Partial Differential Equations

WiSe 2021/22

Lecturer: Prof. Robert Haller-Dintelmann

Set in LaTeX by Jens Biermann
based on the lecture notes for the course "Partial Differential Equations" in 2018/19 written by Jonas Tibke, Johannes Becker, Carsten Litzinger and Jens Biermann

Version of January 7, 2022

Contents

1 Introduction

From the theory of ordinary differential equations we are familiar with equations of the form

$$y''(t) + y(t) = \sin(t).$$

In contrast to an ordinary differential equation, the terms in a partial differential equation may contain terms that depend on derivatives with respect to several variables, for example

$$u_{xx}(x, y) + u_{yy}(x, y) = \cos(xy).$$

We will see that the theory of PDEs is much more complicated than the theory of ODEs. The strength of the theory for ODEs rests in big part on the availability of the fundamental theorem of calculus, which allows the reformulation of a differential equation as an integral equation:

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \iff y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s.$$

In the case of PDEs this approach breaks down and there is no idea how to circumvent this.

Example 1.1 (Modelisation of the heat equation). Let $\Omega \subseteq \mathbb{R}^3$ be filled with some material and $u\colon \mathbb{R}^+ \times \Omega \to \mathbb{R}$ be the function that maps (t, x) to the temperature at $x \in \Omega$ at time $t > 0$. Furthermore, $u_0(x)$ denotes the initial temperature.

The physical description of a heat flow in such a material depends on material parameters such as the density ρ, the specific heat capacity c, and the heat conductivity k. For our model, we will assume that all of these parameters are constant. This is not realistic in general but simplifies our model and gives good approximations in cases where these quantities do not depend too much on the temperature.

Consider the energy balance in some ball $B \subseteq \Omega$. By conservation of energy, the temporal change of thermal energy is the sum of the heat flux $F(t, x)$ through ∂B and the gain/loss by sources/sinks. We observe for the thermal energy E in B that

$$E = \int_B \rho \cdot c \cdot u(t, x)\,\mathrm{d}x.$$

Thus we get for its temporal change

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_B \rho \cdot c \cdot u(t, x)\,\mathrm{d}x = -\int_{\partial B} F(t, x) \cdot \nu(x)\,\mathrm{d}\sigma(x) + \int_B q(t, x)\,\mathrm{d}x,$$

where q models the gain/loss of heat in B and ν denotes the outer unit normal vector to the boundary of B.

As an explanation for the minus sign in front of the boundary integral note:

$$F \cdot \nu > 0\colon \text{loss of heat}$$
$$F \cdot \nu < 0\colon \text{gain of heat}$$

In order to continue, we have to know in which way F depends on u. The easiest model is *Fourier's Law*, i.e. $F = -k \cdot \nabla_x u(t,x)$. Inserting this simplifies our equation to

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_B \rho \cdot c \cdot u(t,x) \, \mathrm{d}x = \int_{\partial B} k \cdot \nabla_x u(x,t) \cdot \nu(x) \, \mathrm{d}\sigma(x) + \int_B q(t,x) \, \mathrm{d}x.$$

By Gauß Theorem this equals

$$= \int_B k \operatorname{div}_x \nabla_x u(t,x) \, \mathrm{d}x + \int_B q(t,x) \, \mathrm{d}x,$$

so we obtain

$$\int_B \left(\rho \cdot c \cdot \frac{\partial}{\partial t} u(t,x) - k \, \Delta_x \, u(t,x) - q(t,x) \right) \mathrm{d}x = 0$$

for all balls $B \subseteq \Omega$. Thus,

$$\rho c \frac{\partial}{\partial t} u(t,x) - k \, \Delta_x \, u(t,x) = q(t,x), \qquad (t,x) \in (0,\infty) \times \Omega,$$

which we call the inhomogeneous heat or diffusion equation. In the case $f \equiv 0$, it is called the homogeneous heat/diffusion equation.

In mathematics, constants are often assumed to be 1[1] and we drop the units, which yields the condensed usual formulation of the heat equation:

$$\boxed{\frac{\partial}{\partial t} u - \Delta \, u = f.}$$

We now establish the general notion of a PDE.

Notation 1.2. For $k \in \mathbb{N}$ and $u \in \mathrm{C}^k(\mathbb{R}^d)$, we use the notation

$$\mathrm{D}^k \, u(x) = (\mathrm{D}^\alpha \, u(x))_{|\alpha|=k}$$

for the vector of all derivatives up to order k.

Definition 1.3. Let $\Omega \subseteq \mathbb{R}^d$ open and $k \in \mathbb{N}$. Then

a) a PDE of order k is given by

$$F(\mathrm{D}^k \, u(x), \mathrm{D}^{k-1} \, u(x), \ldots, \nabla u(x), u(x), x) = 0, \qquad (*)$$

where $F\colon \mathbb{R}^{d^k} \times \mathbb{R}^{d^{k-1}} \times \ldots \times \mathbb{R}^d \times \mathbb{R} \times \Omega \to \mathbb{R}$ is a given function and $u\colon \Omega \to \mathbb{R}$ is unknown.

[1] At least if the exact value of a constant does not play a decisive role.

b) u is a classical solution of $(*)$, if all derivatives of u appearing in $(*)$ exist, are continuous and $(*)$ is fulfilled.

Definition 1.4. Let $\Omega \subseteq \mathbb{R}^d$ and $k \in \mathbb{N}$. Then $(*)$ is called

a) linear, if it is of the form

$$\sum_{|\alpha| \le k} a_\alpha(x) \, \mathrm{D}^\alpha \, u(x) = f(x), \qquad x \in \Omega,$$

with $a_\alpha, f \colon \Omega \to \mathbb{R}$ given. If $f \equiv 0$, we call this a homogeneous linear differential equation, otherwise an inhomogeneous linear differential equation.

b) semi-linear, if it is of the form

$$\sum_{|\alpha| = k} a_\alpha(x) \, \mathrm{D}^\alpha \, u(x) + G(\mathrm{D}^{k-1} \, u(x), \ldots, \nabla u(x), u(x), x) = 0, \qquad x \in \Omega,$$

i.e. the highest order term is linear.

c) quasi-linear, if it is of the form

$$\begin{aligned} &\sum_{|\alpha| = k} a_\alpha(D^{k-1} u(x), \ldots, \nabla u(x), u(x), x) \, \mathrm{D}^\alpha \, u(x) \\ &\quad + G(\mathrm{D}^{k-1} \, u(x), \ldots, \nabla u(x), u(x), x) \\ &= 0, \qquad x \in \Omega, \end{aligned}$$

i.e., the highest order terms enter linearly.

d) non-linear, else.

Definition 1.5 (Classification of quasi-linear PDE of order 2)**.** Consider

$$\sum_{i,j=1}^{d} a_{ij}(\nabla u(x), u(x), x) \partial_i \partial_j u(x) = f(\nabla u(x), u(x), x) \tag{$\triangle$}$$

with $a_{ij}, f \colon \mathbb{R}^d \times \mathbb{R} \times \Omega \to \mathbb{R}$. Then the coefficients form a function-valued matrix $A \coloneqq (a_{ij})_{i,j=1}^d$. Assuming that the involved functions are twice continuously differentiable, this matrix can be chosen to be symmetric due to Schwarz' theorem. Then the PDE $(\triangle)$ and the coefficient matrix A are called

a) elliptic, if all eigenvalues of A are non-zero and of the same sign on all of $\mathbb{R}^d \times \mathbb{R} \times \Omega$.

b) parabolic, if exactly one eigenvalue of A is zero, while all other eigenvalues are non-zero with same sign.

c) hyperbolic, if all eigenvalues are non-zero, and exactly one of opposite sign than all the other.

Example 1.6. a) Heat equation:

$$u_t(t,x) - \Delta u(t,x) = 0$$

is a linear parabolic equation, since its coefficient matrix is

$$A(p,u,(t,x)) = \begin{pmatrix} 0 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{pmatrix}.$$

b) Poisson equation:

$$-\Delta u(x) = f(x)$$

with coefficient matrix

$$A(p,u,x) = \begin{pmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{pmatrix},$$

which is a linear elliptic equation.

c) Wave equation:

$$u_{tt}(t,x) - \Delta u(t,x) = 0$$

The wave equation looks very similar to the heat equation but in fact behaves very differently, as we will see in Chapter 3. The corresponding coefficient matrix is

$$A(p,u,(t,x)) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & -1 \end{pmatrix}$$

and thus it is linear and hyperbolic.

d) Reaction-diffusion equation:

$$u_t(t,x) - \Delta u(t,x) = f(u(t,x))$$

This equation is used to describe substances that mix and enter chemical reactions with each other. The dependence of f on u can intuitively be understood as the idea that the speed of reaction depends on the concentration of substances.

These are of semi-linear parabolic type.

e) Minimal surface equation:

$$-\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+\|\nabla u\|^2}}\right) = f$$

This equation is typically found in the study of minimal surfaces and can be used to describe the shape and behaviour of e.g. soap films.

It is quasi-linear and elliptic.

Remark. Hyperbolic PDEs behave quite differently than elliptic or parabolic PDEs, so in fact most experts in the topic of PDEs are specialised in either hyperbolic PDEs or the other types.

Remark 1.7. We consider the heat equation $u_t - \Delta u = f$ on $(0,\infty)\times\Omega$. For uniqueness of solutions we obviously need additional knowledge about the described system. Typically this information is given in the shape of:

a) Initial conditions: The status of the system is given at some starting time. Usually this time will be taken as 0 which yields the initial condition $u(0,x) = u_0(x)$.

b) Boundary conditions: There is a huge variety of sensible boundary conditions. Here are some of the most commonly used:

- Perfect insulation: Although it is physically not possible, we often assume that the boundary is perfectly insulated, i.e. no heat can enter or leave through the boundary. In other words, for all $t > 0$ and for all $x \in \partial\Omega$ we have

  $$0 = F(t,x)\nu(x) \overset{\text{Fourier}}{=} -k\nabla u(t,x)\cdot\nu(x).$$

  This leads to the Neumann boundary condition

  $$\frac{\partial u}{\partial \nu}(t,x) = 0.$$

- Often we want to fix a prescribed value at the boundary $\partial\Omega$. Mathematically we can write this as follows: for all times $t > 0$ and all $x \in \partial\Omega$ we have

  $$u(t,x) = g(t,x),$$

  where the function g is given data. We call this the Dirichlet boundary condition.

- Non-perfect insulation: This condition describes a mixture of the upper two. Here, the heat flux over the boundary is proportional to the difference of the temperature on the inside and on the outside. We get

  $$F(t,x)\cdot\nu(x) = \underbrace{\gamma(t,x)}_{\text{given coefficient}}\cdot(u(t,x) - \underbrace{g(t,x)}_{\substack{\text{given boundary}\\ \text{condition}}}).$$

This leads to the so-called Robin boundary condition

$$\frac{\partial u}{\partial \nu}(t,x) + \frac{\gamma(t,x)}{k}\left(u(t,x) - g(t,x)\right) = 0.$$

In this lecture we will look at several special cases of mainly elliptic PDEs on more or less general domains and try to find conditions under which we can guarantee existence and/or uniqueness of solutions. In detail, we will look at

- the heat equation and wave equation on $\mathbb{R}^d$ and will derive explicit solutions.
- the Poisson equation on general $\Omega \subseteq \mathbb{R}^d$.

  Here classical solutions seem to be out of reach. In fact, the theory of PDEs was arduous while mathematicians chased after classical solutions. This led to relaxing the requirements on solutions, i.e. to develop a definition of "solutions" that need not be differentiable in the classical sense. For this, we will examine weak solutions, i.e. PDEs in the L^2-setting.

  There are several reasons to work in L^p-spaces instead of spaces of continuous functions. The most important ones are that the norm is differentiable and not only continuous and that L^p-spaces are reflexive while C^k-spaces are not. Last but not least for $p = 2$ we have the rich Hilbert space structure, that we will heavily exploit.
- regularity of weak solutions.

  Once we have weak solutions, we want to know how regular they are and under which conditions they are in fact classical solutions.
- L^p-theory of weak solutions.

  Initially, weak solutions are defined in an L^2-setting. In this final part of the course we want to transfer the concept to p different from 2.

2 The heat equation on $\mathbb{R}^d$

We will start our analysis of PDEs with one example of a parabolic linear equation: the *heat equation*. Our goal for this chapter is to explicitly solve the heat equation on the whole space $\mathbb{R}^d$. At first we will take a look at the equation itself to derive the so-called fundamental solution. This fundamental solution will allow us to solve the heat equation if the adequate data in form of initial values is provided. As we have seen in the previous chapter, the homogeneous heat equation is given by

$$\boxed{u_t(t,x) - \Delta u(t,x) = 0, \qquad (t,x) \in (0,\infty) \times \mathbb{R}^d.} \tag{HE}$$

We cannot help but have certain expectations for the behaviour of an equation called the heat equation. Indeed, we will see expected behaviour but also some completely unexpected properties of its solutions.

While calculating a possible solution, we will often not make sure that our mathematical operations are guaranteed to be allowed in a strict mathematical sense. However, these dark magic derivations will give us an idea for a solution, which in turn can be shown to solve the equation. Our goal will be to reduce the general problem to an ODE, since we know how to solve those.

Lemma 2.1 (Parabolic scaling)**.** *Let u be a classical solution of* (HE)*. Then for all $\lambda \in \mathbb{R}$ the function*

$$w(t,x) := u(\lambda^2 t, \lambda x)$$

is a classical solution as well.

This natural scaling, namely $|x| \sim \sqrt{t}$, is an innate property of the heat equation that somehow reflects that we have one derivative in time and two in space. It also shows that solutions of the heat equation cannot be unique and that initial values are mandatory to obtain uniqueness.

Proof. We show that w solves (HE) and look at

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} w(t,x) - \Delta w(t,x) &= \lambda^2 \frac{\partial u}{\partial t}(\lambda^2 t, \lambda x) - \lambda^2 \Delta u(\lambda^2 t, \lambda x) \\ &= \lambda^2 \underbrace{\left(u_t(\lambda^2 t, \lambda x) - \Delta u(\lambda^2 t, \lambda x)\right)}_{=0} = 0. \end{aligned}$$

□

Next we will focus on finding radially symmetric solutions, i.e. solutions of the form $u(t, |x|)$.[1] This might seem a bit ambitious, but since our space is somehow homogeneous and radially symmetric, it seems quite plausible for such a solution to exist. Once again, our intuition leads us here: given a distribution of heat in a homogeneous space, it seems natural to assume that it spreads equally in all directions, i.e. admits radial symmetry. Our approach will be to use the previously shown scaling invariance, combining it with our aim of finding a radially symmetric solution, which yields the following approach of a so-called "self-similar" solution:

$$u(t, |x|) = \lambda^\alpha u(\lambda t, \lambda^\beta |x|),\ \lambda \in \mathbb{R}.$$

Here, we give ourselves some freedom by inserting parameters α and β to find a self-similar solution for suitably chosen parameters, i.e. it basically looks the same everywhere modulo scaling. Setting $\lambda := \frac{1}{t}$ we conclude

$$u(t, |x|) = t^{-\alpha} u(1, t^{-\beta} |x|) =: t^{-\alpha} v(y)$$

for $y := \frac{|x|}{t^\beta}$. The advantage of this is, that the resulting function $v\colon \mathbb{R} \to \mathbb{R}$ will give rise to an ordinary differential equation.

In the following calculations we will use our freedom to choose suitable α and β. For all $x \in \mathbb{R}^d \setminus \{0\}$ and for all $t > 0$ and $j \in \{1, \dots, d\}$ the derivatives of u can be represented as

$$\begin{aligned} u_{x_j}(t, |x|) &= t^{-\alpha} v'(y) \frac{x_j}{t^\beta |x|}, \\ u_{x_j x_j}(t, |x|) &= t^{-\alpha} \left[v''(y) \cdot \frac{x_j^2}{t^{2\beta} |x|^2} + v'(y) \cdot \frac{1}{t^\beta} \left(\frac{1}{|x|} - \frac{x_j^2}{|x|^3} \right) \right]. \end{aligned}$$

This might look awful for now, but luckily the terms get a lot nicer once we actually sum up the spatial derivatives of order two for the Laplacian. Before that we still need to compute one more derivative:

$$u_t(t, |x|) = -\alpha t^{-\alpha-1} v(y) - \beta t^{-\alpha} v'(y) \cdot \frac{|x|}{t^{\beta+1}}.$$

We now insert these derivatives into (HE). Notice that $\sum_{j=1}^d x_j^2 = |x|^2$. This leads to

$$\begin{aligned} 0 &= -\alpha t^{-1-\alpha} v(y) - \beta t^{-\alpha} v'(y) \cdot \frac{|x|}{t^{\beta+1}} - t^{-\alpha} \left(v''(y) \cdot \frac{1}{t^{2\beta}} + v'(y) \cdot \frac{1}{t^\beta} \cdot \frac{d-1}{|x|} \right) \\ &= -t^{-\alpha-1} \Bigg(\alpha v(y) + \beta v'(y) \underbrace{t \cdot \frac{|x|}{t^{\beta+1}}}_{=y} + t v''(y) \cdot \frac{1}{t^{2\beta}} + t v'(y) \cdot \frac{1}{t^\beta} \cdot \frac{d-1}{|x|} \Bigg) \\ &= -t^{-\alpha-1} \left(t^{1-2\beta} v''(y) + \beta y v'(y) + (d-1) \frac{v'(y)}{y} \cdot t^{1-2\beta} + \alpha v(y) \right). \end{aligned}$$

[1]This is a bit of an abuse of notation, since u was assumed to have $d+1$ variables, while this new u only has 2. Nonetheless, it should be clear what is meant here.

We observe that in the last step there are no instances of x left. We set $\beta = \frac{1}{2}$ and get

$$0 = v''(y) + \frac{1}{2} y v'(y) + (d-1)\frac{v'(y)}{y} + \alpha v(y)$$

for $y > 0$ which finally is an ODE. Multiplying with y^{d-1} does the trick

$$0 = v''(y) \cdot y^{d-1} + \frac{1}{2} y^d v'(y) + (d-1) y^{d-2} v'(y) + \alpha v(y) y^{d-1}.$$

Magically, we find both terms of the derivative of $y^{d-1}v'(y)$ in this equation. Choosing $\alpha = \frac{d}{2}$, we rewrite this as:

$$\begin{aligned} 0 &= (y^{d-1}v'(y))' + \frac{1}{2}(y^d v(y))', \\ &= \left(y^{d-1}v'(y) + \frac{1}{2} y^d v(y)\right)', \end{aligned}$$

so we obtain

$$y^{d-1}v'(y) + \frac{1}{2} y^d v(y) = \text{const.}$$

If v is rapidly decreasing, the constant must be 0. Then

$$y^{d-1}v'(y) = -\frac{1}{2} y^d v(y),$$

or simpler

$$v'(y) = -\frac{1}{2} y v(y).$$

This ODE can be solved quite easily. The solution is $v(y) = c\mathrm{e}^{-\frac{y^2}{4}}$, so

$$u(t, |x|) = t^{-\alpha} v\left(\frac{|x|}{t^{\frac{1}{2}}}\right) = c \cdot t^{-\frac{d}{2}} \mathrm{e}^{-\frac{|x|^2}{4t}}.$$

Finally, we have found a solution: It is the Gaussian kernel for the normal distribution. This specific solution can be used to determine general solutions of the heat equation, which explains the following name.

Definition 2.2. The fundamental solution of (HE) is

$$\Phi(t, x) = \frac{1}{(4\pi t)^{\frac{d}{2}}} \mathrm{e}^{-\frac{|x|^2}{4t}}.$$

It is called the heat kernel or Gaussian kernel.

Remark. This is just one solution of the heat equation. From the point of physical application it might not be special but, as we will see in Theorem 2.4, it allows us to derive solutions for arbitrary initial values via convolution with the initial data.

Let us collect some important properties of the heat kernel.

Proposition 2.3. *a) The function Φ is smooth, i.e. $\Phi \in \mathrm{C}^\infty((0,\infty)\times\mathbb{R}^d)$.*

b) For all $\varepsilon > 0$ it holds that

$$\lim_{t\to 0+} \Phi(t,x) = 0$$

uniformly on $\mathbb{R}^d \setminus B_\varepsilon(0)$.[2]

c) $\lim_{t\to 0+} \Phi(t,0) = \infty$.

d) $\partial_t\Phi - \Delta\,\Phi = 0$ on $(0,\infty)\times\mathbb{R}^d$, i.e. the Gaussian kernel actually is a solution.

e) $\int_{\mathbb{R}^d} \Phi(t,x)\,\mathrm{d}x = 1$ for all $t > 0$.

Proof. a) This follows immediately.

b) For $|x| > \varepsilon$ it holds that

$$|\Phi(t,x)| = \frac{1}{(4\pi t)^{\frac{d}{2}}}\mathrm{e}^{\frac{-|x|^2}{4t}} \le \frac{1}{(4\pi t)^{\frac{d}{2}}}\mathrm{e}^{\frac{-\varepsilon^2}{4t}},$$

which tends to 0, as t does.

c) Obvious.

d) This can be proven by a simple calculation.

e) Let us calculate

$$\begin{aligned}\int_{\mathbb{R}^d} \Phi(x,t)\,\mathrm{d}x &= \frac{1}{(4\pi t)^{d/2}}\int_{\mathbb{R}^d} \mathrm{e}^{-\frac{|x|^2}{4t}}\,\mathrm{d}x \\ &\overset{y=\frac{x}{\sqrt{4t}}}{=} \frac{1}{(4\pi t)^{d/2}}\int_{\mathbb{R}^d} \mathrm{e}^{-|y|^2}(4t)^{\frac{d}{2}}\,\mathrm{d}y \\ &= \frac{1}{\pi^{d/2}}\int_{\mathbb{R}^d} \mathrm{e}^{-|y|^2}\,\mathrm{d}y = 1.\end{aligned}$$

□

Theorem 2.4. *Let $u_0 \in \mathrm{BC}(\mathbb{R}^d)$*[3]*. Then for*

$$u(t,x) := \begin{cases}(\Phi(t,\cdot) * u_0)(x), & (t,x)\in(0,\infty)\times\mathbb{R}^d,\\ u_0(x), & x\in\mathbb{R}^d,\ t=0,\end{cases}$$

[2] Here and lateron $B_r(x_0)$ denotes the open ball with radius r and midpoint x_0.

[3] The space of bounded continuous functions. It is a Banach space if equipped with the supremum norm.

we have $u \in \mathrm{C}^\infty((0,\infty)\times\mathbb{R}^d)\cap\mathrm{C}([0,\infty)\times\mathbb{R}^d)$ *and* u *solves*

$$\begin{cases} u_t(t,x) = \Delta u(t,x), & t>0, x\in\mathbb{R}^d,\\ u(0,x) = u_0(x), & x\in\mathbb{R}^d.\end{cases}$$

This problem is usually called the Cauchy problem for the heat equation.

Proof. 1[st] **step:** We want to show $u \in \mathrm{C}^\infty((0,\infty)\times\mathbb{R}^d)$ and compute its derivatives by interchanging the order of differentiation and integration. To see that this is okay, first notice that $\Phi \in \mathrm{C}^\infty((0,\infty)\times\mathbb{R}^d)$ and is integrable by Proposition 2.3 a) and d). What we are missing now is a uniform bound for its derivatives of arbitrary order.

Let $\delta\in(0,1)$. By induction over $|\alpha|$ we can show that for all $\alpha\in\mathbb{N}_0^{d+1}$ we find $c(\delta,\alpha)\geq 0$ such that for all $(t,x)\in[\delta,1/\delta]\times\mathbb{R}^d$ the derivatives of order $|\alpha|$ are bounded by

$$|\mathrm{D}^\alpha\,\Phi(t,x)| \leq c(\delta,\alpha)\cdot(1+|x|^m)\,\mathrm{e}^{-\frac{\delta|x|^2}{4}} =: \Psi(x),$$

where m depends on $|\alpha|$. This implies

$$|\mathrm{D}^\alpha\,\Phi(t,x-\cdot)u_0| \leq \Psi(x-\cdot)\|u_0\|_\infty$$

and the left-hand side is integrable on $\mathbb{R}^d$.

Thus, by differentiation of parameter integrals we obtain $u\in\mathrm{C}^\infty((\delta,1/\delta)\times\mathbb{R}^d)$ with

$$\mathrm{D}^\alpha\,u(t,x) = \int_{\mathbb{R}^d}\mathrm{D}^\alpha\,\Phi(t,x-y)u_0(y)\,\mathrm{d}y = \left(\mathrm{D}^\alpha\,\Phi(t,\cdot)*u_0\right)(x)$$

for all $t\in(\delta,1/\delta)$ and all $x\in\mathbb{R}^d$. Since δ was arbitrary, the claim follows.

2[nd] **step:** Next we show that u solves (HE):

$$\begin{aligned} u_t-\Delta\,u &= \frac{\mathrm{d}}{\mathrm{d}t}(\Phi(t,\cdot)*u_0)(x) - \Delta\left[(\Phi(t,\cdot)*u_0)(x)\right]\\ &\overset{1^{\text{st}}\text{ step}}{=} (\Phi_t(t,\cdot)*u_0)(x) - ((\Delta\,\Phi)(t,\cdot)*u_0)(x)\\ &= (\underbrace{(\Phi_t-\Delta\,\Phi)}_{=0}(t,\cdot)*u_0)(x)\\ &\overset{\text{Prop. 2.3(c)}}{=} 0.\end{aligned}$$

3[rd] **step:** Lastly, it remains to show that u attains the initial conditions continuously, i.e. for all $x_0\in\mathbb{R}^d$ we have $u(t,x)\to u_0(x_0)$ for $(t,x)\to(0,x_0)$.

Let $\varepsilon>0$ and $x_0\in\mathbb{R}^d$. Since u_0 is continuous, we obtain existence of a $\delta>0$ such that for all $x\in B_\delta(x_0)$ we have

$$|u_0(x)-u_0(x_0)| < \frac{\varepsilon}{2}.$$

Hence, for all $t > 0$ and for all $x \in B_{\frac{\delta}{2}}(x_0)$ we have

$$\begin{aligned}
|u(t,x) - u_0(x_0)| &= \left| \int_{\mathbb{R}^d} \Phi(t, x-y) u_0(y) \, \mathrm{d}y - u_0(x_0) \underbrace{\int_{\mathbb{R}^d} \Phi(t, x-y) \, \mathrm{d}y}_{=1} \right| \\
&= \left| \int_{\mathbb{R}^d} \Phi(t, x-y)(u_0(y) - u_0(x_0)) \, \mathrm{d}y \right| \\
&\leq \int_{\mathbb{R}^d \setminus B_\delta(x_0)} \Phi(t, x-y) \, |u_0(y) - u_0(x_0)| \, \mathrm{d}y + \int_{B_\delta(x_0)} \dots \, \mathrm{d}y \\
&\leq 2 \|u_0\|_\infty \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d \setminus B_\delta(x_0)} \mathrm{e}^{-\frac{|x-y|^2}{4t}} \, \mathrm{d}y + \frac{\varepsilon}{2} \underbrace{\int_{\mathbb{R}^d} \Phi(t, x-y) \, \mathrm{d}y}_{=1} .
\end{aligned}$$

Moreover, we compute

$$\begin{aligned}
|y - x_0| &\leq |y - x| + |x - x_0| \\
&\overset{x \in B_{\frac{\delta}{2}}(x_0)}{\leq} |y - x| + \frac{\delta}{2} \\
&\overset{y \notin B_\delta(x_0)}{\leq} |y - x| + \frac{|y - x_0|}{2} .
\end{aligned}$$

This estimate allows us to estimate $|y - x|$ from below by $\left| \frac{y - x_0}{2} \right| \leq |y - x|$. Using this, we obtain

$$\begin{aligned}
|u(t,x) - u_0(x_0)| &\leq \frac{\varepsilon}{2} + \|u_0\|_\infty \frac{1}{(4\pi t)^{\frac{d}{2}}} \int_{\mathbb{R}^d \setminus B_\delta(x_0)} \mathrm{e}^{-\frac{|y-x_0|^2}{16t}} \, \mathrm{d}y \\
&= \frac{\varepsilon}{2} + \|u_0\|_\infty \frac{1}{(4\pi)^{\frac{d}{2}}} \underbrace{\int_{\mathbb{R}^d \setminus B_{\frac{\delta}{\sqrt{t}}}(0)} \mathrm{e}^{-\frac{|z|^2}{16}} \, \mathrm{d}z}_{\to 0 \text{ for } t \to 0}
\end{aligned}$$

where we substituted $z = \frac{y - x_0}{\sqrt{t}}$. This implies $u(t,x) \to u_0(x_0)$ for $(t,x) \to (0, x_0)$. □

Remark. a) For $t \to 0$, the function $\Phi(t, \cdot)$ approximates the Dirac δ-Distribution δ_0. It is an approximation of unity and even a mollifier.

b) If $u_0 \in \mathrm{BC}(\mathbb{R}^d)$ satisfies $u_0 \geq 0$ and $u_0 \not\equiv 0$, then the solution u stays positive for all times $t > 0$. We have

$$u(t,x) = \int_{\mathbb{R}^d} \Phi(t, x-y) u_0(y) \, \mathrm{d}y > 0$$

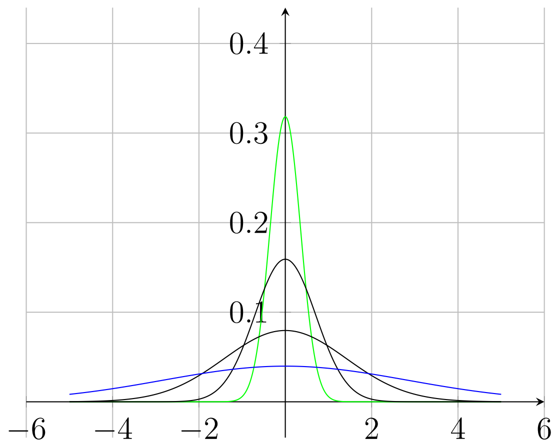

Figure 2.1: The fundamental solution of the heat equation for various parameters t

for all $x \in \mathbb{R}^d$, so the heat equation is positivity preserving. But it has the even stronger property of being positivity improving. Even for an initial value with compact support, u will nonetheless immediately[4] be strictly larger than 0 at *any* point $x \in \mathbb{R}^d$.

In other words: Our model for heat flows has infinite propagation speed. Once again, we are reminded, that our mathematical model does not reflect reality; it is but an approximation, that is simple enough to allow for a rich mathematical investigation and sufficiently accurate on small scales. On large scales, however, more sophisticated tools are required.

c) For an initial value $u_0 \in \mathrm{BC}(\mathbb{R}^d) \cap \mathrm{L}^1(\mathbb{R}^d)$, we find that

$$\begin{aligned} \|u(t,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^d)} &= \operatorname{esssup}_{x\in\mathbb{R}^d} \left| \frac{1}{(4\pi t)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \underbrace{\mathrm{e}^{-\frac{|x-y|^2}{4t}}}_{\leq 1} u_0(y)\,\mathrm{d}y \right| \\ &\leq \frac{1}{(4\pi t)^{\frac{d}{2}}} \int_{\mathbb{R}^d} |u_0(y)|\,\mathrm{d}y \\ &= \frac{\|u_0\|_{\mathrm{L}^1(\mathbb{R}^d)}}{(4\pi t)^{\frac{d}{2}}}, \end{aligned}$$

which tends to 0, as $t \to \infty$.

d) Finite speed of propagation is fulfilled for the porous medium equation which is given by

$$u_t = \Delta(u^m) \left(= m u^{m-1} \Delta u + m(m-1) u^{m-2} |\nabla u|^2\right)$$

for some $m > 1$. It is a quasilinear, degenerate parabolic equation. Here the m describes the conductivity of the material.

[4]This means for any $t > 0$.

e) In Theorem 2.4 we have seen an effect that is typical for linear, parabolic equations. It states that a solution is smooth for all times $t > 0$. This means that even if we start with a rough initial value, a smoothing effect immediately takes place. We can observe this, for example, if we put a snowball into water. It is this feature that limits parabolic equations to the description of only non-reversible processes.

3 The wave equation on $\mathbb{R}^d$

As it happens so often, mathematical theories arise as the attempt of capturing and describing real world systems, problems and their solutions. This is especially true for partial differential equations, which are frequently used to describe physical systems because they depend on the exertion of fundamental forces, which necessitate a description of acceleration or growth, mathematically represented as derivatives. One of the most prominent examples presents itself as the wave equation, which is given by

$$\boxed{u_{tt}(t,x) - \Delta\, u(t,x) = 0, \qquad (t,x) \in (0,\infty) \times \mathbb{R}^d.} \tag{WE}$$

We will mainly focus on the wave equation in the cases of dimensions 1 ($\mathbb{R}$ and $\mathbb{R}_+$), 2 and 3, for which we will derive explicit solutions in this chapter. This, however, is not as big a restriction as it seems, as the qualitative behaviour of the wave equation depends mainly on the parity of the underlying space dimension.

3.1 The wave equation on $\mathbb{R}$

In one dimension the wave equation is given by

$$\begin{cases} u_{tt} - u_{xx} = 0, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ u(0,x) = u_0(x),\ u_t(0,x) = u_1(x), & x \in \mathbb{R}, \end{cases} \tag{WE$_1$}$$

with sufficiently smooth initial data u_0, $u_1 \colon \mathbb{R} \to \mathbb{R}$. We first observe that we can factorise the involved derivatives according to

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right) u(t,x) = \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) \underbrace{\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) u(t,x)}_{=: v(t,x)},$$

which holds for all $(t,x) \in (0,\infty) \times \mathbb{R}$. Then the so-defined function v satisfies

$$v_t + v_x = 0. \tag{3.1}$$

This is a special case of the (one-dimensional) Transport Equation

$$\begin{cases} w_t(t,x) + b w_x(t,x) = f(t,x), & (t,x) \in (0,\infty) \times \mathbb{R}, \\ w(0,x) = w_0(x), & x \in \mathbb{R}, \end{cases} \tag{TE}$$

for $b \in \mathbb{R}$, $f \in \mathrm{C}^{0;1}((0,\infty)\times\mathbb{R})$[1] and $w_0 \in \mathrm{C}^1(\mathbb{R})$.

For future needs we solve this equation in its full generality, for which the following observation will come in handy: Let w be a solution of the transport equation (TE). Then it holds that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s} w(s, x - b(t-s)) &= w_t(s, x-b(t-s)) + w_x(s, x-b(t-s)) \cdot b \\ &= f(s, x-b(t-s)).\end{aligned}$$

Using the fundamental theorem of calculus we get

$$\begin{aligned}w(t,x) - w(0, x-bt) &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} w(s, x-b(t-s))\,\mathrm{d}s \\ &= \int_0^t f(s, x-b(t-s))\,\mathrm{d}s.\end{aligned}$$

Thus, we obtain

$$w(t,x) = w_0(x-bt) + \int_0^t f(s, x-b(t-s))\,\mathrm{d}s. \tag{$\star$}$$

Indeed, this formula already states the correct and unique solution. We gather and prove this result and the desired regularity properties in the following Lemma.

Lemma 3.1. *The transport equation* (TE) *has a unique solution* $w \in \mathrm{C}^1((0,\infty)\times\mathbb{R}) \cap \mathrm{C}([0,\infty)\times\mathbb{R})$ *given by* ($\star$).

Proof. The uniqueness of the solution follows directly from our above calculation. So it remains to check the claimed regularity and that the function given in ($\star$) is a solution indeed.

Regularity: By assumption, w_0 is a C^1 function on $\mathbb{R}$. Since $f \in \mathrm{C}^{0,1}((0,\infty)\times\mathbb{R})$ the integral is in $\mathrm{C}^1((0,\infty)\times\mathbb{R})$ as well. Moreover, for $t \to 0+$, we conclude $w(t,x) \to w_0(x)$. This shows the claimed regularity.

Solution: Let us now check that the function acquired through the above formula actually is a solution. We calculate (remember the trick from Analysis II how to differentiate parameter integrals with the variable appearing in the limits and the integrand!)

$$\begin{aligned}\partial_t w(t,x) + b\partial_x w(t,x) &= -bw_0'(x-bt) + f(t,x) + \int_0^t \partial_x f(s, x-b(t-s))(-b)\,\mathrm{d}s \\ &\qquad + b\left(w_0'(x-bt) + \int_0^t \partial_x f(s, x-b(t-s))\,\mathrm{d}s\right) \\ &= f(t,x),\end{aligned}$$

[1] Here, $\mathrm{C}^{0;1}$ denotes the space of all functions that are continuous in the first and continuously differentiable in the second argument

which shows that we have indeed found a solution. □

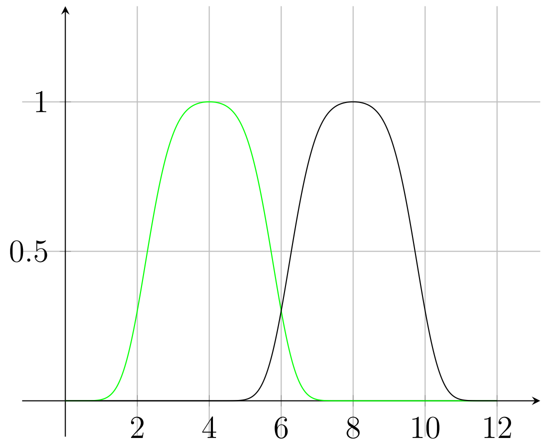

Figure 3.1: Visualisation of a solution to the transport equation for $f = 0$.

The above helps us in our endeavour of solving the wave equation by allowing us to subdivide the task into two simpler problems: Firstly, we need to solve the arised transport equation $v_t + v_x = 0$ for v and secondly, v and u are connected via another transport equation, as by definition $u_t - u_x = v$.[2]

By Lemma 3.1 for $b = 1$ and $f = 0$ we get that

$$v(t,x) = v(0, x-t) = v_0(x-t) \qquad \text{for } (t,x) \in (0,\infty) \times \mathbb{R}.$$

To deal with (WE$_1$), we observe that

$$v_0(x-t) = v(t,x) = u_t(t,x) - u_x(t,x),$$

which is the aforementioned second transport equation, to which we apply Lemma 3.1 with $b = -1$ and $f(t,x) = v_0(x-t)$ to calculate

$$u(t,x) = u_0(x+t) + \int_0^t v_0(x + 1(t-s) - s)\,\mathrm{d}s.$$

Setting $\sigma := x + t - 2s$ we conclude

$$\begin{aligned}
&= u_0(x+t) + \int_{x+t}^{x-t} v(0,\sigma)(-\frac{1}{2})\,\mathrm{d}\sigma \\
&= u_0(x+t) + \frac{1}{2}\int_{x-t}^{x+t} \left(u_t(0,\sigma) - u_x(0,\sigma)\right)\,\mathrm{d}\sigma \\
&= u_0(x+t) + \frac{1}{2}\int_{x-t}^{x+t} u_1(\sigma)\,\mathrm{d}\sigma - \frac{1}{2}\int_{x-t}^{x+t} u_0'(\sigma)\,\mathrm{d}\sigma \\
&= u_0(x+t) - \frac{1}{2}u_0(x+t) + \frac{1}{2}u_0(x-t) + \frac{1}{2}\int_{x-t}^{x+t} u_1(\sigma)\,\mathrm{d}\sigma.
\end{aligned}$$

[2] This is why we considered the more general transport equation, as we need it with arbitrary right-hand sides and coefficient -1.

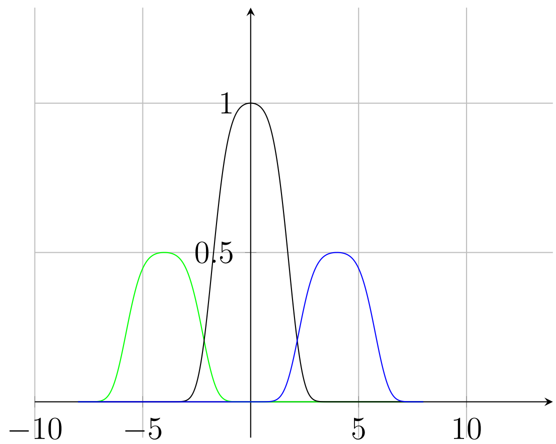

Figure 3.2: Visualisation of the wave equation for $u_1 = 0$

This leads to the final solution

$$u(t,x) = \frac{1}{2}u_0(x+t) + \frac{1}{2}u_0(x-t) + \frac{1}{2}\int_{x-t}^{x+t} u_1(s)\,\mathrm{d}s. \tag{3.2}$$

Interestingly, this solution still reflects the fact that it was obtained by solving a transport equation: The first and second term describe a motion of initial data to the left and to the right, respectively. The integral includes initial data for the velocity. If we start with no starting velocity, we see that the initial data propagates symmetrically in both directions, cf. Figure 3.2

Theorem 3.2 (Formula of d'Alembert)**.** *Let* $u_0 \in \mathrm{C}^2(\mathbb{R})$ *and* $u_1 \in \mathrm{C}^1(\mathbb{R})$*. Then* u *given by* (3.2) *is* $\mathrm{C}^2([0,\infty) \times \mathbb{R})$ *and is the unique solution to* (WE_1)*.*

Proof. The uniqueness and the regularity follow immediately from the formula since $u_0 \in \mathrm{C}^2(\mathbb{R})$ and $u_1 \in \mathrm{C}^1(\mathbb{R})$ but we integrate it once. Additionally one can calculate that u solves the equation. □

In contrast to the heat equation, we can observe by d'Alembert's formula that the solution of the wave equation only has a finite speed of propagation. This is since the amplitude of the wave in a point x at time t depends only on the initial values in the points $x+t$ and $x-t$ and the speed of the wave between those two points. Therefore, the value of $u(t,x)$ only depends on the wave before in the interval $(x-t, x+t)$.

Remark 3.3. a) For $f, g \in \mathrm{C}^2(\mathbb{R})$ let $y(t,x) := f(x+t) + g(x-t)$. Then y solves $y_{tt} = y_{xx}$ since

$$y_{tt}(t,x) = f''(x+t) + g''(x-t) = y_{xx}(t,x).$$

The contribution of f and g can be interpreted as some masses going left and right, respectively.

b) For $u_0 \in C^k(\mathbb{R})$ and $u_1 \in C^{k-1}(\mathbb{R})$ our solution u will be k-times continuously differentiable on $[0,\infty) \times \mathbb{R}$ but higher order derivatives do not exist in general. This means that we have no universal smoothing effect like in the case of the heat equation. We refer to this as propagation of singularities.

c) Notice that the formula even makes sense for more general initial conditions. Especially ones that are not differentiable.

Remark. Parabolic PDEs model irreversible processes, hyperbolic PDEs model reversible processes. The latter is linked to the propagation of singularities: If it would be possible to smooth a singularity away, the reversed process would have to create them.

3.2 The wave equation on $\mathbb{R}_+$

The interpretation of a wave equation that has a bound in the spatial variables is that of a wave hitting a wall in one direction. To reflect this, we equip the equation with a boundary condition that forces the solution to describe the wave's behaviour around the point of impact accordingly. One way of formulating the equation is as follows:

$$\begin{cases} u_{tt}(t,x) - u_{xx}(t,x) = 0, & (t,x) \in (0,\infty)^2 \\ u(t,0) = 0, & t \in (0,\infty) \\ u(0,x) = u_0(x),\ u_t(0,x) = u_1(x), & x \in (0,\infty). \end{cases} \tag{WE$_{1,+}$}$$

Here, we equipped the equation with the Dirichlet boundary condition, i.e. we force the values to be 0 at the boundary. A fruitful technique for solving this is to extend the initial condition to the whole space $\mathbb{R}$ via an odd extension and derive a new equation on $\mathbb{R}$ which we can already solve with the d'Alembert formula.

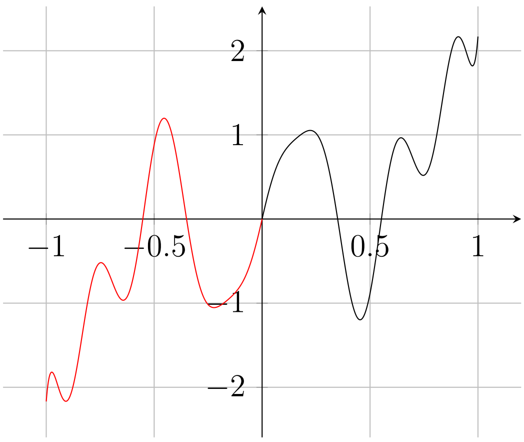

Figure 3.3: Visualisation of an odd extension

Notation 3.4. For $f \in \mathrm{C}([0,\infty))$, we set

$$\tilde{f}(x) := \begin{cases} f(x), & x \geq 0 \\ -f(-x), & x < 0. \end{cases}$$

Lemma 3.5. *Let $k \in \mathbb{N}$, $f \in \mathrm{C}^k([0,\infty))$. Then we have the equivalence*

$$\tilde{f} \in \mathrm{C}^k((-\infty,\infty)) \Leftrightarrow f^{(m)}(0) = 0 \text{ for all even } m \leq k.$$

Proof. Let $m \leq k$. Since $\tilde{f} \in \mathrm{C}^k(\mathbb{R} \setminus \{0\})$, we only look at its derivatives in 0. We compute

$$\begin{aligned} \tilde{f}^{(m)}(0+) &= f^{(m)}(0), \\ \tilde{f}^{(m)}(0-) &= -(-1)^m f^{(m)}(0) \\ &= (-1)^{m+1} f^{(m)}(0). \end{aligned}$$

Hence, $\tilde{f}^{(m)}$ is continuous in 0 if and only if $f^{(m)}(0) = 0$ for all even m. □

Theorem 3.6. *Let $u_0 \in \mathrm{C}^2([0,\infty))$ and $u_1 \in \mathrm{C}^1([0,\infty))$, such that $u_0(0) = 0$, $u_0''(0) = 0$ and $u_1(0) = 0$.*[3] *Define*

$$u(t,x) := \begin{cases} \frac{1}{2}\left(u_0(x+t) + u_0(x-t)\right) + \frac{1}{2}\int_{x-t}^{x+t} u_1(s)\,\mathrm{d}s & \text{for } x \geq t \geq 0 \\ \frac{1}{2}\left(u_0(x+t) - u_0(-x+t)\right) + \frac{1}{2}\int_{-x+t}^{x+t} u_1(s)\,\mathrm{d}s, & \text{for } t \geq x \geq 0. \end{cases} \qquad (*)$$

Then $u \in \mathrm{C}^2([0,\infty) \times [0,\infty))$ and it solves (WE$_{1,+}$) with initial values u_0 and u_1.

Proof. By Lemma 3.5 and the compatibility conditions, we immediately see that $\tilde{u}_0 \in \mathrm{C}^2(\mathbb{R})$ and $\tilde{u}_1 \in \mathrm{C}^1(\mathbb{R})$. Moreover, we obtain by d'Alembert's formula that

$$u^*(t,x) := \frac{1}{2}\left(\tilde{u}_0(x+t) + \tilde{u}_0(x-t)\right) + \frac{1}{2}\int_{t-x}^{t+x} \tilde{u}_1(s)\,\mathrm{d}s, \quad (t,x) \in [0,\infty) \times \mathbb{R}$$

is the unique solution of

$$\begin{cases} u^*_{tt} - u^*_{xx} &= 0 \\ u^*(0,x) &= \tilde{u}_0(x) \\ u^*_t(0,x) &= \tilde{u}_1(x) \end{cases}$$

and $u^* \in \mathrm{C}^2([0,\infty) \times \mathbb{R})$.

Define $u := u^*|_{[0,\infty)\times(0,\infty)} \in \mathrm{C}^2([0,\infty) \times (0,\infty))$. Then $u_{tt} - u_{xx} = 0$ for $(t,x) \in (0,\infty) \times (0,\infty)$. The function u obeys the Dirichlet conditions, because $\tilde{u}_0$ and $\tilde{u}_1$ are odd, so for all $t > 0$ it holds

$$u(t,0) = u^*(t,0) = \frac{1}{2}[\tilde{u}_0(t) + \tilde{u}_0(-t)] + \frac{1}{2}\int_{-t}^{t} \tilde{u}_1(s)\,\mathrm{d}s = 0.$$

[3] These so called compatibility conditions make sense, as they force the initial conditions to also respect the restrictions on the boundary.

To show the validity of the initial condition, first notice that for all $x > 0$ we have

$$u(0,x) = u^*(0,x) = \frac{1}{2}[\tilde{u}_0(x) + \tilde{u}_0(x)] + 0 = \tilde{u}_0(x) = u_0(x).$$

In addition, for all $t > 0$ we obtain

$$u_t(t,x) = u_t^*(t,x) = \frac{1}{2}[\tilde{u}_0'(x+t) - \tilde{u}_0'(x-t)] + \frac{1}{2}[\tilde{u}_1(x+t) + \tilde{u}_1(x-t)].$$

This implies that for $t \searrow 0$ we have

$$u_t(0,x) = \frac{1}{2}[\tilde{u}_0'(x) - \tilde{u}_0'(x)] + \frac{1}{2}[\tilde{u}_1(x) + \tilde{u}_1)(x)] = \tilde{u}_1(x) = u_1(x).$$

The correctness of the formula ($*$) follows from a direct calculation. □

3.3 The wave equation in $\mathbb{R}^3$

The next case of the wave equation we will tackle is the case of three spatial dimensions. In this case it has the form

$$\begin{cases} u_{tt} - \Delta u = 0, & (t,x) \in (0,\infty) \times \mathbb{R}^3, \\ u(0,x) = u_0(x), u_t(0,x) = u_1(x), & x \in \mathbb{R}^3 \end{cases} \tag{WE$_3$}$$

with u_0 and u_1 given for the initial conditions.

The key tool to compute the solution in a point x at a time t is evaluating the mean value on spheres around the point. As we will see, this reduces the problem to the case of the wave equation on $\mathbb{R}_+$, which we just so happen to have solved previously.

Lemma 3.7. *Let $u \in \mathrm{C}^2((0,\infty) \times \mathbb{R}^3)$ be a solution for the three-dimensional wave equation (WE$_3$). For all $x \in \mathbb{R}^3$ we set*

$$U(x,t,r) \coloneqq \frac{1}{|\partial\,\mathrm{B}_r(x)|} \int_{\partial\,\mathrm{B}_r(x)} u(t,y)\,\mathrm{d}S(y), \quad t \geq 0, r > 0$$

the spherical mean of u.

Then $\hat{U}(x,t,r) = r \cdot U(x,t,r)$ solves (WE$_{1,+}$) for the initial conditions

$$\hat{u}_0(x,r) \coloneqq r \cdot \frac{1}{|\partial\,\mathrm{B}_r(x)|} \int_{\partial\,\mathrm{B}_r(x)} u_0(y)\,\mathrm{d}S(y)$$
$$\hat{u}_1(x,r) \coloneqq r \cdot \frac{1}{|\partial\,\mathrm{B}_r(x)|} \int_{\partial\,\mathrm{B}_r(x)} u_1(y)\,\mathrm{d}S(y).$$

Proof. Exercise. □

Theorem 3.8 (Kirchhoff's formula)**.** *Let $u_0 \in \mathrm{C}^3(\mathbb{R}^3)$ and $u_1 \in \mathrm{C}^2(\mathbb{R}^3)$. Then (WE$_3$) has a unique solution $u \in \mathrm{C}^2([0,\infty) \times \mathbb{R}^3)$ given by*

$$u(t,x) = \begin{cases} \dfrac{1}{|\partial B_t(x)|} \displaystyle\int_{\partial B_t(x)} (tu_1(y) + u_0(y) + \nabla u_0 \cdot (y-x)) \,\mathrm{d}S(y), & t > 0, x \in \mathbb{R}^3, \\ u_0(x), & t = 0, x \in \mathbb{R}^3. \end{cases}$$

Proof. Write down $\hat{U}(x,t,r)$ by d'Alembert's formula and use $u(x,t) = \lim_{r \searrow 0} \frac{\hat{U}(x,t,r)}{r}$. □

Remark. This result has an easy visualisation: Basically the formula says that in every point x, only the values on the boundary of the ball with radius t have an impact on the value in x at time t. If we think of the expansion of sound waves, this is natural. At time t we are only able to hear what happened at time 0 in a point p, if the distance between p and x is exactly t.[4]

3.4 The wave equation in $\mathbb{R}^2$

In this section we will use the solution for the three-dimensional wave equation to solve the two-dimensional wave equation. The method we use is called Hadamard's method of descent.

Let $u \in \mathrm{C}^2([0,\infty) \times \mathbb{R}^2)$ be a solution for the two-dimensional wave equation. Then $\hat{u}(t,x_1,x_2,x_3) \coloneqq u(t,x_1,x_2)$ solves (WE$_3$) together with the initial conditions

$$\hat{u}_0(x_1,x_2,x_3) = u_0(x_1,x_2)$$
$$\hat{u}_1(x_1,x_2,x_3) = u_1(x_1,x_2).$$

Now we can use Kirchhoff's formula setting $x_3 = 0$ to derive Poisson's formula.

Theorem 3.9 (Poisson's formula)**.** *Let $u_0 \in \mathrm{C}^3(\mathbb{R}^2)$ and $u_1 \in \mathrm{C}^2(\mathbb{R}^2)$. Then the wave equation in two dimensions has a unique solution $u \in \mathrm{C}^2([0,\infty) \times \mathbb{R}^2)$ given by*

$$u(t,x) \coloneqq \begin{cases} \dfrac{1}{2}\dfrac{1}{|\mathrm{B}_t(x)|} \displaystyle\int_{\mathrm{B}_t(x)} \dfrac{tu_0(y) + t^2 u_1(y) + t\nabla u_0(y) \cdot (y-x)}{(t^2 - |x-y|^2)^{\frac{1}{2}}} \,\mathrm{d}y, & t > 0, x \in \mathbb{R}^2 \\ u_0(x), & t = 0, x \in \mathbb{R}^2. \end{cases}$$

Remark 3.10. We see a huge difference to the formula in the three-dimensional case: In contrast to the boundaries of the balls in Kirchhoff's formula in Theorem 3.8, the whole disc influences our result. Using the same example as before implies the following: In every point x we are able to hear something that happened in point p if it happened *before* $t = |x-p|$, not only *at* $t = |x-p|$. So, in particular, if we hear something at some point, we will be able to hear it forever. We are lucky to live in an odd-dimensional world, because we humans rely on our ability to pinpoint the location of an object based on hearing its sound.

[4]Here, we implicitly assume normalised propagation speed.

We note that similar formulae exist in higher dimensions, based on the dimension's parity.

4 Maximum principle for harmonic functions

Up to now we have looked at the heat equation as a typical example for parabolic PDEs and at the wave equation as an example for hyperbolic PDEs. From now on until the end of the lecture we will focus on elliptic ones. As a prototypical example, we look at the Laplace equation

$$-\Delta u = 0 \text{ in } \Omega \subseteq \mathbb{R}^d \tag{LE}$$

and the Poisson equation (for given f)

$$-\Delta u = f \text{ in } \Omega \subseteq \mathbb{R}^d. \tag{PE}$$

We notice an important difference to the equations we have discussed before: Throughout the whole chapter we will consider equations on arbitrary open and connected domains $\Omega \subseteq \mathbb{R}^d$, instead of the whole space.

Definition 4.1. A function $u \in \mathrm{C}^2(\Omega)$ is called

a) harmonic, if $-\Delta u = 0$ in Ω;

b) subharmonic, if $-\Delta u \leq 0$ in Ω;

c) superharmonic, if $-\Delta u \geq 0$ in Ω.

Remark 4.2. Note that $u \in \mathrm{C}^2(\Omega)$ is harmonic if and only if it is subharmonic and superharmonic.

The following result from Integration Theory will be useful in the sequel:

Lemma 4.3 (Polar coordinates)**.** *Let $x_0 \in \mathbb{R}^d$, $R > 0$ and $f \in \mathrm{C}^1\big(\overline{\mathrm{B}_R(x_0)}\big)$. Then we have*

$$\begin{aligned}\int_{\mathrm{B}_R(x_0)} f(y)\,\mathrm{d}y &= \int_0^R \int_{\partial \mathrm{B}_r(x_0)} f(y)\,\mathrm{d}S(y)\,\mathrm{d}r \\ &= \int_0^R r^{d-1} \int_{\partial \mathrm{B}_1(0)} f(x_0 + rz)\,\mathrm{d}S(z)\,\mathrm{d}r,\end{aligned}$$

where S denotes the surface measure.

Theorem 4.4 (Mean value formula)**.** *Let $u \in \mathrm{C}^2(\Omega)$. Then the following statements are equivalent:*

a) u is harmonic / subharmonic / superharmonic.

b) We have

$$u(x)\begin{Bmatrix} = \\ \leq \\ \geq \end{Bmatrix}\frac{1}{d\omega_d R^{d-1}}\int_{\partial\,\mathrm{B}_R(x)} u(y)\,\mathrm{d}S(y)$$

for all $x \in \Omega$ and for all $R > 0$ such that $\overline{B_R(x)} \subseteq \Omega$, Here ω_d denotes the volume of the d-dimensional unit ball.

c) We have

$$u(x)\begin{Bmatrix} = \\ \leq \\ \geq \end{Bmatrix}\frac{1}{\omega_d R^{d}}\int_{\mathrm{B}_R(x)} u(y)\,\mathrm{d}y$$

for all $x \in \Omega$ and for all $R > 0$ such that $\overline{\mathrm{B}_R(x)} \subseteq \Omega$.

Proof. **a)** $\implies$ **b):** Let $\overline{\mathrm{B}_R(x)} \subseteq \Omega$, $r \in (0, R]$ and define

$$\varphi(r) \coloneqq \frac{1}{d\omega_d r^{d-1}}\int_{\partial\,\mathrm{B}_r(x)} u(y)\,\mathrm{d}S(y) = \frac{1}{d\omega_d}\int_{\partial\,\mathrm{B}_1(0)} u(x+rz)\,\mathrm{d}S(z).$$

Then $\varphi \in \mathrm{C}^1((0, R])$ and

$$\varphi'(r) = \frac{1}{d\omega_d}\int_{\partial\,\mathrm{B}_1(0)} \nabla u(x+rz)\cdot z\,\mathrm{d}S(z).$$

Setting $y \coloneqq x + rz$ this is

$$= \frac{1}{d\omega_d r^{d-1}}\int_{\partial\,\mathrm{B}_r(x)} \nabla u(y)\underbrace{\frac{y-x}{r}}_{=\nu(y)}\,\mathrm{d}S(y),$$

where ν is the unit normal vector. We further get by the Gauß Theorem

$$= \frac{1}{d\omega_d r^{d-1}}\int_{\mathrm{B}_r(x)} \underbrace{\operatorname{div}\nabla u(y)}_{\Delta\, u(y)}\,\mathrm{d}y\begin{Bmatrix} = \\ \geq \\ \leq \end{Bmatrix}0.$$

Furthermore,

$$\begin{aligned}
\lim_{r\searrow 0}\varphi(r) &= \lim_{r\searrow 0}\frac{1}{d\omega_d}\int_{\partial\,\mathrm{B}_1(0)} u(x+rz)\,\mathrm{d}S(z) \\
&= \frac{u(x)}{S(\partial\,\mathrm{B}_1(0))}\int_{\partial\,\mathrm{B}_1(0)} 1\,\mathrm{d}S(z) \\
&= u(x).
\end{aligned}$$

So, $u(x) = \varphi(0)\begin{Bmatrix} = \\ \leq \\ \geq \end{Bmatrix}\varphi(R)$ which implies b).

b) $\implies$ c): Let $\overline{\mathrm{B}_R(x)} \subseteq \Omega$. For all $r \in (0, R]$ the hypothesis b) implies

$$d r^{d-1} u(x) \begin{Bmatrix} = \\ \leq \\ \geq \end{Bmatrix} \frac{1}{\omega_d} \int_{\partial \mathrm{B}_r(x)} u(y) \,\mathrm{d}S(y)$$

Integrating this expression gives

$$R^d u(x) = \int_0^R d r^{d-1} u(x) \,\mathrm{d}r \begin{Bmatrix} = \\ \leq \\ \geq \end{Bmatrix} \frac{1}{\omega_d} \int_0^R \int_{\partial \mathrm{B}_r(x)} u(y) \,\mathrm{d}S(y) \,\mathrm{d}r$$
$$\overset{\text{Lemma 4.3}}{=} \frac{1}{\omega_d} \int_{\mathrm{B}_R(x)} u(y) \,\mathrm{d}y.$$

Hence, c) follows.

c) $\implies$ a): We use a proof by contradiction. Assume that there is some $u \in \mathrm{C}^2(\Omega)$ and some $x_0 \in \Omega$ with

$$-\Delta u(x_0) \begin{Bmatrix} \neq \\ > \\ < \end{Bmatrix} 0.$$

In the case of harmonic u we assume w.l.o.g. that $-\Delta u(x_0) > 0$. The case of $-\Delta u(x_0) < 0$ can be treated analogously. By continuity of Δu, there is some $R > 0$ with $\overline{\mathrm{B}_R(x_0)} \subseteq \Omega$ and $-\Delta u(x) \begin{Bmatrix} > \\ > \\ < \end{Bmatrix} 0$ for all $x \in \overline{\mathrm{B}_R(x_0)}$. For φ as above this yields

$$\varphi'(r) = \frac{1}{d\omega_d r^{d-1}} \int_{\mathrm{B}_r(x_0)} \Delta u(y) \,\mathrm{d}y \begin{Bmatrix} < \\ < \\ > \end{Bmatrix} 0.$$

This implies

$$u(x) = \varphi(0) \begin{Bmatrix} > \\ > \\ < \end{Bmatrix} \varphi(r) = \frac{1}{d\omega_d r^{d-1}} \int_{\partial \mathrm{B}_r(x_0)} u(y) \,\mathrm{d}S(y).$$

We get

$$R^d u(x) = \int_0^R d r^{d-1} u(x) \,\mathrm{d}r \begin{Bmatrix} > \\ > \\ < \end{Bmatrix} \frac{1}{\omega_d} \int_0^R \int_{\partial \mathrm{B}_r(x_0)} u(y) \,\mathrm{d}S(y) \,\mathrm{d}r.$$

Here, Lemma 4.3 gives us

$$= \frac{1}{\omega_d} \int_{\mathrm{B}_R(x_0)} u(y) \,\mathrm{d}y,$$

which is a contradiction to c). □

Theorem 4.5 (Strong maximum principle, version 1)**.** *Let* $u \in \mathrm{C}^2(\Omega)$*.*

a) If u *is subharmonic and* $u(x) = \sup_{y\in\Omega} u(y)$ *for some* $x \in \Omega$*, then* u *is constant.*

b) If u *is superharmonic and* $u(x) = \inf_{y\in\Omega} u(y)$ *for some* $x \in \Omega$*, then* u *is constant.*

In particular, if u *is harmonic and attains either a maximum or a minimum in* Ω*, then* u *is constant.*

Proof. a) Let u be subharmonic and $x \in \Omega$ a point where u attains its maximum $u(x) = S := \sup_{y\in\Omega} u(y) < \infty$ and define $M := \{y \in \Omega\colon u(y) = S\}$.

Since $x \in M$ we know that $M \neq \emptyset$.

M **closed in** Ω: Since S is closed, $M = u^{-1}(\{S\})$ and u is continuous, we conclude that M is closed.

M **open in** Ω: Let $z \in M$ and $r > 0$ such that $\overline{\mathrm{B}_r(z)} \subseteq \Omega$. Then

$$\begin{aligned} S = u(z) &\overset{\text{Thm. 4.4}}{\leq} \frac{1}{\omega_d r^d} \int_{\mathrm{B}_r(z)} \underbrace{u(y)}_{\leq S} \,\mathrm{d}y \\ &\leq S \cdot \frac{1}{\omega_d r^d} \int_{\mathrm{B}_r(z)} \mathrm{d}y = S. \end{aligned}$$

This means that all the inequalities are in fact equalities. This gives us that u is constantly s on a ball around the point z, i.e. $\mathrm{B}_r(z) \subseteq M$.

Ω is connected and we have shown that M is open, closed and non-empty, so $M = \Omega$, i.e. $u(y) = S$ for all $y \in \Omega$.

b) Analogously. □

Corollary 4.6 (Weak maximum principle, comparison principle)**.** *Let* Ω *be bounded and* $u, v \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$*. The we have*

a) If u *is subharmonic, then*

$$\sup_{y\in\overline{\Omega}} u(y) = \sup_{y\in\partial\Omega} u(y).$$

a') If u *is superharmonic, then*

$$\inf_{y\in\overline{\Omega}} u(y) = \inf_{y\in\partial\Omega} u(y).$$

b) If $-\Delta u \leq -\Delta v$ *in* Ω *and* $u \leq v$ *on* $\partial\Omega$*, then* $u \leq v$ *on* $\overline{\Omega}$*.*

Proof. a, a') Let u be subharmonic. Since $u \in \mathrm{C}(\overline{\Omega})$, there exists $x_0 \in \overline{\Omega}$ such that u attains its maximum in x_0, i.e. $u(x_0) = \max_{x\in\overline{\Omega}} u(x)$. If $x_0 \in \partial\Omega$, we are done. Otherwise the claim follows using Theorem 4.5. In the case of superharmonic u, the proof can be done analogously.

b) By assumption, it holds $-\Delta(u-v) \leq 0$ in Ω and $u - v \leq 0$ on $\partial\Omega$. Hence, part a) implies $u - v \leq 0$ on $\overline{\Omega}$. □

Corollary 4.7 (Uniqueness of the Dirichlet problem)**.** *For every $f \in C(\Omega)$ and every $g \in C(\partial\Omega)$ the Dirichlet problem for the Poisson equation*

$$\begin{cases} -\Delta u &= f \text{ in } \Omega \\ \quad u &= g \text{ on } \partial\Omega \end{cases}$$

has at most one solution.

Proof. Let $u, v \in C^2(\Omega) \cap C(\overline{\Omega})$ be solutions of the problem above. Then $-\Delta u = f \leq f = -\Delta v$ in Ω and $u = g \leq g = v$ on $\partial\Omega$ and likewise for u and v switched. Hence Corollary 4.6 yields $u \leq v$ and likewise $v \leq u$ on $\overline{\Omega}$. Together we have $u = v$ on $\overline{\Omega}$. □

We have seen that if there exists a solution for the Dirichlet problem, the solution is unique. The obvious next question to ask is if a solution to this problem exists, or what conditions guarantee existence. Before we turn to this question though, we will tackle the question of uniqueness for solutions to the inhomogeneous Neumann problem. In order to write this down, we have to suppose from now on that $\partial\Omega$ is smooth enough to define a normal vector.

The inhomogeneous Poisson equation with Neumann boundary condition is given by

$$\begin{cases} -\Delta u &= f, \text{ in } \Omega \\ \dfrac{\partial u}{\partial \nu} &= g, \text{ on } \partial\Omega. \end{cases} \tag{inh NL}$$

Assuming the problem is solvable, is the solution unique? Let us assume that both u and v solve (inh NL). Then by linearity of the Laplacian and the normal derivative, we obtain

$$-\Delta(u-v) = -\Delta u + \Delta v = f - f = 0$$

and

$$\frac{\partial(u-v)}{\partial \nu} = \frac{\partial u}{\partial \nu} - \frac{\partial v}{\partial \nu} = g - g = 0.$$

Hence, (inh NL) has a unique solution if the homogeneous Neumann problem

$$\begin{cases} -\Delta u &= 0, \text{ in } \Omega \\ \dfrac{\partial u}{\partial \nu} &= 0, \text{ on } \partial\Omega \end{cases} \tag{hom NL}$$

has only 0 as a solution. Obviously though, (hom NL) is at least solved by all constant functions, so a solution to (inh NL) can at best be unique up to constants. Indeed, Hopf's Lemma will allow us to show that (hom NL) does not possess any other solutions aside from the constant functions.

Theorem 4.8 (Hopf's lemma, strong maximum principle, version 2)**.** *Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$ be subharmonic and let $x_0 \in \partial\Omega$ such that*

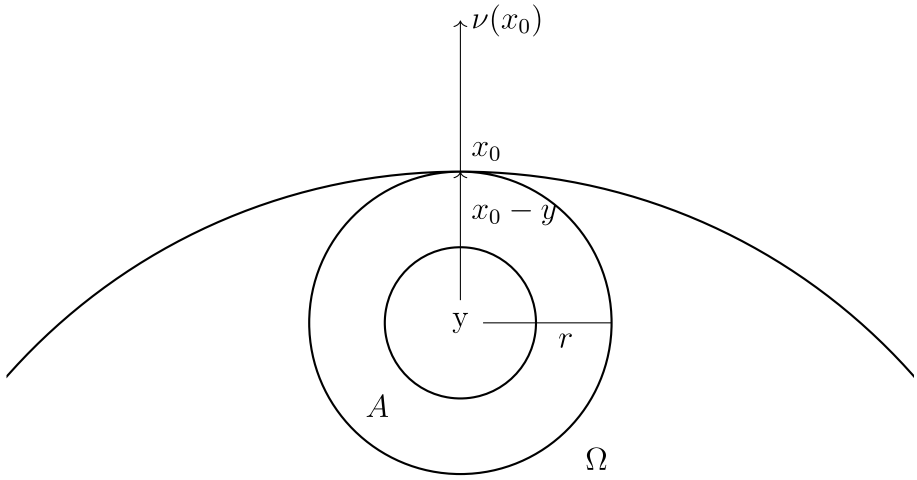

Figure 4.1: Illustration of the set A as defined in the proof of Hopf's lemma 4.8

a) $\nabla u \in \mathrm{C}^1(\Omega, \mathbb{R}^d)$ *has a continuous extension to* x_0*, i.e. the limit*

$$\nabla u(x_0) := \lim_{\Omega \ni x \to x_0} \nabla u(x)$$

exists,

b) $u(x_0) > u(x)$ *for all* $x \in \Omega$*,*

c) there exists $y \in \Omega$ *and* $r > 0$ *such that* $\mathrm{B}_r(y) \subseteq \Omega$ *and* $x_0 \in \partial \mathrm{B}_r(y)$*. ("interior ball condition").*

Then we have $\nabla u(x_0) \cdot \nu(x_0) > 0$*, where* $\nu(x_0)$ *is the unit outer normal vector of* $\mathrm{B}_r(y)$ *at* x_0*.*

Remark. The third requirement arises from a geometrical view. It assures that there is enough space inside Ω close to x_0.
The result of the theorem means that the angle between $\nabla u(x_0)$ and $\nu(x_0)$ is smaller than $\frac{\pi}{2}$. This means that $\nabla u(x_0)$ has to point outwards of Ω.

Proof of Theorem 4.8. Proving that $\nabla u(x_0) \cdot \nu(x_0) \geq 0$ is quite easy. Observe

$$\begin{aligned} \nabla u(x_0) \cdot \nu(x_0) &= \lim_{h \searrow 0} \frac{u(x_0 - h\nu(x_0)) - u(x_0)}{-h} \\ &\overset{(b)}{\geq} \lim_{h \searrow 0} \frac{0}{-h} = 0. \end{aligned}$$

It remains to prove that $\nabla u(x_0) \cdot \nu(x_0) \neq 0$.

Consider $A := \mathrm{B}_r(y) \setminus \mathrm{B}_{\frac{r}{2}}(y)$, cf. Figure 4.1. We claim that there exists a function $h \in \mathrm{C}^2(\overline{A})$ such that

(i) $h = 0$ on $\partial \mathrm{B}_r(y)$ and $h > 0$ in A.

(ii) $-\Delta h \leq 0$ in A.

(iii) $\nabla h(x_0) \cdot \nu(x_0) < 0$.

To show this, we define $h(x) := \mathrm{e}^{-\alpha|x-y|^2} - \mathrm{e}^{-\alpha r^2}$ for some $\alpha > 0$ which we will choose later. Then $h \in \mathrm{C}^2(\overline{A})$ and the following properties hold:

(i) For all $x \in A$ we obtain $|x - y| < r$. This gives us $h(x) > \mathrm{e}^{-\alpha r^2} - \mathrm{e}^{-\alpha r^2} = 0$. In addition, for all $x \in \partial \mathrm{B}_r(y)$ we have $|x - y| = r$, so $h(x) = 0$.

(ii) We compute

$$\begin{aligned}\Delta h(x) &= \sum_{i=1}^{d} \partial_i \left(\mathrm{e}^{-\alpha|x-y|^2}(-2)\alpha(x_i - y_i) \right) \\ &= \sum_{i=1}^{d} \left(\mathrm{e}^{-\alpha|x-y|^2} 4\alpha^2 (x_i - y_i)^2 - 2\alpha \mathrm{e}^{-\alpha|x-y|^2} \right) \\ &= 2\alpha \mathrm{e}^{-\alpha|x-y|^2} \sum_{i=1}^{d} \big(2\alpha(x_i - y_i)^2 - 1 \big) \\ &= 2\alpha \mathrm{e}^{-\alpha|x-y|^2} \big(2\alpha \underbrace{|x - y|^2}_{\geq \frac{r}{2}} - d \big).\end{aligned}$$

Now our definition of A implies

$$\Delta h(x) \geq 2\alpha \mathrm{e}^{-\alpha|x-y|^2} \left(2\alpha \frac{r^2}{4} - d \right) \geq 0,$$

if we choose $\alpha > \frac{2d}{r^2}$.

(iii) We observe that

$$\nabla h(x_0) \cdot \nu(x_0) = -2\alpha e^{-\alpha|x_0-y|^2} \underbrace{(x_0 - y) \cdot \nu(x_0)}_{>0} < 0.$$

This is true since $x_0 - y$ equals (up to renorming) the outer unit normal vector $\nu(x_0)$.

We have $u(x) < u(x_0)$ for all $x \in A$. Since $\partial \mathrm{B}_{\frac{r}{2}}(y) \subseteq A$ is compact, there exists $\varepsilon > 0$ such that

$$g(x) := u(x) - u(x_0) + \varepsilon h(x) \leq 0$$

for all $x \in \partial \mathrm{B}_{\frac{r}{2}}(y)$. On $\partial \mathrm{B}_r(y)$, the function h vanishes, so for $x \in \partial \mathrm{B}_r(y)$ we also have

$$g(x) = u(x) - u(x_0) + \varepsilon h(x) = u(x) - u(x_0) \leq 0,$$

since u is continuous on $\overline{\Omega}$. Using $\partial A = \partial\, \mathrm{B}_{\frac{r}{2}}(y) \cup \partial\, \mathrm{B}_r(y)$, we obtain this inequality for all $x \in \partial A$. Since u is subharmonic, we have

$$-\Delta\, g(x) = -\Delta(u(x) - u(x_0) + \varepsilon h(x)) = -\Delta\, u - \varepsilon\, \Delta\, h \overset{(ii)}{\leq} 0 \text{ in } A.$$

So, g is subharmonic, too. Corollary 4.6 gives us $g \leq 0$ in $\overline{A}$. Furthermore we have

$$g(x_0) = u(x_0) - u(x_0) + \varepsilon h(x_0) = \varepsilon h(x_0) = 0$$

by (i) and since $x_0 \in \partial\, \mathrm{B}_r(y)$. Therefore, g attains its maximum in the point x_0 and hence

$$\nabla g(x_0) \cdot \nu(x_0) = (\nabla u(x_0) + \varepsilon \nabla h(x_0)) \cdot \nu(x_0) \geq 0.$$

Ultimately this yields

$$\nabla u(x_0) \cdot \nu(x_0) \geq -\varepsilon \nabla h(x_0) \cdot \nu(x_0) \overset{(iii)}{>} 0.$$ □

Corollary 4.9. *Let Ω be bounded such that it admits a normal vector and satisfies the interior ball condition in every $x \in \partial\Omega$. If $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$ solves (hom NL), then u is constant.*

Proof. Assume that u is not constant. Since $\partial\Omega$ is compact, there exists $x_0 \in \partial\Omega$ such that $u(x_0) = \sup_{y \in \partial\Omega} u(y)$. By the strong maximum principle in Theorem 4.5 we obtain $u(x) < u(x_0)$ for all $x \in \Omega$. Using Theorem 4.8, this leads to

$$\frac{\partial u}{\partial \nu}(x_0) = \nabla u(x_0) \cdot \nu(x_0) > 0.$$

Hence, u does not fulfill the Neumann boundary condition, so u is no solution for (hom NL). □

The maximum principle not only holds for the Laplace operator, but for a wide variety of elliptic differential operators that we introduce now.

Definition 4.10. Let $a_{ij}, b_j, c \in \mathrm{L}^\infty(\Omega) \cap \mathrm{C}(\Omega)$ for $i, j \in \{1, \dots, d\}$ such that

$$A(x) = (a_{ij}(x))_{i,j=1}^d \in \mathbb{R}^{d \times d}$$

is symmetric and uniformly positive definite, i.e. there exists some $\alpha_0 > 0$ such that for all $\xi \in \mathbb{R}^d$ and all $x \in \Omega$

$$\langle \xi, A(x)\xi \rangle = \xi^{\mathrm{T}} A(x)\xi \geq \alpha_0 \, |\xi|^2 \qquad \text{(ellipticity condition)}$$

is satisfied. Then

$$Lu(x) := -\sum_{i,j=1}^d a_{ij}(x)\partial_i\partial_j u(x) + \sum_{j=1}^d b_j(x)\partial_j u(x) + c(x)u(x)$$

is called elliptic differential operator of second order.

Remark 4.11. a) $A(x) = I, b = 0, c = 0$ yields the (negative) Laplace operator.

b) Constant coefficients: Let a_{ij} be constants $b = 0, c = 0$. Furthermore, let $Q \in \mathbb{R}^{d\times d}$ be invertible, $y := Qx$ and $v(y) := u(Q^{-1}y)$. Then

$$\sum_{i,j=1}^{d} a_{ij}\partial_i\partial_j u(x) = \sum_{i,j=1}^{d} \tilde{a}_{ij}\partial_i\partial_j(x)v(y)$$

for $(\tilde{a}_{ij})_{i,j} = \tilde{A} = QAQ^{\mathrm{T}}$.

A is positive definite, so there exists an orthogonal matrix $S \in \mathbb{R}^{d\times d}$ such that $SAS^{\mathrm{T}} = D = \mathrm{diag}(\lambda_1, \dots, \lambda_d)$ where $\lambda_1, \dots, \lambda_d$ are the eigenvalues of A. Set $Q := D^{-\frac{1}{2}}S$. Then

$$\tilde{A} = QAQ^{\mathrm{T}} = D^{-\frac{1}{2}}SAS^{\mathrm{T}}D^{-\frac{1}{2}} = D^{-\frac{1}{2}}DD^{-\frac{1}{2}} = I.$$

This shows that in this case we can use our theory we have developed for the Laplace operator.

Theorem 4.12 (Weak maximum principle)**.** *Let Ω be bounded, L be an elliptic differential operator and $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$. Then*

a) If $c = 0$ and $Lu \le 0$ in Ω, then $\sup_{\overline{\Omega}} u = \sup_{\partial\Omega} u$.

a') If $c = 0$ and $Lu \ge 0$ in Ω, then $\inf_{\overline{\Omega}} u = \inf_{\partial\Omega} u$.

b) If $c \ge 0$ and $Lu \le 0$ in Ω, then $\sup_{\overline{\Omega}} u \le \sup_{\partial\Omega} u^+$, where $u^+ := \max\{u, 0\}$ is the positive part of u.

b') If $c \ge 0$ and $Lu \ge 0$ in Ω, then $\inf_{\overline{\Omega}} u \ge \inf_{\partial\Omega} u^-$, where $u^- := \min\{u, 0\}$ is the negative part of u.[1]

Proof. Exercise. □

Corollary 4.13. *Let Ω be bounded, L an elliptic differential operator with $c \ge 0$. Then it holds*

a) Uniqueness: For $f \in \mathrm{C}(\Omega), g \in \mathrm{C}(\partial\Omega)$, there exists at most one solution $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$ to

$$\begin{cases} Lu &= f \text{ in } \Omega \\ \quad u &= g \text{ on } \partial\Omega \end{cases} \tag{Ell PDE}$$

b) Comparison principle: If $v, w \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$ fulfill $Lv \le Lw$ on Ω and $v \le w$ on $\partial\Omega$, then $v \le w$ on $\overline{\Omega}$.

[1] This notation is dangerous, in contrast to the integration theory course u^- is a *negative* function.

c) *Sandwich Lemma: Let $f \in \mathrm{C}(\Omega), g \in \mathrm{C}(\partial\Omega)$ and $u, v, w \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$ such that*

$$\begin{cases} Lv & \leq f \leq Lw \text{ in } \Omega \\ v & \leq g \leq w \text{ on } \partial\Omega \end{cases}$$

and u a solution to (Ell PDE)*, then $v \leq u \leq w$ on $\overline{\Omega}$.*

Remark 4.14. The condition $c \geq 0$ in Theorem 4.12 cannot be omitted. To see this, consider for some $\lambda > 0$ the equation

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

where $\Omega = (0,1) \times (0,1) \subseteq \mathbb{R}^2$. So, we have $c(x) = -\lambda < 0$. Take $k, l \in \mathbb{N}_0$, but not $l = k = 0$, and set $u(x_1, x_2) \coloneqq \sin(k\pi x_1) \cdot \sin(l\pi x_2)$ where $(x_1, x_2) \in \overline{\Omega}$, cf. Figure 4.2. Then $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$ and it is easy to compute that u is a solution to the equation above for $\lambda_{k,l} \coloneqq (k^2 + l^2)\pi^2$. But we know that there is also the zero solution which means that we do not have uniqueness and hence the maximum principle cannot hold.

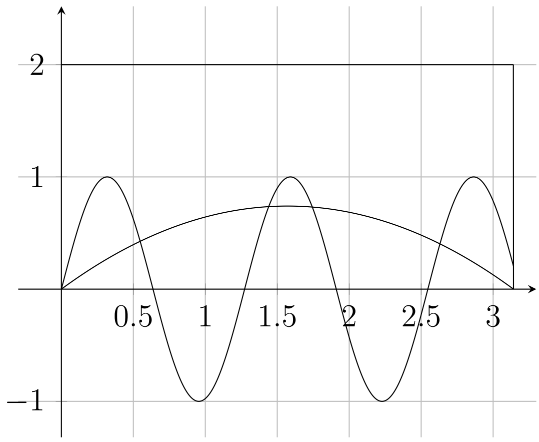

Figure 4.2: Visualisation of u

5 L^2-Theory: Introduction and Motivation

For this chapter, let $\Omega \subseteq \mathbb{R}^d$ be an open, connected and bounded set and $f \in \mathrm{C}(\overline{\Omega})$. Consider the Poisson equation

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

This equation also has roots in physics. It arises from minimizing the energy functional

$$E(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x.$$

Theorem 5.1 (Dirichlet principle)**.** *Let Ω have smooth boundary and define*

$$V := \{v \in \mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega})^{1} \colon v = 0 \text{ on } \partial\Omega\}.$$

Then $u \in V$ solves (PE)*, if and only if $E(u) = \min\{E(v) \colon v \in V\}$.*

In order to prove this theorem, we need the following fundamental lemma of calculus of variations:

Proposition 5.2 (Fundamental lemma of calculus of variations)**.** *For $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ we have the following equivalence:*

$$\int_\Omega u(x)\varphi(x)\,\mathrm{d}x = 0 \text{ for all } \varphi \in \mathcal{C}^\infty_\mathrm{c}(\Omega) \iff u = 0 \text{ a.e. on } \Omega.$$

Proof. **"$\Longleftarrow$":** Clear.

"$\Longrightarrow$": Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ such that for all $\varphi \in \mathcal{C}^\infty_\mathrm{c}(\Omega)$ it holds that

$$\int_\Omega u(x)\varphi(x)\,\mathrm{d}x = 0.$$

For $K \subseteq \Omega$ compact, we define

$$f(x) := \begin{cases} \operatorname{sign}(u(x)) & \text{if } x \in K \\ 0 & \text{if } x \notin K. \end{cases}$$

[1] No, this is not a mistake. The term $\dfrac{\partial u}{\partial \nu}$ appears later in the proof after using Green's formula. Therefore, we need u to be differentiable on $\partial\Omega$.

Then $u(x)f(x) = |u(x)|$ for almost every $x \in K$. Let $(\eta_\varepsilon)_{\varepsilon>0}$ be a standard mollifier and $\varphi_\varepsilon := \eta_\varepsilon * f$. Then

- $\varphi_\varepsilon \in \mathcal{C}_c^\infty(\mathbb{R}^d)$
- $\operatorname{supp}(f) \subseteq K \Subset \Omega$, so $\operatorname{supp}(\varphi_\varepsilon) \subseteq \Omega$ for ε small enough
- $\varphi_{\frac{1}{n}} \overset{n\to\infty}{\longrightarrow} f$ in $L^1(\Omega)$, so there exists a subsequence $\left(\varphi_{\frac{1}{n_k}}\right)$ with $\varphi_{\frac{1}{n_k}} \overset{k\to\infty}{\longrightarrow} f$ a.e. in Ω.

We aim for

$$\begin{aligned} 0 &= \lim_{k\to\infty} \int_\Omega u(x)\varphi_{\frac{1}{n_k}}(x)\,\mathrm{d}x \\ &\overset{?}{=} \int_\Omega u(x)f(x)\,\mathrm{d}x \\ &= \int_K \underbrace{|u(x)|}_{\geq 0}\,\mathrm{d}x \\ &\implies u = 0 \text{ a.e. on } K. \end{aligned}$$

For "?", we only have to show that $\varphi_{\frac{1}{n_k}}$ is bounded independently of k, since $\varphi_{\frac{1}{n_k}}$ has compact support for every $n_k \in \mathbb{N}$ and $u \in L^1_{\mathrm{loc}}(\Omega)$. If we have shown this, we can apply the dominated convergence theorem to obtain "?".

$$\begin{aligned} \left|\varphi_{\frac{1}{n_k}}(x)\right| &\leq \int_\Omega \eta_{\frac{1}{n_k}}(x-y)\,|f(y)|\,\mathrm{d}y \\ &\leq \underbrace{\|f\|_\infty}_{\leq 1} \int_{\mathbb{R}^d} \eta_{\frac{1}{n_k}}(x-y)\,\mathrm{d}y \\ &\leq 1. \end{aligned}$$

Using this argument for the compact subsets

$$K_n := \{x \in \Omega\colon\ \operatorname{dist}(x, \partial\Omega) \geq \frac{1}{n}, |x| \leq n\}$$

of Ω proves the assertion since $\bigcup_n K_n = \Omega$ and countable unions of null sets are null sets. □

Remark. Note that the proof of this Proposition does need neither the boundedness nor the connectivity of Ω, so the fundamental lemma of calculus of variations remains true for Ω an arbitrary open subset of $\mathbb{R}^d$.

Proof of Theorem 5.1. **"$\Longrightarrow$":** Let $u \in V$ solve (PE). Then for all $v \in V$ and for all $t \in \mathbb{R}$ we have

$$\begin{aligned} E(u+tv) &= \frac{1}{2}\int_\Omega (\nabla u + t\nabla v)(\nabla u + t\nabla v)\,\mathrm{d}x - \int_\Omega fu\,\mathrm{d}x - t\int_\Omega fv\,\mathrm{d}x \\ &= E(u) + \frac{1}{2}2t\int_\Omega \nabla u\nabla v\,\mathrm{d}x + \underbrace{\frac{t^2}{2}\int_\Omega |\nabla v|^2\,\mathrm{d}x}_{\geq 0} - t\int_\Omega fv\,\mathrm{d}x \\ &\geq E(u) + t\int_\Omega \nabla u\nabla v\,\mathrm{d}x - t\int_\Omega fv\,\mathrm{d}x \end{aligned}$$

Now we integrate by parts using Green's formula and get

$$\begin{aligned} &\geq E(u) + t\int_\Omega (-\Delta u)v\,\mathrm{d}x + t\int_{\partial\Omega} \underbrace{v}_{=0}\frac{\partial u}{\partial \nu}\,\mathrm{d}S - t\int_\Omega fv\,\mathrm{d}x \\ &= E(u) + t\int_\Omega \underbrace{(-\Delta u - f)}_{=0}\,v\,\mathrm{d}x \\ &= E(u). \end{aligned}$$

Thus, for all $w \in V$ we obtain

$$E(w) = E(u + 1(\underbrace{w-u}_{\in V})) \geq E(u),$$

so u minimizes E in V.

"$\Longleftarrow$": Let $u \in V$ with $E(u) = \min_{v\in V} E(v)$. Then for all $t \in \mathbb{R}$ and all $v \in \mathcal{C}_c^\infty(\Omega)$ we have that $u + tv \in V$ and as before we calculate

$$e(t) := E(u+tv) = E(u) + t\left(\int_\Omega (-\Delta u)v\,\mathrm{d}x - \int_\Omega fv\,\mathrm{d}x\right) + \frac{t^2}{2}\int_\Omega |\nabla v|^2\,\mathrm{d}x.$$

By assumption, $e'(0) = 0$, so for all $v \in \mathcal{C}_c^\infty(\Omega)$ we have

$$\int_\Omega (-\Delta u - f)v\,\mathrm{d}x = e'(0) = 0.$$

Thus, Proposition 5.2 implies $-\Delta u - f = 0$, so u solves (PE). □

Remark. **Question**: Is there always some $u \in V$ such that $E(u) = \inf_{v\in V} E(v)$?

This question was negatively answered by Weierstraß in 1869. The problem in modern language is the following: One can show that every minimizing sequence for E, i.e. $(v_n) \subseteq V$ such that $E(v_n) \to \inf_{v\in V} E(v)$ for $n \to \infty$, is a Cauchy sequence with respect to $\|v\|_V := \|\nabla v\|_{L^2(\Omega)}$ but $(V, \|\cdot\|_V)$ is not complete.

With today's insights it seems appropriate to look for a complete space $Z \supseteq V$. But at the end of the nineteenth century, notions as completions were not known, so the problem was formulated differently by Courant and Hilbert around 1900:

If $u \in V$ solves (PE), then for all $v \in V$, we have

$$\int_\Omega (-\Delta u) \cdot v \,\mathrm{d}x = \int_\Omega f v \,\mathrm{d}x.$$

Applying Green's formula we obtain

$$\int_\Omega \nabla u \nabla v \,\mathrm{d}x - \int_{\partial\Omega} v \cdot \frac{\partial u}{\partial \nu} \,\mathrm{d}S(x) = \int_\Omega f v \,\mathrm{d}x.$$

Hence, if u is a solution for (PE), then u has to fulfill

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega f \cdot v \,\mathrm{d}x \text{ for all } v \in V. \tag{5.1}$$

This is the so called "weak formulation" of our problem and we will now reformulate this in an abstract functional analytic language. For this we would like to consider a space like

$$\mathrm{H} = \mathrm{H}_0^1(\Omega) = \{u \in \mathrm{L}^2(\Omega) \colon \nabla u \in \mathrm{L}^2(\Omega) \text{ and } u = 0 \text{ on } \partial\Omega\},$$

but of course, $\partial\Omega$ is a null set, so we cannot express the condition $u = 0$ on $\partial\Omega$ in this way. In addition, we have to define what ∇u means.

Ignoring these problems for the moment we note that for a given function f, the map $\mathrm{H} \ni v \mapsto \int_\Omega f v \,\mathrm{d}x$ on the right hand side of our weak formulation is a linear form F on H. In the same way we can interpret the left hand side $\mathrm{H} \times \mathrm{H} \ni (u, v) \mapsto \int_\Omega \nabla u \nabla v \,\mathrm{d}x$ as a bilinear form a on H. So the weak formulation may be rephrased as: Given a bilinear form $a\colon \mathrm{H} \times \mathrm{H} \to \mathbb{R}$ and a linear form $F\colon \mathrm{H} \to \mathbb{R}$, find some $u \in \mathrm{H}$ with

$$a(u, v) = F(v) \text{ for all } v \in \mathrm{H}. \tag{V}$$

This is called a variational equation.

We see that to fill in the gaps in the above plan we will need both integration theory ($\mathrm{L}^2, \mathrm{H}, \ldots$) and functional analysis ((bi)linear forms,...). The following tools to solve this sort of problems should be known from the course on functional analysis.

Definition 5.3. Let H be a real (complex) Hilbert space with scalar product $(\cdot, \cdot)$ and let $a\colon H \times H \to \mathbb{R}$ ($\mathbb{C}$) be a bilinear (sesquilinear) form. Then a is called

a) continuous, if there exists $C \geq 0$ with $|a(u, v)| \leq C \|u\|_H \cdot \|v\|_H$ for all $u, v \in H$.

b) coercive, if there exists $\alpha_0 > 0$ such that $\mathrm{Re}(a(u, u)) \geq \alpha_0 \|u\|_H^2$ for all $u \in H$.

Theorem 5.4 (Lax-Milgram)**.** *Let H be a Hilbert space, let $a\colon H \times H \to \mathbb{R}$ ($\mathbb{C}$) be a continuous and coercive bilinear (sesquilinear) form with constants C and α_0 from Definition 5.3 and let F be a continuous linear form on H. Then the variational problem (V) has a unique solution $u \in H$ and it holds $\|u\|_H \leq \frac{1}{\alpha_0} \|F\|_{H'}$. In particular the solution operator $S : H' \to H$, $F \mapsto u$ is linear and bounded with $\|S\| \leq \frac{1}{\alpha_0}$.*

Furthermore, there is a unique $A \in \mathcal{L}(H)$ such that $a(u, v) = (u, Av)$ for all $u, v \in H$, A is invertible, $\|A\| \leq C$ and $\|A^{-1}\| \leq \frac{1}{\alpha_0}$.

6 Sobolev spaces

Consider the space

$$\mathrm{H} = \mathrm{H}_0^1(\Omega) = \{u \in \mathrm{L}^2(\Omega) \colon \nabla u \in \mathrm{L}^2(\Omega) \text{ and } u = 0 \text{ on } \partial\Omega\}$$

of the previous chapter. We still have various problems to solve, i.e.

- What does ∇u mean for an L^2 function u, or more general, what does differentiability mean for a function $u \in \mathrm{L}^2(\Omega)$ (or even for $u \in \mathrm{L}^p(\Omega)$)?

- Since $\partial\Omega$ is a null set, how can we assign boundary values to $u \in H$?

In this chapter, we will deal with the problem of differentiability. Clearly, a new concept of differentiability is needed, so-called *weak derivatives*. This will lead to *Sobolev spaces*, which are the right function spaces to work with in this context.

Let $\Omega \subseteq \mathbb{R}^d$ be open. In order to define weak derivatives, recall (5.1) from the previous chapter.

Definition 6.1. Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ and $\alpha \in \mathbb{N}_0^d$ a multi-index. We say that u has an α^{th} weak derivative if there is some $v \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ such that for all $\varphi \in \mathcal{C}_\mathrm{c}^\infty(\Omega)$ we have

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

In this case, we define $\mathrm{D}^\alpha u \coloneqq v$.

Proposition 6.2 (Uniqueness weak derivative and comparability with classical ones)**.** *Let $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ and $\alpha \in \mathbb{N}_0^d$ be a multi-index.*

a) If $\mathrm{D}^\alpha u$ exists, it is uniquely determined almost everywhere in Ω.

b) If $u \in \mathrm{C}^{|\alpha|}(\Omega)$, then $\mathrm{D}^\alpha u$ exists and is equal to the classical derivative.

Proof. a) Let $v, w \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ be α^{th} weak derivatives of u. Then we know for all $\varphi \in \mathcal{C}_\mathrm{c}^\infty(\Omega)$:

$$\int_\Omega (v - w)\varphi = (-1)^{|\alpha|} \int_\Omega (u - u) \,\mathrm{D}^\alpha \varphi = 0.$$

By Proposition 5.2 this implies $v - w = 0$ almost everywhere in Ω, so $v = w$ almost everywhere in Ω.

b) If $u \in \mathrm{C}^{|\alpha|}$, then for all $\varphi \in \mathcal{C}_c^\infty(\Omega)$ it holds

$$\int_\Omega u(x)\,\mathrm{D}^\alpha \varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega \mathrm{D}^\alpha u(x)\varphi(x)\,\mathrm{d}x$$

by integration by parts. Thus, the classical derivative satisfies the definition for the weak derivative. By part a), the classical derivative and the weak derivative coincide. □

Example. Having introduced the concept of weak derivatives we are now able to differentiate functions like $|x|$ in a weak sense. Since $|x|$ is differentiable on $\mathbb{R} \setminus \{0\}$, the weak derivative has to coincide with $\operatorname{sign}(x)$ on $\mathbb{R} \setminus \{0\}$. Since $\{0\}$ is a null set, we can extend it by an arbitrary value in 0, for example by 0 to obtain $(|x|)' = \operatorname{sign}(x)$. This satisfies the condition for weak derivatives, since

$$\int_\mathbb{R} \operatorname{sign}(x)\varphi(x)\,\mathrm{d}x = -\int_\mathbb{R} |x| \cdot \varphi'(x)\,\mathrm{d}x$$

for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$.

Having seen weak derivatives we can define the so-called Sobolev spaces.

Definition 6.3. Let $k \in \mathbb{N}$ and $p \in [1, \infty]$. Then we define the Sobolev spaces

a) We set

$$\mathrm{W}^{k,p}(\Omega) \coloneqq \{u \in \mathrm{L}^p(\Omega) \colon \mathrm{D}^\alpha u \text{ exists and } \mathrm{D}^\alpha u \in \mathrm{L}^p(\Omega) \text{ for all } |\alpha| \le k\}$$

and equip it with the norm

$$\|u\|_{\mathrm{W}^{k,p}(\Omega)} \coloneqq \Big(\sum_{|\alpha|\le k} \|\mathrm{D}^\alpha u\|_{\mathrm{L}^p(\Omega)}^p\Big)^{\frac{1}{p}} \text{ for } p \in [1,\infty),$$

respectively

$$\|u\|_{\mathrm{W}^{k,\infty}(\Omega)} \coloneqq \max_{|\alpha|\le k} \|\mathrm{D}^\alpha u\|_{\mathrm{L}^\infty(\Omega)}.$$

For $p = 2$ we define $\mathrm{H}^k(\Omega) \coloneqq \mathrm{W}^{k,2}(\Omega)$.

b) $\mathrm{W}_0^{k,p}(\Omega) \coloneqq \overline{\mathcal{C}_c^\infty(\Omega)}^{\|\cdot\|_{\mathrm{W}^{k,p}(\Omega)}}$. For $p = 2$ we define $\mathrm{H}_0^k(\Omega) \coloneqq \mathrm{W}_0^{k,2}(\Omega)$.

c) $\mathrm{W}_{\mathrm{loc}}^{k,p}(\Omega) \coloneqq \bigcap_{V\Subset\Omega} \mathrm{W}^{k,p}(V)$. For $p = 2$ we again denote:

$$\mathrm{H}_{\mathrm{loc}}^k(\Omega) \coloneqq \mathrm{W}_{\mathrm{loc}}^{k,2}(\Omega).$$

We say that $(u_n) \subseteq \mathrm{W}_{\mathrm{loc}}^{k,p}(\Omega)$ converges to $u \in \mathrm{W}_{\mathrm{loc}}^{k,p}(\Omega)$ if $u_n \to u$ in $\mathrm{W}^{k,p}(V)$ for all $V \Subset \Omega$.

Remark 6.4. a) As for L^p-functions, functions in $\mathrm{W}^{k,p}(\Omega)$ are equivalence classes. $f \in \mathrm{W}^{k,p}(\Omega)$ continuous means that f has a continuous representative.

b) We have $u \in \mathrm{W}_0^{k,p}(\Omega)$ if and only if there exists a sequence $(u_n) \subseteq \mathcal{C}_c^\infty(\Omega)$ such that $u_n \xrightarrow{n\to\infty} u$ in $\mathrm{W}^{k,p}(\Omega)$.

Theorem 6.5. *Let $k \in \mathbb{N}$ and $p \in [1,\infty]$. Then $\left(\mathrm{W}^{k,p}(\Omega), \|\cdot\|_{\mathrm{W}^{k,p}(\Omega)}\right)$ and $\left(\mathrm{W}_0^{k,p}(\Omega), \|\cdot\|_{\mathrm{W}^{k,p}(\Omega)}\right)$ are*

a) Banach spaces.

b) separable, if $p < \infty$.

c) reflexive, if $1 < p < \infty$.

d) Hilbert spaces, if $p = 2$ with the scalar product

$$\langle u, v\rangle_{\mathrm{H}^k(\Omega)} := \sum_{|\alpha|\le k} \langle \mathrm{D}^\alpha u, \mathrm{D}^\alpha v\rangle_{\mathrm{L}^2(\Omega)}\,.$$

Proof. a) We leave it to the reader to verify that $\mathrm{W}^{k,p}(\Omega)$ is a normed vector space for every $p \in [1,\infty]$. To show the completeness of $\mathrm{W}^{k,p}(\Omega)$ consider a Cauchy sequence $(u_n)_n$ in $\mathrm{W}^{k,p}(\Omega)$. Then for all $|\alpha| \le k$ and for all $n, m \in \mathbb{N}$ we calculate

$$\begin{aligned}\|\mathrm{D}^\alpha u_n - \mathrm{D}^\alpha u_m\|_{\mathrm{L}^p(\Omega)} &\le \Big(\sum_{|\beta|\le k} \|\mathrm{D}^\beta u_n - \mathrm{D}^\beta u_m\|_{\mathrm{L}^p(\Omega)}^p\Big)^{\frac{1}{p}} \\ &= \|u_n - u_m\|_{\mathrm{W}^{k,p}(\Omega)}\,.\end{aligned}$$

Hence, for all $|\alpha| \le k$, the sequence $(\mathrm{D}^\alpha u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathrm{L}^p(\Omega)$. Since $\mathrm{L}^p(\Omega)$ is complete, for all $|\alpha| \le k$ there exists a function $u_\alpha \in \mathrm{L}^p(\Omega)$ such that $\mathrm{D}^\alpha u_n \to u_\alpha$ in $\mathrm{L}^p(\Omega)$. Set $u := u_{(0,0,\dots,0)} \in \mathrm{L}^p(\Omega)$. Then for all $|\alpha| \le k$ and for all $\varphi \in \mathcal{C}_c^\infty(\Omega)$ we obtain

$$\begin{aligned}\int_\Omega u(x)\,\mathrm{D}^\alpha \varphi(x)\,\mathrm{d}x &= \lim_{n\to\infty} \int_\Omega u_n(x)\,\mathrm{D}^\alpha \varphi(x)\,\mathrm{d}x \\ &= \lim_{n\to\infty} (-1)^{|\alpha|} \int_\Omega \mathrm{D}^\alpha u_n(x)\varphi(x)\,\mathrm{d}x \\ &= (-1)^{|\alpha|} \int_\Omega \lim_{n\to\infty} \mathrm{D}^\alpha u_n(x)\varphi(x)\,\mathrm{d}x \\ &= (-1)^{|\alpha|} \int_\Omega u_\alpha(x)\varphi(x)\,\mathrm{d}x.\end{aligned}$$

Hence, $\mathrm{D}^\alpha u = u_\alpha \in \mathrm{L}^p(\Omega)$ for all $|\alpha| \le k$. Thus, $u \in \mathrm{W}^{k,p}(\Omega)$ and

$$\|u_n - u\|_{\mathrm{W}^{k,p}(\Omega)} = \Big(\sum_{|\alpha|\le k} \|\mathrm{D}^\alpha u_n - u_\alpha\|_{\mathrm{L}^p(\Omega)}^p\Big)^{\frac{1}{p}} \to 0$$

for $n \to \infty$ since $\mathrm{D}^\alpha u_n \to u_\alpha$ in $\mathrm{L}^p(\Omega)$.

b),c),d) The idea is to identify the Sobolev space $\mathrm{W}^{k,p}(\Omega)$ as a closed subspace of some space that behaves like $\mathrm{L}^p(\Omega)$, since this space satisfies the claims b), c) and d). Notice that, due to $\mathrm{W}_0^{k,p}(\Omega)$ being a closed subspace of $\mathrm{W}^{k,p}(\Omega)$, the same argument yields the claim for this space.

In order to execute this plan, consider the mapping $\mathrm{W}^{k,p}(\Omega) \ni u \mapsto (\mathrm{D}^\alpha u)_{|\alpha| \le k} \in \mathrm{L}^p(\Omega)^N$, where N is the number of multiindices with length less or equal to k. This mapping is continuous and hence $\mathrm{W}^{k,p}(\Omega)$ is a closed subspace of $\mathrm{L}^p(\Omega)^N$ for some $N \in \mathbb{N}$. □

Proposition 6.6 (Product rule)**.** *Let $p, q, r \in [1, \infty]$ with $1/p + 1/q = 1/r$, $f \in \mathrm{W}^{1,p}(\Omega)$ and $g \in \mathrm{W}^{1,q}(\Omega)$. Then $fg \in \mathrm{W}^{1,r}(\Omega)$ and $\partial_j(fg) = \partial_j f \cdot g + f \cdot \partial_j g$.*

Remark. As usual, iterating the product rule gives the usual Leibniz rule for higher derivatives.

Proposition 6.7 (Chain rule)**.** *Let $D \subseteq \mathbb{R}^d$ open and $\Phi : D \to \Omega$ a C^1-diffeomorphism such that $D\Phi$ and $D\Phi^{-1}$ are bounded functions.*[1] *If $p \in [1, \infty]$ and $f \in \mathrm{W}^{1,p}(\Omega)$, then $f \circ \Phi \in \mathrm{W}^{1,p}(D)$ and*

$$\partial_j(f \circ \Phi) = (\nabla f \circ \Phi) \cdot \partial_j \Phi.$$

Theorem 6.8 (Poincaré inequality)**.** *Let Ω be bounded, $p \in [1, \infty)$ and $r = \operatorname{diam}(\Omega)$. Then*

$$\|u\|_{\mathrm{L}^p(\Omega)} \le r \, \|\nabla u\|_{\mathrm{L}^p(\Omega)}$$

for all $u \in \mathrm{W}_0^{1,p}(\Omega)$.

Proof. W.l.o.g. let $\Omega \subseteq [0, r]^d$. Let $u \in \mathcal{C}_c^\infty(\Omega)$ and $\tilde{u}$ its extension by 0 to $\mathbb{R}^d$. Then $\tilde{u} \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and for all $x \in \Omega$ it holds

$$\begin{aligned} u(x) &= \tilde{u}(x) \\ &= \underbrace{\tilde{u}(0, x_2, \ldots, x_d)}_{=0} + \int_0^{x_1} \partial_1 \tilde{u}(t, x_2, \ldots, x_d) \, \mathrm{d}t. \end{aligned}$$

Thus,

$$\begin{aligned} |u(x)|^p &\le \left(\int_0^{x_1} |\partial_1 \tilde{u}(t, x_2, \ldots, x_d)| \, \mathrm{d}t \right)^p \\ &\le \left(\int_0^{r} 1 \cdot |\partial_1 \tilde{u}(t, x_2, \ldots, x_d)| \, \mathrm{d}t \right)^p \end{aligned}$$

[1] This assures $\partial_j(f \circ \Phi) \in \mathrm{W}^{1,p}(D)$. The problem in this case is integrability, which can now be shown via Hölder's inequality.

By Hölder's inequality we obtain for the Hölder conjugate q of p (i.e. $\frac{1}{p}+\frac{1}{q}=1$)

$$
\begin{aligned}
&\leq \left[\left(\int_0^r 1^q \,\mathrm{d}t\right)^{\frac{1}{q}} \left(\int_0^r \left|\partial_1 \tilde{u}(t,x_2,\dots,x_d)\right|^p \,\mathrm{d}t\right)^{\frac{1}{p}}\right]^p \\
&= r^{\frac{p}{q}} \int_0^r \left|\partial_1 \tilde{u}(t,x_2,\dots,x_d)\right|^p \,\mathrm{d}t.
\end{aligned}
$$

This implies

$$
\begin{aligned}
\int_0^r |\tilde{u}(x)|^p \,\mathrm{d}x_1 &\leq \int_0^r r^{\frac{p}{q}} \int_0^r \left|\partial_1 \tilde{u}(t,x_2,\dots,x_d)\right|^p \,\mathrm{d}t\,\mathrm{d}x_1 \\
&= r^{1+\frac{p}{q}} \int_0^r \left|\partial_1 \tilde{u}(t,x_2,\dots,x_d)\right|^p \,\mathrm{d}t \\
&= r^p \int_0^r \left|\partial_1 \tilde{u}(t,x_2,\dots,x_d)\right|^p \,\mathrm{d}t.
\end{aligned}
$$

This yields the desired estimate by integration over the whole cube, which, once again, uses the fact that the support of $\tilde{u}$ is compactly contained in Ω.

$$
\begin{aligned}
\|u\|_{\mathrm{L}^p(\Omega)}^p &= \int_0^r \dots \int_0^r |\tilde{u}(x)|^p \,\mathrm{d}x_1 \dots \mathrm{d}x_d \\
&\leq r^p \int_0^r \dots \int_0^r \int_0^r \left|\partial_1 \tilde{u}(t,x_2,\dots,x_d)\right|^p \,\mathrm{d}t\,\mathrm{d}x_2 \dots \mathrm{d}x_d \\
&= r^p \left\|\partial_1 \tilde{u}\right\|_{\mathrm{L}^p([0,r]^d)}^p \\
&\leq r^p \left\|\nabla \tilde{u}\right\|_{\mathrm{L}^p((0,r)^d)}^p \\
&= r^p \left\|\nabla u\right\|_{\mathrm{L}^p(\Omega)}^p .
\end{aligned}
$$

Finally, for all $u \in \mathcal{C}_{\mathrm{c}}^\infty(\Omega)$ we obtain

$$
\|u\|_{\mathrm{L}^p(\Omega)} \leq r \cdot \|\nabla u\|_{\mathrm{L}^p(\Omega)} .
$$

For $u \in \mathrm{W}_0^{1,p}(\Omega)$ take $(u_n) \subseteq \mathcal{C}_{\mathrm{c}}^\infty(\Omega)$ with $u_n \to u$ in $\mathrm{W}^{1,p}(\Omega)$. Then

$$
\|u\|_{\mathrm{L}^p(\Omega)} = \lim_{n\to\infty} \|u_n\|_{\mathrm{L}^p(\Omega)} \leq \lim_{n\to\infty} r\,\|\nabla u_n\|_{\mathrm{L}^p(\Omega)} = r\,\|\nabla u\|_{\mathrm{L}^p(\Omega)} . \qquad \square
$$

Remark 6.9. a) • There are no conditions on $\partial\Omega$.

- The boundedness condition is not necessary in this shape. All the proof required was the existence of *one* coordinate in which the domain is bounded. Typical examples include strips, tubes,...
- In the case of $\Omega = \mathbb{R}^d$ the theorem fails.
- It is not valid in $\mathrm{W}^{1,p}(\Omega)$, even for bounded Ω. To see this, consider $u = 1$. This implies $\mathrm{W}_0^{1,p}(\Omega) \subsetneq \mathrm{W}^{1,p}(\Omega)$ for bounded Ω.

b) On $\mathrm{W}_0^{1,p}(\Omega)$, we can define a norm by $|||u|||_{\mathrm{W}_0^{1,p}(\Omega)} := \|\nabla u\|_{\mathrm{L}^p(\Omega)}$. This norm is equivalent to

$$\|u\|_{\mathrm{W}^{1,p}(\Omega)} = \Big(\underbrace{\|u\|_{\mathrm{L}^p(\Omega)}^p}_{\leq r\|\nabla u\|_{\mathrm{L}^p(\Omega)}^p} + \|\nabla u\|_{\mathrm{L}^p(\Omega)}^p \Big)^{\frac{1}{p}}.$$

Example 6.10. Consider the form $a\colon \mathrm{H}_0^1(\Omega) \times \mathrm{H}_0^1(\Omega) \to \mathbb{R}$ defined by

$$a(u,v) := \int_\Omega \nabla u \nabla v \,\mathrm{d}x.$$

Then we obtain the following:

- It holds

$$\begin{aligned} |a(u,v)| = \left| \int_\Omega \nabla u \nabla v \right| &\overset{\text{CS}}{\leq} \|\nabla u\|_{\mathrm{L}^2(\Omega)} \cdot \|\nabla v\|_{\mathrm{L}^2(\Omega)} \\ &\leq \Big(\|u\|_{\mathrm{L}^2(\Omega)}^2 + \|\nabla u\|_{\mathrm{L}^2(\Omega)}^2 \Big)^{\frac{1}{2}} \cdot \Big(\|v\|_{\mathrm{L}^2(\Omega)}^2 + \|\nabla v\|_{\mathrm{L}^2(\Omega)}^2 \Big)^{\frac{1}{2}} \\ &= \|u\|_{\mathrm{H}^1(\Omega)} \cdot \|v\|_{\mathrm{H}^1(\Omega)}, \end{aligned}$$

  so a is continuous.

- $\operatorname{Re} a(u,u) = \int_\Omega \|\nabla u\|^2 \,\mathrm{d}x = |||u|||_{\mathrm{H}_0^1(\Omega)}^2 \geq \alpha_0 \|u\|_{\mathrm{H}_0^1(\Omega)}^2$, so a is coercive.

In particular we can apply Lax-Milgram and obtain for every $f \in \mathrm{L}^2(\Omega)$ a unique $u \in \mathrm{H}_0^1(\Omega)$ with $a(u,v) = \int_\Omega fv \,\mathrm{d}x$ for all $v \in \mathrm{H}_0^1(\Omega)$.

The above gives a weak solution for every bounded open set, no matter how abstruse it is. Often these solutions will not be differentiable or even not continuous in a classical sense. This will present itself to us as the problem of regularity.

Another issue is fulfillment of boundary conditions. In what sense does such a solution actually attain the required boundary values? To tackle this issue, we will approximate our Sobolev functions by smooth functions that allow us to talk about values on the boundary and then pass to limits.

7 Approximation of Sobolev functions

Working in Sobolev spaces and with weak derivatives can get quite difficult. Hence, we would like to approximate Sobolev functions by some "nicer" functions, which makes the proofs in the following chapters work. The usual tool we are trying to use is convolution with mollifiers. As before, let $\Omega \subseteq \mathbb{R}^d$ be an open set. For a first result, we start with a inner approximation i.e. an approximation if we stay away from the boundary $\partial\Omega$.

Theorem 7.1. *Let $k \in \mathbb{N}, 1 \leq p < \infty$, $u \in \mathrm{W}^{k,p}(\Omega)$, $\tilde{u} \in \mathrm{L}^p(\mathbb{R}^d)$ the extension of u by zero and $(\eta_\varepsilon)_{\varepsilon>0}$ be a mollifier. Furthermore, set $\Omega_\varepsilon := \{x \in \Omega\colon \operatorname{dist}(x, \partial\Omega) > \varepsilon\}$.*

a) For all $\varepsilon > 0$ the function $u_\varepsilon := \tilde{u} * \eta_\varepsilon$ is a C^∞ function and for all $|\alpha| \leq k$ it holds that $\mathrm{D}^\alpha u_\varepsilon(x) = (\eta_\varepsilon * \mathrm{D}^\alpha u)(x) = (\mathrm{D}^\alpha u)_\varepsilon$ for all $x \in \Omega_\varepsilon$.

b) $u_\varepsilon \to u$ as $\varepsilon \to 0$ in $\mathrm{W}^{k,p}_{\mathrm{loc}}(\Omega)$.

Proof. a) From integration theory we know that $u_\varepsilon \in \mathrm{C}^\infty(\mathbb{R}^d)$ and $\mathrm{D}^\alpha u_\varepsilon(x) = (\mathrm{D}^\alpha \eta_\varepsilon * \tilde{u})(x)$ for all $x \in \mathbb{R}^d$. For $x \in \Omega_\varepsilon$ and $y \in \Omega$ consider $\Phi_x(y) := \eta_\varepsilon(x - y)$. Since $\operatorname{supp}(\Phi_x) \subseteq \overline{\mathrm{B}_\varepsilon(x)} \subseteq \Omega$, it holds $\Phi_x \in \mathcal{C}^\infty_{\mathrm{c}}(\Omega)$. Hence,

$$\begin{aligned}
\mathrm{D}^\alpha u_\varepsilon(x) &= \int_\Omega \mathrm{D}^\alpha_x \eta_\varepsilon(x-y)u(y)\,\mathrm{d}y \\
&= (-1)^{|\alpha|} \int_\Omega \mathrm{D}^\alpha_y \underbrace{\eta_\varepsilon(x-y)}_{=\Phi_x(y)} u(y)\,\mathrm{d}y \\
&= \int_\Omega \eta_\varepsilon(x-y)\,\mathrm{D}^\alpha u(y)\,\mathrm{d}y \\
&= (\mathrm{D}^\alpha u)_\varepsilon(x).
\end{aligned}$$

b) For $V \Subset \Omega$ there is $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$ it holds that $V \subset \Omega_\varepsilon \subset \Omega$. Hence, for all $|\alpha| \leq k$, we find that

$$\mathrm{D}^\alpha u_\varepsilon(x) = (\mathrm{D}^\alpha u)_\varepsilon \to \mathrm{D}^\alpha u$$

in $\mathrm{L}^p(V)$. This implies that

$$\|u - u_\varepsilon\|^p_{\mathrm{W}^{k,p}(V)} = \sum_{|\alpha|\leq k} \|\mathrm{D}^\alpha u - (\mathrm{D}^\alpha u)_\varepsilon\|^p_{\mathrm{L}^p(V)} \to 0 \text{ as } \varepsilon \searrow 0. \qquad \square$$

Remark. For $u \in \mathrm{W}^{k,p}(\Omega)$ there exists a sequence $(u_n) \subseteq \mathrm{W}^{k,p}(\Omega) \cap \mathrm{C}^\infty(\Omega)$ such that $u_n \to u$ in $\mathrm{W}^{k,p}(\Omega)$. We will prove this statement later in this chapter in the Meyers and Serrin Theorem 7.3.

Lemma 7.2 (Partition of unity)**.** *Let $(U_j)_{j\in\mathbb{N}}$ be a locally finite covering of Ω, i.e.*

- $U_j \Subset \Omega$ *for all* $j \in \mathbb{N}$.
- $\Omega = \bigcup_{j\in\mathbb{N}} U_j$.
- *For all compact sets* $K \subseteq \Omega$, *the set* $\{j \in \mathbb{N} \colon U_j \cap K \neq \emptyset\}$ *is finite.*

Then there is a partition of unity, i.e. $\varphi_j \in \mathcal{C}_c^\infty(\Omega)$ *for* $j \in \mathbb{N}$ *such that*

- $\operatorname{supp}(\varphi_j) \subseteq U_j$ *for all* $j \in \mathbb{N}$;
- $0 \leq \varphi_j \leq 1$ *for all* $j \in \mathbb{N}$;
- $\sum_{j=1}^\infty \varphi_j = 1$ in Ω.

Proof. 1^{st} **step:** There exist $V_j \Subset U_j$, $j \in \mathbb{N}$, such that for all $m \in \mathbb{N}$ it holds

$$\left(\bigcup_{j<m} V_j\right) \cup \left(\bigcup_{j\geq m} U_j\right) = \Omega.$$

We prove this claim by induction over m. The case $m = 1$ is trivial. For the induction step suppose that the claim holds for some $m \in \mathbb{N}$. Since U_m is open, we conclude that $\partial U_m \cap U_m = \emptyset$ and by the induction hypothesis the collection $\{V_1, \ldots, V_{m-1}, U_{m+1}, U_{m+2}, \ldots\}$ is an open covering of ∂U_m. By the compactness of ∂U_m there exists $r > 0$ with

$$Z_r \coloneqq \{x \in \mathbb{R}^d \colon \operatorname{dist}(x, \partial U_m) < r\} \subseteq \left(\bigcup_{j<m} V_j\right) \cup \left(\bigcup_{j>m} U_j\right).$$

Set $V_m \coloneqq U_m \setminus \overline{Z_{\frac{r}{2}}}$. Then V_m is open, $V_m \Subset \Omega$ and

$$\Omega = \left(\bigcup_{j<m} V_j\right)\cup\left(\bigcup_{j\geq m} U_j\right) = \left(\bigcup_{j<m} V_j\right)\cup V_m \cup\left(\bigcup_{j>m} U_j\right) = \left(\bigcup_{j<m+1} V_j\right)\cup\left(\bigcup_{j\geq m+1} U_j\right)$$

which finishes the induction and proves the claim.

2^{nd} **step (Construction of φ_j)** Choose E_j, $j \in \mathbb{N}$, such that $V_j \Subset E_j \Subset U_j$ for all $j \in \mathbb{N}$ and set

$$\varepsilon_j \coloneqq \frac{1}{2} \min\{\operatorname{dist}(\partial E_j, V_j), \operatorname{dist}(\partial E_j, \partial U_j)\} > 0$$

and $\tilde{\varphi}_j \coloneqq (\chi_{E_j})_{\varepsilon_j}$. Then for all $j \in \mathbb{N}$ we obtain

- $\tilde{\varphi}_j \in \mathcal{C}_c^\infty(\Omega)$ with $\operatorname{supp}(\tilde{\varphi}_j) \subseteq U_j$,
- $0 \leq \tilde{\varphi}_j \leq 1$,
- $\tilde{\varphi}_j = 1$ on V_j,

- $\sum_{j\in\mathbb{N}} \tilde{\varphi}_j > 0$ in Ω.

Now, set

$$\varphi_j := \frac{\tilde{\varphi}_j}{\sum_{j\in\mathbb{N}} \tilde{\varphi}_j}, \ j \in \mathbb{N}.$$

Then

- $\varphi_j \in \mathcal{C}_c^\infty(\Omega)$ with $\operatorname{supp}(\varphi_j) \subseteq U_j$,
- $0 \le \varphi_j \le 1$,
- $\sum_{j\in\mathbb{N}} \varphi_j = \frac{\sum_{j\in\mathbb{N}} \tilde{\varphi}_j}{\sum_{j\in\mathbb{N}} \tilde{\varphi}_j} = 1$ in Ω. □

Using this partition of unity we can prove a theorem of Meyers and Serrin that states that we can approximate Sobolev functions by C^∞ functions.

Theorem 7.3 (Meyers and Serrin)**.** *Let $k \in \mathbb{N}$ and $1 \le p < \infty$. Then* $\mathrm{C}^\infty(\Omega) \cap \mathrm{W}^{k,p}(\Omega)$ *is dense in* $\mathrm{W}^{k,p}(\Omega)$.

Proof. For $j \in \mathbb{N}$ set

$$U_j := \{x \in \Omega\colon \operatorname{dist}(x, \partial\Omega) > \frac{1}{j} \text{ and } |x| < j\}.$$

Then

- $\Omega = \bigcup_{j=1}^\infty U_j$,
- For all $j \in \mathbb{N}$ it holds $\overline{U_j} \subseteq U_{j+1}$.

Hence, the sets defined by $V_1 := U_5, V_j := U_{j+4} \setminus \overline{U_{j+1}}$ for $j \ge 2$ form a locally finite covering. Let φ_j, $j \in \mathbb{N}$, be the corresponding partition of unity.

Now, let $Z_1 := U_6$, $Z_j := U_{j+5} \setminus \overline{U_j}$ for $j \ge 2$. Then $V_j \Subset Z_j \Subset \Omega$ for all $j \in \mathbb{N}$. In addition, let $\varepsilon > 0$ and $u \in \mathrm{W}^{k,p}(\Omega)$. Our aim is to find $v_\varepsilon \in \mathrm{C}^\infty(\Omega) \cap \mathrm{W}^{k,p}(\Omega)$ such that

$$\|u - v_\varepsilon\|_{\mathrm{W}^{k,p}(\Omega)} < \varepsilon.$$

For all $j \in \mathbb{N}$ we have

- $\varphi_j u \in \mathrm{W}^{k,p}(\Omega)$,
- $\operatorname{supp}(\varphi_j u) \subseteq V_j$.

Let $(\eta_\varepsilon)_{\varepsilon>0}$ be a mollifier and for all $j \in \mathbb{N}$ let $\varepsilon_j > 0$ such that

- $u_j := \eta_{\varepsilon_j} * (\varphi_j u) \in \mathcal{C}_c^\infty(Z_j)$.
- $$\|\eta_{\varepsilon_j} * (\varphi_j u) - \varphi_j u\|_{\mathrm{W}^{k,p}(\Omega)} = \|\eta_{\varepsilon_j} * (\varphi_j u) - \varphi_j u\|_{\mathrm{W}^{k,p}(Z_j)} < \frac{\varepsilon}{2^j} \qquad (*)$$

Now $v_\varepsilon := \sum_{j\in\mathbb{N}} u_j$ satisfies

- $v_\varepsilon \in \mathrm{C}^\infty(\Omega)$.
- For all $V \Subset \Omega$ it holds

$$\begin{aligned}
\|v_\varepsilon - u\|_{\mathrm{W}^{k,p}(V)} &= \left\| \sum_{j\in\mathbb{N}} u_j - \sum_{j\in\mathbb{N}} \varphi_j u \right\|_{\mathrm{W}^{k,p}(V)} \\
&\leq \sum_{j\in\mathbb{N}} \|u_j - \varphi_j u\|_{\mathrm{W}^{k,p}(V)} \\
&\leq \sum_{j\in\mathbb{N}} \|u_j - \varphi_j u\|_{\mathrm{W}^{k,p}(Z_j)} \\
&\overset{(*)}{\leq} \sum_{j\in\mathbb{N}} \frac{\varepsilon}{2^j} \\
&= \varepsilon.
\end{aligned}$$

With this, we obtain by the monotone convergence theorem

$$\begin{aligned}
\|v_\varepsilon - u\|_{\mathrm{W}^{k,p}(\Omega)} &= \left(\sum_{|\alpha|\leq k} \int_\Omega |\mathrm{D}^\alpha (v_\varepsilon - u)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} \\
&= \lim_{j\to\infty} \left(\sum_{|\alpha|\leq k} \int_{U_j} |\mathrm{D}^\alpha (v_\varepsilon - u)| \, \mathrm{d}x \right)^{\frac{1}{p}} \\
&= \lim_{j\to\infty} \|v_\varepsilon - u\|_{\mathrm{W}^{k,p}(U_j)} \leq \varepsilon.
\end{aligned}$$

□

Definition 7.4. Let $\Omega \subseteq \mathbb{R}^d$ be a domain. We say that Ω satisfies the segment condition if for all $x \in \partial\Omega$ there exists a neighbourhood $U_x \subseteq \mathbb{R}^d$ of x and $y_x \in \mathbb{R}^d \setminus \{0\}$ such that for all $z \in \overline{\Omega} \cap U_x$ and for all $t \in (0,1)$ it holds $z + t y_x \in \Omega$.

Remark. The segment condition may look a little cryptic. What we try to avoid is that Ω is located on both sides of the boundary, like in the example given in Figure 7.1.

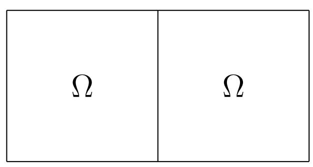

Figure 7.1: Ω does not fulfill the segment condition

We will give the next theorem without proof.

Theorem 7.5. *Let $\Omega \subseteq \mathbb{R}^d$ be a domain satisfying the segment condition. Furthermore, let $k \in \mathbb{N}$ and $1 \leq p < \infty$. Then the set*

$$\{\varphi|_\Omega \colon \varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)\}$$

is dense in $\mathrm{W}^{k,p}(\Omega)$.

Corollary 7.6. *For all $k \in \mathbb{N}$ and all $1 \leq p < \infty$, the space $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is dense in $\mathrm{W}^{k,p}(\mathbb{R}^d)$ and we have $\mathrm{W}_0^{k,p}(\mathbb{R}^d) = \mathrm{W}^{k,p}(\mathbb{R}^d)$.*

Proof. This is an immediate consequence of Theorem 7.5. □

8 Extensions and Traces

Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain and $u \in \mathrm{W}^{k,p}(\Omega)$. The first aim of this chapter is to find an extension $\tilde{u} \in \mathrm{W}^{k,p}(\mathbb{R}^d)$ such that $\tilde{u} = u$ on Ω. We will see that this is only possible if the boundary $\partial\Omega$ is nice enough.

Definition 8.1. Let $m \in \mathbb{N}_0$. We say that Ω is of class C^m, if for all $x_0 \in \partial\Omega$ (upon relabeling and reorienting coordinate axes), there exists an open neighbourhood U of x_0, an open set $O \subseteq \mathbb{R}^{d-1}$ and $a \in \mathrm{C}^m(O, \mathbb{R})$ such that (with $y = (y_1, \dots, y_d) = (y', y_d)$) $\partial\Omega \cap U = \{(y', a(y')) \colon y' \in O\}$ and $\Omega \cap U = \{y \in U \colon y' \in O \text{ and } y_d > a(y')\}$.

Notation. We denote $\partial\Omega \in \mathrm{C}^m$ if $\partial\Omega$ is of class C^m.

Remark 8.2.

a) For Ω bounded we only need finitely many (U, O, a) to achieve a covering of $\partial\Omega$.

b) If $\partial\Omega$ is of class C^m for $m \geq 1$, then Ω satisfies the segment condition[1]. We will use this later when proving the trace theorem.

c) Choosing U and O properly we can always have that there are $\delta, \beta > 0$ such that

$$\begin{aligned} O &= \{y' \in \mathbb{R}^{d-1} \colon \; |y'| < \delta\}, \\ U &= \{y \in \mathbb{R}^d \colon \; |y_d - a(y)| < \beta, y' \in O\} \text{ and} \\ \Omega \cap U &= \{y \in \mathbb{R}^d \colon (a(y') < y_d < a(y') + \beta, y' \in O\}. \end{aligned}$$

Notation 8.3. We denote

$$\begin{aligned} \mathbb{R}^d_+ &\coloneqq \{x \in \mathbb{R}^d \colon x_d > 0\}, & \Omega_+ &\coloneqq \Omega \cap \mathbb{R}^d_+, \\ \mathbb{R}^d_- &\coloneqq \{x \in \mathbb{R}^d \colon x_d < 0\}, & \Omega_- &\coloneqq \Omega \cap \mathbb{R}^d_-, \\ \mathbb{R}^d_0 &\coloneqq \{x \in \mathbb{R}^d \colon x_d = 0\}, & \Omega_0 &\coloneqq \Omega \cap \mathbb{R}^d_0. \end{aligned}$$

Reminder. A function $\phi \colon U \to V$ is called C^m-diffeomorphism if ϕ is bijective, $\phi \in \mathrm{C}^m(\overline{U}, \mathbb{R}^d)$, $\phi^{-1} \in \mathrm{C}^m(\overline{V}, \mathbb{R}^d)$ and $\det(\mathrm{D}\,\phi) \neq 0$ in $\overline{U}$.

Lemma 8.4. *Let Ω be bounded, $m \in \mathbb{N}$, $\partial\Omega \in \mathrm{C}^m$ and $x_0 \in \partial\Omega$ with (U, O, a) as in Definition 8.1. Consider $\phi(x) \coloneqq (x', x_d - a(x'))$ for $x = (x', x_d) \in U$ and $\psi(y) \coloneqq (y', y_d + a(y'))$ for $y \in V \coloneqq \phi(U)$. Then*

- *$\phi \in \mathrm{C}^m(U, \mathbb{R}^d)$ and $\psi \in \mathrm{C}^m(V, \mathbb{R}^d)$ are both injective.*

[1] In fact, every regularity for $\partial\Omega$ which guarantees the existence of a normal vector is sufficient here.

- $\phi^{-1} = \psi$.
- $\det(\mathrm{D}\,\phi) = \det(\mathrm{D}\,\psi) = 1$.
- $\phi\colon U \to V$ *is a* C^m*-diffeomorphism.*
- $\phi(\Omega \cap U) = V_+$ *and* $\phi(\partial\Omega \cap U) = V_0$.

Proof. The proof is straightforward. □

Proposition 8.5. *Let* $\Omega, V \subseteq \mathbb{R}^d$ *be bounded domains,* $m \in \mathbb{N}$*,* $1 \le p < \infty$ *and* $\phi\colon V \to \Omega$ *be a* C^m*-diffeomorphism. Then*

$$T_\phi\colon\ \mathrm{W}^{m,p}(\Omega) \to \mathrm{W}^{m,p}(V),\ (T_\phi u)(y) := (u \circ \phi)(y)$$

is linear, bijective with $T_\phi^{-1} = T_{\phi^{-1}}$ *and both* T_ϕ *and* T_ϕ^{-1} *are bounded, i.e. there are* $c_1, c_2 \ge 0$ *such that for all* $u \in \mathrm{W}^{m,p}(\Omega)$ *and for all* $v \in \mathrm{W}^{m,p}(V)$ *it holds*

$$\begin{aligned} \|T_\phi u\|_{\mathrm{W}^{m,p}(V)} &\le c_1 \|u\|_{\mathrm{W}^{m,p}(\Omega)}, \\ \left\|{T_\phi}^{-1} v\right\|_{\mathrm{W}^{m,p}(\Omega)} &\le c_2 \|v\|_{\mathrm{W}^{m,p}(V)}. \end{aligned}$$

In the following proof and in general from now on we will freely use the concept of "generic constants", i.e. the letter "C" stands for a finite, non-negative value that may change from occurrence to occurrence but is always independent from any object that is all-quantified in the respective situation.

Proof. We only prove the case $m = 1$, the rest follows by induction.

T_ϕ, ${T_\phi}^{-1}$ linear: is clear.

T_ϕ bounded: Let $u \in \mathrm{W}^{1,p}(\Omega) \cap \mathrm{C}^\infty(\Omega)$. Then

$$\begin{aligned} \|T_\phi u\|^p_{\mathrm{L}^p(V)} &= \int_V |T_\phi u|^p \,\mathrm{d}x = \int_V |u(\phi(x))|^p \,\mathrm{d}x \\ &\overset{y=\phi(x)}{=} \int_\Omega |u(y)|^p \cdot \left|\det(\mathrm{D}\,\phi^{-1})(y)\right| \,\mathrm{d}y \\ &\le C \int_\Omega |u(y)|^p \,\mathrm{d}y \\ &= C\,\|u\|^p_{\mathrm{L}^p(\Omega)}. \end{aligned}$$

Since

$$\partial_j T_\phi u = \partial_j (u \circ \phi) = \sum_{k=1}^d (\partial_k u \circ \phi) \cdot \partial_j \phi_k = \sum_{k=1}^d T_\phi(\partial_k u) \cdot \partial_j \phi_k,$$

we obtain

$$
\begin{aligned}
\|\partial_j T_\phi u\|_{\mathrm{L}^p(V)} &= \left\| \sum_{k=1}^{d} T_\phi(\partial_k u) \cdot \partial_j \phi_k \right\|_{\mathrm{L}^p(V)} \\
&\leq \sum_{k=1}^{d} \|T_\phi(\partial_k u) \cdot \partial_j \phi_k\|_{\mathrm{L}^p(V)} \\
&\leq \|\partial_j \phi\|_\infty \sum_{k=1}^{d} \|T_\phi(\partial_k u)\|_{\mathrm{L}^p(V)} .
\end{aligned}
$$

As before we conclude

$$
\leq C \sum_{k=1}^{d} \|\partial_k u\|_{\mathrm{L}^p(\Omega)} .
$$

Putting everything together yields

$$
\begin{aligned}
\|T_\phi u\|^p_{\mathrm{W}^{1,p}(V)} &= \|T_\phi u\|^p_{\mathrm{L}^p(V)} + \sum_{j=1}^{d} \|\partial_j T_\phi u\|^p_{\mathrm{L}^p(V)} \\
&\leq C \|u\|^p_{\mathrm{L}^p(\Omega)} + C \left(\sum_{k=1}^{d} \|\partial_k u\|_{\mathrm{L}^p(\Omega)} \right)^p
\end{aligned}
$$

and by the equivalence of norms in $\mathbb{R}^d$ we can further estimate by

$$
\begin{aligned}
&\leq C \left(\|u\|^p_{\mathrm{L}^p(\Omega)} + \sum_{k=1}^{d} \|\partial_k u\|^p_{\mathrm{L}^p(\Omega)} \right)^p \\
&= C \|u\|_{\mathrm{W}^{1,p}(\Omega)} .
\end{aligned}
$$

The same holds for $T_{\phi^{-1}}$, as it is the same sort of composition with a diffeomorphism. For general $u \in \mathrm{W}^{1,p}(\Omega)$ the assertion follows by a density argument and Theorem 7.3.

Inverse Furthermore, for all $u \in \mathrm{W}^{1,p}(\Omega)$ and for all $v \in \mathrm{W}^{1,p}(V)$:

$$
T_{\phi^{-1}} T_\phi u = T_{\phi^{-1}}(u \circ \phi) = u \circ \phi \circ \phi^{-1} = u,
$$

and analogously $T_\phi T_{\phi^{-1}} v = v$. □

Theorem 8.6 (Sobolev extension theorem)**.** *Let $m \in \mathbb{N}$, $1 \leq p < \infty$, Ω bounded with $\partial\Omega \in \mathrm{C}^m$ and $\Xi \subseteq \mathbb{R}^d$ be a domain with $\Omega \Subset \Xi$. Then there exists a linear operator $E\colon \mathrm{L}^p(\Omega) \to \mathrm{L}^p(\mathbb{R}^d)$ with the following properties:*

- *For all $u \in \mathrm{L}^p(\Omega)$ it holds, that $Eu = u$ in Ω.*

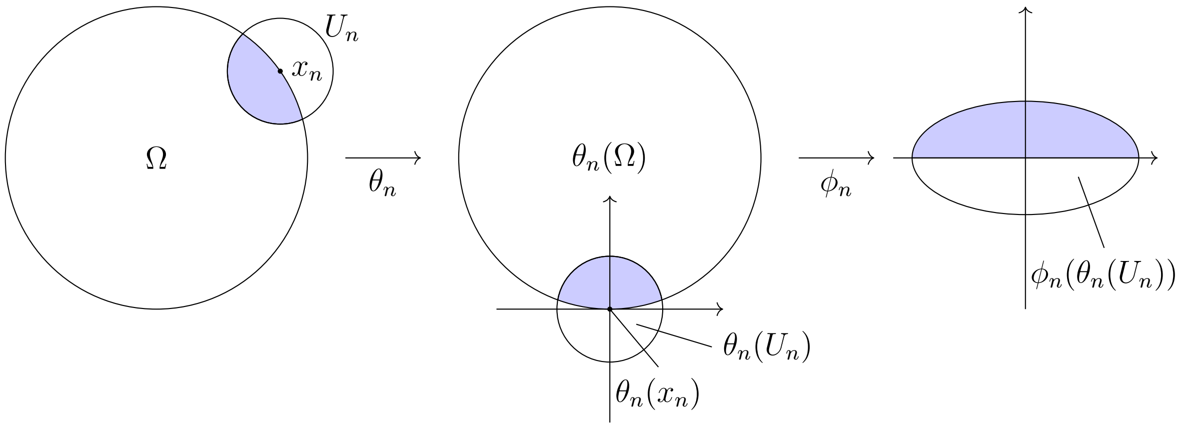

Figure 8.1: Flattening the boundary

- *$E\colon \mathrm{W}^{k,p}(\Omega) \to \mathrm{W}^{k,p}(\mathbb{R}^d)$ for all $1 \le k \le m$.*
- *For all $k \in \{0, 1, \dots, m\}$ and for all $u \in \mathrm{W}^{k,p}(\Omega)$ it holds*
$$\|Eu\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} \le C\,\|u\|_{\mathrm{W}^{k,p}(\Omega)}\,.$$
- *For all $u \in \mathrm{L}^p(\Omega)\colon$ $\operatorname{supp}(Eu) \subseteq \Xi$.*

Remark. a) The assumption $\partial\Omega \in \mathrm{C}^m$ can be weakened, i.e. a Lipschitz boundary is sufficient.

b) A characterisation for boundaries $\partial\Omega$ such that a function $u \in \mathrm{W}^{k,p}(\Omega)$ has an extension operator is still unknown.

Before starting the proof, we give a brief idea of our strategy: We consider a covering (U_n) of $\overline{\Omega}$ with a corresponding partition of unity and start with the special case that u is smooth and the geometry of U_n is flat (i.e. looks like the third picture in Figure 8.1). In this case, we can extend our function by zero to the upper half plane and then do a reflection to the lower half plane. For the general case, we transform U_n to the special case with flat boundary, use the extension there and transform back. This is illustrated in Figure 8.1. As a final step, we use our density results of the previous chapter to obtain an extension for $u \in \mathrm{W}^{k,p}(\Omega)$.

Proof. 1st **step (The flat situation):** We assume smooth u and flat geometry, i.e. for $x_0 \in \partial\Omega$ and some $r > 0$ and $B \coloneqq \mathrm{B}_r(x_0)$ it holds that
$$\Omega \cap B = B_+ \text{ and } \partial\Omega \cap B = B_0$$
Let $u \in \mathrm{C}^m(B_+ \cup B_0)$ and $\operatorname{supp}(u) \subsetneq B_s(x)_+$ for some $s < r$. This means $u = 0$ near $\partial B \cap \overline{\mathbb{R}^d_+}$, so extend by zero to $\overline{\mathbb{R}^d_+}$.[2] Now, we want to extend from $\overline{\mathbb{R}^d_+}$ to $\mathbb{R}^d$ by a higher order reflection:

[2]This extension will again be denoted by u, as per the much loved abuse of notation.

The system of linear equations

$$\sum_{j=1}^{m+1}(-j)^k\lambda_j = 1, \qquad\qquad k = 0, 1, \dots, m \tag{$*$}$$

has a unique solution (the determinant is of Vandermonde type).

Using these $\lambda_1, \dots, \lambda_{m+1}$, we define

$$\tilde{E}u(x) := \begin{cases} u(x), & x_d \geq 0, \\ \sum\limits_{j=1}^{m+1} \lambda_j u(x', -jx_d), & x_d < 0, \end{cases}$$

and for $\alpha \in \mathbb{N}^d$ with $1 \leq |\alpha| \leq m$

$$\tilde{E}_\alpha u(x) := \begin{cases} u(x), & x_d \geq 0, \\ \sum\limits_{j=1}^{m+1} (-j)^{\alpha_d} \lambda_j u(x', -jx_d), & x_d < 0. \end{cases}$$

Then for $x_d \neq 0$, $\mathrm{D}^\alpha \tilde{E}u = \tilde{E}_\alpha \mathrm{D}^\alpha u$ and the same equation holds for $x_d = 0$:

$$\begin{aligned}
\lim_{x_d \nearrow 0} \mathrm{D}^\alpha(\tilde{E}u)(x', x_d) &= \lim_{x_d \nearrow 0} \tilde{E}_\alpha(\mathrm{D}^\alpha u)(x', x_d) \\
&= \lim_{x_d \nearrow 0} \sum_{j=1}^{m+1} (-j)^{\alpha_d} \lambda_j \, \mathrm{D}^\alpha u(x', -jx_d) \\
&\overset{u \in \mathrm{C}^m}{=} \left(\sum_{j=1}^{m+1} (-j)^{\alpha_d} \lambda_j \right) \mathrm{D}^\alpha u(x', 0) \\
&\overset{(*)}{=} \mathrm{D}^\alpha u(x', 0) \\
&= \lim_{x_d \searrow 0} \mathrm{D}^\alpha u(x', x_d) \\
&= \lim_{x_d \searrow 0} \mathrm{D}^\alpha \tilde{E}u(x', x_d).
\end{aligned}$$

Finally, for all $|\alpha| \leq m$ we obtain with the substitution $z = (x', -jx_d)$

$$\begin{aligned}
\left\| \mathrm{D}^\alpha \tilde{E}u \right\|_{\mathrm{L}^p(\mathbb{R}^d)} &\leq \| \mathrm{D}^\alpha u \|_{\mathrm{L}^p(\mathbb{R}^d_+)} + \left\| \sum_{j=1}^{m+1} (-j)^{\alpha_d} \lambda_j \, \mathrm{D}^\alpha u(z) \right\|_{\mathrm{L}^p(\mathbb{R}^d_-)} \\
&\leq \left(1 + \sum_{j=1}^{m+1} j^{\alpha_d} \, |\lambda_j| \, j^{\frac{1}{p}} \right) \| \mathrm{D}^\alpha u \|_{\mathrm{L}^p(\mathbb{R}^d_+)} \\
&\leq \left(1 + \sum_{j=1}^{m+1} j^{\alpha_d + 1} \, |\lambda_j| \right) \| \mathrm{D}^\alpha u \|_{\mathrm{L}^p(B_+)} .
\end{aligned}$$

All in all, so far we have proven that $\tilde{E}u \in \mathrm{C}^m(\mathbb{R}^d)$, the support of $\tilde{E}u$ is indeed contained in B and

$$\left\| \tilde{E}u \right\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} \leq C \left\| u \right\|_{\mathrm{W}^{k,p}(B_+)}$$

for all $k = 0, 1, \ldots, m$.

2^{nd} **step (Patching together $\partial\Omega$):** Since $\partial\Omega \in \mathrm{C}^m$ is compact, there exist $x_n \in \partial\Omega$, $n \in \{1, \ldots, N\}$, $U_n := B_{r_n}(x_n)$ and $a_n \in \mathrm{C}^m(\mathbb{R}^{d-1})$ such that for an appropriate change of coordinate system described by a C^∞-diffeomorphism θ_n the following identities hold:

$$\begin{aligned} \Omega \cap U_n &= \{x \in U_n \colon \theta_n(x)_d > a_n(\theta_n(x)')\} \\ \partial\Omega \cap U_n &= \{x \in U_n \colon \theta_n(x)_d = a_n(\theta_n(x)')\}. \end{aligned}$$

Consider again as in Lemma 8.4

$$\begin{aligned} \phi_n(y) &= (y', y_d - a_n(y')), \quad & y \in \theta_n(U_n) \\ \phi_n^{-1}(y) &= (y', y_d + a_n(y')), \quad & y \in \phi_n(\theta_n(U_n)) =: V_n. \end{aligned}$$

and choose $U_0 \Subset \Omega$, such that $\overline{\Omega} \subseteq \bigcup_{n=0}^{N} U_n$.

Then $U_0, U_1, \ldots, U_N$ is a locally finite cover of Ω.

Let $\varphi_n \in \mathcal{C}_c^\infty(U_n)$ for $n = 0, 1, \ldots, N$ be a corresponding partition of unity.

3^{rd} **step (Localizing smooth u):** Let $u \in \mathrm{C}^m(\overline{\Omega})$ and $n \in \{1, \ldots, N\}$. Then $\varphi_n u \in \mathrm{C}^m(\overline{U_n} \cap \overline{\Omega})$ with $\operatorname{supp}(\varphi_n u) \Subset U_n$.

Thus, $v_n := T_{\phi_n{}^{-1}} T_{\theta_n{}^{-1}}(\varphi_n u) = (\varphi_n u) \circ \theta_n{}^{-1} \circ \phi_n{}^{-1} \in \mathrm{C}^m(\overline{V_{n,+}})$ with $\operatorname{supp}(v_n) \Subset V_n$ and $\operatorname{supp}(v_n) \subseteq V_{n,+} \cup V_{n,0}$.

Now, we can apply the 1^{st} step to v_n, which yields $\tilde{E}v_n \in \mathrm{C}^m(\mathbb{R}^d) \cap \mathrm{W}^{k,p}(\mathbb{R}^d)$ for $k = 1, \ldots, m$ and $\operatorname{supp}(\tilde{E}v_n) \subseteq V_n$.

Set $u_n := T_{\theta_n} T_{\phi_n}(\tilde{E}v_n) \in \mathrm{W}^{k,p}(\mathbb{R}^d)$. Then in $\Omega \cap U_n$ it holds

$$u_n = T_{\theta_n} T_{\varphi_n}(\tilde{E} \underbrace{T_{\phi_n{}^{-1}} T_{\theta_n{}^{-1}}(\varphi_n u)}_{\text{in } V_{n,+}}) = T_{\theta_n} T_{\phi_n} T_{\phi_n{}^{-1}} T_{\theta_n{}^{-1}}(\varphi_n u) = \varphi_n u.$$

Thus, Proposition 8.5 and the results of the first step imply

$$\begin{aligned} \|u_n\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} &\leq C \left\| \tilde{E} T_{\phi_n{}^{-1}} T_{\theta_n{}^{-1}}(\varphi_n u) \right\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} \\ &\leq C \left\| T_{\phi_n{}^{-1}} T_{\theta_n{}^{-1}}(\varphi_n u) \right\|_{\mathrm{W}^{k,p}(V_{n,+})} \end{aligned}$$

$$\leq C \|\phi_n u\|_{\mathrm{W}^{k,p}(\Omega\cap U_n)}$$
$$\leq C \|u\|_{\mathrm{W}^{k,p}(\Omega\cap U_n)} .$$

4^{th} **step (Definition of E):** Let $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}(\psi) \subseteq \Xi$ and $\psi = 1$ in Ω. For $u \in \mathrm{C}^m(\overline{\Omega})$ set

$$Eu := \psi \cdot \sum_{n=0}^{N} u_n$$

where $u_0 := \varphi_0 \cdot u$. We have

- $Eu \in \mathrm{W}^{k,p}(\mathbb{R}^d)$ and $\operatorname{supp}(Eu) \subseteq \Xi$.
- $\|Eu\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} \leq C \left\|\sum_{n=0}^{N} u_n\right\|_{\mathrm{W}^{k,p}(\mathbb{R}^d)} \leq C \|u\|_{\mathrm{W}^{k,p}(\Omega)}$.
- In Ω it holds $Eu = \psi \cdot \sum_{n=0}^{\mathbb{N}} u_n = \sum_{n=0}^{N} \varphi_n u = u$.

5^{th} **step (Extension to $u \in \mathrm{W}^{k,p}(\Omega)$):** The mapping

$$E\colon (\mathrm{C}^m(\overline{\Omega}), \|\cdot\|_{\mathrm{W}^{k,p}(\Omega)}) \to \mathrm{W}^{k,p}(\mathbb{R}^d)$$

is linear and bounded and $\mathrm{C}^m(\overline{\Omega})$ is dense in $\mathrm{W}^{k,p}(\Omega)$, so E has a unique linear bounded extension to $E\colon \mathrm{W}^{k,p}(\Omega) \to \mathrm{W}^{k,p}(\mathbb{R}^d)$. □

The following theorem answers the question how to assign boundary values to a Sobolev function.

Theorem 8.7 (Trace Theorem). *Let $1 \leq p < \infty$, let Ω be bounded with boundary of class $\partial\Omega \in \mathrm{C}^1$ and equip $\partial\Omega$ with the surface measure. Then there exists a linear and bounded operator*

$$\mathrm{Tr}\colon \mathrm{W}^{1,p}(\Omega) \to \mathrm{L}^p(\partial\Omega)$$

such that $\mathrm{Tr}(u) = u|_{\partial\Omega}$ for all $u \in \mathrm{W}^{1,p}(\Omega) \cap \mathrm{C}(\overline{\Omega})$. This operator is called the trace operator.

Proof. We only give a sketch of the proof.

1^{st} **step:** The first step is to deal with the situation of a flat boundary, as we did in the proof of Theorem 8.6. Let $x_0 \in \partial\Omega$ and $B = \mathrm{B}_r(x_0) \subseteq \mathbb{R}^d$ with $\Omega \cap B = B_+$ and $\partial\Omega \cap B = B_0$. In addition, let $u \in \mathrm{C}^1(B_+ \cup B_0)$ with $\operatorname{supp} u \subseteq \overline{B_s(x_0)}_+$ for some $s \in (0, r)$. Then for $B' := \{x' \in \mathbb{R}^{d-1} : (x', 0) \in B_0\}$ it holds

$$\int_{\partial\Omega\cap B} |u|^p \,\mathrm{d}S = \int_{B'} |u(x', 0)|^p \,\mathrm{d}x'.$$

For $x' \in B'$ let $y(x') > 0$ such that $(x', y(x')) \in \partial B$. This yields

$$\begin{aligned}
&= \int_{B'} \left| - \int_0^{y(x')} \partial_d u(x', t)\, \mathrm{d}t + \underbrace{u(x', y(x'))}_{=0} \right|^p \mathrm{d}x' \\
&\leq \int_{B'} \left(\int_0^{y(x')} 1 \cdot |\partial_d u(x', t)|\ \mathrm{d}t \right)^p \mathrm{d}x'
\end{aligned}$$

and using Hölder's inequality, we get

$$\begin{aligned}
&\leq \int_{B'} \left(\int_0^{y(x')} 1^{p'}\ \mathrm{d}t \right)^{\frac{p}{p'}} \left(\int_0^{y(x')} |\partial_d u(x', t)|^p\ \mathrm{d}t \right)^{\frac{p}{p}} \mathrm{d}x' \\
&= \int_{B'} \underbrace{y(x')^{\frac{p}{p'}}}_{\leq r} \int_0^{y(x')} |\partial_d u(x', t)|^p\ \mathrm{d}t\, \mathrm{d}x' \\
&\leq r^{\frac{p}{p'}} \int_{B_+} |\partial_d u(x)|^p\ \mathrm{d}x \\
&\leq C \left\| u \right\|^p_{\mathrm{W}^{1,p}(B_+)} .
\end{aligned}$$

2^{nd} **step:** In the second step, we do a localisation. For $u \in \mathrm{C}^1(\overline{\Omega})$ we consider the functions $\varphi_n u$. Then $T_{\phi_n{}^{-1}}(\varphi_n u)$ satisfies the assumptions of the first step and putting all the emerging Gram's determinants into a constant, yields

$$\int_{\partial\Omega \cap U_n} |\varphi_n u|^p\ \mathrm{d}S \leq C \int_{B_0} \left| T_{\phi_n^{-1}}(\varphi_n u)(x', 0) \right|^p\ \mathrm{d}x'.$$

By step 1 we obtain

$$\begin{aligned}
&\leq C \left\| T_{\phi_n{}^{-1}}(\varphi_n u) \right\|^p_{\mathrm{W}^{1,p}(B_+)} \leq C \left\| \varphi_n u \right\|^p_{\mathrm{W}^{1,p}(\Omega \cap U_n)} \\
&\leq C \left\| u \right\|^p_{\mathrm{W}^{1,p}(\Omega)} .
\end{aligned}$$

This leads to the estimate

$$\left\| u \right\|_{\mathrm{L}^p(\partial\Omega)} = \left\| \sum_{n=1}^N \varphi_n u \right\|_{\mathrm{L}^p(\partial\Omega)} \leq \sum_{n=1}^N \left\| \varphi_n u \right\|_{\mathrm{L}^p(\partial\Omega \cap U_n)} \leq C \left\| u \right\|_{\mathrm{W}^{1,p}(\Omega)} .$$

3^{rd} **step:** Since $\mathrm{Tr}\colon \left(\mathrm{C}^1(\overline{\Omega}), \left\| \cdot \right\|_{\mathrm{W}^{1,p}(\Omega)} \right) \to \mathrm{L}^p(\partial\Omega)$ is a bounded linear operator and $\mathrm{C}^1(\overline{\Omega})$ is dense in $\mathrm{W}^{1,p}(\Omega)$ with respect to the $\left\| \cdot \right\|_{\mathrm{W}^{1,p}(\Omega)}$-norm by Theorem 7.5, Tr extends to a bounded linear operator on $\mathrm{W}^{1,p}(\Omega)$. □

Remark. In fact for $u \in \mathrm{W}^{1,p}(\Omega)$ one even gets that $\mathrm{Tr}(u) \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega)$, whatever that is. This means that, as a rule of thumb, considering the boundary values of a function, one loses $\frac{1}{p}$ of smoothness.

Corollary 8.8. *Let $k \in \mathbb{N}$, Ω be bounded with $\partial\Omega \in \mathrm{C}^k, 1 \leq p < \infty$ and $u \in \mathrm{W}^{k,p}(\Omega)$. Then*

$$u \in \mathrm{W}_0^{k,p}(\Omega) \iff \mathrm{Tr}(\mathrm{D}^\alpha u) = 0 \text{ for all } |\alpha| \leq k-1.$$

In particular, $\mathrm{W}_0^{1,p}(\Omega) = \{u \in \mathrm{W}^{1,p}(\Omega) \colon \mathrm{Tr}(u) = 0\}$.

Proof. We only show the implication from left to right. The second one is provable, but it is a little tricky and we want to save time.

For $k = 1$ we consider $u \in \mathrm{W}_0^{1,p}(\Omega)$. Then there exists a sequence $(u_n)_n \subseteq \mathcal{C}_\mathrm{c}^\infty(\Omega)$ such that $u_n \to u$ in $\mathrm{W}^{1,p}(\Omega)$. Hence, for all $n \in \mathbb{N}$ it holds $\mathrm{Tr}(u_n) = 0$ and

$$\|\mathrm{Tr}(u)\|_{\mathrm{L}^p(\partial\Omega)} = \|\mathrm{Tr}(u) - \mathrm{Tr}(u_n)\|_{\mathrm{L}^p(\partial\Omega)} = \|\mathrm{Tr}(u - u_n)\|_{\mathrm{L}^p(\partial\Omega)} \leq C\,\|u - u_n\|_{\mathrm{W}^{1,p}(\Omega)} \to 0$$

as $n \to \infty$. Therefore, $\|\mathrm{Tr}(u)\|_{\mathrm{L}^p(\partial\Omega)} = 0$, so $\mathrm{Tr}(u) = 0$ almost everywhere on $\partial\Omega$.

For general k, let $u \in \mathrm{W}_0^{k,p}(\Omega)$ and set $(u_n) \subseteq \mathcal{C}_\mathrm{c}^\infty(\Omega)$ such that $u_n \to u$ in $\mathrm{W}^{k,p}(\Omega)$. Let $|\alpha| \leq k-1$.

Then $\mathrm{D}^\alpha u_n \to \mathrm{D}^\alpha u$ in $\mathrm{W}^{1,p}(\Omega)$. Since derivatives of $\mathcal{C}_\mathrm{c}^\infty$-functions are again $\mathcal{C}_\mathrm{c}^\infty$, this yields that $\mathrm{D}^\alpha u \in \mathrm{W}_0^{1,p}(\Omega)$.

Employing the case of $k = 1$ now yields that $\mathrm{Tr}(\mathrm{D}^\alpha u) = 0$ a.e. on $\partial\Omega$. □

9 The Rellich Theorem

For the regularity of the solutions we get via the Lax-Milgram Lemma we only know that they are in H^1. Often this is not satisfactory, for instance if the geometric setup is smooth, we would expect classical solutions. Hence, if we have a weak solution, we want to show additional regularity of the solution such that the weak solution is in fact continuous or even a classical solution. In the context of Sobolev spaces, this can be achieved via embeddings.

Definition 9.1. Let X, Y be Banach spaces with $X \subseteq Y$.

a) X is continuously embedded into Y, denoted as $X \hookrightarrow Y$ if there exists some $C \geq 0$ with $\|u\|_Y \leq C \|u\|_X$ for all $u \in X$.

b) A set $\mathcal{F} \subseteq Y$ is called precompact in Y if every sequence $(f_n)_n \subseteq \mathcal{F}$ has a (in Y) convergent subsequence.

c) X is compactly embedded into Y, denoted as $X \hookrightarrow\hookrightarrow Y$, if X is continuously embedded into Y and all bounded $\mathcal{F} \subseteq X$ are precompact in Y.

Notation. Let $D \subseteq \mathbb{R}^d$ be measurable and $1 \leq p \leq \infty$. For $f \in L^p(D)$ we denote

- its extension to the whole space by zero by $\tilde{f} \in L^p(\mathbb{R}^d)$.
- for some $h \in \mathbb{R}^d$ the translation by h with $(\tau_h f)(x) := \tilde{f}(x+h),\ x \in \mathbb{R}^d$.

Our aim is to show that $W_0^{1,p}(Omega)$ is compactly embedded into $L^p(\Omega)$ and that the same is true for $W^{1,p}(Omega)$ if the boundary of Ω is of class C^1. In order to do this, we will need the following campctness criterion for L^p-spaces that we state without proof. It is in the spirit of the Arzelá-Ascoli Theorem and it can indeed be proved with the help of this result.

Proposition 9.2 (Kolmogorov-Riesz-Fréchet)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be bounded and measurable and $1 \leq p < \infty$. Then $\mathcal{F} \subseteq L^p(\Omega)$ is precompact if and only if*

a) $\mathcal{F}$ is bounded and

b) $\lim_{|h|\to 0} \left(\sup_{f \in \mathcal{F}} \left\| \tau_h f - \tilde{f} \right\|_{L^p(\Omega)} \right) = 0.$

Theorem 9.3 (Rellich's theorem)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain and $1 \leq p < \infty$. Then*

a) $\mathrm{W}_0^{1,p}(\Omega) \hookrightarrow \mathrm{L}^p(\Omega)$.

b) If $\partial\Omega \in \mathrm{C}^1$, then $\mathrm{W}^{1,p}(\Omega) \hookrightarrow \mathrm{L}^p(\Omega)$.

Proof. Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$. Then for all $x \in \mathbb{R}^d$ we have

$$\begin{aligned}
|\tau_h\varphi(x) - \varphi(x)| &= |\varphi(x+h) - \varphi(x)| \\
&= \left|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}[\varphi(x+th)]\,\mathrm{d}t\right| \\
&= \left|\int_0^1 \nabla\varphi(x+th)\cdot h\,\mathrm{d}t\right| \\
&\leq \int_0^1 |\nabla\varphi(x+th)|\cdot|h|\;\mathrm{d}t
\end{aligned}$$

which implies

$$\begin{aligned}
\|\tau_h\varphi - \varphi\|_{\mathrm{L}^p(\Omega)}^p &= \int_\Omega |\tau_h\varphi - \varphi|^p\;\mathrm{d}x \\
&\leq \int_\Omega \left(\int_0^1 1\cdot|\nabla\varphi(x+th)|\cdot|h|\;\mathrm{d}t\right)^p\;\mathrm{d}x \\
&\leq |h|^p \int_\Omega \left[\left(\int_0^1 1^{p'}\;\mathrm{d}t\right)^{\frac{1}{p'}}\cdot\left(\int_0^1 |\nabla\varphi(x+th)|^p\;\mathrm{d}t\right)^{\frac{1}{p}}\right]^p\;\mathrm{d}x \\
&\leq |h|^p \int_\Omega\int_0^1 |\nabla\varphi(x+th)|^p\;\mathrm{d}t\,\mathrm{d}x \\
&= |h|^p \int_0^1\int_{\mathbb{R}^d} |\nabla\varphi(x+th)|^p\;\mathrm{d}x\,\mathrm{d}t \\
&= |h|^p \int_0^1\int_{\mathbb{R}^d} |\nabla\varphi(y)|^p\;\mathrm{d}y\,\mathrm{d}t \\
&= |h|^p \int_{\mathbb{R}^d} |\nabla\varphi(y)|^p\;\mathrm{d}y \\
&\leq |h|^p\,\|\nabla\varphi\|_{\mathrm{L}^p(\mathbb{R}^d)}^p\,.
\end{aligned}$$

(b) Let $u \in \mathrm{W}^{1,p}(\Omega)$. Our goal is to use the Kolmogorov-Riesz-Fréchet compactness criterion. Hence we want to show that 9.2(b) holds for all bounded $\mathcal{F} \subseteq \mathrm{W}^{1,p}(\Omega)$.

The Extension Theorem 8.6 and the fact that $\mathrm{W}_0^{1,p}(\mathbb{R}^d) = \mathrm{W}^{1,p}(\mathbb{R}^d)$ imply: For $u \in \mathrm{W}^{1,p}(\Omega)$ there is a sequence $(\varphi_n)_n \subseteq \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $\varphi_n \to E(u)$ in $\mathrm{W}^{1,p}(\mathbb{R}^d)$.

Thus, by the continuity of the translation we obtain

$$\begin{aligned} \|\tau_h u - \tilde{u}\|_{\mathrm{L}^p(\Omega)} &= \lim_{n\to\infty} \|\tau_h \varphi_n - \varphi_n\|_{\mathrm{L}^p(\Omega)} \\ &\leq \lim_{n\to\infty} |h| \cdot \|\nabla \varphi_n\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &= |h| \cdot \|\nabla E(u)\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\leq |h| \cdot \|E(u)\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} \\ &\leq C\,|h| \cdot \|u\|_{\mathrm{W}^{1,p}(\Omega)} . \end{aligned}$$

If $\mathcal{F} \subseteq \mathrm{W}^{1,p}(\Omega)$ is bounded, then

- $\mathcal{F}$ is also bounded in $\mathrm{L}^p(\Omega)$,
- for all $u \in \mathcal{F}$ by the above it holds

$$\|\tau_h u - \tilde{u}\|_{\mathrm{L}^p(\Omega)} \leq C \cdot |h| \cdot M \xrightarrow{h\to 0} 0,$$

  where $M := \sup_{u\in\mathcal{F}} \|u\|_{\mathrm{W}^{1,p}(\Omega)}$.

Hence, part b) in Proposition 9.2 is fulfilled and $\mathcal{F}$ is precompact in $\mathrm{L}^p(\Omega)$.

(a) To prove the first assertion we just repeat the same proof but without the need of the Extension Theorem. □

As a corollary, we can prove an analogon to the Poincaré inequality (cf. Theorem 6.8) for another subspace of $\mathrm{W}^{1,p}(\Omega)$. Again we somehow have to avoid the constant functions.

Corollary 9.4 (Poincaré-Friedrichs inequality). *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with $\partial\Omega \in \mathrm{C}^1$ and $1 \leq p < \infty$. Then there exists some $C \geq 0$ such that for all $u \in \mathrm{W}^{1,p}(\Omega)$ it holds*

$$\int_\Omega |u|^p \,\mathrm{d}x \leq C \cdot \left(\int_\Omega |\nabla u|^p \,\mathrm{d}x + \left| \int_\Omega u \,\mathrm{d}x \right|^p \right).$$

Remark. The Poincaré-Friedrichs inequality is especially powerful for $u \in \mathrm{W}^{1,p}(\Omega)$ with $\int_\Omega u = 0$. In this case, the result is the Poincaré inequality.

Proof. Assume for a contradiction that the inequality is false. Then there exists a sequence of functions $u_n \in \mathrm{W}^{1,p}(\Omega)$ such that

$$\int_\Omega |u_n|^p \,\mathrm{d}x > n \cdot \left(\int_\Omega |\nabla u_n|^p \,\mathrm{d}x + \left| \int_\Omega u_n \,\mathrm{d}x \right|^p \right)$$

and $\|u_n\|_{\mathrm{L}^p(\Omega)} = 1$ for all $n \in \mathbb{N}$ (if not use $\frac{u_n}{\|u_n\|_{\mathrm{L}^p(\Omega)}}$ instead). This implies

$$\int_\Omega |\nabla u_n|^p \,\mathrm{d}x + \left|\int_\Omega u_n \,\mathrm{d}x\right|^p < \frac{1}{n} \text{ for all } n \in \mathbb{N},$$

so $\|u_n\|_{\mathrm{W}^{1,p}(\Omega)} = \left(\|u_n\|^p_{\mathrm{L}^p(\Omega)} + \|\nabla u_n\|^p_{\mathrm{L}^p(\Omega)}\right)^{\frac{1}{p}} \leq \left(1 + \frac{1}{n}\right)^{\frac{1}{p}} \leq 2^{\frac{1}{p}}$ for all $n \in \mathbb{N}$. Hence, by Theorem 9.3, there exists a subsequence u_{n_k} such that $u_{n_k} \to u$ in $\mathrm{L}^p(\Omega)$ and a subsubsequence $u_{n_{k_l}}$ which converges almost everywhere to some $u \in \mathrm{L}^p(\Omega)$. Therefore it holds

- $\|u\|_{\mathrm{L}^p(\Omega)} = \lim_{k\to\infty} \|u_{n_k}\|_{\mathrm{L}^p(\Omega)} = 1$.
- Since
$$\left|\int_\Omega u \,\mathrm{d}x\right| = \lim_{l\to\infty} \left|\int_\Omega u_{n_{k_l}} \,\mathrm{d}x\right| < \lim_{l\to\infty} \left(\frac{1}{n_{k_l}}\right)^{\frac{1}{p}} = 0,$$
we obtain
$$\int_\Omega u \,\mathrm{d}x = 0.$$
- $\|\nabla u_{n_k}\|_{\mathrm{L}^p(\Omega)} \leq \left(\frac{1}{n_k}\right)^{\frac{1}{p}} \to 0$ for $k \to \infty$. Let $\varphi \in \mathcal{C}^\infty_{\mathrm{c}}(\Omega)$. Then it holds
$$\int_\Omega u \partial_j \varphi \,\mathrm{d}x = \lim_{k\to\infty} \int_\Omega u_{n_k} \partial_j \varphi \,\mathrm{d}x = -\lim_{k\to\infty} \int_\Omega \partial_j u_{n_k} \varphi \,\mathrm{d}x.$$
Since
$$\left|\int_\Omega \partial_j u_{n_k} \varphi \,\mathrm{d}x\right| \leq \int_\Omega |\partial_j u_{n_k} \varphi| \,\mathrm{d}x \overset{\text{Hölder}}{\leq} \|\nabla u_{n_k}\|_{\mathrm{L}^p(\Omega)} \|\varphi\|_{\mathrm{L}^{p'}(\Omega)} \overset{k\to\infty}{\longrightarrow} 0,$$
we obtain $-\lim_{k\to\infty} \int_\Omega \partial_j u_{n_k} \varphi \,\mathrm{d}x = 0 = \int_\Omega 0 \cdot \varphi \,\mathrm{d}x$, so u is weakly differentiable and the weak gradient ∇u equals 0. Thus, u is constant.

The properties $\int_\Omega u \,\mathrm{d}x = 0$ and u constant together imply that $u = 0$ and this is a contradiction to $\|u\|_{\mathrm{L}^p(\Omega)} = 1$. □

Corollary 9.5. *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain, $\partial\Omega \in \mathrm{C}^1, 1 \leq p < \infty$ and*

$$\mathrm{L}^p_0(\Omega) := \left\{u \in \mathrm{L}^p(\Omega) \colon \int_\Omega u \,\mathrm{d}x = 0\right\}.$$

Then on $\mathrm{W}^{1,p}(\Omega) \cap \mathrm{L}^p_0(\Omega)$ the map $|u|_{1,p} := \|\nabla u\|_{\mathrm{L}^p(\Omega)}$ is a norm on $\mathrm{W}^{1,p}(\Omega) \cap \mathrm{L}^p_0(\Omega)$ that is equivalent to $\|\cdot\|_{\mathrm{W}^{1,p}(\Omega)}$.

Remark. In the literature, also Corollary 9.5 is often referred to as the Poincaré-Friedrich inequality.

Proof. For all $u \in \mathrm{W}^{1,p}(\Omega) \cap \mathrm{L}_0^p(\Omega)$ it holds

$$
\begin{aligned}
|u|_{1,p} &= \|\nabla u\|_{\mathrm{L}^p(\Omega)} \\
&\leq \|u\|_{\mathrm{W}^{1,p}(\Omega)} \\
&= \left(\|u\|_{\mathrm{L}^p}^p + \|\nabla u\|_{\mathrm{L}^p}^p\right)^{\frac{1}{p}} \\
&\overset{\text{Cor.9.4}}{\leq} \left[C \left(\|\nabla u\|_{\mathrm{L}^p(\Omega)}^p + \underbrace{\left| \int_\Omega u \,\mathrm{d}x \right|^p}_{=0} \right) + \|\nabla u\|_{\mathrm{L}^p(\Omega)}^p \right]^{\frac{1}{p}} \\
&= C \, \|\nabla u\|_{\mathrm{L}^p(\Omega)} \\
&= C \, |u|_{1,p} \, .
\end{aligned}
$$

□

10 Weak solutions to elliptic boundary value problems

Let $\Omega \subseteq \mathbb{R}^d$ be a domain.

Definition 10.1. Let $A \in \mathrm{L}^\infty(\Omega, \mathbb{R}^{d\times d}), b, c \in \mathrm{L}^\infty(\Omega, \mathbb{R}^d)$ and $e \in \mathrm{L}^\infty(\Omega)$. Then

a) we call

$$\begin{aligned} Lu &:= -\sum_{j,k=1}^{d} \partial_j \left[a_{jk}\partial_k u\right](x) - \sum_{j=1}^{d} \partial_j(b_j u) + \sum_{k=1}^{d} c_k \partial_k u + eu \\ &= -\operatorname{div}(A \cdot \nabla u) - \operatorname{div}(bu) + c \cdot \nabla u + eu \end{aligned}$$

differential operator of 2^{nd} order in divergence form.

b) L is called uniformly strongly elliptic, if there exists $\alpha_0 > 0$ such that for almost all $x \in \Omega$ and for all $\xi \in \mathbb{R}^d$ it holds

$$\langle A(x)\xi, \xi\rangle = \sum_{j,k=1}^{d} a_{jk}\xi_j\xi_k \geq \alpha_0 \left|\xi\right|^2 .$$

This is called Ellipticity condition.

c) We can generalise the definition of uniform strong ellipticity to the complex valued case, i.e. $A \in \mathrm{L}^\infty(\Omega, \mathbb{C}^{d\times d})$ and $b, c \in \mathrm{L}^\infty(\Omega, \mathbb{C}^d)$. Here we have to require that for all $\xi \in \mathbb{C}^d$ it holds

$$\operatorname{Re}\left(\langle A(x)\xi, \xi\rangle\right) = \operatorname{Re}\left(\sum_{j,k=1}^{d} a_{jk}\xi_j\overline{\xi_k}\right) \geq \alpha_0 \left|\xi\right|^2 .$$

Remark 10.2. a) On $\mathrm{H}^1(\Omega) \times \mathrm{H}^1(\Omega)$, the operator L corresponds to a sesquilinear form. Formally, this means

$$\begin{aligned} \int_\Omega Lu\bar{v}\,\mathrm{d}x &= \int_\Omega -\operatorname{div}(A \cdot \nabla u) \cdot \bar{v}\,\mathrm{d}x - \int_\Omega \operatorname{div}(bu) \cdot \bar{v}\,\mathrm{d}x + \int_\Omega c \cdot \nabla u \cdot \bar{v}\,\mathrm{d}x + \int_\Omega eu\bar{v}\,\mathrm{d}x \\ &= \int_\Omega \langle A\nabla u, \nabla v\rangle\,\mathrm{d}x + \int_\Omega \langle bu, \nabla v\rangle\,\mathrm{d}x + \int_\Omega c \cdot \nabla u \cdot \bar{v}\,\mathrm{d}x + \int_\Omega eu\bar{v}\,\mathrm{d}x \\ &=: a(u, v). \end{aligned}$$

This is a well-defined sequilinear form on $\mathrm{H}^1(\Omega) \times \mathrm{H}^1(\Omega)$.

b) If $A, b \in \mathrm{C}^1$, then L can be rewritten as differential operator of 2^{nd} order in non-divergence form (see Definition 4.10). The advantage of the non-divergence form is that the maximum principle 4.12 holds, whereas the advantage of the divergence form is that we can use the form methods from Chapter 5.

c) In the real case, we can look at the symmetrised version of A by forming $\frac{1}{2}(A(x) + A^{\mathrm{T}}(x))$. This is uniformly positive definite almost everywhere in Ω.

Remark. We give a motivation why this approach is important. All around us, we see examples where we cannot suppose the coefficient functions to be continuous. For example, consider a table like in room 234 in the Mathebau. It consists of a wooden plate but metal legs. Wood and metal have in general quite different heat conductivities, so the heat conductivity function is not continuous but has a jump at the interface. Thus, in non-divergence form we cannot even write down the heat equation for such a table.

In Definition 10.1, we do not require the coefficient functions to be continuous anymore, so we can apply it to problems like this.

Definition 10.3. a) For $\mathbb{K} = \mathbb{R}$ we define

$$\mathrm{H}^{-1}(\Omega) \coloneqq (\mathrm{H}_0^1(\Omega))' = \{F\colon H_0^1(\Omega) \to \mathbb{R} \text{ linear and bounded}\}$$

together with the norm $\|F\|_{\mathrm{H}^{-1}(\Omega)} \coloneqq \sup\{|F(v)| : v \in \mathrm{H}_0^1(\Omega), \|v\|_{\mathrm{H}^1(\Omega)} = 1\}$.

b) For $\mathbb{K} = \mathbb{C}$ we define

$$\mathrm{H}^{-1}(\Omega) \coloneqq \{F\colon \mathrm{H}_0^1(\Omega) \to \mathbb{C} \mid F \text{ is antilinear and bounded}\},$$

where antilinear means that a function F satisfies

$$F(\lambda u + \mu v) = \bar{\lambda} F(u) + \bar{\mu} F(v)$$

for all $\lambda, \mu \in \mathbb{C}$ and all $u, v \in \mathrm{H}_0^1$. As norm we choose the same norm as in the first part.

Example 10.4. a) $\mathrm{L}^2(\Omega) \hookrightarrow \mathrm{H}^{-1}(\Omega)$: Let $f \in \mathrm{L}^2(\Omega)$. Then for all $v \in \mathrm{H}_0^1(\Omega)$,

$$F(v) = F_f(v) = \int_\Omega f \bar{v} \,\mathrm{d}x$$

is a bounded (anti-)linear functional. Indeed by the Cauchy-Schwarz inequality

$$|F(v)| = \left| \int_\Omega f \cdot \bar{v} \,\mathrm{d}x \right| \le \|f\|_{\mathrm{L}^2(\Omega)} \cdot \|v\|_{\mathrm{L}^2(\Omega)} \le \|f\|_{\mathrm{L}^2(\Omega)} \cdot \|v\|_{\mathrm{H}^1(\Omega)}$$

which implies $\|F\|_{\mathrm{H}^{-1}(\Omega)} \le \|f\|_{\mathrm{L}^2(\Omega)}$.

b) Let $f \in \mathrm{L}^2(\Omega)$. Then $\partial_j f$ defined by

$$(\partial_j f)(v) := -\int_\Omega f \overline{\partial_j v}\,\mathrm{d}x, \qquad v \in \mathrm{H}_0^1(\Omega)$$

is in $\mathrm{H}^{-1}(\Omega)$. This is the so-called "distributional derivative" of f. Note that in general, $\partial_j f$ is not a function, we will come back to this in Chapter 15. Indeed,

$$\partial_j f(v) = \left| -\int_\Omega f \partial_j v\,\mathrm{d}x \right| \le \|f\|_{\mathrm{L}^2(\Omega)} \|\partial_j v\|_{\mathrm{L}^2(\Omega} \le \|f\|_{\mathrm{L}^2(\Omega)} \cdot \|v\|_{\mathrm{H}^1(\Omega)}$$

which implies $\|\partial_j f\|_{\mathrm{H}^{-1}(\Omega)} \le \|f\|_{\mathrm{L}^2(\Omega)}$.

Notation. For $F \in \mathrm{H}^{-1}(\Omega)$ and $v \in \mathrm{H}_0^1(\Omega)$ it is common to write $\langle F, v\rangle$ instead of $F(v)$. In general, this is not really a scalar product, but it behaves like one in most senses. To stress that it is not a scalar product, one can also write $\langle F, v\rangle_{\mathrm{H}^{-1},\mathrm{H}_0^1}$.

Definition 10.5 (Weak formulation of the Dirichlet problem)**.** Let Ω be a bounded domain, $d \ge 2, L$ a uniformly strongly elliptic differential operator in divergence form, a the corresponding sesquilinear form and $F \in \mathrm{H}^{-1}(\Omega)$.

a) $u \in \mathrm{H}_0^1(\Omega)$ is called a weak solution to the Dirichlet problem for L

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

if u solves $a(u, v) = F(v)$ for all $v \in \mathrm{H}_0^1(\Omega)$.

b) Let $\varphi \in \mathrm{H}^1(\Omega)$. Then $u \in \mathrm{H}^1(\Omega)$ is called weak solution to the inhomogeneous Dirichlet problem for L

$$\begin{cases} Lu &= F \text{ in } \Omega \\ \quad u &= \varphi \text{ on } \partial\Omega \end{cases}$$

if $u - \varphi \in \mathrm{H}_0^1(\Omega)$ and $a(u - \varphi, v) = F(v) - a(\varphi, v)$ is fulfilled for all $v \in \mathrm{H}_0^1(\Omega)$.

Remark 10.6. The boundary values $u = 0$ are understood in the trace sense and are incorporated into the space of test functions $\mathrm{H}_0^1(\Omega)$.

Theorem 10.7 (Gårding's inequality)**.** *With the notation from above there exists*

$$\lambda_0 := \frac{1}{2\alpha_0}\left(\|b\|_{\mathrm{L}^\infty(\Omega)} + \|c\|_{\mathrm{L}^\infty(\Omega)}\right)^2 + \|\mathrm{Re}(e)_-\|_{\mathrm{L}^\infty(\Omega)} \ge 0,$$

such that for all $u \in \mathrm{H}^1(\Omega)$ it holds

$$\mathrm{Re}(a(u,u)) + \lambda_0 \|u\|_{\mathrm{L}^2(\Omega)}^2 \ge \frac{\alpha_0}{2} |u|_{1,2}^2,$$

where as before $|u|_{1,2} := \|\nabla u\|_{\mathrm{L}^2(\Omega)}$.

Proof. Let $u \in \mathrm{H}^1(\Omega)$. We will try to estimate the real part of the sesquilinear form:

- We obtain

$$\begin{aligned}\operatorname{Re}\left(\int_\Omega \langle A\nabla u, \nabla u\rangle\right) &= \int_\Omega (\langle A\nabla u, \nabla u\rangle)\,\mathrm{d}x \\ &\overset{\text{Ellipticity}}{\geq} \int_\Omega \alpha_0 \left|\nabla u\right|^2 \,\mathrm{d}x = \alpha_0 \left|u\right|_{1,2}^2 .\end{aligned}$$

- Now, we rather crudely estimate the real part against the absolute value and use Young's inequality to make the first order term arbitrarily small. This is a common trick to absorb the first order term into the second order term. We obtain

$$\begin{aligned}&\operatorname{Re}\left[\int_\Omega u\,\langle b, \nabla u\rangle \,\mathrm{d}x + \int_\Omega c\cdot\nabla u\cdot\bar{u}\,\mathrm{d}x + \int_\Omega eu\bar{u}\,\mathrm{d}x\right] \\ &\geq -\left|\int_\Omega u\,\langle b, \nabla u\rangle \,\mathrm{d}x + \int_\Omega c\cdot\nabla u\cdot\bar{u}\,\mathrm{d}x\right| + \int_\Omega [\operatorname{Re}(e)]_- \left|u\right|^2 \,\mathrm{d}x \\ &\geq -\left(\left\|b\right\|_{\mathrm{L}^\infty} \frac{1}{\varepsilon}\left\|u\right\|_{\mathrm{L}^2} \varepsilon \left\|\nabla u\right\|_{\mathrm{L}^2} + \left\|c\right\|_{\mathrm{L}^\infty}\left\|u\right\|_{\mathrm{L}^2}\left\|\nabla u\right\|_{\mathrm{L}^2}\right) \\ &\qquad - \left\|(\operatorname{Re}(e))_-\right\|_{\mathrm{L}^\infty}\left\|u\right\|_{\mathrm{L}^2}^2 \\ &\overset{\text{Young}}{\geq} -\left(\left\|b\right\|_{\mathrm{L}^\infty} + \left\|c\right\|_{\mathrm{L}^\infty}\right)\left(\frac{\varepsilon^2}{2}\left|u\right|_{1,2}^2 + \frac{1}{2\varepsilon^2}\left\|u\right\|_{\mathrm{L}^2}^2\right) - \left\|(\operatorname{Re}(e)_-\right\|_{\mathrm{L}^\infty}\left\|u\right\|_{\mathrm{L}^2}^2 .\end{aligned}$$

  Altogether, this yields

$$\begin{aligned}\operatorname{Re}(a(u,u)) \geq& \alpha_0 \left|u\right|_{1,2}^2 - \left(\left\|b\right\|_\infty + \left\|c\right\|_\infty\right)\left(\frac{\varepsilon^2}{2}\left|u\right|_{1,2}^2 + \frac{1}{2\varepsilon^2}\left\|u\right\|_{\mathrm{L}^2}^2\right) \\ &- \left\|(\operatorname{Re}(e))_-\right\|_{\mathrm{L}^\infty}\cdot\left\|u\right\|_{\mathrm{L}^2}^2 \\ =& \left[(\alpha_0 - \frac{\varepsilon^2}{2}(\left\|b\right\|_{\mathrm{L}^\infty} + \left\|c\right\|_{\mathrm{L}^\infty})\right]\left|u\right|_{1,2}^2 \\ &- \left[\frac{1}{2\varepsilon^2}(\left\|b\right\|_{\mathrm{L}^\infty}) + \left\|c\right\|_{\mathrm{L}^\infty} + \left\|(\operatorname{Re}(e))_-\right\|_{\mathrm{L}^\infty}\right]\left\|u\right\|_{\mathrm{L}^2}^2 .\end{aligned}$$

  Choosing $\varepsilon^2 := \dfrac{\alpha_0}{\left\|b\right\|_{\mathrm{L}^\infty} + \left\|c\right\|_{\mathrm{L}^\infty}}$ we can estimate this further by

$$\geq \frac{\alpha_0}{2}\left|u\right|_{1,2}^2 - \lambda_0\left\|u\right\|_{\mathrm{L}^2}^2 .$$

□

Remark. Compared to the amount of work one has to put in to develop a theory of classical solutions, the involved calculations have been comparatively simple. Basically, the nice property of coercivity for the high-order terms allows us to deal with the noise invoked by the terms of orders one and zero.

Notice that we have completely dropped the assumptions on the border of Ω: The set could be the Koch snowflake for all we care.

Corollary 10.8. *a) For all* $\lambda \geq \lambda_0 := \frac{1}{2\alpha_0}\left(\|b\|_{\mathrm{L}^\infty(\Omega)} + \|c\|_{\mathrm{L}^\infty(\Omega)}\right)^2 + \|\mathrm{Re}(e)_-\|_{\mathrm{L}^\infty(\Omega)} \geq 0$ *the sesquilinear form* $a_\lambda(u,v) := a(u,v) + \lambda \langle u, v\rangle_{\mathrm{L}^2(\Omega)}$ *is coercive and continuous in* $\mathrm{H}^1_0(\Omega)$.

b) If $b = c = 0$ *and* $\mathrm{Re}(e) \geq 0$, *then* a *is itself coercive.*

Proof. **a):** For all $u, v \in \mathrm{H}^1_0(\Omega)$ we have that

$$
\begin{aligned}
|a_\lambda(u,v)| &\leq \left|\int_\Omega \langle A\nabla u, \nabla v\rangle \,\mathrm{d}x\right| + \left|\int_\Omega u\,\langle b, \nabla v\rangle \,\mathrm{d}x\right| \\
&\qquad + \left|\int_\Omega c\cdot\nabla u\cdot\bar{v}\,\mathrm{d}x\right| + \left|\int_\Omega eu\bar{v}\,\mathrm{d}x\right| + |\lambda|\left|\int_\Omega u\bar{v}\,\mathrm{d}x\right| \\
&\leq \|A\|_{\mathrm{L}^\infty}\|\nabla u\|_{\mathrm{L}^2}\|\nabla v\|_{\mathrm{L}^2} + \|b\|_{\mathrm{L}^\infty}\|u\|_{\mathrm{L}^2}\|\nabla v\|_{\mathrm{L}^2} \\
&\quad + \|c\|_{\mathrm{L}^\infty}\|\nabla u\|_{\mathrm{L}^2}\|v\|_{\mathrm{L}^2} + \|e\|_{\mathrm{L}^\infty}\|u\|_{\mathrm{L}^2}\|v\|_{\mathrm{L}^2} + |\lambda|\,\|u\|_{\mathrm{L}^2}\|v\|_{\mathrm{L}^2} \\
&\leq C\cdot\|u\|_{\mathrm{H}^1}\|v\|_{\mathrm{H}^1}.
\end{aligned}
$$

Hence, the form a_λ is continuous. Moreover, for all $u \in \mathrm{H}^1_0(\Omega)$ and for all $\lambda \geq \lambda_0$ we have

$$
\begin{aligned}
\mathrm{Re}(a_\lambda(u,u)) &= \mathrm{Re}(a(u,u)) + \lambda\,\mathrm{Re}(u,u) \\
&\overset{\text{Gårding}}{\geq} \frac{\alpha_0}{2}|u|_{1,2}^2 - \lambda_0\|u\|_{\mathrm{L}^2}^2 + \lambda\|u\|_{\mathrm{L}^2}^2 \\
&\geq \frac{\alpha_0}{2}|u|_{1,2}^2 = \frac{\alpha_0}{4}\|\nabla u\|_{\mathrm{L}^2}^2 + \frac{\alpha_0}{4}\|\nabla u\|_{\mathrm{L}^2}^2 \\
&\overset{\text{Poincaré}}{\geq} \frac{\alpha_0}{4}\|\nabla u\|_{\mathrm{L}^2}^2 + C\frac{\alpha_0}{4}\|u\|_{\mathrm{L}^2}^2 \\
&\geq C\|u\|_{\mathrm{H}^1}^2.
\end{aligned}
$$

b): is clear. □

We can now use Corollary 10.8 and the Lax-Milgram lemma 5.4 to prove the following result:

Theorem 10.9 (Existence of weak solutions to inhomogeneous problem)**.** *Imposing the above assumptions on domain, coefficients, ..., we have that for all* $f \in \mathrm{H}^{-1}(\Omega)$, *for all* $\varphi \in \mathrm{H}^1(\Omega)$ *and for all* $\lambda \in \{\mu \in \mathbb{C}\colon\ \mathrm{Re}(\mu) \geq \lambda_0\}$ *the inhomogeneous Dirichlet problem*

$$
\begin{cases} Lu + \lambda u &= f \text{ in } \Omega \\ \qquad\quad u &= \varphi \text{ on } \partial\Omega \end{cases} \tag{IDP}
$$

has a unique weak solution in $\mathrm{H}^1(\Omega)$. *Furthermore, the solution depends continuously on the data, i.e. there exists* $C \geq 0$ *independent of* u, f, φ, λ *such that*

$$\|u\|_{\mathrm{H}^1} \leq C(\|f\|_{H^{-1}} + (1+|\lambda|)\, \|\varphi\|_{\mathrm{H}^1}).$$

Proof. The weak formulation of (IDP) is

$$a_\lambda(u-\varphi, v) := a(u-\varphi, v) + \lambda(u-\varphi, v) = \underbrace{\langle f, v\rangle - a(\varphi, v) - \lambda(\varphi, v)}_{=\left\langle \tilde{F}, v\right\rangle}$$

for all $v \in \mathrm{H}^1_0(\Omega)$. We can interpret the right hand side as the application of some linear form $\tilde{F}$ to v. Since

$$\begin{aligned}\left|\left\langle \tilde{F}, v\right\rangle\right| &\leq \|f\|_{\mathrm{H}^{-1}} \|v\|_{\mathrm{H}^1} + C\, \|\varphi\|_{\mathrm{H}^1} \|v\|_{\mathrm{H}^1} + |\lambda|\, \|\varphi\|_{\mathrm{L}^2} \|v\|_{\mathrm{L}^2} \\ &\leq C\left(\|f\|_{\mathrm{H}^{-1}} + \|\varphi\|_{\mathrm{H}^1} + |\lambda|\, \|\varphi\|_{\mathrm{L}^2}\right) \|v\|_{\mathrm{H}^1},\end{aligned}$$

this linear form is continuous with

$$\left\|\tilde{F}\right\|_{\mathrm{H}^{-1}} \leq \|f\|_{\mathrm{H}^{-1}} + C\, \|\varphi\|_{\mathrm{H}^1} + |\lambda| \cdot \|\varphi\|_{\mathrm{L}^2}.$$

Let $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) \geq \lambda_0$. Then by Corollary 10.8

$$\begin{aligned}\mathrm{Re}[a_\lambda(u,u)] &= \mathrm{Re}(a(u,u) + \lambda_0(u,u)) + \mathrm{Re}((\lambda-\lambda_0)(u,u)) \\ &\geq C\frac{\alpha_0}{2}\, \|u\|^2_{\mathrm{H}^1(\Omega)} \underbrace{+(\mathrm{Re}\,\lambda - \lambda_0)}_{\geq 0} \|u\|^2_{\mathrm{L}^2} \geq C\frac{\alpha_0}{2}\, \|u\|^2_{\mathrm{H}^1(\Omega)}\end{aligned}$$

i.e. a_λ is continuous and coercive in $\mathrm{H}^1_0(\Omega)$. The Lax-Milgram Theorem now gives the claim with

$$\|u-\varphi\|_{\mathrm{H}^1} \leq C\left\|\tilde{F}\right\|_{\mathrm{H}^{-1}}.$$

This yields

$$\begin{aligned}\|u\|_{\mathrm{H}^1} &\leq \|u-\varphi\|_{\mathrm{H}^1} + \|\varphi\|_{\mathrm{H}^1} \\ &\leq C\left\|\tilde{F}\right\|_{\mathrm{H}^{-1}} + \|\varphi\|_{\mathrm{H}^1} \\ &\leq C\left(\|f\|_{\mathrm{H}^{-1}} + C\, \|\varphi\|_{\mathrm{H}^1} + |\lambda|\, \|\varphi\|_{\mathrm{L}^2}\right) + \|\varphi\|_{\mathrm{H}^1} \\ &\leq C\left(\|f\|_{\mathrm{H}^{-1}} + (1+|\lambda|)\, \|\varphi\|_{\mathrm{H}^1}\right).\end{aligned}$$

□

Remark. If $b = c = 0$ and $e \geq 0$, we can choose $\lambda_0 = 0$ to solve the inhomogeneous Dirichlet problem.

Now, we focus on the homogeneous case. For $f \in \mathrm{H}^{-1}(\Omega)$ and $\lambda \in \mathbb{C}$ consider

$$\begin{cases} L_\lambda u = Lu + \lambda u & = f \text{ in } \Omega \\ \qquad\qquad\quad u & = 0 \text{ on } \partial\Omega. \end{cases} \tag{DP$_\lambda$}$$

By Theorem 10.9 for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) \geq \lambda_0$ there is a unique weak solution $u \in \mathrm{H}^1_0(\Omega)$ for all $f \in \mathrm{H}^{-1}(\Omega)$.

Proposition 10.10. *Let*

$$L_\lambda^{-1} \colon \begin{cases} \mathrm{H}^{-1}(\Omega) & \to \mathrm{H}^1_0(\Omega) \\ \qquad f & \mapsto u \end{cases}$$

be the solution operator for (DP_λ). *For* $\mathrm{Re}(\lambda) \geq \lambda_0$, *the operator* L_λ^{-1} *is linear and continuous and its restriction* $L_\lambda^{-1}|_{\mathrm{H}^1_0(\Omega)}$ *is compact.*

Proof. **Linearity:** Let $f, g \in \mathrm{H}^{-1}(\Omega)$ and $\alpha, \beta \in \mathbb{C}$. We want to show that $\alpha L_\lambda^{-1} f + \beta L_\lambda^{-1} g$ is a solution to (DP_λ) with right hand side $\alpha f + \beta g$. For all $v \in \mathrm{H}^1_0(\Omega)$ we obtain

$$\begin{aligned} a_\lambda(\alpha L_\lambda^{-1} f + \beta L_\lambda^{-1} g, v) &= \alpha a_\lambda(L_\lambda^{-1} f, v) + \beta a_\lambda(L_\lambda^{-1} g, v) \\ &= \alpha \langle f, v \rangle + \beta \langle g, v \rangle = \langle \alpha f + \beta g, v \rangle \end{aligned}$$

which shows that $\alpha L_\lambda^{-1} f + \beta L_\lambda^{-1} g$ is the weak solution to (DP_λ) with right-hand side $\alpha f + \beta g$. Uniqueness yields $\alpha L_\lambda^{-1} f + \beta L_\lambda^{-1} g = L_\lambda^{-1}(\alpha f + \beta g)$.

Continuity: We have

$$\left\| L_\lambda^{-1} f \right\|_{\mathrm{H}^1} = \|u\|_{\mathrm{H}^1} \overset{\text{Thm. 10.9}}{\leq} C \, \|f\|_{\mathrm{H}^{-1}} .$$

Compactness: Fix a ball $B \subseteq \mathbb{R}^d$ with $\Omega \Subset B$. Then

$$\mathrm{H}^1_0(\Omega) \overset{\text{extend by 0}}{\hookrightarrow} \mathrm{H}^1_0(B) \overset{\text{Rellich}}{\hookrightarrow} \mathrm{L}^2(B) \overset{\text{restrict}}{\hookrightarrow} \mathrm{L}^2(\Omega) \overset{\text{inclusion}}{\hookrightarrow} \mathrm{H}^{-1}(\Omega) \overset{L_\lambda^{-1}}{\to} \mathrm{H}^1_0(\Omega),$$

so $L_\lambda^{-1}|_{\mathrm{H}^1_0(\Omega)}$ is compact. □

Theorem 10.11 (Existence Theorem for weak solutions, homogeneous case, extended version)**.** *There is (at most) a sequence* $(\lambda_k) \subseteq \{\mu \in \mathbb{C} : \mathrm{Re}(\mu) < \lambda_0\}$ *with* $\lim_{k\to\infty} |\lambda_k| \to \infty$ *such that for all* $\lambda \in \mathbb{C} \setminus \{\lambda_k \colon k \in \mathbb{N}\}$ *and for all* $f \in \mathrm{H}^{-1}(\Omega)$, ($\mathrm{DP}_\lambda$) *has a unique weak solution* $u \in \mathrm{H}^1_0(\Omega)$ *and* $L_\lambda^{-1} \colon \mathrm{H}^{-1}(\Omega) \to \mathrm{H}^1_0(\Omega)$ *is linear and continuous.*

Proof. By Proposition 10.10, the operator $L_{\lambda_0}^{-1} \colon \mathrm{H}^1_0(\Omega) \to \mathrm{H}^1_0(\Omega)$ is compact. From functional analysis we know that the eigenvalues of a compact operator form a (null-)sequence $\mu_k \subseteq \mathbb{C} \setminus \{0\}$ of eigenvalues of finite multiplicity and with finite-dimensional eigenspaces of $L_{\lambda_0}^{-1}$ and (μ_k) can only accumulate at 0.

The formal idea is the following: Rewrite $L_{\lambda_0}^{-1} = (L + \lambda_0)^{-1}$ to obtain for $\lambda \neq \lambda_0$ that

$$\begin{aligned} Lu + \lambda u = (L + \lambda_0)u + (\lambda - \lambda_0)u &= f \\ \iff \; u + (\lambda - \lambda_0) L_{\lambda_0}^{-1} u &= L_{\lambda_0}^{-1} f \\ \iff \; \frac{1}{\lambda - \lambda_0} u + L_{\lambda_0}^{-1} u &= \frac{1}{\lambda - \lambda_0} L_{\lambda_0}^{-1} f. \end{aligned}$$

In essence, we have succesfully reduced our problem of solving a PDE to an eigenvalue problem for a bounded linear operator. Hence,

$$u = \left(\frac{1}{\lambda - \lambda_0} + L_{\lambda_0}^{-1}\right)^{-1} \frac{1}{\lambda - \lambda_0} L_{\lambda_0}^{-1} f$$

whenever $\frac{1}{\lambda_0 - \lambda}$ is not one of the μ_k.

Now we want to justify these formal calculations. For all $g, h \in \mathrm{H}_0^1(\Omega)$, since $L_{\lambda_0}^{-1}$ is the solution operator, we have

$$a_{\lambda_0}\left(L_{\lambda_0}^{-1} g, h\right) = \langle g, h\rangle = (g, h). \tag{$*$}$$

Thus, u is a weak solution of (DP$_\lambda$) with right-hand side $f \in \mathrm{H}^{-1}(\Omega)$ if and only if for all $v \in \mathrm{H}_0^1(\Omega)$ it holds that

$$\begin{aligned}
\langle f, v\rangle &= a_\lambda(u, v) = a(u, v) + \lambda(u, v) \\
&= a_{\lambda_0}(u, v) + (\lambda - \lambda_0)(u, v) \\
&\overset{(*)}{=} a_{\lambda_0}(u, v) + (\lambda - \lambda_0) a_{\lambda_0}\left(L_{\lambda_0}^{-1} u, v\right) \\
&= a_{\lambda_0}\left(u + (\lambda - \lambda_0) L_{\lambda_0}^{-1} u, v\right),
\end{aligned}$$

which is the case if and only if $L_{\lambda_0}^{-1} f = u + (\lambda - \lambda_0) L_{\lambda_0}^{-1} u$, since the form a_{λ_0} is coercive. Dividing by $(\lambda - \lambda_0)$ yields the final formula

$$\frac{1}{\lambda - \lambda_0} L_{\lambda_0}^{-1} f = \frac{1}{\lambda - \lambda_0} u + L_{\lambda_0}^{-1} u.$$

We wrap the proof up by using this formula: For $k \in \mathbb{N}$ set $\lambda_k := \lambda_0 - \frac{1}{\mu_k}$. Let $\lambda \in \mathbb{C} \setminus \{\lambda_k \colon k \in \mathbb{N}\}$. Then $\frac{1}{\lambda - \lambda_0} \neq \mu_k$ since

$$\frac{1}{\lambda_0 - \lambda} = \mu_k \iff \lambda_0 - \lambda = \frac{1}{\mu_k} \iff \lambda = \lambda_0 - \frac{1}{\mu_k} = \lambda_k.$$

So $L_{\lambda_0}^{-1} - \frac{1}{\lambda_0 - \lambda}$ is invertible on $\mathrm{H}_0^1(\Omega)$ and we have that

$$u := \left(L_{\lambda_0}^{-1} - \frac{1}{\lambda_0 - \lambda}\right)^{-1} \frac{1}{\lambda - \lambda_0} L_{\lambda_0}^{-1} f$$

is the unique solution of (DP$_\lambda$).

Furthermore,

$$L_\lambda^{-1} = \frac{1}{\lambda - \lambda_0} \Bigg(\underbrace{L_{\lambda_0}^{-1} - \frac{1}{\lambda_0 - \lambda}}_{\in \mathcal{L}(\mathrm{H}_0^1, \mathrm{H}_0^1)}\Bigg)^{-1} \underbrace{L_{\lambda_0}^{-1}}_{\in \mathcal{L}(\mathrm{H}^{-1}, \mathrm{H}_0^1)}$$

is linear and continuous as operator from $\mathrm{H}^{-1}(\Omega)$ to $\mathrm{H}_0^1(\Omega)$. □

This proof relies heavily on the fact that $L_{\lambda_0}^{-1}$ is a compact operator. Luckily, the solution operator behaves well enough, so that the right-hand side $f \in \mathrm{H}^{-1}(\Omega)$ gets mapped to a function in $\mathrm{H}_0^1(\Omega)$.

However, explicitly calculating the exceptional points remains difficult. In the following remark we make use of the fact that the eigenvalues have finite multiplicities and finite-dimensional eigenspaces.

Remark 10.12. Looking at the exceptional points λ_k, we notice that up to some point for all $\lambda \neq \lambda_0$ our calculations go through as well, i.e. we have $Lu + \lambda u = f$ if and only if

$$(\mu - L_{\lambda_0}^{-1})u = \mu L_{\lambda_0}^{-1} f, \qquad \text{where } \mu = \frac{1}{\lambda_0 - \lambda}.$$

This means that $Lu + \lambda u = f$ is uniquely solvable if and only if $\frac{1}{\lambda_0 - \lambda} \neq \mu_k$, i.e. if and only if λ is none of the λ_k. Furthermore, for all k the numbers $-\lambda_k$ are eigenvalues of L with finite dimensional eigenspaces.[1]

Definition 10.13. Let $a\colon \mathrm{H}_0^1(\Omega) \times \mathrm{H}_0^1(\Omega) \to \mathbb{C}$ be a sesquilinear form. Then the numerical range of a is given by

$$\mathrm{num}_0(a) = \{a(u,u)\colon u \in \mathrm{H}_0^1(\Omega),\ \|u\|_{\mathrm{L}^2(\Omega)} = 1\}.$$

Remark 10.14. The set $\mathrm{num}_0(a)$ is convex in $\mathbb{C}$, but it is neither open nor closed in general.

Proposition 10.15. *The set of negative exceptional points $\{-\lambda_k\colon k \in \mathbb{N}\}$ is contained in the closure of the numerical range, i.e.*

$$\{-\lambda_k\colon k \in \mathbb{N}\} \subseteq \overline{\mathrm{num}_0(a)}.$$

Proof. Let $\lambda \notin \overline{\mathrm{num}_0(a)}$. Then $\mathrm{dist}\left(\lambda, \overline{\mathrm{num}_0(a)}\right) > 0$. Let $u \in \mathrm{H}_0^1(\Omega) \setminus \{0\}$ and $w := \frac{u}{\|u\|_{\mathrm{L}^2(\Omega)}}$. Then

$$\begin{aligned} |a_{-\lambda}(u,u)| &= |a(u,u) - \lambda(u,u)| \\ &= \left| a(\|u\|_{\mathrm{L}^2(\Omega)}\, w, \|u\|_{\mathrm{L}^2(\Omega)}\, w) - \lambda\, \|u\|_{\mathrm{L}^2(\Omega)}^2 \right| \\ &= \|u\|_{\mathrm{L}^2(\Omega)}^2\, |\underbrace{a(w,w)}_{\in \mathrm{num}_0(a)} - \lambda| \\ &\geq \underbrace{\mathrm{dist}(\lambda, \overline{\mathrm{num}_0(a)})}_{>0}\, \|u\|_{\mathrm{L}^2(\Omega)}^2 . \end{aligned}$$

[1] Usually, we can only say that λ_k is a spectral element, here, however, we know that $\sigma(L)$ consists of eigenvalues only.

At a first glance, this might look like a coercivity estimate, but estimating the L^2-norm by the H^1-norm is the wrong direction. However, the estimate yields injectivity, as we will see:

If u solves the homogeneous problem

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

then $u \in H_0^1(\Omega)$ and for all $v \in H_0^1(\Omega)$ we have

$$\begin{aligned} & a_{-\lambda}(u, v) = \langle 0, v\rangle = 0 \\ & \implies 0 = |a_{-\lambda}(u,u)| \geq \operatorname{dist}(\lambda, \overline{\operatorname{num}_0(a)}) \, \|u\|^2_{L^2(\Omega)} \\ & \implies \|u\|^2_{L^2(\Omega)} = 0 \\ & \implies u = 0. \end{aligned}$$

By Remark 10.12 we obtain that $L_{-\lambda}$ is invertible, i.e. (DP_λ) is uniquely solvable for all $f \in H^{-1}(\Omega)$. Hence, λ cannot be one of the $-\lambda_k$. □

Corollary 10.16. *For all $-\lambda \in \mathbb{C} \setminus \overline{\operatorname{num}_0(a)}$ and for all $f \in H^{-1}(\Omega)$ there is a unique weak solution $u \in H_0^1(\Omega)$ to (DP_λ). Moreover, there is a constant $C > 0$ independent of u, λ and f such that*

a) $\|u\|_{H^1} \leq C \cdot \|f\|_{H^{-1}(\Omega)}$.

b) For all $f \in L^2(\Omega)$ it holds that $\|u\|_{L^2(\Omega)} \leq \frac{1}{\operatorname{dist}(\lambda, \overline{\operatorname{num}_0(a)})} \|f\|_{L^2(\Omega)}$.[2]

Proof. a) Follows by Theorem 10.11 and Proposition 10.15

b) The proof of Proposition 10.15 gives

$$\|u\|^2_{L^2(\Omega)} \operatorname{dist}(\lambda, \overline{\operatorname{num}_0(a)}) \leq |a_{-\lambda}(u,u)| = |\langle f, u\rangle| = |(f,u)| \leq \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}.$$ □

Example 10.17. For $L = -\Delta$, i.e. $A = I, b = c = 0, e = 0$ we can estimate the

[2] The explicit knowledge of how the constant depends on λ might not look very useful or interesting at first. However, in the theory of parabolic equations strong results are available in cases in which the corresponding elliptic equation is solvable and information about precisely this relationship is available.

numerical range rather easily:

$$
\begin{aligned}
\operatorname{num}_0(a) &= \Big\{ a(u,u) \colon u \in \mathrm{H}^1_0(\Omega), \|u\|_{\mathrm{L}^2(\Omega)} = 1 \Big\} \\
&= \left\{ \int_\Omega \langle I \nabla u, \nabla u \rangle \, \mathrm{d}x \colon u \in \mathrm{H}^1_0(\Omega), \|u\|_{\mathrm{L}^2(\Omega)} = 1 \right\} \\
&= \left\{ \int_\Omega |\nabla u|^2 \, \mathrm{d}x \colon u \in \mathrm{H}^1_0(\Omega) \colon \; \|u\|_{\mathrm{L}^2(\Omega)} = 1 \right\} \\
&\subseteq (0, \infty),
\end{aligned}
$$

so

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

is uniquely solvable in the weak sense for all $\lambda \in \mathbb{C} \setminus [0, \infty)$. By Gårding's inequality 10.7, we get this result even for all $\lambda \in \mathbb{C} \setminus (0, \infty)$ since $\lambda_0 = 0$.

11 Neumann and natural boundary condition

In the classical case, the Neumann problem was given by

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

However, by now we are dealing with arbitrary bounded domains Ω, so we have completely dropped the smoothness requirements on $\partial\Omega$ that previously guaranteed the existence of a normal vector in every point, hence we must find other means of expressing the desired behaviour on the boundary. But even if we assume the existence of a normal vector, the above condition is a pointwise condition on an L^2-function, which is not meaningful, as the function does not have a trace in general. We will see that we can consider the operators on the whole space $(\mathrm{H}^1(\Omega))'$ and will obtain a natural boundary condition for free.

Definition 11.1. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain,

$$Lu = -\operatorname{div}(A \cdot \nabla u) + c \cdot \nabla u + eu$$

be an elliptic operator[1] with $A, c, e \in \mathrm{L}^\infty$ and $f \in (\mathrm{H}^1(\Omega))'$ an antilinear functional on $\mathrm{H}^1(\Omega)$.[2] Define $a\colon \mathrm{H}^1(\Omega) \times \mathrm{H}^1(\Omega) \to \mathbb{C}$ as before. Then we call $u \in \mathrm{H}^1(\Omega)$ a weak solution to the Neumann problem

$$\begin{cases} Lu &= f, \text{ in } \Omega \\ \nu \cdot A\nabla u &= 0, \text{ on } \partial\Omega, \end{cases} \tag{NP}$$

if for all $v \in \mathrm{H}^1(\Omega)$ it holds that $a(u,v) = \langle f, v\rangle_{(\mathrm{H}^1)', \mathrm{H}^1}$. Furthermore, $\nu \cdot A\nabla u$ is called the conormal derivative of u.

Remark 11.2. a) The PDE is only a formal writing. In general, only the variational problem can be expressed in a meaningful way.

[1] We require $b = 0$ for simplicity.
[2] Note that this is not the space $\mathrm{H}^{-1}(\Omega)$.

b) For the Laplace case, i.e. $A = I$, $\nu \cdot A\nabla u = \dfrac{\partial u}{\partial \nu}$ is the normal derivative.

Lemma 11.3. *Let $A \in \mathrm{C}^1(\overline{\Omega}, \mathbb{C}^{d\times d}), c \in \mathrm{C}(\Omega, \mathbb{C}^d), e \in \mathrm{C}(\Omega, \mathbb{C}), f \in \mathrm{C}(\overline{\Omega})$ and $\partial\Omega \in \mathrm{C}^1$. If $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$ is a weak solution to (NP), then u solves (NP) in the classical sense.*

Proof. Since Ω is bounded, we have $\mathrm{C}^1(\overline{\Omega}) \subseteq \mathrm{H}^1(\Omega)$ and for $f \in \mathrm{C}(\overline{\Omega})$

$$w \mapsto \int_\Omega f\overline{w} \in (\mathrm{H}^1(\Omega))'.$$

Let $w \in \mathrm{C}^1(\overline{\Omega})$ and denote by ν the outer unit normal to $\partial\Omega$. Then by integrating by parts we obtain

$$\begin{aligned}
\int_\Omega f\overline{w}\,\mathrm{d}x &= \langle f, w\rangle_{(\mathrm{H}^1)',\mathrm{H}^1} = a(u,w) \\
&= \int_\Omega \langle A\nabla u, \nabla w\rangle\,\mathrm{d}x + \int_\Omega (c\cdot\nabla u\overline{w} + eu\overline{w})\,\mathrm{d}x \\
&= -\int_\Omega \operatorname{div}(A\nabla u)\overline{w}\,\mathrm{d}x + \int_{\partial\Omega} \langle A\nabla u, \nu\rangle\,\overline{w}\,\mathrm{d}S + \int_\Omega (c\cdot\nabla u\cdot\overline{w} + eu\overline{w})\,\mathrm{d}x \\
&= \int_\Omega Lu\cdot\overline{w}\,\mathrm{d}x + \int_{\partial\Omega} \langle A\nabla u, \nu\rangle\,\overline{w}\,\mathrm{d}S.
\end{aligned}$$

Hence, for all $w \in \mathcal{C}_c^\infty(\Omega)$, we obtain

$$\int_\Omega f\overline{w}\,\mathrm{d}x = \int_\Omega Lu\cdot\overline{w}\,\mathrm{d}x$$

and we conclude for all $w \in \mathcal{C}_c^\infty(\Omega)$ that

$$\int_\Omega (Lu - f)\overline{w}\,\mathrm{d}x = 0.$$

By the fundamental theorem of calculus of variations 5.2, this implies $Lu = f$ in $\mathrm{L}^2(\Omega)$. Since Lu and f are continuous, this also means $Lu = f$. Then

$$\int_{\partial\Omega} \langle A\nabla u, \nu\rangle\,\overline{w}\,\mathrm{d}S = 0$$

for all $w \in \mathrm{C}^1(\overline{\Omega})$, so $\nu\cdot A\nabla u = 0$ on $\partial\Omega$. □

Remark 11.4. a) In this sense one "solves" the Neumann problem even if the conormal derivative is not defined.

b) A greater class of test functions (i.e. $\mathrm{H}^1(\Omega)$ instead of $\mathrm{H}^1_0(\Omega)$) gives the boundary condition for free. For this reason this boundary condition is called "natural" boundary condition.

c) For $c = 0, e = 0$ the problem above simplifies to

$$\begin{cases} Lu &= -\operatorname{div}(A\nabla u) = f, \text{ in } \Omega \\ \nu \cdot A\nabla u &= 0, \text{ on } \partial\Omega. \end{cases} \tag{1}$$

This problem cannot have a unique solution, since if u is a solution, then also $u+C$ where $C \in \mathbb{C}$ is a constant is a solution. In other words: $\lambda = 0$ is an eigenvalue with eigenfunction 1 of L. There is also another way to see this. This time we start by looking at the right hand side. u is a weak solution to (1) if and only if

$$a(u, v) = \int_\Omega \langle A\nabla u, \nabla v\rangle \, \mathrm{d}x = \langle f, v\rangle$$

for all $v \in \mathrm{H}^1(\Omega)$. Choosing $v = 1$ this yields

$$\langle f, 1\rangle = a(u, 1) = \int_\Omega \langle A\nabla u, \nabla 1\rangle \, \mathrm{d}x = 0,$$

the so called "compatibility condition". For $f \in \mathrm{L}^2(\Omega)$, this means $\int_\Omega f \cdot 1 \, \mathrm{d}x = 0$, i.e. f has to be orthogonal to the constant 1-function, otherwise a solution cannot exist. We introduce the notation

$$\mathrm{L}^2_0(\Omega) := \left\{ f \in \mathrm{L}^2(\Omega) \colon \int_\Omega f \, \mathrm{d}x = 0 \right\}.$$

Definition 11.5. Let $a\colon \mathrm{H}^1(\Omega) \times \mathrm{H}^1(\Omega) \to \mathbb{C}$ be a sesquilinear form. Then the numerical range of a is defined by $\operatorname{num}(a) := \{a(u, u) \colon u \in \mathrm{H}^1(\Omega), \|u\|_{\mathrm{L}^2(\Omega)} = 1\}$.

Theorem 11.6. *Let $\partial\Omega \in \mathrm{C}^1$ Then it holds:*

a) There is (at most) a sequence $(\lambda_k) \subseteq \overline{\operatorname{num}(a)} \cap \{\operatorname{Re}(\mu) \leq \lambda_0\}$ such that all $-\lambda_k$ are eigenvalues of L with finite-dimensional eigenspaces with $|\lambda_k| \to \infty$ as $k \to \infty$, such that for all $\lambda \in \mathbb{C} \setminus \{\lambda_k \colon k \in \mathbb{N}\}$ the operator

$$L + \lambda := L_\lambda \colon \ \mathrm{H}^1(\Omega) \to (\mathrm{H}^1(\Omega))', \ u \mapsto a_\lambda(u, \cdot)$$

(In a formal sense this coincides with $\langle L_\lambda u, \cdot\rangle$.) is a continuous linear and invertible operator with compact inverse, so in particular for all $f \in (\mathrm{H}^1(\Omega))'$ and for all $\lambda \in \mathbb{C} \setminus \{\lambda_k \colon k \in \mathbb{N}\} \supseteq \mathbb{C} \setminus \operatorname{num}(a)$ there is a unique weak solution $u \in \mathrm{H}^1(\Omega)$ to

$$\begin{cases} Lu + \lambda u &= f \text{ in } \Omega \\ \nu \cdot A\nabla u &= 0 \text{ on } \partial\Omega. \end{cases} \tag{NP_λ}$$

b) *For all* $f \in \mathrm{L}^2(\Omega)$ *and for all* $\lambda \in \mathbb{C} \setminus \overline{\mathrm{num}(a)}$, *the solution* $u \in \mathrm{H}^1(\Omega)$ *to* (NP_λ) *satisfies*

$$\|u\|_{\mathrm{L}^2(\Omega)} \leq \frac{1}{\mathrm{dist}(\lambda, \mathrm{num}(a))} \cdot \|f\|_{\mathrm{L}^2(\Omega)} .$$

c) *If* $c = 0, e = 0$, *then* $\lambda = 0$ *is an eigenvalue of* $L_0 = L$ *and* (NP_0) *has a solution* $u \in \mathrm{H}^1(\Omega)$ *if and only if* $f \in \{F \in (\mathrm{H}^1(\Omega))' \colon F(1) = 0\}$. *In this case,* u *is unique if we require* $u \in \mathrm{H}^1(\Omega) \cap \mathrm{L}^2_0(\Omega)$.

Proof. Most of the arguments needed in this proof are very similar to the procedure to find a solution of the Dirichlet problem, so we won't do all the details again. The main idea is the following:

By Gårdings inequality 10.7, $\mathrm{Re}(a_{\lambda_0}(u,u)) \geq \frac{\alpha_0}{2} |u|^2_{1,2} = \|\nabla u\|^2_{\mathrm{L}^2}(\Omega)$. Thus,

$$\begin{aligned}
\mathrm{Re}(a_{\lambda_0+\varepsilon}(u,u)) &= a_{\lambda_0}(u,u) + \varepsilon \|u\|^2_{\mathrm{L}^2(\Omega)} \\
&\geq \frac{\alpha_0}{2} \|\nabla u\|^2_{\mathrm{L}^2(\Omega)} + \varepsilon \|u\|^2_{\mathrm{L}^2(\Omega)} \\
&\geq \min\left(\frac{\alpha_0}{2}, \varepsilon\right) \|u\|^2_{\mathrm{H}^1(\Omega)} .
\end{aligned}$$

Hence, a_{λ_0+1} is coercive, so Lax-Milgram implies that $L^{-1}_{\lambda_0+1} \colon (\mathrm{H}^1(\Omega))' \to \mathrm{H}^1(\Omega)$ is linear and continuous. The restriction of $L^{-1}_{\lambda_0+1}$ to $\mathrm{H}^1(\Omega)$ is compact since $\mathrm{H}^1(\Omega) \hookrightarrow \mathrm{L}^2(\Omega)$ by Rellich[3]. For parts a) and b), the rest of the proof is as before.

For part c) note that $c = 0, e = 0$ implies $\lambda_0 = 0$, so

$$\mathrm{Re}(a(u,u)) \geq \frac{\alpha_0}{2} |u|^2_{1,2} \geq C \|u\|^2_{\mathrm{H}^1(\Omega)} .$$

By the Poincaré-Friedrichs inequality on $\mathrm{H}^1(\Omega) \cap \mathrm{L}^2_0(\Omega)$. Hence, a is coercive on $\mathrm{H}^1(\Omega) \cap L^2_0(\Omega)$. □

[3] Note that $\partial\Omega$ was required to be C^1.

12 Difference quotients of Sobolev functions

We know already that convergence of difference quotients yields differentiability. Thus, one can hope that convergence of difference quotients *in* $L^1(\Omega)$ gives weak differentiability. This is actually true and we will show it in this chapter. Let $\Omega \subseteq \mathbb{R}^d$ open and $\{e_j : j \in \{1, \dots, d\}\}$ denote the standard basis of $\mathbb{R}^d$.

Notation 12.1. Let $u \in L^1_{\mathrm{loc}}(\Omega)$, $j \in \{1, \dots, d\}, x \in \Omega$ and $h \in \mathbb{R} \setminus \{0\}$ such that $x + he_j \in \Omega$. Then we define the difference quotient in direction e_j of u by

$$\delta_h^j u(x) \coloneqq \frac{u(x + he_j) - u(x)}{h}.$$

Lemma 12.2. *Let $u, v \in L^1_{\mathrm{loc}}(\Omega)$ such that $u \cdot v \in L^1_{\mathrm{loc}}(\Omega)$. Then*

a) $\delta_h^j (uv)(x) = \left[\delta_h^j u(x)\right] \cdot v(x) + u(x + he_j)\delta_h^j v(x)$ *for* $x \in \Omega$.

b) If $V \Subset \Omega$ and $\operatorname{supp}(v) \subseteq V$, then for all $h \in \mathbb{R} \setminus \{0\}$ which satisfy the conditions $|h| < \frac{1}{2} \operatorname{dist}(\operatorname{supp}(v), \partial V)$ and $|h| < \frac{1}{2} \operatorname{dist}(V, \partial\Omega)$ it holds that

$$\int_V u(\delta_h^j v) \, \mathrm{d}x = -\int_V (\delta_{-h}^j u) \cdot v \, \mathrm{d}x.$$

Proof. a) We have

$$\begin{aligned}
\delta_h^j (uv)(x) &= \frac{u(x + he_j)v(x + he_j) - u(x + he_j)v(x)}{h} + \frac{u(x + he_j)v(x) - u(x)v(x)}{h} \\
&= u(x + he_j)\delta_h^j v(x) + \delta_h^j u(x)v(x).
\end{aligned}$$

b) Set $\tilde{V} \coloneqq V + he_j$. Then $\tilde{V} \subseteq \Omega$ and $\operatorname{supp}(v) \subseteq V \cap \tilde{V}$. Hence, we obtain

$$\begin{aligned}
\int_V u\delta_h^j v \, \mathrm{d}x &= \int_V u(x)\frac{v(x + he_j) - v(x)}{h} \, \mathrm{d}x \\
&= \int_V \frac{u(x)v(x + he_j)}{h} \, \mathrm{d}x - \int_V \frac{u(x)v(x)}{h} \, \mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&\overset{y=x+he_j}{=} \int_{\tilde{V}} \frac{u(y-he_j)v(y)}{h}\,\mathrm{d}y - \int_V \frac{u(x)v(x)}{h}\,\mathrm{d}x \\
&= \int_V \frac{u(y-he_j)v(y)}{h}\,\mathrm{d}y - \int_V \frac{u(x)v(x)}{h}\,\mathrm{d}x \\
&= -\int_V \frac{u(x-he_j)v(x)-u(x)v(x)}{-h}\,\mathrm{d}x \\
&= -\int_V (\delta^j_{-h}v)(x)u(x)\,\mathrm{d}x. \qquad \square
\end{aligned}
$$

Remark. For $\Omega = \mathbb{R}^d, V = \mathbb{R}^d$ all $h \in \mathbb{R}\setminus\{0\}$ are possible, provided $u \cdot v \in \mathrm{L}^1(\mathbb{R}^d)$.

Lemma 12.3. *Let $p \in [1,\infty)$, $u \in \mathrm{W}^{1,p}(\Omega), V \Subset \Omega$ and $0 < |h| < \operatorname{dist}(V,\partial\Omega)$. Then $\delta^j_h u \in \mathrm{L}^p(V)$ and*

$$\left\|\delta^j_h u\right\|_{\mathrm{L}^p(V)} \le \|\nabla u\|_{\mathrm{L}^p(\Omega)}$$

for all $j \in \{1,\dots,d\}$.

Proof. Let $u \in \mathrm{C}^1(\Omega)\cap \mathrm{W}^{1,p}(\Omega)$. Then for $x \in V$ we obtain

$$
\begin{aligned}
\left|\delta^j_h u(x)\right|^p &= \left|\frac{u(x+he_j)-u(x)}{h}\right|^p \\
&= \frac{1}{|h|^p}\left|\int_0^h \partial_j u(x_1,\dots,x_{j-1},x_j+te_j,x_{j+1},\dots,x_d)\,\mathrm{d}t\right|^p \\
&\le \frac{1}{|h|^p}\left|\int_0^h \left|1\cdot\partial_j u(x+te_j)\right|\,\mathrm{d}t\right|^p.
\end{aligned}
$$

Using Hölder's inequality we can further estimate by

$$
\begin{aligned}
&\le \frac{1}{|h|^p}\left[\left|\int_0^h 1^{p'}\,\mathrm{d}t\right|^{\frac{1}{p'}}\left|\int_0^h |\partial_j u(x+te_j)|^p\,\mathrm{d}t\right|^{\frac{1}{p}}\right]^p \\
&= \frac{1}{|h|^p}|h|^{\frac{p}{p'}}\left|\int_0^h |\partial_j u(x+te_j)|^p\,\mathrm{d}t\right|,
\end{aligned}
$$

so we obtain since $\frac{p}{p'} - p = -1$,

$$
\begin{aligned}
\int_V \left|\delta^j_h u(x)\right|^p \mathrm{d}x &\le \frac{1}{|h|}\int_V\int_{[0,h]} |\partial_j u(x+te_j)|^p\,\mathrm{d}t\,\mathrm{d}x \\
&\overset{\text{Fubini}}{=} \frac{1}{|h|}\int_{[0,h]} \underbrace{\int_V |\partial_j u(x+te_j)|^p\,\mathrm{d}x}_{\le \|\nabla u\|^p_{\mathrm{L}^p(\Omega)}}\,\mathrm{d}t \\
&= \frac{1}{|h|}|h|\cdot\|\nabla u\|^p_{\mathrm{L}^p(\Omega)} = \|\nabla u\|^p_{\mathrm{L}^p(\Omega)}.
\end{aligned}
$$

Now, the claim for general $u \in \mathrm{W}^{1,p}(\Omega)$ follows by density. $\square$

Reminder 12.4 (Banach-Alaoglu). Let X be a separable Banach space and $(x'_n) \subseteq X'$ bounded. Then (x'_n) has a weak-$*$-convergent subsequence (x'_{n_k}), i.e. there is $x' \in X'$ such that $x'_{n_k}(x) \to x'(x)$ for all $x \in X$. Furthermore, it holds

$$\|x'\|_{X'} \leq \liminf_{k\to\infty} \left\|x'_{n_k}\right\|_{X'}.$$

Theorem 12.5. *Let $p \in (1,\infty], u \in \mathrm{L}^p(\Omega)$. Suppose there exists $K \geq 0$ such that for all $h \in \mathbb{R} \setminus \{0\}$ and all $V \Subset \Omega$ with $|h| < \frac{1}{2}\operatorname{dist}(V,\partial\Omega)$ we have $\delta_h^j u \in \mathrm{L}^p(V)$ and $\left\|\delta_h^j u\right\|_{\mathrm{L}^p(V)} \leq K$. Then the weak derivative $\partial_j u$ exists, $\partial_j u \in \mathrm{L}^p(\Omega)$ and $\|\partial_j u\|_{\mathrm{L}^p(\Omega)} \leq K$.*

Proof. Let $V \Subset \Omega$ and $0 < h_0 < \frac{1}{2}\operatorname{dist}(V,\partial\Omega)$. Then the set $\{\delta_h^j u \colon 0 < |h| < h_0\}$ is bounded in $\mathrm{L}^p(V) = (\mathrm{L}^{p'}(V))'$ for $\frac{1}{p} + \frac{1}{p'} = 1$. Therefore, Reminder 12.4 implies that there exists $(h_n) \in (-h_0, h_0)$ and $v_j \in \mathrm{L}^p(V)$ such that $h_n \to 0$ and $\delta_{h_n}^j u \overset{*}{\rightharpoonup} v_j$ for $n \to \infty$. In particular, for all $\varphi \in \mathcal{C}_c^\infty(V) \subseteq \mathrm{L}^{p'}(V)$ we have

$$\begin{aligned}\int_V v_j \varphi \,\mathrm{d}x &= \lim_{n\to\infty} \int_V \left(\delta_{h_n}^j u\right) \varphi \,\mathrm{d}x \\ &\overset{\text{Lem. 12.2}}{=} - \lim_{n\to\infty} \int_V u \delta_{-h_n}^j \varphi \,\mathrm{d}x \\ &= - \int_V u \partial_j \varphi \,\mathrm{d}x.\end{aligned}$$

This means that the weak derivative $\partial_j u$ exists, $\partial_j u = v_j \in \mathrm{L}^p(V)$ and

$$\|\partial_j u\|_{\mathrm{L}^p(V)} = \|v_j\|_{\mathrm{L}^p(V)} \leq \liminf_{n\to\infty} \left\|\partial_{h_n}^j u\right\|_{\mathrm{L}^p(V)} \leq K.$$

Since $V \Subset \Omega$ was arbitrary, take

$$V_m := \left\{x \in \Omega \colon \operatorname{dist}(x,\partial\Omega) > \frac{1}{m}, |x| < m\right\},$$

with $u = \lim_{m\to\infty} u \cdot \chi_{V_m}$. By Fatou's lemma, we get $\partial_j u \in \mathrm{L}^p(\Omega)$ and $\|\partial_j u\|_{\mathrm{L}^p(\Omega)} \leq K$. $\square$

13 Regularity theorems

By definition weak solutions are not better than H^1 in regularity. It is a natural question, whether the solution is somehow better, if the data of the problem (right hand side, boundary of the domain, coefficients) is nicer. We will first address interior regularity, i.e. we keep away from the boundary of the domain.

Theorem 13.1 (Interior regularity)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be open, $V \Subset \Omega$ and L be a uniformly elliptic differential operator as before. Suppose $A, b, c, e \in \mathrm{L}^\infty(\Omega)$ and $A, b \in \mathrm{C}^{0,1}(\overline{V})$[1]. In addition, let $u \in \mathrm{H}^1_0(\Omega)$ (or $u \in \mathrm{H}^1(\Omega)$ for the Neumann problem) be a weak solution to the Dirichlet (or Neumann) problem $Lu = f$ in Ω for $f \in \mathrm{H}^{-1}(\Omega)$ ($f \in (\mathrm{H}^1(\Omega))'$). Let in addition $f \in \mathrm{L}^2(V)$. Then for any $W \Subset V$, it holds $u \in \mathrm{H}^2(W)$ and there exists a constant $C \geq 0$ independent of u and f such that*

$$\|u\|_{\mathrm{H}^2(W)} \leq C \left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{L}^2(V)} \right).$$

Proof. We will only prove the Dirichlet case and split the proof into five parts.

Step 1: Preparations. Let $a\colon \mathrm{H}^1_0(\Omega) \times \mathrm{H}^1_0(\Omega) \to \mathbb{C}$ be the associated form to L. Consider the principal part

$$a^*(u, v) = \int_\Omega \langle A\nabla u, \nabla v\rangle \,\mathrm{d}x$$

of a. Choose $\eta \in \mathcal{C}^\infty_c(V)$ with $0 \leq \eta \leq 1$ and $\eta = 1$ on $W \Subset V$. Let $j \in \{1, \dots, d\}$ and $0 < |h| < \frac{1}{2}\min\{\mathrm{dist}(\mathrm{supp}(\eta), \partial V), \mathrm{dist}(V, \partial\Omega)\}$. For $x \in V$ set $v(x) := -\left[\delta^j_{-h}(\eta^2 \delta^j_h u)\right](x)$. Then it holds:

- $\mathrm{supp}(v) \Subset V$.
- Applying the Leibniz rule leads to horrible computations showing that $v \in \mathrm{H}^1(V)$.

Together, this yields $v \in \mathrm{H}^1_0(V)$, so it can be extended by 0 to an $\mathrm{H}^1_0(\Omega)$ function, which we will again denote by v. Note that this is a feasible test function for the form a.

[1] i.e. the space of Lipschitz-continuous functions.

Step 2: Estimating the principal part. For u, v as above we have

$$\begin{aligned}
a^*(u,v) &= \int_\Omega \langle A\nabla u, \nabla v\rangle \,\mathrm{d}x \\
&= -\int_V \left\langle A\nabla u, \nabla\left(\delta^j_{-h}(\eta^2\delta^j_h u)\right)\right\rangle \,\mathrm{d}x \\
&= -\int_V \left\langle A\nabla u, \delta^j_{-h}\nabla(\eta^2\delta^j_h u)\right\rangle \,\mathrm{d}x \\
&\overset{\text{Lem. 12.2(b)}}{=} \int_V \left\langle \delta^j_h(A\nabla u), \eta^2\nabla(\delta^j_h u) + 2\eta\nabla\eta\delta^j_h u\right\rangle \,\mathrm{d}x \\
&\overset{\text{Lem. 12.2(a)}}{=} \int_V \left\langle A\nabla(\delta^j_h u), \eta^2\nabla(\delta^j_h u)\right\rangle \,\mathrm{d}x + \int_V \left\langle A\nabla(\delta^j_h u), 2\eta\nabla\eta\delta^j_h u\right\rangle \,\mathrm{d}x \\
&\quad + \int_V \left\langle (\delta^j_h A)\nabla u(\cdot + he_j), (\eta^2\nabla(\delta^j_h u) + 2\eta\nabla\eta\delta^j_h u)\right\rangle \,\mathrm{d}x \\
&= \int_V \left\langle A\eta\delta^j_h\nabla u, \eta\delta^j_h\nabla u\right\rangle \,\mathrm{d}x + \int_V \left\langle A\eta\delta^j_h\nabla u, 2\nabla\eta\delta^j_h u\right\rangle \,\mathrm{d}x \\
&\quad + \int_V \left\langle (\delta^j_h A)\nabla u(\cdot + he_j), (\eta^2\nabla(\delta^j_h u) + 2\eta\nabla\eta\delta^j_h u)\right\rangle \,\mathrm{d}x.
\end{aligned}$$

By coercivity and continuity of a, we estimate

$$\begin{aligned}
\operatorname{Re}(a^*(u,v)) &\geq \int_V \alpha_0 \left|\eta\delta^j_h\nabla u\right|^2 \,\mathrm{d}x - 2\left\|A\right\|_{\mathrm{L}^\infty(V)} \left\|\eta\delta^j_h\nabla u\right\|_{\mathrm{L}^2(V)} \left\|\nabla\eta\right\|_{\mathrm{L}^\infty(V)} \left\|\delta^j_h u\right\|_{\mathrm{L}^2(\operatorname{supp}(\eta))} \\
&\quad - \left\|A\right\|_{\mathrm{C}^{0,1}(V)} \left\|\nabla u\right\|_{\mathrm{L}^2(V)} \cdot \Big(\left\|\eta\right\|_{\mathrm{L}^\infty(V)} \left\|\eta\cdot\delta^j_h\nabla u\right\|_{\mathrm{L}^2(V)} \\
&\quad + 2\left\|\eta\right\|_{\mathrm{L}^\infty(V)} \left\|\nabla\eta\right\|_{\mathrm{L}^\infty(V)} \left\|\delta^j_h u\right\|_{\mathrm{L}^2(\operatorname{supp}(\eta))} \Big) \\
&\overset{\text{Lem. 12.3}}{\geq} \alpha_0 \left\|\eta\delta^j_h\nabla u\right\|^2_{\mathrm{L}^2(V)} - C\left\|\eta\cdot\delta^j_h\nabla u\right\|_{\mathrm{L}^2(V)} \left\|\nabla u\right\|_{\mathrm{L}^2(V)} - C\left\|\nabla u\right\|^2_{\mathrm{L}^2(V)} \\
&\overset{\text{Young}}{\geq} \alpha_0 \left\|\eta\delta^j_h\nabla u\right\|^2_{\mathrm{L}^2(V)} - \varepsilon\left\|\eta\delta^j_h\nabla u\right\|^2_{\mathrm{L}^2(V)} - C(\varepsilon)\left\|\nabla u\right\|^2_{\mathrm{L}^2(\Omega)} \\
&\geq \frac{\alpha_0}{2}\left\|\eta\delta^j_h\nabla u\right\|^2_{\mathrm{L}^2(V)} - C\left\|\nabla u\right\|^2_{\mathrm{L}^2(\Omega)}.
\end{aligned}$$

Step 3: Estimating the rest. We have

$$\langle f, v\rangle = a(u,v) = a^*(u,v) + \int_\Omega \langle bu, \nabla v\rangle \,\mathrm{d}x + \int_\Omega c\cdot\nabla u\cdot\bar{v}\,\mathrm{d}x + \int_\Omega eu\bar{v}\,\mathrm{d}x.$$

Hence,

$$a^*(u,v) = \int_V f\bar{v}\,\mathrm{d}x - \int_V \langle bu, \nabla v\rangle \,\mathrm{d}x - \int_V c\cdot\nabla u\cdot\bar{v}\,\mathrm{d}x - \int_V eu\bar{v}\,\mathrm{d}x.$$

Moreover, we obtain

$$\begin{aligned}\left|\int_V (f - c\cdot\nabla u - eu)\bar{v}\,\mathrm{d}x\right| &\le \Big(\|f\|_{\mathrm{L}^2(V)} + \|c\|_{\mathrm{L}^\infty(V)}\|\nabla u\|_{\mathrm{L}^2(V)} + \|e\|_{\mathrm{L}^\infty(V)}\|u\|_{\mathrm{L}^2(V)}\Big)\\ &\quad\cdot \left\|\delta^j_{-h}(\eta^2\delta^j_h u)\right\|_{\mathrm{L}^2(V)}\\ &\overset{\text{Lem. 12.3}}{\le} C\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right)\left\|\nabla(\eta^2\delta^j_h u)\right\|_{\mathrm{L}^2(V)}.\end{aligned}$$

Since

$$\begin{aligned}\left\|\nabla(\eta^2\delta^j_h u)\right\|_{\mathrm{L}^2(V)} &\le \left\|\eta^2\nabla(\delta^j_h u)\right\|_{\mathrm{L}^2(V)} + 2\left\|\eta\nabla\eta\delta^j_h u\right\|_{\mathrm{L}^2(\operatorname{supp}\eta)}\\ &\le \left\|\eta\nabla(\delta^j_h u)\right\|_{\mathrm{L}^2(V)} + 2\left\|\delta^j_h u\right\|_{\mathrm{L}^2(\operatorname{supp}\eta)}\|\nabla\eta\|_{\mathrm{L}^\infty(V)}\\ &\overset{\text{Lem. 12.3}}{\le} \left\|\eta\nabla(\delta^j_h u)\right\|_{\mathrm{L}^2(V)} + C\|\nabla u\|_{\mathrm{L}^2(V)}\\ &\le \left\|\eta\nabla(\delta^j_h u)\right\|_{\mathrm{L}^2(V)} + C\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right),\end{aligned}$$

we can once again apply Young's inequality to see

$$\begin{aligned}\left|\int_V (f - c\cdot\nabla u - eu)\bar{v}\,\mathrm{d}x\right| &\le C\Big(\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right)\left\|\eta\nabla(\delta^j_h u)\right\|_{\mathrm{L}^2(V)}\\ &\qquad + \left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right)^2\Big)\\ &\le \varepsilon\left\|\eta\nabla(\delta^j_h u)\right\|^2_{\mathrm{L}^2(V)} + C(\varepsilon)\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right)^2\\ &\le \frac{\alpha_0}{8}\left\|\eta\nabla(\delta^j_h u)\right\|^2_{\mathrm{L}^2(V)} + C\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right)^2.\end{aligned}$$

Analogously we can estimate

$$\left|\int_V \langle bu, \nabla u\rangle\,\mathrm{d}x\right| \le \frac{\alpha_0}{8}\left\|\eta\nabla(\delta^j_h u)\right\|^2_{\mathrm{L}^2(V)} + C\|u\|^2_{\mathrm{H}^1(V)}.$$

Step 4: $u \in \mathrm{H}^2(W)$**.** For all $j \in \{1,\dots,d\}$ and for all $h \neq 0$ small enough it holds

$$\begin{aligned}\frac{\alpha_0}{2}\left\|\delta^j_h(\nabla u)\right\|^2_{\mathrm{L}^2(W)} &= \frac{\alpha_0}{2}\left\|\eta\delta^j_h(\nabla u)\right\|^2_{\mathrm{L}^2(W)}\\ &\overset{\text{Part 2}}{\le} \operatorname{Re}(a^*(u,v)) + C\|\nabla u\|^2_{\mathrm{L}^2(V)}\\ &= \operatorname{Re}\left(\int_V (f - c\cdot\nabla u - eu)\bar{v}\,\mathrm{d}x - \int_V \langle bu, \nabla v\rangle\,\mathrm{d}x\right) + C\|\nabla u\|^2_{\mathrm{L}^2(V)}\\ &\overset{\text{Part 3}}{\le} \frac{\alpha_0}{4}\left\|\eta\nabla\left(\delta^j_h u\right)\right\|^2_{\mathrm{L}^2(V)} + C\left(\|f\|^2_{\mathrm{L}^2(V)} + \|u\|^2_{\mathrm{H}^1(V)}\right) + C\|u\|^2_{\mathrm{H}^1(V)}\\ &\le \frac{\alpha_0}{4}\left\|\eta\nabla\left(\delta^j_h u\right)\right\|^2_{\mathrm{L}^2(V)} + C\left(\|f\|^2_{\mathrm{L}^2(V)} + \|u\|^2_{\mathrm{H}^1(V)}\right)\end{aligned}$$

Hence, we obtain

$$\frac{\alpha_0}{4}\left\|\delta_h^j(\nabla u)\right\|_{\mathrm{L}^2(W)} \le C\left(\|f\|^2_{\mathrm{L}^2(V)} + \|u\|^2_{\mathrm{H}^1(V)}\right),$$

Thus, $u \in \mathrm{H}^2(W)$ by Theorem 12.5 with

$$\begin{aligned}\|u\|_{\mathrm{H}^2(W)} &\le \sqrt{\frac{4}{\alpha_0}C\left(\|f\|^2_{\mathrm{L}^2(V)} + \|u\|^2_{\mathrm{H}^1(V)}\right)}\\ &\le C\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{H}^1(V)}\right).\end{aligned}$$

Step 5: Replace $\|u\|_{\mathrm{H}^1(V)}$ by $\|u\|_{\mathrm{L}^2(V)}$. Let $W \Subset W_0 \Subset V$ and if necessary modify η to have $\eta|_{W_0} \equiv 1$. As in the previous part, we have

$$\|u\|_{\mathrm{H}^2(W)} \le C\left(\|f\|_{\mathrm{L}^2(W_0)} + \|u\|_{\mathrm{H}^1(W_0)}\right).$$

Then take $v = \eta^2 u \in \mathrm{H}^1_0(\Omega)$ as test function to obtain

$$a(u, \eta^2 u) = \langle f, \eta^2 u\rangle = \int_V f\eta^2\bar{u}\,\mathrm{d}x.$$

Similarly to step 2 it holds

$$\mathrm{Re}(a^*(u, \eta^2 u)) \ge \frac{\alpha_0}{2}\|\eta\nabla u\|^2_{\mathrm{L}^2(V)} - C\|u\|^2_{\mathrm{L}^2(V)}$$

and as before we can estimate

$$\frac{\alpha_0}{2}\|\eta\nabla u\|^2_{\mathrm{L}^2(V)} \le \frac{\alpha_0}{4}\|\eta\nabla u\|^2_{\mathrm{L}^2(V)} + C\left(\|f\|^2_{\mathrm{L}^2(V)} + \|u\|^2_{\mathrm{L}^2(V)}\right),$$

which yields

$$\begin{aligned}\|\nabla u\|_{\mathrm{L}^2(W_0)} = \|\eta\nabla u\|_{\mathrm{L}^2(W_0)} &\le \|\eta\nabla u\|_{\mathrm{L}^2(V)}\\ &\le \sqrt{C\left(\|f\|^2_{\mathrm{L}^2(V)} + \|u\|^2_{\mathrm{L}^2(V)}\right)}\\ &\le C\left(\|f\|_{\mathrm{L}^2(V)} + \|u\|_{\mathrm{L}^2(V)}\right).\end{aligned}$$

□

Theorem 13.2 (Global regularity)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with $\partial\Omega \in \mathrm{C}^{1,1}$*[2]*, A, $b \in \mathrm{C}^{0,1}(\overline{\Omega})$, $c, e \in \mathrm{L}^\infty(\Omega)$ with right-hand side $f \in \mathrm{L}^2(\Omega)$.*

If $u \in \mathrm{H}^1_0(\Omega)$ is a weak solution to $Lu = f$, i.e. $a(u,v) = \int_\Omega f\overline{v}$ for all $v \in \mathrm{H}^1_0(\Omega)$, then $u \in \mathrm{H}^2(\Omega)$ and

$$\|u\|_{\mathrm{H}^2(\Omega)} \le C\cdot\left(\|f\|_{\mathrm{L}^2(\Omega)} + \|u\|_{\mathrm{L}^2(\Omega)}\right).$$

[2] i.e. the space of continuously differentiable functions where the derivative is Lipschitz continuous.

If furthermore the solution is unique and L is continuously invertible then we even get

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

An analogous result holds for weak solutions $u \in \mathrm{H}^1(\Omega)$, i.e. $a(u,v) = \int_\Omega f\bar{v}$ for all $v \in \mathrm{H}^1(\Omega)$.

Corollary 13.3. *If in addition to the assumptions in Theorem 13.2 we assume $A, b \in \mathrm{C}^{k,1}(\overline{\Omega})$, $c, e \in \mathrm{C}^{k-1,1}(\overline{\Omega})$, $f \in \mathrm{H}^k(\Omega)$ and $\partial\Omega \in \mathrm{C}^{k+1,1}$ for some $k \in \mathbb{N}$, then $u \in \mathrm{H}^{k+2}(\Omega)$ and*

$$\|u\|_{\mathrm{H}^{k+2}(\Omega)} \leq C \left(\|f\|_{\mathrm{H}^k(\Omega)} + \|u\|_{\mathrm{L}^2(\Omega)} \right)$$

(or $\|u\|_{\mathrm{H}^{k+2}(\Omega)} \leq C \|f\|_{\mathrm{H}^k(\Omega)}$).

Remark 13.4. a) If the assumptions of Corollary 13.3 hold for all $k \in \mathbb{N}$, then

$$u \in \bigcap_{k \in \mathbb{N}} \mathrm{H}^k(\Omega).$$

However, what does it mean for a function to have arbitrarily many weak derivatives? We will see this in the next chapter.

b) If $u \in \mathrm{H}^2(\Omega)$ is a solution to $Lu = f$, then for all $v \in \mathcal{C}_c^\infty(\Omega)$, integrating by parts gives us

$$-\int_\Omega \operatorname{div}(A\nabla u)\bar{v}\,\mathrm{d}x - \int_\Omega \operatorname{div}(bu)\bar{v}\,\mathrm{d}x + \int_\Omega c \cdot \nabla u \bar{v}\,\mathrm{d}x + \int_\Omega eu\bar{v}\,\mathrm{d}x = \int_\Omega f\bar{v}\,\mathrm{d}x.$$

By the fundamental theorem of calculus of variation we obtain

$$Lu = -\operatorname{div}(A\nabla u) - \operatorname{div}(bu) + c \cdot \nabla u + eu = f$$

in $\mathrm{L}^2(\Omega)$. Hence, we can really apply L to the solution u and u solves the equation. This is called a strong solution.

14 Sobolev embeddings

Definition 14.1. Let $\Omega \subseteq \mathbb{R}^d$ be open and $\alpha \in (0,1]$. Then

a) $f : \Omega \to \mathbb{R}$ is called Hölder-continuous with exponent α, if the seminorm

$$[f]_{\alpha,\Omega} \coloneqq \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|^\alpha}$$

is finite.

b) we define

- the Hölder space

$$\mathrm{C}^{0,\alpha}(\Omega) \coloneqq \{f\colon \Omega \to \mathbb{R}\colon [f]_{\alpha,K} < \infty \text{ for all compact sets } K \subseteq \Omega\}.$$

- the Hölder space

$$\mathrm{C}^{0,\alpha}(\overline{\Omega}) \coloneqq \{f\colon \overline{\Omega} \to \mathbb{R}\colon f \in \mathrm{C}(\overline{\Omega}) \text{ and } [f]_{\alpha,\Omega} = [f]_{\alpha,\overline{\Omega}} < \infty\}.$$

with the norm

$$\|f\|_{\mathrm{C}^{0,\alpha}(\overline{\Omega})} \coloneqq \|f\|_\infty + [f]_{\alpha,\Omega}.$$

- For $k \in \mathbb{N}$ we define

$$\mathrm{C}^{k,\alpha}(\overline{\Omega}) \coloneqq \{f \in \mathrm{C}^k(\overline{\Omega})\colon \forall\, |\beta| = k\colon\ \mathrm{D}^\beta f \in \mathrm{C}^{0,\alpha}(\overline{\Omega})\}$$

together with the norm

$$\begin{aligned}\|f\|_{\mathrm{C}^{k,\alpha}(\overline{\Omega})} &\coloneqq \|f\|_{\mathrm{C}^k(\overline{\Omega})} + \sum_{|\beta|=k} [\mathrm{D}^\beta f]_{\alpha,\Omega} \\ &= \sum_{|\beta|\leq k} \left\|\mathrm{D}^\beta f\right\|_\infty + \sum_{|\beta|=k} [\mathrm{D}^\beta f]_{\alpha,\Omega}.\end{aligned}$$

Remark 14.2. a) The space $\mathrm{C}^{0,1}(\overline{\Omega})$ is the space of all Lipschitz-continuous functions on Ω.

b) One can show that for all $k \in \mathbb{N}$ and for all $\alpha \in (0,1]$, the space $\mathrm{C}^{k,\alpha}(\Omega)$ is a Banach space with respect to the $\mathrm{C}^{k,\alpha}$-norm.

Our goal in this section will be to somehow extract information about classical derivatives from knowledge on weak derivatives. In particular we are looking for embeddings

$$\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{C}^{m,\alpha}(\overline{\Omega})$$

for suitable choices of k, p, m and α. This means estimating the $\mathrm{C}^{m,\alpha}$-norm of a function u against its $\mathrm{W}^{k,p}(\Omega)$-norm to get a continuous embedding. Such an embedding allows us to trade weak regularity for classical regularity.

Additionally, we are looking for embeddings of the kind

$$\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{L}^{q}(\Omega)$$

for suitable k, p and q, i.e. $\|u\|_{\mathrm{L}^q(\Omega)} \leq C \cdot \|u\|_{\mathrm{W}^{k,p}(\Omega)}$. This will give us the possibility to trade weak regularity in exchange for higher integrability.

Definition 14.3. For $p \in [1, d)$ we define the Sobolev conjugated index to p by

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

Example. For $p = 2$ we have the following pairs of Sobolev conjugated indices:

d	2	3	4	5
2^*	∞	6	4	$\frac{10}{3}$

We notice that the indices converge to p as the dimension approaches infinity and that we always have $p^* > p$. We calculate $\frac{1}{p^*} = \frac{d-p}{dp} = \frac{1}{p} - \frac{1}{d}$. Similar to the Hölder condition $\frac{1}{p} + \frac{1}{p'} = 1$ we therefore have $\frac{1}{p} - \frac{1}{p^*} = \frac{1}{d}$.

Theorem 14.4 (Gagliardo-Nirenberg-Sobolev inequality)**.** *Let $p \in [1, d)$. Then there exists $C \geq 0$ such that*

$$\|u\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \leq C \cdot \|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} \qquad (*)$$

for all $u \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R}^d)$. Moreover, $\mathrm{W}^{1,p}(\mathbb{R}^d) \hookrightarrow \mathrm{L}^{p^*}(\mathbb{R}^d)$ and $(*)$ holds for all $u \in \mathrm{W}^{1,p}(\mathbb{R}^d)$.

Proof. **Step 1: $p = 1$.** In this case, we have $p^* = \frac{d}{d-1}$.

Let $u \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R}^d)$. Since $\operatorname{supp}(u)$ is compact, for $x \in \mathbb{R}^d$ and $j \in \{1, \dots, d\}$ we obtain

$$u(x) = \int_{-\infty}^{x_j} \partial_j u(x_1, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d) \, \mathrm{d}y_j.$$

Thus,

$$|u(x)| \leq \int_{-\infty}^{x_j} |\partial_j u(x_1, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| \, \mathrm{d}y_j$$
$$\leq \int_{-\infty}^{\infty} |\nabla u(x_1, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| \, \mathrm{d}y_j.$$

Hence,

$$|u(x)|^{\frac{d}{d-1}} = \left(|u(x)|^{\frac{1}{d-1}}\right)^d$$
$$\leq \prod_{j=1}^{d} \left(\int_{-\infty}^{\infty} |\nabla u(x_1, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| \, \mathrm{d}y_j\right)^{\frac{1}{d-1}}.$$

Integrating this expression leads to

$$\int_{-\infty}^{\infty} |u(x)|^{\frac{d}{d-1}} \, \mathrm{d}x_1 \leq \int_{-\infty}^{\infty} \prod_{j=1}^{d} \left(\int_{-\infty}^{\infty} |\nabla u(x_1, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| \, \mathrm{d}y_j\right)^{\frac{1}{d-1}} \mathrm{d}x_1.$$

If we look closely we notice that the first factor in the product does not depend on the integration variable x_1. Hence it can be pulled in front of the integral:

$$= \left(\int_{-\infty}^{\infty} |\nabla u(y_1, x_2, \dots, x_d)| \, \mathrm{d}y_1\right)^{\frac{1}{d-1}}$$
$$\cdot \int_{-\infty}^{\infty} \prod_{j=2}^{d} \left(\int_{-\infty}^{\infty} |\nabla u(x_1, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_d)| \, \mathrm{d}y_j\right)^{\frac{1}{d-1}} \mathrm{d}x_1.$$

This expression can be estimated by using the following generalised version of Hölder's inequality for more than two factors:

$$\int |f_1 f_2 \cdot \ldots \cdot f_m| \leq \prod_{k=1}^{m} \|f_k\|_{\mathrm{L}^{p_k}}, \quad \text{when } \sum_{k=1}^{m} \frac{1}{p_k} = 1.$$

Here, we choose $p_k = \frac{1}{d-1}$ and $m = d - 1$ and obtain

$$\int_{-\infty}^{\infty} |u(x)|^{\frac{d}{d-1}} \, \mathrm{d}x_1 \leq \left(\int_{-\infty}^{\infty} |\nabla u(y_1, x_2, \dots, x_d)| \, \mathrm{d}y_1\right)^{\frac{1}{d-1}}$$
$$\prod_{j=2}^{d} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u(x_1, \dots, y_j, \dots, x_d)| \, \mathrm{d}y_j \, \mathrm{d}x_1\right)^{\frac{1}{d-1}}.$$

Now let us integrate with respect to the second variable x_2 to get in the same way

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u(x)|^{\frac{d}{d-1}}\,\mathrm{d}x_1\,\mathrm{d}x_2 &\le \int_{-\infty}^{\infty}\Bigg[\left(\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}y_1\right)^{\frac{1}{d-1}}\\ &\qquad \prod_{j=2}^{d}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}x_1\,\mathrm{d}y_j\right)^{\frac{1}{d-1}}\Bigg]\mathrm{d}x_2\\ &= \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}x_1\,\mathrm{d}y_2\right)^{\frac{1}{d-1}}\int_{-\infty}^{\infty}\Bigg[\left(\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}y_1\right)^{\frac{1}{d-1}}\\ &\qquad \prod_{j=3}^{d}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}x_1\,\mathrm{d}y_j\right)^{\frac{1}{d-1}}\Bigg]\mathrm{d}x_2\\ &\le \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}x_1\,\mathrm{d}y_2\right)^{\frac{1}{d-1}}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}y_1\,\mathrm{d}x_2\right)^{\frac{1}{d-1}}\\ &\qquad \prod_{j=3}^{d}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}y_j\right)^{\frac{1}{d-1}}.\end{aligned}$$

Iterating yields

$$\begin{aligned}\|u\|_{\mathrm{L}^{p^*}}^{\frac{d}{d-1}} &= \int_{\mathbb{R}^d}|u(x)|^{\frac{d}{d-1}}\,\mathrm{d}x\\ &\le \prod_{j=1}^{d}\left(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|\nabla u|\,\mathrm{d}x_1\ldots\mathrm{d}x_{j-1}\,\mathrm{d}y_j\,\mathrm{d}x_{j+1}\ldots\mathrm{d}x_d\right)^{\frac{1}{d-1}}\\ &= \left(\int_{\mathbb{R}^d}|\nabla u(x)|\,\mathrm{d}x\right)^{\frac{d}{d-1}}.\end{aligned}$$

Step 2: $1 < p < d$**.** Let $u \in \mathcal{C}_{\mathrm{c}}^{\infty}(\mathbb{R}^d)$. For $\gamma > 1$ consider $v := |u|^{\gamma}$. Then $v \in \mathcal{C}_{\mathrm{c}}^{1}(\mathbb{R}^d)$ and by part 1 it holds

$$\begin{aligned}\left(\int_{\mathbb{R}^d}(|u|^{\gamma})^{\frac{d}{d-1}}\,\mathrm{d}x\right)^{\frac{d-1}{d}} &\le \int_{\mathbb{R}^d}|\nabla(|u|^{\gamma})|\,\mathrm{d}x\\ &= \int_{\mathbb{R}^d}\gamma\,|u|^{\gamma-1}\,|\nabla u|\,\mathrm{d}x\\ &\overset{\text{Hölder}}{\le} \gamma\left(\int_{\mathbb{R}^d}|u|^{(\gamma-1)\frac{p}{p-1}}\right)^{\frac{p-1}{p}}\left(\int_{\mathbb{R}^d}|\nabla u|^{p}\,\mathrm{d}x\right)^{\frac{1}{p}}.\end{aligned}$$

Our goal is to choose γ in such a way that $(\gamma-1)\frac{p}{p-1} = \gamma\cdot\frac{d}{d-1}$. To satisfy this equation we define $\gamma := \frac{p(d-1)}{d-p} > 1$. Then $\frac{\gamma d}{d-1} = \frac{pd}{d-p} = p^*$ and

$$(\gamma-1)\cdot\frac{p}{p-1} = \frac{pd-p-d+p}{d-p}\cdot\frac{p}{p-1} = \frac{pd-d}{d-p}\cdot\frac{p}{p-1} = \frac{dp}{d-p} = p^*.$$

Hence,

$$\|u\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)}^{\frac{p^*(d-1)}{d}} = \left(\int_{\mathbb{R}^d} (|u|^\gamma)^{\frac{d}{d-1}} \, \mathrm{d}x\right)^{\frac{d-1}{d}}$$
$$\leq \gamma \|u\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)}^{p^* \frac{p-1}{p}} \|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} .$$

Since $\dfrac{p^*(d-1)}{d} - p^* \dfrac{p-1}{p} = \gamma - (\gamma - 1) = 1$ this means

$$\|u\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \leq \gamma \|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} .$$

Step 3: Density argument. Let $u \in \mathrm{W}^{1,p}(\mathbb{R}^d)$ and $(u_n) \subseteq \mathcal{C}_\mathrm{c}^\infty(\mathbb{R}^d)$ such that $u_n \to u$ in $\mathrm{W}^{1,p}(\mathbb{R}^d)$ (which is $\mathrm{W}_0^{1,p}(\mathbb{R}^d)$ by Corollary 7.6). By switching to a subsequence, we can also assume that $u_n \to u$ almost everywhere in $\mathbb{R}^d$, so

$$\|u\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} \lim_{n\to\infty} |u_n(x)|^{p^*} \, \mathrm{d}x\right)^{\frac{1}{p^*}}$$
$$\overset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \|u_n\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \leq \liminf_{n\to\infty} \gamma \|\nabla u_n\|_{\mathrm{L}^p}$$
$$= \gamma \|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} .$$

In the last step note that because of what we have seen in Step 2, the expression actually converges, not just the limes inferior. □

We note, that our proof relied heavily on the fact that we assumed the full space $\mathbb{R}^d$ as domain. There is no hope of simply adjusting it a bit to rectify this limitation. Luckily, in combination with our results on continuous extension operators, it can still be used in quite general contexts, as we can extend $\mathrm{W}^{k,p}$-functions on bounded domains given a sufficiently smooth boundary and then apply the estimate to the extension.

Theorem 14.5 (Sobolev embedding for $p < d$)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain, $p \in [1, d)$ and $q \in [1, p^*]$. Then*

a) There exists $C \geq 0$ such that for all $u \in \mathrm{W}_0^{1,p}(\Omega)$ it holds $\|u\|_{\mathrm{L}^q(\Omega)} \leq C \|\nabla u\|_{\mathrm{L}^p(\Omega)}$. In particular $\mathrm{W}_0^{1,p}(\Omega) \hookrightarrow \mathrm{L}^q(\Omega)$ is a continuous embedding.

b) If $\partial\Omega \in \mathrm{C}^1$, then $\mathrm{W}^{1,p}(\Omega) \hookrightarrow \mathrm{L}^q(\Omega)$.[1]

[1] Actually, we only require a continuous extension operator, which is guaranteed by the condition on the boundary.

Proof. a) Let $u \in \mathrm{W}_0^{1,p}(\Omega)$ and $(u_n) \subseteq \mathcal{C}_c^\infty(\Omega)$ with $u_n \to u$ in $\mathrm{W}^{1,p}(\Omega)$ and $u_n \to u$ almost everywhere in Ω. We can extend u_n by 0 to $\tilde{u}_n \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. By Theorem 14.4, for all $n, m \in \mathbb{N}$ we obtain

$$\|\tilde{u}_n - \tilde{u}_m\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \leq C \|\nabla \tilde{u}_n - \nabla \tilde{u}_m\|_{\mathrm{L}^p(\mathbb{R}^d)},$$

so $(\tilde{u}_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathrm{L}^{p^*}(\mathbb{R}^d)$. Let $v \in \mathrm{L}^{p^*}(\mathbb{R}^d)$ be its limit.

By convergence almost everywhere we get $u = v \in \mathrm{L}^{p^*}(\Omega)$ and $u_n \to u$ in $\mathrm{L}^{p^*}(\Omega)$. Furthermore,

$$\begin{aligned}\|u\|_{\mathrm{L}^{p^*}(\Omega)} &= \lim_{n\to\infty} \|u_n\|_{\mathrm{L}^{p^*}(\Omega)} = \lim_{n\to\infty} \|\tilde{u}_n\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \\ &\leq C \cdot \lim_{n\to\infty} \|\nabla \tilde{u}_n\|_{\mathrm{L}^p(\mathbb{R}^d)} = C \cdot \lim_{n\to\infty} \|\nabla u_n\|_{\mathrm{L}^p(\Omega)} = C \cdot \|\nabla u\|_{\mathrm{L}^p(\Omega)},\end{aligned}$$

where we have used the Gagliardo-Nirenberg-Sobolev inequality to obtain the inequality.

The domain Ω is bounded, so for $q \in [1, p^*)$ we have an embedding $\mathrm{L}^{p^*}(\Omega) \hookrightarrow \mathrm{L}^q(\Omega)$ which implies

$$\|u\|_{\mathrm{L}^q(\Omega)} \leq C \|u\|_{\mathrm{L}^{p^*}(\Omega)} \leq C \|\nabla u\|_{\mathrm{L}^p(\Omega)}.$$

b) By assumption, the boundary $\partial\Omega$ is of class C^1, so we have a continuous extension operator $E\colon \mathrm{W}^{1,p}(\Omega) \to \mathrm{W}^{1,p}(\mathbb{R}^d)$. Let $u \in \mathrm{W}^{1,p}(\Omega)$. Then $Eu \in \mathrm{W}^{1,p}(\mathbb{R}^d)$ and by Theorem 14.4, $\mathrm{W}^{1,p}(\mathbb{R}^d) \hookrightarrow \mathrm{L}^{p^*}(\mathbb{R}^d)$. This gives

$$\begin{aligned}\|u\|_{\mathrm{L}^{p^*}(\Omega)} &\leq \|Eu\|_{\mathrm{L}^{p^*}(\mathbb{R}^d)} \\ &\overset{\text{Thm. 14.4}}{\leq} C \|\nabla(Eu)\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\leq C \|Eu\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} \\ &\leq C \|u\|_{\mathrm{W}^{1,p}(\Omega)},\end{aligned}$$

where the last step uses the continuity of E. For general $q \in [1, p^*]$, we can do the same proof as above. □

As $p \to d$, the Sobolev conjugate $p^* = \frac{dp}{d-p} \to \infty$, so we could hope that $\mathrm{W}^{1,d}(\Omega) \hookrightarrow \mathrm{L}^\infty(\Omega)$ is a continuous embedding. However, this fails as the following example shows:

Example 14.6. Let $d \geq 2$, $\Omega = \mathrm{B}_1(0)$ and consider

$$u(x) = \ln\left(\ln\left(\frac{\mathrm{e}}{|x|}\right)\right).$$

Then $u \in \mathrm{W}^{1,d}(\Omega)$, but $u \notin \mathrm{L}^\infty(\Omega)$.

Proposition 14.7 (Sobolev embedding for $p = d$)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain, $q \in [1, \infty)$. Then*

a) $\mathrm{W}_0^{1,d}(\Omega) \hookrightarrow \mathrm{L}^q(\Omega)$ *and there exists $C > 0$ such that for all $u \in \mathrm{W}_0^{1,d}(\Omega)$ we have*

$$\|u\|_{\mathrm{L}^q(\Omega)} \le C \|\nabla u\|_{\mathrm{L}^d(\Omega)}.$$

b) If $\partial\Omega \in \mathrm{C}^1$, then $\mathrm{W}^{1,d}(\Omega) \hookrightarrow \mathrm{L}^q(\Omega)$.

c) If $d = 1$ and $\Omega = (a, b) \subseteq \mathbb{R}$, then

$$\mathrm{W}^{1,d}((a, b)) \hookrightarrow \mathrm{C}([a, b]) \hookrightarrow \mathrm{L}^\infty((a, b)).$$

Proof. We only prove part a). The proof of b) can be done analogously and part c) remains as exercise.

Let $q \in [1, \infty)$ and choose $p \in (1, d)$ with $p^* > q$. Since $\mathrm{L}^d(\Omega) \hookrightarrow \mathrm{L}^p(\Omega)$, we have for all $u \in \mathrm{W}_0^{1,d}(\Omega)$:

$$\|u\|_{\mathrm{L}^q(\Omega)} \overset{\text{Thm. } 14.5}{\le} C \|\nabla u\|_{\mathrm{L}^p(\Omega)} \le C \|\nabla u\|_{\mathrm{L}^d(\Omega)}. \quad \square$$

As it turns out, the regularity improves, the bigger p becomes. One might hope that once p overtakes the space dimension, we are blessed by continuity. Indeed, this hope is justified, as the next theorem shows.

Theorem 14.8 (Morrey's inequality)**.** *Let $p > d$ and define $\alpha := 1 - \frac{d}{p}$ or $\alpha := 1$ if $p = \infty$. Then there is a constant $C \ge 0$ such that for all $u \in \mathrm{W}^{1,p}(\mathbb{R}^d) \cap \mathrm{C}^1(\mathbb{R}^d)$ we can estimate the Hölder norm*

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

Proof. **The case $p = \infty$:** Let $u \in \mathrm{C}^1(\mathbb{R}^d) \cap \mathrm{W}^{1,\infty}(\mathbb{R}^d)$. Then by the mean value theorem for all $x, y \in \mathbb{R}^d$ it holds

$$|u(x) - u(y)| \le \|\nabla u\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \cdot |x - y|,$$

which implies that the Hölder-part of the $\mathrm{C}^{0,1}$-norm can be estimated accordingly:

$$\begin{aligned}\|u\|_{\mathrm{C}^{0,1}(\mathbb{R}^d)} &= \|u\|_{\mathrm{L}^\infty(\mathbb{R}^d)} + \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{|x - y|} \\ &\le \|u\|_{\mathrm{L}^\infty(\mathbb{R}^d)} + \|\nabla u\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \\ &\le 2 \|u\|_{\mathrm{W}^{1,\infty}(\mathbb{R}^d)},\end{aligned}$$

where, in the last step, we made use of the fact, that the 1-norm and the ∞-norm on $\mathbb{R}^2$ are equivalent, which is just a more eloquent way of saying that we can estimate the sum of two positive numbers by two times their maximum.

An integral estimate: We want to show that there is a constant $C \geq 0$ such that for all $u \in \mathrm{C}^1(\mathbb{R}^d)$, for all $x \in \mathbb{R}^d$ and for all $r > 0$ it holds[2]

$$\frac{1}{|\mathrm{B}_r(x)|} \int_{\mathrm{B}_r(x)} |u(x) - u(y)| \, \mathrm{d}y \leq C \int_{\mathrm{B}_r(x)} \frac{|\nabla u(y)|}{|x-y|^{d-1}} \, \mathrm{d}y. \tag{$*$}$$

Let $u \in \mathrm{C}^1(\mathbb{R}^d), x \in \mathbb{R}^d$ and $r > 0$. For $w \in \partial B_1(0)$ and $s \in (0, r)$ we have

$$\begin{aligned} |u(x+sw) - u(x)| &= \left| \int_0^s \frac{\mathrm{d}}{\mathrm{d}t} u(x+tw) \, \mathrm{d}t \right| \\ &= \left| \int_0^s \nabla u(x+tw) \cdot w \, \mathrm{d}t \right| \\ &\overset{|w|=1}{\leq} \int_0^s |\nabla u(x+tw)| \, \mathrm{d}t. \end{aligned}$$

This implies

$$\begin{aligned} \int_{\partial \mathrm{B}_1(0)} |u(x+sw) - u(x)| \, \mathrm{d}S(w) &\leq \int_0^s t^{d-1} \int_{\partial \mathrm{B}_1(0)} \frac{|\nabla u(x+tw)|}{t^{d-1}} \, \mathrm{d}S(w) \, \mathrm{d}t \\ &= \int_0^s t^{d-1} \int_{\partial B_1(0)} \frac{|\nabla u(x+tw)|}{|x+tw-x|^{d-1}} \, \mathrm{d}S(w) \, \mathrm{d}t \\ &\overset{y=x+tw}{=} \int_0^s t^{d-1} \int_{\partial \mathrm{B}_t(x)} \frac{|\nabla u(y)|}{|y-x|^{d-1}} \frac{\mathrm{d}S(y)}{t^{d-1}} \, \mathrm{d}t \\ &= \int_0^s \int_{\partial \mathrm{B}_t(x)} \frac{|\nabla u(y)|}{|y-x|^{d-1}} \, \mathrm{d}S(y) \, \mathrm{d}t \\ &= \int_{\mathrm{B}_s(x)} \frac{|\nabla u(y)|}{|x-y|^{d-1}} \, \mathrm{d}y \\ &\leq \int_{\mathrm{B}_r(x)} \frac{|\nabla u(y)|}{|x-y|^{d-1}} \, \mathrm{d}y. \end{aligned}$$

[2]There is no immediately obvious reason for why this estimate will prove to be helpful. However, we want to estimate Hölder-preimages, which correspond to the integrand of the left-hand side. Additionally, because we are in the case of $p < \infty$, our norm involves integration as well.

This yields

$$\begin{aligned}\int_{\mathrm{B}_r(x)} |u(x)-u(y)|\,\mathrm{d}y &= \int_0^r s^{d-1}\int_{\partial B_1(0)} |u(x+sw)-u(x)|\,\mathrm{d}S(w)\,\mathrm{d}s\\ &\le \int_0^r s^{d-1}\int_{\mathrm{B}_r(x)} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\,\mathrm{d}y\,\mathrm{d}s\\ &= \frac{1}{d}\cdot r^d \int_{B_r(x)} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\,\mathrm{d}y.\end{aligned}$$

Dividing by the volume of the ball on the left hand side yields

$$\frac{1}{|\mathrm{B}_r(x)|}\int_{\mathrm{B}_r(x)} |u(x)-u(y)|\,\mathrm{d}y \le C\int_{\mathrm{B}_r(x)} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\,\mathrm{d}y.$$

Estimating the L^∞-norm: We want to show $\|u\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \le C\|u\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)}$. Since

$$\frac{d-1}{d} = 1-\frac{1}{d} < 1-\frac{1}{p} = \frac{p-1}{p}$$

we obtain $\dfrac{(d-1)p}{p-1} < d$, so for all $x\in\mathbb{R}^d$ it holds

$$\int_{\mathrm{B}_1(x)} |x-y|^{-\frac{(d-1)p}{p-1}}\,\mathrm{d}y = d\omega_d\int_0^1 r^{-\frac{(d-1)p}{p-1}} r^{d-1}\,\mathrm{d}r$$

where we made use of the fact, that the integrand is rotationally symmetric to rewrite the integral over the ball as an integral over the unit interval. Since $d-1-(d-1)\frac{p}{p-1} > d-1-d = -1$, this integral is finite, so

$$\int_{\mathrm{B}_1(x)} |x-y|^{-\frac{(d-1)p}{p-1}}\,\mathrm{d}y = C(d,p) < \infty.$$

Hence, for all $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}|u(x)| &= \frac{1}{|\mathrm{B}_1(x)|}\int_{\mathrm{B}_1(x)} |u(x)|\,\mathrm{d}y\\ &\le \frac{1}{|\mathrm{B}_1(x)|}\left(\int_{\mathrm{B}_1(x)} |u(x)-u(y)|\,\mathrm{d}y + \int_{\mathrm{B}_1(x)} |u(y)|\,\mathrm{d}y\right)\\ &\overset{(*)}{\le} C\int_{\mathrm{B}_1(x)} \frac{|\nabla u(y)|}{|y-x|^{d-1}}\,\mathrm{d}y + |\mathrm{B}_1(x)|^{\frac{1}{p'}-1}\|u\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\overset{\text{Hölder}}{\le} C\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)}\left(\int_{B_1(x)} |x-y|^{-(d-1)p'}\,\mathrm{d}y\right)^{\frac{1}{p'}} + |B_1(x)|^{-\frac{1}{p}}\|u\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\le C\cdot C(d,p)\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} + C\|u\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\le C\|u\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)}.\end{aligned}$$

Estimating the Hölder seminorm: We claim

$$[u]_{\alpha,\mathbb{R}^d} \coloneqq \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\alpha} \leq C\, \|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)}\,.$$

Let $x, y \in \mathbb{R}^d$ with $x \neq y$, $r \coloneqq |x-y| > 0$ and $W \coloneqq B_r(x) \cap B_r(y)$. Then we estimate

$$\begin{aligned}
|u(x)-u(y)| &= \frac{1}{|W|}\int_W |u(x)-u(y)|\,\mathrm{d}z \\
&\leq \frac{1}{\left|B_{\frac{r}{2}}\left(\frac{x+y}{2}\right)\right|}\left(\int_{B_r(x)} |u(x)-u(z)|\,\mathrm{d}z + \int_{B_r(y)} |u(z)-u(y)|\,\mathrm{d}z\right) \\
&\leq 2^d\Bigg(\frac{1}{|B_r(x)|}\int_{B_r(x)} |u(x)-u(z)|\,\mathrm{d}z \\
&\qquad + \frac{1}{|B_r(y)|}\int_{B_r(y)} |u(z)-u(y)|\,\mathrm{d}z\Bigg) \\
&\overset{(*)}{\leq} C\left[\int_{B_r(x)} \frac{|\nabla u(z)|}{|x-z|^{d-1}}\,\mathrm{d}z + \int \frac{|\nabla u(z)|}{|z-y|^{d-1}}\,\mathrm{d}z\right] \\
&\leq C\Bigg[\left(\int_{B_r(x)} |\nabla u(z)|^p\,\mathrm{d}z\right)^{\frac{1}{p}}\left(\int_{B_r(x)} |x-z|^{-(d-1)p'}\,\mathrm{d}z\right)^{\frac{1}{p'}} \\
&\qquad + \left(\int_{B_r(y)} |\nabla u(z)|^p\,\mathrm{d}z\right)^{\frac{1}{p}}\left(\int_{B_r(y)} |z-y|^{-(d-1)p'}\,\mathrm{d}z\right)^{\frac{1}{p'}}\Bigg] \\
&\leq C\,\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} \cdot C\left(\int_0^r s^{-(d-1)p'} s^{d-1}\,\mathrm{d}s\right)^{\frac{1}{p'}} \\
&\leq C\,\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)} \cdot r^{1-\frac{d}{p}}.
\end{aligned}$$

Dividing by $r^{1-\frac{d}{p}} = |x-y|^{1-\frac{d}{p}}$ on both sides yields

$$[u]_{\alpha,\mathbb{R}^d} = [u]_{1-\frac{d}{p},\mathbb{R}^d} = \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|u(x)-u(y)|}{|x-y|^{1-\frac{d}{p}}} \leq C\,\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^d)}\,. \qquad \square$$

Theorem 14.9 (Sobolev embeddings, $p > d$)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain, $p \in (d,\infty]$ and define $\alpha \coloneqq 1 - \dfrac{d}{p}$ or $\alpha \coloneqq 1$ for $p = \infty$. Then it holds*

a) $\mathrm{W}^{1,p}(\mathbb{R}^d) \hookrightarrow \mathrm{C}^{0,\alpha}(\mathbb{R}^d)$.

b) $\mathrm{W}^{1,p}_0(\Omega) \hookrightarrow \mathrm{C}^{0,\alpha}(\overline{\Omega})$ *and* $\|u\|_{\mathrm{C}^{0,\alpha}(\overline{\Omega})} \leq C\,\|\nabla u\|_{\mathrm{L}^p(\Omega)}$.

c) If $\partial\Omega \in \mathrm{C}^1$, then $\mathrm{W}^{1,p}(\Omega) \hookrightarrow \mathrm{C}^{0,\alpha}(\overline{\Omega})$.

Proof. a) Let $u \in \mathrm{W}^{1,p}(\mathbb{R}^d)$ and $(u_n)_{n\in\mathbb{N}} \subseteq \mathrm{C}^1(\mathbb{R}^d) \cap \mathrm{W}^{1,p}(\mathbb{R}^d)$ with $u_n \to u$ in $\mathrm{W}^{1,p}(\mathbb{R}^d)$ and almost everywhere on $\mathbb{R}^d$. We want to show that this sequence does not only converge in $\mathrm{W}^{1,p}(\mathbb{R}^d)$, but also with regard to the $\mathrm{C}^{0,\alpha}$-norm.

For all $m, n \in \mathbb{N}$ we have

$$\|u_n - u_m\|_{\mathrm{C}^{0,\alpha}(\mathbb{R}^d)} \overset{\text{Morrey}}{\leq} C\,\|u_n - u_m\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)}\,,$$

so (u_n) is a Cauchy sequence in $\mathrm{C}^{0,\alpha}(\mathbb{R}^d)$. Let $v \in \mathrm{C}^{0,\alpha}(\mathbb{R}^d)$ be its limit. Since $u_n \to u$ almost everywhere in $\mathbb{R}^d$, we get $v = u$ almost everywhere in $\mathbb{R}^d$ and

$$\|u\|_{\mathrm{C}^{0,\alpha}(\mathbb{R}^d)} = \lim_{n\to\infty} \|u_n\|_{\mathrm{C}^{0,\alpha}(\mathbb{R}^d)} \overset{\text{Morrey}}{\leq} C \lim_{n\to\infty} \|u_n\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} = C\,\|u\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)}\,.$$

b) We have a norm estimate on the whole space. To get the desired result for arbitrary bounded domains Ω, we once again make use of our knowledge about extension operators.

Let $u \in \mathrm{W}^{1,p}_0(\Omega)$ and $(u_n) \subseteq \mathcal{C}^\infty_\mathrm{c}(\Omega)$ with $u_n \to u$ in $\mathrm{W}^{1,p}(\Omega)$.

Extend u_n by 0 to $\tilde{u}_n \in \mathcal{C}^\infty_\mathrm{c}(\mathbb{R}^d)$. By the first part this yields $\tilde{u}_n \to v$ in $\mathrm{C}^{0,\alpha}(\mathbb{R}^d)$ and in $\mathrm{W}^{1,p}(\mathbb{R}^d)$. Then $u = v$ in Ω and

$$\|u\|_{\mathrm{C}^{0,\alpha}(\overline{\Omega})} = \|v\|_{\mathrm{C}^{0,\alpha}(\mathbb{R}^d)} \overset{\text{Morrey}}{\leq} C\,\|v\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} = C\,\|u\|_{\mathrm{W}^{1,p}(\Omega)} \overset{\text{Poincaré}}{\leq} C\,\|\nabla u\|_{\mathrm{L}^p(\Omega)}\,.$$

where we have made explicit use of the fact, that we can estimate the norm of u by the norm of its gradient, as they are equivalent by the Poincaré inequality.

c) We use the extension operator again to obtain

$$\|u\|_{\mathrm{C}^{0,\alpha}(\overline{\Omega})} \leq \|E(u)\|_{\mathrm{C}^{0,\alpha}(\mathbb{R}^d)} \overset{(a)}{\leq} C\,\|E(u)\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} \leq C\,\|u\|_{\mathrm{W}^{1,p}(\Omega)}\,. \qquad \square$$

As it turns out, the hard part was proving the desired properties in the case of the whole space. Then our toolbox enabled us to transfer the results to the case of bounded domains.

Remark. Often, it is easier to find solutions in the Hilbert space case, i.e. $p = 2$. For the theory of Sobolev embeddings, this is not the best case, since we already saw that we obtain better results for increasing p. Thus, we want to find methods to deduce solutions for $p > 2$ from the $p = 2$-solution. This will be an important topic later in this course.

Another thing to note is, that if p is larger than d only by a small amount, α is still very small, so we only get limited regularity in terms of continuity. The larger p, the more continuity we can expect.

Until now we have seen that there are strong links between weak differentiability and integrability. However, we are interested in the connection between weak differentiability and classical differentiability. So far we have only considered one weak derivative,

Theorem 14.10 (Embeddings for $\mathrm{W}^{k,p}(\Omega)$)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with $\partial\Omega \subseteq \mathrm{C}^1$. Furthermore, let $p, q \in [1, \infty], k \in \mathbb{N}$ and $l \in \mathbb{N}_0$ with $k > l$.*

a) If $k - \frac{d}{p} < l$ and $k - \frac{d}{p} \geq l - \frac{d}{q}$, then

$$\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{W}^{l,q}(\Omega).$$

b) If $k - \frac{d}{q} = l$ and $q \in [1, \infty)$, then

$$\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{W}^{l,q}(\Omega).$$

c) If $\frac{d}{p} \in \mathbb{N}$ and $l = k - \frac{d}{p} - 1$, then

$$\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{C}^{l,\alpha}(\bar{\Omega})$$

for all $\alpha \in (0, 1)$. If $\frac{d}{p} \notin \mathbb{N}$ and $l = k - \left\lfloor \frac{d}{p} \right\rfloor - 1$, then

$$\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{C}^{l,\alpha}(\bar{\Omega})$$

for $\alpha \in \left(0, 1 - \left(\frac{d}{p} - \left\lfloor \frac{d}{p} \right\rfloor\right)\right]$.

Proof. Let $m \in \mathbb{N}, r \in [1, d)$. Then $u \in \mathrm{W}^{m,r}(\Omega)$ implies $\mathrm{D}^\beta u \in \mathrm{W}^{1,r}(\Omega)$ for all $|\beta| \leq m-1$. Hence, Theorem 14.5 implies $\mathrm{D}^\beta u \in \mathrm{L}^s(\Omega)$ for all $s \in [1, r^*]$ and for all $|\beta| \leq m-1$. In addition, by the equivalence of norms, we get

$$\|u\|_{\mathrm{W}^{m-1,s}(\Omega)} \leq C \sum_{|\beta| \leq m-1} \left\|\mathrm{D}^\beta u\right\|_{\mathrm{L}^s(\Omega)}.$$

The continuity of the Sobolev embedding in Theorem 14.5 allows us to estimate the norms of the weak derivatives:

$$\|u\|_{\mathrm{W}^{m-1,s}(\Omega)} \leq C \sum_{|\beta| \leq m-1} \left\|\mathrm{D}^\beta u\right\|_{\mathrm{W}^{1,r}(\Omega} \leq C \|u\|_{\mathrm{W}^{m,r}(\Omega)}.$$

Iterating this and using the various Sobolev Embeddings from Theorems 14.5, 14.9 and 14.10 yields the claims. □

Remark 14.11. We want to take a look at several useful special cases:

$d = 1$:

$$H^1(\Omega) \hookrightarrow C^{0,\frac{1}{2}}(\overline{\Omega})$$
$$H^2(\Omega) \hookrightarrow C^{1,\frac{1}{2}}(\overline{\Omega}).$$

$d = 2$: $H^1(\Omega) \hookrightarrow L^q(\Omega)$ for all $q \in [1, \infty)$ and

$$W^{1,p}(\Omega) \hookrightarrow C^{0,\alpha}(\overline{\Omega})$$

for $p > 2$ and $\alpha = 1 - \frac{2}{p}$. In addition we have

$$H^2(\Omega) \hookrightarrow C^{0,\beta}(\overline{\Omega})$$

for all $\beta \in (0, 1)$ and

$$H^2(\Omega) \hookrightarrow W^{1,q}(\Omega)$$

for all $q \in [1, \infty)$.

$d = 3$:

$$H^1(\Omega) \hookrightarrow L^q(\Omega)$$

for all $q \in [1, 6]$ and

$$H^2(\Omega) \hookrightarrow C^{0,\frac{1}{2}}(\overline{\Omega}).$$

The next theorem completes what we already know about compactness of Sobolev embeddings. The philosophy is that whenever you have a Sobolev embedding and you are not in the limiting case, the embedding is compact if you give up another ε in smoothness. We assume the boundary of Ω to be of class C^1, which actually is not the most general case.

Theorem 14.12 (Rellich-Kondrachov)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be bounded with $\partial\Omega \in C^1$. Let $p, q \in [1, \infty], k \in \mathbb{N}, l \in \mathbb{N}_0, k > l$.*

a) If $l - \frac{d}{q} < k - \frac{d}{p} < l$, then

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow W^{l,q}(\Omega).$$

b) If $k - \frac{d}{p} > l + \alpha$ for some $\alpha \in (0, 1)$, then

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow C^{l,\alpha}(\overline{\Omega}).[3]$$

[3]The reason behind this is, that for compact Ω, the Hölder spaces are nested compactly, i.e. $C^{0,\alpha} \hookrightarrow\hookrightarrow C^{0,\beta}$, if $\beta < \alpha$.

Remark 14.13. a) Let $u \in \mathrm{H}^2_{\mathrm{loc}}(\Omega)$ be a solution to $Lu = f$ in Ω with Dirichlet or Neumann boundary condition. If $d = 2$ then $u \in \mathrm{C}^{0,\alpha}(\Omega)$ for all $\alpha \in (0,1)$ and if $u \in \mathrm{H}^2(\Omega)$, then $u \in \mathrm{C}^{0,\alpha}(\overline{\Omega})$.

In the case of $d = 3$, $u \in \mathrm{C}^{0,\frac{1}{2}}$, and if $u \in H^2(\Omega)$, then $u \in \mathrm{C}^{0,\frac{1}{2}}(\overline{\Omega})$.

b) If the coefficients of L are all C^∞ functions, $\partial\Omega$ is nice and $f \in \mathrm{C}^\infty$, then $u \in \mathrm{C}^\infty(\overline{\Omega})$.

An example for such a problem is $-\Delta u = \lambda u$ with $u = 0$ on $\partial\Omega$. If $\partial\Omega$ is smooth enough, then $u \in \mathrm{C}^\infty(\overline{\Omega})$.

15 Tempered distributions

In this chapter, we once again weaken the concept of differentiability. Therefore, we introduce a more general concept of functions, so-called distributions, which are linear functionals on a certain space of test functions. We will look at the precise definition throughout this chapter. In our case, this space will be the Schwartz space $\mathcal{S}(\mathbb{R}^d)$, so as a reminder we define this space first.

Reminder 15.1. a) A function $\varphi \in \mathrm{C}^\infty(\mathbb{R}^d)$ is called rapidly decreasing, if for all multi-indices $\alpha, \beta \in \mathbb{N}_0^d$, the Schwartz seminorms

$$p_{\alpha,\beta}(\varphi) \coloneqq \sup_{x\in\mathbb{R}^d} \left|x^\alpha D^\beta \varphi(x)\right|$$

are finite.

b) We define the Schwartz space

$$\mathcal{S}(\mathbb{R}^d) \coloneqq \{\varphi \in \mathrm{C}^\infty(\mathbb{R}^d) \colon \varphi \text{ rapidly decreasing}\}.$$

Lemma 15.2. *Let $\varphi \in \mathrm{C}^\infty(\mathbb{R}^d)$. Then $\varphi \in \mathcal{S}(\mathbb{R}^d)$ if and only if for all $j \in \mathbb{N}_0$ it holds*

$$d_j(\varphi) \coloneqq \sup_{|\alpha|=j} \left\|(1+|\cdot|^2)^j D^\alpha \varphi\right\|_\infty < \infty.$$

Proof. **"$\Longrightarrow$":** For all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = j$ it holds

$$\begin{aligned}\left\|(1+|\cdot|^2)^j D^\alpha \varphi\right\|_\infty &= \sup_{x\in\mathbb{R}^d} (1+|x|^2)^j \left|D^\alpha \varphi(x)\right| \\ &\le |\text{Polynomial in } x| \cdot |D^\alpha \varphi(x)| < \infty.\end{aligned}$$

"$\Longleftarrow$": We have

$$p_{\alpha,\beta}(\varphi) = \sup_{x\in\mathbb{R}^d} \left|x^\alpha D^\beta \varphi(x)\right| \le \sup_{x\in\mathbb{R}^d} |x|^{|\alpha|} \cdot \left|D^\beta \varphi(x)\right|.$$

If we choose j large enough, we can achieve

$$\le \sup_{x\in\mathbb{R}^d} (1+|x|^2)^j \left|D^\beta \varphi(x)\right|.$$

□

Definition 15.3. Let $(\varphi_n) \subseteq \mathcal{S}(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then (φ_n) converges to φ in $\mathcal{S}(\mathbb{R}^d)$, if $p_{\alpha,\beta}(\varphi_n - \varphi) \to 0$ for all $\alpha, \beta \in \mathbb{N}_0^d$ as $n \to \infty$.

Remark 15.4. a) Lemma 15.2 implies that convergence in $\mathcal{S}(\mathbb{R}^d)$ is equivalent to the condition $d_j(\varphi_n - \varphi) \to 0$ for all $j \in \mathbb{N}_0$ as $n \to \infty$.

b) The family $\{d_j \colon j \in \mathbb{N}_0\}$ generates a topology corresponding to this convergence which is metrisable with metric

$$d(\varphi, \psi) := \sum_{j=0}^{\infty} \frac{2^{-j} d_j(\varphi - \psi)}{1 + d_j(\varphi - \psi)}$$

for $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$. Additionally, seen in this way the space $(\mathcal{S}(\mathbb{R}^d), d)$ is a complete metric space. It ia s so-called Fréchet space.

c) If $\varphi_n \to \varphi$ in $\mathcal{S}(\mathbb{R}^d)$, then it holds

- $D^\alpha \varphi_n \to D^\alpha \varphi$ uniformly in $\mathbb{R}^d$ for all $\alpha \in \mathbb{N}_0^d$.
- $D^\alpha \varphi_n \to D^\alpha \varphi$ in $\mathrm{L}^p(\mathbb{R}^d)$ for all $\alpha \in \mathbb{N}_0^d$ and for all $1 \le p \le \infty$.

Example 15.5. Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and $(h_n) \subseteq \mathbb{R}^d$ such that $h_n \to 0$. Define a sequence $\varphi_n(x) := \varphi(x + h_n)$. Then for all $j \in \mathbb{N}$, for all $x \in \mathbb{R}^d$ and for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = j$ it holds

$$\begin{aligned}
(1 + |x|^2)^j \, |D^\alpha (\varphi(x + h_n) - \varphi(x))| &= (1 + |x|^2)^j \, |D^\alpha \varphi(x + h_n) - D^\alpha \varphi(x)| \\
&\overset{\text{mean value thm.}}{\le} (1 + |x|^2)^j \, |\nabla D^\alpha \varphi(\xi)| \, |h_n| \\
&\le C \, |h_n| \to 0, \qquad (n \to \infty).
\end{aligned}$$

Hence, $\varphi_n \to \varphi$ in $\mathcal{S}(\mathbb{R}^d)$.

Before we define tempered distributions, we want to give a brief motivation for the definition. The idea is that every $f \in \mathrm{L}^1(\mathbb{R}^d)$ can be considered to be a linear form on $\mathrm{BUC}(\mathbb{R}^d)$[1] via

$$T_f := \begin{cases} \mathrm{BUC}(\mathbb{R}^d) & \to \mathbb{K} \\ \qquad \varphi & \mapsto \int_{\mathbb{R}^d} f \varphi. \end{cases}$$

Note that T_f is continuous since $|T_f(\varphi)| = \left|\int_{\mathbb{R}^d} f\varphi\right| \le \|f\|_{\mathrm{L}^1(\mathbb{R}^d)} \|\varphi\|_\infty$. This means that $\mathrm{L}^1(\mathbb{R}^d) \hookrightarrow (\mathrm{BUC}(\mathbb{R}^d))'$.[2] As mentioned earlier, distributions are elements of the dual space $\mathcal{D}'$ for $\mathcal{D}$ "sufficiently small" and compatible with $\varphi \mapsto \int_{\mathbb{R}^d} f\varphi$.

[1] This is the space of bounded and uniformly continuous functions on $\mathbb{R}^d$.
[2] This space also contains elements which are no functions.

Definition 15.6. a) Consider

$$\mathcal{S}'(\mathbb{R}^d) := \{T\colon \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}\colon T \text{ linear and continuous}\}$$

Then $T \in \mathcal{S}'(\mathbb{R}^d)$ is called tempered distribution.

b) Let $(T_n) \subseteq \mathcal{S}'(\mathbb{R}^d)$ be a sequence of tempered distributions and $T \in \mathcal{S}'(\mathbb{R}^d)$. Then T_n converges to T in $\mathcal{S}'(\mathbb{R}^d)$ if

$$\langle T_n, \varphi\rangle := T_n(\varphi) \to T(\varphi) =: \langle T, \varphi\rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

Lemma 15.7. *Let $T\colon \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ be linear. Then the following statements are equivalent:*

a) T is a tempered distribution.

b) There exist $C \geq 0$ and $k \in \mathbb{N}_0$ such that $|\langle T, \varphi\rangle| \leq C \sum_{j=0}^{k} d_j(\varphi)$.

c) There exist $C \geq 0$ and $k, m \in \mathbb{N}_0$ such that $|\langle T, \varphi\rangle| \leq C \sum_{\substack{|\alpha|\leq k \\ |\beta|\leq m}} p_{\alpha,\beta}(\varphi)$.

Proof. Exercise. □

Example 15.8. a) Let $f\colon \mathbb{R}^d \to \mathbb{C}$ be measurable and such that $(1 + |\cdot|^2)^{-j} f \in \mathrm{L}^1(\mathbb{R}^d)$ for some $j \in \mathbb{N}_0$ (we call such an f slowly increasing). Then $T_f(\varphi) := \int_{\mathbb{R}^d} f\varphi$, $\varphi \in \mathcal{S}(\mathbb{R}^d)$, is a tempered distribution, since

$$|\langle T, \varphi\rangle| \leq \int_{\mathbb{R}^d} \underbrace{\frac{|f|}{(1 + |\cdot|^2)^j}}_{\in \mathrm{L}^1(\mathbb{R}^d)} \underbrace{(1 + |\cdot|^2)^j\, |\varphi|}_{\in \mathrm{L}^\infty(\mathbb{R}^d)} \leq C d_j(\varphi).$$

b) In particular, part a) implies $\mathrm{L}^p(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$ for all $1 \leq p \leq \infty$. However, note that we identify $f \in \mathrm{L}^p(\mathbb{R}^d)$ with the corresponding distribution $T_f \in \mathcal{S}'(\mathbb{R}^d)$. Note also that $f \mapsto T_f$ is injective.

c) For all $x \in \mathbb{R}^d$ we have that

$$\delta_x\colon \begin{cases} \mathcal{S}(\mathbb{R}^d) & \to \mathbb{C} \\ \quad\varphi & \mapsto \varphi(x) \end{cases}$$

is in $\mathcal{S}'(\mathbb{R}^d)$, since $|\langle \delta_x, \varphi\rangle| = |\varphi(x)| \leq \|\varphi\|_\infty = d_0(\varphi)$. For the special case of $x = 0$, the distribution δ_0 is called (Dirac) delta distribution.

d) Define

$$\left\langle \mathrm{p.\,v.}\,\frac{1}{x}, \varphi \right\rangle := \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \varphi(x)\cdot \frac{1}{x}\,\mathrm{d}x$$

for $\varphi \in \mathcal{S}(\mathbb{R})$. This is the so-called principle value distribution. One can show that it is a tempered distribution.

Definition 15.9. a) Let $T \in \mathcal{S}'(\mathbb{R}^d), \psi \in \mathcal{S}(\mathbb{R}^d)$ and p be a polynomial. Then we define

$$\langle \psi \cdot T, \varphi \rangle := \langle T, \psi \cdot \varphi \rangle$$
$$\langle p \cdot T, \varphi \rangle := \langle T, p \cdot \varphi \rangle$$

for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. It is easy to see that $\psi \cdot T$ and $p \cdot T$ are well-defined and tempered distributions.

b) Let $T \in \mathcal{S}'(\mathbb{R}^d)$ and $j \in \{1, \dots, d\}$. Then the j-th partial distributional derivative of T is given by

$$\langle \partial_j T, \varphi \rangle := -\langle T, \partial_j \varphi \rangle$$

for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. This is again a tempered distribution.

Remark 15.10. a) We can easily extend the above definition to general derivatives. For $\alpha \in \mathbb{N}_0^d$ we have

$$\langle D^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle .$$

b) The definition of distributional derivatives is motivated by integration by parts. Keep in mind that distributions shall be a generalisation of functions. Consider $f \in \mathcal{S}(\mathbb{R}^d)$. Then $f \in \mathrm{L}^p(\mathbb{R}^d)$ for $1 \le p \le \infty$, so we can define the tempered distribution T_f as in Example 15.8. Now we can use integration by parts to see

$$\partial_j T_f(\varphi) = \int_{\mathbb{R}^d} \partial_j f(x)\varphi(x)\,\mathrm{d}x = -\int_{\mathbb{R}^d} f(x)\partial_j\varphi(x)\,\mathrm{d}x$$

for $\varphi \in \mathcal{S}$. Hence, if we want to define a distributional derivative, it should match the above calculation.

c) If we have $f \in \mathrm{W}^{1,p}(\mathbb{R}^d)$, then obviously $T_f \in \mathcal{S}'(\mathbb{R}^d)$ and $\partial_j T_f = T_{\partial_j f}$, so the distributional derivative coincides with the weak derivative.

Example 15.11. Earlier in this course when we introduced weak derivatives, we have already seen as an example that $f(x) = |x|, x \in \mathbb{R}$, is weakly differentiable with weak derivative $f'(x) = \mathrm{sign}(x)$. However, we also mentioned that f is not twice weakly differentiable. Since $\mathrm{sign}(x)$ is slowly increasing, we have $T_{f'} = T_{\mathrm{sign}} \in \mathcal{S}'(\mathbb{R}^d)$. Thus, we can compute the distributional derivative for $T_{f'}$ as

$$\begin{aligned}\langle DT_{f'}, \varphi \rangle = -\langle \mathrm{sign}, \varphi' \rangle &= \int_{-\infty}^{0} \varphi'(x)\,\mathrm{d}x - \int_0^\infty \varphi'(x)\,\mathrm{d}x \\ &= \varphi(0) + \varphi(0) = 2\varphi(0) = \langle 2\delta_0, \varphi \rangle .\end{aligned}$$

Roughly speaking, this implies "$f'' = 2\delta_0$".

Definition 15.12. a) Let $g \in \mathrm{C}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. Then

$$\tilde{\tau}_x g(y) := g(x-y), \qquad y \in \mathbb{R}^d$$

is called the translation of g by x.

b) For $T \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$ we define the convolution by

$$(T * \varphi)(x) := \langle T, \tilde{\tau}_x \varphi \rangle \qquad x \in \mathbb{R}^d.$$

Remark. a) Note that the convolution of a tempered distribution and a function is again a *function.*

b) If $T \in \mathcal{S}'(\mathbb{R}^d)$ is of the form T_f for some f such that this expression is meaningful, the distributional convolution coincides with the convolution known from integration theory as

$$(T_f * \varphi)(x) = \langle T_f, \tilde{\tau}_x \varphi \rangle = \int_{\mathbb{R}^d} f(y) \tilde{\tau}_x \varphi(y) \,\mathrm{d}y = \int_{\mathbb{R}^d} f(y) \varphi(x-y) \,\mathrm{d}y = (f * \varphi)(x).$$

Theorem 15.13. *For all $T \in \mathcal{S}'(\mathbb{R}^d)$ and for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ we have $T * \varphi \in \mathrm{C}^\infty(\mathbb{R}^d)$ and it holds*

$$\partial_j (T * \varphi) = (\partial_j T) * \varphi = T * (\partial_j \varphi).$$

Proof. **Step 1: $T * \varphi$ is continuous:** Let $x, h \in \mathbb{R}^d$. Then

$$\tilde{\tau}_{x+h} \varphi(y) = \varphi(x+h-y) \xrightarrow{h \to 0} \varphi(x-y) = \tilde{\tau}_x \varphi(y) \quad \text{ in } \mathcal{S}(\mathbb{R}^d).$$

Compare this step to Example 15.5 for more detail. Hence,

$$(T * \varphi)(x+h) = \langle T, \tilde{\tau}_{x+h} \varphi \rangle \xrightarrow{h \to 0} \langle T, \tilde{\tau}_x \varphi \rangle = (T * \varphi)(x).$$

Step 2: $\partial_j (T * \varphi)$ is continuous with $\partial_j (T * \varphi) = T * (\partial_j \varphi)$: Let $h \in \mathbb{R} \setminus \{0\}$ and denote by e_j the j-th unit vector. Then, as above,

$$\begin{aligned} \frac{1}{h} \left(\tilde{\tau}_{x+he_j} \varphi - \tilde{\tau}_x \varphi \right)(y) &= \frac{1}{h} (\varphi(x + he_j - y) - \varphi(x-y)) \\ &\xrightarrow{h \to 0} \partial_j \varphi(x-y) = \tilde{\tau}_x \partial_j \varphi(y) \quad \text{ in } \mathcal{S}(\mathbb{R}^d). \end{aligned}$$

Hence,

$$\begin{aligned} \partial_j (T * \varphi)(x) &= \lim_{h \to 0} \frac{1}{h} \left(\langle T, \tilde{\tau}_{x+he_j} \varphi \rangle - \langle T, \tilde{\tau}_x \varphi \rangle \right) \\ &= \lim_{h \to 0} \left\langle T, \frac{1}{h} \left(\tilde{\tau}_{x+he_j} \varphi - \tilde{\tau}_x \varphi \right) \right\rangle \\ &= \langle T, \tilde{\tau}_x \partial_j \varphi \rangle = T * (\partial_j \varphi)(x) \end{aligned}$$

and since $\partial_j \varphi \in \mathcal{S}(\mathbb{R}^d)$, we get that $\partial_j (T * \varphi)$ is continuous by step 1.

Step 3: $\partial_j(T * \varphi) = (\partial_j T) * \varphi$**:** We have

$$[\partial_j (\tilde{\tau}_x \varphi)](y) = (\partial_j (\varphi(x - \cdot))(y) = -(\partial_j \varphi)(x - y) = -\tilde{\tau}_x(\partial_j \varphi). \qquad (*)$$

Thus, Step 2 implies

$$\begin{aligned}\partial_j(T * \varphi)(x) &= (T * (\partial_j \varphi))(x) = \langle T, \tilde{\tau}_x \partial_j \varphi \rangle \\ &\overset{(*)}{=} -\langle T, \partial_j (\tilde{\tau}_x \varphi) \rangle = \langle \partial_j T, \tilde{\tau}_x \varphi \rangle = ((\partial_j T) * \varphi)(x).\end{aligned}$$

□

Example 15.14. For $f \in \mathcal{S}(\mathbb{R}^d)$ it holds

$$(\delta_0 * f)(x) = \langle \delta_0, \tilde{\tau}_x f \rangle = (\tilde{\tau}_x f)(0) = f(x - 0) = f(x).$$

Thus, we see that convoluting with δ_0 does nothing, or δ_0 is the neutral element with respect to convolution.

In the next step, we want to define the Fourier transform for tempered distributions. Since for $f, g \in \mathcal{S}(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} \hat{f} g = \int_{\mathbb{R}^d} f \hat{g},$$

the following definition is natural.

Definition 15.15. Let $T \in \mathcal{S}'(\mathbb{R}^d)$. Then the Fourier transform $\hat{T} = \mathcal{F}T$ is defined by

$$\left\langle \hat{T}, \varphi \right\rangle := \langle T, \hat{\varphi} \rangle, \quad \varphi \in \mathcal{S}(\mathbb{R}^d).$$

Theorem 15.16. *The Fourier transform* $\mathcal{F}\colon \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ *is a continuous isomorphism and its inverse* $\mathcal{F}^{-1}$ *is given by*

$$\langle \check{T}, \varphi \rangle = \langle \mathcal{F}^{-1} T, \varphi \rangle := \langle T, \check{\varphi} \rangle, \quad T \in \mathcal{S}'(\mathbb{R}^d),\ \varphi \in \mathcal{S}(\mathbb{R}^d).$$

Furthermore, for $\psi \in \mathcal{S}(\mathbb{R}^d)$ *it holds* $\widehat{T_\psi} = T_{\hat{\psi}}$.

Proof. Linearity is straightforward, so we show that $\hat{T} \in \mathcal{S}'(\mathbb{R}^d)$ for an arbitrary tempered distribution T. Let $(\varphi_n) \subseteq \mathcal{S}(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$ be such that (φ_n) coverges to φ in $\mathcal{S}(\mathbb{R}^d)$. Then by the continuity of the Fourier transform in $\mathcal{S}(\mathbb{R}^d)$ the sequence $(\hat{\varphi}_n)$ converges to $\hat{\varphi}$ in $\mathcal{S}(\mathbb{R}^d)$. This implies

$$\left\langle \hat{T}, \varphi_n \right\rangle = \langle T, \hat{\varphi}_n \rangle \longrightarrow \langle T, \hat{\varphi} \rangle = \left\langle \hat{T}, \varphi \right\rangle$$

for $n \to \infty$ and this means $\hat{T} \in \mathcal{S}'(\mathbb{R}^d)$.

In order to show that the Fourier transform is contiunuous, let $(T_n) \subseteq \mathcal{S}'(\mathbb{R}^d)$ with $T_n \to T$ in $\mathcal{S}'(\mathbb{R}^d)$ be given. Then for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ we have

$$\lim_{n\to\infty} \left\langle \hat{T}_n, \varphi \right\rangle = \lim_{n\to\infty} \langle T_n, \hat{\varphi} \rangle = \langle T, \hat{\varphi} \rangle = \left\langle \hat{T}, \varphi \right\rangle .$$

Thus $(\hat{T}_n)$ converges to $\hat{T}$ in $\mathcal{S}'(\mathbb{R}^d)$.

The invertibility of the Fourier transform is also inherited from the Schwartz space, since for all $T \in \mathcal{S}'(\mathbb{R}^d)$ and all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ it holds

$$\left\langle \mathcal{F}\mathcal{F}^{-1}T, \varphi \right\rangle = \left\langle \mathcal{F}^{-1}T, \mathcal{F}\varphi \right\rangle = \left\langle T, \mathcal{F}^{-1}F\varphi \right\rangle = \langle T, \varphi \rangle ,$$

so $\mathcal{F}\mathcal{F}^{-1}$ is the identity on $\mathcal{S}'(\mathbb{R}^d)$. Analogously one shows $\mathcal{F}^{-1}\mathcal{F} = \mathrm{id}$.

Finally, let $\psi \in \mathcal{S}(\mathbb{R}^d)$ be given. Then for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$

$$\left\langle \widehat{T_\psi}, \varphi \right\rangle = \langle T_\psi, \hat{\varphi} \rangle = \int_{\mathbb{R}^d} \psi \hat{\varphi} = \int_{\mathbb{R}^d} \hat{\psi} \varphi = \left\langle T_{\hat{\psi}}, \varphi \right\rangle ,$$

so $\widehat{T_\psi} = T_{\hat{\psi}}$. □

Example 15.17. Consider once more the delta distribution δ_0. Then we calculate for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$

$$\left\langle \hat{\delta}_0, \varphi \right\rangle = \langle \delta_0, \hat{\varphi} \rangle = \hat{\varphi}(0) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi i 0 \cdot x} \varphi(x) \,\mathrm{d}x = \int_{\mathbb{R}^d} \varphi(x) \,\mathrm{d}x = \langle 1, \varphi \rangle$$

This implies $\hat{\delta}_0 = 1$.

The usual formulae for manipulating Fourier transforms all can be transferred to tempered distributions. We collect the most important ones in the following lemma. The proof remains as an exercise.

Lemma 15.18. *For all $T \in \mathcal{S}'(\mathbb{R}^d)$ and all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ it holds*

a) $\widehat{T * \varphi} = \hat{T} \cdot \hat{\varphi}$ *and* $\widehat{T \cdot \varphi} = \hat{T} * \hat{\varphi}$.

b) $\widehat{\partial^\alpha T} = (2\pi \mathrm{i} \xi)^\alpha \hat{T}$.

c) $\partial^\alpha \hat{T} = \mathcal{F}((-2\pi \mathrm{i} x)^\alpha T)$.

16 Marcinkiewicz Interpolation Theorem

In this chapter we prove a theorem that allows to transfer continuity properties of operators from one L^p space to another. This is a special case of so called 'interpolation' results. For all this chapter let (X, Σ, μ) and $(Y, \mathcal{S}, \nu)$ be σ-finite measure spaces.

Definition 16.1. Let $T\colon D(T) \to \mathrm{L}^0(Y)$ a mapping with $D(T)$ a subspace of $\mathrm{L}^0(X)$ and let $1 \le p, q \le \infty$.

a) The map T is called sublinear, if for all admissible f, g it holds $|T(f+g)| \le |Tf| + |Tg|$.

b) We say that T is of strong type (p,q), if $\mathrm{L}^p(X) \subset D(T)$ and there is $C \ge 0$ such that $\|Tf\|_{\mathrm{L}^q(Y)} \le C\,\|f\|_{\mathrm{L}^p(X)}$ for all $f \in \mathrm{L}^p(X)$.

c) If $q < \infty$, we say that T is of weak type (p,q), if $\mathrm{L}^p(X) \subset D(T)$ and there is some $C \ge 0$ such that for all $f \in \mathrm{L}^p(X)$ and all $\lambda > 0$, it holds

$$\nu\big(\{y \in Y : |Tf(y)| > \lambda\}\big) \le \left(\frac{C\,\|f\|_{\mathrm{L}^p(X)}}{\lambda}\right)^q.$$

Finally, T is called of weak type (p,∞), if it is of strong type (p,∞).

Note that for linear operators being of strong type just means continuity.

We show that being of weak type is indeed the weaker notion.

Lemma 16.2. *Let $1 \le p, q \le \infty$ and let T be as in the above definition. If T is of strong type (p,q) then T is of weak type (p,q) with the same constant.*

Proof. In the case $q = \infty$ there is nothing to prove. For all the other cases let $f \in \mathrm{L}^p(X)$ and $\lambda > 0$. Then by Tchebychev's inequality and since T is of strong type, we get

$$\nu\big(\{y \in Y : |Tf(y)| > \lambda\}\big) \le \frac{1}{\lambda^q}\,\|Tf\|^q_{\mathrm{L}^q(Y)} \le \left(\frac{C\,\|f\|_{\mathrm{L}^p(X)}}{\lambda}\right)^q. \qquad \square$$

Before we can formulate the main theorem of this chapter, we need one more small notation. For a measurable function $f : X \to \mathbb{K}$ and some $\lambda > 0$ we set

$$f_\lambda \coloneqq f \cdot \mathbf{1}_{\{|f| \le \lambda\}} \quad \text{and} \quad f^\lambda \coloneqq f \cdot \mathbf{1}_{\{|f| > \lambda\}}$$

to the effect that $f = f_\lambda + f^\lambda$.

Theorem 16.3 (Interpolation Theorem of Marcinkiewicz)**.** *Let $D(T)$ be a subspace of $\mathrm{L}^0(X)$ such that for all $f \in D(T)$ also the function f_λ is in $D(T)$ and let $T : D(T) \to \mathrm{L}^0(Y)$ be sublinear. If for $1 \le p_0 < p_1 \le \infty$ the map T is of weak type (p_0, p_0) with constant C_0 and of weak type (p_1, p_1) with constant C_1, then T is of strong type (p, p) for all $p \in (p_0, p_1)$ and the following estimates hold:*

$$\|Tf\|_{\mathrm{L}^p(Y)} \le 2 \left(\frac{p}{p - p_0} + \frac{p_1}{p_1 - p} \right)^{1/p} C_0^{\frac{p_0}{p}(1 - \frac{p - p_0}{p_1 - p_0})} C_1^{\frac{p_1}{p}\frac{p - p_0}{p_1 - p_0}} \|f\|_{\mathrm{L}^p(X)}, \quad \textit{in the case } p_1 < \infty$$

and

$$\|Tf\|_{\mathrm{L}^p(Y)} \le 2 \left(\frac{p}{p - p_0} \right)^{1/p} C_0^{\frac{p_0}{p}} C_1^{1 - \frac{p_0}{p}} \|f\|_{\mathrm{L}^p(X)} \quad \textit{for } p_1 = \infty.$$

For the proof of this theorem it is good to rememeber the following identity from integration theory: For all $1 \le p < \infty$ and all $f \in \mathrm{L}^p(X)$

$$\|f\|_{\mathrm{L}^p(X)}^p = p \int_0^\infty \mu(\{x \in X : |f(x)| > \lambda\}) \lambda^{p-1} \, \mathrm{d}\lambda. \tag{16.1}$$

Proof. We first treat the case of $p_1 < \infty$.

Let $f \in D(T) \cap \mathrm{L}^p(X)$ and $\lambda > 0$. We introduce some constant $\gamma > 0$ to be chosen suitably later. By the hypotheses on $D(T)$ also f_λ and f^λ are in this space and we get by the sublinearity of T

$$|Tf| = |T(f_\lambda + f^\lambda)| \le |Tf_\lambda| + |Tf^\lambda|.$$

This implies that

$$\begin{aligned} \{y \in Y : |Tf(y)| > 2\gamma\lambda\} &\subseteq \{y \in Y : |Tf_\lambda(y)| + |Tf^\lambda(y)| > 2\gamma\lambda\} \\ &\subseteq \{y \in Y : |Tf_\lambda(y)| > \gamma\lambda\} \cup \{y \in Y : |Tf^\lambda(y)| > \gamma\lambda\}. \end{aligned}$$

Since T is of weak type (p_0, p_0) and (p_1, p_1), we conclude

$$\begin{aligned} \nu\big(\{y \in Y : |Tf(y)| > 2\gamma\lambda\}\big) &\le \nu\big(\{y \in Y : |Tf_\lambda(y)| > \gamma\lambda\}\big) + \nu\big(\{y \in Y : |Tf^\lambda(y)| > \gamma\lambda\}\big) \\ &\le \left(\frac{C_0 \left\|f^\lambda\right\|_{\mathrm{L}^{p_0}(X)}}{\gamma\lambda} \right)^{p_0} + \left(\frac{C_1 \|f_\lambda\|_{\mathrm{L}^{p_1}(X)}}{\gamma\lambda} \right)^{p_1}. \end{aligned}$$

Using (16.1) three times, we use this to estimate the p-norm of Tf:

$$\begin{aligned}
&2^{-p} \|Tf\|_{\mathrm{L}^p(Y)}^p \\
&= 2^{-p} p \int_0^\infty \nu\big(\{y \in Y : |Tf(y)| > \tau\}\big) \tau^{p-1} \,\mathrm{d}\tau \\
&\overset{\tau=2\gamma\lambda}{=} p\gamma^p \int_0^\infty \nu\big(\{y \in Y : |Tf(y)| > 2\gamma\lambda\}\big) \lambda^{p-1} \,\mathrm{d}\lambda \\
&\le p\gamma^p \int_0^\infty \left(\frac{C_0 \left\|f^\lambda\right\|_{\mathrm{L}^{p_0}(X)}}{\gamma\lambda} \right)^{p_0} \lambda^{p-1} \,\mathrm{d}\lambda + p\gamma^p \int_0^\infty \left(\frac{C_1 \left\|f_\lambda\right\|_{\mathrm{L}^{p_1}(X)}}{\gamma\lambda} \right)^{p_1} \lambda^{p-1} \,\mathrm{d}\lambda \\
&= C_0^{p_0} p \gamma^{p-p_0} \int_0^\infty \left\|f^\lambda\right\|_{\mathrm{L}^{p_0}(X)}^{p_0} \lambda^{p-p_0-1} \,\mathrm{d}\lambda + C_1^{p_1} p \gamma^{p-p_1} \int_0^\infty \left\|f_\lambda\right\|_{\mathrm{L}^{p_1}(X)}^{p_1} \lambda^{p-p_1-1} \,\mathrm{d}\lambda \\
&= C_0^{p_0} p p_0 \gamma^{p-p_0} \int_0^\infty \int_0^\infty \mu\big(\{x \in X : |f^\lambda(x)| > \tau\}\big) \tau^{p_0-1} \,\mathrm{d}\tau \lambda^{p-p_0-1} \,\mathrm{d}\lambda \\
&\qquad + C_1^{p_1} p p_1 \gamma^{p-p_1} \int_0^\infty \int_0^\infty \mu\big(\{x \in X : |f_\lambda(x)| > \tau\}\big) \tau^{p_1-1} \,\mathrm{d}\tau \lambda^{p-p_1-1} \,\mathrm{d}\lambda \\
&=: C_0^{p_0} \gamma^{p-p_0} I + C_1^{p_1} \gamma^{p-p_1} II.
\end{aligned}$$

We estimate I and II separately. For II we first note, that f_λ is always smaller than λ, so we can replace the upper limit in the inner integral by λ. Furthermore it holds $|f_\lambda| \le |f|$, so we find

$$\{x \in X : |f_\lambda(x)| > \tau\} \subseteq \{x \in X : |f(x)| > \tau\}.$$

This yields

$$II \le p p_1 \int_0^\infty \int_0^\lambda \mu\big(\{x \in X : |f(x)| > \tau\}\big) \lambda^{p-p_1-1} \tau^{p_1-1} \,\mathrm{d}\tau \,\mathrm{d}\lambda.$$

We apply the Tonelli Theorem and obtain, noting that $p < p_1$ by hypotheses,

$$\begin{aligned}
&= p p_1 \int_0^\infty \int_\tau^\infty \mu\big(\{x \in X : |f(x)| > \tau\}\big) \lambda^{p-p_1-1} \tau^{p_1-1} \,\mathrm{d}\lambda \,\mathrm{d}\tau \\
&= p p_1 \int_0^\infty \mu\big(\{x \in X : |f(x)| > \tau\}\big) \tau^{p_1-1} \frac{1}{p-p_1} \lambda^{p-p_1} \Big|_{\lambda=\tau}^{\lambda=\infty} \,\mathrm{d}\tau \\
&= \frac{p p_1}{p_1 - p} \int_0^\infty \mu\big(\{x \in X : |f(x)| > \tau\}\big) \tau^{p_1-1} \tau^{p-p_1} \,\mathrm{d}\tau.
\end{aligned}$$

Relying once more on (16.1), we find

$$= \frac{p_1}{p_1 - p} \|f\|_{\mathrm{L}^p(X)}^p .$$

We turn to the estimate of I. It holds

$$\{x \in X : |f^\lambda(x)| > \tau\} = \begin{cases} \{x \in X : |f(x)| > \tau\}, & \text{if } \tau > \lambda \\ \{x \in X : |f(x)| > \lambda\}, & \text{if } \tau \le \lambda. \end{cases}$$

Thus

$$I = pp_0 \int_0^\infty \left[\int_0^\lambda \mu\big(\{x \in X : |f(x)| > \lambda\}\big)\tau^{p_0-1}\,\mathrm{d}\tau + \int_\lambda^\infty \mu\big(\{x \in X : |f(x)| > \tau\}\big)\tau^{p_0-1}\,\mathrm{d}\tau\right] \lambda^{p-p_0-1}\,\mathrm{d}\lambda.$$

Calculating the inner integral in the first part and applying Tonelli in the second we find

$$= pp_0 \int_0^\infty \frac{1}{p_0}\tau^{p_0}\Big|_{\tau=0}^{\tau=\lambda} \mu\big(\{x \in X : f(x) > \lambda\}\big)\lambda^{p-p_0-1}\,\mathrm{d}\lambda + pp_0 \int_0^\infty \int_0^\tau \lambda^{p-p_0-1}\,\mathrm{d}\lambda \mu\big(\{x \in X : |f(x)| > \tau\}\big)\tau^{p_0-1}\,\mathrm{d}\tau.$$

Since $p_0 < p$ we may calculate the inner integral in the second part and get

$$\begin{aligned} &= p\int_0^\infty \mu\big(\{x \in X : |f(x)| > \lambda\}\big)\lambda^{p-1}\,\mathrm{d}\lambda + pp_0 \int_0^\infty \frac{1}{p-p_0}\tau^{p-p_0}\mu\big(\{x \in X : |f(x)| > \tau\}\big)\tau^{p_0-1}\,\mathrm{d}\tau \\ &= \left(1 + \frac{p_0}{p-p_0}\right)\|f\|^p_{\mathrm{L}^p(X)} = \frac{p-p_0+p_0}{p-p_0}\|f\|^p_{\mathrm{L}^p(X)} = \frac{p}{p-p_0}\|f\|^p_{\mathrm{L}^p(X)}. \end{aligned}$$

Putting everything together, we conclude that

$$\|Tf\|^p_{\mathrm{L}^p(Y)} \le 2^p \left(C_0^{p_0}\gamma^{p-p_0}\frac{p}{p-p_0} + C_1^{p_1}\gamma^{p-p_1}\frac{p_1}{p_1-p}\right)\|f\|^p_{\mathrm{L}^p(X)}.$$

Now, choose γ in such a way that $C_0^{p_0}\gamma^{p-p_0} = C_1^{p_1}\gamma^{p-p_1}$, i.e. $\gamma = (C_1^{p_1}/C_0^{p_0})^{\frac{1}{p_1-p_0}}$. Then the content of the parantheses in the above constant transforms to

$$\left(\frac{p}{p-p_0} + \frac{p_1}{p_1-p}\right) C_0^{p_0(1-\frac{p-p_0}{p_1-p_0})} C_1^{p_1\frac{p-p_0}{p_1-p_0}}$$

and we find that T is of strong type (p,p) with the asserted constant.

We turn towards the case of $p_1 = \infty$.

Let $f \in D(T)$ and $\lambda > 0$ be given and set $\gamma := 1/(2C_1)$. Then since T is of strong type (∞,∞) by hypotheses,

$$\|Tf_{\gamma\lambda}\|_{\mathrm{L}^\infty(Y)} \le C_1 \|f_{\gamma\lambda}\|_{\mathrm{L}^\infty(X)} \le C_1\gamma\lambda = \frac{\lambda}{2}.$$

This means that $|Tf_{\gamma\lambda}| \le \lambda/2$ almost everywhere and thus

$$\begin{aligned} \nu\big(\{y \in Y : |Tf(y)| > \lambda\}\big) &\le \nu\big(\{y \in Y : |Tf_{\gamma\lambda}(y)| + |Tf^{\gamma\lambda}(y)| > \lambda\}\big) \\ &\le \nu\big(\{y \in Y : |Tf_{\gamma\lambda}(y)| > \lambda/2\}\big) + \nu\big(\{y \in Y : |Tf^{\gamma\lambda}(y)| > \lambda/2\}\big) \\ &= \nu\big(\{y \in Y : |Tf^{\gamma\lambda}(y)| > \lambda/2\}\big). \end{aligned}$$

We use this to estimate the L^p-norm of Tf in the following way:

$$\begin{aligned}\|Tf\|_{\mathrm{L}^p(Y)}^p &= p\int_0^\infty \nu\big(\{y\in Y : |Tf(y)| > \lambda\}\big)\lambda^{p-1}\,\mathrm{d}\lambda\\ &\le p\int_0^\infty \nu\big(\{y\in Y : |Tf^{\gamma\lambda}(y)| > \lambda/2\}\big)\lambda^{p-1}\,\mathrm{d}\lambda.\end{aligned}$$

Now the weak (p_0,p_0)-estimate enters to the effect that

$$\begin{aligned}&\le p\int_0^\infty \left(\frac{C_0\left\|f^{\gamma\lambda}\right\|_{\mathrm{L}^{p_0}(X)}}{\frac{\lambda}{2}}\right)^{p_0}\lambda^{p-1}\,\mathrm{d}\lambda\\ &= (2C_0)^{p_0}p\int_0^\infty \left\|f^{\gamma\lambda}\right\|_{\mathrm{L}^{p_0}(X)}^{p_0}\lambda^{p-p_0-1}\,\mathrm{d}\lambda\\ &= (2C_0)^{p_0}p\gamma^{p_0-p}\int_0^\infty \left\|f^{\lambda}\right\|_{\mathrm{L}^{p_0}(X)}^{p_0}\lambda^{p-p_0-1}\,\mathrm{d}\lambda.\end{aligned}$$

Looking into the proof in the case $p_1<\infty$, we find that this is exactly

$$= (2C_0)^{p_0}\gamma^{p_0-p}I.$$

Since p_1 is not used in the earlier estimate of I, we can just copy this and obtain

$$\le (2C_0)^{p_0}\gamma^{p_0-p}\frac{p}{p-p_0}\|f\|_{\mathrm{L}^p(X)}^p.$$

Finally, putting in $\gamma = 1/(2C_1)$ we end up with

$$= (2C_0)^{p_0}(2C_1)^{p-p_0}\frac{p}{p-p_0}\|f\|_{\mathrm{L}^p(X)}^p$$

and taking the pth root, this is

$$\|Tf\|_{\mathrm{L}^p(Y)} \le 2C_0^{\frac{p_0}{p}}C_1^{1-\frac{p_0}{p}}\left(\frac{p}{p-p_0}\right)^{\frac{1}{p}}\|f\|_{\mathrm{L}^p(X)}. \qquad \square$$

This theorem can be applied to the following sublinear map, that we will need in the next chapter.

Definition 16.4. a) If $Q\subseteq\mathbb{R}^d$ is a cube, we will write $\ell(Q)$ for its sidelength.

b) We consider the grid of unit cubes

$$\mathcal{D}_0 := \{Q\subseteq\mathbb{R}^d : Q \text{ closed cube, } \ell(Q)=1 \text{ and all vertices of } Q \text{ are in } \mathbb{Z}^d\}$$

and for all $k\in\mathbb{Z}$ we set $\mathcal{D}_k := \{2^kQ : Q\in\mathcal{D}_0\}$. Then

$$\mathcal{D} := \bigcup_{k\in\mathbb{Z}}\mathcal{D}_k$$

is the set of all dyadic cubes in $\mathbb{R}^d$.

c) For $x \in \mathbb{R}^d$ we set $\mathcal{D}(x) := \{Q \in \mathcal{D} : x \in Q^\circ\}$. Then the (dyadic) Hardy-Littlewood maximal function or (dyadic) Hardy-Littlewood maximal operator M is defined by

$$(\mathrm{M}f)(x) := \sup_{Q \in \mathcal{D}(x)} \frac{1}{\lambda_d(Q)} \int_Q |f|, \qquad x \in \mathbb{R}^d,$$

for all $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$.

Remark 16.5. The dyadic cubes have the following important property: If $Q_1, Q_2 \in \mathcal{D}$ are such that $Q_1^\circ \cap Q_2^\circ \neq \emptyset$, then $Q_1 \subseteq Q_2$ or $Q_2 \subseteq Q_1$, i.e. if two such cubes have non-disjoint interior, than they are contained one in another.

We collect some properties of the dyadic maximal operator.

Theorem 16.6. *a)* M *is sublinear and* $\mathrm{M}f$ *is mesurable for every* $f \in \mathrm{L}^1(\mathbb{R}^d)$.

b) For every $f \in \mathrm{L}^1(\mathbb{R}^d)$ *it holds* $|f| \leq \mathrm{M}f$ *almost everywhere.*

c) M *is of weak type* $(1,1)$.

d) M *is of strong type* (p,p) *for all* $1 < p \leq \infty$.

Proof. a) Sublinearity is straightforward and measurability of $\mathrm{M}f$ follows, once we show that $(\mathrm{M}f)^{-1}((a,\infty))$ is open for every $a \in \mathbb{R}$. In order to do so, let $x \in (\mathrm{M}f)^{-1}((a,\infty))$. Then $\mathrm{M}f(x) > a$, which means that there is some dyadic cube Q with $x \in Q^\circ$ such that

$$\frac{1}{\lambda_d(Q)} \int_Q |f| > a.$$

Since x is in the interior of Q, there is some ball B around x that is completely contained in Q and for all $y \in B$ we consequently have

$$\mathrm{M}f(y) \geq \frac{1}{\lambda_d(Q)} \int_Q |f| > a,$$

so $B \subseteq (\mathrm{M}f)^{-1}((a,\infty))$.

b) In order to prove this, recall the Lebesgue Differentiation Theorem: Let $f \in \mathrm{L}^1(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ and (Q_n) be a sequence of cubes such that $x \in Q_n$ for all $n \in \mathbb{N}$ and $\bigcap_{n\in\mathbb{N}} Q_n = \{x\}$. Then

$$f(x) = \lim_{n\to\infty} \frac{1}{\lambda_d(Q_n)} \int_{Q_n} f(y)\,\mathrm{d}y$$

almost everywhere in $\mathbb{R}^d$. Using this, we get

$$\begin{aligned}
|f(x)| &= \lim_{n\to\infty} \left| \frac{1}{\lambda_d(Q_n)} \int_{Q_n} f(y)\,\mathrm{d}y \right| \\
&\le \lim_{n\to\infty} \sup_{Q\in D(x)} \frac{1}{\lambda_d(Q)} \int_Q |f(y)|\ \mathrm{d}y \\
&= \lim_{n\to\infty} (Mf)(x) \\
&= (Mf)(x)
\end{aligned}$$

almost everywhere in $\mathbb{R}^d$.

c) Exercise

d) One finds that M is of strong type (∞,∞) just by estimating for every $f \in \mathrm{L}^\infty(\mathbb{R}^d)$ and all $x \in \mathbb{R}^d$

$$|\mathrm{M}f(x)| = \sup_{Q\in\mathcal{D}(x)} \frac{1}{\lambda_d(Q)} \int_Q |f| \le \sup_{Q\in\mathcal{D}(x)} \frac{1}{\lambda_d(Q)} \int_Q \|f\|_{\mathrm{L}^\infty(\mathbb{R}^d)} = \|f\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \,.$$

So, the claim follows from the Marcinkiewicz Interpolation Theorem. $\square$

17 The Calderón-Zygmund Decomposition

We introduce an important class of linear operators, that are given by convolution with a so called kernel. These appear frequently as solution operators to PDEs.

Definition 17.1. a) A linear operator $T\colon \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is called a convolution operator, if there is some $K \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \{0\})$ such that for all $f, \varphi \in \mathcal{S}$ with $\operatorname{supp}(f) \cap \operatorname{supp}(\varphi) = \emptyset$ it holds

$$\langle Tf, \varphi\rangle = \int_{\mathbb{R}^d \times \mathbb{R}^d} K(x-y) f(y) \varphi(x) \,\mathrm{d}(x,y).$$

In this case K is called convolution kernel of T.

b) A convolution kernel K satisfies the Hörmander condition, if

$$\sup_{y \in \mathbb{R}^d \setminus \{0\}} \int_{|x|>2|y|} \big|K(x-y) - K(x)\big| \,\mathrm{d}x =: C_H < \infty.$$

c) If T is a convolution operator with a kernel K that satisfies the Hörmander condition, and if T may be continuously extended to a bounded linear operator on $\mathrm{L}^2(\mathbb{R}^d)$ such that

$$Tf(x) = \int_{\mathbb{R}^d} K(x-y) f(y) \,\mathrm{d}y$$

for all $f \in \mathrm{L}^2(\mathbb{R}^d)$, then T is called a Calderón-Zygmund operator.

The goal of this chapter will be to prove the following theorem.

Theorem 17.2. *Every Calderón-Zygmund operator T can be continuously extended to a bounded operator on $L^p(\mathbb{R}^d)$ for all $1 < p < \infty$ and for all $f \in \mathrm{L}^p(\mathbb{R}^d)$ it holds*

$$\|Tf\|_{\mathrm{L}^p(\mathbb{R}^d)} \le \begin{cases} C(C_{\mathrm{L}^2} + C_H)^{\frac{2-p}{p}} C_{\mathrm{L}^2}^{\frac{2p-2}{p}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d)}, & \text{if } p \le 2, \\ C(C_{\mathrm{L}^2} + C_H)^{\frac{p-2}{p}} C_{\mathrm{L}^2}^{\frac{2}{p}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d)}, & \text{if } p > 2. \end{cases}$$

Here C_{L^2} denotes the boundedness constant of T on $\mathrm{L}^2(\mathbb{R}^d)$ and C is independent of T.

The outline of the proof is as follows: A Calderón-Zygmund operator is of strong type $(2,2)$, so it is also of weak type $(2,2)$. The main work will be to prove that it is of weak type $(1,1)$. Then the Marcinkiewicz Interpolation Theorem gives the claim for all $p \in (1,2]$. For p bigger than two, one observes that the adjoint of a Calderón-Zygmund operator is again of the same class. Then the claim follows from a duality argument.

For the proof, that every Calderón-Zygmund operator is of weak type $(1,1)$, we will introduce a powerful machinery. Recall the notation concerning dyadic cubes from the preceding chapter.

Lemma 17.3 (Whitney decomposition)**.** *Let $E \subsetneq \mathbb{R}^d$ be open. Then there is a selection $(Q_j)_{j\in\mathbb{N}}$ of dyadic cubes with the following properties:*

a) They have pairwise disjoint interiors, i.e. $Q_j^\circ \cap Q_k^\circ = \emptyset$ for all $j \neq k$,

b) $\bigcup_{j\in\mathbb{N}} Q_j = E$ and

c) For all $j \in \mathbb{N}$ it holds $\sqrt{d}\ell(Q_j) \leq \operatorname{dist}(Q_j, E^{\mathrm{c}}) \leq 4\sqrt{d}\ell(Q_j)$.

Proof. For every $x \in E$, the distance of x to E^{c} is strictly positive, so, setting

$$E_k := \left\{x \in E : 2\sqrt{d}2^k < \operatorname{dist}(x, E^{\mathrm{c}}) \leq 4\sqrt{d}2^k\right\}$$

for every $k \in \mathbb{Z}$, we get $E = \bigcup_{k\in\mathbb{Z}} E_k$. As a first candidate for our Whitney decomposition we consider

$$\mathcal{Q} := \bigcup_{k\in\mathbb{Z}} \{Q \in \mathcal{D}_k : Q \cap E_k \neq \emptyset\}.$$

Then $\mathcal{Q}$ fulfills already b) and c), as we will show now.

In order to show that $\mathcal{Q}$ satisfies c), let $Q \in \mathcal{Q}$. Then, by the construction of $\mathcal{Q}$, there is some $k \in \mathbb{Z}$ with $Q \in \mathcal{D}_k$ and there exists an $x \in Q \cap E_k$. This implies $2\sqrt{d}2^k \leq \operatorname{dist}(x, E^{\mathrm{c}}) \leq 4\sqrt{d}2^k$ and $\ell(Q) = 2^k$, so we find

$$\begin{aligned}\sqrt{d}\ell(Q) = \sqrt{d}2^k = 2\sqrt{d}2^k - \sqrt{d}2^k &\leq \operatorname{dist}(x, E^c) - \sqrt{d}\ell(Q)\\ &\leq \operatorname{dist}(Q, E^c) + \operatorname{dist}(x, \partial Q) - \sqrt{d}\ell(Q) \leq \operatorname{dist}(Q, E^c) + \operatorname{diam}(Q) - \operatorname{diam}(Q)\\ &= \operatorname{dist}(Q, E^c)\\ &\leq \operatorname{dist}(x, E^c) \leq 4\sqrt{d}2^k = 4\sqrt{d}\ell(Q).\end{aligned}$$

We show that the cubes of $\mathcal{Q}$ have property b). We just saw that $\operatorname{dist}(Q, E^{\mathrm{c}}) \geq \sqrt{d}\ell(Q) > 0$ for all $Q \in \mathcal{Q}$, so every cube in $\mathcal{Q}$ is contained in E, which implies $\bigcup_{Q\in\mathcal{Q}} Q \subseteq E$.

For the converse inclusion let $x \in E$ and choose $k \in \mathbb{Z}$ such that $x \in E_k$. Since the cubes in $\mathcal{D}_k$ cover all of $\mathbb{R}^d$, there is some $Q_x \in \mathcal{D}_k$ with $x \in Q_x$. Now, Q_x is a cube in $\mathcal{D}_k$ with $x \in Q_x \cap E_k$, so $Q_x \in \mathcal{Q}$ and this implies $E \subseteq \bigcup_{Q\in\mathcal{Q}} Q$.

The only problem about $\mathcal{Q}$ is, that its cubes do not have pairwise disjoint interiors. There are too many cubes in $\mathcal{Q}$, but it will turn out that all of these superfluous cubes are not needed to accomplish b), so we will throw them out now.

For this, we first prove, that for every $Q \in \mathcal{Q}$ there is some maximal $\widetilde{Q} \in \mathcal{Q}$ with $Q \subseteq \widetilde{Q}$. Let $Q \in \mathcal{Q}$. Pick some point $x \in Q^\circ$. Then $x \in E$ and for all $Q_* \in \mathcal{Q}$ with $Q \subseteq Q_*$ it holds

$$\ell(Q_*) \leq \frac{\operatorname{dist}(Q_*, E^c)}{\sqrt{d}} \leq \frac{\operatorname{dist}(x, E^c)}{\sqrt{d}}.$$

This means that all cubes in $\mathcal{Q}$ that contain Q are bounded in size. So, for our cube Q, we can choose $\widetilde{Q}$ to be a cube with maximal size in $\mathcal{Q}$ that contains Q.

Finally, we can define our Whitney decomposition as any enumeration $(Q_j)_{j\in\mathbb{N}}$ of the set $\{\widetilde{Q} : Q \in \mathcal{Q}\}$. Then for this choice of cubes c) is still satisfied, while a) is now forced by Remark 16.5: If $Q_k^\circ \cap Q_j^\circ \neq \emptyset$ for some $j, k \in \mathbb{N}$, then one of these cubes is contained in the other. So, either $j = k$ or the inclusion is strict. And the latter case would be in contradiction to the maximality of the cubes $(Q_j)_{j\in\mathbb{N}}$.

It remains to assure that we have not discarded too many cubes, that is b) is still valid. For this, let $x \in E$. Then by our considerations above, there is some cube $Q \in \mathcal{Q}$ containing x. This means that x is also in the corresponding cube $\widetilde{Q}$, and so it is contained in a Whitney cube. □

Theorem 17.4 (Calderón-Zygmund decomposition)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^d)$ and $\lambda > 0$. Then there exist at most countably many cubes $Q_j \in \mathcal{D}$, $j \in \mathbb{N}$, with pairwise disjoint interiors and there are functions $g, b \in \mathrm{L}^1(\mathbb{R}^d)$ with $f = g + b$, such that the following assertions hold for some $C \geq 0$ that is independent of f and λ*

a) $g \in \mathrm{L}^\infty(\mathbb{R}^d)$ *with* $\|g\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \leq C\lambda$,

b) $\|g\|_{\mathrm{L}^1(\mathbb{R}^d)} \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}$,

c) $b = \sum_{j\in\mathbb{N}} b_j$ *with* $\operatorname{supp}(b_j) \subseteq Q_j$ *for all* $j \in \mathbb{N}$,

d) $\int_{\mathbb{R}^d} b_j = 0$ *for all* $j \in \mathbb{N}$,

e) $\|b_j\|_{\mathrm{L}^1(\mathbb{R}^d)} \leq 2 \int_{Q_j} |f|$ *for all* $j \in \mathbb{N}$, *and*

f) $\sum_{j\in\mathbb{N}} \lambda_d(Q_j) \leq \frac{C}{\lambda} \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}$.

Remark 17.5. The parameter λ in this theorem is often refered to as the *height* of the Calderón-Zygmund decomposition. Furthermore, the function g is often called the *good* function and b the *bad* function, whence the notation. The idea is that g contains the nice, bounded parts of f, while b contains the singularities. The payoff of the decomposition is that there is a good control on the support of the bad function. This then helps in the estimates as we will see.

Proof of Theorem 17.4. Let $f \in \mathrm{L}^1(\mathbb{R}^d)$ and $\lambda > 0$. Invoking the Hardy-Littlewood maximal function, we consider the set $E_\lambda := \{x \in \mathbb{R}^d : \mathrm{M}f(x) > \lambda\}$. Since the maximal operator is of weak type $(1,1)$ by Threorem 16.6, this set has finite measure:

$$\lambda_d(E_\lambda) = \lambda_d(\{x \in \mathbb{R}^d : |\mathrm{M}f(x)| > \lambda\}) \le \frac{C}{\lambda} \|f\|_{\mathrm{L}^1(\mathbb{R}^d)} < \infty.$$

In particular E_λ is not the whole space. Furthermore, $E_\lambda = (\mathrm{M}f)^{-1}((\lambda,\infty))$ and we have already shown in the proof of Theorem 16.6 that this set is open. According to Lemma 17.3 we can therefore take $(Q_j)_{j\in\mathbb{N}} \subseteq \mathcal{D}$ as the Whitney decomposition of E_λ.

We can immediately infer f) thanks to

$$\sum_{j\in\mathbb{N}} \lambda_d(Q_j) = \lambda_d(E_\lambda) \le \frac{C}{\lambda} \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}.$$

For every $j \in \mathbb{N}$ we set

$$b_j := \mathbf{1}_{Q_j}\left(f - \frac{1}{\lambda_d(Q_j)}\int_{Q_j} f\right), \qquad b := \sum_{j\in\mathbb{N}} b_j, \qquad \text{and} \quad g := f - b.$$

Note that b is well defined, since the ocurring sum is locally finite. This definition immediately estblishes c), as well as $f = g + b$.

In a next step d) follows from

$$\int_{\mathbb{R}^d} b_j = \int_{Q_j}\left(f - \frac{1}{\lambda_d(Q_j)}\int_{Q_j} f\right) = \int_{Q_j} f - \frac{\lambda_d(Q_j)}{\lambda_d(Q_j)}\int_{Q_j} f = 0$$

and e) is also easy:

$$\|b_j\|_{\mathrm{L}^1(\mathbb{R}^d)} = \int_{Q_j}\left|f - \frac{1}{\lambda_d(Q_j)}\int_{Q_j} f\right| \le \int_{Q_j} |f| + \frac{\lambda_d(Q_j)}{\lambda_d(Q_j)}\int_{Q_j} |f| = 2\int_{Q_j} |f|.$$

We turn to the proof of b). If $x \in Q_j^\circ$ for some $j \in \mathbb{N}$, then x is in no other Whitney cube, so

$$g(x) = f(x) - b(x) = f(x) - b_j(x) = f(x) - f(x) + \frac{1}{\lambda_d(Q_j)}\int_{Q_j} f = \frac{1}{\lambda_d(Q_j)}\int_{Q_j} f.$$

On the other hand, if $x \in E_\lambda^c$, then x is in no Whitney cube, and we simply have $g(x) = f(x)$. Since, the boundaries of all Whitney cubes together form a null set, we have

$$\begin{aligned}\|g\|_{L^1(\mathbb{R}^d)} &= \int_{E_\lambda} |g| + \int_{E_\lambda^c} |g| = \sum_{j\in\mathbb{N}} \int_{Q_j^\circ} |g| + \int_{E_\lambda^c} |f| = \sum_{j\in\mathbb{N}} \int_{Q_j^\circ} \left| \frac{1}{\lambda_d(Q_j)} \int_{Q_j} f \right| + \int_{E_\lambda^c} |f| \\ &\le \sum_{j\in\mathbb{N}} \int_{Q_j} |f| + \int_{E_\lambda^c} |f| = \int_{\mathbb{R}^d} |f| = \|f\|_{L^1(\mathbb{R}^d)},\end{aligned}$$

which is b).

It remains to prove a). Let $x \in E_\lambda$ and choose $j \in \mathbb{N}$ such that $x \in Q_j$. By the properties of the Whitney cubes it holds $\operatorname{dist}(Q_j, E_\lambda^c) \le 4\sqrt{d}\ell(Q_j)$. This means that $5\sqrt{d}Q_j$, i.e. the cube with the same midpoint as Q_j but sidelength $5\sqrt{d}\ell(Q_j)$, intersects E_λ^c. Now, choose $d_0 \in \mathbb{N}$ with $2^{d_0-1} \le 5\sqrt{d} \le 2^{d_0}$. Then also the even larger cube $2^{d_0}Q_j$ intersects E_λ^c, so we can pick some z from this intersection. Since $2^{d_0}Q$ is a dyadic cube that contains z, we find

$$|g(x)| = \frac{1}{\lambda_d(Q_j)} \left| \int_{Q_j} f \right| \le \frac{1}{\lambda_d(Q_j)} \int_{Q_j} |f| \le \frac{2^{dd_0}}{\lambda_d(2^{d_0}Q_j)} \int_{2^{d_0}Q_j} |f| \le 2^{dd_0} \mathrm{M}f(z) \le C\lambda,$$

where the last inequality is true, since $z \in E_\lambda^c$.

For $x \in E_\lambda^c$, it holds $\mathrm{M}f(x) \le \lambda$ and $g(x) = f(x)$, so we find with the help of Theorem 16.6 b) for almost all $x \in E_\lambda^c$

$$|g(x)| = |f(x)| \le \mathrm{M}f(x) \le \lambda. \qquad \square$$

We have now collected all the necessary tools to attack the main task in taming Calderón-Zygmund operators.

Lemma 17.6. *Every Calderón-Zygmund operator T is of weak type $(1, 1)$.*

Proof. We will show that there is some constant $C \ge 0$ independent of T such that for all $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and all $\lambda > 0$ the estimate

$$\lambda_d\big(\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\}\big) \le \frac{C}{\lambda}(C_{L^2} + C_H) \|f\|_{L^1(\mathbb{R}^d)} \tag{17.1}$$

is valid. For given $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, $\lambda > 0$ and some $\gamma > 0$ to be chosen later, we apply the Calderón-Zygmund decomposition at height $\gamma\lambda$, i.e. we decompose $f = g + b$ with functions $g, b \in L^1(\mathbb{R}^d)$ such that the properties a) to f) of the Calderón-Zygmund decomposition hold with λ replaced by $\gamma\lambda$. Property a) tells us that $g \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, so by the interpolation inequality we even have $g \in \mathrm{L}^2(\mathbb{R}^d)$. Then also for b there is no other choice as to ly in $\mathrm{L}^2(\mathbb{R}^d)$. This means, that we can apply T to these two functions and $Tf = Tg + Tb$.

In the same manner as in the proof of the Marcinkiewicz interpolation Theorem we split up the set we wish to control:

$$\begin{aligned}\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\} &\subseteq \{x \in \mathbb{R}^d : |Tg(x)| + |Tb(x)| > \lambda\} \\ &\subseteq \{x \in \mathbb{R}^d : |Tg(x)| > \lambda/2\} \cup \{x \in \mathbb{R}^d : |Tb(x)| > \lambda/2\}.\end{aligned}$$

This means that

$$\begin{aligned}&\lambda_d\big(\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\}\big) \\ \leq &\lambda_d\big(\{x \in \mathbb{R}^d : |Tg(x)| > \lambda/2\}\big) + \lambda_d\big(\{x \in \mathbb{R}^d : |Tb(x)| > \lambda/2\}\big) =: I + II\end{aligned}$$

and we will estimate both addends I and II seperately.

As one would guess, the easy part is the one involving the good function, so we start with the estimate of I. Here it suffices to apply the Tchebychev inequality and the boundedness of T on $\mathrm{L}^2(\mathbb{R}^d)$ together with some properties of the good function. In detail it holds

$$\begin{aligned}I = \lambda_d\big(\{x \in \mathbb{R}^d : |Tg(x)| > \lambda/2\}\big) &\leq \frac{C}{\lambda^2}\left\|Tg\right\|^2_{\mathrm{L}^2(\mathbb{R}^d)} \leq \frac{CC^2_{\mathrm{L}^2}}{\lambda^2}\left\|g\right\|^2_{\mathrm{L}^2(\mathbb{R}^d)} \\ = \frac{CC^2_{\mathrm{L}^2}}{\lambda^2}\int_{\mathbb{R}^d}|g|^2 &\overset{\text{a)}}{\leq} \frac{CC^2_{\mathrm{L}^2}}{\lambda^2}\int_{\mathbb{R}^d}\gamma\lambda|g| = \frac{CC^2_{\mathrm{L}^2}\gamma}{\lambda}\|g\|_{\mathrm{L}^1(\mathbb{R}^d)} \overset{\text{b)}}{\leq} \frac{CC^2_{\mathrm{L}^2}\gamma}{\lambda}\left\|f\right\|_{\mathrm{L}^1(\mathbb{R}^d)}.\end{aligned}$$

Let's turn to the estimate of the bad function. As before, we denote the kernel of T by K. For every $j \in \mathbb{N}$ we denote the midpoint of the cube Q_j by x_j. For all $x \in \mathbb{R}^d \backslash 2\sqrt{d}Q_j$, we have $x \notin \operatorname{supp}(b_j)$ and thus, using property d) of the Calderón-Zygmund decomposition

$$\begin{aligned}Tb_j(x) &= \int_{\mathbb{R}^d} K(x-y)b_j(y)\,\mathrm{d}y - K(x-x_j)\int_{\mathbb{R}^d} b_j(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d}\big(K(x-y) - K(x-x_j)\big)b_j(y)\,\mathrm{d}y.\end{aligned}$$

This implies with Tonelli's Theorem and the support properties of b_j

$$\begin{aligned}\int_{\mathbb{R}^d\backslash 2\sqrt{d}Q_j}|Tb_j(x)|\,\mathrm{d}x &= \int_{\mathbb{R}^d\backslash 2\sqrt{d}Q_j}\left|\int_{\mathbb{R}^d}\big(K(x-y) - K(x-x_j)\big)b_j(y)\,\mathrm{d}y\right|\mathrm{d}x \\ &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d\backslash 2\sqrt{d}Q_j}\big|K(x-y) - K(x-x_j)\big|\,\mathrm{d}x|b_j(y)|\,\mathrm{d}y \\ &= \int_{Q_j}\int_{\mathbb{R}^d\backslash 2\sqrt{d}Q_j}\big|K(x-y) - K(x-x_j)\big|\,\mathrm{d}x|b_j(y)|\,\mathrm{d}y.\end{aligned}$$

Centering the inner integral around the origin, i.e. substituting $x' = x - x_j$, this may be rewritten as

$$= \int_{Q_j} \int_{(\mathbb{R}^d \setminus 2\sqrt{d}Q_j)-x_j} \big|K(x' - (y - x_j)) - K(x')\big| \,\mathrm{d}x' |b_j(y)| \,\mathrm{d}y.$$

If $y \in Q_j$ and $x' \in (\mathbb{R}^d \setminus 2\sqrt{d}Q_j) - x_j$, then $|x'| \geq \sqrt{d}\ell(Q_j)$ and $|y - x_j| \leq \frac{1}{2}\sqrt{d}\ell(Q_j)$, so $|x'| \geq \sqrt{d}\ell(Q_j) \geq 2|y - x_j|$. This implies

$$\leq \int_{Q_j} \int_{\{|x'| \geq 2|y - x_j|\}} \big|K(x' - (y - x_j)) - K(x')\big| \,\mathrm{d}x' |b_j(y)| \,\mathrm{d}y$$

and we are exactly in the situation of the Hörmander condition. This results in

$$\leq C_H \int_{Q_j} |b_j(y)| \,\mathrm{d}y = C_H \|b_j\|_{\mathrm{L}^1(\mathbb{R}^d)} \leq 2C_H \int_{Q_j} |f|,$$

where the last estimate stems on e).

We split up the estimate of II into the part inside the cubes $2\sqrt{d}Q_j$, $j \in \mathbb{N}$, and outside of these and apply the Tchebychev inequality to the second part:

$$\begin{aligned} II &= \lambda_d\big(\{x \in \mathbb{R}^d : |Tb(x)| > \lambda/2\}\big) \\ &\leq \lambda_d\left(\bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j\right) + \lambda_d\left(\left\{x \in \mathbb{R}^d \setminus \bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j : |Tb(x)| > \lambda/2\right\}\right) \\ &\leq \lambda_d\left(\bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j\right) + \frac{C}{\lambda} \int_{\mathbb{R}^d \setminus \bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j} |Tb|. \end{aligned}$$

Since

$$\lambda_d\left(\bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j\right) \leq \sum_{j \in \mathbb{N}} \lambda_d(2\sqrt{d}Q_j) \leq 2^d d^{d/2} \sum_{j \in \mathbb{N}} \lambda_d(Q_j) \overset{\text{f)}}{\leq} \frac{C}{\gamma\lambda} \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}$$

and since by the above calculations

$$\begin{aligned} \int_{\mathbb{R}^d \setminus \bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j} |Tb| &\leq \int_{\mathbb{R}^d \setminus \bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j} \sum_{k \in \mathbb{N}} |Tb_k| = \sum_{k \in \mathbb{N}} \int_{\mathbb{R}^d \setminus \bigcup_{j \in \mathbb{N}} 2\sqrt{d}Q_j} |Tb_k| \\ &\leq \sum_{k \in \mathbb{N}} \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q_k} |Tb_k| \leq 2C_H \sum_{k \in \mathbb{N}} \int_{Q_k} |f| \leq 2C_H \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}, \end{aligned}$$

we find

$$II \leq \frac{C}{\lambda}\left(\frac{1}{\gamma} + C_H\right) \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}.$$

Altogether, we found until now

$$\lambda_d\big(\{x \in \mathbb{R}^d : |Tf(x)| > \lambda\}\big) \leq \frac{C}{\lambda}\left(C_{\mathrm{L}^2}^2 \gamma + \frac{1}{\gamma} + C_H\right) \|f\|_{\mathrm{L}^1(\mathbb{R}^d)}.$$

Setting $\gamma = C_{L^2}^{-1}$ yields the estimate in (17.1). □

Having done the main work, we can now harvest.

Proof of Theorem 17.2. Let T be a Calderón-Zygmund operator. Then by definition T is of strong type $(2,2)$ with constant C_{L^2} and T is of weak type $(1,1)$ by Lemma 17.6, where the constant is given by $C(C_{\mathrm{L}^2}+C_H)$, cf. (17.1). The Marcinkiewicz interpolation result thus yields that T is of strong type (p,p) for all $p \in (1,2]$ with constant

$$\begin{aligned}
&2\left(\frac{p}{p-p_0}+\frac{p_1}{p_1-p}\right)^{1/p} C_0^{\frac{p_0}{p}(1-\frac{p-p_0}{p_1-p_0})} C_1^{\frac{p_1}{p}\frac{p-p_0}{p_1-p_0}} \\
&= 2\left(\frac{p}{p-1}+\frac{2}{2-p}\right)^{1/p} (C_{\mathrm{L}^2}+C_H)^{\frac{1}{p}(1-(p-1))} C_{\mathrm{L}^2}^{\frac{2}{p}(p-1)} \\
&= C(p)(C_{\mathrm{L}^2}+C_H)^{\frac{2-p}{p}} C_{\mathrm{L}^2}^{\frac{2p-2}{p}}.
\end{aligned}$$

Since T is a linear operator, this is already the claim for this range of p's.

Let $p > 2$ and consider the adjoint T^* of T in $L^2(\mathbb{R}^d)$. This is also a Calderón-Zygmund operator with the same constants C_H and C_{L^2} (exercise!), so it is bounded in $L^{p'}(\mathbb{R}^d)$ with constant $C(C_{\mathrm{L}^2}+C_H)^{\frac{2-p'}{p'}} C_{\mathrm{L}^2}^{\frac{2p'-2}{p'}}$. This means that $T = T^{**} = (T^*)^*$ is bounded in $L^p(\mathbb{R}^d)$ with the same constant. Since $\frac{2-p'}{p'} = \frac{p-2}{p}$ and $\frac{2p'-2}{p'} = \frac{2}{p}$, the claim follows. ☐

18 Fourier multipliers

In the whole space $\mathbb{R}^d$ the Fourier transform is a powerful tool to solve PDEs. If one looks for instance at the resolvent problem for the Laplace operator $\lambda u - \Delta u = f$ on $\mathbb{R}^d$, then applying the Fourier transfrom yields the algebraic equation $\lambda\hat{u}(\xi) + |\xi|^2\hat{u}(\xi) = \hat{f}(\xi)$, which leads to the formula

$$\hat{u}(\xi) = \frac{1}{\lambda + |\xi|^2}\hat{f}(\xi)$$

for $\hat{u}$. Applying the inverse Fourier transform, we formally get a representation of the solution u as

$$u = \mathcal{F}^{-1}\left(\frac{1}{\lambda + |\cdot|^2}\mathcal{F}f\right).$$

In this chapter we want to give calculations like this a precise meaning. We investigate operators of the form $Tf = \mathcal{F}^{-1}(m\mathcal{F}f)$ for given functions m. The main goal will be to know for which functions m the corresponding operator is continuous on $\mathrm{L}^p(\mathbb{R}^d)$, as this will be the key to an L^p-theory of elliptic problems like the one discussed above. Note that these operators are closely related to convolution operators, as by the usual rules for the Fourier transform it holds

$$Tf = \mathcal{F}^{-1}(m\mathcal{F}f) = (\mathcal{F}^{-1}m) * (\mathcal{F}^{-1}\mathcal{F}f) = K * f, \quad \text{where } K := \mathcal{F}^{-1}m.$$

So, one aim of this chapter will be to give criteria for the multiplier m such that the corresponding operator T gets a Calderón-Zygmund operator in order to apply our results of the preceding section.

Let $m \in \mathrm{L}^\infty(\mathbb{R}^d)$. Then for every $f \in \mathcal{S}(\mathbb{R}^d)$, we have $m\mathcal{F}f \in \mathrm{L}^\infty(\mathbb{R}^d)$, so this function is in particular in $\mathcal{S}'(\mathbb{R}^d)$ and we can perform the inverse Fourier transform, yielding some element of $\mathcal{S}'(\mathbb{R}^d)$. This justifies the following definition.

Definition 18.1. Let $m \in \mathrm{L}^\infty(\mathbb{R}^d)$. Then $T : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ given by $Tf := \mathcal{F}^{-1}(m\hat{f})$, $f \in \mathcal{S}(\mathbb{R}^d)$, is called Fourier multiplier and the function m is called symbol of the Fourier multiplier.

As usual the easy case is $p = 2$. Just by the Plancherel Thereom, every Fourier multiplier can be extended to a bounded operator in $L^2(\mathbb{R}^d)$: For all $f \in \mathcal{S}(\mathbb{R}^d)$ it holds

$$\left\|\mathcal{F}^{-1}m\hat{f}\right\|_{\mathrm{L}^2(\mathbb{R}^d)} = \left\|m\hat{f}\right\|_{\mathrm{L}^2(\mathbb{R}^d)} \leq \|m\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \left\|\hat{f}\right\|_{\mathrm{L}^2(\mathbb{R}^d)} = \|m\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \|f\|_{\mathrm{L}^2(\mathbb{R}^d)}.$$

For the other p's, the exendability question is much more involved. One obvious strategy is to link the Fourier multipliers to the Calderón-Zygmund operators, what we will do now.

Theorem 18.2 (Hörmander-Mikhlin multiplier Theorem). *Let $d_* := \lfloor \frac{d}{2} \rfloor + 1$ and let $m \in L^\infty(\mathbb{R}^d)$ be d_* times continuously differentiable on $\mathbb{R}^d \setminus \{0\}$. If*

$$\sup_{\xi \neq 0} |\xi|^{|\alpha|} |D^\alpha m(\xi)| =: M_\alpha < \infty \quad \text{for all } \alpha \in \mathbb{N}_0^d \text{ with } |\alpha| \leq d_*, \tag{18.1}$$

then the Fourier multiplier given by m, i.e. $Tf := \mathcal{F}^{-1} m \hat{f}$, $f \in \mathcal{S}(\mathbb{R}^d)$, is a Calderón-Zygmund operator. In particular it can be continuously extended to $\mathrm{L}^p(\mathbb{R}^d)$ for all $1 < p < \infty$. Finally, there is some constant C independent of m, such that for all $f \in \mathrm{L}^p(\mathbb{R}^d)$ it holds

$$\|Tf\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq CM \|f\|_{\mathrm{L}^p(\mathbb{R}^d)},$$

where $M := \sup_{|\alpha| \leq d_*} M_\alpha$.

Proof. We already saw that T can be extended to a bounded operator on $L^2(\mathbb{R}^d)$. In order to have a Calderón-Zygmund operator it remains to show, that $W := \mathcal{F}^{-1} m \in \mathcal{S}'(\mathbb{R}^d)$ on $\mathbb{R}^d \backslash \{0\}$ is given by an integral kernel K that satsfies the Hörmander condition.

First step: Getting started. Take some $\eta_0 \in \mathrm{C}_c^\infty(\mathbb{R})$ with $\eta_0 \geq 0$, $\operatorname{supp}(\eta_0) \subseteq [\frac{1}{2}, 2]$ and $\eta_0 = 1$ on $[\frac{3}{4}, \frac{7}{4}]$. For all $t \in \mathbb{R} \setminus \{0\}$ there is some $j \in \mathbb{Z}$ with $\frac{3}{4} \leq 2^{-j} t \leq \frac{7}{4}$, so $\sum_{j \in \mathbb{Z}} \eta_0(2^{-j} t) > 0$. Set

$$\eta(\xi) := \frac{\eta_0(|\xi|)}{\sum_{j \in \mathbb{Z}} \eta_0(2^{-j} |\xi|)}, \quad \xi \in \mathbb{R}^d \setminus \{0\}$$

and $\eta(0) = 0$. Then $\eta \in \mathrm{C}_c^\infty(\mathbb{R}^d)$, its support is contained in the annulus $\{\xi \in \mathbb{R}^d : \frac{1}{2} \leq |\xi| \leq 2\}$ and for all $\xi \in \mathbb{R}^d \setminus \{0\}$ we have

$$\begin{aligned} \sum_{k \in \mathbb{Z}} \eta(2^{-k} \xi) &= \sum_{k \in \mathbb{Z}} \frac{\eta_0(|2^{-k} \xi|)}{\sum_{j \in \mathbb{Z}} \eta_0(2^{-j} |2^{-k} \xi|)} = \sum_{k \in \mathbb{Z}} \frac{\eta_0(2^{-k} |\xi|)}{\sum_{j \in \mathbb{Z}} \eta_0(2^{-(j+k)} |\xi|)} \\ &= \sum_{k \in \mathbb{Z}} \frac{\eta_0(2^{-k} |\xi|)}{\sum_{\ell \in \mathbb{Z}} \eta_0(2^{-\ell} |\xi|)} = \frac{\sum_{k \in \mathbb{Z}} \eta_0(2^{-k} |\xi|)}{\sum_{\ell \in \mathbb{Z}} \eta_0(2^{-\ell} |\xi|)} = 1. \end{aligned} \tag{18.2}$$

Using this magic function η, for every $j \in \mathbb{Z}$ we define new functions m_j as $m_j(\xi) := m(\xi) \eta(2^{-j} \xi)$, $\xi \in \mathbb{R}^d$, and corresponding distributions $K_j := \check{m}_j$. As η has compact support away from the origin, m_j is in $\mathrm{C}^{d_*}(\mathbb{R}^d)$ and has compact support. In particular m_j is in $\mathrm{L}^1(\mathbb{R}^d)$, so K_j is even a bounded function for all $j \in \mathbb{Z}$.

We claim that $\sum_{j=-N}^N K_j$ for $N \to \infty$ converges to $\check{m}$ in $\mathcal{S}'(\mathbb{R}^d)$. Indeed for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ we have

$$\left\langle \sum_{j=-N}^N K_j, \varphi \right\rangle = \left\langle \sum_{j=-N}^N \check{m}_j, \varphi \right\rangle = \left\langle \sum_{j=-N}^N m_j, \check{\varphi} \right\rangle = \int_{\mathbb{R}^d} m(\xi) \sum_{j=-N}^N \eta(2^{-j} \xi) \check{\varphi}(\xi) \, \mathrm{d}\xi.$$

By (18.2) the integrand converges pointwise almost everywhere to $m\check{\varphi}$ and since $\check{\varphi}$ is a Schwartz function and the other functions are bounded, the Lebesgue Theorem gives convergence of the integral to $\langle m, \check{\varphi}\rangle = \langle \check{m}, \varphi\rangle$ for $N \to \infty$.

Second step: A useful estimate. We show that there is some constant $C \geq 0$, such that

$$\sup_{j\in\mathbb{Z}} \int_{\mathbb{R}^d} |K_j(x)|(1+2^j|x|)^{1/4}\,\mathrm{d}x \leq CM,$$

where M is as defined in the theorem.

Using the Cauchy-Schwarz inequality, we find for all $j \in \mathbb{Z}$

$$\begin{aligned}\int_{\mathbb{R}^d} |K_j(x)|(1+2^j|x|)^{1/4}\,\mathrm{d}x &= \int_{\mathbb{R}^d} |K_j(x)|(1+2^j|x|)^{d_*}(1+2^j|x|)^{1/4-d_*}\,\mathrm{d}x \\ &\leq \left(\int_{\mathbb{R}^d} |K_j(x)|^2(1+2^j|x|)^{2d_*}\,\mathrm{d}x\right)^{1/2}\left(\int_{\mathbb{R}^d}(1+2^j|x|)^{1/2-2d_*}\,\mathrm{d}x\right)^{1/2}.\end{aligned}$$

Since $\frac{1}{2} - 2d_* = \frac{1}{2} - 2\lfloor\frac{d}{2}\rfloor - 2 \leq \frac{1}{2} - 2(\frac{d}{2} - \frac{1}{2}) - 2 = -\frac{1}{2} - d < -d$, the second integral is finite and satisfies

$$\int_{\mathbb{R}^d}(1+2^j|x|)^{1/2-2d_*}\,\mathrm{d}x = \int_{\mathbb{R}^d}(1+|y|)^{1/2-2d_*}2^{-jd}\,\mathrm{d}y = C2^{-jd}.$$

In order to estimate the first integral we first note that

$$(1+2^j|x|)^{d_*} = \sum_{k=0}^{d_*}\binom{d_*}{k}2^{jk}|x|^k \leq C\sum_{|\gamma|\leq d_*}\left|(2^jx)^\gamma\right| = C\sum_{|\gamma|\leq d_*}2^{j|\gamma|}|x^\gamma|.$$

Putting this together yields

$$\begin{aligned}\int_{\mathbb{R}^d} |K_j(x)|(1+2^j|x|)^{1/4}\,\mathrm{d}x &\leq C2^{-jd/2}\left(\int_{\mathbb{R}^d}|K_j(x)|^2\Big(\sum_{|\gamma|\leq d_*}2^{j|\gamma|}|x^\gamma|\Big)^2\,\mathrm{d}x\right)^{1/2} \\ &= C2^{-jd/2}\left\|x\mapsto \sum_{|\gamma|\leq d_*}2^{j|\gamma|}K_j(x)x^\gamma\right\|_{\mathrm{L}^2(\mathbb{R}^d)} \\ &\leq C2^{-jd/2}\sum_{|\gamma|\leq d_*}2^{j|\gamma|}\left\|x\mapsto K_j(x)x^\gamma\right\|_{\mathrm{L}^2(\mathbb{R}^d)}.\end{aligned}$$

Using properties of the Fourier transform and the Plancherel Theorem, we rewrite this last norm:

$$\begin{aligned}\left\|x\mapsto K_j(x)x^\gamma\right\|_{\mathrm{L}^2(\mathbb{R}^d)} &= \left\|x\mapsto x^\gamma\check{m}_j(x)\right\|_{\mathrm{L}^2(\mathbb{R}^d)} = \left\|\mathcal{F}^{-1}(D^\gamma m_j)\right\|_{\mathrm{L}^2(\mathbb{R}^d)} \\ &= \left\|D^\gamma m_j\right\|_{\mathrm{L}^2(\mathbb{R}^d)}.\end{aligned}$$

The next task is to estimate this remaining norm and to link it to the hypotheses on m. In order to do so, we plug in the definition of m_j and apply the Leibniz rule:

$$\begin{aligned}\int_{\mathbb{R}^d} \left|D^\gamma m_j(\xi)\right|^2 \mathrm{d}\xi &= \int_{\mathbb{R}^d} \left|D^\gamma\big(m\eta(2^{-j}\cdot)\big)(\xi)\right|^2 \mathrm{d}\xi \\ &= \int_{\mathbb{R}^d} \left|\sum_{\alpha\le\gamma}\binom{\gamma}{\alpha} D^\alpha m(\xi) D^{\gamma-\alpha}\eta(2^{-j}\xi)2^{-j|\gamma-\alpha|}\right|^2 \mathrm{d}\xi \\ &\le \int_{\mathbb{R}^d} \left(\sum_{\alpha\le\gamma}\binom{\gamma}{\alpha} |D^\alpha m(\xi)|\left|D^{\gamma-\alpha}\eta(2^{-j}\xi)\right|2^{-j|\gamma|}2^{j|\alpha|}\right)^2 \mathrm{d}\xi.\end{aligned}$$

Next, we substitute $\tau = 2^{-j}\xi$.

$$= \int_{\mathbb{R}^d} \left(\sum_{\alpha\le\gamma}\binom{\gamma}{\alpha} |D^\alpha m(2^j\tau)|\left|D^{\gamma-\alpha}\eta(\tau)\right|2^{-j|\gamma|}2^{j|\alpha|}\right)^2 2^{jd}\,\mathrm{d}\tau.$$

Investing the support properties of η and the hypotheses on m, we get

$$\begin{aligned}&\le 2^{jd}\int_{\frac{1}{2}\le|\tau|\le 2}\left(\sum_{\alpha\le\gamma}\binom{\gamma}{\alpha} M_\alpha |2^j\tau|^{-|\alpha|}\left\|D^{\gamma-\alpha}\eta\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)} 2^{-j|\gamma|}2^{j|\alpha|}\right)^2 \mathrm{d}\tau \\ &\le CM^2 2^{jd}\int_{\frac{1}{2}\le|\tau|\le 2}\left(\sum_{\alpha\le\gamma}\binom{\gamma}{\alpha}|\tau|^{-|\alpha|}2^{-j|\gamma|}\right)^2 \mathrm{d}\tau.\end{aligned}$$

Finally $|\tau|^{-|\alpha|} \le (1/2)^{-|\alpha|} = 2^{|\alpha|} \le 2^{|\gamma|} \le 2^{d_*}$ can be put into the constant, together with the remaining binomial coefficients and the volume of the annulus. This leaves us with

$$\le CM^2 2^{jd} 2^{-2j|\gamma|}.$$

Putting everything together we have proved

$$\begin{aligned}\int_{\mathbb{R}^d} |K_j(x)|(1+2^j|x|)^{1/4}\,\mathrm{d}x &\le C2^{-jd/2}\sum_{|\gamma|\le d_*} 2^{j|\gamma|}\left\|D^\gamma m_j\right\|_{\mathrm{L}^2(\mathbb{R}^d)} \\ &\le C2^{-jd/2}\sum_{|\gamma|\le d_*} 2^{j|\gamma|} M 2^{jd/2}2^{-j|\gamma|} = CM\end{aligned}$$

for every $j \in \mathbb{Z}$.

Third step: The same procedure again. Doing analogous calculations for the partial derivatives of K_j, one finds

$$\sup_{j\in\mathbb{Z}} 2^{-j}\int_{\mathbb{R}^d}\left|\nabla K_j(x)\right|(1+2^j|x|)^{1/4}\,\mathrm{d}x \le CM.$$

The only point, where one has to pay attention, is when applying the Plancherel Theorem. There one has for the k'th partial derivative $\partial_k K_j(x)x^\gamma = \partial_k \check{m}_j(x)x^\gamma$, which via Plancherel leads to the norm of $D^\gamma(\xi_k m_j(\xi)) = D^\gamma(m(\xi)\xi_k\eta(2^{-j}\xi))$. For the rest of the argument, the ξ_k is then grouped together with the function η and when performing the substitution that introduces τ, the additional factor 2^{-j} pops up.

Fourth step: Definition of the kernel K. We show that for almost all $x \in \mathbb{R}^d \setminus \{0\}$ the series $\sum_{j\in\mathbb{Z}} K_j(x)$ is absolutely convergent. For all $x \in \mathbb{R}^d$, it holds

$$
\begin{aligned}
|K_j(x)| = |\check{m}_j(x)| &= \left|\mathcal{F}^{-1}\big(m(\xi)\eta(2^{-j}\xi)\big)\right| = \left|\int_{\mathbb{R}^d} \mathrm{e}^{2\pi \mathrm{i} x\xi} m(\xi)\eta(2^{-j}\xi)\,\mathrm{d}\xi\right| \\
&\le \|m\|_{L^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} \left|\eta(2^{-j}\xi)\right| \mathrm{d}\xi = \|m\|_{L^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} |\eta(\tau)| 2^{jd}\,\mathrm{d}\tau \\
&= C 2^{jd} \|m\|_{L^\infty(\mathbb{R}^d)}\,.
\end{aligned}
$$

So, for $j \to -\infty$ the sequence $(|K_j(x)|)$ is majorised by a geometric series, which means that $\sum_{j\le 0} K_j(x)$ is absolutely convergent.

We now deal with the positive summation indices j. For this, we note that thanks to the estimate in the second step we have for every $\delta > 0$

$$
(1+2^j\delta)^{1/4} \int_{|x|\ge\delta} |K_j(x)|\,\mathrm{d}x \le \int_{|x|\ge\delta} |K_j(x)|(1+2^j|x|)^{1/4}\,\mathrm{d}x \le CM.
$$

This means that for all $\delta > 0$

$$
\int_{|x|\ge\delta} \sum_{j>0} |K_j(x)|\,\mathrm{d}x = \sum_{j>0} \int_{|x|\ge\delta} |K_j(x)|\,\mathrm{d}x \le CM \sum_{j>0} \frac{1}{(1+2^j\delta)^{1/4}} < C_\delta M,
$$

since the last sum converges for every $\delta > 0$. This implies that $\sum_{j>0} |K_j|$ is in $\mathrm{L}^1(\mathbb{R}^d \setminus B_\delta(0))$ for every $\delta > 0$. Being in L^1 means in particular that the function is almost everywhere finite. Shrinking δ to zero, this entrails that $\sum_{j>0} |K_j|$ is almost everywhere finite. Altogether we have proved that $\sum_{j\in\mathbb{Z}} K_j(x)$ is abolsutely convergent almost everywhere in $\mathbb{R}^d$. This allows us to define $K := \sum_{j\in\mathbb{Z}} K_j$ and from the results of the first step we know that $K = \check{m}$ in $\mathcal{S}'(\mathbb{R}^d)$. Additionally, we get $K \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \{0\})$ by the above calculation. However, note that for the negative indices, the series $\sum_{j\le 0} |K_j(x)|$ was finite for every $x \in \mathbb{R}^d$, so it is in $\mathrm{L}^\infty(\mathbb{R}^d)$ which implies that it is even in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$.

Fifth step: K satisfies the Hörmander condition. For all $y \in \mathbb{R}^d \setminus \{0\}$ we estimate

$$
\begin{aligned}
\int_{|x|\ge 2|y|} \left|K(x-y) - K(x)\right| \mathrm{d}x &= \int_{|x|\ge 2|y|} \left|\sum_{j\in\mathbb{Z}} \big(K_j(x-y) - K_j(x)\big)\right| \mathrm{d}x \\
&\le \sum_{j\in\mathbb{Z}} \int_{|x|\ge 2|y|} \left|K_j(x-y) - K_j(x)\right| \mathrm{d}x.
\end{aligned}
$$

We choose $J \in \mathbb{Z}$ with the property $2^{-J} \leq |y| < 2^{-J+1}$ and we split the sum over j into the parts with $j > J$ and $j \leq J$, respectively.

For the large indices we use a brute force triangular inequality:

$$\begin{aligned}\sum_{j>J} \int_{|x|\geq 2|y|} \big|K_j(x-y) - K_j(x)\big| \, \mathrm{d}x &\leq \sum_{j>J} \int_{|x|\geq 2|y|} \big(|K_j(x-y)| + |K_j(x)|\big) \, \mathrm{d}x \\ &= \sum_{j>J} \left(\int_{|x|\geq 2|y|} |K_j(x-y)| \, \mathrm{d}x + \int_{|x|\geq 2|y|} |K_j(x)| \, \mathrm{d}x \right).\end{aligned}$$

As $|x| \geq 2|y|$, for the difference $x - y$, we still have $|x - y| \geq |y|$, so we can enlarge both integrals to get

$$\leq \sum_{j>J} 2 \int_{|z|\geq |y|} |K_j(z)| \, \mathrm{d}z.$$

In view of the estimate in the second step we calculate

$$\begin{aligned} &= 2 \sum_{j>J} \int_{|z|\geq|y|} \frac{|K_j(z)|(1+2^j|z|)^{1/4}}{(1+2^j|z|)^{1/4}} \, \mathrm{d}z \\ &\leq 2 \sum_{j>J} \frac{1}{(1+2^j|y|)^{1/4}} \int_{|z|\geq|y|} |K_j(z)|(1+2^j|z|)^{1/4} \, \mathrm{d}z \\ &\leq CM \sum_{j>J} \frac{1}{(1+2^j|y|)^{1/4}} \\ &\leq CM \sum_{j>J} \frac{1}{(1+2^j 2^{-J})^{1/4}} \\ &= CM \sum_{k>0} \frac{1}{(1+2^k)^{1/4}} = CM, \end{aligned}$$

independently of y.

For the indices $j \leq J$, we base our calculations on the estimate in the third step. In order to do so, we let appear a gradient of K_j by writing the difference as an integral:

$$\begin{aligned}\sum_{j\leq J} \int_{|x|\geq 2|y|} \big|K_j(x-y) - K_j(x)\big| \, \mathrm{d}x &= \sum_{j\leq J} \int_{|x|\geq 2|y|} \left| \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\big(K_j(x-ty)\big) \, \mathrm{d}t \right| \mathrm{d}x \\ &\leq \sum_{j\leq J} \int_{|x|\geq 2|y|} \int_0^1 \big|(-y) \cdot \nabla K_j(x-ty)\big| \, \mathrm{d}t \, \mathrm{d}x.\end{aligned}$$

We use that $|y| \leq 2^{-J+1}$ and $(1+2^j|x-ty|)^{1/4} \geq 1$, we forget about the restriction on x and, finally, we apply the estimate from the third step:

$$\leq \int_0^1 \sum_{j \leq J} \int_{\mathbb{R}^d} 2^{-J+1} \big|\nabla K_j(x-ty)\big| (1+2^j|x-ty|)^{1/4} \,\mathrm{d}x\,\mathrm{d}t$$
$$\leq CM \int_0^1 \sum_{j \leq J} 2^{-J+1} 2^j \,\mathrm{d}t = CM \sum_{k \leq 0} 2^{k+1} = CM.$$

Altogether this proves the Hörmander condition for K with $C_H = CM$.

Sixth step: The norm estimate. We now know that T is a Calderón-Zygmund operator with L^2 norm bound $C_{\mathrm{L}^2} = M$ and its constant C_H in the Hörmander condition is also given by M. So, by Theorem 17.2 its L^p norm bound for $1 < p \leq 2$ is given by

$$C(C_{\mathrm{L}^2} + C_H)^{\frac{2-p}{p}} C_{\mathrm{L}^2}^{\frac{2p-2}{p}} = C(2M)^{\frac{2-p}{p}} M^{\frac{2p-2}{p}} = CM^{\frac{2-p+2p-2}{p}} = CM$$

with a constant C that is independent of the multiplier m.

For $p > 2$ we get the same:

$$C(C_{\mathrm{L}^2} + C_H)^{\frac{p-2}{p}} C_{\mathrm{L}^2}^{\frac{2}{p}} = CM^{\frac{p-2+2}{p}} = CM. \qquad \square$$

Example 18.3. We consider again the symbol that was already mentionned in the beginning of this chapter. For $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$ we set $m_\lambda(\xi) := \frac{1}{\lambda + |\xi|^2}$, $\xi \in \mathbb{R}^d$. We show that this symbol and some of its close relatives fulfil the Mikhlin condition (18.1) and thus give rise to Fourier multipliers that are bounded on all L^p spaces.

We first note that $\mathrm{Re}(\lambda + |\xi|^2) = \mathrm{Re}(\lambda) + |\xi|^2$ and that both contributions on the right hand side are positive. Thus

$$\big|\lambda + |\xi|^2\big| = \sqrt{(\mathrm{Re}(\lambda) + |\xi|^2)^2 + \mathrm{Im}(\lambda)^2}$$

can be estimated from below by $|\lambda|$, as well as by $|\xi|^2$.

Let $a \in \{0,1,2\}$. We have for all $|\xi| \leq \sqrt{|\lambda|}$, using the estimate from below by $|\lambda|$,

$$|\xi|^a |m_\lambda(\xi)| = \frac{|\xi|^a}{\big|\lambda + |\xi|^2\big|} \leq \frac{\sqrt{|\lambda|}^a}{|\lambda|} = \frac{1}{|\lambda|^{1-\frac{a}{2}}}.$$

Using the estimate from below by $|\xi|^2$ and oberving that $2 - a \geq 0$, we find for $|\xi| \geq \sqrt{|\lambda|}$

$$|\xi|^a |m_\lambda(\xi)| = \frac{|\xi|^a}{\big|\lambda + |\xi|^2\big|} \leq \frac{|\xi|^a}{|\xi|^2} = \frac{1}{|\xi|^{2-a}} \leq \frac{1}{|\lambda|^{1-\frac{a}{2}}}.$$

Consequently, this estimate is true for every $\xi \in \mathbb{R}^d$.

In order to verify the Mikhlin condition (18.1) we have to estimate arbitrary derivatives of m_λ. In doing so, we will use the following calculation. Let $\ell \in \{1, 2, \dots, d\}$. Then

$$\partial_\ell m_\lambda(\xi) = \partial_\ell \left(\frac{1}{\lambda + |\xi|^2} \right) = \frac{-2\xi_\ell}{(\lambda + |\xi|^2)^2} = -2\xi_\ell m_\lambda(\xi)^2. \tag{18.3}$$

Having this at hand, we can prove the following claim.

Claim: Let $\alpha \in \mathbb{N}_0^d$ be a multiindex with $|\alpha| \leq 2$. For all $\beta \in \mathbb{N}_0^d$ the function $D^\beta(\xi^\alpha m_\lambda(\xi))$ is a linear combination of terms of the form

$$\xi^\gamma m_\lambda(\xi)^{r+1}$$

with a multiindex $\gamma \in \mathbb{N}_0^d$ and $r \in \mathbb{N}_0$ such that $2r - |\gamma| = |\beta| - |\alpha|$.

We do an induction over the length of the multiindex β. For $|\beta| = 0$ everything is fine with $\gamma = \alpha$ and $r = 0$. So let $\beta \in \mathbb{N}_0^d$ be some multiindex, let $\ell \in \{1, 2, \dots, d\}$ be given and consider $\tilde{\beta} = \beta + \mathrm{e}_\ell$, where e_ℓ is the ℓth unit vector. Then $|\tilde{\beta}| = |\beta| + 1$ and by the induction hypotheses $D^{\tilde{\beta}}(\xi^\alpha m_\lambda(\xi)) = \partial_\ell D^\beta(\xi^\alpha m_\lambda(\xi))$ is a linear combination of terms of the form

$$\partial_\ell\big(\xi^\gamma m_\lambda(\xi)^{r+1}\big)$$

We use the product rule to rewrite this as

$$\partial_\ell\big(\xi^\gamma m_\lambda(\xi)^{r+1}\big) = \partial_\ell(\xi^\gamma) m_\lambda(\xi)^{r+1} + \xi^\gamma (r+1) m_\lambda(\xi)^r \partial_\ell m_\lambda(\xi)$$

If $\gamma_\ell = 0$ the first summand vanishes, so in this case, we do not have to care about it. If $\gamma_\ell > 0$ we continue with the help of (18.3):

$$\begin{aligned} &= \gamma_\ell \xi^{\gamma - \mathrm{e}_\ell} m_\lambda(\xi)^{r+1} - 2(r+1)\xi^\gamma m_\lambda(\xi)^r \xi_\ell m_\lambda(\xi)^2 \\ &= \gamma_\ell \xi^{\gamma - \mathrm{e}_\ell} m_\lambda(\xi)^{r+1} - 2(r+1)\xi^{\gamma + \mathrm{e}_\ell} m_\lambda(\xi)^{r+2} \end{aligned}$$

This is again a linear combination of terms that fulfil the claim for $\tilde{\beta}$, since for the first

$$2r - |\gamma - \mathrm{e}_\ell| = 2r - |\gamma| + 1 = |\beta| - |\alpha| + 1 = |\tilde{\beta}| - |\alpha|$$

and for the second

$$2(r+1) - |\gamma + \mathrm{e}_\ell| = 2r + 2 - |\gamma| - 1 = |\beta| + 1 - |\alpha| = |\tilde{\beta}| - |\alpha|.$$

With this claim proved, we can now show that for every $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq 2$ the function $\xi^\alpha m_\lambda(\xi)$ satisfies the Mikhlin condition. For every multiindex $\beta \in \mathbb{N}_0^d$ the claim tells us that $|\xi|^{|\beta|}\big|D^\beta(\xi^\alpha m_\lambda(\xi))\big|$ is a linear combination of terms of the form $|\xi|^{|\beta|}\big|\xi^\gamma m_\lambda(\xi)^{r+1}\big|$ with $2r - |\gamma| = |\beta| - |\alpha|$. Each of these terms may be estimated by

$$\begin{aligned} |\xi|^{|\beta|}\big|\xi^\gamma m_\lambda(\xi)^{r+1}\big| &\leq |\xi|^{|\beta|}|\xi|^{|\gamma|}|m_\lambda(\xi)|^{r+1}| = |\xi|^{|\beta|+|\gamma|}|m_\lambda(\xi)|^{r+1} \\ &= |\xi|^{2r+|\alpha|}|m_\lambda(\xi)|^r|m_\lambda(\xi)| = \big(|\xi|^2|m_\lambda(\xi)|\big)^r|\xi|^{|\alpha|}|m_\lambda(\xi)| \\ &\leq 1^r \frac{1}{|\lambda|^{1-\frac{|\alpha|}{2}}} = \frac{1}{|\lambda|^{1-\frac{|\alpha|}{2}}}, \end{aligned}$$

where in the last step we used the estimate obtained in the beginning of this example.

We show an application of this example to some function spaces.

Definition 18.4. Let $s > 0$ and $p \in (1, \infty)$ and consider the polynomial $p(\xi) := 1 + |\xi|^2$, $\xi \in \mathbb{R}^d$. The space

$$H^{s,p}(\mathbb{R}^d) := \left\{ f \in \mathrm{L}^p(\mathbb{R}^d) : \mathcal{F}^{-1}(p^{s/2}\hat{f}) \in \mathrm{L}^p(\mathbb{R}^d) \right\}$$

is called Bessel potential space of order s. The corresponding norm is

$$\|f\|_{H^{s,p}(\mathbb{R}^d)} := \left\| \mathcal{F}^{-1}(p^{s/2}\hat{f}) \right\|_{\mathrm{L}^p(\mathbb{R}^d)}.$$

It is not too difficult to show, that the so defined Bessel potential spaces are Banach spaces. We will use our knowledge from the foregoing example to show a relationship to the Sobolev spaces.

Proposition 18.5. $H^{2,p}(\mathbb{R}^d) = W^{2,p}(\mathbb{R}^d)$ *with equivalent norms.*

Proof. For the Bessel potential space we are in the special case $s = 2$, so the exponent of p is just 1.

"$\supseteq$" Let $f \in W^{2,p}(\mathbb{R}^d)$. Then $(1 - \Delta)f \in \mathrm{L}^p(\mathbb{R}^d)$ and since

$$p(\xi)\hat{f}(\xi) = (1 + |\xi|^2)\hat{f}(\xi) = \mathcal{F}((1 - \Delta)f)(\xi),$$

we have

$$\mathcal{F}^{-1}(p\hat{f}) = (1 - \Delta)f \in \mathrm{L}^p(\mathbb{R}^d)$$

and

$$\begin{aligned} \|f\|_{H^{2,p}(\mathbb{R}^d)} &= \left\| \mathcal{F}^{-1}(p\hat{f}) \right\|_{\mathrm{L}^p(\mathbb{R}^d)} = \|(1 - \Delta)f\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\leq \|f\|_{\mathrm{L}_p(\mathbb{R}^d)} + \sum_{j=1}^{d} \left\| \partial_j^2 f \right\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq C \|f\|_{W^{2,p}(\mathbb{R}^d)}. \end{aligned}$$

"$\subseteq$" Let $f \in H^{2,p}(\mathbb{R}^d)$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq 2$. Then with the notation of the foregoing example

$$\begin{aligned} \|D^\alpha f\|_{\mathrm{L}^p(\mathbb{R}^d)} &= \left\| \mathcal{F}^{-1}\xi^\alpha \hat{f} \right\|_{\mathrm{L}^p(\mathbb{R}^d)} = \left\| \mathcal{F}^{-1} \frac{\xi^\alpha}{p(\xi)} p(\xi)\hat{f} \right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &= \left\| \mathcal{F}^{-1}\xi^\alpha m_1(\xi)\mathcal{F}\mathcal{F}^{-1}(p\hat{f}) \right\|_{\mathrm{L}^p(\mathbb{R}^d)}. \end{aligned}$$

Since f is in the Bessel potential space of order 2 the function $\mathcal{F}^{-1}(p\hat{f})$ is in $\mathrm{L}^p(\mathbb{R}^d)$. Furthermore, $\xi^\alpha m_1(\xi)$ fulfills the Mikhlin condition, so the corresponding Fourier multiplier is a bounded operator on $\mathrm{L}^p(\mathbb{R}^d)$. This implies

$$\|D^\alpha f\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq C \left\| \mathcal{F}^{-1}(p\hat{f}) \right\|_{\mathrm{L}^p(\mathbb{R}^d)} = C \|f\|_{H^{s,p}(\mathbb{R}^d)} .$$

Consequently, $f \in W^{2,p}(\mathbb{R}^d)$ and the $W^{2,p}(\mathbb{R}^d)$-norm can be controlled by the $H^{2,p}(\mathbb{R}^d)$-norm. □

Remark 18.6. a) This is only a special case of a more general result. In fact for all $k \in \mathbb{N}$ it holds $H^{k,p}(\mathbb{R}^d) = W^{k,p}(\mathbb{R}^d)$. For even k this can be shown more or less like above, considering the symbol $\xi^\alpha m_1(\xi)^{k/2}$ that turns out to satisfy the Mikhlin condition for all $|\alpha| \leq k$. For odd k the prove is more involved, mainly because then $(1 + |\xi|^2)^{k/2}$ is not a polynomial.

b) For $\Omega \subseteq \mathbb{R}^d$ one defines the corresponding Bessel potential space by restriction as

$$H^{s,p}(\Omega) := \left\{ f|_\Omega : f \in H^{s,p}(\mathbb{R}^d) \right\}$$

with the quotient norm

$$\|f\|_{H^{s,p}(\Omega)} = \inf\left\{ \|g\|_{H^{s,p}(\mathbb{R}^d)} : g \in H^{s,p}(\mathbb{R}^d) \text{ with } f|_\Omega = g \right\}.$$

This also is a Banach space and as long as there is a continuous extension operator from $W^{k,p}(\Omega)$ to $W^{k,p}(\mathbb{R}^d)$ the equality $H^{k,p}(\Omega) = W^{k,p}(\Omega)$ remains valid. In general this is false!

19 Elliptic boundary value problems in $\mathrm{L}^p(\Omega)$

Let $\Omega \subseteq \mathbb{R}^d$ be open and bounded with a C^2-boundary. On Ω we want to solve a second-order elliptic problem in non-divergence form with Dirichlet boundary conditions, where the coefficients are bounded measurable functions and the coefficients of the leading order part are uniformly continuous up to the boundary. That is, we consider coefficient functions $a \in C(\overline{\Omega}, \mathbb{R}^{d\times d})$, $b \in L^\infty(\Omega, \mathbb{R}^d)$ and $c \in L^\infty(\Omega, \mathbb{R})$ such that a is symmetric and fulfills the following ellipticity condition: There exists some $\kappa_0 > 0$, called the ellipticity constant, such that

$$\xi^T a(x)\xi \geq \kappa_0 |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d \text{ and all } x \in \overline{\Omega}.$$

With these coefficients we define the formal differential operator

$$Lu(x) := \sum_{j,k=1}^{d} a_{jk}(x)\partial_j\partial_k u(x) + \sum_{j=1}^{d} b_j(x)\partial_j u(x) + c(x)u(x)$$

and for a given $f \in \mathrm{L}^p(\Omega)$ we aim at solutions $u \in W^{2,p}(\Omega)$ of

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

for suitable values of λ.

We start with easier problems and consider first the case $\Omega = \mathbb{R}^d$ (no boundary condition), $b = 0$, $c = 0$ and a constant. So for the time being let $a = (\alpha_{jk})_{jk} \in \mathbb{R}^{d\times d}$ be a positive definite matrix. Then we define the unbounded operator $(A_{p,\mathbb{R}^d}, D(A_{p,\mathbb{R}^d}))$ in $\mathrm{L}^p(\mathbb{R}^d)$ with $D(A_{p,\mathbb{R}^d}) = W^{2,p}(\mathbb{R}^d)$ and

$$A_{p,\mathbb{R}^d} u(x) := Au(x) := \sum_{j,k=1}^{d} \alpha_{jk}\partial_j\partial_k u(x)$$

for $u \in D(A_{p,\mathbb{R}^d})$.

Proposition 19.1. *For all $f \in \mathrm{L}^p(\mathbb{R}^d)$ and all $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$ there is a unique $u \in W^{2,p}(\mathbb{R}^d)$ such that $\lambda u - A_{p,\mathbb{R}^d} u = f$ and for all $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq 2$ it holds*

$$\left\| D^\beta u \right\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \left\| f \right\|_{\mathrm{L}^p(\mathbb{R}^d)},$$

where the constant C depends on the matrix a only via its norm and the ellipticity constant.

Remark 19.2. The norm estimate in the above proposition means that for large $|\lambda|$

$$\begin{aligned}\left\|R(\lambda, A_{p,\mathbb{R}^d})\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d))} &\leq \frac{C}{|\lambda|},\\ \left\|R(\lambda, A_{p,\mathbb{R}^d})\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d),W^{1,p}(\mathbb{R}^d))} &\leq \frac{C}{\sqrt{|\lambda|}} \quad \text{and}\\ \left\|R(\lambda, A_{p,\mathbb{R}^d})\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d),W^{2,p}(\mathbb{R}^d))} &\leq C.\end{aligned}$$

This is a typical behaviour for nice elliptic operators of second order.

Proof of Proposition 19.1. We do this proof only for $A = \Delta$, i.e. for a being the identity matrix. If $u \in W^{2,p}(\mathbb{R}^d)$ solves $(\lambda - \Delta)u = f$, then by applying the Fourier transform one finds $\lambda\hat{u} + |\cdot|^2\hat{u} = \hat{f}$ or equivalently

$$\hat{u} = \frac{1}{\lambda + |\cdot|^2}\hat{f} = m_\lambda \hat{f}$$

with the notation m_λ from Example 18.3. This leads us to set $u := \mathcal{F}^{-1}(m_\lambda \hat{f})$. If we can show that for this u we have $u \in W^{2,p}(\mathbb{R}^d)$, then $u \in D(A_{p,\mathbb{R}^d})$ and

$$\lambda u - A_{p,\mathbb{R}^d}u = (\lambda - \Delta)\mathcal{F}^{-1}(m_\lambda \hat{f}) = \mathcal{F}^{-1}((\lambda + |\cdot|^2)m_\lambda \hat{f}) = \mathcal{F}^{-1}\hat{f} = f,$$

so we have indeed found a solution. Uniqueness follows by the above considerations and the injectivity of the Fourier transform.

In Example 18.3 we have seen that m_λ fulfills the Mikhlin condition, so the corresponding Fourier multiplier is continuous on $\mathrm{L}^p(\mathbb{R}^d)$. Hence, since f is in $\mathrm{L}^p(\mathbb{R}^d)$, we know that $u \in \mathrm{L}^p(\mathbb{R}^d)$ as well. Even more is true, since for all multiindices $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq 2$ the same example shows that

$$D^\beta u = D^\beta \mathcal{F}^{-1}(m_\lambda \hat{f}) = i^{|\beta|}\mathcal{F}^{-1}(\xi^\beta m_\lambda \hat{f}) \in \mathrm{L}^p(\mathbb{R}^d),$$

we know that $u \in W^{2,p}(\mathbb{R}^d)$ and, finally, the corresponding norm estimates in this example yield

$$\left\|D^\beta u\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq \left\|\mathcal{F}^{-1}\xi^\beta m_\lambda \mathcal{F}\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d))} \|f\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq \frac{1}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d)}. \qquad \square$$

To include general positive definite matrices a one has to adapt the estimates in Example 18.3. This can be done along the same lines, but it is definitly more tedious. Doing these estimates one finds that the appearing constant C in the proposition depends on the matrix a only via its norm and the ellipticity constant.

In a next step we include the lower order terms and small deviations from constant matrices in the highest order terms by a perturbation argument. Let $(\alpha_{jk})_{jk} \in \mathbb{R}^{d\times d}$ be some positive definite matrix and define A as above. Then we split up our differential operator L in the following way:

$$
\begin{aligned}
Lu(x) &= \sum_{j,k=1}^{d} a_{jk}(x)\partial_j\partial_k u(x) + \sum_{j=1}^{d} b_j(x)\partial_j u(x) + c(x)u(x) \\
&= \sum_{j,k=1}^{d} \alpha_{jk}\partial_j\partial_k u(x) + \sum_{j,k=1}^{d} (a_{jk}(x) - \alpha_{jk})\partial_j\partial_k u(x) + \sum_{j=1}^{d} b_j(x)\partial_j u(x) + c(x)u(x) \\
&= Au(x) + Bu(x).
\end{aligned}
$$

Corresponding to this splitting we consider the corresponding unbounded differential operators $L_{p,\mathbb{R}^d}$, $A_{p,\mathbb{R}^d}$ and $B_{p,\mathbb{R}^d}$ on $\mathrm{L}^p(\mathbb{R}^d)$ with $D(A_{p,\mathbb{R}^d}) = D(B_{p,\mathbb{R}^d}) = D(L_{p,\mathbb{R}^d}) = W^{2,p}(\mathbb{R}^d)$. For these we have the following result.

Proposition 19.3. *There exists some $\varepsilon > 0$ and some $\lambda_0 > 0$ such that the following holds: If $a \in \mathrm{L}^\infty(\mathbb{R}^d, \mathbb{R}^{d\times d})$ satisfies $\|a_{jk} - \alpha_{jk}\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \le \varepsilon$ for all $j,k = 1,\dots,d$ and $\lambda \in \mathbb{C}$ satsifies $\mathrm{Re}(\lambda) \ge \lambda_0$, then for every $f \in \mathrm{L}^p(\mathbb{R}^d)$ the equation*

$$(\lambda - L_{p,\mathbb{R}^d})u = (\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d})u = f$$

has a unique solution $u \in W^{2,p}(\mathbb{R}^d)$ with

$$\left\|D^\beta u\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \le \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d)}$$

for all multiindices $\beta \in \mathbb{N}_0^d$ with $|\beta| \le 2$.

Proof. Let $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$ be given. We want to find a continuous inverse of $\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d}$. It is easily shown that $\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d}$ is a continuous linear map from $W^{2,p}(\mathbb{R}^d)$ to $\mathrm{L}^p(\mathbb{R}^d)$. Furthermore, from Proposition 19.1 we know that $\lambda - A_{p,\mathbb{R}^d}$ is continuously invertible in $\mathrm{L}^p(\mathbb{R}^d)$ with

$$\left\|D^\beta(\lambda - A_{p,\mathbb{R}^d})^{-1}\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d))} \le \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \tag{19.1}$$

for all $|\beta| \le 2$. This yields for all $u \in W^{2,p}(\mathbb{R}^d)$

$$(\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d})u = \big(\mathrm{I} - B_{p,\mathbb{R}^d}(\lambda - A_{p,\mathbb{R}^d})^{-1}\big)(\lambda - A_{p,\mathbb{R}^d})u.$$

If we can guarantee that

$$\left\|B_{p,\mathbb{R}^d}(\lambda - A_{p,\mathbb{R}^d})^{-1}\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d))} \le 1/2, \tag{19.2}$$

the Neumann series gives us that $\mathrm{I} - B_{p,\mathbb{R}^d}(\lambda - A_{p,\mathbb{R}^d})^{-1}$ is invertible in $\mathrm{L}^p(\mathbb{R}^d)$ with the norm of the inverse bounded by 2 and this means that

$$(\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d})^{-1} = (\lambda - A_{p,\mathbb{R}^d})^{-1}\big(\mathrm{I} - B_{p,\mathbb{R}^d}(\lambda - A_{p,\mathbb{R}^d})^{-1}\big)^{-1}.$$

as a continuous inverse in $\mathrm{L}^p(\mathbb{R}^d)$. Furthermore, for $u \coloneqq (\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d})^{-1} f$ we then get for every $|\beta| \le 2$ the norm estimate

$$\begin{aligned}\left\|D^\beta u\right\|_{\mathrm{L}^p(\mathbb{R}^d)} &= \left\|D^\beta(\lambda - A_{p,\mathbb{R}^d} - B_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\le \left\|D^\beta(\lambda - A_{p,\mathbb{R}^d})^{-1}\right\|_{\mathcal{L}(\mathrm{L}^p(\mathbb{R}^d))} \left\|\big(\mathrm{I} - B_{p,\mathbb{R}^d}(\lambda - A_{p,\mathbb{R}^d})^{-1}\big)^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\le \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \left\|f\right\|_{\mathrm{L}^p(\mathbb{R}^d)},\end{aligned}$$

where we used (19.1) in the last step. So $u \in W^{2,p}(\mathbb{R}^d)$ and we have found our unique solution with the correct norm estimate.

Consequently, it remains to prove (19.2). In doing so, we will find ε and λ_0. For all $f \in \mathrm{L}^p(\mathbb{R}^d)$ it holds by the definition of $B_{p,\mathbb{R}^d}$, the triangle inequality and the hypotheses on the coefficients

$$\begin{aligned}&\left\|B_{p,\mathbb{R}^d}(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\le \sum_{j,k=1}^d \left\|(a_{jk} - \alpha_{jk})\partial_j\partial_k(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} + \sum_{j=1}^d \left\|b_j\partial_j(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\qquad + \left\|c(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\le \varepsilon \sum_{j,k=1}^d \left\|\partial_j\partial_k(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} + \sum_{j=1}^d \left\|b_j\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \left\|\partial_j(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\qquad + \left\|c\right\|_{\mathrm{L}^\infty(\mathbb{R}^d)} \left\|(\lambda - A_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)}.\end{aligned}$$

Investing again (19.1) we find

$$\le C\varepsilon \left\|f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} + \frac{C}{\sqrt{|\lambda|}} \left\|f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} + \frac{C}{|\lambda|} \left\|f\right\|_{\mathrm{L}^p(\mathbb{R}^d)}.$$

Now choosing, ε small enough and λ_0 big enough we can indeed force the desired estimate by

$$\le \frac{1}{2} \left\|f\right\|_{\mathrm{L}^p(\mathbb{R}^d)}.$$

And this finishes the proof. □

We now start to introduce some boundary values, but we still keep the geometry rather simple by considering the half space case $\Omega = \mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x_d > 0\}$. Then $\partial\Omega = \{x \in \mathbb{R}^d : x_d = 0\}$, so we have a flat boundary and are in a good situation for a reflection argument. On $\partial\Omega$ we impose Dirichlet boundary conditions, so we consider the problem

$$\begin{cases} \lambda u - Lu = f, & \text{in } \mathbb{R}^d_+, \\ \qquad\quad u = 0, & \text{on } \partial\mathbb{R}^d_+, \end{cases}$$

where L is the same formal differential operator as before. Again we start with constant coefficients and set $b = 0$ and $c = 0$. So, let $(\alpha_{jk})_{jk} \in \mathbb{R}^{d\times d}$ be a positive definite matrix and consider $A_{p,\mathbb{R}^d_+}$ to be the operator in $\mathrm{L}^p(\mathbb{R}^d_+)$ with $D(A_{p,\mathbb{R}^d_+}) = W^{2,p}(\mathbb{R}^d_+) \cap W^{1,p}_0(\mathbb{R}^d_+)$ and

$$A_{p,\mathbb{R}^d_+} u = Au = \sum_{j,k=1}^{d} \alpha_{jk} \partial_j \partial_k u \tag{19.3}$$

for $u \in D(A_{p,\mathbb{R}^d_+})$.

Proposition 19.4. *Let $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$. Then for all $f \in \mathrm{L}^p(\mathbb{R}^d_+)$ the problem*

$$\begin{cases} \lambda u - Au = f, & \textit{in } \mathbb{R}^d_+, \\ \qquad\quad u = 0, & \textit{on } \partial\mathbb{R}^d_+, \end{cases} \tag{19.4}$$

has a unique solution $u \in D(A_{p,\mathbb{R}^d_+})$ and for all multiindices $\beta \in \mathbb{N}^d_0$ with $|\beta| \leq 2$ we have

$$\left\| D^\beta u \right\|_{\mathrm{L}^p(\mathbb{R}^d_+)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \left\| f \right\|_{\mathrm{L}^p(\mathbb{R}^d_+)}.$$

Proof. We have seen in the proof of Proposition 19.1 that $(\lambda - A_{p,\mathbb{R}^d})^{-1}$ is given by a symbol satsfying the Mikhlin condition, so this is a Calderón-Zygmund operator with a locally integrable integral kernel. From this we infer, that for every $f \in \mathrm{C}^\infty_c(\mathbb{R}^d)$ also $(\lambda - A_{p,\mathbb{R}^d})^{-1} f$ is a smooth function on $\mathbb{R}^d$.

Now let $f \in C^\infty_c(\mathbb{R}^d_+)$ be given and define $\tilde{f}$ as the odd extension of f to $\mathbb{R}^d$ by

$$\tilde{f}(x', x_d) := \begin{cases} f(x', x_d), & \text{for } x_d > 0 \\ 0, & \text{for } x_d = 0, \\ -f(x', -x_d), & \text{for } x_d > 0. \end{cases}$$

Since the support of f is away from the boundary also the extension $\tilde{f}$ is smooth and has a compact support, so $\tilde{f} \in \mathrm{C}^\infty_c(\mathbb{R}^d)$. So, by the reasoning above and Proposition 19.1, the funtion $\tilde{u} := (\lambda - A_{p,\mathbb{R}^d})^{-1} \tilde{f}$ is in $\mathrm{C}^\infty(\mathbb{R}^d) \cap W^{2,p}(\mathbb{R}^d)$ and using the unique solvability of the whole space problem, one finds that $\tilde{u}$ is an odd function as well. This entrails

$\tilde{u}(x',0) = 0$ for all $x' \in \mathbb{R}^{d-1}$, so $u := \tilde{u}|_{\mathbb{R}^d_+}$ is in $W^{2,p}(\mathbb{R}^d_+) \cap W^{1,p}_0(\mathbb{R}^d_+) = D(A_{p,\mathbb{R}^d_+})$ and it solves (19.4). Finally for $|\beta| \leq 2$ it holds

$$\left\|D^\beta u\right\|_{\mathrm{L}^p(\mathbb{R}^d_+)} \leq \left\|D^\beta \tilde{u}\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \left\|\tilde{f}\right\|_{\mathrm{L}^p(\mathbb{R}^d)} = \frac{2C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d_+)}.$$

The rest is a density argument: We have shown that $D^\alpha(\lambda - A_{p,\mathbb{R}^d_+})^{-1}$ maps $\mathrm{C}^\infty_c(\mathbb{R}^d_+)$ to $\mathrm{L}^p(\mathbb{R}^d_+)$ with the right norm estimate, so we can extend theses operators continuously to all of $\mathrm{L}^p(\mathbb{R}^d_+)$ with the same estimates. □

The next step is to repeat the reasoning in the proof of Proposition 19.3 to also incorporate lower order terms and small deviations from the constant coefficients in the highest order into the results for the half space. This proof is done by a purely functional analytic reasoning, mainly based on the Neumann series, so it carries over without difficulties. We will therefore not repeat the proof, but just state the result. For this we define the $\mathrm{L}^p(\mathbb{R}_+)$ realization $L_{p,\mathbb{R}^d_+}$ of our formal differential operator L by $D(L_{p,\mathbb{R}^d_+}) = W^{2,p}(\mathbb{R}^d_+) \cap W^{1,p_0}(\mathbb{R}^d_+)$ and $L_{p,\mathbb{R}^d_+}u = Lu$ for $u \in D(L_{p,\mathbb{R}^d_+})$.

Proposition 19.5. *There exists $\varepsilon > 0$ and $\lambda_0 \geq 0$ such that the following holds: If $a \in \mathrm{L}^\infty(\mathbb{R}^d_+, \mathbb{R}^{d\times d})$ satisfies $\|a_{jk} - \alpha_{jk}\|_{\mathrm{L}^\infty(\mathbb{R}^d_+)} \leq \varepsilon$ for all $j,k = 1,\dots,d$ and $\lambda \in \mathbb{C}$ satisfies $\mathrm{Re}(\lambda) \geq \lambda_0$, then for every $f \in \mathrm{L}^p(\mathbb{R}^d_+)$ the equation $(\lambda - L_{p,\mathbb{R}^d_+})u = f$ has a unique solution $u \in D(L_{p,\mathbb{R}^d_+})$ with*

$$\left\|D^\beta u\right\|_{\mathrm{L}^p(\mathbb{R}^d_+)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d_+)}$$

for all multiindices $\beta \in \mathbb{N}^d_0$ with $|\beta| \leq 2$.

From now on let $\Omega \subseteq \mathbb{R}^d$ be a bounded open set with a C^2-boundary. We recall that the definition of a C^2-boundary means that for every $x_0 \in \partial\Omega$, there is an open neighbourhood U of x_0 such that in a suitable coordninate system centered at x_0 the boundary of Ω can be represented as the graph of a C^2-function h. This leads to a C^2-diffeomeorphism Φ that locally around x_0 flattens the boundary and brings us locally to a half space situation, cf. Lemma 8.4 and Figure 8.1.

In the local coordinate system around x_0 this diffeomeorphism was given by $\Phi(x) = (x', x_d - h(x'))$, to the effect that

$$J_\Phi(x) = \begin{pmatrix} \mathrm{I} & 0 \\ \nabla h(x')^T & 1 \end{pmatrix}.$$

By picking the local coordinate system appropriately, one can adjust that $\nabla h(x_0') = 0$, so without loss of generality in the sequel we can always assume that $J_\Phi(x_0)$ equals the identity matrix.

For this localised situation we prove a little lemma.

Lemma 19.6. *Let $\alpha = (\alpha_{jk})_{jk} \in \mathbb{R}^{d\times d}$ be a positive definite matrix, let A again be as in (19.3) and let $\Phi : U \to V$ be a C^2-diffeomorphism as above. Assume in addition that for some $\varepsilon > 0$ it holds $\left\|J_\Phi^T \alpha J_\Phi - \alpha\right\|_{L^\infty(U)} \le \varepsilon$. Then there is a differential operator $\tilde{A}$ given by*

$$\tilde{A}\tilde{v}(y) = \sum_{j,k=1}^{d} \tilde{a}_{jk}(y)\partial_j\partial_k\tilde{v}(y) + \sum_{j=1}^{d} \tilde{b}_j(y)\partial_j\tilde{v}(y), \quad y \in \mathbb{R}^d,$$

such that $Av = [\tilde{A}(v\circ\Phi^{-1})]\circ\Phi$ for all $v \in W^{2,p}(\Omega\cap U)$ and for the coefficients it holds $\tilde{a} \in L^\infty(\mathbb{R}^d, \mathbb{R}^{d\times d})$ with $\|\tilde{a} - \alpha\|_{L^\infty(\mathbb{R}^d,\mathbb{R}^{d\times d})} \le \varepsilon$ and $\tilde{b} \in L^\infty(\mathbb{R}^d, \mathbb{R}^{d\times d})$.

Proof. Set $\tilde{v} := v\circ\Phi^{-1}$. Then

$$\begin{aligned}
Av(x) = A(\tilde{v}\circ\Phi)(x) &= \sum_{j,k=1}^{d} \alpha_{jk}\partial_j\partial_k(\tilde{v}(\Phi(x))) = \sum_{j,k=1}^{d} \alpha_{jk}\partial_j\left[\left(\nabla\tilde{v}(\Phi(x))\right)\cdot\partial_k\Phi(x)\right]\\
&= \sum_{j,k=1}^{d} \alpha_{jk}\partial_j \sum_{\ell=1}^{d} \partial_\ell\tilde{v}(\Phi(x))\partial_k\Phi_\ell(x)\\
&= \sum_{j,k=1}^{d} \alpha_{jk} \sum_{\ell=1}^{d}\left[\nabla(\partial_\ell\tilde{v})(\Phi(x))\cdot\partial_j\Phi(x)\partial_k\Phi_\ell(x) + \partial_\ell\tilde{v}(\Phi(x))\partial_j\partial_k\Phi_\ell(x)\right]\\
&= \sum_{j,k=1}^{d} \alpha_{jk} \sum_{\ell=1}^{d}\left[\sum_{m=1}^{d}(\partial_m\partial_\ell\tilde{v})(\Phi(x))\partial_j\Phi_m(x)\partial_k\Phi_\ell(x) + \partial_\ell\tilde{v}(\Phi(x))\partial_j\partial_k\Phi_\ell(x)\right]\\
&= \sum_{m,\ell=1}^{d}\left[\left(J_\Phi(x)^T\alpha J_\Phi(x)\right)_{ml}\partial_m\partial_\ell\tilde{v}(\Phi(x)) + \sum_{\ell=1}^{d}\sum_{j,k=1}^{d}\alpha_{jk}\partial_j\partial_k\Phi_\ell(x)\partial_\ell\tilde{v}(\Phi(x))\right.\\
&=: \sum_{m,\ell=1}^{d} \tilde{a}_{m\ell}(\Phi(x))\partial_m\partial_\ell\tilde{v}(\Phi(x)) + \sum_{\ell=1}^{d}\tilde{b}_\ell(\Phi(x))\partial_\ell\tilde{v}(\Phi(x)).
\end{aligned}$$

Setting

$$\tilde{a}(y) := \begin{cases} J_\Phi(\Phi^{-1}(y))^T\alpha J_\Phi(\Phi^{-1}(y)), & y \in V,\\ \alpha, & y \in \mathbb{R}^d\setminus V,\end{cases}$$

and

$$\tilde{b}(y) := \begin{cases} \left(\sum_{j,k=1}^{d}\alpha_{jk}\partial_j\partial_k\Phi_\ell(\Phi^{-1}(y))\right)_{\ell=1,\dots,d}, & y \in V,\\ 0, & y \in \mathbb{R}^d\setminus V,\end{cases}$$

by the above calculation we indeed have $Av = (\tilde{A}\tilde{v}) \circ \Phi$. Furthermore, we also find $\tilde{a} \in \mathrm{L}^\infty(\mathbb{R}^d_+, \mathbb{R}^{d\times d})$ with

$$\|\tilde{a} - \alpha\|_{\mathrm{L}^\infty(\mathbb{R}^d, \mathbb{R}^{d\times d})} \leq \left\|(J_\Phi \circ \Phi^{-1})^T \alpha (J_\Phi \circ \Phi^{-1}) - \alpha\right\|_{\mathrm{L}^\infty(V, \mathbb{R}^{d\times d})}$$
$$= \left\|J_\Phi^T \alpha J_\Phi - \alpha\right\|_{\mathrm{L}^\infty(U, \mathbb{R}^{d\times d})} \leq \varepsilon. \qquad \square$$

Now we are ready to formulate the main result.

Theorem 19.7. *Let $\Omega \subseteq \mathbb{R}^d$ be open and bounded with a C^2-boundary and let $a \in \mathrm{C}(\overline{\Omega}, \mathbb{R}^{d\times d})$ be symmetric and elliptic, $b \in \mathrm{L}^\infty(\Omega, \mathbb{R}^d)$ and $c \in \mathrm{L}^\infty(\Omega)$. Then there exists a $\lambda_0 \geq 0$ such that for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) \geq \lambda_0$ and every $f \in \mathrm{L}^p(\Omega)$ the problem*

$$\begin{cases} \lambda u - Lu = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{19.5}$$

has a unique solution $u \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ and

$$\left\|D^\beta u\right\|_{\mathrm{L}^p(\Omega)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\Omega)}$$

for every multiindex $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq 2$.

Proof. For every $x_0 \in \overline{\Omega}$ we "freeze the coefficients" and consider the operators

$$A_{x_0} u(x) := \sum_{j,k=1}^{d} a_{jk}(x_0) \partial_j \partial_k u(x).$$

These are of purely second order with constant elliptic coefficient matrix, as the operators A in all our previous considerations. Since a is bouneded on $\overline{\Omega}$ and elliptic, all of these matrices have a uniform norm bound and a uniform ellipticity constant.

Thus, by Proposition 19.1 for all $x_0 \in \overline{\Omega}$ and all $\mathrm{Re}(\lambda) > 0$ the operator $(\lambda - (A_{x_0})_{p,\mathbb{R}^d})^{-1}$ is bounded from $\mathrm{L}^p(\mathbb{R}^d)$ to $W^{2,p}(\mathbb{R}^d)$ and for all $f \in \mathrm{L}^p(\mathbb{R}^d)$ and $|\beta| \leq 2$ the estimate

$$\left\|D^\beta (\lambda - (A_{x_0})_{p,\mathbb{R}^d})^{-1} f\right\|_{\mathrm{L}^p(\mathbb{R}^d)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d)} \tag{19.6}$$

is valid with a constant C that is independent of x_0.

Let $\varepsilon > 0$. We will choose a suitable value for this number later. For the moment we note that thanks to the uniform continuity of a on the compact set $\overline{\Omega}$ there exists some $\tilde{\delta}_\varepsilon > 0$ such that for all $x_0 \in \overline{\Omega}$ and all $x \in B_{\tilde{\delta}_\varepsilon}(x_0)$ it holds $|a(x) - a(x_0)| \leq \varepsilon$.

First step: Localisation. Let $x_0 \in \partial\Omega$. We choose an open neighbourhood $U \subseteq \mathbb{R}^d$ of x_0 that allows for a C^2-diffeomorphism Φ that locally flattens the boundary and satisfies $J_\Phi(x_0) = I$. Since J_Φ is still continuously differentiable, there exists some $\delta_{\varepsilon,x_0} \in (0, \tilde{\delta}_\varepsilon)$ such that $\left\| J_\Phi^T a(x_0) J_\Phi - a(x_0) \right\|_{L^\infty(B_{\delta_{\varepsilon,x_0}})} \leq \varepsilon$ and, at the same time the matrix $J_\Phi(x)^T a(x_0) J_\Phi(x)$ stays elliptic with ellipticity constant $\kappa_0/2$ for all $x \in B_{\delta_{\varepsilon,x_0}}(x_0)$.

Now, $\{B_{\delta_{\varepsilon,x_\ell}}(x_0) : x_0 \in \partial\Omega\}$ is an open covering of the compact set $\partial\Omega$. So, there are $x_1, \dots, x_n \in \partial\Omega$ such that already the balls $U_\ell := B_{\delta_{\varepsilon,x_\ell}}(x_\ell)$, $\ell = 1, \dots, n$, form an open covering of $\partial\Omega$. Setting $\delta_\varepsilon := \min_{\ell=1}^n \delta_{\varepsilon,x_\ell}$ we note that $\delta_\varepsilon \leq \tilde{\delta}_\varepsilon$ and $\delta_\varepsilon \leq \delta_{\varepsilon,x_\ell}$ for all $\ell = 1, \dots, n$.

We have now a suitable localisation of the boundary of Ω, so let us deal with the rest of this set. For this we cover $\Omega \setminus U_\ell$ with a locally finite collection of balls $U_\ell := B_{\delta_\varepsilon}(y_\ell)$, $\ell = n+1, n+2, \dots, n+N$, such that $U_\ell \subseteq \Omega$ for all such ℓ.

Finally we take a quadratic partition of unity φ_ℓ, $\ell = 1, \dots, n+N$, subordinated to our localisation, i.e. we choose functions $\varphi_\ell \in C_c^\infty(\mathbb{R}^d)$ such that

- $\operatorname{supp}(\varphi_\ell) \subseteq U_\ell$ for $\ell = 1, \dots, n+N$,
- $0 \leq \varphi_\ell \leq 1$,
- $\sum_{\ell=1}^{n+N} \varphi_\ell^2 = 1$ on $\overline{\Omega}$.

(To construct such a partition, one starts with a "usual" partition of unity $(\psi_\ell)_{\ell=1,\dots,n+N}$ and sets $\varphi_\ell = \psi_\ell (\sum_{k=1}^{n+N} \psi_k^2)^{-1/2}$.)

Second step: Local approximative solutions. Let $g \in L^p(\Omega)$ be given. and also denote by $g \in L^p(\mathbb{R}^d)$ its extension by zero.

For the balls in the interior of Ω, i.e. for $\ell = n+1, \dots, n+N$ we set

$$u_\ell := \big(\lambda - (A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g).$$

Then $u_\ell \in W^{2,p}(\mathbb{R}^d)$ and for all $|\beta| \leq 2$ we have thanks to (19.6)

$$\left\| D^\beta u_\ell \right\|_{L^p(\mathbb{R}^d)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \left\| \varphi_\ell g \right\|_{L^p(\mathbb{R}^d)}.$$

For our boundary balls, for every x_ℓ, $\ell = 1, \dots, n$, we consider the operator A_{x_ℓ} and associate the corresponding operator $\tilde{A}_{x_\ell}$ given by Lemma 19.6. By our choice of $\delta_{\varepsilon,x_\ell}$ we have assured, that the hypotheses of this Lemma are fulfilled with our ε. So, taking ε small enough, the operators $\tilde{A}_{x_\ell}$ all satisfy the assumptions of Proposition 19.5. Thus for λ_0 big enough and all $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) \geq \lambda_0$ the operator $\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}$ is invertible in $\mathrm{L}^p(\mathbb{R}^d_+)$ and for all $f \in \mathrm{L}^p(\mathbb{R}^d_+)$ it holds

$$\left\| D^\beta \big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1} f \right\|_{\mathrm{L}^p(\mathbb{R}^d_+)} \leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|f\|_{\mathrm{L}^p(\mathbb{R}^d_+)} .$$

Denoting the C^2-diffeomorphism associated to U_ℓ by Φ_ℓ, we now set

$$u_\ell := \big[\big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1}((\varphi_\ell g) \circ \Phi_\ell^{-1})\big] \circ \Phi_\ell =: S_{\ell,\lambda} g.$$

Then for $\ell = 1, \dots, n$

$$\begin{aligned}
\left\| D^\beta u_\ell \right\|_{\mathrm{L}^p(U_\ell)} &= \left\| D^\beta (S_{\lambda,\ell} g) \right\|_{\mathrm{L}^p(U_\ell)} \\
&= \left\| D^\beta \big[\big[\big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1} \big((\varphi g) \circ \Phi_\ell^{-1}\big)\big] \circ \Phi_\ell\big] \right\|_{\mathrm{L}^p(U_\ell)} \\
&\leq C \left\| D^\beta \big[\big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1} \big((\varphi g) \circ \Phi_\ell^{-1}\big)\big] \right\|_{\mathrm{L}^p(\mathbb{R}^d_+)} \\
&\leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \left\| (\varphi_\ell g) \circ \Phi^{-1} \right\|_{\mathrm{L}^p(\mathbb{R}^d_+)} \\
&\leq \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|\varphi_\ell g\|_{\mathrm{L}^p(\mathbb{R}^d)} = \frac{C}{|\lambda|^{1-\frac{|\beta|}{2}}} \|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)} .
\end{aligned}$$

Third Step: Global approximative solution. Let $g \in \mathrm{L}^p(\Omega)$ be given. Using the functions u_ℓ, $\ell = 1 \dots, n+N$, constructed in the preceding step, we set

$$v_g := \sum_{\ell=1}^{n+N} \varphi_\ell u_\ell.$$

Then it holds

$$\begin{aligned}
(\lambda - L) v_g &= \lambda v_g - L \sum_{\ell=1}^{n+N} \varphi_\ell u_\ell \\
&= \lambda \sum_{\ell=1}^{n+N} \varphi_\ell u_\ell - \sum_{\ell=1}^{n+N} \big(\varphi_\ell L u_\ell + (L(\varphi_\ell u_\ell) - \varphi_\ell L u_\ell\big) \\
&= \sum_{\ell=1}^{n+N} \big(\varphi_\ell (\lambda - L) u_\ell + [L, \varphi_\ell] u_\ell\big),
\end{aligned}$$

where $[L, \varphi_\ell]$ is a common shorthand writing for the *commutator* $[L, \varphi_\ell] u := L(\varphi_\ell u) - \varphi_\ell L u$. We introduce our constant coefficient operators A_{x_ℓ} and A_{y_ℓ} and get

$$\begin{aligned}
&= \sum_{\ell=1}^{n} \big(\varphi_\ell (\lambda - A_{x_\ell}) u_\ell + \varphi_\ell (A_{x_\ell} - L) u_\ell + [L, \varphi_\ell] u_\ell\big) \\
&\qquad + \sum_{\ell=n+1}^{N} \big(\varphi_\ell (\lambda - A_{y_\ell}) u_\ell + \varphi_\ell (A_{y_\ell} - L) u_\ell + [L, \varphi_\ell] u_\ell\big).
\end{aligned}$$

By definition of u_ℓ, we immediately get $(\lambda - A_{y_\ell})u_\ell = \varphi_\ell g$ for $\ell = n+1,\dots,N$. For $\ell = 1,\dots,n$, the operator $\tilde{A}_{x_\ell}$ was carefully chosen to provide us with

$$\begin{aligned}
(\lambda - A_{x_\ell})u_\ell &= \lambda u_\ell \circ \Phi_\ell^{-1} \circ \Phi_\ell - [\tilde{A}_{x_\ell}(u_\ell \circ \Phi_\ell^{-1})] \circ \Phi_\ell \\
&= \big[(\lambda - \tilde{A}_{x_\ell})(u_\ell \circ \Phi_\ell^{-1})\big] \circ \Phi_\ell \\
&= \big[(\lambda - \tilde{A}_{x_\ell})\big(\lambda - \big(\tilde{A}_{x_\ell}\big)_{p,\mathbb{R}^d}\big)^{-1}((\varphi_\ell g) \circ \Phi_\ell^{-1})\big] \circ \Phi_\ell \\
&= \varphi_\ell g.
\end{aligned}$$

So, we can continue our calculation to obtain

$$\begin{aligned}
(\lambda - L)v_g &= \sum_{\ell=1}^{n}\big(\varphi_\ell^2 g + \varphi_\ell(A_{x_\ell} - L)u_\ell + [L,\varphi_\ell]u_\ell\big) + \sum_{\ell=n+1}^{n+N}\big(\varphi_\ell^2 g + \varphi_\ell(A_{y_\ell} - L)u_\ell + [L,\varphi_\ell]u_\ell\big) \\
&= g + \sum_{\ell=1}^{n}\big(\varphi_\ell(A_{x_\ell} - L)u_\ell + [L,\varphi_\ell]u_\ell\big) + \sum_{\ell=n+1}^{n+N}\big(\varphi_\ell(A_{y_\ell} - L)u_\ell + [L,\varphi_\ell]u_\ell\big) \\
&=: g + T_\lambda g.
\end{aligned}$$

Summing up, we did not solve $\lambda v - Lv = g$, but we have produced some error $T_\lambda g$ and our aim in the next step will be to correct for this. Before doing so, we want to make sure that, even if our function is not the searched solution, it at least has the right boundary behaviour, i.e. that $v_g = 0$ on $\partial\Omega$.

No point of the boundary of Ω lies in one of the interior balls U_ℓ for $\ell = n+1,\dots,n+N$, so on the boundary of Ω we have

$$v_g = \sum_{\ell=1}^{n+N}\varphi_\ell u_\ell = \sum_{\ell=1}^{N}\varphi_\ell u_\ell = \sum_{\ell=1}^{n}\varphi_\ell\big[\big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1}((\varphi_\ell g) \circ \Phi_\ell^{-1})\big] \circ \Phi_\ell$$

By the definition of $(\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}$, it holds $\big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1}((\varphi_\ell g) \circ \Phi_\ell^{-1}) \in D((\tilde{A}_{x_\ell})_{p,\mathbb{R}^d}) = W^{2,p}(\mathbb{R}^d_+) \cap W_0^{1,p}(\mathbb{R}^d_+)$, so it vanishes on $\partial\mathbb{R}^d_+$. This means in turn that the function $u_\ell = \big[\big(\lambda - (\tilde{A}_{x_\ell})_{p,\mathbb{R}^d_+}\big)^{-1}((\varphi_\ell g) \circ \Phi_\ell^{-1})\big] \circ \Phi_\ell$ vanishes on $\Phi^{-1}(V \cap \mathbb{R}^d_+) = U \cap \partial\Omega$ and we get

$$v_g|_{\partial\Omega} = \sum_{\ell=1}^{n}\varphi_\ell u_\ell|_{\partial\Omega} = 0.$$

Fourth step: Estimate of T_λ. Our aim is to invert $I + T_\lambda$ in $L^p(\Omega)$. So, in view of the Neumann series, we have to show that the norm of T_λ can be made smaller than 1. The operator T_λ consists of four terms:

$$\begin{aligned}
T_\lambda g &= \sum_{\ell=1}^{n} \varphi_\ell (A_{x_\ell} - L) S_{\lambda,\ell} g + \sum_{\ell=1}^{n} [L, \varphi_\ell] S_{\lambda,\ell} g \\
&\quad + \sum_{\ell=n+1}^{n+N} \varphi_\ell (A_{y_\ell} - L)\big(\lambda - (A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g) + \sum_{\ell=n+1}^{n+N} [L, \varphi_\ell]\big(\lambda - (A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g) \\
&=: \sum_{\ell=1}^{n} I_\ell + \sum_{\ell=1}^{n} II_\ell + \sum_{\ell=n+1}^{n+N} III_\ell + \sum_{\ell=n+1}^{n+N} IV_\ell.
\end{aligned}$$

Before starting to estimate the four terms, we take a closer look at the commutator. For every $\varphi \in C_c^\infty(\mathbb{R}^d)$ and all $w \in W^{2,p}(\Omega)$ we have

$$\begin{aligned}
[L, \varphi]w &= \sum_{j,k=1}^{d} \big[a_{jk}\partial_j\partial_k(\varphi w) - \varphi a_{jk}\partial_j\partial_k w\big] + \sum_{j=1}^{d} \big[b_j\partial_j(\varphi w) - \varphi b_j \partial_j w\big] + c\varphi w - \varphi c w \\
&= \sum_{j,k=1}^{d} a_{jk}\big[\partial_j\partial_k\varphi w + \partial_j\varphi\partial_k w + \partial_k\varphi\partial_j w + \varphi\partial_j\partial_k w - \varphi\partial_j\partial_k w\big] \\
&\qquad\qquad + \sum_{j=1}^{d} b_j\big[\partial_j\varphi w + \varphi\partial_j w - \varphi\partial_j w\big] \\
&= \sum_{j,k=1}^{d} a_{jk}\big[\partial_j\partial_k\varphi w + \partial_j\varphi\partial_k w + \partial_k\varphi\partial_j w\big] + \sum_{j=1}^{d} b_j\partial_j\varphi w \\
&= \sum_{j=1}^{d}\sum_{k=1}^{d} a_{jk}\big[\partial_j\varphi\partial_k w + \partial_k\varphi\partial_j w\big] + \Big(\sum_{j=1}^{d}\sum_{k=1}^{d} a_{jk}\partial_j\partial_k\varphi + b_j\partial_j\varphi\Big)w \\
&= \sum_{j=1}^{d}\Big(\sum_{k=1}^{d}(a_{jk} + a_{kj})\partial_k\varphi\Big)\partial_j w + \Big(\sum_{j=1}^{d}\sum_{k=1}^{d} a_{jk}\partial_j\partial_k\varphi + b_j\partial_j\varphi\Big)w \\
&= \sum_{j=1}^{d} \hat{b}_j\partial_j w + \hat{c}w.
\end{aligned}$$

with some coefficients $\hat{b}_j$, $j = 1, \dots, d$, and $\hat{c}$ that are bounded and measurable on Ω with a compact support that is contained in the support of φ. The important point here is, that the commutator of the *second* order differential operator L with the multiplication by φ is only a *first* order differential operator.

Having this at hand, we can now estimate the four terms of T_λ. For the first one we use that on the support of φ_ℓ, so in one of our localisation charts, the second order coefficients of L only deviate very little from the constant coefficients of A_{x_ℓ}. In detail we find for every $\ell = 1, \dots, n$

$$\begin{aligned}
\|I_\ell\|_{\mathrm{L}^p(\Omega)} &= \|\varphi_\ell(A_{x_\ell} - L)S_{\lambda,\ell}g\|_{L^p(\Omega)} = \|\varphi_\ell(A_{x_\ell} - L)S_{\lambda,\ell}g\|_{L^p(U_\ell)} \\
&= \left\| \sum_{j,k=1}^{d} \varphi_\ell(\alpha_{jk} - a_{jk})\partial_j\partial_k(S_{\lambda,\ell}g) - \sum_{j=1}^{d} \varphi_\ell b_j \partial_\ell(S_{\lambda,\ell}g) - \varphi_\ell c S_{\lambda,\ell}g \right\|_{\mathrm{L}^p(U_\ell)} \\
&\leq \sum_{j,k=1}^{d} \|\alpha_{jk} - a_{jk}\|_{\mathrm{L}^\infty(U_\ell)} \|\partial_j\partial_k(S_{\lambda,\ell}g)\|_{\mathrm{L}^p(U_\ell)} + \sum_{j=1}^{d} \|b_j\|_{\mathrm{L}^\infty(\Omega)} \|\partial_j(S_{\lambda,\ell}g)\|_{\mathrm{L}^p(U_\ell)} \\
&\qquad + \|c\|_{\mathrm{L}^\infty(\Omega)} \|S_{\lambda,\ell}g\|_{\mathrm{L}^p(U_\ell)}
\end{aligned}$$

Note that by our choice of the balls U_ℓ, the difference $\alpha_{jk} - a_{jk}$ stays smaller than ε on this set. Relying on the estimates on $S_{\lambda,\ell}$ in the second step, we continue

$$\leq C\Big(\sum_{j,k=1}^{d} \varepsilon \|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)} + \frac{1}{\sqrt{|\lambda|}} \|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)} + \frac{1}{|\lambda|} \|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}\Big).$$

For the second term we use our knowledge on the commutator $[L, \varphi]$. We get with the coefficients of the commutator defined above

$$\|II_\ell\|_{\mathrm{L}^p(\Omega)} = \|[L, \varphi_\ell]S_{\lambda,\ell}g\|_{\mathrm{L}^p(\Omega)} = \left\| \sum_{j=1}^{d} \hat{b}_j \partial_j(S_{\lambda,\ell}g) + \hat{c}S_{\lambda,\ell}g \right\|_{\mathrm{L}^p(\Omega)}$$

Using the support properties of $\hat{b}$ and $\hat{c}$ and the estimates for $S_{\lambda,\ell}$ established in Step 2, we continue

$$\begin{aligned}
&\leq \sum_{j=1}^{d} \left\| \hat{b}_j \partial_j(S_{\lambda.\ell}g) \right\|_{\mathrm{L}^p(U_\ell)} + \|\hat{c}S_{\lambda,\ell}g\|_{\mathrm{L}^p(U_\ell)} \\
&\leq \sum_{j=1}^{d} \left\| \hat{b}_j \right\|_{\mathrm{L}^\infty(U_\ell)} \|\partial_j(S_{\lambda,\ell}g)\|_{\mathrm{L}^p(U_\ell)} + \|\hat{c}\|_{\mathrm{L}^\infty(U_\ell)} \|S_{\lambda,\ell}g\|_{\mathrm{L}^p(U_\ell)} \\
&\leq C\Big(\sum_{j=1}^{d} \frac{1}{\sqrt{|\lambda|}} \|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)} + \frac{1}{|\lambda|} \|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}\Big)
\end{aligned}$$

By analogous calculations as for the first and second term, the third and fourth term allow for similar estimates by

$$\begin{aligned}\|III_\ell\|_{\mathrm{L}^p(\Omega)} &= \Big\|\varphi_\ell(A_{y_\ell}-L)\big(\lambda-(A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g)\Big\|_{\mathrm{L}^p(\Omega)}\\ &\le \sum_{j,k=1}^d \|\alpha_{jk}-a_{jk}\|_{\mathrm{L}^\infty(U_\ell)}\Big\|\partial_j\partial_k\big(\lambda-(A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g)\Big\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\qquad+\sum_{j=1}^d\|b_j\|_{\mathrm{L}^\infty(\Omega)}\Big\|\partial_j\big(\lambda-(A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g)\Big\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\qquad+\|c\|_{\mathrm{L}^\infty(\Omega)}\Big\|\big(\lambda-(A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g)\Big\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\le C\Big(\sum_{j,k=1}^d\varepsilon\|\varphi_\ell g\|_{\mathrm{L}^p(\mathbb{R}^d)}+\frac{1}{\sqrt{|\lambda|}}\|\varphi_\ell g\|_{\mathrm{L}^p(\mathbb{R}^d)}+\frac{1}{|\lambda|}\|\varphi_\ell g\|_{\mathrm{L}^p(\mathbb{R}^d)}\Big)\\ &= C\Big(\sum_{j,k=1}^d\varepsilon\|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}+\frac{1}{\sqrt{|\lambda|}}\|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}+\frac{1}{|\lambda|}\|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}\Big).\end{aligned}$$

and

$$\begin{aligned}\|IV_\ell\|_{\mathrm{L}^p(\Omega)} &\le \sum_{j=1}^d\Big\|\hat{b}_j\Big\|_{\mathrm{L}^\infty(U_\ell)}\Big\|\partial_j\big(\lambda-(A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g)\Big\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\qquad+\|\hat{c}\|_{\mathrm{L}^\infty(U_\ell)}\Big\|\big(\lambda-(A_{y_\ell})_{p,\mathbb{R}^d}\big)^{-1}(\varphi_\ell g)\Big\|_{\mathrm{L}^p(\mathbb{R}^d)}\\ &\le C\Big(\sum_{j=1}^d\frac{1}{\sqrt{|\lambda|}}\|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}+\frac{1}{|\lambda|}\|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}\Big).\end{aligned}$$

Putting everything together we found for every $g\in\mathrm{L}^p(\Omega)$

$$\|T_\lambda g\|_{\mathrm{L}^p(\Omega)}\le C\Big(\varepsilon+\frac{1}{\sqrt{|\lambda|}}+\frac{1}{|\lambda|}\Big)\sum_{\ell=1}^{n+N}\|\varphi_\ell g\|_{\mathrm{L}^p(U_\ell)}\le C\Big(\varepsilon+\frac{1}{\sqrt{|\lambda|}}+\frac{1}{|\lambda|}\Big)\|g\|_{\mathrm{L}^p(\Omega)}.$$

Note that $n+N$ may depend on the choice of ε. However, the covering where the φ_l have their support can be estimated by Lebesgue's covering number which is only dependent on d.

Fifth step: Exakt solution. Let $\varepsilon>0$ so small and $\lambda_0\ge 0$ so large that all arguments up to now go through and, additionally, we have $C(\varepsilon+1/\sqrt{|\lambda|}+1/|\lambda|)\le 1/2$ for all $\mathrm{Re}(\lambda\ge\lambda_0$ in the end of the fourth step. Then by the Neumann series $I+T_\lambda$ is an invertible operator in $\mathrm{L}^p(\Omega)$ with norm bounded by 2.

Now, let $f\in\mathrm{L}^p(\Omega)$ and some $\lambda\in\mathbb{C}$ with $\mathrm{Re}(\lambda)\ge\lambda_0$ be given. We set $g:=(I+T_\lambda)^{-1}f$ and define v_g as in the third step. Then $g\in\mathrm{L}^p(\Omega)$ with $\|g\|_{\mathrm{L}^p(\Omega)}\le 2\|f\|_{\mathrm{L}^p(\Omega)}$ and by the results of the third step we have

$$(\lambda - L)v_g = g + T_\lambda g = (I + T_\lambda)(I + T_\lambda)^{-1} f = f.$$

Furthermore v_g is zero on $\partial\Omega$. So, $u \coloneqq v_g = v_{(I+T_\lambda)^{-1}f}$ indeed solves our elliptic boundary value problem (19.5). Finally, for every $|\beta| \leq 2$ we estimte

$$\left\| D^\beta u \right\|_{\mathrm{L}^p(\Omega)} = \left\| D^\beta \sum_{\ell=1}^{n+N} \varphi_\ell u_\ell \right\|_{\mathrm{L}^p(\Omega)} \leq \sum_{\ell=1}^{n+N} \left\| D^\beta (\varphi_\ell u_\ell) \right\|_{\mathrm{L}^p(U_\ell)}.$$

All derivatives of φ_ℓ up to order 2 and for all ℓ can be bounded in L^∞-norm by some uniform constant, so we can estimate this by the Leibniz rule as

$$\leq C \sum_{\ell=1}^{n+N} \sum_{\gamma \leq \beta} \| D^\gamma u_\ell \|_{\mathrm{L}^p(U_\ell)}$$

and, investing the estimates obtained in the second step, this yields

$$\leq C \sum_{\ell=1}^{n+N} \sum_{\gamma \leq \beta} \frac{1}{|\lambda|^{1-\frac{|\gamma|}{2}}} \| \varphi_\ell g \|_{\mathrm{L}^p(U_\ell)}.$$

Since we talk about large $|\lambda|$, the worst exponent of $|\lambda|$ is the one for the biggest value of $|\gamma|$, which is $|\beta|$. This finally produces

$$\leq C \sum_{\ell=1}^{n+N} \frac{1}{|\lambda|^{1-\frac{1}{|\beta|}}} \| \varphi_\ell g \|_{\mathrm{L}^p(U_\ell)} \leq C \frac{1}{|\lambda|^{1-\frac{1}{|\beta|}}} \| g \|_{\mathrm{L}^p(\Omega)} \leq C \frac{1}{|\lambda|^{1-\frac{1}{|\beta|}}} \| f \|_{\mathrm{L}^p(\Omega)}$$

and we are done. □

Index