CHAPTER 3

Ellipticity: Laplace, Poisson and diffusion equations

1. Toolbox: Sobolev spaces

1.1. Definitions. Let us start, as a reference point, by defining the standard scale of C^k spaces (interpolating with Hölder regularity in between integer indices):

Definition 1.1. *We define C^θ for $\theta \in \mathbb{R}_+$ as a subspace $C^{[\theta]}$, with $[\theta]$ the integer part of θ, where the $[\theta]$-order derivatives are $(\theta - [\theta])$ Hölder continuous. In formula this means:*

$$\|u\|_{C^\theta} := \sum_{|\alpha| \le [\theta]} \|\partial^\alpha u\|_{L^\infty(\mathbb{R}^\ell)} + \sum_{|\alpha| = [\theta]} \sup_{x \neq y \in \mathbb{R}^\ell} \frac{|\partial^\alpha u(x) - \partial^\alpha u(y)|}{|x-y|^{\theta - [\theta]}}.$$

Comparatively, Sobolev spaces are tools that allow measuring regularity by means of *integrals* (as opposed to pointwise as for C^k spaces):

Definition 1.2. *We define the Sobolev space $W^{s,p}(\mathbb{R}^\ell)$ on $\mathbb{R}^\ell$, for $s \in \mathbb{N}$, as a subset of $L^p(\mathbb{R}^\ell)$ by building the completion of the vector space $C_c^\infty(\mathbb{R}^\ell)$ (infinitely differentiable with compact support) endowed with the norm*

$$\|g\|_{W^{s,p}(\mathbb{R}^\ell)} := \left(\sum_{|\alpha| \le s} \|\partial_x^\alpha g\|^2_{L^p(\mathbb{R}^\ell)} \right)^{\frac{1}{2}}.$$

It means:

$$W^{s,p}(\mathbb{R}^\ell) = \overline{C_c^\infty(\mathbb{R}^\ell)}^{\|\cdot\|_{W^{s,p}(\mathbb{R}^\ell)}} \subset L^p(\mathbb{R}^\ell).$$

We write $H^s(\mathbb{R}^\ell) = W^{s,2}(\mathbb{R}^\ell)$ in the case $p = 2$.

Remark 1.3. *In the case $p = 2$ we can give two other definitions:*

A second definition is: $g \in L^2(\mathbb{R}^\ell)$ belongs to $H^s(\mathbb{R}^\ell)$ iff there is a constant $C > 0$ so that

$$\forall \varphi \in C_c^\infty(\mathbb{R}^\ell), \ \forall |\alpha| \le s, \quad \left| \int_{\mathbb{R}^\ell} g(x) \partial_x^\alpha \varphi(x) \, \mathrm{d}x \right| \le C \|\varphi\|_{L^2(\mathbb{R}^\ell)}$$

and the smaller such constant is precisely the $H^s(\mathbb{R}^\ell)$ norm of g (this is the definition used by Leray).

A third definition can be obtained by Fourier calculus: $g \in H^s(\mathbb{R}^\ell) \subset L^2(\mathbb{R}^\ell)$ iff there is a constant $C > 0$ so that

$$\left(\int_{\mathbb{R}^\ell} |\hat{g}(\xi)|^2 (1 + |\xi|^2)^{\frac{s}{2}} \, \mathrm{d}\xi \right)^{\frac{1}{2}} \le C$$

where $\hat{g}$ is the Fourier transform of g, and the smaller such constant is precisely the $H^s(\mathbb{R}^\ell)$ norm of g. This definition allows to consider any $s \in \mathbb{R}_+$ (another way to define H^s for non-integer s would be to use the interpolation theory, see the mid-term presentations).

Exercise 24. *Check that all the three previous definitions are equivalent for $H^s(\mathbb{R}^\ell)$ and provide a Hilbert space, which is dense in $L^2(\mathbb{R}^\ell)$ (hint: use the approximation of the unit for the last point).*

Remark 1.4. *For $g \in W^{1,p}(\mathbb{R}^\ell)$ with $s \geq 1$, we define thus a **generalised derivative** $\nabla g \in L^p(\mathbb{R}^\ell)$ as the limit in $L^p(\mathbb{R}^\ell)$ of ∇g^n where $g^n \in C_c^\infty(\mathbb{R}^\ell)$ approximates g in $W^{1,p}(\mathbb{R}^\ell)$. In general this generalised derivative is **not** related to g by the standard differential calculus, that is in a pointwise. However it satisfies the integration by parts as follows:*

$$\forall \varphi \in C_c^\infty(\mathbb{R}^\ell), \quad \int_{\mathbb{R}^\ell} \nabla g \varphi \,\mathrm{d}x = - \int_{\mathbb{R}^\ell} g \nabla \varphi \,\mathrm{d}x.$$

We can also give a variant of the previous definition when only the highest-order derivatives are considered in the norm:

Definition 1.5. *We can now define the Sobolev* homogeneous *norm $\dot{H}^s(\mathbb{R}^\ell)$ for $s \in \mathbb{N}$ as the subspace of $L^2(\mathbb{R}^\ell)$ obtained by closing $C_c^\infty(\mathbb{R}^\ell)$ within $L^2(\mathbb{R}^\ell)$ for the semi-norm*

$$\|g\|_{\dot{H}^s(\mathbb{R}^\ell)} := \left(\sum_{|\alpha|=s} \|\partial_x^\alpha g\|^2_{L^2(\mathbb{R}^\ell)} \right)^{\frac{1}{2}}$$

or equivalently

$$\forall \varphi \in C_c^\infty(\mathbb{R}^\ell),\ \forall |\alpha| = s, \quad \left| \int_{\mathbb{R}^\ell} g(x) \partial_x^\alpha \varphi(x) \,\mathrm{d}x \right| \leq \|g\|_{\dot{H}^s(\mathbb{R}^\ell)} \|\varphi\|_{L^2(\mathbb{R}^\ell)}$$

or equivalently

$$\|g\|_{\dot{H}^s(\mathbb{R}^\ell)} := \left(\int_{\mathbb{R}^\ell} |\hat{g}(\xi)|^2 |\xi|^s \,\mathrm{d}\xi \right)^{\frac{1}{2}}.$$

Exercise 25. *In general this formula provides only a* semi-norm*, however since we restrict to functions in $L^2(\mathbb{R}^\ell)$, prove that it is a norm due to the integrability restriction.*

1.2. Sobolev inequalities. This fundamental tool relates the smoothness measured by integrals in the Sobolev spaces H^s to the usual C^k spaces where the smoothness is measured in pointwise form:

$$\forall s \neq (\ell/2)\mathbb{N},\ s \in \mathbb{N},\ s > \ell/2,\ \exists C > 0, \quad \|u\|_{C^{s-\ell/2}(\mathbb{R}^\ell)} \leq C \|u\|_{H^s(\mathbb{R}^\ell)}.$$

Let us go through the proof of this landmark result in analysis. We start with dimension $\ell = 1$ where the proof is quite simple:

Proposition 1.6. *We have $H^1(\mathbb{R}) \subset C^{1/2}(\mathbb{R})$ and there is $C > 0$ so that*

$$\forall u \in H^1(\mathbb{R}), \quad \|u\|_{C^{1/2}(\mathbb{R})} \leq C \|u\|_{H^1(\mathbb{R})}.$$

Proof of Proposition 1.6. Consider $u \in C_c^\infty(\mathbb{R})$ during the proof and finally argue by density. We have $(u|u|)' = 2|u|u'$ and thus

$$\begin{aligned} |u|(x)u(x) &= 2\int_{-\infty}^{x} |u|(y)u'(y)\,\mathrm{d}y \\ \Longrightarrow u(x)^2 &\lesssim \|u\|_{L^2(\mathbb{R})}\|u'\|_{L^2(\mathbb{R})} \\ \Longrightarrow \|u\|_{L^\infty(\mathbb{R})} &\lesssim \left(\|u\|_{L^2(\mathbb{R})}^2 + \|u'\|_{L^2(\mathbb{R})}^2\right)^{1/2} = \|u\|_{H^1(\mathbb{R})}. \end{aligned}$$

Moreover we can estimate variations as

$$\begin{aligned} u(x) - u(y) &= \int_x^y u'(z)\,\mathrm{d}z \\ \Longrightarrow |u(x)-u(y)| &\lesssim |x-y|^{1/2}\|u'\|_{L^2(\mathbb{R})} \\ \Longrightarrow \sup_{x\neq y} \frac{|u(x)-u(y)|}{|x-y|^{1/2}} &\lesssim \|u\|_{H^1(\mathbb{R})} \end{aligned}$$

which concludes the proof. □

In general dimension we shall prove the slightly weaker following result, which is sufficient for our study of ellipticity later.

Proposition 1.7. *For $k, s \in \mathbb{N}$, $s > k + \ell/2$, we have $C^k(\mathbb{R}^\ell) \subset H^s(\mathbb{R}^\ell)$, and there is a constant $C > 0$ so that*

$$\forall\, u \in H^s(\mathbb{R}^\ell), \quad \|u\|_{C^k(\mathbb{R}^\ell)} \leq C\|u\|_{H^s(\mathbb{R}^\ell)}.$$

Observe that the case $\ell = 1$ is proved by the previous proposition. In higher dimension, the proof is longer and will go through the so-called Sobolev-Gagliardo-Nirenberg inequality.

Proposition 1.8. *Assume $\ell > p$. We have $W^{1,p}(\mathbb{R}^\ell) \subset L^{p^*}(\mathbb{R})$ with $p^* := p\ell/(\ell-p)$ and there is $C > 0$ so that*

$$\forall\, u \in W^{1,p}(\mathbb{R}^\ell), \quad \|u\|_{L^{p^*}(\mathbb{R}^\ell)} \leq C\|u\|_{W^{1,p}(\mathbb{R}^\ell)}.$$

Remark 1.9. *Observe that it implies a control of all L^q norms for $q \in [p, p^*]$ by Hölder inequality.*

Proof of Proposition 1.8. We prove an intermediate result, the Sobolev-Gagliardo-Nirenberg inequality, that will allow us to use the one-dimensional argument in an "average way" on all variables:

Lemma 1.10. *Let $\ell \geq 2$ and $f_1, \dots, f_\ell \in L^{\ell-1}(\mathbb{R}^{\ell-1})$. For any $1 \leq i \leq \ell$ we denote $\tilde{x}_i = (x_1, \dots, x_{i-1}, x_i, \dots, x_\ell)$ (removing the i-th component), and $f(x) := f_1(\tilde{x}_1)\cdots f_\ell(\tilde{x}_\ell)$. Then $f \in L^1(\mathbb{R}^\ell)$ with*

$$\|f\|_{L^1(\mathbb{R}^\ell)} \leq \prod_{i=1}^{\ell} \|f_i\|_{L^{\ell-1}(\mathbb{R}^{\ell-1})}.$$

Proof of Lemma 1.10. The case $\ell = 2$ is clear: $f(x) = f_1(x_2) f_2(x_1)$ and

$$\int_{x_1,x_2} |f| = \left(\int_{x_2} |f_1| \right) \left(\int_{x_1} |f_2| \right).$$

The case $\ell = 3$ is obtained by applying three times the Cauchy-Schwarz inequality:

The general case is obtained by induction. Suppose the case $\ell \geq 2$ is true. Then write $f = f_{\ell+1}(\tilde{x}_{\ell+1}) F(x)$, $F(x) = f_1(\tilde{x}_1) \cdots f_\ell(\tilde{x}_\ell)$ and

$$\int_{x_1,\dots x_\ell} |f(\cdot, x_{\ell+1})| \leq \|f_{\ell+1}\|_{L^\ell(\mathbb{R}^\ell)} \|F(\cdot, x_{\ell+1})\|_{L^{\ell/(\ell-1}(\mathbb{R}^\ell)}.$$

We then apply the case ℓ to $f_1^{\ell/(\ell-1)}(\cdot, x_{\ell+1}) \cdots f_\ell^{\ell/(\ell-1)}(\cdot, x_{\ell+1})$ with $x_{\ell+1}$ fixed:

$$\begin{aligned}
\int_{x_1,\dots x_\ell} |f(\cdot, x_{\ell+1})| &\leq \|f_{\ell+1}\|_{L^\ell(\mathbb{R}^\ell)} \left(\prod_{i=1}^{\ell} \left\| f_i^{\ell/(\ell-1)}(\cdot, x_{\ell+1}) \right\|_{L^{\ell-1}(\mathbb{R}^{\ell-1})} \right)^{(\ell-1)/\ell} \\
&= \|f_{\ell+1}\|_{L^\ell(\mathbb{R}^\ell)} \left(\prod_{i=1}^{\ell} \|f_i(\cdot, x_{\ell+1})\|_{L^\ell(\mathbb{R}^{\ell-1})} \right).
\end{aligned}$$

We finally integate in $x_{\ell+1}$ to get

$$\begin{aligned}
\|f\|_{L^1(\mathbb{R}^{\ell+1})} &\leq \|f_{\ell+1}\|_{L^\ell(\mathbb{R}^\ell)} \int_{x_{\ell+1}} \left(\prod_{i=1}^{\ell} \|f_i(\cdot, x_{\ell+1})\|_{L^\ell(\mathbb{R}^{\ell-1})} \right) \mathrm{d}x_{\ell+1} \\
&\leq \|f_{\ell+1}\|_{L^\ell(\mathbb{R}^\ell)} \prod_{i=1}^{\ell} \left(\int_{x_{\ell+1}} \|f_i(\cdot, x_{\ell+1})\|^\ell_{L^\ell(\mathbb{R}^{\ell-1})} \mathrm{d}x_{\ell+1} \right)^{1/\ell} \\
&\leq \|f_{\ell+1}\|_{L^\ell(\mathbb{R}^\ell)} \prod_{i=1}^{\ell} \|f_i\|_{L^\ell(\mathbb{R}^\ell)}
\end{aligned}$$

which proves the case $\ell + 1$ and concludes the proof. □

We now go back to the proof of the proposition with the lemma at hand. We define $v := |u|^{t-1} u$ with $\partial v = t|u|^{t-1} \partial u$, for any partial derivative ∂. We compute on v for any $1 \leq i \leq \ell$:

$$\begin{aligned}
|v(x)| &\leq \left| \int_{-\infty}^{x_i} \frac{\partial v}{\partial x_i}(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_\ell) \, \mathrm{d}t \right| \\
&\leq \int_{-\infty}^{+\infty} \left| \frac{\partial v}{\partial x_i}(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_\ell) \right| \mathrm{d}t =: f_i(\tilde{x}_i).
\end{aligned}$$

We deduce that $|v|^{\ell/(\ell-1)} \leq \prod_{i=1}^{\ell} f_i^{1/(\ell-1)}$ and by using the lemma:

$$\|v\|_{L^{\ell/(\ell-1)}(\mathbb{R}^\ell)} \leq \left(\prod_{i=1}^{\ell} \|f_i\|^{1/(\ell-1)}_{L^1(\mathbb{R}^{\ell-1})} \right)^{\frac{\ell-1}{\ell}} \leq \prod_{i=1}^{\ell} \left\| \frac{\partial v}{\partial x_i} \right\|^{\frac{1}{\ell}}_{L^1(\mathbb{R}^\ell)}$$

which implies

$$\begin{aligned}
\|u\|^t_{L^{t\ell/(\ell-1)}(\mathbb{R}^\ell)} &\lesssim t\prod_{i=1}^{\ell}\left\||u|^{t-1}\frac{\partial u}{\partial x_i}\right\|^{1/\ell}_{L^1(\mathbb{R}^\ell)} \\
&\lesssim \prod_{i=1}^{\ell}\left(\|u\|^{t-1}_{L^{p'(t-1)}(\mathbb{R}^\ell)}\left\|\frac{\partial u}{\partial x_i}\right\|_{L^p(\mathbb{R}^\ell)}\right)^{1/\ell} \\
\|u\|^t_{L^{t\ell/(\ell-1)}(\mathbb{R}^\ell)} &\lesssim \|u\|^{t-1}_{L^{p'(t-1)}(\mathbb{R}^\ell)}\|\nabla u\|_{L^p(\mathbb{R}^\ell)}
\end{aligned}$$

where $p' := p/(p-1)$. We then choose (in a unique way) t so that the Lebesgue coefficients match: $t\ell/(\ell-1) = p'(t-1) = p^*$, which gives the result of the statement. □

Proposition 1.11. *Assume $\ell = p$. We have $W^{1,p}(\mathbb{R}^\ell) \subset L^q(\mathbb{R})$ for any $q \in [p,+\infty)$ and for any such q there is $C_q > 0$ so that*

$$\forall\, u \in W^{1,p}(\mathbb{R}^\ell), \quad \|u\|_{L^q(\mathbb{R}^\ell)} \le C_q\|u\|_{W^{1,p}(\mathbb{R}^\ell)}.$$

Proof of Proposition 1.11. We define $v := |u|^{t-1}u$ with $\partial v = t|u|^{t-1}\partial u$ for any partial derivative ∂. We perform the same calculation as above based on the Sobolev-Gagliardo-Nirenberg lemma to get:

$$\begin{aligned}
&\|u\|^t_{L^{t\ell/(\ell-1)}(\mathbb{R}^\ell)} \le t\|u\|^{t-1}_{L^{(t-1)\ell/(\ell-1)}(\mathbb{R}^\ell)}\|\nabla u\|_{L^\ell(\mathbb{R}^\ell)} \\
\Longrightarrow\ &\|u\|_{L^{t\ell/(\ell-1)}(\mathbb{R}^\ell)} \lesssim \|u\|_{L^{(t-1)\ell/(\ell-1)}(\mathbb{R}^\ell)} + \|\nabla u\|_{L^\ell(\mathbb{R}^\ell)}.
\end{aligned}$$

By interpolation it implies for any $t \ge \ell$:

$$\|u\|_{L^{t\ell/(\ell-1)}(\mathbb{R}^\ell)} \lesssim \|u\|_{L^\ell(\mathbb{R}^\ell)} + \|\nabla u\|_{L^\ell(\mathbb{R}^\ell)}$$

which concludes the proof. □

Proposition 1.12. *Assume $\ell < p$. We have $W^{1,p}(\mathbb{R}^\ell) \subset C^{1-\ell/p}(\mathbb{R})$ and there is $C > 0$ so that*

$$\forall\, u \in W^{1,p}(\mathbb{R}^\ell), \quad \|u\|_{C^{1-\ell/p}(\mathbb{R}^\ell)} \le C\|u\|_{W^{1,p}(\mathbb{R}^\ell)}.$$

Proof of Proposition 1.12. We first prove the Hölder regularity. Consider an open cube with size $r > 0$ containing, say, 0. Call $\bar u$ the average of u on that cube. Then

$$\begin{aligned}
|\bar u - u(0)| &\le \frac{1}{|Q|}\int_Q |u(x) - u(0)|\,\mathrm{d}x \\
&\le \frac{r}{|Q|}\int_Q\int_0^1\sum_{i=1}^{\ell}\left|\frac{\partial u}{\partial x_i}(tx)\right|\mathrm{d}t\,\mathrm{d}x \\
&\le \frac{r}{|Q|}\int_0^1 t^{-\ell}\left(\int_{tQ}\sum_{i=1}^{\ell}\left|\frac{\partial u}{\partial x_i}(y)\right|\mathrm{d}y\right)\mathrm{d}t \\
&\le \frac{r}{|Q|}\int_0^1 t^{-\ell}\left(\sum_{i=1}^{\ell}\left\|\frac{\partial u}{\partial x_i}(y)\right\|_{L^p(Q)}|tQ|^{1/p'}\right)\mathrm{d}t
\end{aligned}$$

$$\lesssim \|\nabla u\|_{L^p(Q)} r^{1-\ell+\ell/p'} \int_0^1 t^{-\ell+\ell/p'} \, \mathrm{d}t = \|\nabla u\|_{L^p(Q)} r^{1-\ell/p} \int_0^1 t^{-\ell/p} \, \mathrm{d}t$$
$$\lesssim \frac{r^{1-\ell/p}}{1-\ell/p} \|\nabla u\|_{L^p(Q)}.$$

Since the point zero was arbitrary we have

$$\forall\, x \in Q, \quad |\bar{u} - u(x)| \lesssim \frac{r^{1-\ell/p}}{1-\ell/p} \|\nabla u\|_{L^p(Q)}.$$

This proves the Hölder regularity with index $1 - \ell/p$ by triangular inequality. Finally the L^∞ control is obtained as follows: any point $x \in \mathbb{R}^\ell$ belongs to a cube Q as above and then

$$|u(x)| \lesssim |\bar{u}| + \frac{r^{1-\ell/p}}{1-\ell/p} \|\nabla u\|_{L^p(Q)} \lesssim \|u\|_{L^p(Q)} + \|\nabla u\|_{L^p(Q)}.$$

□

Proof of Proposition 1.7. We apply successively the previous propositions (see example classes for more discussion on the technical aspects). Observe that as long as $p < \ell$ we continue applying the first proposition, which results into the loss of one derivative and the Lebesgue exponent p increasing by the transformation $\varphi(p) = p\ell/(\ell - p) > p$. This transformation maps $[\ell/(k+1), \ell/k)$ to $[\ell/k, \ell/(k-1))$ for $k \geq 2$ and $[\ell/2, \ell)$ to $[\ell, +\infty)$. Therefore starting from $p = 2$, the number of necessary iteration to reach C^0 is s so that $2 > \ell/s$ (in the borderline case one must use the second proposition once). Finally we conclude by applying the third proposition once $p > \ell$. □

1.3. Sobolev spaces on an open set.

Definition 1.13. *We consider $\mathcal{U}$ a bounded and open set of $\mathbb{R}^\ell$ with smooth boundary $\partial\mathcal{U}$.*

We define the Sobolev space $W^{s,p}(\mathcal{U})$ on $\mathcal{U}$, for $s \in \mathbb{N}$ and $p \in [1+\infty]$, as a subset of $L^p(\mathcal{U})$ by building the completion of the vector space $C^\infty(\mathbb{R}^\ell)$ (infinitely differentiable) endowed with the norm

$$\|g\|_{W^{s,p}(\mathcal{U})} := \left(\sum_{|\alpha| \leq s} \|\partial_x^\alpha g\|^2_{L^p(\mathcal{U})} \right)^{\frac{1}{2}}.$$

It means:

$$W^{s,p}(\mathcal{U}) = \overline{C^\infty(\mathcal{U})}^{\|\cdot\|_{W^{s,p}(\mathbb{R}^\ell)}} \subset L^p(\mathcal{U}).$$

We write $H^s(\mathcal{U}) = W^{s,2}(\mathcal{U})$ in the case $p = 2$.

We also define the Sobolev space $W^{s,p}_0(\mathcal{U})$ on $\mathcal{U}$, for $s \in \mathbb{N}$ and $p \in [1+\infty]$, as a subset of $L^p(\mathcal{U})$ by building the completion of the vector space $C^\infty_c(\mathbb{R}^\ell)$ (infinitely differentiable with compact support included in $\mathcal{U}$) endowed with the same norm $W^{s,p}(\mathcal{U})$. It means:

$$W^{s,p}_0(\mathcal{U}) = \overline{C^\infty_c(\mathcal{U})}^{\|\cdot\|_{W^{s,p}(\mathbb{R}^\ell)}} \subset L^p(\mathcal{U}).$$

We write $H^s_0(\mathcal{U}) = W^{s,2}_0(\mathcal{U})$ in the case $p = 2$.

EXERCISE 26. *(1) Show that $H^1_0(\mathcal{U})$ is a Hilbert space, included in $H^1(\mathcal{U})$.*

(2) Show that on this subspace the homogeneous semi-norm $\dot{H}^1$ is a norm thanks to the Poincaré inequality (see later in Subsection 1.4).

(3) Show also that for any $u, v \in H^1_0(\mathcal{U})$ and any first-order partial derivative ∂ we have

$$\int_{\mathcal{U}} (\partial u) v \, \mathrm{d}x = - \int_{\mathcal{U}} u (\partial v) \, \mathrm{d}x.$$

(Actually check that it is enough that only one of the two functions u and v is in $H^1_0(\mathcal{U})$, while the other one can be merely in $H^1(\mathcal{U})$.)
Hint: Use an approximation argument by $C^\infty_c(\mathcal{U})$ functions.

The Sobolev inequalities extend to the case of a smooth bounded domain (see the example classes for a proof):

$$\forall\, s \neq (\ell/2)\mathbb{N}, \ s \in \mathbb{N}, \ s > \ell/2, \ \exists C > 0, \quad \|u\|_{C^{s-\ell/2}(\mathcal{U})} \leq C \|u\|_{H^s(\mathcal{U})}.$$

Important warning. In the Borel-Lebesgue integration theory, functions in L^2 are only defined up to redefinition on a set of measure 0. Thus, the restriction of an L^2 function to a point or any hypersurface with non-zero codimension is meaningless. When we now consider solutions in spaces of generalised functions, this is an important thing not to be forgotten: the space should have enough regularity for the boundary (and/or initial) conditions to make sense.

On the one hand, we have seen with the *Sobolev inequalities* that the H^s regularity implies the $C^{s-\ell/2-0}$ regularity. Obviously our boundary conditions make sense for continuous functions, therefore the regularity H^s with $s > \ell/2$ would clearly guarantee that these boundary conditions make sense. However in dimension $\ell \geq 2$ this seems to prevent using H^1...

On the other hand, another family of inequality, the so-called *trace inequalities*, allows one to restrict H^s functions to $H^{s-1/2}$ functions on a smooth codimension-1 hypersurface. We can thus characterize $H^1_0(\mathcal{U})$ as the subset of $H^1(\mathcal{U})$ consisting of functions such that, extending to an $H^1(\mathbb{R}^\ell)$, which is always possible, the trace on $\partial\mathcal{U}$ vanishes as an L^2 function. More generally trace inequalities allow one to restrict H^s functions to $H^{s-d/2}$ functions on a codimension-d hypersurface. When $d = \ell$ we recover the numerology of Sobolev inequalities below: being continuous is similar to the fact that restriction to points (full codimension subset) is well-defined.

1.4. Poincaré inequality.

THEOREM 1.14 (Poincaré's inequality with Dirichlet conditions). *Let $\mathcal{U} \subset \mathbb{R}^\ell$ be a open bounded set such that $\partial\mathcal{U}$ is smooth. Then there exists $C_{\mathcal{U}} > 0$ (only depending on $\mathcal{U}$) such that the following holds. Let $u \in C^1(\overline{\mathcal{U}})$ such that $u = 0$ on $\partial\mathcal{U}$, then*

$$\left(\int_{\mathcal{U}} u(x)^2 \, \mathrm{d}x \right)^{1/2} \leq C_{\mathcal{U}} \left(\int_{\mathcal{U}} |\nabla u(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

PROOF OF THEOREM 1.14. Consider any point $x = (x_1, \dots, x_\ell)$ in $\mathcal{U}$, and let $\bar{x}_1$ be such that $\bar{x} = (\bar{x}_1, x_2, \dots, x_\ell) \in \partial\mathcal{U}$ (see Figure 1.4).

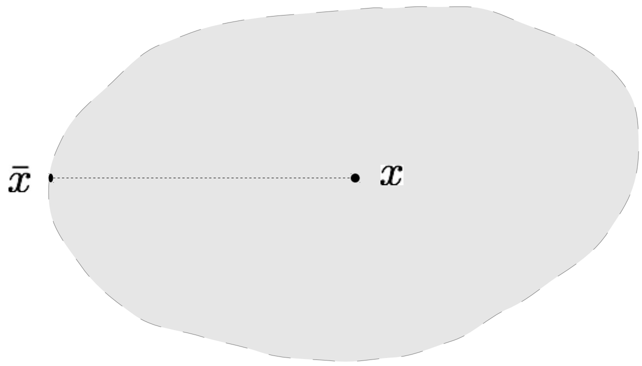

FIGURE 1.1. Illustration of the proof.

Then we have

$$u(x) = \int_{\bar{x}_1}^{x_1} \partial_{x_1} u(y, x_2, \ldots, x_\ell) \, \mathrm{d}y$$

and we estimate

$$\int_{\mathcal{U}} u(x)^2 \, \mathrm{d}x = \int_{\mathcal{U}} \left(\int_{\bar{x}_1}^{x_1} \partial_{x_1} u(y, x_2, \ldots, x_\ell) \, \mathrm{d}y \right)^2 \, \mathrm{d}x$$
$$\leq D_1(\mathcal{U})^2 \int_{\mathcal{U}} |\partial_{x_1} u(x)|^2 \, \mathrm{d}x \leq D_1(\mathcal{U})^2 \int_{\mathcal{U}} |\nabla u(x)|^2 \, \mathrm{d}x$$

which concludes the proof. We have denoted here by $D_1(\mathcal{U})$ the length of the greatest interval along the first axis included in $\mathcal{U}$. □

EXERCISE 27. *Show that this inequality remains true under the more general condition that $\mathcal{U}$ is bounded along one of its direction only.*

REMARK 1.15. *Note that there are also* Poincaré's inequalities *in the whole space, provided the reference measure γ has some strong decay (essentially at least exponential) and regularity properties. The most famous example is the gaussian case $\gamma(x) = e^{-|x|^2}$:*

$$\left(\int_{\mathbb{R}^\ell} \left| u(x) - \int_{\mathbb{R}^\ell} u(y)\gamma(y) \, \mathrm{d}y \right|^2 \gamma(x) \, \mathrm{d}x \right)^{1/2} \leq C_\gamma \left(\int_{\mathbb{R}^\ell} |\nabla u(x)|^2 \gamma(x) \, \mathrm{d}x \right)^{1/2}.$$

The proof is more involved than the one above, see for instance the 2011 exam paper of the course on kinetic theory for intermediate steps.

2. What is ellipticity?

2.1. The notion of ellipticity. We say that a linear differential operator P of order k defined on an open set $\mathcal{U}$ is *elliptic* at $x \in \mathcal{U}$ if $\sigma_d(x,\xi) \neq 0$ for all $\xi \in \mathbb{R}^\ell \setminus \{0\}$. We say that it is *elliptic* on $\mathcal{U}$ if it is elliptic for all x in $\mathcal{U}$. Equivalently, a linear differential operator P is elliptic if all hypersurfaces are non-characteristic.

EXAMPLE 2.1. *For a first order operator written as*

$$P = \sum_{i=1}^{\ell} b_i(x)\partial_i + b_0(x)$$

the definition of ellipticity can never be satisfied as soon as $\ell \geq 2$ since $\sigma_p(x,\xi)$ cancels on $\mathbf{b}(x)^\perp$ with $\mathbf{b}(x) = (b_1(x), \dots, b_\ell(x))$. But $\ell \geq 2$ is required for a PDE, unless we consider an ODE.

EXAMPLE 2.2. *Consider now a second order linear operator written as*

$$P = -\sum_{i,j=1}^{\ell} a_{ij}(x)\partial^2_{ij} + \textit{ lower order terms...}$$

A first remark is that since $\partial^2_{ij} = \partial^2_{ji}$ for C^2 functions we can assume w.l.o.g. that the matrix $A := (a_{ij})_{ij}$ is symmetric (replacing it by $\bar{a}_{ij} = (a_{ij} + a_{ji})/2$ if necessary). Then the ellipticity means that

$$\forall\, \xi \in \mathbb{R}^\ell, \qquad \sum_{i,j=1}^{\ell} a_{ij}(x)\xi_i\xi_j = \xi^T A(x)\xi \neq 0$$

Hence by continuity and connectedness the sign is constant on $\mathbb{R}^\ell \setminus \{0\}$, and by convention (changing A to $-A$ if necessary), we are thus reduced to the case where $A(x)$ is a positive definite matrix, i.e. all its eigenvalues are positive. The (opposite of the) Laplacian operator $-\Delta$ is thus the most simple example of elliptic operator.

The most characteristic feature of ellipticity is so-called *elliptic regularity*. Let us first do some heuristics. The motivation for this can be found precisely in the above properties of the symbol. In the case of a constant coefficient differential operator of the form

$$P = \sum_{|\alpha|=k} a_\alpha \partial^\alpha_x,$$

if $\sigma_P(x,\xi) = \sigma_P(\xi) = 0$, then for all $\lambda \in \mathbb{R}$ we deduce that

$$u_\lambda(x) = e^{i\lambda x\cdot\xi}$$

is a solution to $Pu_\lambda(x) = 0$. In particular, as $\lambda \to \infty$, one can construct more and more oscillatory solutions. In the general, non constant-coefficient case, if $\sigma_p(x_0,\xi) = 0$ at some point x_0, then $u_\lambda(x)$ for large λ can still be viewed as an approximate highly oscillatory solution near x_0. Thus, operators P which fail to be elliptic have highly oscillatory solutions (or approximate solutions) of $Pu = 0$, and this intuition can be "localised" in a region. It turns out that conversely, the condition of ellipticity "prohibits" highly oscillatory behaviour for solutions of $Pu = f$ if oscillation are not present on the right hand side. This is at the heart of elliptic regularity.

2.2. Elliptic regularity in the whole Euclidean space. One can distinguish between qualitative and quantitative versions of the *elliptic regularity principle*. A qualitative version is the statement that if f is C^∞ in a neighborhood of a point $x \in \mathcal{U}$ in an open set $\mathcal{U}$, and u satisfies $Pu = f$, then u is C^∞ in a neighbourhood of x as well. Already, we should have been a bit careful when we said "u satisfies". What kind of function are we assuming u to be a priori? To keep this discussion elementary, let us say that u is a *classical solution*, that is to say, it has at least the order of differentiability appearing in the equation, i.e. it is C^2, and let us first consider $\mathcal{U} = \mathbb{R}^\ell$. To summarise:

Proposition 2.3 (Qualitative statement of elliptic regularity). *Let $u : \mathbb{R}^\ell \to \mathbb{R}$ be C^2 with $\Delta u = f$, where f is C^∞ on $\mathbb{R}^\ell$. Then u is C^∞ on $\mathbb{R}^\ell$.*

The proof of this proposition makes use of new notions, which are different from the previous chapters, and which can be organised along the following method:

(1) The use of *Sobolev spaces* to measure regularity by means of integral controls, and *Sobolev inequalities* in order to recover C^k regularity.
(2) The proof of a (fundamental) *a priori estimate*, taking advantage of the positivity of the principal in order to show that a given number of derivatives on f controls a higher number of derivatives on u.
(3) The justification of the a priori estimate by a *regularization argument* (painful technically but necessary to the rigor of the proof).

Step 1 was discussed in the previous section, let us go through steps 2 and 3.

2.2.1. *The key a priori esimate.* The estimate is at the core the elliptic regularity principle. Assume that $\Delta u = f$ is satisfied in $\mathbb{R}^\ell$, and that $u, f \in C^\infty_c(\mathbb{R}^\ell)$. This means that we argue *a priori* by assuming all the necessary regularity and decay at infinity, i.e. decay in both real and Fourier variables.

We now claim that

$$\forall\, s \in \mathbb{N}, \quad \|u\|_{\dot H^{s+2}(\mathbb{R}^\ell)} = \|f\|_{\dot H^s(\mathbb{R}^\ell)}. \tag{2.1}$$

The proof is elementary but fundamental; it is a simple example of "energy method".

We square and integrate the equation:

$$\int_{\mathbb{R}^\ell} |f|^2 \,\mathrm{d}x = \int_{\mathbb{R}^\ell} (\Delta u)^2 \,\mathrm{d}x$$

and we perform integration by parts on the RHS:

$$\int_{\mathbb{R}^\ell} (\Delta u)^2 \,\mathrm{d}x =_{def} \int_{\mathbb{R}^\ell} \sum_{i,j=1}^{\ell} \partial^2_{ii} u \partial^2_{jj} u \,\mathrm{d}x =_{IBP} \int_{\mathbb{R}^\ell} \sum_{i,j=1}^{\ell} \partial^2_{ij} u \partial^2_{ij} u \,\mathrm{d}x = \|u\|_{\dot H^2(\mathbb{R}^\ell)}$$

which proves the claim for $s = 0$. Now for a general $s \in \mathbb{N}$ observe that if we differentiate the equation by ∂^α_x for some $|\alpha| \le s$ we have

$$\Delta(\partial^\alpha_x u) = \partial^\alpha_x f$$

which means that the function $\partial^\alpha_x u$ satisfies the same elliptic equation as u. We are using a fundamental fact here:

$$[\Delta, \partial^\alpha_x] := \Delta \partial^\alpha_x - \partial^\alpha_x \Delta = 0$$

i.e. the commutation between the operator defining the equation and the operator according to which we want to estimate the regularity. We hence readily deduce by performing the same energy estimate as before on $\partial^\alpha u$ that

$$\|\partial_x^\alpha u\|_{\dot{H}^2(\mathbb{R}^\ell)} = \|\partial_x^\alpha f\|_{L^2(\mathbb{R}^\ell)}$$

and we conclude the proof of the claim for $s \in \mathbb{N}$ by summing over all $|\alpha| = s$.

2.2.2. *Justification.* As a third and last step, we can now complete the proof of the elliptic regularity by *justifying and making rigorous* the latter a priori estimate. In deriving (2.1), we have used underlying differentiability assumptions, i.e. a qualitative fact. But in fact, these assumptions can be proven from the "quantitative" statement that appears to rely on it. In other words, we shall see a first instance of the important fact that as soon as objects exists in one side of an a priori identity, they will exist in the other side. Roughly speaking, an a priori control of a quantity by a finite estimate will imply that the quantity indeed exists, i.e. *"existence follows from the estimate"*.

Let us illustrate this idea. For instance, suppose we know only a priori that u is a classical solution, i.e. C^2, whereas f is assumed to be C^∞. Consider an *approximation of the unit*, i.e. consider $\chi \in C_c^\infty(\mathbb{R}^\ell)$, $0 \le \chi \le 1$ with compact support in $B(0,1)$ and so that $\chi = 1$ on $B(0,1/4)$ and $\int \chi = 1$, and then scale it as $\chi_\epsilon(x) := \epsilon^{-\ell}\chi(\epsilon^{-1}x)$ for $\epsilon > 0$ (meant to be small), and as $\chi_R = \chi(Rx)$, $R > 0$ (meant to be large). Then define

$$\text{(localisation)} \qquad \tilde{u} := u\chi_R, \quad \tilde{f} := \Delta\tilde{u}$$

and

$$\text{(regularisation)} \qquad u_\epsilon := \tilde{u} \star \chi_\epsilon, \quad f_\epsilon = \tilde{f} \star \chi_\epsilon$$

where $\star$ denotes convolution, which satisfy

$$\forall\, \epsilon > 0, \quad \Delta u_\epsilon = f_\epsilon.$$

Observe that

$$\tilde{f} = u(\Delta\chi_R) + (\Delta u)\chi_R + 2\nabla u \cdot \nabla\chi_R = u(\Delta\chi_R) + f\chi_R + 2\nabla u \cdot \nabla\chi_R$$

is H^{s-1} (with compact support) as soon as $u \in H^s$ in $B(0, R+1)$.

Observe that u_ϵ has support in $B(0, R+\epsilon)$ and is C^∞ thanks to the convolution, and $f_\epsilon \in C_c^\infty$ clearly as well. Observe also that $\tilde{u} \in L^2(\mathbb{R}^\ell)$ thanks to the compact support and the L^∞. We now apply the previous a priori estimate for $s \in \mathbb{N}$ on differences of two mollified solutions:

$$\|u_{\epsilon_1} - u_{\epsilon_2}\|_{\dot{H}^{s+2}(\mathbb{R}^\ell)} \le \|f_{\epsilon_1} - f_{\epsilon_2}\|_{\dot{H}^s(\mathbb{R}^\ell)}.$$

By induction assume that $\tilde{u} \in H^{s+1}$, then $\tilde{f} \in H^s$. Then let us prove that the RHS is converging to zero. It is enough to prove

$$\lim_{\epsilon\to 0} \left\| f_\epsilon - \tilde{f} \right\|_{\dot{H}^s(\mathbb{R}^\ell)} = 0$$

and by taking ∂_x^α, $|\alpha| = s$, it is enough to prove

$$\lim_{\epsilon\to 0} \|g_\epsilon - g\|_{L^2(\mathbb{R}^\ell)} = 0$$

for a general $g \in L^2(\mathbb{R}^\ell)$. We then write

$$g_\epsilon(x) - g(x) = \int_{\mathbb{R}^\ell} g(x-y)\chi_\epsilon(y)\,\mathrm{d}y - g(x) = \int_{\mathbb{R}^\ell} (g(x-y)-g(x))\chi_\epsilon(y)\,\mathrm{d}y$$

where we have used $\int \chi_\epsilon = 1$. We then take the L^2 norm:

$$\|g_\epsilon - g\|_{L^2(\mathbb{R}^\ell)} \lesssim \sup_{|y|\le\epsilon} \|g(\cdot - y) - g\|_{L^2(\mathbb{R}^\ell)} \left(\int_{\mathbb{R}^\ell} \chi_\epsilon(y)\,\mathrm{d}y\right) = \sup_{|y|\le\epsilon} \|g(\cdot - y) - g\|_{L^2(\mathbb{R}^\ell)}$$

and we finally use the fundamental fact of the theory of Lebesgue integration (on which all mollification arguments rely on), which is the *continuity* of the translation operator in all L^p spaces, $p \in [1,+\infty)$:

$$\sup_{|y|\le r\epsilon} \|g(\cdot - y) - g\|_{L^2(\mathbb{R}^\ell)} \xrightarrow{\epsilon\to 0} 0,$$

which concludes the proof.

We now go back to our original problem. As a consequence of the fact that the RHS goes to zero, we deduce that the sequence u_ϵ has the Cauchy property in $\dot{H}^{s+2}(\mathbb{R}^\ell)$, and since $\tilde{u} \in L^2(\mathbb{R}^\ell)$, it is also converging in $L^2(\mathbb{R}^\ell)$ and we deduce that it has the Cauchy property in $H^{s+2}(\mathbb{R}^\ell)$. Since we already know that $u_\epsilon \to \tilde{u}$ in $L^2(\mathbb{R}^\ell)$, the uniqueness of the limit almost everywhere shows that $\tilde{u} \in H^{s+2}$. We deduce therefore that if $u \in H^s$ in $B(0,R+1)$ then $u\chi_R \in H^{s+1}$. Since this is true for any $R > 0$ we deduce that by induction on s that $\tilde{u} \in \cap_{s\ge 0} H^s$ on any ball $B(0,R)$, and finally by Sobolev inequalities, we get that $u \in C^\infty$ on any ball, and thus $u \in C^\infty(\mathbb{R}^\ell)$.

Remark 2.4. *As one can see, the full argument of justification of the a priori estimate is (1) long (and in particular significantly longer than the proof of the a priori estimate itself!), (2) technically tedious. It is rarely performed in full details in research papers and often postponed when working "heuristically" or "intuitively" on a problem, however it is important to perfectly understand and check these justification arguments, in order to provide fully rigorous proofs.*

2.3. Localisation of energy estimates and elliptic regularity in open sets. We shall now present a very simple by useful refinement of the previous argument: the *localisation* of the a priori estimate in order to obtain the regularity in any neighbourhood.

Proposition 2.5 (Qualitative local statement of elliptic regularity). *Let $\mathcal{U} \subset \mathbb{R}^\ell$ be a bounded smooth open set, and $u : \mathcal{U} \to \mathbb{R}$ be C^2 with $\Delta u = f$ in $\mathcal{U}$, where f is C^∞ on $\mathcal{U}$. Then u is C^∞ on $\mathcal{U}$.*

Remark 2.6. *This statement stresses the fact, in fact already present in the previous proof, that the elliptic regularity is a* local *phenomenon, in the sense that it is a consequence of the PDE in any neighborhood, independently of what happens outside of this neighborhood.*

Proof of Proposition 2.5. Consider a base point $x_0 \in \mathcal{U}$ and η so that $B(x_0, 4\eta) \subset \mathcal{U}$. Then define

$$\tilde{u}(x) = u(x)\chi_\eta(x - x_0)$$

and

$$\tilde{f} = \Delta\tilde{u}, \quad u_\epsilon = \tilde{u} \star \chi_\epsilon,\ \epsilon \in (0,\eta), \quad f_\epsilon = \tilde{f} \star \chi_\epsilon,\ \epsilon \in (0,\eta).$$

We have that $\tilde{f}$ satisfies the equation

$$\begin{aligned}\tilde{f}(x) &= \Delta(u(x)\chi_\eta(x-x_0)) \\ &= \Delta u(x)\chi_\eta(x-x_0) + u(x)\Delta\chi_\eta(x-x_0) + 2\nabla u(x)\cdot\nabla\chi_\eta(x-x_0)\end{aligned}$$

and therefore $\tilde{f} = f$ on $B(x_0, \eta/4)$ since $\chi = 1$ on $B(0, 1/4)$. Observe that if $u \in H^s$ in $B(x_0, 4\eta)$ then $\tilde{f} \in H^{s+1}$. Moreover $\tilde{f}$ and $\tilde{u}$ have support in $B(x_0, \eta) \subset \mathcal{U}$. Then we can perform the same argument as above and show that $\tilde{u} \in H^{s+1}$. Since the point x_0 was arbitrary, and η can vary (with the inclusion condition), we can induct on s to prove that $u \in C^\infty(\mathcal{U})$. □

2.4. Ill-posedness of the Cauchy problem. Let us consider the issue of solving the Cauchy problem for the Poisson equation, with the Cauchy surface $\Gamma = \{x_1 = 0\}$, i.e. taking the first variable as time. The Cauchy problem then writes

$$\begin{cases} \Delta u = f, \quad x = (x_1, \dots, x_\ell) \in \mathbb{R}^\ell \\ u(0, x_2, \dots, x_\ell) = u_0(x_2, \dots, x_\ell), \\ \partial_{x_1} u(0, x_2, \dots, x_\ell) = u_1(x_2, \dots, x_\ell) \end{cases} \tag{2.2}$$

where u_0 and u_1 are the boundary data. We know from the Cauchy-Kovalevskaya Theorem that if the data u_0 and u_1 are analytic near the origin, there exists a unique analytic solution u of the Cauchy problem (2.2) in a neighborhood of the origin. On the other hand, if we replace the analyticity assumption on u_0, u_1, with the assumption that $u_0 \in C^k$, $u_1 \in C^{k-1}$, but $u_0 \notin C^{k+1}$, then elliptic regularity implies that there cannot exist a C^2 solution u satisfying (2.2). More precisely:

PROPOSITION 2.7. *Assume that $u_1 \in C^{k-1}$ and u_0 is C^k but not C^{k+1} in a neighbourhood of 0, for some $k \geq 2$, and f is a smooth (C^∞) function in this neighbourhood. Then for no neighbourhood of 0 does there exist a C^2 solution u of (2.2).*

PROOF OF PROPOSITION 2.7. We apply the proposition 2.5: if u is a C^2 solution to (2.2) in a neighborhood $\mathcal{U}$ of the origin, then $u \in C^\infty(\mathcal{U})$, which contradicts $u_0 \notin C^{k+1}$. □

REMARK 2.8. *As we shall see, this behaviour is very different from that of the wave equation. The above non-existence result can be thought of as a qualitative statement of ill-posedness, that is to say, saying that there does not exist a solution of a suitable Cauchy problem.*

REMARK 2.9. *Taken alone, it could be (mis)interpreted as merely saying that one should never consider non-analytic functions. In fact this is a hint that the problem is overdetermined, we shall later in the chapter* remove *one of the boundary condition and consider the Dirichlet problem.*

It turns out that like the statement of elliptic regularity, this qualitative statement can be related to a quantitative statement, namely that even if the solution exists, the C^k_{loc} norm of the solution can grow as fast as wanted, and the same holds no matter what other way one tries to measure regularity, for instance Sobolev spaces H^s, etc.

PROPOSITION 2.10. *Given any constant $B > 0$ and any $k \geq l \geq 0$, there exists an analytic solution u to (2.2) with $f = 0$ such that*

$$\|u(1, \cdot)\|_{C^l} \geq B\left(\|u(0, \cdot)\|_{C^k} + \|\partial_{x_1} u(0, \cdot)\|_{C^k}\right)$$

and similarly when C^k, C^l are replaced by $H^k(B_R)$, $H^l(B_R)$ for some given ball $B(0, R)$ with any $R > 0$.

PROOF OF PROPOSITION 2.10. The proof can be performed going back to the example of Hadamard we already discussed (dropping all variables $x_3, \ldots, x_\ell$):

$$\partial^2_{x_1} u + \partial^2_{x_2} u = 0, \quad u(0, x_2) = 0, \quad \partial_{x_1} u(0, x_2) = a_\omega \sin(\omega x_2)$$

for some parameter $\omega > 0$, whose solution is explicitely given by

$$u(x_1, x_2) = \frac{a_\omega}{\omega} \sinh(\omega x_1) \cos(\omega x_2).$$

Then if we choose $a_\omega = \exp(-\sqrt{\omega})$ and $\omega \to \infty$, we see that the initial data at $x_1 = 0$ goes to zero in any C^k or H^k norm, whereas the solution at $x_1 = 1$ goes to infinity in any C^ℓ or H^ℓ norm. $\square$

EXERCISE 28. *Explain how to extend this statement with an* analytic *right hand side f in the Poisson equation $\Delta u = f$.*

COROLLARY 2.11. *As a consequence, if we define $\mathcal{D}$ the set of analytic (u_0, u_1) for which $u(1, \cdot)$ is defined, then for all $k \geq l \geq 0$, the map $(u_0, u_1) \mapsto u(1, \cdot)$ is not continuous from $\mathcal{D} \cap C^k$ to C^l (or from $\mathcal{D} \cap H^k(B_R)$ to $H^l(B_R)$, for any given $R > 0$).*

It was Hadamard who first really understood this and formulated the notion of well-posedness, which states not only that one can solve the initial value problem in suitable family of function classes, but also, that the relevant map be continuous. Note that allowing $k > l$ gives a little bit of room for well-posedness type statements. For the wave equation, we shall see that in Sobolev spaces we have continuity with $k = l$, whereas in C^k spaces we must indeed take $k > l$.

One should already view this as a small victory for mathematical analysis in illuminating a fundamental physical principle. This difference between elliptic (parabolic equations too will fit this as we shall see) and hyperbolic equations is a deep physical fact with many implications to the notion of time, the admissibility of theories, etc. On the other hand, it is completely swept under the rug when looking only at analytic solutions.

3. Toolbox: Hilbert space analysis

3.1. Hilbert spaces. A (real or complexed) normed vector space $\mathcal{H}$ is a Hilbert space if:

- The norm satisfies the parallelogram identity

$$\forall\, x, y \in \mathcal{H}, \quad \frac{\|x + y\|^2 + \|x - y\|^2}{2} = \|x\|^2 + \|y\|^2.$$

Exercise 29. *Show that this identity is equivalent to the existence of a scalar product $\langle \cdot, \cdot \rangle$ compatible with the norm. A vector space $\mathcal{H}$ having this property but not being complete is sometimes called a* pre-Hilbert space.

Hint: Use the so-called polarization identity *which is for real Hilbert space*

$$\langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$$

and for complex Hilbert spaces

$$\langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right).$$

- The space should be *complete* when endowed with the distance given by the norm.
- (This last property may or may not be included in the standard definition) The space should be *separable*, i.e. there should be a countable subset $S \subset \mathcal{H}$ which is dense in $\mathcal{H}$.

Assuming these three properties, one can speak of *the* (complex or real) Hilbert space, as all such spaces possess a *Hilbertian base*, i.e. a countable orthonormal family $(e_n)_{n\geq 0}$ so that $\mathrm{Span}\{e_n, n \geq 0\}$ is dense in $\mathcal{H}$, and are therefore isomorphic to $\ell^2(\mathbb{R})$ (or $\ell^2(\mathbb{C})$).

Exercise 30. *Prove the last claim.*

Exercise 31. *Show with the following example how to construct a* non-separable *Hilbert space: consider $\mathcal{H}$ the subset of $L^2_{loc}(\mathbb{R})$ functions (square-integrable on any compact subset) with finite norm for the scalar product*

$$\langle f, g \rangle := \lim_{R\to\infty} \frac{1}{R}\int_{-R}^{R} f(x)g(x)\,\mathrm{d}x.$$

Then $\mathcal{H}$ has a Hilbert space structure, but one can check that for any $a \in \mathbb{R}$, the function $S_a : x \mapsto \sin(ax)$ belongs to $\mathcal{H}$ (with norm 1) and that

$$\forall\, a, a' \in \mathbb{R}, \quad \langle S_a, S_{a'} \rangle = \delta_{a=a'}.$$

This produces a non-countable orthonormal family.

3.2. The Riesz (Hilbert space) representation Theorem. [1]

The following result generalises the intuition of Euclidean spaces:

Proposition 3.1 (Projection on a closed convex set). *Consider $(H, \langle \cdot, \cdot \rangle)$ a Hilbert space, $C \subset \mathcal{H}$ a closed non-empty convex subset of $\mathcal{H}$, and $x \in \mathcal{H}$. Then there exists a* unique *$P_C(x) \in C$ which realizes the following minimization problem*

$$\|P_c(x) - x\| = \min_{y\in C} \|y - x\|.$$

Moreover the application $P_C : \mathcal{H} \to C$ is 1-Lipschitz.

[1] Warning: there are two main theorems usually called "Riesz representation Theorem" that should be not confused, the other one being that of representation of positive linear functionals on $C_c(X)$ (X locally compact Hausdorff – i.e. separated– space) by regular Borel measures.

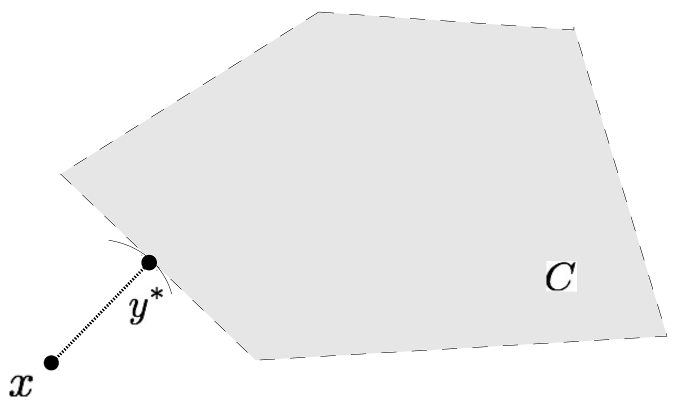

FIGURE 3.1. Geometric summary of the proof.

PROOF OF PROPOSITION 3.1. Set $x = 0$ w.l.o.g. If $0 \in C$ then we are done. If not we denote the minimum value by I, and we consider a minimizing sequence $y_n \in C$, $n \in \mathbb{N}$, so that $\|y_n\| \to I \geq 0$.

The sequence is Cauchy:

$$\|y_m - y_n\|^2 = 2\|y_m\|^2 + 2\|y_n\|^2 - \|y_m + y_n\|^2 = 2\|y_m\|^2 + 2\|y_n\|^2 - 4\left\|\frac{y_m + y_n}{2}\right\|^2$$

and since $(y_m + y_n)/2 \in C$ by convexity, we have

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

and

$$0 \leq \|y_m - y_n\|^2 = 2\|y_m\|^2 + 2\|y_n\|^2 - 4\left\|\frac{y_m + y_n}{2}\right\|^2$$
$$\leq 2\|y_m\|^2 + 2\|y_n\|^2 - 4I^2 \xrightarrow{m,n\to\infty} 0.$$

This show the existence of the limit $y_n \to y_\infty$ by using the completeness of the space. By closedness of C then we have $y_\infty \in C$, which implies the existence of a minimizer and also that $I > 0$ since $0 \notin C$.

Let us show that the minimizer is unique. The intuitive reason is that a scalar product structure implies the *strict convexity* of the unit ball of the space, and therefore there are no non-trivial segment at constant distance from the origin. Let us argue by contradiction and suppose that there are two distinct minimizers $y^1_\infty \neq y^2_\infty$, $y^1_\infty, y^2_\infty \in C$.

First proof: consider the mid-point $y^3_\infty = (y^1_\infty + y^2_\infty)/2$ and write (by expanding)

$$I^2 \leq \left\|y^3_\infty\right\|^2 = \frac{I^2}{2} + \frac{1}{2}\left\langle y^1_\infty, y^2_\infty\right\rangle \quad \Longrightarrow \quad I^2 \leq \left\langle y^1_\infty, y^2_\infty\right\rangle.$$

This implies since $I^2 = \|y^1_\infty\|^2 = \|y^2_\infty\|^2$ that

$$\langle y^1_\infty, y^2_\infty - y^1_\infty \rangle \geq 0, \quad \langle y^1_\infty - y^2_\infty, y^2_\infty \rangle \geq 0, \quad \Longrightarrow \quad \left\| y^1_\infty - y^2_\infty \right\| = 0.$$

Second proof: We have for any $t \in [0,1]$ that $ty^1_\infty + (1-t)y^2_\infty \in C$ and therefore

$$I^2 \leq \left\| ty^1_\infty + (1-t)y^2_\infty \right\|^2 = \left\| y^2_\infty \right\|^2 + 2t\langle y^2_\infty, y^1_\infty - y^2_\infty \rangle + t^2 \left\| y^1_\infty - y^2_\infty \right\|^2.$$

Using that $\left\| y^2_\infty \right\|^2 = I^2$ we deduce that

$$\forall\, t \in [0,1], \quad 2t\langle y^2_\infty, y^1_\infty - y^2_\infty \rangle + t^2 \left\| y^1_\infty - y^2_\infty \right\|^2 \geq 0.$$

Taking $t \to 0^+$ this imposes $\langle y^2_\infty, y^1_\infty - y^2_\infty \rangle \geq 0$. But a symmetric argument would yield $\langle y^1_\infty, y^2_\infty - y^1_\infty \rangle \geq 0$. Summing both inequalities gives

$$0 \leq \langle y^2_\infty, y^1_\infty - y^2_\infty \rangle + \langle y^1_\infty, y^2_\infty - y^1_\infty \rangle = \langle y^2_\infty - y^1_\infty, y^1_\infty - y^2_\infty \rangle = -\left\| y^1_\infty - y^2_\infty \right\|^2$$

which shows uniqueness.

Complement: arguing similarly we can now prove the Lipschitz regularity. Consider $t \in (0,1)$, two points $x_1, x_2 \in H$ and their projections $P_C(x_1), P_C(x_2) \in C$, and we write down the two inequalities

$$\begin{aligned} \|P_C(x_2) - x_2\|^2 &\leq \|tP_C(x_1) + (1-t)P_C(x_2) - x_2\|^2 \\ &= \|P_C(x_2) - x_2\|^2 + t^2 \|P_C(x_2) - P_C(x_1)\|^2 + 2t \langle P_C(x_2) - x_2, P_C(x_1) - P_C(x_2) \rangle \end{aligned}$$

$$\begin{aligned} \|P_C(x_1) - x_1\|^2 &\leq \|tP_C(x_2) + (1-t)P_C(x_1) - x_1\|^2 \\ &= \|P_C(x_1) - x_1\|^2 + t^2 \|P_C(x_2) - P_C(x_1)\|^2 + 2t \langle P_C(x_1) - x_1, P_C(x_2) - P_C(x_1) \rangle \end{aligned}$$

which use the fact that $tP_C(x_1) + (1-t)P_C(x_2) \in C$, $tP_C(x_2) + (1-t)P_C(x_1) \in C$. We deduce

$$t^2 \|P_C(x_2) - P_C(x_1)\|^2 + 2t \langle P_C(x_2) - x_2, P_C(x_1) - P_C(x_2) \rangle \geq 0$$
$$t^2 \|P_C(x_2) - P_C(x_1)\|^2 + 2t \langle P_C(x_1) - x_1, P_C(x_2) - P_C(x_1) \rangle \geq 0$$

and by summing the two inequalities

$$2t^2 \|P_C(x_2) - P_C(x_1)\|^2 \geq 2t \langle P_C(x_2) - x_2 + x_1 - P_C(x_1), P_C(x_2) - P_C(x_1) \rangle$$

which implies

$$\|P_C(x_2) - P_C(x_1)\| \leq \frac{1}{1-t} \|x_2 - x_1\|$$

and we finally let $t \to 0$. □

Let us now prove the Riesz representation theorem (in the case of a real Hilbert space, the complex case being similar).

Theorem 3.2 (Riesz representation Theorem). *Consider a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ and a continuous linear form $g : \mathcal{H} \to \mathbb{R}$. Then there is a unique $y \in H$ such that*

$$\forall\, x \in \mathcal{H}, \quad g(x) = \langle x, y \rangle.$$

Moreover the map $g \in \mathcal{H}' \mapsto y \in \mathcal{H}$ is linear and continuous (in fact isometric)

$$|||g|||_{\mathcal{L}(\mathcal{H},\mathbb{R})} = \|y\|_{\mathcal{H}}$$

where the LHS in the last equation denotes the canonical operator norm

$$|||g|||_{\mathcal{L}(\mathcal{H},\mathbb{R})} := \sup_{\|x\|_{\mathcal{H}}=1} |g(x)|.$$

Proof of Theorem 6.2. If $g = 0$ then $y = 0$ is the only solution and we are done. If g is non-zero, then the vector structure implies that $g(\mathcal{H}) = \mathbb{R}$, and there is $y \in \mathcal{H}$ so that $g(y) = 1$. Consider the non-empty set

$$C_g = \{y \in \mathcal{H}, \quad g(y) = 1\}.$$

It is convex by linearity of g, and closed by continuity of g. We then define $y^* = P_C(0)$ from the previous proposition, which satisfies

$$\|y^*\| = \min_{y \in C_g} \|y\|.$$

Since $y^* \in C_g$ we have $y^* \neq 0$. Let us prove that $y^* \perp \mathrm{Null}(g)$. Consider $x \in \mathrm{Null}(g)$ and $t \in R$, then $g(y^* + tx) = 1$ hence $y^* + tx \in C_g$. Therefore

$$\forall\, t \in \mathbb{R}, \quad \|y^* + tx\|^2 \geq \|y^*\|^2$$

which implies

$$\forall\, t \in \mathbb{R}, \quad 2t\langle y^*, x\rangle + t^2\|x\|^2 \geq 0$$

and therefore $\langle y^*, x\rangle = 0$ by taking $t \to 0^-$ or $t \to 0^+$.

Finally for any $x \in \mathcal{H}$ we have the decomposition

$$\forall\, x \in \mathcal{H}, \quad (x - g(x)y^*) \in \mathrm{Null}(g) \subset (y^*)^\perp$$

and therefore $g(x) = \langle x, y^*\rangle / \|y^*\|^2$, which conclude the proof with $y^{**} := y^*/\|y^*\|^2$. The identity on the norm follows from Cauchy-Schwarz' inequality, and testing it with $x = y/\|y\|$. □

4. The Dirichlet problem for the Poisson equation

Since as we have just seen, the Cauchy problem is ill-posed, what is the "correct" way of obtaining solutions to $\Delta u = f$? It turns out that the "correct" problem to study is the so-called *Dirichlet problem*, i.e. solving $\Delta u = f$ on a bounded smooth open set $\mathcal{U} \subset \mathbb{R}^\ell$ with prescribed boundary values on the boundary $\partial\mathcal{U}$. A hint that this is correct is provided by the fact that we can estimate a priori all solutions of the Dirichlet problem, as we shall see. Let us first consider the simplest case, when the prescribed boundary values are assumed to vanish. Consider a priori some u which satisfies

$$\begin{aligned} &\Delta u(x) = f(x) \text{ for any } x = (x_1, \dots, x_\ell) \in \mathcal{U}, \\ &u \in C^2(\mathcal{U}) \cap C^1(\overline{\mathcal{U}}) \text{ and } u(x) = 0 \text{ on } x \in \partial\mathcal{U}. \end{aligned}$$

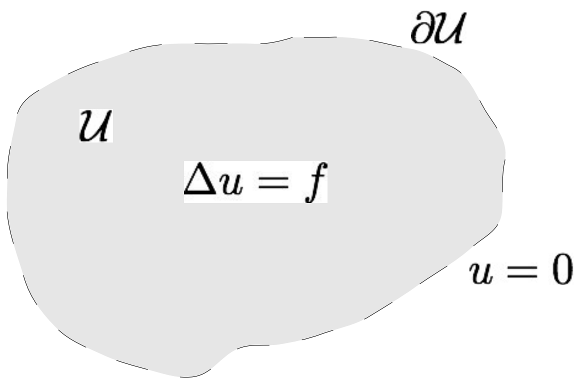

FIGURE 4.1. The Dirichlet problem.

4.1. The key a priori estimate. Again we shall start by highlighting the key a priori estimate that we can obtain on this problem, and then (shortly) justify rigorously the proofs that can be drawn from it. As before we argue *a priori*, assuming that, as we said, $u \in C^2(\mathcal{U}) \cap C^1(\overline{\mathcal{U}})$ and $u = 0$ on $\partial\mathcal{U}$. However since we do not prescribe anything on the first derivative[2], we need to establish en estimate that does not depend boundary integrals of the gradient on $\partial\mathcal{U}$.

We multiply the equation by u and integrate to obtain

$$\int_{\mathcal{U}} (\Delta u) u \, \mathrm{d}x = \int_{\mathcal{U}} uf \, \mathrm{d}x.$$

Integrating by parts (in view of the boundary conditions[3]), we get

$$\int_{\mathcal{U}} |\nabla u|^2 \, \mathrm{d}x = -\int_{\mathcal{U}} uf \, \mathrm{d}x \leq \|u\|_{L^2(\mathcal{U})} \|f\|_{L^2(\mathcal{U})}.$$

We now use Theorem 1.14 (Poincaré's inequality): $\mathcal{U} \subset \mathbb{R}^\ell$ is open smooth bounded set such hence there exists $C_{\mathcal{U}} > 0$ (only depending on $\mathcal{U}$) such that for any $u \in C^1(\overline{\mathcal{U}})$ such that $u = 0$ on $\partial\mathcal{U}$

$$\left(\int_{\mathcal{U}} u(x)^2 \, \mathrm{d}x \right)^{1/2} \leq C_{\mathcal{U}} \left(\int_{\mathcal{U}} |\nabla u(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

We apply this inequality to deduce

$$\|u\|^2_{L^2(\mathcal{U})} \leq C^2_{\mathcal{U}} \int_{\mathcal{U}} |\nabla u|^2 \, \mathrm{d}x \leq C^2_{\mathcal{U}} \|u\|_{L^2(\mathcal{U})} \|f\|_{L^2(\mathcal{U})}$$

[2]We know from the previous section that it would be a bad idea to prescribe both u and its normal gradient on some Cauchy surface, and the Dirichlet problem consists in removing the condition on the gradient. However there is another important class of boundary conditions, the Neumann boundary conditions, where some orthogonality condition is prescribed on the gradient at $\partial\mathcal{U}$, but in this case *nothing* is prescribed on the value of u.

[3]Note in particular why $C^1(\overline{\mathcal{U}})$ is natural.

which implies

$$\|u\|_{L^2(\mathcal{U})} \leq C_{\mathcal{U}}^2 \|f\|_{L^2(\mathcal{U})}.$$

But now we can build up on this first estimate, and obtain more: by boostraping the information on the L^2 norm of u into the first a priori estimate we obtain

$$\int_{\mathcal{U}} |\nabla u|^2 \, \mathrm{d}x \leq \|u\|_{L^2(\mathcal{U})} \|f\|_{L^2(\mathcal{U})} \leq C_{\mathcal{U}}^2 \|f\|_{L^2(\mathcal{U})}^2.$$

Combining the two last inequalities we can write

$$\|u\|_{H^1(\mathcal{U})}^2 \leq \left(C_{\mathcal{U}}^4 + C_{\mathcal{U}}^2\right) \|f\|_{L^2(\mathcal{U})}^2.$$

As a consequence of this discussion we can prove

PROPOSITION 4.1. *Suppose $u, v \in C^2(\mathcal{U}) \cap C^1(\overline{\mathcal{U}})$ satisfy $\Delta u = \Delta v = f$ on $\mathcal{U}$ with $u = v = g$ on $\partial\mathcal{U}$. Then $u = v$.*

PROOF OF PROPOSITION 4.1. The proof follows from the previous estimate applied to the solution $w = u - v$ which solves $\Delta w = 0$ on $\mathcal{U}$ with $w = 0$ on $\partial\mathcal{U}$:

$$\|w\|_{H^1(\mathcal{U})}^2 \leq 2C_{\mathcal{U}} \|0\|_{L^2(\mathcal{U})}^2 = 0$$

from our previous estimate, which shows by continuity that $w = 0$ everywhere. $\square$

This last proposition solves the problem of uniqueness, but leaves open that of existence and continuity according to the data, which are the object of the next subsections.

4.2. Existence of weak (generalised) solutions. Weak formulations are an important tool for the analysis of PDEs that permit the transfer of concepts of *linear algebra* to solve the problems. In a weak formulation, an equation is no longer required to hold in the classical sense (and this is not even well defined) and has instead weak solutions only with respect to certain "test vectors" or "test functions". This is equivalent to formulating the problem to require a solution in the sense of a distribution. We introduce a formulation for weak solutions for the Poisson equation and show how to construct solutions using Riesz representation Theorem. (We will then explain how to construct solutions to more general elliptic problems with the help of the *Lax-Milgram theorem.*)

Let us first define the notion of weak solutions. Assume that $\Delta u = f$ with $u \in C^2(\mathcal{U}) \cap C^1(\overline{\mathcal{U}})$ and $u = 0$ on $\partial\mathcal{U}$ then for any $v \in C^2(\mathcal{U}) \cap C^1(\overline{\mathcal{U}})$ with $v = 0$ on $\partial\mathcal{U}$ we have

$$\langle\langle u, v\rangle\rangle := \int_{\mathcal{U}} \nabla u \cdot \nabla v \, \mathrm{d}x = -\int_{\mathcal{U}} (\Delta u) v \, \mathrm{d}x = -\int_{\mathcal{U}} f v \, \mathrm{d}x$$

where we have denoted by $\langle\langle \cdot, \cdot \rangle\rangle$ the scalar product associated with the homogeneous norm $\dot{H}^1(\mathcal{U})$, keeping the notation $\langle \cdot, \cdot \rangle$ for the usual $L^2(\mathcal{U})$ scalar product. Observe crucially that the objects in the LHS and RHS of this statement still make sense as soon as $u, v \in H_0^1(\mathcal{U})$.

DEFINITION 4.2. *We call* generalised (or weak) solution *a function $u \in H_0^1(\mathcal{U})$ such that*

$$\forall\, v \in H_0^1(\mathcal{U}), \quad \langle\langle u, v\rangle\rangle = -\langle f, v\rangle.$$

Note that we have an equivalent characterization that $u \in H^1_0(\mathcal{U})$ *and*

$$\langle u, \Delta v\rangle = \langle f, v\rangle$$

for all $v \in C^\infty_c(\mathcal{U})$ *smooth of compact support in* $\mathcal{U}$*. This latter equality is the statement that* u *is a distributional solution of* $\Delta u = f$*.*

EXERCISE 32. *Prove the equivalence in the definition.*

REMARK 4.3. *Note the important idea behind this reformulation: the boundary conditions have been enforced-encoded* in the functional space *itself.*

We can now state and prove the existence theorem:

THEOREM 4.4. *Let* $\mathcal{U} \subset \mathbb{R}^\ell$ *and open set, and* $f \in L^2(\mathcal{U})$*. Then there exists a unique* $u \in H^1_0(\mathcal{U})$ *such that* u *is a weak solution of* $\Delta u = f$*, in the sense defined above.*

PROOF OF THEOREM 4.4. The proof is a straightfoward application of the Riesz representation Theorem: we consider the following linear form on $H^1_0(\mathcal{U})$:

$$\forall\, v \in H^1_0(\mathcal{U}), \quad g(v) := -\langle f, v\rangle$$

which is continuous by Cauchy-Schwarz and Poincaré's inequalities

$$\forall\, v \in H^1_0(\mathcal{U}), \quad |g(v)| = |\langle f, v\rangle| \le \|f\|_{L^2(\mathcal{U})}\|v\|_{L^2(\mathcal{U})} \lesssim \|f\|_{L^2(\mathcal{U})}\|v\|_{\dot H^1(\mathcal{U})}.$$

Then the Riesz representation theorem applied in the Hilbert space $H^1_0(\mathcal{U})$ endowed with the norm $\dot H^1(\mathcal{U})$, shows that there is a unique $u \in H^1_0(\mathcal{U})$ so that

$$\forall\, v \in H^1_0(\mathbb{R}^\ell), \quad g(v) = \langle\langle u, v\rangle\rangle$$

which concludes the proof. $\square$

REMARK 4.5. *Observe moreover that in the previous statement the solution map* $\mathfrak{S} : f \mapsto u$ *is continuous from* $L^2(\mathcal{U})$ *to* $H^1_0(\mathcal{U})$ *since*

$$\|g\|_{H^1_0(\mathcal{U})^*} = \sup_{\|v\|_{\dot H^1(\mathcal{U})}=1} |g(v)| \lesssim \|f\|_{L^2(\mathcal{U})}$$

and $g \mapsto u$ *is an isometry in the representation theorem.*

EXERCISE 33. *In fact prove that* $\mathfrak{S}$ *is even continuous from* $H^{-1}_0(\mathcal{U})$*, the dual of* $H^1_0(\mathcal{U})$ *for the* $L^2(\mathcal{U})$ *scalar product* $\langle \cdot, \cdot\rangle$*, to* $H^1_0(\mathcal{U})$*:*

$$\|f\|_{H^{-1}_0(\mathcal{U})} := \sup_{\|v\|_{H^1_0(\mathcal{U})}=1} \langle f, v\rangle_{L^2(\mathcal{U})}.$$

5. Toolbox: compactness tools in PDEs

5.1. Weak vs strong compactness. Let us first give some recalls. We consider a separable Hilbert space $\mathcal{H}$.

DEFINITION 5.1. *The* weak compactness *of a sequence* $u_n \in \mathcal{H}$ *in a Hilbert space* $\mathcal{H}$ *means that there is a subsequence* $u_{\varphi(n)}$ *so that*

$$\forall\, v \in \mathcal{H}, \quad \langle u_{\varphi(n)}, v\rangle \to \langle u_\infty, v\rangle$$

for some $u_\infty \in \mathcal{H}$*.*

The strong compactness *of a sequence* $u_n \in \mathcal{H}$ *in a Hilbert space* $\mathcal{H}$ *means that there is a subsequence* $u_{\varphi(n)}$ *so that*

$$\left\| u_{\varphi(n)} - u_\infty \right\|_{\mathcal{H}} \to 0$$

for some $u_\infty \in \mathcal{H}$.

The weak topology *on* $\mathcal{H}$ *is the topology induced by continuous linear forms on* $\mathcal{H}$*, i.e. the dual* $\mathcal{H}^*$*, which turns out to be isometric to* $\mathcal{H}$ *by Riesz representation theorem.*

EXERCISE 34. *The unit ball of* $\mathcal{H}$ *is strongly compact iff* $\mathcal{H}$ *is finite dimensional. Show that in this case the weak and strong topologies coincide.*

Hint: If $\mathcal{H}$ *is finite dimensional the implication is Bolzano-Weierstrass. Else argue by contradiction and construct a sequence with no cluster points (with distances uniformly bounded from below between all pairs).*

EXERCISE 35. *Prove however that the unit ball is always weakly compact.*

Hint: consider a Hilbertian base and perform a diagonal process.

EXERCISE 36. *Show that the weak topology cannot be associated to any metric on the whole space* $\mathcal{H}$*. Show however that the unit ball of* $\mathcal{H}$*, endowed with the weak topology, can be endowed with a metric.*

Hint for the first part: Show that an open set for the weak topology always contains a whole line.

In a non-reflexive Banach space, the situation is more complicated, there are two non-strong topologies: the weak and weak-* topologies, and the Banach-Alaoglu theorem gives compactness of the unit ball of $\mathcal{H}^*$ for weak-* topology. The most general form relies upon Tychonov's theorem and the axiom of choice.

EXERCISE 37. *Show that the Banach-Alaoglu theorem for the dual of a* separable space *can be proved without using the axiom of choice, thanks to a countable diagonal argument.*

Note that in PDEs we almost always have spaces with the separability property and we thus almost never need the axiom of choice. The space of Schwartz' distributions may suggest the opposite, but any concrete problem lives in a subset of the set of distributions with countable basis of neighborhoods (e.g. tempered distributions).

EXERCISE 38. *What are the conditions for a sequence to be weakly-* compact in* $L^p(\mathbb{R})$, $p \in (1,\infty)$*? in* $L^1(\mathbb{R})$*?*

Let us now turn to the question of *strong compactness*. Here the "root" of all theorems is Arzelà-Ascoli's Theorem:

THEOREM 5.2 (Arzelà-Ascoli's Theorem). *A sequence* f_n *of continuous real functions on a compact Hausdorff space* X *is relatively compact in the topology induced by the uniform norm if and only if it is (1) equicontinuous: for any* $\varepsilon > 0$ *and* $x \in X$*, there is* V *neighborhood of* x *so that*

$$\forall\, n \geq 1,\ \forall\, y \in V, \quad |f_n(x) - f_n(y)| \leq \varepsilon,$$

and (2) pointwise bounded: for any $x \in X$

$$\sup_{n \geq 0} |f_n(x)| < +\infty.$$

REMARK 5.3. *Observe that the boundedness assumption is reminiscent of weak compactness statements. However the equicontinuity is an assumption of uniform regularity along the sequence that proscribes the possibility of infinite oscillatory behaviors as n goes to infinity. As a general principle, the main obstacle to weak compactness is the divergence, and the main obstacle to a weak compactness being strong is* oscillations *(i.e. divergence in Fourier variable) which can be ruled out by uniform regularity assumption. The Arzelà-Ascoli Theorem is in fact the origin of most strong compactness theorems in analysis.*

PROOF OF THEOREM 5.2. Construct a countable dense subset Y of X, then show convergence of a subsequence of $f_n(x)$ for any $x \in Y$, and find a subsequence convergence for all points of Y by a diagonal argument. This defines a limit f on Y. The equicontinuity implies then the continuity of this limit on Y. It then can extended by density to a continuous function on X. The uniform convergence follows from the equicontinuity and the pointwise convergence on Y. □

EXERCISE 39. *Give example of sequences converging weakly but not strongly. Show that the weak limit of a product is not in general the product of the weak limits. Show that weak convergence plus convergence of the norm implies strong convergence.*

EXERCISE 40. *Search in textbooks some sufficient conditions to strong compactness of a set of functions of $L^p(\mathbb{R})$, $p \in (1, +\infty)$?*

We shall conclude this subsection with the following fundamental compactness theorem:

THEOREM 5.4 (Rellich-Kondrachov Compactness Theorem). *Assume $\mathcal{U} \subset \mathbb{R}^\ell$ is a bounded open set with $\partial\mathcal{U}$ smooth then*

$$H^1(\mathcal{U}) \subset\subset L^2(\mathcal{U})$$

which means that the canonical inclusion is compact.

PROOF OF THEOREM 5.4. Let us only sketch the proof as it requires tools from Sobolev inequalities. We admit here the extension process of functions of $H^1(\mathcal{U})$ to $H^1(\mathbb{R}^\ell)$ with compact support.

In dimension $\ell = 1$, we know from Morrey's inequality (cf. mid-term assignements) that $H^1(\mathcal{U}) \subset C^{1/2-0}(\overline{\mathcal{U}})$. Then the classical Arzelà-Ascoli Theorem on the compact set $\mathcal{U}$ concludes the proof, as the Hölder regularity implies the equicontinuity, and L^∞ bound implies the required boundedness.

In higher dimension $\ell \geq 2$, the Sobolev inequalities show the inclusion and only the compactness of the inclusion remains to be proved. Then one considers a sequence $u_n \in H^1(\mathcal{U})$ uniformly bounded, extends the function to $\mathbb{R}^\ell$ with a common support K. Next one shows that a uniformly bounded sequence in $H^1(\mathbb{R}^\ell)$ can be uniformly approximated in $L^2(\mathcal{U})$ by the convolution with an approximation of the unit. Finally apply Arzelà-Ascoli Theorem on K on each approximation to show its compactness in $L^2(K)$, and use a standard diagonal argument. □

5.2. Compact operators. We introduce an important class of operators that are useful generalisation of finite ranked operators (matrices).

Definition 5.5. *We consider a bounded linear operator $\mathfrak{K} : \mathcal{H} \to \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space. This operator is said* compact *if it maps the unit ball into a relatively compact set (for the strong topology).*

Remark 5.6. *Note that the boundedness is in fact implied by the second part of the definition.*

The key result we shall need, and at the starting point of the Fredholm theory, is:

Proposition 5.7. *Given a compact operator $\mathfrak{K}$ on a Hilbert space H, then* $\mathrm{Id} + \mathfrak{K}$ *has closed range, finite dimensional kernel and finite dimensional cokernel (the cokernel is the orthogonal complement of the range).*

Proof. First, the null space of $\mathrm{Id} + \mathfrak{K}$ has a compact unit ball by the compactness of $\mathfrak{K}$: any sequence g_n with $\|g_n\| \le 1$ and $g_n + \mathfrak{K}g_n = 0$ satisfies that $\mathfrak{K}g_n$ has a converging subsequence by the compactness of $\mathfrak{K}$, which implies that g_n has a converging subsequence. This implies by Riesz theorem that $\mathrm{Ker}(\mathrm{Id} + \mathfrak{K})$ is finite dimensional.

Second, let us prove that $(\mathrm{Id} + \mathfrak{K})$ is coercive on the orthogonal of its kernel: there is $\lambda_{\mathfrak{K}} > 0$ so that $\|g + \mathfrak{K}g\| \ge \lambda_{\mathfrak{K}} \|g\|$ for $g \in \mathrm{Ker}(\mathrm{Id} + \mathfrak{K})^\perp$. It is proved by contradiction: consider $g_n \in \mathrm{Ker}(\mathrm{Id} + \mathfrak{K})^\perp$ a sequence so that $\|g_n\| = 1$ and $\|g_n + \mathfrak{K}(g_n)\| \to 0$ as n goes to infinity. Then using on the one hand the weak compactness of the unit ball and on the other hand the compactness of the operator $\mathfrak{K}$, there is a subsequence $g_{\varphi(n)}$ so that $g_{\varphi(n)} \rightharpoonup g$ (weak convergence) and $\mathfrak{K}g_{\varphi(n)}$ is strongly converging. Since the weak convergence of $g_{\varphi(n)}$ implies $\mathfrak{K}g_{\varphi(n)} \rightharpoonup \mathfrak{K}g$ and the weak and strong limits are the same, we deduce that $\mathfrak{K}g_{\varphi(n)} \to \mathfrak{K}g = -g$ (strong convergence) and then finally $g_{\varphi(n)} \to g \in \mathrm{Ker}(\mathrm{Id} + \mathfrak{K})$ (strong convergence). But by weak limit we also have $g \in \mathrm{Ker}(\mathrm{Id} + \mathfrak{K})^\perp$, which implies that $g = 0$. Moreover the strong convergence implies $\|g\| = 1$, which yields the desired contradiction.

Third let us prove that the range of $\mathrm{Id} + \mathfrak{K}$ is closed. Consider a sequence $g_n + \mathfrak{K}g_n \to h$. We can decompose $g_n = g_n^1 + g_n^2 \in \mathrm{Ker}(\mathrm{Id} + \mathfrak{K})^\perp + \mathrm{Ker}(\mathrm{Id} + \mathfrak{K})$. From the previous point, we have that g_n^1 is a bounded sequence, with $g_n^1 + \mathfrak{K}g_n^1 \to h$. There is a subsequence $g_{\varphi(n)}$ so that $\mathfrak{K}g_{\varphi(n)}$ is converging (in the strong topology), which implies that $g_{\varphi(n)}$ is strongly converging as well, say to some g. Finally by continuity of $\mathfrak{K}$ we deduce $g + \mathfrak{K}g = h \in \mathrm{Range}(\mathrm{Id} + \mathfrak{K})$.

Fourth let us prove that the adjoint of a compact operator is also compact. Let us recall that the adjoint operator $\mathfrak{K}^*$ of $\mathfrak{K}$ is defined by the formula

$$\forall\, f, g \in H, \quad \langle \mathfrak{K}f, g \rangle = \langle f, \mathfrak{K}^* g \rangle$$

and Riesz' representation theorem. Consider a bounded sequence $g_n \in H$. Then $\mathfrak{K}^* g_n$ is bounded and there is a subsequence $g_{\varphi(n)}$ so that $g_{\varphi(n)} \rightharpoonup g \in H$ and $\mathfrak{K}^* g_{\varphi(n)} \rightharpoonup \mathfrak{K}^* g \in H$. Then

$$\left\| \mathfrak{K}^* g_{\varphi(n)} - h \right\|^2 = \langle \mathfrak{K}^2 (g_{\varphi(n)} - g), g_{\varphi(n)} - g \rangle.$$

Then by the compactness of $\mathfrak{K}$, we have $\mathfrak{K}^2(g_{\varphi(n)} - g) \to 0$ (strong convergence), which implies that $\left\| \mathfrak{K}^* g_{\varphi(n)} - h \right\| \to 0$ and concludes the proof of the fact that $\mathfrak{K}^*$ is compact.

Then applying the points 1-2-3 above to $(\mathrm{Id} + \mathfrak{K})^* = \mathrm{Id} + \mathfrak{K}^*$, we deduce that the cokernel is also finite dimensional, since we recall that

$$\mathrm{Range}(\mathrm{Id} + \mathfrak{K})^\perp = \mathrm{Ker}(\mathrm{Id} + \mathfrak{K}^*), \quad \mathrm{Range}(\mathrm{Id} + \mathfrak{K}^*)^\perp = \mathrm{Ker}(\mathrm{Id} + \mathfrak{K}).$$

This concludes the proof. Observe also that $H_0^1(\mathcal{U})$ is the direct sum of $\mathrm{Ker}(\mathrm{Id}+\mathfrak{K}^*)$ and $\mathrm{Range}(\mathrm{Id}+\mathfrak{K})$ (the latter being closed). □

Let us now study the dimensions of $\mathrm{Ker}(\mathrm{Id}+\mathfrak{K})$ and $\mathrm{Coker}(\mathrm{Id}+\mathfrak{K})$ and prove that they are the same.

Proposition 5.8. *Given a compact operator $\mathfrak{K}$ on a Hilbert space H, then one has*

$$\mathrm{dimKer}(\mathrm{Id}+\mathfrak{K}) = \mathrm{dimCoker}(\mathrm{Id}+\mathfrak{K}) < +\infty.$$

Proof. The proof relies on a *continuation argument* and the following more general *claim*: if a bounded operator $\mathfrak{T}$ has closed range, finite dimensional kernel and cokernel, and satisfies $\dim(\mathrm{Ker}(\mathfrak{T})) = \dim(\mathrm{Coker}(\mathfrak{T}))$, then all these properties are preserved under small enough perturbation (for the norm of bounded operators).

The proof of this claim is a reduction to the finite dimensional case. We decompose the space as $C \oplus \mathrm{Ker}(\mathfrak{T})$ and $\mathrm{Range}(\mathfrak{T}) \oplus D$ with $\dim(\mathrm{Ker}(\mathfrak{T})) = \dim(D)$. The operator $\mathfrak{T}$ writes

$$\mathfrak{T} = \begin{pmatrix} \tilde{\mathfrak{T}} & 0 \\ 0 & 0 \end{pmatrix}$$

in this decomposition, and we consider a small perturbation P which writes

$$\mathfrak{P} = \begin{pmatrix} \mathfrak{P}_{11} & \mathfrak{P}_{12} \\ \mathfrak{P}_{21} & \mathfrak{P}_{22} \end{pmatrix}$$

in this decomposition. Then we define the two invertible operators

$$\mathfrak{G} = \begin{pmatrix} \mathrm{Id} & -(\tilde{\mathfrak{T}}+\mathfrak{P}_{11})^{-1}\mathfrak{P}_{12} \\ 0 & \mathrm{Id} \end{pmatrix}, \quad \mathfrak{H} = \begin{pmatrix} \mathrm{Id} & 0 \\ -\mathfrak{P}_{21}(\tilde{\mathfrak{T}}+\mathfrak{P}_{11})^{-1} & \mathrm{Id} \end{pmatrix}$$

and we write

$$\mathfrak{H}(\mathfrak{T}+\mathfrak{P})\mathfrak{G} = \begin{pmatrix} \tilde{\mathfrak{T}}+\mathfrak{P}_{11} & 0 \\ 0 & -\mathfrak{P}_{21}(\tilde{\mathfrak{T}}+\mathfrak{P}_{11})^{-1}\mathfrak{P}_{12}+\mathfrak{P}_{22} \end{pmatrix}.$$

Let us denote $\mathfrak{A} := -\mathfrak{P}_{21}(\tilde{\mathfrak{T}}+\mathfrak{P}_{11})^{-1}\mathfrak{P}_{12}+\mathfrak{P}_{22}$ which goes from $\mathrm{Ker}(\mathfrak{T})$ to D. Since these two spaces have the same (finite) dimension we have from the rank nullity theorem

$$\dim(\mathrm{Ker}(\mathfrak{A})) = \dim(\mathrm{Coker}(\mathfrak{A})).$$

Since $\mathfrak{G}$, $\mathfrak{H}$ are invertible, and, for $\mathfrak{P}$ small enough, $\tilde{\mathfrak{T}}+\mathfrak{P}_{11}$ is invertible, we deduce that

$$\dim(\mathrm{Ker}(\mathfrak{T})) = \dim(\mathrm{Ker}(\mathfrak{A})), \quad \dim(\mathrm{Coker}(\mathfrak{T})) = \dim(\mathrm{Coker}(\mathfrak{A}))$$

which concludes the proof of the claim.

Now going back to the operator $\mathrm{Id}+\mathfrak{K}$ we consider the path $\mathrm{Id}+t\mathfrak{K}$, $t \in [0,1]$ and perform a continuation argument. The interval $[0,t_0)$, $t_0 > 0$, for which the properties above are satisfied is non-empty and open from the previous claim. But since $\mathrm{Id}+t_0\mathfrak{K}$ still has finite range and finite dimensional kernel and cokernel (as $t_0\mathfrak{K}$ is compact), the same argument could be performed around t_0 as well, which shows that $t_0 = 1$, and therefore concludes the proof. □

6. General elliptic equations and Lax-Milgram Theorem

Let us consider now the more general boundary-value problem

$$\begin{cases} \mathfrak{P}u = f & \text{in } \mathcal{U} \\ u = g & \text{on } \partial\mathcal{U}. \end{cases} \tag{6.1}$$

where $\mathcal{U} \subset \mathbb{R}^\ell$ is open, bounded and smooth, $\mathfrak{P}$ denotes a second-order linear elliptic partial operator, and g is smooth enough to be the trace of an $H^1(\mathcal{U})$ function.

6.1. Divergence and non-divergence forms. The operator $\mathfrak{P}$ can be given in two different forms. The first one is

$$\mathfrak{P}u := -\sum_{i,j=1}^{\ell} \partial_j \left(a_{ij}(x)\partial_i u\right) + \sum_{i=1}^{\ell} b_i(x)\partial_i u + c(x)u \tag{6.2}$$

and is called the *divergence form*, as the leading-order part of the operator writes $\nabla \cdot U$ for some $U(x) \in \mathbb{R}^\ell$.

The second formulation is

$$\mathfrak{P}u = -\sum_{i,j=1}^{\ell} a_{ij}(x)\partial^2_{ij} u + \sum_{i=1}^{\ell} b_i(x)\partial_i u + c(x)u \tag{6.3}$$

and is called the *non-divergence form* for obvious reasons.

Remark 6.1. *Note the minus sign in front of the second-order term, which is convenient for reminding that the ellipticity is related to* coercivity *(positivity) of the Fourier symbol.*

Exercise 41. *Assume that the coefficients a_{ij}, b_i, c are C^1 and show that an operator given in divergence form can be rewritten in non-divergence form, and vice-versa.*

Both forms are useful. The divergence form is usually better adapted to energy estimates which uses integration by parts, Hilbertian methods, and more generally integral arguments. The non-divergence form is better adapted to pointwise arguments and maximum principles.

6.2. Ellipticity and reduction of the problem. We consider in the rest of this section the operator $\mathfrak{P}$ as above in divergence form and we assume (1) w.l.o.g. the symmetry condition

$$\forall\, i,j \in \{1,\dots,\ell\}, \quad a_{ij} = a_{ij}$$

and (2) the uniform ellipticity condition

$$\forall\, x \in \mathcal{U}, \quad A(x) \geq \alpha > 0, \quad A(x) = (a_{ij})_{1 \leq i,j \leq \ell}\,.$$

We also assume that the coefficients satisfy $a_{ij}, b_i, c \in L^\infty(\mathcal{U})$. Let us denote $B(x) = (b_1(x),\dots,b_\ell(x))^\perp$. We consider the equation $\mathfrak{P}u = f$ in $\mathcal{U}$ with some $f \in L^2(\mathcal{U})$, together with the boundary condition $u = g$ on $\partial\mathcal{U}$ with some $g \in C^0(\partial\mathcal{U})$.

Observe that we can reduce to the case of zero boundary conditions: consider $w \in H^1(\mathbb{R}^\ell)$ so that w restricts to g on $\partial\mathcal{U}$. Then $\tilde{u} = u - w$ solves a similar elliptic

PDE for the same matrix $A(x)$ and some other RHS $\tilde{f}$. We therefore assume that $g = 0$ in the sequel.

6.3. The key a priori estimate. Assuming that $u \in C^2(\mathcal{U}) \cap C^1(\bar{\mathcal{U}})$, we multiply the equation by u and integrate it over $\mathcal{U}$:

$$\int_{\mathcal{U}} (\nabla u)^T A(x)(\nabla u)\,\mathrm{d}x + \int_{\mathcal{U}} u(x)B(x)\cdot\nabla u\,\mathrm{d}x + \int_{\mathcal{U}} cu^2\,\mathrm{d}x = \int_{\mathcal{U}} fu\,\mathrm{d}x$$

and we deduce that

$$\alpha \int_{\mathcal{U}} |\nabla u|^2\,\mathrm{d}x \le \|B\|_\infty \|u\|_{L^2(\mathcal{U})}\|\nabla u\|_{L^2(\mathcal{U})} + \|c\|_\infty \|u\|^2_{L^2(\mathcal{U})} + \|u\|_{L^2(\mathcal{U})}\|f\|_{L^2(\mathcal{U})}$$
$$\le \frac{\alpha}{2}\|\nabla u\|^2_{L^2(\mathcal{U})} + \|f\|^2_{L^2(\mathcal{U})} + C\|u\|^2_{L^2(\mathcal{U})}$$

for some constant $C > 0$ (possibly large). We deduce that

$$\|\nabla u\|^2_{L^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})}.$$

(Recall that the sign $\lesssim$ means that the inequality holds with a constant unimportant for the argument and not depending on the quantities in the inequality). Hence we see again a gain of regularity, however now the uniqueness does not follow immediately from the a priori estimate. We shall first deal with the case where the coercivity is recovered in the sense that one can establish the a priori estimate $\|\nabla u\|^2_{L^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})}$. This first case will be solved with the help of Lax-Milgram Theorem. We will then consider the general case thanks to the Fredholm theory.

6.4. The Lax-Milgram Theorem. Consider a linear form $\mathfrak{L}$ and a bilinear form (not necessarily symmetric!) $\mathfrak{B}$ on $H^1_0(\mathcal{U})$. Then we consider the problem of finding $u \in H^1_0(\mathcal{U})$ so that

$$\forall\, v \in H^1_0(\mathcal{U}), \quad \mathfrak{B}(u,v) = \mathfrak{L}(v). \tag{6.4}$$

THEOREM 6.2 (Lax-Milgram). *Assume in* (6.4) *that* $\mathfrak{L}$ *and* $\mathfrak{B}$ *are bounded as linear (resp. bilinear) forms on* $H^1_0(\mathcal{U})$. *Assume moreover that* $\mathfrak{B}$ *is* coercive*: there is* $\lambda > 0$ *so that*

$$\forall\, v \in H^1_0(\mathcal{U}), \quad B(v,v) \ge \lambda \|v\|_{H^1(\mathcal{U})}.$$

Then there exists a unique solution $u \in H^1_0(\mathcal{U})$ *to* (6.4). *The solution map* $\mathfrak{S} : \mathfrak{L} \mapsto u$ *is moreover continuous from* $H^{-1}_0(\mathcal{U})$ *to* $H^1_0(\mathcal{U})$.

PROOF OF THEOREM 6.2. For any given $u \in H^1_0(\mathcal{U})$, the linear form $v \in H^1_0(\mathcal{U}) \mapsto \mathfrak{B}(u,v) \in \mathbb{R}$ is continuous, and therefore the Riesz representation theorem shows that there is a (unique) u^* so that

$$\forall\, v \in H^1_0(\mathcal{U}), \quad \mathfrak{B}(u,v) = \langle\langle u^*, v\rangle\rangle.$$

Moreover since from the theorem the map $u \mapsto u^*$ is linear and continuous we can represent it as $u^* = \mathfrak{T}u$ for some bounded operator $\mathfrak{T}$ on $H^1_0(\mathcal{U})$. We also represent $\mathfrak{L}$ by $l \in H^1_0(\mathcal{U})$ by the same theorem

$$\forall\, v \in H^1_0(\mathcal{U}), \quad \mathfrak{L}(v) = \langle\langle l, v\rangle\rangle$$

and therefore the equation on u rewrites as

$$\forall\, v \in H_0^1(\mathcal{U}), \quad \langle\langle \mathfrak{T} u, v \rangle\rangle = \langle\langle l, v \rangle\rangle$$

which is equivalent to $\mathfrak{T} u = l \in H_0^1(\mathcal{U})$. It remains to prove that $\mathfrak{T}$ is invertible in $H_0^1(\mathcal{U})$ to conclude the proof (defining finally $u = \mathfrak{T}^{-1} l$).

The operator $\mathfrak{T}$ satisfies the following

$$\forall\, v \in H_0^1(\mathcal{U}), \quad \|v\|^2_{H^1(\mathcal{U})} \le \lambda \mathfrak{B}(v,v) \le \lambda \|\mathfrak{T} v\|_{H^1(\mathcal{U})} \|v\|_{H^1(\mathcal{U})}$$

which implies $\|\mathfrak{T} v\|_{H^1(\mathcal{U})} \ge \lambda^{-1} \|v\|_{H^1(\mathcal{U})}$ for any non-zero $v \in H_0^1(\mathcal{U})$. This proves the injectivity. This also proves that the range of $\mathfrak{T}$ is closed: if $\mathfrak{T} v_n$ is converging it is Cauchy and therefore v_n is Cauchy and converging to v_∞ since the space is complete, and by continuity of $\mathfrak{T}$, $\mathfrak{T} v_n \to \mathfrak{T} v_\infty$. Finally if the range of $\mathfrak{T}$ is not $H_0^1(\mathcal{U})$, because of its closedness there is a non-zero $w \perp \mathrm{Range}(\mathfrak{T})$ which contradicts $\langle\langle \mathfrak{T} w, w \rangle\rangle = \mathfrak{B}(w,w) > 0$. □

REMARK 6.3. *Note that if $\mathfrak{B}$ is symmetric: $\mathfrak{B}(u,v) = \mathfrak{B}(v,u)$, then a simpler proof can be devised: defining a scalar product out of $\mathfrak{B}$ one observes that it endows $H_0^1(\mathcal{U})$ with a norm equivalent to $H^1(\mathcal{U})$, for which Riesz representation theorem can be immediately applied. Lax-Milgram theorem is really meaningful only in the non-symmetric case.*

6.5. Existence of weak solutions assuming global coercivity. We go back to the problem and assume the operator $\mathfrak{P}$ is so that the following stronger a priori estimate yields

$$\|u\|_{H^1(\mathcal{U})} \lesssim \|u\|_{L^2(\mathcal{U})}.$$

EXERCISE 42. *Check that is always possible to satisfy this condition by taking the coefficients b_i and c small enough in terms of the constant λ so that $A \ge \lambda$.*

Then we want to apply Lax-Milgram theorem with

$$\begin{cases} \mathfrak{B}(u,v) := \displaystyle\int_{\mathcal{U}} (\nabla u)^T A(x) (\nabla v)\,\mathrm{d}x + \int_{\mathcal{U}} v B(x) \cdot \nabla u \,\mathrm{d}x + \int_{\mathcal{U}} c u v \,\mathrm{d}x, \\ \mathfrak{L} v := \langle f, v \rangle. \end{cases}$$

The definition of $\mathfrak{B}$ corresponds to $\langle \mathfrak{P} u, v \rangle$ after integration by parts. The fact that $\mathfrak{B}$ is bounded in $H^1(\mathcal{U})$ is clear from $a_{ij} \in L^\infty(\mathcal{U})$. The fact that $\mathfrak{L}$ is bounded is clear from $f \in L^2(\mathcal{U})$ and Poincaré's inquality (note here that in the definition of $\mathfrak{L}$ we use the standard $L^2(\mathcal{U})$ scalar product).

The coercivity assumption means in a more precise form that

$$\forall\, v \in H_0^1(\mathcal{U}), \quad \mathfrak{B}(v,v) \ge \lambda \|v\|^2_{H^1(\mathcal{U})}$$

for some constant $\lambda > 0$. We can then apply the Lax-Milgram theorem and conclude the proof.

6.6. Existence of weak solutions without the global coercivity. We now consider the same framework but only assume the (strict uniform) ellipticity

$$A(x) \geq \lambda_e > 0$$

on the second-order part of the operator $\mathfrak{P}$, but no more global coercivity on $\mathfrak{P}$. Let us denote

$$\mathfrak{P}_e = \sum_{i,j=1}^{\ell} a_{ij}(x) \partial^2_{ij}$$

the second-order part of the operator P.

6.6.1. *A first attempt.* The previous study show that the equation

$$\begin{cases} \mathfrak{P}_e u = f & \text{in } \mathcal{U} \\ u = 0 & \text{on } \partial \mathcal{U}. \end{cases}$$

admits a unique solution $u \in H^1_0(\mathcal{U})$, which depends continuously on the linear form $v \mapsto \langle f, v \rangle$ on $H^1_0(\mathcal{U})$. The norm of this linear form is the $H^{-1}_0(\mathcal{U})$ norm of f. We have hence inverted the operator $\mathfrak{P}_e$ (with the Dirichlet conditions) and shown that

$$\left\| \mathfrak{P}_e^{-1} f \right\|_{H^1_0(\mathcal{U})} \lesssim \| f \|_{H^{-1}_0(\mathcal{U})}.$$

However the complete operator $\mathfrak{P}$ does not have the necessary coercivity for inverting it in a similar manner. A first idea could be to *factorize* $\mathfrak{P}$ by $\mathfrak{P}_e$:

$$\mathfrak{P} = \mathfrak{P}_e \left(\mathrm{Id} + \mathfrak{P}_e^{-1} \mathfrak{B} \cdot \nabla + \mathfrak{P}_e^{-1} c \right)$$

and use compactness properties of $\mathfrak{K} := \mathfrak{P}_e^{-1} \mathfrak{B} \cdot \nabla + \mathfrak{P}_e^{-1} c$. This would require more regularity on the coefficients and we shall follow a simpler and more optimal approach.

6.6.2. *A simpler and more optimal approach.* The operator $\mathfrak{P}$ can be made coercive by a zero-order simple modification, and we can then factorize by the modified operator.

Proposition 6.4. *There is $\lambda_0 > 0$ so that $\mathfrak{P} + \lambda_0 \mathrm{Id}$ is coercive in the sense*

$$\forall \, v \in H^1_0(\mathcal{U}), \quad \langle (P + \lambda_0 \mathrm{Id}) v, v \rangle \geq \tilde{\lambda} \| v \|^2_{H^1(\mathcal{U})}$$

for some $\tilde{\lambda} > 0$.

Proof of Proposition 6.4. We perform the same a priori energy estimate

$$\begin{aligned} \langle (\mathfrak{P} + \lambda_0 \mathrm{Id}) v, v \rangle &\geq \lambda_e \| v \|^2_{\dot{H}^1(\mathcal{U})} - \langle B \cdot \nabla v, v \rangle - \langle c v, v \rangle + \lambda_0 \| v \|^2_{L^2(\mathcal{U})} \\ &\geq \lambda_e \| v \|^2_{\dot{H}^1(\mathcal{U})} - \| B \|_\infty \frac{\lambda_e}{2 \| B \|_\infty} \| v \|^2_{\dot{H}^1(\mathcal{U})} - \frac{\| B \|_\infty}{2 \lambda_e} \| v \|^2_{L^2(\mathcal{U})} \\ &\quad - \| c \|_\infty \| v \|^2_{L^2(\mathcal{U})} + \lambda_0 \| v \|^2_{L^2(\mathcal{U})} \\ &\geq \frac{\lambda_e}{2} \| v \|^2_{\dot{H}^1(\mathcal{U})} + \left(\lambda_0 - \frac{\| B \|_\infty}{2 \lambda_e} - \| c \|_\infty \right) \| v \|^2_{L^2(\mathcal{U})} \\ &\geq \frac{\lambda_e}{2} \| v \|^2_{\dot{H}^1(\mathcal{U})} + \bar{\lambda} \| v \|^2_{L^2(\mathcal{U})} \end{aligned}$$

for some $\bar{\lambda} > 0$, when choosing λ_0 large enough, which concludes the proof. ☐

Therefore from the previous subsection we can invert $(\mathfrak{P}+\lambda_0\mathrm{Id})$ with the Dirichlet conditions. We now write the following factorisation

$$\mathfrak{P} = (\mathfrak{P}+\lambda_0\mathrm{Id}) - \lambda_0\mathrm{Id} = (\mathfrak{P}+\lambda_0)\left[\mathrm{Id} - \lambda_0(\mathfrak{P}+\lambda_0\mathrm{Id})^{-1}\right] = (\mathfrak{P}+\lambda_0)\left[\mathrm{Id} - \mathfrak{K}\right]$$

where we denote $\mathfrak{K} := -\lambda_0(\mathfrak{P}+\lambda_0\mathrm{Id})^{-1}$.

We shall show later that the operator $\mathfrak{K}$ maps $L^2(\mathcal{U})$ to $H^1_0(\mathcal{U})$ and is a compact operator in $L^2(\mathcal{U})$, i.e. it maps the closed unit ball inside a compact set. Let us assume this for now and let us prove a criterion for the existence and uniqueness of solutions.

Observe that the equation is now *equivalent to finding* $u \in L^2(\mathcal{U})$ so that

$$u + \mathfrak{K}u = F, \quad F := (\mathfrak{P}+\lambda_0\mathrm{Id})^{-1} f \in H^1_0(\mathcal{U}).$$

Indeed any such $u \in L^2(\mathcal{U})$ will in fact be in $H^1_0(\mathcal{U})$ since both $\mathfrak{K}u$ and F belong to $H^1_0(\mathcal{U})$. Let us analyse t From the results in the previous section, we now that the operators of the form $\mathrm{Id}+\mathfrak{K}$ with $\mathfrak{K}$ compact have finite dimensional kernel and co-kernel, with equal dimension. It remains to prove the compactness of $\mathfrak{K} = -\lambda_0(P+\lambda_0\mathrm{Id})^{-1}$ on $L^2(\mathcal{U})$.

We consider a sequence $u_n \in L^2(\mathcal{U})$ with $\|u_n\|_{L^2(\mathcal{U})} \le 1$, $n \ge 0$. Let us denote $v_n := \mathfrak{K}u_n \in H^1_0(\mathcal{U})$. Then we have

$$(\mathfrak{P}+\lambda_0\mathrm{Id})v_n = -\lambda_0 u_n$$

and we can perform the following estimate, which takes advantage of the coercivity estimate on $(\mathfrak{P}+\lambda_0\mathrm{Id})$:

$$\|v_n\|_{H^1_0(\mathcal{U})} \lesssim \|u_n\|_{L^2(\mathcal{U})}.$$

It proves the compactness of the sequence v_n by the Rellich-Kondrachov theorem in the previous subsection.

By combining the previous results we obtain

Theorem 6.5 (Weak solutions to the general Dirichlet problem). *We assume that the second-order elliptic operator $\mathfrak{P}$ has coefficients a_{ij}, b_i, $c \in L^\infty(\mathcal{U})$ in divergence form, and satisfies the uniform ellipticity condition on $\mathcal{U}$. We consider the problem*

$$\begin{cases} \mathfrak{P}u = f & \text{in } \mathcal{U} \\ u = 0 & \text{on } \partial\mathcal{U}. \end{cases} \tag{6.5}$$

where $f \in H^{-1}_0(\mathbb{R}^\ell)$.

Then there are two exclusive possible cases:

(I) *either there is a **unique** weak solution $u \in H^1_0(\mathcal{U})$ to this elliptic problem ($\dim \mathrm{Ker}(\mathrm{Id}+\mathfrak{K}) = 0$ above),*

(II) *or else there is at least one non-zero weak solution $u \in H^1_0(\mathcal{U})$ to the* homogeneous *problem ($\dim \mathrm{Ker}(\mathrm{Id}+\mathfrak{K}) \neq 0$ above)*

$$\begin{cases} \mathfrak{P}u = 0 & \text{in } \mathcal{U} \\ u = 0 & \text{on } \partial\mathcal{U}. \end{cases} \tag{6.6}$$

Remark 6.6. *As said before one could consider the more general Dirichlet condition $u = g$ on $\partial\mathcal{U}$ where g is the restriction of an $H^1(\mathbb{R}^\ell)$ function to $\partial\mathcal{U}$.*

REMARK 6.7. *In the situation* **(II)** *the dimension of the subspace N of weak* homogeneous *solutions to* (6.6) *is the same as the dimension of the subspace N^* of weak homogeneous solutions to the adjoint problem*

$$\begin{cases} \mathfrak{P}^* u = 0 & \text{in } \mathcal{U} \\ u = 0 & \text{on } \partial\mathcal{U}. \end{cases}$$

REMARK 6.8. *In the situation* **(II)***, the necessary and sufficient condition for the existence of a solution to the original problem* (6.5) *is*

$$\forall\, v \in N^* \in H_0^1(\mathcal{U}), \quad \langle f, v \rangle = 0.$$

REMARK 6.9. *Observe finally that the reduction to the homogeneous problem is not simply a consequence of linearity due to the boundary conditions.*

7. Regularity study of elliptic equations

We now address the question as to whether a *weak* solution to $\mathfrak{P}u = f$ in $\mathcal{U}$ is smooth. This is the *regularity problem* for weak solutions, which we had already encountered for classical solutions in the whole space. We must distinguish between *interior* regularity and *boundary* regularity. We shall start with a warm-up on the key a priori estimate and some comments on the question of the approximation argument.

7.1. A priori estimate and approximation argument. Let us first consider the Poisson equation, and note that our energy estimate holds for weak solutions. For, choosing $v = u$ in the definition of a weak solution, we have

$$\|u\|^2_{\dot{H}^1(\mathcal{U})} = \langle\langle u, u \rangle\rangle \le \|f\|_{L^2(\mathcal{U})} \|u\|_{L^2(\mathcal{U})} \lesssim \|f\|_{L^2(\mathcal{U})} \|u\|_{\dot{H}^1(\mathcal{U})}$$

where we have used the Poincaré inequality, and therefore by dividing by the $\dot{H}^1(\mathcal{U})$ of u we get

$$\|u\|^2_{\dot{H}^1(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})}.$$

But if we remember the previous a priori estimate we used in the regularity study, it is reasonable to expect that, *away from boundary*, we would have the stronger estimate

$$\|u\|^2_{\dot{H}^2(\mathcal{V})} \lesssim \|f\|^2_{L^2(\mathcal{U})}$$

where $\mathcal{V}$ is an open set with $\overline{\mathcal{V}} \subset \mathcal{U}$. This leaves open the question of what happens at the boundary. However we shall see that the control of *tangential derivatives* combined with the differential relation provided by the PDE and the fact that $\partial\mathcal{U}$ is *non-characteristic* will allow to control the normal derivative. Then our previous estimate indeed *extends* up to the boundary, and we have

$$\|u\|^2_{\dot{H}^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})}.$$

Finally it is not hard to convince oneself formally that one should expect the same a priori estimate to hold at *any* order of regularity:

$$\forall\, s \in \mathbb{N}, \quad \|u\|^2_{\dot{H}^{s+2}(\mathcal{U})} \lesssim \|f\|^2_{\dot{H}^s(\mathcal{U})}.$$

Let us now consider the more general case of an elliptic equation in the form (6.5) with a second-order elliptic operator $\mathfrak{P}$ with coefficients a_{ij}, b_i, $c \in L^\infty(\mathcal{U})$ in divergence form, and which satisfies the uniform ellipticity condition on $\mathcal{U}$. We assume that we are in the situation **(I)** in the theorem 6.5 above, i.e. there is no non-zero solution to the homogeneous problem. Then the previous proof shows that the *solution map* $\mathcal{S} : L^2(\mathcal{U}) \to H^1_0(\mathcal{U})$ is well-defined, linear and continuous. In particular the $L^2(\mathcal{U})$ norm of u is controlled by that of f. This allows to "fix" the following incomplete a priori estimate that we have already seen:

$$\begin{aligned}\|u\|^2_{\dot H^1(\mathcal{U})} = \langle\langle u, u\rangle\rangle \le \langle \mathfrak{P} u, u\rangle + \|u\|^2_{L^2(\mathcal{U})} \\ \le C_{\mathcal{U}} \|f\|_{L^2(\mathcal{U})}\|u\|_{L^2(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})} \\ \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})}\end{aligned}$$

in the sense that the RHS in now fulling under control (in the last inequality we use the continuity of the solution map) and *one can use the estimate to gain further knowledge on the higher regularity of the solution.* In other words the inductive structure of the a priori estimate is still preserved.

Remark 7.1. *Actually the discussion on the last equation is slightly artificial since the solution map already provides the H^1_0 regularity. However the idea is important for higher regularity, and leads to the next a priori estimate.*

Going back the full elliptic regularity estimate again now, we expect that, *away from boundary*, we would have the stronger estimate

$$\|u\|^2_{\dot H^2(\mathcal{V})} \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|_{\dot H^1(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})}$$

where $\mathcal{V}$ is an open set with $\overline{\mathcal{V}} \subset \mathcal{U}$, and we have used the previous estimate in the last inequality. Then using the fact that the boundary is non-characteristic for an elliptic PDE we would then expect formally to extend it to

$$\|u\|^2_{\dot H^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})} \lesssim \|f\|^2_{L^2(\mathcal{U})}.$$

Finally it is again not hard to convince oneself formally that one should expect the same a priori estimate to hold at *any* order of regularity:

$$\forall\, s \in \mathbb{N}, \quad \|u\|^2_{\dot H^{s+2}(\mathcal{U})} \lesssim \|f\|^2_{H^s(\mathcal{U})} + \|u\|^2_{\dot H^s(\mathcal{U})} \lesssim \|f\|^2_{H^s(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})} \lesssim \|f\|^2_{H^s(\mathcal{U})}.$$

This "starting point", i.e. solving the equation in some "ground functional space" was provided by the careful study of the kernel of a Fredholm operator. Once obtained however the a priori estimate can now be used again, as it shows a *gain of two derivative* and can be used "inductively".

Now the question is again: how to transform these formal arguments into rigorous ones? Observe that the previous calculations do **not** constitute a proof, as they assume the regularity of u. We need again some *approximation argument.* The point is to derive analytic estimates from the structural, algebraic assumption of ellipticity.

We already, in the beginning of this chapter, presented a classical such approximation argument based on the convolution by a mollifier. The proofs in the rest of this section could be performed with this approximation, which is a good exercise. We shall present another interesting approximation argument, based on the *discretisation of the differentiation process.*

Definition 7.2. *We introduce the so-called* difference quotients*. For each coordinate i, and each $h \in \mathbb{R} \setminus \{0\}$, we define*

$$D_i^h u := \frac{u(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_\ell) - u(x_1, \dots, x_\ell)}{h}$$

Let us prove that (1) if $u \in H^s(\mathbb{R}^\ell)$, then so does $D_i^h u$, (2) conversely if $D_i^h u$ is uniformly bounded in $L^2(\mathbb{R}^\ell)$ as $h \to 0$ then $\partial_i u \in L^2(\mathbb{R}^\ell)$. And similarly for functions which are supported inside some open set $\mathcal{V}$ whose closure is contained in $\mathcal{U}$, as long as h is sufficiently small. Since classical derivatives commute with the difference quotients, it is enough to prove the result for $s = 0$.

Proposition 7.3 (Manipulation of the difference quotient). *First for $u \in H^1(\mathcal{U})$, there is $C_1 > 0$ so that*

$$\forall\, |h| \in (0, h_0), \quad \left\| D_i^h u \right\|_{L^2(\mathcal{V})} \le C_1 \|\partial_i u\|_{L^2(\mathcal{U})}$$

where $h_0 \in (0, dist(\overline{\mathcal{V}}, \partial\mathcal{U}))$ is so that $D_i^h u$ is well-defined on $\mathcal{V}$.

Second assume that for $u \in L^2(\mathcal{V})$ with support included in $\mathcal{V}$ there is a constant $C_2 > 0$ so that

$$\forall\, |h| \in (0, h_0), \quad \left\| D_i^h u \right\|_{L^2(\mathcal{U})} \le C_2,$$

then $u \in H^1(\mathcal{V})$ with

$$\|\partial_i u\|_{L^2(\mathcal{V})} \le C_2.$$

Proof of Proposition 7.3. We consider first smooth functions and then argue by density. The first part relies on the Taylor integral formula:

$$u(x + he_i) - u(x) = \int_0^1 \partial_i u(x + the_i) h \, \mathrm{d}t$$

and therefore

$$\int_{\mathcal{V}} \left| D_i^h \right|^2 \mathrm{d}x \lesssim |h_0| \int_0^1 \left(\int_{\mathcal{V}} |\partial_i u(x + te_i)|^2 \, \mathrm{d}x \right) \mathrm{d}t \lesssim \int_{\mathcal{U}} |\partial_i u(y)|^2 \, \mathrm{d}y.$$

For the second part, we first note the following *discrete integration-by-parts formula*: for $u \in L^2(\mathcal{V})$ with support included in $\mathcal{V}$ and $\varphi \in C_c^\infty(\mathcal{V})$

$$\int_{\mathcal{V}} u(x) \left[\frac{\varphi(x + he_i) - \varphi(x)}{h} \right] \mathrm{d}x = - \int_{\mathcal{V}} \left[\frac{u(x - he_i) - u(x)}{h} \right] \varphi(x) \, \mathrm{d}x$$

which means

$$\int_{\mathcal{V}} u(x) D_i^h \varphi(x) \, \mathrm{d}x = - \int_{\mathcal{V}} D_i^{-h} u(x) \varphi(x) \, \mathrm{d}x$$

The assumption means that

$$\sup_{|h| \in (0, h_0)} \left\| D_i^{-h} u \right\|_{L^2(\mathcal{V})} < \infty$$

and therefore (weak compactness of the unit ball) there is a subsequence $h_k \to 0$ so that $D_i^{-h_k} u \rightharpoonup v_i$ for some $v_i \in L^2(\mathcal{V})$ with

$$\|v_i\|_{L^2(\mathcal{V})} \le \sup_{|h| \in (0, h_0)} \left\| D_i^{-h} u \right\|_{L^2(\mathcal{V})}.$$

But then if $\varphi \in C_c^\infty(\mathcal{V})$ we obviously have $D_i^h \varphi \to \partial_i \varphi$ with uniform (and L^2!) convergence, and therefore

$$\int_{\mathcal{V}} u \partial_i \varphi \, \mathrm{d}x = \int_{\mathcal{V}} u \partial_i \varphi \, \mathrm{d}x = \lim_{h_k \to 0} \int_{\mathcal{V}} u D_i^{h_k} \varphi \, \mathrm{d}x.$$

But the RHS is also

$$\int_{\mathcal{V}} u D_i^{h_k} \varphi \, \mathrm{d}x = - \int_{\mathcal{V}} D_i^{-h_k} u \varphi \, \mathrm{d}x \xrightarrow{h_k \to 0} - \int_{\mathcal{V}} v_i \varphi \, \mathrm{d}x$$

from the weak convergence. This implies that

$$\forall \, \varphi \in C_c^\infty(\mathcal{V}), \quad \int_{\mathcal{V}} u \partial_i \varphi \, \mathrm{d}x = - \int_{\mathcal{U}} v_i \varphi \, \mathrm{d}x.$$

This shows from the definition of generalised derivatives that the generalised derivative $\partial_i u = v_i$ exists and is $L^2(\mathcal{V})$, which concludes the proof. $\square$

7.2. Interior regularity. We first consider the case $s = 0$ in the gain of regularity, with $f \in L^2(\mathcal{U})$ and the usual other assumptions, plus $a_{ij}, b_i, c \in C^1(\mathcal{U}) \cap L^\infty(\mathcal{U})$. Consider the situation **(I)** and a solution u to the elliptic problem $\mathfrak{P} u = f$ in $H^1(\mathcal{U})$.

Let us show that $u \in H^2_{loc}(\mathcal{U})$. This is the space of locally H^2 functions, i.e. that are H^2 on any compact included in $\mathcal{U}$. Therefore we want to show that for any open subset $\mathcal{V} \subset\subset \mathcal{U}$ (i.e. $\overline{\mathcal{V}} \subset \mathcal{U}$) we have the estimate

$$\|u\|^2_{H^2(\mathcal{V})} \le C \left(\|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{L^2(\mathcal{U})} \right)$$

where the constant $C > 0$ depends on $\mathcal{U}$, $\mathcal{V}$, and the coefficients of L.

Remark 7.4. *The boundary conditions are not required in the proof here, as we shall stay away from the boundary, i.e. we do not require $u \in H^1_0(\mathcal{U})$ here but only $u \in H^1(\mathcal{U})$ and solves the PDE in weak sense: it is enough that u is* a *(we do not care about uniqueness here) solution to $\mathfrak{P} u = f$ in the open set $\mathcal{U}$ in weak sense, and that $a_{ij}, b_i, c \in C^1(\mathcal{U})$ (no $L^\infty(\mathcal{U})$ is required on the coefficients in the regularity study of the interior, unless one is interested in how the estimates "degenerate" at the boundary; similarly for higher regularity we will only need $a_{ij}, b_i, c \in C^k(\mathcal{U}) \dots$).*

Remark 7.5. *Observe also that it allows to make sense of the PDE in the almost everywhere sense, as it shows that in the open set $\mathcal{U}$, a distributional second-order derivative exists in $L^2_{loc}(\mathcal{U})$, and by integration by parts one can recover $\mathfrak{P} u = f$ almost everywhere from the weak form.*

Choose an open set $\mathcal{W}$ with $\mathcal{V} \subset\subset \mathcal{W} \subset\subset \mathcal{U}$ (the symbol $\subset\subset$ means "included in a compact set included in"). Define a localisation smooth function $\zeta \in C_c^\infty(\mathcal{V})$ with support in $\mathcal{U}$ and which is one on $\mathcal{V}$ and zero outside $\mathcal{W} \subset\subset \mathcal{U}$, and the following test function

$$v := -D_i^{-h}(\zeta^2 D_i^h u)$$

for any index i. We then use the weak formulation

$$\int_{\mathcal{U}} (\nabla u)^T A(x) (\nabla v) \, \mathrm{d}x = \int_{\mathcal{U}} \tilde{f} v \, \mathrm{d}x, \quad \tilde{f}(x) := f(x) - \nabla \cdot (B(x) u) - c(x) u(x).$$

From the construction of weak solution in $H_0^1(\mathcal{U})$ and the assumption on the coefficients b_i, c, we know that $\tilde{f} \in L^2(\mathcal{V})$. We know calculate for the LHS

$$\begin{aligned}
\int_{\mathcal{U}} (\nabla u)^T A(x)(\nabla v)\,\mathrm{d}x &= -\int_{\mathcal{U}} (\nabla u)^T A(x)(\nabla D_i^{-h}(\zeta^2 D_i^h u))\,\mathrm{d}x \\
&= -\int_{\mathcal{U}} (\nabla u)^T A(x)(D_i^{-h}\nabla(\zeta^2 D_i^h u))\,\mathrm{d}x \\
= \int_{\mathcal{U}} (D_i^h \nabla u)^T A(x)(\nabla(\zeta^2 D_i^h u))\,\mathrm{d}x &+ \int_{\mathcal{U}} (\nabla u)^T D_i^h A(x)(\nabla(\zeta^2 D_i^h u))\,\mathrm{d}x \\
= \int_{\mathcal{U}} \zeta^2 (D_i^h \nabla u)^T A(x)(\nabla(D_i^h u))\,\mathrm{d}x &+ \int_{\mathcal{U}} 2(D_i^h \nabla u)^T A(x)(\nabla\zeta)\zeta(D_i^h u)\,\mathrm{d}x \\
+ \int_{\mathcal{U}} \zeta^2 (\nabla u)^T D_i^h A(x)(\nabla(D_i^h u))\,\mathrm{d}x &+ 2\int_{\mathcal{U}} (\nabla u)^T D_i^h A(x)(\nabla\zeta)\zeta D_i^h u\,\mathrm{d}x \\
&=: I_1 + I_2 + I_3 + I_4
\end{aligned}$$

where we have used the commutation $\nabla D_i^h = D_i^h \nabla$, and the distributivity $D_i^h(uv) = (D_i^h u)v + u(D_i^h v)$.

We calculate for the first term, using the ellipticity assumption,

$$I_1 \geq \lambda_e \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \mathrm{d}x.$$

We calculate for the second term, using Cauchy-Schwarz' inequality and the bounds on ζ,

$$\begin{aligned}
I_2 &\geq -\varepsilon \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \mathrm{d}x - \frac{C}{\varepsilon} \int_{\mathcal{U}} |\nabla\zeta|^2 \left|D_i^h u\right|^2 \mathrm{d}x \\
&\geq -\varepsilon \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \mathrm{d}x - \frac{C}{\varepsilon} \|u\|_{H^1(\mathcal{W})}^2
\end{aligned}$$

for any $\varepsilon > 0$ and some constant $C > 0$. We calculate similarly for the third term, using Cauchy-Schwarz' inequality,

$$\begin{aligned}
I_3 &\geq -\varepsilon \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \mathrm{d}x - \frac{C}{\varepsilon} \int_{\mathcal{U}} \zeta^2 \left|D_i^h u\right|^2 \mathrm{d}x \\
&\geq -\varepsilon \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \mathrm{d}x - \frac{C}{\varepsilon} \|u\|_{H^1(\mathcal{W})}^2.
\end{aligned}$$

Finally we calculate for the last term, using the bounds on ζ,

$$I_4 \geq -C\|u\|_{H^1(\mathcal{W})}^2.$$

Choosing $\varepsilon = \lambda_e/3$ we deduce that

$$I_1 + I_2 + I_3 + I_4 \geq \frac{\lambda_e}{3} \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \mathrm{d}x - C' \|u\|_{H^1(\mathcal{W})}^2$$

for some other constant $C' > 0$.

Going back to the weak formulation, we now estimate the RHS:

$$\int_{\mathcal{U}} \tilde{f} v\,\mathrm{d}x \lesssim \left\|\tilde{f}\right\|_{L^2(\mathcal{W})} \|v\|_{L^2(\mathcal{W})}$$

$$\leq C_\varepsilon \|f\|^2_{L^2(\mathcal{W})} + C_\varepsilon \|u\|_{H^1(\mathcal{W})} + \varepsilon \int_{\mathcal{W}} \zeta^2 \left| D_i^{-h} D_i^h u \right|^2 \mathrm{d}x$$

for any $\varepsilon > 0$ and some corresponding constant $C_\varepsilon > 0$. We deduce that

$$\begin{aligned}\int_{\mathcal{V}} \left| D_i^h \nabla u \right|^2 \mathrm{d}x &\leq \int_{\mathcal{U}} \zeta^2 \left| D_i^h \nabla u \right|^2 \mathrm{d}x \\ &\leq C'_\varepsilon \|f\|^2_{L^2(\mathcal{W})} + C'_\varepsilon \|u\|^2_{H^1(\mathcal{W})} + \varepsilon \int_{\mathcal{W}} \zeta^2 \left| D_i^{-h} D_i^h u \right|^2 \mathrm{d}x.\end{aligned}$$

Finally using the property of the difference quotient we deduce

$$\|\partial_i \nabla u\|^2_{L^2(\mathcal{V})} \leq C'_\varepsilon \|f\|^2_{L^2(\mathcal{W})} + C'_\varepsilon \|u\|^2_{H^1(\mathcal{W})} + \varepsilon \left\|\nabla^2 u\right\|_{L^2(\mathcal{W})}.$$

and since it is true for any $i = 1, \dots, \ell$,

$$\left\|\nabla^2 u\right\|^2_{L^2(\mathcal{V})} \leq C''_\varepsilon \|f\|^2_{L^2(\mathcal{W})} + C''_\varepsilon \|u\|^2_{H^1(\mathcal{W})} + \varepsilon \left\|\nabla^2 u\right\|_{L^2(\mathcal{W})}.$$

By taking ε small enough it concludes the proof of

$$\left\|\nabla^2 u\right\|^2_{L^2(\mathcal{V})} \lesssim \|f\|^2_{L^2(\mathcal{W})} + \|u\|^2_{H^1(\mathcal{W})}.$$

Let us now explain how to prove *higher-order interior regularity estimates.* We argue by induction with the induction assumption

$$(\mathscr{H}_m) \qquad \left\{ \begin{array}{l} a_{ij}, b_i, c \in C^{m+1}(\mathcal{U}) \\[1ex] f \in H^m_{loc}(\mathcal{U}) \\[1ex] \mathfrak{P} u = f \quad \text{in any} \quad \mathcal{W} \subset\subset \mathcal{U} \end{array} \right\} \quad \Longrightarrow \quad u \in H^{m+2}_{loc}(\mathcal{U})$$

on $m \in \mathbb{N}$.

The initialisation of the induction $(\mathscr{H}_0)$ was proven above. Assume $(\mathscr{H}_m)$ holds with $m \geq 1$, and assume that $a_{ij}, b_i, c \in C^{m+2}_B(\mathcal{U})$ and $f \in H^{m+1}(\mathcal{U})$. Let us prove that $u \in H^{m+3}_{loc}(\mathcal{U})$. Consider any first order derivative ∂_i and estimate the weak formulation with $-\partial_i \tilde{v}$, $\tilde{v} \in C^\infty_c(\mathcal{U})$ (with support included in $\mathcal{W} \subset\subset \mathcal{U}$).

Let us recall the weak formulation

$$\forall\, v \in H^1_0(\mathcal{U}), \quad \mathfrak{B}(u, v) = \mathfrak{L} v$$

with

$$\left\{ \begin{array}{l} \mathfrak{B}(u,v) = \displaystyle\int_{\mathcal{U}} (\nabla u)^T A(x) (\nabla v) \,\mathrm{d}x + \int_{\mathcal{U}} v B(x) \cdot \nabla u \,\mathrm{d}x + \int_{\mathcal{U}} c u v \,\mathrm{d}x, \\[2ex] \mathfrak{L} v = \langle f, v \rangle. \end{array} \right.$$

and let us also recall that is enough to check this formulation for $v \in C^\infty_c(\mathcal{U})$ by approximation of $H^1_0(\mathcal{U})$ functions.

An easy calculation by integration by parts shows that

$$\begin{aligned}\int_{\mathcal{U}} (\nabla \partial_i u)^T A(x) (\nabla \tilde{v}) \,\mathrm{d}x + \int_{\mathcal{U}} \tilde{v} B(x) \cdot \nabla \partial_i u \,\mathrm{d}x + \int_{\mathcal{U}} c \partial_i u \tilde{v} \,\mathrm{d}x \\ = \int_{\mathcal{U}} \left(\partial_i f - \partial_i c u - \partial_i B(x) \cdot \nabla u \right) \tilde{v} \,\mathrm{d}x\end{aligned}$$

which means that $\tilde{u} = \partial_i u$ satisfies

$$\mathfrak{B}(\tilde{u}, \tilde{v}) = \tilde{\mathfrak{L}}\tilde{v}, \quad \tilde{\mathfrak{L}}w := \int_{\mathcal{U}} \tilde{f} w \, \mathrm{d}x, \quad \tilde{f} := \partial_i f - \partial_i c u - \partial_i B(x) \cdot \nabla u.$$

Since $m \geq 1$, we know that $u \in H^3(\mathcal{W})$, and therefore the equation translates into

$$\mathfrak{P}\tilde{u} = \tilde{f}$$

in weak sense on $\mathcal{W}$. Since $\tilde{f} \in H^m(\mathcal{W})$ we deduce from the previous study that $\tilde{u} = \partial_i u \in H^{m+2}(\mathcal{W})$. Since it is true for any first-order partial derivative ∂_i and any $\mathcal{W} \subset\subset \mathcal{U}$ we deduce that $u \in H^{m+3}_{loc}(\mathcal{U})$. This concludes the proof of the induction.

We deduce from this subsection the following property, similar to that we have proved for classical solutions in the whole plane: if $a_{ij}, b_i, c, f \in C^\infty(\mathcal{U})$ and $u \in H^1_0(\mathcal{U})$ is a weak solution to $Pu = f$, with the uniform ellipticity assumption, then $u \in C^\infty(\mathcal{U})$. Observe that it says nothing so far about the possible singularities of u at the boundary, only that any such possible singularities do not propagate into the interior.

7.3. Boundary regularity. We first consider again the gain of $H^2(\mathcal{U})$ regularity, but in the whole domain, i.e. including the boundary. We consider $a_{ij}, b_i, c \in C^1(\overline{\mathcal{U}})$, and assume that $\partial\mathcal{U}$ is C^2. As became clear from the previous subsection, the regularity study is an *a priori* study. This notion is simple but extremely important: this means that the argument applies to any solutions that satisfies a minimal set of assumptions. To be concrete we proved the interior regularity as soon as $a_{ij}, b_i, c, f \in C^\infty(\mathcal{U})$ and $u \in H^1(\mathcal{U})$ solves in the weak sense $\mathfrak{P}u = f$ in the open set $\mathcal{U}$.

In the case of the boundary regularity we now need to impose, on the contrary to the previous case, some boundary conditions on u, and we shall therefore assume that $u \in H^1_0(\mathcal{U})$. We also need now some estimates on the coefficients and the forcing term f that are uniform when approaching the boundary. Recall that if, *furthermore*, we assume that u is *the* solution to the Dirichlet problem given by the situation **(I)** in the study we made before, then we have the *additional* information

$$\|u\|_{H^1_0(\mathcal{U})} \lesssim \|f\|_{L^2(\mathcal{U})}.$$

We now want to show that $u \in H^2(\mathcal{U})$ with the precise estimate

$$\|u\|_{H^2(\mathcal{U})} \lesssim \|f\|_{L^2(\mathcal{U})} + \|u\|_{H^1(\mathcal{U})}.$$

From the previous interior regularity study it is clear that it is enough to prove the result in neighborhoods of the boundary.

We first consider the case where $\mathcal{U} = B(0,1) \cap \mathbb{R}^\ell_+$ is a *half-ball*, with $\mathbb{R}^\ell_+ = \{x_\ell \geq 0\}$. Consider the subset $\mathcal{V} = B(0,1/2) \cap \mathbb{R}^\ell_+$ and some smooth localisation function ζ that is 1 on $\mathcal{V}$ and zero outside $B(0,1)$, with $0 \leq \zeta \leq 1$.

For any coordinate $i \in \{1, \dots, \ell-1\}$ that is tangential to the boundary, we consider the test function $v := -D_i^{-h}(\zeta^2 D_i^h u)$, which belongs to $H^1_0(\mathcal{U})$. Indeed heuristically

$$v(x) = -\frac{1}{h^2}\left(\zeta^2(x - he_i)(u(x) - u(x - he_k)) - \zeta^2(x)(u(x + he_k) - u(x))\right).$$

and the RHS cancels at $x_\ell = 0$. Rigorously we must prove that there is $v_n \to v$ with $v_n \in C^\infty_c(\mathcal{U})$: indeed a sequence $u_n \to u$, $u_n \in C^\infty_c(\mathcal{U})$ certainly exists for $u \in H^1_0(\mathcal{U})$, and then one checks that

$$v_n(x) := -\frac{1}{h^2}\left(\zeta^2(x-he_i)(u_n(x)-u_n(x-he_k)) - \zeta^2(x)(u_n(x+he_k)-u_n(x))\right)$$

belongs to $C_c^\infty(\mathcal{U})$ with $v_n \to v$ in $H^1(\mathcal{U})$.

Then reproducing the same key calculation performed for the interior regularity (relying on the uniform ellipticity estimate and the bound on the coefficients) we easily get

$$\int_{\mathcal{V}} \left|D_i^h \nabla u\right|^2 \, \mathrm{d}x \le \int_{\mathcal{U}} \zeta^2 \left|D_i^h \nabla u\right|^2 \, \mathrm{d}x \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{H^1(\mathcal{U})}$$

(observe that here we use bounds on the coefficients up to the boundary) which shows that

$$\forall\, i \in \{1, \dots, \ell-1\}, \quad \|\partial_i \nabla u\|^2_{L^2(\mathcal{V})} \lesssim \|f\|^2_{L^2(\mathcal{U})} + \|u\|^2_{H^1(\mathcal{U})}.$$

We still need to estimate the derivative along the last coordinate ∂_ℓ. But then observe that the uniform ellipticity $A(x) \ge \lambda_e > 0$ on the symmetric matrix $A(x)$ implies that (consider the quadratic form at the vector e_ℓ) that $a_{\ell\ell} \ge \lambda_e > 0$. Hence using the following equality on $\mathcal{V}$

$$a_{\ell\ell}\partial^2_{\ell\ell} u = -\sum_{j=1}^{\ell-1}\sum_{i=1}^{\ell} \partial_j(a_{ij}(x)\partial_i u(x)) - (\partial_\ell a_{\ell\ell})\partial_\ell u - B(x)\cdot\nabla u - cu$$

between $L^2(\mathcal{V})$ functions, we deduce that $\partial^2_{\ell\ell} u \in L^2(\mathcal{V})$ (using the bounds on the coefficients), which finally shows that

$$\|u\|_{H^2(\mathcal{V})} \lesssim \|f\|_{L^2(\mathcal{U})} + \|u\|_{H^1(\mathcal{U})}.$$

Let us consider the case of a general boundary shape. For the sake of simplicity we now rewrite the PDE in *non-divergence* form

$$\sum_{i,j=1}^{\ell} a_{ij}(x)\partial^2_{ij} u + \sum_{i=1}^{\ell} b_i(x)\partial_i u + cu = f,$$

which changes only the coefficients b_i, c, still denoted with the same name by a slight abuse of notation. Consider a base point $x_0 \in \partial\mathcal{U}$ on the boundary. Using the regularity assumption on the boundary we have (for r small enough so that any sphere with radius with $0 < r' \le r$ intersects $\partial\mathcal{U}$ as a connected $\ell-2$-hypersurface)

$$\mathcal{U} \cap B(x_0, r) = \{x_\ell > \varphi(x_1, \dots, x_{\ell-1})\} \cap B(x_0, r)$$

where $\varphi : \mathbb{R}^{\ell-1} \to \mathbb{R}$ is C^2. We now perform the same kind of change of variable as in the proof of the Cauchy-Kovalevskaya theorem (however with less regularity): $\Phi \in C^2$ maps $\mathcal{U} \cap B(x_0, r)$ to

$$\tilde{\mathcal{U}}_{x_0} := \{y \in B(0, \tilde{r}) \ : \ y_\ell > 0\}$$

where furthermore the normal vectors to $\partial\mathcal{U}$ are mapped to e_ℓ, and we write Ψ for the inverse of Φ. This maps the PDE to a new PDE

$$\tilde{\mathfrak{P}}\tilde{u} = \tilde{f}$$

for a second-order linear operator $\tilde{\mathfrak{P}}$ of the same form, with $\tilde{u}(y) = u(\Psi(y))$, $\tilde{f} = f(\Psi(y))$, $\tilde{c}(y) = c(\Psi(y))$, and (using the chain rule)

$$\begin{cases} \tilde{a}_{kl}(y) := \displaystyle\sum_{i,j=1}^{\ell} a_{ij}(\Psi(y))(\partial_i \Phi)_k(\Psi(y))(\partial_j \Phi)_l(\Psi(y)), \\ \\ \tilde{b}_k(y) := \displaystyle\sum_{i=1}^{\ell} b_i(\Psi(y))(\partial_i \Phi)_k(\Psi(y)). \end{cases}$$

Exercise 43. *Check the previous transformation and prove that it preserves the regularity and pointwise bounds from above on the coefficients.*

Let us now prove that the new operator $\tilde{\mathfrak{P}}$ is uniformly elliptic. Consider any point $y \in \tilde{\mathcal{U}}_{x_0}$, and any $\xi \in \mathbb{R}^\ell$ and calculate

$$\begin{aligned} \tilde{\sigma}_p(x,\xi) &= \sum_{k,l=1}^{\ell} \tilde{a}_{kl}(y)\xi_k\xi_l = \sum_{k,l=1}^{\ell}\sum_{i,j=1}^{\ell} a_{ij}(\Psi(y))(\partial_i\Phi)_k(\Psi(y))\xi_k(\partial_j\Phi)_l(\Psi(y))\xi_l \\ &= \sum_{i,j=1}^{\ell} a_{ij}(\Psi(y))(\Xi D\Phi(\Psi(y)))_i(\Xi D\Phi(\Psi(y)))_j \end{aligned}$$

where we denoted $D\Phi(x) = ((\partial_i\Phi)_j)_{j,i}$ the $\ell \times \ell$ usual jacobian matrix, and $\Xi = (\xi_1, \dots, \xi_\ell)$. Let us denote $\Theta = \Xi D\Phi(\Psi(y))$. This relation can be inverted as $\Xi = \Theta D\Psi(y)$. Now using the uniform bounds on the differential we have $|\Theta|^2 \ge C_{\mathcal{V}',\Phi}|\Xi|^2$ and we deduce

$$\tilde{\sigma}_p(x,\xi) \ge \lambda_e|\Theta|^2 \ge \lambda_e C_{\mathcal{V}',\Phi}|\Xi|^2$$

which allows to apply the previous study of the half-ball and concludes the proof of the regularity in $\mathcal{V}$ by using the transformation Ψ.

Finally since $\partial\mathcal{U}$ is compact we can use a finite covering by small enough balls, and taking the worst constant, it concludes the proof.

Exercise 44. *Prove the higher-order boundary regularity statement: for $a_{ij}, b_i, c \in C^{m+1}(\overline{\mathcal{U}})$, a C^{m+2} boundary $\partial\mathcal{U}$ and $f \in H^m(\mathcal{U})$, any weak solution $u \in H^1_0(\mathcal{U})$ to $Pu = f$ is $H^{m+2}(\mathcal{U})$ with*

$$\|u\|_{H^{m+2}(\mathcal{U})} \lesssim \|f\|_{H^m(\mathcal{U})} + \|u\|_{L^2(\mathcal{U})}.$$

8. Maximum principles for elliptic equations

We now see a last fundamental property of elliptic (and parabolic!) equations, that of maximum principles. The idea is now to quantify in a *pointwise* (instead of *integral* as in the energy estimates) manner the "negative" aspect of the operator. To be more precise we want to exploit the following basis remark: if $u \in C^2$ attains a maximum $x_0 \in \mathcal{U}$ in the interior of the domain of definition of u, then $\nabla u(x_0) = 0$ and $\nabla^2 u(x_0) \le 0$ (i.e. the symmetric matrix is matrix $\nabla^2 u$ is non-positive at x_0). It is clear that we will need then that u is a classical solution in C^2 to exploit this idea. We know however from the previous study than any weak solution is indeed in C^2 for a right-hand side f regular enough.

We consider a second-order differential operator in *non-divergence* form

$$Pu = -\sum_{i,j=1}^{\ell} a_{ij}(x)\partial^2_{ij}u(x) + B(x)\cdot\nabla u(x) + c(x)u(x)$$

and we assume the uniform ellipticity condition, as well as $a_{ij}, b_i, c \in C^0(\overline{\mathcal{U}})$ (and the symmetry w.l.o.g. of $A = (a_{ij})$).

8.1. Weak maximum principle with no zero-order term.

PROPOSITION 8.1. *Let us first assume that $c = 0$. Then the weak version of the maximum principle states that for any $u \in C^2(\mathcal{U}) \cap C^0(\overline{\mathcal{U}})$ so that $Pu \leq 0$ (resp. $Pu \geq 0$) in $\mathcal{U}$, then the maximum (resp. minimum) of u on $\overline{\mathcal{U}}$ is attained at the boundary $\partial\mathcal{U}$.*

REMARK 8.2. *Such functions are called* subsolutions *(resp.* supersolutions*), just like for ordinary differential inequalities.*

Let us prove this property. We obviously only need to study subsolutions thanks to the transformation $u \to -u$. We shall reduce to the case where $Pu < 0$ on $\mathcal{U}$ by an important *lifting argument* (see later). So let us now assume that $u \in C^2(\mathcal{U}) \cap C^0(\overline{\mathcal{U}})$ with $Pu < 0$ on $\mathcal{U}$.

In a first stage, to see the naked idea of the proof, simplify a bit further and assume that $P = -\Delta + B(x)\cdot\nabla$. Then argue by contradiction: assume there is a point $x_0 \in \mathcal{U}$ so that u is maximum at x_0. Basic calculus of variation then shows that $\Delta u(x_0) \leq 0$ and $\nabla u(x_0) = 0$ which implies that $Pu(x_0) \geq 0$ and therefore a contradiction. In the general setting the matrix $A(x_0)$ can be diagonalised by some orthogonal $\ell \times \ell$-matrix Q:

$$A(x_0) = Q^T \mathrm{Diag}(\lambda_1, \dots, \lambda_\ell) Q$$

with $\lambda_i \geq \lambda_e$ for any $i \in \{1, \dots, \ell\}$. Define the new variable $y = x_0 + Q(x - x_0)$ by a "rotation" around x_0. Then at x_0:

$$\sum_{i,j=1}^{\ell} a_{ij}(x)\partial^2_{x_i x_j}u(x_0)$$
$$= \sum_{i,j=1}^{\ell}\sum_{k,l=1}^{\ell} a_{ij}(x) Q_{ki} Q_{lj} \partial^2_{y_k y_l} u(x_0) = \sum_{k=1}^{\ell} \lambda_k \partial^2_{y_k y_k} u(x_0) \leq 0$$

and we deduce again that $Pu(x_0) \geq 0$, which contradicts the assumption.

Let us now consider the *lifting argument.* This is a further instance of approximation: the key thing is that the property we are searching for is *stable* by uniform convergence approximation, whereas the closure of our assumption $Pu < 0$ under this convergence includes $Pu \leq 0$.

More precisely we define $u_\varepsilon(x) = u(x) + \epsilon e^{\kappa x_1}$ with $\epsilon > 0$ meant to be small, and κ meant to be large. Then prove that $Pu_\epsilon < 0$ for κ large enough:

$$-\sum_{i,j=1}^{\ell} a_{ij}(x)\partial^2_{ij}u_\epsilon(x) + \sum_{i=1}^{\ell} b_i(x)\partial_i u_\epsilon(x)$$

$$= Pu(x) + \epsilon\kappa e^{\kappa x_1}\left(-\kappa a_{11}(x) + b_1(x)\right) \leq \epsilon\kappa e^{\kappa x_1}(-\kappa\lambda_e + \|b_1\|_\infty)$$

which yields the required control by choosing $\kappa \geq 2\|b_1\|_\infty/\lambda_e$.

Then we deduce, using the previous step, that

$$\forall\, \epsilon > 0, \quad \max_{\overline{\mathcal{U}}} u_\epsilon = \max_{\partial\mathcal{U}} u_\epsilon$$

and pass to the limit $\varepsilon \to 0$ to conclude the proof.

8.2. Weak maximum principle with a signed zero-order term. Now let us consider the case where we include a zero-order term c in the formula for P. In order to fit with the idea that the operator should push "up" (resp. down) when at a maximum (resp. minimum), we assume that $c \geq 0$, as the maximum (resp. minimum) will only be considered for non-negative (resp. non-positive) values.

Proposition 8.3. *Assume $Pu \leq 0$ (resp. $Pu \geq 0$) and $c \geq 0$ in $\mathcal{U}$, then*

$$\max_{\overline{\mathcal{U}}} u \leq \max_{\partial\mathcal{U}} u_+ \qquad \left(\textit{resp.} \quad \min_{\overline{\mathcal{U}}} u \geq -\max_{\partial\mathcal{U}} u_- \right)$$

where $u_+ \geq 0$ (resp. $u_- \geq 0$) denotes the non-negative (resp. non-positive) part of u.

The proof is straightforward by considering the region where $u \geq 0$. Consider w.l.o.g. the case of a subsolution $Pu \leq 0$ and define the open set $\mathcal{V} \subset \mathcal{U}$ by

$$\mathcal{V} := \{x \in \mathcal{U} \ : \ u(x) > 0\}.$$

If $\mathcal{V} = \emptyset$ we are done, else observe that in $\mathcal{V}$ we have $\tilde{P}u := Pu - cu \leq 0$ and $\tilde{P}$ has no zero-order term. We can therefore apply the previous result to conclude.