Partial Differential Equations 2
Variational Methods *

Martin Brokate †

Contents

*Lecture Notes, Summer Term 2016

†Technical University of Munich, Department of Mathematics

1 Variational Methods: Some Basics

Equations and minimization. The solution of equations is related to minimization. Suppose we want to find $u \in \mathbb{R}^n$ with

$$F(u) = 0\,, \quad F : \mathbb{R}^n \to \mathbb{R}^n\,.$$

If we can find a function $J : \mathbb{R}^n \to \mathbb{R}$ such that $F = \nabla J$, and if we can prove that J has a minimizer u,

$$J(u) = \min J(v)\,, \quad v \in \mathbb{R}^n\,,$$

then we know that $F(u) = \nabla J(u) = 0$. Conversely, suppose we want to find a minimizer u of J. If we can find $u \in \mathbb{R}^n$ with $\nabla J(u) = 0$, then u is a candidate for a minimizer of J. If in addition J is convex, all such candidates are minimizers; otherwise, second order conditions may be used.

Dirichlet's principle for the Laplace operator. We consider the Dirichlet problem for the Poisson equation in an open bounded set $\Omega \subset \mathbb{R}^n$,

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega \\ u &= g \qquad \text{auf } \partial\Omega. \end{aligned} \tag{1.1}$$

We also consider the functional ("$|\cdot|$" denotes the Euclidean norm in $\mathbb{R}^n$)

$$J(v) = \int_\Omega \frac{1}{2}|\nabla v(x)|^2 - f(x)v(x)\,dx\,. \tag{1.2}$$

The functional J and the boundary value problem (1.1) are related as follows.

Assume that $f : \Omega \to \mathbb{R}$ and $g : \partial\Omega \to \mathbb{R}$ are continuous, let

$$K = \{v : v \in C^2(\overline{\Omega}),\, v(x) = g(x) \text{ for all } x \in \partial\Omega\}\,. \tag{1.3}$$

Let u be a minimizer of J on K, that is,

$$u \in K\,, \quad J(u) = \min_{v \in K} J(v)\,. \tag{1.4}$$

Let $\varphi \in C_0^\infty(\Omega)$ be given. We have $u + \lambda\varphi \in K$ for all $\lambda \in \mathbb{R}$. We consider

$$\tilde{J} : \mathbb{R} \to \mathbb{R}\,, \quad \tilde{J}(\lambda) = J(u + \lambda\varphi)\,. \tag{1.5}$$

We have

$$\begin{aligned} \tilde{J}(\lambda) &= \int_\Omega \frac{1}{2}|\nabla(u + \lambda\varphi)(x)|^2 - f(x)(u + \lambda\varphi)(x)\,dx \\ &= J(u) + \lambda \int_\Omega \langle \nabla u(x), \nabla\varphi(x)\rangle - f(x)\varphi(x)\,dx + \lambda^2 \int_\Omega \frac{1}{2}|\nabla\varphi(x)|^2\,dx\,. \end{aligned}$$

Since 0 is a minimizer of the differentiable function $\tilde{J}$, we get

$$0 = \tilde{J}'(0) = \int_\Omega \langle \nabla u(x), \nabla\varphi(x)\rangle - f(x)\varphi(x)\,dx = \int_\Omega (-\Delta u(x) - f(x))\varphi(x)\,dx\,.$$

(Here we have used partial integration and the fact that $\varphi = 0$ on $\partial\Omega$.) Since $\varphi \in C_0^\infty(\Omega)$ can be chosen arbitrarily, it follows that (we will prove this later)

$$-\Delta u(x) = f(x)\,, \quad \text{for all } x \in \Omega. \tag{1.6}$$

Thus, every minimizer of J on K solves (1.1). Conversely, let $u \in C^2(\overline{\Omega})$ be a solution of (1.1), so in particular $u \in K$. Let $v \in K$ be arbitrary. Then

$$\int_\Omega (-\Delta u(x) - f(x))(u(x) - v(x))\, dx = 0\,.$$

Since $u - v = 0$ on $\partial\Omega$, partial integration gives

$$\int_\Omega \langle \nabla u(x), \nabla u(x) - \nabla v(x) \rangle \; dx - \int_\Omega f(x))(u(x) - v(x))\, dx = 0\,.$$

It follows that (we drop the argument "x" of the functions)

$$\int_\Omega |\nabla u|^2 - fu\, dx = \int_\Omega \langle \nabla u, \nabla v \rangle - fv\, dx \le \int_\Omega \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla v|^2 - fv\, dx\,, \tag{1.7}$$

since for arbitrary vectors $y, z \in \mathbb{R}^n$ we have

$$|\langle y, z \rangle| \le |y| \cdot |z| \le \frac{1}{2}|y|^2 + \frac{1}{2}|z|^2\,.$$

Subtracting $(1/2)\int_\Omega |\nabla u|^2\, dx$ on both sides of (1.7), we arrive at

$$J(u) \le J(v)\,,$$

so u is a minimizer of J on K.

The equivalence

$$u \text{ solves (1.1)} \qquad \Leftrightarrow \qquad u \text{ is a minimizer of } J \text{ on } K \tag{1.8}$$

is called **Dirichlet's principle**. In this context, the equation

$$-\Delta u = f$$

is called the **Euler equation** or **Euler-Lagrange equation** corresponding to the functional J in (1.2).

More general, this approach relates functionals of the form

$$J(v) = \int_\Omega L(x, v(x), \nabla v(x))\, dx \tag{1.9}$$

to differential equations. These are ordinary differential equations in the case $n = 1$, $\Omega \subset \mathbb{R}$, and partial differential equations in the case $n > 1$.

Minimizing functionals of the form (1.9) constitutes the basic problem of the **calculus of variations**.

The Dirichlet principle yields equivalence, but **not** existence.

The existence problem is hard in general, because the domain of the functional J is a subset of a function space, which has infinite dimension. Therefore, functional analysis comes into play. Using its methods to prove existence of a minimizer is called the **direct method** of the calculus of variations. That such a minimizer solves the Euler equation is then proved as above by setting to zero the derivatives of J in all directions $\varphi \in C_0^\infty(\Omega)$.

Proving existence of a minimizer of J in (1.9) by proving first the existence of a solution to the Euler equation is called the **indirect method**. If J is not convex, however, solutions of the Euler equation for J in general need to have additional properties in order to be minimizers of J (as is the case in finite dimensions, where second derivatives of J play a role).

Quadratic minimization: An abstract form of Dirichlet's principle. We consider the following situation.

Problem 1.1 (Quadratic minimization problem)
Let V be a vector space, $K \subset V$ convex, let $a : V \times V \to \mathbb{R}$ be a bilinear form and $F : V \to \mathbb{R}$ a linear form. We consider the problem

$$\min_{v \in K} J(v)\,, \quad J(v) = \frac{1}{2}a(v,v) - F(v)\,. \tag{1.10}$$

□

Proposition 1.2 *Consider the situation of Problem 1.1 and assume in addition that a is symmetric and positive definite. Then J is strictly convex, and for $u \in K$ we have*

$$J(u) = \min_{v \in K} J(v) \qquad \Leftrightarrow \qquad a(u, v-u) \geq F(v-u) \quad \forall v \in K\,. \tag{1.11}$$

There exists at most one $u \in K$ with this property.

Proof: Let $u, h \in V$ be arbitrary with $h \neq 0$. We define

$$J_{u,h} : \mathbb{R} \to \mathbb{R}\,, \quad J_{u,h}(\lambda) = J(u + \lambda h)\,.$$

We have

$$J_{u,h}(\lambda) = \frac{1}{2}a(u+\lambda h, u+\lambda h) - F(u+\lambda h) = J(u) + \lambda(a(u,h) - F(h)) + \frac{\lambda^2}{2}a(h,h)\,. \tag{1.12}$$

Since $a(h,h) > 0$, the quadratic function $J_{u,h}$ is strictly convex, and so is J. Therefore J has at most one minimizer.
"$\Leftarrow$": Let $v \in K$. Setting $h = v - u$, from (1.12) and (1.11) it follows that

$$J(v) - J(u) = J(u+h) - J(u) = (a(u,h) - F(h)) + \frac{1}{2}a(h,h) \geq 0\,.$$

"$\Rightarrow$": Let $v \in K$. Again setting $h = v - u$, we have $u + \lambda h \in K$ for $0 \leq \lambda \leq 1$ (since K is convex). Thus

$$0 \leq J_{u,h}(\lambda) - J_{u,h}(0) = \lambda(a(u,h) - F(h)) + \frac{\lambda^2}{2}a(h,h)\,. \tag{1.13}$$

Dividing by λ and passing to the limit $\lambda \downarrow 0$ yields $0 \le a(u,h) - F(h)$. $\Box$

The inequality system (one inequality for each $v \in K$)

$$u \in K\,, \qquad a(u, v-u) \ge F(v-u) \quad \forall v \in K\,, \tag{1.14}$$

is called a **variational inequality**. As (1.13) shows, the variational inequality says that the directional derivative of J in a minimizer is nonnegative for those directions $h = v-u$ which point into the "admissible set" K.

Corollary 1.3 *In the situation of Proposition 1.2 assume that K is an affine subspace of V, that is, $K = v_0 + U$ with $v_0 \in V$, U subspace of V. Then we have*

$$J(u) = \min_{v \in K} J(v) \qquad \Leftrightarrow \qquad a(u,w) = F(w) \quad \forall w \in U\,. \tag{1.15}$$

Proof: If $u \in K$ we have $U = \{v - u : v \in K\}$ and therefore

$$\begin{aligned} a(u, v-u) \ge F(v-u) \quad \forall v \in K \qquad &\Leftrightarrow \qquad a(u,w) \ge F(w) \quad \forall w \in U \\ &\Leftrightarrow \qquad a(u,w) = F(w) \quad \forall w \in U\,, \end{aligned}$$

the latter equivalence holds because $w \in U$ implies $-w \in U$. $\Box$

The system of equations

$$u \in v_0 + U\,, \qquad a(u,v) = F(v) \quad \forall v \in U\,, \tag{1.16}$$

is called a **variational equation**.

The situation (1.1) – (1.4), which we considered at the beginning, becomes a special case of (1.16) if we set

$$a(u,v) = \int_\Omega \langle \nabla u(x), \nabla v(x) \rangle \; dx\,, \quad F(v) = \int_\Omega f(x) v(x)\, dx\,. \tag{1.17}$$

The variational equation then reads

$$\int_\Omega \langle \nabla u(x), \nabla v(x) \rangle \; dx = \int_\Omega f(x) v(x)\, dx\,, \quad \text{for all } v \text{ in } U. \tag{1.18}$$

(So far we did not specify the space V and the subspace U.)

If one starts from the Poisson equation

$$-\Delta u = f\,, \quad \text{in } \Omega, \tag{1.19}$$

one may arrive at the variational formulation (1.18) immediately by multiplying (1.19) on both sides with functions v which are zero on $\partial\Omega$, and performing partial integration. In this manner one bypasses the functional J.

The variational equation. In order to obtain existence, the algebraic vector space structure by itself does not suffice. One uses the setting of normed spaces. Thus, topological aspects enter the picture.

Definition 1.4 (Bilinear form, continuity and ellipticity)
Let V be a normed space. A bilinear form $a : V \times V \to \mathbb{R}$ is called ***continuous*** *if there exists a $C_a > 0$ such that*

$$|a(u,v)| \leq C_a \|u\| \, \|v\| \,, \quad \textit{for all } u, v \in V \,, \tag{1.20}$$

it is called V-elliptic if there exists a $c_a > 0$ such that

$$a(v,v) \geq c_a \|v\|^2 \,, \quad \textit{for all } v \in V \,. \tag{1.21}$$

□

As an immediate consequence of this definition we see that a V-elliptic bilinear form is positive definite.

In order to obtain existence, completeness of the normed space is required.

Proposition 1.5 (Variational equation, solvability)
Let $(V, \|\cdot\|)$ be a Banach space, let $a : V \times V \to \mathbb{R}$ be a symmetric, continuous and V-elliptic bilinear form, let $F : V \to \mathbb{R}$ be linear and continuous. Then there exists a unique solution $u \in V$ of the variational equation

$$a(u,v) = F(v) \qquad \forall v \in V \,. \tag{1.22}$$

Proof: For all $v \in V$ we have

$$c_a \|v\|^2 \leq a(v,v) \leq C_a \|v\|^2 \,.$$

Since $c_a > 0$,

$$\langle u, v \rangle_a = a(u,v) \,, \quad \|v\|_a = \sqrt{a(v,v)} \,,$$

defines a scalar product whose associated norm $\|\cdot\|_a$ is equivalent to the original norm $\|\cdot\|$. Therefore, $(V, \langle \cdot, \cdot \rangle_a)$ is complete and thus a Hilbert space. Moreover, $F : (V, \|\cdot\|_a) \to \mathbb{R}$ is continuous. The representation theorem of Riesz (from functional analysis) states that there is a unique $u \in V$ such that

$$\langle u, v \rangle_a = F(v) \,.$$

□

Corollary 1.6 *The solution u of the variational equation (1.22) in Proposition 1.5 satisfies*

$$\|u\| \leq \frac{1}{c_a} \|F\| \,. \tag{1.23}$$

Proof: Inserting u for v in (1.22) yields

$$c_a \|u\|^2 \leq a(u,u) = F(u) \leq \|F\| \cdot \|u\| \,.$$

□

Proposition 1.5 and Corollary 1.6 state that, under the given assumptions, the solution of the variational equation is a **well-posed problem** in the sense that its solution exists, is unique and depends continuously upon the data (the right hand side specified by F). The latter property has to interpreted as follows: the solution operator $S : V^* \to V$ which maps $F \in V^*$ (the dual space of V) to the solution u of (1.22), is continuous. Indeed, it is linear (this follows immediately from (1.22)), and according to (1.23) we have

$$\|S(F)\| \le \frac{\|F\|}{c_a}, \quad \text{therefore} \quad \|S\| \le \frac{1}{c_a}.$$

The variational inequality. In Proposition 1.5 we have assumed that the bilinear form is symmetric. It turns out that this assumption is superfluous. We present this extension in an even more general context, namely, for the variational inequality.

Proposition 1.7 (Variational inequality, solvability)
Let V be a Banach space, let $K \subset V$ be closed, convex and nonempty. Let $a : V \times V \to \mathbb{R}$ be a continuous V-elliptic bilinear form with

$$a(v,v) \ge c_a \|v\|^2 \qquad \forall v \in V, \tag{1.24}$$

let $F : V \to \mathbb{R}$ be linear and continuous. Then the variational inequality

$$a(u, v-u) \ge F(v-u) \qquad \forall\, v \in K, \tag{1.25}$$

has a unique solution $u \in K$. Moreover, if $\tilde{u} \in K$ is the solution of (1.25) corresponding to a linear and continuous $\tilde{F} : V \to \mathbb{R}$ in place of F, we have

$$\|u - \tilde{u}\| \le \frac{1}{c_a} \|F - \tilde{F}\|. \tag{1.26}$$

Proof: We first prove (1.26). This also implies the uniqueness. Assume that $u, \tilde{u}$ are solutions corresponding to $F, \tilde{F}$. Then

$$\begin{aligned} a(u, \tilde{u} - u) &\ge F(\tilde{u} - u), \\ a(\tilde{u}, u - \tilde{u}) &\ge \tilde{F}(u - \tilde{u}). \end{aligned}$$

Adding these inequalities yields

$$a(u - \tilde{u}, \tilde{u} - u) \ge (F - \tilde{F})(\tilde{u} - u), \tag{1.27}$$

so

$$0 \le c_a \|u - \tilde{u}\|^2 \le a(u - \tilde{u}, u - \tilde{u}) \le \|F - \tilde{F}\| \|u - \tilde{u}\|. \tag{1.28}$$

This shows that (1.26) holds. To prove existence we first consider the special case where a is symmetric. By Proposition 1.2 it suffices to show that the associated quadratic functional

$$J(v) = \frac{1}{2} a(v,v) - F(v)$$

has a minimizer on K. In order to prove this, we see that

$$J(v) \ge \frac{1}{2} c_a \|v\|^2 - \|F\| \|v\| = \left(\sqrt{\frac{c_a}{2}} \|v\| - \sqrt{\frac{1}{2c_a}} \|F\| \right)^2 - \frac{1}{2c_a} \|F\|^2 \tag{1.29}$$

holds for all $v \in V$, therefore J is bounded from below. Define

$$d = \inf_{v \in K} J(v)\,, \tag{1.30}$$

and let $(u_n)_{n\in\mathbb{N}}$ be a sequence in K such that

$$d \le J(u_n) \le d + \frac{1}{n}\,. \tag{1.31}$$

The sequence $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, because

$$\begin{aligned} c_a\|u_n - u_m\|^2 &\le a(u_n - u_m, u_n - u_m) \\ &= 2a(u_n, u_n) + 2a(u_m, u_m) - 4a\left(\frac{1}{2}(u_n + u_m), \frac{1}{2}(u_n + u_m)\right) \\ &= 4J(u_n) + 4J(u_m) - 8J\left(\frac{1}{2}(u_n + u_m)\right) \\ &\le 4\left(d + \frac{1}{n}\right) + 4\left(d + \frac{1}{m}\right) - 8d \\ &\le 4\left(\frac{1}{n} + \frac{1}{m}\right)\,. \end{aligned} \tag{1.32}$$

As V is a Banach space, there exists a $u \in V$ such that $u_n \to u$. Since K is closed, we have $u \in K$. Since J is continuous, we have $J(u) = d$. This concludes the proof in the case where a is symmetric.

Now let a be arbitrary. We utilize a continuation argument. We consider the family of bilinear forms

$$a_t(u, v) = a_0(u, v) + tb(u, v)\,, \quad t \in [0, 1]\,, \tag{1.33}$$

where

$$a_0(u, v) = \frac{1}{2}(a(u, v) + a(v, u))\,, \tag{1.34}$$

$$b(u, v) = \frac{1}{2}(a(u, v) - a(v, u))\,. \tag{1.35}$$

Then a_0 is symmetric and $a_1 = a$. The bilinear forms a_t are continuous and V-elliptic, and (1.24) holds for a_t with the same constant c_a as for a, because $a_t(u, u) = a(u, u)$ holds for all $u \in V$. As a_0 is symmetric, it now suffices to prove the following claim:

If the variational inequality

$$a_\tau(u, v - u) \ge G(v - u)\,, \qquad \forall\, v \in K\,, \tag{1.36}$$

has a unique solution $u \in K$ for every $G \in V^*$, then the same is true if we replace τ by t with

$$\tau \le t \le \tau + \frac{c_a}{2C_a}\,. \tag{1.37}$$

We prove this claim. Let t be given, satisfying (1.37). We only have to prove the existence of a solution. Let $G : V \to \mathbb{R}$ be linear and continuous. We look for a solution $u_t \in K$ satisfying

$$a_t(u_t, v - u_t) \ge G(v - u_t)\,, \qquad \forall\, v \in K\,. \tag{1.38}$$

For this purpose, we consider for arbitrary given $w \in V$ the variational inequality

$$a_\tau(u, v-u) \geq F_w(v-u)\,, \qquad \forall\, v \in K\,, \tag{1.39}$$

where

$$F_w(v) = G(v) - (t-\tau) b(w, v)\,. \tag{1.40}$$

Let $T : V \to K$ be the mapping which associates to each $w \in V$ the solution $u \in K$ of (1.39). T is a contraction, since

$$\begin{aligned} \|Tw_1 - Tw_2\| &\leq \frac{1}{c_a} \|F_{w_1} - F_{w_2}\| \leq \frac{1}{c_a} |t-\tau| \sup_{\|v\|=1} |b(w_1 - w_2, v)| \\ &\leq \frac{1}{c_a} |t-\tau| C_a \|w_1 - w_2\| \leq \frac{1}{2} \|w_1 - w_2\|\,. \end{aligned} \tag{1.41}$$

By Banach's fixed point theorem, T has a unique fixed point which we denote by u_t. Then $u_t \in K$, and for all $v \in K$ we have

$$\begin{aligned} a_t(u_t, v-u_t) &= a_\tau(u_t, v-u_t) + (t-\tau) b(u_t, v-u_t) \\ &\geq F_{u_t}(v-u_t) + (t-\tau) b(u_t, v-u_t) = G(v-u_t)\,. \end{aligned}$$

This finishes the proof. □

Corollary 1.8 (Lax-Milgram theorem)
Let V be a Banach space, let $a : V \times V \to \mathbb{R}$ be a continuous V-elliptic bilinear form, let $F : V \to \mathbb{R}$ be linear and continuous. Then the variational equation

$$a(u, v) = F(v) \qquad \forall\, v \in V\,, \tag{1.42}$$

has a unique solution $u \in V$.

Proof: We apply Proposition 1.7 with $K = V$. As in Corollary 1.3, the assertion follows from the equivalences

$$\begin{aligned} a(u, v-u) \geq F(v-u) \quad \forall\, v \in V \qquad &\Leftrightarrow \qquad a(u,v) \geq F(v) \quad \forall\, v \in V \\ &\Leftrightarrow \qquad a(u,v) = F(v) \quad \forall\, v \in V\,. \end{aligned} \tag{1.43}$$

□

Assume that one wants to solve a given linear partial differential equation. This usually determines the form of the associated bilinear form a. If one wants to apply the Lax-Milgram theorem to prove existence and uniqueness, one has to choose the Banach space V such that a is continuous and V-elliptic. The spaces V for which this works are the **Sobolev spaces**.

2 Sobolev Spaces: Definition

In this section Ω always denotes an open subset of $\mathbb{R}^n$.

Weak derivatives. Let $u \in C^1(\Omega)$ and $\varphi \in C_0^\infty(\Omega)$. Partial integration yields

$$\int_\Omega \partial_i u(x)\varphi(x)\,dx = -\int_\Omega u(x)\partial_i\varphi(x)\,dx\,, \quad 1 \le i \le n\,. \tag{2.1}$$

The right side of this equation is well-defined for functions u whose restriction to the support of φ (a compact subset of Ω) is integrable. We thus consider the vector space

$$L^1_{\mathrm{loc}}(\Omega) = \{v : v|K \in L^1(K) \text{ for all compact subsets } K \subset \Omega\}\,. \tag{2.2}$$

The functions in $L^1_{\mathrm{loc}}(\Omega)$ are called **locally integrable**.

Continuous functions $u : \Omega \to \mathbb{R}$ are locally integrable since they are bounded on every compact subset of Ω; they may or may not be integrable on Ω. For example,

$$u(x) = \frac{1}{x}\,, \quad u : (0,1) \to \mathbb{R}\,,$$

belongs to $L^1_{\mathrm{loc}}(0,1)$, but not to $L^1(0,1)$. We remark that $L^p(\Omega) \subset L^1_{\mathrm{loc}}(\Omega)$ for all $p \in [1,\infty]$, no matter whether Ω is bounded or not.

For locally integrable functions one can use (2.1) in order to define a generalization of the classical derivative.

Definition 2.1 (Weak partial derivative)
Let $u \in L^1_{loc}(\Omega)$. A function $w \in L^1_{loc}(\Omega)$ which satisfies

$$\int_\Omega w(x)\varphi(x)\,dx = -\int_\Omega u(x)\partial_i\varphi(x)\,dx\,, \quad \textit{for all } \varphi \in C_0^\infty(\Omega), \tag{2.3}$$

*is called an **i-th weak partial derivative** of u. We also denote it by $\partial_i u$.* □

The notation "$\partial_i u$" for weak derivatives makes sense only if they are uniquely determined. This we will see later in Corollary 2.8.

For $u \in C^1(\Omega)$, the classical partial derivatives $\partial_i u$ of u are also weak derivatives of u.

Consider the example

$$u : \mathbb{R} \to \mathbb{R}\,, \quad u(x) = |x|\,. \tag{2.4}$$

For $\varphi \in C_0^\infty(\mathbb{R})$ we have

$$\begin{aligned}
-\int_{\mathbb{R}} |x|\varphi'(x)\,dx &= \int_{-\infty}^0 x\varphi'(x)\,dx - \int_0^\infty x\varphi'(x)\,dx \\
&= x\varphi(x)\Big|_{x=-\infty}^{x=0} - \int_{-\infty}^0 \varphi(x)\,dx - x\varphi(x)\Big|_{x=0}^{x=\infty} + \int_0^\infty \varphi(x)\,dx \\
&= \int_{\mathbb{R}} (\operatorname{sign} x)\varphi(x)\,dx\,,
\end{aligned}$$

therefore the weak derivative of u exists, and

$$u' = \mathrm{sign}\ . \tag{2.5}$$

This equality has to be understood as an equality in $L^1_{\mathrm{loc}}(\mathbb{R})$, that is,

$$u'(x) = \mathrm{sign}\,(x)\,, \quad \text{a.e. in } \mathbb{R}. \tag{2.6}$$

For this example, the standard pointwise derivative of u exists in all points $x \neq 0$, and is a.e. equal to the weak derivative.

As a second example, consider the **Heaviside funktion** $H : \mathbb{R} \to \mathbb{R}$,

$$H(x) = \begin{cases} 1\,, & x > 0\,, \\ 0\,, & x \leq 0\,. \end{cases} \tag{2.7}$$

For $\varphi \in C_0^\infty(\mathbb{R})$ we have

$$-\int_{\mathbb{R}} H(x)\varphi'(x)\,dx = -\int_0^\infty \varphi'(x)\,dx = \varphi(0)\,. \tag{2.8}$$

If $w \in L^1_{\mathrm{loc}}(\mathbb{R})$ were a weak derivative of H, we would have

$$\int_{\mathbb{R}} w(x)\varphi(x)\,dx = \varphi(0)\,, \quad \text{for all } \varphi \in C_0^\infty(\mathbb{R}). \tag{2.9}$$

But this is not possible (see an exercise), thus the Heaviside function does not have a weak derivative in the sense of Definition 2.1.

One can of course consider the pointwise derivative of H, which is zero except at $x = 0$. However, this function is a.e. equal to zero, so (when it appears under an integral sign) it does not distinguish between H and a constant function. But one can use (2.8) to generalize the concept of a weak derivative even further and consider the linear functional $T : C_0^\infty(\mathbb{R}) \to \mathbb{R}$ defined by

$$T(\varphi) = \varphi(0) \tag{2.10}$$

as a generalized derivative of H. This is the starting point of **distribution theory**; the corresponding derivative is also called **distributional derivative**. The theory of distributions was initiated in the '40s of the previous century, the main proponent being Laurent Schwartz (1915 – 2002).

Approximation by smooth functions. Given a function $u : \Omega \to \mathbb{R}$, there are various ways to define smooth approximations u^ε of u with $u^\varepsilon \to u$ as $\varepsilon \downarrow 0$. Here we mainly use approximation by convolution.

Let us recall the definition of the convolution of two functions u and v,

$$(u * v)(x) = \int_{\mathbb{R}^n} u(x - y)v(y)\,dy\,. \tag{2.11}$$

It is defined for $u, v \in L^1(\mathbb{R}^n)$ and yields a function $u * v \in L^1(\mathbb{R}^n)$. It has the properties

$$u * v = v * u\,, \quad \|u * v\|_1 \leq \|u\|_1 \cdot \|v\|_1\,. \tag{2.12}$$

We want to approximate a given function u by functions $\eta_\varepsilon * u$; they will turn out to be smooth if η_ε is smooth. In order to achieve this, we define

$$\psi : \mathbb{R} \to \mathbb{R}\,, \quad \psi(t) = \begin{cases} \exp\left(-\frac{1}{t}\right)\,, & t > 0\,, \\ 0\,, & t \le 0\,, \end{cases} \tag{2.13}$$

$$\tilde{\psi} : \mathbb{R} \to \mathbb{R}\,, \quad \tilde{\psi}(r) = \psi(1 - r^2)\,, \tag{2.14}$$

$$\eta_1 : \mathbb{R}^n \to \mathbb{R}\,, \quad \eta_1(x) = \alpha\tilde{\psi}(|x|)\,, \tag{2.15}$$

where $\alpha > 0$ is chosen such that

$$\int_{\mathbb{R}^n} \eta_1(x)\,dx = 1\,. \tag{2.16}$$

For given $\varepsilon > 0$ we define the functions

$$\eta_\varepsilon : \mathbb{R}^n \to \mathbb{R}\,, \quad \eta_\varepsilon(x) = \frac{1}{\varepsilon^n}\eta_1\left(\frac{x}{\varepsilon}\right)\,. \tag{2.17}$$

These functions are radially symmetric (that is, they only depend upon $|x|$), and we have

$$\eta_\varepsilon \in C_0^\infty(\mathbb{R}^n)\,, \quad \mathrm{supp}\,(\eta_\varepsilon) = K(0;\varepsilon)\,, \quad \eta_\varepsilon \ge 0\,, \quad \int_{\mathbb{R}^n} \eta_\varepsilon(x)\,dx = 1\,. \tag{2.18}$$

The function η_1, or the family $(\eta_\varepsilon)_{\varepsilon>0}$, is called a **mollifier**.

Given $u : \Omega \to \mathbb{R}$, by $\tilde{u}$ we denote its extension to $\mathbb{R}^n$,

$$\tilde{u}(x) = \begin{cases} u(x)\,, & x \in \Omega\,, \\ 0\,, & x \notin \Omega\,. \end{cases} \tag{2.19}$$

Given $u \in L^1_{\mathrm{loc}}(\Omega)$, we define functions $u^\varepsilon : \mathbb{R}^n \to \mathbb{R}$ by

$$u^\varepsilon = \eta_\varepsilon * \tilde{u}\,. \tag{2.20}$$

We then have

$$u^\varepsilon(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)\tilde{u}(y)\,dy = \int_\Omega \eta_\varepsilon(x-y)u(y)\,dy\,, \quad \text{for all } x \in \mathbb{R}^n. \tag{2.21}$$

We define the ε-neighbourhood of a subset U of $\mathbb{R}^n$ by

$$U_\varepsilon = \{x : x \in \mathbb{R}^n,\, \mathrm{dist}\,(x, U) < \varepsilon\}\,. \tag{2.22}$$

Since $\eta_\varepsilon(x-y) = 0$ if $|x-y| \ge \varepsilon$, we see from (2.21) that

$$\mathrm{supp}\,(u^\varepsilon) \subset \overline{\Omega}_\varepsilon\,, \quad \text{if } u \in L^1_{\mathrm{loc}}(\Omega). \tag{2.23}$$

To denote higher order partial derivatives we use multi-indices. Let

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$$

be a multi-index. We set

$$\partial^\alpha = \partial_1^{\alpha_1}\partial_2^{\alpha_2}\cdots\partial_n^{\alpha_n}\,, \quad |\alpha| = \sum_{j=1}^n \alpha_j\,. \tag{2.24}$$

Proposition 2.2 *Let $\Omega \subset \mathbb{R}^n$ be open, $u \in L^p(\Omega)$, $1 \le p \le \infty$, $\varepsilon > 0$. Then we have*

$$u^\varepsilon \in C^\infty(\mathbb{R}^n)\,, \quad \operatorname{supp}(u^\varepsilon) \subset \overline{\Omega}_\varepsilon\,, \tag{2.25}$$

$$\partial^\alpha u^\varepsilon(x) = \int_\Omega \partial^\alpha \eta_\varepsilon(x-y)u(y)\,dy\,, \quad \textit{for all } x \in \mathbb{R}^n. \tag{2.26}$$

$$\|u^\varepsilon\|_{L^p(\mathbb{R}^n)} \le \|u\|_{L^p(\Omega)}\,. \tag{2.27}$$

Proof: Let α be a multi-index. Fix $x_0 \in \mathbb{R}^n$. For $x, y \in \mathbb{R}^n$ with $|x - x_0| \le 1$ we have we have

$$|u(y)\partial^\alpha \eta_\varepsilon(x-y)| \le \begin{cases} \|\partial^\alpha \eta_\varepsilon\|_\infty \cdot |u(y)|\,, & |y - x_0| \le 1+\varepsilon\,, \\ 0\,, & \text{otherwise,} \end{cases}$$

since $|x-y| > \varepsilon$ if $|y - x_0| > 1 + \varepsilon$. The function on the right side is integrable and does not depend on x. Therefore, we can interchange ∂^α with the convolution integral in the neighbourhood of every point $x_0 \in \mathbb{R}^n$. This proves (2.25) and (2.26). Let us now prove (2.27). In the case $p = \infty$, (2.27) follows from the estimate

$$|u^\varepsilon(x)| \le \int_\Omega \eta_\varepsilon(x-y)|u(y)|\,dy \le \|u\|_\infty \cdot \int_\Omega \eta_\varepsilon(x-y)\,dy \le \|u\|_\infty\,, \quad x \in \mathbb{R}^n\,.$$

In the case $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, we obtain for all $x \in \mathbb{R}^n$, using Hölder's inequality,

$$\begin{aligned} |u^\varepsilon(x)| &= \left|\int_\Omega \eta_\varepsilon(x-y)u(y)\,dy\right| \le \int_\Omega |u(y)|(\eta_\varepsilon(x-y))^{\frac{1}{p}}(\eta_\varepsilon(x-y))^{\frac{1}{q}}\,dx \\ &\le \left(\int_\Omega |u(y)|^p \eta_\varepsilon(x-y)\,dx\right)^{\frac{1}{p}} \underbrace{\left(\int_\Omega \eta_\varepsilon(x-y)\,dx\right)^{\frac{1}{q}}}_{\le 1}\,. \end{aligned} \tag{2.28}$$

From this we obtain, now for $1 \le p < \infty$

$$\begin{aligned} \int_{\mathbb{R}^n} |u^\varepsilon(x)|^p\,dx &\le \int_{\mathbb{R}^n}\int_\Omega |u(y)|^p \eta_\varepsilon(x-y)\,dy\,dx = \int_\Omega |u(y)|^p \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)\,dx\,dy \\ &\le \int_\Omega |u(y)|^p\,dx\,. \end{aligned}$$

□

Definition 2.3
Let $\Omega \subset \mathbb{R}^n$ be open. We say that $U \subset \mathbb{R}^n$ is ***compactly embedded in Ω****, if $\overline{U}$ is compact and $\overline{U} \subset \Omega$. We write*

$$U \subset\subset \Omega\,. \tag{2.29}$$

□

Let us denote by

$$C_0(\Omega) = \{v : v \in C(\Omega),\, \operatorname{supp}(v) \subset\subset \Omega\} \tag{2.30}$$

the set of all continuous functions on Ω whose support is a compact subset of Ω. We have $C_0(\Omega) \subset L^p(\Omega)$ for all $p \in [1, \infty]$.

Proposition 2.4 *The set $C_0(\mathbb{R}^n)$ is a dense subspace of $L^p(\mathbb{R}^n)$ for all $p \in [1,\infty)$, that is, for every $u \in L^p(\mathbb{R}^n)$ and every $\delta > 0$ there exists a $v \in C_0(\mathbb{R}^n)$ such that $\|u - v\|_p \leq \delta$.*

Proof: One possibility is to consider first the special case $u = 1_B$, where B is an arbitrary measurable subset of Ω, using the regularity of the Lebesgue measure. Another possibility is to use Lusin's theorem (which says that measurable functions can be approximated uniformly by continuous functions on subsets of "almost full" measure). This might be discussed in the exercises. □

Proposition 2.5 *Let $u \in C_0(\mathbb{R}^n)$. Then $u^\varepsilon \to u$ uniformly in $\mathbb{R}^n$.*

Proof: We have for all $x \in \mathbb{R}^n$

$$\begin{aligned}|u^\varepsilon(x) - u(x)| &= \Big|\int_{\mathbb{R}^n} \eta_\varepsilon(x-y)(u(y)-u(x))\,dy\Big| \leq \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)|u(y)-u(x)|\,dy \\ &\leq \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)\,dy \cdot \max_{|z-y|\leq\varepsilon} |u(z)-u(y)| = \max_{|z-y|\leq\varepsilon} |u(z)-u(y)| \;\to 0\end{aligned}$$

as $\varepsilon \to 0$, since u is uniformly continuous on the compact set $\operatorname{supp}(u)$. □

Proposition 2.6 *Let $u \in L^p(\Omega)$, $1 \leq p < \infty$. Then we have $u^\varepsilon \to u$ in $L^p(\Omega)$ for $\varepsilon \to 0$.*

Proof: Let $\delta > 0$. We extend u to $\tilde{u} \in L^p(\mathbb{R}^n)$, setting $\tilde{u}$ to 0 outside Ω. According to Proposition 2.4 we can find $v \in C_0(\mathbb{R}^n)$ with $\|\tilde{u} - v\|_{L^p(\mathbb{R}^n)} \leq \delta$, thus $\|u - v\|_{L^p(\Omega)} \leq \delta$. Then

$$\begin{aligned}\|u - u^\varepsilon\|_{L^p(\Omega)} &\leq \|u - v\|_{L^p(\Omega)} + \|v - v^\varepsilon\|_{L^p(\Omega)} + \|v^\varepsilon - u^\varepsilon\|_{L^p(\Omega)} &(2.31)\\ &\leq 2\delta + \|v - v^\varepsilon\|_{L^p(\Omega)}\,, &(2.32)\end{aligned}$$

because

$$v^\varepsilon - u^\varepsilon = \eta_\varepsilon * v - \eta_\varepsilon * u = (v-u)^\varepsilon$$

and, by Proposition 2.2

$$\|(v-u)^\varepsilon\|_{L^p(\Omega)} \leq \|v - \tilde{u}\|_{L^p(\mathbb{R}^n)} \leq \delta\,.$$

Since $v \in C_0(\mathbb{R}^n)$, we have $v^\varepsilon \to v$ uniformly by Proposition 2.5, so $v^\varepsilon \to v$ in $L^p(\Omega)$ and therefore

$$\limsup_{\varepsilon\to 0} \|u - u^\varepsilon\|_{L^p(\Omega)} \leq 2\delta\,.$$

As $\delta > 0$ was arbitrary, the assertion follows. □

The next result (or a variant of it) is also called the **fundamental lemma of the calculus variations**.

Proposition 2.7 *Let $\Omega \subset \mathbb{R}^n$ be open, let $u \in L^1_{loc}(\Omega)$, assume that*

$$\int_\Omega u(x)\varphi(x)\,dx = 0\,, \quad \textit{for all } \varphi \in C_0^\infty(\Omega). \tag{2.33}$$

Then $u = 0$ a.e. in Ω.

Proof: Let $x \in \Omega$ be arbitrary. If ε is small enough, the support of the function $\varphi(y) = \eta_\varepsilon(x-y)$ is compactly embedded in Ω, so $\varphi \in C_0^\infty(\Omega)$. From (2.33) it follows that

$$u^\varepsilon(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)u(y)\,dy = \int_\Omega u(y)\varphi(y)\,dy = 0\,.$$

Consequently,

$$\lim_{\varepsilon\downarrow 0} u^\varepsilon(x) = 0\,, \quad \text{for all } x \in \Omega.$$

On the other hand, for every ball $B \subset\subset \Omega$ we have $u \in L^1(B)$, so $u^\varepsilon \to u$ in $L^1(B)$ by Proposition 2.6. Thus $u = 0$ a.e. in B and therefore $u = 0$ a.e. in Ω. □

Corollary 2.8 *Weak partial derivatives of a function $u \in L^1_{loc}(\Omega)$ are uniquely determined.*

Proof: If $w, \tilde{w} \in L^1_{\text{loc}}(\Omega)$ satisfy

$$\int_\Omega w(x)\varphi(x)\,dx = -\int_\Omega u(x)\partial_i\varphi(x)\,dx = \int_\Omega \tilde{w}(x)\varphi(x)\,dx$$

for all $\varphi \in C_0^\infty(\Omega)$, Proposition 2.7 implies that $w - \tilde{w} = 0$ a.e. in Ω. □

Sobolev spaces. We have already defined weak derivatives of first order via the partial integration rule. The same can be done for higher derivatives.

Definition 2.9 *Let $u \in L^1_{loc}(\Omega)$, α a multi-index. A function $w \in L^1_{loc}(\Omega)$ which satisfies*

$$\int_\Omega w(x)\varphi(x)\,dx = (-1)^{|\alpha|}\int_\Omega u(x)\partial^\alpha\varphi(x)\,dx\,, \quad \textit{for all } \varphi \in C_0^\infty(\Omega), \tag{2.34}$$

is called an ***α-th weak partial derivative*** *of u and denoted by $\partial^\alpha u$.* □

Again, it follows from the fundamental lemma of the calculus of variations (Proposition 2.7) that there can be at most one α-th weak partial derivative of u.

Definition 2.10 (Sobolev space)
Let $\Omega \subset \mathbb{R}^n$ be open, $k \in \mathbb{N}$, $1 \le p \le \infty$. We define

$$W^{k,p}(\Omega) = \{v : v \in L^p(\Omega),\ \partial^\alpha v \in L^p(\Omega) \textit{ for all } |\alpha| \le k\}\,. \tag{2.35}$$

□

In particular, we have

$$W^{0,p}(\Omega) = L^p(\Omega)\,. \tag{2.36}$$

The spaces $W^{k,p}(\Omega)$ are vector spaces. Indeed, if $u, v \in L^1_{\text{loc}}(\Omega)$ have weak derivatives $\partial^\alpha u$ and $\partial^\alpha v$, we have

$$\int_\Omega (\beta\partial^\alpha u(x) + \gamma\partial^\alpha v(x))\varphi(x)\,dx = (-1)^{|\alpha|}\int_\Omega (\beta u(x) + \gamma v(x))\partial^\alpha\varphi(x)\,dx$$

for all $\beta, \gamma \in \mathbb{R}$ and all $\varphi \in C_0^\infty(\Omega)$. Therefore, $\partial^\alpha(\beta u + \gamma v)$ exists as a weak partial derivative, and

$$\partial^\alpha(\beta u + \gamma v) = \beta\partial^\alpha u + \gamma\partial^\alpha v\,. \tag{2.37}$$

Proposition 2.11 *Let $\Omega \subset \mathbb{R}^n$ be open, $k \in \mathbb{N}$, $1 \le p \le \infty$. The space $W^{k,p}(\Omega)$ is a Banach space when equipped with the norm*

$$\|v\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha|\le k} \|\partial^\alpha v\|_p^p \right)^{\frac{1}{p}} = \left(\sum_{|\alpha|\le k} \int_\Omega |\partial^\alpha v(x)|^p \, dx \right)^{\frac{1}{p}} , \quad 1 \le p < \infty , \tag{2.38}$$

$$\|v\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha|\le k} \|\partial^\alpha v\|_\infty , \quad p = \infty . \tag{2.39}$$

Proof: Consider first $p < \infty$. The triangle inequality holds because

$$\begin{aligned} \|u+v\|_{W^{k,p}(\Omega)} &= \left(\sum_{|\alpha|\le k} \|\partial^\alpha u + \partial^\alpha v\|_p^p \right)^{\frac{1}{p}} \le \left(\sum_{|\alpha|\le k} (\|\partial^\alpha u\|_p + \|\partial^\alpha v\|_p)^p \right)^{\frac{1}{p}} \\ &\le \left(\sum_{|\alpha|\le k} \|\partial^\alpha u\|_p^p \right)^{\frac{1}{p}} + \left(\sum_{|\alpha|\le k} \|\partial^\alpha v\|_p^p \right)^{\frac{1}{p}} = \|u\|_{W^{k,p}(\Omega)} + \|v\|_{W^{k,p}(\Omega)} . \end{aligned}$$

All other properties of the norm follow immediately from the definitions. It remains to prove the completeness. Let $(u_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $W^{k,p}(\Omega)$. Due to

$$\|\partial^\alpha u_n - \partial^\alpha u_m\|_p \le \|u_n - u_m\|_{W^{k,p}(\Omega)}$$

the sequences $(\partial^\alpha u_n)_{n\in\mathbb{N}}$ are Cauchy sequences in $L^p(\Omega)$ for all multi-indices $|\alpha| \le k$. Therefore there exist $u \in L^p(\Omega)$, $u_\alpha \in L^p(\Omega)$ satisfying

$$u_n \to u , \quad \partial^\alpha u_n \to u_\alpha \tag{2.40}$$

in $L^p(\Omega)$. Let now $\varphi \in C_0^\infty(\Omega)$. We have

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

therefore $\partial^\alpha u = u_\alpha$ for all $|\alpha| \le k$ and thus $u \in W^{k,p}(\Omega)$. For the case $p = \infty$ the proofs are analogous. □

We have seen that for $u \in L^p(\Omega)$, the functions u^ε defined by

$$u^\varepsilon(x) = \int_\Omega \eta_\varepsilon(x-y)u(y)\, dy$$

converge to u in $L^p(\Omega)$. This does not carry over to $W^{k,p}(\Omega)$; near the boundary of Ω, the situation is more complicated. This is related to the fact that for $u \in W^{k,p}(\Omega)$, the extension $\tilde{u} = u$ on Ω, $\tilde{u} = 0$ outside of Ω, in general does not belong to $W^{k,p}(\Omega)$ (compare the example of the Heaviside function). However, if we stay away from the boundary, this problem does not arise.

Consider open sets $\Omega, U \subset \mathbb{R}^n$ with $U \subset\subset \Omega$. Then

$$\operatorname{dist}(U, \partial\Omega) := \inf_{x\in U, y\in\partial\Omega} |x-y| > 0 , \tag{2.41}$$

since $\overline{U}$ is compact. (In the case $\Omega = \mathbb{R}^n$ we set $\operatorname{dist}(U,\partial\Omega) = \infty$.)

Lemma 2.12 *Let $\Omega, U \subset \mathbb{R}^n$ be open, $U \subset\subset \Omega$. let $u \in W^{k,p}(\Omega)$, $1 \le p < \infty$. Then $u^\varepsilon \in C^\infty(U) \cap W^{k,p}(U)$ holds for all ε satisfying $0 < \varepsilon < \operatorname{dist}(U, \partial\Omega)$. Moreover, $u^\varepsilon \to u$ in $W^{k,p}(U)$ for $\varepsilon \to 0$.*

Proof: Let $\varepsilon < \operatorname{dist}(U, \partial\Omega)$. By Proposition 2.2 we have $u^\varepsilon \in C^\infty(U)$. Moreover, for all $x \in U$

$$\begin{aligned}\partial^\alpha u^\varepsilon(x) &= \int_\Omega \partial^\alpha \eta_\varepsilon(x-y)u(y)\,dy = (-1)^{|\alpha|}\int_\Omega \partial_y^\alpha \eta_\varepsilon(x-y)u(y)\,dy \\ &= \int_\Omega \eta_\varepsilon(x-y)\partial^\alpha u(y)\,dy = (\eta_\varepsilon * \partial^\alpha u)(x)\,.\end{aligned}$$

Since $\partial^\alpha u \in L^p(\Omega)$, by Proposition 2.2 we get $\partial^\alpha u^\varepsilon \in L^p(U)$. From Proposition 2.6 we get that $\partial^\alpha u^\varepsilon \to \partial^\alpha u$ in $L^p(U)$. □

While in general it does not hold that $u^\varepsilon \to u$ in $W^{k,p}(\Omega)$, we nevertheless can construct smooth approximations in $W^{k,p}(\Omega)$ from the interior.

Proposition 2.13 *Let $\Omega \subset \mathbb{R}^n$ be open, let $v \in W^{k,p}(\Omega)$, $1 \le p < \infty$. Then there exists a sequence $(v_n)_{n\in\mathbb{N}}$ in $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ with $v_n \to v$ in $W^{k,p}(\Omega)$.*

Proof: We first consider the case $\Omega \ne \mathbb{R}^n$. We define

$$U_j = \{x : x \in \Omega, \operatorname{dist}(x, \partial\Omega) > \frac{1}{j} \text{ and } |x| < j\}\,, \quad V_j = U_{j+3} \setminus \overline{U}_{j+1}\,, \quad j \ge 1\,, \tag{2.42}$$

and $V_0 = U_3$. We have

$$\Omega = \bigcup_{j=0}^\infty V_j\,.$$

Let $(\beta_j)_{j\ge 0}$ be a partition of unity for Ω with

$$0 \le \beta_j \le 1\,, \quad \beta_j \in C_0^\infty(V_j)\,, \quad \sum_{j=0}^\infty \beta_j = 1\,. \tag{2.43}$$

Since $v \in W^{k,p}(\Omega)$, one also has $\beta_j v \in W^{k,p}(\Omega)$ (see the exercises), and $\operatorname{supp}(\beta_j v) \subset V_j$. Let now $\delta > 0$ be arbitrary. We choose $\varepsilon_j > 0$ sufficiently small so that for

$$w_j = \eta_{\varepsilon_j} * (\beta_j v)\,, \quad W_j = U_{j+4} \setminus \overline{U}_j\,, \quad j \ge 1\,, \quad W_0 = U_4\,,$$

we have from Lemma 2.12, applied to $\beta_j v$ and W_j in place of u and U,

$$\operatorname{supp}(w_j) \subset W_j\,, \tag{2.44}$$

$$\|w_j - \beta_j v\|_{W^{k,p}(\Omega)} = \|w_j - \beta_j v\|_{W^{k,p}(W_j)} \le 2^{-(j+1)}\delta\,. \tag{2.45}$$

We set

$$w = \sum_{j=0}^\infty w_j\,.$$

By construction, on each W_j only finitely many summands are nonzero. Therefore we have $w \in C^\infty(\Omega)$, as $w_j \in C^\infty(\Omega)$ for each j according to Lemma 2.12. It now follows that

$$\|w - v\|_{W^{k,p}(\Omega)} = \left\| \sum_{j=0}^{\infty} w_j - \sum_{j=0}^{\infty} \beta_j v \right\|_{W^{k,p}(\Omega)} \leq \sum_{j=0}^{\infty} \|w_j - \beta_j v\|_{W^{k,p}(\Omega)} \leq \delta \sum_{j=0}^{\infty} 2^{-(j+1)} = \delta \,.$$

As $\delta > 0$ was arbitrary, the assertion follows.

In the case $\Omega = \mathbb{R}^n$ the proof proceeds analogously, with the choice $U_j = B(0, j)$, the open ball around 0 with radius j. □

Density results of this kind can be used to extend to Sobolev spaces properties known to hold for smooth functions. For example, we know that

$$\partial_i(uv) = (\partial_i u)v + u\partial_i v \,, \quad u, v \in C^1(\Omega) \,. \tag{2.46}$$

Consider now $u \in W^{1,p}(\Omega)$, $v \in C^1(\Omega)$. According to Proposition 2.13, choose $u_n \in W^{1,p}(\Omega) \cap C^\infty(\Omega)$ such that $u_n \to u$ in $W^{1,p}(\Omega)$. Let $\varphi \in C_0^\infty(\Omega)$ be arbitrary. We have

$$-\int_\Omega u_n v \partial_i \varphi \, dx = \int_\Omega \partial_i(u_n v)\varphi \, dx = \int_\Omega ((\partial_i u_n)v + u_n \partial_i v)\varphi \, dx \,.$$

Since $u_n \to u$ and $\partial_i u_n \to \partial_i u$ in $L^p(\Omega)$ as $n \to \infty$, we obtain

$$-\int_\Omega u v \partial_i \varphi \, dx = \int_\Omega ((\partial_i u)v + u \partial_i v)\varphi \, dx \,.$$

This shows that $\partial_i(uv)$ exists as a weak derivative and that (2.46) holds. If moreover v and its derivatives are bounded, then $uv \in W^{1,p}(\Omega)$.

Definition 2.14 *Let $\Omega \subset \mathbb{R}^n$ be open, let $1 \leq p < \infty$, $k \in \mathbb{N}$. We define $W_0^{k,p}(\Omega) \subset W^{k,p}(\Omega)$ by*

$$W_0^{k,p}(\Omega) = \overline{C_0^\infty(\Omega)} \,, \tag{2.47}$$

the closure being taken with respect to the norm in $W^{k,p}(\Omega)$. □

$W_0^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$ and therefore itself a Banach space, if equipped with the norm of $W^{k,p}(\Omega)$. $W_0^{k,p}(\Omega)$ is a function space whose elements are zero on $\partial\Omega$ in a certain weak sense; it is not clear at this point whether the statement "$v = 0$ on $\partial\Omega$" makes sense for a general $v \in W^{k,p}(\Omega)$, because equality in $W^{k,p}(\Omega)$ means equality almost everywhere, and the statement "$v = 0$ a.e. on $\partial\Omega$" does not give any information when $\partial\Omega$ has measure 0 in $\mathbb{R}^n$ (which usually is the case).

Definition 2.15 *Let $\Omega \subset \mathbb{R}^n$ be open, let $k \in \mathbb{N}$. We define*

$$H^k(\Omega) = W^{k,2}(\Omega) \,, \quad H_0^k(\Omega) = W_0^{k,2}(\Omega) \,. \tag{2.48}$$

□

Proposition 2.16 *Let $\Omega \subset \mathbb{R}^n$ be open, $k \in \mathbb{N}$. The spaces $H^k(\Omega)$ and $H_0^k(\Omega)$ are Hilbert spaces when equipped with the scalar product*

$$\langle u, v \rangle_{H^k(\Omega)} = \sum_{|\alpha| \le k} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2(\Omega)} = \sum_{|\alpha| \le k} \int_\Omega \partial^\alpha u(x) \cdot \partial^\alpha v(x) \, dx \,. \tag{2.49}$$

We have

$$\|v\|_{W^{k,2}(\Omega)} = \sqrt{\langle v, v \rangle_{H^k(\Omega)}} \,, \quad v \in H^k(\Omega) \,. \tag{2.50}$$

Proof: The properties of the scalar product (2.49) immediately follow from the corresponding properties of the scalar product in $L^2(\Omega)$. Obviously we have (2.50), and by Proposition 2.11 the space $H^k(\Omega)$ is complete. □

Setting

$$|v|_{H^k(\Omega)} = \left(\sum_{|\alpha| = k} \int_\Omega |\partial^\alpha v(x)|^2 \, dx \right)^{\frac{1}{2}} \,, \tag{2.51}$$

one obtains a seminorm on $H^k(\Omega)$ with

$$|v|_{H^k(\Omega)} \le \|v\|_{H^k(\Omega)} \,. \tag{2.52}$$

For $k > 0$, this is not a norm since $|v|_{H^k(\Omega)} = 0$ whenever v is a polynomial of degree less than k.

The following inequality (2.53) is known as the **Poincaré-Friedrichs inequality**.

Proposition 2.17
Let $\Omega \subset [-R, R]^n$ be open. Then for every $v \in H_0^1(\Omega)$ we have

$$\int_\Omega |v(x)|^2 \, dx \le 4R^2 \int_\Omega |\partial_j v(x)|^2 \, dx \,, \quad 1 \le j \le n \,, \tag{2.53}$$

and therefore with $C = 4R^2$

$$\int_\Omega |v(x)|^2 \, dx \le C \int_\Omega |\nabla v(x)|^2 \, dx \,. \tag{2.54}$$

Proof: Consider first $v \in C_0^\infty(\Omega)$. Extending v by zero on all of $\mathbb{R}^n$ we have $v \in C_0^\infty(\mathbb{R}^n)$. Let $x = (x_1, \dots, x_n) \in \Omega$. Then we have

$$v(x) = \underbrace{v(-R, x_2, \dots, x_n)}_{=0} + \int_{-R}^{x_1} \partial_1 v(t, x_2, \dots, x_n) \, dt \,,$$

and therefore, by the Cauchy-Schwarz inequality,

$$\begin{aligned} |v(x)|^2 &= \left(\int_{-R}^{x_1} 1 \cdot \partial_1 v(t, x_2, \dots, x_n) \, dt \right)^2 \le \int_{-R}^{x_1} 1^2 \, dt \cdot \int_{-R}^{x_1} (\partial_1 v(t, x_2, \dots, x_n))^2 \, dt \\ &\le 2R \int_{-R}^{R} (\partial_1 v(t, x_2, \dots, x_n))^2 \, dt \,. \end{aligned}$$

Since the right hand side does not depend on x_1, it follows that

$$\int_{-R}^{R} |v(x)|^2\, dx_1 \le 4R^2 \int_{-R}^{R} (\partial_1 v(t, x_2, \dots, x_n))^2\, dt\,. \tag{2.55}$$

Integrating with respect to the other coordinates $x_2, \dots, x_n$ yields

$$\int_\Omega |v(x)|^2\, dx \le 4R^2 \int_\Omega (\partial_1 v(x))^2\, dx\,,$$

Since we may choose any other j in places of 1, (2.53) is proved for $v \in C_0^\infty(\Omega)$. For arbitrary $v \in H_0^1(\Omega)$, we choose a sequence $(v_k)_{k\in\mathbb{N}}$ in $C_0^\infty(\Omega)$ with $v_k \to v$ in $H_0^1(\Omega)$. Then (2.53) holds for v_k and, passing to the limit $k \to \infty$, for v also. □

Proposition 2.18 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then there exists a $C > 0$ such that*

$$\|v\|_{H^k(\Omega)} \le C|v|_{H^k(\Omega)}\,, \quad \textit{for all } v \in H_0^k(\Omega). \tag{2.56}$$

Proof: Let $v \in H_0^k(\Omega)$, let α be a multi-index with $|\alpha| = j-1$, $1 \le j \le k$. Then we have $\partial^\alpha v \in H_0^1(\Omega)$, and it follows from Proposition 2.17 that

$$\|\partial^\alpha v\|_2^2 \le 4R^2 \|\partial_1 \partial^\alpha v\|_2^2\,.$$

Consequently,

$$\sum_{|\alpha|=j-1} \|\partial^\alpha v\|_2^2 \le 4R^2 \sum_{|\alpha|=j-1} \|\partial_1 \partial^\alpha v\|_2^2 \le 4R^2 \sum_{|\alpha|=j} \|\partial^\alpha v\|_2^2\,,$$

so

$$|v|_{H^{j-1}(\Omega)}^2 \le 4R^2 |v|_{H^j(\Omega)}^2\,,$$

and finally

$$\|v\|_{H^k(\Omega)}^2 = \sum_{j=0}^{k} |v|_{H^j(\Omega)}^2 \le \sum_{j=0}^{k} (4R^2)^{k-j} |v|_{H^k(\Omega)}^2\,.$$

□

Corollary 2.19 *The space $H_0^k(\Omega)$ becomes a Hilbert space when equipped with the scalar product corresponding to $|\cdot|_{H^k(\Omega)}$, and $|\cdot|_{H^k(\Omega)}$ defines a norm on $H_0^k(\Omega)$ which is equivalent to $\|\cdot\|_{H^k(\Omega)}$.* □

With an analogous proof, one can also obtain the Poincaré-Friedrichs inequality for an exponent $p \in (1, \infty)$ instead of 2,

$$\|v\|_{L^p(\Omega)} \le C\|\nabla v\|_{L^p(\Omega)}\,, \quad \text{for all } v \in W_0^{1,p}(\Omega). \tag{2.57}$$

This yields an equivalent norm on $W_0^{1,p}(\Omega)$ in the same manner as in Corollary 2.19.

The optimal (that is, smallest possible) C in (2.57) is called **Poincaré constant**, it depends on Ω and p. It is not easy to determine it.

3 Elliptic Boundary Value Problems

We want to solve the boundary value problem

$$Lu = f\,, \quad \text{in } \Omega, \tag{3.1}$$
$$u = 0\,, \quad \text{on } \partial\Omega. \tag{3.2}$$

Here, $\Omega \subset \mathbb{R}^n$ is open, $f : \Omega \to \mathbb{R}$, and L is an elliptic differential operator of the form

$$\begin{aligned} Lu &= -\sum_{i,j=1}^{n} \partial_j(a_{ij}(x)\partial_i u) + \sum_{i=1}^{n} b_i(x)\partial_i u + c(x)u \\ &= -\mathrm{div}\,(A(x)^T\nabla u) + \langle b(x), \nabla u\rangle + c(x)u\,. \end{aligned} \tag{3.3}$$

One says that L is written in **divergence form**. The variational formulation of (3.1), (3.2) is obtained multiplying (3.1) by an arbitary test function $\varphi \in C_0^\infty(\Omega)$ and integrating over Ω. After partial integration, the left side of (3.1) then becomes

$$\begin{aligned} \int_\Omega Lu(x)\varphi(x)\,dx = \int_\Omega \sum_{i,j=1}^{n} a_{ij}(x)\partial_i u(x)\partial_j\varphi(x)\,dx &+ \int_\Omega \sum_{i=1}^{n} b_i(x)\partial_i u(x)\varphi(x)\,dx \\ &+ \int_\Omega c(x)u(x)\varphi(x)\,dx\,. \end{aligned}$$

Thus, the corresponding bilinear form is given by

$$\begin{aligned} a(u,v) &= \int_\Omega \sum_{i,j=1}^{n} a_{ij}(x)\partial_i u(x)\partial_j v(x)\,dx + \int_\Omega \sum_{i=1}^{n} b_i(x)\partial_i u(x)v(x)\,dx \\ &\quad + \int_\Omega c(x)u(x)v(x)\,dx \\ &= \int_\Omega \langle \nabla u(x), A(x)\nabla v(x)\rangle\;dx + \int_\Omega \langle b(x), \nabla u(x)\rangle\, v(x)\,dx + \int_\Omega c(x)u(x)v(x)\,dx\,. \end{aligned} \tag{3.4}$$

It is symmetric if A is symmetric and $b = 0$.

Assumption 3.1 *We assume that there exists $a_* > 0$ such that*

$$\xi^T A(x)\xi = \sum_{i,j=1}^{n} \xi_i a_{ij}(x)\xi_j \ge a_*|\xi|^2\,, \quad \textit{for all } \xi \in \mathbb{R}^n,\ x \in \Omega. \tag{3.5}$$

*(In this case, the differential operator L is called **uniformly elliptic**.) We moreover assume that $a_{ij}, b_i, c \in L^\infty(\Omega)$ for all i, j.* □

Definition 3.2 (Weak solution)
*Assume that 3.1 holds and that $f \in L^2(\Omega)$. A function $u \in H_0^1(\Omega)$ is called a **weak solution** of the boundary value problem (3.1), (3.2) if u solves the variational equation*

$$a(u,v) = F(v)\,, \quad \textit{for all } v \in H_0^1(\Omega), \tag{3.6}$$

where a is defined by (3.4) and

$$F(v) = \int_\Omega f(x)v(x)\,dx\,. \tag{3.7}$$

□

Lemma 3.3 *Let $\Omega \subset \mathbb{R}^n$ be open, assume that 3.1 holds. Then (3.4) defines a continuous bilinear form on $H^1(\Omega)$. Moreover, (3.7) defines a continuous linear form on $H^1(\Omega)$ if $f \in L^2(\Omega)$.*

Proof: Setting

$$a^* = n \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 1\le i,j\le n}} |a_{ij}(x)| \,, \quad b^* = n \operatorname*{ess\,sup}_{x\in\Omega} |b(x)|$$

we get for all $u, v \in H^1(\Omega)$

$$\begin{aligned} |a(u,v)| &\le \int_\Omega |\langle \nabla u(x), A(x)\nabla v(x)\rangle| \, dx + \int_\Omega |\langle b(x), \nabla u(x)\rangle \, v(x)| \, dx + \int_\Omega |c(x)u(x)v(x)| \, dx \\ &\le a^* \|\nabla u\|_2 \, \|\nabla v\|_2 + b^* \|\nabla u\|_2 \, \|v\|_2 + \|c\|_\infty \|u\|_2 \, \|v\|_2 \\ &\le (a^* + b^* + \|c\|_\infty) \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \,. \end{aligned}$$

□

Let us now consider the special case $b = c = 0$, that is, the boundary value problem becomes

$$-\sum_{i,j} \partial_j (a_{ij}(x) \partial_i u) = f \,, \quad \text{in } \Omega, \tag{3.8}$$

$$u = 0 \,, \quad \text{on } \partial\Omega. \tag{3.9}$$

Proposition 3.4 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $f \in L^2(\Omega)$, assume 3.1. Then (3.8), (3.9) has a unique weak solution $u \in H_0^1(\Omega)$, and*

$$\|u\|_{H^1(\Omega)} \le C \|F\| \,. \tag{3.10}$$

Here, the norm $\|F\|$ of the functional (3.7) is taken in the dual space of $H_0^1(\Omega)$, and C does not depend on F.

Proof: In order to apply Lax-Milgram, Proposition 1.8, in view of Lemma 3.3 it remains to show that a is $H_0^1(\Omega)$-elliptic. According to our assumptions we have

$$\begin{aligned} a(v,v) &= \int_\Omega \langle \nabla v(x), A(x) \nabla v(x) \rangle \, dx \ge \int_\Omega a_* |\nabla v(x)|^2 \, dx = a_* |v|^2_{H^1(\Omega)} \\ &\ge a_* C_0 \|v\|^2_{H^1(\Omega)} \,, \end{aligned}$$

the last inequality is a consequence of the Poincaré-Friedrichs inequality, see Proposition 2.18. □

Since

$$|F(v)| \le \int_\Omega |f(x)v(x)| \, dx \le \|f\|_2 \, \|v\|_2 \le \|f\|_2 \, \|v\|_{H^1(\Omega)} \,, \tag{3.11}$$

we have

$$\|F\| \le \|f\|_2 \,. \tag{3.12}$$

Therefore, from (3.10) we obtain the estimate

$$\|u\|_{H^1(\Omega)} \le C \|f\|_2 \,. \tag{3.13}$$

We now consider the general form of the operator L from (3.3), that is, the first-order and zero-order terms are also included.

Lemma 3.5 *Let $\Omega \subset \mathbb{R}^n$ be open, assume 3.1. Then there exists $c_* \geq 0$ such that the bilinear form (3.4) satisfies*

$$a(v,v) \geq \frac{a_*}{2}\|\nabla v\|_2^2 - c_*\|v\|_2^2\,, \tag{3.14}$$

for all $v \in H^1(\Omega)$.

This inequality (or a variant of it) is called **Gårding's inequality**.

Proof: We have, with b^* as in the proof of Lemma 3.3,

$$\begin{aligned} a_*\|\nabla v\|_2^2 &\leq \int_\Omega \langle \nabla v(x), A(x)\nabla v(x)\rangle \; dx \\ &= a(v,v) - \int_\Omega \langle b(x), \nabla v(x)\rangle \, v(x) + c(x)v(x)^2 \, dx \\ &\leq a(v,v) + b^* \int_\Omega |\nabla v(x)| \, |v(x)| \, dx + \|c\|_\infty \|v\|_2^2\,. \end{aligned}$$

We have

$$\int_\Omega |\nabla v(x)| \, |v(x)| \, dx = \int_\Omega \left(\sqrt{\varepsilon}|\nabla v(x)|\right)\left(\frac{1}{\sqrt{\varepsilon}}|v(x)|\right) \, dx \leq \frac{\varepsilon}{2}\|\nabla v\|_2^2 + \frac{1}{2\varepsilon}\|v\|_2^2\,.$$

Setting $\varepsilon = a_*/b^*$, we get

$$\frac{a_*}{2}\|\nabla v\|_2^2 \leq a(v,v) + \left(\frac{(b^*)^2}{2a_*} + \|c\|_\infty\right)\|v\|_2^2\,.$$

From this, the assertion follows, setting $c_* = (b^*)^2/2a_* + \|c\|_\infty$. □

The bilinear form a is in general not $H_0^1(\Omega)$-elliptic; it is if (3.5) holds for $c_* = 0$. However, the bilinear form

$$a_\mu(u,v) = a(u,v) + \mu \, \langle u, v\rangle_{L^2(\Omega)} \tag{3.15}$$

is even $H^1(\Omega)$-elliptic for $\mu > \mu_0 := c_*$ since we then have by virtue of Lemma 3.5

$$a_\mu(v,v) \geq \frac{a_*}{2}\|\nabla v\|_2^2 + (\mu - c_*)\|v\|_2^2\,.$$

The bilinear form a_μ is associated to the boundary value problem

$$Lu + \mu u = f\,, \quad \text{in } \Omega, \tag{3.16}$$

$$u = 0\,, \quad \text{on } \partial\Omega, \tag{3.17}$$

the operator L being taken from (3.3). The equation $Lu + \mu u = f$ is also called the **Helmholtz equation**. The foregoing considerations together with the Lax-Milgram theorem 1.8 yield the following result.

Proposition 3.6 *Let $\Omega \subset \mathbb{R}^n$ be open, assume 3.1, let $f \in L^2(\Omega)$. Then there exists a $\mu_0 \in \mathbb{R}$ such that, for every $\mu > \mu_0$, the boundary value problem (3.16), (3.17) has a unique weak solution $u \in H_0^1(\Omega)$.* □

Note that μ_0 may be negative if Gårding's inequality holds for some $c_* < 0$.

The derivation of the variational formulation (3.1) – (3.4) shows that every classical solution of the boundary value problem is a weak solution (we refrain from presenting an exact formulation of this issue). On the other hand, the existence and uniqueness results of the variational theory only yield weak solutions. One then asks whether these weak solutions possess additional smoothness properties (thus, possibly, one might arrive at classical solutions). Such results are called **regularity results**.

Let us come back to the boundary value problem

$$-\Delta u = f\,, \quad \text{in } \Omega, \tag{3.18}$$

$$u = 0\,, \quad \text{auf } \partial\Omega. \tag{3.19}$$

Let us assume for the moment that u is a solution such that $\Delta u \in L^2(\Omega)$ and the following partial integration is valid:

$$\begin{aligned}\int_\Omega \Delta u(x)^2\,dx &= \int_\Omega \sum_{i,j=1}^n \partial_i^2 u(x)\cdot \partial_j^2 u(x)\,dx = \int_\Omega \sum_{i,j=1}^n \partial_j\partial_i u(x)\cdot\partial_i\partial_j u(x)\,dx \\ &= \int_\Omega \sum_{i,j=1}^n (\partial_i\partial_j u(x))^2\,dx\,.\end{aligned}$$

This would imply that

$$|u|^2_{H^2(\Omega)} = \|\Delta u\|_2^2 = \|f\|_2^2\,. \tag{3.20}$$

One therefore may hope that a solution u of (3.18), (3.19) has to be an element of $H^2(\Omega)$ if the right side f belongs to $L^2(\Omega)$.

An estimate like

$$|u|_{H^2(\Omega)} \le C\|f\|_2$$

obtained as above from manipulation of the problem equations is called an **a priori estimate**. This is an important heuristic technique; in a second step (usually the difficult part) one has to prove that the solution actually has the asserted regularity.

For our elliptic boundary value problem, one way of proving H^2 regularity is to replace Δu by a difference quotient approximation for ∇u and to obtain the existence of weak second derivatives from a compactness argument. Our function spaces have infinite dimension; it turns out that the appropriate notion is weak compactness.

We are interested in the relation between weak derivatives and difference quotients. Let $v : \Omega \to \mathbb{R}$ and $h \neq 0$ be given. We consider the **difference quotient**

$$(D_j^h v)(x) = \frac{v(x+he_j) - v(x)}{h}\,, \quad 1 \le j \le n\,. \tag{3.21}$$

Here, e_j denotes the j-th unit vector. We set

$$D^h v = (D_1^h v, \ldots, D_n^h v)\,. \tag{3.22}$$

Lemma 3.7 *Let $\Omega, U \subset \mathbb{R}^n$ be open, let $U \subset\subset \Omega$, $p \in [1,\infty)$. Then*

$$\|D_j^h v\|_{L^p(U)} \le \|\partial_j v\|_{L^p(\Omega)}\,, \quad 1 \le j \le n\,, \tag{3.23}$$

for all $v \in W^{1,p}(\Omega)$ and all $h \neq 0$ with $|h| < \operatorname{dist}(U, \partial\Omega)$.

Proof: First, assume that $v \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$. For all $x \in U$ and $|h| < \operatorname{dist}(U, \partial\Omega)$ we get

$$|v(x + he_j) - v(x)| = \left| \int_0^1 \partial_j v(x + the_j) h \, dt \right| \le |h| \int_0^1 |\partial_j v(x + the_j)| \, dt \,,$$

thus

$$\begin{aligned} \int_U |D_j^h v(x)|^p \, dx &\le \int_U \left(\int_0^1 |\partial_j v(x + the_j)| \, dt \right)^p \, dx \le \int_U \int_0^1 |\partial_j v(x + the_j)|^p \, dt \, dx \\ &= \int_0^1 \int_U |\partial_j v(x + the_j)|^p \, dx \, dt \le \int_0^1 \|\partial_j v\|_{L^p(\Omega)}^p \, dt = \|\partial_j v\|_{L^p(\Omega)}^p \,. \end{aligned}$$

Let now $v \in W^{1,p}(\Omega)$ be arbitrary. According to Proposition 2.13 we choose a sequence $(v_k)_{k \in \mathbb{N}}$ in $C^\infty(\Omega) \cap W^{1,p}(\Omega)$ satisfying $v_k \to v$ in $W^{1,p}(\Omega)$. For fixed $h \neq 0$, (3.23) holds for all v_k. Since $D_j^h v_k \to D_j^h v$ in $W^{1,p}(U)$ as well as $\partial_j v_k \to \partial_j v$ in $L^p(\Omega)$, passing to the limit $k \to \infty$ we obtain (3.23). □

Conversely, we want to derive the existence of a weak derivative $\partial_j v$ from the uniform boundedness of the difference quotients $D_j^h v$. We need a result of functional analysis: if $(f_k)_{k \in \mathbb{N}}$ is a bounded sequence in $L^p(\Omega)$, $1 < p < \infty$, then there exists a subsequence $(f_{k_m})_{m \in \mathbb{N}}$ which converges weakly to some $f \in L^p(\Omega)$. This means that

$$\lim_{m \to \infty} \int_\Omega f_{k_m}(x) \varphi(x) \, dx = \int_\Omega f(x) \varphi(x) \, dx \,, \quad \text{for all } \varphi \in L^q(\Omega), \tag{3.24}$$

where q is the exponent dual to p, $\frac{1}{p} + \frac{1}{q} = 1$. Moreover, the weak limit has the property

$$\|f\|_{L^p(\Omega)} \le \liminf_{m \to \infty} \|f_{k_m}\|_{L^p(\Omega)} \,. \tag{3.25}$$

We write

$$f_{k_m} \rightharpoonup f$$

for weak convergence.

Proposition 3.8 *Let $\Omega, U \subset \mathbb{R}^n$ be open, $U \subset\subset \Omega$, $1 < p < \infty$, $v \in L^p(\Omega)$, $j \in \{1, \ldots, n\}$. Assume that there exists $C > 0$ and $h_0 \le \operatorname{dist}(U, \partial\Omega)$ such that*

$$\|D_j^h v\|_{L^p(U)} \le C \tag{3.26}$$

for all $h \neq 0$, $|h| < h_0$. Then we have $\partial_j v \in L^p(U)$, and

$$\|\partial_j v\|_{L^p(U)} \le C \,. \tag{3.27}$$

Proof: Let $\varphi \in C_0^\infty(U)$ be arbitrary. For $|h| < h_0$ we have

$$\begin{aligned} \int_\Omega v(x) D_j^h \varphi(x) \, dx &= \int_\Omega v(x) \frac{\varphi(x + he_j) - \varphi(x)}{h} \, dx = - \int_\Omega \frac{v(x) - v(x - he_j)}{h} \varphi(x) \, dx \\ &= - \int_\Omega (D_j^{-h} v)(x) \varphi(x) \, dx \,. \end{aligned}$$

We choose a sequence $(h_m)_{m\in\mathbb{N}}$ such that $h_m \to 0$, $h_m \neq 0$, and a $w \in L^p(U)$ satisfying

$$D_j^{-h_m} v \rightharpoonup w\,.$$

Since $D_j^{h_m}\varphi \to \partial_j\varphi$ uniformly in Ω, for sufficiently small h it follows that

$$\begin{aligned}\int_U v(x)\partial_j\varphi(x)\,dx &= \int_\Omega v(x)\partial_j\varphi(x)\,dx = \lim_{m\to\infty}\int_\Omega v(x)(D_j^{h_m}\varphi)(x)\,dx\\ &= \lim_{m\to\infty}\left[-\int_\Omega (D_j^{-h_m}v)(x)\varphi(x)\,dx\right] = -\int_\Omega w(x)\varphi(x)\,dx = -\int_U w(x)\varphi(x)\,dx\,,\end{aligned}$$

so $w = \partial_j v$. □

We come back to the question of the regularity of solutions of the elliptic equation

$$Lu = f\,,\quad \text{in } \Omega, \tag{3.28}$$

where L is the operator given in divergence form as in (3.3). We present a result on **interior regularity**, that is, on the regularity of a weak solution in parts of Ω away from the boundary. Such a weak solution satisfies

$$a(u,v) = \int_\Omega f(x)v(x)\,dx\,,\quad \text{for all } v \in H_0^1(\Omega), \tag{3.29}$$

where a is the bilinear form given by (3.4).

Proposition 3.9
Let Ω, U be open, $U \subset\subset \Omega$, let $f \in L^2(\Omega)$. In addition to 3.1, assume that $a_{ij} \in C^1(\Omega)$ for $1 \le i,j \le n$. Let $u \in H^1(\Omega)$ be a weak solution of $Lu = f$ in the sense of (3.29). Then we have $u \in H^2(U)$, and

$$\|u\|_{H^2(U)} \le C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)})\,. \tag{3.30}$$

Here, C depends only on Ω, U and on the coefficients of L.

Proof: The idea of the proof is to estimate $\|D_k^h \nabla u\|_{L^2(\Omega)}$ and then apply Proposition 3.8 to conclude that the weak second derivatives of u exist and that (3.30) holds.

The variational equation (3.29) uses variations v defined on Ω. In order to obtain results on U, one uses the technique of localization by cut-off functions. For this purpose, we choose open sets $W, Y \subset \mathbb{R}^n$ such that

$$U \subset\subset W \subset\subset Y \subset\subset \Omega\,. \tag{3.31}$$

Let $\zeta \in C_0^\infty(\Omega)$ such that

$$0 \le \zeta \le 1\,,\quad \zeta|U = 1\,,\quad \zeta|(\Omega\setminus W) = 0\,. \tag{3.32}$$

Such a ζ is called a **cut-off function**. For every $v \in H_0^1(\Omega)$ we have

$$\int_\Omega \langle \nabla u(x), A(x)\nabla v(x)\rangle\,dx = \int_\Omega [f(x) - \langle b(x), \nabla u(x)\rangle - c(x)u(x)]v(x)\,dx\,. \tag{3.33}$$

Let $h \neq 0$, $|h|$ sufficiently small, let $k \in \{1, \ldots, n\}$. As a test function in (3.33) we choose

$$v = -D_k^{-h}(\zeta^2 D_k^h u)\,. \tag{3.34}$$

This function replaces $-\Delta u$ as a test function. The difference quotient is defined as before,

$$(D_k^h u)(x) = \frac{1}{h}(u(x + he_k) - u(x)) \tag{3.35}$$

First, we estimate the left side of (3.33) from below. We have

$$\begin{aligned}\int_\Omega \langle \nabla u(x), A(x)\nabla v(x)\rangle \,dx &= -\sum_{i,j=1}^n \int_\Omega a_{ij}(x)\partial_i u(x)\partial_j \left[D_k^{-h}(\zeta^2 D_k^h u)(x)\right]\,dx \\ &= \sum_{i,j=1}^n \int_\Omega \left(D_k^h(a_{ij}\partial_i u)\right)(x)\cdot(\partial_j(\zeta^2 D_k^h u))(x)\,dx \\ &= A_1 + A_2\,,\end{aligned} \tag{3.36}$$

where we set

$$A_1 = \sum_{i,j=1}^n \int_\Omega a_{ij}(x + he_k)(D_k^h(\partial_i u))(x)\cdot\zeta^2(x)(D_k^h(\partial_j u))(x)\,dx\,, \tag{3.37}$$

and

$$\begin{aligned}A_2 = &\sum_{i,j=1}^n \int_\Omega a_{ij}(x + he_k)(D_k^h(\partial_i u))(x)\cdot(D_k^h u)(x)\cdot 2\zeta(x)\partial_j\zeta(x)\,dx \\ &+ \sum_{i,j=1}^n \int_\Omega \partial_i u\cdot D_k^h a_{ij}\cdot[\zeta^2 D_k^h(\partial_j u) + 2\zeta(\partial_j\zeta)(D_k^h u)]\,dx\,.\end{aligned} \tag{3.38}$$

Since L is uniformly elliptic, we get (see 3.1)

$$A_1 = \int_\Omega \left\langle (\zeta(x)D_k^h\nabla u)(x), A(x + he_k)\zeta(x)(D_k^h\nabla u)(x)\right\rangle\,dx \geq a_* \int_\Omega |(D_k^h\nabla u)(x)|^2\zeta^2(x)\,dx\,. \tag{3.39}$$

Next, we want to estimate A_2 from above. As $a_{ij} \in C^1(\Omega)$, we have $a_{ij} \in C^1(\overline{U})$ and

$$|D_k^h a_{ij}(x)| \leq \|\partial_k a_{ij}\|_\infty\,. \tag{3.40}$$

In the following computations will appear constants C_i; they depend on U, W, Y, Ω, a_{ij}, b_i, c, ζ, but not on f, u and h. For A_2 we get, because $\zeta = 0$ outside of W,

$$|A_2| \leq C_1 \int_W \zeta(x)\left[|D_k^h\nabla u(x)|\,|D_k^h u(x)| + |D_k^h\nabla u(x)|\,|\nabla u(x)| + |D_k^h u(x)|\,|\nabla u(x)|\right]\,dx\,. \tag{3.41}$$

Therefore, for every $\varepsilon > 0$ we get

$$\begin{aligned}|A_2| \leq \varepsilon C_1 &\int_\Omega \zeta^2(x)|D_k^h\nabla u(x)|^2\,dx + \frac{C_1}{\varepsilon}\int_W |D_k^h u(x)|^2 + |\nabla u(x)|^2\,dx \\ &+ C_1\int_W |D_k^h u(x)|^2 + |\nabla u(x)|^2\,dx\,.\end{aligned}$$

Setting $\varepsilon = a_*/(2C_1)$, we obtain

$$|A_2| \le \frac{a_*}{2} \int_\Omega \zeta^2(x) |D_k^h \nabla u(x)|^2 \, dx + C_2 \int_W |D_k^h u(x)|^2 + |\nabla u(x)|^2 \, dx \,. \tag{3.42}$$

By Lemma 3.7 we have

$$\int_W |D_k^h u(x)|^2 \, dx \le \int_Y |\nabla u(x)|^2 \, dx \,. \tag{3.43}$$

Altogether, the left side of (3.33) can be estimated as

$$\int_\Omega \langle \nabla u(x), A(x) \nabla v(x) \rangle \, dx \ge \frac{a_*}{2} \int_\Omega \zeta^2(x) |D_k^h \nabla u(x)|^2 \, dx - C_3 \int_Y |\nabla u(x)|^2 \, dx \,. \tag{3.44}$$

Secondly, we consider the right side of (3.33) and denote it by B. Because $v = 0$ outside of Y, we have

$$|B| \le C_4 \int_Y (|f| + |\nabla u| + |u|) |v| \, dx \,. \tag{3.45}$$

Using once more Lemma 3.7, since $\zeta = 0$ outside of W, it follows that

$$\begin{aligned}
\int_Y |v|^2 \, dx &= \int_W |D_k^{-h}(\zeta^2 D_k^h u)|^2 \, dx \le \int_\Omega |\nabla(\zeta^2 D_k^h u)|^2 \, dx \\
&= \int_W 2\zeta^4 |D_k^h \nabla u|^2 + 8\zeta^2 |\nabla \zeta|^2 |D_k^h u|^2 \, dx \\
&\le \int_W 2\zeta^2 |D_k^h \nabla u|^2 \, dx + C_5 \int_Y |\nabla u|^2 \, dx \,.
\end{aligned}$$

It follows that

$$|B| \le \varepsilon \int_\Omega 2\zeta^2 |D_k^h \nabla u|^2 \, dx + \frac{C_6}{\varepsilon} \int_Y f^2 + |\nabla u|^2 + u^2 \, dx + C_5 \int_Y |\nabla u|^2 \, dx \,. \tag{3.46}$$

Setting $\varepsilon = a_*/8$, we get

$$|B| \le \frac{a_*}{4} \int_\Omega \zeta^2 |D_k^h \nabla u|^2 \, dx + C_7 \int_Y f^2 + |\nabla u|^2 + u^2 \, dx \,. \tag{3.47}$$

All in all, we now have obtained from (3.33) the estimate

$$\frac{a_*}{4} \int_U |D_k^h \nabla u|^2 \, dx \le \frac{a_*}{4} \int_\Omega \zeta^2 |D_k^h \nabla u|^2 \, dx \le C_8 \int_Y f^2 + |\nabla u|^2 + u^2 \, dx \,. \tag{3.48}$$

It now follows from Proposition 3.8, since C_8 does not depend on h, that

$$\partial_k \nabla u \in L^2(U) \,, \quad \|\partial_k \nabla u\|_{L^2(U)}^2 \le C_9 \int_Y f^2 + |\nabla u|^2 + u^2 \, dx \,, \tag{3.49}$$

and therefore $u \in H^2(U)$, as k was arbitrary. Finally, we want to estimate

$$\int_Y |\nabla u|^2 \, dx$$

from above. Choose $\zeta \in C_0^\infty(\Omega)$ such that $\zeta|Y = 1$. We test the variational equation with $v = \zeta^2 u$. This yields as above (but simpler)

$$\begin{aligned} a_* \int_\Omega \zeta^2 |\nabla u|^2 \, dx &\le \int_\Omega \left\langle \nabla u, A\zeta^2 \nabla u \right\rangle \, dx = \int_\Omega \langle \nabla u, A\nabla v\rangle \, dx - \int_\Omega \langle \nabla u, 2Au\zeta\nabla\zeta\rangle \, dx \\ &\le C_{10} \int_\Omega (|f| + |\nabla u| + |u|)|u|\zeta \, dx \\ &\le \frac{a_*}{2} \int_\Omega \zeta^2 |\nabla u|^2 \, dx + C_{11} \int_\Omega f^2 + u^2 \, dx \, . \end{aligned}$$

Consequently,

$$\int_Y |\nabla u|^2 \, dx \le C_{12} \int_\Omega f^2 + u^2 \, dx \, . \tag{3.50}$$

Combining (3.49) and (3.50) we obtain

$$|u|^2_{H^2(U)} \le C_{13}(\|f\|^2_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega)})$$

and then, again because of (3.50), the assertion. □

Note that the boundary behaviour of the solution u does not matter for this result.

Proposition 3.9 shows that, in the given situation, the solution u gains two orders of differentiability with respect to the right side f. (This is optimal, since one certainly cannot expect more.)

This also applies when $f \in H^k(\Omega)$.

Proposition 3.10 *In the situation of Proposition 3.9, assume in addition that for some $k \in \mathbb{N}$ we have $a_{ij}, b_i, c \in C^{k+1}(\Omega)$ and $f \in H^k(\Omega)$. Then $u \in H^{k+2}(U)$ and*

$$\|u\|_{H^{k+2}(U)} \le C(\|f\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)}) \, , \tag{3.51}$$

where C depends only on k, Ω, U and on the coefficients of L.

Proof: One uses induction with respect to k. Proposition 3.9 both yields the induction basis $k = 0$ and the essential tool for the induction step. It is carried out in detail in the book L. Evans (Partial Differential Equations), section 6.3. □

4 Boundary Conditions, Traces

Inhomogeneous Dirichlet boundary conditions. The only boundary condition we have discussed so far is "$u = 0$ on $\partial\Omega$", the so-called homogeneous Dirichlet boundary condition. Let us now consider the problem with an inhomogeneous Dirichlet boundary condition,

$$\begin{aligned} -\Delta u &= f\,, \quad \text{in } \Omega, \\ u &= g\,, \quad \text{on } \partial\Omega. \end{aligned} \tag{4.1}$$

In the homogeneous case we did not treat "$u = 0$ on $\partial\Omega$" as a constraining equation, but instead incorporated it into the underlying function space by requiring a solution u to belong to $H_0^1(\Omega)$. This can be extended to (4.1) as follows. Assume that $g \in H^1(\Omega)$ is given. Instead of "$u = g$ on $\partial\Omega$" we impose the condition

$$u - g \in H_0^1(\Omega)\,. \tag{4.2}$$

As $H_0^1(\Omega) = \overline{C_0^\infty(\Omega)}$ by definition, this means that there exists a sequence $\varphi_n \in C_0^\infty(\Omega)$ with $\varphi_n \to u - g$ in $H_0^1(\Omega)$. As before, the weak formulation of the differential equation reads

$$\int_\Omega f\varphi\,dx = \int_\Omega (-\Delta u)\varphi\,dx = \int_\Omega \langle \nabla u, \nabla\varphi\rangle\;dx\,, \quad \text{for all } \varphi \in C_0^\infty(\Omega),$$

or, with a and F as before,

$$a(u,\varphi) = F(\varphi)\,, \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{4.3}$$

This is equivalent to

$$a(u,v) = F(v)\,, \quad \text{for all } v \in H_0^1(\Omega), \tag{4.4}$$

because for every $v \in H_0^1(\Omega)$ we can find $\varphi_n \in C_0^\infty(\Omega)$ with $\varphi_n \to v$ in $H_0^1(\Omega)$, and we may pass to the limit in (4.3).

The variational formulation (4.2), (4.4) can be rewritten in terms of $w = u - g$. We then have the problem to find w,

$$w \in H_0^1(\Omega)\,, \quad a(w,v) = F(v) - a(g,v)\,, \quad \text{for all } v \in H_0^1(\Omega). \tag{4.5}$$

By the results of the previous chapter, this problem has a unique solution $w \in H_0^1(\Omega)$.

These considerations apply to the more general problem

$$\begin{aligned} Lu &= f\,, \quad \text{in } \Omega, \\ u &= g\,, \quad \text{on } \partial\Omega. \end{aligned} \tag{4.6}$$

Proposition 4.1 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $f \in L^2(\Omega)$, $g \in H^1(\Omega)$, assume 3.1. Then (4.6) has a unique weak solution $u \in H^1(\Omega)$, and*

$$\|u\|_{H^1(\Omega)} \le C(\|f\|_{L^2(\Omega)} + \|g\|_{H^1(\Omega)}) \tag{4.7}$$

for some constant C independent from f and g.

Proof: From Proposition 3.4, applied to (4.5), we get that

$$\|w\|_{H^1(\Omega)} \le C_1(C_2\|f\|_{L^2(\Omega)} + C_3\|g\|_{H^1(\Omega)})$$

Since $\|u\|_{H^1(\Omega)} \le \|w\|_{H^1(\Omega)} + \|g\|_{H^1(\Omega)}$, the assertion follows. □

The procedure above gives rise to the question: How are functions $g : \partial\Omega \to \mathbb{R}$ related to functions $g \in H^1(\Omega)$?

If $v : \overline{\Omega} \to \mathbb{R}$ is continuous, then its restrictions to Ω and $\partial\Omega$ are continuous. One would like to define an operator γ which maps $H^1(\Omega)$ to a suitable class of functions defined on $\partial\Omega$, such that for continuous functions it coincides with their restrictions.

Traces on hyperplanes. One first considers a function $v \in C_0^\infty(\mathbb{R}^n)$ and its restriction to the hyperplane $\mathbb{R}^{n-1} \times \{0\}$. This defines a function

$$\gamma v : \mathbb{R}^{n-1} \to \mathbb{R}\,, \quad (\gamma v)(x') = v(x', 0)\,, \quad x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}\,, \tag{4.8}$$

and an operator

$$\gamma : C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^{n-1})\,. \tag{4.9}$$

The function γv is called the **trace** of v on the hyperplane $\mathbb{R}^{n-1} \times \{0\}$, and the operator γ is called a **trace operator**.

Let X be a Banach space of functions over $\mathbb{R}^n$ such that $C_0^\infty(\Omega)$ is dense in X. One wants to find a Banach space Y of functions over $\mathbb{R}^{n-1}$ such that γ can be extended to a linear and continuous operator from X to Y. It turns out that for $X = H^1(\mathbb{R}^n)$ one can choose $Y = L^2(\mathbb{R}^{n-1})$. However, $\gamma(H^1(\mathbb{R}^n))$ turns out to be a proper subset of $L^2(\mathbb{R}^{n-1})$; in order to describe the class of functions which are traces of functions in H^1, one has to introduce Sobolev spaces of fractional order.

One way to define those spaces makes use of the Fourier transform. For $v \in C_0^\infty(\Omega)$ the Fourier transform is defined by

$$(\mathcal{F}v)(\xi) = \hat{v}(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} v(x) e^{-i\langle x, \xi\rangle}\, dx\,, \quad \xi \in \mathbb{R}^n\,. \tag{4.10}$$

It has the properties

$$\begin{aligned} \widehat{\partial_j v}(\xi) &= i\xi_j \hat{v}(\xi)\,, \\ \widehat{x_j v}(\xi) &= i\partial_j \hat{v}(\xi)\,. \end{aligned} \tag{4.11}$$

Since

$$\int_{\mathbb{R}^n} u(x)v(x)\, dx = \int_{\mathbb{R}^n} \hat{u}(\xi)\hat{v}(\xi)\, d\xi\,, \quad u, v \in C_0^\infty(\Omega)\,, \tag{4.12}$$

it can be extended to an isometry $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.

Lemma 4.2 *For every $k \in \mathbb{N}$ there exists $C_k > 0$ such that*

$$\frac{1}{C_k}(1 + |\xi|^2)^k \le \sum_{|\alpha| \le k} \xi^{2\alpha} \le C_k(1 + |\xi|^2)^k\,, \quad \textit{for all } \xi \in \mathbb{R}^n. \tag{4.13}$$

Proof: Exercise. □

Proposition 4.3
Let $k \in \mathbb{N}$. *Then*

$$\langle u, v\rangle_{k,\mathcal{F}} = \int_{\mathbb{R}^n} (1+|\xi|^2)^k \hat{u}(\xi)\overline{\hat{v}(\xi)}\, d\xi \tag{4.14}$$

defines a scalar product on $C_0^\infty(\mathbb{R}^n)$. *The associated norm*

$$\|v\|_{k,\mathcal{F}}^2 = \int_{\mathbb{R}^n} (1+|\xi|^2)^k |\hat{v}(\xi)|^2\, d\xi \tag{4.15}$$

on $C_0^\infty(\mathbb{R}^n)$ *is equivalent to the Sobolev norm* $\|\cdot\|_{H^k(\mathbb{R}^n)}$, *that is, there exists* $C_k > 0$ *such that*

$$\frac{1}{C_k}\int_{\mathbb{R}^n} (1+|\xi|^2)^k |\hat{v}(\xi)|^2\, d\xi \le \|v\|_{H^k(\mathbb{R}^n)}^2 \le C_k \int_{\mathbb{R}^n} (1+|\xi|^2)^k |\hat{v}(\xi)|^2\, d\xi \tag{4.16}$$

for all $v \in C_0^\infty(\Omega)$.

Proof: Let $v \in C_0^\infty(\mathbb{R}^n)$. Since

$$\int_{\mathbb{R}^n} |v(x)|^2\, dx = \int_{\mathbb{R}^n} |\hat{v}(\xi)|^2\, d\xi\,, \tag{4.17}$$

we get

$$\begin{aligned}
\|v\|_{H^k(\mathbb{R}^n)}^2 &= \sum_{|\alpha|\le k} \int_{\mathbb{R}^n} |\partial^\alpha v(x)|^2\, dx = \sum_{|\alpha|\le k} \int_{\mathbb{R}^n} |\widehat{\partial^\alpha v}(\xi)|^2\, d\xi \\
&= \sum_{|\alpha|\le k} \int_{\mathbb{R}^n} |\xi^\alpha \hat{v}(\xi)|^2\, d\xi = \int_{\mathbb{R}^n} |\hat{v}(\xi)|^2 \sum_{|\alpha|\le k} \xi^{2\alpha}\, d\xi\,.
\end{aligned}$$

From Lemma 4.2 we now obtain (4.16). □

Proposition 4.4 (Sobolev space with fractional norm)
Let $s \ge 0$. *Then*

$$\langle u, v\rangle_{s,\mathcal{F}} = \int_{\mathbb{R}^n} (1+|\xi|^2)^s \hat{u}(\xi)\overline{\hat{v}(\xi)}\, d\xi\,, \tag{4.18}$$

$$\|v\|_{s,\mathcal{F}} = \sqrt{\langle v, v\rangle_{s,\mathcal{F}}}\,, \tag{4.19}$$

defines a scalar product on the space

$$H^s(\mathbb{R}^n) = \{v : v \in L^2(\mathbb{R}^n),\ \|v\|_{s,\mathcal{F}} < \infty\}\,, \tag{4.20}$$

and $H^s(\mathbb{R}^n)$ *is a Hilbert space.*

The spaces $H^s(\mathbb{R}^n)$ are called **Bessel potential spaces**.

Proof: We consider $X = L^2(\mathbb{R}^n; \mu)$ with the measure

$$\mu = f\lambda\,, \quad f(\xi) = (1+|\xi|^2)^s\,,$$

that is,

$$\int_{\mathbb{R}^n} w\,d\mu = \int_{\mathbb{R}^n} (1+|\xi|^2)^s w(\xi)\,d\xi\,.$$

X is a Hilbert space, and $H^s(\mathbb{R}^n) = \{v : v \in L^2(\mathbb{R}^n)\,,\ \hat{v} \in L^2(\mathbb{R}^n;\mu)\}$. Since the Fourier transform is linear, $H^s(\mathbb{R}^n)$ is a vector space, and

$$\langle u, v\rangle_{s,\mathcal{F}} = \langle \hat{u}, \hat{v}\rangle_X\ , \quad \|v\|_{s,\mathcal{F}} = \|\hat{v}\|_X\,,$$

defines a scalar product and its associated norm on $H^s(\mathbb{R}^n)$. For $v \in H^s(\mathbb{R}^n)$ we have $\|v\|_{L^2} = \|\hat{v}\|_{L^2} \le \|\hat{v}\|_X = \|v\|_{s,\mathcal{F}}$. Let $\{v_m\}$ be a Cauchy sequence in $H^s(\mathbb{R}^n)$. Then $\{\hat{v}_m\}$ is a Cauchy sequence in X, and $\{v_m\}$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. Therefore, there exist $w \in X$ and $v \in L^2(\mathbb{R}^n)$ such that $\hat{v}_m \to w$ in X and $v_m \to v$ in $L^2(\mathbb{R}^n)$, thus moreover $\hat{v}_m \to w$ in $L^2(\mathbb{R}^n)$ and $\hat{v}_m \to \hat{v}$ in $L^2(\mathbb{R}^n)$. It follows that $w = \hat{v}$ and therefore $v \in H^s(\mathbb{R}^n)$. □

Lemma 4.5 *Let $k \in \mathbb{N}$. Then $C_0^\infty(\mathbb{R}^n)$ is dense in $H^k(\mathbb{R}^n)$ with respect to $\|\cdot\|_{H^k(\mathbb{R}^n)}$, that is, we have $H^k(\mathbb{R}^n) = H_0^k(\mathbb{R}^n)$.*

Proof: Since $C^\infty(\mathbb{R}^n) \cap H^k(\mathbb{R}^n)$ is dense in $H^k(\mathbb{R}^n)$ by Proposition 2.13, it suffices to find for every $v \in C^\infty(\mathbb{R}^n) \cap H^k(\mathbb{R}^n)$ a sequence $(v_m)_{m\in\mathbb{N}}$ in $C_0^\infty(\mathbb{R}^n)$ such that $v_m \to v$ in $H^k(\mathbb{R}^n)$. To this purpose, choose $\psi \in C_0^\infty(\mathbb{R}^n)$ such that $\psi(x) = 1$ for $|x| \le 1$ and $\psi(x) = 0$ for $|x| \ge 2$. We set

$$v_m(x) = \psi\left(\frac{x}{m}\right) v(x)\,. \tag{4.21}$$

Then for all multi-indices α with $|\alpha| \le k$ we have

$$|\partial^\alpha(v - v_m)(x)| = \left|\partial^\alpha\left((1-\psi(\frac{x}{m}))v(x)\right)\right| \quad \begin{cases} \le C\sum_{|\beta|\le|\alpha|} |\partial^\beta v(x)|\,, & |x| > m\,,\\ = 0\,, & |x| \le m\,,\end{cases}$$

where C does not depend on m. It follows that

$$\int_{\mathbb{R}^n} |\partial^\alpha(v - v_m)(x)|^2\,dx \le C \sum_{|\beta|\le|\alpha|} \int_{|x|>m} |\partial^\beta v(x)|^2\,dx \to 0$$

for $m \to \infty$, since $\partial^\beta v \in L^2(\mathbb{R}^n)$ for all $|\beta| \le k$. □

Lemma 4.6 *For $s \ge 0$ the space $C_0^\infty(\mathbb{R}^n)$ is dense in $H^s(\mathbb{R}^n)$ with respect to $\|\cdot\|_{s,\mathcal{F}}$.*

Proof: This is proved, for example, in the book of Wloka: Partial Differential Equations (in German and in English), on p.96 of the German edition. □

With the aid of Lemma 4.5 and Lemma 4.6 one can prove that, for integer k, the two definitions of $H^k(\mathbb{R}^n)$ (based on weak derivatives, or on the Fourier transform) yield the same space and that the corresponding norms are equivalent on $H^k(\mathbb{R}^n)$ (and not only on the subspace $C_0^\infty(\mathbb{R}^n)$, as was obtained in Proposition 4.3). The proof is analogous to that of Proposition 4.4.

Lemma 4.7 *For $0 \le s \le t$ we have*

$$H^t(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)\,. \tag{4.22}$$

The embedding $j : H^t(\mathbb{R}^n) \to H^s(\mathbb{R}^n)$ is continuous.

Proof: From $(1+|\xi|^2)^s \le (1+|\xi|^2)^t$ we obtain $\|v\|_{s,\mathcal{F}} \le \|v\|_{t,\mathcal{F}}$. $\square$

These results now enable us to define the trace operator γ on $H^s(\mathbb{R}^n)$ such that

$$\gamma v : \mathbb{R}^{n-1} \to \mathbb{R}\,, \quad (\gamma v)(x') = v(x',0)\,, \quad x' = (x_1,\ldots,x_{n-1}) \in \mathbb{R}^{n-1}\,, \tag{4.23}$$

for smooth functions v.

Proposition 4.8 (Trace theorem)
Let $s \in \mathbb{R}$, $s > \frac{1}{2}$. There exists a unique linear and continuous mapping

$$\gamma : H^s(\mathbb{R}^n) \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1}) \tag{4.24}$$

such that (4.23) holds for all $v \in C_0^\infty(\mathbb{R}^n)$.

Proof: It suffices to prove that there exists $C > 0$ such that

$$\|\gamma\varphi\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \le C\|\varphi\|_{H^s(\mathbb{R}^n)}\,, \quad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^n), \tag{4.25}$$

because then γ is linear and continuous on the dense subset $C_0^\infty(\mathbb{R}^n)$ of $H^s(\mathbb{R}^n)$, and therefore can be extended uniquely to a linear continuous mapping on $H^s(\mathbb{R}^n)$, by a theorem of functional analysis.

Let $\varphi \in C_0^\infty(\mathbb{R}^n)$. We define

$$g(x') = \varphi(x',0)\,, \quad x' \in \mathbb{R}^{n-1}\,. \tag{4.26}$$

We fix $x' \in \mathbb{R}^{n-1}$ and define

$$\psi(x_n) = \varphi(x',x_n)\,, \quad x_n \in \mathbb{R}\,. \tag{4.27}$$

Applying the Fourier transform in $\mathbb{R}$, we get

$$\hat{\psi}(\xi_n) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(x',x_n) e^{-ix_n\xi_n}\,dx_n\,, \tag{4.28}$$

thus

$$\begin{aligned} g(x') = \varphi(x',0) = \psi(0) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{\psi}(\xi_n) e^{i0\xi_n}\,d\xi_n \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(x',x_n) e^{-ix_n\xi_n}\,dx_n\,d\xi_n\,. \end{aligned}$$

Applying the Fourier transform in $\mathbb{R}^{n-1}$ yields

$$\begin{aligned}\hat{g}(\xi') &= (2\pi)^{-\frac{n-1}{2}} \int_{\mathbb{R}^{n-1}} g(x') e^{-i\langle x',\xi'\rangle}\, dx' \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} \varphi(x',x_n) e^{-ix_n\xi_n} e^{-i\langle x',\xi'\rangle}\, dx'\, dx_n\, d\xi_n \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{\varphi}(\xi',\xi_n)\, d\xi_n\,,\end{aligned}$$

in the last line the Fourier transform has been applied in $\mathbb{R}^n$. We continue with

$$\begin{aligned}\|g\|^2_{s-\frac{1}{2},\mathcal{F}} &= \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{s-\frac{1}{2}} |\hat{g}(\xi')|^2\, d\xi' = \frac{1}{2\pi} \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{s-\frac{1}{2}} \left| \int_{\mathbb{R}} \hat{\varphi}(\xi',\xi_n)\, d\xi_n \right|^2 d\xi' \\ &= \frac{1}{2\pi} \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{s-\frac{1}{2}} \left| \int_{\mathbb{R}} \hat{\varphi}(\xi',\xi_n)(1+|\xi|^2)^{\frac{s}{2}}(1+|\xi|^2)^{-\frac{s}{2}}\, d\xi_n \right|^2 d\xi' \\ &\le \frac{1}{2\pi} \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{s-\frac{1}{2}} \int_{\mathbb{R}} |\hat{\varphi}(\xi',\xi_n)|^2 (1+|\xi|^2)^s\, d\xi_n \cdot \int_{\mathbb{R}} (1+|\xi|^2)^{-s}\, d\xi_n\, d\xi'\,.\end{aligned}$$

Now we obtain, for $s > \frac{1}{2}$ as assumed,

$$\begin{aligned}\int_{\mathbb{R}} (1+|\xi|^2)^{-s}\, d\xi_n &= \int_{\mathbb{R}} (1+|\xi'|^2+\xi_n^2)^{-s}\, d\xi_n \\ &= (1+|\xi'|^2)^{-s+\frac{1}{2}} \underbrace{\int_{\mathbb{R}} (1+y^2)^{-s}\, dy}_{=:c(s)} \qquad \Big(\text{substitution } \xi_n = (1+|\xi'|^2)^{\frac{1}{2}} y\Big)\,.\end{aligned}$$

It now follows (since $c(s) < \infty$ because of $s > \frac{1}{2}$) that

$$\|g\|^2_{s-\frac{1}{2},\mathcal{F}} \le \frac{c(s)}{2\pi} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} |\hat{\varphi}(\xi',\xi_n)|^2 (1+|\xi|^2)^s\, d\xi_n\, d\xi' = \frac{c(s)}{2\pi} \|\varphi\|^2_{s,\mathcal{F}}\,.$$

□

Proposition 4.8 means that when one restricts a function, one loses one half (generalized) order of differentiation for each dimension. Restricting a function in $H^k(\mathbb{R}^n)$ to $\mathbb{R}^{n-2}$, one thus obtains a function in $H^{k-1}(\mathbb{R}^{n-2})$.

Corollary 4.9 *For $k \in \mathbb{N}$, $k \ge 1$, the trace operator $\gamma : H^k(\mathbb{R}^n) \to H^{k-1}(\mathbb{R}^{n-1})$ is well-defined, linear and continuous.*

Proof: This is a direct consequence of 4.8 and Lemma 4.7, the latter applied with $s = k-1$ and $t = k - \frac{1}{2}$. □

If one is satisfied with the corollary (whose direct proof is simpler than that of the proposition above), one can avoid fractional spaces. However, because of the following result Proposition 4.8 precisely describes the range of the trace operator, that is, for which functions g on $\mathbb{R}^{n-1}$ one can find a Sobolev function whose trace coincides with g.

Proposition 4.10 *The trace operator γ considered in Proposition 4.8 is surjective.*

Proof: Given $g \in C_0^\infty(\mathbb{R}^{n-1})$ one can define a function v by an explicit formula and prove that $\gamma v = g$ with similar computations than those in the proof of Proposition 4.8. For details we refer again to the book of Wloka. $\square$

Sobolev spaces over the boundary of a region. Unlike spaces of continuous functions (which can be defined directly on arbitrary subsets of $\mathbb{R}^n$), spaces of integrable functions on lower-dimensional sets require some consideration; the integrals over the boundary in which one is interested are surface integrals. In order that they are meaningful, the boundary has to have a certain regularity. Besides addressing that aspect, the following definition does not allow Ω to lie on both sides of its boundary.

Definition 4.11 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded.*
*(i) Let $x \in \partial\Omega$. A **local representation of Ω near x** consists of an open set $U \subset \mathbb{R}^n$ with $x \in U$ and of a mapping $\alpha : B \to U$, where $B = \{y : |y| < 1\}$ is the unit ball in $\mathbb{R}^n$, such that $\alpha(0) = x$, α is bijective, and*

$$\alpha(y) = \alpha(y_1, \ldots, y_n) \in \begin{cases} \partial\Omega\,, & \text{if } y_n = 0, \\ \Omega\,, & \text{if } y_n > 0, \\ \mathbb{R}^n \setminus (\Omega \cup \partial\Omega)\,, & \text{if } y_n < 0. \end{cases} \tag{4.29}$$

*(ii) Let $k \in \mathbb{N}$. A local representation (α, U) of Ω near x has **regularity C^k** (or **Lipschitz**, respectively), if α and α^{-1} are C^k-functions (or Lipschitz functions, respectively).*
*(iii) The open set Ω has **regularity C^k** (or **Lipschitz**, respectively), if there exist finitely many points $x^j \in \partial\Omega$, $j \in J$, and local representations (α_j, U_j) near x^j of corresponding regularity such that*

$$\partial\Omega \subset \bigcup_{j \in J} U_j\,. \tag{4.30}$$

*The family $(\alpha_j, U_j)_{j \in J}$ is called a **C^k** resp. **Lipschitz representation of Ω**.*

(4.30) means that $(U_j)_{j \in J}$ forms an open covering of $\partial\Omega$.

A side remark: One can extend this definition to open sets which are unbounded; in (iii), one allows for infinitely many points x^j and requires that every $x \in \partial\Omega$ lies in a ball which intersects only finitely many of the U_j. A covering $(U_j)_{j \in J}$ of $\partial\Omega$ having this property is called **locally finite**.

In order to define Sobolev spaces (in particular, L^p spaces) over $\partial\Omega$, one uses a two-step procedure: localization and "flattening out". In the first step, we choose a finite open covering $(U_j)_{j \in J}$ of $\partial\Omega$ and a C^∞-partition of unity $(\beta_j)_{j \in J}$ for that covering, that is,

$$\beta_j \in C_0^\infty(U_j)\,, \quad 0 \le \beta_j \le 1\,, \quad \sum_{j \in J} \beta_j(x) = 1 \quad \text{for all } x \in \partial\Omega.$$

For functions $v : \partial\Omega \to \mathbb{R}$ we then have

$$\operatorname{supp}(\beta_j v) \subset\subset U_j\,, \quad v = \sum_{j \in J} \beta_j v\,.$$

For the second step, let $(\alpha_j, U_j)_{j \in J}$ be a representation of Ω according to Definition 4.11, and $(\beta_j)_{j \in J}$ be a partition of unity as above. We consider the functions

$$(\beta_j v) \circ \alpha_j : B \cap (\mathbb{R}^{n-1} \times \{0\})\,.$$

They have compact support, because each $\beta_j v$ has compact support and each α_j^{-1} is continuous. Next, we identify $B \cap (\mathbb{R}^{n-1} \times \{0\})$ with the unit ball $\hat{B}$ in $\mathbb{R}^{n-1}$. In this manner, we obtain functions

$$(\beta_j v) \circ \alpha_j : \hat{B} \to \mathbb{R}\,.$$

Extending by 0 outside of $\hat{B}$ as usual, we end up with functions

$$(\beta_j v) \circ \alpha_j : \mathbb{R}^{n-1} \to \mathbb{R}\,. \tag{4.31}$$

The idea now is to use function spaces for these functions in order to construct function spaces for functions $v : \partial\Omega \to \mathbb{R}$. One wants to define

$$H^s(\partial\Omega) = \{v : \quad v : \partial\Omega \to \mathbb{R}\,,\, (\beta_j v) \circ \alpha_j \in H^s(\mathbb{R}^{n-1}) \text{ for all } j \in J\} \tag{4.32}$$

and equip it with the scalar product

$$\langle u, v\rangle_{H^s(\partial\Omega)} = \sum_{j\in J} \langle \tilde{u}_j, \tilde{v}_j\rangle_{H^s(\mathbb{R}^{n-1})}\,. \tag{4.33}$$

where $\tilde{u}_j = (\beta_j u) \circ \alpha_j$ and $\tilde{v}_j = (\beta_j v) \circ \alpha_j$. For $s = 0$ this becomes

$$\langle u, v\rangle_{L^2(\partial\Omega)} = \sum_{j\in J} \int_{\mathbb{R}^{n-1}} \tilde{u}_j(x)\tilde{v}_j(x)\,dx\,. \tag{4.34}$$

This construction depends on the choice of the representation (α_j, U_j) and of the partition of unity (β_j). The question is whether the resulting spaces are independent from these choices. It turns out that they are, provided Ω has high enough regularity.

Proposition 4.12 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Assume that Ω has regularity C^k, let $0 \le s \le k$. Then (4.32) defines a Hilbert space $H^s(\partial\Omega)$ when equipped with the scalar product (4.33). Different choices of (α_j, U_j) and of (β_j) yield the same space, and the norms generated by the corresponding scalar products are equivalent.*

Proof: We will not present the proof of this result and refer to the book of Wloka. $\square$

The procedure which defines $H^s(\partial\Omega)$ via $H^s(\mathbb{R}^{n-1})$ can be also be used to obtain a trace theorem on $H^s(\partial\Omega)$ from the trace theorem on $H^s(\mathbb{R}^n)$, Theorem 4.8. We restrict ourselves to the case where s is integer.

Proposition 4.13 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with regularity C^k, $k \in \mathbb{N}$, $k \ge 1$. Then the trace operator $\gamma : H^k(\Omega) \to H^{k-\frac{1}{2}}(\partial\Omega)$ is well-defined, linear and continuous.*

As in the case of a hyperplane, the trace operator ist first defined for smooth functions and then extended to the Sobolev space by continuity. For the first step we need the following result.

Proposition 4.14 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with regularity C^k, $k \in \mathbb{N}$, $k \ge 1$. Then*

$$C_0^\infty(\mathbb{R}^n)|\Omega = \{v|\Omega : v \in C_0^\infty(\mathbb{R}^n)\}$$

is a dense subspace of $H^k(\Omega)$.

Proof: Again we will not prove this result. $\square$

We also need the following result. Here, B denotes the unit ball in $\mathbb{R}^n$ as before, and $B_+ = \{(y', y_n) \in B : y_n > 0\}$.

Proposition 4.15 *Let $v \in C^k(\overline{B_+})$ with* $\operatorname{supp}(v) \subset\subset B$. *Then there exists an extension $w \in C_0^k(B)$ of v and a $c > 0$ such that*

$$\|w\|_{H^k(B)} \le c\|v\|_{H^k(B_+)} . \tag{4.35}$$

The constant c can be chosen independent from v.

Proof: We present the case $k = 1$. We set $w = v$ on $\overline{B_+}$ and have to define w on $B_- = \{(y', y_n) \in B : y_n < 0\}$. We define for $y_n < 0$

$$w(y', y_n) = a_1 v(y', -y_n) + a_2 v(y', -\frac{y_n}{2}) \tag{4.36}$$

and determine $a_1, a_2 \in \mathbb{R}^n$ such that ("0–" denotes the limit from below)

$$w(y', 0-) = v(y', 0) , \quad \partial_n w(y', 0-) = \partial_n v(y', 0) . \tag{4.37}$$

The equations in (4.37) are equivalent to

$$a_1 + a_2 = 1 , \quad -a_1 - \frac{a_2}{2} = 1 .$$

Therefore, setting $a_1 = -3$ and $a_2 = 4$ we obtain an extension $w \in C_0^k(B) \subset H^k(B)$ of v, and the integrals of w and its derivatives over B_- can be estimated against integrals of v over B_+ using (4.36).

For $k > 1$ one replaces (4.36) by a similar ansatz with $k + 1$ coefficients. $\square$

Proof of Proposition 4.13. Assume that $v \in H^k(\Omega)$ is a function which is a restriction of an element of $C_0^\infty(\mathbb{R}^n)$. Let (α_j, U_j) be a C^k representation of Ω with a corresponding partition of unity (β_j), let $B_+ = \{(y', y_n) \in B : y_n > 0\}$. Set $\overline{v}_j = (\beta_j v) \circ \alpha_j$. Then $\operatorname{supp}(\overline{v}_j) \subset\subset B$ and $\overline{v}_j \in C^k(\overline{B_+})$. According to Proposition 4.15 we choose an extension $\tilde{v}_j \in C_0^k(B)$ of $\overline{v}_j$ with

$$\|\tilde{v}_j\|_{H^k(B)} \le c\|\overline{v}_j\|_{H^k(B_+)} , \tag{4.38}$$

where c only depends on k. Extending by zero we get $\tilde{v}_j \in C_0^k(\mathbb{R}^n) \subset H^k(\mathbb{R}^n)$. We consider its trace

$$\tilde{\gamma}\tilde{v}_j : \mathbb{R}^{n-1} \to \mathbb{R}$$
$$(\tilde{\gamma}\tilde{v}_j)(y') = \tilde{v}_j(y', 0) = \overline{v}_j(x', 0) = ((\beta_j v) \circ \alpha_j)(y', 0) , \quad \text{for all } (y', 0) \in B.$$

By Proposition 4.8 we have

$$\|\tilde{\gamma}\tilde{v}_j\|_{H^{k-1/2}(\mathbb{R}^{n-1})} \le C_1 \|\tilde{v}_j\|_{H^k(\mathbb{R}^n)} \tag{4.39}$$

for some constant C_1 independent from v. Moreover,

$$\|\tilde{v}_j\|_{H^k(\mathbb{R}^n)} \le c\|\tilde{v}_j\|_{H^k(\mathbb{R}^n_+)} \tag{4.40}$$

according to Proposition 4.15. Furthermore,

$$\|\tilde{v}_j\|_{H^k(\mathbb{R}^n_+)} = \|\tilde{v}_j\|_{H^k(B_+)} = \|(\beta_j v)\circ\alpha_j\|_{H^k(B_+)} \leq C_2\|(\beta_j v)\|_{H^k(U_j\cap\Omega)} \leq C_3\|v\|_{H^k(\Omega)}\,. \quad (4.41)$$

In view of all these estimates, we finally obtain, setting $s = k - 1/2$,

$$\|\gamma v\|^2_{H^s(\partial\Omega)} = \sum_{j\in J} \|(\beta_j v)\circ\alpha_j\|^2_{H^s(\mathbb{R}^{n-1})} = \sum_{j\in J} \|\tilde{\gamma}_j\tilde{v}_j\|^2_{H^s(\mathbb{R}^{n-1})} \leq C\|v\|^2_{H^k(\Omega)}$$

for some constant C independent from v. Because the restrictions of functions from $C_0^\infty(\mathbb{R}^n)$ are dense in $H^k(\Omega)$ by Proposition 4.14, the trace operator γ can be uniquely extended to a linear continuous $\gamma : H^k(\Omega) \to H^{k-1/2}(\partial\Omega)$. □

A proposition analogous to 4.13 can be obtained for the case where the domain Ω has only Lipschitz regularity (as, for example, when Ω is a polyhedron). One also obtains the inclusions $H^t(\partial\Omega) \subset H^s(\partial\Omega)$ for $s \leq t$, by applying Lemma 4.7 to the representations. As a consequence:

Corollary 4.16 *Let $\Omega \subset \mathbb{R}^n$ be open and have C^1 regularity. Then the trace operator $\gamma : H^1(\Omega) \to L^2(\partial\Omega)$ is well-defined, linear and continuous.* □

Proposition 4.17 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with regularity C^k, $k \in \mathbb{N}$, $k \geq 1$. Then the trace operator $\gamma : H^k(\Omega) \to H^{k-\frac{1}{2}}(\partial\Omega)$ has a linear and continuous right inverse, that is, there exists a linear and continuous $E : H^{k-\frac{1}{2}}(\partial\Omega) \to H^k(\Omega)$ such that $\gamma E v = v$ for all $v \in H^{k-\frac{1}{2}}(\partial\Omega)$. In particular, γ is surjective.*

Proof: See the book of Wloka, Chapter 8. □

We come back to the inhomogeneous Dirichlet problem

$$\begin{aligned} Lu &= f\,, \quad \text{in } \Omega, \\ u &= g\,, \quad \text{on } \partial\Omega. \end{aligned} \tag{4.42}$$

and improve Proposition 4.1 as follows.

Proposition 4.18 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with regularity C^1, let $f \in L^2(\Omega)$, $g \in H^{1/2}(\partial\Omega)$, assume 3.1. Then (4.42) has a unique weak solution $u \in H^1(\Omega)$, and*

$$\|u\|_{H^1(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}) \tag{4.43}$$

for some constant C independent from f and g.

Proof: We apply Proposition 4.17 with $k = 1$ and set $\tilde{g} = Eg$. Due to Proposition 4.1, there is a unique solution $u \in H^1(\Omega)$ with $u - \tilde{g} \in H^1_0(\Omega)$, so $\gamma u = \gamma\tilde{g} = g$ and

$$\|u\|_{H^1(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|\tilde{g}\|_{H^1(\Omega)}) \leq C(\|f\|_{L^2(\Omega)} + \|E\|\|g\|_{H^{1/2}(\partial\Omega)})\,.$$

□

Neumann boundary conditions. We consider the problem

$$-\Delta u = f\,, \quad \text{in } \Omega, \tag{4.44}$$
$$\partial_\nu u = g\,, \quad \text{on } \partial\Omega. \tag{4.45}$$

The boundary condition (4.45) is called **Neumann boundary condition** or simply Neumann condition.

Testing with $\varphi \in C^\infty(\Omega)$ we get

$$\begin{aligned}\int_\Omega f(x)\varphi(x)\,dx &= \int_\Omega -\Delta u(x)\varphi(x)\,dx = \int_\Omega \langle \nabla u(x), \nabla\varphi(x)\rangle\,dx - \int_{\partial\Omega} \partial_\nu u(\xi)\varphi(\xi)\,dS(\xi)\\ &= \int_\Omega \langle \nabla u(x), \nabla\varphi(x)\rangle\,dx - \int_{\partial\Omega} g(\xi)\varphi(\xi)\,dS(\xi)\,. \end{aligned}\tag{4.46}$$

The right side involves a surface integral over the boundary. We recall from vector analysis that the surface integral is defined via a C^1 representation (α_j, U_j) of Ω and a corresponding partition of unity (β_j); for suitable functions $v : \partial\Omega \to \mathbb{R}$ one defines

$$\int_{\partial\Omega} v(\xi)\,dS(\xi) = \sum_{j\in J} \int_{\partial\Omega} (\beta_j v)(\xi)\,dS(\xi)\,, \tag{4.47}$$

and

$$\int_{\partial\Omega} (\beta_j v)(\xi)\,dS(\xi) = \int_{\mathbb{R}^{n-1}} (\beta_j v)(\alpha_j(y'))\sqrt{G_j(y')}\,dy'\,, \tag{4.48}$$

where G_j is the Gram determinant of the Jacobian of α_j,

$$G_j(y') = \det(D\alpha_j(y')^T D\alpha_j(y'))\,, \quad y' \in \mathbb{R}^{n-1}\,. \tag{4.49}$$

We consider the function space $L^1(\partial\Omega)$ of all functions $v : \partial\Omega \to \mathbb{R}$ such that $\tilde v_j = (\beta_j v)\circ\alpha_j$ belongs to $L^1(\mathbb{R}^{n-1})$ for all $j \in J$. This is a Banach space with the norm

$$\|v\|_{L^1(\partial\Omega)} = \sum_{j\in J} \|\tilde v_j\|_{L^1(\mathbb{R}^{n-1})}\,. \tag{4.50}$$

As in the case of $H^s(\partial\Omega)$ one can prove that different representations lead to the same space with different but equivalent norms.

Lemma 4.19 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, having C^1 regularity. Let $g \in L^2(\partial\Omega)$. Then*

$$F(v) = \int_{\partial\Omega} v(\xi)g(\xi)\,dS(\xi) \tag{4.51}$$

defines a linear and continuous functional $F : L^2(\partial\Omega) \to \mathbb{R}$.

Proof: Using a representation of Ω one checks that

$$\begin{aligned}|F(v)| &\le \int_{\partial\Omega} |v(\xi)g(\xi)|\,dS(\xi) \le \Big(\int_{\partial\Omega} |g(\xi)|^2\,dS(\xi)\Big)^{1/2}\Big(\int_{\partial\Omega} |v(\xi)|^2\,dS(\xi)\Big)^{1/2}\\ &= \|g\|_{L^2(\partial\Omega)}\|v\|_{L^2(\partial\Omega)}\,.\end{aligned}$$

□

We consider the Neumann problem for the operator $-\Delta + \mu I$,

$$-\Delta u + \mu u = f\,, \quad \text{in } \Omega, \tag{4.52}$$
$$\partial_\nu u = g\,, \quad \text{on } \partial\Omega. \tag{4.53}$$

According to the computation (4.46), its variational formulation reads as follows. We want to find $u \in H^1(\Omega)$ such that

$$a(u,v) = F(v)\,, \quad \text{for all } v \in H^1(\Omega), \tag{4.54}$$

where

$$a(u,v) = \int_\Omega \langle \nabla u, \nabla v\rangle + \mu uv\,dx\,, \quad F(v) = \int_\Omega fv\,dx + \int_{\partial\Omega} g\cdot(\gamma v)\,dS(\xi)\,. \tag{4.55}$$

The rightmost integral involves the trace of v on $\partial\Omega$.

Proposition 4.20 *Let $\Omega \subset \mathbb{R}^n$ open and bounded, having C^1 regularity, let $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega)$ and $\mu > 0$. Then the boundary value problem (4.52), (4.53) has a unique weak solution $u \in H^1(\Omega)$, and*

$$\|u\|_{H^1(\Omega)} \le C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)})\,, \tag{4.56}$$

where the constant C is independent from f and g.

Proof: Because

$$a(v,v) \ge \min\{1,\mu\}\|v\|^2_{H^1(\Omega)}\,,$$

the bilinear form a is $H^1(\Omega)$-elliptic, and a is continuous by Lemma 3.3. Using the trace theorem and Lemma 4.19, we obtain

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

Thus, F is linear and continuous. The assertion now follows from Lax-Milgram. □

In the homogeneous case $g = 0$, the boundary condition "$\partial_\nu u = 0$ on $\partial\Omega$" is also called a **natural boundary condition**. It arises "naturally" when we minimize on $H^1(\Omega)$ the quadratic functional

$$J(v) = \frac{1}{2}a(v,v) - \int_\Omega fv\,dx\,, \tag{4.57}$$

where $a(v,v) = \int_\Omega |\nabla v|^2\,dx$. In contrast to that, the Dirichlet boundary condition "$u = 0$ on $\partial\Omega$" is "enforced" through a suitable choice of a subspace (namely, $H^1_0(\Omega)$) of $H^1(\Omega)$. Therefore, the Dirichlet condition is also termed a "forced boundary condition".

We now pose the question: What is the natural boundary condition which arises if we minimize (4.57) on $H^1(\Omega)$ with the more general bilinear form

$$a(u,v) = \int_\Omega \sum_{i,j=1}^n a_{ij}(x)\partial_i u(x)\partial_j v(x)\,dx + \int_\Omega c(x)u(x)v(x)\,dx\,. \tag{4.58}$$

Since we minimize, we assume A to be symmetric.

For the following computation we assume that all functions involved are sufficiently smooth. We get

$$\int_\Omega a_{ij}\partial_i u \partial_j v \, dx = -\int_\Omega \partial_j(a_{ij}\partial_i u) \cdot v \, dx + \int_{\partial\Omega} a_{ij}\partial_i u \cdot v\nu_j dS \,, \tag{4.59}$$

thus

$$a(u,v) = \int_\Omega (Lu)(x)v(x) \, dx + \int_{\partial\Omega} (Bu)(\xi)v(\xi) \, dS(\xi) \,, \tag{4.60}$$

where

$$Lu = -\mathrm{div}\,(A^T \nabla u) + cu \tag{4.61}$$

as in (3.3), and

$$(Bu)(\xi) = \sum_{i,j=1}^n a_{ij}(\xi)(\partial_i u)(\xi)\nu_j(\xi) \,, \quad \xi \in \partial\Omega \,. \tag{4.62}$$

We consider the variational formulation

$$a(u,v) = F(v) = \int_\Omega fv \, dx + \int_{\partial\Omega} gv \, dS \,. \tag{4.63}$$

It follows that

$$\int_\Omega (Lu)(x)v(x) \, dx = \int_\Omega f(x)v(x) \, dx \,, \quad \text{for all } v \in C_0^\infty(\Omega), \tag{4.64}$$

since for $v \in C_0^\infty(\Omega)$ the boundary integrals vanish. Therefore,

$$Lu = f \,, \quad \text{in } \Omega. \tag{4.65}$$

Inserting (4.65) into (4.63), we now obtain

$$\int_{\partial\Omega} (Bu)(\xi)v(\xi) \, dS(\xi) = \int_{\partial\Omega} g(\xi)v(\xi) \, dS(\xi) \,, \tag{4.66}$$

for all $v \in H^1(\Omega)$. Since $\gamma(H^1(\Omega)) = H^{1/2}(\partial\Omega)$ is dense in $L^2(\partial\Omega)$, it follows that

$$Bu = g \,, \quad \text{on } \partial\Omega. \tag{4.67}$$

The minimizer $u \in H^1(\Omega)$ of (4.57) satisfies (4.63) with $g = 0$. Thus, the natural boundary condition associated with the bilinear form a is

$$Bu = 0 \tag{4.68}$$

with B from (4.62).

Compatibility conditions. The assumption $\mu > 0$ in Proposition 4.20 is essential. Let us return to the Neumann problem

$$-\Delta u = f \,, \quad \text{in } \Omega, \tag{4.69}$$

$$\partial_\nu u = g \,, \quad \text{on } \partial\Omega \tag{4.70}$$

with the associated variational formulation

$$\int_\Omega \langle \nabla u(x), \nabla\varphi(x)\rangle \, dx = \int_\Omega f(x)\varphi(x)\, dx + \int_{\partial\Omega} g(\xi)\varphi(\xi)\, dS(\xi)\,, \quad \varphi \in C_0^\infty(\Omega)\,. \tag{4.71}$$

Assume that u solves (4.69), (4.70). This solution is not unique, because all functions $u + c$, $c \in \mathbb{R}$ being an arbitrary constant, are also solutions. On the other hand, setting $\varphi = 1$ in (4.71) we obtain

$$\int_\Omega f(x)\, dx + \int_{\partial\Omega} g(y)\, dS(y) = 0\,. \tag{4.72}$$

This means that a solution can exist only if the data f and g satisfy an additional **compabitibility condition**, namely (4.72).

This issue corresponds to the situation which arises when solving linear equations in finite dimensions,

$$Ax = b\,, \quad b \in \mathbb{R}^n\,,\, A \in \mathbb{R}^{(n,n)}\,. \tag{4.73}$$

There we have the dimension formula

$$\dim(\ker A) = n - \dim(\operatorname{im} A) = \dim(\operatorname{coker} A)\,, \tag{4.74}$$

and (4.73) is uniquely solvable if and only if $\dim(\ker A) = 0$.

This has been generalized to linear equations in infinite dimensional spaces since around 1900, by Fredholm for linear integral equations and later by Riesz and Schauder for linear equations in Hilbert spaces. The main result is the so-called **Fredholm alternative**. But this is not the subject of this course.

5 Homogenization: Introduction

A good reference is the book of D. Cioranescu and P. Donato, An introduction to homogenization.

We consider

$$Lu = f \,, \tag{5.1}$$

where

$$Lu = -\sum_{i,j=1}^{n} \partial_j (a_{ij}(x)\partial_i u) = -\mathrm{div}\,(A(x)^T \nabla u)\,. \tag{5.2}$$

When describing real processes with partial differential equations, often the form of the operator L is determined by the underlying model (for example, a conservation law), while the coefficients a_{ij} represent properties of the medium where the process takes place (heat conduction coefficients, elasticity modulus, electric conductivity etc.). In the case of constant coefficients a_{ij} one speaks of a **homogeneous medium**, otherwise of an **inhomogeneous medium**. Inhomogeneities may be of different types. For example, if the medium represented by Ω consists of different materials $M_1, \ldots, M_K$ which take up the parts $\Omega_1, \ldots, \Omega_K$ of Ω and are characterized by coefficients $a_{ij,k}$, $k \in \{1, \ldots, K\}$, one may set

$$a_{ij}(x) = \sum_{k=1}^{K} a_{ij,k} 1_{\Omega_k}(x)\,, \tag{5.3}$$

where 1_{Ω_k} denotes the characteristic function of Ω_k.

In this chapter we consider the situation where inhomogeneities appear on two "scales". For example, when one wants to describe biomechanical processes in the human bone, the macroscale is in the centimeter range, whereas the porous structure (which determines the mechanical strength of the bone) has typical lengths in the range of $10 - 100\ \mu m$. This leads to material parameters

$$a_{ij}^{\varepsilon}(x)\,,$$

where ε is small.

Often one is mainly interested in macroscopic properties. One then wants to take the microstructure into account only to the extent it influences the macroscopic properties. One hopes that this can be described by some averaging procedure. The question is: what is the correct way to average ?

We consider a one-dimensional example. Let $Y = (0, l)$ be a reference interval in $\mathbb{R}$, let $a \in L^\infty(0, l)$. Setting

$$a(y + l) = a(y)\,, \quad y \in \mathbb{R}\,, \tag{5.4}$$

we obtain a l-periodic function $a \in L^\infty(\mathbb{R})$. We define now, for a given $\varepsilon > 0$,

$$a^\varepsilon(x) = a\left(\frac{x}{\varepsilon}\right)\,, \tag{5.5}$$

and consider in $\Omega = (d_1, d_2)$ the boundary value problem

$$-\partial_x (a^\varepsilon(x)\partial_x u_\varepsilon) = f\,, \quad x \in \Omega\,, \tag{5.6}$$

$$u_\varepsilon(d_1) = u_\varepsilon(d_2) = 0\,. \tag{5.7}$$

The coefficient a^ε varies with the period εl. We assume that there exist constants $c_a, C_a > 0$ such that

$$c_a \le a(y) \le C_a\,, \quad \text{for all } y \in Y. \tag{5.8}$$

We have

$$c_a \le a^\varepsilon(x) \le C_a\,, \quad \text{for all } x \in \Omega. \tag{5.9}$$

By the theorem of Lax-Milgram, the elliptic boundary value problem (5.6), (5.7) has a unique weak solution $u_\varepsilon \in H_0^1(\Omega)$. Now there arises the question whether u_ε converges to some u_0 in some sense, and whether such a u_0 can be obtained as the solution of a suitable "averaged boundary value problem".

We already know that

$$\|u_\varepsilon\|_{H_0^1(\Omega)} \le \frac{1}{c_a}\|f\|_{L^2(\Omega)}\,, \tag{5.10}$$

therefore $(u_\varepsilon)_{\varepsilon>0}$ is bounded in $H_0^1(\Omega)$. Consequently, $(u_\varepsilon)_{\varepsilon>0}$ and $(\partial_x u_\varepsilon)_{\varepsilon>0}$ are bounded in $L^2(\Omega)$. Therefore, there exists weakly convergent subsequences

$$u_{\varepsilon_k} \rightharpoonup u_0 \quad \text{in } L^2(\Omega), \quad \partial_x u_{\varepsilon_k} \rightharpoonup w_0 = \partial_x u_0 \quad \text{in } L^2(\Omega). \tag{5.11}$$

We now define

$$\xi_\varepsilon(x) = a^\varepsilon(x)\partial_x u_\varepsilon(x)\,, \quad x \in \Omega\,. \tag{5.12}$$

By (5.6),

$$-\partial_x \xi_\varepsilon = f\,, \quad \text{in } \Omega. \tag{5.13}$$

The family (ξ_ε), too, is bounded in $H_0^1(\Omega)$, due to (5.13) and because $|\xi_\varepsilon(x)| \le C_a|\partial_x u_\varepsilon(x)|$, so

$$\|\xi_\varepsilon\|_{L^2(\Omega)} \le C_a\|\partial_x u_\varepsilon\|_{L^2(\Omega)} \le \frac{C_a}{c_a}\|f\|_{L^2(\Omega)}\,. \tag{5.14}$$

Therefore there exists $\xi_0 \in H_0^1(\Omega)$ such that (after having passed to a suitable subsequence)

$$\xi_{\varepsilon_k} \rightharpoonup \xi_0 \quad \text{in } L^2(\Omega), \quad \partial_x \xi_{\varepsilon_k} = -f = \partial_x \xi_0 \quad \text{in } L^2(\Omega). \tag{5.15}$$

Since $\Omega = (d_1, d_2)$ is one-dimensional, every bounded subset of $H_0^1(\Omega)$ satisfies the assumptions of the theorem of Ascoli and Arzela. Therefore, for a subsequence

$$\xi_{\varepsilon_k} \to \xi_0 \quad \text{uniformly in } C[d_1, d_2]. \tag{5.16}$$

We now investigate the relation between u_0 and ξ_0. By (5.12) we have

$$\partial_x u_\varepsilon(x) = \frac{1}{a^\varepsilon(x)}\xi_\varepsilon(x)\,, \quad x \in \Omega\,. \tag{5.17}$$

It follows from (5.9) that

$$\frac{1}{C_a} \le \frac{1}{a^\varepsilon} \le \frac{1}{c_a}\,, \tag{5.18}$$

so $(1/a^\varepsilon)$ is bounded in $L^\infty(\Omega)$. As we will prove later, for yet another subsequence

$$\frac{1}{a^{\varepsilon_k}} \overset{*}{\rightharpoonup} \beta\,, \quad \beta = \frac{1}{l}\int_0^l \frac{1}{a(y)}\,dy\,. \tag{5.19}$$

It then follows (as we also will see later)

$$\frac{1}{a^{\varepsilon_k}}\xi_{\varepsilon_k} \rightharpoonup \beta\xi_0\,, \quad \text{in } L^2(\Omega). \tag{5.20}$$

Since $\partial_x u_{\varepsilon_k} \rightharpoonup \partial_x u_0$ in $L^2(\Omega)$ due to (5.11), from (5.17) and (5.20) we obtain

$$\partial_x u_0(x) = \frac{1}{l}\int_0^l \frac{1}{a(y)}\,dy \cdot \xi_0(x)\,. \tag{5.21}$$

We now define

$$a^0 = \left(\frac{1}{l}\int_0^l \frac{1}{a(y)}\,dy\right)^{-1}\,. \tag{5.22}$$

From (5.21) and (5.15) it follows that

$$-\partial_x(a^0\partial_x u_0) = f\,, \quad x \in \Omega\,, \tag{5.23}$$

as well as

$$u(d_1) = u(d_2) = 0\,. \tag{5.24}$$

Since $H_0^1(d_1, d_2) \subset C[d_1, d_2]$, the boundary conditions (5.24) hold in the classical sense (no trace theorem is needed).

Thus, it has turned out that the "homogenization limit" u_0 solves the boundary value problem (5.23), (5.24), which has the constant coefficient a^0 given by (5.22). Moreover, the solution u_0 of this boundary value problem is unique since $a^0 \geq c_a$. Since u_0 is unique, we furthermore obtain that

$$u_{\varepsilon_k} \rightharpoonup u_0 \quad \text{in } H_0^1(\Omega) \tag{5.25}$$

for every sequence $\varepsilon_k \to 0$, by the "convergence principle": if every subsequence of a given sequence has a convergent subsequence, and if those limits are identical, then the whole sequence converges to this limit.

Since a^0 is constant, one even obtains an explicit formula for u_0, namely

$$u_0(x) = -\frac{1}{a^0}\int_0^x\int_0^t f(s)\,ds\,dt + \frac{x}{a^0}\int_0^1\int_0^t f(s)\,ds\,dt\,. \tag{5.26}$$

We consider the example

$$(0, l) = (0, 1)\,, \quad a(y) = 1 + y\,. \tag{5.27}$$

We have

$$\int_0^1 \frac{1}{a(y)}\,dy = \ln(1+y)\Big|_0^1 = \ln 2\,, \quad a^0 = \frac{1}{\ln 2}\,, \tag{5.28}$$

but on the other hand

$$\int_0^1 a(y)\,dy = \frac{3}{2}\,. \tag{5.29}$$

Thus, the homogenization coefficient a^0 is **not** the integral mean of a. The same is true for the example

$$(0, l) = (0, 1)\,, \quad a(y) = \begin{cases} 1\,, & 0 < y < \frac{2}{3} \\ 2\,, & \frac{2}{3} < y\,. \end{cases} \tag{5.30}$$

Here we get

$$\int_0^1 \frac{1}{a(y)}\,dy = \frac{2}{3}\cdot 1 + \frac{1}{3}\cdot\frac{1}{2} = \frac{5}{6}\,, \quad a^0 = \frac{6}{5}\,, \tag{5.31}$$

but

$$\int_0^1 a(y)\,dy = \frac{2}{3}\cdot 1 + \frac{1}{3}\cdot 2 = \frac{4}{3}\,. \tag{5.32}$$

Discontinuous coefficients. We return to the multidimensional situation. Before considering the homogenization problem, we want to take a look at the meaning of a weak solution of an elliptic problem when the coefficients are discontinuous, as in (5.3). If the right side f in (5.1) is smooth, one cannot expect that $\nabla u(x)$ is continuous in points where $A(x)$ is discontinuous. We consider the special case where Ω decomposes according to

$$\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2\,, \quad \Omega_1 \cap \Omega_2 = \emptyset\,, \tag{5.33}$$

where Ω_1, Ω_2 are open subsets of a bounded open $\Omega \subset \mathbb{R}^n$. We assume moreover that $\partial\Omega_1, \partial\Omega_2$ are sufficiently smooth. Let $u \in H_0^1(\Omega)$ be a weak solution of

$$-\mathrm{div}\,(A(x)^T \nabla u) = f \quad \text{in } \Omega \tag{5.34}$$

with $u = 0$ on $\partial\Omega$, assume moreover that $u \in C(\Omega)$ and $u \in C^2(\Omega_i)$ for $i = 1, 2$, and that the derivatives ∇u in Ω_i can be continuously extended to $\partial\Omega_i$. We denote these extensions by ∇u_i. We do **not** assume that ∇u_1 and ∇u_2 coincide on $\partial\Omega_1 \cap \partial\Omega_2$. Let moreover $A : \Omega \to \mathbb{R}^{(n,n)}$ be continuously differentiable on Ω_i with continuous extensions A_1 and A_2 to $\partial\Omega_i$. Then we have for all test functions $\varphi \in C_0^\infty(\Omega)$

$$\int_\Omega f(x)\varphi(x)\,dx = \int_\Omega \nabla u(x)^T A(x) \nabla\varphi(x)\,dx = \sum_{i=1}^2 \int_{\Omega_i} \nabla u(x)^T A(x)\nabla\varphi(x)\,dx =$$

$$= -\sum_{i=1}^2 \int_{\Omega_i} \mathrm{div}\,(A(x)^T \nabla u(x))\varphi(x)\,dx + \sum_{i=1}^2 \int_{\partial\Omega_i} \nabla u_i(\xi)^T A_i(\xi)\nu_i(\xi)\varphi(\xi)\,dS(\xi)\,.$$

Due to (5.34) and since $\varphi = 0$ on $\partial\Omega$,

$$\sum_{i=1}^2 \int_{\partial\Omega_1\cap\partial\Omega_2} \nabla u_i(\xi)^T A_i(\xi)\nu_i(\xi)\varphi(\xi)\,dS(\xi) = 0\,, \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{5.35}$$

Since the unit outer normals ν_i to $\partial\Omega_i$ satisfy $\nu_1 = -\nu_2$ on $\partial\Omega_1 \cap \partial\Omega_2$, setting $\nu = \nu_1$ it follows that

$$\int_{\partial\Omega_1\cap\partial\Omega_2} \nabla u_1(\xi)^T A_1(\xi)\nu(\xi)\varphi(\xi)\,dS(\xi) = \int_{\partial\Omega_1\cap\partial\Omega_2} \nabla u_2(\xi)^T A_2(\xi)\nu(\xi)\varphi(\xi)\,dS(\xi)\,, \tag{5.36}$$

for all $\varphi \in C_0^\infty(\Omega)$. Thus

$$\left\langle A_1(x)^T \nabla u_1(x), \nu(x)\right\rangle = \left\langle A_2(x)^T \nabla u_2(x), \nu(x)\right\rangle\,, \quad \text{a.e. on } \partial\Omega_1 \cap \partial\Omega_2. \tag{5.37}$$

This means that the normal component of the flux $A(x)^T\nabla u(x)$ is continuous across the discontinuity surface $\partial\Omega_1 \cap \partial\Omega_2$ of the coefficients.

We conclude: If in a given situation to be described (for example, a diffusion process) the flux of some quantity u has this conservation property (namely, there is no flux production on the surface), then the concept of a weak solution for u correctly reflects this property.

6 Averages and weak convergence

We recall that by definition the sequence $(v_n)_{n\in\mathbb{N}}$ in $L^p(\Omega)$ weakly converges to a $v \in L^p(\Omega)$ if

$$\lim_{n\to\infty} \int_\Omega v_n(x)\varphi(x)\,dx = \int_\Omega v(x)\varphi(x)\,dx\,, \quad \text{for all } \varphi \in L^q(\Omega), \tag{6.1}$$

where

$$\frac{1}{p} + \frac{1}{q} = 1\,. \tag{6.2}$$

The case $p = \infty$: a sequence $(v_n)_{n\in\mathbb{N}}$ in $L^\infty(\Omega)$ is said to be weakly star convergent to $v \in L^\infty(\Omega)$ if (6.2) holds with $q = 1$.

A subset I of $\mathbb{R}^n$ having the form

$$I = \prod_{i=1}^n (a_i, b_i)\,, \quad \text{resp.} \quad I = \prod_{i=1}^n [a_i, b_i]$$

is called **open** resp. **closed interval** in $\mathbb{R}^n$.

Proposition 6.1 *Let $\Omega \subset \mathbb{R}^n$ be open, let $(v_n)_{n\in\mathbb{N}}$ be a sequence in $L^p(\Omega)$, $1 < p \le \infty$, let $v \in L^p(\Omega)$. Then the following are equivalent:*

(i) $(v_n)_{n\in\mathbb{N}}$ weakly converges to v (in the case $p = \infty$: $(v_n)_{n\in\mathbb{N}}$ weakly star converges to v).

(ii) $(v_n)_{n\in\mathbb{N}}$ is bounded in $L^p(\Omega)$ (that is, the set $\{\|v_n\|_{L^p(\Omega)} : n \in \mathbb{N}\}$ is bounded), and

$$\lim_{n\to\infty} \int_I v_n(x)\,dx = \int_I v(x)\,dx \tag{6.3}$$

for all intervals $I \subset \Omega$.

Proof: "(i)⇒(ii)": Every weakly convergent (weakly star convergent, resp.) sequence is bounded, by a theorem of functional analysis. Setting $\varphi = \chi_I$ in (6.1) we obtain (6.3).

"(ii)⇒(i)": By (6.3), (6.1) holds for all $\varphi = \chi_I$, I interval, and therefore also for all functions of the form

$$\varphi = \sum_{k=1}^m \alpha_k \chi_{I_k}\,, \quad \alpha_k \in \mathbb{R}\,. \tag{6.4}$$

Let now $\varphi \in L^q(\Omega)$ be arbitrary, let q be the dual exponent from (6.2), let $\delta > 0$. Since the subspace of functions of the form (6.4) is dense in $L^q(\Omega)$, there exists a simple function φ_δ such that

$$\|\varphi - \varphi_\delta\|_q \le \delta\,.$$

We have

$$\int_\Omega (v_n - v)\varphi\,dx = \int_\Omega (v_n - v)\varphi_\delta\,dx + \int_\Omega (v_n - v)(\varphi - \varphi_\delta)\,dx\,.$$

We choose n_0 such that

$$\left|\int_\Omega (v_n - v)\varphi_\delta\,dx\right| \le \delta$$

for all $n \geq n_0$. Then it follows that

$$\left| \int_\Omega (v_n - v)\varphi \, dx \right| \leq \delta + \|v_n - v\|_p \, \|\varphi - \varphi_\delta\|_q \leq C\delta \, ,$$

for all $n \geq n_0$, where C is independent from n_0 and δ. □

For $p = 1$, (ii) does not imply (i).

Let Y be an interval in $\mathbb{R}^n$ of the form

$$Y = \prod_{i=1}^{n} (0, l_i) \, , \quad l_i > 0 \, . \tag{6.5}$$

Definition 6.2 *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called **Y-periodic** if*

$$f(x + l_i e_i) = f(x) \tag{6.6}$$

holds for all $x \in \mathbb{R}^n$, $1 \leq i \leq n$, where e_i denotes the i-th unit vector. □

We immediately see that every Y-periodic function satisfies

$$f\left(x + \sum_{i=1}^{n} k_i l_i e_i \right) = f(x) \, , \quad x \in \mathbb{R}^n \, , \, k_i \in \mathbb{Z} \, . \tag{6.7}$$

Every $f : Y \to \mathbb{R}$ can be extended to a Y-periodic function $f : \mathbb{R}^n \to \mathbb{R}$ which is uniquely determined except for its values on the null set given by the translation of ∂Y according to (6.7).

Lemma 6.3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Y-periodic, $f|Y \in L^1(Y)$. Then*

$$\int_{y_0 + Y} f(y) \, dy = \int_Y f(y) \, dy \tag{6.8}$$

holds for all $y_0 \in \mathbb{R}^n$.

Proof: It suffices to consider the case $y_0 = ce_i$, $c \in \mathbb{R}$. We define

$$g(y_i) = \int_{\prod_{j \neq i}(0, l_j)} f(y_1, \ldots, y_n) \, dy_1 \cdots dy_{i-1} \, dy_{i+1} \cdots dy_n \, .$$

We then have

$$\begin{aligned} \int_{y_0 + Y} f(y) \, dy &= \int_c^{c + l_i} g(y_i) \, dy_i = \left(\int_c^{l_i} + \int_{l_i}^{c + l_i} \right) g(y_i) \, dy_i = \left(\int_c^{l_i} + \int_0^{c} \right) g(y_i) \, dy_i \\ &= \int_0^{l_i} g(y_i) \, dy_i = \int_Y f(y) \, dy \, . \end{aligned}$$

□

For the function defined by

$$f_\varepsilon(x) = f\left(\frac{x}{\varepsilon} \right)$$

it follows from the substitution formula and from Lemma 6.3, since f_ε is (εY)-periodic,

$$\int_{\varepsilon(y_0+Y)} f_\varepsilon(x)\,dx = \int_{\varepsilon Y} f_\varepsilon(x)\,dx = \varepsilon^n \int_Y f(y)\,dy \tag{6.9}$$

for all $y_0 \in \mathbb{R}^n$ and all $\varepsilon > 0$.

We investigate the convergence behaviour of (f_ε) as $\varepsilon \to 0$. As an example we consider

$$Y = (0,1)\,, \quad f : Y \to \mathbb{R}\,, \quad f(y) = \sin(2\pi y)\,. \tag{6.10}$$

For every interval $I = [a,b] \subset \mathbb{R}$ we obtain

$$\int_a^b f_\varepsilon(x)\,dx = \int_a^b \sin\left(2\pi\frac{x}{\varepsilon}\right)\,dx = -\frac{\varepsilon}{2\pi}\cos\left(2\pi\frac{x}{\varepsilon}\right)\Big|_a^b \to 0$$

for $\varepsilon \to 0$, and from Proposition 6.1 it follows that $f_\varepsilon \rightharpoonup 0$ in $L^p(\Omega)$ for $1 < p < \infty$ resp. $f_\varepsilon \overset{*}{\rightharpoonup} 0$ in $L^\infty(\Omega)$ for every open $\Omega \subset \mathbb{R}$. On the other hand,

$$\|f_\varepsilon\|^2_{L^2(\Omega)} = \frac{b-a}{2} + \frac{\varepsilon}{8\pi}\left[-\sin\left(\frac{4\pi b}{\varepsilon}\right) + \sin\left(\frac{4\pi a}{\varepsilon}\right)\right] \to \frac{b-a}{2}$$

for $\varepsilon \to 0$, thus f_ε does not converge strongly to 0.

Notation 6.4 (Mean value)
Let $\Omega \subset \mathbb{R}^n$ be measurable, $|\Omega| := \mathrm{meas}\,(\Omega) < \infty$. We define the mean value of $f \in L^1(\Omega)$ by

$$M_\Omega(f) = \frac{1}{|\Omega|}\int_\Omega f(y)\,dy\,. \tag{6.11}$$

□

Proposition 6.5 (Weak convergence to the mean value)
Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $Y = \prod_{i=1}^n (0, l_i)$, let $p \in [1,\infty]$, let $f : \mathbb{R}^n \to \mathbb{R}$ be a Y-periodic function with $f \in L^p(Y)$. Then, in the case $p < \infty$ we have

$$f_\varepsilon \rightharpoonup M_Y(f) \quad \text{in } L^p(\Omega), \text{ for } \varepsilon \to 0, \tag{6.12}$$

and in the case $p = \infty$

$$f_\varepsilon \overset{*}{\rightharpoonup} M_Y(f) \quad \text{in } L^\infty(\Omega), \text{ for } \varepsilon \to 0. \tag{6.13}$$

In each case, for every interval $I \subset \mathbb{R}^n$ there exists a constant c which is independent from f such that

$$\|f_\varepsilon\|_{L^p(I)} \le c\|f\|_{L^p(Y)}\,, \tag{6.14}$$

for every $\varepsilon > 0$.

Proof: At first, assume that

$$I = (a,b) = \prod_{i=1}^n (a_i, b_i) \tag{6.15}$$

is an arbitrary open interval in $\mathbb{R}^n$. Starting in the corner a with the interval $a + \varepsilon Y$, we fill in I with adjacent intervals of the form

$$J_\varepsilon = x + \varepsilon Y\,, \quad x_i = a_i + j_i \varepsilon l_i\,, \quad 0 \le j_i \le k_i^\varepsilon\,, \tag{6.16}$$

where $k_i^\varepsilon \in \mathbb{N}$ with

$$\varepsilon l_i k_i^\varepsilon \le b_i - a_i \le \varepsilon l_i (k_i^\varepsilon + 1)\,. \tag{6.17}$$

Let A^ε be the set of all such J_ε with $J_\varepsilon \subset I$, then

$$|A^\varepsilon| = \prod_{i=1}^n k_i^\varepsilon\,, \quad \lim_{\varepsilon\to 0} \varepsilon^n |A^\varepsilon| = \lim_{\varepsilon \to 0} \prod_{i=1}^n \varepsilon k_i^\varepsilon = \prod_{i=1}^n \frac{b_i - a_i}{l_i} = \frac{|I|}{|Y|}\,. \tag{6.18}$$

Let B^ε be the set of all remaining J_ε. Then

$$|B^\varepsilon| = \sum_{i=1}^n \prod_{j\neq i} k_j^\varepsilon\,, \quad \lim_{\varepsilon\to 0} \varepsilon^n |B^\varepsilon| = \lim_{\varepsilon\to 0} \varepsilon \sum_{i=1}^n \prod_{j\neq i} \varepsilon k_j^\varepsilon = 0\,. \tag{6.19}$$

After these preliminary considerations we prove that the set $(f_\varepsilon)_{\varepsilon>0}$ is bounded in $L^p(\Omega)$ and that (6.14) holds. For $p = \infty$ we obviously have $\|f_\varepsilon\|_\infty = \|f\|_\infty$. Let $p \in [1, \infty)$. Using (6.9) we obtain that, for an arbitrary open interval I in $\mathbb{R}^n$,

$$\begin{aligned}
\|f_\varepsilon\|^p_{L^p(I)} &\le \sum_{J_\varepsilon \in A^\varepsilon} \int_{J_\varepsilon} |f_\varepsilon(x)|^p \, dx + \sum_{J_\varepsilon \in B^\varepsilon} \int_{J_\varepsilon} |f_\varepsilon(x)|^p \, dx \\
&= |A^\varepsilon| \varepsilon^n \int_Y |f(y)|^p \, dy + \underbrace{|B^\varepsilon|\varepsilon^n}_{\to 0} \int_Y |f(y)|^p \, dy \\
&\to \frac{|I|}{|Y|} \int_Y |f(y)|^p \, dy\,.
\end{aligned} \tag{6.20}$$

This implies the boundedness (choose I with $\Omega \subset I$ and moreover (6.14). We now prove the asserted convergence for the case $p > 1$ by applying Proposition 6.1. For an arbitrary interval $I \subset \Omega$ we have, analogous to (6.20),

$$\begin{aligned}
\int_I f_\varepsilon(x) \, dx &= \sum_{J_\varepsilon \in A^\varepsilon} \int_{J_\varepsilon} f_\varepsilon(x) \, dx + \sum_{J_\varepsilon \in B^\varepsilon} \int_{J_\varepsilon \cap I} f_\varepsilon(x) \, dx \\
&\to \frac{|I|}{|Y|} \int_Y f(y) \, dy = |I| M_Y(f) = \int_I M_Y(f) \, dx\,.
\end{aligned}$$

It remains to prove the convergence in the case $p = 1$. This is done by reduction to the case $p > 1$. Let I be a fixed interval in $\mathbb{R}^n$ with $\Omega \subset I$, let $\eta > 0$ be arbitrary. We choose $f^\eta \in L^2(Y)$ with

$$\|f - f^\eta\|_{L^1(Y)} \le \eta\,, \tag{6.21}$$

and extend f^η Y-periodically to $\mathbb{R}^n$. We have

$$(f - f^\eta)_\varepsilon(x) = (f - f^\eta)\left(\frac{x}{\varepsilon}\right) = (f_\varepsilon - f^\eta_\varepsilon)(x)\,. \tag{6.22}$$

Let $\varphi \in L^\infty(\Omega)$ be arbitrary. We have

$$\int_\Omega (f_\varepsilon(x) - M_Y(f))\varphi(x)\,dx = \int_\Omega (f_\varepsilon(x) - f_\varepsilon^\eta(x))\varphi(x)\,dx + \int_\Omega (f_\varepsilon^\eta(x) - M_Y(f^\eta))\varphi(x)\,dx + \int_\Omega (M_Y(f^\eta) - M_Y(f))\varphi(x)\,dx\,. \tag{6.23}$$

The first integral on the right side can be estimated, using (6.14) and (6.22), by

$$\begin{aligned}\left|\int_\Omega (f_\varepsilon(x) - f_\varepsilon^\eta(x))\varphi(x)\,dx\right| &\le \|\varphi\|_\infty \|f_\varepsilon - f_\varepsilon^\eta\|_{L^1(I)} \le c\|\varphi\|_\infty \|f - f^\eta\|_{L^1(Y)} \\ &\le c\|\varphi\|_\infty\,\eta\,.\end{aligned} \tag{6.24}$$

For the third integral we obtain

$$\left|\int_\Omega (M_Y(f^\eta) - M_Y(f))\varphi(x)\,dx\right| \le \|\varphi\|_\infty |\Omega| \frac{1}{|Y|}\int_Y |f^\eta(y) - f(y)|\,dy \le \|\varphi\|_\infty |\Omega| \frac{1}{|Y|}\eta\,. \tag{6.25}$$

We have $\varphi \in L^2(\Omega)$ since Ω is bounded. We apply the convergence result which we already proved for $p = 2$, and obtain

$$\lim_{\varepsilon\to 0} \int_\Omega (f_\varepsilon^\eta(x) - M_Y(f^\eta))\varphi(x)\,dx = 0\,. \tag{6.26}$$

Combining (6.24) – (6.26) we get

$$\lim_{\varepsilon\to 0} \int_\Omega (f_\varepsilon(x) - M_Y(f))\varphi(x)\,dx = 0\,. \tag{6.27}$$

□

Definition 6.6 (Compact embedding)
Let $(X, \|\cdot\|_X)$, $(Z, \|\cdot\|_Z)$ be Banach spaces with $X \subset Z$. We say that X is ***compactly embedded*** *in Z if every bounded subset of X is relatively compact in Z. We write*

$$X \subset\subset Z\,. \tag{6.28}$$

□

This is equivalent to: the canonical embedding $j : X \to Z$ is a compact mapping (see functional analysis).

Proposition 6.7 (Compact embedding in Sobolev space)
Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then

$$W_0^{1,p}(\Omega) \subset\subset L^p(\Omega)\,, \quad \textit{for all } p \in [1,\infty]. \tag{6.29}$$

If Ω has regularity C^1, then

$$W^{1,p}(\Omega) \subset\subset L^p(\Omega)\,, \quad \textit{for all } p \in [1,\infty]. \tag{6.30}$$

Proof: See for example the books of Evans, Gilbarg/Trudinger, Wloka. $\square$

Corollary 6.8 *In the situation of Proposition 6.7 it holds: Every bounded sequence in $W_0^{1,p}(\Omega)$ resp. $W^{1,p}(\Omega)$ has a subsequence which converges in the norm of L^p. Every sequence which converges weakly (resp. weak star) in $W_0^{1,p}(\Omega)$ resp. $W^{1,p}(\Omega)$ also converges in the norm of L^p.* $\square$

Proposition 6.9 (Poincaré inequality, mean values)
Let $\Omega \subset \mathbb{R}^n$ be bounded, open and connected, let Ω have C^1 regularity, let $p \in [1, \infty]$. Then there exists a $C > 0$ such that

$$\|v - M_\Omega(v)\|_{L^p(\Omega)} \le C \|\nabla v\|_{L^p(\Omega)} , \quad \textit{for all } v \in W^{1,p}(\Omega). \tag{6.31}$$

Proof: We assume that (6.31) does not hold. Then there exists a sequence $(v_k)_{k\in\mathbb{N}}$ in $W^{1,p}(\Omega)$ such that

$$\|v_k - M_\Omega(v_k)\|_{L^p(\Omega)} > k\|\nabla v_k\|_{L^p(\Omega)} . \tag{6.32}$$

Setting

$$w_k = \frac{v_k - M_\Omega(v_k)}{\|v_k - M_\Omega(v_k)\|_{L^p(\Omega)}} , \tag{6.33}$$

we see that

$$M_\Omega(w_k) = 0 , \quad \|w_k\|_{L^p(\Omega)} = 1 , \tag{6.34}$$

$$\|\nabla w_k\|_{L^p(\Omega)} = \frac{\|\nabla v_k\|_{L^p(\Omega)}}{\|v_k - M_\Omega(v_k)\|_{L^p(\Omega)}} < \frac{1}{k} . \tag{6.35}$$

Thus, the sequence $(w_k)_{k\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega)$. By Corollary 6.8 there exists a subsequence $(w_{k_m})_{m\in\mathbb{N}}$ and a $w \in W^{1,p}(\Omega)$ such that

$$w_{k_m} \to w \quad \text{in } L^p(\Omega). \tag{6.36}$$

It follows that

$$M_\Omega(w) = 0 , \quad \|w\|_{L^p(\Omega)} = 1 . \tag{6.37}$$

Let $\varphi \in C_0^\infty(\Omega)$, then because of (6.36) and (6.35) we have for all j

$$\int_\Omega w(x)\partial_j\varphi(x)\,dx = \lim_{m\to\infty} \int_\Omega w_{k_m}(x)\partial_j\varphi(x)\,dx = -\lim_{m\to\infty} \int_\Omega \partial_j w_{k_m}(x)\varphi(x)\,dx = 0 ,$$

thus $\nabla w = 0$ in Ω. Since Ω is connected, the function w has to be constant, a contradiction to (6.37). $\square$

7 Periodic boundary conditions

In the next chapter we will consider the problem for periodic homogenization in several space dimensions ($\Omega \subset \mathbb{R}^n$, $n > 1$). It will turn out that, as in the one-dimensional case, the homogenized coefficients are constant (that is, they do not depend on $x \in \Omega$). It will also turn out that in order to characterize them, one has to solve an elliptic boundary problem on the reference interval $Y = \prod_{i=1}^n (0, l_i)$ with periodic boundary conditions.

We consider the boundary value problem

$$Lu = -\sum_{i,j=1}^n \partial_j (a_{ij}(y)\partial_i u) = f\,, \quad \text{in } Y\,, \tag{7.1}$$

$$u \quad Y\text{-periodic.} \tag{7.2}$$

If u is a solution, then $u + c$ also is a solution, for arbitary constants c.

Assumption 7.1 *Let $a_{ij} \in L^\infty(Y)$ for all i, j, let L in (7.1) be uniformly elliptic with ellipticity constant $c_a > 0$, that is,*

$$\xi^T A(y)\xi \geq c_a|\xi|^2\,, \quad \text{for all } y \in Y \text{ and all } \xi \in \mathbb{R}^n. \tag{7.3}$$

□

Definition 7.2 (Sobolev space of periodic functions)
Setting

$$C^\infty_{per}(Y) = \{v|Y : v \in C^\infty(\mathbb{R}^n),\ v \text{ is } Y\text{-periodic}\}\,, \tag{7.4}$$

we define $H^1_{per}(Y)$ to be the closure of $C^\infty_{per}(Y)$, considered as a subspace of $H^1(Y)$, with respect to the H^1 norm. □

The interval Y has $2n$ faces of dimension $n - 1$, namely for each i the opposite faces

$$S_i^0 = \{y : y_i = 0\,,\ y_j \in [0, l_j] \text{ for } j \neq i\}\,, \quad S_i^1 = \{y : y_i = l_i\,,\ y_j \in [0, l_j] \text{ for } j \neq i\}\,. \tag{7.5}$$

The periodicity condition

$$v(x + l_i e_i) = v(x)\,, \quad x \in S_i^0\,, \tag{7.6}$$

for $v \in C^\infty_{\text{per}}(Y)$ extends to functions $v \in H^1_{\text{per}}(Y)$ in the sense of traces,

$$(\gamma v)(x + l_i e_i) = (\gamma v)(x)\,, \quad \text{for a.e. } x \in S_i^0\,. \tag{7.7}$$

since the trace operator is continuous w.r.t. the H^1 norm.

In the following proposition, by v_p we denote the Y-periodic extension to $\mathbb{R}^n$ of a function $v : Y \to \mathbb{R}$, in order to make understanding easier.

Due to (7.7), for a given $v \in H^1_{\text{per}}(Y)$ the trace γv_p is well-defined on the grid

$$G = \bigcup \{x + \partial Y : x = \sum_{i=1}^n j_i l_i e_i,\ j_i \in \mathbb{Z}\}\,, \tag{7.8}$$

that is, almost everywhere on each face of dimension $n - 1$.

Proposition 7.3 *Let $v \in H^1_{per}(Y)$. Then we have $v_p|\Omega \in H^1(\Omega)$ for every bounded open set $\Omega \subset \mathbb{R}^n$, and*

$$\partial_i v_p = (\partial_i v)_p \,, \quad 1 \le i \le n \,. \tag{7.9}$$

Proof: Let $\varphi \in C_0^\infty(\Omega)$. We cover supp (φ) by an open interval I that consists of translated intervals

$$Y_k = y^k + Y \,, \quad y_i^k = j_i l_i \,, \quad j_i \in \mathbb{Z} \,, \tag{7.10}$$

such that

$$\text{supp}\,(\varphi) \subset \Omega \subset I \,, \quad \overline{I} = \left(\bigcup_k \overline{Y_k} \right) \,. \tag{7.11}$$

We have

$$-\int_\Omega v_p(y)\partial_i\varphi(y)\,dy = -\sum_k \int_{Y_k} v_p(y)\partial_i\varphi(y)\,dy \,. \tag{7.12}$$

Transformation to Y and partial integration yields

$$\begin{aligned}\int_{Y_k} v_p(y)\partial_i\varphi(y)\,dy &= \int_Y v(y)\partial_i\varphi(y+y^k)\,dy \\ &= -\int_Y (\partial_i v)(y)\varphi(y+y^k)\,dy + \int_{\partial Y} (\gamma v)(\xi)\varphi(\xi+y^k)\nu_i(\xi)\,dS(\xi) \,.\end{aligned}$$

Therefore,

$$-\int_\Omega v_p(y)\partial_i\varphi(y)\,dy = \sum_k \int_{Y_k} (\partial_i v)_p(y)\varphi(y)\,dy - \sum_k \int_{\partial Y_k} (\gamma v_p)(\xi)\varphi(\xi)\nu_i(\xi)\,dS(\xi) \,. \tag{7.13}$$

The second sum equals zero because on those faces of ∂Y_k which also belong to ∂I we have $\varphi = 0$, and because all other faces appear in exactly two boundary terms with opposite sign. It follows that

$$-\int_\Omega v_p(y)\partial_i\varphi(y)\,dy = \int_\Omega (\partial_i v)_p(x)\varphi(x)\,dx \,. \tag{7.14}$$

Since φ was arbitrary, the assertion follows. □

Lemma 7.4 *Let $v, w \in H^1_{per}(Y)$. Then*

$$\int_Y v(y)\partial_i w(y)\,dy = -\int_Y w(y)\partial_i v(y)\,dy \tag{7.15}$$

holds for all i. In particular (set $w = 1$)

$$\int_Y \nabla v(y)\,dy = 0 \,. \tag{7.16}$$

Proof: Since v and w are periodic, the boundary terms on opposite faces cancel each other out, so

$$\int_{\partial Y} [(\gamma v)(\gamma w)\nu](\xi)\,dS(\xi) = 0 \,.$$

$\square$

Since the solutions of (7.1), (7.2) can be unique only up to a constant, one considers the space

$$\tilde{H}(Y) = H^1_{\mathrm{per}}(Y)/\mathbb{R}\,. \tag{7.17}$$

The elements of $\tilde{H}(Y)$ are equivalence classes $[v]$ of functions $v \in H^1_{\mathrm{per}}(Y)$ satisfying

$$[u] = [v] \quad \Leftrightarrow \quad u - v \text{ is constant} \quad \Leftrightarrow \quad \nabla u = \nabla v\,. \tag{7.18}$$

The canonical norm in $\tilde{H}(Y)$ is the quotient norm

$$\|[v]\|_{\tilde{H}(Y)} = \inf_{w\in[v]} \|w\|_{H^1(Y)} = \inf_{c\in\mathbb{R}} \|v - c\|_{H^1(Y)}\,. \tag{7.19}$$

Proposition 7.5 *The mapping $[v] \mapsto \|\nabla v\|_{L^2(Y)}$ defines a norm on $\tilde{H}(Y)$, so that $\tilde{H}(Y)$ becomes a Banach space. Moreover, there exists $C > 0$ such that*

$$\|\nabla v\|_{L^2(Y)} \le \|[v]\|_{\tilde{H}(Y)} \le \|v - M_Y(v)\|_{H^1(Y)} \le C\|\nabla v\|_{L^2(Y)} \tag{7.20}$$

for all $v \in H^1(Y)$.

Proof: We have

$$\|\nabla v\|_{L^2(Y)} = \inf_{c\in\mathbb{R}} \|\nabla(v - c)\|_{L^2(Y)} \le \|[v]\|_{\tilde{H}(Y)} \le \|v - M_Y(v)\|_{H^1(Y)} \le C\|\nabla v\|_{L^2(Y)}\,.$$

The last inequality as well as the existence of C follow from Poincaré inequality for mean values (Proposition 6.9). That $\tilde{H}(Y)$ is a Banach space when equipped with the by (7.20) equivalent quotient norm (7.19), is a consequence of a general result of functional analysis. $\square$

In the following, for elements of $\tilde{H}(Y)$ we will also write v instead of $[v]$.

We return to the boundary value problem

$$Lu = -\sum_{i,j=1}^{n} \partial_j(a_{ij}(y)\partial_i u) = f\,, \quad \text{in } Y, \tag{7.21}$$

$$u \quad Y\text{-periodic.} \tag{7.22}$$

We consider the variational formulation: Find $u \in \tilde{H}(Y)$ such that

$$a(u, v) = F(v)\,, \quad \text{for all } v \in \tilde{H}(Y). \tag{7.23}$$

As before we set

$$a(u, v) = \int_Y \nabla u(y)^T A(y)\nabla v(y)\,dy\,, \quad F(v) = \int_Y f(y)v(y)\,dy\,. \tag{7.24}$$

Due to (7.18), the bilinear form a is well-defied on $\tilde{H}(Y)$. This does not apply to F unless we require in addition that

$$\int_Y f(y)\,dy = 0\,. \tag{7.25}$$

Indeed, then

$$\int_Y f(y)v(y)\,dy = \int_Y f(y)(v(y) - c)\,dy\,, \quad \text{for all } c \in \mathbb{R}. \tag{7.26}$$

Proposition 7.6 (Unique solvability, periodic bounddary conditions) *Let assumption 7.1 hold. Then the variational problem (7.23) has a unique solution $u \in \tilde{H}(Y)$ for every $f \in L^2(Y)$ that satisfies (7.25), and*

$$\|u\|_{\tilde{H}(Y)} \le \frac{C}{c_a}\|f\|_{L^2(Y)}\,, \tag{7.27}$$

where C denotes the constant from the Poincaré inequality for mean values.

Proof: The bilinear form a is continuous and $\tilde{H}(Y)$-elliptic. For the right side F we have for all v

$$\begin{aligned}|F(v)| = \left|\int_Y f(y)v(y)\,dy\right| = \left|\int_Y f(y)(v(y) - M_Y(v))\,dy\right| \le \|f\|_{L^2(Y)}\,\|v - M_Y(v)\|_{L^2(Y)} \\ \le \|f\|_{L^2(Y)}C\|\nabla v\|_{L^2(Y)} \le C\|f\|_{L^2(Y)}\|v\|_{\tilde{H}(Y)}\,,\end{aligned}$$

where again we have used Proposition 6.9. The linear form F therefore is continuous, too. The assertion now follows from Lax-Milgram. □

We may remove the multivalued character of the solution of the boundary value problem by a suitable normalization. We consider

$$Lu = -\sum_{i,j=1}^{n} \partial_j(a_{ij}(y)\partial_i u) = f\,, \quad \text{in } Y, \tag{7.28}$$

$$u \quad Y\text{-periodic}, \tag{7.29}$$

$$M_Y(u) = 0\,. \tag{7.30}$$

Corollary 7.7 *Under the assumptions of Proposition 7.6, the boundary value problem (7.28) – (7.30) has a unique weak solution $u \in H^1_{per}(Y)$.*

Proof: Proposition 7.6 yields a unique solution $[u]$ in $H^1_{\mathrm{per}}(Y)$. From this equivalence class we choose the unique function u which satisfies (7.30). □

Alternatively one may proceed as follows. We define

$$H_M(Y) = \{v : v \in H^1_{\mathrm{per}}(Y),\, M_Y(v) = 0\}\,. \tag{7.31}$$

$H_M(Y)$ is a closed subspace of $H^1_{\mathrm{per}}(Y)$, because $v \mapsto M_Y(v)$ is continuous, and therefore a Banach space.

Lemma 7.8 *The assignment $v \mapsto \|\nabla v\|_{L^2(Y)}$ defines a norm on $H_M(Y)$ which is equivalent to the H^1 norm.*

Proof: By Proposition 6.9,

$$\|v\|_{L^2(Y)} \le C\|\nabla v\|_{L^2(Y)}\,, \quad \text{for all } v \in H_M(Y),$$

for a suitable constant C. This yields the assertion. □

A variational formulation of the periodic boundary value problem in the space $H_M(Y)$ is the following.

$$\text{Find } u \in H_M(Y) \text{ such that } a(u,v) = F(v) \text{ for all } v \in H_M(Y). \tag{7.32}$$

Here, $F : H_M(Y) \to \mathbb{R}$ is a continuous linear functional, and a is given by (7.24). We consider an F of the form

$$F(v) = \int_Y \langle h(y), \nabla v(y)\rangle \, dy\,, \quad h : Y \to \mathbb{R}^n\,. \tag{7.33}$$

Proposition 7.9 (Unique solvability, second version)
Let assumption 7.1 hold. Then the variational problem (7.32), (7.33) has a unique solution $u \in H_M(Y)$ for each $h \in L^2(Y)^n$, and

$$\|\nabla u\|_{L^2(Y)} \le \frac{1}{c_a}\|h\|_{L^2(Y)^n}\,. \tag{7.34}$$

Proof: This proposition, too, is a direct consequence of Lax-Milgram. We use the norm from Lemma 7.8. □

Let $h \in H^1_{\text{per}}(Y)$. Then we have, according to Lemma 7.4 applied with $w = h_i$,

$$F(v) = \int_Y \langle h(y), \nabla v(y)\rangle \, dy = -\int_Y v(y)(\operatorname{div} h)(y)\, dy\,. \tag{7.35}$$

The corresponding boundary value problem is

$$Lu = -\sum_{i,j=1}^n \partial_j(a_{ij}(y)\partial_i u) = -\operatorname{div} h\,, \quad \text{in } Y, \tag{7.36}$$

$$u \quad Y\text{-periodic}, \tag{7.37}$$

$$M_Y(u) = 0\,. \tag{7.38}$$

One may ask whether the periodic extension u_p of the solution u of (7.32), (7.33) constitutes a weak solution of the partial differential equation with coefficients A_p and right side h_p, extended by periodicity from A and h. This is true in the following sense.

Proposition 7.10 *Let assumption 7.1 hold. If $u \in H_M(Y)$ solves (7.32), (7.33) for $h \in L^2(Y)^n$, then*

$$\int_{\mathbb{R}^n} \nabla u_p(x)^T A_p(x) \nabla\varphi(x)\, dx = \int_{\mathbb{R}^n} \langle h_p(x), \nabla\varphi(x)\rangle \, dx \tag{7.39}$$

for all $\varphi \in C_0^\infty(\mathbb{R}^n)$.

Proof: Let $\varphi \in C_0^\infty(\mathbb{R}^n)$ be arbitrary. Let $(Y_k)_{k\in K}$ be a finite open covering of $\operatorname{supp}(\varphi)$ with translated intervals $Y_k = y^k + Y$, let $(\psi_k)_{k\in K}$ be an associated C^∞ partition of unity.

For each $k \in K$ we define the function $(\varphi\psi_k)_p$ as the Y-periodic extension of $(\varphi\psi_k)|Y_k$. Then

$$\int_{\mathbb{R}^n} \nabla u_p(x)^T A_p(x) \nabla \varphi(x)\, dx = \sum_{k\in K} \int_{Y_k} \nabla u_p(x)^T A_p(x) \nabla(\varphi\psi_k)(x)\, dx$$
$$= \sum_{k\in K} \int_Y \nabla u(x)^T A(x) \nabla(\varphi\psi_k)_p(x)\, dx = \sum_{k\in K} \int_Y \langle h(x), \nabla(\varphi\psi_k)_p(x)\rangle \; dx$$
$$= \sum_{k\in K} \int_{Y_k} \langle h_p(x), \nabla(\varphi\psi_k)(x)\rangle \; dx = \int_{\mathbb{R}^n} \langle h_p(x), \nabla\varphi(x)\rangle \; dx\,.$$

□

8 Homogenization: The multidimensional case

On an open and bounded set $\Omega \subset \mathbb{R}^n$ we consider the boundary value problem

$$-\sum_{i,j=1}^{n} \partial_j (a_{ij}^{\varepsilon}(x) \partial_i u_\varepsilon) = f\,, \quad \text{in } \Omega, \tag{8.1}$$

$$u_\varepsilon = 0\,, \quad \text{on } \partial\Omega. \tag{8.2}$$

The coefficients are given by

$$a_{ij}^{\varepsilon}(x) = a_{ij}\left(\frac{x}{\varepsilon}\right)\,, \quad 1 \le i,j \le n\,, \tag{8.3}$$

Written in matrix-vector form,(8.1), (8.3) becomes

$$-\mathrm{div}\,(A^\varepsilon(x)^T \nabla u) = f\,, \quad A^\varepsilon(x) = A\left(\frac{x}{\varepsilon}\right)\,. \tag{8.4}$$

Again, let

$$Y = \prod_{i=1}^{n} (0, l_i) \tag{8.5}$$

be a fixed reference interval.

Assumption 8.1 *Let $a_{ij} \in L^\infty(\mathbb{R}^n)$ be Y-periodic and uniformly elliptic with ellipticity constant c_a.* □

As we know already, under this assumption the boundary value problem (8.1), (8.2) has a unique weak solution $u_\varepsilon \in H_0^1(\Omega)$ for every $f \in L^2(\Omega)$,

$$\int_\Omega \nabla u_\varepsilon(x)^T A^\varepsilon(x) \nabla v(x)\, dx = \int_\Omega f(x) v(x)\, dx\,, \quad \text{for all } v \in H_0^1(\Omega). \tag{8.6}$$

We want to find a matrix A^0 of coefficients (the "homogenized coefficient matrix") such that, as $\varepsilon \to 0$, the solutions u_ε of (8.1), (8.2) converge to the solution u_0 of

$$-\sum_{i,j=1}^{n} \partial_j (a_{ij}^{0}(x) \partial_i u_0) = f\,, \quad \text{in } \Omega, \tag{8.7}$$

$$u_0 = 0\,, \quad \text{on } \partial\Omega. \tag{8.8}$$

It turns out that the matrix A^0 arises as the result of averaging the solutions of certain boundary value problems on Y with periodic boundary conditions, in the following way.

For a given $\lambda \in \mathbb{R}^n$ we look for the weak solution $\chi_\lambda \in H_M(Y)$ of the boundary value problem (in variational formulation)

$$\tilde{a}(\chi_\lambda, v) = F_\lambda(v)\,, \quad \text{for all } v \in H_M(Y), \tag{8.9}$$

where

$$\tilde{a}(u,v) = \int_Y \nabla u(y)^T A(y)^T \nabla v(y)\, dy = \int_Y \langle A(y)\nabla u(y), \nabla v(y)\rangle\, dy\,, \tag{8.10}$$

$$F_\lambda(v) = \int_Y \lambda^T A(y)^T \nabla v(y)\, dy = \int_Y \langle A(y)\lambda, \nabla v(y)\rangle\, dy\,. \tag{8.11}$$

According to Proposition 7.9, this problem has a unique solution $\chi_\lambda \in H_M(Y)$. The associated "standard" formulation is

$$-\mathrm{div}\,(A(y)\nabla\chi_\lambda) = -\mathrm{div}\,(A(y)\lambda)\,, \tag{8.12}$$

$$\chi_\lambda \text{ } Y\text{-periodic, } M_Y(\chi_\lambda) = 0. \tag{8.13}$$

If the matrix A of coefficients has weak derivatives as a function of y, then the right side of (8.12) defines a function on Y. If this is not the case, for example when A is discontinuous, one can understand (8.12) only in the weak sense (8.9) – (8.11), and F_λ in (8.11) has to be understood as an element the dual space of $H_0^1(Y)$, which is denoted by $H^{-1}(Y)$.

Having obtained χ_λ, we define another auxiliary function w_λ,

$$w_\lambda(y) = -\chi_\lambda(y) + \lambda^T y\,. \tag{8.14}$$

It will turn out that the homogenized coefficient matrix can be computed by the equation

$$A^0\lambda = M_Y(A\nabla w_\lambda)\,, \tag{8.15}$$

that is, we obtain the ith column of A^0 by setting $\lambda = e_i$ in (8.15). (Since the mapping $\lambda \mapsto M_Y(A\nabla w_\lambda)$ is linear, (8.15) then holds for all $\lambda \in \mathbb{R}^n$.)

In particular, the functions a_{ij}^0 in (8.7) are **constant**, since the matrix A^0 does not depend on x.

Theorem 8.2 (Periodic homogenization, convergence)
Let assumption 8.1 hold, let the constant matrix A^0 be given by (8.15), let $f \in L^2(\Omega)$. Then A^0 is uniformly elliptic, the boundary value problem (8.7), (8.8) has a unique weak solution $u_0 \in H_0^1(\Omega)$, and the solutions u_ε of (8.1), (8.2) satisfy

$$u_\varepsilon \rightharpoonup u_0 \quad \text{in } H_0^1(\Omega), \tag{8.16}$$

$$(A^\varepsilon)^T\nabla u_\varepsilon \rightharpoonup (A^0)^T\nabla u_0 \quad \text{in } L^2(\Omega)^n. \tag{8.17}$$

Proof: This will be developed in the remainder of this chapter. □

Lemma 8.3 (Weak convergence of products)
(i) Let $(v_k)_{k\in\mathbb{N}}$, $(w_k)_{k\in\mathbb{N}}$ be sequences in $L^2(\Omega)$ such that

$$v_k \to v \quad \text{in } L^2(\Omega), \quad w_k \rightharpoonup w \quad \text{in } L^2(\Omega). \tag{8.18}$$

Then

$$\int_\Omega v_k(x)w_k(x)\,dx \to \int_\Omega v(x)w(x)\,dx\,. \tag{8.19}$$

(i) Let $(v_k)_{k\in\mathbb{N}}$ be a sequence in $L^\infty(\Omega)$, $(w_k)_{k\in\mathbb{N}}$ a sequence in $L^2(\Omega)$ such that

$$v_k \to v \quad \text{uniformly}, \quad w_k \rightharpoonup w \quad \text{in } L^2(\Omega). \tag{8.20}$$

Then

$$v_k w_k \rightharpoonup vw \quad \text{in } L^2(\Omega). \tag{8.21}$$

Proof: Exercise. $\square$

We define

$$\xi_\varepsilon(x) = (A^\varepsilon(x))^T \nabla u_\varepsilon(x)\,, \quad x \in \Omega\,. \tag{8.22}$$

We have $\xi_\varepsilon \in L^2(\Omega)^n$, since $A^\varepsilon(x)$ is uniformly bounded in x,

Lemma 8.4 *There exist $u_0 \in H_0^1(\Omega)$ and $\xi_0 \in L^2(\Omega)^n$ such that for some subsequence*

$$u_\varepsilon \rightharpoonup u_0 \quad in\ H_0^1(\Omega), \quad u_\varepsilon \to u_0 \quad in\ L^2(\Omega), \tag{8.23}$$

$$\xi_\varepsilon \rightharpoonup \xi_0 \quad in\ L^2(\Omega)^n. \tag{8.24}$$

Moreover,

$$\int_\Omega f(x)v(x)\,dx = \int_\Omega \langle \xi_\varepsilon, \nabla v\rangle\,dx = \int_\Omega \langle \xi_0, \nabla v\rangle\,dx \tag{8.25}$$

for all $v \in H_0^1(\Omega)$ and all $\varepsilon > 0$.

Proof: Due to Lax-Milgram, for a suitable constant C

$$\|u_\varepsilon\|_{H_0^1(\Omega)} \le \frac{C}{c_a}\|f\|_{L^2(\Omega)} \tag{8.26}$$

for all $\varepsilon > 0$. The set $(u_\varepsilon)_{\varepsilon>0}$ is thus bounded in $H_0^1(\Omega)$. Consequently, there exists a sequence (u_{ε_n}) that weakly converges to some $u_0 \in H_0^1(\Omega)$. Due to Corollary 6.8, $u_{\varepsilon_n} \to u_0$ strongly in $L^2(\Omega)$. As $A^\varepsilon(x)$ is bounded uniformly with respect to x and ε, the set $(\xi_\varepsilon)_{\varepsilon>0}$ is bounded in $L^2(\Omega)^n$. Passing to a subsequence, we also obtain (8.24). The left equation in (8.25) is the same as (8.6), the right equation follows from the weak convergence $\xi_\varepsilon \rightharpoonup \xi_0$. $\square$

We define

$$w_\lambda^\varepsilon(x) = \varepsilon w_\lambda\left(\frac{x}{\varepsilon}\right) = \lambda^T x - \varepsilon\chi_\lambda\left(\frac{x}{\varepsilon}\right)\,. \tag{8.27}$$

Lemma 8.5 *Set $\Lambda(x) = \lambda^T x$. Then*

$$w_\lambda^\varepsilon \rightharpoonup \Lambda \quad in\ H^1(\Omega), \quad w_\lambda^\varepsilon \to \Lambda \quad in\ L^2(\Omega), \tag{8.28}$$

the latter for a suitable sequence $\varepsilon_n \to 0$.

Proof: The functions

$$x \mapsto \chi_\lambda\left(\frac{x}{\varepsilon}\right)$$

weakly converge to $M_Y(\chi_\lambda) = 0$ in $L^2(\Omega)$ as $\varepsilon \to 0$ by Proposition 6.5. It follows that $w_\lambda^\varepsilon \rightharpoonup \Lambda$ in $L^2(\Omega)$. Moreover,

$$\nabla w_\lambda^\varepsilon(x) = \lambda - \nabla\chi_\lambda\left(\frac{x}{\varepsilon}\right)\,.$$

The function $y \mapsto \nabla\chi_\lambda(y)$ is Y-periodic. Therefore, again by Proposition 6.5

$$\nabla w_\lambda^\varepsilon \rightharpoonup \lambda - M_Y(\nabla\chi_\lambda) \quad \text{in } L^2(\Omega)^n.$$

By Lemma 7.4 we get $M_Y(\nabla\chi_\lambda) = 0$. The first assertion in (8.28) is now proved, the second follows from the compactness of the embedding $H^1(\Omega) \subset\subset L^2(\Omega)$. $\square$

We now define

$$\eta_\lambda^\varepsilon(x) = A^\varepsilon(x)\nabla w_\lambda^\varepsilon(x)\,. \tag{8.29}$$

Lemma 8.6 *We have*

$$\eta_\lambda^\varepsilon \rightharpoonup A^0\lambda \quad \text{in } L^2(\Omega)^n, \tag{8.30}$$

and

$$\int_\Omega \langle \eta_\lambda^\varepsilon(x), \nabla v(x)\rangle \, dx = 0 \,, \quad \text{for all } v \in H_0^1(\Omega). \tag{8.31}$$

Proof: From

$$\eta_\lambda^\varepsilon(x) = A\left(\frac{x}{\varepsilon}\right) \nabla w_\lambda \left(\frac{x}{\varepsilon}\right)$$

we obtain, by virtue of Proposition 6.5,

$$\eta_\lambda^\varepsilon \rightharpoonup M_Y(A\nabla w_\lambda) = A^0\lambda \quad \text{in } L^2(\Omega)^n.$$

Let now $\varphi \in C_0^\infty(\Omega)$ be arbitrary. We set

$$\varphi_\varepsilon(y) = \varphi(\varepsilon y)\,.$$

Then $\varphi_\varepsilon \in C_0^\infty(\mathbb{R}^n)$. We apply Proposition 7.10 to the variational problem (8.9) and obtain

$$\int_{\mathbb{R}^n} \langle A(y)\nabla \chi_\lambda(y), \nabla\varphi_\varepsilon(y)\rangle \, dy = \int_{\mathbb{R}^n} \langle A(y)\lambda, \nabla\varphi_\varepsilon(y)\rangle \, dy\,,$$

therefore

$$\begin{aligned} 0 &= \int_{\mathbb{R}^n} \langle A(y)\nabla w_\lambda(y), \nabla\varphi_\varepsilon(y)\rangle \, dy = \int_{\mathbb{R}^n} \left\langle A\left(\frac{x}{\varepsilon}\right) \nabla w_\lambda\left(\frac{x}{\varepsilon}\right), \nabla\varphi_\varepsilon\left(\frac{x}{\varepsilon}\right)\right\rangle \, dx \\ &= \int_\Omega \langle \eta_\lambda^\varepsilon(x), \nabla\varphi(x)\rangle \, dx\,. \end{aligned}$$

Since $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, the assertion follows. $\square$

Proof of Theorem 8.2, the convergence part. We now prove that (8.16) and (8.17) hold for every convergent subsequence. Let $\varphi \in C_0^\infty(\Omega)$ be arbitrary. As a test function in (8.25) we use $v = \varphi w_\lambda^\varepsilon$ and obtain (we omit the argument x)

$$\begin{aligned} \int_\Omega f\varphi w_\lambda^\varepsilon \, dx &= \int_\Omega \langle \xi_\varepsilon, \nabla(\varphi w_\lambda^\varepsilon)\rangle \, dx \\ &= \int_\Omega \langle \xi_\varepsilon, \nabla w_\lambda^\varepsilon\rangle \, \varphi \, dx + \int_\Omega \langle \xi_\varepsilon, \nabla\varphi\rangle \, w_\lambda^\varepsilon \, dx\,. \end{aligned} \tag{8.32}$$

In (8.31) we use $v = \varphi u_\varepsilon$ as a test function and obtain

$$0 = \int_\Omega \langle \eta_\lambda^\varepsilon, \nabla(\varphi u_\varepsilon)\rangle \, dx = \int_\Omega \langle \eta_\lambda^\varepsilon, \nabla u_\varepsilon\rangle \, \varphi \, dx + \int_\Omega \langle \eta_\lambda^\varepsilon, \nabla\varphi\rangle \, u_\varepsilon \, dx\,. \tag{8.33}$$

We look at the scalar products in the first integrals on the right sides of both (8.32) and (8.33). Both factors in these scalar products are weakly convergent only, so that we cannot pass to the limits directly. But according to the definition of ξ_ε and η_λ^ε,

$$\langle \xi_\varepsilon, \nabla w_\lambda^\varepsilon\rangle = \left\langle A^{\varepsilon T}\nabla u_\varepsilon, \nabla w_\lambda^\varepsilon\right\rangle = \langle \nabla u_\varepsilon, A^\varepsilon \nabla w_\lambda^\varepsilon\rangle = \langle \nabla u_\varepsilon, \eta_\lambda^\varepsilon\rangle \,. \tag{8.34}$$

We subtract (8.33) from (8.32) and obtain because of (8.34)

$$\int_\Omega \langle \xi_\varepsilon, \nabla\varphi\rangle\, w_\lambda^\varepsilon\, dx - \int_\Omega \langle \eta_\lambda^\varepsilon, \nabla\varphi\rangle\, u_\varepsilon\, dx = \int_\Omega f\varphi w_\lambda^\varepsilon\, dx\,. \tag{8.35}$$

Now we may pass to the limit. By Lemma 8.4, Lemma 8.5 and Lemma 8.3, we get

$$\int_\Omega \langle \xi_\varepsilon, \nabla\varphi\rangle\, w_\lambda^\varepsilon\, dx \to \int_\Omega \langle \xi_0, \nabla\varphi\rangle\, \Lambda\, dx \tag{8.36}$$

for a suitable subsequence. By Lemma 8.4, Lemma 8.6 and Lemma 8.3,

$$\int_\Omega \langle \eta_\lambda^\varepsilon, \nabla\varphi\rangle\, u_\varepsilon\, dx \to \int_\Omega \left\langle A^0\lambda, \nabla\varphi\right\rangle u_0\, dx \tag{8.37}$$

for a subsequence, and moreover

$$\int_\Omega f\varphi w_\lambda^\varepsilon\, dx \to \int_\Omega f\varphi\Lambda\, dx\,. \tag{8.38}$$

Thus we obtain

$$\int_\Omega \langle \xi_0, \nabla\varphi\rangle\, \Lambda\, dx - \int_\Omega \left\langle A^0\lambda, \nabla\varphi\right\rangle u_0\, dx = \int_\Omega f\varphi\Lambda\, dx\,. \tag{8.39}$$

In (8.25) we now use $v = \varphi\Lambda$ as a test function and get

$$\int_\Omega f\varphi\Lambda\, dx = \int_\Omega \langle \xi_0, \nabla(\varphi\Lambda)\rangle\; dx = \int_\Omega \langle \xi_0, \nabla\varphi\rangle\, \Lambda\, dx + \int_\Omega \langle \xi_0, \lambda\rangle\, \varphi\, dx\,. \tag{8.40}$$

Inserting (8.40) into (8.39) we get, since $A^0\lambda$ is a constant vector,

$$\int_\Omega \langle \xi_0, \lambda\rangle\, \varphi\, dx = -\int_\Omega \left\langle A^0\lambda, \nabla\varphi\right\rangle u_0\, dx = \int_\Omega \left\langle A^0\lambda, \nabla u_0\right\rangle \varphi\, dx\,, \tag{8.41}$$

for all $\varphi \in C_0^\infty(\Omega)$. This implies

$$\langle \xi_0, \lambda\rangle = \left\langle A^0\lambda, \nabla u_0\right\rangle = \left\langle (A^0)^T\nabla u_0, \lambda\right\rangle\,. \tag{8.42}$$

Since (8.42) holds for all $\lambda \in \mathbb{R}^n$, it follows that

$$\xi_0 = (A^0)^T\nabla u_0\,. \tag{8.43}$$

□

We now consider another periodic auxiliary problem. It differs from von (8.9) – (8.12) only in that the matrix A is replaced by the matrix A^T. For a given $\lambda \in \mathbb{R}^n$ we want to find $\hat\chi_\lambda \in H_M(Y)$ such that

$$a(\hat\chi_\lambda, v) = \hat F_\lambda(v)\,, \quad \text{for all } v \in H_M(Y), \tag{8.44}$$

where $a(u,v) = \int_Y \nabla u^T A\nabla v\, dy$ and

$$\hat F_\lambda(v) = \int_Y \lambda^T A(y)\nabla v(y)\, dy\,. \tag{8.45}$$

We set

$$\hat w_\lambda(y) = -\hat\chi_\lambda(y) + \lambda^T y\,. \tag{8.46}$$

Lemma 8.7 *For all $\lambda \in \mathbb{R}^n$,*

$$(A^0)^T \lambda = M_Y(A^T \nabla \hat{w}_\lambda) \,. \tag{8.47}$$

Proof: Let $\lambda, \mu \in \mathbb{R}^n$ be arbitrary. According to the definition of A^0 we then have

$$\mu^T (A^0)^T \lambda = \lambda^T A^0 \mu = \lambda^T M_Y(A \nabla w_\mu) = \lambda^T M_Y(A\mu) - \lambda^T M_Y(A \nabla \chi_\mu) \,. \tag{8.48}$$

Moreover,

$$\lambda^T M_Y(A\mu) = M_Y(\lambda^T A \mu) = M_Y(\mu^T A^T \lambda) = \mu^T M_Y(A^T \lambda) \,. \tag{8.49}$$

In order to transform the rightmost term in (8.48) we use the variational equations for χ_λ and $\hat{\chi}_\lambda$. We compute

$$\begin{aligned} |Y| \lambda^T M_Y(A \nabla \chi_\mu) &= \int_Y \lambda^T A(y) \nabla \chi_\mu(y) \, dy = \int_Y \nabla \hat{\chi}_\lambda(y)^T A(y) \nabla \chi_\mu(y) \, dy \\ &= \int_Y \nabla \chi_\mu(y)^T A(y)^T \nabla \hat{\chi}_\lambda(y) \, dy = \int_Y \mu^T A(y)^T \nabla \hat{\chi}_\lambda(y) \, dy \\ &= |Y| \mu^T M_Y(A^T \nabla \hat{\chi}_\lambda) \,. \end{aligned}$$

Thus we obtain

$$\mu^T (A^0)^T \lambda = \mu^T M_Y(A^T \lambda - A^T \nabla \hat{\chi}_\lambda) = \mu^T M_Y(A^T \nabla \hat{w}_\lambda) \,.$$

□

Corollary 8.8 *The homogenized matrix A^0 satisfies*

$$(A^T)^0 = (A^0)^T \,. \tag{8.50}$$

If A is symmetric, then so is A^0.

Proof: Replacing A by A^T in the definition of the auxiliary periodic problem we obtain

$$(A^T)^0 \lambda = M_Y(A^T \nabla \hat{w}_\lambda) \,.$$

Now Lemma 8.7 implies (8.50). If A is symmetric,

$$(A^0)^T = (A^T)^0 = A^0 \,.$$

□

Setting $\lambda = e_j$ we obtain formulas for the elements a^0_{ij} from the equation defining A^0. We set $\chi_j = \chi_{e_j}$, $w_j = w_{e_j}$, so

$$\chi_j(y) + w_j(y) = e_j^T y = y_j \,. \tag{8.51}$$

Lemma 8.9 *We have*

$$a^0_{ij} = M_Y(e_i^T A \nabla w_j) = M_Y(a_{ij}) - M_Y\left(\sum_{k=1}^n a_{ik} \partial_k \chi_j \right) \,. \tag{8.52}$$

Proof: We have

$$
\begin{aligned}
a^0_{ij} &= e_i^T A^0 e_j = e_i^T M_Y(A\nabla w_j) = M_Y(e_i^T A\nabla w_j)\\
&= M_Y(e_i^T A e_j) - M_Y(e_i^T A\nabla\chi_j) = M_Y(a_{ij}) - M_Y\left(\sum_{k=1}^n a_{ik}\partial_k\chi_j\right).
\end{aligned}
$$

□

Thus, the difference $A^0 - M_Y(A)$ is a matrix whose elements are $-M_Y(e_i^T A\nabla\chi_j)$. Due to this, the function χ_j are called **correctors**.

Lemma 8.10 *We have*

$$a^0_{ij} = M_Y(\nabla w_j^T A^T \nabla w_i)\,. \tag{8.53}$$

Proof: We use χ_i as a test function in (8.9) for $\lambda = e_j$,

$$\tilde{a}(\chi_j,\chi_i) = F_j(\chi)\,, \quad F_j := F_{e_j}\,.$$

Then

$$\int_Y \nabla\chi_j^T A^T \nabla\chi_i\,dy = \int_Y e_j^T A^T\nabla\chi_i\,dy\,,$$

therefore

$$0 = M_Y(\nabla w_j^T A^T \nabla\chi_i)\,. \tag{8.54}$$

From Lemma 8.9 we obtain

$$a^0_{ij} = M_Y(e_i^T A\nabla w_j) = M_Y(\nabla w_j^T A^T e_i)\,. \tag{8.55}$$

Subtracting (8.54) from (8.55) we arrive at (8.53). □

Lemma 8.11 *The matrix A^0 is positive definite.*

Proof: Let $\xi \in \mathbb{R}^n$ be arbitrary. Using Lemma 8.10 we get

$$\xi^T A^0 \xi = \sum_{i,j=1}^n \xi_i a^0_{ij}\xi_j = \sum_{i,j=1}^n \xi_i M_Y(\nabla w_j^T A^T\nabla w_i)\xi_j = \sum_{i,j=1}^n M_Y(\xi_j\nabla w_j^T A^T \nabla w_i \xi_i)\,. \tag{8.56}$$

We set

$$\zeta(y) = \sum_{k=1}^n \xi_k w_k(y)\,. \tag{8.57}$$

From (8.56) it follows that

$$\xi^T A^0\xi = M_Y(\nabla\zeta^T A^T\nabla\zeta) = M_Y(\nabla\zeta^T A\nabla\zeta) \geq 0\,, \tag{8.58}$$

since by assumption for A

$$\nabla\zeta^T(y)A(y)\nabla\zeta(y) \geq 0\,, \quad \text{a.e. in } Y. \tag{8.59}$$

Let us assume that there exists a $\xi \neq 0$ with $\xi^T A^0 \xi = 0$. Then it follows from (8.58) and (8.59) that

$$\nabla\zeta^T(y)A(y)\nabla\zeta(y) = 0\,, \quad \text{a.e. in } Y. \tag{8.60}$$

Since A is uniformly elliptic, $\nabla\zeta = 0$ a.e. in Y, and moreover, since $M_Y(\nabla\chi_k) = 0$ by Lemma 7.4,

$$0 = M_Y(\nabla\zeta) = \sum_{k=1}^n \xi_k M_Y(\nabla w_k) = \sum_{k=1}^n \xi_k M_Y(e_k - \nabla\chi_k) = \sum_{k=1}^n \xi_k M_Y(e_k) = \sum_{k=1}^n \xi_k e_k$$
$$= \xi\,,$$

which contradicts the assumption $\xi \neq 0$. □

Completion of the proof of Theorem 8.2. For all $\xi \in \mathbb{R}^n$ with $\xi \neq 0$ we have

$$\xi^T A^0 \xi = |\xi|^2 \frac{\xi^T}{|\xi|} A^0 \frac{\xi}{|\xi|} \geq c_0|\xi|^2\,, \tag{8.61}$$

where

$$c_0 := \min_{|e|=1} e^T A^0 e > 0 \tag{8.62}$$

due to Lemma 8.11 and because the unit ball in $\mathbb{R}^n$ is compact. Therefore, A^0 is uniformly elliptic. Therefore, the solution u_0 of (8.7), (8.8) is unique by Lax-Milgram.

According to what we have proved so far: If $(\varepsilon_k)_{k\in\mathbb{N}} \to 0$, then there exists a subsequence $(\varepsilon_{k_m})_{m\in\mathbb{N}}$ satisfying

$$u_{\varepsilon_{k_m}} \rightharpoonup u_0\,.$$

But now the "convergence principle" implies that $u_{\varepsilon_k} \rightharpoonup u_0$ for every sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ which converges to 0. Thus, Theorem 8.2 is completely proved. □

We now apply Theorem 8.2 to two situations in which we can explicitly compute the homogenized matrix A^0.

First, we reproduce for the one dimensional case the result which we already have obtained in a previous chapter. We consider

$$-\partial_x\left(a\left(\frac{x}{\varepsilon}\right)\partial_x u_\varepsilon\right) = f\,, \quad \text{in } \Omega = (d_1, d_2), \tag{8.63}$$

$$u_\varepsilon(d_1) = u_\varepsilon(d_2) = 0\,. \tag{8.64}$$

On the periodicity interval $Y = (0, l_1)$ of a, the auxiliary problem becomes ($\lambda = 1$)

$$\int_Y \chi'(y)a(y)v'(y)\,dy = \int_Y a(y)v'(y)\,dy\,, \quad \text{for all } v \in H_M(Y), \tag{8.65}$$

which has a unique solution $\chi \in H_M(Y)$. For arbitrary $v \in C_0^\infty(Y)$ we have $v - M_Y(v) \in H_M(Y)$, thus

$$\int_Y (\chi'(y)a(y) - a(y))v'(y)\,dy = 0\,. \tag{8.66}$$

It follows that

$$\chi'(y)a(y) - a(y) = c \in \mathbb{R}\,, \tag{8.67}$$

since by (8.66) the weak derivative of the left side of (8.67) is equal to zero. Moreover,

$$\chi'(y) = \frac{c}{a(y)} + 1\,, \quad w'(y) = 1 - \chi'(y) = -\frac{c}{a(y)}\,. \tag{8.68}$$

By the definition of a^0 we get

$$a^0 = M_Y(aw') = M_Y(-c) = -c\,. \tag{8.69}$$

Since χ is Y-periodic,

$$0 = \int_Y \chi'(y)\,dy = l_1 + l_1 c M_Y\left(\frac{1}{a}\right)\,, \tag{8.70}$$

therefore

$$c = -\left(M_Y\left(\frac{1}{a}\right)\right)^{-1}\,, \quad a^0 = \left(M_Y\left(\frac{1}{a}\right)\right)^{-1}\,. \tag{8.71}$$

As a second example we consider a layered two-dimensional medium, that is, we have $n = 2$, $Y = Y_1 \times Y_2 = (0, l_1) \times (0, l_2)$, and A depends only upon y_1, $\partial_2 A = 0$. The problem now becomes

$$-\sum_{i,j=1}^{2} \partial_j \left(a_{ij}\left(\frac{x_1}{\varepsilon}\right)\partial_i u_\varepsilon(x_1, x_2)\right) = f(x_1, x_2)\,, \quad \text{in } \Omega \subset \mathbb{R}^2, \tag{8.72}$$

$$u_\varepsilon = 0\,, \quad \text{on } \partial\Omega. \tag{8.73}$$

We compute the coefficients of A^0. The periodic problem

$$\int_Y \nabla\chi_j^T A^T \nabla v\,dy = \int_Y e_j^T A^T \nabla v\,dy\,, \quad \text{for all } v \in H_M(Y), \tag{8.74}$$

has unique solutions χ_j for $j = 1, 2$. We want to prove that they depend upon y_1 only, and that they can be obtained from solving a suitable one-dimensional problem. For this purpose we consider the unique solution $\tilde{\chi}_j \in H_M(Y_1)$ of the periodic problem

$$\int_{Y_1} \tilde{\chi}_j'(y_1) a_{11}(y_1) \tilde{v}'(y_1)\,dy_1 = \int_{Y_1} a_{1j}(y_1)\tilde{v}'(y_1)\,dy_1\,, \quad \text{for all } \tilde{v} \in H_M(Y_1). \tag{8.75}$$

Let $v \in H_M(Y)$, let

$$\tilde{v}(y_1) = \int_{Y_2} v(y_1, y_2)\,dy_2\,. \tag{8.76}$$

Then

$$\tilde{v} \in H_M(Y_1)\,, \quad \tilde{v}'(y_1) = \int_{Y_2} \partial_1 v(y_1, y_2)\,dy_2\,. \tag{8.77}$$

We now define

$$\chi_j(y_1, y_2) = \tilde{\chi}_j(y_1)\,. \tag{8.78}$$

For arbitrary $v \in H_M(Y)$ we get

$$\int_{Y_2} \partial_2 v(y_1, y_2)\,dy_2 = v(y_1, l_2) - v(y_1, 0) = 0\,. \tag{8.79}$$

From (8.75) – (8.77) it follows that

$$\begin{aligned}\int_Y \nabla\chi_j^T A^T \nabla v\,dy &= \int_Y \tilde{\chi}_j'(y_1)[a_{11}(y_1)\partial_1 v(y_1,y_2) + a_{21}(y_1)\partial_2 v(y_1,y_2)]\,dy \\ &= \int_{Y_1} \tilde{\chi}_j'(y_1)a_{11}(y_1)\int_{Y_2}\partial_1 v(y_1,y_2)\,dy_2\,dy_1 \\ &= \int_{Y_1} a_{1j}(y_1)\int_{Y_2}\partial_1 v(y_1,y_2)\,dy_2\,dy_1 = \int_Y e_j^T A^T \nabla v\,dy\,.\end{aligned}$$

Therefore, the function $\chi_j \in H_M(Y)$ defined by (8.78) indeed is the solution of (8.74). We now can compute $\tilde{\chi}_j$ in analogy to the one-dimensional case. We have

$$\tilde{\chi}_j'(y_1)a_{11}(y_1) - a_{1j}(y_1) = c_j\,, \quad j = 1,2\,. \tag{8.80}$$

Since $\tilde{\chi}_j$ is periodic,

$$0 = \int_{Y_1} \tilde{\chi}_j'(y_1)\,dy_1 = c_j M_{Y_1}\left(\frac{1}{a_{11}}\right) + M_{Y_1}\left(\frac{a_{1j}}{a_{11}}\right)\,. \tag{8.81}$$

We obtain the coefficients of A^0 from Lemma 8.9. We have

$$a_{ij}^0 = M_Y(a_{ij}) - M_Y(a_{i1}\partial_1\chi_j + a_{i2}\partial_2\chi_j) = M_Y(a_{ij} - a_{i1}\tilde{\chi}_j')\,. \tag{8.82}$$

We consider the case $j = 1$. According to (8.80),

$$1 - \tilde{\chi}_1' = -\frac{c_1}{a_{11}}\,, \quad c_1 = -\left(M_{Y_1}\left(\frac{1}{a_{11}}\right)\right)^{-1}\,, \tag{8.83}$$

therefore

$$a_{11}^0 = M_Y(a_{11} - a_{11}\tilde{\chi}_1') = M_Y(-c_1) = -c_1 = \left(M_{Y_1}\left(\frac{1}{a_{11}}\right)\right)^{-1}\,, \tag{8.84}$$

$$a_{21}^0 = M_Y(a_{21} - a_{21}\tilde{\chi}_1') = M_Y\left(-a_{21}\frac{c_1}{a_{11}}\right) = a_{11}^0 M_{Y_1}\left(\frac{a_{21}}{a_{11}}\right)\,. \tag{8.85}$$

In the case $j = 2$ we get from (8.80) and (8.81)

$$\tilde{\chi}_2' = \frac{c_2}{a_{11}} + \frac{a_{12}}{a_{11}}\,, \quad c_2 = -a_{11}^0 M_{Y_1}\left(\frac{a_{12}}{a_{11}}\right)\,. \tag{8.86}$$

This finally yields

$$a_{12}^0 = M_Y(a_{12} - a_{11}\tilde{\chi}_2') = M_Y(-c_2) = -c_2 = a_{11}^0 M_{Y_1}\left(\frac{a_{12}}{a_{11}}\right)\,, \tag{8.87}$$

and

$$\begin{aligned}a_{22}^0 &= M_Y(a_{22} - a_{21}\tilde{\chi}_2') = M_Y\left(a_{22} - a_{21}\left(\frac{c_2}{a_{11}} + \frac{a_{12}}{a_{11}}\right)\right) \\ &= M_{Y_1}\left(a_{22} - a_{21}\frac{a_{12}}{a_{11}}\right) + a_{11}^0 M_{Y_1}\left(\frac{a_{12}}{a_{11}}\right) M_{Y_1}\left(\frac{a_{21}}{a_{11}}\right)\,.\end{aligned} \tag{8.88}$$

In particular we obtain that A^0 is diagonal if so is A.

9 Monotone Problems

We consider the elliptic boundary value problem

$$\begin{aligned} -\mathrm{div}\,(a(\nabla u)) &= f\,, \quad \text{in } \Omega, \\ u(x) &= 0\,, \quad \text{on } \partial\Omega. \end{aligned} \tag{9.1}$$

The function $a : \mathbb{R}^n \to \mathbb{R}^n$ is given. If a is linear, the problem is linear, and if the matrix representing a is positive definite, we have a linear elliptic problem. The special case $a(y) = y$ corresponds to the equation $-\Delta u = f$. If a is nonlinear, one speaks of a **quasilinear** problem; performing differentiation in $\mathrm{div}\,(a(\nabla u(x)))$ with respect to x leads to an expression which is linear with respect to the derivative of highest (in this case second) order, but nonlinear with respect to the lower order derivatives.

The variational formulation of (9.1) becomes

$$\int_\Omega \langle a(\nabla u(x)), \nabla v(x)\rangle \; dx = \int_\Omega f(x)v(x)\,dx\,, \quad \text{for all } v \in H_0^1(\Omega), \tag{9.2}$$

as in the linear case it is obtained by testing with v and partial integration.

We may rewrite (9.2) as an operator equation

$$Au = F\,. \tag{9.3}$$

Here, $A : V \to V^*$ with $V = H_0^1(\Omega)$,

$$\langle Au, v\rangle = \int_\Omega \langle a(\nabla u(x)), \nabla v(x)\rangle \; dx\,, \tag{9.4}$$

and $F \in V^*$,

$$\langle F, v\rangle = \int_\Omega f(x)v(x)\,dx\,.$$

When the function a is nonlinear, the operator A is nonlinear. Thus, (9.3) represents a nonlinear operator equation.

In this section we consider the situation when a satisfies the condition

$$\langle a(y) - a(z), y - z\rangle \geq 0\,, \quad \text{for all } y, z \in \mathbb{R}^n. \tag{9.5}$$

If a is linear, (9.5) means that the associated matrix $M \in \mathbb{R}^{(n,n)}$ is positive semidefinite. This occurs for example when a is the derivative of a convex quadratic function $J : \mathbb{R}^n \to \mathbb{R}$. Then M equals D^2J, the Hessian of J.

Definition 9.1 (Monotone Operator)
Let V be a normed space. An operator $A : V \to V^*$ is called ***monotone*** *if*

$$\langle Au - Av, u - v\rangle \geq 0\,, \quad \textit{for all } u, v \in V\,, \tag{9.6}$$

and ***strictly monotone****, if this inequality is strict whenever $u \neq v$.*

Definition 9.2 *Let V be a normed space. An operator $A : V \to V^*$ is called* ***coercive*** *if*

$$\lim_{\|v\|\to+\infty} \frac{\langle Av, v\rangle}{\|v\|} = +\infty\,. \tag{9.7}$$

For solving $Au = F$, let us consider first the finite dimensional case.

Proposition 9.3 (Brouwer's fixed point theorem)
Let $K \subset \mathbb{R}^n$ be compact and convex, $f : K \to K$ continuous, $K \neq \emptyset$. Then f has a fixed point $u \in K$, that is, $F(u) = u$.

Proof: This is a fundamental result of analysis, not proved here. □

Corollary 9.4 *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be continuous, let $R > 0$ such that*

$$\langle Av, v\rangle \geq 0\,, \quad \text{for all } v \text{ with } |v| = R. \tag{9.8}$$

Then A has a zero u (that is, $Au = 0$) satisfying $|u| \leq R$.

Proof: By contradiction. Let us assume that no such zero exists. Then

$$f(v) = -R\frac{Av}{|Av|}$$

defines a continous mapping $f : B_R \to B_R$, where B_R denotes the closed ball around 0 with radius R. According to Brouwer's fixed point theorem there exists an $u \in B_R$ with $f(u) = u$. By definition of f we have $|u| = |f(u)| = R$. By virtue of (9.8),

$$0 \leq \langle Au, u\rangle = \langle Au, f(u)\rangle = -R|Au|\,,$$

thus $Au = 0$, a contradiction. □

Proposition 9.5 *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be continuous and coercive. Then the equation $Au = F$ has a solution $u \in \mathbb{R}^n$ for every $F \in \mathbb{R}^n$.*

Proof: Assume first that $F = 0$. We apply Corollary 9.4. Since A is coercive, there exists $R > 0$ such that (9.8) holds. Therefore, $Au = 0$ has a solution $u \in \mathbb{R}^n$. For arbitrary $F \in \mathbb{R}^n$ we consider $A_F : \mathbb{R}^n \to \mathbb{R}^n$, $A_F v = Av - F$. A_F is continuous; it is also coercive: We have

$$\langle A_F v, v\rangle = \langle Av, v\rangle - \langle F, v\rangle \geq \langle Av, v\rangle - \|F\|\|v\|\,,$$

therefore

$$\frac{\langle A_F v, v\rangle}{\|v\|} \geq \frac{\langle Av, v\rangle}{\|v\|} - \|F\|\,,$$

and since A is coercive, it follows that A_F is coercive. □

We now consider the infinite dimensional case. By approximation, we reduce it to the finite dimensional case.

In view of applying the results to partial differential equations, one wants to keep the continuity assumptions on A to a minimum.

Definition 9.6 (Hemicontinuity) *Let V be a normed space. An operator $A : V \to V^*$ is called **hemicontinuous** if, for all $u, v, w \in V$, the mapping*

$$t \mapsto \langle A(u + tv), w\rangle \tag{9.9}$$

is continuous on $[0, 1]$.

Lemma 9.7 *Let V be a Banach space, $A : V \to V^*$. Then A is monotone if and only if for all $u, v \in V$ the mapping*

$$t \mapsto \langle A(u + tv), v\rangle \tag{9.10}$$

is monotone nondecreasing on $[0, 1]$.

Proof: Exercise.

Proposition 9.8 *Let V be a Banach space, $A : V \to V^*$ monotone. Then A is locally bounded, that is, for every $u \in V$ there exists a neighbourhood U of u such that $A(U)$ is a bounded subset of V^*.*

Proof: Let us assume that A is not locally bounded. Then there exists an $u \in V$ and a sequence (u_n) in V such that

$$u_n \to u\,, \quad \|Au_n\| \to \infty\,. \tag{9.11}$$

We define

$$c_n = 1 + \|Au_n\|\,\|u_n - u\|\,. \tag{9.12}$$

We want to show that the sequence $c_n^{-1}Au_n$ is bounded in V^*. To this purpose, let $v \in V$ be arbitrary. We have

$$0 \le \langle A(u + v) - A(u_n), u + v - u_n\rangle\ , \tag{9.13}$$

consequently

$$\begin{aligned}\frac{1}{c_n}\langle A(u_n), v\rangle &\le \frac{1}{c_n}\left(\langle A(u_n), u_n - u\rangle + \langle A(u + v), u + v - u_n\rangle\right)\\ &\le 1 + \frac{1}{c_n}\|A(u + v)\|\,\|u + v - u_n\| \le M(v)\end{aligned} \tag{9.14}$$

with a constant $M(v)$ which does not depend on n. Applying the same argument for $-v$ in place of v, we obtain

$$\sup_{n\in\mathbb{N}}\left|\frac{1}{c_n}\langle Au_n, v\rangle\right| \le \max\{M(v), M(-v)\} < \infty\,. \tag{9.15}$$

The principle of uniform boundedness (the theorem of Banach and Steinhaus from functional analysis) implies that

$$\sup_{n\in\mathbb{N}}\frac{1}{c_n}\|Au_n\| =: C < \infty\,. \tag{9.16}$$

It follows that

$$\|Au_n\| \le Cc_n = C(1 + \|Au_n\|\,\|u - u_n\|)\,, \quad n \in \mathbb{N}\,, \tag{9.17}$$

therefore

$$(1 - C\|u - u_n\|)\|Au_n\| \le C\,, \quad n \in \mathbb{N}\,, \tag{9.18}$$

and so $\|Au_n\| \le 2C$ if $\|u - u_n\| \le 1/2C$. This contradicts (9.11). □

Corollary 9.9 *Let V be a Banach space, $A : V \to V^*$ monotone, let (u_n) be a sequence which converges in the norm of V. Then the sequence (Au_n) is bounded in V^*.*

Proof: This is a direct consequence of Proposition 9.8. $\square$

Corollary 9.10 *Let V be a Banach space, $A : V \to V^*$ monotone, $K \subset V$ bounded, let $C > 0$ such that*

$$\langle Au, u\rangle \le C\,, \quad \textit{for all } u \in K. \tag{9.19}$$

Then $A(K)$ is bounded in V^*.

Proof: For $\varepsilon > 0$ sufficiently small, it follows from Proposition 9.8, setting there $u = 0$,

$$\sup_{\|v\|\le\varepsilon} \|Av\| =: c < \infty\,. \tag{9.20}$$

Therefore, for arbitrary $u \in K$ we then have, since $0 \le \langle Au - Av, u - v\rangle$,

$$\begin{aligned} \|Au\| &= \sup_{\|v\|\le 1} \langle Au, v\rangle = \sup_{\|v\|\le\varepsilon} \frac{1}{\varepsilon} \langle Au, v\rangle \\ &\le \sup_{\|v\|\le\varepsilon} \frac{1}{\varepsilon}\big(\langle Au, u\rangle + \langle Av, v\rangle - \langle Av, u\rangle \big) \\ &\le \frac{1}{\varepsilon}(C + c\varepsilon + cC_K)\,, \quad C_K := \sup_{u\in K} \|u\|\,. \end{aligned} \tag{9.21}$$

$\square$

Proposition 9.11 *Let V be a reflexive Banach space, $A : V \to V^*$ monotone. Then the following are equivalent:*

(i) *A ist hemicontinuous.*

(ii) *For all $u \in V$ and all $b \in V^*$ we have: If*

$$\langle b - Av, u - v\rangle \ge 0\,, \quad \textit{for all } v \in V, \tag{9.22}$$

then $Au = b$.

(iii) *For all $u \in V$ and $b \in V^*$ we have: If (u_n) is a sequence in V satisfying*

$$u_n \rightharpoonup u\,, \quad Au_n \rightharpoonup b\,, \quad \limsup_{n\to\infty} \langle Au_n, u_n\rangle \le \langle b, u\rangle\,, \tag{9.23}$$

then $Au = b$.

(iv) *A is demicontinuous, that is, for all $u \in V$ we have: If (u_n) is a sequence in V with $u_n \to u$, then $Au_n \rightharpoonup Au$.*

The implication "(i)⇒(iii)" (or a variant of this) is called the "Minty trick".

Proof: "(i)⇒(ii)": Let $u \in V$, $b \in V^*$ such that (9.22) holds. Setting $v = u - tw$ it follows that

$$\langle b - A(u - tw), tw\rangle \ge 0\,, \quad \forall\, w \in V\,,\ t > 0\,, \tag{9.24}$$

therefore

$$\langle b - A(u - tw), w \rangle \geq 0 \,, \quad \forall\, w \in V \,,\, t > 0 \,. \tag{9.25}$$

Passing to the limit $t \downarrow 0$ we obtain, since A is hemicontinuous,

$$\langle b - A(u), w \rangle \geq 0 \,, \quad \forall\, w \in V \,. \tag{9.26}$$

Since w was arbitrary, $Au = b$ follows.
"(ii)⇒(iii)": Let (u_n) be a sequence which satisfies (9.23), let $v \in V$. Then

$$0 \leq \langle Au_n - Av, u_n - v \rangle = \langle Au_n, u_n \rangle - \langle Au_n, v \rangle - \langle Av, u_n - v \rangle \,, \tag{9.27}$$

and therefore

$$\begin{aligned} 0 &\leq \limsup_{n\to\infty} (\langle Au_n, u_n \rangle - \langle Au_n, v \rangle - \langle Av, u_n - v \rangle) \\ &= \limsup_{n\to\infty} \langle Au_n, u_n \rangle - \lim_{n\to\infty} \langle Au_n, v \rangle - \lim_{n\to\infty} \langle Av, u_n - v \rangle \\ &\leq \langle b, u \rangle - \langle b, v \rangle - \langle Av, u - v \rangle \\ &= \langle b - Av, u - v \rangle \,. \end{aligned} \tag{9.28}$$

Since v was arbitrary, $b = Au$ follows from assumption (ii).
"(iii)⇒(iv)": Let (u_n) be a sequence in V with $u_n \to u$. Due to Corollary 9.9, (Au_n) is bounded in V^*. Let (u_{n_k}) be a subsequence such that (Au_{n_k}) converges weakly in V^*; such a subsequence exists since bounded subsets of reflexive spaces have this property. Let $Au_{n_k} \rightharpoonup b$. Since (u_{n_k}) converges in the norm of V, it follows that

$$\lim_{k\to\infty} \langle Au_{n_k}, u_{n_k} \rangle = \langle b, u \rangle \,, \tag{9.29}$$

thus $b = Au$ by assumption (iii). Consequently, the limit of every such weakly convergent subsequence (Au_{n_k}) is the same, namely Au. By the convergence principle, $Au_n \rightharpoonup Au$ holds for the whole sequence.
"(iv)⇒(i)": Let $u, v, w \in V$, $t_n \to t$ in $[0, 1]$. Then $u + t_n v \to u + tv$ in V, thus $A(u + t_n v) \rightharpoonup A(u + tv)$ in V^* by assumption (iv). Therefore,

$$\lim_{n\to\infty} \langle A(u + t_n v), w \rangle = \langle A(u + tv), w \rangle \,. \tag{9.30}$$

Thus A is hemicontinuous. □

We now present a frame for finite-dimensional approximations of equations in infinite-dimensional spaces.

A normed space X is called **separable** if there exists a countable subset M of X which is dense in X, that is, $\overline{M} = X$.

Proposition 9.12 *Let V be a separable normed space with* $\dim(V) = \infty$. *Then there exists a sequence $(w_n)_{n\in\mathbb{N}}$ in V such that*

$$\dim(V_n) = n \,, \quad V_n := \operatorname{span}\{w_1, \dots, w_n\} \,, \tag{9.31}$$

and

$$V = \overline{\bigcup_{k=1}^{\infty} V_k} \,. \tag{9.32}$$

Proof: Let $M = \{u_n : n \in \mathbb{N}\}$ be a countable subset of V with $\overline{M} = V$. We obtain the sequence (w_n) by removing all those u_n for which $u_n \in \operatorname{span}\{u_1, \ldots, u_{n-1}\}$. The remaining elements are linearly independent; there are infinitely many of them, since otherwise $\operatorname{span}(M)$ would be finite dimensional and hence closed, contradicting the fact that $\overline{M} = V$. □

Let V be a separable Banach space. We consider the equation

$$Au = F\,, \tag{9.33}$$

in the dual V^* of V, where $A : V \to V^*$ and $F \in V^*$. An element $u \in V$ solves (9.33) if and only if it is a solution of the associated variational equation

$$\langle Au, v\rangle = \langle F, v\rangle\ , \quad \text{for all } v \in V. \tag{9.34}$$

We consider a finite dimensional approximation of (9.34), namely

$$\langle Au_n, v\rangle = \langle F, v\rangle\ , \quad \text{for all } v \in V_n. \tag{9.35}$$

Here, V_n is a subspace of V with $\dim(V_n) = n$, and we look for solutions $u_n \in V_n$.

If $\{w_1, \ldots, w_n\}$ is a basis of V_n, then $u_n \in V_n$ solves (9.35) if and only if u_n solves

$$\langle Au_n, w_k\rangle = \langle F, w_k\rangle\ , \quad 1 \le k \le n\,. \tag{9.36}$$

This is a finite-dimensional problem: We want to determine n unknowns, namely coefficients of u_n with respect to the basis $\{w_1, \ldots, w_n\}$, from the n equations in (9.36).

The equations (9.36) are called **Galerkin equations**. The method which consists in choosing these u_n as approximations for the solution u of (9.33) is called the **Galerkin method**. The following theorem is also called the main theorem for monotone operators.

Theorem 9.13 (Browder and Minty)
Let V be a reflexive separable Banach space, let $A : V \to V^*$ be monotone, hemicontinuous and coercive. Then the equation

$$Au = F \tag{9.37}$$

has a solution $u \in V$ for every $F \in V^*$.

Proof: It suffices to consider the case $F = 0$ zu betrachten. The general case is reduced to this replacing A by A_F with $A_F u = Au - F$. Then A_F, too, is monotone, hemicontinuous and coercive (to see that A_F is coercive, one argues as in the proof of Proposition 9.5).

According to Proposition 9.12 we choose a sequence (w_n) in V satisfying

$$\dim(V_n) = n\,, \quad V_n := \operatorname{span}\{w_1, \ldots, w_n\}\,, \quad V = \overline{\bigcup_{k=1}^{\infty} V_k}\,. \tag{9.38}$$

At first, we prove that, for every $n \in \mathbb{N}$, the Galerkin equations

$$\langle Au_n, w_k\rangle = 0\,, \quad 1 \le k \le n\,, \tag{9.39}$$

have a solution $u_n \in V_n$. Let $j_n : \mathbb{R}^n \to V_n$ the linear isomorphism defined by

$$j_n(x) = \sum_{k=1}^{n} x_k w_k \,, \tag{9.40}$$

let $A_n : \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$(A_n x)_k = \langle A j_n(x), w_k \rangle \,, \quad 1 \le k \le n \,. \tag{9.41}$$

Since A is demicontinuous by Proposition 9.11(iv), all $(A_n)_k$ and thus all A_n are continuous. For $x \in \mathbb{R}^n$ we have

$$\langle A_n x, x \rangle = \sum_{k=1}^{n} (A_n x)_k x_k = \sum_{k=1}^{n} \langle A j_n(x), w_k \rangle \, x_k = \langle A j_n(x), j_n(x) \rangle \,.$$

Since $\|j_n(x)\| \to \infty$ if $|x| \to \infty$, A_n is coercive because A is coercive. By Proposition 9.5, the equation $A_n x = 0$ has a solution $x \in \mathbb{R}^n$, so $u_n = j(x) \in V_n$ is a solution of the Galerkin equations (9.39). Equivalently,

$$\langle A u_n, v \rangle = 0 \,, \quad \text{for all } v \in V_n. \tag{9.42}$$

Since $u_n \in V_n$ we have

$$\langle A u_n, u_n \rangle = 0 \,. \tag{9.43}$$

As A is coercive, the sequence (u_n) is bounded in V. Corollary 9.10 now implies that $(A u_n)$ is bounded in V^*. From (9.42) we obtain

$$\lim_{n \to \infty} \langle A u_n, v \rangle = 0 \,, \quad \text{for all } v \in \bigcup_{k=1}^{\infty} V_k \,. \tag{9.44}$$

Because $(A u_n)$ is bounded, it follows from (9.44) that $A u_n \rightharpoonup 0$ in V^* (exercise). Let now (u_{n_k}) be a weakly convergent subsequence of (u_n) with $u_{n_k} \rightharpoonup u \in V$. We apply Proposition 9.11(iii), setting $b = 0$, to the subsequence $u_{n_k})$ and obtain $Au = 0$. $\square$

One can prove that the set of solutions of $Au = F$ is a convex, closed and bounded subset of V. Since V is assumed to be reflexive, one concludes that the set of solutions is weakly compact.

Proposition 9.14 *Let the assumptions of Theorem 9.13 be satisfied, let moreover A be strictly monotone. Then the solution $u \in V$ of $Au = F$ is uniquely determined for every given $F \in V^*$, so $A : V \to V^*$ is bijective. The inverse operator $A^{-1} : V^* \to V$ is strictly monotone and hemicontinuous, and maps bounded sets to bounded sets.*

Proof: If $Au = F = Av$, then $\langle Au - Av, u - v \rangle = 0$, and therefore $u = v$ since A is strictly monotone. Let $F_1, F_2 \in V^*$, set $u_i = A^{-1} F_i$. Then $u_1 \neq u_2$ if and only if $F_1 \neq F_2$, and in this case we have

$$\left\langle F_1 - F_2, A^{-1} F_1 - A^{-1} F_2 \right\rangle = \langle A u_1 - A u_2, u_1 - u_2 \rangle > 0$$

since A is strictly monotone. To prove that A^{-1} is hemicontinous, according to Proposition 9.11 it suffices to show that, for every $F \in V^*$ and every $u \in V$, the property

$$\left\langle u - A^{-1}G, F - G\right\rangle \geq 0\,, \quad \text{for all } G \in V^*, \tag{9.45}$$

implies $A^{-1}F = u$. Thus, let (9.45) be satisfied, let $G \in V^*$. Setting $v = A^{-1}G$ we have

$$\langle u - v, F - Av\rangle \geq 0\,. \tag{9.46}$$

Since A is hemicontinuous and $A^{-1}(V^*) = V$, we conclude that $F = Au$ by applying Proposition 9.11 to A. It remains to show that A^{-1} is bounded. Let $F \in V^*$ with

$$\|F\| \leq C\,, \quad u = A^{-1}F\,. \tag{9.47}$$

It follows that

$$\frac{\langle Au, u\rangle}{\|u\|} = \frac{\langle F, u\rangle}{\|u\|} \leq \|F\| \leq C\,. \tag{9.48}$$

If $\{u : \|Au\| \leq C\}$ were unbounded, there would exist a sequence (u_n) in V with $\|u_n\| \to \infty$ and

$$\frac{\langle Au_n, u_n\rangle}{\|u_n\|} \leq C$$

This contradicts the assumption that A is coercive. □

We now investigate the convergence of the Galerkin method. Let V_n, as in the proof of Theorem 9.13, be an increasing sequence of subspaces of V with

$$V_n = \operatorname{span}\{w_1, \dots, w_n\}\,, \quad \dim(V_n) = n\,, \quad \overline{\bigcup_{k=1}^{\infty} V_k} = V\,. \tag{9.49}$$

Once again we consider the problem: Given $F \in V^*$, determine $u_n \in V_n$ such that

$$\langle Au_n, v\rangle = \langle F, v\rangle\;, \quad \text{for all } v \in V_n. \tag{9.50}$$

Proposition 9.15 *Let the assumptions of Theorem 9.13 hold, let moreover A be strictly monotone, let $F \in V^*$. Then for every $n \in \mathbb{N}$ there exists a unique solution u_n of (9.50). Moreover, $u_n \rightharpoonup u$ in V, where u denotes the unique solution of $Au = F$. If in addition A is uniformly monotone, that is, there exists a constant c with*

$$\langle Av - Aw, v - w\rangle \geq c\|v - w\|^2\,, \quad \textit{for all } v, w \in V, \tag{9.51}$$

then $u_n \to u$ in V.

Proof: The existence of u_n is a consequence of Theorem 9.13. If $u_n, v_n \in V_n$ are two solutions of (9.50), then

$$\langle Au_n, v\rangle = \langle F, v\rangle = \langle Av_n, v\rangle\;, \quad \text{f'ur alle } v \in V_n, \tag{9.52}$$

and since $v = u_n - v_n$ it follows that $\langle Au_n - Av_n, u_n - v_n\rangle = 0$, thus $u_n = v_n$. In the proof of Theorem 9.13 we also have shown that every limit $\tilde{u}$ of a weakly convergent subsequence (u_{n_k}) of (u_n) solves the orginal problem, that is, $A\tilde{u} = F$. The uniqueness result 9.14 implies that $\tilde{u} = u$. From the convergence principle we obtain that $u_n \rightharpoonup u$. Let now (9.51) be satisfied. Since $u_n \rightharpoonup u$, the set (u_n) is bounded in V. Moreover,

$$\langle Au_n, u_n \rangle = \langle F, u_n \rangle \leq \|F\| \sup_{n \in \mathbb{N}} \|u_n\| . \tag{9.53}$$

Therefore, Corollary 9.10 implies that (Au_n) is bounded in V^*. We have

$$\langle Au_n - Au, v \rangle = 0 , \quad \text{for all } v \in V_n. \tag{9.54}$$

We choose a sequence $v_n \in V_n$ with $v_n \to u$ in V. Then

$$\begin{aligned} 0 &\leq c\|u_n - u\|^2 \leq \langle Au_n - Au, u_n - u \rangle = \langle Au_n - Au, v_n - u \rangle \\ &\leq (\|Au_n\| + \|Au\|) \|v_n - u\| \leq C\|v_n - u\| \end{aligned} \tag{9.55}$$

for some constant C which does not depend on n. Therefore, $v_n \to u$ implies that $u_n \to u$. □

Remark: The uniform monotonicity (9.51) already implies coercivity.

We consider the boundary value problem

$$\begin{aligned} -\mathrm{div}\,(a(x, \nabla u)) + a_0(x, u) = f , &\quad \text{in } \Omega, \\ u(x) = 0 , &\quad \text{on } \partial\Omega, \end{aligned} \tag{9.56}$$

which constitutes a slight generalization of Problem (9.1). An example is given by the p-Laplacian operator which results when we set $a(x, \xi) = |\xi|^{p-2}\xi$ with $p \geq 2$. ($p = 2$ yields the Laplace operator.) The variational formulation becomes

$$\int_\Omega \langle a(x, \nabla u(x)), \nabla v(x) \rangle \, dx + \int_\Omega a_0(x, u(x))v(x)\, dx = \int_\Omega f(x)v(x)\, dx , \quad \text{for all } v \in V. \tag{9.57}$$

The space V still has to be fixed. We assume that a satisfies the growth condition

$$|a(x, \xi)| \leq c_0(1 + |\xi|^{p-1}) \tag{9.58}$$

for some $p \in (1, \infty)$ and the Carathéodory condition

$$\begin{aligned} x \mapsto a(x, \xi) &\quad \text{is measurable for all } \xi \in \mathbb{R}^n \\ \xi \mapsto a(x, \xi) &\quad \text{is continuous for almost all } x \in \Omega. \end{aligned} \tag{9.59}$$

We assume that the same is true for a_0 in place of a. The Carathéodory condition implies that the function $x \mapsto a(x, \nabla u(x))$ is measurable if ∇u is measurable. The growth condition (9.58) implies that

$$\begin{aligned} \int_\Omega \langle a(x, \nabla u(x)), \nabla v(x) \rangle \, dx &\leq c_0 \int_\Omega (1 + |\nabla u(x)|^{p-1})|\nabla v(x)|\, dx \\ &\leq c_0(|\Omega|^{1/q} + \|\, |\nabla u|^{p-1}\|_q)\|\nabla v\|_p \end{aligned} \tag{9.60}$$

holds with $1/p + 1/q = 1$, due to H"older's inequality. Because $q(p-1) = p$ we have

$$\|\, |\nabla u|^{p-1}\|_q = (\|\nabla u\|_p)^{q/p} .$$

We conclude that the reflexive and separable Banach space

$$V = W_0^{1,p}(\Omega)$$

is the appropriate space for the variational formulation (9.57). Indeed,

$$\langle Au, v\rangle = \int_\Omega \langle a(x, \nabla u(x)), \nabla v(x)\rangle \, dx + \int_\Omega a_0(x, u(x))v(x)\, dx \tag{9.61}$$

defines an operator $A : V \to V^*$, according to the considerations above (and analogous ones for a_0). This operator is continuous (exercise). It is moreover monotone if we require that

$$\langle a(x, w) - a(x, z), w - z\rangle \geq 0\,, \quad \text{for all } w, z \in \mathbb{R}^n,\ x \in \Omega, \tag{9.62}$$

and the same for a_0 instead of a. If in addition

$$\langle a(x, w) - a(x, z), w - z\rangle \geq c|w - z|^2\,, \quad \text{for all } w, z \in \mathbb{R}^n,\ x \in \Omega \tag{9.63}$$

for some constant $c > 0$, then A is uniformly monotone in the sense of (9.51). (For a_0 we do not require (9.63) to hold, this already follows from the Poincaré inequality in V.) From Proposition 9.15 it now follows:

Proposition 9.16 *Let (9.58), (9.59) and (9.62) be satisfied. Then the quasilinear boundary value problem (9.56) has a unique weak solution in $V = W_0^{1,p}(\Omega)$ for every right hand side $F \in V^*$, and the Galerkin approximations u_n converge to u in the norm of V.* $\square$

10 The Bochner Integral

In this chapter, $[a,b]$ always denotes a compact interval in $\mathbb{R}$. For $A \subset [a,b]$ we denote its characteristic function by χ_A,

$$\chi_A(t) = \begin{cases} 1\,, & t \in A\,, \\ 0\,, & t \notin A\,. \end{cases} \tag{10.1}$$

Definition 10.1 (Simple function) *Let V be a Banach space. A function $u : [a,b] \to V$ is called* ***simple*** *if it has the form*

$$u(t) = \sum_{i=1}^{n} \chi_{A_i}(t) v_i\,, \tag{10.2}$$

where $n \in \mathbb{N}$, $A_i \subset [a,b]$ measurable and $v_i \in V$ for $1 \le i \le n$.

Lemma 10.2 *Let V be a Banach space, $u : [a,b] \to V$ simple. Then there exists a unique representation of u in the form (10.2) satisfying*

$$\bigcup_i A_i = [a,b]\,, \quad A_i \cap A_j = \emptyset \text{ and } v_i \neq v_j \text{ for } i \neq j\,. \tag{10.3}$$

It is called the ***canonical representation*** *of u.*

Proof: Exercise. □

Definition 10.3 (Bochner measurability)
Let V be a Banach space. A function $u : [a,b] \to V$ is called ***Bochner measurable*** *if there exists a sequence of simple functions $u_n : [a,b] \to V$ such that*

$$\lim_{n\to\infty} u_n(t) = u(t) \tag{10.4}$$

for almost all $t \in [a,b]$.

Definition 10.4 *Let V be a Banach space, $u : [a,b] \to V$ a simple function*

$$u(t) = \sum_{i=1}^{n} \chi_{A_i}(t) v_i\,. \tag{10.5}$$

The ***Bochner integral*** *of u is defined as*

$$\int_a^b u(t)\,dt = \sum_{i=1}^{n} \operatorname{meas}(A_i) v_i\,. \tag{10.6}$$

For measurable $A \subset [a,b]$ we define

$$\int_A u(t)\,dt = \int_a^b \chi_A(t) u(t)\,dt\,. \tag{10.7}$$

Definition 10.4 makes sense since the value of the right side of (10.6) does not depend on which representation of u we choose.

As a direct consequence of the definition we obtain that for simple functions $u, v : [a, b] \to V$ and numbers $\alpha, \beta \in \mathbb{R}$

$$\int_a^b \alpha u(t) + \beta v(t)\, dt = \alpha \int_a^b u(t)\, dt + \beta \int_a^b v(t)\, dt\,, \tag{10.8}$$

as well as

$$\left\| \int_a^b u(t)\, dt \right\| \le \int_a^b \|u(t)\|\, dt\,. \tag{10.9}$$

Lemma 10.5 *Let V be a Banach space, $u_n : [a, b] \to V$ a sequence of simple functions satisfying $u_n \to u$ almost everywhere. Then for every $n \in \mathbb{N}$ the function $f : [a, b] \to \mathbb{R}$ defined by*

$$f(t) = \|u_n(t) - u(t)\| \tag{10.10}$$

is measurable.

Proof: We have

$$f(t) = \lim_{m\to\infty} f_m(t)\,, \quad f_m(t) := \|u_n(t) - u_m(t)\|\,, \tag{10.11}$$

and f_m is a simple function for all $m \in \mathbb{N}$. □

Let now $u_n : [a, b] \to V$ be a sequence of simple function with $u_n \to u$ pointwise a.e., satisfying

$$\lim_{n\to\infty} \int_a^b \|u_n(t) - u(t)\|\, dt = 0\,. \tag{10.12}$$

(By Lemma 10.5 the integrand is measurable.) Due to

$$\begin{aligned} \left\| \int_a^b u_n(t)\, dt - \int_a^b u_m(t)\, dt \right\| &\le \int_a^b \|u_n(t) - u_m(t)\|\, dt \\ &\le \int_a^b \|u_n(t) - u(t)\|\, dt + \int_a^b \|u_m(t) - u(t)\|\, dt\,, \end{aligned} \tag{10.13}$$

setting

$$y_n = \int_a^b u_n(t)\, dt \tag{10.14}$$

we obtain a Cauchy sequence $\{y_n\}$ in V. If $v_n : [a, b] \to V$ defines another sequence with the same properties as $\{u_n\}$,

$$\begin{aligned} \left\| \int_a^b v_n(t)\, dt - \int_a^b u_n(t)\, dt \right\| &\le \int_a^b \|v_n(t) - u_n(t)\|\, dt \\ &\le \int_a^b \|v_n(t) - u(t)\|\, dt + \int_a^b \|u_n(t) - u(t)\|\, dt\,. \end{aligned} \tag{10.15}$$

Therefore, the limit of $\{y_n\}$ does not depend on the choice of the sequence $\{u_n\}$.

Definition 10.6 (Bochner integral) *Let $u : [a,b] \to V$. If there exists a sequence of simple functions $u_n : [a,b] \to V$ such that $u_n \to u$ pointwise a.e. and*

$$\lim_{n\to\infty} \int_a^b \|u_n(t) - u(t)\| \, dt = 0 \,, \tag{10.16}$$

*then u is called **Bochner integrable**, and the **Bochner integral** of u is defined by*

$$\int_a^b u(t)\, dt = \lim_{n\to\infty} \int_a^b u_n(t)\, dt \,. \tag{10.17}$$

Lemma 10.7 *Let V be a Banach space, let $u, v : [a,b] \to V$ Bochner integrable and $\alpha, \beta \in \mathbb{R}$. Then $\alpha u + \beta v$ is Bochner integrable, and*

$$\int_a^b \alpha u(t) + \beta v(t)\, dt = \alpha \int_a^b u(t)\, dt + \beta \int_a^b v(t)\, dt \,. \tag{10.18}$$

Proof: This follows directly from the definitions. □

Proposition 10.8 *Let V be a Banach space. A function $u : [a,b] \to V$ is Bochner integrable if and only if u is Bochner measurable and the function $t \mapsto \|u(t)\|$ is integrable. In this case,*

$$\left\| \int_a^b u(t)\, dt \right\| \le \int_a^b \|u(t)\|\, dt \,. \tag{10.19}$$

Proof: "$\Rightarrow$": Let (u_n) be a sequence of simple functions with $u_n \to u$ pointwise a.e. and

$$\int_a^b \|u(t) - u_n(t)\|\, dt = 0 \,. \tag{10.20}$$

Since $\|u_n(t)\| \to \|u(t)\|$ for a.e. $t \in [a,b]$, the function $t \mapsto \|u(t)\|$ is measurable. Then

$$\int_a^b \|u(t)\|\, dt \le \int_a^b \|u(t) - u_n(t)\|\, dt + \int_a^b \|u_n(t)\|\, dt < \infty \,. \tag{10.21}$$

"$\Leftarrow$": Let (u_n) be a sequence of simple functions with $u_n \to u$ pointwise almost everywhere. For any given $\varepsilon > 0$ we define $v_n : [a,b] \to V$ by

$$v_n(t) = \begin{cases} u_n(t)\,, & \text{if } \|u_n(t)\| \le (1+\varepsilon)\|u(t)\|\,, \\ 0\,, & \text{otherwise.} \end{cases} \tag{10.22}$$

v_n is a simple function, since $\{t : \|u_n(t)\| \le (1+\varepsilon)\|u(t)\|\}$ is measurable. For

$$f_n(t) = \|v_n(t) - u(t)\| \tag{10.23}$$

we have $f_n \to 0$ pointwise a.e. and

$$0 \le f_n(t) \le (2+\varepsilon)\|u(t)\| \,. \tag{10.24}$$

From Lebesgue's theorem on dominated convergence we obtain

$$\lim_{n\to\infty} \int_a^b \|v_n(t) - u(t)\| \, dt = \lim_{n\to\infty} \int_a^b f_n(t) \, dt = 0 \,. \tag{10.25}$$

Therefore, u is Bochner integrable. With (10.9) it follows that

$$\left\| \int_a^b v_n(t) \, dt \right\| \le \int_a^b \|v_n(t)\| \, dt \le (1+\varepsilon) \int_a^b \|u(t)\| \, dt \tag{10.26}$$

and thus

$$\left\| \int_a^b u(t) \, dt \right\| = \left\| \lim_{n\to\infty} \int_a^b v_n(t) \, dt \right\| \le (1+\varepsilon) \int_a^b \|u(t)\| \, dt \,. \tag{10.27}$$

Since $\varepsilon > 0$ was arbitrary, (10.19) follows. □

We now consider functions $u : [a,b] \to V$ for which

$$\int_a^b \|u(t)\|^p \, dt < \infty \tag{10.28}$$

holds.

Definition 10.9 *Let V be a Banach space, $1 \le p < \infty$. We define*

$$L^p(a,b;V) = \{[u] \mid u : [a,b] \to V \text{ is Bochner measurable and satisfies (10.28) } \} \,. \tag{10.29}$$

Here, $[u]$ denotes the equivalence class of u with respect to the equivalence relation

$$u \sim v \quad \Leftrightarrow \quad u = v \text{ almost everywhere } . \tag{10.30}$$

Due to Proposition 10.8, $L^1(a,b;V)$ coincides with the vector space of all Bochner integrable functions on $[a,b]$.

Proposition 10.10 *Let V be a Banach space, $1 \le p < \infty$. The space $L^p(a,b;V)$ is a Banach space when equipped with the norm*

$$\|u\|_{L^p(a,b;V)} = \left(\int_a^b \|u(t)\|_V^p \, dt \right)^{\frac{1}{p}} \,. \tag{10.31}$$

If V is a Hilbert space, then $L^2(a,b;V)$ becomes a Hilbert space when equipped with the scalar product

$$\langle u, v \rangle = \int_a^b \langle u(t), v(t) \rangle_V \, dt \,. \tag{10.32}$$

Proof: Omitted. For a given Cauchy sequence, one constructs a limit in the same way as in the scalar case $V = \mathbb{R}$. In order to prove that this limit is Bochner measurable, one uses a characterization of measurability due to Pettis. □

Definition 10.11 *For $u : [a,b] \to \mathbb{R}$ we define*

$$\operatorname*{ess\,sup}_{t\in[a,b]} u(t) = \inf\{M : M \in \mathbb{R},\ u(t) \le M \text{ for almost all } t \in [a,b]\} \,. \tag{10.33}$$

We now consider functions $u : [a, b] \to V$ with values in a Banach space V for which

$$\operatorname*{ess\,sup}_{t\in[a,b]} \|u(t)\|_V < \infty\,. \tag{10.34}$$

Definition 10.12 *Let V be a Banach space. We define*

$$L^\infty(a,b;V) = \{[u] \,|\, u : [a,b] \to V \textit{ is Bochner measurable and (10.34) holds}\}\,. \tag{10.35}$$

Proposition 10.13 *Let V be a Banach space. Then $L^\infty(a,b;V)$ is a Banach space.*

Proof: Again, this is proved in the same manner as in the case $V = \mathbb{R}$. □

Lemma 10.14 *Let V be a Banach space. Then for all $1 \le p \le q \le \infty$ we have*

$$L^q(a,b;V) \subset L^p(a,b;V)\,. \tag{10.36}$$

Proof: As in the case $V = \mathbb{R}$. □

Definition 10.15 *Let V be a Banach space. We define*

$$C([a,b];V) = \{u \,|\, u : [a,b] \to V \textit{ continuous}\}\,. \tag{10.37}$$

Definition 10.16 (Oszillation) *Let V be a Banach space, $u : [a,b] \to V$. We define the oscillation of u by*

$$\operatorname*{osc}_{[a,b]}(u;\delta) = \sup\{\|u(t) - u(s)\| : s,t \in [a,b],\ |t-s| \le \delta\}\,. \tag{10.38}$$

Lemma 10.17 *Let V be a Banach space, $u : [a,b] \to V$ continuous. Then*

$$\lim_{\delta\to 0} \operatorname*{osc}_{[a,b]}(u;\delta) = 0\,. \tag{10.39}$$

Proof: The statement (10.39) is equivalent to the uniform continuity of u. □

For a continuous function $u : [a,b] \to V$ we have

$$\operatorname*{ess\,sup}_{t\in[a,b]} \|u(t)\| = \max_{t\in[a,b]} \|u(t)\|\,, \tag{10.40}$$

since continuity of u implies that $\|u(t)\| \le \operatorname{ess\,sup}_{s\in[a,b]} \|u(s)\|$ holds for all t.

Proposition 10.18 *Let V be a Banach space. Then $C([a,b];V)$ is a Banach space when equipped with the norm*

$$\|u\|_{C([a,b];V)} = \max_{t\in[a,b]} \|u(t)\|\,. \tag{10.41}$$

Moreover, $C([a,b];V)$ can be identified with a closed subspace of $L^\infty(a,b;V)$.

Proof: If $u : [a,b] \to V$ is continuous, it is Bochner measurable: We define a sequence of simple functions $u_n : [a,b] \to V$ by $u_n(t) = u(ih)$, if $t \in [ih, (i+1)h)$, $h = (b-a)/n$ is. Then

$$\|u_n(t) - u(t)\| \le \operatorname*{osc}_{[a,b]}(u; h)\,, \quad h = \frac{b-a}{n}\,, \tag{10.42}$$

therefore $u_n \to u$ uniformly (pointwise convergence would already be sufficient for our purpose). Moreover: Let (u_n) be a sequence in C satisfying $[u_n] \to [u]$ in L^∞. We choose a subset N in $[a,b]$ of zero measure such that $u_n \to u$ uniformly in $M = [a,b] \setminus N$. Then u is continuous on M. For an arbitrary given $t \in N$ we choose a sequence $(t_k)_{k\in\mathbb{N}}$ in M such that $t_k \to t$. Then

$$\|u_n(t) - u_m(t)\| \le \|u_n(t) - u_n(t_k)\| + \|u_n - u_m\|_{L^\infty(a,b;V)} + \|u_m(t_k) - u_m(t)\|\,,$$

thus $(u_n(t))_{n\in\mathbb{N}}$ is a Cauchy sequence. For $t \in N$, let $\tilde{u}(t)$ be the limit of this Cauchy sequence, and set $\tilde{u}(t) = u(t)$ for $t \in M$. Then $\tilde{u} : [a,b] \to V$ is continuous and $[\tilde{u}] = [u]$.
□

11 Linear parabolic equations

A parabolic equation has the form

$$\partial_t u + Lu = f\,, \tag{11.1}$$

where L is an elliptic operator. Let L be given in divergence form as

$$Lu = -\sum_{i,j=1}^{n} \partial_j(a_{ij}(x,t)\partial_i u) + \sum_{i=1}^{n} b_i(x,t)\partial_i u + c(x,t)u\,. \tag{11.2}$$

We consider the corresponding initial-boundary value problem

$$\partial_t u + Lu = f \quad \text{in } \Omega_T = \Omega \times (0,T], \tag{11.3}$$

$$u = 0 \quad \text{on } \partial\Omega \times (0,T), \tag{11.4}$$

$$u(x,0) = u_0(x) \quad \text{for } x \in \Omega. \tag{11.5}$$

We pass to a variational formulation with respect to the space variable x.

Let $v \in C_0^\infty(\Omega)$ be a test function. We multiply both sides of (11.3) with v, integrate over Ω and perform partial integration of the divergence term. This yields

$$\begin{aligned}
&\int_\Omega \partial_t u(x,t)v(x)\,dx + \int_\Omega \sum_{i,j=1}^{n} a_{ij}(x,t)\partial_i u(x,t)\partial_j v(x)\,dx \\
&\qquad + \int_\Omega \sum_{i=1}^{n} b_i(x,t)\partial_i u(x,t)v(x)\,dx + \int_\Omega c(x,t)u(x,t)v(x)\,dx \\
&= \int_\Omega f(x,t)v(x)\,dx\,.
\end{aligned} \tag{11.6}$$

Assumption 11.1 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let a_{ij}, b_i, c be measurable and bounded for all i,j, let $u_0 \in L^2(\Omega)$, $f \in L^2(\Omega_T)$. We assume that $\partial_t + L$ is* ***uniformly parabolic****, that is, there exists $a_* > 0$ such that*

$$\xi^T A(x,t)\xi = \sum_{i,j=1}^{n} \xi_i a_{ij}(x,t)\xi_j \geq a_*|\xi|^2\,, \quad \textit{for all } \xi \in \mathbb{R}^n,\ x \in \Omega,\ t \in (0,T). \tag{11.7}$$

□

"Uniformly parabolic" thus means that L is uniformly elliptic, and the ellipticity constant does not depend on t.

We want to interpret the unknown function u as a function $u : [0,T] \to V$, where V is a suitable Banach space of functions on Ω,

$$u(t) : \Omega \to \mathbb{R}\,, \quad \text{for all } t \in [0,T]. \tag{11.8}$$

The value $(u(t))(x)$ corresponds to $u(x,t)$ in (11.3).

We write the initial-boundary value problem as

$$\frac{\mathrm{d}}{\mathrm{dt}}(u(t),v)_H + a(u(t),v;t) = \langle F(t),v\rangle_V\,, \quad \text{for all } v\in V, \tag{11.9}$$

$$u(0) = u_0\,. \tag{11.10}$$

Here,

$$H = L^2(\Omega)\,, \quad (w,v)_H = \int_\Omega w(x)v(x)\,dx\,, \tag{11.11}$$

$$V = H_0^1(\Omega)\,, \quad F:[0,T]\to V^*\,, \tag{11.12}$$

$\langle F(t),v\rangle_V$ stands for $(F(t))(v)$, and

$$a(w,v;t) = \int_\Omega (\nabla w(x))^T A(x,t)\nabla v(x)\,dx + \int_\Omega b(x,t)^T\nabla w(x)v(x)\,dx \tag{11.13}$$

$$+\int_\Omega c(x,t)w(x)v(x)\,dx\,. \tag{11.14}$$

Since in (11.9)

$$v\mapsto a(u(t),v;t)$$

defines an element of V^*, one has to expect the same for

$$v\mapsto \frac{\mathrm{d}}{\mathrm{dt}}(u(t),v)_H\,.$$

Thus, when considering (11.9), the three spaces V, H, and V^* are involved. It turns out that the notion of an **evolution triple** provides a suitable abstract formulation for this situation.

For $v^*\in V^*$ and $v\in V$, as above we use the notation

$$\langle v^*,v\rangle_V := v^*(v)\,. \tag{11.15}$$

Proposition 11.2 *Let H be a Hilbert space, V a reflexive Banach space over $\mathbb{R}$, let $j:V\to H$ be linear, continuous and injective, let $j(V)$ be dense in H. Then*

$$\langle j^*(h),v\rangle_V = (h,j(v))_H \tag{11.16}$$

defines a linear, continuous and injective mapping $j^*:H\to V^*$ satifying $\|j^*\|\le\|j\|$, and $j^*(H)$ is dense in V^*.

Proof: Let $h\in H$. The right hand side of (11.16) defines, since

$$|(h,j(v))_H| \le \|h\|_H\,\|j(v)\|_H \le \|h\|_H\,\|j\|\,\|v\|_V\,, \tag{11.17}$$

an element $j^*(h)\in V^*$ with

$$\|j^*(h)\|_{V^*}\le \|j\|\,\|h\|_H\,. \tag{11.18}$$

j^* is linear and, because of (11.18), continuous, and $\|j^*\|\le\|j\|$. Let $j^*(h)=0$. Then

$$(h,j(v))_H = 0\,, \quad \text{for all } v\in V\,. \tag{11.19}$$

Let (v_n) be a sequence in V with $j(v_n) \to h$ in H. Then

$$(h,h)_H = (h, \lim_{n\to\infty} j(v_n))_H = \lim_{n\to\infty} (h, j(v_n))_H = 0\,, \tag{11.20}$$

thus $h = 0$. This shows that j^* is injective. It remains to show that $j^*(H)$ is dense in V^*. Let $v^{**} \in V^{**}$ be arbitrary with $v^{**}(j^*(H)) = 0$. It suffices to show that we then must have $v^{**} = 0$. (If $W = \mathrm{cl}\,(j^*(H))$ were a proper subset of V^*, according to the theorem of Hahn-Banach there would exist a $v^{**} \in V^{**}$ with $v^{**}(W) = 0$, but $v^{**} \neq 0$). As V is reflexive, we find a $v \in V$ with

$$\langle v^*, v\rangle_V = v^{**}(v^*)\,, \qquad \text{for all } v^* \in V^*\,. \tag{11.21}$$

We then have

$$0 = v^{**}(j^*(h)) = \langle j^*(h), v\rangle_V = (h, j(v))_H\,, \quad \text{for all } h \in H, \tag{11.22}$$

therefore $j(v) = 0$ and, since j is injective, $v = 0$. Thus, $v^{**} = 0$. □

Corollary 11.3 *In the situation of Proposition 11.2 we moreover have: the mapping $J : V \to V^*$, $J = j^* \circ j$, is linear, continuous and injective, $J(V)$ is dense in V^*, and*

$$\langle Jv, w\rangle_V = \langle Jw, v\rangle_V\,, \qquad \textit{for all } v, w \in V\,. \tag{11.23}$$

Proof: This follows directly from Proposition 11.2 and from the identity

$$\langle Jv, w\rangle_V = (j(v), j(w))_H = (j(w), j(v))_H = \langle Jw, v\rangle_V\,. \tag{11.24}$$

□

We thus have

$$V \xrightarrow{j} H \xrightarrow{j^*} V^* \tag{11.25}$$

with continuous and dense embeddings.

Definition 11.4 *Under the assumptions of Proposition 11.2, (11.25) is called an* ***evolution triple*** *or* ***Gelfand triple****.*

For the parabolic problem (11.3) – (11.5) we set

$$V = H_0^1(\Omega)\,, \quad H = L^2(\Omega)\,, \tag{11.26}$$

and let

$$j : H_0^1(\Omega) \to L^2(\Omega) \tag{11.27}$$

be the canonical embedding defined by $j(v)(x) = v(x)$. The mapping j^* can be interpreted as follows. For $h \in H = L^2(\Omega)$,

$$\langle j^*(h), v\rangle_V = (h, j(v))_H = \int_\Omega h(x)v(x)\,dx\,. \tag{11.28}$$

To the element $j^*(h) \in V^*$ we associate an element $w \in V$ via the Riesz representation theorem and obtain

$$\langle j^*(h), v\rangle_V = \int_\Omega \langle \nabla w(x), \nabla v(x)\rangle \, dx \,. \tag{11.29}$$

Here, we have taken the scalar product in $H_0^1(\Omega)$ to be the L^2 scalar product of the gradients (see the Poincaré inequality). From (11.28) and (11.29) it follows that w is nothing else than the weak solution of the elliptic boundary value problem

$$\begin{aligned} -\Delta w = h &\quad \text{in } \Omega, \\ w = 0 &\quad \text{on } \partial\Omega. \end{aligned} \tag{11.30}$$

The embeddings j and j^* of an evolution triple induce via

$$u \mapsto j \circ u \mapsto j^* \circ j \circ u \tag{11.31}$$

embeddings of the corresponding L^p spaces,

$$L^p(0,T;V) \to L^p(0,T;H) \to L^p(0,T;V^*) \,. \tag{11.32}$$

In what follows we sometimes do not write j and j^* explicitly, that is, when we write "$h = v$" or "$w = v$" with $v \in V$, $h \in H$ and $w \in V^*$, it stands for "$h = j(v)$" resp. "$w = j^*(j(v))$".

Definition 11.5 (Weak time derivative)
*Let $u \in L^1(0,T;V)$. A $w \in L^1(0,T;V^*)$ is called a **weak derivative** of u if*

$$\int_0^T w(t)\varphi(t)\, dt = -\int_0^T J(u(t))\varphi'(t)\, dt \,, \tag{11.33}$$

for all $\varphi \in C_0^\infty(0,T)$. □

The integrals in (11.33) are Bochner integrals.

We want to reduce equations in V and V^* to equations in $\mathbb{R}$. For this, we need the formulas

$$\left\langle v^* , \int_0^T u(t)\, dt \right\rangle_V = \int_0^T \langle v^*, u(t)\rangle_V \, dt \,, \quad v^* \in V^* \,, \, u \in L^1(0,T;V) \,, \tag{11.34}$$

and

$$\left\langle \int_0^T u(t)\, dt \,, \, v \right\rangle_V = \int_0^T \langle u(t), v\rangle_V \, dt \,, \quad v \in V \,, \, u \in L^1(0,T;V^*) \,. \tag{11.35}$$

Proposition 11.6 *Let V be a Banach space. If $v^* \in V^*$ and $u \in L^1(0,T;V)$, then $t \mapsto \langle v^*, u(t)\rangle_V$ is integrable, and (11.34) holds. If $v \in V$ and $u \in L^1(0,T;V^*)$, then $t \mapsto \langle u(t), v\rangle_V$ is integrable, and (11.35) holds.*

Proof: We consider (11.34). If

$$u = \sum_{i=1}^{n} \chi_{A_i} v_i \tag{11.36}$$

is a simple function with values $v_i \in V$, then $t \mapsto \langle v^*, u(t)\rangle_V$ is a simple function, too, and

$$\begin{aligned}\left\langle v^*, \int_0^T u(t)\,dt \right\rangle_V &= \left\langle v^*, \sum_{i=1}^{n} \text{meas}\,(A_i) v_i \right\rangle_V = \sum_{i=1}^{n} \text{meas}\,(A_i)\,\langle v^*, v_i\rangle_V \\ &= \int_0^T \langle v^*, u(t)\rangle_V\, dt\,, \quad v^* \in V^*\,.\end{aligned}$$

Let now $u \in L^1(0,T;V)$ be arbitrary. For $v^* \in V^*$ we have $\langle v^*, u(t)\rangle_V = (v^* \circ u)(t)$. Moreover, $v^* \circ u$ is measurable, since v^* is continuous and u is Bochner measurable. Moreover,

$$\int_0^T |\langle v^*, u(t)\rangle_V|\,dt \le \int_0^T \|v^*\|_{V^*}\|u(t)\|_V\,dt \le \|v^*\|_{V^*}\|u\|_{L^1(0,T;V)}\,.$$

Analogously, we obtain

$$\left|\left\langle v^*, \int_0^T u(t)\,dt \right\rangle_V\right| \le \|v^*\|_{V^*} \left\|\int_0^T u(t)\,dt\right\|_V \le \|v^*\|_{V^*}\|u\|_{L^1(0,T;V)}\,.$$

Therefore, both sides of (11.34) define a linear and continous functional on $L^1(0,T;V)$. They coincide on the dense subspace of the simple functions and are therefore equal.

The proof of (11.35) is analogous. □

The following two results are concerned with the characterization and the uniqueness of the weak derivative.

Lemma 11.7 *Let $V \xrightarrow{j} H \xrightarrow{j^*} V^*$ be an evolution triple, $u \in L^1(0,T;V)$, $w \in L^1(0,T;V^*)$. Then w is the weak derivative of u if and only if*

$$\int_0^T (u(t), v)_H \varphi'(t)\,dt = -\int_0^T \langle w(t), v\rangle_V\, \varphi(t)\,dt\,, \quad \forall\, v \in V,\ \varphi \in C_0^\infty(0,T)\,. \tag{11.37}$$

Proof: Due to (11.35), for all $v \in V$ and all $\varphi \in C_0^\infty(0,T)$ we have

$$\begin{aligned}\int_0^T (j(u(t)), j(v))_H \varphi'(t)\,dt &= \int_0^T \langle \varphi'(t) J(u(t)), v\rangle_V\, dt \\ &= \left\langle \int_0^T J(u(t))\varphi'(t)\,dt\,, v \right\rangle_V\,,\end{aligned} \tag{11.38}$$

and

$$-\int_0^T \langle w(t), v\rangle_V\, \varphi(t)\,dt = \left\langle \int_0^T -w(t)\varphi(t)\,dt\,, v \right\rangle_V\,. \tag{11.39}$$

Therefore, (11.37) is equivalent to

$$\int_0^T J(u(t))\varphi'(t)\,dt = -\int_0^T w(t)\varphi(t)\,dt\,, \quad \text{for all } \varphi \in C_0^\infty(0,T)\,. \tag{11.40}$$

□

Corollary 11.8 *Let $w \in L^2(0,T;V^*)$ be the weak derivative of $u \in L^2(0,T;V)$. Then, for every $v \in V$, the function $t \mapsto \langle w(t), v\rangle_V$ is the weak derivative of the function $t \mapsto (u(t), v)_H$ in $L^2(0,T)$. The same is true if L^2 is replaced by L^1.*

Lemma 11.9 *Let V be a separable Banach space, $w \in L^1(0,T;V^*)$, let*

$$\int_0^T \varphi(t)w(t)\,dt = 0\,, \qquad \text{for all } \varphi \in C_0^\infty(0,T)\,. \tag{11.41}$$

Then $w = 0$ almost everywhere.

Proof: In the special case $V = \mathbb{R} = V^*$, this is the fundamental lemma of the calculus of variations, see Proposition 2.7.

For $v \in V$ and $\varphi \in C_0^\infty(0,T)$ we have

$$0 = \left\langle \int_0^T \varphi(t)w(t)\,dt\,,\, v \right\rangle_V = \int_0^T \varphi(t)\,\langle w(t), v\rangle_V\, dt\,. \tag{11.42}$$

Using the result for $V = \mathbb{R}$, we obtain that for every $v \in V$ there exists a null set $N(v)$ with

$$\langle w(t), v\rangle = 0\,, \qquad \text{for all } t \in (0,T) \setminus N(v)\,. \tag{11.43}$$

Let D be a countable dense subset of V, let

$$N = \bigcup_{v\in D} N(v)\,. \tag{11.44}$$

Then N is a null set, and

$$\langle w(t), v\rangle = 0\,, \qquad \text{for all } t \in (0,T) \setminus N\,,\ v \in D\,. \tag{11.45}$$

Since D is dense in V,

$$\langle w(t), v\rangle = 0\,, \qquad \text{for all } t \in (0,T) \setminus N\,,\ v \in V\,, \tag{11.46}$$

and thus $w(t) = 0$ for all $t \notin N$. □

Proposition 11.10 *Let V be a separable Banach space, $u \in L^2(0,T;V)$. Then there exists at most one weak derivative $w \in L^2(0,T;V^*)$ of u. If it exists, we denote it by u'.*

Proof: If $w_1, w_2 \in L^2(0,T;V^*)$ are weak derivatives of u, it follows from its definition that for $w = w_1 - w_2$ we have

$$\int_0^T w(t)\varphi(t)\,dt = 0 \tag{11.47}$$

for all $\varphi \in C_0^\infty(0,T)$. From Lemma 11.9 we conclude that $w = 0$. □

Remark: The assertions of Lemma 11.9 and of Proposition 11.10 hold for also for non-separable Banach spaces X and $w \in L^1(0,T;X)$.

We return to the parabolic initial-boundary value problem and repeat its formulation as an "abstract" initial value problem

$$\langle u'(t), v\rangle_V + a(u(t), v; t) = \langle F(t), v\rangle_V \,, \quad \text{for all } v \in V, \tag{11.48}$$

$$u(0) = u_0 \,. \tag{11.49}$$

The boundary condition $u = 0$ on $\partial\Omega$ is taken care of by the space $H_0^1(\Omega)$.

We look for a solution u in the space

$$W = \{u : u \in L^2(0, T; V),\, u' \in L^2(0, T; V^*)\} \,, \tag{11.50}$$

which satisfies (11.48) for almost all $t \in (0, T)$.

Proposition 11.11 *Let $V \overset{j}{\to} H \overset{j^*}{\to} V^*$ be an evolution triple. Then we have*

$$W \subset C([0, T]; H) \,. \tag{11.51}$$

Moreover, the rule of partial integration

$$(u(t), v(t))_H - (u(s), v(s))_H = \int_s^t \langle u'(\tau), v(\tau)\rangle_V + \langle v'(\tau), u(\tau)\rangle_V \, d\tau \tag{11.52}$$

holds for all $u, v \in W$ and all $s, t \in [0, T]$.

Proof: See for example chapter IV in the book of Gajewski, Gröger and Zacharias (in German), or Theorem 5.9.3 in Evans, or my lecture notes from the summer term 2013 (in German). □

The inclusion (11.51) has to be understood in the following sense: If $u \in W$, the equivalence classe $[j \circ u] \in L^2(0, T; H)$ of $j \circ u$ contains a continuous function (in fact, this function is then uniquely determined). The left side of (11.52) has to be interpreted in this sense, as continuous functions can be evaluated unambiguously at the points s and t. For the same reason, the initial condition (11.49) is well-defined for functions in W. (For arbitrary functions in $L^2(0, T; V)$, (11.49) is not well-defined.)

Assumption 11.12

(i) *$V \overset{j}{\to} H \overset{j^*}{\to} V^*$ is an evolution triple, V is a separable Banach space with dim$(V) = +\infty$.*

(ii) *$a : V \times V \times (0, T] \to \mathbb{R}$ is a bilinear form for every $t \in (0, T]$, end there exist $c_a, c_h, C_a > 0$ with*

$$a(v, v; t) \geq c_a \|v\|_V^2 - c_h \|j(v)\|_H^2 \,, \quad \textit{for all } v \in V,\, t \in (0, T], \tag{11.53}$$

$$|a(v, w; t)| \leq C_a \|v\| \, \|w\| \,, \quad \textit{for all } v, w \in V,\, t \in (0, T]. \tag{11.54}$$

The mappings $t \mapsto a(v, w; t)$ are measurable for all $v, w \in V$.

(iii) *$u_0 \in H$, $F \in L^2(0, T; V^*)$.*

□

Theorem 11.13 *Under the assumptions in 11.12, the initial-boundary value problem (11.48), (11.49) has a solution $u \in W$.*

The proof consists of a sequence of Lemmas; we always assume that Assumption 11.12 is satisfied.

As in chapter 9 we use Galerkin approximation. The idea is to solve first the initial value problem on finite dimensional subspaces V_n of V, and to obtain the solution on V by a limit passage. The existence of the limit results from a compactness argument. In order for this to work, the sequence of approximate solutions has to be bounded. This will follow from the properties of the bilinear form guaranteed by assumption 11.12. These properties will turn out to hold since $\partial_t + L$ is uniformly parabolic.

The Galerkin approximation is based on a sequence $(w_n)_{n\in\mathbb{N}}$ in V with

$$\dim(V_n) = n\,, \quad V_n := \operatorname{span}\{w_1, \dots, w_n\}\,, \quad V = \operatorname{cl}\left(\bigcup_{n\in\mathbb{N}} V_n\right)\,. \tag{11.55}$$

We want to determine functions $u_n : [0,T] \to V_n$, represented as

$$u_n(t) = \sum_{k=1}^n c_{nk}(t) w_k\,, \quad c_{nk} : [0,T] \to \mathbb{R}\,, \tag{11.56}$$

such that we have, for almost all $t \in (0,T]$,

$$\begin{aligned} (u_n'(t), v)_H + a(u_n(t), v; t) &= \langle F(t), v\rangle_V\,, \qquad \text{for all } v \in V_n\,, \\ u_n(0) &= u_{0n}\,. \end{aligned} \tag{11.57}$$

The initial values $u_{0n} \in V_n$ are uniquely defined (recall that j is injective) by

$$j(u_{0n}) = p(u_0, j(V_n))\,. \tag{11.58}$$

Here, $p(u_0, j(V_n))$ denotes the orthogonal projection of u_0 to the closed subspace $j(V_n)$ of H. This definition ensures that $j(u_{0n}) \to u_0$ in H and that $\|j(u_{0n})\|_H \le \|u_0\|_H$. Let

$$u_{0n} = \sum_{k=1}^n \alpha_{nk} w_k \in V_n\,. \tag{11.59}$$

A system equivalent to (11.57) and (11.58) is given by

$$\sum_{k=1}^n c_{nk}'(t)(w_k, w_i)_H + \sum_{k=1}^n c_{nk}(t) a(w_k, w_i; t) = \langle F(t), w_i\rangle_V\,, \quad 1 \le i \le n\,, \tag{11.60}$$

$$c_{nk}(0) = \alpha_{nk}\,, \qquad 1 \le k \le n\,. \tag{11.61}$$

Lemma 11.14 *The Galerkin equations (11.57), (11.58) have a unique solution $u_n : [0,T] \to V_n$ with $u_n' \in L^2(0,T;V_n)$ and*

$$u_n(t) = u_{0n} + \int_0^t u_n'(s)\,ds\,. \tag{11.62}$$

Proof: The vectors $w_1, \dots, w_n$ are linearly independent in V. Since j is injective, the vectors $j(w_1), \dots, j(w_n)$ are linearly independent in H. The matrix

$$B = (b_{ik}), \quad b_{ik} = (w_k, w_i)_H, \tag{11.63}$$

is invertible (exercise), and we may write (11.60), (11.61) in the form

$$c_n'(t) + B^{-1}\tilde{A}(t)c_n(t) = B^{-1}\tilde{F}(t), \quad c_n(0) = \alpha_n, \tag{11.64}$$

where

$$\tilde{A}(t) = (\tilde{a}_{ik}(t)), \quad \tilde{a}_{ik}(t) = a(w_k, w_i; t), \quad \tilde{F}_i(t) = \langle F(t), w_i \rangle_V . \tag{11.65}$$

By our assumptions, $\tilde{A} \in L^\infty(0, T; \mathbb{R}^{(n,n)})$, and because of

$$|\tilde{F}_i(t)| = |\langle F(t), w_i \rangle_V| \le \|F(t)\|_{V^*} \|w_i\|_V$$

we have $\tilde{F} \in L^2(0, T; \mathbb{R}^n)$. The initial value problem (11.64) has a unique solution $c_n : [0, T] \to \mathbb{R}^n$ with

$$c_n(t) = c_n(0) + \int_0^t c_n'(s)\, ds, \quad c_n' \in L^2(0, T; \mathbb{R}^n), \tag{11.66}$$

according to the Picard-Lindelöf theorem in the version for measurable right hand sides, see for example the book of Walter on ordinary differential equations. $\square$

Lemma 11.15 *There exists a constant $C > 0$, which does not depend on n, such that*

$$\max_{t\in[0,T]} \|u_n(t)\|_H + \|u_n\|_{L^2(0,T;V)} + \|u_n'\|_{L^2(0,T;V^*)} \le C(\|u_0\|_H + \|F\|_{L^2(0,T;V^*)}). \tag{11.67}$$

Proof: Let $n \in \mathbb{N}$ be fixed. We set $v = u_n(t)$ in (11.57) and obtain

$$(u_n'(t), u_n(t))_H + a(u_n(t), u_n(t); t) = \langle F(t), u_n(t) \rangle_V . \tag{11.68}$$

We have

$$\frac{\mathrm{d}}{\mathrm{dt}} \frac{1}{2}(u_n(t), u_n(t))_H = (u_n'(t), u_n(t))_H \tag{11.69}$$

according to the product rule for bilinear forms (applied on the finite dimensional subspace V_n). By our assumption on a,

$$a(u_n(t), u_n(t); t) \ge c_a \|u_n(t)\|_V^2 - c_h \|u_n(t)\|_H^2 . \tag{11.70}$$

We integrate (11.68) over $[0, t]$ and obtain, using (11.69) and (11.70),

$$\begin{aligned} &\frac{1}{2}\|u_n(t)\|_H^2 - \frac{1}{2}\|u_n(0)\|_H^2 + c_a \int_0^t \|u_n(s)\|_V^2\, ds \\ &\qquad \le c_h \int_0^t \|u_n(s)\|_H^2\, ds + \int_0^t \langle F(s), u_n(s) \rangle_V\, ds \end{aligned} \tag{11.71}$$

According to our choice of u_{0n},

$$\|u_{0n}\|_H \le \|u_0\|_H . \tag{11.72}$$

Using Young's inequality, we get

$$\int_0^t \langle F(s), u_n(s)\rangle_V \, ds \le \frac{c_a}{2}\int_0^t \|u_n(s)\|_V^2 \, ds + \frac{1}{2c_a}\int_0^t \|F(s)\|_{V^*}^2 \, ds\,. \tag{11.73}$$

We insert (11.72) and (11.73) into (11.71), multiply by 2 and obtain

$$\begin{aligned} &\|u_n(t)\|_H^2 + c_a \int_0^t \|u_n(s)\|_V^2 \, ds \\ &\qquad \le \|u_0\|_H^2 + 2c_h \int_0^t \|u_n(s)\|_H^2 \, ds + \frac{1}{c_a}\int_0^t \|F(s)\|_{V^*}^2 \, ds\,. \end{aligned} \tag{11.74}$$

For

$$\eta(t) = \|u_n(t)\|_H^2 \tag{11.75}$$

it follows that

$$\eta(t) \le d_0 + 2c_h \int_0^t \eta(s)\, ds\,, \quad d_0 = \|u_0\|_H^2 + \frac{1}{c_a}\int_0^T \|F(s)\|_{V^*}^2 \, ds\,, \quad t \in [0,T]\,. \tag{11.76}$$

Using Gronwall's lemma we conclude that

$$\eta(t) \le d_0 e^{2c_h t}\,, \quad \text{for all } t \in [0,T]. \tag{11.77}$$

Thus, for a suitable constant C,

$$\|u_n(t)\|_H^2 \le C(\|u_0\|_H^2 + \|F\|_{L^2(0,T;V^*)}^2)\,, \quad \text{for all } t \in [0,T]. \tag{11.78}$$

From (11.74) it follows (with C suitably enlarged)

$$\|u_n\|_{L^2(0,T;V)}^2 \le C(\|u_0\|_H^2 + \|F\|_{L^2(0,T;V^*)}^2)\,. \tag{11.79}$$

In order to estimate u_n' in the norm of V^* we consider once more the variational equation

$$\langle u_n'(t), v\rangle_V + a(u_n(t), v; t) = \langle F(t), v\rangle_V\,, \quad \text{for all } v \in V. \tag{11.80}$$

It follows that

$$|\langle u_n'(t), v\rangle_V| \le C_a \|u_n(t)\|_V \|v\|_V + \|F(t)\|_{V^*}\|v\|_V\,, \tag{11.81}$$

thus

$$\|u_n'(t)\|_{V^*} \le C_a \|u_n(t)\|_V + \|F(t)\|_{V^*}\,, \tag{11.82}$$

and therefore

$$\begin{aligned} \|u_n'\|_{L^2(0,T;V^*)}^2 &\le \int_0^T (C_a\|u_n(t)\|_V + \|F(t)\|_{V^*})^2 \, dt \\ &\le 2C_a^2 \|u_n\|_{L^2(0,T;V)}^2 + 2\|F\|_{L^2(0,T;V^*)}^2\,. \end{aligned}$$

The assertion now follows from (11.79). □

Lemma 11.16 *There exists a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ and a $u \in L^2(0,T;V)$ such that u_{n_k} converges weakly in $L^2(0,T;V)$. The limit function u satisfies*

$$\begin{aligned} &-(u_0,v)_H\,\varphi(0) - \int_0^T \langle u(t),v\rangle_V\,\varphi'(t)\,dt + \int_0^T a(u(t),v;t)\varphi(t)\,dt \\ &\qquad = \int_0^T \langle F(t),v\rangle_V\,\varphi(t)\,dt\,, \end{aligned} \tag{11.83}$$

for all $v \in V$ and all $\varphi \in C^1[0,T]$ with $\varphi(T) = 0$.

Proof: Due to Lemma 11.15, (u_n) is bounded in the Hilbert space $L^2(0,T;V)$. Thus, there exists a subsequence (u_{n_k}) which converges weakly to some $u \in L^2(0,T;V)$.

Let $i \in \mathbb{N}$, $v \in V_i$, $\varphi \in C^1[0,T]$ with $\varphi(T) = 0$. For $n \geq i$ we get, because of (11.57),

$$\int_0^T (u_n'(t),v)_H\varphi(t)\,dt + \int_0^T a(u_n(t),v;t)\varphi(t)\,dt = \int_0^T \langle F(t),v\rangle_V\,\varphi(t)\,dt\,. \tag{11.84}$$

Partial integration yields

$$\begin{aligned} &-(u_n(0),v)_H\varphi(0) - \int_0^T \langle u_n(t),v\rangle_V\,\varphi'(t)\,dt + \int_0^T a(u_n(t),v;t)\varphi(t)\,dt \\ &\qquad = \int_0^T \langle F(t),v\rangle_V\,\varphi(t)\,dt\,. \end{aligned} \tag{11.85}$$

Since $u_n(0) = u_{0n} \to u_0$ in H, we obtain

$$\lim_{k\to\infty} -(u_{n_k}(0),v)_H = -(u_0,v)_H\,. \tag{11.86}$$

We investigate the passage to the limit for the two integrals on the left side of (11.85). The assignment

$$z \mapsto \int_0^T \langle z(t),v\rangle_V\,\varphi'(t)\,dt \tag{11.87}$$

defines a linear continuous functional on $L^2(0,T;V)$, since

$$\begin{aligned} \left|\int_0^T \langle J(z(t)),v\rangle_V\,\varphi'(t)\,dt\right| &\leq \int_0^T \|J(z(t))\|_{V^*}\|v\|_V|\varphi'(t)|\,dt \\ &\leq \|j^*\circ j\|\,\|v\|_V\|\varphi'\|_\infty \int_0^T \|z(t)\|_V\,dt \\ &\leq \|j^*\circ j\|\,\|v\|_V\|\varphi'\|_\infty\sqrt{T}\|z\|_{L^2(0,T;V)}\,. \end{aligned} \tag{11.88}$$

The weak convergence $u_{n_k} \rightharpoonup u$ therefore implies that

$$\lim_{k\to\infty}\int_0^T \langle u_{n_k}(t),v\rangle_V\,\varphi'(t)\,dt = \int_0^T \langle u(t),v\rangle_V\,\varphi'(t)\,dt\,. \tag{11.89}$$

Analogously, from

$$\int_0^T a(z(t),v;t)\varphi(t)\,dt \leq C_a\|v\|_V\|\varphi\|_\infty\sqrt{T}\|z\|_{L^2(0,T;V)} \tag{11.90}$$

for all $z \in L^2(0,T;V)$ we obtain that

$$\lim_{k\to\infty} \int_0^T a(u_{n_k}(t), v; t)\varphi(t)\, dt = \int_0^T a(u(t), v; t)\varphi(t)\, dt\,. \tag{11.91}$$

Thus, (11.83) is proved for $v \in V_i$. Since i was arbitrary and $\cup_i V_i$ is dense in V, the assertion follows. □

Lemma 11.17 *The expression*

$$\langle \tilde\alpha(t), v\rangle_V = a(u(t), v; t)\,, \quad v \in V\,, \quad t \in (0,T)\,, \tag{11.92}$$

defines a function $\tilde\alpha \in L^2(0,T;V^*)$ which satisfies

$$\|\tilde\alpha\|_{L^2(0,T;V^*)} \le C_a \|u\|_{L^2(0,T;V)}\,. \tag{11.93}$$

Proof: We have

$$|a(u(t), v; t)| \le C_a \|u(t)\|_V \|v\|_V\,, \quad v \in V\,,$$

for almost all $t \in (0,T)$. Therefore,

$$\tilde\alpha(t) \in V^*\,, \quad \|\tilde\alpha(t)\|_{V^*} \le C_a \|u(t)\|_V$$

for almost all $t \in (0,T)$, and (11.93) follows from

$$\int_0^T \|\tilde\alpha(t)\|_{V^*}^2\, dt \le C_a^2 \int_0^T \|u(t)\|_V^2\, dt\,.$$

□

Proof of Theorem 11.13. According to Lemma 11.16 and Lemma 11.17, for all $\varphi \in C_0^\infty(0,T)$ and all $v \in V$ we have

$$-\int_0^T (u(t), v)_H \varphi'(t)\, dt = -\int_0^T \langle \tilde\alpha(t), v\rangle_V\, \varphi(t)\, dt + \int_0^T \langle F(t), v\rangle_V\, \varphi(t)\, dt\,. \tag{11.94}$$

By Lemma 11.7, u has the weak derivative

$$u' \in L^2(0,T;V^*)\,, \quad u'(t) = -\tilde\alpha(t) + F(t)\,,$$

so

$$\langle u'(t), v\rangle_V + a(u(t), v; t) = \langle F(t), v\rangle_V\,, \quad v \in V\,, \tag{11.95}$$

as claimed, and $u \in W$. From Proposition 11.11 we obtain that $u \in C([0,T];H)$. We now replace in (11.83) the two rightmost integrals according to (11.95) and obtain

$$-(u_0, v)_H \varphi(0) - \int_0^T \langle u(t), v\rangle_V\, \varphi'(t)\, dt = \int_0^T \langle u'(t), v\rangle_V\, \varphi(t)\, dt \tag{11.96}$$

for all $v \in V$ and all $\varphi \in C^1[0,T]$ with $\varphi(T) = 0$. Since, on the other hand, the function $t \mapsto \varphi(t)v$ also is an element of W, we obtain for such functions φ by Proposition 11.11

$$-(u(0), v)_H \varphi(0) = \int_0^T \langle u'(t), \varphi(t)v\rangle_V + \langle \varphi'(t)v, u(t)\rangle_V\, dt\,. \tag{11.97}$$

Choosing in particular a φ with $\varphi(0) = 1$ it follows from (11.96) and (11.97) that

$$(u(0) - u_0, v)_H = 0 \tag{11.98}$$

for all $v \in V$. As $j(V)$ is dense in H, we obtain $u(0) = u_0$. Theorem 11.13 is now completely proved. □

Theorem 11.18 *Under the assumptions 11.12, the initial value problem (11.48), (11.49) has a unique solution $u \in W$. Moreover, there exists a constant $C > 0$ which is independent from u_0 and F such that*

$$\max_{t\in[0,T]} \|u(t)\|_H + \|u\|_{L^2(0,T;V)} + \|u'\|_{L^2(0,T;V^*)} \leq C(\|u_0\|_H + \|F\|_{L^2(0,T;V^*)}) . \tag{11.99}$$

Proof: Existence follows from Proposition 11.13. For every solution $u \in W$ it follows from Proposition 11.11 that for every $t \in [0, T]$

$$\begin{aligned} \frac{1}{2}(u(t), u(t))_H - \frac{1}{2}(u_0, u_0)_H &= \int_0^t \langle u'(s), u(s)\rangle_V \, ds \\ &= -\int_0^t a(u(s), u(s); s)\, ds + \int_0^t \langle F(s), u(s)\rangle_V \, ds . \end{aligned}$$

Thus,

$$\frac{1}{2}\|u(t)\|_H^2 + \int_0^t a(u(s), u(s); s)\, ds = \frac{1}{2}\|u_0\|_H^2 + \int_0^t \langle F(s), u(s)\rangle_V \, ds . \tag{11.100}$$

It follows that

$$\begin{aligned} &\frac{1}{2}\|u(t)\|_H^2 + c_a \int_0^t \|u(s)\|_V^2 \, ds - c_h \int_0^t \|u(s)\|_H^2 \, ds \\ &\qquad \leq \frac{1}{2}\|u_0\|_H^2 + \frac{c_a}{2} \int_0^t \|u(s)\|_V^2 \, ds + \frac{1}{2c_a} \int_0^t \|F(s)\|_{V^*}^2 \, ds . \end{aligned} \tag{11.101}$$

Using Gronwall's lemma, we obtain with an argument analogous to that in the proof of Lemma 11.15

$$\max_{t\in[0,T]} \|u(t)\|_H \leq C(\|u_0\|_H + \|F\|_{L^2(0,T;V^*)}) . \tag{11.102}$$

From (11.101) we also obtain

$$\|u\|_{L^2(0,T;V)} \leq C(\|u_0\|_H + \|F\|_{L^2(0,T;V^*)}) . \tag{11.103}$$

From (11.95) and Lemma 11.17 it follows that

$$\|u'\|_{L^2(0,T;V^*)} \leq C_a\|u\|_{L^2(0,T;V)} + \|F\|_{L^2(0,T;V^*)} . \tag{11.104}$$

From (11.101) and (11.104) we now obtain (11.99) as well as the uniqueness, since to the difference of two solutions we may apply (11.99) with $F = 0$, $u_0 = 0$. □

We return to the initial-boundary value problem

$$\partial_t u + Lu = f \quad \text{in } \Omega_T = \Omega \times (0, T], \tag{11.105}$$

$$u = 0 \quad \text{on } \partial\Omega \times (0, T), \tag{11.106}$$

$$u(x, 0) = u_0(x) \quad \text{for } x \in \Omega. \tag{11.107}$$

As a weak solution of (11.105) – (11.107) we define a solution of the associated initial value problem (11.48), (11.49), which is discussed above. In this case,

$$a(w, v; t) = \int_\Omega (\nabla w(x))^T A(x,t) \nabla v(x)\, dx + \int_\Omega b(x,t)^T \nabla w(x) v(x)\, dx \tag{11.108}$$

$$+ \int_\Omega c(x,t) w(x) v(x)\, dx\,, \tag{11.109}$$

and

$$(F(t))(v) = \int_\Omega f(x,t) v(x)\, dx\,. \tag{11.110}$$

Theorem 11.19 (Unique solvability of the initial-boundary value problem) *Let the assumptions 11.1 be satisfied. The the problem (11.105) – (11.107) has a unique solution $u \in W$.*

Proof: The assertion is a consequence of Proposition 11.18; we only have to check that the assumptions in 11.12 are satisfied. As evolution triple we choose

$$V = H_0^1(\Omega)\,, \quad H = L^2(\Omega)\,, \quad (j(v))(x) = v(x)\,.$$

Since $\partial_t + L$ is uniformly parabolic, Gårding's inequality (Lemma 3.5) implies that

$$a(v, v; t) \geq c_a \|v\|_V^2 - c_h \|v\|_H^2\,, \quad \text{f'ur alle } v \in V,\ t \in (0, T],$$

holds with suitable constants c_a, c_h. Since the coefficients A, b, and c are bounded, it follows that

$$|a(v, w; t)| \leq C_a \|v\|_V\, \|w\|_V\,, \quad \text{for all } v, w \in V,\ t \in (0, T],$$

for a suitable constant C_a. Moreover,

$$\begin{aligned} |(F(t))(v)|^2 &\leq \int_\Omega |f(x,t) v(x)|^2\, dx \leq \int_\Omega |f(x,t)|^2\, dx \cdot \int_\Omega |v(x)|^2\, dx \\ &\leq C_\Omega \int_\Omega |f(x,t)|^2\, dx \cdot \|v\|_V^2\,, \end{aligned}$$

where we have used Poincaré's inequality. It follows that

$$\|F\|^2_{L^2(0,T;V^*)} = \int_0^T \|F(t)\|^2_{V^*}\, dt \leq C_\Omega \int_0^T \int_\Omega |f(x,t)|^2\, dx\, dt \leq C_\Omega \|f\|^2_{L^2(\Omega_T)}\,,$$

therefore $F \in L^2(0, T; V^*)$. □

The method described above reduces the initial-boundary value problem to an initial value problem in the context of an evolution triple. This method can also be applied to other partial differential equations. For example, one can reformulate hyperbolic equations (in particular, the wave equation) as a second order equation (with leading term $u''(t)$) in a form analogous to (11.48), and one can prove a corresponding existence and uniqueness result.

The method of Galerkin approximation can also be used for the numerical solution of a parabolic initial-boundary value problem. One chooses V_n together with a basis $\{w_1, \ldots, w_n\}$ and solves the system

$$\sum_{k=1}^{n} c'_{nk}(t)(j(w_k), j(w_i))_H + \sum_{k=1}^{n} c_{nk}(t) a(w_k, w_i; t) = \langle F(t), w_i \rangle_V \,, \quad 1 \le i \le n \,, \tag{11.111}$$

of n ordinary differential equations for the unknown functions c_{nk}, $1 \le k \le n$, by a suitable numerical method.

12 PDEs in Continuum Mechanics

Space as a continuum. Continuum mechanics describes phenomena in space and time. Space is modeled as $\mathbb{R}^3$, and time is modeled as $\mathbb{R}$. With regard to space, continuum mechanics operates on a macroscopic scale, many magnitudes of order above molecular and atomic length scales. Nevertheless, a function with arguments in $\mathbb{R}^3$ has a mathematical meaning on all (even arbitrarily small) length scales. In view of this contrast, what is the relation between "reality" and mathematical model ?

As an example we consider a porous material, that is, a body consisting of solid material and of cavities (also called pores). Examples are a pile of sand or of rocks, or a cellular .... Whether the cavities are empty or filled (with a gas or a fluid) does not matter at the moment. Let $G \subset \mathbb{R}^d$ be the region in space that we consider, let $P \subset G$ be the region "occupied" by the cavities. Depending on the situation we may have $d = 1, 2,$ or 3. Let $x \in G$ be a point, let

$$W_h(x) = \{z : \|z - x\|_\infty \leq \frac{h}{2}\} \tag{12.1}$$

be a cube completely contained in G with center x and sidelength h. We define

$$\varphi_h(x) = \frac{\mathrm{vol}_d(P \cap W_h(x))}{\mathrm{vol}_d(W_h(x))} = h^{-d}\mathrm{vol}_d(P \cap W_h(x)) \,. \tag{12.2}$$

The function φ_h yields the fraction of cavity volume in the cube $W_h(x)$. We want to replace this function by a function which depends only on x, not on h. This means that we want to define the fraction of cavity volume "at a point x". For this, it does **not** make sense to look for a limit as $h \to 0$; one wants to obtain a function which represents the behaviour for small, but finite values of h.

The **basic assumption** which underlies such a procedure is the following:

> There is an interval I of values of h, in which $h \mapsto \varphi_h(x)$ is approximately constant, at all points x of G.

This basic assumption means that, in this interval, averaging according to (12.2) yields a reasonable approximation to the real situation. The values of h in I are typically big with respect to the fine structure of the porous material. In addition, one requires that I also includes values of h which are small when compared to the size of G. If this is not the case, it would not make much sense to consider a dependence on x. One now models the volume fraction of the cavities (as a variable which contributes to the description of the porous structure) as a function $\varphi : G \to \mathbb{R}$ which approximates φ_h from (12.2) This function is called the **porosity** of the material.

We may rewrite (12.2) equivalently as

$$\mathrm{vol}_d(P \cap W_h(x)) = \varphi_h(x) \cdot h^d = \varphi_h(x)\mathrm{vol}_d(W_h(x)) \,. \tag{12.3}$$

For the continuum model this means that

$$\int_U \varphi(x)\, dx \tag{12.4}$$

is a reasonable approximation for the actual volume $\mathrm{vol}_d(P \cap U)$ of the cavity. We may interpret the function φ as the **density function** of a measure. This measure approximates the measure ν defined by $\nu(U) = \mathrm{vol}_d(P \cap U)$ which yields the actual cavity volume in subsets U of G.

Another example of a variable which is continuous in space is the **mass density**. Let us consider a solid body occupying a region $G \subset \mathbb{R}^d$. Let the cube $W_h(x)$ given by (12.1) have the mass $m(W_h(x))$. Its average per volume is

$$\rho_h(x) = \frac{m(W_h(x))}{\mathrm{vol}_d(W_h(x))}\,. \tag{12.5}$$

In the same manner as above one introduces the continuous mass density $\rho : G \to \mathbb{R}$. Then the mass of a subset $U \subset G$ is taken to be

$$m(U) = \int_U \rho(x)\,dx\,. \tag{12.6}$$

Space and time as continuum. In this case, the density function depends on time, "$\varphi = \varphi(t,x)$". Let us assume that φ models a quantity of dimension $\mathcal{G}$. For example, $\mathcal{G}$ could be kilograms or calories. The density function φ has dimension $\mathcal{G}\mathcal{L}^{-d}$, where d is the dimension of space, and $\mathcal{L}$ is the dimension of length. The total amount of the quantity contained in a region $U \subset \mathbb{R}^d$ at time t is

$$\int_U \varphi(t,x)\,dx\,, \tag{12.7}$$

which has dimension $\mathcal{G}$.

We want to derive a general so-called **balance equation** in space dimension $d = 1$, that is, for regions $U = (a,b)$. We assume that the total amount of the quantity moving (or "flowing") through a given space point $x \in (a,b)$ during the time interval $[t,\tau]$ is given by

$$\int_t^\tau q(s,x)\,ds\,. \tag{12.8}$$

The function q is called **flux**, it is also a density function. It describes the time rate with which the quantity flows through the point x in the positive x-direction, it has the dimension $\mathcal{G}\mathcal{T}^{-1}$, where $\mathcal{T}$ is a dimension of time.

A balance equation describing this process is

$$\int_a^b \varphi(\tau,x)\,dx - \int_a^b \varphi(t,x)\,dx = -\int_t^\tau q(s,b)\,ds + \int_t^\tau q(s,a)\,ds\,. \tag{12.9}$$

This equation means that the difference of the total amount of the quantity within (a,b) between time t and time τ equals the total amount of the quantity which flows through the boundary points a and b during the time interval $[t,\tau]$. This is valid for a process where the "items" of the quantity (e.g. the individual fluid particles) just flow but are neither destroyed nor created.

Let us now assume that the balance equation (12.9) is valid for every space interval (a,b) and every time interval $[t,\tau]$. Dividing by $\tau - t$ yields

$$\int_a^b \frac{\varphi(\tau,x) - \varphi(t,x)}{\tau - t}\,dx = \frac{1}{\tau - t}\int_t^\tau -q(s,b) + q(s,a)\,ds\,. \tag{12.10}$$

Passing to the limit $\tau \to t$ yields

$$\int_a^b \partial_t \varphi(t,x)dx = -q(t,b) + q(t,a)\,. \tag{12.11}$$

Analogously, dividing by $b-a$ passing to the limit $b \to a$ gives

$$\partial_t \varphi(t,a) = -\partial_x q(t,a)\,.$$

Since (12.9) is assumed to hold for all intervals (a,b) and $[t,\tau]$ of interest, we arrive at the partial differential equation

$$\partial_t \varphi(t,x) + \partial_x q(t,x) = 0\,, \quad \text{for all } t,x, \tag{12.12}$$

for the two unknown functions φ and q.

Since this equation is valid for all kinds of processes which conserve a quantity in the sense described above, it is not surprising that it alone is not sufficient to determine the functions φ and q. To describe a specific process, one needs additional properties, for example an equation of the form

$$q = Q(\varphi)\,, \quad \text{that is,} \quad q(t,x) = Q(\varphi(t,x))\,, \tag{12.13}$$

with a given function Q. (Other forms are also possible, for example $q = Q(\nabla\varphi)$.) Such an equation is called a **constitutive equation**. Inserting (12.13) into (12.12) we obtain

$$\partial_t \varphi(t,x) + Q'(\varphi(t,x))\partial_x \varphi(t,x) = 0\,. \tag{12.14}$$

As an example we consider a model for traffic flow. The interval (a,b) corresponds to a section of a road. The quantity to be modelled is the number of cars. Thus, φ describes the time dependent spatial density of the number of cars. We interpret

$$\int_a^b \varphi(t,x)\,dx \quad = \quad \text{number of cars in the section } (a,b) \text{ at time } t,$$

and

$$\int_t^\tau q(s,x)\,ds$$

as the number of cars which pass the point x within the time interval $[t,\tau]$, so that $q(x,t)$ yields the rate (amount per unit of time) at time t.

Kinematics. The so-called kinematics is concerned with the change in time of quantities which are modelled as continuous in space. It is based on the assumption that the quantities are attached to material points ("particles") which move in space. One fixes a reference configuration $\Omega \subset \mathbb{R}^d$. Let a material point occupy the point $X \in \Omega$ at time t_0. Its position at time t is described by a function

$$t \mapsto x(t,X)\,.$$

One assumes that this function satisfies, for all $x \in \Omega$ and all times t of interest,

$$x(t_0, X) = X\,. \tag{12.15}$$

$$(t,X) \mapsto x(t,X) \text{ is continuously differentiable.} \tag{12.16}$$

$$X \mapsto x(t,X) \text{ is bijective between } \Omega \text{ and } x(t,\Omega), \text{ for every fixed } t. \tag{12.17}$$

$$J(t,X) = \det \partial_X x(t,X) > 0\,. \tag{12.18}$$

Here, $\partial_X x(t,X) \in \mathbb{R}^{(d,d)}$ denotes the Jacobi matrix of the derivative with respect to X,

$$(\partial_X x(t,X))_{jk} = \partial_{X_k} x_j(t,X)\,.$$

The values of the quantity density to be modeled can be described either as $\varphi(t,x)$ or as $\Phi(t,X)$, namely

- $\varphi(t,x)$ refers to the value at the material point which occupies the point x at time t ("the quantity is given in **Euler coordinates**"),
- $\Phi(t,X)$ refers to the material point at time t which originally (in the reference configuration at time t_0) occupied the point X ("the quantity is given in **Lagrange coordinates**").

Since we intend to describe with φ and Φ the same quantity, we must have

$$\varphi(t,x(t,X)) = \Phi(t,X)\,. \tag{12.19}$$

The chain rule yields

$$\partial_t\Phi(t,X) = \partial_t\varphi(t,x(t,X)) + \nabla_x\varphi(t,x(t,X))\cdot\partial_t x(t,X)\,. \tag{12.20}$$

This expression describes the rate of change of the quantity attached to a moving material point (characterized by its position X in the reference configuration) at time t.

The velocity of this material point X at time t is given in Lagrange coordinates by

$$V(t,X) = \partial_t x(t,X)\,, \tag{12.21}$$

this is nothing else than the tangent vector to the curve $t \mapsto x(t,X)$. In Euler coordinates it becomes

$$v(t,x) = V(t,X(t,x))\,,\quad v(t,x(t,X)) = V(t,X)\,. \tag{12.22}$$

The mapping $t \mapsto X(t,x)$ specifies from which point of the reference configuration the material point occupying the position x at time t "has started". We have

$$X = X(t,x(t,X))\,,\quad \text{for all } X \in \Omega.$$

Expressing (12.20) in Euler coordinates we obtain the so-called **material derivative**

$$D_t\varphi(t,x) := \partial_t\varphi(t,x) + \nabla_x\varphi(t,x)\cdot v(t,x)\,. \tag{12.23}$$

It also describes the rate of change of the quantity in a moving material point, namely in that point which at time t occupies the position x.

We define

$$\Omega(t) = x(t,\Omega) = \{x(t,X) : X \in \Omega\}\,. \tag{12.24}$$

According to the substitution formula of multidimensional integration,

$$\mathrm{vol}_d(\Omega(t)) = \int_{\Omega(t)} 1\,dx = \int_\Omega J(t,X)\,dX\,. \tag{12.25}$$

Let now $\varphi(t,\cdot)$ be the density per volume of a quantity. Its total amount within the volume $\Omega(t)$ is given by

$$\int_{\Omega(t)} \varphi(t,x)\,dx = \int_\Omega \varphi(t,x(t,X))J(t,X)\,dX = \int_\Omega \Phi(t,X)J(t,X)\,dX\,. \tag{12.26}$$

In order to compute the time rate of change of this total amount within the time dependent volume $\Omega(t)$, we will need the time derivative $\partial_t J(t,X)$ of the determinant of the Jacobian.

Lemma 12.1 *Let I be an open interval, $A : I \to \mathbb{R}^{(d,d)}$ continuously differentiable, let $A(t)$ be invertible for all $t \in I$. Then*

$$\frac{\mathrm{d}}{\mathrm{dt}} \det A(t) = \operatorname{trace}\big(A(t)^{-1}A'(t)\big) \det A(t)\,. \tag{12.27}$$

Proof: We first compute the derivative of $\det : \mathbb{R}^{(d,d)} \to \mathbb{R}$ at a given fixed matrix A. From linear algebra we gather the identity

$$(\det A)I = A \cdot \hat{A}\,, \quad \hat{A} = (\det A)A^{-1}\,. \tag{12.28}$$

Here, $\hat{A}$ is the adjunct of A having the elements

$$\hat{a}_{ij} = (-1)^{i+j} \det A^{(ji)}\,, \tag{12.29}$$

where $A^{(ji)}$ is obtained from A by deleting the j-th row and the i-th column. From (12.28) we see that the diagonal elements $\det A$ of $A\hat{A}$ satisfy

$$\det A = \sum_{k=1}^d a_{ik}\hat{a}_{ki}\,, \quad 1 \le i \le d\,. \tag{12.30}$$

Due to (12.29), the elements $\hat{a}_{ki}$ do not depend on the elements a_{ij} of the i-th row of A. It follows that

$$\partial_{a_{ij}}(\det A) = \hat{a}_{ji}\,, \quad 1 \le i,j \le d\,, \tag{12.31}$$

and furthermore

$$\begin{aligned} ((D\det)(A))(H) &= \sum_{i,j=1}^d \partial_{a_{ij}}(\det A)h_{ij} = \sum_{i,j=1}^d \hat{a}_{ji}h_{ij} = \sum_{j=1}^d\sum_{i=1}^d \hat{a}_{ji}h_{ij} \\ &= \operatorname{trace}(\hat{A}H)\,, \quad \text{for all } H \in \mathbb{R}^{(d,d)}. \end{aligned} \tag{12.32}$$

Using the chain rule we now obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}} \det A(t) &= ((D\det)(A(t)))A'(t) = \operatorname{trace}(\hat{A}(t)A'(t)) \\ &= \operatorname{trace}\big(A(t)^{-1}A'(t)\big) \det A(t)\,. \end{aligned} \tag{12.33}$$

$\square$

Proposition 12.2 (Euler's expansion formula)
Let (12.15) – (12.18) hold, let moreover $(t,X) \mapsto \partial_t x(t,X)$ be continuously differentiable. Then

$$\partial_t J(t,X) = (\operatorname{div} v)(t,x(t,X))J(t,X)\,. \tag{12.34}$$

Proof: We apply Lemma 12.1 to

$$A(t) = \partial_X x(t,X)\,.$$

For the elements of $A(t)$ we get

$$\begin{aligned} a'_{ij}(t) &= \partial_t \partial_{X_j} x_i(t,X) = \partial_{X_j}\partial_t x_i(t,X) = \partial_{X_j} V_i(t,X) = \partial_{X_j} v_i(t,x(t,X)) \\ &= \sum_{k=1}^d \partial_k v_i(t,x(t,X))\partial_{X_j} x_k(t,X)\,. \end{aligned} \tag{12.35}$$

It follows that

$$\begin{aligned} \operatorname{trace}\big(A(t)^{-1}A'(t)\big) &= \sum_{i,j=1}^d (A^{-1}(t))_{ji} a'_{ij}(t) = \sum_{i,k=1}^d \sum_{j=1}^d (A^{-1}(t))_{ji}\partial_k v_i(t,x(t,X)) a_{kj}(t) \\ &= \sum_{i,k=1}^d \delta_{ki}\partial_k v_i(t,x(t,X)) = (\operatorname{div} v)(t,x(t,X))\,, \end{aligned} \tag{12.36}$$

and finally

$$\partial_t J(t,X) = \frac{\mathrm{d}}{\mathrm{dt}} \det A(t) = \operatorname{trace}\big(A(t)^{-1}A'(t)\big)\det A(t) = (\operatorname{div} v)(t,x(t,X))J(t,X)\,.$$

□

Corollary 12.3 (Volume change)
The rate of change of the volume of $\Omega(t)$ is given by

$$\frac{\mathrm{d}}{\mathrm{dt}}\operatorname{vol}\Omega(t) = \int_{\Omega(t)} \operatorname{div} v(t,x)\,dx\,. \tag{12.37}$$

Proof: We again use the substitution formula of multidimensional integration. We have

$$\operatorname{vol}\Omega(t) = \int_\Omega J(t,X)\,dX\,,$$

therefore

$$\frac{\mathrm{d}}{\mathrm{dt}}\operatorname{vol}\Omega(t) = \int_\Omega (\operatorname{div} v)(t,x(t,X))J(t,X)\,dX = \int_{\Omega(t)} \operatorname{div} v(t,x)\,dx\,.$$

□

We now can compute the rate of change of the total amount of a quantity described by φ within a time dependent volume $\Omega(t)$.

Proposition 12.4 (Transport theorem of Reynolds)
Let (12.15) – (12.18) be satisfied, let moreover $(t,X)\mapsto \partial_t x(t,X)$ as well as $(t,x)\mapsto \varphi(t,x)$ be continuously differentiable. Then

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_{\Omega(t)} \varphi(t,x)\,dx = \int_{\Omega(t)} \partial_t\varphi(t,x) + (\operatorname{div}(\varphi v))(t,x)\,dx\,. \tag{12.38}$$

Proof: From Proposition 12.2 we obtain, using the chain rule and the substitution rule

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{dt}}\int_{\Omega(t)}\varphi(t,x)\,dx &= \frac{\mathrm{d}}{\mathrm{dt}}\int_{\Omega}\varphi(t,x(t,X))J(t,X)\,dX\\ &= \int_{\Omega}\Bigg[\partial_t\varphi(t,x(t,X)) + \sum_{k=1}^{d}\partial_k\varphi(t,x(t,X))V_k(t,X)\\ &\qquad + \varphi(t,x(t,X))(\operatorname{div} v)(t,x(t,X))\Bigg]J(t,X)\,dX\\ &= \int_{\Omega(t)}\partial_t\varphi(t,x) + (\operatorname{div}(\varphi v))(t,x)\,dx\,.\end{aligned}$$

□

As an exemplary application of the transport theorem we consider the balance equation for the mass. In this case $\varphi = \rho$ represents the mass density, it has the dimension $\mathcal{M}\mathcal{L}^{-d}$ where $\mathcal{M}$ is the dimension of mass. We assume that **mass conservation** holds, that is, the total amount of mass within $\Omega(t)$ is constant as a function of t,

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_{\Omega(t)}\rho(t,x)\,dx = 0\,. \tag{12.39}$$

The transport theorem implies that

$$\int_{\Omega(t)}\partial_t\rho(t,x) + (\operatorname{div}(\rho v))(t,x)\,dx = 0\,. \tag{12.40}$$

Let now $U \subset \Omega(t)$ be open, let us define

$$\Omega_U = X(t,U) = \{X : x(t,X) \in U\}\,.$$

Applying the transport theorem to Ω_U we obtain, since $U = \Omega_U(t)$,

$$\int_{U}\partial_t\rho(t,x) + (\operatorname{div}(\rho v))(t,x)\,dx = 0\,. \tag{12.41}$$

Since $U \subset \Omega(t)$ is an arbitrary open set, we can conclude (for example, if $\partial_t\rho$ and $\operatorname{div}(\rho v)$ are continuous)

$$\partial_t\rho + \operatorname{div}(\rho v) = 0 \tag{12.42}$$

in $\Omega(t)$. This equation is called the **continuity equation**.

Conservation laws, balance equations. If instead of mass density we also consider other quantities φ, additional mechanisms usually come into play. Propagation of heat, for instance, arises not only from transport but also from diffusion. Moreover, heat is generated when other types of energy (for example, mechanical or electromagnetic energy) are transformed into heat. Other mechanical and electromagnetic quantities are related, too (e.g. in piezoelectricity).

In order to analyze all those dependencies, one has to understand how to express the "flow" of a quantity through a given $(d-1)$-dimensional manifold M in $\mathbb{R}^d$.

We first discuss the concept of **flux** of a quantity. Let us assume that a quantity flows in $\mathbb{R}^d$ in a fixed direction with an intensity which is constant in space and time. The corresponding flux is described by a vector $q \in \mathbb{R}^d$. Its direction

$$\frac{q}{|q|}\,, \quad |q| = \|q\|_2 = \sqrt{\sum_{i=1}^{d} q_i^2}\,.$$

specifies the direction of the flux. Its length $|q|$ specifies the amount of the quantity (with dimension $\mathcal{G}$) which flows per unit of time through a unit cube of dimension $d-1$ perpendicular to the direction of the flux. This unit cube is a unit square in the case $d = 3$, or a unit interval in the case $d = 2$. Thus, the dimension of the flux is given by $\mathcal{G}\mathcal{T}^{-1}\mathcal{L}^{1-d}$, it is a density of a time rate called the **flux density**.

We consider such a flux of constant intensity, given by the vector $q \in \mathbb{R}^d$, through a unit cube $E \in \mathbb{R}^{d-1}$ perpendicular to the flux direction. We imagine a second cube W which is positioned at an angle α with respect to E in such a manner that the same "particles" flow through E and W. We then have

$$1 = \mathrm{vol}_{d-1}(E) = \mathrm{vol}_{d-1}(W) \cdot \cos\alpha\,.$$

The total flux per unit time through E is the same as through W, and is given by $|q|$. The flux densities, however, are different, they are equal to

$$\frac{|q|}{\mathrm{vol}_{d-1}(E)} = |q|\,, \quad \frac{|q|}{\mathrm{vol}_{d-1}(W)} = |q|\cos\alpha\,,$$

respectively. Let $n \in \mathbb{R}^d$ be the unit normal to W in the flux direction. Then the angle between q and n is equal to α, and

$$\langle q, n\rangle = |q|\cos\alpha \tag{12.43}$$

specifies the flux density with respect to W.

We now turn to the general situation, a flux through a given $(d-1)$ dimensional manifold M in $\mathbb{R}^d$. In the case $d = 3$, M is an (open or closed) surface, we denote its unit normal in the point $\xi \in M$ by $n(\xi)$. The flux is described by a function $q : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$. According to the explanations above, the function

$$\xi \mapsto \langle q(t, \xi), n(\xi)\rangle \tag{12.44}$$

represents the time-dependent flux density on M, it has the dimension $\mathcal{G}\mathcal{T}^{-1}\mathcal{L}^{1-d}$. The integral (it is a surface integral)

$$\int_M \langle q(t, \xi), n(\xi)\rangle \; dS(\xi) \tag{12.45}$$

represents the total flux through M of the considered quantity in form of a time rate, with dimension $\mathcal{G}\mathcal{T}^{-1}$. The total amount of the quantity flowing through M within the time interval $[t_1, t_2]$ is equal to

$$\int_{t_1}^{t_2} \int_M \langle q(t, \xi), n(\xi)\rangle \; dS(\xi)\, dt\,. \tag{12.46}$$

Let now $U \subset \mathbb{R}^d$ be open with boundary $M = \partial U$, assumed to be a two-dimensional manifold. By $n(\xi)$ we denote the exterior unit normal, that is the vector normal to ∂U in ξ pointing to the exterior of U. We can now formulate a **conservation law** by

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_U \varphi(t,x)\, dx = - \int_{\partial U} \langle q(t,\xi), n(\xi) \rangle \; dS(\xi)\,. \tag{12.47}$$

This means that the total amount within U of the quantity considered changes only by flow through the boundary of U. The minus sign arises because the scalar product $\langle q, n\rangle$ is negative for inflow and positive for outflow. Equation (12.47) is also called the **integral form** of the conservation law.

The divergence theorem states that

$$\int_{\partial U} \langle q(t,\xi), n(\xi) \rangle \; dS(\xi) = \int_U \operatorname{div} q(t,x)\, dx\,. \tag{12.48}$$

Therefore, for all t

$$\int_U \partial_t \varphi(t,x) + \operatorname{div} q(t,x)\, dx = 0\,. \tag{12.49}$$

If the conservation law (12.47) holds for "arbitrary domains" U, we may deduce from (12.47) the **differential form** of the conservation law,

$$\partial_t \varphi(t,x) + \operatorname{div} q(t,x) = 0\,, \quad \text{for all } x, t. \tag{12.50}$$

In the one-dimensional case ($d = 1$) we have obtained this equation already in (12.12).

When considering the flux of a quantity, one distinguishes between its **convective** and **nonconvective** part. The convective part refers to the situation, discussed in the subsection on kinematics above, where the quantity φ is attached to particles moving with velocity v. In (12.41), where $\varphi = \rho$, we have deduced from the transport theorem that

$$\int_U (\partial_t \varphi + \operatorname{div}(\varphi v))(t,x)\, dx = 0$$

holds for all U. The convective part q_K of the flux is therefore given by (compare with (12.49))

$$q_K = \varphi v\,. \tag{12.51}$$

An example for the nonconvective part q_{NK} is provided by the mechanism of diffusion. In the simplest case we have

$$q_{NK} = -\lambda \nabla \varphi\,, \quad \lambda > 0\,. \tag{12.52}$$

This form is attained by Fourier's law (for the heat flux) and by Fick's law (for a substance dissolved in another substance, like a pollutant in water). Let us consider a situation where no convection is present, thus $v = 0$. From (12.50) we obtain

$$\partial_t \varphi - \operatorname{div}(\lambda \nabla \varphi) = 0\,.$$

If λ is constant, we may move λ in front of the divergence and obtain, since $\operatorname{div}(\nabla\varphi) = \Delta\varphi$ the equation

$$\partial_t \varphi = \lambda \Delta \varphi\,. \tag{12.53}$$

This equation represents the simplest parabolic equation and, due to its origin, is called **diffusion equation** or **heat equation**.

In addition to flowing through the boundary, the total amount of a quantity within a region may also change by supply or drain in the interior. Examples are gravity (which acts on all "mass particles" of a solid body), chemical reactions or radiation (for example, heating by microwaves). When their time rate of change is described by a time-dependent density function z, the total rate of change within a volume U due to this action becomes

$$\int_U z(t,x)\,dx\,. \tag{12.54}$$

Considering all those terms we obtain a **general balance equation** in integral form,

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_U \varphi\,dx = -\int_{\partial U}\langle \varphi v, n\rangle\;dS - \int_{\partial U}\langle q_{NK}, n\rangle\;dS + \int_U z\,dx\,. \tag{12.55}$$

The functions appearing in (12.55) depend on t, x, or ξ respectively. Using the divergence theorem, we get

$$\int_U \partial_t\varphi\,dx = -\int_U \operatorname{div}\left(\varphi v + q_{NK}\right)dx + \int_U z\,dx\,. \tag{12.56}$$

As above, this yields the differential form

$$\partial_t\varphi + \operatorname{div}\left(\varphi v + q_{NK}\right) = z\,. \tag{12.57}$$

In (12.55) – (12.57), the convective part of the flux is represented by the term $\operatorname{div}(\varphi q)$. In many applications, one speaks of "flux" meaning the nonconvective part of the flux (this does not make a difference when no convection is present) and writes q instead of q_{NK}.

So far, the quantity φ has been a **volume density** with dimension $\mathcal{G}\mathcal{L}^{-d}$. Its total amount within a volume U equals

$$\int_U \varphi(t,x)\,dx\,. \tag{12.58}$$

However, it is often more convenient to deal with quantities related to mass, these have the dimension $\mathcal{G}\mathcal{M}^{-1}$. A quantity ψ related to mass is connected to a volumetric quantity φ of the same type by

$$\varphi = \rho\psi\,. \tag{12.59}$$

Here, ρ is the mass density as before with dimension $\mathcal{M}\mathcal{L}^{-d}$. The general balance equation (12.55) then becomes

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_U \rho\psi\,dx = -\int_{\partial U}\langle \rho\psi v, n\rangle\;dS - \int_{\partial U}\langle q_{NK}, n\rangle\;dS + \int_U \rho\zeta\,dx\,, \tag{12.60}$$

in integral form; one then writes the rate of the internal supply in the form $z = \rho\zeta$. The differential formulation is

$$\partial_t(\rho\psi) + \operatorname{div}\left(\rho\psi v + q_{NK}\right) = \rho\zeta\,. \tag{12.61}$$

A density related to mass is also called **specific density**.

We now investigate several special situations and exhibit the equations which result from the general balance equation.

We have already considered the mass balance. Here, $\psi = 1$ resp. $\varphi = \rho$, $q = 0$ and $\zeta = 0$. Next, we mention a so-called **multicomponent system**. Here we have J different substances, for example chemical compounds. In the continuum model these are characterized by their mass-related **concentrations**. $c_j : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, $1 \le j \le J$, taking values in $[0, 1]$. The function c_j specifies the relative proportion (or fraction) of the j-th component as a function of (t, x). The total amount of the j-th component within the volume U is thus equal to

$$\int_U (\rho_j c_j)(t, x)\, dx\,.$$

For each component we obtain a balance equation. We have $\psi = c_j$ and therefore, in the differential formulation,

$$\partial_t(\rho_j c_j) + \operatorname{div}(\rho_j c_j v + q_j) = \rho_j \zeta_j\,, \quad 1 \le j \le J\,. \tag{12.62}$$

Moreover it must be the case that

$$\sum_{j=1}^{J} c_j = 1\,. \tag{12.63}$$

If the components interact, for example due to a chemical reaction, additional terms appear on the right side of (12.62).

Momentum balance. Here $\varphi = \rho v$ and $\psi = v$, thus we deal with vector-valued quantities. The expression

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_U \rho v\, dx \tag{12.64}$$

represents the rate of change of the total momentum within the volume U. For $d = 3$, the general balance equation (12.55) becomes a system of 3 equations involving the 4 scalar quantities ρ and $\psi_i = v_i$, $i = 1, 2, 3$. Gravity induces a momentum influx within U, thus (on the surface of the earth)

$$\zeta = \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix}, \quad g = 9.8067 \frac{\mathrm{m}}{\mathrm{sec}^2}\,. \tag{12.65}$$

The first two space coordinates represent a horizontal plane, the third the vertical direction.

The nonconvective momentum flux arises from transmission by forces in the form of stresses (in particular, pressure forces). These forces are described by the stress tensor $\sigma(t, x) \in \mathbb{R}^{(3,3)}$. The function

$$x \mapsto \sigma(t, x) n(x)\,, \quad (\sigma n)_i = \sum_{j=1}^{3} \sigma_{ij} n_j\,,$$

represents the surface density of the force vector acting on a surface with normal n. The nonconvective flux thus becomes

$$q_{NK,i} = -\sigma_i\,,$$

where σ_i denotes the i-th row of the stress tensor σ. The special case of pure pressure corresponds to

$$\sigma = -pI\,, \quad \sigma_{ij} = -p\delta_{ij}\,, \tag{12.66}$$

where $p(t,x)$ is a scalar quantity (the pressure). The momentum balance equation thus has the form

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_U \rho v_i\,dx = -\int_{\partial U} \langle \rho v_i v, n\rangle\; dS + \int_{\partial U} \langle \sigma_i, n\rangle\; dS + \int_U \rho\zeta_i\,dx\,, \quad 1\le i\le 3\,. \tag{12.67}$$

The differential formulation becomes

$$\partial_t(\rho v_i) + \mathrm{div}\,(\rho v_i v - \sigma_i) = \rho\zeta_i\,, \quad 1\le i\le 3\,, \tag{12.68}$$

or, in the case of pure pressure,

$$\partial_t(\rho v_i) + \mathrm{div}\,(\rho v_i v) + \partial_i p = \rho\zeta_i\,, \quad 1\le i\le 3\,. \tag{12.69}$$

We do not discuss the balance equation for the angular momentum; we only remark that from it one derives that the stress tensor is symmetric, that is $\sigma_{ij} = \sigma_{ji}$ holds for all i,j.

Incompressible flow. Let us consider a flow of an incompressible fluid at constant temperature. Then we may assume ρ to be constant. The mass balance $\partial_t\rho + \mathrm{div}\,(\rho v) = 0$ becomes

$$\mathrm{div}\, v = 0\,. \tag{12.70}$$

From the momentum balance (12.68) we then obtain

$$\rho\partial_t v_i + \rho\,\mathrm{div}\,(v_i v) - \mathrm{div}\,\sigma_i - \rho\zeta_i = 0\,, \quad 1\le i\le 3\,. \tag{12.71}$$

Using the product rule we get

$$\mathrm{div}\,(v_i v) = v_i\,\mathrm{div}\,(v) + \langle\nabla v_i, v\rangle = \langle\nabla v_i, v\rangle = \sum_{j=1}^3 v_j\partial_j v_i\,. \tag{12.72}$$

Combining this with (12.68), we obtain

$$\partial_t v_i + \sum_{j=1}^3 v_j\partial_j v_i - \frac{1}{\rho}\mathrm{div}\,\sigma_i = \zeta_i\,. \tag{12.73}$$

Incompressible flow, pure pressure. This applies, for example, to water flow at constant temperature. As above, we have $\mathrm{div}\,\sigma_i = -\partial_i p$. We can write (12.73) in vector form as

$$\partial_t v + (v\cdot\nabla)v = -\frac{1}{\rho}\nabla p + \zeta\,. \tag{12.74}$$

Here, one uses the notation

$$(v\cdot\nabla)v = \big(\sum_{j=1}^3 v_j\partial_j\big)v\,. \tag{12.75}$$

Let us consider the even more special case when the flow is stationary, that is, v and p depend only on x but not on t. We also assume that ζ arises from gravity according to (12.65). We define the scalar quantity

$$\varphi(x) = \frac{1}{2}\rho\,\langle v(x), v(x)\rangle + p(x) + \rho g x_3\,. \tag{12.76}$$

We have

$$\partial_i\varphi(x) = \rho\,\langle v(x), \partial_i v(x)\rangle + \partial_i p(x) + \rho g\delta_{i3}\,. \qquad \text{(Kronecker delta)} \tag{12.77}$$

Since

$$\sum_{i=1}^{3} v_i\,\langle v, \partial_i v\rangle = \sum_{i,j=1}^{3} v_i v_j \partial_i v_j = \sum_{i,j=1}^{3} v_i v_j \partial_j v_i = \sum_{i=1}^{3} v_i \sum_{j=1}^{3} v_j \partial_j v_i\,,$$

we obtain, using (12.71) and (12.72)

$$\langle v, \nabla\varphi\rangle = \sum_{i=1}^{3} v_i\partial_i\varphi = \sum_{i=1}^{3} v_i\Big(\rho\sum_{j=1}^{3} v_j\partial_j v_i + \partial_i p - \rho\zeta_i\Big) = 0\,. \tag{12.78}$$

The expression $\langle v(x), \nabla\varphi(x)\rangle$ is nothing else than the material derivative of φ (note that φ does not depend on t). Therefore,

$$\frac{1}{2}\rho\|v(x)\|^2 + p(x) + gx_3 = \text{const} \tag{12.79}$$

holds along streamlines of the flow (that is, curves along which the material particles move). This result, or equivalently equation (12.78), is called the **Bernoulli equation**. It means that the pressure is small where the velocity is big.

Incompressible viscous flow. This describes the case with internal friction, that is, adjacent particles with different velocity exert forces on each other. The stress tensor then contains other terms in addition to the pressure term. The simplest case arises when the addition term is proportional to the "velocity gradient" Dv, the Jacobian of v with respect to x, and when

$$\sigma = 2\mu\frac{Dv + (Dv)^T}{2} - pI = \mu(Dv + (Dv)^T) - pI\,. \tag{12.80}$$

According to (12.71), one has to compute $\operatorname{div}\sigma_i$. We have

$$\operatorname{div}(Dv)_i = \sum_j \partial_j (Dv)_{ij} = \sum_j \partial_j\partial_j v_i$$

as well as, using incompressibility,

$$\operatorname{div}(Dv)^T_{ij} = \sum_j \partial_j\partial_i v_j = \partial_i \operatorname{div} v = 0\,,$$

so

$$\operatorname{div}\sigma_i = \mu\Delta v_i - \partial_i p\,, \quad 1 \le i \le 3\,. \tag{12.81}$$

Now (12.73) becomes

$$\partial_t v_i + \sum_{j=1}^{3} v_j \partial_j v_i = \frac{\mu}{\rho}\Delta v_i - \frac{1}{\rho}\partial_i p + \zeta_i\,, \quad 1 \le i \le 3\,. \tag{12.82}$$

These are the **Navier-Stokes equations**. In vector form they are written as

$$\partial_t v + (v\cdot\nabla)v = \frac{\mu}{\rho}\Delta v - \frac{1}{\rho}\nabla p + \zeta\,. \tag{12.83}$$

Energy balance. The energy balance is a scalar equation. Here, ψ is the specific (that is, mass-related) energy density

$$\psi = u + \frac{1}{2}|v|^2\,, \tag{12.84}$$

where u stands for the specific density of the internal energy and $\frac{1}{2}|v|^2$ represents the specific density of the kinetic energy. The rate of change of the total energy within the volume U is given by

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_U \rho\left(u + \frac{1}{2}|v|^2\right)\,dx\,. \tag{12.85}$$

The internal energy may for example include

- the kinetic energy of the fluctuation of atoms and molecules (measured on the macroscopic scale as temperature),
- the potential energy between atoms or molecules depending on their distance from each other (for example, the elastic energy which accompanies the deformation of solid bodies),
- the chemical bond energy,
- the nuclear energy

The nonconvective energy flux consists of the mechanical power (work per time) and the heat flux,

$$q_{NK} = -\sigma v + q_W\,. \tag{12.86}$$

The energy inflow consists of the power due to gravity and of the absorbed specific heat radiation ζ_W,

$$\zeta = -g v_3 + \zeta_W\,. \tag{12.87}$$

The energy balance becomes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{dt}}\int_U \rho\left(u + \frac{1}{2}|v|^2\right)\,dx &= -\int_{\partial U} \rho(u + \frac{1}{2}|v|^2)\,\langle v, n\rangle\,dS \\ &\quad + \int_{\partial U}\langle \sigma v, n\rangle\,dS - \int_{\partial U}\langle q_W, n\rangle\,dS + \int_U \rho\zeta\,dx\,.\end{aligned} \tag{12.88}$$

This balance is called the **first law of thermodynamics**. After several transformations one obtains the partial differential equation

$$\partial_t(\rho u) + \operatorname{div}(\rho u v + q_W) = \sigma : Dv + \rho\zeta\,, \tag{12.89}$$

where

$$\sigma : Dv = \sum_{i,j=1}^{3} \sigma_{ij}\, \partial_j v_i \,.$$