u] + \gamma \|u\|^2_{L^2(\Omega)} \quad \text{for all } u \in H^1_0(\Omega).$$

Proof. (i) We readily check that

$$\begin{aligned} |B[u,v]| \;\le\; & \sum_{i,j=1}^d \|a^{ij}\|_{L^\infty(\Omega)} \|\partial_i u\|_{L^2(\Omega)} \|\partial_i v\|_{L^2(\Omega)} + \sum_{i=1}^d \|b^i\|_{L^\infty(\Omega)} \|\partial_i u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ & + \|c\|_{L^\infty(\Omega)} \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \le \alpha \|u\|_{H^1_0(\Omega)} \|v\|_{H^1_0(\Omega)}, \end{aligned}$$

for $\alpha = \sum_{i,j=1}^d \|a^{ij}\|_{L^\infty(\Omega)} + \sum_{i=1}^d \|b^i\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}$.

(ii) In view of the ellipticity condition,

$$\begin{aligned} \theta \int_\Omega |Du|^2 \;\le\; & \int_\Omega \sum_{i,j=1}^d a^{ij} \partial_i u \partial_j u = B[u,u] - \int_\Omega \Big(\sum_{i=1}^d b^i \partial_i u u + c u^2 \Big) \\ \le\; & B[u,u] + \sum_{i=1}^d \|b^i\|_{L^\infty(\Omega)} \|Du\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} + \|c\|_{L^\infty(\Omega)} \|u\|^2_{L^2(\Omega)}. \end{aligned}$$

Then, using Young's inequality[4] with ε

$$\|Du\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \le \varepsilon \|Du\|^2_{L^2(\Omega)} + \frac{1}{4\varepsilon^2} \|u\|^2_{L^2(\Omega)}.$$

Therefore, taking $\varepsilon > 0$ such that $\varepsilon \sum_{i=1}^d \|b^i\|_{L^\infty(\Omega)} \le \frac{\theta}{2}$, we arrive to

$$\frac{\theta}{2} \|Du\|^2_{L^2(\Omega)} \le B[u,u] + M \|u\|^2_{L^2(\Omega)}$$

for some appropriate constant M. In addition, we recall from Poincaré's inequality that $\|Du\|_{L^2(\Omega)} \ge c \|u\|_{L^2(\Omega)}$ for some constant $c > 0$, from where the result easily follows for appropriate constants $\beta > 0$ and $\gamma \ge 0$. □

[4] *Young's inequality with ε*, also known as *Peter-Paul inequality*, valid for every $\varepsilon > 0$, says that $ab \le \varepsilon a^2 + \frac{b^2}{4\varepsilon}$. The second name refers to the fact that tighter control of the first term is achieved at the cost of losing some control of the second one – one must "rob Peter to pay Paul". This inequality follows immediately from Young's inequality with $p = q = 2$.

Remark. The inequality in (ii) is known as Gårding's inequality.

Corollary 2.4.2. *Let Ω be a bounded open subset of $\mathbb{R}^d$. There is a number $\gamma \geq 0$ such that for each $\mu \geq \gamma$ and all $f \in H^{-1}(\Omega)$ there exists a unique weak solution $u \in H_0^1(\Omega)$ of the boundary-value problem*

$$Lu + \mu u = f \text{ in } \Omega, \quad u = 0 \text{ in } \partial\Omega.$$

Proof. Take γ from Proposition 2.4.1. Let $\mu \geq \gamma$ and define then the bilinear form

$$B_\mu[u, v] = B[u, v] + \mu(u, v)_{L^2(\Omega)}, \quad u, v \in H_0^1(\Omega),$$

which corresponds to the operator $L_\mu u := Lu + \mu u$. Then B_μ satisfies the hypotheses of Lax-Migram's Theorem, hence the result. □

Examples. (a) Let Ω be an arbitrary open subset of $\mathbb{R}^d$. Given $f \in H^{-1}(\Omega)$ and $c \in L^\infty(\Omega)$ we consider the elliptic problem

$$-\Delta u + cu = f \text{ in } \Omega, \quad u = 0 \text{ in } \partial\Omega. \tag{2.11}$$

The associated bilinear form $B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is given by

$$B[u, v] = \int_\Omega (Du \cdot Dv + cuv).$$

We already know from Proposition 2.4.1 that B is continuous. On the other hand, since $B[u, u] = \int_\Omega (|Du|^2 + cu^2)$, if $c_0 := \operatorname{essinf} c > 0$, then B is also coercive, and by Lax-Milgram's theorem the problem (2.11) has a unique weak solution.

If Ω is moreover bounded, we can relax the hypothesis on c. Indeed, thanks to Poincaré's inequality, we know that $\left(\int_\Omega |Du|^2\right)^{1/2}$ is equivalent to the $H_0^1(\Omega)$ norm when Ω is bounded. Therefore, if $c \geq 0$,

$$B[u, u] \geq \int_\Omega |Du|^2 \geq \beta \|u\|^2_{H_0^1(\Omega)}$$

for some $\beta > 0$, and B is coercive.

We can go even further. Let $\mu = \mu(\Omega) > 0$ be the best constant in Poincaré's inequality, so that

$$\|u\|_{L^2(\Omega)} \leq \mu \|Du\|_{L^2(\Omega)}.$$

Assume that $0 > c_0 > -1/\mu^2$. Then

$$B[u, u] \geq \int_\Omega (|Du|^2 + c_0 u^2) \geq \|Du\|^2_{L^2(\Omega)} + c_0 \mu^2 \|Du\|^2_{L^2(\Omega)} = (1 + c_0 \mu^2) \|Du\|^2_{L^2(\Omega)},$$

and B is again coercive.

Note that B is in this case symmetric. Hence, the unique weak solution of problem (2.11) given by Lax-Milgram's method under the conditions on c that guarantee that B is coercive is the unique minimizer of the functional $J : H_0^1(\Omega) \to \mathbb{R}$ given by

$$J(u) = \frac{1}{2} \int_\Omega (|Du|^2 + cu^2) - \langle f, u \rangle.$$

(b) We consider now a bounded open set Ω, and an elliptic operator with associated bilinear form

$$B[u,v] = \int_\Omega \Big(Du\cdot Dv + \sum_{i=1}^d b^i \partial_i u v \Big).$$

Continuity is once more provided by Proposition 2.4.1. On the other hand,

$$B[u,u] \geq \|Du\|^2_{L^2(\Omega)} - \sum_{i=1}^d \|b^i\|_{L^\infty(\Omega)} \|\partial_i u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}.$$

Let $\nu := \max_{1\leq i\leq d} \|b^i\|_{L^\infty}$. Since, by Peter-Paul inequality,

$$\|u\|_{L^2(\Omega)} \|\partial_i u\|_{L^2(\Omega)} \leq \varepsilon \|u\|^2_{L^2(\Omega)} + \frac{\|\partial_i u\|^2_{L^2(\Omega)}}{4\varepsilon}$$

for any $\varepsilon > 0$, then

$$\begin{aligned} -\sum_{i=1}^d \|b^i\|_{L^\infty(\Omega)} \|\partial_i u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} &\geq -\nu \sum_{i=1}^d \Big(\varepsilon \|u\|^2_{L^2(\Omega)} + \frac{\|\partial_i u\|^2_{L^2(\Omega)}}{4\varepsilon} \Big) \\ &= -\nu d \varepsilon \|u\|^2_{L^2(\Omega)} - \frac{\nu}{4\varepsilon} \|Du\|^2_{L^2(\Omega)}. \end{aligned}$$

Therefore, using Poincaré's inequality,

$$B[u,u] \geq \Big(1 - \frac{\nu}{4\varepsilon}\Big) \|Du\|^2_{L^2(\Omega)} - \nu d \varepsilon \|u\|^2_{L^2(\Omega)} \geq \Big(1 - \frac{\nu}{4\varepsilon} - \nu d\varepsilon \mu^2\Big) \|Du\|^2_{L^2(\Omega)}.$$

We take now $\varepsilon > 0$ maximizing $1 - \frac{\nu}{4\varepsilon} - \nu d \varepsilon \mu^2$, that is, $\varepsilon = \frac{1}{2\sqrt{d}\mu}$, which gives a value $1 - \nu\sqrt{d}\mu$. Therefore, if $\nu < \frac{1}{\sqrt{d}\mu}$ we obtain the desired coercivity.

(c) We consider now the case in which the matrix a^{ij} is not the identity. Let

$$B[u,v] = \int_\Omega \Big(\sum_{i,j} a^{ij} \partial_i u \partial_j v + cuv \Big).$$

Applying the ellipticity we obtain

$$B[u,u] \geq \int_\Omega \Big(\lambda |\nabla u|^2 + cu^2 \Big),$$

from where we can argue as in example (a).

(c) Let L be such that its associated bilinear form B is coercive. Let $\overline{b^i} \in \mathbb{R}$, $i = 1, \dots, d$, and $\overline{L}u := Lu + \sum_{i=1}^d \overline{b^i} \partial_i u$. Then L makes $\overline{B}$ coercive. Indeed, if $u \in C^\infty_{\rm c}(\Omega)$, then

$$(\overline{B} - B)[u,u] = \sum_{i=1}^d \overline{b^i} \int_\Omega \partial_i u u = \sum_{i=1}^d \overline{b^i} \int_\Omega \partial_i \left(\frac{u^2}{2} \right) = 0,$$

so that $\overline{B}[u,u] = B[u,u]$. The result for general $u \in H^1_0(\Omega)$ follows by approximation.

(d) We consider now the system

$$-\Delta u_1 + u_2 = f_1 \text{ in } \Omega, \quad -\Delta u_2 - u_1 = f_2 \text{ in } \Omega, \quad u_1 = u_2 = 0 \text{ in } \partial\Omega.$$

We will work in the Hilbert space $H = H_0^1(\Omega)\times H_0^1(\Omega)$. We multiply the first equation by $v_1 \in H_0^1(\Omega)$ and the second one by $v_2 \in H_0^1(\Omega)$, integrate by parts and add the equations to obtain

$$B\Big[\begin{pmatrix}u_1\\u_2\end{pmatrix},\begin{pmatrix}v_1\\v_2\end{pmatrix}\Big] = \Big\langle\begin{pmatrix}f_1\\f_2\end{pmatrix},\begin{pmatrix}v_1\\v_2\end{pmatrix}\Big\rangle,$$

where

$$B\left[\begin{pmatrix}u_1\\u_2\end{pmatrix},\begin{pmatrix}v_1\\v_2\end{pmatrix}\right] = \int_\Omega (Du_1\cdot Dv_1 + u_2v_1 + Du_2\cdot Dv_2 - u_1v_2),$$
$$\left\langle\begin{pmatrix}f_1\\f_2\end{pmatrix},\begin{pmatrix}v_1\\v_2\end{pmatrix}\right\rangle = \int_\Omega \Big(f_1v_1 + f_2v_2\Big).$$

Since, thanks to Poincaré's inequality,

$$\begin{aligned}B\left[\begin{pmatrix}u_1\\u_2\end{pmatrix},\begin{pmatrix}u_1\\u_2\end{pmatrix}\right] &= \int_\Omega (|Du_1|^2+|Du_2|^2) = \|Du_1\|^2_{L^2(\Omega)} + \|Du_2\|^2_{L^2(\Omega)}\\ &\geq \beta\|u_1\|^2_{H_0^1(\Omega)} + \beta\|u_2\|^2_{H_0^1(\Omega)} = \beta\left\|\begin{pmatrix}u_1\\u_2\end{pmatrix}\right\|_H^2\end{aligned}$$

for some $\beta>0$, the bilinear form is coercive. It is also continuous, since

$$\begin{aligned}\left|B\left[\begin{pmatrix}u_1\\u_2\end{pmatrix},\begin{pmatrix}v_1\\v_2\end{pmatrix}\right]\right| &\leq \|Du_1\|_{L^2(\Omega)}\|Dv_1\|_{L^2(\Omega)} + \|u_2\|_{L^2(\Omega)}\|v_1\|_{L^2(\Omega)}\\ &\quad + \|Du_2\|_{L^2(\Omega)}\|Dv_2\|_{L^2(\Omega)} + \|u_1\|_{L^2(\Omega)}\|v_2\|_{L^2|(\Omega)}\\ &\leq 4\left\|\begin{pmatrix}u_1\\u_2\end{pmatrix}\right\|_H\left\|\begin{pmatrix}v_1\\v_2\end{pmatrix}\right\|_H.\end{aligned}$$

(e) We consider now a fourth order problem, namely

$$-\Delta^2 u = f \text{ in } \Omega,\quad u = \partial_\nu u = 0 \text{ in } \partial\Omega,$$

where $\partial_\nu u = Du\cdot\nu$ (directional derivative in the direction of the exterior unit normal ν). Sea $H_0^2(\Omega)$ be the closure of $C_c^\infty(\Omega)$ in H^2. Note that if $u\in H_0^2(\Omega)$ then u and $\partial_\nu u$ are both 0 almost everywhere in $\partial\Omega$ in the sense of traces.

Multiply the equation by $v\in H_0^2(\Omega)$ and integrate by parts. Then

$$\begin{aligned}\int_\Omega fv &= \int_\Omega(\Delta^2u)v = -\int_\Omega D(\Delta u)\cdot Dv = -\sum_{i,j=1}^d\int_\Omega \partial_{iij}u\partial_j v\\ &= \sum_{i,j=1}^d\int_\Omega \partial_{ii}u\partial_{jj}v = \int_\Omega \Delta u\Delta v.\end{aligned}$$

Thus, $B[u,v] = \int_\Omega fv$, where $B[u,v] = \int_\Omega \Delta u\Delta v$. To check that the bilinear form is coercive we observe that

$$\begin{aligned}\|\Delta u\|^2_{L^2(\Omega)} &= \int_\Omega\sum_{i,j=1}^d \partial_{ii}u\partial_{jj}u = -\int_\Omega\sum_{i,j}\partial_i u\partial_{ijj}u = \int_\Omega\sum_{i,j=1}^d\partial_{ij}u\partial_{ij}u\\ &= \int_\Omega\Big(\sum_{i,j=1}^d\partial_{ij}u\Big)^2 = \|D^2u\|^2_{L^2}.\end{aligned}$$

Coercivity then follows from Poincaré's inequality, since $\|u\|_{L^2(\Omega)} \leq c\|Du\|_{L^2(\Omega)}$ and $\|Du\|_{L^2(\Omega)} \leq C_1\|D^2u\|_{L^2(\Omega)}$.

2.5 Other boundary conditions

In this section we will give through a series of examples an idea of how to deal with boundary conditions different from Dirichlet homogeneous conditions.

Non-homogeneous Dirichlet conditions. We consider the problem

$$Lu = f \text{ in } \Omega, \quad u = g \text{ in } \partial\Omega,$$

where L is a second order elliptic operator in divergence form. If there exists a function $\overline{u} \in H^1(\Omega)$ such that $T\overline{u} = g$, then the change of variables $v = u - \overline{u}$ transforms the problem into one with homogeneous Dirichlet boundary conditions.

$$Lv = f - L\overline{u} \text{ in } \Omega, \quad v = 0 \text{ in } \partial\Omega. \tag{2.12}$$

Notice that $L\overline{u} \in H^{-1}(\Omega)$. Hence, if the bilinear form associated with L is continuous and coercive, and $f \in H^{-1}(\Omega)$, Lax-Milgram's theorem guarantees the existence of a unique weak solution v to (2.12), which in turn provides us with a unique weak solution $u = \overline{u} + v$ to our original problem (this is in fact a way to define the notion of weak solution for the original problem).

Thus, we have to face the problem of finding $\overline{u} \in H^1(\Omega)$ such that $T\overline{u} = g$. Let us remind that the trace operator $T : H^1(\Omega) \to L^2(\partial\Omega)$ is linear and bounded. However, it is not onto. Its image is the Sobolev space of fractional order $H^{1/2}(\partial\Omega)$, that can be characterized as

$$H^{1/2}(\partial\Omega) = \left\{ v \in L^2(\partial\Omega) : \int_{\partial\Omega}\int_{\partial\Omega} \frac{|v(x) - v(y)|^2}{|x-y|^d} d\sigma(x) d\sigma(y) < \infty \right\}.$$

Homogeneous Neumann conditions. Neumann boundary conditions consist in prescribing the exterior normal derivative (which in applications corresponds to a flux) at the boundary.

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with smooth boundary. We consider the *homogeneous* Neumann problem

$$-\Delta u + cu = f \text{ in } \Omega, \quad \partial_\nu u = 0 \text{ in } \partial\Omega, \tag{2.13}$$

Let $c, f \in C(\Omega)$. A classical solution to (2.13) is a function $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ satisfying $-\Delta u + cu = f$ everywhere in Ω and $\partial_\nu u = 0$ everywhere in $\partial\Omega$.

Given a classical solution, if we multiply the equation by a test function $v \in C^1(\overline{\Omega})$ and integrate by parts we obtain

$$(f, v) = \int_\Omega (Du \cdot Dv + cuv).$$

By density, this equality will hold for all $v \in H^1(\Omega)$. In order for this expression to make sense we only need $u \in H^1(\Omega)$. Therefore, we define a weak solution to (2.13) as a function $u \in H^1(\Omega)$ such that $B[u, v] = (f, v)$ for all $v \in H^1(\Omega)$ where $B : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ is the bilinear form given by $B[u, v] = \int_\Omega (Du \cdot Dv + cuv)$. Let us remark that we have proved that a classical solution is a weak solution.

If we deal with weak solutions we can relax the conditions on the data f and c. Indeed, in order for the weak formulation to make sense we only need $c \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$. In fact, we may even take $f \in (H^1(\Omega))^*$, in which case the weak formulation consists in finding $u \in H^1(\Omega)$ such that $B[u,v] = \langle f, v\rangle$ for all $v \in H^1(\Omega)$.

We have already seen in Section 3 that if $\operatorname{essinf}_\Omega c > 0$, then B is continuous and coercive. Hence, by Lax-Milgram's Theorem, the problem has a unique weak solution for any data $f \in (H^1(\Omega))^*$ and $c \in L^\infty(\Omega)$ satisfying the above mentioned positivity condition. The unique weak solution is the unique minimizer in $H^1(\Omega)$ of the functional

$$J(v) = \frac{1}{2}\int_\Omega (|Dv|^2 + cv^2 - fv).$$

Let us remark a big difference between homogeneous Dirichlet and Neumann boundary conditions. The first ones are *essential* and have to be forced by asking the solutions to belong to the functional space $H_0^1(\Omega)$), while Neumann conditions are *natural*: they do not have to be imposed in the functional space since minimization in $H^1(\Omega)$ already tries to make derivatives small,

Non-homogeneous Neumann conditions. Let $\Omega \subset \mathbb{R}^d$ be again a bounded domain with smooth boundary. We consider the *non-homogeneous* Neumann problem

$$-\Delta u + cu = f \text{ in } \Omega, \quad \frac{\partial u}{\partial \nu} = g \text{ in } \partial\Omega. \tag{2.14}$$

If f, g and c are smooth, then a classical solution to (2.14) is a function $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ satisfying $-\Delta u + cu = f$ everywhere in Ω and $\partial_\nu u = g$ everywhere in $\partial\Omega$.

Given a classical solution, if we multiply the equation by a test function $v \in C^1(\overline{\Omega})$ and integrate by parts we obtain

$$\int_\Omega fv = \int_\Omega (Du \cdot Dv + cuv) - \int_{\partial\Omega} gv.$$

By density, this equality will hold for all $v \in H^1(\Omega)$. In order for this expression to make sense we only need $u \in H^1(\Omega)$. Therefore, we define a weak solution to (2.13) as a function $u \in H^1(\Omega)$ such that $B[u,v] = F(v)$ for all $v \in H^1(\Omega)$ where $B : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ is the bilinear form given by $B[u,v] = \int_\Omega (Du \cdot Dv + cuv)$ and $F : H^1(\Omega) \to \mathbb{R}$ is the linear functional given by $F(v) = \int_\Omega fv + \int_{\partial\Omega} gv$. Let us remark that we have proved that a classical solution is a weak solution.

If we deal with weak solutions we can relax the conditions on the data f, c and g. Indeed, in order for the weak formulation to make sense we only need $c \in L^\infty(\Omega)$, $f \in L^2(\Omega)$ (or even $f \in (H^1(\Omega))^*$), and $g \in L^2(\partial\Omega)$.

If $\operatorname{essinf}_\Omega c > 0$, then B is continuous and coercive. Hence, to have existence and uniqueness it is enough to check that F is continuous. This is an easy consequence of the trace inequality,

$$|F(v)| \le \|f\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)}\|v\|_{L^2(\partial\Omega)} \le (\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)})\|v\|_{H^1(\Omega)}.$$

Neumann condition for general elliptic operators. Once more Ω is a bounded open subset of $\mathbb{R}^d$ with smooth boundary. Let L be a general elliptic operator in divergence form

$$Lu = -\sum_{i,j=1}^d \partial_j(a^{ij}(x)\partial_i u) + \sum_{i=1}^d b^i(x)\partial_i u + c(x)u.$$

As we have already mentioned, Neumann conditions consist in prescribing the flux at the boundary. The flux has to do with the operator, and in this case is given by $\partial_{\nu_A} u = \sum_{i,j=1}^d a^{ij}\partial_i u \nu_j$. Thus, we consider the problem

$$Lu = f \text{ in } \Omega, \quad \partial_{\nu_A} u = 0 \text{ in } \partial\Omega.$$

A classical solution satisfies

$$B[u,v] = (f,v)_{L^2(\Omega)} \quad \text{for all } v \in H^1(\Omega) \tag{2.15}$$

(we are assuming $f \in L^2(\Omega)$, though more general right-hand sides are allowed), where

$$B[u,v] := \int_\Omega \Big(\sum_{i,j=1}^d a^{ij}\partial_i u \partial_j v + \sum_i b^i \partial_i u v + cuv \Big).$$

To check this, multiply the equation by $v \in C^1(\overline{\Omega})$, integrate by parts and use the boundary condition, and then use a density argument. A weak solution is a function $u \in H^1(\Omega)$ satisfying (2.15).

Neumann conditions for a non-coercive bilinear form. Let us now consider the Neumann problem

$$-\Delta u = f \text{ in } \Omega, \quad \partial_\nu u = 0 \text{ in } \partial\Omega,$$

where $\Omega \subset \mathbb{R}^d$ is bounded, connected and smooth, and $f \in L^2(\Omega)$. The bilinear form associated to the operator, $B[u,v] = \int_\Omega Du \cdot Dv$ is not coercive in this case.

Remarks. (a) Weak solutions, if they exist, are not unique. Indeed, if we add a constant to a solution we get another solution. The same is true for classical solutions.

(b) Any two weak solutions differ by a constant. Indeed, if $u_1, u_2 \in H^1(\Omega)$ solve $B[u,v] = F(v)$ for all $v \in H^1(\Omega)$, then $B[u_1 - u_2, v] = 0$ for all $v \in H^1(\Omega)$. Taking $v = u_1 - u_2$ we get that $u_1 - u_2$ is a constant.

(c) If u is a classical solution, then, using Gauss' Divergence Theorem,

$$\int_\Omega f = -\int_\Omega \Delta u = -\int_{\partial\Omega} \partial_\nu u = 0.$$

Thus, there is no hope to obtain a classical solution unless the compatibility condition $\int_\Omega f = 0$ is satisfied.

In order to eliminate the lack of uniqueness due to the possibility of adding constants, we will work in the subspace $H = \{u \in H^1(\Omega) : \int_\Omega u = 0\}$. Observe that adding a constant to an element of H takes us out of H.

The set H is the kernel of the linear and bounded application $G : H^1(\Omega) \to \mathbb{R}$ given by $G(u) = \int_\Omega u$. Hence H is a closed subspace of $H^1(\Omega)$, and therefore a Hilbert space. Is $B[u,v] = \int_\Omega Du \cdot Dv$ coercive in H? The answer is yes, thanks to Poincaré-Wirtinger's inequality. Indeed, since $\|u - \fint_\Omega u\|_{L^2(\Omega)} \le C\|Du\|_{L^2(\Omega)}$ for every $u \in H^1(\Omega)$, then $\|u\|_{L^2(\Omega)} \le C\|Du\|_{L^2(\Omega)}$ for all $u \in H$, and we conclude that $\|Du\|_{L^2(\Omega)}$ is a norm in H equivalent to the $H^1(\Omega)$-norm. Hence, B is coercive in H. The continuity of B is trivi

On the other hand, given $f \in L^2(\Omega)$, the functional $F : H \to \mathbb{R}$ given by $F(v) : \int_\Omega fv$ is linear and bounded. Hence, by Lax-Milgram's theorem, there exists a unique function $u \in H$ such that $B[u,v] = F(v)$ for all $v \in H$.

Exercise. If $f \in L^2(\Omega)$ there is a weak solution even if f does not satisfy the compatibility condition. What problem are we solving in this case?

Robin conditions. In this case we prescribe a value at the boundary for a certain combination of u and its normal derivative,

$$-\Delta u = f \text{ in } \Omega, \quad u + \partial_\nu u = 0 \text{ in } \partial\Omega.$$

Multiplying by a test function $v \in H^1(\Omega)$ and integrating by parts we get

$$\int_\Omega fv = \int_\Omega Du \cdot Dv - \int_{\partial\Omega} \partial_\nu uv.$$

Thus, if we impose the boundary condition we arrive to $B[u,v] = \int_\Omega fv$ for all $v \in H^1(\Omega)$, where $B[u,v] = \int_\Omega Du \cdot Dv + \int_{\partial\Omega} uv$. Note that the boundary condition appears now in the bilinear form. The continuity of B follows easily from the trace inequality, and the coercivity from the inequality

$$\|u\|_{L(\Omega)} \leq (\|u\|_{L^p(\Gamma)} + \|Du\|_{L^p(\Omega)}),$$

(with $p = 2$, and $\Gamma = \partial\Omega$) valid for functions in $W^{1,p}(\Omega)$, if $\Gamma \subset \partial\Omega$ has nonzero ($(d-1)$-dimensional) measure. This inequality was proved in problem 2b of worksheet number 4.

Mixed conditions. In this case we prescribe Dirichlet conditions in part of the boundary and Neumann conditions in the complement,

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ in } \Gamma_1, \quad \frac{\partial u}{\partial \nu} = 0 \text{ in } \Gamma_2$$

where $\partial\Omega = \Gamma_1 \cup \Gamma_2$, $\Gamma_1 \cap \Gamma_2 = \emptyset$, Γ_1 open in $\partial\Omega$ and $\Gamma_1 \neq \emptyset$.

Dirichlet conditions are essential, and have to be imposed through the functional space. Thus, we consider $H = \{u \in H^1(\Omega) : u = 0 \text{ a.e. in } \Gamma_1\}$, which is a closed subspace of $H^1(\Omega)$, and hence a Hilbert space. Let us check this assertion more carefully.

Let $(u_k)_{k=1}^\infty$ be a sequence in H converging in $H^1(\Omega)$ to some function u. From the trace inequality we have

$$\|Tu\|_{L^2(\Gamma_1)} \leq \|Tu_k - Tu\|_{L^2(\Gamma_1)} \leq \|T(u_k - u)\|_{L^2(\partial\Omega)} \leq \|u_k - u\|_{H^1(\Omega)} \to 0.$$

Hence $Tu = 0$ a.e. in Γ_1, and therefore $u \in H$, which proves the claim.

As for the weak notion of solution, multiplying by a test function $v \in H$, integrating by parts, and imposing the Neumann boundary condition plus the fact that $v \in H$ to get rid of the boundary term, we get that $B[u,v] = \int_\Omega fv$ for all $v \in H$. Coercivity comes again from problem 2b in worsheet 4.

Periodic boundary conditions. To simplify things, let us present this kind of boundary conditions in a one-dimensional setting,

$$-u'' + u = f \text{ in } \Omega = (0,1), \quad u(0) = u(1), \quad u'(0) = u'(1).$$

Take a test function $v \in H^1(\Omega)$. After integration by parts we get

$$\int_0^1 fv = \int_0^1 (u'v' + uv) - u'(1)v(1) + u'(0)v(0) = \int_0^1 (u'v' + uv) - u'(1)(v(1) - v(0)).$$

Thus, if v belongs to the space $H = \{u \in H^1(\Omega) : u(0) = u(1)\}$, then u satisfies $B[u,v] = \int_\Omega fv$, where $B[u,v] = \int_0^1 (u'v' + uv)$. It is trivial to check that B is continuous and coercive. Hence, if we prove that H is a closed subspace of $H^1(\Omega)$, Lax-Milgram's theorem will yield the existence of a unique solution $u \in H$.

Let $(u_k)_{k=1}^\infty$ be a sequence in H converging in $H^1(\Omega)$ to some function u. Then

$$\begin{aligned}|u(1) - u(0)| &= |u(1) - u_k(1) + u_k(1) - u_k(0) + u_k(0) - u(0)| \\ &= |u(1) - u_k(1) + u_k(0) - u(0)| = \left|\int_0^1 (u - u_k)'\right| \\ &\le \int_0^1 |u - u_k|' \le \|(u - u_k)'\|_{L^2(\Omega)} \le \|u - u_k\|_{H^1(\Omega)} \to 0.\end{aligned}$$

Hence $u(1) = u(0)$, which proves the claim.

Let us remark that the conditions on the function had to be included in the Hilbert space (they are essential), in contrast with the conditions on the derivatives (which are natural).

2.6 Regularity

Let us go back again to the homogeneous Dirichlet problem (HDP). Along this section we always assume that Ω is $\mathbb{R}^d$, $\mathbb{R}^d_+ = \mathbb{R}^{d-1} \times \mathbb{R}_+$ or a bounded domain. We assume moreover that L is a uniformly elliptic operator having the divergence form (2.5), and that its associated bilinear form is continuous and coercive, so that we have a unique weak solution $u \in H_0^1(\Omega)$ of problem (HDP) if $f \in L^2(\Omega)$. Is it true that if the data of the problem are smooth, then u is smooth? The answer is *yes*. To make this statement more precise, we will as necessary make various additional assumptions about the smoothness of the coefficients a^{ij}, b^i, and c, the right-hand side function f, and the domain Ω.

Let us first give conditions to have H^2 regularity.

Theorem 2.6.1 (H^2 **regularity).** *Assume*

$$a^{ij} \in C^1(\overline{\Omega}), \quad b^i, c \in L^\infty(\Omega), \quad f \in L^2(\Omega), \quad \partial\Omega \in C^2.$$

If $u_0 \in H_0^1(\Omega)$ is the unique weak solution to (HDP)*, then $u \in H^2(\Omega)$ and we have the estimate $\|u\|_{H^2(\Omega)} \le C\|f\|_{L^2(\Omega)}$, the constant C depending only on Ω and the coefficients of L.*

Let us explain the heuristics of the proof in the somewhat simpler case of problem

$$-\Delta u + u = f \quad \text{in } \mathbb{R}^d.$$

Remember that a weak solution of this problem is a function $u \in H^1(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} (Du \cdot Dv + uv) = \int_{\mathbb{R}^d} fv \quad \text{for all } v \in H^1(\mathbb{R}^d).$$

Take $v = \partial_{ii} u$ as a test function. Then, assuming that all the computations are justified,

$$\int_{\mathbb{R}^d} f\partial_{ii}u = \int (Du \cdot D\partial_{ii}u + u\partial_{ii}u) = -\int_{\mathbb{R}^d} (|D\partial_i u|^2 + (\partial_i u)^2) = -\|\partial_i u\|^2_{H^1(\mathbb{R}^d)}.$$

Therefore

$$\|\partial_i u\|^2_{H^1(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)} \|\partial_{ii} u\|_{L^2(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)} \|\partial_i u\|_{H^1(\mathbb{R}^d)},$$

and we conclude that $\|\partial_i u\|_{H^1(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)}$.

The problem with the above argument is that in order to use $\partial_{ii} u$ as a test function, we would need $u \in H^3$, a regularity that is not available. To circunvent this difficulty, we follow the *method of translations of L. Nirenberg*, which takes as test functions *difference quotients* instead of derivatives.

Definition 2.6.1. Let $u \in L^1_{\text{loc}}(\Omega)$ and $V \subset\subset \Omega$.

(a) The i^{th}-*difference quotient* of u of size h is

$$D_i^h u(x) = \frac{u(x + he_i) - u(x)}{h}, \quad i \in \{1, \dots, d\},$$

for $x \in V$ and $h \in \mathbb{R}$, $0 < |h| < \operatorname{dist}(V, \partial\Omega)$.

(b) $D^h u := (D_1^h u, \dots, D_d^h u)$.

The key point is that Sobolev spaces can be characterized in terms of incremental quotients.

Theorem 2.6.2 (Difference quotients and weak derivatives).

(i) Let $p \in [1, \infty)$. For each $V \subset\subset \Omega$ there exists a constant C such that

$$\|D^h u\|_{L^p(V)} \le C \|Du\|_{L^p(\Omega)} \quad \textit{for all } u \in W^{1,p}(\Omega) \textit{ and } h \in \mathbb{R}^d,\ 0 < |h| < \frac{1}{2} \operatorname{dist}(V, \partial\Omega).$$

(ii) Let $u \in L^p(V)$, $p \in (1, \infty)$. If there exists a constant C such that

$$\|D^h u\|_{L^p(V)} \le C \quad \textit{for all } h \in \mathbb{R},\ 0 < |h| < \frac{1}{2} \operatorname{dist}(V, \partial\Omega), \tag{2.16}$$

then $u \in W^{1,p}(V)$, and $\|Du\|_{L^p(V)} \le C$.

Remark. (ii) is false for $p = 1$ (find a counterexample!).

Proof. (i) Assume u is smooth. For each $x \in V$, $i = 1, \dots, d$, and $0 < |h| < \frac{1}{2} \operatorname{dist}(V, \partial\Omega)$ we have

$$u(x + he_i) - u(x) = \int_0^1 \frac{d}{dt}(u(x + the_i))\, dt = h \int_0^1 \partial_i u(x + the_i)\, dt,$$

so that

$$\frac{|u(x + he_i) - u(x)|}{|h|} \le \int_0^1 |Du(x + the_i)|\, dt.$$

Therefore,

$$\begin{aligned}\int_V |D^h u|^p &\le C \sum_{i=1}^d \int_V \int_0^1 |Du(x + the_i)|^p\, dt dx = C \sum_{i=1}^d \int_0^1 \int_V |Du(x + the_i)|^p\, dx dt \\ &\le C \sum_{i=1}^d \int_0^1 \|Du\|^p_{L^p(\Omega)} \le Cd \|Du\|^p_{L^p(\Omega)}.\end{aligned}$$

This estimate holds for all smooth u, and also by an approximation argument (check it!), for all $u \in W^{1,p}(\Omega)$.

(ii) Assume now that (2.16) holds for some constant C. Choose $i \in \{1, \dots, d\}$, $\phi \in C_c^\infty(V)$ and note for small enough h that

$$\int_V u(x) \left(\frac{\phi(x + he_i) - \phi(x)}{h} \right) dx = - \int_V \left(\frac{u(x) - u(x - he_i)}{h} \right) \phi(x)\, dx.$$

This equality, that follows just from a change of variables, can be written as

$$\int_V u(D_i^h \varphi) = - \int_V (D_i^{-h} u)\phi,$$

which is known as the "integration-by-parts" formula for difference quotients. Estimate (2.16) implies

$$\sup_h \|D_i^{-h} u\|_{L^p(V)} < \infty.$$

Therefore, since $p \in (1, p)$ and hence $L^p(V)$ is reflexive, there exist a function $v_i \in L^p(V)$ and a subsequence $(h_k)_{k=1}^\infty$, $h_k \to 0$, such that

$$D_i^{-h_k} u \rightharpoonup v_i \quad \text{weakly in } L^p(V).$$

But then

$$\int_V u\partial_i\phi = \int_\Omega u\partial_i\phi = \lim_{h_k \to 0} \int_\Omega u D_i^{h_k}\phi = - \lim_{h_k \to 0} \int_V (D_i^{-h_k} u)\phi = - \int_V v_i\phi = - \int_\Omega v_i\phi.$$

Thus, $v_i = \partial_i u$, $i = 1, \dots, d$, in the weak sense, and so $\partial_i u \in L^p(V)$, $i = 1, \dots, d$. As $u \in L^p(V)$, we deduce therefore that $u \in W^{1,p}(V)$. □

Proof of Theorem 2.6.1. In order to simplify the exposition, we will only give the proof for the equation $-\Delta u + u = f$, since this case already contains the main ideas, while keeping technicalities to a minimum. On the other hand we will only consider the domains $\Omega = \mathbb{R}^d$ and $\Omega = \mathbb{R}_+^d$. Once these two cases are well understood, other domains can be treated using a partition of unity and straightening the boundary.

The idea is to use as test function in the weak formulation the second difference quotient

$$D_i^{-h} D_i^h u(x) = \frac{u(x + he_i) + u(x - he_i) - 2u(x)}{h^2}, \quad h \neq 0,$$

which is an approximation of $\partial_{ii} u(x)$ if u is smooth.

Les us start by considering the case $\Omega = \mathbb{R}^d$. Taking $v = D_i^{-h} D_i^h u$ in the weak formulation we get

$$\int_{\mathbb{R}^d} f D_i^{-h} D_i^h u = \int_{\mathbb{R}^d} \Big(DuD(D_i^{-h} D_i^h u) + u D_i^{-h} D_i^h u \Big).$$

Integrating by parts the difference quotients, we obtain

$$\int_{\mathbb{R}^d} f D_i^{-h} D_i^h u = - \int_{\mathbb{R}^d} \Big(|D(D_i^h u)|^2 + |D_i^h|^2 \Big).$$

Thus,

$$\begin{aligned}\|D_i^h u\|^2_{H^1(\mathbb{R}^d)} &= -\int_\Omega f D_i^{-h} D_i^h u \le \|f\|_{L^2(\mathbb{R}^d)} \|D_i^{-h} D_i^h u\|_{L^2(\mathbb{R}^d)} \\ &\le C\|f\|_{L^2(\mathbb{R}^d)} \|D(D_i^h u)\|_{L^2(\mathbb{R}^d)} \le C\|f\|_{L^2(\mathbb{R}^d)} \|D_i^h u\|_{H^1(\mathbb{R}^d)},\end{aligned}$$

and hence $\|D_i^h u\|_{H^1(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)}$. Therefore, $\|D_i^h Du\|_{L^2(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)}$, and hence, thanks to Theorem 2.6.2, $\|\partial_i Du\|_{L^2(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)}$, $i = 1, ..., d$; in particular, $u \in H^2(\mathbb{R}^d)$, and $\|D^2 u\|_{L^2(\mathbb{R}^d)} \le \|f\|_{L^2(\mathbb{R}^d)}$. On the other hand, if we take u as test function in the definition of weak solution we get

$$\|Du\|^2_{L^2(\mathbb{R}^d)} + \|u\|^2_{L^2(\mathbb{R}^d)} = \int_{\mathbb{R}^d} fu \le \|f\|_{L^2(\mathbb{R}^d)} \|u\|_{L^2(\mathbb{R}^d)} \le \frac{1}{2}\|f\|^2_{L^2(\mathbb{R}^d)} + \frac{1}{2}\|u\|_{L^2(\mathbb{R}^d)},$$

and hence $\|u\|_{H^1_0(\Omega)} \le \|f\|^2_{L^2(\mathbb{R}^d)}$. Therefore $\|u\|_{H^2(\Omega)} \le C\|f\|_{L^2(\Omega)}$.

We consider next $\Omega = \mathbb{R}^{d-1}_+$. We take again as test function $D_i^{-h} D_i^h u$, but only for $i = 1, \dots, d-1$. Argueing as in the case of the whole space we get $\|\partial_i Du\|_{L^2(\mathbb{R}^d_+)} \le C\|f\|_{L^2(\mathbb{R}^d_+)}$. We still have to control $\partial_{ii} u$. To this aim we use the equation. We have

$$\int_{\mathbb{R}^d_+} \partial_d u \partial_d \varphi = \int_{\mathbb{R}^d_+} \Big(-\sum_{i=1}^{d-1} \partial_i u \partial_i \varphi - u\varphi + f\varphi \Big) = \int_{\mathbb{R}^d_+} \Big(\sum_{i=1}^{d-1} \partial_{ii} u - u + f \Big) \varphi \quad \text{for all } \varphi \in C^\infty_{\rm c}(\Omega).$$

Therefore, by the definition of weak derivative, $\partial_{dd} u = -\sum_{i=1}^{d-1} \partial_{ii} u + u - f \in L^2(\mathbb{R}^d_+)$ in the weak sense, and hence

$$\|\partial_{dd} u\|_{L^2(\mathbb{R}^d)} \le (d-1)\|f\|_{L^2(\mathbb{R}^d)} + \|u\|_{L^2(\mathbb{R}^d)} + \|f\|_{L^2(\mathbb{R}^d)} \le C\|f\|_{L^2(\mathbb{R}^d)}.$$

The estimate $\|u\|_{H^2 L^2(\mathbb{R}^d)} \le C\|f\|_{L^2}$ follows easily. □

We observe that under the hypotheses of Theorem 2.6.1 the weak solution u of (HDP) satisfies $Lu = f$ almost everywhere in Ω. It is hence a *strong solution*: all the derivatives appearing in the equation are $L^1_{\rm loc}$ functions, and the equation is satisfied almost everywhere.

Theorem 2.6.3 (Higher regularity). *Let m be a nonnegative integer, and assume*

$$a^{ij} \in C^{m+1}(\overline{\Omega}), \quad b^i, c \in C^m(\overline{\Omega}), \quad f \in H^m(\Omega), \quad \partial\Omega \in C^{m+2}.$$

If $u_0 \in H^1_0(\Omega)$ is the unique weak solution to (HDP), *then $u \in H^{m+2}(\Omega)$ and we have the estimate $\|u\|_{H^{m+2}(\Omega)} \le C\|f\|_{H^m(\Omega)}$, the constant C depending only on Ω and the coefficients of L.*

The proof is by induction, and follows the same lines as the proof of the previous theorem, hence we omit it. You can find a detailed proof in [4, Section 6.3].

Theorem 2.6.4 (Infinite differentiability). *Assume*

$$a^{ij}, b^i, c \in C^\infty(\overline{\Omega}), \quad f \in C^\infty(\Omega), \quad \partial\Omega \in C^\infty.$$

If $u_0 \in H^1_0(\Omega)$ is the unique weak solution to (HDP), *then $u \in C^\infty(\overline{\Omega})$.*

Proof. According to the previous theorem, $u \in H^m(\Omega)$ for each $m \in \mathbb{N}$. Thus, Theorem 10.3 in Chapter 1 implies that $u \in C^k(\overline{\Omega})$ for each $k \in \mathbb{N}$. □

Remark. If $u \in H^1(\Omega)$ satisfies the equation, but not necessarily the boundary conditions, the above mentioned global regularity results (valid in the whole Ω) fail in general, but we still have local regularity results, valid in any $\Omega' \subset\subset \Omega$. These results do not require smoothness of the boundary, or regularity of the coefficients up to the boundary. See [4, Section 6.3] for the details.

Thanks to these regularity results we will be able to show that if the data of the problem are smooth, then weak solutions are smooth, and hence classical. Let us check this in a couple of examples.

Examples. (a) *Dirichlet problem.* Let $u \in H_0^1(\Omega)$ be the classical solution to (HDP). Assume that the coefficients a^{ij}, b^i and c, and the domain Ω are smooth. Let $f \in H^m(\Omega)$, $m > d/2$. The regularity result shows that $u \in H^{m+2}(\Omega)$, hence $\partial_{ij}u \in H^m(\Omega)$. Thus, from the Sobolev embeddings, since $2m > d$, we get that

$$\partial_{ij}u \in C^{m-[d/2]-1,\gamma}(\overline{\Omega}) \subset C^{0,\gamma}(\overline{\Omega}), \quad \gamma = 1 - \{d/2\}.$$

Therefore, $u \in C^{2,\gamma}(\overline{\Omega})$. In particular, $u \in C^2(\Omega) \cap C(\overline{\Omega})$. Since $u \in H_0^1(\Omega)$, we moreover have $u = 0$ at $\partial\Omega$. We can now integrate by parts, and we get

$$\int_\Omega (Lu - f)\varphi = 0 \quad \text{for all } \varphi \in C_c^\infty,$$

hence $Lu = f$ everywhere in Ω, and we conclude that u is a classical solution.

(b) *Neumann problem.* Let u be a weak solution to the Neumann problem

$$-\Delta u + u = f \quad \text{in } \Omega, \qquad \partial_\nu u = 0 \quad \text{in } \partial\Omega.$$

That is, $u \in H^1(\Omega)$ satisfies

$$\int_\Omega (Du \cdot Dv + uv) = \int_\Omega fv \text{for all } v \in H^1(\Omega).$$

It can be easily checked that the same proof that we gave for the Dirichlet problem works in this case, and we get that $f \in L^2(\Omega)$ implies $H^2(\Omega)$, while $f \in H^m(\Omega)$ implies $u \in H^{m+2}(\Omega)$.

As in example (a), if $f \in H^m(\Omega)$, $m > d/2$, then $u \in C^{2,\gamma}(\overline{\Omega})$. In particular $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$. We can now integrate by parts, and we obtain

$$\int_\Omega (-\Delta u + u - f)\varphi + \int_{\partial\Omega} \partial_\nu u\varphi = 0 \quad \text{for all } \varphi \in C^1(\overline{\Omega}).$$

In particular, for all $v \in C_c^\infty(\Omega)$, $\int_\Omega(-\Delta u + u - f)\varphi = 0$, and hence $-\Delta u + u = f$ everywhere in Ω. Once we know this, we have that $\int_{\partial\Omega} \partial_\nu u\varphi = 0$ for all $\varphi \in C^1(\overline{\Omega})$, which implies that $\partial_\nu u = 0$ in $\partial\Omega$.

Summarizing, u is a classical solution of the Neumann problem.

2.7 Maximum Principle

The Maximum Principle is a very useful tool that admits a number of formulations. We present here a simple form, and only for the Dirichlet problem for a particular equation in a bounded domain. But the technique that we will use can be adapted to deal with other equations and boundary conditions.

Theorem 2.7.1 (Maximum Principle for the Dirichlet problem). *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. Let $f \in L^2(\Omega)$ and $u \in H^1(\Omega) \cap C(\overline{\Omega})$ satisfy*[5]

$$\int_\Omega (Du \cdot Dv + uv) = \int_\Omega fv \quad \textit{for all } v \in H_0^1(\Omega).$$

Then for all $x \in \Omega$

$$\min\{\inf_{\partial\Omega} u, \inf_{\Omega} f\} \le u(x) \le \max\{\sup_{\partial\Omega} u, \sup_{\Omega} f\}.$$

(Here and in the following, $\sup = \operatorname{esssup}$, $\inf = \operatorname{essinf}$*.)*

Proof. We use *Stampacchias's truncation method.* Fix a function $G \in C^1(\mathbb{R})$ such that:

- $|G'(s)| \le M$ for all $s \in \mathbb{R}$ for some $M \in \mathbb{R}_+$;
- G is strictly increasing on $\mathbb{R}_+$;
- $G(s) = 0$ for all $s \le 0$.

Set $K = \max\{\sup_{\partial\Omega} u, \sup_\Omega f\}$ and assume $K < \infty$ (otherwise there is nothing to prove). Let $v = G(u-K)$. Then $v \in H^1(\Omega)$. Moreover, since $u \le \sup_{\partial\Omega} u \le K$ at $\partial\Omega$, then $v = 0$ at $\partial\Omega$, and hence $v \in H_0^1(\Omega)$. Thus, we can use this v as a test function in the equation. Hence, since $Dv = G'(u-K)Du$ almost everywhere, after subtracting $\int_\Omega G(u-K)K$ on both sides of the equation we get

$$0 \le \int_\Omega \big(G'(u-K)|Du|^2 + G(u-K)(u-K)\big) = \int_\Omega (f-K)G(u-K) \le 0,$$

where we have used that $tG(t) \ge 0$ for all $t \in \mathbb{R}$. Hence (remember that u is assumed to be continuous), $G(u-K)(u-K) = 0$ everywhere in Ω, which means that either $u = K$, or $G(u-K) = 0$, which in turn implies that $u \le K$.

For the lower bound, argue with $-u$. □

[5] If Ω is of class C^1 one can remove the assumption $u \in C(\overline{\Omega})$ by invoking the theory of traces to give a meaning to $u\big|_{\partial\Omega}$. One can also remove this assumption if $u \in H_0^1(\Omega)$. It is a good exercise to remove this assumption in these two cases.

Chapter 3

Semilinear elliptic equations

The aim of this chapter is to introduce the student to variational techniques which allow to study semilinear elliptic equations. The model problem that we have in mind is

$$-\Delta u = f(u) \quad \text{in } \Omega, \qquad u = 0 \quad \text{in } \partial\Omega.$$

The idea is to do differential calculus in Banach spaces.

3.1 Gâteaux and Fréchet differentiability

Let X be a real Banach space, and $I : X \to \mathbb{R}$ a continuous functional (not necessarily linear).

Definition 3.1.1. The *directional derivative* of I at $x \in X$ in the direction of $y \in X$ is

$$\partial_y I[x] = \lim_{\varepsilon \to 0} \frac{I[x + \varepsilon y] - I[x]}{\varepsilon}$$

when this limit exists.

Definition 3.1.2. The functional I is *Gâteaux differentiable* at $x \in X$ if there exists $L_x \in X^*$ such that

$$\partial_y I[x] = \langle L_x, y \rangle \quad \text{for all } y \in X.$$

Thus, in order to determine if I is Gâteaux differentiable we have to check first if $\partial_y I[x]$ exists for all $y \in X$, and then whether this quantity is linear and bounded as a function of y.

Notation. If I is Gâteaux differentiable we write $L_x = I'[x]$, and $I'[x] \in X^*$ is known as the Gáteaux derivative of I at x.

Definition 3.1.3. The functional I is *(Fréchet) differentiable* at $x \in X$ if there exists $A \in X^*$ such that

$$I[x + y] = I[x] + Ay + o(\|y\|_X) \quad \text{as } \|y\|_X \to 0.$$

If I is differentiable at x, then $A = I'[x]$. Indeed, since

$$\frac{I[x + \|y\|_X \frac{y}{\|y\|_X}] - I[x]}{\|y\|_X} = \frac{I[x+y] - I[x]}{\|y\|_X} = A\left(\frac{y}{\|y\|_X}\right) + \frac{o(\|y\|_X)}{\|y\|_X},$$

passing to the limit we get $\partial_{y/\|y\|_X} I[x] = A(y/\|y\|_X)$, from where the result follows.

Definition 3.1.4. We say that $I \in C^1(X)$ if $I' : X \to X^*$ is continuous.

Definition 3.1.5. A point $x \in X$ is a critical point of I if $I'[x] = 0$ in X^*, that is, if

$$I'[x]y = 0 \quad \text{for all } y \in X. \tag{3.1}$$

Equation (3.1) is known as the Euler-Lagrange equation for the functional I.

Example. Let $\Omega \subset \mathbb{R}^d$ be a bounded smooth domain and $f \in L^2(\Omega)$. We consider the functional $I : H_0^1(\Omega) \to \mathbb{R}$ given by

$$I[u] = \int_\Omega \left(\frac{1}{2}|Du|^2 - fu\right).$$

Let $u, v \in H_0^1(\Omega)$. Then,

$$\begin{aligned}\partial_v I[u] &= \lim_{t\to 0} \frac{I[u+tv] - I[u]}{t} = \lim_{t\to 0} \frac{1}{t}\left(\frac{t^2}{2}\int_\Omega |Dv|^2 + t\int_\Omega Du\cdot Dv - t\int_\Omega fv\right)\\ &= \int_\Omega (Du\cdot Dv - fv).\end{aligned}$$

Given $u \in X$, the application $v \to \partial_v I[u]$ is obviously linear. Moreover, Hölder's inequality implies that it is also bounded,

$$|\partial_v I[u]| \le \|Du\|_{L^2(\Omega)}\|Dv\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \le (\|Du\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)})\|v\|_{H_0^1(\Omega)}.$$

Hence, I is Gâteaux differentiable at every $u \in H_0^1(\Omega)$, and

$$I'[u]v = \langle I'[u], v\rangle = \int_\Omega (Du\cdot Dv - fv).$$

Let us now check that I is Fréchet differentiable. Indeed,

$$|I[u+v] - I[u] - I'[u]v| = \frac{1}{2}\int_\Omega |Dv|^2 \le \frac{1}{2}\|v\|^2_{H_0^1(\Omega)} = o(\|v\|_{H_0^1(\Omega)}).$$

We finally check that $I \in C^1(H_0^1(\Omega))$. Let $u, v, w \in H_0^1(\Omega)$. Then,

$$\begin{aligned}|(I'[u] - I'[v])w| &= \left|\int_\Omega (Du - Dv)\cdot Dw\right| \le \|D(u-v)\|_{L^2(\Omega)}\|Dw\|_{L^2(\Omega)}\\ &\le \|u-v\|_{H_0^1(\Omega)}\|w\|_{H_0^1(\Omega)}.\end{aligned}$$

Hence,

$$\|I'[u] - I'[v]\|_{H^{-1}(\Omega)} = \sup_{w\neq 0} \frac{|(I'[u] - I'[v])w|}{\|w\|_{H_0^1(\Omega)}} \le \|v - w\|_{H_0^1(\Omega)},$$

and we conclude that I' is continuous (and even Lipschitz continuous). In particular, $I \in C^1(H_0^1(\Omega))$.

Notice that if $u \in H_0^1(\Omega)$ is a critical point of I, then it satisfies the Euler-Lagrange equation (in weak form)

$$\int_\Omega (Du \cdot Dv - fv) = 0 \quad \text{for all } v \in H_0^1(\Omega).$$

It is thus a weak solution to the Dirichlet problem

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{in } \partial\Omega. \tag{3.2}$$

Remark. If a functional is C^1, then it is Fréchet differentiable. The proof, which is left like an exercise, uses the Mean Value Theorem for functionals (prove it also!).

3.2 Variational approach

As we have seen in the previous chapter, one can easily prove the existence of a unique weak solution for problem (3.2) by means, for example, of Lax-Milgram's theorem. However, this is a linear technique, and will not work to deal with the nonlinear problems we have in mind. Hence, we follow another approach: the so-called *variational method*. The idea is to identify a functional having as Euler-Lagrange equation the one that we are trying to analyze. If we are able to find a critical point of the functional, we will have found a weak solution to our problem.

How will we find critical points? A first idea is to find extremal points, minimizing or maximizing the functional.

Proposition 3.2.1. *Let X be a Banach space, and let $I : X \to \mathbb{R}$ be Gâteaux differentiable in $u \in X$. If $I[u] = \min_{v\in X} I[v]$ then $I'[u] = 0$.*

Proof. Let $\varphi \in X$, $t \in \mathbb{R}$. Define $g(t) = I[u + t\varphi]$. Notice that $u + t\varphi \in X$. Then, $g(t) \geq g(0)$. Therefore,

$$0 = g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \lim_{t\to 0} \frac{I[u + t\varphi] - I[u]}{t} = I'[u]\varphi.$$

□

Remark. The argument is local. The result is also true if u is a minimum point in $B_\delta(u)$ for some $\delta > 0$.

Let us consider now a nonlinear example.

Proposition 3.2.2. *Let $\Omega \subset \mathbb{R}^d$ be bounded and smooth. We consider the functional $I : H_0^1(\Omega) \to \mathbb{R}$ given by*

$$I[u] = \int_\Omega F(x, u(x))\, dx, \quad F(x, u) = \int_0^u f(x, s)\, ds,$$

where $f : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$ *is a continuous function satisfying the growth condition*

$$|f(x,s)| \le A(1+|s|^p), \quad p \in [1,(d+2)/(d-2)] \text{ if } d \ge 3, \quad p \in [1,\infty) \text{ if } d = 2.$$

No growth condition is needed if $d = 1$. *Then* I *is Gâteaux differentiable and*

$$I'[u]v = \int_\Omega f(x,u(x))v(x)\,dx. \tag{3.3}$$

Remark. This will be used to find weak solutions to

$$-\Delta u(x) = f(x,u(x)), \quad x \in \Omega, \qquad u(x) = 0, \quad x \in \partial\Omega,$$

which are critical points of the functional $I[u] = \frac{1}{2}\int_\Omega |Du|^2 - \int_\Omega F(x,u(x))\,dx$.

Proof. We only give it for $d \ge 3$. It is a good exercise to do it for $d = 1, 2$.

STEP 1: *The functional is well defined.* Notice that

$$|F(x,u)| = \left|\int_0^u f(x,s)\,ds\right| \le \int_0^{|u|} |f(x,s)|\,ds \le A\int_0^{|u|}(1+|s|^p) \le A(1+|u|^p)|u|.$$

Thus, it is enough to check that $|u|+|u|^{p+1} \in L^1(\Omega)$. But, thanks to Sobolev's embedding we have $u \in L^{2^*}(\Omega)$, where $2^* = 2d/(d-2) \ge p+1$ if $p \le (d+2)/(d-2)$. Hence, using Hölder's inequality we have the desired inclusion.

STEP 2: *Computation of the directional derivative.* Let $u, v \in H_0^1(\Omega)$. We have

$$I'[u]v = \lim_{t\to 0}\frac{I[u+tv]-I[u]}{t} = \lim_{t\to 0}\int_\Omega \frac{F(x,u(x)+tv(x)) - F(x,u(x))}{t}\,dx.$$

We want to pass to the limit inside the integral. Let us check that we can apply the Dominated Convergence Theorem. Indeed, since $(a+b)^p \le 2^p(a^p+b^p)$ for all $a, b > 0$,

$$\begin{aligned}
\frac{|F(x,u(x)+tv(x)) - F(x,u(x))|}{t} &= \left|\frac{1}{t}\int_0^1 \frac{d}{ds}F(x,u(x)+stv(x))\,ds\right| \\
&= \left|\int_0^1 f(x,u(x)+stv(x))v(x)\,ds\right| \\
&\le A\int_0^1 (1+|u(x)+stv(x)|^p)|v(x)|\,ds \\
&\le C\int_0^1 (1+|u(x)|^p+|v(x)|^p)|v(x)|\,ds \\
&= C\Big(|v(x)| + |u(x)||v(x)| + |v(x)|^{p+1}\Big).
\end{aligned}$$

Thanks to Sobolev's embedding we know that v and $|v|^{p+1}$ belong to $L^1(\Omega)$. Hence, it is enough to check that $|u|^p|v| \in L^1(\Omega)$. This follows easily from Hölder's inequality and Sobolev's embedding. Indeed,

$$\int_\Omega |u|^p|v| \le \left(\int_\Omega |u|^{p+1}\right)^{\frac{p}{p+1}}\left(\int_\Omega |v|^{p+1}\right)^{\frac{1}{p+1}} < \infty.$$

We can therefore pass to the limit to obtain $\partial_v I[u] = \int_\Omega f(x, u(x))v(x)\,dx$.

STEP 3: *I is Gâteaux differentiable.* It is trivially checked that $v \to \partial_v I[u]$ is linear in v. It is enough then to check that this application is bounded. Indeed, using Hölder's inequality and Sobolev's embedding

$$
\begin{aligned}
|\partial_v I[u]| &\le \int_\Omega |f(x,u(x))||v(x)|\,dx \le \left(\int_\Omega |f(x,u(x))|^{\frac{2d}{d+2}}\,dx\right)^{\frac{d+2}{2d}} \|v\|_{2^*} \\
&\le C\left(\int_\Omega (1+|u|^p)^{\frac{2d}{d+2}}\,dx\right)^{\frac{d+2}{2d}} \|v\|_{H_0^1(\Omega)} \\
&\le C\left(\int_\Omega \left(1+|u|^{\frac{2dp}{d+2}}\right)dx\right)^{\frac{d+2}{2d}} \|v\|_{H_0^1(\Omega)}.
\end{aligned}
$$

Hence, it is enough to prove that $|u|^{\frac{2dp}{d+2}} \in L^1(\Omega)$, which is the case, since $\frac{2dp}{d+2} \le \frac{2d}{d-2}$ (this is equivalent to $p \le \frac{d+2}{d-2}$). □

Exercise. Study wether the functional I in the above proposition is Fréchet differentiable and C^1.

3.3 Direct method of the Calculus of Variations

Let X be a real Banach space, and $I : X \to \mathbb{R}$ a continuous functional. Is there a minimum? We don't even know whether there is an infimum. Hence we start be giving a hypothesis to guarantee the existence of a finite infimum: *coercivity.*

Definition 3.3.1. We say that the functional I is *coercive* if there exist constants $\alpha > 0$ and $\beta \in \mathbb{R}$ such that

$$I[u] \ge \alpha\|u\|_X - \beta \quad \text{for all } u \in X.$$

Notice that coercivity implies that

$$m := \inf\{I[u] : u \in X\} > -\infty.$$

Remark. One can find in the literature other definitions of coercivity, for example,

$$I[u] \ge \alpha\|u\|_X^2 - \beta \quad \text{for all } u \in X \quad \text{for some } \alpha > 0,\ \beta \in \mathbb{R},$$

which is more restrictive than the one we have given, or

$$I[u_k] \to \infty \quad \text{if } \|u_k\| \to \infty,$$

which is less restrictive. All of them serve for our purposes.

Let $(u_k)_{k=1}^\infty$ be a minimizing sequence, that is, a sequence of elements of X such that $I[u_k] \to m$. If k is big enough, then $I[u_k] \le m+1$. Hence, using the coercivity,

$$\alpha\|u_k\|_X - \beta \le m+1,$$

which implies that $\|u_k\|_X \le C < \infty$. Now, if X is *reflexive*, then there exists a subsequence that converges weakly in X, $u_{k_j} \rightharpoonup u$ in X. However, even if I is continuous, we do not have in general $I[u_{k_j}] \to I[u]$ (weak convergence is not enough). Hence, we will need to require something else from I.

Definition 3.3.2. A functional $I : X \to \mathbb{R}$ is *sequentially weakly lower semicontinuous* if

$$I[u] \le \liminf_{k\to\infty} I[u_k] \quad \text{if } u_k \rightharpoonup u \text{ in } X.$$

Thus, if $(u_{k_j})_{j=1}^{\infty}$ is a weakly convergent subsequence of the minimizing sequence, and I is sequentially weakly lower semicontinuous, then

$$m \le I[u] \le \liminf_{j\to\infty} I[u_{k_j}] = m;$$

that is, $I[u] = m$, and hence I achieves its minimum at $u \in X$. We have therefore proved the following result.

Theorem 3.3.1. *Let X be a reflexive real Banach space, and $I : X \to \mathbb{R}$ a coercive and sequentially weakly lower semicontinuous functional. Then, there exists $u \in X$ such that $I[u] = \min_{v\in X} I[v]$.*

Proposition 3.3.2. *Norms in a Hilbert space are sequentially weakly lower semicontinuous.*

Proof. Let $u_k \rightharpoonup u$ in X. Then, on the one hand, using Cauchy-Schwarz's inequality, $|\langle u_k, u\rangle| \le \|u_k\|_X \|u\|_X$. On the other hand, from the weak convergence, $|\langle u_k, u\rangle| \to \|u\|_X^2$. We conclude that

$$\|u\|_X^2 \le \|u\|_X \liminf_{k\to\infty} \|u_k\|_X.$$

If $\|u\|_X \ne 0$, dividing by $\|u\|_X$ we get the desired result, $\|u\|_X \le \liminf_{k\to\infty} \|u_k\|_X$. If $\|u\|_X = 0$ the result is trivial. □

Remark. Let $\mathcal{M} \subset X$ be weakly closed and non-empty. The proof of the theorem can be repeated to show that there exists $u \in \mathcal{M}$ such that $I[u] = \inf\{I[u] : u \in \mathcal{M}\}$.

Theorem 3.3.3 (Mazur's theorem). *Let X be a reflexive Banach space. A closed convex subset of X is weakly closed.*

The proof is an application of Hahn-Banach's Theorem.

Corollary 3.3.4. *Let X be a reflexive Banach space, and $\mathcal{M} \subset X$ a non-empty closed convex set. If I is coercive and sequentially weakly lower semicontinuous, then there exists $u \in \mathcal{M}$ such that $I[u] = \inf\{I[u] : u \in \mathcal{M}\}$.*

Definition 3.3.3. The functional $I : X \to \mathbb{R}$ is *convex* if $I[\lambda x + (1-\lambda)y] \le \lambda I[x] + (1-\lambda)I[y]$ for all $x, y \in X$ and $\lambda \in [0,1]$. If equality holds only for $\lambda = 0$ and $\lambda = 1$, then I is said to be *strictly convex.*

Proposition 3.3.5. *If I is strictly convex, then I has at most one minimum.*

Proof. If $I[x] = \underbrace{\min_{v\in X} I[v]}_{m} = I[y]$, then

$$m \le I\left[\frac{x+y}{2}\right] < \frac{I[x]+I[y]}{2} = m,$$

a contradiction. □

Proposition 3.3.6. *Let $I : X \to \mathbb{R}$ be Gâteaux differentiable. The following statements are equivalent.*

(i) *I is convex.*
(ii) *$I[y] \geq I[x] + I'[x](y-x)$ for all $x, y \in X$.*
(iii) *$(I'[y] - I'[x])(y-x) \geq 0$ for all $x, y \in X$ (monotonicity of I').*

Remark. There is an analogous result to characterize strict convexity.

Proof. (i)⇒(ii): Let $\lambda \in (0,1)$. Using the convexity of I we get

$$\begin{aligned} I[x+\lambda(y-x)] - I[x] &= I[(1-\lambda)x + \lambda y] - I[x] \leq (1-\lambda)I[x] + \lambda I[y] - I[x] \\ &= \lambda(I[y] - I[x]). \end{aligned}$$

Therefore,

$$\frac{I[x+\lambda(y-x)] - I[x]}{\lambda} \leq I[y] - I[x].$$

Passing to the limit as $\lambda \to 0^+$ we get $I'[x](y-x) \leq I[y] - I[x]$, and the result follows.
(ii)⇒(i): We have, using (ii) twice,

$$\begin{aligned} I[x] &\geq I[\lambda x + (1-\lambda)y] + I'[\lambda x + (1-\lambda)y](x - (\lambda x + (1-\lambda)y)), \\ I[y] &\geq I[\lambda x + (1-\lambda)y] + I'[\lambda x + (1-\lambda)y](y - (\lambda x + (1-\lambda)y)). \end{aligned}$$

Adding the first inequality multiplied by λ to the second one multiplied by $1-\lambda$ we get

$$\lambda I[x] + (1-\lambda)I[y] \geq I[\lambda x + (1-\lambda)y].$$

(ii)⇒(iii): We use (ii) twice to obtain the inequalities

$$\begin{aligned} I[y] - I[x] &\geq I'[x](y-x), \\ I[x] - I[y] &\geq I'[y](x-y). \end{aligned}$$

Adding them we get $0 \geq (I'[x] - I'[y])(y-x)$, which is equivalent to (iii).
(iii)⇒(ii): For any positive λ we define $\phi(\lambda) = I[x+\lambda(y-x)]$. Notice that $\phi'(\lambda) = I'[x+\lambda(y-x)](y-x)$, $\phi'(0) = I'[x](y-x)$. Hence, using (iii),

$$\phi'(\lambda) - \phi'(0) = (I'[x+\lambda(y-x)] - I'[x])(y-x) = \frac{1}{\lambda}(I'[x+\lambda(y-x)] - I'[x])(x+\lambda(y-x)-x) \geq 0.$$

Integrating the inequality $\phi'(\lambda) \geq \phi'(0)$ we get $\phi(\lambda) \geq \phi(0) + \phi'(0)\lambda$, that is,

$$I[x+\lambda(y-x)] \geq I[x] + I'[x](y-x)\lambda.$$

Defining $z := x + \lambda(y-x)$ we have $(y-x)\lambda = z - x$, and hence $I[z] \geq I[x] + I'[x](z-x)$. □

Proposition 3.3.7. *If $I : X \to \mathbb{R}$ is sequentially lower semicontinuous and convex then it is sequentially weakly lower semicontinuous. In particular, if I is continuous and convex it is sequentially weakly lower semicontinuous.*

You can find a proof, which is based in a lemma due to Mazur, for example in [7].

Proposition 3.3.8 (Generalized Dominated Convergence Theorem). *Let Ω be a measurable sets. Let $(f_j)_{j=1}^\infty$ and $(h_j)_{j=1}^\infty$ be sequences in $L^1(\Omega)$ such that $f_j \to f$ almost everywhere in Ω, $|f_j| \le h_j$, $h_j \to h$ almost everywhere in Ω and in $L^1(\Omega)$. Then $f \in L^1(\Omega)$ and $\int_\Omega f_j \to \int_\Omega f$.*

Proof. The sequence $(h_j - f_j)_{j=1}^\infty$ converges almost everywhere in Ω to $h - f$. Since the functions in the sequence are nonnegative, we may apply Fatou's Lemma to obtain

$$\begin{aligned}\int_\Omega h - \int_\Omega f &= \int_\Omega (h-f) \le \liminf_{j\to\infty} \int_\Omega (h_j - f_j) \\ &= \lim_{j\to\infty}\int_\Omega h_j - \limsup_{j\to\infty}\int_\Omega f_j = \int_\Omega h - \limsup_{j\to\infty}\int_\Omega f_j.\end{aligned}$$

Therefore,

$$\limsup_{j\to\infty}\int_\Omega f_j \le \int_\Omega f.$$

On the other hand, the sequence $(h_j + f_j)_{j=1}^\infty$ converges almost everywhere in Ω to $h + f$. Since the functions in the sequence are nonnegative, we may apply Fatou's Lemma to obtain

$$\begin{aligned}\int_\Omega h + \int_\Omega f &= \int_\Omega (h+f) \le \liminf_{j\to\infty} \int_\Omega (h_j + f_j) \\ &= \lim_{j\to\infty}\int_\Omega h_j + \liminf_{j\to\infty}\int_\Omega f_j = \int_\Omega h + \liminf_{j\to\infty}\int_\Omega f_j.\end{aligned}$$

Therefore,

$$\int_\Omega f \le \liminf_{j\to\infty}\int_\Omega f_j.$$

Summarizing,

$$\limsup_{j\to\infty}\int_\Omega f_j \le \int_\Omega f \le \liminf_{j\to\infty}\int_\Omega f_j,$$

and hence

$$\lim_{j\to\infty}\int_\Omega f_j = \int_\Omega f.$$

□

Example. We consider the semilinear problem

$$-\Delta u(x) = f(x, u(x)), \quad x \in \Omega, \qquad u(x) = 0, \quad x \in \partial\Omega.$$

A weak solution to this problem is a function $u \in H_0^1(\Omega)$ such that

$$\int_\Omega Du \cdot Dv - \int_\Omega f(x, u(x))v(x)\,dx = 0 \quad \text{for all } v \in H_0^1(\Omega).$$

This is the Euler-Lagrange equation of the functional

$$I[u] = \underbrace{\frac{1}{2}\int_\Omega |Du|^2}_{I_1[u]} - \underbrace{\int_\Omega F(x, u(x))\,dx}_{I_2[u]}, \quad \text{where } F(x,u) = \int_0^u f(x,s)\,ds,$$

which is Gâteaux differentiable if $|f(x,s)| \le A(1+|s|^p)$, $p \in [1,(d+2)/(d-2)]$ if $d \ge 3$, $p \in [1,\infty]$ if $d=2$, no growth condition if $d=1$.

Is this functional sequentially weakly lower semicontinuous? This is clear for $I_1[u] = \frac{1}{2}\|u\|^2_{H^1_0(\Omega)}$, since it is a multiple of the square of the norm in a Hilbert space. As for I_2, it is also sequentially weakly lower semicontinuous under the additional restriction $p \in [1,(d+2)/(d-2))$ if $d \ge 3$, $p \in [1,\infty)$ if $d=2$. Let us prove it for the case $d \ge 3$. It is a good exercise to do the cases $d = 1, 2$.

Let $u_k \rightharpoonup u$ in $H^1_0(\Omega)$. Then $\|u_k\|_{H^1_0(\Omega)} \le C$. Therefore, using Rellich-Kondrachov's Compactness Theorem, there is a subsequence $(u_{k_j})_{j=1}^\infty$ such that $u_{k_j} \to u$ in $L^{p+1}(\Omega)$ (the restriction on p is used here to guarantee that $p+1 < 2^*$) and $u_{k_j} \to u$ almost everywhere in Ω. Hence, $F(x,u_{k_j}(x)) \to F(x,u(x))$ for almost every $x \in \Omega$. On the other hand,

$$|F(x,u_{k_j}(x))| \le A(|u_{k_j}(x)| + |u_{k_j}(x)|^{p+1}) \to A(|u(x)| + |u(x)|^{p+1}) \in L^1(\Omega)$$

both almost everywhere in Ω and in $L^1(\Omega)$. Hence, we can apply the Generalized Dominated Convergence Theorem to obtain that

$$I_2[u_{k_j}] = \int_\Omega F(x,u_{k_j}(x))\,dx \to \int_\Omega F(x,u(x))\,dx = I_2[u].$$

By using a contradiction argument (do it!) it is easy that this convergence is not restricted to a subsequence. Hence I_2, and therefore I, is sequentially weakly lower semicontinuous.

What about coercivity? We will give several sufficient conditions guaranteing it. Let us recall that

$$I[u] = \frac{1}{2}\|u\|^2_{H^1_0(\Omega)} - \int_\Omega F(x,u(x))\,dx.$$

Condition 1: $F(x,u(x)) \le a(x)$ for some $a \in L^1(\Omega)$.

Indeed, under this condition $I[u] \ge \frac{1}{2}\|u\|^2_{H^1_0(\Omega)} - \|a\|_{L^1(\Omega)}$.

Condition 2: $F(x,u(x)) \le \gamma|u(x)|^q + a(x)$ for some $a \in L^1(\Omega)$, $\gamma \ge 0$ and $q \in (1,2)$.

Indeed, under these conditions

$$I[u] = \frac{1}{2}\|u\|^2_{H^1_0(\Omega)} - \gamma\|u\|^q_{L^q(\Omega)} - \|a\|_{L^1(\Omega)} \ge \frac{1}{2}\|u\|^2_{H^1_0(\Omega)} - C\|u\|^q_{H^1_0(\Omega)} - \|a\|_{L^1(\Omega)},$$

where we are using that $\|u\|_{L^q(\Omega)} \le C\|u\|_{L^2(\Omega)} \le C\|u\|_{H^1_0(\Omega)}$, since Ω is bounded. Therefore, $I[u] = \alpha\|u\|^2_{H^1_0(\Omega)} - \beta$ for some $\alpha > 0$, because

$$\|u\|^2_{H^1_0(\Omega)}\left(\frac{1}{2} - \frac{C}{\|u\|^{2-q}_{H^1_0(\Omega)}}\right) \ge \begin{cases} \frac{1}{4}\|u\|^2_{H^1_0(\Omega)} & \text{if } \|u\|_{H^1_0(\Omega)} \ge (4C)^{\frac{1}{2-q}}, \\ \frac{1}{2}\|u\|^2_{H^1_0(\Omega)} - C(4C)^{\frac{q}{2-q}} & \text{if } \|u\|_{H^1_0(\Omega)} \le (4C)^{\frac{1}{2-q}}. \end{cases}$$

Condition 3: $F(x,u(x)) \le \gamma|u(x)|^2 + a(x)$ for some $a \in L^1(\Omega)$, and $\gamma \ge 0$ small.

Indeed, using Poincaré's inequality $\|u\|_{L^2(\Omega)} \le C\|u\|_{H^1_0(\Omega)}$, we have

$$I[u] \ge \frac{1}{2}\|u\|^2_{H^1_0(\Omega)} - \gamma\|u\|^2_{L^2(\Omega)} - \|a\|_{L^1(\Omega)} \ge \big(\frac{1}{2} - \gamma C^2\big)\|u\|^2_{H^1_0(\Omega)} - \|a\|_{L^1(\Omega)},$$

which means that I is coercive if if $\gamma < 1/(2C^2)$.

3.4 Optimization with restrictions. Lagrange multipliers. Eigenvalues

We start by reminding a well-known result valid for finite dimensional spaces. Let $F, G \in C^1(\mathbb{R}^d;\mathbb{R})$, and $\mathcal{C} = \{x \in \mathbb{R}^d : G(x) = 0\}$. If $x_0 \in \mathcal{C}$ is such that $F(x_0) = \min_{\mathcal{C}} F$, then, either $G'(x_0) = 0$, or there exists a value $\mu \in \mathbb{R}$ such that $F'(x_0) = \mu G'(x_0)$. The value μ is said to be a *Lagrange multiplier*.

What happens in a general Banach space? As we will see, if we state it properly, the result is still true.

We start by proving a technical lemma.

Lemma 3.4.1. *Let X be a real Banach space and $F, G \in C^1(X;\mathbb{R})$. Let $x_0 \in X$. If there exist $v, w \in X$ such that*

$$(F'[x_0]v)(G'[x_0]w) \neq (F'[x_0]w)(G'[x_0]v),$$

then F cannot have an extremal point at x_0 even if we restrict to the level set $\mathcal{C} = \{x \in X : G[x] = G[x_0]\}$.

Proof. Fix $v, w \in X$. For every $s, t \in \mathbb{R}$ we define the real-valued functions

$$f(s,t) = F(x_0 + sv + tw), \qquad g(s,t) = G(x_0 + sv + tw).$$

Since $F, G \in C^1(X;\mathbb{R})$, it is easy to check (do it!) that $f, g \in C^1(\mathbb{R}^2;\mathbb{R})$. A simple computation yields

$$\partial_s f(0,0) = F'[x_0]v, \qquad \partial_s g(0,0) = G'[x_0]v,$$
$$\partial_t f(0,0) = F'[x_0]w, \qquad \partial_t g(0,0) = G'[x_0]w.$$

Hence, $|\partial(f,g)/\partial(s,t)|(0,0) \neq 0$. Applying the Inverse Function Theorem we conclude that x_0 cannot be an extremal point of $F\big|_{\mathcal{C}}$. □

Theorem 3.4.2 (Lagrange multipliers). *Let X be a real Banach space, $F, G \in C^1(X;\mathbb{R})$, and $\mathcal{C} = \{x \in X : G[x] = 0\}$. If $x_0 \in \mathcal{C}$ is a local extremal point of $F\big|_{\mathcal{C}}$, then at least one of the following alternatives holds:*

(a) *$G'[x_0]v = 0$ for all $v \in X$;*

(b) *there exists a value $\mu \in \mathbb{R}$ such that $F'[x_0]v = \mu G'[x_0]v$ for all $v \in X$.*

The value μ is said to be a Lagrange multiplier.

Proof. If (a) does not hold, there exists $w \in X$ such that $G'[x_0]w \neq 0$. By the previous lemma, necessarily

$$(F'[x_0]v)(G'[x_0]w) = (F'[x_0]w)(G'[x_0]v).$$

Hence, for all $v \in X$ we have $F'[x_0]v = \mu G'[x_0]v$, with $\mu = (F'[x_0]w)/(G'[x_0]w)$. □

As an application of this theorem we will now study the eigenvalue problem for the Laplacian with homogeneous Dirichlet boundary conditions.

Let $\Omega \subset \mathbb{R}^d$ be bounded, connected, and smooth. We are looking for non-trivial (weak) solutions u to

$$-\Delta u = \lambda u \quad \text{in } \Omega, \qquad u = 0 \quad \text{in } \partial\Omega.$$

As we will see, such solutions only exist for certain values λ, the *eigenvalues*. The solution u is said to be an *eigenfunction* associated to the eigenvalue λ. Thus, we are looking for pairs (λ, u), $u \not\equiv 0$, $u \in H_0^1(\Omega)$ such that

$$\int_\Omega Du \cdot Dv = \lambda \int_\Omega uv \quad \text{for all } v \in H_0^1(\Omega).$$

We remark that if u is an eigenfunction associated to the eigenvalue λ, then all the functions in the one-dimensional subspace $\{\alpha u\}$, $\alpha \in \mathbb{R}$ are also eigenfunctions associated to λ. To choose one among all of them, we look for eigenfunctions satisfying $\|u\|_{L^2(\Omega)} = 1$. This is the restriction that will lead to a Lagrange multiplier (the eigenvalue)! Let us follow this approach.

We consider $F, G \in C^1(H_0^1(\Omega); \mathbb{R})$ given by $F[u] = \|Du\|_2^2$, $G[u] = \|u\|^2 - 1$. We are looking for a function $u \in H_0^1(\Omega)$ minimizing F in $\mathcal{C} := \{u \in H_0^1(\Omega) : G[u] = 0\} = \{u \in H_0^1(\Omega) : \|u\|_2 = 1\}$. We already know that

$$F'[u]v = 2\int_\Omega Du \cdot Dv, \qquad G'[u]v = 2\int_\Omega uv.$$

Hence, if F has a minimum at some $\bar{u} \in \mathcal{C}$, since

$$G'[\bar{u}]\bar{u} = 2\int_\Omega |\bar{u}|^2 = 2 \neq 0,$$

the Lagrange Multiplier's Theorem will imply that there exists a value λ such that $F'[\bar{u}]v = \lambda G'[\bar{u}]v$ for all $v \in H_0^1(\Omega)$. Thus, the pair $(\lambda, \bar{u})$ will be a solution pair to the eigenvalue problem.

Let us prove then that F achieves its minimum in $\mathcal{C}$. Let $I := \inf\{F[u] : u \in \mathcal{C}\} \geq 0$. Let $(u_k)_{k=1}^\infty$, $u_k \in \mathcal{C}$, be a minimizing sequence in $\mathcal{C}$, $F[u_k] \to I$. We have $\|Du_k\|_2 \leq K$. Hence, thanks to Poincaré's inequality, $\|u_k\|_{H_0^1(\Omega)} \leq K$. We can now apply Rellich-Kondrachov's compactness theorem to obtain a subsequence $(u_{k_j})_{j=1}^\infty$ such that $u_{k_j} \to \bar{u}$ in $L^2(\Omega)$, $u_{k_j} \rightharpoonup \bar{u}$ in $H_0^1(\Omega)$. The strong convergence in $L^2(\Omega)$ yields

$$\|\bar{u}\|_2 = \lim_{j\to\infty} \|u_{k_j}\|_2 = 1.$$

Hence, $\bar{u} \in \mathcal{C}$.

On the other hand, since the norm in a Hilbert space is sequentially weakly lower semicontinuous,

$$\begin{aligned} I \leq F[\bar{u}] &= \|\bar{u}\|_{H_0^1(\Omega)}^2 - \|\bar{u}\|_{L^2(\Omega)}^2 \leq \liminf_{j\to\infty} \|u_{k_j}\|_{H_0^1(\Omega)}^2 - \lim_{j\to\infty} \|u_{k_j}\|_{L^2(\Omega)}^2 \\ &\leq \liminf_{j\to\infty} \left(\|u_{k_j}\|_{H_0^1(\Omega)}^2 - \|u_{k_j}\|_{L^2(\Omega)}^2\right) = \liminf_{j\to\infty} F[u_{k_j}] = I, \end{aligned}$$

so that $F[\bar{u}] = I$, as desired.

Observe that $I = \int_\Omega |D\bar{u}|^2 = \lambda \int_\Omega |\bar{u}|^2 = \lambda$. Therefore,

$$\lambda = \inf \left\{ \int_\Omega |Du|^2 : u \in H_0^1(\Omega), \|u\|_2 = 1 \right\} = \inf \left\{ \int_\Omega |Du|^2 / \int_\Omega u^2 : u \in H_0^1(\Omega), u \not\equiv 0 \right\}.$$

The quotient $\int_\Omega |Du|^2 / \int_\Omega u^2$ is known as Rayleigh's quotient. Thus, we have

$$\lambda \int_\Omega u^2 \leq \int_\Omega |Du|^2 \quad \text{for all } v \in H_0^1(\Omega).$$

We recognize here Poincaré's inequality. We conclude that the best constant in Poincaré's inequality is given by the *first*[1] eigenvalue. Moreover, equality in this inequality is achieved by $\bar{u}$.

Notation. We denote the pair (λ, u) that we have just obtained by (λ_1, u_1).

Observe that $\lambda_1 > 0$. Indeed, since $\lambda_1 = \int_\Omega |D\bar{u}|^2$, if $\lambda_1 = 0$ we have that $\bar{u}$ is a constant, which has to be 0, since $\bar{u} \in H_0^1(\Omega)$. Therefore, $\bar{u} \equiv 0$, a contradiction, since $\|\bar{u}\|_1 = 1$.

Are there other essentially different (eigenfunctions which are not multiples of u_1) solution pairs?

Assume that there is another solution pair, (λ_2, u_2), $\|u_2\|_2 = 1$. We have

$$(\lambda_2 - \lambda_1) \int_\Omega u_1 u_2 = \lambda_2 \int_\Omega u_1 u_2 - \lambda_1 \int_\Omega u_1 u_2 = \int_\Omega Du_1 \cdot Du_2 - \int_\Omega Du_1 \cdot Du_2 = 0,$$

where we have used that both (λ_1, u_1) and (λ_2, u_2) are solution pairs. Therefore, if $\lambda_1 \neq \lambda_2$, then $\int_\Omega u_1 u_2 = 0$. In other words, $(u_1, u_2)_{L^2(\Omega)} = 0$. We have thus proved the following result.

Proposition 3.4.3. *The eigenfunctions of the Laplacian corresponding to different eigenvalues are orthogonal in $L^2(\Omega)$.*

Even if $\lambda_2 = \lambda_1$, which in principle is possible, we look for an eigenfunction u_2 which is linearly independent of u_1. Hence, it seems sensible to look for the second eigenfunction in the orthogonal complement of the subspace generated by u_1. Let us formalize this idea.

Let $X_1 = \{v \in H_0^1(\Omega) : (v, u_1)_{L^2(\Omega)} = 0\}$. Notice that X_1 is the kernel of the linear and continuous functional on the Hilbert space $H_0^1(\Omega)$ given by $v \to (v, u_1)_{L^2(\Omega)}$. Hence, X_1 is a closed subspace of $H_0^1(\Omega)$, and therefore it is a Hilbert space with the norm inherited from $H_0^1(\Omega)$. As before, let $F[u] = \|Du\|_2^2$. Let

$$\lambda_2 = \inf \{F[v] : v \in X_1, \|v\|_2 = 1\} = \inf \left\{ \int_\Omega |Dv|^2 / \int_\Omega v^2 : v \in X_1 \right\}.$$

Since $X_1 \subset H_0^1(\Omega)$, $\lambda_2 \geq \lambda_1$. Repeating what we did to obtain the pair (λ_1, u_1), it is easily checked that λ_2 is attained at some $u_2 \in X_1$.

[1]In the next pages it will be clear why we denote the eigenvalue that we have just obtained as the *first* one.

Recursively, given the solution pairs (λ_j, u_j), $j = 1, \dots, n$, let

$$X_n = \left\{v \in H_0^1(\Omega) : (v, u_j)_{L^2(\Omega)} = 0, j = 1, \dots, n\right\}, \quad \text{and}$$
$$\lambda_{n+1} = \inf\left\{F[v] : v \in X_n, \|v\|_2 = 1\right\} = \inf\left\{\int_\Omega |Dv|^2 / \int_\Omega v^2 : v \in X_n\right\}.$$

Then $\lambda_{n+1} \geq \lambda_n$, and λ_{n+1} is achieved at some $u_{n+1} \in X_{n+1}$. Notice that $(u_k, u_j)_{L^2(\Omega)} = \delta_{kj}$, so that we have an orthonormal set in $L^2(\Omega)$. Moreover, $(Du_k, Du_j)_{L^2(\Omega)} = \lambda_k \delta_{kj}$, so that $(u_k, u_j)_{H_0^1(\Omega)} = (\lambda_k + 1)\delta_{kj}$.

Remark. We do not know yet whether this procedure generates all the eigenvalues and eigenfunctions.

We claim now that $\lambda_k \to \infty$. Otherwise, λ_k is a bounded sequence, and hence, since $\|u_k\|^2_{H_0^1(\Omega)} = \lambda_k + 1$, the sequence $(u_k)_{k=1}^\infty$ would be bounded. Therefore, Rellich-Kondrachov's compactness theorem implies that the sequence has a convergent in $L^2(\Omega)$ subsequence $(u_{k_j})_{j=1}^\infty$. It is then a Cauchy sequence, $\|u_{k_j} - u_{k_l}\|_{L^2(\Omega)} \to 0$ as $j, l \to \infty$. However, since the sequence is orthonormal in $L^2(\Omega)$, we have $\|u_{k_j} - u_{k_l}\|_{L^2(\Omega)} = 2$ if $k \neq l$, a contradiction.

Corollary 3.4.4. *Each λ_k can appear in the list generated by this procedure only a finite number of times.*

We now check that the family $(u_k)_{k=1}^\infty$ is a Hilbert orthonormal basis of $L^2(\Omega)$.

Theorem 3.4.5. *Let $(u_k)_{k=1}^\infty$ be the orthonormal family of eigenfunctions of the Laplacian generated in the procedure described above. Given $f \in L^2(\Omega)$, there exists a sequence $(\alpha_k)_{k=1}^\infty$ of real numbers such that $f = \sum_{k=1}^\infty \alpha_k u_k$ in $L^2(\Omega)$, that is,*

$$\left\| f - \sum_{k=1}^N \alpha_k u_k \right\|_{L^2(\Omega)} \to 0 \quad \text{as } N \to \infty.$$

Proof. The sequence $(\alpha_k)_{k=1}^\infty$ will be given by $\alpha_k = (f, u_k)_{L^2(\Omega)}$.

Along the proof we denote $\rho_N = f - \sum_{k=1}^N \alpha_k u_k$ (the remainder).

For a start we assume that $f \in H_0^1(\Omega)$. Because of the orthonormality of the sequence $(u_k)_{k=1}^\infty$, we have $(\rho_N, u_k)_{L^2(\Omega)} = 0$, $k = 1, \dots, N$. Thus, $\rho_N \in X_N$. Taking $\rho_N \in H_0^1(\Omega)$ as test function in the PDE satisfied by the function u_k we get

$$\int_\Omega D\rho_N \cdot Du_k = \lambda_k \int_\Omega \rho_N u_k = 0, \quad k = 1, \dots, N.$$

Therefore,

$$\|Df\|_2^2 = \|D\rho_N + D(\sum_{k=1}^N \alpha_k u_k)\|_2^2 = \|D\rho_N\|_2^2 + \sum_{k=1}^N \alpha_k^2 \|Du_k\|_2^2 = \|D\rho_N\|_2^2 + \sum_{k=1}^N \alpha_k^2 \lambda_k,$$

which implies that $\|D\rho_N\|_2^2 \leq \|Df\|_2^2 = C$. Since moreover $\rho_N \in X_N$, then $\lambda_{N+1}\|\rho_N\|_2^2 \leq \|D\rho_N\|_2^2 \leq C$. We now use that $\lambda_{N+1} \to \infty$ as $N \to \infty$ to conclude that $\|\rho_N\|_2 \to 0$.

To prove the result for a general $f \in L^2(\Omega)$ we use a density argument. Let us recall that $H_0^1(\Omega)$ is dense in $L^2(\Omega)$. Thus, given $\varepsilon > 0$, there is a function $f_\varepsilon \in H_0^1(\Omega)$ such that $\|f - f_\varepsilon\|_2 \le \varepsilon$. Let $\alpha_{k,\varepsilon} = (f_\varepsilon, u_k)_{L^2(\Omega)}$. Then, using Cauchy-Schwarz's inequality,

$$|\alpha_{k,\varepsilon} - \alpha_k| = |(f_\varepsilon - f, u_k)_{L^2(\Omega)}| \le \|f_\varepsilon - f\|_2 \le \varepsilon.$$

On the other hand,

$$\|f - \sum_{k=1}^N \alpha_k u_k\|_2 \le \|f - f_\varepsilon\|_2 + \|f_\varepsilon - \sum_{k=1}^N \alpha_{k,\varepsilon} u_k\|_2 + \|\sum_{k=1}^N (\alpha_{k,\varepsilon} - \alpha_k) u_k\|_2.$$

Therefore, since, thanks to Bessel's inequality[2] we have

$$\|\sum_{k=1}^N (\alpha_{k,\varepsilon} - \alpha_k) u_k\|_2 = \left(\sum_{k=1}^N (\alpha_{k,\varepsilon} - \alpha_k)^2\right)^{1/2} \le \|f_\varepsilon - f\|_2.$$

We conclude that

$$\|f - \sum_{k=1}^N \alpha_k u_k\|_2 \le 2\|f - f_\varepsilon\|_2 + \|f_\varepsilon - \sum_{k=1}^N \alpha_{k,\varepsilon} u_k\|_2.$$

Letting $N \to \infty$ we finally obtain

$$\limsup_{N\to\infty} \|f - \sum_{k=1}^N \alpha_k u_k\|_2 \le 2\|f - f_\varepsilon\|_2 \le 2\varepsilon.$$

We conclude by letting $\varepsilon \to 0$. □

This theorem guarantees that all eigenvalues and eigenfunctions are generated in the recursive minimization process described above. Indeed, if (λ, u) $u \not\equiv 0$, is a solution pair with $\lambda \ne \lambda_n$ for all $n \in \mathbb{N}$, then, since $u \in H_0^1(\Omega) \subset L^2(\Omega)$,

$$u = \sum_{k=1}^\infty (u, u_k)_{L^2(\Omega)} u_k \quad \text{in } L^2(\Omega).$$

But eigenfunctions corresponding to different eigenvalues are orthogonal, and we conclude that $u = 0$, a contradiction.

If u is an eigenfunction associated to an eigenvalue λ_n generated in the minimization process, we get that

$$u = \sum_{k:\lambda_k = \lambda_n} (u, u_k)_{L^2(\Omega)} u_k,$$

and hence u belongs to the eigenspace generated by the eigenfunctions generated in the minimization process associated to the eigenvalue λ_n. We conclude that the dimension of the eigenspace associated to any eigenvalue λ is finite, and coincides with the number of eigenvalues obtained in the minimization process that coincide with λ.

Conveniently normalized the above sequence of eigenfunctions is also a basis of $H_0^1(\Omega)$.

[2] Let H be a real Hilbert space, and $(v_k)_{k=1}^\infty$ an orthonormal sequence in H. Given $g \in H$, if $\beta_k = (g, v_k)$, then $\sum_{k=1}^\infty \beta_k^2 \le \|g\|^2$. This is *Bessel's inequality*. The proof follows easily from $\|g - \sum_{k=1}^N \beta_k v_k\|^2 \ge 0$ (check it!).

Proposition 3.4.6. *The sequence $(u_n/\lambda_n^{1/2})_{n=1}^\infty$, with (λ_n, u_n) as above, is a Hilbert orthornomal basis of $H_0^1(\Omega)$ (with the escalar product $(u,v) = \int_\Omega Du \cdot Dv$).*

Proof. The system is clearly orthornomal. Hence it is enough to check that it is *maximal*[3].

Let $f \in H_0^1(\Omega)$ be such that $(f, u_n/\lambda_n^{1/2}) = 0$ for all $n \in \mathbb{N}$. Since

$$0 = \int_\Omega Df \cdot D\frac{u_n}{\lambda_n^{1/2}} = \lambda_n^{1/2}\int_\Omega fu_n \quad \text{for all } n \in \mathbb{N},$$

and $\lambda_n > 0$, we conclude that $(f, u_n)_{L^2(\Omega)} = 0$ for all $n \in \mathbb{N}$. But $(u_n)_{n=1}^\infty$ is a Hilbert orthormal basis in $L^2(\Omega)$, hence a maximal system in $L^2(\Omega)$, and therefore $f = 0$. We conclude that $(u_n/\lambda_n^{1/2})_{n=1}^\infty$ is maximal in $H_0^1(\Omega)$. □

Corollary 3.4.7. *If $f \in H_0^1(\Omega)$, then $f = \sum_{k=1}^\infty (f, u_k)_{L^2(\Omega)} u_k$ in the sense of $H_0^1(\Omega)$, that is,*

$$\left\| f - \sum_{k=1}^N (f, u_k)_{L^2(\Omega)} u_k \right\|_{H_0^1(\Omega)} \to 0 \quad \text{as } N \to \infty.$$

Proof. We are using the scalar product $(u,v) = \int_\Omega Du \cdot Dv$.

Let $\mu_k = (f, u_k/\lambda_k^{1/2})$, $k \in \mathbb{N}$. From the previous proposition, $f = \sum_{k=1}^\infty \mu_k \frac{u_k}{\lambda_k^{1/2}}$ in the sense of $H_0^1(\Omega)$. But

$$\mu_k = (f, u_k/\lambda_k^{1/2}) = \frac{1}{\lambda_k^{1/2}}\int_\Omega Df \cdot Du_k = \lambda_k^{1/2}\int_\Omega fu_k = \lambda_k^{1/2}(f, u_k)_{L^2(\Omega)},$$

from where the result follows. □

Remark. Using Parseval's identity we get

$$\int_\Omega |Df|^2 = \|f\|^2_{H_0^1(\Omega)} = \sum_{k=1}^\infty |(f, u_k/\lambda_k^{1/2})|^2 = \sum_{k=1}^\infty \lambda_k |(f, u_k)_{L^2(\Omega)}|^2$$

for all $f \in H_0^1(\Omega)$.

Theorem 3.4.8 (Smoothness of eigenfunctions). *If $\partial\Omega \in C^\infty$, then all eigenfunctions of the homogeneous Dirichlet problem for the Laplacian belong to $C^\infty(\overline{\Omega})$.*

Proof. Since $-\Delta u = \lambda u$, and $\lambda u \in H_0^1(\Omega)$, the regularity theorem that we saw in Chapter 2 yields $u \in H^3(\Omega)$. But then, the same theorem implies $u \in H^5(\Omega)$. Iterating the argument we get that $u \in H^m(\Omega)$ for all $m \in \mathbb{N}$, and Morrey's theorem yields the result. □

Remark. If the boundary is not smooth, we still get $u \in C^\infty(\Omega)$. But regularity up to the boundary requires some smoothness of the boundary.

We now show that eigenfunctions associated to the first eigenvalue have a strict sign.

[3] An orthonormal system $(v_n)_{n=1}^\infty$ in a Hilber space H is said to be *maximal* if the only function $f \in H$ satisfying $(f, v_n)_H = 0$ is $f = 0$. It is well known that an orthonormal system is a Hilbert basis if and only if it is maximal.

Theorem 3.4.9. *Let $\Omega \subset \mathbb{R}^d$ be bounded and connected. Let $u \not\equiv 0$ be a weak solution to*

$$-\Delta u = \lambda_1 u \quad \text{in } \Omega, \qquad u = 0 \quad \text{in } \partial\Omega, \tag{3.4}$$

with $\lambda_1 = \inf\{\int_\Omega |Du|^2 / \int_\Omega u^2 : u \in H_0^1(\Omega), u \not\equiv 0\}$. Then, either $u > 0$ or $u < 0$ in Ω.

Proof. We assume without loss of generality that $\|u\|_2 = 1$. Let $u^+ = \max(u, 0)$, $u^- = -\min(u, 0)$. Let

$$\alpha = \int_\Omega (u^+)^2, \quad \beta = \int_\Omega (u^-)^2.$$

Then, $\alpha + \beta = 1$, and $u^\pm \in H_0^1(\Omega)$, with

$$Du^+ = \begin{cases} Du, & \text{a.e. in } \{u \geq 0\}, \\ 0, & \text{a.e. in } \{u \leq 0\}, \end{cases} \qquad Du^- = \begin{cases} 0, & \text{a.e. in } \{u \geq 0\}, \\ -Du, & \text{a.e. in } \{u \leq 0\}. \end{cases}$$

We have

$$\lambda_1 = \int_\Omega |Du|^2 = \int_\Omega |Du^+|^2 + \int_\Omega |Du^-|^2 \geq \lambda_1 \int_\Omega (u^+)^2 + \lambda_1 \int_\Omega (u^-)^2 = (\alpha + \beta)\lambda_1 = \lambda_1.$$

Hence, the inequality is an equality, and therefore

$$\int_\Omega |Du^+|^2 = \lambda_1 \int_\Omega (u^+)^2, \qquad \int_\Omega |Du^-|^2 = \lambda_1 \int_\Omega (u^-)^2.$$

Therefore, both u^+ and u^- solve (3.4) in the weak sense. Hence, $u^\pm \in C^\infty(\Omega)$. We conclude that $|u| = u^+ + u^- \in C^\infty(\Omega)$, and $-\Delta|u| = \lambda_1|u|$ in the classical sense. Since $\lambda_1|u| \geq 0$, $|u| \geq 0$ is superharmonic in Ω, $-\Delta|u| \geq 0$ in Ω. Therefore, $|u|(x) \geq ⨍_{B_r(x)} |u|$ if $B_r(x) \subset \Omega$. Hence, if there is some $\bar{x} \in \Omega$ such that $|u|(\bar{x}) = 0$, then $0 = |u|(\bar{x}) \geq ⨍_{B_r(\bar{x})} |u| \geq 0$, which implies that $|u| \equiv 0$ in $B_r(\bar{x})$. We have thus proved that the set $\mathcal{A} = \{x \in \Omega : u(x) = 0\}$ is open. Since u is continuous, it is also closed. Therefore, since Ω is connected, either $\mathcal{A} = \Omega$, or $\mathcal{A} = \emptyset$. The first option is not possible, since $u \not\equiv 0$, hence the result. □

Remark. In the course of the proof we have proved the *Strong Maximum Principle* for superharmonic functions: if $u \in C^2(\Omega)$ is nonnegative and superharmonic in a connected set Ω, then it is strictly positive in Ω, unless it is identically 0.

We now prove that the first eigenvalue is simple.

Theorem 3.4.10 (The first eigenvalue is simple). *Let $\Omega \subset \mathbb{R}^d$ be bounded and connected. Let $u \not\equiv 0$ be a weak solution to* (3.4) *with $\lambda_1 = \inf\{\int_\Omega |Du|^2 / \int_\Omega u^2 : u \in H_0^1(\Omega), u \not\equiv 0\}$. Then, $u = \alpha u_1$ for some $\alpha \in \mathbb{R}$, $\alpha \neq 0$.*

Proof. Since u_1 has a strict sign in Ω, we know that $\int_\Omega u_1 \neq 0$. Therefore, there exists a value $\chi \in \mathbb{R}$, $\chi \neq 0$, such that $\int_\Omega u = \chi \int_\Omega u_1$ (remember that also u has a strict sign). Then $\int_\Omega (u - \chi u_1) = 0$. But this is not possible if $u - \chi u_1 \not\equiv 0$, since $u - \chi u_1$ is also a weak solution to (3.4) and hence has a strict sign in Ω if it is not identically 0. □

Let us now apply the above theory to the study of a nonlinear problem.

Let $\Omega \subset \mathbb{R}^d$ be bounded and smooth, and $a \in C^\infty(\overline{\Omega})$. We consider the Dirichlet problem

$$-\Delta u = a|u|^{p-1}u \quad \text{in } \Omega, \qquad u = 0 \quad \text{in } \partial\Omega. \tag{3.5}$$

This problem has a classical solution, namely $u = 0$. Are there others (maybe weak)? If $a \le 0$ this is not possible. Indeed, using the solution itself as test function we get

$$\int_\Omega (|Du|^2 - a|u|^{p+1}) = 0.$$

Therefore, if $a \le 0$, then $Du = 0$ almost everywhere in Ω. Hence, u is constant in each connected component of Ω. Since $u \in H_0^1(\Omega)$, then it has to be identically 0. What happens if $a > 0$ in Ω? (We do not consider here the interesting case in which a changes sign.)

Let $F, G : H_0^1(\Omega) \to \mathbb{R}$ be given by $F[u] = \frac{1}{2}\int_\Omega |Du|^2$, $G[u] = \left(\int_\Omega a|u|^{p+1}\right) - 1$. We want to minimize F in the 0-level set of G, $\mathcal{C} = \{u \in H_0^1(\Omega) : G[u] = 0\}$. Notice that the elements of $\mathcal{C}$ are nontrivial.

From now on we assume that $d \ge 3$ (it is a good exercise to analyze the cases $d = 1, 2$). Since $H_0^1(\Omega) \subset L^{2^*}(\Omega) = L^{\frac{2d}{d-2}}(\Omega)$, then G is well defined if $2 \le p + 1 \le \frac{2d}{d-2}$, that is, if $p \in [1, (d+2)/(d-2)]$. Under this restriction we already know that $F, G \in C^1(X; \mathbb{R})$.

It is easy to check that $\mathcal{C} \ne \emptyset$. Indeed, take any $\bar{u} \in H_0^1(\Omega)$, $\bar{u} \not\equiv 0$. Then

$$G[\lambda\bar{u}] = \lambda^{p+1}\int_\Omega a|\bar{u}|^{p+1} - 1 = 0$$

if we take $\lambda = (\int_\Omega a|\bar{u}|^{p+1})^{-1/(p+1)}$. Therefore, $I = \inf\{F[u] : u \in \mathcal{C}\} \ge 0$. Let $(u_k)_{k=1}^\infty$ be a minimizing sequence in $\mathcal{C}$, $u_k \in \mathcal{C}$, $F[u_k] \to I$. Then, $I \le F[u_k] \le I + 1$ if k is large, from where we deduce that $\|u_k\|_{H_0^1(\Omega)} \le K < \infty$. Since $H_0^1(\Omega)$ is reflexive, there are a subsequence $(u_{k_j})_{j=1}^\infty$ and a function $u \in H_0^1(\Omega)$ such that $u_{k_j} \rightharpoonup u$ in $H_0^1(\Omega)$. Since F is sequentially weakly lower semicontinuous (it is a norm in a Hilbert space), the weak convergence implies that $F[u] = I$. On the other hand, if $p \in [1, (d+2)/(d-2))$, then $p + 1 \in [2, 2d/(d-2))$. Hence, Rellich-Kondrachov's compactness theorem implies that there is a subsequence $(u_{k_{j_l}})_{j=1}^\infty$ such that $u_{k_{j_l}} \to u$ in $L^{p+1}(\Omega)$, and hence $G[u_{k_{j_l}}] \to G[u]$. Therefore, since $G[u_{k_{j_l}}] = 0$, $G[u] = 0$, that is, $u \in \mathcal{C}$. Notice that

$$G'[u]u = (p+1)\int_\Omega a|u|^{p+1} > 0.$$

Hence, Lagrange multipliers' theorem implies the existence of a constant $\mu \in \mathbb{R}$ such that $F'[u]v = \mu G'[u]v$ for all $v \in H_0^1(\Omega)$, that is,

$$\int_\Omega Du \cdot Du = \mu(p+1)\int_\Omega a|u|^{p-1}uv \quad \text{for all } v \in H_0^1(\Omega).$$

Let $w = \tau u$. Then,

$$\frac{1}{\tau}\int_\Omega Dw \cdot Dv = \frac{\mu(p+1)}{\tau^p}\int_\Omega |w|^{p-1}wv.$$

Taking $\tau = (\mu(p+1))^{1/(p-1)}$ (this requires $p > 1$), we have $\frac{\mu(p+1)}{\tau^{p-1}} = 1$, and hence $w \in H_0^1(\Omega)$ is a weak solution to (3.5). We still have to check that it is nontrivial which will be immediate if $\mu > 0$. To show that we take $v = u$ as test function and we obtain $\int_\Omega |Du|^2 = \mu(p+1)\int_\Omega a|u|^{p+1}$. This implies that $\mu \geq 0$. Moreover, if $\mu = 0$, then $Du = 0$ almost everywhere in Ω and hence u is constant in each connected component of Ω. Since $u \in H_0^1(\Omega)$, then $u \equiv 0$, a contradiction with $u \in \mathcal{C}$.

Summarizing, we have proved that problem (3.5) with $a > 0$ in Ω has a nontrivial weak solution if $p \in (1, (d+2)/(d-2))$.

Remark. The upper restriction is not technical. Indeed, by means of the so-called Derrick-Pohozaev's identity one can prove that the problem does not admit a nontrivial solution if $a \equiv 1$, $p > (d+2)/(d-2)$ and the domain is a ball.

3.5 Saddle points: Mountain Pass Theorem

Besides local or global minima and maxima, there are other types of critical points: saddle-points.

Example. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x^2 - y^2$. The point $(0, 0)$ is a critical point of f, that is, $Df(0, 0) = 0$. However, it is neither a local maximum nor a local minimum. The graph of f looks like a horse saddle, hence $(0, 0)$ is called a saddle-point. If we move along the graph in the direction of the x-axis, $(0, 0)$ is the minimum, while if we move along the graph in the direction of the y-axis, $(0, 0)$ is the maximum. For this reason, we also call $(0, 0)$ a mini-max critical point of f.

Idea to find saddle-points. Given two points separated by a *mountain range*, there is a *mountain pass* that allows to go from one of the points to the other "climbing the less possible".

Let us formulate this idea mathematically. Let X be a real Banach space and let $I : X \to \mathbb{R}$ be a C^1 functional such that

(i) $I[0] = 0$;

(ii) there exist constants $\rho, \alpha > 0$ such that $I\Big|_{\partial B_\rho(0)} \geq \alpha$;

(iii) there exists $\bar{x} \in X \setminus \overline{B_\rho(0)}$ such that $I[\bar{x}] \leq 0$.

We consider a path connecting 0 and $\bar{x}$,

$$\gamma : [0, 1] \to X \quad \text{continuous, such that } \gamma(0) = 0,\ \gamma(1) = \bar{x}.$$

Since γ has to cross $\partial B_\rho(0)$,

$$\max_{t\in[0,1]} I[g(t)] \geq \alpha.$$

Let $\Gamma = \{\gamma \in C([0, 1]; X) : \gamma(0) = 0, \gamma(1) = \bar{x}\}$. We consider

$$c := \inf_{\gamma\in\Gamma} \max_{t\in[0,1]} I[\gamma(t)] \geq \alpha.$$

In finite dimensional spaces, the critical value c is achieved for some $\gamma_0 \in \Gamma$ and $t_0 \in [0, 1]$. What is more important, $x_0 := \gamma_0(t_0)$ is a critical point of I, $I'[x_0] = 0$ (the tangent plane at a mountain pass is horizontal). The proof uses strongly that bounded sequences have a convergent subsequence. This is no longer true in infinite dimensional Banach spaces. In order to be able to follow this approach, we need to require that the functional satisfies some compactness condition.

Definition 3.5.1. We say that the functional I satisfies the *Palais-Smale condition* (in short, (PS)), if any sequence $(u_k)_{k=1}^{\infty}$ in X for which $I[u_k]$ is bounded and $I'[u_k] \to 0$ in X^* as $k \to \infty$ possesses a subsequence converging strongly in X.

Notation. Sequences like the ones appearing in the definition are known as *Palais-Smale sequences* (in short, PS sequence).

Example. Let $\Omega \subset \mathbb{R}^d$ be bounded and smooth. Given $f \in L^2(\Omega)$, we consider the functional $I : H_0^1(\Omega) \to \mathbb{R}$ given by $I[u] = \frac{1}{2}\int_\Omega |Du|^2 - \int_\Omega fu$. We want to check that I satisfies the condition (PS). We will take $(u,v) = \int_\Omega Du \cdot Dv$ as inner product in $H_0^1(\Omega)$.

Let $(u_k)_{k=1}^{\infty}$ be a PS sequence. On the one hand, $C \geq I[u_k] = \frac{1}{2}\|u_k\|^2_{H_0^1(\Omega)} - \int_\Omega fu_k$. On the other hand, using the inequalities of Hölder, Poincaré, and Peter-Paul,

$$\int_\Omega fu_k \leq \|f\|_{L^2(\Omega)}\|u_k\|_{L^2(\Omega)} \leq C\|u_k\|_{H_0^1(\Omega)} \leq \varepsilon\|u_k\|^2_{H_0^1} + C_\varepsilon.$$

Therefore, $C \geq I[u_k] \geq \left(\frac{1}{2} - \varepsilon\right)\|u_k\|^2_{H_0^1(\Omega)} - C_\varepsilon$. Taking $\varepsilon = 1/4$, we conclude that $\|u_k\|_{H_0^1(\Omega)} \leq K < \infty$. Using Rellich-Kondrachov's compactness theorem and the reflexivity of the space, we get that $(u_k)_{k=1}^{\infty}$ has a subsequence $(u_{k_j})_{j=1}^{\infty}$ such that $u_{k_j} \to u$ in $L^2(\Omega)$, $u_{k_j} \rightharpoonup u$ in $H_0^1(\Omega)$. This is not enough for our purposes. But we have not yet used that $I'[u_k] \to 0$ in $H^{-1}(\Omega)$.

Let $v \in H_0^1(\Omega)$. Then

$$\left|\int_\Omega (Du_{k_j} \cdot Dv - fv)\right| = |I'[u_{k_j}]v| \leq \|I'[u_{k_j}]\|_{H^{-1}(\Omega)}\|v\|_{H_0^1(\Omega)}.$$

Taking $v = u_{k_j}$, this yields

$$\left|\int_\Omega (|Du_{k_j}|^2 - fu_{k_j})\right| = |I'[u_{k_j}]u_{k_j}| \leq \|I'[u_{k_j}]\|_{H^{-1}(\Omega)}\|u_{k_j}\|_{H_0^1(\Omega)} \leq K\|I'[u_{k_j}]\|_{H^{-1}(\Omega)} \to 0.$$

If we take $v = u$ instead, we get

$$\left|\int_\Omega (Du_{k_j} \cdot Du - fu)\right| \leq \|I'[u_{k_j}]\|_{H^{-1}(\Omega)}\|u\|_{H_0^1(\Omega)} \to 0.$$

But, because of the weak convergence, we know that

$$\int_\Omega Du_{k_j} \cdot Du \to \int_\Omega |Du|^2,$$

and we conclude that $\int_\Omega (|Du|^2 - fu) = 0$. Since

$$\left|\|u_{k_j}\|^2_{H_0^1(\Omega)} - \|u\|^2_{H_0^1(\Omega)}\right| \leq \left|\int_\Omega (|Du_{k_j}|^2 - fu_{k_j})\right| + \left|\int_\Omega (fu - |Du|^2)\right| + \left|\int_\Omega f(u_{k_j} - u)\right|,$$

the above information plus the strong convergence of u_{k_j} in $L^2(\Omega)$ yields $\|u_{k_j}\|_{H_0^1(\Omega)} \to \|u\|_{H_0^1(\Omega)}$. Weak convergence plus convergence of the norms finally yields the desired strong convergence in $H_0^1(\Omega)$. Indeed,

$$\begin{aligned}\|u_{k_j} - u\|^2_{H_0^1(\Omega)} &= (u_{k_j} - u, u_{k_j} - u)_{H_0^1(\Omega)} = \|u_{k_j}\|^2_{H_0^1(\Omega)} - 2(u_{k_j}, u)_{H_0^1(\Omega)} + \|u\|^2_{H_0^1(\Omega)}\\ &\to \|u\|^2_{H_0^1(\Omega)} - 2\|u\|^2_{H_0^1(\Omega)} + \|u\|^2_{H_0^1(\Omega)} = 0.\end{aligned}$$

Hence condition (PS) holds.

Using the (PS) compactness condition, Ambrosetti and Rabinowitz established in [1] the following well-known theorem on the existence of a mini-max critical point.

Theorem 3.5.1 (Mountain Pass Theorem). *Let X be a real Banach space. If $I \in C^1(X;\mathbb{R})$ satisfies conditions* (i)–(iii) *and* (PS)*, then I has a critical point x_0 for which*

$$I[x_0] = \inf_{\gamma\in\Gamma} \max t \in [0,1] I[\gamma(t)].$$

We don't give here the proof of the theorem. The interested student can find it, for instance, in [3].

Remark. $I[x_0] \geq \alpha > 0$ implies that $x_0 \not\equiv 0$. So, this method gives us nontrivial solutions.

Example. We consider once more the Dirichlet problem (3.5) with $a \in C(\overline{\Omega})$, $a > 0$ in $\overline{\Omega}$. We know that $u \equiv 0$ is a classical solution. We want to find a second weak solution, that is, a nontrivial function $u \in H_0^1(\Omega)$ such that

$$\int_\Omega (Du \cdot Dv - a|u|^{p-1}uv) = 0 \quad \text{for all } v \in H_0^1(\Omega).$$

We will do this by means of the Mountain Pass Theorem. As we will check later, this is the Euler-Lagrange equation of the functional $I : H_0^1(\Omega) \to \mathbb{R}$, $I[u] = I_1[u] - I_2[u]$, where $I_1, I_2 : H_0^1(\Omega) \to \mathbb{R}$ are given by $I_1[u] = \frac{1}{2}\|Du\|^2_{H_0^1(\Omega)}$, $I_2[u] = \frac{1}{p+1}\int_\Omega a|u|^{p+1}$.

We will do the analysis for $d \geq 3$. It is good exercise to do the computations for the cases $d = 1, 2$.

Step 1: The functional is well defined. Let $p \geq 1$ be such that $p+1 \leq 2^* = \frac{2d}{d-2}$. Then, applying Hölder's inequality and the GNS Sobolev embedding,

$$\int_\Omega a|u|^{p+1} \leq \|a\|_{L^\infty(\Omega)} \int_\Omega |u|^{p+1} \leq C\|u\|^{p+1}_{2^*} \leq C\|u\|^{p+1}_{H_0^1(\Omega)}.$$

Therefore, the functional is well defined if $1 \leq p \leq \frac{2d}{d-2} - 1 = \frac{d+2}{d-2}$.

Step 2: Computation of directional derivatives. Let $u, v \in H_0^1(\Omega)$. Then, $\partial_v I[u] = \lim_{\varepsilon\to 0} \frac{I[u+\varepsilon v]-I[u]}{\varepsilon}$ if the limit exists. We have on the one hand

$$\partial_v I_1[u] = \lim_{\varepsilon\to 0} \frac{\frac{1}{2}\int_\Omega |Du|^2 + \varepsilon\int_\Omega Du\cdot Dv + \frac{\varepsilon^2}{2}\int_\Omega |Dv|^2 - \frac{1}{2}\int_\Omega |Du|^2}{\varepsilon} = \int_\Omega Du\cdot Dv.$$

On the other hand

$$\partial_v I_2[u] = \lim_{\varepsilon\to 0} \frac{\int_\Omega \left(a(|u+\varepsilon v|^{p+1} - |u|^{p+1})\right)}{(p+1)\varepsilon}.$$

Let $g(t) = \frac{|t|^{p+1}}{p+1}$ with $p \geq 1$. Then $g'(t) = |t|^{p-1}t$ and $g''(t) = p|t|^{p-1}$. Therefore, performing a Taylor expansion around u,

$$\frac{|u+\varepsilon v|^{p+1}}{p+1} = \frac{|u|^{p+1}}{p+1} + |u|^{p-1}u\varepsilon v + \frac{p|u+\theta\varepsilon v|^{p-1}\varepsilon^2 v^2}{2}$$

for some $\theta \in [0,1]$. Hence,

$$\partial_v I_2[u] = \int_\Omega (a|u|^{p-1}uv) + \lim_{\varepsilon\to 0}\frac{\varepsilon p}{2}\int_\Omega \left(a|u+\theta\varepsilon v|^{p-1}v^2\right) = \int_\Omega (a|u|^{p-1}uv)$$

if we can bound $\int_\Omega \left(a|u+\theta\varepsilon v|^{p-1}v^2\right)$ by a constant independent of ε, $|\varepsilon| \leq 1$. To obtain the required bound we use that for all $\alpha, \beta, r > 0$ we have $(\alpha+\beta)^r \leq 2^r(a^r+b^r)$. Therefore,

$$\int_\Omega \left(a|u+\theta\varepsilon v|^{p-1}v^2\right) \leq \|a\|_{L^\infty(\Omega)}2^{p-1}\int_\Omega \left((|u|^{p-1}+|v|^{p-1})v^2\right) \leq C\Big(\int_\Omega |u|^{p-1}v^2+\int_\Omega |v|^{p+1}\Big).$$

On the one hand, if $p \leq \frac{d+2}{d-2}$ then $p+1 \leq 2^*$. Hence, Sobolev's embedding yields $\int_\Omega |v|^{p+1} \leq C\big(\int_\Omega |v|^{2^*}\big)^{\frac{p+1}{2^*}} \leq C\|v\|^{p+1}_{H^1_0(\Omega)}$. On the other hand, using Hölder's inequality twice, and Sobolev's embedding,

$$\int_\Omega |u|^{p-1}v^2 \leq \|u\|^{p-1}_{\frac{d(p-1)}{2}}\|v\|^2_{2^*} \leq C\|u\|^{p-1}_{2^*}\|v\|^2_{2^*} \leq C\|u\|^{p-1}_{H^1_0(\Omega)}\|v\|^2_{H^1_0(\Omega)},$$

which completes the desired bound.

We conclude that $\partial_v I[u] = \int_\Omega (Du\cdot Dv - a|u|^{p-1}uv)$.

Step 3: *I* is Gâteaux differentiable. Euler-Lagrange equation. We have to check that for every $u \in H^1_0(\Omega)$ fixed the application $v \mapsto \partial_v I[u]$ is linear and bounded (hence continuous). Linearity is immediate. As for boundedness,

$$\begin{aligned}
|\partial_v I_1[u]| &= \left|\int_\Omega Du\cdot Dv\right| \leq \|Du\|_2\|Dv\|_2 = \|u\|_{H^1_0(\Omega)}\|v\|_{H^1_0(\Omega)},\\
|\partial_v I_2[u]| &= \left|\int_\Omega a|u|^{p-1}uv\right| \leq \|a\|_\infty\int_\Omega |u|^p|v| \leq C\|u\|^p_{\frac{2pd}{d+2}}\|v\|_{2^*}\\
&\leq C\|u\|^p_{2^*}\|v\|_{2^*} \leq C\|u\|^p_{H^1_0(\Omega)}\|v\|_{H^1_0(\Omega)}.
\end{aligned}$$

We conclude that $v \mapsto \partial_v I[u]$ is bounded, and hence that I is Gâteaux differentiable, with $I'[u]v = \partial_v I[u] = \int_\Omega (Du\cdot Dv - a|u|^{p-1}uv)$. Then, the Euler-Lagrange equation associated to I is

$$\int_\Omega (Du\cdot Dv - a|u|^{p-1}uv) = 0 \quad \text{for all } v \in H^1_0(\Omega),$$

$(u \in H^1_0(\Omega))$ which is nothing but the weak formulation of problem (3.5).

Step 4: *I* is Fréchet differentiable. A straightforward computation shows that

$$I_1[u+v] - I_1[u] - I_1'[u]v = \frac{1}{2}\|Dv\|_2^2 = \frac{1}{2}\|v\|^2_{H^1_0(\Omega)} = o(\|v\|_{H^1_0(\Omega)}) \quad \text{as } \|v\|_{H^1_0(\Omega)} \to 0.$$

On the other hand,

$$I_2[u+v] - I_2[u] - I_2'[u]v = \frac{1}{p+1}\int_\Omega \left(a(|u+v|^{p+1} - |u|^{p+1}\right) - \int_\Omega \left(a|u|^{p-1}uv\right).$$

Doing a Taylor expansion as in Step 2, and arguing as there,

$$\begin{aligned}I_2[u+v]-I_2[u]-I_2'[u]v&=\frac{\varepsilon p}{2}\int_\Omega\Big(a|u+\theta v|^{p-1}v^2\Big)\le C\Big(\int_\Omega|u|^{p-1}v^2+\int_\Omega|v|^{p+1}\Big)\\&\le C\big(\|u\|^{p-1}_{H_0^1(\Omega)}\|v\|^2_{H_0^1(\Omega)}+\|v\|^{p+1}_{H_0^1(\Omega)}\big)=o(\|v\|_{H_0^1(\Omega)})\end{aligned}$$

as $\|v\|_{H_0^1(\Omega)}\to 0$.

Since both I_1 and I_2 are Fréchet differentiable, the same is true for $I=I_1-I_2$.

Step 5: $I\in C^1(H_0^1(\Omega);\mathbb{R})$. Let $u,v,w\in H_0^1(\Omega)$ Then,

$$|(I_1'[u]-I_1'[v])w|=\left|\int_\Omega D(u-v)\cdot Dw\right|\le\|D(u-v)\|_2\|Dw\|_2=\|u-v\|_{H_0^1(\Omega)}\|w\|_{H_0^1(\Omega)}.$$

Therefore, $\|I_1'[u]-I_1'[v]\|_{H^{-1}(\Omega)}\le\|u-v\|_{H_0^1(\Omega)}$, which shows that $I_1\in C^1(H_0^1(\Omega);\mathbb{R})$.

As for I_2,

$$|(I_2'[u]-I_2'[v])w|=\left|\int_\Omega\big(a(|u|^{p-1}u-|v|^{p-1}v)w\big)\right|\le\|a\|_\infty\int_\Omega\big(\big||u|^{p-1}u-|v|^{p-1}v\big||w|\big).$$

Then, using the Mean Value Theorem, we have for some $\theta\in[0,1]$

$$\begin{aligned}\int_\Omega\big(\big||u|^{p-1}u-|v|^{p-1}v\big||w|\big)&=p\int_\Omega\big(|\theta u+(1-\theta)v|^{p-1}|u-v||w|\big)\\&\le C\int_\Omega\big((|u|^{p-1}+|v|^{p-1})|u-v||w|\big)\\&\le C\|(|u|^{p-1}+|v|^{p-1})|u-v|\|_{\frac{2d}{d+2}}\|w\|_{2^*}\\&\le C\||u|^{p-1}+|v|^{p-1}\|_{\frac{d}{2}}\|u-v\|_{2^*}\|w\|_{2^*}\\&\le C(\||u|^{p-1}\|_{\frac{d}{2}}+\||v|^{p-1}\|_{\frac{d}{2}})\|u-v\|_{2^*}\|w\|_{2^*}\\&\le C(\|u\|^{p-1}_{2^*}+\|v\|^{p-1}_{2^*})\|u-v\|_{2^*}\|w\|_{2^*}\\&\le C(\|u\|^{p-1}_{H_0^1(\Omega)}+\|v\|^{p-1}_{H_0^1(\Omega)})\|u-v\|_{H_0^1(\Omega)}\|w\|_{H_0^1(\Omega)}.\end{aligned}$$

Therefore, $\|I_2'[u]-I_2'[v]\|_{H^{-1}(\Omega)}\le C(\|u\|^{p-1}_{H_0^1(\Omega)}+\|v\|^{p-1}_{H_0^1(\Omega)})\|u-v\|_{H_0^1(\Omega)}$, which shows that $I_2\in C^1(H_0^1(\Omega);\mathbb{R})$. We conclude that $I\in C^1(H_0^1(\Omega);\mathbb{R})$

Step 6: Geometric conditions.

6.1. It is trivial to check that $I[0]=0$.

6.2. Using Sobolev's embedding we get

$$I[v]=\frac12\|v\|^2_{H_0^1(\Omega)}-\int_\Omega(a|v|^{p+1})\ge\frac12\|v\|^2_{H_0^1(\Omega)}-\|a\|_\infty\|v\|^{p+1}_{p+1}\ge\frac12\|v\|^2_{H_0^1(\Omega)}-C\|v\|^{p+1}_{H_0^1(\Omega)}.$$

Let $f(\rho)=\frac12\rho^2-C\rho^{p+1}$. If $p>1$, the cuadratic term dominates for small values of ρ, $f(\rho)\ge\frac14\rho^2$ if $\rho\in[0,\rho_0]$ for some small $\rho_0>0$. We conclude that there exists a constant $\bar\rho\in(0,\rho_0)$ such that $I\Big|_{\partial B_{\bar\rho}(0)}\ge\alpha$ for some constant $\alpha>0$.

6.3. Let us fix some nontrivial $\bar u\in H_0^1(\Omega)$. Then,

$$I[\lambda\bar u]=\frac{\lambda^2}{2}\|\bar u\|^2_{H_0^1(\Omega)}-\lambda^{p+1}\int_\Omega(a|\bar u|^{p+1})<0$$

if λ is positive and large enough. We can moreover take λ large so that $\|\lambda\bar{u}\|_{H_0^1(\Omega)} > \bar{\rho}$.

STEP 7: (PS) CONDITION. Let $(u_k)_{k=1}^\infty$ be a PS sequence, that is, $I[u_k]$ is bounded and $I'[u_k] \to 0$ in $H^{-1}(\Omega)$. We have to prove that there is a strongly convergent subsequence.

7.1. *The sequence is bounded.* We have

$$C \geq I[u_k] = \frac{1}{2}\|u_k\|^2_{H_0^1(\Omega)} - \frac{1}{p+1}\int_\Omega \big(a|u_k|^{p+1}\big).$$

Thus,

$$\frac{1}{2}\|u_k\|^2_{H_0^1(\Omega)} \leq \frac{1}{p+1}\int_\Omega \big(a|u_k|^{p+1}\big) + C. \tag{3.6}$$

On the other hand, since $I'[u_k] \to 0$ in $H^{-1}(\Omega)$,

$$\left|\int_\Omega \big(Du_k \cdot Dv - a|u_k|^{p-1}u_k v\big)\right| = |I'[u_k]v| \leq \|I'[u_k]\|_{H^{-1}(\Omega)}\|v\|_{H_0^1(\Omega)} \leq \|v\|_{H_0^1(\Omega)}$$

for all k large and $v \in H_0^1(\Omega)$. Therefore, taking $v = u_k$,

$$\left|\|u_k\|^2_{H_0^1(\Omega)} - \int_\Omega \big(a|u_k|^{p+1}\big)\right| \leq \|u_k\|_{H_0^1(\Omega)},$$

and hence

$$\int_\Omega \big(a|u_k|^{p+1}\big) \leq \|u_k\|^2_{H_0^1(\Omega)} + \|u_k\|_{H_0^1(\Omega)}.$$

Combining this estimate with (3.6) we obtain

$$\left(\frac{1}{2} - \frac{1}{p+1}\right)\|u_k\|^2_{H_0^1(\Omega)} - \frac{1}{p+1}\|u_k\|_{H_0^1(\Omega)} \leq C.$$

Therefore, if $p > 1$, necessarily $\|u_k\|_{H_0^1(\Omega)} \leq K < \infty$.

7.2. *Strongly convergent subsequence.* Since $\|u_k\|_{H_0^1(\Omega)} \leq K$, there are a subsequence $(u_{k_j})_{j=1}^\infty$ and a function $u \in H_0^1(\Omega)$ such that $u_{k_j} \rightharpoonup u$ in $H_0^1(\Omega)$, $u_{k_j} \to u$ in $L^{p+1}(\Omega)$. We are assuming that $p < (d+2)/(d-2)$. Hence $p+1 < 2^*$.

We claim that $a|u_{k_j}|^{p-1}u_{k_j}$ converges in $H^{-1}(\Omega)$ towards $a|u|^{p-1}u$. Indeed, using the Mean Value Theorem we get

$$\begin{aligned}
\left|\int_\Omega \big(a(|u_k|^{p-1}u_{k_j} - |u|^{p-1}u)v\big)\right| &\leq \|a\|_\infty\, p\int_\Omega \big(|v|(|u_{k_j}|^{p-1} + |v|^{p-1})|u_{k_j} - u|\big)\\
&\leq C\|v(|u_{k_j}|^{p-1} + |v|^{p-1})\|_{\frac{p+1}{p}}\|u_{k_j} - u\|_{p+1}\\
&\leq C\||u_{k_j}|^{p-1} + |v|^{p-1}\|_{\frac{p+1}{p-1}}\|v\|_{p+1}\|u_{k_j} - u\|_{p+1}\\
&\leq C(\|u_{k_j}\|^{p-1}_{p+1} + \|v\|^{p-1}_{p+1})\|v\|_{p+1}\|u_{k_j} - u\|_{p+1}\\
&\leq C(\|u_{k_j}\|^{p-1}_{H_0^1(\Omega)} + \|v\|^{p-1}_{H_0^1(\Omega)})\|v\|_{H_0^1(\Omega)}\|u_{k_j} - u\|_{H_0^1(\Omega)}\\
&\leq C\|u_{k_j} - u\|_{H_0^1(\Omega)},
\end{aligned}$$

(the constant depends on v, but not on k_j), which proves the claim.

We now observe that

$$I'[u_{k_j}]v = \int_\Omega \big(Du\cdot Dv - a|u_{k_j}|^{p-1}u_{k_j}v\big) = \langle -\Delta u_{k_j} - a|u_{k_j}|^{p-1}u_{k_j}, v\rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)},$$

that is, $-\Delta u_{k_j} = I'[u_{k_j}]+a|u_{k_j}|^{p-1}u_{k_j}$ in $H^{-1}(\Omega)$. But we know that $\Delta : H_0^1(\Omega)\to H^{-1}(\Omega)$ is an isometry. Therefore, using the claim,

$$u_{k_j} = -\Delta^{-1}(I'[u_{k_j}]) - \Delta^{-1}(a|u_{k_j}|^{p-1}u_{k_j}) \to -\Delta^{-1}(a|u|^{p-1}u) \quad \text{in } H_0^1(\Omega),$$

so that the subsequence is strongly convergent.

Conclusion. The hypotheses of the Mountain Pass Theorem are fulfilled, and hence the functional has a nontrivial critical point.

Bibliography

[1] Ambrosetti, A.; Rabinowitz, P. H. *Dual variational methods in critical point theory and applications.* J. Functional Analysis 14 (1973), 349–381.

[2] Brezis, H. "Functional analysis, Sobolev spaces and partial differential equations". Universitext. Springer, New York, 2011. ISBN: 978-0-387-70913-0.

[3] Chen, W.; Li, C. "Methods on nonlinear elliptic equations". AIMS Series on Differential Equations & Dynamical Systems, 4. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2010. ISBN: 978-1-60133-006-2; 1-60133-006-5.

[4] Evans, L. C. "Partial differential equations". Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010. ISBN: 978-0-8218-4974-3.

[5] Folland, G. B. "Real analysis. Modern techniques and their applications". Second edition. Pure and Applied Mathematics (New York). A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999. ISBN: 0-471-31716-0.

[6] Leoni, G. "A first course in Sobolev spaces. Second edition". Graduate Studies in Mathematics, 181. American Mathematical Society, Providence, RI, 2017. ISBN: 978-1-4704-2921-8.

[7] Peral, I. Curso "Métodos variacionales y ecuaciones en derivadas parciales". Universidad de Almería, 1998. Disponible en `http://matematicas.uam.es/ ireneo.peral/ALMERIA1.pdf`.

[8] Spivak, M. "Calculus on manifolds. A modern approach to classical theorems of advanced calculus". W. A. Benjamin, Inc., New York-Amsterdam 1965.