C4.3 Functional Analytic Methods for PDEs

Luc Nguyen

Michaelmas 2019

This set of lecture notes builds upon Gregory Seregin's lecture notes who taught the course in previous years. The following literature was also used (either for this set of notes, or for my predecessor's):

L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, 19, American Mathematical Society, 2010.

E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics, 14, American Mathematical Society, 2001.

H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer, 2011.

R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Dekker, 1977.

R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Pure and Applied Mathematics 140. Elsevier/Academic Press, 2003.

P. D. Lax, *Functional Analysis*, Wiley, 2002.

D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Classics in Mathematics, Springer, 2001.

Preface

In this set of lecture notes, we will be concerned with linear partial differential equations of the form

$$Lu = -\partial_i(a_{ij}\partial_j u) + b_i\partial_i u + cu = f + \partial_i g_i \text{ in } \Omega. \quad (1)$$

Here Ω is a domain in $\mathbb{R}^n$, $u : \Omega \to \mathbb{R}$ is the unknown, $(a_{ij}) = (a_{ji})$, (b_i) and c are given coefficients, f and g_i are given sources, and repeated indices are summed from 1 to n. The coefficients (a_{ij}) are assumed to be uniformly elliptic, i.e. there exists $\Lambda > 1$ such that

$$\frac{1}{\Lambda}|\xi|^2 \le a_{ij}(x)\xi_i\xi_j \le \Lambda|\xi|^2 \text{ for all } x \in \Omega, \xi \in \mathbb{R}^n.$$

In order to solve (1), one needs to supplement it with a boundary condition. Here we will only consider an important boundary condition called the Dirichlet boundary condition

$$u = u_0 \text{ on } \partial\Omega \quad (2)$$

where u_0 is a given function.

When the coefficients and the sources are sufficiently nice, a classical solution to (1)-(2) is a function $u \in C^2(\Omega) \cap C(\bar{\Omega})$ such that (1)-(2) are satisfied in the usual sense.

If we multiply (2) by a function $\varphi \in C^1(\bar{\Omega})$ with $\varphi = 0$ on $\partial\Omega$ and integrate by parts over Ω, we get

$$\int_\Omega [a_{ij}\partial_j u\partial_i\varphi + b_i\partial_i u\varphi + cu\varphi] = \int_\Omega [f\varphi - g_i\partial_i\varphi]. \quad (3)$$

It is important to note that (3) makes sense for $u \in C^1(\Omega)$ which is in contrast with (1) which requires two derivatives. In fact, all it requires are that u and $\partial_i u$ are integrable. Now if $u \in C^1(\Omega)$ is such that (3) holds for all $\varphi \in C^1(\bar{\Omega})$ with $\varphi = 0$ on $\partial\Omega$, we say that u is a *weak* solution to (1).

The introduction of weak solutions is not merely a methodological matter. In many physical applications, be it linear like (1) or nonlinear, classical solutions need not exist. For example, in problems arising in composite materials, the coefficients a_{ij} does not have to be even continuous, and the notion of classical solutions to (1) becomes obscured.

The so-called *variational approach* to partial differential equation (of the kind (1)-(2)) roughly consists of 3 stages:

- One makes precise the notion of weak solutions, and in particular the functional spaces – Sobolev spaces in this course – in which solutions live.
- One establishes existence (and uniqueness) of weak solutions.
- One studies if weak solutions have better regularity than what was preset in the definition of weak solutions. For example, one would like to understand if, for nice coefficients and sources, are weak solutions to (1)-(2) classical?

As this course is an introduction to the field, I have no intention of being thorough. In fact, I have deliberately cut out or over-simplified a number of important topics to better illustrate other important points. For a more complete treatment, students are encouraged to consult the texts mentioned at the beginning of this set of lecture notes.

Contents

Chapter 1

Lebesgue Spaces

1.1 Definition of Lebesgue spaces

Let E be a measurable subset of $\mathbb{R}^n$. For $1 \le p < \infty$, we let $\mathcal{L}^p(E)$ denote the space of measurable functions $f : E \to \mathbb{R}$ for which $\int_E |f|^p\,dx$ is finite, i.e.

$$\mathcal{L}^p(E) = \Big\{ f : E \to \mathbb{R} \Big|\ f \text{ is measurable on } E \text{ and } \int_E |f|^p\,dx < \infty \Big\}.$$

We let $L^p(E)$ denote the set of all equivalence classes in $\mathcal{L}^p(E)$ under the equivalence relation

$$f \sim g \text{ if } f = g \text{ a.e. in } E. \tag{1.1}$$

Functions belonging to $L^p(E)$ is sometimes referred to as p-integrable functions.

When it is clear from the context what E is, we will write $\mathcal{L}^p$ and L^p in place of $\mathcal{L}^p(E)$ and $L^p(E)$, respectively. Let

$$\|f\|_{L^p(E)} = \Big[\int_E |f|^p\,dx \Big]^{1/p} \qquad (1 \le p < \infty),$$

so that $L^p(E)$ consists of [equivalence classes of] measurable functions f for which $\|f\|_{L^p(E)}$ is finite.

When $p = \infty$, we define $L^\infty(E)$ as follows. For a measurable set E of positive measure and a measurable function f defined on E, define the essential supremum of f on E by

$$\operatorname*{ess\,sup}_E f = \inf\{c > 0 : f \le c \text{ a.e. in } E\}.$$

A measurable function f is said to be essentially bounded, or simply bounded, on E if $\operatorname{ess\,sup}_E |f|$ is finite. The set of all essentially bounded measurable functions on E

is denoted by $\mathcal{L}^\infty(E)$. The set of equivalence classes of $\mathcal{L}^\infty(E)$ under the equivalence relation (1.1) is denoted by $L^\infty(E)$.

For simplicity, instead of saying equivalent classes in $L^p(E)$, we will call them 'functions' in $L^p(E)$.

The set of measurable functions f which belongs to $L^p(K)$ for any compact set $K \subset E$ is denoted by $L^p_{loc}(E)$.

Theorem 1.1.1. *Suppose that $1 \le p \le \infty$. For all $f, g \in L^p(E)$ and $\lambda \in \mathbb{R}$, we have that $f + \lambda g \in L^p(E)$. In order words, $L^p(E)$ is a vector space.*

Proof. Exercise. □

1.2 Hölder's inequality and Minkowski's inequality

Theorem 1.2.1 (Hölder's inequality)**.** *If $1 \le p \le \infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$, then $\|fg\|_{L^1(E)} \le \|f\|_{L^p(E)}\|g\|_{L^{p'}(E)}$.*

In the above, we use the convention that, when f does not belong to $L^p(E)$ or g does not belong to $L^{p'}(E)$, the right hand side of Hölder's inequality is assumed to take the value ∞. Also, in the special case that $p = q = 2$, we have Cauchy-Schwarz' inequality: $\|fg\|_{L^1(E)} \le \|f\|_{L^2(E)}\|g\|_{L^2(E)}$.

Proof. When $p = 1$ or $p = \infty$, the inequality is obvious. If $\|f\|_{L^p} = 0$ or $\|g\|_{L^{p'}} = 0$, then $fg = 0$ a.e. and so the conclusion is also obvious. We assume henceforth that $1 < p < \infty$, $\|f\|_{L^p} > 0$ and $\|g\|_{L^{p'}} > 0$.

Consider first the case in which $\|f\|_{L^p} = 1$ and $\|g\|_{L^{p'}} = 1$. Using Young's inequality $|fg| \le \frac{1}{p}|f|^p + \frac{1}{p'}|g|^{p'}$, we have

$$\int_E |fg| \le \int_E \frac{1}{p}|f|^p + \frac{1}{p'}|g|^{p'} = \frac{1}{p}\|f\|^p_{L^p} + \frac{1}{p'}\|g\|^{p'}_{L^{p'}} = \frac{1}{p} + \frac{1}{p'} = 1 = \|f\|_{L^p}\|g\|_{L^{p'}}.$$

In the general case, let $\tilde{f} = \frac{1}{\|f\|_{L^p}} f$ and $\tilde{g} = \frac{1}{\|g\|_{L^{p'}}} g$ so that $\|\tilde{f}\|_{L^p} = 1$ and $\|\tilde{g}\|_{L^{p'}} = 1$. By the above, we have $\|\tilde{f}\tilde{g}\|_{L^1} \le 1$, which gives precisely $\|fg\|_{L^1} \le \|f\|_{L^p}\|g\|_{L^{p'}}$. □

Theorem 1.2.2 (Minkowski's inequality)**.** *If $1 \le p \le \infty$, then $\|f + g\|_{L^p(E)} \le \|f\|_{L^p(E)} + \|g\|_{L^p(E)}$.*

Again, in this inequality, if f or g does not belong to $L^p(E)$, the right hand side is assumed to take the value ∞.

Proof. If $p = 1$ or $p = \infty$, the conclusion is obvious. Suppose that $1 < p < \infty$. Using Hölder's inequality we have

$$\int_E |f||f+g|^{p-1} \le \|f\|_{L^p} \||f+g|^{p-1}\|_{L^{\frac{p}{p-1}}} = \|f\|_{L^p} \|f+g\|_{L^p}^{p-1}.$$

Likewise,

$$\int_E |g||f+g|^{p-1} \le \|g\|_{L^p} \|f+g\|_{L^p}^{p-1}.$$

Summing up the two estimate we then have

$$\|f+g\|_{L^p}^p = \int_E |f||f+g|^{p-1} + \int_E |g||f+g|^{p-1} \le (\|f\|_{L^p} + \|g\|_{L^p}) \|f+g\|_{L^p}^{p-1}.$$

If $\|f+g\|_{L^p} = 0$, there is nothing to prove. Otherwise, we can divide both side by $\|f+g\|_{L^p}^{p-1}$ to get the conclusion. □

1.3 Banach space properties

Recall that a set X is called a Banach space (over $\mathbb{R}$) if it satisfies the following properties

(1) (Linearity) X is a vector space.

(2) (Norm) X is a normed space, i.e. there is a map $x \mapsto \|x\|$ from X into $[0, \infty)$ such that

   (i) $\|x\| = 0$ if and only if $x = 0$.

   (ii) $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{R}, x \in X$.

   (iii) $\|x + y\| \le \|x\| + \|y\|$ for all $x, y \in X$.

(3) (Completeness) X is complete with respect to its norm, i.e. every Cauchy sequence in X converges in X.

1.3.1 Completeness

Theorem 1.3.1 (Riesz-Fischer's theorem)**.** *If $1 \le p \le \infty$, then $L^p(E)$ is a Banach space with norm $\| \cdot \|_{L^p(E)}$.*

Proof. Properties (1), (2)(i) and (ii) are clear. Property (2)(iii) is precisely Minkowski's inequality. Let us prove (3), i.e. the completeness of L^p. Suppose that (f_k) is a Cauchy sequence in L^p. We need to show that f_k converges in L^p to some $f \in L^p$.

Consider first the case $p = \infty$. For every k, m, we have that $|f_k - f_m| \leq \|f_k - f_m\|_{L^\infty}$ except for a set of measure zero, which we denote by $Z_{k,m}$. Let Z be the union of all those $Z_{k,m}$'s. Then Z has measure zero and $|f_k - f_m| \leq \|f_k - f_m\|_{L^\infty}$ in $E \setminus Z$ for all k and m. It follows that f_k converges uniformly in $E \setminus Z$ to some measurable function f. Now, for any k, we have

$$|f_k - f| < \sup_{m \geq k} \|f_k - f_m\|_{L^\infty} \text{ in } E \setminus Z.$$

Since f_k is essentially bounded and the right hand side is bounded (in fact can be made arbitrarily small for large k), we have that f is essential bounded, i.e. $f \in L^\infty$. Also, sending $k \to \infty$ in the above inequality also shows that $\|f_k - f\|_{L^\infty} \to 0$, i.e. f_k converges to f in L^∞.

We now consider the case $1 \leq p < \infty$. For any $\varepsilon > 0$ we have that

$$|\{x \in E : |f_k(x) - f_m(x)| > \varepsilon\}| \leq \frac{1}{\varepsilon^p} \int_E |f_k(x) - f_m(x)|^p = \frac{1}{\varepsilon^p} \|f_k(x) - f_m(x)\|_{L^p}^p,$$

and so

$$\lim_{k,m \to \infty} |\{x \in E : |f_k(x) - f_m(x)| > \varepsilon\}| = 0 \text{ for every } \varepsilon > 0.$$

By a result from integration theory, this implies that f_k converges in measure to some measurable function f. Furthermore, there is a subsequence f_{k_j} which converges to f a.e. in E.

We next show that $\|f_k - f\|_{L^p} \to 0$ as $k \to \infty$. Indeed, fix any $\delta > 0$, and select K such that $\|f_{k_j} - f_k\|_{L^p} < \delta$ for $k_j, k > K$. Letting $j \to \infty$ and using Fatou's lemma, we have for every $k > K$ that

$$\|f - f_k\|_{L^p}^p = \int_E |f - f_k|^p \leq \liminf_{j \to \infty} \int_E |f_{k_j} - f_k|^p \leq \delta^p.$$

We hence have $\|f_k - f\|_{L^p} \to 0$ as $k \to \infty$. Now, by Minkowski's inequality, we have $\|f\|_{L^p} \leq \|f_k\|_{L^p} + \|f - f_k\|_{L^p} < \infty$, and so $f \in L^p$. This completes the proof. □

1.3.2 Dual spaces

Proposition 1.3.2 (Converse to Hölder's inequality)**.** *Let f be measurable on E. If $1 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$, then*

$$\|f\|_{L^p(E)} = \sup\Big\{ \int_E fg : g \in L^{p'}(E), \|g\|_{L^{p'}(E)} \leq 1 \text{ and } fg \text{ is integrable}\Big\}.$$

Proof. Call the supremum on the right hand side α. Then $\alpha \leq \|f\|_{L^p}$ by Hölder's inequality. We proceed to prove the opposite inequality. If $\|f\|_{L^p} = 0$, the result is obvious. We henceforth assume that $\|f\|_{L^p} > 0$.

Case 1: $0 < \|f\|_{L^p} < \infty$.

Case 1(a): $1 \leq p < \infty$. Let

$$g_0(x) = \mathrm{sign} f(x)|f(x)|^{p-1}\|f\|_{L^p}^{-(p-1)}.$$

Then $g_0 \in L^{p'}$, $\|g_0\|_{L^{p'}} = 1$ and so $\alpha \geq \int_E fg = \|f\|_{L^p}$, as desired.

Case 1(b): $p = \infty$.

For $\varepsilon > 0$, let $E_\varepsilon = \{x \in E : |x| \leq \varepsilon^{-1} \text{ and } |f(x)| \geq \|f\|_{L^\infty} - \varepsilon\}$, which has positive measure. Let $g_0(x) = \frac{1}{|E_\varepsilon|}\mathrm{sign} f(x)\chi_{E_\varepsilon}(x)$. Then $\|g_0\|_{L^1} = 1$ and $\alpha \geq \int_E fg_0 \geq \|f\|_{L^\infty} - \varepsilon$. Sending $\varepsilon \rightarrow$ we obtain $\alpha \geq \|f\|_{L^\infty}$.

Case 2: $\|f\|_{L^p} = \infty$. Let

$$f_k(x) = \begin{cases} 0 & \text{if } |x| > k, \\ \min(|f(x)|, k) & \text{if } |x| \leq k. \end{cases}$$

Then $f_k \in L^p$ and $\|f_k\|_{L^p} \rightarrow \|f\|_{L^p} = \infty$ by the monotone convergence theorem. By Case 1, we have that $\|f_k\|_{L^p} = \int_E f_k g_k$ for some non-negative g_k with $\|g_k\|_{L^{p'}} = 1$. As $|f| \geq f_k \geq 0$, it follows that

$$\int_E |f|g_k \geq \int_E f_k g_k = \|f_k\|_{L^p} \rightarrow \infty.$$

Let $\tilde{g}_k(x) = \mathrm{sign} f(x) g_k(x)$, we thus have

$$\alpha \geq \int_E f\tilde{g}_k = \int_E f g_k \rightarrow \infty,$$

and so $\alpha = \infty$ as desired. □

Recall that if X is a (real) normed vector space with norm $\|\cdot\|$, then its dual space X^* is defined as the space of all bounded linear functional $T : X \rightarrow \mathbb{R}$, which is a normed space with norm

$$\|T\|_* = \sup\{|Tx| : x \in X, \|x\| = 1\}.$$

We have the following characterisation of dual spaces of Lebesgue space, which we will not prove.

Theorem 1.3.3 (Riesz' representation theorem). *Let $1 \leq p < \infty$ and $p' = \frac{p}{p-1}$. Then there is an isometric isomorphism $\pi : (L^p(E))^* \to L^{p'}(E)$ so that*

$$Tg = \int_E \pi(T)g \text{ for all } g \in L^p(E).$$

A consequence of the above result is that $L^p(E)$ is reflexive for $1 < p < \infty$.

Remark 1.3.4. *The dual space of $L^\infty(E)$ is **NOT** $L^1(E)$.*

1.3.3 Separability

Theorem 1.3.5. *For $1 \leq p < \infty$, the space $L^p(E)$ is separable, i.e. it has a countable dense subset.*

This theorem will be proven later when we consider dense subsets of L^p spaces.

1.3.4 Weak/Weak* convergence

Definition 1.3.6. *Let X be a normed vector space and X^* its dual.*

(i) We say that a sequence (x_n) in X converges weakly to some $x \in X$ if $Tx_n \to Tx$ for all $T \in X^*$. We write $x_n \rightharpoonup x$.

(ii) We say that a sequence (T_n) in X' converges weakly to some $T \in X^*$ if $T_n x \to Tx$ for all $x \in X$. We write $T_n \rightharpoonup^* T$.*

We have the following important theorems on weak and weak* convergence.

Theorem 1.3.7 (Weak sequential compactness in reflexive Banach spaces). *Every bounded sequence in a reflexive Banach space has a weakly convergent subsequence.*

Theorem 1.3.8 (Helly's theorem on weak* sequential compactness in duals of separable Banach spaces). *Every bounded sequence in the dual of a separable Banach space has a weakly* convergent subsequence.*

Applying the above result to Lebesgue spaces (noting that $L^p(E)$ is reflexive for $1 < p < \infty$ by Riesz' representation theorem (Theorem 1.3.3) and is separable for $1 \leq p < \infty$ by Theorem 1.3.5), we obtain:

Theorem 1.3.9. *Assume that $1 < p \leq \infty$ and (f_k) is a bounded sequence in $L^p(E)$. Then there exists a subsequence f_{k_j} and a function $f \in L^p$ such that*

$$\int_E f_{k_j} g \to \int_E fg \text{ for all } g \in L^{p'}(E).$$

(In other words, $f_{k_j} \rightharpoonup f$ in L^p if $1 < p < \infty$ or $f_{k_j} \rightharpoonup^* f$ in L^p if $p = \infty$.)

We sum up in the following table:

	Reflexivity	Separability	Dual Space	Sequential compactness of the closed unit ball
L^p $1 < p < \infty$	Yes	Yes	$L^{p'}$	Weak and weak*
L^1	No	Yes	L^∞	Neither
L^∞	No	No	Strictly larger than L^1	Weak*

1.4 Hilbert space properties

Recall that a set H is called a Hilbert space (over $\mathbb{R}$) if it satisfies the following properties

(1) (Linearity) H is a vector space.

(2) (Inner product) H is an inner product space, i.e. there is a map $(x, y) \mapsto \langle x, y\rangle$ from $X \times X$ into $\mathbb{R}$ such that

   (i) $\langle x_1 + \lambda x_2, y\rangle = \langle x_1, y\rangle + \lambda\langle x_2, y\rangle$ for all $\lambda \in \mathbb{R}, x_1, x_2, y \in H$,

   (ii) $\langle x, y\rangle = \langle y, x\rangle$ for all $x, y \in H$,

   (iii) $\langle x, x\rangle \geq 0$ for all $x \in H$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

(3) (Completeness) H is complete with respect to its associated norm $\|x\| = \sqrt{\langle x, x\rangle}$.

Theorem 1.4.1. *The space $L^2(E)$ is a Hilbert space with inner product*

$$\langle f, g\rangle = \int_E fg.$$

Proof. Exercise. □

1.5 Density

In this subsection, we consider subsets of L^p which are dense in L^p.

1.5.1 Approximation by simple functions

Theorem 1.5.1. *Let $1 \le p < \infty$. The set of all p-integrable simple functions is dense in $L^p(E)$.*

Recall that a measurable function is called simple if it assumes only a finite number of values, all of which are finite.

Proof. Fix some $f \in L^p$. We need to show that there is a sequence of p-integrable simple function (f_k) such that $f_k \to f$ in L^p. By splitting $f = f^+ - f^-$, it suffices to consider the case that f is non-negative. In this case, a result from integration theory asserts that there is a *non-decreasing* sequence (f_k) of simple functions such that $f_k \to f$ a.e. Now we have $|f_k - f|^p \le |f|^p$ for all k and $|f_k - f|^p \to 0$ a.e. The dominated convergence theorem then implies that $\int_E |f_k - f|^p \to 0$, i.e. $f_k \to f$ in L^p. $\square$

In the next result, we consider a class of dyadic cubes whose construction is as follows: One considers a lattice of $\mathbb{R}^n$ of size 1 and the corresponding set K_0 of closed cubes with edge of length 1 and vertices at those lattice point. By bysecting each cube in K_0 one obtains 2^n subcubes of edge length $\frac{1}{2}$. The set of all these subcubes is denoted as K_1. By repeating this process, one obtains finer set of cubes K_m of cubes of edge length 2^{-m}, each of which is a subcube of a cube in K_{m-1} and contains 2^n non-overlapping smaller cubes in K_{m+1}. The union of all these K_m's is called a class of dyadic cubes.

Theorem 1.5.2. *Let $1 \le p < \infty$. The set of all finite rational linear combinations of characteristic functions of cubes belonging to a fixed class of dyadic cubes is dense in $L^p(\mathbb{R}^n)$.*

Proof. Fix a class of dyadic cubes of $\mathbb{R}^n$ and let $\mathscr{F}$ denote the set of all finite linear combinations of a characteristic functions of those cubes. In view of Theorem 1.5.1 and of $\mathbb{Q}$ in $\mathbb{R}$, to show that $\mathscr{F}$ is dense in L^p, it suffices to show that characteristic functions of a measurable set of finite measure belongs to the closure of $\mathscr{F}$.

Indeed, since any open set can be written as a countable union of non-overlapping dyadic cubes, characteristic functions of open sets belong to the closure of $\mathscr{F}$. Now if E is a measurable set of finite measure, then we can select $U_k \supset E$ such that $|U_k \setminus E| < \frac{1}{k}$ so that $\chi_{U_k} \to \chi_E$ in L^p. As $\chi_{U_k} \in \bar{\mathscr{F}}$, it follows that $\chi_E \in \bar{\mathscr{F}}$, as wanted. $\square$

We have a couple of applications.

Proof of Theorem 1.3.5. Let $\mathscr{F}$ be the dense subset of $L^p(\mathbb{R}^n)$ in Theorem 1.5.2. Note that $\mathscr{F}$ is countable, and so this proves the theorem when $E = \mathbb{R}^n$.

For general E, let $\tilde{\mathscr{F}}$ be the set of restrictions to E of functions in $\mathscr{F}$. Then $\tilde{\mathscr{F}}$ is countable and dense in $L^p(E)$. □

Theorem 1.5.3 (Continuity in L^p). *If $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p < \infty$, then*

$$\lim_{|y|\to 0} \|f(\cdot + y) - f(\cdot)\|_{L^p(\mathbb{R}^n)} = 0.$$

It should be clear that the above statement is false for $p = \infty$.

Proof. We will only give a sketch. Details are left as an exercise. Let $\mathscr{A}$ denote the set of functions f in L^p such that $\|f(\cdot + y) - f(\cdot)\|_{L^p} \to 0$ as $|y| \to 0$. Using Minkowski's inequality, it can be shown that

(i) $\mathscr{A}$ is a vector subspace of L^p, i.e. finite linear combinations of members of $\mathscr{A}$ belongs to $\mathscr{A}$.

(ii) $\mathscr{A}$ is closed in L^p, i.e. if (f_k) is a sequence in $\mathscr{A}$ and $f_k \to f$ in L^p, then $f \in \mathscr{A}$.

Now, by direct computation, $\chi_E \in \mathscr{A}$ if E is a cube. Hence, by (i), $\mathscr{A}$ contains the set in Theorem 1.5.2, which is dense in L^p. By (ii), $\mathscr{A} = L^p$. □

1.5.2 Convolution

Let f and g be measurable functions on $\mathbb{R}^n$. The convolution $f * g$ of f and g is defined by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)\, dy$$

wherever the integral converges.

Theorem 1.5.4 (Young's convolution theorem). *Let p, q and r satisfy $1 \leq p, q, q \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ and $\|f * g\|_{L^r(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)}$.*

Proof. We will only deal with the case $q = 1$ and $r = p$. The general case is left as an exercise.

Note that $|f * g| \leq |f| * |g|$, we may assume without loss of generality that f and g are non-negative.

<u>Case 1:</u> $p = 1$.

First note that as g is measurable, the function $G(x, y) = g(x - y)$ defines a measurable function on $\mathbb{R}^n \times \mathbb{R}^n$. Thus, as f and g are both non-negative, the integral

$$I = \int_{\mathbb{R}^n \times \mathbb{R}^n} f(y)g(x - y)\, dy\, dx$$

is well-defined. Furthermore, by Tonelli's theorem,

$$\begin{aligned}\|f * g\|_{L^1} &= \int_{\mathbb{R}^n} \Big\{ \int_{\mathbb{R}^n} f(y)g(x-y)\,dy \Big\} dx \\ &= I = \int_{\mathbb{R}^n} \Big\{ \int_{\mathbb{R}^n} g(x-y)\,dx \Big\} f(y)dy \\ &= \Big\{ \int_{\mathbb{R}^n} g(x-y)\,dx \Big\} \Big\{ \int_{\mathbb{R}^n} f(y)dy \Big\} = \|f\|_{L^1}\|g\|_{L^1},\end{aligned}$$

which proves the theorem.

Case 2: $p = \infty$. We have

$$|f * g(x)| = \int_{\mathbb{R}^n} f(y)g(x-y)\,dy \leq \|f\|_{L^\infty} \int_{\mathbb{R}^n} g(x-y)dy = \|f\|_{L^\infty}\|g\|_{L^1}.$$

This also proves the theorem.

Case 3: $1 < p < \infty$. In this case we write

$$|f * g(x)| = \int_{\mathbb{R}^n} [f(y)g(x-y)^{\frac{1}{p}}][g(x-y)^{\frac{1}{p'}}]\,dy.$$

Applying Hölder's inequality, we obtain

$$|f * g(x)| \leq \Big\{ \int_{\mathbb{R}^n} f(y)^p g(x-y)\,dy \Big\}^{\frac{1}{p}} \Big\{ \int_{\mathbb{R}^n} g(x-y)\,dy \Big\}^{\frac{1}{p'}} = |f^p * g(x)|^{\frac{1}{p}} \|g\|_{L^1}^{\frac{1}{p'}}$$

and so

$$\|f * g\|_{L^p} \leq \|f^p * g(x)\|_{L^1}^{\frac{1}{p}} \|g\|_{L^1}^{\frac{1}{p'}}.$$

Now recall that $f^p \in L^1$ and so we have from Case 1 that

$$\|f^p * g\|_{L^1} \leq \|f\|_{L^p}^p \|g\|_{L^1}.$$

The conclusion is readily seen from the above two inequalities. □

For $k = 0, 1, \ldots$, let $C^k(\mathbb{R}^n)$ denotes the space of functions on $\mathbb{R}^n$ whose partial derivatives up to and including those of order k exist and are continuous. Let $C_c^k(\mathbb{R}^n)$ denote the set of functions in $C^k(\mathbb{R}^n)$ which have compact supports.

Lemma 1.5.5. *If $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$ and $g \in C_c^k(\mathbb{R}^n)$ for some $k \geq 0$, then $f * g \in C^k(\mathbb{R}^n)$ and*

$$\partial^\alpha (f * g)(x) = (f * \partial^\alpha g)(x)$$

for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $|\alpha| := \alpha_1 + \ldots + \alpha_n \leq k$.

Proof. Let us first consider the case $k = 0$: Suppose that g is continuous and compactly supported. We will show that $f * g$ is continuous. Indeed, for $z \in \mathbb{R}^n$, we have

$$\begin{aligned} |f * g(x+z) - f * g(x)| &= \Big| \int_{\mathbb{R}^n} f(y)g(x+z-y)\,dy - \int_{\mathbb{R}^n} f(y)g(x-y)\,dy \Big| \\ &= \Big| \int_{\mathbb{R}^n} f(x-u)g(u+z)\,du - \int_{\mathbb{R}^n} f(x-u)g(u)\,du \Big| \\ &\le \int_{\mathbb{R}^n} |f(x-u)||g(u+z) - g(u)|\,du. \end{aligned}$$

Using Hölder's inequality, this gives

$$|f * g(x+z) - f * g(x)| \le \|f(x-\cdot)\|_{L^p} \|g(\cdot + z) - g(\cdot)\|_{L^{p'}} = \|f\|_{L^p} \|g(\cdot + z) - g(\cdot)\|_{L^{p'}}.$$

Now as g is continuous and compactly supported, g is uniformly continuous. Hence for every given $\varepsilon > 0$, we can select $\delta > 0$ such that $\|f\|_{L^p} \|g(\cdot + z) - g(\cdot)\|_{L^{p'}} \le \varepsilon$. The continuity of $f * g$ follows.

Now consider the case $k = 1$. We showed above that $f * g$ is continuous. Consider the partial derivative $\partial_{x_1} f * g$ at some fixed point x. We have

$$\frac{1}{t}[(f * g)(x + te_1) - f * g(x)] = \int_{\mathbb{R}^n} f(y) \frac{g(x - y + te_1) - g(x-y)}{t}\,dy$$

As g has compact support and x is a fixed point, the integrand on the right hand side of the above identity vanishes outside of a compact set, say K. Since $f \in L^p(\mathbb{R}^n)$ and K has bounded measure, $f \in L^1(K)$. Since g is differentiable, we have, as $t \to 0$, that $\frac{g(x-y+te_1) - g(x-y)}{t}$, as a function of y, converges uniformly to $\partial_{x_1} g(x-y)$ and is bounded by some large constant M in K. An application of the dominated convergence theorem thus gives

$$\lim_{t \to 0} \frac{1}{t}[(f * g)(x + te_1) - f * g(x)] = \int_{\mathbb{R}^n} f(y) \partial_{x_1} g(x-y) = (f * \partial_{x_1} g)(x).$$

Therefore $\partial_{x_1}(f * g)$ exists and is equal to $f * \partial_{x_1} g$, which is continuous by the case $k = 0$. Clearly the same conclusion hold for other partial derivatives, which conclude the case $k = 1$.

Finally, applying the case $k = 1$ repeatedly, we obtain the conclusion for $k \ge 2$. □

1.5.3 Approximation of identity

A family of kernels $\{\varrho_\varepsilon : \varepsilon > 0\}$ with the property that $f * \varrho_\varepsilon \to f$ as $\varepsilon \to 0$ in some suitable sense is called an approximation of identity.

Theorem 1.5.6 (Approximation of identity)**.** *Let ϱ be a non-negative function in $C_c^\infty(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} \varrho = 1$. For $\varepsilon > 0$, let*

$$\varrho_\varepsilon(x) = \frac{1}{\varepsilon^n} \varrho\Big(\frac{x}{\varepsilon}\Big).$$

If $f \in C(\mathbb{R}^n)$, then $f * \varrho_\varepsilon$ converges uniformly on compact subsets of $\mathbb{R}^n$ to f.

Note that we have $\int_{\mathbb{R}^n} \varrho_\varepsilon = 1$ for every ε. A family (ϱ_ε) as in the statement is call a family of mollifiers, and the family $(f * \varrho_\varepsilon)$ is called a regularization of f by mollification.

Proof. Exercise. □

Theorem 1.5.7 (Approximation of identity)**.** *Let ϱ be a non-negative function in $L^1(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} \varrho = 1$. For $\varepsilon > 0$, let*

$$\varrho_\varepsilon(x) = \frac{1}{\varepsilon^n} \varrho\Big(\frac{x}{\varepsilon}\Big).$$

If $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p < \infty$, then

$$\lim_{\varepsilon \to 0} \| f * \varrho_\varepsilon - f \|_{L^p(\mathbb{R}^n)} = 0.$$

Proof. Let $f_\varepsilon = f * \varrho_\varepsilon$. As $\int_{\mathbb{R}^n} \varrho_\varepsilon = 1$, we have

$$f(x) = f(x) \int_{\mathbb{R}^n} \varrho_\varepsilon(y)\, dy.$$

Hence

$$\begin{aligned} |f_\varepsilon(x) - f(x)| &\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)||\varrho_\varepsilon(y)| dy \\ &= \int_{\mathbb{R}^n} |f(x-y) - f(x)||\varrho_\varepsilon(y)|^{\frac{1}{p}} |\varrho_\varepsilon(y)|^{\frac{1}{p'}}\, dy. \end{aligned}$$

Applying Hölder's inequality, this gives

$$\begin{aligned} |f_\varepsilon(x) - f(x)| &\leq \Big\{ \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p |\varrho_\varepsilon(y)|\, dy \Big\}^{\frac{1}{p}} \Big\{ \int_{\mathbb{R}^n} |\varrho_\varepsilon(y)|\, dy \Big\}^{\frac{1}{p'}} \\ &= \Big\{ \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p |\varrho_\varepsilon(y)|\, dy \Big\}^{\frac{1}{p}}. \end{aligned}$$

It follows that

$$\|f_\varepsilon - f\|_{L^p}^p \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p |\varrho_\varepsilon(y)|\,dy\,dx.$$

In particular, if we let $\delta(y) := \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p\,dx$, then, in view of Tonelli's theorem,

$$\|f_\varepsilon - f\|_{L^p}^p \le \int_{\mathbb{R}^n} \delta(y)\varrho_\varepsilon(y)\,dy. \tag{1.2}$$

Now, for given $\eta > 0$, using the continuity property in L^p (Theorem 1.5.3), we can find $r > 0$ such that $\delta(y) < \eta/2$ for $|y| \le r$. Note also that δ is bounded: $\delta(y) \le 2^p\|f\|_{L^p}^p$ for all y. Hence

$$\begin{aligned}\|f_\varepsilon - f\|_{L^p}^p &\le \frac{\eta}{2}\int_{\{|y|\le r\}} \varrho_\varepsilon(y)\,dy + \|\delta\|_{L^\infty}\int_{|y|>r} \varrho_\varepsilon(y)\,dy \\ &\le \frac{\eta}{2} + \|\delta\|_{L^\infty}\int_{|y|>r/\varepsilon} \varrho(y)\,dy.\end{aligned}$$

As $\varrho \in L^1$, the last integral goes to zero as $\varepsilon \to 0$. Hence there is some $\bar\varepsilon$ (depending on η) such that $\|f_\varepsilon - f\|_{L^p}^p < \eta$ for any $\varepsilon < \bar\varepsilon$. The conclusion follows. □

1.5.4 Approximation by smooth functions

Theorem 1.5.8. *For $1 \le p < \infty$, the space $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.*

Proof. Pick an arbitrary non-negative kernel $\varrho \in C_c^\infty(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varrho = 1$. For $\varepsilon > 0$, let

$$\varrho_\varepsilon(x) = \frac{1}{\varepsilon^n}\varrho\Big(\frac{x}{\varepsilon}\Big).$$

Let $f \in L^p$, $1 \le p < \infty$. Fix some $\eta > 0$. We would like to find some $f_\eta \in C_c^\infty(\mathbb{R}^n)$ such that $\|f_\eta - f\|_{L^p} < \eta$.

First, select g and h in L^p such that $f = g + h$, g has compact support and $\|h\|_{L^p} < \eta/2$, e.g. by letting $g = f\chi_{\{|x|<R\}}$ for some suitably large R.

Let $g_\varepsilon = g * \varrho_\varepsilon$. As g and ϱ_ε have compact supports, so does g_ε. By Lemma 1.5.5, $g_\varepsilon \in C^\infty(\mathbb{R}^n)$, hence $g_\varepsilon \in C_c^\infty(\mathbb{R}^n)$. By Theorem 1.5.7, $g_\varepsilon \to g$ in L^p. So we can select some small ε such that $\|g_\varepsilon - g\|_{L^p} < \eta/2$. By Minkowski's inequality, this gives

$$\|g_\varepsilon - f\|_{L^p} \le \|g_\varepsilon - g\|_{L^p} + \|g - f\|_{L^p} < \eta.$$

We can now conclude with the choice $f_\eta = g_\varepsilon$. □

Theorem 1.5.9. *For $1 \le p < \infty$, the space $C^\infty(E)$ is dense in $L^p(E)$.*

Here $C^\infty(E)$ is the space of restrictions to E of functions which are smooth on some open sets containing $\bar{E}$.

Proof. For every $f \in L^p(E)$, we define $\bar{f} : \mathbb{R}^n \setminus \mathbb{R}$ by $\bar{f}(x) = f(x)$ for $x \in E$ and $\bar{f}(x) = 0$ if $x \notin E$. Then $\bar{f} \in L^p(\mathbb{R}^n)$. By Theorem 1.5.8, there is a sequence $(\bar{f}_k) \subset C_c^\infty(\mathbb{R}^n)$ which converges to $\bar{f}$ in $L^p(\mathbb{R}^n)$. A desired sequence of approximations for f is given by $(\bar{f}_k|_E)$. □

1.6 A criterion for strong pre-compactness

A set A in a normed vector space X is called pre-compact if every sequence in A has a sub-sequence which converges in X.

Recall the following theorem concerning pre-compactness in the space of continuous functions.

Theorem 1.6.1 (Ascoli-Arzelà's theorem)**.** *Let K be a compact subset of $\mathbb{R}^n$. Suppose that a subset $\mathscr{F}$ of $C(K)$ satisfies*

(1) (Boundedness) $\sup_{f\in\mathscr{F}} \|f\|_{C(K)} < \infty$,

(2) (Equi-continuity) For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $f \in \mathscr{F}$ and $x, y \in K$ with $|x - y| < \delta$.

Then $\mathscr{F}$ is pre-compact in $C(K)$.

The following is an analogue in L^p spaces.

Theorem 1.6.2 (Kolmogorov-Riesz-Fréchet's theorem)**.** *Let $1 \le p < \infty$ and Ω be an open subset of $\mathbb{R}^n$. Suppose that a subset $\mathscr{F}$ of $L^p(\Omega)$ satisfies*

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

(2) For every $\varepsilon > 0$, there exists $\delta > 0$ such that $\|\tilde{f}(\cdot + y) - \tilde{f}(\cdot)\|_{L^p(\Omega)} < \varepsilon$ for all $f \in \mathscr{F}$ and $|y| < \delta$, where $\tilde{f}$ is the extension by zero of f to the whole of $\mathbb{R}^n$.

Then, for every bounded open subset ω of Ω such that $\bar{\omega} \subset \Omega$, the set $\mathscr{F}|_\omega$ of restrictions to ω of functions in $\mathscr{F}$ is pre-compact in $L^p(\omega)$.

Proof. Without loss of generality we may assume that Ω is bounded. We need to show that every sequence of $\mathscr{F}|_\omega$ admits a convergent subsequence.

For $f \in \mathscr{F}$, let $\tilde{f}$ be the extension (by zero) of f to the whole of $\mathbb{R}^n$ by letting $\tilde{f}(x) = 0$ for $x \in \mathbb{R}^n \setminus \Omega$. Let $\widetilde{\mathscr{F}} = \{\tilde{f} : f \in \mathscr{F}\}$.

Note that the set $\widetilde{\mathscr{F}}$ is bounded in both $L^p(\mathbb{R}^n)$ and $L^1(\mathbb{R}^n)$.

Let (ϱ_η) be a family of mollifiers such that the support of ϱ_η is contained in $B_\eta(0)$.

For $\tilde{f} \in \widetilde{\mathscr{F}}$, Let $\tilde{f}_j = \varrho_{\frac{1}{j}} * \tilde{f}$. We will use the following two properties of the approximants $\tilde{f}_j$:

(P1) Recall from the proof of Theorem 1.5.7 (cf. (1.2)) the estimate

$$\|\tilde{f}_j - \tilde{f}\|^p_{L^p(\mathbb{R}^n)} \le \int_{\mathbb{R}^n} \|\tilde{f}(\cdot + y) - \tilde{f}(\cdot)\|^p_{L^p(\mathbb{R}^n)} \varrho_{\frac{1}{j}}(y)\, dy.$$

Keeping in mind that $\int_{\mathbb{R}^n} \varrho_{\frac{1}{j}} = 1$, we thus have, with the notation in property (2), that

$$\|\tilde{f}_j - \tilde{f}\|^p_{L^p(\omega)} \le \varepsilon \text{ for all } \tilde{f} \in \widetilde{\mathscr{F}} \text{ and } j > \frac{1}{\delta}.$$

(P2) Next we show that, for each fixed j, the set $\mathscr{F}_j = \{\tilde{f}_j|_\omega : \tilde{f} \in \widetilde{\mathscr{F}}\}$ satisfies the condition of Ascoli-Arzelà's theorem. Indeed, by Young's convolution inequality,

$$\|\tilde{f}_j\|_{L^\infty(\mathbb{R}^n)} \le \|\tilde{f}\|_{L^1(\mathbb{R}^n)} \|\varrho_{\frac{1}{j}}\|_{L^\infty(\mathbb{R}^n)} \le C(j).$$

Also,

$$\begin{aligned} |\tilde{f}_j(x) - \tilde{f}_j(y)| &\le \int_{\mathbb{R}^n} \Big|\varrho_{\frac{1}{j}}(x-z) - \varrho_{\frac{1}{j}}(y-z)\Big| |\tilde{f}(z)|\, dz \\ &\le \|\varrho_{\frac{1}{j}}\|_{Lip(\mathbb{R}^n)} |x-y| \|\tilde{f}\|_{L^1(\mathbb{R}^n)} \le C(j)|x-z|. \end{aligned}$$

Now, if (g_k) is a sequence in $\mathscr{F}$, we construct a convergent subsequence as follows. For $l = 1$, select j_1 so that, by (P1), $\|\varrho_{\frac{1}{j_1}} * \tilde{f} - \tilde{f}\|_{L^p(\omega)} < \frac{1}{1}$ for all $f \in \mathscr{F}$. Then by (P2) and Ascoli-Arzelà's theorem, we can select a subsequence $(g_{k_p^{(1)}})$ so that $\varrho_{\frac{1}{j_1}} * g_{k_r^{(1)}}$ is convergent in $C(\omega)$ and hence in $L^p(\omega)$. Proceed inductively, we select for $l \ge 2$, some $j_l > j_{l-1}$ such that $\|\varrho_{\frac{1}{j_l}} * \tilde{f} - \tilde{f}\|_{L^p(\omega)} < \frac{1}{l}$ for all $f \in \mathscr{F}$ and a subsequence $(g_{k_p^{(l)}})$ of $(g_{k_p^{(l-1)}})$ so that $\varrho_{\frac{1}{j_l}} * g_{k_p^{(l)}}$ is convergent in $L^p(\omega)$. Let $(g_{k_l}) = (g_{k_l^{(l)}})$ be the diagonal subsequence, which is a subsequence of all previously constructed subsequences and so satisfies that, for every r, $\|\rho_{\frac{1}{j_r}} * g_{k_l} - g_{k_l}\|_{L^p(\omega)} < \frac{1}{r}$ for all $l > r$ and $\rho_{\frac{1}{j_r}} * g_{k_l}$ is convergent in $L^p(\omega)$. It follows that, for every r, we can select N_r sufficiently large so that $\|g_{k_l} - g_{k_s}\|_{L^p(\omega)} < \frac{3}{r}$ for all $l, s > N_r$. Hence (g_{k_l}) is Cauchy and hence convergent in $L^p(\omega)$. □

Chapter 2

Sobolev Spaces

Throughout this chapter, Ω is a domain in $\mathbb{R}^n$.

2.1 Weak derivatives

Definition 2.1.1. *Let $f \in L^1_{loc}(\Omega)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index. A function $g \in L^1_{loc}(\Omega)$ is said to be a weak α-derivative of f if*

$$\int_\Omega f \partial^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_\Omega g \varphi \, dx \text{ for all } \varphi \in C_c^\infty(\Omega). \tag{2.1}$$

We write $g = \partial^\alpha f$ in the weak sense.

In the above definition, the function φ is called a *test function*.

Remark 2.1.2. *In Definition 2.1.1, one can use $C_c^{|\alpha|}(\Omega)$ in place of $C_c^\infty(\Omega)$ for the space of test functions. This is because if $\varphi \in C_c^{|\alpha|}(\Omega)$, then $\varrho_n * \varphi \to \varphi$ in $C^{|\alpha|}$ for a suitable sequence of mollifiers (ϱ_n). The assertion then follows by applying dominated convergence theorem.*

Example 2.1.3. *If $u \in C^k(\Omega)$, then its classical derivatives $\partial^\alpha u$ are weak derivatives for $|\alpha| \leq k$.*

Example 2.1.4. *Let $I = (-1, 1)$ and $u(x) = |x|$. Its weak first derivative is $u'(x) = sign(x)$.*

Example 2.1.5. *Let $I = (-1, 1)$ and $u(x) = sign(x)$. Then u has no weak derivatives.*

Lemma 2.1.6. *Let $f \in L^1_{loc}(\Omega)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index. The weak α-derivative of f, if exists, is uniquely defined up to a set of measure zero.*

An equivalent statement is as follows.

Lemma 2.1.7 (Fundamental lemma of the Calculus of Variations). *Let $g \in L^1_{loc}(\Omega)$. If $\int_\Omega g\varphi\,dx = 0$ for all $\varphi \in C_c^\infty(\Omega)$, then $g = 0$ a.e. in Ω.*

Proof. We will only give a sketch and leave the details as exercise. By using an exhaustion by compact subsets, it is enough to consider the case $g \in L^1(\Omega)$ and Ω is bounded. By density $\int_\Omega g\varphi = 0$ for all $\varphi \in C_c(\Omega)$. Select a continuous function $h \in C_c(\Omega)$ such that $\|g - h\|_{L^1}$ is as small as one prefers. Using Tietze-Uryhsohn's theorem, take a continuous function $\varphi \in C_c(\Omega)$ which take values 1 on $\{h \geq \delta\}$ and -1 on $\{h \leq -\delta\}$. All this will imply that $\|g\|_{L^1}, \|h\|_{L^1}$, $\int_\Omega h\varphi$ and $\int_\Omega gh (= 0!)$ are about the same, modulo small errors which can be made as small as one wishes, and so the conclusion follows. □

2.2 Definition of Sobolev spaces

Definition 2.2.1. *For $k \geq 0$ and $1 \leq p \leq \infty$, the Sobolev space $W^{k,p}(\Omega)$ is the set of all functions in $L^p(\Omega)$ whose weak partial derivatives up to and including order k exist and belong also to $L^p(\Omega)$. For $p = 2$, we also write $H^k(\Omega)$ for $W^{k,2}(\Omega)$.*

When the context makes clear what Ω is, we write $W^{k,p}$ and H^k in place of $W^{k,p}(\Omega)$ and $H^k(\Omega)$.

We equip $W^{k,p}(\Omega)$ with the norm

$$\|u\|_{W^{k,p}(\Omega)} = \Big[\sum_{|\alpha| \leq k} \|\partial^\alpha u\|^p_{L^p(\Omega)}\Big]^{\frac{1}{p}}.$$

For $p = 2$, we equip $W^{k,2}(\Omega) = H^k(\Omega)$ with the inner product

$$\langle u, v\rangle_{W^{k,2}(\Omega)} = \sum_{|\alpha| \leq k} \langle \partial^\alpha u, \partial^\alpha v\rangle_{L^2(\Omega)}.$$

Theorem 2.2.2. *For $k \geq 0$ and $1 \leq p \leq \infty$, $W^{k,p}(\Omega)$ is a Banach space. When $p = 2$, $W^{k,2}(\Omega)$ is a Hilbert space.*

Proof. We will only show completeness; the proofs of other properties are routine.

Suppose that (u_m) is a Cauchy sequence in $W^{k,p}$. Then, for $|\alpha| \leq k$, $(\partial^\alpha u_m)$ is Cauchy in L^p and hence converges to some $v_\alpha \in L^p$. Set $u = v_{(0,\dots,0)}$.

Recalling that

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

we can send $m \to \infty$ to obtain

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

This shows that u belongs to $W^{k,p}$ with weak derivatives $\partial^\alpha u = v_\alpha$, which further implies that $\|u_m - u\|_{W^{k,p}} \to 0$. Hence (u_m) is convergent. □

We make the following useful observation from the proof:

Remark 2.2.3. *If $(u_m) \subset L^p(\Omega)$ converges strongly in L^p to u and if, for some multi-index α, $(\partial^\alpha u_m) \subset L^p(\Omega)$ converges strongly in L^p to v, then v is the α-weak derivative of u. If $p < \infty$, the conclusion continues to hold if the strong convergence is relaxed to weak convergence.*

Definition 2.2.4. *For $k \geq 0$ and $1 \leq p < \infty$, the Sobolev space $W_0^{k,p}(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$, i.e. $u \in W_0^{k,p}(\Omega)$ if and only if there is a sequence $(u_m) \subset C_c^\infty(\Omega)$ such that $u_m \to u$ in $W^{k,p}(\Omega)$. For $p = 2$, we also write $H_0^k(\Omega)$ for $W_0^{k,2}(\Omega)$.*

We interpret $W^{k,p}(\Omega)$ as the subspace of $W^{k,p}(\Omega)$ such that "$\partial^\alpha u = 0$ on $\partial\Omega$" for $|\alpha| \leq k-1$. The sense in which this property is understood will be made precise later on.

We now list some elementary properties of Sobolev spaces.

Proposition 2.2.5. *Assume that $u, v \in W^{k,p}(\Omega)$ and $|\alpha| \leq k$. Then*

(i) $\partial^\alpha u \in W^{k-|\alpha|,p}(\Omega)$ and $\partial^\beta(\partial^\alpha u) = \partial^{\alpha+\beta} u$ for $|\beta| \leq k - |\alpha|$.

(ii) $\partial^\alpha(\lambda u + v) = \lambda\partial^\alpha u + \partial^\alpha v$ for all $\lambda \in \mathbb{R}$.

(iii) If Ω' is an open subset of Ω, then $u \in W^{k,p}(\Omega')$.

(iv) (Leibnitz' rule) If $\zeta \in C_c^\infty(\Omega)$, then $\zeta u \in W^{k,p}(\Omega)$ and

$$\partial^\alpha(\zeta u) = \sum_{0 \leq \beta_i \leq \alpha_i} \frac{\alpha_1! \ldots \alpha_n!}{\beta_1!(\alpha_1 - \beta_1)! \ldots \beta_n!(\alpha_n - \beta_n)!}\partial^{(\beta_1,\ldots,\beta_n)}\zeta\partial^{(\alpha_1-\beta_1,\ldots,\alpha_n-\beta_n)}u.$$

.

Proof. Exercise. □

Proposition 2.2.6 (Integration by parts)**.** *Let $u \in W^{k,p}(\Omega)$ and $v \in W_0^{k,p'}(\Omega)$ with $k \geq 0$, $1 < p \leq \infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$. Then*

$$\int_\Omega \partial^\alpha u v\,dx = (-1)^{|\alpha|}\int_\Omega u\partial^\alpha v\,dx \text{ for all } |\alpha| \leq k.$$

Proof. As $v \in W_0^{k,p'}$, there exists $v_m \in C_c^\infty(\Omega)$ such that $v_m \to v$ in $W^{k,p'}$. The conclusion follows by sending $m \to \infty$ in the identity $\int_\Omega \partial^\alpha u v_m = (-1)^{|\alpha|}\int_\Omega u\partial^\alpha v_m$. □

2.3 Approximation by smooth functions

2.3.1 Weak derivative and convolution

We fix a non-negative function $\varrho \in C_c^\infty(B_1(0))$ such that $\int_{\mathbb{R}^n} \varrho = 1$ and define for $\varepsilon > 0$ the mollifiers $\varrho_\varepsilon(x) = \frac{1}{\varepsilon^n}\varrho(x/\varepsilon)$ as usual.

Lemma 2.3.1. *Assume $f \in W^{k,p}(\mathbb{R}^n)$ for some $k \geq 0$ and $1 \leq p < \infty$, then $f * \varrho_\varepsilon \in C^\infty(\mathbb{R}^n)$ and*

$$\partial^\alpha(f * \varrho_\varepsilon) = \partial^\alpha f * \varrho_\varepsilon \text{ in } \mathbb{R}^n \text{ for any } |\alpha| \leq k.$$

Proof. From Lemma 1.5.5, we know that $f * \varrho_\varepsilon \in C^\infty$ and $\partial^\alpha(f * \varrho_\varepsilon) = f * \partial^\alpha \varrho_\varepsilon$. Hence

$$\begin{aligned}\partial^\alpha(f * \varrho_\varepsilon)(x) &= \int_{\mathbb{R}^n} \partial_x^\alpha \varrho_\varepsilon(x-y) f(y) dy \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \partial_y^\alpha \varrho_\varepsilon(x-y) f(y) dy.\end{aligned}$$

Using the definition of weak derivatives in the last integral, we obtain

$$\partial^\alpha(f * \varrho_\varepsilon)(x) = (-1)^{|\alpha|}(-1)^{|\alpha|} \int_{\mathbb{R}^n} \varrho_\varepsilon(x-y) \partial_y^\alpha f(y) dy,$$

from which the conclusion follows. □

2.3.2 Approximation by smooth functions

An immediate consequence of Lemma 2.3.1 and Theorem 1.5.7 is the following approximation result.

Theorem 2.3.2 (Approximation by smooth functions)**.** *Assume that $u \in W^{k,p}(\mathbb{R}^n)$ for some $k \geq 0$, $1 \leq p < \infty$ and let $u_\varepsilon := u * \varrho_\varepsilon$. Then $u_\varepsilon \in C^\infty(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$ and u_ε converges to u in $W^{k,p}(\mathbb{R}^n)$ as $\varepsilon \to 0$.*

For Sobolev spaces on domains of $\mathbb{R}^n$, we also have

Theorem 2.3.3 (Meyers-Serrin's theorem on global approximation by smooth functions)**.** *Let $k \geq 0$ and $1 \leq p < \infty$. For every $u \in W^{k,p}(\Omega)$ there exist a sequence $(u_m) \subset C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that u_m converges to u in $W^{k,p}(\Omega)$.*

We would like now to understand if functions in $W^{k,p}(\Omega)$ can be approximated by functions in $C^\infty(\bar{\Omega})$, i.e. if $C^\infty(\bar{\Omega})$ is dense in $W^{k,p}(\Omega)$. For $k = 0$, we knew this is true. It turns out that for $k \geq 1$, this is not always true, for example when Ω is a disk in $\mathbb{R}^2$ with one small line segment removed. We thus need to restrict our attention to some suitable class of domains.

Definition 2.3.4. *Let $\Omega \subset \mathbb{R}^n$ be a domain.*

(i) $\partial\Omega$ is said to be Lipschitz (or C^m) if for every $x_0 \in \partial\Omega$ there exists a radius $r_0 > 0$ such that, after a relabeling of coordinate axes if necessary,

$$\Omega \cap B_{r_0}(x_0) = \{x \in B_{r_0}(x_0) : x_n > \gamma(x_1, \dots, x_{n-1})\}$$

for some Lipschitz (or C^m) function γ.

(ii) Ω is said to satisfy the segment condition if every $x_0 \in \partial\Omega$ has a neighborhood U_{x_0} and a non-zero vector y_{x_0} such that if $z \in \bar{\Omega} \cap U_{x_0}$, then $z + ty_{x_0} \in \Omega$ for all $t \in (0, 1)$.

It can be shown that when $\partial\Omega$ is Lipschitz, Ω satisfies the segment condition.

We state without proof the following result.

Theorem 2.3.5 (Global approximation by functions smooth up to the boundary)**.** *Suppose $k \geq 1$ and $1 \leq p < \infty$. If Ω satisfies the segment condition, then the set of restrictions to Ω of functions in $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\Omega)$.*

Corollary 2.3.6. *For $k \geq 1$ and $1 \leq p < \infty$, the space $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$. In order words $W^{k,p}(\mathbb{R}^n) = W_0^{k,p}(\mathbb{R}^n)$.*

2.4 Extension

Lemma 2.4.1. *Assume that $k \geq 0$ and $1 \leq p < \infty$. If $u \in W_0^{k,p}(\Omega)$, then its extension by zero $\bar{u}$ to $\mathbb{R}^n$ belongs to $W_0^{k,p}(\mathbb{R}^n)$.*

Proof. Exercise. □

We have the following result:

Theorem 2.4.2 (Stein's extension theorem)**.** *Assume that Ω is a bounded Lipschitz domain. Then there exists a linear operator sending functions defined a.e. in Ω to functions defined a.e. in $\mathbb{R}^n$ such that for every $k \geq 0$, $1 \leq p < \infty$ and $u \in W^{k,p}(\Omega)$ it hold that $Eu = u$ a.e. and*

$$\|Eu\|_{W^{k,p}(\mathbb{R}^n)} \leq C_{k,p,\Omega}\|u\|_{W^{k,p}(\Omega)}$$

We call E a total extension for Ω and Eu an extension of u to $\mathbb{R}^n$.

2.5 Traces

Of importance in the study of partial differential equations is the determination of boundary values. If $u \in C(\bar{\Omega})$, then $u|_{\partial\Omega}$ makes sense in the usual way. If u is a Sobolev function, u needs not be continuous and typically is defined only a.e. in Ω. Since $\partial\Omega$ has (n-dimensional Lebesgue) measure zero, this begs for a study of what 'the restriction of u to $\partial\Omega$' might mean. This will be taken care of by the notion of *trace operator*.

Theorem 2.5.1. *Let $k \geq 1$ and $1 \leq p < \infty$, and assume that Ω is bounded and $\partial\Omega$ is $C^{k-1,1}$ regular. Then there exists a linear operator $T : W^{k,p}(\Omega) \to L^p(\partial\Omega)$ such that*

(i) $Tu = u|_{\partial\Omega}$ if $u \in W^{k,p}(\Omega) \cap C(\bar{\Omega})$,

(ii) T is bounded, i.e. $\|Tu\|_{L^p(\partial\Omega)} \leq C_{k,p,\Omega}\|u\|_{W^{k,p}(\Omega)}$.

The operator T is called the trace operator.

Recall that $C^{k-1,1}$ functions are those whose partial derivatives up to and including order $k-1$ exist and are Lipschitz.

"Proof". To avoid technical difficulties, we will only consider a simple setting where $k = 1$, $\partial\Omega$ contains a flat piece $\hat{\Gamma} = \{(x', 0) : |x'| < 2r\} \subset \{x_n = 0\}$, Ω contains $B^+_{2r}(0) := \{(x', x_n) : |x'| < 2r, x_n > 0\}$, and where we will only be concerned with the trace of u on $\Gamma = \{(x', 0) : |x'| < r\} \subset \{x_n = 0\}$. We will show that

$$\|u\|_{L^p(\Gamma)} \leq C_{p,\Omega,\hat{\Gamma}}\|u\|_{W^{1,p}(\Omega)} \text{ for all } u \in C^1(\bar{\Omega}). \tag{2.2}$$

Once this is established, we can define a local trace operator as follows. For $u \in C^\infty(\bar{\Omega})$, we let $T_\Gamma u = u|_\Gamma$. As $C^\infty(\bar{\Omega})$ is dense in $W^{1,p}(\Omega)$ (by Theorem 2.3.5), estimate (2.2) allows us to define $T_\Gamma u$ for all $u \in W^{1,p}(\Omega)$ and T_Γ is a bounded linear operator from $W^{1,p}(\Omega)$ into $L^p(\Gamma)$.

To prove (2.2), fix a smooth function $\zeta \in C^\infty_c(B_{2r}(0))$ such that $\zeta \equiv 1$ in $B_r(0)$. Consider the function ζu. We have

$$\begin{aligned}\int_\Gamma |u|^p\,dx' &\leq \int_{\hat{\Gamma}} \zeta|u|^p\,dx' = -\int_{\hat{\Gamma}} \Big[\int_0^{\sqrt{4r^2-|x'|^2}} \partial_{x_n}(\zeta|u|^p)\,dx_n\Big]\,dx' \\ &= -\int_{B^+_{2r}(0)} \partial_{x_n}(\zeta|u|^p)\,dx \leq C\int_{B^+_{2r}(0)} [|u|^p + |Du||u|^{p-1}]\,dx,\end{aligned}$$

where here and below C is some generic constant which will always be independent of u. Using Young's inequality, $|a||b|^{p-1} \leq \frac{1}{p}|a|^p + \frac{p-1}{p}|b|^p$, we thus have

$$\int_\Gamma |u|^p\,dx' \leq C\int_{B^+_{2r}(0)} [|u|^p + |Du|^p]\,dx, \leq C\|u\|_{W^{1,p}(\Omega)},$$

as desired. □

We have the following characterization of $W_0^{1,p}$ in terms of trace.

Theorem 2.5.2 (Trace-zero functions in $W^{1,p}$)**.** *Let $1 \le p < \infty$, and Ω be a bounded Lipschitz domain. Suppose that $u \in W^{1,p}(\Omega)$. Then $u \in W_0^{1,p}(\Omega)$ if and only if $Tu = 0$.*

One direction is very easy: If $u \in W_0^{1,p}$, then there is some $u_m \in C_c^\infty(\Omega)$ such that $u_m \to u$ in $W^{1,p}$. Clearly $Tu_m = 0$ and so by continuity of T, $Tu = 0$. We omit the difficult proof of the converse.

2.A Distributions and distributional derivatives

Let $\mathscr{D}(\Omega) = C_c^\infty(\Omega)$, called the space of test functions, be endowed with the following notion of convergence (i.e. topology): For $(\varphi_m) \subset \mathscr{D}(\Omega)$ and $\varphi \in \mathscr{D}(\Omega)$, we say that $\varphi_m \to \varphi$ in $\mathscr{D}(\Omega)$ if there exists a compact set K such that all φ_m's and φ are supported in K and $\partial^\alpha \varphi_m \to \partial^\alpha \varphi$ uniformly in K for every multi-indices α.

Clearly $\mathscr{D}(\Omega)$ is a linear vector space. A functional $T : \mathscr{D}(\Omega) \to \mathbb{R}$ is said to be continuous if it is continuous with respect to the above topology, i.e. if $\varphi_m \to \varphi$ in $\mathscr{D}(\Omega)$, then $T\varphi_m \to T\varphi$.

Definition 2.A.1. *A continuous linear functional from $\mathscr{D}(\Omega)$ into $\mathbb{R}$ is called a distribution. The set of all distributions is denoted by $\mathscr{D}'(\Omega)$.*

Example 2.A.2. *Every function $f \in L^1_{loc}(\Omega)$ defines a canonical distribution T_f by*

$$T_f(\varphi) = \int_\Omega f\varphi.$$

If a distribution T equals to T_f for some $f \in L^1_{loc}(\Omega)$, we say that T is a regular distribution. We say that a regular distribution T belongs to $L^p(\Omega)$ (or $L^p_{loc}(\Omega)$) if $T = T_f$ for some $f \in L^p(\Omega)$ (or $f \in L^p_{loc}(\Omega)$).

Lemma 2.A.3. *Suppose that $f, g \in L^1_{loc}(\Omega)$. Then $T_f = T_g$ if and only if $f = g$ a.e.*

Proof. It is clear that if $f = g$ a.e. then $T_f = T_g$. Conversely if $T_f = T_g$ then $\int_\Omega (f-g)\varphi = 0$ for all $\varphi = \mathscr{D}(\Omega) = C_c^\infty(\Omega)$. We knew that this implies $f = g$ a.e. (cf. Lemma 2.1.6). □

Definition 2.A.4. *Let $T \in \mathscr{D}'(\Omega)$ and α be a multi-index. The distributional α-derivative of T is defined by*

$$\partial^\alpha T(\varphi) = (-1)^{|\alpha|} T(\partial^\alpha \varphi).$$

In particular, every distribution has partial derivatives up to any order. Clearly if $g \in L^1_{loc}(\Omega)$ is a weak α-derivative of $f \in L^1_{loc}(\Omega)$, then T_g is the distributional α-derivative of T_f. In this way, the Sobolev space $W^{k,p}(\Omega)$ comprises of functions in $L^p(\Omega)$ whose distributional partial derivatives up to and including order k also belong to $L^p(\Omega)$.

Chapter 3

Embedding Theorems

If X_1 and X_2 are two Banach spaces, we say that X_1 is embedded in X_2 if $X_1 \subset X_2$. We write $X_1 \hookrightarrow X_2$. We say that the embedding is continuous if there exists a constant C such that $\|u\|_{X_2} \le C\|u\|_{X_1}$ for all $u \in X_1$. To keep the discussion simple, whenever we use the term 'an embedding', we mean 'a continuous embedding'. When X_1 is embedded in X_2, we say that the embedding is compact if bounded subsets of X_1 are pre-compact in X_2.

A major account for the usefulness of Sobolev spaces in analysis, in particular the study of differential equations, is their embedding characteristics. We will consider embeddings of $W^{k,p}(\Omega)$ into

(i) Lebesgue spaces $L^q(\Omega)$,

(ii) Hölder spaces $C^\gamma(\Omega)$.

3.1 Gagliardo-Nirenberg-Sobolev's inequality

In this section, we assume $1 \le p < n$. We are interested in establishing an inequality of the form

$$\|u\|_{L^q(\mathbb{R}^n)} \le C\|Du\|_{L^p(\mathbb{R}^n)} \text{ for all } u \in W^{1,p}(\mathbb{R}^n), \tag{3.1}$$

where the constant C may depend on n and p but is independent of u. When this holds, it clearly follows that $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$.

A simple but deep(!) scaling argument shows that if (3.1) holds, then q must equals to $\frac{np}{n-p}$. To see this fix a function $u \in C_c^\infty(\mathbb{R}^n) \subset W^{1,p}(\mathbb{R}^n)$. For $\lambda > 0$, let $u_\lambda(x) = u(\lambda x)$, which is also of compact support. Then (3.1) gives that

$$\|u_\lambda\|_{L^q(\mathbb{R}^n)} \le C\|Du_\lambda\|_{L^p(\mathbb{R}^n)} \text{ for all } \lambda > 0. \tag{3.2}$$

A direct computation gives $\|u_\lambda\|^q_{L^q(\mathbb{R}^n)} = \lambda^{-n}\|u\|^q_{L^q(\mathbb{R}^n)}$ and $\|Du_\lambda\|^p_{L^p(\mathbb{R}^n)} = \lambda^{p-n}\|Du\|^p_{L^p(\mathbb{R}^n)}$. Plugging into (3.2) we obtain

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\lambda^{1-\frac{n}{p}+\frac{n}{q}}\|Du\|_{L^p(\mathbb{R}^n)}. \tag{3.3}$$

Now if $1 - \frac{n}{p} + \frac{n}{q} \neq 0$, the right hand side of (3.3) can be made arbitrarily small by sending λ either to 0 or ∞, which is impossible when $u \not\equiv 0$. We thus have that $1 - \frac{n}{p} + \frac{n}{q} = 0$, i.e. $q = \frac{np}{n-p}$.

Definition 3.1.1. *For $1 \leq p < n$, we call the number $p^* = \frac{np}{n-p}$ the Sobolev conjugate of p.*

Note that $p^* > p$.

Theorem 3.1.2 (Gagliardo-Nirenberg-Sobolev's inequality)**.** *Assume $1 \leq p < n$. Then there exists a constant $C_{n,p}$ such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C_{n,p}\|Du\|_{L^p(\mathbb{R}^n)} \text{ for all } u \in W^{1,p}(\mathbb{R}^n). \tag{3.4}$$

Proof. We will only give the proof for $p = 1$. The more general case can be established by applying the case $p = 1$ to $|u|^\gamma$ for some suitable $\gamma > 1$ and is left as an exercise.

We knew from Corollary 2.3.6 that $W^{1,1}(\mathbb{R}^n) = W_0^{1,1}(\mathbb{R}^n)$ and so $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{1,1}(\mathbb{R}^n)$. It thus suffices to consider $u \in C_c^\infty(\mathbb{R}^n)$.

Since u has compact support, we have for every x that

$$u(x) = \int_{-\infty}^{x_1} \partial_{x_1} u(y_1, x_2, \ldots, x_n)\, dy_1.$$

This implies

$$|u(x)| \leq \int_{-\infty}^{\infty} |Du(y_1, x_2, \ldots, x_n)|\, dy_1.$$

Similar estimates hold for other variables. Multiplying all these estimates and taking $(n-1)$-th root yields

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^n \Big[\int_{-\infty}^{\infty} |Du(x_1, \ldots, y_i, \ldots, x_n)|\, dy_i\Big]^{\frac{1}{n-1}}.$$

Integrating in x_1 yields

$$\begin{aligned}\int_{-\infty}^{\infty} |u(x)|^{\frac{n}{n-1}}\, dx_1 &\leq \int_{-\infty}^{\infty} \prod_{i=1}^n \Big[\int_{-\infty}^{\infty} |Du(x_1, \ldots, y_i, \ldots, x_n)|\, dy_i\Big]^{\frac{1}{n-1}} dx_1\\ &\leq \Big[\int_{-\infty}^{\infty} |Du(y_1, x_2, \ldots, x_n)|\, dy_1\Big]^{\frac{1}{n-1}} \times\\ &\qquad \times \int_{-\infty}^{\infty} \prod_{i=2}^n \Big[\int_{-\infty}^{\infty} |Du(x_1, \ldots, y_i, \ldots, x_n)|\, dy_i\Big]^{\frac{1}{n-1}} dx_1.\end{aligned}$$

Applying Hölder's inequality to the last integral yields

$$\int_{-\infty}^{\infty} |u(x)|^{\frac{n}{n-1}}\,dx_1 \le \Big[\int_{-\infty}^{\infty} |Du(y_1,x_2,\dots,x_n)|\,dy_1\Big]^{\frac{1}{n-1}} \times \\ \times \prod_{i=2}^{n}\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(x_1,\dots,y_i,\dots,x_n)|\,dy_i\,dx_1\Big]^{\frac{1}{n-1}}.$$

Integrating in x_2 yields

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u(x)|^{\frac{n}{n-1}}\,dx_1\,dx_2 \le \Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(x_1,y_2,x_3,\dots,x_n)|\,dy_2\,dx_1\Big]^{\frac{1}{n-1}} \times \\ \times \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty} |Du(y_1,x_2,\dots,x_n)|\,dy_1\Big]^{\frac{1}{n-1}} \times \\ \times \prod_{i=3}^{n}\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(x_1,\dots,y_i,\dots,x_n)|\,dy_i\,dx_1\Big]^{\frac{1}{n-1}}\,dx_2.$$

Applying Hölder's inequality then yields

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u(x)|^{\frac{n}{n-1}}\,dx_1\,dx_2 \le \Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(x_1,y_2,x_3,\dots,x_n)|\,dy_2\,dx_1\Big]^{\frac{1}{n-1}} \times \\ \times \Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(y_1,x_2,\dots,x_n)|\,dy_1\,dx_2\Big]^{\frac{1}{n-1}} \times \\ \times \prod_{i=3}^{n}\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(x_1,\dots,y_i,\dots,x_n)|\,dy_i\,dx_1\,dx_2\Big]^{\frac{1}{n-1}}.$$

Proceeding in this way with other variables, we eventually obtain

$$\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}}\,dx \le \prod_{i=1}^{n}\Big[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} |Du(x_1,\dots,y_i,\dots,x_n)|\,dx_1\dots dy_2\dots dx_n\Big]^{\frac{1}{n-1}} \\ = \int_{\mathbb{R}^n} |Du|\,dx,$$

which proves (3.4) for $p = 1$ (with $C = 1$.) □

Theorem 3.1.3 (Gagliardo-Nirenberg-Sobolev's inequality)**.** *Assume that Ω is a bounded Lipschitz domain and $1 \le p < n$. Then, for every $q \in [1, p^*]$, there exists $C_{n,p,q,\Omega}$ such that*

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

Proof. Let E be the extension operator in Stein's extension theorem (Theorem 2.4.2). By Gagliardo-Nirenberg-Sobolev's inequality, we have that

$$\|u\|_{L^{p*}(\Omega)} \leq \|Eu\|_{L^{p*}(\mathbb{R}^n)} \leq C\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

As Ω has finite measure, we also have that $\|u\|_{L^q(\Omega)} \leq C\|u\|_{L^{p^*}(\Omega)}$ and so the conclusion follows. □

Remark 3.1.4. *When Ω is bounded and $p = n$, we have that $W^{1,n}(\Omega) \hookrightarrow W^{1,s}(\Omega)$ for any $1 \leq s < n$ and so $W^{1,n}(\Omega) \hookrightarrow L^q(\Omega)$ for any $1 \leq q < \infty$. It turns out that $W^{1,n}(\Omega)$ does not embed into $L^\infty(\Omega)$ unless $n = 1$. For example, for $n \geq 2$, the function $u(x) = \ln\ln(1 + \frac{1}{|x|})$ belongs to $W^{1,n}(B_1(0))$ but is clearly unbounded.*

3.2 Friedrichs' inequality

Theorem 3.2.1 (Friedrichs' inequality)**.** *Assume that Ω is a bounded open set and $1 \leq p < \infty$. Then, there exists $C_{p,\Omega}$ such that*

$$\|u\|_{L^p(\Omega)} \leq C_{p,\Omega}\|Du\|_{L^p(\Omega)} \text{ for all } u \in W_0^{1,p}(\Omega).$$

Note that only the derivatives of u appear on the right hand side.

Remark 3.2.2. *By Friedrichs' inequality, for bounded open Ω, on $W_0^{1,p}(\Omega)$, the norm $\|Du\|_{L^p(\Omega)}$ is equivalent to $\|u\|_{W^{1,p}(\Omega)}$.*

Proof. We may assume that Ω is contain in the slab $S := \{(x', x_n) : 0 < x_n < L\}$. Since $C_c^\infty(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, we only need to consider $u \in C_c^\infty(\Omega)$. Extend u to be zero in $\mathbb{R}^n \setminus \Omega$ so that $u \in C_c^\infty(\mathbb{R}^n)$. We have

$$|u(x)| \leq \int_0^{x_n} |\partial_n u(x', t)|\,dt$$

and so, by Hölder's inequality,

$$|u(x)|^p \leq \Big[\int_0^{x_n} |\partial_n u(x', t)|\,dt\Big]^p \leq x_n^{p-1}\int_0^{x_n} |Du(x', t)|^p\,dt.$$

It follows that

$$\|u\|_{L^p(\Omega)}^p = \int_{\mathbb{R}^{n-1}}\int_0^L |u(x', x_n)|^p\,dx_n\,dx' \leq \int_{\mathbb{R}^{n-1}}\int_0^L x_n^{p-1}\int_0^{x_n} |Du(x', t)|^p\,dt\,dx_n\,dx'.$$

Interchanging order of differentiation yields

$$\begin{aligned}\|u\|^p_{L^p(\Omega)} &\le \int_0^L x_n^{p-1} \int_{\mathbb{R}^{n-1}} \int_0^{x_n} |Du(x',t)|^p \, dt \, dx' \, dx_n \\ &\le \int_0^L x_n^{p-1} \|Du\|^p_{L^p(\Omega)} \, dx_n = \frac{1}{p} L^p \|Du\|^p_{L^p(\Omega)},\end{aligned}$$

which concludes the proof with $C_{p,\Omega} = L\, p^{-\frac{1}{p}}$. □

Theorem 3.2.3 (Friedrichs-type inequality)**.** *Assume that Ω is a bounded open set and $1 \le p < n$. Suppose that $1 \le q \le p^*$ if $p < n$, $1 \le p < \infty$ if $p = n$, and $1 \le q \le \infty$ if $p > n$. Then there exists $C_{n,p,q,\Omega}$ such that*

$$\|u\|_{L^q(\Omega)} \le C_{n,p,q,\Omega} \|Du\|_{L^p(\Omega)} \text{ for all } u \in W_0^{1,p}(\Omega).$$

Proof. Since $C_c^\infty(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, we only need to consider $u \in C_c^\infty(\Omega)$. Extend u to be zero in $\mathbb{R}^n \setminus \Omega$ so that $u \in C_c^\infty(\mathbb{R}^n)$.

If $p < n$, then by Gagliardo-Nirenberg-Sobolev's inequality, $\|u\|_{L^{p^*}(\Omega)} \le C\|Du\|_{L^p(\Omega)}$. But, as Ω has finite measure, $\|u\|_{L^q(\Omega)} \le C\|u\|_{L^{p^*}(\Omega)}$ and so the conclusion follows.

If $p > n$, then by Morrey's inequality and Friedrichs' inequality, $\|u\|_{L^\infty(\Omega)} \le C\|u\|_{W^{1,p}(\Omega)} \le C\|Du\|_{L^p(\Omega)}$. The conclusion also follows.

The case $p = n$ is left as an exercise. □

Remark 3.2.4. *In some literature, Friedrichs' and Friedrichs-type inequalities are sometimes referred to as Poincaré's inequality. Other Poincaré-type inequalities will be considered later in Section 3.5.*

3.3 Morrey's inequality

In this section, we assume $p > n$. We will show that if $u \in W^{1,p}(\Omega)$, then it is Hölder continuous.

Definition 3.3.1. *Let $\alpha \in (0,1]$. A function $u : \Omega \to \mathbb{R}$ is said to be α-Hölder continuous if*

$$[u]_{C^{0,\alpha}(\Omega)} := \sup\left\{\frac{|u(x)-u(y)|}{|x-y|^\alpha} : x \ne y \in \Omega\right\} < \infty.$$

The space of all α-Hölder continuous functions on Ω is denoted by $C^{0,\alpha}(\Omega)$ or simply $C^\alpha(\Omega)$. It can be made a Banach space with the norm

$$\|u\|_{C^{0,\alpha}(\Omega)} = \|u\|_{C^0(\Omega)} + [u]_{C^{0,\alpha}(\Omega)}.$$

Theorem 3.3.2 (Morrey's inequality)**.** *Assume that* $n < p \leq \infty$. *Then every* $u \in W^{1,p}(\mathbb{R}^n)$ *has a* $(1-\frac{n}{p})$*-Hölder continuous representative. Furthermore there exists a constant* $C_{n,p}$ *such that*

$$\|u\|_{C^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \leq C_{n,p}\|u\|_{W^{1,p}(\mathbb{R}^n)}. \tag{3.5}$$

Remark 3.3.3. *In view of Ascoli-Arzelà's theorem (Theorem 1.6.1), we thus have for* $p > n$ *that the embedding* $W^{1,p}(\Omega) \hookrightarrow C^{0,\beta}(\Omega)$ *is compact for every* $0 < \beta < 1 - \frac{n}{p}$.

We will use the following lemma.

Lemma 3.3.4. *There holds*

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

for all $u \in C^1(\mathbb{R}^n)$ *and balls* $B_r(x) \subset \mathbb{R}^n$.

Proof. First, for every $\theta \in \partial B_1(0)$ and $s \in (0, r)$, we have

$$|u(x+s\theta) - u(x)| \leq \int_0^s |\frac{d}{dt} u(x+t\theta)|\, ds \leq \int_0^s |Du(x+t\theta)|\, ds.$$

Integrating over θ and using Tonelli's theorem give

$$\begin{aligned}\int_{\partial B_1(0)} |u(x+s\theta) - u(x)|\, d\theta &\leq \int_0^s \int_{\partial B_1(0)} |Du(x+t\theta)|\, d\theta\, dt \\ &= \int_0^s \int_{\partial B_t(x)} |Du(y)| \frac{dS(y)}{t^{n-1}}\, dt = \int_{B_s(x)} \frac{|Du(y)|}{|y-x|^{n-1}}\, dy.\end{aligned}$$

Now multiplying both sides by s^{n-1} and integrating over s, we get

$$\begin{aligned}\int_{B_r(x)} |u(y) - u(x)|\, dy &= \int_0^r \int_{\partial B_1(0)} |u(x+s\theta) - u(x)|\, d\theta s^{n-1} ds \\ &\leq \int_{B_r(x)} \frac{|Du(y)|}{|y-x|^{n-1}}\, dy \int_0^r s^{n-1}\, ds = \frac{1}{n} r^n \int_{B_r(x)} \frac{|Du(y)|}{|y-x|^{n-1}}\, dy,\end{aligned}$$

which gives the lemma. □

Proof of Theorem 3.3.2. Case 1: $p \in (n, \infty)$.

Suppose for the moment that (3.5) has been proved for $u \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$. Now if $u \in W^{1,p}(\mathbb{R}^n)$, then by Theorem 2.3.2, there exists $u_m \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$ such that $u_m \to u$ in $W^{1,p}$. Applying (3.5) to $u_m - u_{m'}$, we see that the sequence

u_m is Cauchy in $C^{0,1-\frac{n}{p}}$ and so converges (uniformly) to some $\bar{u} \in C^{0,1-\frac{n}{p}}$. But as u_m converges a.e. to u (due to the $W^{1,p}$ convergence), we have that $u = \bar{u}$ a.e., i.e. u has a continuous representative. Returning to inequality (3.5) for u_m and sending $m \to \infty$, we see that (3.5) also hold for u.

From the above discussion, it suffices to prove (3.5) for $u \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$, i.e. we need to show

$$|u(x)| \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)} \text{ for a.e.} x \in \mathbb{R}^n, \tag{3.6}$$

and

$$|u(x) - u(y)| \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}|x-y|^{1-\frac{n}{p}} \text{ for a.e.} x, y \in \mathbb{R}^n. \tag{3.7}$$

Applying Lemma 3.3.4 to u on $B_r(x)$ we have

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

Using Hölder's inequality on the right hand side we get

$$\begin{aligned}\int_{B_r(x)} |u(y) - u(x)|\,dy &\leq \frac{r^n}{n}\|Du\|_{L^p(B_1(x))}\Big[\int_{B_r(x)} \frac{1}{|y-x|^{\frac{(n-1)p}{p-1}}}\,dy\Big]^{\frac{p-1}{p}}\\ &= C_n r^n\|Du\|_{L^p(B_r(x))}\Big[\int_0^r s^{n-1-\frac{(n-1)p}{p-1}}\,ds\Big]^{\frac{p-1}{p}}.\end{aligned}$$

As $p > n$, we have that $\frac{(n-1)p}{p-1} < n$ and so the integral in the square bracket converges. We thus have

$$\int_{B_r(x)} |u(y) - u(x)|\,dy \leq C_{n,p}\|Du\|_{L^p(B_r(x))} r^{\frac{n(p-1)}{p}+1}. \tag{3.8}$$

Now, note that

$$|u(x)| \leq \int_{B_1(x)} |u(y) - u(x)|\,dy + \int_{B_1(x)} |u(y)|dy.$$

Thus, by applying (3.8) to estimate the first term and Hölder's inequality to estimate the second term, we obtain

$$|u(x)| \leq C_{n,p}[\|Du\|_{L^p(B_1(x))} + \|u\|_{L^p(B_1(x)}] \leq C_{n,p}\|u\|_{W^{1,p}(\mathbb{R}^n)},$$

which is (3.6).

We turn to (3.7). Pick some arbitrary $x \neq y$ and let $r = |x-y|$. Set $W = B_r(x) \cap B_r(y) \neq \emptyset$. We have

$$\begin{aligned}|u(x)-u(y)| &\leq \frac{1}{|W|}\int_W |u(x)-u(z)|dz + \frac{1}{|W|}\int_W |u(y)-u(z)|dz\\ &\leq \frac{1}{|W|}\int_{B_r(x)} |u(x)-u(z)|dz + \frac{1}{|W|}\int_{B_r(y)} |u(y)-u(z)|dz\end{aligned}$$

Now as $|W| = C_n r^n$, estimate (3.7) is readily seen from the above inequality and (3.8).

<u>Case 2:</u> $p = \infty$.

We will only give a sketch. Details are left as exercise.

Suppose that $u \in W^{1,\infty}(\mathbb{R}^n)$. Then $u \in W^{1,t}_{loc}(\mathbb{R}^n)$ for any $t < \infty$. In particular, using extension theorems and Case 1, we have that u has a continuous representative (see also Theorem 3.3.5 below), which we henceforth assume to coincide with u.

By approximating u by functions in $C^0_{loc}(\mathbb{R}^n) \cap W^{1,t}_{loc}(\mathbb{R}^n)$, we can show that the conclusion of Lemma 3.3.4 holds for u. We hence have

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

We can now follows the proof of Case 1 to obtain (3.8), and hence (3.6) and (3.7). □

For bounded domain we have:

Theorem 3.3.5 (Morrey's inequality). *Suppose that $n < p \leq \infty$ and Ω is a bounded Lipschitz domain. Then every $u \in W^{1,p}(\Omega)$ has a $(1-\frac{n}{p})$-Hölder continuous representative and*

$$\|u\|_{C^{0,1-\frac{n}{p}}(\Omega)} \leq C_{n,p,\Omega}\|u\|_{W^{1,p}(\Omega)}.$$

Proof. The theorem follows from Morrey's inequality for $\mathbb{R}^n$ and by mean of extension. Details are left as exercise (cf. Theorem 3.1.3). □

3.4 Rellich-Kondrachov's compactness theorem

Theorem 3.4.1 (Rellich-Kondrachov's compactness theorem). *Let Ω be a bounded Lipschitz domain and $1 \leq p \leq \infty$. Let $1 \leq q < p^*$ when $p < n$, $1 \leq q < \infty$ when $p = n$, and $1 \leq p \leq \infty$ when $p > n$. Then the embedding $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ is compact, i.e. every bounded sequence in $W^{1,p}(\Omega)$ contains a subsequence which converges in $L^q(\Omega)$.*

Remark 3.4.2. *(i) For $1 \leq p < n$, the embedding $W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$ is in fact not compact.*

(ii) For every $1 \leq p \leq \infty$, the embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is always compact.

We will only consider the case $q = p < \infty$ where no prior knowledge of Gagliardo-Nirenberg-Sobolev's or Morrey's inequalities is needed. We will need the following lemma (compare Theorem 1.5.3).

Lemma 3.4.3. *Let $1 \leq p < \infty$. For every $v \in W^{1,p}(\mathbb{R}^n)$ and $y \in \mathbb{R}^n$, it holds that*

$$\|v(y + \cdot) - v(\cdot)\|_{L^p(\mathbb{R}^n)} \leq |y| \|Dv\|_{L^p(\mathbb{R}^n)}.$$

Let us assume for now the above lemma and proceed with the proof of Rellich-Kondrachov's theorem.

Proof of Theorem 3.4.1 when $1 \leq p = q < \infty$. Suppose that (u_m) is bounded in $W^{1,p}(\Omega)$. We need to construct a subsequence (u_{m_j}) which converges in $L^p(\Omega)$.

Let $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ be an extension operator (which exists due to Stein's extension theorem; Theorem 2.4.2). Fix some large ball B_R such that $\bar{\Omega} \subset B_R$ and select a cut-off function $\zeta \in C_c^\infty(B_R)$ such that $\zeta \equiv 1$ in Ω. It is easy to check that the map $u \mapsto \tilde{E}u := \zeta E u$ is also an extension of $W^{1,p}(\Omega)$ to $W^{1,p}(\mathbb{R}^n)$. Thus replacing E by $\tilde{E}$ if necessary, we may assume that Eu has support in B_R for every u.

Let $v_m = Eu_m$. To conclude, we show that (v_m) is pre-compact in $L^p(B_R)$ by using Kolmogorov-Riesz-Fischer's theorem (Theorem 1.6.2). It is clear that (v_m) is bounded in $L^p(B_{2R})$. Also, by Lemma 3.4.3, we have

$$\|v_m(y + \cdot) - v_m(\cdot)\|_{L^p(\mathbb{R}^n)} \leq |y| \|Dv_m\|_{L^p(\mathbb{R}^n)}.$$

As (Dv_m) is bounded in $L^p(\mathbb{R}^n)$, we can find for every $\varepsilon > 0$ some $\delta > 0$ so that

$$\sup_m \|v_m(y + \cdot) - v_m(\cdot)\|_{L^p(\mathbb{R}^n)} \leq \varepsilon \text{ whenever } |y| < \delta. \tag{3.9}$$

Applying Kolmogorov-Riesz-Fischer's theorem, we obtain the conclusion. □

Proof of Lemma 3.4.3. By density (Theorem 2.3.2), it suffices to show the stated inequality for $v \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$. We have

$$|v(y + x) - v(x)| \leq \int_0^1 |\frac{d}{dt} v(ty + x)|\, dt \leq |y| \int_0^1 |Dv(ty + x)|\, dt.$$

Thus

$$\|v(y + \cdot) - v(\cdot)\|_{L^p(\mathbb{R}^n)}^p \leq |y|^p \int_{\mathbb{R}^n} \Big[\int_0^1 |Dv(ty + x)|\, dt\Big]^p\, dx.$$

Applying Hölder's inequality to the integral inside the square brackets we get

$$\|v(y+\cdot)-v(\cdot)\|^p_{L^p(\mathbb{R}^n)} \le |y|^p \int_{\mathbb{R}^n}\int_0^1 |Dv(ty+x)|^p\,dt\,dx.$$

Interchanging the order of integration we obtain

$$\|v(y+\cdot)-v(\cdot)\|^p_{L^p(\mathbb{R}^n)} \le |y|^p \int_0^1\int_{\mathbb{R}^n} |Dv(ty+x)|^p\,dx\,dt = |y|^p\|Dv\|^p_{L^p(\mathbb{R}^n)},$$

as desired. □

3.5 Poincaré's inequality

In the following, for a given integrable function $u:\Omega\to\mathbb{R}$, we denote by $\bar{u}_\Omega$ the constant

$$\bar{u}_\Omega := \frac{1}{|\Omega|}\int_\Omega u$$

Theorem 3.5.1 (Poincaré's inequality)**.** *Suppose that $1\le p\le\infty$ and Ω is a bounded Lipschitz domain. There exists a constant $C_{n,p,\Omega}>0$ such that*

$$\|u-\bar{u}_\Omega\|_{L^p(\Omega)} \le C_{n,p,\Omega}\|Du\|_{L^p(\Omega)} \text{ for all } u\in W^{1,p}(\Omega).$$

Note that only the derivative of u appears on the right hand side.

Proof. Suppose by contradiction that the conclusion fails. Then we can find $(u_m)\subset W^{1,p}$ such that

$$\|u_m-(\bar{u}_m)_\Omega\|_{L^p(\Omega)} > m\|Du_m\|_{L^p(\Omega)}.$$

In particular, $\|u_m-(\bar{u}_m)_\Omega\|_{L^p(\Omega)}>0$. Set

$$v_m = \frac{u_m-(\bar{u}_m)_\Omega}{\|u_m-(\bar{u}_m)_\Omega\|_{L^p(\Omega)}}$$

so that $\|v_m\|_{L^p(\Omega)}=1$, $(\bar{v}_m)_\Omega=0$ and $\|Dv_m\|_{L^p(\Omega)}<\frac{1}{m}$. This implies that (v_m) is bounded in $W^{1,p}(\Omega)$. By the Rellich-Kondrachov's compactness theorem, after extracting a subsequence if necessary, we may assume that (v_m) converges strongly in $L^p(\Omega)$ to some $v\in L^p(\Omega)$.

As (v_m) converges to v strongly in $L^p(\Omega)$, we have

(i) $\|v\|_{L^p(\Omega)}=\lim\|v_m\|_{L^p(\Omega)}=1$, and

(ii) $\bar{v}_\Omega=\lim(\bar{v}_m)_\Omega=0$.

As Dv_m converges strongly to 0 in $L^p(\Omega)$, it follows that v is weakly differentiable with $Dv = 0$. Hence

(iii) $v \equiv$ constant a.e. in Ω.

Clearly (i), (ii) and (iii) amount to a contradiction. $\square$

Chapter 4

Functional Analytic Methods for PDEs

We now turn to the PDEs part of the course. We will consider linear, second-ordered partial differential equations of the form

$$Lu := -\partial_i(a_{ij}\partial_j u) + b_i\partial_i u + cu = f + \partial_i g_i \text{ in } \Omega. \tag{4.1}$$

Here Ω is a domain in $\mathbb{R}^n$, $u : \Omega \to \mathbb{R}$ is the unknown, $(a_{ij}) = (a_{ji})$, (b_i) and c are given coefficients, f and g_i are given sources, and repeated indices are summed from 1 to n.

Equation (4.1) can be written in a more compact form $Lu = -\text{div}(aDu) + b \cdot Du + cu = f + \text{div}g$ where $a = (a_{ij})$ is an $n \times n$ matrix and $b = (b_i)$ and $g = (g_i)$ are vectors. For this reason, (4.1) is called an equation in divergence form. Equations in non-divergence form takes the form

$$\tilde{L}u = -a_{ij}\partial_i\partial_j u + b_i\partial_i u + cu = f + \partial_i g_i \text{ in } \Omega. \tag{4.2}$$

Clearly, when a_{ij} is differentiable, one can recast an equation in divergence form as one in non-divergence form and vice versa. But this is not always possible for less regular coefficients.

In this course, we will only deal with equations in divergence form.

4.1 Dirichlet boundary value problem for second-ordered elliptic equations

Definition 4.1.1. *Let $a = (a_{ij}) : \Omega \to \mathbb{R}^{n\times n}$ be symmetric and have measurable entries.*

(i) *We say that a is* elliptic

$$a_{ij}(x)\xi_i\,\xi_j \geq 0 \text{ for all } \xi \in \mathbb{R}^n \text{ and a.e. } x \in \Omega.$$

(In other words, a is non-negative definite a.e. in Ω.)

(ii) *We say that a is* strictly elliptic *if there exists a constant $\lambda > 0$ such that*

$$a_{ij}(x)\xi_i\,\xi_j \geq \lambda|\xi|^2 \text{ for all } \xi \in \mathbb{R}^n \text{ and a.e. } x \in \Omega.$$

(iii) *We say that a is* uniformly elliptic *if there exist constants $0 < \lambda \leq \Lambda < \infty$ such that*

$$\lambda|\xi|^2 \leq a_{ij}(x)\xi_i\,\xi_j \leq \Lambda|\xi|^2 \text{ for all } \xi \in \mathbb{R}^n \text{ and a.e. } x \in \Omega.$$

Note that if a is uniformly elliptic, then $a_{ij} \in L^\infty(\Omega)$.
In this set of notes, we will assume that

$$a_{ij}, b_i, c \text{ belongs to } L^\infty(\Omega) \text{ and are given, and } (a_{ij}) \text{ is uniformly elliptic.}$$

The Dirichlet boundary value problem for L is to find for given sources f and g and a given boundary data u_0 a function u satisfying

$$\begin{cases} Lu = f + \partial_i g_i & \text{in } \Omega, \\ u = u_0 & \text{on } \partial\Omega. \end{cases} \tag{BVP}$$

Definition 4.1.2. *Suppose $a \in C^1(\Omega), b, c \in C(\Omega)$. For a given $f \in C(\Omega)$, $g \in C^1(\Omega)$ and $u_0 \in C(\partial\Omega)$, a function $u \in C^2(\Omega) \cap C(\bar{\Omega})$ is called a* classical solution to *the Dirichlet boundary value problem* (BVP) *if it satisfies* (BVP) *in the usual sense.*

Now if u is a classical solution for (BVP), we can multiply the equation $Lu = f$ by a smooth test function $\varphi \in C_c^\infty(\Omega)$ and integrate over Ω (and by parts) to obtain

$$\int_\Omega [a_{ij}\partial_j u \partial_i \varphi + b_i \partial_i u \varphi + cu\varphi] = \int_\Omega [f\varphi - g_i \partial_i \varphi].$$

By mean of approximation, the above identity holds true for $\varphi \in H_0^1(\Omega)$, and the identity make sense for u belonging to $H^1(\Omega)$. This motivates the following definition.

Definition 4.1.3. *Let $a, b, c \in L^\infty(\Omega)$, $f \in L^2(\Omega)$, $g \in L^2(\Omega)$ and $u_0 \in H^1(\Omega)$.*

(i) *The bilinear from $B(\cdot,\cdot)$ associated with the operator L defined in* (4.1) *is*

$$B(u, v) = \int_\Omega [a_{ij}\partial_j u \partial_i v + b_i \partial_i uv + cuv] \qquad u, v \in H^1(\Omega).$$

(ii) *We say that* $u \in H^1(\Omega)$ *is a* weak solution *(or* generalized solution*) to the equation* $Lu = f + \partial_i g_i$ *in* Ω *if*

$$B(u, \varphi) = \langle f, \varphi\rangle - \langle g_i, \partial_i \varphi\rangle \text{ for all } \varphi \in H_0^1(\Omega),$$

where $\langle \cdot, \cdot\rangle$ *denotes the inner product of* $L^2(\Omega)$*. When this holds, we say interchangeably that* u *satisfies* $Lu = f + \partial_i g_i$ *in* Ω *in the weak sense.*

(iii) *We say that* $u \in H^1(\Omega)$ *is a* weak solution *(or* generalized solution*) to the Dirichlet boundary value problem* (BVP) *if* $u - u_0 \in H_0^1(\Omega)$ [1] *and if* $Lu = f + \partial_i g_i$ *in the weak sense.*

Note that in the above definition, the boundary data u_0 is given as a function belong to $H^1(\Omega)$. In particular, it is defined on all of Ω. This is merely a technical point and can be taken care of by introducing appropriate functional spaces on $\partial\Omega$ which is ignored in this course.

We have the following estimates for the bilinear form B.

Theorem 4.1.4 (Energy estimates)**.** *Suppose that* $a, b, c \in L^\infty(\Omega)$*,* a *is uniformly elliptic,* L *is as in* (4.1) *and* B *is its associated bilinear form. There exists some large constant* $C > 0$ *such that*

$$|B(u, v)| \leq C\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}, \tag{4.3}$$

$$\frac{\lambda}{2}\|u\|^2_{H_1(\Omega)} \leq B[u, u] + C\|u\|^2_{L^2(\Omega)}. \tag{4.4}$$

Here λ *is the constant appearing in the definition of ellipticity of* a*.*

Proof. The proof of (4.3) is easy and left as an exercise. Let us prove (4.4). By the strict ellipticity and Cauchy-Schwarz' inequality, we have

$$\begin{aligned}
\lambda\|Du\|^2_{L^2(\Omega)} &\leq \int_\Omega a_{ij}\partial_i u \partial_j u = B(u, u) - \int_\Omega [b_i \partial_i u u + cu^2]\\
&\leq B(u, u) + \|b\|_{L^\infty(\Omega)}\|Du\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} + \|c\|_{L^\infty(\Omega)}\|u\|^2_{L^2(\Omega)}\\
&\leq B(u, u) + \frac{1}{2}\lambda\|Du\|^2_{L^2(\Omega)} + \frac{1}{2\lambda}\|b\|^2_{L^\infty(\Omega)}\|u\|^2_{L^2(\Omega)} + \|c\|_{L^\infty(\Omega)}\|u\|^2_{L^2(\Omega)}.
\end{aligned}$$

It follows that

$$\frac{1}{2}\lambda\|Du\|^2_{L^2(\Omega)} \leq B(u, u) + \Big[\frac{1}{2\lambda}\|b\|^2_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\Big]\|u\|^2_{L^2(\Omega)},$$

from which the conclusion follows. □

[1] This would be the same as saying that the traces of u and of u_0 agree on $\partial\Omega$ when $\partial\Omega$ is sufficiently regular.

4.2 Existence theorems

4.2.1 Existence via the direct method in the calculus of variations

In some cases, the Dirichlet boundary value problem (BVP) can be solved by a variational approach. Let us illustrate this in the case $b \equiv 0$ and $c \geq 0$.

We will need the following result from functional analysis.

Theorem 4.2.1 (Mazur). *Let K be a closed convex subset of a normed vector space X, (x_n) be a sequence of points in K converging weakly to x. Then $x \in K$.*

We prove:

Theorem 4.2.2 (Existence via direct minimization). *Suppose that $a, c \in L^\infty(\Omega)$, a is uniformly elliptic, $c \geq 0$ a.e. in Ω, $b \equiv 0$ and L is as in* (4.1). *Then for every $f \in L^2(\Omega)$, $g \in L^2(\Omega)$ and $u_0 \in H^1(\Omega)$, the Dirichlet boundary value problem* (BVP) *has a unique weak solution $u \in H^1(\Omega)$.*

Proof. The key point is that the problem (BVP) is related to the following so-called variational energy

$$I[u] = \frac{1}{2}B(u,u) - \langle f, u\rangle + \langle g_i, \partial_i u\rangle.$$

We will show that the solution to (BVP) is the unique minimizer or I on $X := \{u \in H^1(\Omega) : u - u_0 \in H^1_0(\Omega)\}$.

Let $\alpha = \inf_X I \in [-\infty, I[u_0]]$. Then we can pick $u_m \in X$ such that $I[u_m] \to \alpha$.

<u>Step 1:</u> We show that the sequence (u_m) is bounded in $H^1(\Omega)$.

Indeed, we have by strict ellipticity and the non-negativity of c that

$$\begin{aligned}\lambda\|Du_m\|^2_{L^2(\Omega)} &\leq \int_\Omega a_{ij}\partial_i u \partial_j u \leq B(u,u)\\ &\leq 2I[u_m] + 2\langle f, u_m\rangle - 2\langle g_i, \partial_i u_m\rangle\\ &\leq 2I[u_m] + 2\|f\|_{L^2(\Omega)}\|u_m\|_{L^2(\Omega)} + \frac{2}{\lambda}\|g\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|Du_m\|_{L^2(\Omega)},\end{aligned}$$

we we have used Cauchy-Schwarz' inequality. As $I[u_m] \to \alpha \leq I[u_0]$, we thus have that $(I[u_m])$ is bounded. Hence we can find some C such that

$$\|Du_m\|^2_{L^2(\Omega)} \leq C + C\|u_m\|_{L^2(\Omega)}. \tag{4.5}$$

By Minkowski's inequality, this implies

$$\|D(u_m - u_0)\|^2_{L^2(\Omega)} \leq \|Du_0\|^2_{L^2(\Omega)} + C + C\|u_m\|_{L^2(\Omega)} \leq C + C\|u_m\|_{L^2(\Omega)}.$$

By Friedrichs' inequality (Theorem 3.2.1), this implies

$$\|u_m - u_0\|^2_{L^2(\Omega)} \le C + C\|u_m\|_{L^2(\Omega)},$$

and so by Minkowski's inequality,

$$\|u_m\|^2_{L^2(\Omega)} \le \|u_0\|^2_{L^2(\Omega)} + C + C\|u_m\|_{L^2(\Omega)} \le C + C\|u_m\|_{L^2(\Omega)}. \tag{4.6}$$

Putting together (4.5) and (4.6) we conclude Step 1.

Step 2: The subconvergence of (u_m) to a minimizer of $I|_X$.

Since $H^1(\Omega)$ is reflexive, the bounded sequence (u_m) has a weakly convergent subsequence. We still denote this subsequence (u_m) and say $u_m \rightharpoonup u$ in $H^1(\Omega)$.

We also have that $u_m - u_0 \rightharpoonup u - u_0$. Note that $H^1_0(\Omega)$ is closed (by definition) and convex. By Mazur's theorem, $H^1_0(\Omega)$ is weakly closed, and so $u - u_0 \in H^1_0(\Omega)$, i.e. $u \in X$.

We claim that $\liminf I[u_m] \ge I[u]$ (and so $I[u] = \alpha$, i.e. u minimizes $I|_X$). By the weak convergence of u_m and Du_m to u and Du, respectively, in $L^2(\Omega)$, we have that $\langle f, u_m\rangle \to \langle f, u\rangle$ and $\langle g_i, \partial_i u_m\rangle \to \langle g_i, \partial_i u\rangle$. Thus it suffices to show that

$$\liminf_{m\to\infty} B(u_m, u_m) \ge B(u, u). \tag{4.7}$$

To this end, we use the explicit form of B:

$$\begin{aligned} B(u_m, u_m) - B(u, u) &= \int_\Omega [a_{ij}\partial_i u_m \partial_j u_m + c u_m^2] - \int_\Omega [a_{ij}\partial_i u \partial_j u + c u_m^2] \\ &= \int_\Omega [a_{ij}\partial_i (u_m - u)\partial_j (u_m - u) + c(u_m - u)^2] \\ &\quad + \int_\Omega [a_{ij}\partial_i (u_m - u)\partial_j u + a_{ij}\partial_i u \partial_j (u_m - u) + 2c(u_m - u)u]. \end{aligned}$$

The first integral on the right hand side is non-negative due to the ellipticity. The second integral converges to zero as $D(u_m - u) \rightharpoonup 0$ and $(u_m - u) \rightharpoonup 0$ in $L^2(\Omega)$. This proves (4.7). So we have $\alpha = \liminf I[u_m] \ge I[u]$. As $u \in X$ and $\alpha = \inf_X I$ it follows that $I[u] = \alpha = \inf_X I$, which concludes Step 2.

Step 3: We show that u is a weak solution to the problem (BVP).

As $u - u_0 \in H^1_0(\Omega)$, we only need to show that $B(u, \varphi) = \langle f, \varphi\rangle$ for all $\varphi \in H^1_0(\Omega)$. Indeed, if $\varphi \in H^1_0(\Omega)$, then $u + t\varphi \in X$ and so $I[u] \le I[u + t\varphi]$ for every $t \in \mathbb{R}$. It is clear that the map $t \mapsto I[u+t\varphi]$ is differentiable and so

$$\begin{aligned}
0 &= \frac{d}{dt}\Big|_{t=0} I[u+t\varphi] \\
&= \frac{d}{dt}\Big|_{t=0} \int_\Omega [\frac{1}{2}a_{ij}\partial_i(u+t\varphi)\partial_j(u+t\varphi) + \frac{1}{2}c(u+t\varphi)^2 - f(u+t\varphi) + g_i\partial_i(u+t\varphi)] \\
&= \int_\Omega [\frac{1}{2}a_{ij}\partial_i u\partial_j\varphi + \frac{1}{2}a_{ij}\partial_i\varphi\partial_i u + cu\varphi - f\varphi + g_i\partial_i\varphi] \\
&\overset{a_{ij}=a_{ji}}{=} \int_\Omega [a_{ij}\partial_i u\partial_j\varphi + cu\varphi - f\varphi + g_i\partial_i\varphi] \\
&= B(u,\varphi) - \langle f,\varphi\rangle + \langle g_i, \partial_i\varphi\rangle,
\end{aligned}$$

which gives the required identity.

Step 4: We prove the uniqueness: If $\bar{u}$ is also a weak solution to (BVP), then $\bar{u} = u$ a.e.

Indeed, as $B(u,\varphi) = \langle f,\varphi\rangle - \langle g_i,\partial_i\varphi\rangle = B(\bar{u},\varphi)$ for all $\varphi \in H_0^1(\Omega)$, we thus have that $B(u-\bar{u},\varphi) = 0$ for all $\varphi \in H_0^1(\Omega)$. As $u - \bar{u} \in H_0^1(\Omega)$ it follows that $B(u-\bar{u}, u-\bar{u}) = 0$. Hence, by the ellipticity and the non-negativity of c, this implies that

$$\lambda\|D(u-\bar{u})\|^2_{L^2(\Omega)} \leq B(u-\bar{u}, u-\bar{u}) = 0,$$

and so $\|D(u-\bar{u})\|_{L^2(\Omega)} = 0$. By Friedrichs' inequality (Theorem 3.2.1), this then gives $\|u-\bar{u}\|_{L^2(\Omega)} = 0$, and so $u = \bar{u}$ a.e., which concludes the proof. □

4.2.2 Fredholm alternative

For more general coefficients, problem (BVP) does not always have a solution.

Example 4.2.3. *Let $\Omega = (0,\pi) \subset \mathbb{R}$, $L = -\frac{d^2}{dx^2} - 1$, $u_0 = 0$. If the problem* (BVP) *has a weak solution, then $\int_0^\pi f(x)\sin x\,dx = 0$. For if $u \in H_0^1(0,\pi)$ is a weak solution, then*

$$\int_0^\pi f(x)\sin x\,dx = \int_0^\pi [u'(x)(\sin x)' - u(x)\,\sin x]\,dx = \int_0^\pi u(x)[-(\sin x)'' - \sin x]\,dx = 0.$$

We will see that this is also a sufficient condition for existence.

Definition 4.2.4. *Let $Lu = -\partial_i(a_{ij}\partial_j u) + b_i u_i + cu$. The formal adjoint L^* of L is defined as the operator*

$$L^* v = -\partial_i(a_{ij}\partial_j v) - \partial_i(b_i v) + cv.$$

We say that $L^*v = f + \partial_i g_i$ in Ω in the weak sense if

$$B(\varphi, v) = \langle \varphi, f\rangle - \langle \partial_i \varphi, g_i\rangle \text{ for all } \varphi \in H^1_0(\Omega).$$

where B is the bilinear form associated to L and $\langle\cdot,\cdot\rangle$ is the inner product of $L^2(\Omega)$.

Note that if u is a solution to (BVP) and $v \in H^1_0(\Omega)$ satisfies $L^*v = 0$, then, as $u - u_0 \in H^1_0(\Omega)$,

$$\langle f, v\rangle - \langle g_i, \partial_i v\rangle \overset{Lu=f+\partial_i g_i}{=} B(u,v) = B(u_0, v) + B(u - u_0, v) \overset{L^*v=0}{=} B(u_0, v).$$

We will see now that this is the main 'obstacle' for existence and uniqueness.

Theorem 4.2.5 (Fredholm alternative)**.** *Suppose that Ω is a bounded Lipschitz domain. Suppose that $a, b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in* (4.1).

(i) We have the dichotomy: either

$$\begin{cases} \text{For each } f \in L^2(\Omega),\ g \in L^2(\Omega) \text{ and } u_0 \in H^1(\Omega), \text{ there exists a unique weak solution } u \in H^1(\Omega) \text{ to the boundary value problem (BVP),} \end{cases} \tag{4.8}$$

or

$$\begin{cases} \text{There exists a non-trivial weak solution } 0 \not\equiv w \in H^1(\Omega) \text{ to the homogeneous problem} \\ \qquad \begin{cases} Lu = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \qquad \text{(Hom)} \end{cases} \tag{4.9}$$

(ii) In case (4.9) *holds, the space N of weak solutions to* (Hom) *is a finite dimensional subspace of $H^1_0(\Omega)$. Furthermore, the dimension of N is equal to the dimension of the space $N^* \subset H^1_0(\Omega)$ of weak solutions to*

$$\begin{cases} L^*v = 0 & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega. \end{cases} \tag{Hom*}$$

(iii) Finally, the boundary value problem (BVP) *has a solution if and only if*

$$B(u_0, v) = \langle f, v\rangle - \langle g_i, \partial_i v\rangle \text{ for all } v \in N^*.$$

We will only the pursue the proof of (i) and omit that of (ii) and (iii). Part (i) can be equivalently restated as follows.

Theorem 4.2.6 (Uniqueness implies existence)**.** *Suppose that Ω is a bounded Lipschitz domain. Suppose that $a, b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in* (4.1)*. If the only weak solution to* (Hom) *is the trivial solution, then for every $f \in L^2(\Omega)$, $g \in L^2(\Omega)$ and $u_0 \in H^1(\Omega)$, the boundary value problem* (BVP) *has a unique weak solution $u \in H^1(\Omega)$.*

An immediate consequence of this theorem is the following (which is stronger than Theorem 4.2.2).

Theorem 4.2.7. *Suppose that Ω is a bounded Lipschitz domain. Suppose that $a, b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in* (4.1)*. If the bilinear form B associated to L is* coercive*, i.e. there is a constant $C > 0$ such that*

$$B(w, w) \geq C\|w\|^2_{L^2(\Omega)} \text{ for all } w \in C_c^\infty(\Omega),$$

then the boundary value problem (BVP) *has a unique solution for every $f \in L^2(\Omega)$, $g \in L^2(\Omega)$ and $u_0 \in H^1(\Omega)$.*

Let us start with some functional analytic preliminaries.

Definition 4.2.8. *Let H be a Hilbert space. An bounded linear operator $K : H \to H$ is said to be* compact *if K maps bounded subset of H into pre-compact subsets of H.*

Theorem 4.2.9 (Projection theorem)**.** *If Y is a closed subspace of a Hilbert space H, then Y and $Y^\perp$ are complementary subspaces: $H = Y \oplus Y^\perp$, i.e. every $x \in H$ can be decomposed uniquely as a sum of a vector in Y and in $Y^\perp$.*

Theorem 4.2.10 (Fredholm alternative)**.** *Let H be a Hilbert space and $K : H \to H$ be a compact bounded linear operator. Then we have the dichotomy that either $I - K$ is invertible or Ker$(I - K)$ is non-trivial.*

Proof. (*–Not for examination–*) Suppose that $\text{Ker}(I - K) = 0$. To conclude, we need to show that $V = \text{Im}(I - K)$ is the whole of H. Suppose by contradiction that V is a proper subspace of H.

Step 1: We show that V is closed.

Suppose that $(u_m) \subset H$ is such that $v_m = (I - K)(u_m) \in V$ converges to some $x \in H$. We need to show that $x \in V$.

We claim that (u_m) is bounded. Otherwise, there is a subsequence (u_{m_j}) with $\|u_{m_j}\| \to \infty$. Let $\tilde{u}_{m_j} = \frac{u_{m_j}}{\|u_{m_j}\|}$ and $\tilde{v}_{m_j} = (I - K)\tilde{u}_{m_j} = \frac{v_{m_j}}{\|u_{m_j}\|}$. Note that as (v_m) is convergent, $\tilde{v}_{m_j} \to 0$. On the other hand, as $(\tilde{u}_{m_j})$ is bounded and K is compact, we can assume after passing to a subsequence if necessary that $K\tilde{u}_{m_j}$ converges to some $y \in H$. It follows that $\tilde{u}_{m_j} = \tilde{v}_{m_j} + K\tilde{u}_{m_j}$ converges to y. We hence have on one hand that $\|y\| = 1$ (due to $\|\tilde{u}_{m_j}\| = 1$) and on the other hand that $(I-K)y = 0$ (as the common limit of $(I-K)(\tilde{u}_{m_j}) = \tilde{v}_{m_j}$). These contradicts one another as $\text{Ker}(I-K) = 0$. The claim is proved.

As (u_m) is bounded and K is compact, there is a subsequence such that Ku_{m_j} converges to some $z \in H$. It follows that $u_{m_j} = v_{m_j} + Ku_{m_j} \to x + z$ and so $x = (I-K)(x+z) \in V$. This finishes Step 1.

Step 2: Let $V_0 = H$ and define inductively $V_{m+1} = (I-K)(V_m)$. We show that each V_{m+1} is a proper closed subspace of V_m, $m \geq 0$.

For $m = 0$, this follows from the contradiction hypothesis that $V_1 = V$ is a proper subspace of H and Step 1 that $V_1 = V$ is closed. Assume that the statement has been proved for some $m \geq 0$. We need to show that V_{m+2} is a proper closed subspace of V_{m+1}.

Note that $I - K$ maps V_{m+1} into V_{m+1} and so K maps V_{m+1} into V_{m+1}. Since V_{m+1} is a closed subspace of H, it is a Hilbert space, and so Step 1 applied to the compact map $K|_{V_{m+1}}$ shows that $V_{m+2} = (I-K)(V_{m+1})$ is a closed subspace of V_{m+1}.

Next, as V_{m+1} is a proper subspace of V_m, we can pick $u \in V_m \setminus V_{m+1}$. Now if we had $V_{m+2} = V_{m+1}$, then as $(I-K)u \in V_{m+1} = V_{m+2}$ we could find $w \in V_m$ such that $(I-K)u = (I-K)^2 w$. As $\text{Ker}(I-K) = 0$, this would imply $u = (I-K)w \subset (I-K)(V_m) = V_{m+1}$, which would be a contradiction. Hence V_{m+2} is a proper subspace of V_{m+1}.

Step 3: We conclude using the projection theorem.

From Step 2, we have a nested sequence of proper closed subspaces $H = V_0 \supset V_1 \supset V_2 \supset \ldots$ By the projection theorem (Theorem 4.2.9), we can decompose V_m into direct sum of orthogonal complementary closed subspaces $V_m = V_{m+1} \oplus W_{m+1}$ where $W_{m+1} = \{w \in V_m : \langle v, w\rangle = 0 \ \forall\ v \in V_{m+1}\}$.

Select $w_m \in W_{m+1}$ such that $\|w_{m+1}\| = 1$. As K is compact (Kw_m) has a convergent subsequence. To reach a contradiction, we shows that (Kw_m) has no Cauchy subsequence.

Fix $m > l$. Then $w_m \in W_{m+1} \subset V_{l+1}$, $(I-K)w_l \in (I-K)(V_l) = V_{l+1}$ and $(I-K)w_m \in (I-K)(V_m) = V_{m+1} \subset V_{l+1}$. It follows that

$$Kw_l - Kw_m = \underbrace{(I-K)w_m - (I-K)w_l - w_m}_{\in V_{l+1}} + \underbrace{w_l}_{\in W_{l+1}},$$

and so, by Pythagoras' theorem, $\|Kw_l - Kw_m\| \geq \|w_l\| = 1$. Hence (Kw_m) has no Cauchy subsequence. The proof is complete. □

Proof of Theorem 4.2.6. Step 0: Reduction to the case $u_0 \equiv 0$.

Note that the problem (BVP) can be recast as a problem for $\tilde{u} = u - u_0$ as follows:

$$\begin{cases} L\tilde{u} = \tilde{f} + \partial_i \tilde{g}_i & \text{in } \Omega, \\ \tilde{u} = 0 & \text{on } \partial\Omega. \end{cases}$$

where $\tilde{f} = (f - b_i \partial_i u_0 - c u_0)$ and $\tilde{g}_i = g_i + a_{ij} \partial_j u_0$. Thus it is enough to consider the case $u_0 \equiv 0$, which we will assume from now on.

Step 1: Consideration of the top order operator L_{top} defined by $L_{top} u = -\partial_i (a_{ij} \partial_j u)$.

We knew from Theorem 4.2.2 that for every $f \in L^2(\Omega)$ and $g \in L^2(\Omega)$, the problem

$$\begin{cases} L_{top} u = f + \partial_i g_i & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{BVP$_{top}$}$$

has a unique solution $u \in H_0^1(\Omega)$. We denote this solution as $A(f, g)$ so that A defines a linear operator from $L^2(\Omega) \times (L^2(\Omega))^n$ into $H_0^1(\Omega)$. Also, as $u \in H_0^1(\Omega)$, we can use it as a test function in the weak formulation of (BVP$_{top}$) to obtain

$$B_{top}(u, u) \leq \langle f, u \rangle - \langle g_i, \partial_i u \rangle \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}) \|u\|_{H^1(\Omega)},$$

where B_{top} is the bilinear form associated with L_{top}. By ellipticity, we have $B_{top}(u, u) \geq \lambda \|Du\|^2_{L^2(\Omega)}$. Thus, in view of Friedrichs' inequality (Theorem 3.2.1), we have

$$\|u\|^2_{H^1(\Omega)} \leq \|Du\|^2_{L^2(\Omega)} \leq C B_{top}(u, u) \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}) \|u\|_{H^1(\Omega)},$$

and so

$$\|A(f, g)\|_{H^1(\Omega)} = \|u\|_{H^1(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

This shows that A is a bounded operator.

Step 2: We recast (BVP) as an equation in the form $(I - K)u = x$ where K is a linear operator from $H_0^1(\Omega)$ into itself.

Observe that (BVP) is equivalent to

$$\begin{cases} L_{top} u = (f - b_i \partial_i u - c u) + \partial_i g_i & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

So $u \in H_0^1(\Omega)$ is a weak solution to (BVP) if and only if

$$u = A(f - b_i \partial_i u - c u, g).$$

We now define $K : H_0^1(\Omega) \to H_0^1(\Omega)$ by

$$K u = A(-b_i \partial_i u - c u, 0).$$

and let $x = A(f, g) \in H_0^1(\Omega)$. Clearly, as A is bounded linear, so is K. We are thus led to show that $(I - K)u = x$ has a unique solution u, given that the kernel of $I - K$ is trivial.

Step 3: In view of the Fredholm alternative (Theorem 4.2.10), to conclude it suffices to show that K is compact, i.e. for every bounded sequence $(u_m) \subset H_0^1(\Omega)$, there is a subsequence u_{m_j} such that (Ku_{m_j}) is convergent.

As $H^1(\Omega)$ is reflexive and (u_m) and (Ku_m) are bounded, we may assume after passing to a subsequence that (u_m) and (Ku_m) converges weakly in H^1 to some $u \in H_0^1(\Omega)$ and $w \in H_0^1(\Omega)$. In addition, by Rellich-Kondrachov's theorem, we may also assume that (u_m) converges strongly in L^2 to u.

We claim that $w = Ku$. Indeed, since $Ku_m = A(-b_i\partial_i u_m - cu_m, 0)$, we have $B(Ku_m, \varphi) = \langle -b_i\partial_i u_m - cu_m, \varphi\rangle$ for all $\varphi \in H_0^1(\Omega)$. The weak convergence of (u_m), (Du_m), and (Ku_m) to u, Du and w, respectively, in L^2 thus implies that $B(w, \varphi) = \langle -b_i\partial_i u - cu, \varphi\rangle$ for all $\varphi \in H_0^1(\Omega)$. This means that $w = A(-b_i\partial_i u - cu, 0) = Ku$, as claim.

Let $\mathring{u}_m = u_m - u$. Then $K\mathring{u}_m = A(-b_i\partial_i \mathring{u}_m - c\mathring{u}_m, 0)$ and so $B(K\mathring{u}_m, \varphi) = \langle -b_i\partial_i \mathring{u}_m - c\mathring{u}_m, \varphi\rangle$ for all $\varphi \in H_0^1(\Omega)$. In particular, for $\varphi = K\mathring{u}_m$, we have

$$B(K\mathring{u}_m, K\mathring{u}_m) = \langle \underbrace{-b_i\partial_i \mathring{u}_m - c\mathring{u}_m}_{\to 0 \text{ in } L^2}, \underbrace{K\mathring{u}_m}_{\to 0 \text{ in } L^2} \rangle \to 0.$$

On the other hand, by ellipticity and Friedrichs' inequality (Theorem 3.2.1),

$$B(K\mathring{u}_m, K\mathring{u}_m) \geq \lambda \|DK\mathring{u}_m\|_{L^2(\Omega)}^2 \geq \frac{1}{C}\|K\mathring{u}_m\|_{H^1(\Omega)}^2.$$

It follows that $K\mathring{u}_m \to 0$ in H^1, i.e. (Ku_m) converges strongly in H^1 to Ku. This shows that K is compact and concludes the proof. □

4.2.3 Spectrum of elliptic differential operators under Dirichlet boundary condition

In this section, we restrict our attention to the case that $g \equiv 0$ and $u_0 \equiv 0$.

Theorem 4.2.11 (Spectrum of elliptic operators)**.** *Suppose that Ω is a bounded Lipschitz domain. Suppose that $a, b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in* (4.1)*. Then there exists an at most countable set $\Sigma \subset \mathbb{R}$ such that the boundary value problem*

$$\begin{cases} Lu = \lambda u + f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{EBVP}$$

has a unique solution if and only if $\lambda \notin \Sigma$*. Furthermore, if* Σ *is infinite then* $\Sigma = \{\lambda_k\}_{k=1}^{\infty}$ *with*

$$\lambda_1 \leq \lambda_2 \leq \ldots \to \infty.$$

The set Σ is called the (real) spectrum of the operator L.

The heart of the theorem above is the following general result for compact operators, whose proof is omitted.

Theorem 4.2.12 (Spectrum of compact operators)**.** *Let* H *be a Hilbert space of infinite dimension,* $K : H \to H$ *be a compact bounded linear operator and* $\sigma(K)$ *be its spectrum (i.e. the set of* $\lambda \in \mathbb{C}$ *such that* $\lambda I - K$ *is not invertible). Then*

(i) 0 *belongs to* $\sigma(K)$*.*

(ii) $\sigma(K) \setminus \{0\} = \sigma_p(K) \setminus \{0\}$*, i.e.* $\lambda I - K$ *has non-trivial kernel for* $\lambda \in \sigma(K) \setminus \{0\}$*.*

(iii) $\sigma(K) \setminus \{0\}$ *is either finite or an infinite sequence tending to* 0*.*

Proof of Theorem 4.2.11. By Theorem 4.1.4 there exists some large $\mu > 0$ such that the operator $L_\mu u = Lu + \mu u$ has a coercive bilinear form $B_\mu(u, v) = B(u, v) + \mu\langle u, v\rangle$. By Theorem 4.2.7, the problem

$$\begin{cases} L_\mu u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

is uniquely solvable for every $f \in L^2(\Omega)$. Call the solution Kf so that K is a bounded linear map from $L^2(\Omega)$ into itself. Note that as $K(L^2(\Omega)) \subset H_0^1(\Omega)$ and $H_0^1(\Omega)$ is compactly embedded in $L^2(\Omega)$, K is a compact operator.

Now let Σ be the set of $\lambda \in \mathbb{R}$ such that (EBVP) is not always uniquely solvable. By the Fredholm alternative, $\lambda \in \Sigma$ if and only if the problem

$$\begin{cases} Lu = \lambda u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

has a non-trivial solution. In other words, this means that the equation $K((\mu+\lambda)u) = u$ has a non-trivial solution. The conclusion then follows from Theorem 4.2.12. □

4.3 Regularity theorems

4.3.1 Differentiable leading coefficients

We will now turns to the study of regularity. We have

Theorem 4.3.1 (Interior H^2 regularity)**.** *Suppose that $a \in C^1(\Omega)$, $b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in (4.1). Suppose that $f \in L^2(\Omega)$. If $u \in H^1(\Omega)$ satisfies $Lu = f$ in Ω in the weak sense then $u \in H^2_{loc}(\Omega)$, and for any open ω such that $\bar{\omega} \subset \Omega$ we have*

$$\|u\|_{H^2(\omega)} \leq C(\|f\|_{L^2(\Omega)} + \|u\|_{H^1(\Omega)})$$

where the constant C depends only on $n, \Omega, \omega, a, b, c$.

Theorem 4.3.2 (Global H^2 regularity)**.** *Suppose that $a \in C^1(\bar{\Omega})$, $b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in (4.1) and that $\partial\Omega$ is C^2 regular. Suppose that $f \in L^2(\Omega)$. If $u \in H^1_0(\Omega)$ satisfies $Lu = f$ in Ω in the weak sense then $u \in H^2(\Omega)$ and*

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$$

where the constant C depends only on n, Ω, a, b, c.

Theorem 4.3.3 (Global C^∞ regularity)**.** *Suppose that $a, b, c \in C^\infty(\bar{\Omega})$, a is uniformly elliptic, and L is as in (4.1) and that $\partial\Omega$ is C^∞ regular. Suppose that $f \in C^\infty(\Omega)$. If $u \in H^1_0(\Omega)$ satisfies $Lu = f$ in Ω in the weak sense then $u \in C^\infty(\Omega)$.*

To understand better the idea, let us focus on the proof of Theorem 4.3.1 in the simplest but nevertheless important case $a = (\delta_{ij})$, $b \equiv 0$, $c \equiv 0$, i.e. $L = -\Delta$, and Ω is the ball B_2 and ω is the ball B_1.

We start with an important auxiliary result.

Lemma 4.3.4. *Suppose that $u \in C^\infty_c(\mathbb{R}^n)$. Then*

$$\|D^2u\|_{L^2(\mathbb{R}^n)} = \|\Delta u\|_{L^2(\mathbb{R}^n)}.$$

Proof. The proof is a direct computation using integration by parts. We compute

$$\begin{aligned}\|D^2u\|^2_{L^2(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \partial_i\partial_j u \partial_i\partial_j u = -\int_{\mathbb{R}^n} \partial_j u \partial_j \partial_i^2 u \\ &= \int_{\mathbb{R}^n} \partial_j^2 u \partial_i^2 u = \|\Delta u\|^2_{L^2(\mathbb{R}^n)},\end{aligned}$$

which is exactly what we claimed. □

Proof of Theorem 4.3.1 in the above simple setting. Step 1: Reduction to regular estimates for solutions which vanish near $\partial\Omega$.

Fix a cut-off function $\zeta \in C^\infty_c(B_2)$ such that $\zeta \equiv 1$ in B_1. Let $w := \zeta u$. We claim that satisfies $-\Delta w = (\zeta f - D\zeta \cdot Du) - \partial_i(u\partial_i\zeta)$ in B_2 in the weak sense, i.e.

$$\int_{B_2} Dw \cdot Dv = \int_{B_2} [(\zeta f - D\zeta \cdot Du)v + uD\zeta \cdot Dv] \text{ for all } v \in H^1_0(B_2).$$

Using $w = \zeta u$, we see that this is equivalent to

$$\int_{B_2} \zeta Du \cdot Dv = \int_{B_2} (\zeta f - D\zeta \cdot Du)v \text{ for all } v \in H_0^1(B_2),$$

which upon rearranging term is equivalent to

$$\int_{B_2} Du \cdot D(\zeta v) = \int_{B_2} f(\zeta v) \text{ for all } v \in H_0^1(B_2),$$

As $\zeta v \in H_0^1(B_2)$ and $-\Delta u = f$ in B_2 in the weak sense, this latter identity holds true, whence the original identity.

Now if the conclusion has been established for functions which vanish near the boundary, then such estimate applies to w. Hence $w \in H^2(B_1)$ and

$$\|w\|_{H^2(B_1)} \leq C(\|(\zeta f - D\zeta \cdot Du) - \partial_i(u\partial_i\zeta)\|_{L^2(B_2)} + \|w\|_{H^1(B_2)}) \leq C(\|f\|_{L^2(B_2)} + \|u\|_{H^1(B_2)}),$$

which gives the desired estimate.

Step 2: Reduction to a priori estimate on the whole space and conclusion of proof.

Suppose that $u \in H_0^1(B_2)$ vanishes near ∂B_2 and satisfies $-\Delta u = f$ in B_2 in the weak sense for some $f \in L^2(B_2)$. Extend u to be zero outside of B_2.

Fix a non-negative function $\varrho \in C_c^\infty(B_1)$ with $\int_{\mathbb{R}^n} \varrho = 1$ and let $\varrho_\varepsilon(x) = \varepsilon^{-n}\varrho(x/\varepsilon)$ be the usual mollifiers. Set $u_\varepsilon = \varrho_\varepsilon * u$ and $f_\varepsilon = \varrho_\varepsilon * f$. Then $u_\varepsilon, f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$.

We claim that $-\Delta u_\varepsilon = f_\varepsilon$ in $\mathbb{R}^n$. By Lemma 2.3.1, we know that

$$\partial_i u_\varepsilon = \varrho_\varepsilon * \partial_i u.$$

We hence use Fubini's theorem to compute for $v \in C_c^\infty(B_2)$ that

$$\begin{aligned}
\int_{B_2} Du_\varepsilon \cdot Dv &= \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n} \varrho_\varepsilon(x-y)\partial_{y_i} u(y)\,dy\Big]\partial_{x_i} v(x)\,dx\\
&= \int_{\mathbb{R}^n} \partial_{y_i} u(y)\Big[\int_{\mathbb{R}^n} \varrho_\varepsilon(x-y)\partial_{x_i} v(x)\,dx\Big]\,dy\\
&= -\int_{\mathbb{R}^n} \partial_{y_i} u(y)\Big[\int_{\mathbb{R}^n} \partial_{x_i}\varrho_\varepsilon(x-y) v(x)\,dx\Big]\,dy\\
&= \int_{\mathbb{R}^n} \partial_{y_i} u(y)\Big[\int_{\mathbb{R}^n} \partial_{y_i}\varrho_\varepsilon(x-y) v(x)\,dx\Big]\,dy\\
&= \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n} \partial_{y_i} u(y)\partial_{y_i}\varrho_\varepsilon(x-y)\,dy\Big] v(x)\,dx\\
&= \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n} f(y)\varrho_\varepsilon(x-y)\,dy\Big] v(x)\,dx\\
&= \int_{\mathbb{R}^n} f_\varepsilon(x) v(x)\,dx.
\end{aligned}$$

Since $v \in C_c^\infty(\mathbb{R}^n)$ is arbitrary, this shows that $-\Delta u_\varepsilon = f_\varepsilon$ in $\mathbb{R}^n$ the weak sense. As both u_ε and f_ε are smooth, we thus have that $-\Delta u_\varepsilon = f_\varepsilon$ in the classical sense, as claimed.

We are now in position to apply Lemma 4.3.4. We have

$$\|D^2 u_\varepsilon\|_{L^2(\mathbb{R}^n)} = \|\Delta u_\varepsilon\|_{L^2(\mathbb{R}^n)} = \|f_\varepsilon\|_{L^2(\mathbb{R}^n)}$$

By Young's convolution inequality, we thus have

$$\|D^2 u_\varepsilon\|_{L^2(\mathbb{R}^n)} \leq \|f\|_{L^2(B_2)}.$$

This implies on the one hand that, along a subsequence, $(D^2 u_\varepsilon)$ converges weakly to some $A \in L^2(\mathbb{R}^n)$ with $\|A\|_{L^2(\mathbb{R}^n)} \leq \|f\|_{L^2(B_2)}$. Since we also knew that (u_ε) converges strongly to u in $H^1(\mathbb{R}^n)$ (by Theorem 2.3.2), can send $\varepsilon \to 0$ in the identity

$$\int_{\mathbb{R}^n} u_\varepsilon \partial_i \partial_j v = \int_{\mathbb{R}^n} \partial_i \partial_j u_\varepsilon v$$

to see that u admits weak second derivatives in and $D^2 u = A \in L^2(\mathbb{R}^n)$.

We have thus shown $u \in H^2(\mathbb{R}^n)$ and and $\|D^2 u\|_{L^2(\mathbb{R}^n)} \leq \|f\|_{L^2(B_2)}$, from which the assertion follows. □

Let us now briefly indicate how the results in the case $L = -\Delta$ can lead to results the case of variable coefficients. First of all, the case when a is a constant matrix can be reduced to the case of the Laplacian by a change of variable. The case of variable coefficients is treated using the so-called method of freezing coefficients. If x_0 is a given point in Ω, let $a^0_{ij} = a_{ij}(x_0)$ and $L^0 u = -\partial_i(a^0_{ij}\partial_j u)$. Then the equation $Lu = f$ can be re-expressed as

$$L^0 u = -(a^0_{ij} - a_{ij})\partial_i \partial_j u + \partial_i a_{ij} \partial_j u - b_i \partial_i u - cu + f$$

Now if the global regular estimate for L^0 has been established, then we will have, after suitably cutting off the solution so that u is compactly supported in ω as in Step 1 above, that

$$\begin{aligned} \|u\|_{H^2(\omega)} &\leq C\| - (a^0_{ij} - a_{ij})\partial_i \partial_j u + \partial_i a_{ij} \partial_j u - b_i \partial_i u - cu + f\|_{L^2(\omega)} \\ &\leq C \sup_\omega |a^0_{ij} - a_{ij}| \|D^2 u\|_{L^2(\omega)} + C\|u\|_{H^1(\omega)} + \|f\|_{L^2(\omega)}. \end{aligned}$$

Now if ω is chosen sufficiently small from the start so that $C \sup_\omega |a^0_{ij} - a_{ij}|$ is smaller than 1 (which is possible since a is continuous), then the term containing second derivative on the right hand side above can be absorbed into the left hand side, yielding the desired estimate. The case of general non-small ω is treated by using a finite cover of small balls.

4.3.2 Bounded measurable leading coefficients

We conclude this set of lecture notes with the following remarkable result which only requires that the coefficients are measurable.

Theorem 4.3.5 (De Giorgi-Moser-Nash's theorem)**.** *Suppose that $a, b, c \in L^\infty(\Omega)$, a is uniformly elliptic, and L is as in* (4.1)*. If $u \in H^1(\Omega)$ satisfies $Lu = f$ in Ω in the weak sense for some $f \in L^q(\Omega)$ with $q > \frac{n}{2}$, then u is locally Hölder continuous, and for any open ω such that $\bar{\omega} \subset \Omega$ we have*

$$\|u\|_{C^{0,\alpha}(\omega)} \leq C(\|f\|_{L^q(\Omega)} + \|u\|_{H^1(\Omega)})$$

where the constant C depends only on $n, \Omega, \omega, a, b, c$. and the Hölder exponent α depends only on n, Ω, ω, a.

Let us remark that the fact that the coefficients a is discontinuous renders the method of freezing coefficients inapplicable. No matter how small the subdomain ω is, the coefficients a_{ij} can be as jumpy as one would like them to be and so the character of solutions to such operator is far different from that for operators with constant coefficients. In fact, if in the above theorem, if the coefficients a_{ij} are α-Hölder continuous and if $q > n$, it can be shown that the solution u will then have β-Hölder continuous derivatives for any $\beta \in (0, \min(\alpha, 1 - \frac{n}{q}))$.

To keep the discussion more transparent we will only consider the case that $b \equiv 0$, $c \equiv 0$ and $f \equiv 0$, Ω is the ball B_2 and ω is the ball B_1. We will be content with establishing only local L^∞ bound of solutions which are already known to be bounded, i.e. we are turning a qualitative property (boundedness) into a quantitative property (an actual bound for its L^∞-norm). Such estimates are referred to as *a priori* estimates. A careful adaptation of the argument will in fact remove the boundedness assumption, but we will not pursue here.

Theorem 4.3.6. *Suppose that $a \in L^\infty(B_2)$, a is uniformly elliptic, $b \equiv 0$, $c \equiv 0$ and L is as in* (4.1)*. If $u \in H^1(B_2) \cap L^\infty(B_2)$ satisfies $Lu = 0$ in B_2 in the weak sense, then*

$$\|u\|_{L^\infty(B_1)} \leq C\|u\|_{L^2(B_2)}$$

where the constant C depends only on n, a.

We will use the so-called Moser iteration method. When $u \in H^1(B_2) \cap L^\infty(B_2)$, the chain rule will give that $u^p \in H^1(B_2)$ for any $p > 1$. In particular, we can obtain estimates by using cut-off versions of powers of u as test functions, in a way similar to how we obtained energy estimates.

Proof. (*–Not for examination–*) To illustrate the main ideas while avoiding technicality, we assume an artificial condition that $u > 0$.

Let $\zeta \in C_c^\infty(B_2)$. Fix some $p \geq 1$ for the moment. Using $\zeta^2 u^p$ as test function (note that this makes sense as $u > 0$), we have

$$0 = B(u, \zeta u^p) = \int_{B_2} a_{ij}\partial_j u \partial_i(\zeta^2 u^p) = \int_{B_2} [p\zeta^2 u^{p-1} a_{ij}\partial_j u \partial_j u + 2\zeta u^p a_{ij}\partial_j u \partial_i \zeta].$$

Thus by using ellipticity on the second term and Cauchy-Schwarz' inequality on the second term, we have

$$\int_{B_2} p\zeta^2 u^{p-1}|Du|^2 \leq C\int_{B_2} u^{p+1}|D\zeta|^2.$$

This implies that

$$\int_{B_2} \zeta^2 |Du^{\frac{p+1}{2}}|^2 \leq Cp\int_{B_2} u^{p+1}|D\zeta|^2$$

and so

$$\int_{B_2} |D(\zeta u^{\frac{p+1}{2}})|^2 \leq \int_{B_2} 2[\zeta^2|Du^{\frac{p+1}{2}}|^2 + u^{p+1}|D\zeta|^2] \leq Cp\int_{B_2} u^{p+1}[\zeta^2 + |D\zeta|^2]$$

By the Friedrichs-type inequality (Theorem 3.2.3), we hence have with $\chi = \frac{n}{n-2}$ that

$$\Big[\int_{B_2} |\zeta u^{\frac{p+1}{2}}|^{2\chi}\Big]^{\frac{1}{\chi}} \leq Cp\int_{B_2} u^{p+1}[\zeta^2 + |D\zeta|^2]. \tag{4.10}$$

Now, if $1 \leq r_2 < r_1 \leq 2$, we can select $\zeta \in C_c^\infty(B_{r_1})$ such that $\zeta \equiv 1$ in B_{r_2} and $|D\zeta| \leq \frac{C}{r_1 - r_2}$, where C is a universal constant (the reason why this ζ exists is left as an exercise). Using this in (4.10) we obtain

$$\begin{aligned}\Big[\int_{B_{r_2}} |u|^{(p+1)\chi}\Big]^{\frac{1}{\chi}} &\leq \Big[\int_{B_2} |\zeta u^{\frac{p+1}{2}}|^{2\chi}\Big]^{\frac{1}{\chi}} \\ &\leq Cp\int_{B_2} |u|^{p+1}[\zeta^2 + |D\zeta|^2] \leq \frac{Cp}{(r_1 - r_2)^2}\int_{B_{r_1}} |u|^{p+1}.\end{aligned}$$

In other words, we have

$$\|u\|_{L^{(p+1)\chi}(B_{r_2})} \leq \Big[\frac{C(p+1)}{(r_1 - r_2)^2}\Big]^{\frac{1}{p+1}} \|u\|_{L^{p+1}(B_{r_1})}. \tag{4.11}$$

Roughly speaking, as we are shrinking the domain, we get better in integrability. An inequality of this kind is called a reversed Hölder's inequality.

One would like to somehow send $p \to \infty$ in (4.11) to obtain an L^∞ bound in the limit. As one does this, one would need to use a sequence of nested ball $B_{r_1} \supset B_{r_2} \supset \ldots \supset B_1$. A possible obstacle then occurs: on the right hand side of (4.11), the difference of the radii occurs on the denominator and this is goes to zero along the sequence of nested balls. The key point to observe here is that at the same time, this factor is raised to the $\frac{1}{p+1}$ power, and $\frac{1}{p+1}$ is going to zero.

Let us now detail the above scheme. We start with $r_1 = 2$, $r_2 = 1 + 2^{-1}$, $p_1 = 1$, $p_2 = 2\chi - 1$. Then (4.11) gives

$$\|u\|_{L^{2\chi}(B_{r_2})} \leq \left[\frac{C}{2^{-2}}\right]^{\frac{1}{2}} \|u\|_{L^2(B_{r_1})}.$$

Then we let $r_3 = 1 + 2^{-2}$, $p_3 = 2\chi^2 - 1$ so that

$$\|u\|_{L^{2\chi^2}(B_{r_3})} \leq \left[\frac{C}{2^{-4}}\right]^{\frac{1}{2\chi}} \|u\|_{L^{2\chi}(B_{r_2})}.$$

Proceeding in this way with $r_k = 1 + 2^{-k+1}$ and $p_k = 2\chi^{k-1} - 1$ we have

$$\|u\|_{L^{2\chi^{k-1}}(B_{r_k})} \leq \left[\frac{C}{2^{-2(k-1)}}\right]^{\frac{1}{2\chi^{k-2}}} \|u\|_{L^{2\chi^{k-2}}(B_{r_{k-1}})}.$$

Putting together these estimates we get

$$\begin{aligned}\|u\|_{L^{2\chi^{k-1}}(B_{r_k})} &\leq \prod_{j=2}^{k} \left[\frac{C}{2^{-2(j-1)}}\right]^{\frac{1}{2\chi^{j-2}}} \|u\|_{L^2(B_{r_1})} \\ &\leq C^{\frac{1}{2}\sum_{j=2}^{k}\chi^{-(j-2)}} 2^{\sum_{j=2}^{k}(j-1)\chi^{-(j-2)}} \|u\|_{L^2(B_{r_1})}.\end{aligned}$$

As the sums $\sum_{j\geq 2}\chi^{-(j-2)}$ and $\sum_{j\geq 2}(j-1)\chi^{-(j-2)}$ converge, we can now safely send $k \to \infty$ to obtain

$$\|u\|_{L^\infty(B_1)} \leq C\|u\|_{L^2(B_2)},$$

as desired. □