Partial differential equations

Nathanael Skrepek

July 26, 2024

Contents

Introduction **v**

1 Motivation **1**

2 Distributions **3**
- 2.1 Fundamental lemma of calculus of variation . . . 3
- 2.2 Test functions . . . 8
- 2.3 Fundamental solution . . . 16
- 2.4 Partition of unity . . . 20
- 2.5 Stokes' theorem . . . 21

3 Sobolev spaces **25**
- 3.1 Weak derivative and Sobolev spaces . . . 25
- 3.2 Weak formulation . . . 30
- 3.3 The negative Sobolev space . . . 32
- 3.4 Galerkin method . . . 35

4 Strongly continuous semigroups **39**
- 4.1 Strongly continuous semigroups . . . 40
- 4.2 Hille–Yosida theorem . . . 48
- 4.3 Contraction semigroups . . . 51
- 4.4 Mild solutions . . . 55

A Locally convex topology **57**
- A.1 Basics . . . 57

B Lipschitz domains and boundaries **59**
- B.1 Definition . . . 59
- B.2 Outer normal vector . . . 61
- B.3 Independence of the charts . . . 63
- B.4 Some auxiliary lemmas . . . 65

C Linear relations **67**
C.1 Basics 67
C.2 Adjoint linear relations 69

Bibliography **75**

Introduction

In this lecture we will develop a functional analytic approach to linear partial differential equations. We will introduce three approaches to linear partial differential equations:

- The *distributional formulation* that leads to *funamental solutions*.
- The *weak formulation* that leads to *Sobolev spaces* and the Lax–Milgram theorem for existence and uniqueness of solutions.
- The notion of *strongly continuous semigroups* for dynamic (time dependent) systems, which gives existence and uniqueness results in form of the Hille–Yosida theorem.

Our focus will be on the semigroup theory, which opens the field of the analysis of a lot of physcially motivated systems. Questions that arises in this context are long-time behaviour, stability, controllability, observability, etc. These pursuing questions are unfortunately beyond the scope of this lecture, but we will present the gateway drug.

For fundamentals in analysis in German we want to mention the very good references [8, 9].

Chapter 1

Motivation

Let us regard the following setting. We have a matrix $A \in \mathbb{C}^{m\times m}$, an initial state $x_0 \in \mathbb{C}^m$ and the following ordinary differential equation

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} x(t) &= Ax(t), \quad t \geq 0, \\ x(0) &= x_0. \end{aligned}$$

Here the space $\mathbb{C}^m$ is called the *state space* and elements of the state space are called *states*. The unknown is the function $x\colon \mathbb{R}_+ \to \mathbb{C}^m$. The solution to this problem is given by the matrix exponential

$$x(t) = \mathrm{e}^{tA} x_0.$$

Example 1.0.1. Let us regard the following system that is derived from a mass spring equation, where $m > 0$ is the mass and $k > 0$ is the spring constant:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \underbrace{\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}}_{=x(t)} &= \underbrace{\begin{bmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}}_{=x(t)}, \quad t \geq 0, \\ x(0) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \end{aligned}$$

In order to really calculate the matrix exponential we have to find out the eigenvalues and eigenvectors of the matrix A. Solving $\det(A - \lambda) = 0$ gives $\lambda = \pm\mathrm{i}\sqrt{\frac{k}{m}}$. The corresponding eigenvectors are given by

$$v_1 = \begin{bmatrix} -\mathrm{i} \\ \sqrt{km} \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} \mathrm{i} \\ \sqrt{km} \end{bmatrix}.$$

We denote the matrix composed of the normalized eigenvectors by V, i.e.,

$V = \frac{1}{\sqrt{1+km}}\left[\begin{smallmatrix} v_1 & v_2 \end{smallmatrix}\right]$. Hence, we can diagonalize A as

$$A = V \underbrace{\begin{bmatrix} \mathrm{i}\sqrt{\frac{k}{m}} & 0 \\ 0 & -\mathrm{i}\sqrt{\frac{k}{m}} \end{bmatrix}}_{=:D} V^{-1}$$

the matrix exponential is given by

$$\begin{aligned} \mathrm{e}^{tA} &= \mathrm{e}^{tVDV^{-1}} = \sum_{k=0}^{\infty} \frac{t^k (VDV^{-1})^k}{k!} = V \sum_{k=0}^{\infty} \frac{t^k D^k}{k!} V^{-1} \\ &= V \begin{bmatrix} \mathrm{e}^{t\mathrm{i}\sqrt{\frac{k}{m}}} & 0 \\ 0 & \mathrm{e}^{-t\mathrm{i}\sqrt{\frac{k}{m}}} \end{bmatrix} V^{-1}. \end{aligned}$$

Thus, the solution of the differential equation is

$$x(t) = V \begin{bmatrix} \mathrm{e}^{t\mathrm{i}\sqrt{\frac{k}{m}}} & 0 \\ 0 & \mathrm{e}^{-t\mathrm{i}\sqrt{\frac{k}{m}}} \end{bmatrix} V^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$ ✧

However, we want to solve an equation like the heat equation

$$\begin{aligned} \frac{\partial}{\partial t} x(t,\zeta) &= \frac{\partial^2}{\partial \zeta^2} x(t,\zeta) \\ x(0,\zeta) &= x_0(\zeta) \end{aligned} \tag{1.1}$$

(with possible additional boundary conditions). In order to mimic the previous approach we fix a function space, e.g., $\mathrm{L}^p(a,b)$ or $\mathrm{C}^2(a,b)$. In particular we will work most of the time in $\mathrm{L}^2(a,b)$, because this is even a Hilbert space. Now we make the following change of perspective and regard the function

$$x\colon \mathbb{R}_+ \times (a,b) \to \mathbb{C}$$

as

$$x\colon \mathbb{R}_+ \to \mathrm{L}^2(a,b) \quad \text{by} \quad x(t) := x(t,\cdot),$$

i.e., $x(t)(\zeta) = x(t,\zeta)$. This means that instead of regarding the function in time and space simultaneously we regard the function "only" in time, but for every point in time the function does not map into $\mathbb{C}$, but in $\mathrm{L}^2(a,b)$, i.e., $x(t)$ is again a function. We can interprete this as looking for every point in time at all positions simultaneously.

By this reinterpretation the differential operator $\frac{\partial^2}{\partial \zeta^2}$ is a linear operator on $\mathrm{L}^2(a,b)$, which we will denote as A. Consequently, we rewrite (1.1) as

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} x(t) &= Ax(t), \\ x(t) &= x_0. \end{aligned}$$

Hence, we would like to construct something like the matrix exponential e^{tA} for a linear operator A.

Chapter 2

Distributions

We will introduce a concept that allows us to regard derivatives even for non-smooth functions. Distributions are very carefully covered in [7].

2.1 Fundamental lemma of calculus of variation

The fundamental lemma of calculus of variation is important to justify that our generalized notion of derivatives is well-defined.

Recall the definition of the convolution

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,\mathrm{d}\lambda(y)\,.$$

Proposition 2.1.1. *Let $f \in \mathrm{L}^p(\mathbb{R}^d)$ for $p \in [1, +\infty)$ and $t \in \mathbb{R}^d$. Then we define f_y by $f_y(x) = f(x+y)$. For fixed $y \in \mathbb{R}^d$ the mapping $f \mapsto f_y$ is an isometric isomorphism. Moreover, the argument translation mapping*

$$T_f \colon \left\{ \begin{array}{ccc} \mathbb{R}^d & \to & \mathrm{L}^p(\mathbb{R}^d), \\ y & \mapsto & f_y, \end{array} \right.$$

is uniformly continuous, i.e., for all $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\|y_1 - y_2\| \leq \delta \quad \Rightarrow \quad \|T_f(y_1) - T_f(y_2)\| \leq \epsilon$$

(independent of y_1 and y_2).

Proof. Clearly, $f \mapsto f_y$ is linear. Moreover,

$$\begin{aligned} \|f_y\|_{\mathrm{L}^p}^p &= \int_{\mathbb{R}^d} |f_y(x)|^p\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} |f(x+y)|^p\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} |f(x)|^p\,\mathrm{d}\lambda(x) \\ &= \|f\|_{\mathrm{L}^p}^p, \end{aligned}$$

which implies that the mapping is also isometric. Its inverse is given by $f \mapsto f_{-y}$.

Recall that set of step functions of the form

$$\sum_{i=1}^{m} \alpha_i \mathbb{1}_{R_i}, \tag{2.1}$$

where $\alpha_i \in \mathbb{C}$ and every R_i is a rectangular cuboid of the form $\prod_{j=1}^{d}(a_j, b_j]$, is dense in $\mathrm{L}^p(\mathbb{R}^d)$.

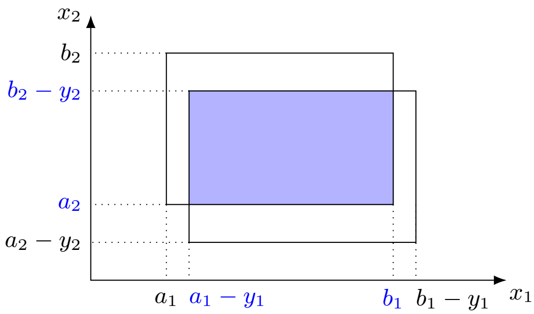

Figure 2.1: Intersection of R and $R - y$

For $R = \prod_{j=1}^{d}(a_j, b_j]$ and $f = \mathbb{1}_R$ we have $f_y(x) = \mathbb{1}_R(x+y) = \mathbb{1}_{R-y}(x)$, where $R - y = \prod_{j=1}^{d}(a_j - y_j, b_j - y_j]$. Hence,

$$\begin{aligned} \|f - T_f(y)\|_{\mathrm{L}^p}^p = \|f - f_y\|_{\mathrm{L}^p}^p &= \int |\mathbb{1}_R - \mathbb{1}_{R-y}| \,\mathrm{d}\lambda \\ &= \lambda\big(R \setminus (R-y) \,\dot{\cup}\, (R-y) \setminus R\big) \\ &= \lambda(R) - \lambda(R \cap (R-y)) + \lambda(R-y) - \lambda((R-t) \cap R). \end{aligned} \tag{2.2}$$

Recall that the Lebesgue measure λ is translation invariant. Therefore, $\lambda(R) = \lambda(R - y)$. Moreover,

$$\begin{aligned} R \cap (R-y) &= \prod_{j=1}^{d}(a_j, b_j] \cap \prod_{j=1}^{d}(a_j - y_j, b_j - y_j] \\ &= \prod_{j=1}^{d}\big(a_j - \min(0, y_j), b_j - \max(0, y_j)\big] \end{aligned}$$

and $\lambda(R \cap (R-y)) = \prod_{j=1}^{d}(b_j - a_j - |y_j|)$ for $|y_j| \le (b_j - a_j)$. Hence, we continue (2.2):

$$\begin{aligned} \|f - T_f(y)\|_{\mathrm{L}^p}^p &= 2\lambda(R) - 2\lambda(R \cap (R-y)) \\ &= 2\prod_{j=1}^{d}(b_j - a_j) - 2\prod_{j=1}^{d}(b_j - a_j - |y_j|). \end{aligned}$$

Consequently $T_f(y)$ converges to f for $y \to 0$. If f is a step function of the form $f = \sum_{i=1}^m \alpha_i \mathbb{1}_{R_i}$, then $T_f = \sum_{i=1}^m \alpha T_{\mathbb{1}_{R_i}}$, which implies that also $T_f(y)$ converges to f for $y \to 0$.

For $f \in \mathrm{L}^p(\mathbb{R}^d)$ arbitrary, there exists a sequence $(s_l)_{l\in\mathbb{N}}$ of step functions (of the form (2.1)) that converges to f. For an arbitrary $\epsilon > 0$ there exists an $l_0 \in \mathbb{N}$ such that $\|f - s_l\| \le \epsilon$ for all $l \ge l_0$. For fixed $l \ge l_0$ there exists a $y_0 \in \mathbb{R}^d$ such that $\|s_l - T_{s_l}(y)\| \le \epsilon$ for all $\|y\|_\infty \le \|y_0\|_\infty$. Consequently,

$$\begin{aligned}\|f - T_f(y)\| &= \|f - s_l + s_l - T_{s_l}(y) + T_{s_l}(y) - T_f(y)\| \\ &\le \|f - s_l\| + \|s_l - (s_l)_y\| + \underbrace{\|(s_l)_y - f_y\|}_{=\|s_l - f\|} \le 3\epsilon,\end{aligned}$$

which implies the continuity of T_f at 0. Note that $\|T_f(y_1) - T_f(y_2)\| = \|f - T_f(y_2 - y_1)\|$. Hence, the uniform continuity follows from the continuity in 0. ❑

The next example is very important. We will use the function that we will construct often subsequently. Moreover, it gives an explicit example of a C^∞ function with compact support.

Example 2.1.2. Let us regard the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} \exp(\frac{1}{-x}), & x > 0, \\ 0, & x \le 0. \end{cases}$$

Clearly, f is arbitrarily often differentiable in every $x \neq 0$. Hence, in order to show that $f \in \mathrm{C}^\infty(\mathbb{R})$ we only have to check the differentiability at 0. Recall that for every polynomial p we have $\lim_{x\to+\infty} \exp(-x)p(x) = 0$ and therefore $\lim_{x\to 0+} \exp(-\frac{1}{x})p(\frac{1}{x}) = 0$.

For $x > 0$ we have $\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \exp(\frac{1}{-x})\frac{1}{x^2}$. By induction and the product rule for derivatives we conclude $\frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x) = \exp(\frac{1}{-x})p_n(\frac{1}{x})$ for a suitable polynomial p_n. Consequently,

$$\lim_{x\to 0+} \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x) = 0 = \lim_{x\to 0-} \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x),$$

which implies that $f \in \mathrm{C}^\infty(\mathbb{R})$.

We define the function $\hat\rho\colon \mathbb{R}^d \to \mathbb{R}$ by

$$\hat\rho(x) = \begin{cases} \exp(\frac{1}{\|x\|_2^2 - 1}), & \|x\| < 1, \\ 0, & \text{else.} \end{cases}$$

We have the identity $\hat\rho(x) = f(1 - \|x\|_2^2)$. Since f and $x \mapsto 1 - \|x\|_2^2$ are C^∞, we conclude $\hat\rho \in \mathrm{C}^\infty(\mathbb{R}^d)$. Moreover, $\operatorname{supp}\hat\rho = \overline{\mathrm{B}}_1(0)$, which implies $\hat\rho \in \mathring{\mathrm{C}}^\infty(\mathbb{R}^d)$. ✧

Lemma 2.1.3. *Let $(\rho_n)_{n\in\mathbb{N}}$ be sequence in $\mathrm{L}^\infty(\mathbb{R}^d)$ with compact support $\operatorname{supp}\rho_n$ such that $\rho_n \geq 0$, $\|\rho_n\|_{\mathrm{L}^1} = 1$ and $\sup_{x\in\operatorname{supp}\rho_n}\|x\|_2 \to 0$ for $n \to \infty$. Then for every $f \in \mathrm{L}^p(\mathbb{R}^d)$ for $p \in [1,\infty)$ we have $f * \rho_n \in \mathrm{L}^\infty(\mathbb{R}^d)\cap\mathrm{L}^p(\mathbb{R}^d)$, where $\|f*\rho_n\|_{\mathrm{L}^p} \leq \|f\|_{\mathrm{L}^p}$ and*

$$\lim_{n\to\infty}\|f - f*\rho_n\|_{\mathrm{L}^p} = 0.$$

Proof. Note that $\rho_n \in \mathrm{L}^q(\mathbb{R}^d)$ for all $q \in [1,\infty]$. Hence, $f*\rho_n \in \mathrm{L}^\infty(\mathbb{R}^d)$ follows immediately from Hölder's inequality:

$$|(f*\rho_n)(x)| = \left|\int_{\mathbb{R}^d} f(y)\rho(x-y)\,\mathrm{d}y\right| \leq \|f\|_{\mathrm{L}^p}\|\rho\|_{\mathrm{L}^q}.$$

Let $\frac{1}{p}+\frac{1}{q} = 1$. Then

$$\begin{aligned}
&\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d} f(y)\rho_n(x-y)\,\mathrm{d}y\right|^p \mathrm{d}x \\
&\qquad\leq \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|f(y)|\rho_n(x-y)^{\frac{1}{p}}\rho(x-y)^{\frac{1}{q}}\,\mathrm{d}y\right)^p \mathrm{d}x \\
&\qquad\overset{\text{Hölder}}{\leq} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(y)|^p\rho_n(x-y)\,\mathrm{d}y\Bigg(\underbrace{\int\rho(x-y)\,\mathrm{d}y}_{=1}\Bigg)^{\frac{p}{q}}\mathrm{d}x \\
&\qquad\overset{\text{Fubini}}{=} \int_{\mathbb{R}^d}|f(y)|^p\underbrace{\int_{\mathbb{R}^d}\rho_n(x-y)\,\mathrm{d}x}_{=1}\,\mathrm{d}y = \|f\|_{\mathrm{L}^p}^p
\end{aligned}$$

shows that $f*\rho_n \in \mathrm{L}^p(\mathbb{R}^d)$. Finally,

$$\begin{aligned}
\|f - f*\rho_n\|_{\mathrm{L}^p}^p &= \int_{\mathbb{R}^d}\left|f(x) - \int_{\mathbb{R}^d} f(x-y)\rho_n(y)\,\mathrm{d}y\right|^p\mathrm{d}x \\
&= \int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\big(f(x)-f(x-y)\big)\rho_n(y)\,\mathrm{d}y\right|^p\mathrm{d}x \\
&\leq \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\big|f(x)-f(x-y)\big|\rho_n(y)^{\frac{1}{p}}\rho_n(y)^{\frac{1}{q}}\,\mathrm{d}y\right)^p\mathrm{d}x \\
&\overset{\text{Hölder}}{\leq} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(x)-f(x-y)|^p\rho_n(y)\,\mathrm{d}y\,\mathrm{d}x \\
&\overset{\text{Fubini}}{=} \int_{\operatorname{supp}\rho_n}\rho_n(y)\int_{\mathbb{R}^d}|f(x)-f(x-y)|^p\,\mathrm{d}x\,\mathrm{d}y \\
&\leq \sup_{y\in\operatorname{supp}\rho_n}\|f - f_{-y}\|_{\mathrm{L}^p}^p
\end{aligned}$$

finishes the proof as f_{-y} converges uniformly to f for $y \to 0$. ❏

Remark 2.1.4. Note that such a sequence $(\rho_n)_{n\in\mathbb{N}}$ exists. Let us regard the function $\hat{\rho}$ from Example 2.1.2. Then we define $\rho := \frac{1}{\|\hat{\rho}\|_{\mathrm{L}^1}}\hat{\rho}$ and

$$\rho_\epsilon(x) := \epsilon^{-d}\rho\Big(\frac{x}{\epsilon}\Big).$$

Clearly, $\rho_\epsilon(x) \geq 0$, it is also not hard to see that $\operatorname{supp}\rho_\epsilon = \overline{\mathrm{B}}_\epsilon(0)$ and $\|\rho_\epsilon\|_{\mathrm{L}^1} = 1$. Hence, $(\rho_{\frac{1}{n}})_{n\in\mathbb{N}}$ is such a sequence. The function ρ is called *mollifier.* ✧

Note that for $f \in \mathrm{L}^p(\mathbb{R}^d)$ we have

$$\frac{\partial}{\partial x_i}(f * \rho_\epsilon) = f * \frac{\partial}{\partial x_i}\rho_\epsilon$$

by dominated convergence. Hence, if ρ is the mollifier from Remark 2.1.4, then $f * \rho_\epsilon$ is in $\mathrm{C}^\infty(\mathbb{R}^d)$. Moreover, for $x \notin \operatorname{supp} f + \mathrm{B}_\epsilon(0)$, i.e., $\|x - y\| > \epsilon$ for all $y \in \operatorname{supp} f$, we have

$$(f * \rho_\epsilon)(x) = \int_{\mathbb{R}^d} f(y)\rho_\epsilon(x-y)\,\mathrm{d}y = \int_{\operatorname{supp} f} f(y)\rho_\epsilon(x-y)\,\mathrm{d}y = 0.$$

Corollary 2.1.5. *Let $\Omega \subseteq \mathbb{R}^d$ open. Then $\mathring{\mathrm{C}}^\infty(\Omega)$ is dense in $\mathrm{L}^p(\Omega)$ for all $p \in [1, +\infty)$.*

Proof. Every $f \in \mathrm{L}^p(\Omega)$ can be extended to $f \in \mathrm{L}^p(\mathbb{R}^d)$ by defining $f(x) = 0$ for $x \notin \Omega$. The compact sets

$$K_n := \overline{\mathrm{B}}_n(0) \setminus \mathrm{B}_{\frac{1}{n}}(\Omega^\complement) = \left\{ x \in \mathbb{R}^d \,\middle|\, \operatorname{dist}(x, \Omega^\complement) \geq \frac{1}{n}, \|x\| \leq n \right\}$$

cover Ω. Hence $f\mathbb{1}_{K_n}$ converges to f. Moreover, for the mollifier from Remark 2.1.4 the function $f\mathbb{1}_{K_n} * \rho_{\frac{1}{k}}$ is $\mathrm{C}^\infty(\mathbb{R}^d)$ and has compact support in Ω for $k \geq 2n$ and converges to $f\mathbb{1}_{K_n}$ for $k \to \infty$. ❑

Lemma 2.1.6 (Fundamental lemma of calculus of variations)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be open. If $g \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ is such that*

$$\int_\Omega g\phi\,\mathrm{d}\lambda = 0 \quad \textit{for all} \quad \phi \in \mathring{\mathrm{C}}^\infty(\Omega),$$

then $g = 0$ λ-almost everywhere.

Proof. For $\psi \in \mathring{\mathrm{C}}^\infty(\Omega)$ the product $g\psi$ is integrable. Moreover, $\int_\Omega g\psi\phi\,\mathrm{d}\lambda = 0$ even for every $\phi \in \mathrm{C}^\infty(\mathbb{R}^d)$ ($\psi\phi \in \mathring{\mathrm{C}}^\infty(\Omega)$). We extend $g\psi$ to $\mathbb{R}^d$ by $g\psi(x) = 0$ for all $x \in \mathbb{R}^d \setminus \Omega$. Let ρ be the mollifier from Remark 2.1.4 that establishes the sequence $(\rho_{\frac{1}{n}})_{n\in\mathbb{N}}$. Then by Lemma 2.1.3 we have

$$\lim_{n\to\infty} (g\psi) * \rho_{\frac{1}{n}} = g\psi \quad \text{in} \quad \mathrm{L}^1(\mathbb{R}^d).$$

Since $\rho_{\frac{1}{n}} \in \mathring{\mathrm{C}}^\infty(\mathbb{R}^d)$, we have

$$\big((g\psi) * \rho_{\frac{1}{n}}\big)(x) = \int_\Omega g(y) \big[\underbrace{\psi(y)\rho_{\frac{1}{n}}(x-y)}_{\in \mathring{\mathrm{C}}^\infty(\Omega)}\big] \,\mathrm{d}y = 0,$$

and consequently $g\psi = 0$ almost everywhere. Since for every $x \in \Omega$ the function $\psi_x := \rho_{\frac{1}{n}}(\cdot - x)$ is strictly positive on the neighborhood $\mathrm{B}_{\frac{1}{2n}}(x)$, we conclude from $g\psi_x = 0$ that $g = 0$. ❑

2.2 Test functions

Let Ω be an open subset of $\mathbb{R}^d$. Before we can introduce distributions we have to introduce the space of *test functions* on Ω. We define

$$\mathcal{D}(\Omega) := \mathring{\mathrm{C}}^\infty(\Omega) = \{\phi \in \mathrm{C}^\infty(\Omega) \,|\, \operatorname{supp}\phi \text{ is compact in } \Omega\}.$$

We use the notation $\mathcal{D}(\Omega)$ instead of $\mathring{\mathrm{C}}^\infty(\Omega)$, because we will endow this space with a special topology. Note that $\mathcal{D}(\Omega) = \mathring{\mathrm{C}}^\infty(\Omega)$ is not empty as translations and rescaled versions of $\hat{\rho}$ from Example 2.1.2 are in $\mathcal{D}(\Omega)$. Furthermore by Corollary 2.1.5 $\mathcal{D}(\Omega)$ is even dense in $\mathrm{L}^p(\Omega)$ for $p \in [1, +\infty)$.

For a multi-index $\alpha \in \mathbb{N}_0^d$ we define

$$|\alpha| := \sum_{i=1}^d \alpha_i \quad \text{and} \quad \mathrm{D}^\alpha\phi := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} \phi.$$

For $\Omega \subseteq \mathbb{R}^d$ open and $K \subseteq \Omega$ we define

$$\mathcal{D}_K(\Omega) := \{\phi \in \mathcal{D}(\Omega) \,|\, \operatorname{supp}\phi \subseteq K\} \subseteq \mathcal{D}(\Omega).$$

On $\mathcal{D}_K(\Omega)$ we can define the seminorms

$$p_\alpha(\phi) := \|\mathrm{D}^\alpha\phi\|_\infty = \sup_{x\in\Omega} |\mathrm{D}^\alpha\phi(x)|.$$

Note that p_0 is obviously a norm. It can be even shown that every p_α is a norm on $\mathcal{D}_K(\Omega)$. Hence, the family $(p_\alpha)_{\alpha\in\mathbb{N}_0^n}$ of seminorms is separating, i.e., for all $\phi \neq 0$ there exists a p_α such that $p_\alpha(\phi) \neq 0$. Therefore, we denote the locally convex topology on $\mathcal{D}_K(\Omega)$ that is induced by this family of seminorms by $\mathcal{T}_K$.

Alternative we can introduce the following norms

$$\|\phi\|_m := \sup_{|\alpha|\le m} \sup_{x\in\Omega} |\mathrm{D}^\alpha\phi(x)| = \sup_{|\alpha|\le m} p_\alpha(\phi), \quad m \in \mathbb{N}_0.$$

These norms induce the same topology on $\mathcal{D}_K(\Omega)$. Moreover, we have

$$\|\phi\|_0 \le \|\phi\|_1 \le \|\phi\|_2 \le \cdots \le \|\phi\|_m \le \|\phi\|_{m+1} \le \cdots .$$

Definition 2.2.1. A *Fréchet space* is a topological space $(X, \mathcal{T})$, whose topology is induced by a metric dist, which makes the metric space (X, dist) complete. ✧

Theorem 2.2.2. *The vector space $\mathcal{D}_K(\Omega)$ endowed with the previously defined topology $\mathcal{T}_K$ is a Fréchet space.*

Proof. Since $\|\cdot\|_n$ is a norm, we know that $(\phi, \psi) \mapsto \|\phi - \psi\|_n$ is a metric. Therefore,

$$\mathrm{dist}(\phi, \psi) := \sum_{n=0}^{\infty} 2^{-n} \frac{\|\phi - \psi\|_n}{1 + \|\phi - \psi\|_n}$$

is also a metric. Recall that the topology $\mathcal{T}_K$ is generated by the sets

$$V(\phi, \|\cdot\|_n, \epsilon) = \{\psi \in \mathcal{D}_K(\Omega) \mid \|\phi - \psi\|_n < \epsilon\}.$$

First we show that the ball $\mathrm{B}_{2^{-n}\frac{\epsilon}{1+\epsilon}}(\phi)$ (w.r.t. dist) is contained in $V(\phi, \|\cdot\|_n, \epsilon)$ for every $n \in \mathbb{N}_0$. For $\psi \in \mathrm{B}_{2^{-n}\frac{\epsilon}{1+\epsilon}}(\phi)$ we have

$$2^{-n} \frac{\|\psi - \phi\|_n}{1 + \|\psi - \phi\|_n} \leq \sum_{k=0}^{\infty} 2^{-k} \frac{\|\psi - \phi\|_k}{1 + \|\psi - \phi\|_k} < 2^{-n} \frac{\epsilon}{1 + \epsilon},$$

which immediately implies

$$\|\psi - \phi\|_n < \epsilon.$$

On the other hand for $\epsilon > 0$ we choose $n_0 \in \mathbb{N}_0$ such that $2^{-n_0} < \frac{\epsilon}{2}$, then $V(\phi, \|\cdot\|_n, \frac{\epsilon}{4})$ is contained in $\mathrm{B}_\epsilon(\phi)$ for $n \geq n_0$, because

$$\mathrm{dist}(\phi, \psi) \leq \sum_{k=0}^{n_0} 2^{-k} \frac{\|\phi - \psi\|_k}{1 + \|\phi - \psi\|_k} + \sum_{k=n_0+1}^{\infty} 2^{-k} \leq \frac{\epsilon}{4} \sum_{k=0}^{n_0} 2^{-k} + 2^{n_0} < \frac{\epsilon}{2} + \frac{\epsilon}{2}$$

for $\psi \in V(\phi, \|\cdot\|_n, \frac{\epsilon}{4})$. Hence, the neighborhoods in both topologies are the same and consequently the topologies coincide.

Let $(\phi_n)_{n\in\mathbb{N}}$ be a Cauchy sequence (w.r.t. dist). Then $(\phi_n)_{n\in\mathbb{N}}$ is a Cauchy sequence w.r.t. $\|\cdot\|_0 = \|\cdot\|_\infty$ and consequently $(\phi_n)_{n\in\mathbb{N}}$ converges w.r.t. $\|\cdot\|_0$ and its limit ϕ is continuous. Note that by the same argument also every $(\mathrm{D}^\alpha \phi_n)_{n\in\mathbb{N}}$ converges. In particular we denote the limit of $\frac{\partial}{\partial x_i}\phi_n$ by ψ_i and by the fundamental theorem of calculus

$$\begin{aligned}\phi(x + te_i) = \lim \phi_n(x + te_i) &= \lim \phi_n(x) + \int_0^t \frac{\partial}{\partial x_i}\phi_n(x + se_i)\,\mathrm{d}s \\ &= \phi(x) + \int_0^t \psi_i(x + se_i)\,\mathrm{d}s,\end{aligned}$$

where e_i is the unit vector that points in the i-th directions. Hence, $\psi_i = \frac{\partial}{\partial x_i}\phi$ and $\phi \in \mathrm{C}^1(\Omega)$. By induction we even have $\phi \in \mathrm{C}^\infty(\Omega)$ and since $\operatorname{supp}\phi_n \subseteq K$ we have also have $\operatorname{supp}\phi \subseteq K$. Thus, $\phi \in \mathcal{D}_K(\Omega)$. ❑

Remark 2.2.3. Note that a linear functional $\Lambda\colon \mathcal{D}_K(\Omega) \to \mathbb{C}$ is continuous w.r.t. $\mathcal{T}_K$, if and only if

$$\forall \epsilon > 0\ \exists \delta > 0, n \in \mathbb{N}_0 \quad \text{such that} \quad \Lambda\Big(V(0, \|\cdot\|_n, \delta)\Big) \subseteq \mathrm{B}_\epsilon(0).$$

This is equivalent to

$$\exists n \in \mathbb{N}_0, c > 0 \quad \text{such that} \quad |\Lambda(\phi)| \leq c\|\phi\|_n\ \forall \phi \in \mathcal{D}_K(\Omega). \qquad \diamond$$

Remark 2.2.4. A complex measure μ can be seen as an element of the dual space of $\mathcal{D}_K(\Omega)$ (representation theorem of Riesz-Markov-Kakutani) by

$$\Lambda_\mu(\phi) := \int_\Omega \phi \,\mathrm{d}\mu\,.$$

Moreover, also a locally integrable function f ($f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$) can be seen as an element of $\mathcal{D}'_K(\Omega)$ by

$$\Lambda_f(\phi) := \int_\Omega f \cdot \phi \,\mathrm{d}\lambda\,. \qquad \diamond$$

Note that for $f \in \mathrm{C}^1(\Omega)$ we have $\Lambda_f \in \mathcal{D}'_K(\Omega)$. Moreover, $\frac{\partial}{\partial x_i} f \in \mathrm{C}(\Omega)$ and consequently $\Lambda_{\frac{\partial}{\partial x_i} f} \in \mathcal{D}'_K(\Omega)$. Now the integration by parts formula gives

$$\Lambda_f\Big(\frac{\partial}{\partial x_i}\phi\Big) = \int_\Omega f \cdot \frac{\partial}{\partial x_i}\phi \,\mathrm{d}\lambda = -\int_\Omega \frac{\partial}{\partial x_i} f \cdot \phi \,\mathrm{d}\lambda = -\Lambda_{\frac{\partial}{\partial x_i} f}(\phi).$$

Inspired by the integration by parts formula we define the derivative for elements in $\mathcal{D}'_K(\Omega)$.

Definition 2.2.5. Let $\Lambda \in \mathcal{D}'_K(\Omega)$ and $\alpha \in \mathbb{N}_0^d$. Then we define

$$(\mathrm{D}^\alpha \Lambda)(\phi) := (-1)^{|\alpha|}\Lambda(\mathrm{D}^\alpha \phi). \qquad \diamond$$

Lemma 2.2.6. *Let μ be a complex measure on the Borel sets of a compact set K. Then μ is the derivative of a continuous function f on K in the distributional sense, i.e., there exists an $\alpha \in \mathbb{N}_0^d$ such that $\Lambda_\mu(\phi) = \mathrm{D}^\alpha \Lambda_f(\phi)$ for all $\phi \in \mathcal{D}_K(\Omega)$.*

Proof. For a compact $K \subseteq \Omega$ we find a rectangular cuboid $\prod_{i=1}^d (a_i, +\infty)$ that contains K. Hence, by the fundamental theorem of calculus every $\phi \in \mathcal{D}_K(\Omega)$ can be written as

$$\phi(x) = \int_{\mathbb{R}^d} \mathbb{1}_{\prod(a_i, x_i)}(y) \mathrm{D}^{\alpha^1}\phi(y)\,\mathrm{d}y,$$

where $\alpha^1 = (1, \dots, 1)$. This gives us the following identity

$$\begin{aligned}\Lambda_\mu(\phi) &= \int_K \phi(x)\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbb{1}_{\prod(a_i,x_i)}(y)\mathrm{D}^{\alpha^1}\phi(y)\,\mathrm{d}y\,\mathrm{d}\mu(x)\\ &= \int_{\mathbb{R}^d}\underbrace{\int_{\mathbb{R}^d} \mathbb{1}_{\prod(a_i,x_i)}(y)\,\mathrm{d}\mu(x)}_{=:F(y)}\,\mathrm{D}^{\alpha^1}\phi(y)\,\mathrm{d}y\,.\end{aligned}$$

Note that F is a bounded and measurable function. We can repeat the same argument for the measure $F(y)\,\mathrm{d}y$ and $\mathrm{D}^{\alpha^1}\phi$ and obtain

$$\Lambda_\mu = \int_{\mathbb{R}^d} f(z)\underbrace{\mathrm{D}^{\alpha^1}\mathrm{D}^{\alpha^1}}_{=\mathrm{D}^{\alpha^2}}\phi(z)\,\mathrm{d}z,$$

where $f(z) := \int_{\mathbb{R}^d}\mathbb{1}_{\prod(a_i,y_i)}(z)F(y)\,\mathrm{d}y$ and $\alpha^2 = (2, \dots, 2)$. Clearly, f is continuous as an antiderivative. ❑

Theorem 2.2.7. *For every $\Lambda \in \mathcal{D}'_K(\Omega)$ there exists an $\alpha \in \mathbb{N}_0^d$ and an $f \in \mathrm{C}(K)$ such that $\Lambda = \mathrm{D}^\alpha\Lambda_f$, i.e., $\Lambda(\phi) = (-1)^{|\alpha|}\Lambda_f(\mathrm{D}^\alpha\phi)$ for all $\phi \in \mathcal{D}_K(\Omega)$.*

Proof. Before we show the actual claim we show an estimate for $\|\cdot\|_n$. By the mean value theorem we have

$$|\phi(x)| \le M_K \max_{x\in K}\left|\frac{\partial}{\partial x_i}\phi(x)\right| = M_K\left\|\frac{\partial}{\partial x_i}\phi\right\|_\infty,$$

where M_K is the diameter of K. By induction we obtain

$$|\phi(x)| \le M_K^{|\alpha|}\|\mathrm{D}^\alpha\phi\|_\infty \quad \text{for} \quad \alpha \in \mathbb{N}_0^d.$$

Hence, for $\alpha^n := (n, \dots, n)$, we have

$$|\mathrm{D}^\alpha\phi(x)| \le M_K^{nd-|\alpha|}\|\mathrm{D}^{\alpha^n}\phi\|_\infty$$

for every $\alpha \in \mathbb{N}_0^d$ with $\alpha_i \le n$ for every $i \in \{1, \dots, d\}$. Consequently,

$$\|\phi\|_n = \max_{|\alpha|\le n}\|\mathrm{D}^\alpha\phi\|_\infty \le \underbrace{\max_{|\alpha|\le n} M_K^{nd-|\alpha|}}_{=:C_{n,K}}\|\mathrm{D}^{\alpha^n}\phi\|_\infty. \tag{2.3}$$

Since Λ is continuous, there exists a $c > 0$ and an $n \in \mathbb{N}_0$ such that

$$|\Lambda(\phi)| \le c\|\phi\|_n \le cC_{n,K}\|\mathrm{D}^{\alpha^n}\phi\|_\infty. \tag{2.4}$$

Note that $\phi \mapsto \mathrm{D}^{\alpha^n}\phi$ is a injective mapping from $\mathcal{D}_K(\Omega)$ to $\mathcal{D}_K(\Omega)$ (a consequence of (2.3)). Hence the mapping

$$\tilde{\Lambda}\colon \left\{\begin{array}{rcl}\mathrm{D}^{\alpha^n}\mathcal{D}_K(\Omega) & \to & \mathbb{C},\\ \mathrm{D}^{\alpha^n}\phi & \mapsto & \Lambda(\phi),\end{array}\right.$$

is well-defined and by (2.4) even continuous. By Hahn-Banach we can continuously and linearly extend $\tilde{\Lambda}$ to $\mathrm{C}(K)$. By the representation theorem of Riesz-Markov-Kakutani, there exists a measure μ such that the extension of $\tilde{\Lambda}$ can be written as $\int_K \psi \,\mathrm{d}\mu$ for $\psi \in \mathrm{C}(K)$. Therefore, by Lemma 2.2.6 we have

$$\begin{aligned}\Lambda(\phi) = \tilde{\Lambda}(\mathrm{D}^{\alpha^n}\phi) = \int_K \mathrm{D}^{\alpha^n}\phi \,\mathrm{d}\mu = \Lambda_\mu(\mathrm{D}^{\alpha^n}\phi) \overset{L.2.2.6}{=} \mathrm{D}^{\beta}\Lambda_f(\mathrm{D}^{\alpha^n}\phi)\\ = \mathrm{D}^{\alpha^n+\beta}\Lambda_{(-1)^{|\alpha^n|}f}(\phi).\end{aligned}$$

for an $f \in \mathrm{C}(K)$. ❑

We have

$$\mathcal{D}(\Omega) = \bigcup_{K\subseteq\Omega \text{ compact}} \mathcal{D}_K(\Omega)$$

and we will endow $\mathcal{D}(\Omega)$ with the finest locally convex topology such that all inclusion mappings $\iota_K\colon \mathcal{D}_K(\Omega) \to \mathcal{D}(\Omega), f \mapsto f$ are continuous.

Definition 2.2.8. We define P as the set of all seminorms p on $\mathcal{D}(\Omega)$ that satisfy

$$\forall K \subseteq \Omega \text{ compact } \exists c > 0, n \in \mathbb{N}_0 : p(\phi) \le c\|\phi\|_n \ \forall \phi \in \mathcal{D}_K(\Omega).$$

Finally, we define the topology $\mathcal{T}$ on $\mathcal{D}(\Omega)$ as the locally convex topology that is generated by P. ✧

Clearly, $\{\|\cdot\|_n \,|\, n \in \mathbb{N}_0\}$ is a subset of P. However, the additional seminorms in P put more weight on the area close to the boundary of Ω, as they only need to be bounded by $\|\cdot\|_n$ on compact sets.

Theorem 2.2.9. *Let $\mathcal{T}$ be the topology on the test functions $\mathcal{D}(\Omega)$ from Definition 2.2.8. Then the following holds:*

(i) *$\mathcal{T}$ is the finest locally convex topology on $\mathcal{D}(\Omega)$ such that $\iota_K\colon \mathcal{D}_K(\Omega) \to \mathcal{D}(\Omega)$, $\phi \mapsto \phi$ is continuous for all $K \subseteq \Omega$ compact.*

(ii) *$\mathcal{T}_K$ equals the relative topology of $\mathcal{T}$ on $\mathcal{D}_K(\Omega)$.*

(iii) *Let $(X, \mathcal{T}_X)$ be another locally convex topological vector space. Then a linear mapping $f\colon (\mathcal{D}(\Omega), \mathcal{T}) \to (X, \mathcal{T}_X)$ is continuous, if and only if $f \circ \iota_K$ is continuous for all $K \subseteq \Omega$ compact.*

Proof. We show the assertion in the listed order.

(i) Let $\tilde{\mathcal{T}}$ be locally convex topology on $\mathcal{D}(\Omega)$ such that all ι_K are continuous and $\mathcal{T} \subseteq \tilde{\mathcal{T}}$ ($\tilde{\mathcal{T}}$ is finer than $\mathcal{T}$). Then there exists a set $\tilde{P}$ of seminorms that induces $\tilde{\mathcal{T}}$. The continuity of $\iota_K\colon (\mathcal{D}_K(\Omega), \mathcal{T}_K) \to (\mathcal{D}(\Omega), \tilde{\mathcal{T}})$ implies that

$$\iota_K^{-1}(V(0, \tilde{p}, 1)) = \{\psi \in \mathcal{D}_K(\Omega) \,|\, \tilde{p}(\psi) < 1\} \supseteq V(0, \|\cdot\|_n, \epsilon)$$

for an $n \in \mathbb{N}_0$ and an $\epsilon > 0$ or equivalently $\|\psi\|_n < \epsilon \Rightarrow \tilde{p}(\psi) < 1$ for $\psi \in \mathcal{D}_K(\Omega)$. Hence, rescaling yields

$$\tilde{p}\left(\frac{\frac{\epsilon}{2}\psi}{\|\psi\|_n}\right) < 1 \quad \text{or equivalently} \quad \tilde{p}(\psi) < \frac{2}{\epsilon}\|\psi\|_n$$

for all $\psi \in \mathcal{D}_K(\Omega)$, which implies $\tilde{p} \in P$ and $\tilde{\mathcal{T}} = \mathcal{T}$.

(ii) The continuity of ι_K implies that $\mathcal{T}_K$ is finer than $\iota_K^{-1}(\mathcal{T}) = \mathcal{T}\big|_{\mathcal{D}_K(\Omega)}$ ($\mathcal{T}_K \supseteq \mathcal{T}\big|_{\mathcal{D}_K(\Omega)}$), where $\mathcal{T}\big|_{\mathcal{D}_K(\Omega)} = \{O \cap \mathcal{D}_K(\Omega) \mid O \in \mathcal{T}\}$.

On the other hand

$$\begin{aligned} V_{\mathcal{D}_K(\Omega)}(\phi, \|\cdot\|_n, \epsilon) &= \{\psi \in \mathcal{D}_K(\Omega) \mid \|\phi - \psi\|_n < \epsilon\} \\ &= \{\psi \in \mathcal{D}(\Omega) \mid \|\phi - \psi\|_n < \epsilon\} \cap \mathcal{D}_K(\Omega) \\ &= V_{\mathcal{D}(\Omega)}(\phi, \|\cdot\|_n, \epsilon) \cap \mathcal{D}_K(\Omega) \in \mathcal{T}\big|_{\mathcal{D}_K(\Omega)}, \end{aligned}$$

because $\|\cdot\|_n$ is also in P. Since $\mathcal{T}_K$ is generated by $V_{\mathcal{D}_K(\Omega)}(\phi, \|\cdot\|_n, \epsilon)$ we conclude $\mathcal{T}_K \subseteq \mathcal{T}\big|_{\mathcal{D}_K(\Omega)}$.

(iii) Clearly, if f is continuous, then also $f \circ \iota_K$ is continuous.

Hence, it is left to show that the continuity of f follows from the continuity of $f \circ \iota_K$: Let Q denote a set of seminorms that induces $\mathcal{T}_X$. By the linearity of f it follows immediately that $q \circ f$ is seminorm on $\mathcal{D}(\Omega)$ (it is not obvious that $q \circ f \in P$, we will show this). Moreover, $q \circ f \circ \iota_K$ is a seminorm on $\mathcal{D}_K(\Omega)$, which is even continuous by continuity of $f \circ \iota_K$. Hence, there exists an $n \in \mathbb{N}_0$ and a $c > 0$ such that

$$q \circ f(\underbrace{\iota_K \psi}_{=\psi}) \leq c\|\psi\|_n,$$

which implies $q \circ f \in P$ and in particular that $q \circ f$ is continuous. Therefore, the preimage of an open set under $q \circ f$ is open in $\mathcal{D}(\Omega)$. Recall that the preimage of $(-\epsilon, \epsilon)$ under q is $V(0, q, \epsilon)$ and these sets generate $\mathcal{T}_X$. Consequently,

$$\mathcal{T} \ni (q \circ f)^{-1}\Big((-\epsilon, \epsilon)\Big) = f^{-1}\Big(q^{-1}\Big((-\epsilon, \epsilon)\Big)\Big) = f^{-1}\Big(V(0, q, \epsilon)\Big).$$

This implies the continuity of f at 0 and since f is linear also its continuity. ❑

Theorem 2.2.10. *A sequence $(\phi_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(\Omega)$ converges to 0 (w.r.t. $\mathcal{T}$), if and only if*

- *there exists a compact $K \subseteq \Omega$ such that* $\operatorname{supp} \phi_n \subseteq K$ *for all $n \in \mathbb{N}$*
- *and* $\sup_{x\in\Omega}|\mathrm{D}^\alpha \phi_n(x)| \to 0$ *for all $\alpha \in \mathbb{N}_0^d$.*

Proof. We show both implications seperately.

$\boxed{\Leftarrow}$: If there exits a compact K such that $\operatorname{supp}\phi_n \subseteq K$ for all $n \in \mathbb{N}$ and $\sup_{x\in\Omega}|\mathrm{D}^\alpha \phi_n(x)| \to 0$ for all $\alpha \in \mathbb{N}_0^n$, then $(\phi_n)_{n\in\mathbb{N}}$ converges to 0 in $\mathcal{D}_K(\Omega)$ (w.r.t. $\mathcal{T}_K$). Hence, by the continuity of the embedding mapping ι_K the sequence $(\phi_n)_{n\in\mathbb{N}}$ converges also in $\mathcal{D}(\Omega)$ to 0 (w.r.t. $\mathcal{T}$).

$\boxed{\Rightarrow}$: For the reverse implication we assume that $(\phi_n)_{n\in\mathbb{N}}$ converges to 0 w.r.t. $\mathcal{T}$. We define $A_n := \phi_n^{-1}(\mathbb{C} \setminus \{0\})$ and $A := \bigcup_{n\in\mathbb{N}} A_n$. Note that $\overline{A_n} = \operatorname{supp}\phi_n$. We will show that A is relatively compact in Ω ($\overline{A}$ is compact).

Let us assume A is not relatively compact. Then there exists a sequence $(x_j)_{j\in\mathbb{N}}$ in A that has no accumulation point. Since every A_n is relatively compact there are at most finitely many members of $(x_j)_{j\in\mathbb{N}}$ in A_n. Hence, by passing over to subsequences of $(x_j)_{j\in\mathbb{N}}$ and $(\phi_n)_{n\in\mathbb{N}}$ we can assume $\phi_n(x_j) = 0$ for $n < j$ and $\phi_n(x_n) \neq 0$. Inductively we can define a sequence $(c_j)_{j\in\mathbb{N}}$ in $\mathbb{C}$ such that $\sum_{j=1}^n c_j\phi_n(x_j) = 1$. Note that $\phi \mapsto c_j\phi(x_j)$ is continuous for $\phi \in \mathcal{D}_K(\Omega)$ for every compact $K \subseteq \Omega$. We define the mapping $\Lambda(\phi) := \sum_{j=1}^\infty c_j\phi(x_j)$ for every $\phi \in \mathcal{D}(\Omega)$. Note that $\phi \in \mathcal{D}(\Omega)$ has compact support and consequently there are only finitely many members of $(x_j)_{j\in\mathbb{N}}$ in this compact support. Consequently $\Lambda(\phi)$ is just a finite sum. Therefore, $\Lambda \circ \iota_K$ is a continuous mapping and by Theorem 2.2.9 also Λ is continuous. Since $(\phi_n)_{n\in\mathbb{N}}$ converges to 0, we conclude $\Lambda(\phi_n) \to 0$, but by construction we have $\Lambda(\phi_n) = 1$ for all $n \in \mathbb{N}$. Hence, our assumption cannot hold and A is relatively compact.

We define $K = \overline{A}$ and conclude that $(\phi_n)_{n\in\mathbb{N}}$ is also a sequence in $\mathcal{D}_K(\Omega)$. Since the relative topology of $\mathcal{T}$ coincides with $\mathcal{T}_K$, we conclude that $(\phi_n)_{n\in\mathbb{N}}$ converges to 0 w.r.t. $\mathcal{T}_K$ and therefore

$$\|\phi_n\|_m \to 0 \quad \text{for every} \quad m \in \mathbb{N}_0,$$

which implies $\|\mathrm{D}^\alpha\phi_n\|_\infty \to 0$ for every $\alpha \in \mathbb{N}_0^d$. ❑

Definition 2.2.11. We define the space of *distributions* $\mathcal{D}'(\Omega)$ as the (topological) dual space of test functions $\mathcal{D}(\Omega)$. For $\Lambda \in \mathcal{D}'(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$ we define

$$\langle \Lambda, \phi\rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} := \Lambda(\phi).$$

We will write just $\langle \Lambda, \phi\rangle_{\mathcal{D}',\mathcal{D}}$, if Ω is clear and $\langle \Lambda, \phi\rangle$, if it can't be confused with another dual pairing. ✧

Remark 2.2.12. Every $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ can be regarded as a distribution by

$$\langle f, \phi\rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} = \int_\Omega f\phi \,\mathrm{d}\lambda,$$

where λ denotes the Lebesgue measure. If we want to emphasize that we regard f as a distribution we write Λ_f, i.e., $\Lambda_f(\phi) := \int_\Omega f\phi\,\mathrm{d}\lambda$. This is because of the way we constructed the topology on $\mathcal{D}(\Omega)$. Note that if we would have chosen the topology on $\mathcal{D}(\Omega)$ that is generated by the seminorms $\{\|\cdot\|_n \mid n \in \mathbb{N}\}$, then not every $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ would represent a distribution by this integral form.

A distribution that can be represented by an $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ via the previous integral is called *regular*. ✧

Example 2.2.13. Let us regard the pre-construction $\hat{\rho}$ of our standard mollifier ρ_ϵ on $\mathbb{R}$ from Example 2.1.2. This time we regard $\epsilon \to \infty$ instead of $\epsilon \to 0$. Then the function coverges to the constant 1 function. In particular we regard the sequence

$$(\phi_n)_{n\in\mathbb{N}} \coloneqq \Big(\frac{1}{n}\hat{\rho}\big(\tfrac{\cdot}{n}\big)\Big)_{n\in\mathbb{N}}.$$

This sequence converges to 0 w.r.t. $\|\cdot\|_m$ for every $m \in \mathbb{N}$ (but not with respect to $\mathcal{T}$, because there is no compact set that covers the $\operatorname{supp}\phi_n$ for all $n \in \mathbb{N}$ simultaneously). Now the function $f(x) = \mathrm{e}^x$ is certainly in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$, but

$$\int_{\mathbb{R}} \mathrm{e}^x \frac{1}{n}\hat{\rho}_n\big(\tfrac{x}{n}\big)\,\mathrm{d}x \geq \frac{1}{n}\int_0^{\frac{n}{2}} \frac{1}{n}\mathrm{e}^x\mathrm{e}^{-2}\,\mathrm{d}x = \frac{\mathrm{e}^{-2}}{n}(\mathrm{e}^{\frac{n}{2}} - 1) \to +\infty$$

as $n \to +\infty$. Hence f (Λ_f) is not continuous w.r.t. the topology the is induce by $\{\|\cdot\|_n \mid n \in \mathbb{N}\}$. ✧

Inspired by the integration by parts formula we define $\mathrm{D}^\alpha\Lambda$ for a distribution.

Definition 2.2.14. Let $\Lambda \in \mathcal{D}'(\Omega)$ we define the *distributional derivative* $\mathrm{D}^\alpha\Lambda$ pointwise for every $\phi \in \mathcal{D}(\Omega)$ by

$$\langle \mathrm{D}^\alpha\Lambda, \phi\rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} = (-1)^{|\alpha|}\langle \Lambda, \mathrm{D}^\alpha\phi\rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)}.$$ ✧

Note that a distribution is arbitrarily often differentiable (in the distributional sense).

Example 2.2.15. We define the *Heaviside function* $\mathrm{H_f}\colon \mathbb{R} \to \mathbb{R}$ by

$$\mathrm{H_f}(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0. \end{cases}$$

Clearly, $\mathrm{H_f} = \mathbb{1}_{(0,+\infty)}$. Its distributional derivative can be calculated by

$$\langle \mathrm{H_f}', \phi\rangle_{\mathcal{D}'(\mathbb{R}),\mathcal{D}(\mathbb{R})} = -\int_{\mathbb{R}} \mathrm{H_f}\phi'\,\mathrm{d}\lambda = -\int_0^{+\infty} \phi'(x)\,\mathrm{d}x = -\phi(x)\Big|_0^{+\infty} = \phi(0)$$

where $\phi \in \mathcal{D}(\mathbb{R})$. Note that $\delta_0\colon \phi \mapsto \phi(0)$ is continuous and linear, and therefore an element of $\mathcal{D}'(\mathbb{R})$. The function δ_0 is called the *Dirac delta* or *Delta distribution*. ✧

Lemma 2.2.16. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ that converges pointwise to $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ such that for every compact $K \subseteq \Omega$ there exists an integrable function g_K such that $|f_n(x)| \le g_K(x)$ for a.e. $x \in K$ (independent of $n \in \mathbb{N}$). Then f_n converges to f in $\mathcal{D}'(\Omega)$, i.e.,*

$$\lim_{n\to\infty} \langle f_n, \phi\rangle = \langle f, \phi\rangle \quad \text{for all} \quad \phi \in \mathcal{D}(\Omega).$$

Proof. Let $\phi \in \mathcal{D}(\Omega)$ be arbitrary. Then $\operatorname{supp}\phi$ is compact and therefore there exists an integrable function $g_{\operatorname{supp}\phi}$ such that $|f_n(x)| \le g_{\operatorname{supp}\phi}(x)$ for a.e. $x \in \operatorname{supp}\phi$. Hence, by Lebesgue's dominated convergence theorem, we have

$$\begin{aligned}\lim_{n\to\infty} |\langle f_n - f, \phi\rangle| &= \lim_{n\to\infty} \left| \int_\Omega (f_n - f)\phi \,\mathrm{d}\lambda \right| \\ &\le \|\phi\|_\infty \lim_{n\to\infty} \int_{\operatorname{supp}\phi} |f_n - f| \,\mathrm{d}\lambda = 0. \end{aligned}$$ ❑

Lemma 2.2.17. *Let $(\Lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(\Omega)$ that converges to $\Lambda \in \mathcal{D}'(\Omega)$ in $\mathcal{D}'(\Omega)$. Then $\mathrm{D}^\alpha\Lambda_n$ converges to $\mathrm{D}^\alpha\Lambda$ in $\mathcal{D}'(\Omega)$ for every $\alpha \in \mathbb{N}_0^d$.*

Proof. Let $\phi \in \mathcal{D}(\Omega)$. Note that also $\mathrm{D}^\alpha\phi \in \mathcal{D}(\Omega)$. Hence,

$$\langle \mathrm{D}^\alpha(\Lambda_n - \Lambda), \phi\rangle = (-1)^{|\alpha|}\langle \Lambda_n - \Lambda, \mathrm{D}^\alpha\phi\rangle \to 0.$$ ❑

Theorem 2.2.18. *Let $\Lambda\colon \mathcal{D}(\Omega) \to \mathbb{C}$. Then the following is equivalent*

(i) *Λ is continuous, i.e., $\Lambda \in \mathcal{D}'(\Omega)$.*

(ii) *For all $K \subseteq \Omega$ compact $\Lambda \circ \iota_K$ is continuous, i.e., $\Lambda \circ \iota_K \in \mathcal{D}'_K(\Omega)$.*

(iii) *For all sequences $(\phi_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(\Omega)$ that converge to 0 (w.r.t. $\mathcal{T}$) the image sequence $(\Lambda(\phi_n))_{n\in\mathbb{N}}$ converges also to 0.*

Note that the implication (iii) $\Rightarrow$ (i) is not trivial in the sense that the topology $\mathcal{T}$ is not induces by a metric and hence sequential continuity does not imply continuity in general.

Proof. The equivalence of (i) and (ii) is valid by Theorem 2.2.9.

(iii) $\Rightarrow$ (ii): Note that the topology $\mathcal{T}_K$ on $\mathcal{D}_K(\Omega)$ is metrizable. Since every zero sequence $(\phi_n)_{n\in\mathbb{N}}$ in $\mathcal{D}_K(\Omega)$ is also a zero sequence in $\mathcal{D}(\Omega)$ we conclude that $\Lambda \circ \iota_K$ is sequentially continuous and thus continuous.

(i) $\Rightarrow$ (iii): Continuity always implies sequential continuity. ❑

2.3 Fundamental solution

Let us regard a differential operator of order k

$$L\phi = \sum_{|\alpha|\le k} c_\alpha \mathrm{D}^\alpha\phi,$$

where $c_\alpha \in \mathbb{C}$. We could also regard non constant c_α, but we will restrict ourselves to constant coefficients, which simplifies some concepts. Our goal will to solve the differential equation

$$Lu = f$$

for a given f.

Example 2.3.1. The operator of the *Poisson equation* $\Delta u = f$ is

$$L\phi = \partial_1^2\phi + \partial_2^2\phi + \cdots + \partial_d^2\phi = \Delta\phi.$$

Furthermore the differential operator that correspond to the *heat equation* $\partial_t u = \Delta u + f$ is

$$L\phi = \partial_t\phi - \partial_1^2\phi - \partial_2^2\phi - \cdots - \partial_d^2\phi = \partial_t^2\phi - \Delta\phi.$$

Hence, by adding an additional dimension we can also include time derivatives. ✧

Definition 2.3.2. Let $\Omega \subseteq \mathbb{R}^d$ be open, L a differential operator of order k and $f \in \mathcal{D}'(\Omega)$. Then we say:

(i) A function u is a *classical solution* of $Lu = f$, if $f \in \mathrm{C}(\Omega)$, $u \in \mathrm{C}^k(\Omega)$ and $Lu = f$ is pointwise satisfied.

(ii) A distribution $u \in \mathcal{D}'(\Omega)$ is a *distributional solution* of $Lu = f$, if the equation is satisfied in the distributional sense, i.e., $\langle Lu - f, \phi\rangle = 0$ for all $\phi \in \mathcal{D}(\Omega)$. ✧

Definition 2.3.3. Let δ_0 be the delta distribution in 0. Then we say a distributional solution $u_0 \in \mathcal{D}'(\mathbb{R}^d)$ of

$$Lu_0 = \delta_0$$

is a *fundamental solution* of L. ✧

Remark 2.3.4. Note that a fundamental solution u_0 is in general not unique. If there is a $v \in \mathcal{D}'(\Omega)$ such that $Lv = 0$, then clearly also $u_0 + v$ is a fundamental solution. In particular if $c_\alpha = 0$, then $v = c$ for $c \in \mathbb{C}$ is such that $Lv = 0$. ✧

Note that for $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ and $\psi \in \mathrm{C}^\infty(\Omega)$ it is straightforward to define $\psi\Lambda_f$ as $\Lambda_{\psi f}$. However, we can also regard it as

$$\langle \Lambda_{\psi f}, \phi\rangle = \int_\Omega \psi f \phi \,\mathrm{d}\lambda = \int_\Omega f\psi\phi \,\mathrm{d}\lambda = \langle \Lambda_f, \psi\phi\rangle \quad \forall \phi \in \mathcal{D}(\Omega).$$

Thus for a regular distriubtion Λ we have $(\psi\Lambda)(\phi) = \Lambda(\psi\phi)$. This motivates the definition for a general distribution $\Lambda \in \mathcal{D}'(\Omega)$ and $\psi \in \mathrm{C}^\infty(\Omega)$:

$$\langle \psi\Lambda, \phi\rangle := \langle \Lambda, \psi\phi\rangle \quad \forall \phi \in \mathcal{D}(\Omega).$$

Similary we have proceeded with the derivative of a distribution. In general this strategy of defining operations on distribution by applying the "adjoint" on the argument (test function) leeds to the term *formal adjoint.* Note that the part *formal* can be missleading, as it does not mean that there is an extra amount of rigorousness. In fact when we compare it to the L^2 adjoint, we will observe that the *formal adjoint* is less strict.

Definition 2.3.5. Let $A\colon \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ be a linear mapping. If there exists a linear and continuous mapping $B\colon \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ such that

$$\int_\Omega (A\psi)\phi \,\mathrm{d}\lambda = \int_\Omega \psi B\phi \,\mathrm{d}\lambda \quad \forall \psi, \phi \in \mathcal{D}(\Omega),$$

then we say B is the *formal adjoint* of A and for $\Lambda \in \mathcal{D}'(\Omega)$ we define $A\Lambda \coloneqq \Lambda \circ B$, i.e.,

$$\langle A\Lambda, \phi\rangle = \langle \Lambda, B\phi\rangle \quad \forall \phi \in \mathcal{D}(\Omega).$$

We will also denote B as A^*. ✧

Note that $A\Lambda \in \mathcal{D}'(\Omega)$ by the continuity of B.

Example 2.3.6. We have already used the formal adjoint to extend the derivative D^α on the distributions. It can be easily seen, by integration by parts, that the formal adjoint of D^α is given by $(-1)^{|\alpha|}\mathrm{D}^\alpha$.

To show the continuity of D^α we only need to show continuity at 0 (since D^α is linear). Let $(\phi_n)_{n\in\mathbb{N}}$ be a zero sequence in $\mathcal{D}(\Omega)$, i.e., there exists a compact $K \subseteq \Omega$ such that $\operatorname{supp}\phi_n \subseteq K$ for all $n \in \mathbb{N}$ and $\sup_{|\beta|\le m}\|\mathrm{D}^\beta\phi_n\|_\infty \to 0$ for every $m \in \mathbb{N}$. This implies that also $\sup_{|\beta|\le m}\|\mathrm{D}^{\alpha+\beta}\phi_n\|_\infty \to 0$, which gives that also $(\mathrm{D}^\alpha\phi_n)_{n\in\mathbb{N}}$ is zero sequence. Consequently, the mapping $\mathrm{D}^\alpha\colon \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ is continuous. ✧

Example 2.3.7. Let us regard the convolution operator $A\colon \mathcal{D}(\mathbb{R}^d) \to \mathcal{D}(\mathbb{R}^d)$, $\phi \mapsto \phi * f$, where $f \in \mathrm{L}^1(\mathbb{R}^d)$ with compact support. First of all this operator is really well-defined:

- $\phi * f \in \mathrm{C}^\infty(\mathbb{R}^d)$, because $\mathrm{D}^\alpha(\phi * f) = \mathrm{D}^\alpha\phi * f$.
- $\phi * f$ has compact support, because $\operatorname{supp}\phi * f \subseteq \operatorname{supp}\phi + \operatorname{supp} f$.

In order to calculate its formal adjoint A^* we introduce the operator $R\colon g \mapsto g(-\cdot)$. Then the following calculation gives A^*

$$\begin{aligned}
&\int_{\mathbb{R}^d} (A\phi)\psi \,\mathrm{d}\lambda \\
&\quad= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \phi(y)f(x-y)\,\mathrm{d}y\,\psi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \phi(y)f(x-y)\psi(x)\,\mathrm{d}y\,\mathrm{d}x \\
&\quad= \int_{\mathbb{R}^d} \phi(y)\int_{\mathbb{R}^d} \psi(x)f(x-y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d} \phi(y)\int_{\mathbb{R}^d} \psi(x)f(\underbrace{x-y}_{=-(y-x)})\,\mathrm{d}x\,\mathrm{d}y \\
&\quad= \int_{\mathbb{R}^d} \phi(y)(\psi * Rf)(y)\,\mathrm{d}y\,.
\end{aligned}$$

Hence, A^* is basically A again, the only difference that we convolve with Rf instead of f. Thus, the formal adjoint A^* is continuous if and only if A itself is continuous.

In order to show the continuity of A we have to show that every zero sequence $(\phi_n)_{n\in\mathbb{N}}$ is mapped on a zero sequence. Let $(\phi_n)_{n\in\mathbb{N}}$ be a zero sequence then there exists a compact $K \subseteq \mathbb{R}^d$ such that $\operatorname{supp}\phi_n \subseteq K$ for all $n \in \mathbb{N}$ and $\|\mathrm{D}^\alpha \phi_n\|_\infty \to 0$ for all $\alpha \in \mathbb{N}_0^d$. Therefore,

$$\operatorname{supp} A\phi_n = \operatorname{supp}\phi_n * f \subseteq \operatorname{supp}\phi_n + \operatorname{supp} f \subseteq K + \operatorname{supp} f =: K'.$$

Moreover,

$$\|\mathrm{D}^\alpha A\phi_n\|_\infty = \|(\mathrm{D}^\alpha \phi_n) * f\|_\infty \leq \|\mathrm{D}^\alpha \phi_n\|_\infty \|f\|_{\mathrm{L}^1} \to 0,$$

which proves the continuity of A and A^*.

Therefore, A can be extended to $\mathcal{D}'(\mathbb{R}^d)$, i.e., for $\Lambda \in \mathcal{D}'(\mathbb{R}^d)$ we have

$$\langle \Lambda * f, \phi\rangle = \langle \Lambda, \phi * Rf\rangle \quad \forall \phi \in \mathcal{D}(\mathbb{R}^d). \qquad ✧$$

Theorem 2.3.8. *Let u_0 be a fundamental solution of L. Then for every $f \in \mathrm{L}^1(\mathbb{R}^d)$ with compact support the distribution $u_0 * f$ is a distributional solution of*

$$Lu = f.$$

Proof. Note that by our assumptions L has constant coefficients. Hence, $(L^*\phi) * (Rf) = L^*(\phi * (Rf))$. Therefore,

$$\begin{aligned}\langle L(u_0 * f), \phi\rangle &= \langle u_0 * f, L^*\phi\rangle = \langle u_0, (L^*\phi) * (Rf)\rangle \\ &= \langle Lu_0, \phi * (Rf)\rangle = \langle \delta_0, \phi * (Rf)\rangle = (\phi * (Rf))(0) \\ &= \int_{\mathbb{R}^d} \phi(x) f(x)\,\mathrm{d}x = \langle f, \phi\rangle,\end{aligned}$$

which implies $L(u_0 * f) = f$ in the distribtional sense. ❑

Example 2.3.9. Let $\Omega = \mathbb{R}$ and $L = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$. Then a fundamental solution of L is given by

$$u_0(x) = \frac{1}{2}|x|$$

This can be seen by $\frac{\mathrm{d}}{\mathrm{d}x}u_0 = \mathrm{H_f} - \frac{1}{2}$ and $\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{H_f} - \frac{1}{2}) = \delta_0$. ✧

Note that one can extend the convolution to two distribution, where one has compact support. Then we can get a generalization of Theorem 2.3.8. Moreover, it can be shown that every differential operator L (with constant coefficients) has a fundamental solution. This result is called the Ehrenpreis–Malgrange theorem. For these results we refer to [7].

Moreover, if we want to solve a differential equation on a general $\Omega \subseteq \mathbb{R}^d$ with additional boundary conditions similar to the presented approach with fundamental solutions, then this leads to Green's functions, see e.g., [5].

2.4 Partition of unity

We recall the mollifier from Remark 2.1.4 which is constructed from the function in Example 2.1.2, i.e.,

$$\rho_\epsilon(x) = \begin{cases} \epsilon^{-d}\frac{1}{\|\hat{\rho}\|_{\mathrm{L}^1}}\exp\left(\frac{1}{\|\frac{1}{\epsilon}x\|_2^2-1}\right), & \|x\|_2 < \epsilon, \\ 0, & \text{else}, \end{cases}$$

where $\hat{\rho}$ is the function from Example 2.1.2, i.e., $\hat{\rho}(x) = \exp(\frac{1}{\|x\|_2^2-1})$ for $\|x\|_2 < 1$ and 0 else. Note that

$$\rho_\epsilon \in \mathring{\mathrm{C}}^\infty(\mathbb{R}^d), \quad \|\rho_\epsilon\|_{\mathrm{L}^1} = 1 \quad \text{and} \quad \operatorname{supp}\rho_\epsilon \subseteq \mathrm{B}_\epsilon(0).$$

Theorem 2.4.1 (Partition of unity). *Let $K \subseteq \mathbb{R}^d$ be compact and $(U_i)_{i=1}^n$ an open covering of K. Then there are $\alpha_1, \ldots, \alpha_n \in \mathring{\mathrm{C}}^\infty(\mathbb{R}^d)$ such that $\alpha(x) \in [0,1]$, $\operatorname{supp}\alpha_i \subseteq U_i$ and*

$$\sum_{i=1}^n \alpha_n(x) \in [0,1] \quad \textit{for } x \in \mathbb{R}^d \quad \textit{and} \quad \sum_{i=1}^n \alpha_n(x) = 1 \quad \textit{for } x \in K.$$

Proof. Note that for every $x \in K$ there exists an $i(x) \in \{1,\ldots,n\}$ such that $x \in U_{i(x)}$. We choose δ_x small enough such that $\mathrm{B}_{2\delta_x}(x) \subseteq U_{i(x)}$. Clearly also $(\mathrm{B}_{\delta_x}(x))_{x\in K}$ is an open covering of K. Hence, there is there exits a finite subcovering $(\mathrm{B}_{\delta_{x_j}}(x_j))_{j=1}^m$. We define

$$A_1 = \mathrm{B}_{\delta_{x_1}}(x_1) \quad \text{and} \quad A_j = \mathrm{B}_{\delta_{x_j}}(x_j) \setminus \big(A_1 \cup \cdots \cup A_{j-1}\big) \quad \text{for } j \geq 2.$$

Moreover, we set $\delta = \min\{\delta_{x_j} \mid j = 1,\ldots,m\}$. Then

$$A_j + \overline{\mathrm{B}}_\delta(0) \subseteq \mathrm{B}_{\delta_{x_j}}(x_j) + \overline{\mathrm{B}}_\delta(0) \subseteq \mathrm{B}_{2\delta_{x_j}}(x_j) \subseteq U_{i(x_j)}. \tag{2.5}$$

We define $\beta_j = \mathbb{1}_{A_j} * \rho_\delta$. Then $\operatorname{supp}\beta_j \subseteq U_{i(x_j)}$ by (2.5). Then we have

$$\sum_{j=1}^m \beta_j(x) = \sum_{j=1}^m (\rho_\delta * \mathbb{1}_{A_j})(x) = \Big(\rho_\delta * \sum_{j=1}^m \mathbb{1}_{A_j}\Big)(x) = (\rho * \mathbb{1}_{\bigcup A_j})(x).$$

Hence, for $x \in \mathbb{R}^d$ we conclude $\sum_{j=1}^m \beta_j(x) \in [0,1]$. Moreover, for $x \in K$ we have that $\mathrm{B}_\delta(x) \subseteq \bigcup_{j=1}^m A_j$ and therefore $\sum_{j=1}^m \beta_j(x) = 1$.

Finally, we just sort the β_j: We define $i(j)$ by $i(x_j)$ and define

$$\alpha_k = \sum_{j \in i^{-1}(k)} \beta_j.$$

Note that $i^{-1}(k)$ can be empty, but then the corresponding α_k is just 0. ❑

2.5 Stokes' theorem

For this section we want to recall the notion of Lipschitz domains and Lipschitz boundaries. That is we can locally represent $\partial\Omega$ as the graph of a Lipschitz mapping. In particular for every $p \in \partial\Omega$ we find a cylinder $C_{\epsilon,h}(p)$, a Lipschitz mapping a and the corresponding chart k, see Appendix B.

$$\int_{\partial\Omega\cap C_{\epsilon,h}(p)} f \,\mathrm{d}\mu = \int_{B_\epsilon(0)} f\big(k^{-1}(x)\big)\sqrt{1+\|\nabla a(x)\|^2}\,\mathrm{d}\lambda_{d-1}(x)\,.$$

In order to prove the *divergence theorem* or *Gauß's theorem*,

$$\int_\Omega \operatorname{div} f \,\mathrm{d}\lambda = \int_{\partial\Omega} \nu\cdot f\,\mathrm{d}\mu,$$

we will prove locally $\int_\Omega \partial_i\psi\,\mathrm{d}\lambda = \int_{\partial\Omega}\nu_i\psi\,\mathrm{d}\mu$ and then obtain the global result by a partition of unity. Finally, the divergence theorem/Gauß's theorem is just an easy consequence.

Theorem 2.5.1. *Let $\psi \in \mathcal{D}(C_{\epsilon,h}(p))$. Then*

$$\int_{\Omega\cap C_{\epsilon,h}(p)} \partial_i\psi\,\mathrm{d}\lambda = \int_{\partial\Omega\cap C_{\epsilon,h}(p)} \nu_i\psi\,\mathrm{d}\mu$$

for every $i \in \{1,\ldots,d\}$.

We will (without loss of generality) assume that $p = 0$, $W = \left[\begin{smallmatrix} e_1 & \cdots & e_{d-1}\end{smallmatrix}\right]$, $v = e_d$ and therefore the chart k and its inverse k^{-1} is given by

$$k\left(\begin{bmatrix} x_1\\ \vdots\\ x_{d-1}\\ x_d\end{bmatrix}\right) = \begin{bmatrix} x_1\\ \vdots\\ x_{d-1}\end{bmatrix} \quad\text{and}\quad k^{-1}\left(\begin{bmatrix} x_1\\ \vdots\\ x_{d-1}\end{bmatrix}\right) = \begin{bmatrix} x_1\\ \vdots\\ x_{d-1}\\ a\left(\begin{bmatrix} x_1\\ \vdots\\ x_{d-1}\end{bmatrix}\right)\end{bmatrix}.$$

We explain the modification for the general case in Remark 2.5.2

Proof. Let $h \in \mathrm{C}^\infty(\mathbb{R})$ be such that

$$h(\zeta) \in \begin{cases} \{0\}, & \zeta\in(-\infty,0),\\ [0,1], & \zeta\in[0,1],\\ \{1\}, & \zeta\in(1,\infty).\end{cases}$$

Figure 2.2 illustrates the function h (such a function can be constructed by the convolution of $\mathbb{1}_{(0,+\infty)}$ and ρ_ϵ from Remark 2.1.4). We define $h_n(x) \coloneqq h(nx)$, which converges pointwise to the Heaviside function $\mathrm{H_f} = \mathbb{1}_{(0,+\infty)}$. By the second condition of Lipschitz boundaries, we have $x \in \Omega\cap C_{\epsilon,h}(p)$ if and

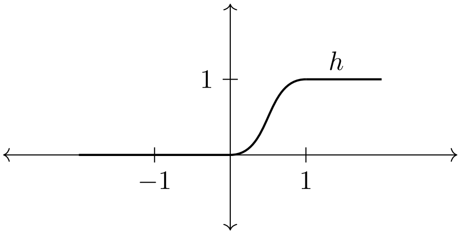

Figure 2.2: The function h

only if $x_d < a(\tilde{x})$, where $\tilde{x} = k(x)$, the projection of x on the first $d-1$ coordinates. Therefore, we can write $\mathbb{1}_\Omega$ for $x \in C_{\epsilon,h}(p)$ as a pointwise limit

$$\mathbb{1}_\Omega(x) = \mathbb{1}_{(0,\infty)}(a(\tilde{x}) - x_d) = \lim_{n\to\infty} h_n(a(\tilde{x}) - x_d) = \lim_{n\to\infty} \tilde{h}_n(x),$$

where $\tilde{h}_n(x) := h_n(a(\tilde{x}) - x_d)$. Hence $\mathbb{1}_\Omega$ regarded as distribution, i.e., as element of $\mathcal{D}'(C_{\epsilon,h})$, is also the limit of $\tilde{h}_n$ (Lemma 2.2.16). The distributional derivative of $\mathbb{1}_\Omega$ can be written as (Lemma 2.2.17)

$$\frac{\partial}{\partial x_i}\mathbb{1}_\Omega = \lim_{n\to\infty}\frac{\partial}{\partial x_i}\tilde{h}_n, \quad \text{where} \quad \frac{\partial}{\partial x_i}\tilde{h}_n(x) = n w_i(\tilde{x}) h'\big(n(a(\tilde{x}) - x_d)\big)$$

and $w(\tilde{x}) = \left[\begin{smallmatrix}\nabla a(\tilde{x})\\ -1\end{smallmatrix}\right] = -\sqrt{1 + \|\nabla a(\tilde{x})\|^2}\,\nu\left(\left[\begin{smallmatrix}\tilde{x}\\ a(\tilde{x})\end{smallmatrix}\right]\right)$ (for a.e. x). For $\psi \in \mathcal{D}(C_{\epsilon,h}(p))$ we have

$$\begin{aligned}
\int_{\Omega\cap C_{\epsilon,h}(p)} \partial_i\psi \,\mathrm{d}\lambda &= -\left\langle \frac{\partial}{\partial x_i}\mathbb{1}_\Omega, \psi\right\rangle_{\mathcal{D}'(C_{\epsilon,h}(p)),\mathcal{D}(C_{\epsilon,h}(p))} \\
&= -\lim_{n\to\infty}\int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}} n w_i(\tilde{x}) h'(n(a(\tilde{x}) - x_d))\psi(\left[\begin{smallmatrix}\tilde{x}\\ x_d\end{smallmatrix}\right])\,\mathrm{d}x_d\,\mathrm{d}\tilde{x} \\
\overset{\substack{\text{change of variables}\\ x_d \mapsto a(\tilde{x}) - x_d}}{=} &\int_{\mathbb{R}^{d-1}} w_i(\tilde{x}) \lim_{n\to\infty}\int_{\mathbb{R}}\left[\frac{\partial}{\partial x_d} h(n x_d)\right]\psi\left(\left[\begin{smallmatrix}\tilde{x}\\ a(\tilde{x}) - x_d\end{smallmatrix}\right]\right)\mathrm{d}x_d\,\mathrm{d}\tilde{x} \\
&= \int_{\mathbb{R}^{d-1}} w_i(\tilde{x})\left\langle \mathrm{H}_\mathrm{f}', \psi\left(\left[\begin{smallmatrix}\tilde{x}\\ a(\tilde{x}) - .\end{smallmatrix}\right]\right)\right\rangle_{\mathcal{D}'(\mathbb{R}),\mathcal{D}(\mathbb{R})}\mathrm{d}\tilde{x} \\
&= \int_{\mathbb{R}^{d-1}} \nu_i\left(\left[\begin{smallmatrix}\tilde{x}\\ a(\tilde{x})\end{smallmatrix}\right]\right)\psi\left(\left[\begin{smallmatrix}\tilde{x}\\ a(\tilde{x})\end{smallmatrix}\right]\right)\sqrt{1 + \|\nabla a(\tilde{x})\|^2}\,\mathrm{d}\tilde{x} \\
&= \int_{\partial\Omega\cap C_{\epsilon,h}(p)} \nu_i\psi\,\mathrm{d}\mu\,.
\end{aligned}$$

❑

Remark 2.5.2. In order to make the previous proof work for the general case we make the following modifications: $\tilde{h}_n(x) = h(n(a(\tilde{x}) - x_v))$, where $\tilde{x} = k(x) = W^\mathsf{T}(x - p)$ and $x_v = v \cdot (x - p)$, which is the replacement of x_d. Then again

$$\frac{\partial}{\partial x_i}\tilde{h}_n(x) = n w_i(\tilde{x}) h'\big(n[a(\tilde{x}) - x_v]\big),$$

where $w(\tilde{x}) = \begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} \nabla a(\tilde{x}) \\ -1 \end{bmatrix} = -\sqrt{1 + \|\nabla a(\tilde{x})\|^2}\, \nu(k^{-1}(\tilde{x}))$. Finally, the integral over $\mathbb{R}^d$ has to be split according to the hyperspace W and its normal vector v. ✧

Theorem 2.5.3 (Stokes' theorem)**.** *Let $\psi \in \mathcal{D}(\mathbb{R}^d)$ and $\Omega \subseteq \mathbb{R}^d$ be open with Lipschitz boundary. Then*

$$\int_\Omega \partial_i \psi \,\mathrm{d}\lambda = \int_{\partial\Omega} \nu_i \psi \,\mathrm{d}\mu$$

for every $i \in \{1, \ldots, d\}$.

Proof. Note that $K = \operatorname{supp}\psi \cap \overline{\Omega}$ is compact. For every $p \in \partial\Omega$ we choose $\epsilon, h > 0$ such that the corresponding cylinder satisfies the conditions for a Lipschitz boundary Definition B.1.1. Hence, the open covering $\Omega \cup \bigcup_{p \in \partial\Omega} C_{\epsilon,h}(p)$ of K has a finite subcovering consisting of $O_0 = \Omega$ and cylinders $O_j = C_{\epsilon_j,h_j}(p_j)$ for $j \in \{1, \ldots, k\}$. We employ a partition of unity and obtain $(\alpha_j)_{j=0}^k$ subordinate to this covering, i.e.,

$$\alpha_i \in \mathcal{D}(O_j), \quad \alpha_j(x) \in [0,1], \quad \text{and} \quad \sum_{j=0}^k \alpha_j(x) = 1 \quad \text{for} \quad x \in \overline{\Omega} \cap K.$$

We define $\psi_j = \alpha_j \psi \in \mathcal{D}(O_j)$. Hence, we have $\psi = \sum_{j=0}^k \psi_j$ and

$$\int_\Omega \partial_i \psi \,\mathrm{d}\lambda = \int_{\Omega \cap K} \partial_i \sum_{j=0}^k \psi_j \,\mathrm{d}\lambda = \sum_{j=0}^k \int_{\Omega \cap O_j} \partial_i \psi_j \,\mathrm{d}\lambda\,.$$

Note that ψ_0 has compact support in $O_0 = \Omega$ and therefore $\int_{\Omega \cap O_0} \partial_i \psi_0 \,\mathrm{d}\lambda = \int_{\mathbb{R}^n} \partial_i \psi_0 \,\mathrm{d}\lambda = 0$. Therefore, by Theorem 2.5.1 we have

$$\sum_{j=0}^k \int_{\Omega \cap O_j} \partial_i \psi_j \,\mathrm{d}\lambda = \sum_{j=1}^k \int_{\Omega \cap O_j} \partial_i \psi_j \,\mathrm{d}\lambda = \sum_{j=1}^k \int_{\partial\Omega \cap O_j} \nu_i \psi_j \,\mathrm{d}\mu = \int_{\partial\Omega} \nu_i \psi \,\mathrm{d}\mu,$$

which proves the claim. ❑

Corollary 2.5.4 (Gauß's theorem)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be open with Lipschitz boundary and $f \in \mathcal{D}(\mathbb{R}^d)^d$. Then*

$$\int_\Omega \operatorname{div} f \,\mathrm{d}\lambda = \int_{\partial\Omega} \nu \cdot f \,\mathrm{d}\mu\,.$$

Proof. Note that $f_i \in \mathcal{D}(\mathbb{R}^d)$. Hence, by Theorem 2.5.3

$$\int_\Omega \operatorname{div} f \,\mathrm{d}\lambda = \sum_{i=1}^d \int_\Omega \partial_i f_i \,\mathrm{d}\lambda = \sum_{i=1}^d \int_{\partial\Omega} \nu_i f_i \,\mathrm{d}\mu = \int_{\partial\Omega} \nu \cdot f \,\mathrm{d}\mu\,. \quad ❑$$

This result can be extended to a more general class of functions by continuity, e.g., $\mathrm{H}^1(\Omega)^d$, if Ω is bounded. Note that for an unbounded Ω this formula cannot be extended to $\mathrm{H}^1(\Omega)^d$ in general as shown in [11, Re. 13.7.4].

Chapter 3

Sobolev spaces

In this chapter we will introduce subspaces of $\mathrm{L}^p(\Omega)$, where every element has a meaningful derivative. Moreover, we will introduce the notion of a weak solution of a differential equation.

3.1 Weak derivative and Sobolev spaces

Since in the following the topology on the test functions is not important, we will just use $\mathring{\mathrm{C}}^\infty(\Omega)$ instead of $\mathcal{D}(\Omega)$. This is the usual notation in this context.

Definition 3.1.1. Let $\Omega \subseteq \mathbb{R}^d$ be an open set, $p \in [1, \infty]$ and $f \in \mathrm{L}^p(\Omega)$. Then we say f has a *weak α derivative*, if $\mathrm{D}^\alpha f \in \mathrm{L}^p(\Omega)$ (in the distributional sense), i.e., there exists a $g \in \mathrm{L}^p(\Omega)$ such that

$$\int_\Omega f \mathrm{D}^\alpha \phi \,\mathrm{d}\lambda = (-1)^\alpha \int_\Omega g\phi \,\mathrm{d}\lambda \quad \forall \phi \in \mathring{\mathrm{C}}^\infty(\Omega).$$

We say that then g is the weak α derivative of f or just $\mathrm{D}^\alpha f = g$. ✧

Definition 3.1.2. Let $\Omega \subseteq \mathbb{R}^d$ be an open set, $p \in [1, \infty]$. We define the *Sobolev space* $\mathrm{W}^{m,p}(\Omega)$ of order $m \in \mathbb{N}$ as

$$\mathrm{W}^{m,p}(\Omega) := \big\{ f \in \mathrm{L}^p(\Omega) \,\big|\, \mathrm{D}^\alpha f \in \mathrm{L}^p(\Omega) \ \forall |\alpha| \leq m \big\}.$$

We equip this space with

$$\|f\|_{\mathrm{W}^{m,p}(\Omega)} := \Bigg(\sum_{|\alpha| \leq m} \|\mathrm{D}^\alpha f\|_{\mathrm{L}^p(\Omega)}^p \Bigg)^{1/p} \quad \text{for} \quad p \in [1, \infty)$$

and with

$$\|f\|_{\mathrm{W}^{m,\infty}(\Omega)} := \max_{|\alpha| \leq m} \|\mathrm{D}^\alpha f\|_{\mathrm{L}^\infty(\Omega)} \quad \text{for} \quad p = \infty.$$

Moreover, for $p = 2$ we define the inner product

$$\langle f, g\rangle_{\mathrm{W}^{m,2}(\Omega)} \coloneqq \sum_{|\alpha|\le m} \langle \mathrm{D}^\alpha f, \mathrm{D}^\alpha g\rangle_{\mathrm{L}^2(\Omega)}.$$

As short notation for this norm we sometimes use $\|\cdot\|_{m,p}$. ✧

Lemma 3.1.3. *Let $\Omega \subseteq \mathbb{R}^d$ be open. Then $\mathrm{W}^{m,p}(\Omega)$ is a Banach space.*

Proof. Note that the product space

$$X \coloneqq \prod_{\substack{\alpha\in\mathbb{N}_0^d\\|\alpha|\le m}} \mathrm{L}^p(\Omega), \quad \|(f_\alpha)_{|\alpha|\le m}\|_{X,\infty} \coloneqq \max_{|\alpha|\le m}\|f_\alpha\|_{\mathrm{L}^p}$$

is a Banach space. Moreover, for $p \ne \infty$ the norm is equivalent to

$$\|(f_\alpha)_{|\alpha|\le m}\|_X \coloneqq \left(\sum_{|\alpha|\le m}\|f_\alpha\|^p_{\mathrm{L}^p(\Omega)}\right)^{1/p}.$$

for $p = \infty$ we define $\|\cdot\|_X \coloneqq \|\cdot\|_{X,\infty}$ We can identify $\mathrm{W}^{m,p}(\Omega)$ as a subspace of X by

$$\iota\colon \begin{cases} \mathrm{W}^{m,p}(\Omega) & \to & X, \\ f & \mapsto & (\mathrm{D}^\alpha f)_{|\alpha|\le m}. \end{cases}$$

This mapping is linear and isometric. The image $Y = \operatorname{ran}\iota$ can be characterized as follows. An element $(f_\alpha)_{|\alpha|\le m} \in X$ is in Y, if and only if,

$$\chi_{\phi,\beta}\big((f_\alpha)_{|\alpha|\le m}\big) \coloneqq \int_\Omega \big(f_{(0,\dots,0)}\mathrm{D}^\beta\phi - (-1)^\beta f_\beta\phi\big)\mathrm{d}\lambda = 0$$

for all $\phi \in \mathring{\mathrm{C}}^\infty(\Omega)$ and $|\beta| \le m$, i.e.,

$$Y = \bigcap_{\substack{\phi\in\mathring{\mathrm{C}}^\infty(\Omega)\\|\beta|\le m}} \ker I_{\phi,\beta}.$$

Note that the mappings $I_{\phi,\beta}\colon X \to \mathbb{C}$ are linear and bounded. Hence, $\ker I_{\phi,\beta}$ is closed and consequently Y as intersection of closed sets is also closed. Since $\mathrm{W}^{m,p}(\Omega)$ is isometric to Y, we conclude that it is a Banach space. ❑

Theorem 3.1.4 (Meyers-Serrin)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be open. Then $\mathrm{C}^\infty(\Omega) \cap \mathrm{W}^{m,p}(\Omega)$ is dense in $\mathrm{W}^{m,p}(\Omega)$ w.r.t. $\|\cdot\|_{m,p}$.*

Proof. First we will show that every $f \in \mathrm{W}^{m,p}(\Omega)$ with compact support can be approximated by $\mathring{\mathrm{C}}^\infty(\Omega)$ functions. Note that $\operatorname{supp}\mathrm{D}^\alpha f \subseteq \operatorname{supp} f$ and therefore as all derivatives of f have compact support. In order to convolve

$\mathrm{D}^\alpha f$ with the mollifier ρ_ϵ from Remark 2.1.4, we extend $\mathrm{D}^\alpha f$ outside of Ω by 0. Recall

$$\mathrm{D}^\alpha(\rho_\epsilon * f) = (\mathrm{D}^\alpha \rho_\epsilon * f) = (\rho_\epsilon * \mathrm{D}^\alpha f),$$

which implies that $\rho_\epsilon * f \in \mathrm{C}^\infty(\mathbb{R}^d)$. By choosing $n_0 > \frac{1}{\operatorname{dist}(\operatorname{supp} f, \partial\Omega)}$ we make sure that

$$\operatorname{supp} \mathrm{D}^\alpha(\rho_{\frac{1}{n}} * f) \subseteq \operatorname{supp} f + \mathrm{B}_{\frac{1}{n}}(0) \subseteq \Omega \quad \text{for} \quad n > n_0,$$

which gives $\rho_{\frac{1}{n}} * f \in \mathring{\mathrm{C}}^\infty(\Omega)$ for $n > n_0$. By Lemma 2.1.3 we have

$$\left\|f * \rho_{\frac{1}{n}} - f\right\|_{\mathrm{W}^{m,p}} = \left(\sum_{|\alpha| \leq m} \left\|(\mathrm{D}^\alpha f) * \rho - \mathrm{D}^\alpha f\right\|_{\mathrm{L}^p}^p\right)^{\frac{1}{p}} \to 0,$$

which shows that we can approximate f with $\mathring{\mathrm{C}}^\infty(\Omega)$ functions.

For a general $f \in \mathrm{W}^{m,p}(\Omega)$ we will use the previous approximation locally. We define the following sets that approximate Ω

$$K_n = \overline{\mathrm{B}}_n(0) \setminus \mathrm{B}_{\frac{1}{n}}(\Omega^\complement) \quad \text{and} \quad \Omega_n = \mathrm{B}_n(0) \setminus \overline{\mathrm{B}}_{\frac{1}{n}}(\Omega^\complement).$$

We have $\Omega_n \subseteq K_n \subseteq \Omega_{n+1} \subseteq K_{n+1} \subseteq \cdots \subseteq \Omega$, where K_n is compact and Ω_n is open. Note that convolving $\mathbb{1}_{K_n + \mathrm{B}_\epsilon(0)}$ with $\rho_{\frac{\epsilon}{2}}$ for $\epsilon > 0$ sufficiently small yields a function h_n that satisfies

$$\operatorname{ran} h_n \subseteq [0,1], \quad h_n\big|_{K_n} = 1 \quad \text{and} \quad h_n \in \mathring{\mathrm{C}}^\infty(\Omega_{n+1}).$$

Therefore, $h_{n+1} - h_n$ vanishes on $K_n \supseteq \Omega_n$, which implies

$$\operatorname{supp}(h_{n+1} - h_n) \subseteq \Omega_{n+2} \setminus \Omega_n \subseteq \Omega_{n+1} \setminus K_{n-1}.$$

Note that $(h_{n+1} - h_n)f \in \mathrm{W}^{m,p}(\Omega)$ and has compact support $\operatorname{supp}(h_{n+1} - h_n)f \subseteq \Omega_{n+2} \setminus K_{n-1}$. Hence by the first part of the proof for a given $\epsilon > 0$ there exists a $k_n \in \mathbb{N}$ such that

$$\left\|(h_{n+1} - h_n)f - \left[(h_{n+1} - h_n)f\right] * \rho_{\frac{1}{k_n}}\right\|_{m,p} \leq \frac{\epsilon}{2^n}. \tag{3.1}$$

Since $\operatorname{supp}(h_{n+1} - h_n)f$ is contained in the open set $\Omega_{n+2} \setminus K_{n-1}$ we can choose $k_n \in \mathbb{N}$ even larger such that

$$\operatorname{supp}(h_{n+1} - h_n)f * \rho_{\frac{1}{k_n}} \subseteq \operatorname{supp}(h_{n+1} - h_n) + \overline{\mathrm{B}}_{\frac{1}{k_n}} \subseteq \Omega_{n+2} \setminus K_{n-1}.$$

Hence, both $(h_{n+1} - h_n)f$ and $(h_{n+1} - h_n)f * \rho_{\frac{1}{k_n}}$ live only on a small stripe that becomes thinner and converges to the boundary of $\partial\Omega$ as n grows. We define $h_0 = 0$ and

$$h_\epsilon := \sum_{n=0}^\infty (h_{n+1} - h_n)f * \rho_{\frac{1}{k_n}}.$$

For $x \in \Omega_\ell \subseteq K_\ell$ we have $\big((h_{n+1} - h_n)f * \rho_{k_n}\big)(x) = 0$, if $\ell \leq n-1$, which gives

$$h_\epsilon(x) = \underbrace{\sum_{n=0}^{\ell+1} \big[(h_{n+1} - h_n)f * \rho_{\frac{1}{k_n}}\big](x)}_{=:h_\epsilon^\ell(x)} \quad \text{for} \quad x \in \Omega_\ell.$$

By a telescoping series argument we have $f\big|_{\Omega_\ell} = h_{\ell+2} f\big|_{\Omega_\ell} = \sum_{n=0}^{\ell+1} (h_{n+1} - h_n) f\big|_{\Omega_\ell}$. Thus, for every $x \in \Omega$ the function h_ϵ is a finite sum of $\mathrm{C}^\infty(\Omega)$ functions and therefore C^∞ in x, which gives $h_\epsilon \in \mathrm{C}^\infty(\Omega)$. In particular, the finite sum h_ϵ^ℓ is even in $\mathring{\mathrm{C}}^\infty(\Omega)$, which implies $h_\epsilon^\ell \in \mathrm{W}^{m,p}(\Omega)$. Hence, for $|\alpha| \leq m$ we have

$$\begin{aligned}
&\left(\sum_{|\alpha| \leq m} \int_{\Omega_\ell} \big|\mathrm{D}^\alpha (f - h_\epsilon)\big|^p \,\mathrm{d}\lambda \right)^{\frac{1}{p}} \\
&\quad \leq \left(\sum_{|\alpha| \leq m} \int_{\Omega} \big|\mathrm{D}^\alpha (h_{\ell+2} f - h_\epsilon^\ell)\big|^p \,\mathrm{d}\lambda \right)^{\frac{1}{p}} = \|h_{\ell+2} f - h_\epsilon^\ell\|_{m,p} \\
&\quad = \left\| \sum_{n=0}^{\ell+1} (h_{n+1} - h_n) f - (h_{n+1} - h_n) f * \rho_{\frac{1}{k_n}} \right\|_{m,p} \\
&\quad \leq \sum_{n=0}^{\ell+1} \underbrace{\left\| (h_{n+1} - h_n) f - (h_{n+1} - h_n) f * \rho_{\frac{1}{k_n}} \right\|_{m,p}}_{\overset{(3.1)}{\leq} \frac{\epsilon}{2^n}} \leq \epsilon.
\end{aligned}$$

By monotone convergence for $\ell \to \infty$ we obtain $f - h_\epsilon \in \mathrm{W}^{m,p}(\Omega)$ and therefore $h_\epsilon \in \mathrm{W}^{m,p}(\Omega)$, and $h_\epsilon \to f$ in $\mathrm{W}^{m,p}(\Omega)$, which finishes the proof. ❑

So far we did not use any regularity assumption on the boundary $\partial\Omega$ of Ω. However, for some results the regularity of the boundary is crucial. Fortunately, we only need to impose relatively mild assumptions on the boundary to obtain the results we desire, namely, Ω should be a Lipschitz domain.

Moreover, since we are mostly interested in the Hilbert space case $p = 2$ we will in the following only regard $p = 2$. In this context it is very usual to denote the Sobolev space $\mathrm{W}^{m,2}(\Omega)$ by $\mathrm{H}^m(\Omega)$, i.e.,

$$\mathrm{H}^m(\Omega) := \mathrm{W}^{m,2}(\Omega).$$

Historically these two notations come from different approaches to the Sobolev space, which turned out to coincide. The space was used $\mathrm{H}^m(\Omega)$ for the closure of $\mathrm{C}^\infty(\Omega) \cap \mathrm{W}^{m,2}(\Omega)$ in $\mathrm{W}^{m,2}(\Omega)$, which coincides with $\mathrm{W}^{m,2}(\Omega)$ by Theorem 3.1.4.

In particular, we will focus on $m = 1$, i.e., $\mathrm{H}^1(\Omega)$, as this is the most relevant case for the following.

We define

$$\mathrm{C}^\infty(\overline{\Omega}) := \left\{ f\big|_\Omega \,\middle|\, f \in \mathring{\mathrm{C}}^\infty(\mathbb{R}^d) \right\},$$

i.e., $f \in \mathrm{C}^\infty(\overline{\Omega})$ is also at the boundary C^∞ in the sense that there exists a C^∞ extension to an open set that contains $\partial\Omega$.

Theorem 3.1.5. *Let $\Omega \subseteq \mathbb{R}^d$ be a Lipschitz domain. Then $\mathrm{C}^\infty(\overline{\Omega})$ is dense in $\mathrm{H}^1(\Omega)$.*

Contrary to $\mathrm{C}^\infty(\Omega)$ the space $\mathrm{C}^\infty(\overline{\Omega})$ is automatically a subspace of $\mathrm{H}^1(\Omega)$ as it only contains bounded functions with bounded derivatives. Hence, it the previous result is really an improvement of Theorem 3.1.4, in the sense that even a smaller set is dense in $\mathrm{H}^1(\Omega)$ under slightly stricter assumptions.

Theorem 3.1.6. *Let Ω be a bounded Lipschitz domain. Then the mapping*

$$\gamma_0 \colon \begin{cases} \mathrm{C}^\infty(\overline{\Omega}) & \to & \mathrm{L}^2(\partial\Omega), \\ \quad f & \mapsto & f\big|_{\partial\Omega}, \end{cases}$$

can be continuously extended to $\mathrm{H}^1(\Omega)$.

Sketch of a proof. We will only prove the assertion locally for flat boundaries. Let Ω be a rectangular cuboid and $f \in \mathrm{C}^\infty(\overline{\Omega})$ such that it vanishes on all flat boundary parts but one. W.l.o.g. this boundary part Γ corresponds

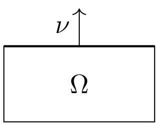

Figure 3.1: Cuboid with normal vector

to the normal vector $\nu\big|_\Gamma = e_d$. Now Stokes' theorem gives

$$\begin{aligned}\int_\Gamma |f|^2 \,\mathrm{d}\mu &= \int_{\partial\Omega} \nu_d |f|^2 \,\mathrm{d}\mu = \int_\Omega \partial_d |f|^2 \,\mathrm{d}\lambda = \int_\Omega \left[(\partial_d f)\overline{f} + f \partial_d \overline{f}\right] \mathrm{d}\lambda \\ &\overset{\text{C.–S.}}{\leq} 2\|f\|_{\mathrm{L}^2(\Omega)} \|\nabla f\|_{\mathrm{L}^2(\Omega)} \leq 2(\|f\|^2_{\mathrm{L}^2(\Omega)} + \|\nabla f\|^2_{\mathrm{L}^2(\Omega)}) = 2\|f\|^2_{\mathrm{H}^1(\Omega)}.\end{aligned}$$

For the general case we employ a partition of unity subordinate to cylinders that admit charts. Then the charts give Lipschitz diffeomorphisms of the cylinders to rectangular cuboids. With the transformed function we can proceed as shown. Finally we apply the inverse transformation to obtain the result. ❑

The range of the boundary trace γ_0 will play an important role for boundary value problems.

Definition 3.1.7. Let Ω be a bounded Lipschitz domain. Then we define

$$\mathrm{H}^{\frac{1}{2}}(\partial\Omega) := \operatorname{ran}\gamma_0 \quad \text{with} \quad \|f\|_{\mathrm{H}^{\frac{1}{2}}(\partial\Omega)} := \inf_{f=\gamma_0 F} \|F\|_{\mathrm{H}^1(\Omega)}. \qquad \diamond$$

Actually, defining $\mathrm{H}^{\frac{1}{2}}(\partial\Omega)$ by the range of γ_0 is a short cut. As the notation already suggests $\mathrm{H}^{\frac{1}{2}}(\partial\Omega)$ is a Sobolev space (with $p = 2$) with fractional order, i.e., it allows fractional derivatives. Defining these fractional Sobolev space is beyond the scope of this lecture. However, it turns out that the range of γ_0 coincides with this fractional Sobolev space and also the norms are equivalent. In fact $\mathrm{H}^{\frac{1}{2}}(\partial\Omega)$ is even a Hilbert space. Moreover, $\mathrm{H}^{\frac{1}{2}}(\partial\Omega)$ is dense in $\mathrm{L}^2(\partial\Omega)$ w.r.t. $\|\cdot\|_{\mathrm{L}^2(\partial\Omega)}$.

3.2 Weak formulation

In this section we are interested in the differential operator

$$(Lu)(x) = -\operatorname{div} A(x)\nabla u(x), \tag{3.2}$$

where $A\colon \Omega \to \mathbb{C}^{d\times d}$ is a measurable matrix-valued function such that there exists a $c > 0$ such that

$$c^{-1}\mathrm{I} \leq A(x) \leq c\mathrm{I} \quad \text{for a.e. } x \in \Omega$$

in the sense of positive definiteness. This implies that A is a.e. invertible and $A \in \mathrm{L}^\infty(\Omega; \mathbb{C}^{d\times d})$ and $A^{-1} \in \mathrm{L}^\infty(\Omega; \mathbb{C}^{d\times d})$. In order to derive the weak formulation of

$$\begin{aligned} Lu &= f, \\ u\big|_{\partial\Omega} &= 0, \end{aligned}$$

for a given $f \in \mathrm{L}^2(\Omega)$ (we will later see that we can even allow slightly more general f), we multiply the equation with the complex conjugated of a test function $\phi \in \mathring{\mathrm{C}}^\infty(\Omega)$ and integrate over Ω. This leads to

$$\int_\Omega Lu(x)\overline{\phi(x)}\,\mathrm{d}x = \int_\Omega f(x)\overline{\phi(x)}\,\mathrm{d}x$$

or in the setting of $\mathrm{L}^2(\Omega)$

$$\langle Lu, \phi\rangle_{\mathrm{L}^2(\Omega)} = \langle f, \phi\rangle_{\mathrm{L}^2(\Omega)}.$$

We will assume that the solution u is sufficiently regular to apply an integration by parts formula:

$$\begin{aligned}
&\int_\Omega Lu(x)\overline{\phi(x)}\,\mathrm{d}x \\
&\qquad = -\int_\Omega \operatorname{div} A(x)\nabla u(x)\overline{\phi(x)}\,\mathrm{d}x \\
&\qquad = \int_\Omega A(x)\nabla u(x)\cdot\nabla\overline{\phi(x)}\,\mathrm{d}x + \int_{\partial\Omega} \nu(x)\cdot A(x)\nabla u(x)\overline{\phi(x)}\,\mathrm{d}\mu(x) \\
&\qquad = \int_\Omega A(x)\nabla u(x)\cdot\nabla\overline{\phi(x)}\,\mathrm{d}x,
\end{aligned}$$

where the boundary integral vanishes because $\phi \in \mathring{\mathrm{C}}^\infty(\Omega)$. Hence, we have

$$\int_\Omega A(x)\nabla u(x)\cdot\overline{\nabla\phi(x)}\,\mathrm{d}x = \int_\Omega f(x)\overline{\phi(x)}\,\mathrm{d}x \quad \forall\phi\in\mathring{\mathrm{C}}^\infty(\Omega).$$

Again in the setting of $\mathrm{L}^2(\Omega)$ this can be written as

$$\langle A\nabla u, \nabla\phi\rangle_{\mathrm{L}^2(\Omega)} = \langle f,\phi\rangle_{\mathrm{L}^2(\Omega)} \quad \forall\phi\in\mathring{\mathrm{C}}^\infty(\Omega). \tag{3.3}$$

In order to make sense of this equation we have to ask $u \in \mathrm{H}^1(\Omega)$. Moreover, we want u to satisfy the boundary condition $u\big|_{\partial\Omega} = 0$, which leads to $u \in \mathring{\mathrm{H}}^1(\Omega)$. Note that (3.3) continuously depends on ϕ w.r.t. $\|\cdot\|_{\mathrm{H}^1(\Omega)}$. Hence, we can equivalently regard ϕ in the closure of $\mathring{\mathrm{C}}^\infty(\Omega)$ in $\mathrm{H}^1(\Omega)$, which is $\mathring{\mathrm{H}}^1(\Omega)$. Finally, this leads to the weak formulation.

Definition 3.2.1 (Weak formulation)**.** Let $L = \operatorname{div} A\nabla$, where A is a measurable matrix-valued function such that $c^{-1} \le A \le c$ for a $c > 0$ and $f \in \mathrm{L}^2(\Omega)$. Then the *weak formulation* of $Lu = f$ in Ω and $u = 0$ on $\partial\Omega$ is: Find a $u \in \mathring{\mathrm{H}}^1(\Omega)$ such that

$$\langle A\nabla u, \nabla\phi\rangle_{\mathrm{L}^2(\Omega)} = \langle f,\phi\rangle_{\mathrm{L}^2(\Omega)} \quad \forall\phi\in\mathring{\mathrm{H}}^1(\Omega). \tag{3.4}$$

We say $u \in \mathring{\mathrm{H}}^1(\Omega)$ is a *weak solution* of the problem, if u satisfies (3.4) ✧

Note that $(u,\phi) \mapsto \langle A\nabla u, \nabla\phi\rangle_{\mathrm{L}^2(\Omega)}$ is a continuous sesquilinear form on $\mathrm{H}^1(\Omega)$ and $\phi \mapsto \langle f,\phi\rangle_{\mathrm{L}^2(\Omega)}$ is a bounded antilinear mapping.

Theorem 3.2.2 (Lax–Milgram)**.** *Let H be a Hilbert space and $b\colon H\times H \to \mathbb{C}$ a bounded sesquilinear form. Then there exists a unique operator $B \in \mathcal{L}_\mathrm{b}(H)$ such that*

$$b(x,y) = \langle Bx, y\rangle_H \quad \forall x,y\in H$$

and $\|b\| = \|B\|$. Moreover, if b is coercive, i.e., there exists a $c > 0$ such that $b(x,x) \ge c\|x\|^2$, then B is positive, invertible and $\|B^{-1}\| \le \frac{1}{c}$.

Proof. Note that for a fixed $x \in H$ the mapping $y \mapsto b(x, y)$ is bounded and antilinear. Hence, by the representation theorem of Fréchet–Riesz there exists a unique z_x such that $b(x, y) = \langle z_x, y\rangle_H$. Since b is linear in the first argument, we have for $x_1, x_2 \in H$

$$\langle z_{x_1+\lambda x_2}, y\rangle_H = b(x_1+\lambda x_2, y) = b(x_1, y)+\lambda b(x_2, y) = \langle z_{x_1}, y\rangle_H+\lambda\langle z_{x_2}, y\rangle_H.$$

Thus, the uniqueness of z_x implies $z_{x_1+\lambda x_2} = z_{x_1} + \lambda z_{x_2}$. Hence, the mapping $B\colon H \to H$, $x \mapsto z_x$ is linear. Moreover,

$$\|B\| = \sup_{\|x\|=\|y\|=1} |\langle Bx, y\rangle_H| = \sup_{\|x\|=\|y\|=1} |b(x, y)| = \|b\|.$$

If b is additionally coercive, then we have

$$c\|x\|^2 \leq \langle Bx, x\rangle \leq \|Bx\|\|x\| \tag{3.5}$$

and consequently $\|Bx\| \geq c\|x\|$. This immediately gives that B is injective and has closed range. Moreover, (3.5) implies $\langle Bx, x\rangle \in \mathbb{R}_+$, which gives that B is self-adjoint and positive. The injectivity and the self-adjointness of B implies

$$\{0\} = \ker B = \ker B^* = (\operatorname{ran} B)^\perp.$$

Hence, $\overline{\operatorname{ran} B} = H$ and since $\operatorname{ran} B$ is closed, we conclude that B is surjective. Using $y = B^{-1}x$ in

$$\|By\| \geq c\|y\|$$

implies $\|B^{-1}\| \leq \frac{1}{c}$. ❑

3.3 The negative Sobolev space

We want to solve the weak formulation

$$\langle A\nabla u, \nabla\phi\rangle_{\mathrm{L}^2(\Omega)} = \langle f, \phi\rangle_{\mathrm{L}^2(\Omega)} \quad \forall \phi \in \mathring{\mathrm{H}}^1(\Omega)$$

for a given $f \in \mathrm{L}^2(\Omega)$. As already mentioned the left-hand-side of this equations is a sesquilinear form on $\mathring{\mathrm{H}}^1(\Omega) \times \mathring{\mathrm{H}}^1(\Omega)$ and the right-hand-side is bounded antilinear mapping. For a Hilbert space H it is very common to identify the (anti)dual space by H itself. However, in our case it comes more natural to choose a different representation for the (anti)dual space. As already discussed the mapping $\phi \mapsto \langle f, \phi\rangle_{\mathrm{L}^2(\Omega)}$ is in H', but f is only in $\mathrm{L}^2(\Omega)$ and therefore in general not in $\mathring{\mathrm{H}}^1(\Omega)$. Hence, every element $f \in \mathrm{L}^2(\Omega)$ represents an element of $(\mathring{\mathrm{H}}^1(\Omega))'$. Moreover, two different elements $f_1, f_2 \in \mathrm{L}^2(\Omega)$ also represent different elements in $(\mathring{\mathrm{H}}^1(\Omega))'$. This can be seen by:

$$\langle f_1 - f_2, \phi\rangle_{\mathrm{L}^2(\Omega)} = 0 \quad \forall \phi \in \mathring{\mathrm{H}}^1(\Omega)$$

implies that $f_1 - f_2 = 0$ by the density of $\mathring{\mathrm{H}}^1(\Omega)$ in $\mathrm{L}^2(\Omega)$. Hence, we can say

$$\mathring{\mathrm{H}}^1(\Omega) \subseteq \mathrm{L}^2(\Omega) \subseteq (\mathring{\mathrm{H}}^1(\Omega))',$$

which can lead to some confusion, as $(\mathring{\mathrm{H}}^1(\Omega))'$ can be identified with $\mathring{\mathrm{H}}^1(\Omega)$, by Riesz representation theorem. This could be used to come to the wrong conclusion that $\mathring{\mathrm{H}}^1(\Omega)$ equals $\mathrm{L}^2(\Omega)$. Therefore, it is important to point out that an identification cannot be replaced with equality in every context. In this particular situation the important detail is that $\mathring{\mathrm{H}}^1(\Omega)$ is identified with $(\mathring{\mathrm{H}}^1(\Omega))'$ with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathrm{H}^1(\Omega)}$ and $\mathrm{L}^2(\Omega)$ is identified with a subspace of $(\mathring{\mathrm{H}}^1(\Omega))'$ with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathrm{L}^2(\Omega)}$. Hence, $g \in \mathring{\mathrm{H}}^1(\Omega) \subseteq \mathrm{L}^2(\Omega)$ does not represent the same element in $(\mathring{\mathrm{H}}^1(\Omega))'$ by these two identification.

Since the inner product of $\mathrm{L}^2(\Omega)$ is much easier to work with, we prefer to characterize the dual space of $\mathring{\mathrm{H}}^1(\Omega)$ with respect to that inner product. This leads to the notion of *Gelfand triples* or *rigged Hilbert spaces*, see e.g., [11, Sec. 2.9] or [2, Sec. 8.1].

Definition 3.3.1. We define the space $\mathrm{H}^{-1}(\Omega)$ as the completion of $\mathrm{L}^2(\Omega)$ with respect to the norm

$$\|f\|_{\mathrm{H}^{-1}(\Omega)} \coloneqq \sup_{g\in\mathring{\mathrm{H}}^1(\Omega)\setminus\{0\}} \frac{|\langle f,g\rangle_{\mathrm{L}^2(\Omega)}|}{\|g\|_{\mathrm{H}^1(\Omega)}}.$$

Moreover we define for $f \in \mathrm{H}^{-1}(\Omega)$ and $g \in \mathring{\mathrm{H}}^1(\Omega)$ the following dual pairing

$$\langle f,g\rangle_{\mathrm{H}^{-1}(\Omega),\mathring{\mathrm{H}}^1(\Omega)} \coloneqq \lim_{n\to\infty}\langle f_n,g\rangle_{\mathrm{L}^2(\Omega)},$$

where $(f_n)_{n\in\mathbb{N}}$ is a sequence in $\mathrm{L}^2(\Omega)$ that converges to f w.r.t. $\|\cdot\|_{\mathrm{H}^{-1}(\Omega)}$. ✧

Note that by definition of the norm of $\mathrm{H}^{-1}(\Omega)$ we have for every $f \in \mathrm{L}^2(\Omega)$

$$\left\|\langle f,\cdot\rangle_{\mathrm{H}^{-1}(\Omega),\mathring{\mathrm{H}}^1(\Omega)}\right\|_{(\mathring{\mathrm{H}}^1(\Omega))'} = \left\|\langle f,\cdot\rangle_{\mathrm{L}^2(\Omega)}\right\|_{(\mathring{\mathrm{H}}^1(\Omega))'} = \|f\|_{\mathrm{H}^{-1}(\Omega)}.$$

Hence, The mapping $\Phi\colon f \mapsto \langle f,\cdot\rangle_{\mathrm{H}^{-1}(\Omega),\mathring{\mathrm{H}}^1(\Omega)}$ is isometric from $\mathrm{L}^2(\Omega)$ to $(\mathring{\mathrm{H}}^1(\Omega))'$, which automatically allows us to extend this mapping isometrically to $\mathrm{H}^{-1}(\Omega)$ by continuity. Consequently, every $f \in \mathrm{H}^{-1}(\Omega)$ represents an element in $(\mathring{\mathrm{H}}^1(\Omega))'$ with the same norm by the dual pairing. The next lemma shows that also the reverse statement holds.

Lemma 3.3.2. *For every $\psi \in (\mathring{\mathrm{H}}^1(\Omega))'$ there exists an $f \in \mathrm{H}^{-1}(\Omega)$ such that*

$$\psi(g) = \langle f,g\rangle_{\mathrm{H}^{-1}(\Omega),\mathring{\mathrm{H}}^1(\Omega)} \quad \forall g \in \mathring{\mathrm{H}}^1(\Omega)$$

and $\|\psi\|_{(\mathring{\mathrm{H}}^1(\Omega))'} = \|f\|_{\mathrm{H}^{-1}(\Omega)}$.

Proof. Note that $\iota\colon \mathring{\mathrm{H}}^1(\Omega) \to \mathrm{L}^2(\Omega)$ is continuous and injective. Moreover, for $f \in \mathrm{L}^2(\Omega)$ we have

$$\langle f, g\rangle_{\mathrm{H}^{-1}(\Omega),\mathring{\mathrm{H}}^1(\Omega)} = \langle f, g\rangle_{\mathrm{L}^2(\Omega)} = \langle f, \iota g\rangle_{\mathrm{L}^2(\Omega)} = \langle \iota^* f, g\rangle_{\mathrm{H}^1(\Omega)}. \tag{3.6}$$

Since the ι is a bounded and injective operator, we have $\overline{\operatorname{ran}\iota^*} = (\ker\iota)^\perp = \mathring{\mathrm{H}}^1(\Omega)$, which means $\operatorname{ran}\iota^*$ is dense. Furthermore, (3.6) implies $\|\iota^* f\|_{\mathrm{H}^1(\Omega)} = \|f\|_{\mathrm{H}^{-1}(\Omega)}$ for $f \in \mathrm{L}^2(\Omega)$. For given $\psi \in (\mathring{\mathrm{H}}^1(\Omega))'$ there exists, by Riesz representation theorem, an $h \in \mathring{\mathrm{H}}^1(\Omega)$ such that $\|\psi\|_{\mathring{\mathrm{H}}^1(\Omega)'} = \|h\|_{\mathrm{H}^1(\Omega)}$ and $\psi(g) = \langle h, g\rangle_{\mathrm{H}^1(\Omega)}$ for all $g \in \mathring{\mathrm{H}}^1(\Omega)$. Since ι^* is isometric and $\operatorname{ran}\iota^*$ is dense in $\mathring{\mathrm{H}}^1(\Omega)$, there exists a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathrm{L}^2(\Omega)$ such that $(\iota^* f_n)_{n\in\mathbb{N}}$ converges to g w.r.t. $\|\cdot\|_{\mathrm{H}^1(\Omega)}$. This also implies that $(f_n)_{n\in\mathbb{N}}$ converges to an $f \in \mathrm{H}^{-1}(\Omega)$ and

$$\begin{aligned}\langle f, g\rangle_{\mathrm{H}^{-1}(\Omega),\mathring{\mathrm{H}}^1(\Omega)} &= \lim_{n\to\infty} \langle f_n, g\rangle_{\mathrm{L}^2(\Omega)} \\ &= \lim_{n\to\infty} \langle \iota^* f_n, g\rangle_{\mathrm{H}^1(\Omega)} = \langle h, g\rangle_{\mathrm{H}^1(\Omega)} = \psi(g)\end{aligned}$$

for all $g \in \mathring{\mathrm{H}}^1(\Omega)$. Finally, $\|\psi\|_{\mathring{\mathrm{H}}^1(\Omega)'} = \|f\|_{\mathrm{H}^{-1}(\Omega)}$ follows from the previous equation. ❑

This lemma tells us that $\mathrm{H}^{-1}(\Omega)$ is a representation of the dual space of $\mathring{\mathrm{H}}^1(\Omega)$. The dual pairing of these two spaces is basically given by the $\mathrm{L}^2(\Omega)$ inner product (up to limits). That is why we say $\mathrm{H}^{-1}(\Omega)$ is the dual space of $\mathring{\mathrm{H}}^1(\Omega)$ with respect to the inner product of $\mathrm{L}^2(\Omega)$. The space $\mathrm{L}^2(\Omega)$ is called the *pivot space* in this context. The triple

$$\mathring{\mathrm{H}}^1(\Omega) \subseteq \mathrm{L}^2(\Omega) \subseteq \mathrm{H}^{-1}(\Omega)$$

is a so-called Gelfand triple.

Theorem 3.3.3 (Friedrich's inequality)**.** *Let Ω be a bounded domain. Then there exists a $C > 0$ such that*

$$\|f\|_{\mathrm{L}^2(\Omega)} \leq C\|\nabla f\|_{\mathrm{L}^2(\Omega)}$$

for all $f \in \mathring{\mathrm{H}}^1(\Omega)$.

Proof. We will show the inequality for $f \in \mathring{\mathrm{C}}^\infty(\Omega)$ and conclude it for $f \in \mathring{\mathrm{H}}^1(\Omega)$ by density. Moreover, we assume that f is real-valued and conclude the general result by splitting f into its real- and imaginary part.

Let $f \in \mathring{\mathrm{C}}^\infty(\Omega)$ be real-valued. Note that the function $x \mapsto x_1$ is in $\mathrm{C}^\infty(\Omega)$. Hence, $x \mapsto x_1 f(x)^2$ is also in $\mathring{\mathrm{C}}^\infty(\Omega)$. The product rule for derivatives gives

$$\int_\Omega |f|^2 \,\mathrm{d}\lambda = \int_\Omega \partial_1\big(x_1 f(x)^2\big)\,\mathrm{d}x - \int_\Omega x_1\partial_1\big(f(x)^2\big)\,\mathrm{d}x\,.$$

Note that

$$\int_\Omega \partial_1(x_1 f(x)^2)\,\mathrm{d}x = \int_\Omega \partial_1(x_1 f(x)^2)1\,\mathrm{d}x = -\int_\Omega x_1 f(x)^2 \partial_1 1\,\mathrm{d}x = 0.$$

By the boundedness of Ω there exists a $C > 0$ such that $2|x_1| \le C$ for $x \in \Omega$. Therefore, we have

$$\|f\|^2_{\mathrm{L}^2(\Omega)} = \int_\Omega |f|^2\,\mathrm{d}\lambda = -\int_\Omega x_1 2 f(x)\partial_1 f(x)\,\mathrm{d}x \le C\|f\|_{\mathrm{L}^2(\Omega)}\|\nabla f\|_{\mathrm{L}^2(\Omega)},$$

which immediately implies $\|f\|_{\mathrm{L}^2(\Omega)} \le C\|\nabla f\|_{\mathrm{L}^2(\Omega)}$. ❑

Corollary 3.3.4. *Let $L = -\operatorname{div} A\nabla$ the differential operator from (3.2). Then for every $f \in \mathrm{L}^2(\Omega)$ the partial differential equation*

$$Lu = f$$

has a unique weak solution $u \in \mathring{\mathrm{H}}^1(\Omega)$ that continuously depends on $f \in \mathrm{L}^2(\Omega)$, i.e., $\|u\|_{\mathrm{H}^1(\Omega)} \le c\|f\|_{\mathrm{L}^2(\Omega)}$.

Proof. This is just a consequence of Lax–Milgram Theorem 3.2.2 and Friedrich's inequality Theorem 3.3.3. ❑

Remark 3.3.5. Note that we can weaken the condtions on A and f such that we only assume that $A \in \mathcal{L}_\mathrm{b}(\mathrm{L}^2(\Omega)^d)$ is a positive and boundedly invertible operator and $f \in \mathrm{H}^{-1}(\Omega)$. Then the result still holds with where u depends continuously on $f \in \mathrm{H}^{-1}(\Omega)$, i.e., $\|u\|_{\mathrm{H}^1(\Omega)} \le c\|f\|_{\mathrm{H}^{-1}(\Omega)}$. ✧

In order the solve a boundary value problem

$$\begin{aligned} Lu &= f, \\ \gamma_0 u &= g, \end{aligned}$$

we will just shift u by a function G that satisfies $\gamma_0 G = g$. Then we have to solve

$$\begin{aligned} L(u - G) &= f - LG, \\ \gamma_0(u - G) &= 0. \end{aligned}$$

By substituting $\tilde{u} = (u - G)$ and $\tilde{f} = f - LG$ we arrive at our original homogeneous problem.

3.4 Galerkin method

The Galerkin method gives an abstract theory for numerical approaches to general problems of the form

$$b(u, v) = f(v) \quad \forall v \in H \tag{3.7}$$

for a sesquilinear form $b\colon H \times H \to \mathbb{C}$ and a (anti)linear functional $f \in H'$, where H is a Hilbert space.

In particular, we are motivated by the previous sections and therefore have $b(u,v) = \langle A\nabla u, \nabla v\rangle_{\mathrm{L}^2(\Omega)}$ and $f(v) = \langle f, v\rangle_{\mathrm{L}^2(\Omega)}$ for $u, v \in \mathring{\mathrm{H}}^1(\Omega)$ and $f \in \mathrm{L}^2(\Omega)$ in mind.

We will regard a family of finite dimensional subspaces $V_h \subseteq H$, where the index $h > 0$ is an indicator of how good V_h approximates H, i.e., the smaller h the larger V_h. In applications h is usually the mesh-size.

The idea is: instead of solving (3.7) for all $v \in H$ we just solve

$$b(u,v) = f(v) \quad \forall v \in V_h. \tag{3.8}$$

We denote the solution to this problem by u_h. Clearly, in general $u_h \neq u$, where u is the solution of (3.7). However, if the spaces V_h are choosen suitably, then $u_h \to u$ for $h \to 0$.

The problem (3.8) can be reduced to linear matrix equation, if we choose a basis of $\{v_1, \ldots, v_k\}$ of V_h: We can represent every vector as $\sum_{i=1}^k \alpha_i v_i$ and by linearity we reduce (3.8) to

$$b(u, v_j) = f(v_j) \quad \forall j \in \{1, \ldots, k\}.$$

Since also the solution $u_h \in V_h$ can be written as $u_h = \sum_{i=1}^k \beta_i v_i$ for unkown $\beta_1, \ldots, \beta_k$, we obtain the equation

$$\sum_{i=1}^k \beta_i b(v_i, v_j) = f(v_j).$$

Hence, we define the matrix $B_h = (b(v_i, v_j))_{i,j=1}^k$ and the vectors $f_h = (f(v_j))_{j=1}^k$ and $\beta = (\beta_i)_{i=1}^k$, which leads to

$$B_h \beta = f_h. \tag{3.9}$$

Proposition 3.4.1. *Let $b\colon H \times H \to \mathbb{C}$ a bounded and coercive sesquilinear form on the Hilbert space H. Then the equations (3.9) and (3.8) admit unique solutions.*

Proof. Note that for $0 \neq \alpha \in \mathbb{C}^k$ and $w = \sum_{i=1}^k \alpha_i v_i$ we have

$$\alpha^{\mathsf{H}} B_h \alpha = \sum_{i,j=1}^k \alpha_i b(v_i, v_j) \alpha_j = b(w,w) \geq c\|w\|_H^2 > 0.$$

Hence, B_h is a positive definite matrix and therefore invertible. ❑

Lemma 3.4.2 (Céa)**.** *Let $b\colon H \times H \to \mathbb{C}$ a bounded and coercive sesquilinear form on the Hilbert space H and $f \in H'$. Moreover, let $u \in H$ be the solution* *of* $b(u,v) = f(u)$ *for all* $v \in H$, V_h *a subspace of* H *and* $u_h \in V_h$ *the solution of* (3.8). *Then*

$$\|u - u_h\|_H \leq \frac{\|b\|}{c} \inf_{v \in V_h} \|u - v\|_H,$$

where $c > 0$ *is the coercivity constant of* b.

Proof. Note that $b(u,v) = f(v)$ and $b(u_h, v) = f(v)$ for all $v \in V_h$. Hence, $b(u - u_h, v) = 0$ for all $v \in V_h$. Since b is coercive and $v - u_h \in V_h$ we obtain

$$\begin{aligned} c\|u - u_h\|_H^2 &\leq b(u - u_h, u - u_h + v - v) \\ &= \underbrace{b(u - u_h, v - u_h)}_{=0} + b(u - u_h, u - v) \\ &\leq \|b\| \|u - u_h\|_H \|u - v\|_H \end{aligned}$$

for every $v \in V_h$. Hence,

$$\|u - u_h\| \leq \inf_{v \in V_h} \frac{\|b\|}{c} \|u - v\|_H.$$ ❑

Proposition 3.4.3. *Let the assumptions of Lemma 3.4.2 be satisfied and additionally*

$$\lim_{h \to 0} \operatorname{dist}(w, V_h) = 0 \quad \forall w \in H.$$

Then

$$\lim_{h \to 0} \|u - u_h\|_H = 0.$$

Recall that $\operatorname{dist}(w, V_h) = \inf_{v \in V_h} \|w - v\|_H$.

Proof. For given $\epsilon > 0$ there exists an $h_0 > 0$ such that $\operatorname{dist}(u, V_h) \leq \epsilon \frac{\|b\|}{c}$ for all $h \leq h_0$. Hence, by Lemma 3.4.2 we obtain

$$\|u - u_h\|_H \leq \frac{\|b\|}{c} \inf_{v \in V_h} \|u - v\|_H \leq \epsilon.$$ ❑

There is a lot more to say about the Galerkin method. So this short introduction to this topic should serve as a gateway drug. The interested reader is referred to [3]

Chapter 4

Strongly continuous semigroups

As motivation we recall Chapter 1, where we looked at the solution of a linear ordinary differential equation. Let A be a $n \times n$ matrix and $x_0 \in \mathbb{C}^n$ any initial vector. Then we regard the following differential equation (Cauchy problem)

$$\begin{aligned} \dot{x}(t) &= Ax(t), \quad t \in [0, +\infty) \\ x(0) &= x_0. \end{aligned}$$

The solution of this equation is given by $x(t) = \mathrm{e}^{tA} x_0$. The exponential function is not only defined for matrices, but also for bounded linear mappings on a Banach space. Hence, this approach to solve differential equations can easily extended to so called abstract Cauchy problems: Let X be a Banach space, A be a bounded linear mapping and $x_0 \in X$. Find a function $x\colon [0, +\infty) \to X$ such that

$$\begin{aligned} \dot{x}(t) &= Ax(t), \quad t \in [0, +\infty) \\ x(0) &= x_0. \end{aligned} \tag{4.1}$$

Again the solution is given by $x(t) = \mathrm{e}^{tA} x_0$.

However, we want to go even further and want to solve this abstract Cauchy problem for unbounded operators. For unbounded operators the exponential function is harder to define or not even possible, but if A satisfies a few conditions we can find something that carries the essence to solve the abstract Cauchy problem.

A very rich source for strongly continuous semigroups is [4].

4.1 Strongly continuous semigroups

Definition 4.1.1. Let X be a Banach space and $T\colon [0,+\infty) \to \mathcal{L}_\mathrm{b}(X)$. We say T is a *strongly continuous semigroup* or C_0*-semigroup*, if

- $T(0) = \mathrm{I}$,
- $T(t+s) = T(t)T(s)$ for all $t, s \in [0,+\infty)$,
- and $t \mapsto T(t)x$ is continuous for every $x \in X$, i.e., T is strongly continuous.

✧

Note that it is actually enough to ask for T is strongly continuous in 0, as $T(t+s) = T(t)T(s)$ then already implies that T is strongly continuous in every $t \in [0,+\infty)$.

By the properties of the exponential function we can see that $T(t) := \mathrm{e}^{tA}$, for $A \in \mathcal{L}_\mathrm{b}(X)$, is a C_0-semigroup.

Lemma 4.1.2. *Let T be a strongly continuous semigroup. Then there exists an $M \geq 1$ and an $\omega \in \mathbb{R}$ such that*

$$\|T(t)\| \leq M\mathrm{e}^{\omega t} \quad \text{for all} \quad t \in [0,+\infty).$$

Proof. First we will show that there is an $\epsilon > 0$ such that $\|T(t)\|$ is uniformly bounded for $t \in [0,\epsilon]$:

Let us assume that this is not true. Then for each $n \in \mathbb{N}$ there exists a $t_n \in [0, \frac{1}{n}]$ such that

$$\|T(t_n)\| \geq n. \tag{4.2}$$

Since $(t_n)_{n\in\mathbb{N}}$ converges to 0 and T is strongly continuous we have $T(t_n)x \to x$ for all $x \in X$. Consequently, the set $\{T(t_n)x \,|\, n \in \mathbb{N}\}$ is bounded in X for every $x \in X$. The principle of uniform boundedness implies that the set $\{T(t_n) \,|\, n \in \mathbb{N}\}$ is bounded in $\mathcal{L}_\mathrm{b}(X)$, which contradicts (4.2). Thus there exists an $\epsilon > 0$ such that $\|T(t)\| \leq M$ on $[0,\epsilon]$.

We can write every $t = n\epsilon + \delta$, where $\delta < \epsilon$ and $n \in \mathbb{N}$ ($n = \lfloor \frac{t}{\epsilon} \rfloor$). This leads to

$$\|T(t)\| = \|T(n\epsilon + \delta)\| = \|T(\epsilon)^n T(\delta)\| \leq M^n M \leq M M^{\frac{t}{\epsilon}} = M\mathrm{e}^{\frac{1}{\epsilon}\ln(M)t}.$$

Defining ω as $\frac{1}{\epsilon}\ln(M)$ finishes the proof. ❑

The definition of a strongly continuous semigroup is motivated by the matrix exponential function $t \mapsto \mathrm{e}^{tA}$ for a matrix (or a bounded linear operator) A. Now we go somehow the opposite direction by deducing a linear operator from the semigroup. In general this will not be a bounded operator.

Definition 4.1.3. Let T be a strongly continuous semigroup on a Banach space X. We define its *infinitesimal generator* by

$$A := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in X \times X \,\middle|\, y = \lim_{t\to 0} \frac{T(t)x - x}{t} \right\}.$$

✧

Note that the infinitesimal generator A is an operator (i.e., $\operatorname{mul} A = \{0\}$), since limits in Hausdorff spaces are unique. So for $x \in \operatorname{dom} A$ we can also write

$$Ax = \lim_{t\to 0} \frac{T(t)x - x}{t}.$$

Lemma 4.1.4. *Let T be a strongly continuous semigroup, A its infinitesimal generator and $x \in \operatorname{dom} A$. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(T(t)x\big) = T(t)Ax = AT(t)x.$$

In particular $T(t)x \in \operatorname{dom} A$, if $x \in \operatorname{dom} A$.

Proof. Note that for fixed t the operator $T(t)$ is continuous. Therefore,

$$\begin{aligned}\lim_{s\to 0+} \frac{T(t+s)x - T(t)x}{s} &= \lim_{s\to 0+} \frac{T(t)T(s)x - T(t)x}{s} \\ &= T(t) \lim_{s\to 0+} \frac{T(s)x - x}{s} = T(t)Ax.\end{aligned}$$

On the other hand, we have to check the limit from the left hand side, which we can rewrite as a right hand side limit

$$\lim_{s\to 0-} \frac{T(t+s)x - T(t)x}{s} = \lim_{s\to 0+} \frac{T(t)x - T(t-s)x}{s}.$$

Hence, we have to check whether the limit agrees with $T(t)Ax$. Note that $T(t) \le M\mathrm{e}^{\omega t}$ (by Lemma 4.1.2) and that T is strongly continuous:

$$\begin{aligned}&\left\| T(t-s)\frac{T(s)x - x}{s} - T(t)Ax \right\| \\ &\quad\le \left\| T(t-s)\frac{T(s)x - x}{s} - T(t-s)Ax \right\| + \|T(t-s)Ax - T(t)Ax\| \\ &\quad\le M\mathrm{e}^{\omega(t-s)} \underbrace{\left\| \frac{T(s)x - x}{s} - Ax \right\|}_{\to 0} + \underbrace{\|T(t-s)Ax - T(t)Ax\|}_{\to 0},\end{aligned}$$

where we used the strong continuity of T for the second summand. Hence, $\frac{\mathrm{d}}{\mathrm{d}t}T(t)x$ exists and equals $T(t)Ax$, which implies $T(t)x \in \operatorname{dom} A$. Finally,

$$\lim_{s\to 0+} \frac{T(t+s)x - T(t)x}{s} = \lim_{s\to 0+} \frac{T(s)T(t)x - T(t)x}{s} = AT(t)x.$$

For the limit from left hand side we obtain the same, since we have already shown that the limits agree. ❑

Lemma 4.1.5. *Let T be strongly continuous semigroup, A its infinitesimal generator and $x \in \mathrm{C}^1([0,+\infty);X)$ such that $x(t) \in \operatorname{dom} A$ for every $t \in [0,+\infty)$. Then we have the following product rule*

$$\frac{\mathrm{d}}{\mathrm{d}t} T(t)x(t) = AT(t)x(t) + T(t)x'(t).$$

Proof. First we will show that for a function $y\colon (-\epsilon,\epsilon) \to X$ that is continuous in 0 we have

$$T(t+s)y(s) \to T(t)y(0).$$

For fixed t and $s \in (-\epsilon,\epsilon)$ we have the estimate $\|T(t+s)\| \le M\mathrm{e}^{\omega(t+\epsilon)} =: C$ from Lemma 4.1.2. Hence,

$$\begin{aligned}
&\|T(t+s)y(s) - T(t)y(0)\| \\
&\qquad \le \|T(t)\|\|T(s)y(s) - y(0)\| \\
&\qquad \le \|T(t)\|\Big(\|T(s)y(s) - T(s)y(0)\| + \|T(s)y(0) - y(0)\|\Big) \\
&\qquad \le \|T(t)\|\Big(\|T(s)\|\|y(s) - y(0)\| + \|T(s)y(0) - y(0)\|\Big) \\
&\qquad \le C^2\|y(s) - y(0)\| + C\|T(s)y(0) - y(0)\|,
\end{aligned}$$

which shows $T(t+s)y(s) \to T(t)y(0)$ for $s \to 0$. For $t = 0$ we only regard $s \in [0,\epsilon)$ and the limit $s \to 0+$.

Finally, we have

$$\begin{aligned}
&\frac{T(t+s)x(t+s) - T(t)x(t)}{s} \\
&\qquad = \frac{T(t+s)x(t+s) - T(t+s)x(t) + T(t+s)x(t) - T(t)x(t)}{s} \\
&\qquad = T(t+s)\frac{x(t+s) - x(t)}{s} + \frac{T(t+s)x(t) - T(t)x(t)}{s} \\
&\qquad \to T(t)x'(t) + AT(t)x(t),
\end{aligned}$$

where we used the first part of the proof for the first summand and Lemma 4.1.4 for the second summand. ❑

Now let T be a strongly continuous semigroup, A its infinitesimal generator and $x_0 \in \operatorname{dom} A$. Then the abstract Cauchy problem

$$\begin{aligned}
\dot{x}(t) &= Ax(t), \quad t \in [0,+\infty), \\
x(0) &= x_0,
\end{aligned}$$

is solved by $x(t) \coloneqq T(t)x_0$, as

$$\dot{x}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\big(T(t)x_0\big) = AT(t)x_0 = Ax(t)$$

and $x(0) = T(0)x_0 = x_0$. The next proposition shows that this is even the only solution. Hence, it is of particular interest to decide whether a given operator A is an infinitesimal generator of a strongly semigroup, because then the corresponding Cauchy problem is solvable.

Proposition 4.1.6. *Let A be a generator of a strongly continuous semigroup T. Then $x := T(\cdot)x_0$ is the unique solution of the corresponding Cauchy problem in $\mathrm{C}^1([0, +\infty); X)$.*

Proof. Let $y \in \mathrm{C}^1([0, +\infty), X)$ be a solution of the Cauchy problem with $y(0) = x_0$. Note that $y(s) \in \operatorname{dom} A$ (otherwise it cannot be plugged in the differential equation) and hence $AT(t)y(s) = T(t)Ay(s)$. For $s > t$ we have, by Lemma 4.1.5,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}T(t)y(s-t) &= AT(t)y(s-t) - T(t)y'(s-t) \\ &= T(t)Ay(s-t) - T(t)Ay(s-t) = 0.\end{aligned}$$

Thus, $t \mapsto T(t)y(s-t)$ is constant and therefore

$$T(s)x_0 = T(s)y(s-s) = T(0)y(s-0) = y(s).$$ ❑

Therefore, it is natural to ask when a linear operator A is an infinitesimal generator of a strongly continuous semigroup. We will characterize this by Theorem 4.2.4. However, the conditions in Theorem 4.2.4 are in general also not that easy to check. Hence, we will give an easy and sufficient condition in Theorem 4.3.10.

We can even extend the solution term for initial conditions that are not in $\operatorname{dom} A$. This will lead to the notion of *mild solutions*, see Definition 4.4.1

Lemma 4.1.7. *Let A be the infinitesimal generator of a strongly continuous semigroup T on a Banach space X and $\lambda \in \mathbb{C}$. Then $T_\lambda \colon [0, +\infty) \to \mathcal{L}_\mathrm{b}(X)$ given by $T_\lambda(t) := \mathrm{e}^{\lambda t}T(t)$ is a strongly continuous semigroup with infinitesimal generator $A + \lambda$.*

Proof. For $t = 0$ we have $T_\lambda(0) = \mathrm{e}^0 T(0) = \mathrm{I}$. The semigroup property follows from the semigroup property of $\mathrm{e}^{\lambda \cdot}$ and T:

$$T_\lambda(t+s) = \mathrm{e}^{\lambda(t+s)}T(t+s) = \mathrm{e}^{\lambda t}T(t)e^{\lambda s}T(s) = T_\lambda(t)T_\lambda(s).$$

By the continuity of $t \mapsto \mathrm{e}^{\lambda t}$ and $t \mapsto T(t)x$ we conclude the continuity of $t \mapsto T_\lambda(t)x$ for every fixed $x \in X$. Hence, T_λ is a strongly continuous semigroup.

By the product rule for semigroups (Lemma 4.1.5 with $x(t) = \mathrm{e}^{\lambda t}x$) we have for $x \in X$

$$\frac{\mathrm{d}}{\mathrm{d}t}(T_\lambda(t)x) = \frac{\mathrm{d}}{\mathrm{d}t}(T(t)\mathrm{e}^{\lambda t}x) = AT(x)\mathrm{e}^{\lambda t}x + \lambda T(t)\mathrm{e}^{\lambda t}x = (A+\lambda)T(t)x.$$

Hence, for $t = 0$ we get that $(A + \lambda)$ is subset of the infinitesimal generator A_λ of T_λ. Applying the same argument to $(T_\lambda)_{-\lambda}(t) = \mathrm{e}^{-\lambda t}T_\lambda(t) = T(t)$ we obtain that $A_\lambda - \lambda$ is a subset of A. Therefore, $A_\lambda = A + \lambda$. ❑

Lemma 4.1.8. *Let A be an infinitesimal generator of a strongly continuous semigroup T on a Banach space X. Then the following holds:*

(i) *For every $x \in X$ and $s, h \geq 0$ we have*

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(T(s)\int_0^h T(t)x\,\mathrm{d}t\Big) = T(s)(T(h)x - x).$$

In particular for $s = 0$ we conclude $\int_0^h T(t)x\,\mathrm{d}t \in \operatorname{dom} A$.

(ii) *For every $t \geq 0$ we have*

$$\begin{aligned} T(t)x - x &= A\int_0^t T(s)x\,\mathrm{d}s, \quad \textit{for } x \in X,\\ \textit{and}\quad T(t)x - x &= \int_0^t T(s)Ax\,\mathrm{d}s, \quad \textit{for } x \in \operatorname{dom} A. \end{aligned}$$

(iii) *A is a densely defined and closed operator.*

Proof. We show the assertion in the listed order:

(i) By the fundamental theorem of calculus we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}\Big(T(s)\int_0^h T(t)x\,\mathrm{d}t\Big) &= \frac{\mathrm{d}}{\mathrm{d}s}\int_0^h T(t+s)x\,\mathrm{d}t = \frac{\mathrm{d}}{\mathrm{d}s}\int_s^{h+s} T(t)x\,\mathrm{d}t\\ &= T(h+s)x - T(s)x = T(s)(T(h)x - x). \end{aligned}$$

This identity for $s = 0$ implies $\int_0^h T(t)x\,\mathrm{d}t \in \operatorname{dom} A$.

(ii) The previous identity for $s = 0$ gives

$$T(h)x - x = \Big(\frac{\mathrm{d}}{\mathrm{d}s}\Big(T(s)\int_0^h T(t)x\,\mathrm{d}t\Big)\Big)\Big|_{s=0} = A\int_0^h T(t)x\,\mathrm{d}t$$

For $x \in \operatorname{dom} A$ we further calculate

$$\begin{aligned} A\int_0^h T(t)x\,\mathrm{d}t &= \lim_{s\to 0}\frac{T(s)\int_0^h T(t)x\,\mathrm{d}t - \int_0^h T(t)x\,\mathrm{d}t}{s}\\ &= \int_0^h \lim_{s\to 0}\frac{T(t+s)x - T(t)x}{s}\,\mathrm{d}t\\ &= \int_0^h T(t)Ax\,\mathrm{d}t, \end{aligned}$$

where we used dominated converges for the interchange of the limit and the integral.

(iii) Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{dom} A$ that converges to $x \in X$ w.r.t. $\|\cdot\|_X$ such that $(Ax_n)_{n\in\mathbb{N}}$ converges to $y \in X$ w.r.t. $\|\cdot\|_X$. By item (ii) we have

$$\frac{T(h)x_n - x_n}{h} = \frac{1}{h}\int_0^h T(t)Ax_n \,\mathrm{d}t$$

By dominated convergence we have for $n \to \infty$:

$$\frac{T(h)x - x}{h} = \frac{1}{h}\int_0^h T(t)y \,\mathrm{d}t\,.$$

Finally, by continuity of $t \mapsto T(t)y$ the right-hand-side converges to y and we obtain

$$\lim_{h\to 0}\frac{T(h)x - x}{h} = y,$$

which shows $x \in \operatorname{dom} A$ and $Ax = y$, and therefore the closedness of A.

For arbitrary $x \in X$ with have, by item (i), that $\int_0^h T(t)x \,\mathrm{d}t \in \operatorname{dom} A$. Again the continuity of $t \mapsto T(t)x$ implies

$$\lim_{h\to 0}\frac{1}{h}\int_0^h T(t)x \,\mathrm{d}t = x.$$

Hence, $\operatorname{dom} A$ is dense in X. ❑

Note that the previous lemma item (ii) says in the language of linear relations that for $x \in X$ and $h \geq 0$

$$\begin{bmatrix} \int_0^h T(t)x \,\mathrm{d}t \\ T(h)x - x \end{bmatrix} \in A.$$

Roughly speaking, if the semigroup converges to 0 for $h \to \infty$ we obtain

$$\begin{bmatrix} \int_0^\infty T(t)x \,\mathrm{d}t \\ -x \end{bmatrix} \in A \quad \text{and in turn} \quad \begin{bmatrix} x \\ -\int_0^\infty T(t)x \,\mathrm{d}t \end{bmatrix} \in A^{-1}$$

Hence,

$$A^{-1}x = -\int_0^\infty T(t)x \,\mathrm{d}t\,.$$

Since we can always modify the semigroup by multiply by $\mathrm{e}^{\lambda t}$ such that the semigroup decays exponentially, we can force this behavior for a modified semigroup.

Lemma 4.1.9. *Let T be a strongly continuous semigroup on a Banach space X and A its infinitesimal generator. Moreover, let $M, \omega \in \mathbb{R}$ be such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$. Then for $\operatorname{Re}\lambda > \omega$ we have $\lambda \in \rho(A)$ and*

$$(A - \lambda)^{-1}x = \int_0^\infty \mathrm{e}^{-\lambda t}T(t)x \,\mathrm{d}t\,.$$

In particular, $\mathbb{C}_{\mathrm{Re}>\omega} \subseteq \rho(A)$*, where* $\mathbb{C}_{\mathrm{Re}>\omega} = \{\lambda \in \mathbb{C} \mid \mathrm{Re}\,\lambda > \omega\}$*. Furthermore, the following estimate holds*

$$\|(A-\lambda)^{-n}\| \leq \frac{M}{(\mathrm{Re}\,\lambda - \omega)^n}$$

for $\lambda \in \mathbb{C}_{\mathrm{Re}>\omega}$*.*

Proof. By Lemma 4.1.7 we know that $\mathrm{e}^{-\lambda\cdot}T(\cdot)$ is a semigroup with generator $A-\lambda$. Hence, Lemma 4.1.8 item (ii) applied on $\mathrm{e}^{-\lambda\cdot}T(\cdot)$ gives

$$\mathrm{e}^{-\lambda h}T(h)x - x = (A-\lambda)\int_0^h \mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t\,.$$

Note that by assumption we have the estimate $\|\mathrm{e}^{-\lambda t}T(t)\| \leq M\mathrm{e}^{(\omega-\mathrm{Re}\,\lambda)t}$. Hence, the semigroup converges exponentially to 0, which implies $\mathrm{e}^{-\lambda h}T(h)x - x \to -x$ and $\int_0^h \mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t \to \int_0^\infty \mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t$ for $h \to +\infty$. Since $A-\lambda$ is closed, we conclude

$$-x = (A-\lambda)\int_0^\infty \mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t$$

or equivalently

$$(A-\lambda)^{-1}x = -\int_0^\infty \mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t\,.$$

Finally, we show the last assertion. For $n=1$ we have

$$\begin{aligned}\|(A-\lambda)^{-1}x\| &\leq \int_0^\infty \|\mathrm{e}^{-\lambda t}T(t)x\|\,\mathrm{d}t \leq \int_0^\infty M\mathrm{e}^{(\omega-\mathrm{Re}\,\lambda)t}\,\mathrm{d}t\|x\| \\ &= \|x\|\frac{M}{\omega-\mathrm{Re}\,\lambda}\mathrm{e}^{(\omega-\mathrm{Re}\,\lambda)t}\Big|_0^\infty = \frac{\|x\|M}{\mathrm{Re}\,\lambda-\omega}.\end{aligned}$$

By induction we can show that

$$(A-\lambda)^{-n}x = \int_0^\infty \cdots \int_0^\infty \mathrm{e}^{-\lambda(t_1+\cdots+t_n)}T(t_1+\cdots+t_n)\,\mathrm{d}t_1\cdots\mathrm{d}t_n\,.$$

By repeating the previous strategy of $n=1$ we obtain

$$\begin{aligned}\|(A-\lambda)^{-n}x\| &\leq M\|x\|\int_0^\infty \cdots \int_0^\infty \mathrm{e}^{(\omega-\mathrm{Re}\,\lambda)(t_1+\cdots+t_n)}\,\mathrm{d}t_1\cdots\mathrm{d}t_n \\ &= M\|x\|\left(\int_0^\infty \mathrm{e}^{(\omega-\lambda)t}\,\mathrm{d}t\right)^n \leq \frac{M\|x\|}{(\mathrm{Re}\,\lambda-\omega)^n}.\end{aligned}$$ ❑

Note that the integral $\int_0^\infty \mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t$ can be seen as the Laplace transform of $T(t)x$. Hence, it is not completely surprising that we get the inverse of $(A-\lambda)$, as this exactly what happens for the Laplace transform of e^{tA}, if A is a matrix (or a bounded operator).

The uniqueness of solutions of the abstract Cauchy problem in Proposition 4.1.6 also implies that the infinitesimal generator uniquely determines the strongly continuous semigroup.

Corollary 4.1.10. *Let T_1 and T_2 be strongly continuous semigroups on a Banach space X with infinitesimal generators A_1 and A_2, respectively. If $A_1 \subseteq A_2$, then $A_1 = B_2$ and $T_1 = T_2$.*

Or in other words: two different strongly semigroups have different infinitesimal generators.

Proof. Let $x \in \operatorname{dom} A$. Then $x_1(t) := T_1(t)x$ and $x_2 := T_2(t)x$ are both solutions of the Cauchy problem $\dot{x} = A_1 x$ and $x(0) = x$, because $A_1 \subseteq A_2$. Hence, $x_1 = x_2$ and therefore $T_1(t)x = T_2(t)x$ for all $t \geq 0$ and $x \in \operatorname{dom} A_1$. Since $\operatorname{dom} A_1$ is dense in X and both $T_1(t)$ and $T_2(t)$ are bounded operators they have to coincide. Hence, also A_1 and A_2 have to coincide. ❑

Proposition 4.1.11. *Let T be a strongly continuous semigroup on a Banach space X and A its infinitesimal generator. Then $T\colon [0,+\infty) \to \mathcal{L}_\mathrm{b}(X)$ is norm continuous (not only strongly continuous), if and only if A is bounded, i.e., $A \in \mathcal{L}_\mathrm{b}(X)$.*

Proof. Recall that for a boundedly invertible operator $P \in \mathcal{L}_\mathrm{b}(X)$ all elements $Q \in \mathcal{L}_\mathrm{b}(X)$ that satisfy $\|P - Q\| \leq \frac{1}{\|P^{-1}\|}$ are boundedly invertible (this is a Neumann series argument).

$\boxed{\Rightarrow}$: Let T be norm continuous. Then the $\mathcal{L}_\mathrm{b}(X)$-valued integral $\int_0^h T(t)\,\mathrm{d}t$ exists and we have

$$\left(\int_0^h T(t)\,\mathrm{d}t\right)x = \int_0^h T(t)x\,\mathrm{d}t \quad \forall x \in X,$$

because point evaluation is continuous and linear (so it commutes with limits and Riemann sums).

By the continuity in 0 there exists a $\delta > 0$ such that $\|\mathrm{I} - T(t)\| \leq \frac{1}{2}$ for all $t \in [0,\delta]$. Hence,

$$\left\|\delta\mathrm{I} - \int_0^\delta T(t)\,\mathrm{d}t\right\| \leq \left\|\int_0^\delta \mathrm{I} - T(t)\,\mathrm{d}t\right\| \leq \int_0^\delta \|\mathrm{I} - T(t)\|\,\mathrm{d}t \leq \frac{\delta}{2}$$

and consequently $\|\mathrm{I} - \frac{1}{\delta}\int_0^\delta T(t)\,\mathrm{d}t\| \leq \frac{1}{2}$. This implies that $\frac{1}{\delta}\int_0^\delta T(t)\,\mathrm{d}t$ is boundedly invertible and in turn also $\int_0^\delta T(t)\,\mathrm{d}t$. By Lemma 4.1.8 item (ii) we have

$$(T(\delta) - \mathrm{I}) = A\int_0^\delta T(t)\,\mathrm{d}t\,.$$

Since $\int_0^\delta T(t)\,\mathrm{d}t$ is boundedly invertible we conclude

$$(T(\delta) - \mathrm{I})\left(\int_0^\delta T(t)\,\mathrm{d}t\right)^{-1} = A$$

and since the left-hand-side is bounded, also A is bounded.

$\boxed{\Leftarrow}$: If A is bounded, then A is the infinitesimal generator of $t \mapsto \mathrm{e}^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}$, which is norm continuous as the series converges absolutely for t in a bounded interval. Since the T and $\mathrm{e}^{\cdot A}$ have the same infinitesimal generator, they coincide by Corollary 4.1.10. ❑

4.2 Hille–Yosida theorem

Lemma 4.2.1. *Let X, Y be Banach spaces, $D \subseteq X$ dense and $(T_i)_{i\in I}$ a net in $\mathcal{L}_\mathrm{b}(X,Y)$ with*

$$\limsup_{i\in I} \|T_i\| < +\infty.$$

If $T_i x$ converges for all $x \in D$, then T_i converges strongly to a $T \in \mathcal{L}_\mathrm{b}(X,Y)$ with $\|T\| \leq \limsup_{i\in I} \|T_i\|$.

Proof. We define Tx for $x \in D$ as $\lim_{i\in I} T_i x$ and we set $C = \limsup_{i\in I}\|T_i\| + 1$. Then we have $\|Tx\| \leq \limsup_{i\in I}\|T_i\|\|x\|$. Hence, we can continuously extend T on X such that $\|T\| \leq \limsup_{i\in I}\|T_i\|$. For every $x \in X$ there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in D that converges to x in X. For given $\epsilon > 0$ we choose $n_0 \in \mathbb{N}$ so large that $\|x_{n_0} - x\| \leq \frac{\epsilon}{C}$ and $i_0 \in I$ so large that $\|T_i x_{n_0} - T x_{n_0}\| \leq \epsilon$ and $\|T_i\| \leq C$ for all $i \geq i_0$. Then we have for all $i \geq i_0$

$$\begin{aligned}
\|T_i x - Tx\| &= \|T_i x - T_i x_{n_0} + T_i x_{n_0} - T x_{n_0} + T x_{n_0} - Tx\| \\
&\leq \|T_i x - T_i x_{n_0}\| + \|T_i x_{n_0} - T x_{n_0}\| + \|T x_{n_0} - Tx\| \\
&\leq \|T_i\|\|x - x_{n_0}\| + \epsilon + \|T\|\|x - x_{n_0}\| \leq 3\epsilon.
\end{aligned}$$

Hence, $T_i x$ converges to Tx for all $x \in X$. ❑

Definition 4.2.2. Let X be a Banach space and A be a densely defined and closed operator on X such that $(\omega, +\infty) \subseteq \rho(A)$ for some $\omega \in \mathbb{R}$. Then we define the *Yosida approximant* of A by

$$A_\lambda \coloneqq -\lambda\big(\lambda(A-\lambda)^{-1} + \mathrm{I}\big) \quad \text{for} \quad \lambda > \omega. \qquad ✧$$

Note that A_λ is bounded, i.e., $A_\lambda \in \mathcal{L}_\mathrm{b}(X)$. Moreover, there is also an alternative representation for the Yosida approximant:

$$\begin{aligned}
A_\lambda &= -\lambda\big(\lambda(A-\lambda)^{-1} + \mathrm{I}\big) = -\lambda\big(\lambda(A-\lambda)^{-1} + (A-\lambda)(A-\lambda)^{-1}\big) \\
&= -\lambda\big(\lambda + (A-\lambda)\big)(A-\lambda)^{-1} = -A\lambda(A-\lambda)^{-1}.
\end{aligned}$$

Heuristically, $(A-\lambda)^{-1}$ behaves like $-\frac{1}{\lambda}$ for large λ, therefore we expect $-\lambda(A-\lambda)^{-1}$ to converge to I in some sense for $\lambda \to +\infty$.

If we restrict the Yosida approximant on the domain of A, then we even have

$$A_\lambda x = -\lambda(A-\lambda)^{-1}Ax \quad \text{for} \quad x \in \operatorname{dom} A.$$

The next result will justify the *approximant* part in the name of the Yosida approximant.

Lemma 4.2.3. *Let X be a Banach space and A be a densely defined and closed operator on X such that $(\omega, +\infty) \subseteq \rho(A)$ for some $\omega \in \mathbb{R}$. Moreover, let $\|(A-\lambda)^{-1}\| \leq \frac{M}{(\lambda-\omega)}$ for some $M > 0$ and $\lambda \in (\omega, \infty)$. Then*

$$\lim_{\lambda\to\infty} A_\lambda x = Ax \quad \textit{for all} \quad x \in \operatorname{dom} A.$$

Note that every infinitesimal generator satisfies the condition of the lemma.

Proof. First we will show that $-\lambda(A-\lambda)^{-1}$ converges strongly to the identity operator I for $\lambda \to +\infty$. For $x \in \operatorname{dom} A$ we have

$$\begin{aligned}
\|-\lambda(A-\lambda)^{-1}x - x\| &= \|-\lambda(A-\lambda)^{-1}x - (A-\lambda)^{-1}(A-\lambda)x\| \\
&= \|(A-\lambda)^{-1}(-\lambda x - (A-\lambda)x)\| = \|(A-\lambda)^{-1}Ax\| \\
&\leq \frac{M}{\lambda-\omega}\|Ax\| \to 0
\end{aligned}$$

Moreover, we have $\limsup_{\lambda\to\infty}\|-\lambda(A-\lambda)^{-1}\| \leq \limsup_{\lambda\to\infty}\frac{M|\lambda|}{(\lambda-\omega)} = M$. Hence, by Lemma 4.2.1 the net $\big(-\lambda(A-\lambda)^{-1}\big)_{\lambda\in(\omega,\infty)}$ converges strongly to I.

Since $A_\lambda x = -\lambda(A-\lambda)^{-1}Ax$, we conclude that $A_\lambda x$ converges to Ax for all $x \in \operatorname{dom} A$. ❏

Theorem 4.2.4 (Hille–Yosida)**.** *Let X be a Banach space and $A\colon \operatorname{dom} A \subseteq X \to X$, $\omega \in \mathbb{R}$ and $M \geq 1$. Then A is a generator of a strongly continuous semigroup $T\colon [0,+\infty) \to \mathcal{L}_{\mathrm{b}}(X)$ with $\|T(t)\| \leq M\mathrm{e}^{\omega t}$, if and only if*

$$(\omega, +\infty) \subseteq \rho(A) \quad \textit{and} \quad \|(A-\lambda)^{-n}\| \leq \frac{M}{(\lambda-\omega)^n}$$

for all $\lambda \in (\omega, +\infty)$ and all $n \in \mathbb{N}$.

Proof. The two conditions are necessary by Lemma 4.1.9. We will show that they are also sufficient in several steps. The main idea is to use the Yosida approximant A_λ of A and show that $t \mapsto \mathrm{e}^{tA_\lambda}$ converges to a semigroup T whose infinitesimal generator is A. Hence, we will assume $(\omega, +\infty) \subseteq \rho(A)$ and $\|(A-\lambda)^{-n}\| \leq \frac{M}{(\lambda-\omega)^n}$ in the following.

1. Step: *Show that the semigroup of the Yosida approximant is uniformly bounded in λ:* The Yosida approximant $A_\lambda = -\lambda^2(A-\lambda)^{-1} - \lambda$ of A is a bounded operator, hence, e^{tA_λ} is well-defined. Moreover,

$$\begin{aligned}
\|\mathrm{e}^{tA_\lambda}\| = \|\mathrm{e}^{-\lambda t}\mathrm{e}^{-t\lambda^2(A-\lambda)^{-1}}\| &\leq \mathrm{e}^{-\lambda t}\sum_{n=0}^{\infty}\frac{\|-t\lambda^2(A-\lambda)^{-1}\|^n}{n!} \\
&\leq \mathrm{e}^{-\lambda t}\sum_{n=0}^{\infty}\frac{t^n\lambda^{2n}}{n!}\frac{M}{(\lambda-\omega)^n} = M\mathrm{e}^{-\lambda t}\mathrm{e}^{\frac{t\lambda^2}{\lambda-\omega}} = M\mathrm{e}^{t\frac{\omega\lambda}{\lambda-\omega}} \\
&\leq M\mathrm{e}^{2t|\omega|},
\end{aligned}$$

if $\lambda > 2\omega$ as then $\frac{\omega\lambda}{\lambda-\omega} \leq 2|\omega|$.

2. Step: *Show that* $\mathrm{e}^{tA_\lambda}x$ *converges for* $\lambda \to \infty$*:* For fixed $x \in X$ and $\tau > 0$ we will show that $\mathrm{e}^{\cdot A_\lambda}x$ is a Cauchy sequence in $\mathrm{C}([0,\tau];X)$. First note

$$\mathrm{e}^{tA_\lambda}x - \mathrm{e}^{tA_\mu}x = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{stA_\lambda}\mathrm{e}^{(1-s)tA_\mu}x\,\mathrm{d}s \quad \text{and}$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{stA_\lambda}\mathrm{e}^{(1-s)tA_\mu}x = \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{st(A_\lambda-A_\mu)+tA_\mu}x = \mathrm{e}^{st(A_\lambda-A_\mu)+tA_\mu}t(A_\lambda - A_\mu)x.$$

Hence,

$$\begin{aligned}
\sup_{t\in[0,\tau]}\|\mathrm{e}^{tA_\lambda}x - \mathrm{e}^{tA_\mu}x\| &= \sup_{t\in[0,\tau]}\left\|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{stA_\lambda}\mathrm{e}^{(1-s)tA_\mu}x\,\mathrm{d}s\right\| \\
&\leq \sup_{t\in[0,\tau]}\int_0^1 \|\underbrace{\mathrm{e}^{st(A_\lambda-A_\mu)+tA_\mu}}_{=\mathrm{e}^{stA_\lambda}\mathrm{e}^{(1-s)tA_\mu}}t(A_\lambda - A_\mu)x\|\,\mathrm{d}s \\
&\leq \sup_{t\in[0,\tau]}\int_0^1 M\mathrm{e}^{2st|\omega|}M\mathrm{e}^{2(1-s)t|\omega|}t\|A_\lambda x - A_\mu x\|\,\mathrm{d}s \\
&\leq \tau M^2\mathrm{e}^{4\tau|\omega|}\|A_\lambda x - A_\mu x\|.
\end{aligned}$$

Since $(A_\lambda x)_{\lambda\geq 0}$ is a convergent net, we conclude that $(\mathrm{e}^{\cdot A_\lambda}x)_{\lambda\geq 0}$ is a Cauchy net in $\mathrm{C}([0,\tau];X)$ and therefore a convergent net. Moreover, the limit is also continuous on $[0,\tau]$. For arbitrary $t \geq 0$ we choose $\tau > t$, which justifies that $\mathrm{e}^{tA_\lambda}x$ converges as $\lambda \to \infty$. We define this limit as

$$T(t)x \coloneqq \lim_{\lambda\to\infty}\mathrm{e}^{tA_\lambda}x.$$

Hence, $T(\cdot)x$ is continuous on $[0,\tau]$ for every τ and therefore continuous on $[0,+\infty)$.

3. Step: *Show that* T *is a semigroup:* By the linearity of the limit we immediately conclude

$$T(t)(x+\alpha y) = \lim_{\lambda\to\infty}\mathrm{e}^{tA_\lambda}(x+\alpha y) = \lim_{\lambda\to\infty}\mathrm{e}^{tA_\lambda}x + \alpha\mathrm{e}^{tA_\lambda}y = T(t)x + \alpha T(t)y.$$

Moreover, from the estimate in 1. Step we obtain

$$\|T(t)x\| \leq \limsup_{\lambda\to\infty}\|\mathrm{e}^{tA_\lambda}x\| \leq \limsup_{\lambda\to\infty} M\mathrm{e}^{t\frac{\omega\lambda}{\lambda-\omega}}\|x\| = M\mathrm{e}^{t\omega}\|x\|.$$

Clearly we have $T(0)x = \lim_{\lambda\to\infty}\mathrm{e}^{0A_\lambda}x = x$. For $t,s \geq 0$ we have

$$\begin{aligned}
&\|T(t)T(s)x - \mathrm{e}^{tA_\lambda}\mathrm{e}^{sA_\lambda}x\| \\
&\quad\leq \|T(t)T(s)x - \mathrm{e}^{tA_\lambda}T(s)x\| + \|\mathrm{e}^{tA_\lambda}T(s)x - \mathrm{e}^{tA_\lambda}\mathrm{e}^{sA_\lambda}x\| \\
&\quad\leq \|T(t)T(s)x - \mathrm{e}^{tA_\lambda}T(s)x\| + M\mathrm{e}^{2t|\omega|}\|T(s)x - \mathrm{e}^{sA_\lambda}x\| \to 0,
\end{aligned}$$

as $\mathrm{e}^{tA_\lambda}x$ converges to $T(t)x$ for every $t \geq 0$ and every $x \in X$. Hence,

$$T(t+s)x = \lim_{\lambda\to\infty} \mathrm{e}^{(t+s)A_\lambda}x = \lim_{\lambda\to\infty} \mathrm{e}^{tA_\lambda}\mathrm{e}^{sA_\lambda}x = T(t)T(s)x,$$

which shows that T is a strongly continuous semigroup such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$.

4. Step: *Show that A is the infinitesimal generator of T:* For $x \in \operatorname{dom} A$ we have

$$\begin{aligned} T(h)x - x &= \lim_{\lambda\to\infty} \mathrm{e}^{hA_\lambda}x - x = \lim_{\lambda\to\infty} \int_0^h \mathrm{e}^{tA_\lambda}A_\lambda x \,\mathrm{d}t \\ &= \lim_{\lambda\to\infty} \int_0^h \mathrm{e}^{tA_\lambda}Ax\,\mathrm{d}t + \int_0^h \underbrace{\mathrm{e}^{tA_\lambda}(A_\lambda - A)x}_{\to 0}\,\mathrm{d}t \\ &= \int_0^h T(t)Ax. \end{aligned}$$

Multiplying this by $\frac{1}{h}$ and taking the limit $h \to 0$ gives

$$\lim_{\lambda\to\infty} \frac{T(h)x - x}{h} = \lim_{\lambda\to\infty} \frac{1}{h}\int_0^h T(t)Ax = Ax. \qquad ❑$$

Theorem 4.2.4 gives a characterization of infinitesimal generators. However, these conditions are often not so easy to check. Fortunately, in many relevant problems things can be simplified.

4.3 Contraction semigroups

In a lot of physical motivated examples the characterization for an infinitesimal generator can be simplified. In particular if we have models that only allow solution that respect certain conservation laws, then the conditions in Theorem 4.2.4 can be reduced. In most of those physical motivated Cauchy problems the norm of the Banach space relates to (or is) the energy of a state (the elements of the Banach spaces are referred to as states). Hence, the semigroup has to satisfy

$$\|T(t)x\| \leq \|x\|,$$

because otherwise energy is generated, which contradicts the assumptions of the model. This leads to the following concept.

Definition 4.3.1. We say a strongly continuous semigroup T is a *contraction semigroup*, if $\|T(t)\| \leq 1$ for all $t \in [0, +\infty)$. ✧

The infinitesimal generator of a contraction semigroup can be much easier characterized. For this we introduce the notion of a dissipative linear relation. In particular we will be interested in maximal dissipative operators.

Most of the following is also possible in general Banach spaces. However, we restrict ourselves to Hilbert spaces, as there the situation is even more accessible.

Definition 4.3.2. Let H be a Hilbert space and A a linear relation on H. Then we say A is *dissipative*, if

$$\operatorname{Re}\langle y, x\rangle_H \le 0 \quad \forall \begin{bmatrix} x \\ y \end{bmatrix} \in A.$$

We say A is *maximally dissipative*, if there is no proper dissipative extension of A, i.e., there is no dissipative B such that $B \supsetneq A$. ✧

If A is an operator, then the condition for dissipativity can be formulated as $\operatorname{Re}\langle Ax, x\rangle_H \le 0$.

Remark 4.3.3. In literature there is also the term *m-dissipative*, where the "m" also originates from maximally, but is defined as: A is m-dissipative, if A is dissipative and $A - \mathrm{I}$ is surjective. For Hilbert spaces that is equivalent to maximally dissipative, but in general not. ✧

Let T be a contraction semigroup on the Hilbert space H and A its infinitesimal generator. Then for $x_0 \in \operatorname{dom} A$ the trajectory $x(t) = T(t)x_0$ solves the corresponding Cauchy problem $\dot{x} = Ax$ and $x(0) = x_0$. For $t < s$ we have

$$\|x(s)\| = \|T(s)x_0\| = \|T(s-t)T(t)x_0\| \le \|T(t)x_0\| = \|x(t)\|.$$

Hence, a solution does not grow. Moreover, we have

$$\begin{aligned} 0 \ge \frac{\mathrm{d}}{\mathrm{d}t}\|x(t)\|^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\langle x(t), x(t)\rangle \\ &= \langle \dot{x}(t), x(t)\rangle + \langle x(t), \dot{x}(t)\rangle = 2\operatorname{Re}\langle Ax(t), x(t)\rangle, \end{aligned} \tag{4.3}$$

which indicates that dissipativity of A is closely related to the contractivity of T.

Corollary 4.3.4. *Let T be a contraction semigroup on a Hilbert space H and A its infinitesimal generator. Then A is a dissipative operator.*

Proof. Let $x_0 \in \operatorname{dom} A$ and $x(t) = T(t)x_0$. Then (4.3) for $t = 0$ gives

$$\operatorname{Re}\langle Ax_0, x_0\rangle \le 0.$$ ❑

Lemma 4.3.5. *Let A be a closed dissipative linear relation on a Hilbert space H and $\lambda \in \mathbb{C}$ such that $\operatorname{Re}\lambda > 0$. Then $A - \lambda$ is injective, $(A-\lambda)^{-1}$ is an operator*[1] *such that $\|(A-\lambda)^{-1}x\| \le \frac{1}{\operatorname{Re}\lambda}\|x\|$ and $\operatorname{ran}(A-\lambda)$ is closed in H.*

[1] It is well-defined, but not necessarily everywhere defined or densely defined.

Proof. Let $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A$ and $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \in (A-\lambda)$ such that $z = y - \lambda x$. Note that $(A - \mathrm{i}\operatorname{Im}\lambda)$ is also a dissipative linear relation and therefore $\operatorname{Re}\langle y - \mathrm{i}\operatorname{Im}\lambda x, x\rangle \leq 0$. Then we have the following inequality

$$\begin{aligned} \|z\|_H^2 &= \|y-\lambda x\|_H^2 = \|y - \mathrm{i}\operatorname{Im}\lambda x - \operatorname{Re}\lambda x\|_H^2 \\ &= \|y - \mathrm{i}\operatorname{Im}\lambda x\|_H^2 \underbrace{- 2(\operatorname{Re}\lambda)\overbrace{\operatorname{Re}\langle y - \mathrm{i}\operatorname{Im}\lambda x, x\rangle_H}^{\leq 0}}_{\geq 0} + |\operatorname{Re}\lambda|^2\|x\|_H^2 \\ &\geq |\operatorname{Re}\lambda|^2\|x\|_H^2. \end{aligned}$$

Hence, $A-\lambda$ is injective and therefore $(A-\lambda)^{-1}$ is an operator such that $\|(A-\lambda)^{-1}z\| \leq \frac{1}{\operatorname{Re}\lambda}\|z\|$.

Let $(\left[\begin{smallmatrix} x_n \\ z_n \end{smallmatrix}\right])_{n\in\mathbb{N}}$ be a sequence in $(A-\lambda)$ such that $(z_n)_{n\in\mathbb{N}}$ converges to $z \in H$. Then the previous inequality implies that also $(x_n)_{n\in\mathbb{N}}$ converges to a limit $x \in H$. Since A is closed (and therefore also $(A-\lambda)$), we conclude that $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \in (A-\lambda)$ and consequently that $\operatorname{ran}(A-\lambda)$ is closed. ❑

Lemma 4.3.6. *If A is a maximally dissipative linear relation on a Hilbert space H, then $(A-1)$ is bijective, i.e.,* $\ker(A-1) = \{0\}$ *and* $\operatorname{ran}(A-1) = H$.

Proof. Note that $(A-1)$ is injective and $\operatorname{ran}(A-1)$ is closed by Lemma 4.3.5. Assume that $(A-1)$ is not surjective. Then there is a non zero $z \in H$ that is orthogonal on $\operatorname{ran}(A-1)$, i.e.,

$$0 = \langle y - x, z\rangle = \langle y, z\rangle - \langle x, z\rangle \quad \text{for all} \quad \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A. \tag{4.4}$$

If $z \in \operatorname{dom} A$, then, by the previous equation and the dissipativity of A, we have for all $\left[\begin{smallmatrix} z \\ w \end{smallmatrix}\right] \in A$

$$\|z\|^2 = \operatorname{Re}\|z\|^2 = \operatorname{Re}\langle z, z\rangle \overset{(4.4)}{=} \operatorname{Re}\langle w, z\rangle \leq 0.$$

Therefore, $z = 0$, which contradicts our assumption $z \neq 0$. On the other hand, if $z \notin \operatorname{dom} A$, then we extend A to $B := \operatorname{span}(A \cup \{\left[\begin{smallmatrix} z \\ -z \end{smallmatrix}\right]\})$, which is again dissipative. This can be seen by using (4.4)

$$\operatorname{Re}\langle \alpha z + x, -\alpha z + y\rangle = -|\alpha|^2\|z\|^2 + \underbrace{\operatorname{Re}(\langle \alpha z, y\rangle - \langle x, \alpha z\rangle)}_{=0} + \operatorname{Re}\langle x, y\rangle \leq 0$$

for $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A$. However, this contradicts the maximal dissipativity of A. Hence, such a z cannot exist. ❑

Remark 4.3.7. Note that a maximally dissipative linear relation A does not have to be an operator even though $A-1$ is bijective. The most degenerated counter example is

$$A := \left\{ \begin{bmatrix} 0 \\ x \end{bmatrix} \,\middle|\, x \in H \right\}.$$

Note that A is bijective and $A - 1 = A$. This linear relation is maximally dissipative, because every proper extension B would violate $B - 1$ being injective, which is necessary for a dissipative linear relation. However, $\operatorname{mul} A = H$. ✧

Corollary 4.3.8. *If A is a maximally dissipative linear relation on a Hilbert space H, then $(A - \lambda)$ for $\lambda \in (0, +\infty)$ is bijective, i.e.,* $\ker(A - \lambda) = \{0\}$ *and* $\operatorname{ran}(A - \lambda) = H$.

Proof. $A - \lambda$ is injective by Lemma 4.3.5. If A is maximally dissipative, then also $\frac{1}{\lambda}A$ is maximally dissipative for $\lambda > 0$. Hence, by Lemma 4.3.6 we obtain

$$H = \operatorname{ran}(\tfrac{1}{\lambda}A - 1) = \tfrac{1}{\lambda}\operatorname{ran}(A - \lambda) = \operatorname{ran}(A - \lambda),$$

which shows the assertion. ❑

Proposition 4.3.9. *Let A be a closed linear relation on a Hilbert space H. Then A is maximally dissipative, if and only if A and A^* are dissipative.*

Proof. Note that if a linear relation B is dissipative, then also B^{-1} is dissipative. Moreover, recall ${B^*}^{-1} = {B^{-1}}^*$ for any linear relation B.

$\boxed{\Rightarrow}$: Let A be maximally dissipative. Then for arbitrary $\epsilon > 0$ the linear relation $A - \epsilon$ is also dissipative and therefore also $(A - \epsilon)^{-1}$. Moreover, $(A - \epsilon)^{-1}$ is a bounded linear operator. Hence, for every $x \in H$

$$0 \geq \operatorname{Re}\langle (A - \epsilon)^{-1}x, x\rangle = \operatorname{Re}\langle x, {(A - \epsilon)^{-1}}^* x\rangle = \operatorname{Re}\langle x, (A^* - \epsilon)^{-1}x\rangle,$$

which leads to the dissipativity of $(A^* - \epsilon)^{-1}$ and in turn to the dissipativity of $(A^* - \epsilon)$. Therefore, for every $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A^*$ we have

$$0 \geq \operatorname{Re}\langle y - \epsilon x, x\rangle \overset{\epsilon \to 0}{\to} \operatorname{Re}\langle y, x\rangle,$$

which shows the dissipativity of A^*.

$\boxed{\Leftarrow}$: If A and A^* are both dissipative, then $(A - 1)$ and $(A^* - 1)$ are both injective and have closed range by Lemma 4.3.5. Hence,

$$\operatorname{ran}(A - 1) = \overline{\operatorname{ran}(A - 1)} = \ker(A^* - 1)^{\perp} = \{0\}^{\perp} = H,$$

which implies $A - 1$ is also surjective and therefore bijective. Every proper dissipative extension B of A would violate the injectivity of $B - 1$ as $A - 1$ is already surjective. Hence, A is maximally dissipative. ❑

Theorem 4.3.10 (Lumer–Phillips)**.** *Let A be a linear operator on a Hilbert space H. Then A is the infinitesimal generator of a contraction semigroup T, if and only if A is maximally dissipative.*

Proof. We show the two directions separately.

$\boxed{\Rightarrow}$: The generator of contraction semigroup is dissipative by Corollary 4.3.4. Moreover, M, ω in Theorem 4.2.4 can be chosen as $M = 1$ and $\omega = 0$. Hence, $(0, \infty) \in \rho(A)$, which implies that $(A-1)$ is surjective. Since every dissipative extension B of A satisfies $(B - 1)$ is injective by Lemma 4.3.5, we conclude that there cannot be a proper extension A, which implies the maximality of A.

$\boxed{\Leftarrow}$: If A is maximally dissipative, then by Lemma 4.3.5 and Lemma 4.3.6 we have $(0, +\infty) \subseteq \rho(A)$ with

$$\|(A - \lambda)^{-1}\| \leq \frac{1}{\lambda},$$

which immediately leads to $\|(A - \lambda)^{-n}\| \leq \frac{1}{\lambda^n}$. Hence, by Theorem 4.2.4 we conclude A is the infinitesimal generator of a strongly continuous semigroup T with $\|T(t)\| \leq e^{0t} = 1$. ❑

Corollary 4.3.11. *Let A be a closed and densely defined linear operator on a Hilbert space H. Then A is the infinitesimal generator of a contraction semigroup T, if and only if A and A^* are dissipative.*

Proof. This is Proposition 4.3.9 with Theorem 4.3.10. ❑

4.4 Mild solutions

Recall the abstract Cauchy problem for an infinitesimal generator A on a Banach space X: For given $x_0 \in X$ find a function $x\colon [0, +\infty) \to \operatorname{dom} A$ that satisfies

$$\begin{aligned} \dot{x}(t) &= Ax(t), \quad t \geq 0, \\ x(0) &= x_0. \end{aligned} \tag{4.5}$$

We have already discussed that for $x_0 \in \operatorname{dom} A$ this problem is uniquely solved by $x(t) \coloneqq T(t)x_0$, where T is the semigroup generated by A.

Interestingly, we can even define solutions of the abstract Cauchy problem (4.5) for initial values x_0 that are not in the domain of the infinitesimal generator A. The idea is to look at the integrated equation

$$\int_0^\tau \dot{x}(t)\,\mathrm{d}t = \int_0^\tau Ax(t)\,\mathrm{d}t$$

For the solution $x(t) = T(t)x_0$ where $x \in \operatorname{dom} A$ we can interchange the integral and the operator A by Lemma 4.1.8 item (ii). Hence, we obtain

$$x(\tau) - x_0 = A \int_0^\tau x(t)\,\mathrm{d}t\,.$$

This formulation does not require that $x(\cdot)$ maps into $\operatorname{dom} A$, but only $\int_0^\tau x(t)\,\mathrm{d}t \in \operatorname{dom} A$, which is automatically satified for $T(\cdot)x_0$ (even if $x_0 \notin \operatorname{dom} A$). This leads to the following definition.

Definition 4.4.1. We say a function $x\colon [0,+\infty) \to X$ is a *mild solution* of the abstract Cauchy problem (4.5), if

$$x(t) - x_0 = A \int_0^t x(s)\,\mathrm{d}s,$$

where we implicitly demand that $\int_0^t x(s)\,\mathrm{d}s \in \operatorname{dom} A$. ✧

Proposition 4.4.2. *The mild solution of (4.5) is uniquely given by $x(\cdot) = T(\cdot)x_0$, where T is the semigroup generated by A.*

Proof. By Lemma 4.1.8 item (ii) we know that $T(\cdot)x_0$ is a mild solution. By linearity it is enough to show that the solution for the initial value $x_0 = 0$ is unique.

Let x be a mild solution for $x_0 = 0$. Then

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(T(t-s)\int_0^s x(r)\,\mathrm{d}r\right) = T(t-s)x(s) + T(t-s)A\int_0^s x(r)\,\mathrm{d}r = 0.$$

Hence, $s \mapsto T(t-s)\int_0^s x(r)\,\mathrm{d}r$ is constant, which implies

$$0 = T(t-0)\int_0^0 x(r)\,\mathrm{d}r = T(t-t)\int_0^t x(r)\,\mathrm{d}r = \int_0^t x(r)\,\mathrm{d}r\,.$$

Since x is a mild solution for $x_0 = 0$ we conclude

$$x(t) = x(t) - x_0 = A\int_0^t x(r)\,\mathrm{d}r = 0$$

and therefore the uniqueness of the solution. ❑

Appendix A

Locally convex topology

In this chapter we want to recall the notion of a locally convex topology on a vector space. In particular we want to present the most important results for our purpose.

A.1 Basics

Definition A.1.1. Let X be a vector space. Then we say $p\colon X \to \mathbb{R}$ is a seminorm, if p satisfies

(i) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$ *(triangle inequality)*,

(ii) $p(\alpha x) = |\alpha| p(x)$ for all $x \in X$, $\alpha \in \mathbb{C}$ *(absolute homogeneity)*. ✧

Hence, the difference between a norm and a seminorm is that a seminorm is not definite, i.e., $p(x) = 0$ does not necessarily imply $x = 0$.

We choose to introduce *locally convex topological vector spaces* and *locally convex topologies* by a result that characterizes them. As *locally convex* in the name already suggests, the usual definition involves convexity of some parts of the topology. However, we take a shortcut via seminorms.

Definition A.1.2 (Locally convex topological vector space)**.** Let X be a vector space and $\mathcal{P}$ a set of seminorms on X, which is separating, i.e.,

$$\bigcap_{p \in \mathcal{P}} p^{-1}\{0\} = \{x \in X \mid p(x) = 0 \text{ for all } p \in \mathcal{P}\} = \{0\}.$$

Then we say the topology $\mathcal{T}$ generated by

$$V(x, p, \epsilon) := \{y \in X \mid p(x-y) \leq \epsilon\}$$

for $x \in X$, $p \in \mathcal{P}$ and $\epsilon > 0$ is a *locally convex topology* on X and X is then called a *locally convex topological vector space* or just a *locally convex space.* ✧

Remark A.1.3. Let X be a locally convex space and $\mathcal{P}$ the set of its seminorms.

(i) Every $V(x,p,\epsilon)$ is convex for $x \in X$, $p \in \mathcal{P}$ and $\epsilon > 0$.

(ii) By construction W is a neighborhood of $x \in X$, if and only if there is a finite subset $M \subseteq \mathcal{P}$ such that

$$V(x,M,\epsilon) := \bigcap_{p \in M} V(x,p,\epsilon) \subseteq W.$$

(iii) Moreover, E is Hausdorff and

$$+\colon X \times X \to X, (x,y) \mapsto x+y$$
$$\cdot\colon \mathbb{C} \times X \to X, (\alpha,x) \mapsto \alpha x$$

are continuous. ✧

Proposition A.1.4. *Let X be a locally convex space and $\mathcal{P}$ its set of seminorms. Then we have the following equivalence for a net $(x_i)_{i\in I}$ in X:*

(i) $x_i \to x$.

(ii) $p(x - x_i) \to 0$ *for every* $p \in \mathcal{P}$.

Proof. $\boxed{\Rightarrow}$: Let $x_i \to x$. Then by definition of convergence this means that for every $V(x,p,\epsilon)$ there exists an i_0 such that $x_i \in V(x,p,\epsilon)$, if $i \geq i_0$. This means $p(x - x_i) \leq \epsilon$, if $i \geq i_0$.

$\boxed{\Leftarrow}$: On the other hand let $p(x - x_i) \to 0$. Hence, for given $\epsilon > 0$ there exists an i_0 such that $p(x - x_i) < \epsilon$. Consequently, $x_i \in V(x,p,\epsilon)$, if $i \geq i_0$. Since sets of the form $V(x,p,\epsilon)$ establish a subbasis for the topology we conclude $x_i \to x$. ❑

Appendix B

Lipschitz domains and boundaries

We want to present the basics for strongly Lipschitz domains and strongly Lipschitz boundaries. Since we don't regard weakly Lipschitz domains and boundaries, we omit the term "strongly" and just call them Lipschitz domains and Lipschitz boundaries.

A rich source for Lipschitz domains and boundaries is, e.g., [6].

B.1 Definition

Definition B.1.1. Let Ω be an open subset of $\mathbb{R}^d$. We say Ω is a *strongly Lipschitz domain*, if for every $p \in \partial\Omega$ there exist $\epsilon, h > 0$, a hyperplane $W = \operatorname{span}\{w_1, \dots, w_{d-1}\}$, where $\{w_1, \dots, w_{d-1}\}$ is an orthonormal basis of W, and a Lipschitz continuous function $a\colon (p + W) \cap \mathrm{B}_\epsilon(p) \to (-\frac{h}{2}, \frac{h}{2})$ such that

$$\begin{aligned} \partial\Omega \cap C_{\epsilon,h}(p) &= \{x + a(x)v \mid x \in (p + W) \cap \mathrm{B}_\epsilon(p)\}, \\ \Omega \cap C_{\epsilon,h}(p) &= \{x + sv \mid x \in (p + W) \cap \mathrm{B}_\epsilon(p), -h < s < a(x)\}, \end{aligned}$$

where v is the normal vector of W and $C_{\epsilon,h}(p)$ is the cylinder $\{x + \delta v \mid x \in (p + W) \cap \mathrm{B}_\epsilon(p), \delta \in (-h, h)\}$.

The boundary $\partial\Omega$ is then called *strongly Lipschitz boundary.* ✧

Note that the condition $|a| < \frac{h}{2}$ is not really necessary, however it reduces technical constructions. If it was not already satisfied, we can force it by shrinking ϵ.

Locally the boundary is given by the graph of a Lipschitz function, see Figure B.1. Therefore, we can define Lipschitz charts on $\partial\Omega$ in the following way. Let p, $C_{\epsilon,h}(p)$, W, v, a be as in Definition B.1.1. We will also denote the matrix that contains the orthonormal basis of W as columns by W, i.e.,

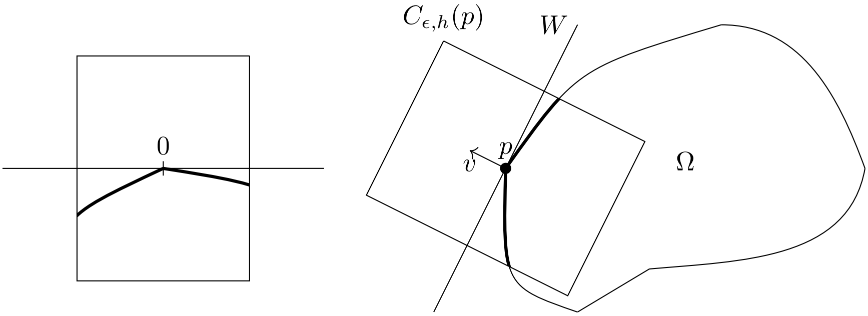

Figure B.1: Lipschitz boundary

$W \in \mathbb{R}^{d\times(d-1)}$. Hence, the mapping $\zeta \mapsto W^{\mathsf{T}}\zeta$ gives the coordinates (w.r.t. the basis $w_1, \dots, w_{d-1}$) of the orthogonal projection of ζ on the hyperplane W. We introduce a *strongly Lipschitz chart* locally at p by

$$k\colon \begin{cases} \partial\Omega \cap C_{\epsilon,h}(p) & \to & \mathrm{B}_\epsilon(0) \subseteq \mathbb{R}^{d-1}, \\ \zeta & \mapsto & W^{\mathsf{T}}(\zeta - p). \end{cases}$$

We say that $\Gamma := \partial\Omega \cap C_{\epsilon,h}(p)$ is the *chart domain* of k. Also every restriction of a chart to an open non-empty $\hat{\Gamma} \subseteq \Gamma$ (w.r.t. the trace topology) is again a chart with chart domain $\hat{\Gamma}$. The corresponding inverse chart is given by

$$k^{-1}\colon \begin{cases} \mathrm{B}_\epsilon(0) \subseteq \mathbb{R}^{d-1} & \to & \partial\Omega \cap C_{\epsilon,h}(p), \\ x & \mapsto & p + \sum_{i=1}^{d-1} x_i w_i + a(p + \sum_{i=1}^{d-1} x_i w_i)v. \end{cases}$$

In the case where k is a "restricted" chart, we have $k^{-1}\colon U \to \hat{\Gamma}$, where U is an open non-empty subset of $\mathrm{B}_\epsilon(0)$ in $\mathbb{R}^{d-1}$. For notational simplicity we just write $a(x)$ instead of $a\big(p + \sum_{i=1}^{d-1} x_i w_i\big)$. By this convention we have $a\colon U \subseteq \mathbb{R}^{d-1} \to \mathbb{R}$ and

$$k^{-1}(x) = p + Wx + a(x)v = p + \begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} x \\ a(x) \end{bmatrix}.$$

Note that in fact W, v and p establish an alternative coordinate system with origin p. Hence, by translation and rotation we can, most of the time, assume (w.l.o.g.) that $W = (e_1, \dots, e_{d-1})$, $v = e_d$ and $p = 0$. This will also better transport the essence of our ideas. In this coordinate system we have

$$k\left(\begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_d \end{bmatrix}\right) = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_{d-1} \end{bmatrix} \quad \text{and} \quad k^{-1}(x) = \begin{bmatrix} x \\ a(x) \end{bmatrix}.$$

However, sometimes it is not entirely obvious that we can reduce the general setting to this situation or the justification that such a reduction is valid is as difficult as working in the general setting in the first place.

Note that k^{-1} is Lipschitz continuous—since a is Lipschitz continuous by assumption—and therefore k^{-1} is a.e. differentiable by Rademacher's theorem, see, e.g., [1, Thm. 2.14]. In particular, $k^{-1} \in \mathrm{W}^{1,\infty}(U)$ and therefore $\mathrm{d}k^{-1}$ is a bounded multiplication operator on $\mathrm{L}^2(U)$. Hence, if we don't write arguments (of functions), then we regard the functions as L^p objects and omit the comment "a.e.".

Let $k\colon \Gamma \to U$ be a strongly Lipschitz chart. The surface measure on $\partial\Omega$ is locally given by

$$\mu(\Upsilon) = \int_{k(\Upsilon)} \sqrt{\det(\mathrm{d}k^{-1})^{\mathsf{T}}\mathrm{d}k^{-1}}\,\mathrm{d}\lambda_{d-1} \quad \text{for} \quad \Upsilon \subseteq \Gamma,$$

where λ_{d-1} is the Lebesgue measure in $\mathbb{R}^{d-1}$. The surface measure is then defined by a partition of $\partial\Omega$. By Lindelöf's lemma there exists a countable partition, see, e.g., [10, Ch. 3 § 4]. If $\partial\Omega$ is bounded then there exists even a finite partition. The surface measure is independent of the partition and the charts, see Proposition B.3.3. Hence, we can switch between the inner products of $\mathrm{L}^2(\Gamma)$ and $\mathrm{L}^2(U)$ by

$$\langle f, g\rangle_{\mathrm{L}^2(\Gamma)} = \left\langle f \circ k^{-1}, \sqrt{\det(\mathrm{d}k^{-1})^{\mathsf{T}}\mathrm{d}k^{-1}}\, g \circ k^{-1}\right\rangle_{\mathrm{L}^2(U)}.$$

We can even simplify the the determinant term by the following lemma.

Lemma B.1.2. *Let k be a Lipschitz chart. Then we have*

$$(\mathrm{d}k^{-1})^{\mathsf{T}}\mathrm{d}k^{-1} = \mathrm{I} + \nabla a(\nabla a)^{\mathsf{T}} \quad \textit{and} \quad \det(\mathrm{d}k^{-1})^{\mathsf{T}}\mathrm{d}k^{-1} = 1 + \|\nabla a\|^2.$$

Proof. Recall that $k(x) = p + Wx + a(x)v$. Hence,

$$\mathrm{d}k^{-1} = W + \begin{bmatrix}\partial_1 av & \partial_2 av & \cdots & \partial_{d-1}av\end{bmatrix}.$$

Since $\begin{bmatrix}W & v\end{bmatrix}$ is a orthogonal matrix we conclude

$$\begin{aligned}(\mathrm{d}k^{-1})^{\mathsf{T}}\mathrm{d}k^{-1} &= \left(W^{\mathsf{T}} + \begin{bmatrix}\partial_1 av^{\mathsf{T}} \\ \vdots \\ \partial_{d-1}av^{\mathsf{T}}\end{bmatrix}\right)\left(W + \begin{bmatrix}\partial_1 av & \cdots & \partial_{d-1}av\end{bmatrix}\right) \\ &= \mathrm{I} + (\nabla a)(\nabla a)^{\mathsf{T}}.\end{aligned}$$

Finally, Lemma B.4.1 gives $\det(\mathrm{d}k^{-1})^{\mathsf{T}}\mathrm{d}k^{-1} = 1 + \|\nabla a\|^2$. ❑

B.2 Outer normal vector

In this section we will show that the outer normal vector can be parameterized by

$$\nu \circ k^{-1} = \frac{1}{\sqrt{1 + \|\nabla a\|^2}} \begin{bmatrix}W & v\end{bmatrix} \begin{bmatrix}-\nabla a \\ 1\end{bmatrix}.$$

First of all note that a and k^{-1} are a.e. differentiable as Lipschitz continuous mappings. Hence, we can characterize the tangential space in almost every $p \in \partial\Omega$ by the columns of $\mathrm{d}k^{-1}$. In particular for a.e. $p \in \partial\Omega$ the range of the matrix $\mathrm{d}k^{-1}(k(p))$ is the tangential space in p on $\partial\Omega$. Lemma B.3.1 shows that this space is independent of the chart k.

Locally in $C_{\epsilon,h}(p)$ we can characterize $\partial\Omega$ implicitly by

$$\zeta \in \partial\Omega \cap C_{\epsilon,h}(p) \quad \Leftrightarrow \quad v \cdot (\zeta - p) - a(k(\zeta)) = 0.$$

If W and v forms the standard basis this means just

$$\zeta_d - a\left(\begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_{d-1} \end{bmatrix}\right) = 0.$$

Lemma B.2.1. *The function*

$$\begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} -\nabla a \\ 1 \end{bmatrix} \circ k.$$

is point wise a.e. normal on the tangential space of $\partial\Omega$.

Proof. Let $F\colon C_{\epsilon,h}(p) \to \mathbb{R}$, $\zeta \mapsto v \cdot (\zeta - p) - a(k(\zeta))$. Then clearly, $F \circ k^{-1} = 0$. Hence, by the chain rule we obtain

$$0 = \mathrm{d}(F \circ k^{-1}) = (\mathrm{d}F) \circ k^{-1} \mathrm{d}k^{-1},$$

which gives that $(\mathrm{d}F)^{\mathsf{T}}\big|_{\partial\Omega\cap C_{\epsilon,h}(p)}$ is orthogonal on every column of $(\mathrm{d}k^{-1}) \circ k$. Note that $\mathrm{d}k = W^{\mathsf{T}}$. Thus, again by the chain rule we obtain

$$\nabla F = v - W(\nabla a) \circ k = \begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} -\nabla a \\ 1 \end{bmatrix} \circ k. \qquad \square$$

Clearly, if we want a unit normal vector, we just have to divide

$$\begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} -\nabla a \\ 1 \end{bmatrix} \circ k$$

by its pointwise norm.

Theorem B.2.2. *The outward unit normal vector on $\partial\Omega$ is locally on $\partial\Omega \cap C_{\epsilon,h}(p)$ given by*

$$\nu(\zeta) = \frac{1}{\sqrt{1 + \|(\nabla a)(k(\zeta))\|^2}} \begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} -(\nabla a)(k(\zeta)) \\ 1 \end{bmatrix}$$

for almost every $\zeta \in \partial\Omega \cap C_{\epsilon,h}(p)$.

Figure B.2 illustrates the proof.

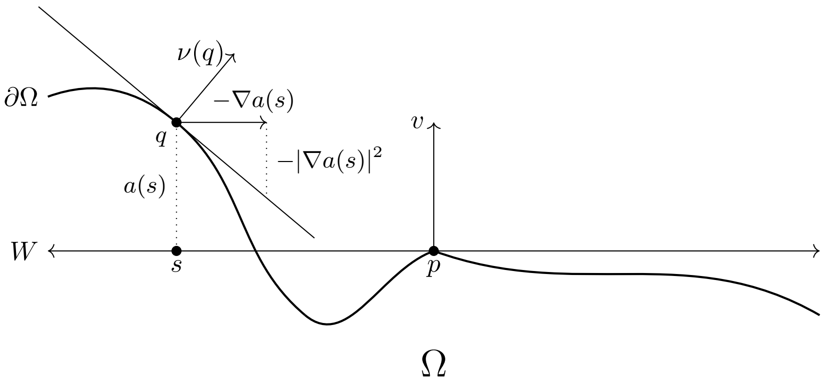

Figure B.2: Outer normal vector

Proof. Let $q \in \partial\Omega \cap C_{\epsilon,h}$ such that the tangential space exists and $s = k(q)$. For an $x \in W \cap \mathrm{B}_\epsilon(p)$ we can express the corresponding point on surface of Ω by

$$p + \begin{bmatrix} W & v \end{bmatrix} \begin{bmatrix} x \\ a(x) \end{bmatrix}.$$

For notational simplicity we assume $p = 0$, and that W and v form the standard basis, which allows us to parameterize $\partial\Omega$ with just $\left[\begin{smallmatrix} x \\ a(x) \end{smallmatrix}\right]$. Since a is differentiable in s, we have

$$\begin{bmatrix} s - \mu\nabla a(s) \\ a(s - \mu\nabla a(s)) \end{bmatrix} = \begin{bmatrix} s \\ a(s) \end{bmatrix} - \mu \begin{bmatrix} \nabla a(s) \\ \|\nabla a(s)\|^2 \end{bmatrix} + \begin{bmatrix} 0 \\ o(\mu) \end{bmatrix}.$$

Hence, for $\mu > 0$ sufficiently small we have $a(s - \mu\nabla g(s)) \leq a(s)$, which implies $\left[\begin{smallmatrix} s-\mu\nabla a(s) \\ a(s) \end{smallmatrix}\right] \notin \Omega$. Consequently,

$$q + \mu\nu(q) = \begin{bmatrix} s \\ a(s) \end{bmatrix} + \mu \begin{bmatrix} -\nabla a(s) \\ 1 \end{bmatrix} \notin \Omega.$$

Therefore, $\nu(q)$ points outward Ω. ❑

B.3 Independence of the charts

Note that for two strongly Lipschitz charts $k_1 \colon \Gamma_1 \to U_1$, $k_2 \colon \Gamma_2 \to U_2$ with overlapping chart domains (i.e., $\Gamma_1 \cap \Gamma_2 \neq \emptyset$) we have that the columns of $\mathrm{d}k_1^{-1}(k_1(\zeta))$ and the columns of $\mathrm{d}k_2^{-1}(k_2(\zeta))$ span the same linear subspace of $\mathbb{R}^d$ for a.e. $\zeta \in \Gamma_1 \cap \Gamma_2$, namely the *tangential space* of $\partial\Omega$ at ζ. The next lemma will specify this.

Lemma B.3.1. *Let $k_1 \colon \Gamma_1 \to U_1$ and $k_2 \colon \Gamma_2 \to U_2$ be strongly Lipschitz charts. Then*

$$\operatorname{ran}\left[\mathrm{d}k_1^{-1}\big(k_1(\zeta)\big)\right] = \operatorname{ran}\left[\mathrm{d}k_2^{-1}\big(k_2(\zeta)\big)\right] \quad \textit{for a.e.} \quad \zeta \in \Gamma_1 \cap \Gamma_2.$$

Moreover,

$$(\mathrm{d}k_1^{-1})^\dagger \circ (k_1 \circ k_2^{-1})\,\mathrm{d}k_2^{-1} = \mathrm{d}(k_1 \circ k_2^{-1}). \tag{B.1}$$

Proof. The first assertion follows from

$$\mathrm{d}k_2^{-1} = \mathrm{d}(k_1^{-1} \circ k_1 \circ k_2^{-1}) = (\mathrm{d}k_1^{-1}) \circ (k_1 \circ k_2^{-1})\,\mathrm{d}(k_1 \circ k_2^{-1}) \tag{B.2}$$

and the fact that $\mathrm{d}(k_1 \circ k_2^{-1})(\zeta)$ is a regular matrix for a.e. $\zeta \in \Gamma_1 \cap \Gamma_2$. Multiplying both side of (B.2) from left with $(\mathrm{d}k_1^{-1})^\dagger \circ (k_1 \circ k_2^{-1})$ implies (B.1). ❑

Lemma B.3.2. *Let $k_1 \colon \Gamma_1 \to U_1$, $k_2 \colon \Gamma_2 \to U_2$ strongly Lipschitz charts. Then for a.e. $\zeta \in \Gamma_1 \cap \Gamma_2$ the following holds*

$$(\mathrm{d}k_1^{-1})(\mathrm{d}k_1^{-1})^\dagger \circ k_1(\zeta) = (\mathrm{d}k_2^{-1})(\mathrm{d}k_2^{-1})^\dagger \circ k_2(\zeta).$$

Proof. Note that by Lemma B.3.1 $\operatorname{ran}[\mathrm{d}k_1^{-1}(k_1(\zeta))] = \operatorname{ran}[\mathrm{d}k_2^{-1}(k_2(\zeta))]$ for a.e. $\zeta \in \Gamma_1 \cap \Gamma_2$. By Lemma B.4.3 $(\mathrm{d}k_1^{-1})(\mathrm{d}k_1^{-1})^\dagger \circ k_1(\zeta)$ is the orthogonal projection on $\operatorname{ran}[\mathrm{d}k_1^{-1}(k_1(\zeta))]$ and $(\mathrm{d}k_2^{-1})(\mathrm{d}k_2^{-1})^\dagger \circ k_2(\zeta)$ is the orthogonal projection on $\operatorname{ran}[\mathrm{d}k_2^{-1}(k_2(\zeta))]$. Since these ranges coincide we conclude the assertion. ❑

Proposition B.3.3. *The surface measure on $\partial\Omega$ is independent of the partition and the charts.*

Proof. It is enough to show that two charts $k_1 \colon \Gamma_1 \to U_1$ and $k_2 \colon \Gamma_2 \to U_2$ with intersecting chart domains define the same surface measure on the intersection $\Gamma_1 \cap \Gamma_2$. The rest can be done by intersecting the two partitions.

We define the mapping

$$T \colon \left\{ \begin{array}{rcl} k_2(\Gamma_1 \cap \Gamma_2) \subseteq U_2 & \to & k_1(\Gamma_1 \cap \Gamma_2) \subseteq U_1, \\ x & \mapsto & (k_1 \circ k_2^{-1})(x), \end{array} \right.$$

which gives a bijective bi-Lipschitz continuous mapping. Note that by the chain rule we have

$$\mathrm{d}k_2^{-1} = \mathrm{d}(k_1^{-1} \circ k_1 \circ k_2^{-1}) = (\mathrm{d}k_1^{-1}) \circ (k_1 \circ k_2^{-1})\mathrm{d}(k_1 \circ k_2^{-1}) = (\mathrm{d}k_1^{-1}) \circ T\mathrm{d}T.$$

Moreover, by properties of the determinant we have

$$
\begin{aligned}
&|\det \mathrm{d}T| \sqrt{\det(\mathrm{d}k_1^{-1} \circ T)^\mathsf{T} (\mathrm{d}k_1^{-1} \circ T)} \\
&\qquad = \sqrt{\det(\mathrm{d}T)^\mathsf{T} \mathrm{d}T} \sqrt{\det(\mathrm{d}k_1^{-1} \circ T)^\mathsf{T} (\mathrm{d}k_1^{-1} \circ T)} \\
&\qquad = \sqrt{\det(\mathrm{d}T)^\mathsf{T} (\mathrm{d}k_1^{-1} \circ T)^\mathsf{T} (\mathrm{d}k_1^{-1} \circ T) \mathrm{d}T} \\
&\qquad = \sqrt{\det((\mathrm{d}k_1^{-1} \circ T)\mathrm{d}T)^\mathsf{T} ((\mathrm{d}k_1^{-1} \circ T)\mathrm{d}T)} \\
&\qquad = \sqrt{\det(\mathrm{d}k_2^{-1})^\mathsf{T} \mathrm{d}k_2^{-1}}.
\end{aligned}
$$

Now for $\Upsilon \subseteq \Gamma_1 \cap \Gamma_2$ we have by change of variables

$$
\begin{aligned}
&\int_{k_1(\Upsilon)} \sqrt{\det(\mathrm{d}k_1^{-1})^\mathsf{T} \mathrm{d}k_1^{-1}} \, \mathrm{d}\lambda_{d-1} \\
&\qquad = \int_{T^{-1}(k_1(\Upsilon))} \sqrt{\det(\mathrm{d}k_1^{-1})^\mathsf{T} \mathrm{d}k_1^{-1}} \circ T |\det \mathrm{d}T| \, \mathrm{d}\lambda_{d-1} \\
&\qquad = \int_{k_2(\Upsilon)} \sqrt{\det(\mathrm{d}k_2^{-1})^\mathsf{T} \mathrm{d}k_2^{-1}} \, \mathrm{d}\lambda_{d-1} \, .
\end{aligned}
$$

Hence, the surface measure $\mu(\Upsilon)$ is independent of the charts. ❑

B.4 Some auxiliary lemmas

Lemma B.4.1. *Let $v \in \mathbb{R}^d$ then*

$$\det(I + vv^\mathsf{T}) = 1 + \|v\|^2.$$

Proof. Note that the determinant of a matrix equals the product of all eigenvalues. Let $b_1, \ldots, b_{d-1}$ denote an orthonormal basis of $\{v\}^\perp$. Then we can easily see that each b_i is an eigenvector of $I + vv^\mathsf{T}$ with eigenvalue 1. Furthermore, $(I + vv^\mathsf{T})v = (1 + \|v\|^2)v$ implies that v is an eigenvector with eigenvalue $1 + \|v\|^2$. Hence, we have found all eigenvalues and consequently the determinant equals $1 + \|v\|^2$. ❑

Lemma B.4.2. *For $w \in \mathbb{C}^3$ with $\|w\| = 1$ the mapping $A\colon v \mapsto (w \times v) \times w$ is the orthogonal projection on the orthogonal complement of* $\operatorname{span}\{w\}$.

Proof. Note that $(w \times v) \times w = -w \times (w \times v)$ and $w \times v = \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix} v$. Therefore,

$$
(w \times v) \times w = - \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}^2 v = \begin{bmatrix} w_2^2 + w_3^2 & -w_1 w_2 & -w_1 w_3 \\ -w_1 w_2 & w_1^2 + w_3^2 & -w_2 w_3 \\ -w_1 w_3 & -w_2 w_3 & w_1^2 + w_2^2 \end{bmatrix} v
$$

Since $\|w\| = 1$ we further have

$$= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} w_1^2 & w_1 w_2 & w_1 w_3 \\ w_1 w_2 & w_2^2 & w_2 w_3 \\ w_1 w_3 & w_2 w_3 & w_3^2 \end{bmatrix} \right) v = (I - ww^\top) v,$$

which shows the claim. ❑

Lemma B.4.3. *Let A be an injective matrix and $A^\dagger = (A^\top A)^{-1} A^\top$ its Moore-Penrose inverse. Then $AA^\dagger$ is the orthogonal projection on* $\operatorname{ran} A$.

Proof. Note that $\ker A^\top = (\operatorname{ran} A)^\perp$, $\ker A = (\operatorname{ran} A^\top)^\perp$, and $\ker A^\dagger = \ker A^\top$. Therefore, $\ker AA^\dagger = \ker A^\top = (\operatorname{ran} A)^\perp$. Moreover,

$$AA^\dagger A = A(A^\top A)^{-1} A^\top A = A,$$

which implies that the $\operatorname{ran} A$ is invariant under $AA^\dagger$. Consequently $AA^\dagger$ is an orthogonal projection on $\operatorname{ran} A$. ❑

Appendix C

Linear relations

In this chapter we will introduce linear relations, which can be seen as a generalization of linear operators or as multi-valued linear operators. Although it may be possible to completely avoid this concept, it is worth to use it, as otherwise proofs can become cumbersome and some interesting links will stay hidden.

Linear relation are treated in details in the first chapter of [2].

C.1 Basics

Definition C.1.1. Let X, Y be two vector spaces over the same scalar field. Then we will call a subspace T of $X \times Y$ a *linear relation* between X and Y. A linear relation between X and X will be called a linear relation on X. ✧

Remark C.1.2. Every linear operator $T\colon X \to Y$ can be identified by a linear relation by considering the graph of T. In fact, if we consider mappings from X to Y as subsets of $X \times Y$ then T is already a linear relation. On the other hand not every linear relation comes from an operator, as $\{0\} \times Y$ demonstrates the most degenerated example. ✧

The statement $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T$ can be interpreted as Tx "=" y. If T comes from a linear operator, then this is also its literal meaning. However, for a general linear relation y is not uniquely determined by x. So from a multi-valued operator perspective this can be interpreted as $y \in Tx$.

Definition C.1.3. For a linear relation T between the vector spaces X and Y we define

- $\operatorname{dom} T \coloneqq \{x \in X \mid \exists\, y \in Y \text{ such that } \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T\}$ the *domain* of T,
- $\operatorname{ran} T \coloneqq \{y \in Y \mid \exists\, x \in X \text{ such that } \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T\}$ the *range* of T,
- $\ker T \coloneqq \{x \in X \mid \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right] \in T\}$ the *kernel* of T,
- $\operatorname{mul} T \coloneqq \{y \in Y \mid \left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right] \in T\}$ the *multi-value-part* of T.

We say T is *single-valued* or a *linear operator*, if $\operatorname{mul} T = \{0\}$. ✧

Remark C.1.4. Every linear relation T which satisfies $\operatorname{mul} T = \{0\}$ can be regarded as a linear mapping T defined on $\operatorname{dom} T$, where $Tx = y$ is well defined by $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T$. ✧

Definition C.1.5. Let X, Y, Z be vector spaces and S, T be linear relations between X and Y, and R a linear relation between Y and Z.

- $S + T := \{\left[\begin{smallmatrix} x \\ y_1 + y_2 \end{smallmatrix}\right] \in X \times Y \mid \left[\begin{smallmatrix} x \\ y_1 \end{smallmatrix}\right] \in S \text{ and } \left[\begin{smallmatrix} x \\ y_2 \end{smallmatrix}\right] \in T\}$,
- $\lambda T := \{\left[\begin{smallmatrix} x \\ \lambda y \end{smallmatrix}\right] \in X \times Y \mid \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T\}$,
- $T^{-1} := \{\left[\begin{smallmatrix} y \\ x \end{smallmatrix}\right] \in Y \times X \mid \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T\}$,
- $RS := \{\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \in X \times Z \mid \exists\, y \in Y \text{ such that } \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in S \text{ and } \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right] \in R\}$. ✧

It is easy to check that the sets defined in the previous definition are also linear relations. Furthermore, if S, T and R are linear operators, then the previous definition coincide with the usual definition of addition, scalar multiplication, inverse and composition.

Definition C.1.6. For a Banach space $(X, \|\cdot\|)$ and a linear relation A on X, we define

- $\rho(A) := \{\lambda \in \mathbb{C} \cup \{\infty\} \mid (A - \lambda)^{-1} \in \mathcal{L}_{\mathrm{b}}(X)\}$ as the *resolvent set*,
- $\sigma(A) := (\mathbb{C} \cup \{\infty\}) \setminus \rho(A)$ as the *spectrum*,
- $\sigma_{\mathrm{p}}(A) := \{\lambda \in \mathbb{C} \cup \{\infty\} \mid \ker(A - \lambda)^{-1} \neq \{0\}\}$ as *point spectrum*, and
- $\mathrm{r}(A) := \{\lambda \in \mathbb{C} \cup \{\infty\} \mid (A - \lambda)^{-1} \in \mathcal{L}_{\mathrm{b}}(\operatorname{ran}(A - \lambda), X)\}$ as the *points of regular type*,

where $\operatorname{ran}(A - \lambda)$ is endowed with the norm of X and we set $(T - \infty)^{-1} := T$ and $\operatorname{ran}(T - \infty) := \operatorname{dom} T$. ✧

Note that the definition of $(A - \infty)^{-1}$ is just to ensure that $\infty \in \sigma(A)$, if A is not bounded.

Let A be a linear relation between the normed spaces X and Y. Then A is a subspace of $X \times Y$, which is a normed space with one of the canonical norms on a product space. Hence, A can be closed as subspace of $X \times Y$.

Definition C.1.7. Let A be a linear relation between the normed spaces X and Y We say A is a *closed linear relation*, if A is closed in $X \times Y$ (as a subspace). A linear operator A from X to Y is called a *closed linear operator*, if A is closed in the sense of linear relations. ✧

Lemma C.1.8. *Let X, Y be normed spaces and A be a linear operator from X to Y (a linear relation between X and Y with* $\operatorname{mul} A = \{0\}$*). Then A is closed, if and only if, whenever a sequence $(x_n)_{n\in\mathbb{N}}$ in* $\operatorname{dom} A$ *converges to x (w.r.t. $\|\cdot\|_X$) and the sequence $(Ax_n)_{n\in\mathbb{N}}$ converges to y (w.r.t. $\|\cdot\|_Y$), then $x \in$* $\operatorname{dom} A$ *and $Ax = y$ ($\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A$).*

The equivalent statement of a closed linear operator in the previous lemma is often used as the definition of a closed linear operator.

Remark C.1.9. Note that every linear relation A between X and Y has a closure $\overline{A}$ in $X \times Y$ which is again a linear relation. However, if A is an operator, then $\overline{A}$ is not necessarily an operator. ✧

Definition C.1.10. We say a linear operator A is *closable*, if $\overline{A}$ is an operator. ✧

Lemma C.1.11. *Let X, Y, Z be Banach spaces, $B \in \mathcal{L}_{\mathrm{b}}(X, Y)$ and A a closed linear relation between Y and Z. Then AB is a closed linear relation.*

C.2 Adjoint linear relations

Definition C.2.1. Let H_1, H_2 be Hilbert spaces and A a linear relation between H_1 and H_2. Then we define the *adjoint linear relation* by

$$A^* := \left\{ \begin{bmatrix} y_2 \\ y_1 \end{bmatrix} \in H_2 \times H_1 \,\middle|\, \langle y_2, x_2 \rangle_{H_2} = \langle y_1, x_1 \rangle_{H_1} \text{ for all } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in A \right\}. \quad ✧$$

Remark C.2.2. Let H_1, H_2 be Hilbert spaces. Then the adjoint of a densely defined linear operator $A \colon H_1 \to H_2$ can be characterized by

$$\begin{bmatrix} y_2 \\ y_1 \end{bmatrix} \in A^* \quad \Leftrightarrow \quad \langle y_2, Ax \rangle_{H_2} = \langle y_1, x \rangle_{H_1} \quad \text{for all} \quad x \in \operatorname{dom} A.$$

This matches the usual definition of a Hilbert space adjoint, if we regard y_1 as $A^* y_2$

$$\langle y_2, Ax \rangle_{H_2} = \langle A^* y_2, x \rangle_{H_1}.$$

In fact we will later see that for a densely defined linear relation its adjoint is an operator. ✧

In the operator case the next lemma is often used as the definition of $\operatorname{dom} A^*$.

Lemma C.2.3. *Let A be an operator ($\operatorname{mul} A = \{0\}$). Then we can characterize the domain of A^* by*

$$x \in \operatorname{dom} A^* \quad \Leftrightarrow \quad \operatorname{dom} A \ni u \mapsto \langle x, Au \rangle_{H_2} \textit{ is continuous w.r.t. } \|\cdot\|_{X_1}.$$

Proof. If $x \in \operatorname{dom} A^*$, then there exists (at least one) $y \in H_1$ such that

$$\langle x, Au \rangle_{H_2} = \langle y, u \rangle_{H_1} \quad \text{for all} \quad u \in \operatorname{dom} A.$$

Hence, $u \mapsto \langle x, Au \rangle_{H_2}$ is bounded by $\|y\|_{H_1}$ and therefore continuous.

If $\phi \colon \operatorname{dom} A \to \mathbb{C}, u \mapsto \langle x, Au \rangle_{H_2}$ is continuous, then we can extend this mapping by continuity on $\overline{\operatorname{dom} A}$. By Hahn-Banach we can further continuously extend this on H_1 (not necessarily uniquely), denoted by $\hat{\phi}$. Since H_1 is a Hilbert space there exists a $y \in Y_1$ such that $\hat{\phi}(\cdot) = \langle y, \cdot \rangle_{H_1}$. Hence,

$$\langle x, Au \rangle_{H_2} = \hat{\phi}(u) = \langle y, u \rangle_{H_1}$$

which implies $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A^*$ and $x \in \operatorname{dom} A^*$. ❑

Lemma C.2.4. *Let H_1, H_2 be Hilbert spaces and let A be a linear relation between H_1 and H_2. Then*

- $(-A)^{-1} = -A^{-1}$,
- $(-A)^{\perp_{H_1 \times H_2}} = -A^{\perp_{H_1 \times H_2}}$ *and*
- $(A^{-1})^{\perp_{H_2 \times H_1}} = (A^{\perp_{H_1 \times H_2}})^{-1}$.

Proof. We show $(-A)^{-1} = -A^{-1}$ by

$$\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in (-A)^{-1} \;\Leftrightarrow\; \left[\begin{smallmatrix} y \\ -x \end{smallmatrix}\right] \in A \;\Leftrightarrow\; \left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right] \in A \;\Leftrightarrow\; \left[\begin{smallmatrix} x \\ -y \end{smallmatrix}\right] \in A^{-1} \;\Leftrightarrow\; \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in -A^{-1}.$$

The second assertion $(-A)^{\perp} = -A^{\perp}$ follows from

$$\begin{aligned} \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in (-A)^{\perp} &\Leftrightarrow \langle \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right], \left[\begin{smallmatrix} u \\ -v \end{smallmatrix}\right] \rangle_{H_1 \times H_2} = 0 \quad \forall \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \\ &\Leftrightarrow \langle \left[\begin{smallmatrix} x \\ -y \end{smallmatrix}\right], \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \rangle_{H_1 \times H_2} = 0 \quad \forall \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \\ &\Leftrightarrow \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in -(A^{\perp}). \end{aligned}$$

Finally, $(A^{-1})^{\perp} = (A^{\perp})^{-1}$ can be seen by

$$\begin{aligned} \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in (A^{-1})^{\perp} &\Leftrightarrow \langle \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right], \left[\begin{smallmatrix} v \\ u \end{smallmatrix}\right] \rangle_{H_1 \times H_2} = 0 \quad \forall \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \\ &\Leftrightarrow \langle \left[\begin{smallmatrix} y \\ x \end{smallmatrix}\right], \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \rangle_{H_2 \times H_1} = 0 \quad \forall \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \\ &\Leftrightarrow \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in (A^{\perp})^{-1}. \end{aligned}$$

❑

Proposition C.2.5. *Let H_1, H_2 be Hilbert spaces and A a linear relation between H_1 and H_2. Then we have the following identities*

$$A^* = ((-A)^{-1})^{\perp} = -(A^{-1})^{\perp} = -(A^{\perp})^{-1}.$$

Moreover, A^* is closed.

Proof. Note that

$$\langle y, u \rangle_{H_1} + \langle x, v \rangle_{H_2} = \langle \left[\begin{smallmatrix} y \\ x \end{smallmatrix}\right], \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \rangle_{H_1 \times H_2}$$

Therefore we can reformulate the condition in the definition of A^*

$$A^* = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in H_2 \times H_1 \;\middle|\; \left\langle \begin{bmatrix} -y \\ x \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix} \right\rangle = 0 \text{ for all } \begin{bmatrix} u \\ v \end{bmatrix} \in A \right\} = (-A^{\perp})^{-1}.$$

The other characterizations follow from Lemma C.2.4. The closedness follows from the closedness of the orthogonal complement. ❑

Lemma C.2.6. *Let H_1, H_2 and H_3 be dual pairs and A a linear relation between H_1 and H_2. Then*

(i) $\operatorname{mul} A^* = (\operatorname{dom} A)^\perp$, $\ker A^* = (\operatorname{ran} A)^\perp$,

(ii) $(BA)^* \supseteq A^*B^*$ *for all linear relations B between H_2 and H_3,*

(iii) $(BA)^* = A^*B^*$ *for all operators $B \in \mathcal{L}_\mathrm{b}(H_2, H_3)$,*

Proof.

(i) By the definition of A^*, we have

$$\begin{aligned} \operatorname{mul} A^* &= \left\{ y \in Y_2 \,\middle|\, \left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right] \in A^* \right\} \\ &= \left\{ y \in Y_2 \,\middle|\, \underbrace{\langle 0, v\rangle}_{=0} = \langle y, u\rangle \text{ for all } \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \right\} = (\operatorname{dom} A)^\perp, \end{aligned}$$

$$\begin{aligned} \ker A^* &= \left\{ x \in Y_1 \,\middle|\, \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right] \in A^* \right\} \\ &= \left\{ x \in Y_1 \,\middle|\, \langle x, v\rangle = \underbrace{\langle 0, u\rangle}_{=0} \text{ for all } \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \right\} = (\operatorname{ran} A)^\perp. \end{aligned}$$

(ii) If $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A^*B^*$, then there exist a $z \in Y_2$ such that $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \in B^*$ and $\left[\begin{smallmatrix} z \\ y \end{smallmatrix}\right] \in A^*$. Moreover,

$$\begin{aligned} \langle x, w\rangle_{H_3} &= \langle z, v\rangle_{H_2} \quad \text{for all} \quad \left[\begin{smallmatrix} v \\ w \end{smallmatrix}\right] \in B, \\ \langle z, v\rangle_{H_2} &= \langle y, u\rangle_{H_1} \quad \text{for all} \quad \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A. \end{aligned}$$

Hence, $\langle x, w\rangle_{H_3} = \langle y, u\rangle_{H_1}$ for all $\left[\begin{smallmatrix} u \\ w \end{smallmatrix}\right] \in BA$ and consequently $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in (BA)^*$.

(iii) Since B is an everywhere defined operator, we can write $BA = \{\left[\begin{smallmatrix} u \\ Bv \end{smallmatrix}\right] \,|\, \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A\}$. Therefore,

$$(BA)^* = \left\{ \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in Y_3 \times Y_1 \,\middle|\, \langle x, Bv\rangle_{H_3} = \langle y, u\rangle_{H_1} \text{ for all } \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A \right\}.$$

If $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in (BA)^*$, then

$$\langle B^*x, v\rangle_{H_2} = \langle x, Bv\rangle_{H_3} = \langle y, u\rangle_{H_1} \quad \text{for all} \quad \left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in A,$$

and in turn $\left[\begin{smallmatrix} B^*x \\ y \end{smallmatrix}\right] \in A^*$. Clearly, we also have $\left[\begin{smallmatrix} x \\ B^*x \end{smallmatrix}\right] \in B^*$. Hence $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in A^*B^*$. ❑

Lemma C.2.7. *Let H_1, H_2 be Hilbert spaces and A a linear relation between H_1 and H_2. Then*

$$A^{**} = \overline{A}.$$

Proof. By the identities in Proposition C.2.5 we have

$$A^{**} = \big(-(A^\perp)^{-1} \big)^* = \Big(\big((A^\perp)^{-1} \big)^{-1} \Big)^\perp = A^{\perp\perp} = \overline{A}. \qquad ❑$$

Definition C.2.8. Let H be a Hilbert space and A a linear relation on H. We call A

- *symmetric*, if $A \subseteq A^*$ and *self-adjoint*, if $A = A^*$.
- *skew-symmetric*, if $A \subseteq -A^*$ and *skew-adjoint*, if $A = -A^*$.
- *normal*, if $AA^* = A^*A$. ✧

Remark C.2.9. If A is symmetric/self-adjoint, then $\mathrm{i}A$ is skew-symmetric/skew-adjoint. Conversely, if A is skew-symmetric/skew-adjoint, then $\mathrm{i}A$ is symmetric/self-adjoint. ✧

Lemma C.2.10. *A self-adjoint operator A, i.e. $A^* = A$ and $\operatorname{mul} A = \{0\}$, is densely defined. A skew-adjoint operator B is densely defined.*

Proof. By Lemma C.2.6 we have

$$\overline{\operatorname{dom} A} = (\operatorname{mul} A)^\perp = \{0\}^\perp = X,$$

which proves the claim.

Clearly, this already implies the result for skew-adjoint operators, as $\mathrm{i}B$ is self-adjoint. ❏

Note that the previous result holds only for self-adjoint operators. There are self-adjoint linear relations that are not densely defined.

Theorem C.2.11 (J. von Neumann)**.** *Let T be a closed and densely defined linear operator from the Hilbert space X to the Hilbert space Y. Then T^*T and TT^* are self-adjoint, and $(\mathrm{I}_X + T^*T)$ and $(\mathrm{I}_Y + TT^*)$ are boundedly invertible.*

Proof. Since $T^* = \left[\begin{smallmatrix} 0 & \mathrm{I}_Y \\ -\mathrm{I}_X & 0 \end{smallmatrix}\right] T^\perp$, we have $T \oplus \left[\begin{smallmatrix} 0 & -\mathrm{I}_X \\ \mathrm{I}_Y & 0 \end{smallmatrix}\right] T^* = X \times Y$. Hence, for $\left[\begin{smallmatrix} h \\ 0 \end{smallmatrix}\right] \in X \times Y$ there are unique $x \in \operatorname{dom} T$ and $y \in \operatorname{dom} T^*$ such that

$$\begin{bmatrix} h \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ Tx \end{bmatrix} + \begin{bmatrix} -T^*y \\ y \end{bmatrix}. \tag{C.1}$$

Consequently, $h = x - T^*y$ and $y = -Tx$, which implies $x \in \operatorname{dom} T^*T$ and

$$h = x + T^*Tx.$$

Because of the uniqueness of the decomposition in (C.1), $x \in \operatorname{dom} T^*T$ is uniquely determined by $h \in X$. Therefore, $(\mathrm{I}_X + T^*T)^{-1}$ is a well-defined and everywhere defined operator.

For $h_1, h_2 \in X$, we define $x_1 := (\mathrm{I}_X + T^*T)^{-1}h_1$ and $x_2 := (\mathrm{I}_X + T^*T)^{-1}h_2$. Then $x_1, x_2 \in \operatorname{dom} T^*T$ and, by the closedness of T, $T^{**} = T$. Hence,

$$\begin{aligned} \langle h_1, (\mathrm{I}_X + T^*T)^{-1}h_2\rangle &= \langle (\mathrm{I}_X + T^*T)x_1, x_2\rangle = \langle x_1, x_2\rangle + \langle T^*Tx_1, x_2\rangle \\ &= \langle x_1, x_2\rangle + \langle Tx_1, Tx_2\rangle = \langle x_1, x_2\rangle + \langle x_1, T^*Tx_2\rangle \\ &= \langle x_1, (\mathrm{I}_X + T^*T)x_2\rangle = \langle (\mathrm{I}_X + T^*T)^{-1}h_1, h_2\rangle, \end{aligned}$$

which yields that $(\mathrm{I}_X + T^*T)^{-1}$ is self-adjoint. Therefore $(\mathrm{I}_X + T^*T)$ and T^*T are also self-adjoint. Moreover, $(\mathrm{I}_X + T^*T)^{-1}$ is bounded as a closed and everywhere defined operator.

By $TT^* = (T^*)^*(T^*)$ the other statements follow by the already shown. ❑

Applying this theorem to $S = \lambda T$ implies that $\mathbb{R}_-$ is contained in the resolvent set of T^*T.

Definition C.2.12. Let X, Y be normed spaces and T a linear operator between X and Y. Then we define the *graph norm* of A by

$$\|x\|_T \coloneqq \sqrt{\|x\|_X^2 + \|Tx\|_Y^2} \quad \text{for} \quad x \in \operatorname{dom} T.$$

If X and Y are Hibert spaces, then we define the *graph inner product*

$$\langle x, y\rangle_T \coloneqq \langle x, y\rangle_X + \langle Tx, Ty\rangle_Y \quad \text{for} \quad x, y \in \operatorname{dom} T. \qquad ✧$$

Note that for Hilbert spaces the graph norm is exactly the norm that is associated with the graph inner product. Moreover, the graph norm is equivalent to

$$\|x\|_{T,p} \coloneqq \begin{cases} \left(\|x\|_X^p + \|Tx\|_Y^p\right)^{1/p}, & \text{if } p \neq \infty, \\ \max(\|x\|_X, \|Tx\|_Y), & \text{if } p = \infty, \end{cases}$$

for all $p \in [1, \infty]$. Hence, if we don't work with Hilbert spaces it is sometimes more convenient to use $\|x\|_{T,1}$ as the graph norm.

Remark C.2.13. Clearly, every linear operator $T\colon \operatorname{dom} T \subseteq X \to Y$ is continuous, if we endow $\operatorname{dom} T$ with the graph norm of T. ✧

Lemma C.2.14. *Let T be a closed and densely defined linear operator from the Hilbert space X to the Hilbert space Y. Then* $\operatorname{dom} T^*T$ *is a* core *of T, i.e.* $\operatorname{dom} T^*T$ *is dense in* $\operatorname{dom} T$ *with respect to the graph norm.*

Proof. Suppose $\operatorname{dom} T^*T$ is not dense in $\operatorname{dom} T$ w.r.t the graph norm of T. Then there exists an $x \in \operatorname{dom} T$ such that

$$0 = \langle x, y\rangle_T = \langle x, y\rangle_X + \langle Tx, Ty\rangle_Y = \langle x, (\mathrm{I} + T^*T)y\rangle_X \quad \forall y \in \operatorname{dom} T^*T,$$

Note that $\operatorname{ran}(\mathrm{I} + T^*T) = X$ by Theorem C.2.11 and hence $x \perp X$, which implies $x = 0$. ❑

Corollary C.2.15. *Let T be a closed and densely defined linear operator from the Hilbert space H_1 to the Hilbert space H_2. Then*

$$\overline{T\big|_{\operatorname{dom} T^*T}} = T.$$

Proof. Let $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in T$, which is equivalent to $x \in \operatorname{dom} T$ and $Tx = y$. Then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in $\operatorname{dom} T^*T$ that converges to x with respect to the graph norm of T by Lemma C.2.14. Hence, $\|x_n - x\|_T = \|x_n - x\|_{H_1} + \|Tx_n - Tx\|_{H_2} \to 0$. This implies $\left(\left[\begin{smallmatrix} x_n \\ Tx_n \end{smallmatrix}\right]\right)$ is a sequence in $T\big|_{\operatorname{dom} T^*T}$ that converges to $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$. ❑

Bibliography

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000. doi:10.1093/oso/9780198502456.001.0001.

[2] J. Behrndt, S. Hassi, and H. de Snoo. *Boundary value problems, Weyl functions, and differential operators*, volume 108 of *Monographs in Mathematics*. Birkhäuser/Springer, Cham, [2020] ©2020. doi:10.1007/978-3-030-36714-5.

[3] D. Braess. *Finite elements*. Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker. doi:10.1017/CBO9780511618635.

[4] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. doi:10.1007/b97696.

[5] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010. doi:10.1090/gsm/019.

[6] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[7] L. Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1990. Distribution theory and Fourier analysis. doi:10.1007/978-3-642-61497-2.

[8] M. Kaltenbäck. *Fundament Analysis*, volume 26 of *Berl. Studienr. Math.* Lemgo: Heldermann Verlag, 2014.

[9] M. Kaltenbäck. *Aufbau Analysis*, volume 27 of *Berl. Studienr. Math.* Lemgo: Heldermann Verlag, 2021.

[10] J.-i. Nagata. *Modern general topology*, volume 33 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1985.

[11] M. Tucsnak and G. Weiss. *Observation and control for operator semigroups*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, 2009. doi:10.1007/978-3-7643-8994-9.