Form Methods for Evolution Equations

Wolfgang Arendt Hendrik Vogt Jürgen Voigt

February 19, 2025

Preface

The theory of forms in Hilbert spaces has its origin in the 1830's, when Johann Peter Gustav Lejeune Dirichlet used the expression $\int_\Omega |\nabla u(x)|^2 \, dx$ to find a function u satisfying the Laplace equation $\Delta u = 0$, with prescribed values of u on the boundary $\partial\Omega$, for suitable domains $\Omega \subseteq \mathbb{R}^3$. In the language introduced by his student Riemann, this task is called the "Dirichlet problem". In his reasoning the expression

$$a(u, v) = \int_\Omega \nabla u(x) \cdot \nabla v(x) \, dx,$$

for suitable functions u, v, played a decisive role; this expression nowadays is called the "classical Dirichlet form".

Forms were also the basic objects in Hilbert's famous articles on integral equations 1904–1910; see [Hil12]. Only later, by the work of Hilbert's students, Hilbert spaces were introduced and operators became the central objects. In 1932, von Neumann redefined the mathematical foundations of quantum mechanics in his book [Neu32a] which is still valid today. There the central objects are unbounded self-adjoint operators. In the same year, Stone [Sto32] used groups of unitary operators to solve the problem of describing the dynamical behaviour in quantum theory. This in turn stimulated research on one-parameter semigroups and led to its key result, the Hille–Yosida theorem from 1948, which characterises the operators A that generate a contractive C_0-semigroup on a Banach space X. Here, A being the generator of a C_0-semigroup means that the initial value problem for the evolution equation

$$u'(t) = Au(t), \qquad u(0) = x$$

is well-posed for $x \in \mathrm{dom}(A)$. Many equations from mathematical physics belong to this type of evolution equations; a prototype is the heat equation with various boundary conditions whose solutions describe how the temperature distribution in the system evolves in time. Equations of the above form, in particular parabolic equations, can be solved and analysed by the theory of one-parameter semigroups. This theory was developed to a large extent by Hille and Phillips [HiPh57], and today quite a few textbooks exist on the subject.

Of special importance are semigroups on Hilbert spaces. From the modelling side they occur in quantum theory by von Neumann's mathematical approach but also in many other physical situations, for example when the energy can be expressed by a Hilbert space norm. But there is also a purely mathematical reason: the theorem of Fréchet–Riesz expressing linear functionals in terms of the scalar product. A more elaborate version is the Lax–Milgram lemma [LaMi54] from 1954 which is formulated for general sesquilinear forms on a Hilbert space. For brevity we will use the term "form" instead of "sesquilinear form". Under some further assumptions, forms defined on a dense subspace of a Hilbert space H have another most intriguing property: they are associated with an operator A in H that generates a C_0-semigroup on H (which is even very regular, namely holomorphic). In the symmetric case this goes back to Friedrichs [Fri34] in 1934. In the general case, the subject "evolution via forms" was developed much later by Kato, starting with his papers

[Kat61b], [Kat61a] and elaborated to a wonderful theory in his book [Kat80], which first appeared in 1966. In parallel, Lions [Lio61] established an equivalent theory using a different language.

Our book is devoted to the study of evolution equations via form methods. We present the theory in the language of Kato, with densely defined forms, and that of Lions, in the spirit of Gelfand triples; in this sense our treatment is "bilingual".

We now describe the contents of the book. In the first three chapters we give an introduction to the theory of C_0-semigroups. We prove the Hille–Yosida theorem and the Euler representation formula for the semigroup. The characterisation of generators of holomorphic semigroups that are contractive on a sector is proved in an easy non-standard way by reducing it to the Hille–Yosida theorem. This approach avoids contour integrals, and using the notion of accretivity we reformulate it as a "complex Lumer–Phillips theorem", Theorem 3.20. In Chapter 4 we present basic analytical tools needed for applications, e.g. distributional derivatives of functions and the Sobolev space $H^1(\Omega)$, and we define the Dirichlet Laplacian in $L_2(\Omega)$.

In Chapter 5 we start with the main subject of the book: the theory of forms. The complex Lumer–Phillips theorem mentioned above leads to the central generation theorem for forms, Corollary 5.11. Basically, this chapter is formulated in Lions' language, but we use the recent j-method introduced in [ArEl12b]: the Hilbert space V on which the form is defined need not be embedded in the Hilbert space H where the generator is defined and where the evolution takes place. This has several advantages, for example it enables us to define the Dirichlet-to-Neumann operator in an elegant way, by using the classical Dirichlet form on the Sobolev space $H^1(\Omega)$ even though the evolution takes place in $L_2(\partial\Omega)$, i.e. on the boundary of Ω; see Chapter 8.

The invariance of a closed convex set under the semigroup can be characterised in an elegant fashion by properties of the form. Such results go back to Beurling and Deny in 1958 [BeDe58], [BeDe59] and were systematically developed by Ouhabaz [Ouh92], [Ouh05]; they are the subject of Chapter 9. Further topics concern interpolation in Chapter 10), and as an application, semigroups generated by elliptic operators under diverse boundary conditions in Chapter 11.

Sectorial forms, whose domain is just a vector space, are introduced in Chapter 12. In principle, this is the formulation of Kato. An interesting aspect of our j-method is that it allows us to avoid the notion of closability of the form; see Remark 12.13(b).

In Chapters 13 and 14 we investigate convergence of semigroups and forms. The convergence results are used in Chapter 15 to treat a topic so far not found in book form: the striking Trotter product formula for the sum of two forms, due to T. Kato for the case of symmetric forms and to B. Simon concerning the generalisation to sectorial forms.

Chapter 16 is devoted to the Stokes operator, acting in an L_2-space of divergence-free vector-valued functions, and to properties of spaces of functions and distributions arising in the treatment of the Navier–Stokes equation.

The last three chapters of the book concern non-autonomous evolution equations. We present Lions' elegant proof of existence and uniqueness of solutions via his representation theorem of Lax–Milgram type. A natural problem is that of maximal regularity, which we solve under some special conditions (Lipschitz continuity in time). The maximal regularity is applied to nonlinear problems in the last chapter.

The reader of the book is expected to have basic knowledge in analysis, functional analysis and Lebesgue integration, as prerequisites. Concerning the organisation and the style of the book a few comments are in order. Each of the 19 chapters is devoted to one special subject, specified in the title of the chapter. It is one of our principles to provide complete information, including proofs, on subjects from other areas that are used in the development of the material and in applications, such as elements of operator theory or vector-valued holomorphic functions. To achieve this aim we insert "interludes" in the chapters where we give full proofs of results exceeding a certain level in analysis and functional analysis. There are also some items that would exceed the frame of an interlude. They are collected in appendices, for example on the Stone–Weierstrass theorem, Hausdorff measure, Maz'ya's inequality (related to the isoperimetric inequality), the spectral theorem for self-adjoint operators, results on singular integrals, and the Brouwer and Schauder fixed point theorems.

The book grew out of the Internet Seminar "Form Methods for Evolution Equations, and Applications" for which the three authors figured as virtual lecturers in the academic year 2014/15. The format "Internet Seminar" was conceived and launched in the late 1990's by our colleague Rainer Nagel, Tübingen, and since then has taken place every year, with a subject in the area of evolution equations and related topics. It consists of a first phase with a series of weekly online lectures, including a discussion forum, followed by a project phase in which groups of three or four students from different countries work on various projects, and a final one-week workshop where the projects are presented and additional talks round up the subject of the ISEM. Our internet seminar was the 18th in the series.

The authors thank the co-organisers of this ISEM18, Ralph Chill and Christian Seifert, for their help and support. Moreover we are grateful to all participants of the internet seminar for their comments and enthusiasm.

Last but not least, we cordially thank our partners Frauke, Nele and Marianne for their support and patience during the realisation of this book.

Ulm, Bremen, Dresden — W. Arendt, H. Vogt, J. Voigt
August 2024 (Hi, hi???)

Contents

Chapter 1

C_0-semigroups

C_0-semigroups serve to describe the time evolution of autonomous linear systems. The objective of the present chapter is to introduce the notion of C_0-semigroups and their generators, and to derive some basic properties. In an "interlude" we provide basic definitions concerning linear operators as well as fundamental facts about integration and differentiation of Banach space valued functions.

1.1 Motivation

Let X be a Banach space, and let $\mathcal{L}(X)$ denote the space of bounded linear operators on X. A C_0-semigroup is a function $T\colon [0,\infty) \to \mathcal{L}(X)$ associated with the solutions of the initial value problem for a linear autonomous differential equation on $[0,\infty)$,

$$u' = Au, \quad u(0) = x.$$

Here, A should be a suitable (usually unbounded) linear operator in X, defined on its domain $\operatorname{dom}(A) \subseteq X$. If $x \in \operatorname{dom}(A)$, then the function $u\colon [0,\infty) \to X$ given by $u(t) := T(t)x$ should be the unique solution of the initial value problem given above. These properties will be studied in more detail in the present chapter.

If the operator A is bounded, then the problem can be treated by the usual methods of ordinary differential equations. The aim of the theory of C_0-semigroups is to describe the solution theory for the case when A is an unbounded operator. Typically, in applications the Banach space will be a space of functions, defined on an open subset of $\mathbb{R}^n$, and the operator A will be a partial differential operator. A prototypical example is the heat equation

$$\partial_t u = \Delta u$$

which will be treated in the context of C_0-semigroups; see e.g. Subsection 4.2.1.

1.2 Definition and some basic properties

Let X be a (real or complex) Banach space. A **one-parameter semigroup** on X is a function $T\colon [0,\infty) \to \mathcal{L}(X)$ satisfying

(i) $T(t+s) = T(t)T(s)$ for all $t, s \geqslant 0$.

If additionally

(ii) $\lim_{t\to 0+} T(t)x = x$ for all $x \in X$,

then T is called a ***C_0*-semigroup** on X (or a 'strongly continuous semigroup').

If T is defined on $\mathbb{R}$ instead of $[0,\infty)$, and property (i) holds for all $t, s \in \mathbb{R}$, then T is called a **one-parameter group**, and if additionally (ii) holds, then T is called a ***C_0*-group**.

1.1 Remarks. (a) Property (i) expresses that T describes the time evolution of an *autonomous* system with state space X, i.e., the law governing the further evolution of a state $T(t)x$ reached at time t depends only on the current state $T(t)x$, but not on t.

(b) Property (i) implies that for $t, s \geqslant 0$ the operators $T(t)$, $T(s)$ commute; also, if $t_1, t_2, \ldots, t_n \geqslant 0$, then $T\big(\sum_{j=1}^n t_j\big) = \prod_{j=1}^n T(t_j)$.

(c) Property (i) implies that $T(0) = T(0)^2$ is a projection.

(d) If T is a C_0-semigroup, then $T(0)x = \lim_{t\to 0+} T(t)T(0)x = \lim_{t\to 0+} T(t)x = x$ for all $x \in X$, i.e. $T(0) = I$, the identity operator in X. △

In property (i) one immediately recognises the functional equation for the exponential function, and in fact this will be our first example for a C_0-group.

1.2 Example. Let $A \in \mathcal{L}(X)$. Then

$$T(t) := \mathrm{e}^{tA} = \sum_{j=0}^{\infty} \frac{1}{j!}(tA)^j$$

converges in $\mathcal{L}(X)$ for all $t \in \mathbb{R}$ and defines a C_0-group T. In fact, $\mathbb{R} \ni t \mapsto T(t) \in \mathcal{L}(X)$ is even continuous with respect to the operator norm.

We leave this as an exercise (see Exercise 1.1). △

1.3 Lemma. *Let T be a one-parameter semigroup on X, and assume that there exists $\delta > 0$ such that $M := \sup_{0\leqslant t<\delta} \|T(t)\| < \infty$. Then there exists $\omega \in \mathbb{R}$ such that*

$$\|T(t)\| \leqslant M\mathrm{e}^{\omega t} \qquad (t \geqslant 0).$$

Proof. There exists $\omega \geqslant 0$ such that $\|T(\delta)\| \leqslant \mathrm{e}^{\omega\delta}$. For $t \geqslant 0$ there exists $n \in \mathbb{N}_0$ such that $n\delta \leqslant t < (n+1)\delta$. The semigroup property (i) implies $T(t) = T(t-n\delta)T(\delta)^n$, and therefore

$$\|T(t)\| \leqslant \|T(t-n\delta)\|\|T(\delta)\|^n \leqslant M\mathrm{e}^{\omega n\delta} \leqslant M\mathrm{e}^{\omega t}.$$ □

1.4 Proposition. *Let T be a C_0-semigroup on X.*

(a) *Then there exist $M \geqslant 0$ and $\omega \in \mathbb{R}$ such that*

$$\|T(t)\| \leqslant M\mathrm{e}^{\omega t} \qquad (t \geqslant 0).$$

(b) *For all $x \in X$ the function $[0,\infty) \ni t \mapsto T(t)x \in X$ is continuous.*

(c) *If T is a C_0-group on X, then there exist $M \geqslant 0$ and $\omega \in \mathbb{R}$ such that*

$$\|T(t)\| \leqslant M\mathrm{e}^{\omega|t|} \qquad (t \in \mathbb{R}).$$

For all $x \in X$ the function $\mathbb{R} \ni t \mapsto T(t)x \in X$ is continuous.

Part (b) of the above proposition states that the function T is 'strongly continuous'; for the definition of this notion and more information we refer to Subsection 2.1.2.

Proof of Proposition 1.4. (a) In view of Lemma 1.3 it is sufficient to show that there exists $\delta > 0$ such that $\sup_{0\leqslant t<\delta} \|T(t)\| < \infty$. Assuming that there is no such δ, we can find a null sequence (t_n) such that $\|T(t_n)\| \to \infty$ as $n \to \infty$. However, for all $x \in X$ the sequence $(T(t_n)x)$ is convergent (to x), by property (ii) of C_0-semigroups. Therefore the uniform boundedness theorem (see e.g. [Yos68; Sect. II.1, Corollary 1] or [Bre11; Theorem 2.2]) implies that $\sup_{n\in\mathbb{N}} \|T(t_n)\| < \infty$, a contradiction.

(b) Let $x \in X$, $t > 0$. Then $T(t+h)x - T(t)x = T(t)(T(h)x - x) \to 0$ as $h \to 0+$, which proves the right-sided continuity of $T(\cdot)x$. In order to prove the left-sided continuity we let $-t \leqslant h < 0$ and write $T(t+h)x - T(t)x = T(t+h)(x - T(-h)x)$. Then we obtain

$$\|T(t+h)x - T(t)x\| \leqslant \Big(\sup_{0\leqslant s\leqslant t} \|T(s)\|\Big)\|x - T(-h)x\| \to 0$$

as $h \to 0-$.

(c) First we show that, given $x \in X$, the orbit $T(\cdot)x$ is continuous. As the restriction of T to $[0,\infty)$ is a C_0-semigroup it follows from (b) that $T(\cdot)x$ is continuous on $(0,\infty)$. Thus $T(\cdot)x = T(a)T(\cdot - a)x$ is continuous on (a,∞) for all $a \in \mathbb{R}$.

As a consequence, the function $[0,\infty) \ni t \mapsto T(-t) \in \mathcal{L}(X)$ is a C_0-semigroup, and therefore satisfies an estimate as in (a). Putting the estimates for the C_0-semigroups $t \mapsto T(t)$ and $t \mapsto T(-t)$ together one obtains the asserted estimate. □

In applications it is sometimes not immediately clear how to prove the strong continuity property (ii) of a one-parameter semigroup, whereas the boundedness condition of Lemma 1.3 is easy to verify. The following condition is useful in such situations.

1.5 Lemma. *Let T be a one-parameter semigroup on X. Assume that* $\sup_{0\leqslant t\leqslant 1} \|T(t)\| < \infty$ *and that there exists a dense subset D of X such that* $\lim_{t\to 0+} T(t)x = x$ *for all $x \in D$. Then T is a C_0-semigroup.*

This lemma is an immediate consequence of the next proposition, a fundamental fact of operator theory.

Let X, Y be Banach spaces. We denote the space of bounded linear operators from X to Y by $\mathcal{L}(X,Y)$. (Writing $\mathcal{L}(X,Y)$ we will always tacitly assume that the two Banach spaces are over the same scalar field $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$.) A sequence (B_n) in $\mathcal{L}(X,Y)$ is called **strongly convergent** to $B \in \mathcal{L}(X,Y)$, abbreviated $B = \text{s-lim}_{n\to\infty} B_n$, if $Bx = \lim_{n\to\infty} B_n x$ for all $x \in X$. The uniform boundedness theorem implies that then the sequence (B_n) is bounded.

1.6 Proposition. *Let X, Y be Banach spaces, and let $(B_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{L}(X,Y)$. Assume that the sequence $(B_n x)_{n\in\mathbb{N}}$ is convergent for all x in a dense subset of X.*

Then $Bx := \lim_{n\to\infty} B_n x$ exists for all $x \in X$, and $B\colon X \to Y$ thus defined is an operator $B \in \mathcal{L}(X,Y)$. In other words, $B_n \to B$ strongly for some $B \in \mathcal{L}(X,Y)$.

Proof. A standard $\varepsilon/3$-argument shows that $(B_n x)$ is a Cauchy sequence in Y, for all $x \in X$, and therefore convergent. The linearity and boundedness of B are then easy to show. □

We note that in applications of this proposition, the limiting operator is often already known, but the pointwise convergence is only known on a dense subset. (In the application to Lemma 1.5, for instance, one has $B = I$.)

If Ω is a set and $A \subseteq \Omega$, then $\mathbf{1}_A$ denotes the **indicator function** (also known as the characteristic function) of A,

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in \Omega \setminus A. \end{cases}$$

1.7 Examples (Right translation on $L_p(\mathbb{R})$, $L_p(-\infty, 0)$, $L_p(0, \infty)$ and $L_p(0, 1)$, for $1 \leqslant p < \infty$). (a) On $L_p(\mathbb{R})$: For $t \in \mathbb{R}$ we define $T(t) \in \mathcal{L}(L_p(\mathbb{R}))$ by

$$T(t)f(x) := f(x - t) \qquad \big(x \in \mathbb{R},\ f \in L_p(\mathbb{R})\big).$$

It is clear that $T(t)$ is an isometric isomorphism for all $t \in \mathbb{R}$. Also, it is easy to show that T is a one-parameter group. Let $a, b \in \mathbb{R}$, $a < b$, and put $f := \mathbf{1}_{[a,b]}$. Then it is easy to see that $T(t)f \to f$ as $t \to 0$ (because $p < \infty$). This carries over to all linear combinations of such indicator functions. Now the linear span D of all these indicator functions is dense in $L_p(\mathbb{R})$. Therefore Lemma 1.5 implies that T is a C_0-group.

(b) On $L_p(-\infty, 0)$: The operator $T(t)$, for $t \geqslant 0$, is defined by

$$T(t)f(x) := f(x - t) \qquad \big(x \in (-\infty, 0),\ f \in L_p(-\infty, 0)\big).$$

In this case the operators $T(t)$ are not isometric for $t > 0$, but they satisfy $\|T(t)\| = 1$. Again it is easy to see that T is a one-parameter semigroup (not a group), and as in (a) one shows that T is a C_0-semigroup.

(c) On $L_p(0, \infty)$: Denote by S the C_0-semigroup of right translations on $L_p(\mathbb{R})$, defined in (a), but with time parameter t restricted to $[0, \infty)$. Consider $L_p(0, \infty)$ as the subspace $\big\{f \in L_p(\mathbb{R});\ f|_{(-\infty,0)} = 0\big\}$ of $L_p(\mathbb{R})$. Clearly the semigroup operators $S(t)$ leave this subspace invariant, and therefore the restriction T of S to this subspace is a C_0-semigroup. The operator $T(t)$ is isometric for all $t \geqslant 0$. However, $T(t)$ is not surjective if $t > 0$.

(d) On $L_p(0, 1)$: Analogously to part (b) one can define the C_0-semigroup S of right translations on $L_p(-\infty, 1)$. Similarly as in (c) one sees that $S(t)$ leaves the subspace $L_p(0, 1)$ of $L_p(-\infty, 1)$ invariant, and therefore the restriction T of S to $L_p(0, 1)$ is a C_0-semigroup. This semigroup has the property that $\|T(t)\| = 1$ for $0 \leqslant t < 1$, and that $T(t) = 0$ for $t \geqslant 1$. Because of the latter property it is called the *nilpotent* right translation semigroup on $L_p(0, 1)$. △

1.3 Interlude: operators, integration and differentiation

1.3.1 Operators

We have already indicated in Section 1.1 that, besides bounded linear operators between normed spaces, one also needs linear operators defined on subspaces. One can (and should) think of an operator A from X to Y as a linear mapping $A\colon \operatorname{dom}(A) \to Y$, where $\operatorname{dom}(A)$ is a subspace of X. For some purposes, however, it is convenient to go back to the original meaning of a mapping as a relation in $X \times Y$ possessing certain additional properties.

Let X, Y be two vector spaces over the same field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. For a **linear relation** A in $X \times Y$, i.e. a subspace of $X \times Y$, we define the **domain** of A,

$$\operatorname{dom}(A) := \{x \in X\,;\ \text{there exists } y \in Y \text{ such that } (x, y) \in A\},$$

the **range** of A,

$$\operatorname{ran}(A) := \{y \in Y\,;\ \text{there exists } x \in X \text{ such that } (x, y) \in A\},$$

and the **kernel** (or **null space**) of A,

$$\ker(A) := \{x \in X\,;\ (x, 0) \in A\}.$$

The linear relation

$$A^{-1} := \{(y, x)\,;\ (x, y) \in A\}$$

in $Y \times X$ is the **inverse (relation)** of A. If B is another linear relation in $X \times Y$, satisfying $A \subseteq B$, then B is called an **extension** of A, and A a **restriction** of B. If the spaces X and Y coincide and A is a linear relation in $X \times X$, then we will also call A a linear relation in X.

In this setting, a **linear operator from X to Y** is a linear relation in $X \times Y$ satisfying

$$A \cap (\{0\} \times Y) = \{(0, 0)\}.$$

Then for all $x \in \operatorname{dom}(A)$ there exists a unique $y \in Y$ such that $(x, y) \in A$, and we will write $Ax = y$. In this sense, A is also a mapping $A\colon \operatorname{dom}(A) \to Y$. As we will only consider linear operators we will mostly drop 'linear' and simply speak of 'operators'. If the spaces X and Y coincide, then we call A an **operator in X**.

Next, let X and Y be Banach spaces. We define a norm on $X \times Y$ by

$$\|(x, y)\|_{X \times Y} := \|x\|_X + \|y\|_Y \qquad ((x, y) \in X \times Y),$$

which makes $X \times Y$ a Banach space. In this context an operator A from X to Y (or a linear relation $A \subseteq X \times Y$) is called **closed** if A is a closed subset of $X \times Y$. An operator A from X to Y is **closable** if its **closure** $\overline{A}$ in $X \times Y$ is also an operator; we refer to Exercises 1.2 and 1.3 for examples in which $\overline{A}$ is not an operator.

For a subspace $D \subseteq \operatorname{dom}(A)$, the **restriction** of A to D is the operator $A|_D := A \cap (D \times Y)$. The set D is called a **core for** A if A is a restriction of the closure of $A|_D$, i.e. $A \subseteq \overline{A|_D}$.

Finally, if A and B are operators from X to Y, then the **sum** of A and B is the operator defined by

$$\operatorname{dom}(A + B) := \operatorname{dom}(A) \cap \operatorname{dom}(B), \qquad (A + B)x := Ax + Bx \quad (x \in \operatorname{dom}(A + B)),$$

or, expressed differently,

$$A + B = \big\{(x, Ax + Bx)\,;\ x \in \operatorname{dom}(A) \cap \operatorname{dom}(B)\big\} \subseteq X \times Y.$$

1.3.2 Integration of continuous functions

Let $a, b \in \mathbb{R}$, $a < b$, and let X be a Banach space. We define the space of **step functions** from $[a, b]$ to X,

$$T([a,b];X) := \operatorname{lin}\{\mathbf{1}_{[s,t]}(\cdot)x\,;\ a \leqslant s \leqslant t \leqslant b,\ x \in X\},$$

where $\mathbf{1}_{[s,t]}(\cdot)x$ denotes the function $r \mapsto \mathbf{1}_{[s,t]}(r)x$ and 'lin' denotes the linear span. (The letter 'T' stands for the German 'Treppenfunktion'.) For $f \in T([a,b];X)$, $f = \sum_{j=1}^{n} \mathbf{1}_{[s_j,t_j]}(\cdot)x_j$, we define the integral

$$\int_a^b f(t)\,\mathrm{d}t := \sum_{j=1}^{n} (t_j - s_j)x_j\,.$$

It is standard to show that this integral is well-defined, that the mapping $f \mapsto \int_a^b f(t)\,\mathrm{d}t$ is linear, and that

$$\Big\|\int_a^b f(t)\,\mathrm{d}t\Big\| \leqslant \int_a^b \|f(t)\|\,\mathrm{d}t \leqslant (b-a)\sup_{a\leqslant t\leqslant b}\|f(t)\| \qquad \big(f \in T([a,b];X)\big). \tag{1.1}$$

In short, the mapping $T([a,b];X) \ni f \mapsto \int_a^b f(t)\,\mathrm{d}t \in X$ is a bounded linear operator, where $T([a,b];X)$ is provided with the supremum norm. This implies that there is a unique continuous extension of the integral to functions in the closure of $T([a,b];X)$ in the Banach space

$$\ell_\infty([a,b];X) := \{f\colon [a,b] \to X\,;\ f \text{ bounded}\},$$

provided with the supremum norm. We denote this closure by $R([a,b];X)$ (the space of 'regulated functions') and observe that $R([a,b];X)$ contains the space of continuous functions $C([a,b];X)$. Moreover, the extension of the integral is linear, and the inequalities (1.1) carry over to all $f \in R([a,b];X)$. For $X = \mathbb{R}$ the functions in $R([a,b];X)$ are Riemann integrable, and the integral defined above is the Riemann integral.

Next we describe how operators act on integrals.

1.8 Theorem. *Let X, Y be Banach spaces, and let $a, b \in \mathbb{R}$, $a < b$.*

(a) *Let $f\colon [a,b] \to X$ be continuous, and let $A \in \mathcal{L}(X,Y)$. Then*

$$A\int_a^b f(t)\,\mathrm{d}t = \int_a^b Af(t)\,\mathrm{d}t.$$

(b) *Let A be a closed operator from X to Y. Let $f\colon [a,b] \to X$ be continuous, $f(t) \in \operatorname{dom}(A)$ for all $t \in [a,b]$, and $t \mapsto Af(t) \in Y$ continuous. Then $\int_a^b f(t)\,\mathrm{d}t \in \operatorname{dom}(A)$, and*

$$A\int_a^b f(t)\,\mathrm{d}t = \int_a^b Af(t)\,\mathrm{d}t.$$

Proof. (a) The equality is clear for step functions and carries over to continuous functions (in fact, even to regulated functions) by continuous extension.

(b) The hypotheses are just a complicated way of expressing that one is given a continuous function $t \mapsto (f(t), g(t)) \in A \subseteq X \times Y$ (where $g(t) = Af(t)$). Because A is a closed subspace of $X\times Y$, and therefore a Banach space, it follows that $\int_a^b (f(t), g(t))\,\mathrm{d}t \in A$. As the canonical projections from $X \times Y$ to X and to Y are bounded linear operators, one concludes from part (a) that

$$\Big(\int_a^b f(t)\,\mathrm{d}t, \int_a^b g(t)\,\mathrm{d}t\Big) = \int_a^b (f(t), g(t))\,\mathrm{d}t \in A,$$

and this proves the assertions. □

The last issue in this interlude is the connection between differentiation and integration, i.e., the fundamental theorem of differential and integral calculus for Banach space valued functions.

1.9 Theorem. *Let X be a Banach space, and let $a, b \in \mathbb{R}$, $a < b$.*
(a) *Let $f \colon [a, b] \to X$ be continuous,*

$$F(t) := \int_a^t f(s)\,\mathrm{d}s \qquad (a \leqslant t \leqslant b).$$

Then F is continuously differentiable, and $F' = f$.
(b) *Let $g \colon [a, b] \to X$ be continuously differentiable. Then*

$$\int_a^b g'(t)\,\mathrm{d}t = g(b) - g(a). \tag{1.2}$$

Proof. (a) is proved in the same way as for scalar-valued functions; cf. Exercise 1.7(a).
(b) For $\eta \in X'$ ($= \mathcal{L}(X, \mathbb{K})$, the dual space of X) the function $\eta \circ g$ is continuously differentiable, and one has $(\eta \circ g)' = \eta \circ g'$. The fundamental theorem of differential and integral calculus for $\mathbb{K}$-valued functions then implies that

$$\begin{aligned}\eta\Big(\int_a^b g'(t)\,\mathrm{d}t\Big) &= \int_a^b \eta(g'(t))\,\mathrm{d}t = \int_a^b (\eta \circ g)'(t)\,\mathrm{d}t \\ &= (\eta \circ g)(b) - (\eta \circ g)(a) = \eta\big(g(b) - g(a)\big).\end{aligned}$$

As this holds for all $\eta \in X'$, equation (1.2) follows from an application of the theorem of Hahn–Banach (see e.g. [Bre11; Section 1.1] or [Voi20; Appendix A]). □

1.4 The generator of a C_0-semigroup

Let X be a Banach space. For a C_0-semigroup T on X we define the **generator** (also called the **infinitesimal generator**) A, an operator in X, by

$$A := \big\{(x, y) \in X \times X\,;\ \tfrac{1}{h}(T(h)x - x) \to y\ (h \to 0+)\big\}.$$

In other words, since $T(0)x = x$, the domain of A consists of those $x \in X$ for which the orbit $t \mapsto T(t)x$ is (right-sided) differentiable at $t = 0$, and the image of x under A is the derivative of this orbit at $t = 0$.

1.10 Example. If $A \in \mathcal{L}(X)$, then A is the generator of the C_0-group $(\mathrm{e}^{tA})_{t\in\mathbb{R}}$; see Example 1.2 and Exercise 1.1. △

Similarly to the property that for the exponential function $t \mapsto \mathrm{e}^{ta}$ the derivative a at 0 determines the function, we will see that the generator determines the C_0-semigroup.

First we derive some fundamental properties of the generator.

1.11 Theorem. *Let T be a C_0-semigroup on X, with generator A. Then:*

(a) *For $x \in \mathrm{dom}(A)$ one has $T(t)x \in \mathrm{dom}(A)$ for all $t \geqslant 0$, the function $t \mapsto T(t)x$ is continuously differentiable on $[0,\infty)$, and*

$$\frac{\mathrm{d}}{\mathrm{d}t}T(t)x = AT(t)x = T(t)Ax \qquad (t \geqslant 0).$$

(b) *For all $x \in X$, $t \geqslant 0$ one has $\int_0^t T(s)x\,\mathrm{d}s \in \mathrm{dom}(A)$,*

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

(c) $\mathrm{dom}(A)$ *is dense in X, and A is a closed operator.*

Before the proof we insert a small fact on strong convergence of operators.

1.12 Lemma. *Let X, Y be Banach spaces, and let (B_n) be a sequence in $\mathcal{L}(X,Y)$, $B_n \to B \in \mathcal{L}(X,Y)$ strongly as $n \to \infty$. Let (x_n) in X, $x_n \to x \in X$ as $n \to \infty$.*

Then $B_n x_n \to Bx$ as $n \to \infty$.

Proof. The uniform boundedness theorem implies that $M := \sup_{n\in\mathbb{N}} \|B_n\| < \infty$. Therefore

$$\begin{aligned}\|Bx - B_n x_n\| &\leqslant \|Bx - B_n x\| + \|B_n(x - x_n)\| \\ &\leqslant \|Bx - B_n x\| + M\|x - x_n\| \to 0 \qquad (n \to \infty).\end{aligned}$$

□

Proof of Theorem 1.11. (a) For $t \geqslant 0$, $h > 0$ one has

$$\frac{1}{h}\big(T(t+h)x - T(t)x\big) = \frac{1}{h}(T(h) - I)T(t)x = T(t)\frac{1}{h}(T(h)x - x).$$

As $h \to 0$, the third of these expressions converges to $T(t)Ax$. Looking at the second term one obtains $T(t)x \in \mathrm{dom}(A)$, and looking at the first term one concludes that $t \mapsto T(t)x$ is right-sided differentiable, with right-sided derivative

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

On the other hand, let $t > 0$, $h \in (0,t)$. Then

$$\frac{1}{-h}\big(T(t-h)x - T(t)x\big) = T(t-h)\frac{1}{h}\big(T(h)x - x\big),$$

and this converges to $T(t)Ax$ as $h \to 0$, by Lemma 1.12.

So we have shown that the continuous function $t \mapsto AT(t)x = T(t)Ax$ is the derivative of $t \mapsto T(t)x$.

(b) For $t \geqslant 0$ we put $f(t) := \int_0^t T(s)x \,\mathrm{d}s$. Then Theorem 1.8(a) implies

$$T(h)f(t) = \int_0^t T(s+h)x \,\mathrm{d}s = \int_h^{t+h} T(s)x \,\mathrm{d}s = f(t+h) - f(h)$$

for all $h \geqslant 0$ (where we have used basic rules of the Riemann integral that are easily seen to be valid for Banach space valued functions as well). By Theorem 1.9(a) it follows that $\frac{\mathrm{d}}{\mathrm{d}h} T(h)f(t)|_{h=0} = f'(t) - f'(0) = T(t)x - x$. Thus we obtain $f(t) \in \operatorname{dom}(A)$, $Af(t) = T(t)x - x$.

(c) Let $x \in X$. Then $h^{-1}\int_0^h T(s)x \,\mathrm{d}s \in \operatorname{dom}(A)$ for all $h > 0$, by part (b). Theorem 1.9(a) implies that $h^{-1}\int_0^h T(s)x \,\mathrm{d}s \to x$ $(h \to 0+)$. This shows that $\operatorname{dom}(A)$ is dense in X.

Let $\big((x_n, y_n)\big)_{n\in\mathbb{N}}$ be a sequence in A, $(x_n, y_n) \to (x, y)$ in $X \times X$. From part (a) and Theorem 1.9(b) we obtain

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

for all $t > 0$, $n \in \mathbb{N}$. For $n \to \infty$ we conclude that

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

and then

$$\frac{1}{t}(T(t)x - x) = \frac{1}{t}\int_0^t T(s)y \,\mathrm{d}s \to y \qquad (t \to 0+).$$

This shows that $(x, y) \in A$. □

Theorem 1.11(a) implies in particular that the function $t \mapsto T(t)x$ solves the initial value problem $u' = Au$, $u(0) = x$. We now show that this solution is unique and conclude that the generator determines the C_0-semigroup.

1.13 Theorem. *Let T be a C_0-semigroup on X, with generator A.*

(a) *Let $b \in (0, \infty]$, and let $u\colon [0, b) \to X$ be continuous, $u(t) \in \operatorname{dom}(A)$ for all $t \in (0, b)$, u differentiable on $(0, b)$, and $u'(t) = Au(t)$ for all $t \in (0, b)$. Then $u(t) = T(t)u(0)$ for all $t \in [0, b)$.*

(b) *Let S be a C_0-semigroup on X, with generator $B \supseteq A$. Then $S = T$, $B = A$.*

Proof. (a) Let $0 < t < b$. Lemma 1.12 implies that the function $[0, t] \ni s \mapsto T(t-s)u(s) \in X$ is continuous, and it is not difficult to see that this function is differentiable on $(0, t]$, with derivative

$$\frac{\mathrm{d}}{\mathrm{d}s} T(t-s)u(s) = -T(t-s)Au(s) + T(t-s)u'(s) = 0$$

(a kind of product rule; see Exercise 1.8). Therefore Theorem 1.9(b) yields

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

(b) Let $x \in \operatorname{dom}(A)$, and put $u(t) := T(t)x$ $(t \geqslant 0)$. Then u satisfies the equation $u'(t) = Au(t) = Bu(t)$ $(t \geqslant 0)$. From part (a) it follows that $u(t) = S(t)u(0) = S(t)x$ for all $t \geqslant 0$. This shows that $S(t) = T(t)$ on $\operatorname{dom}(A)$. As $\operatorname{dom}(A)$ is dense in X one obtains $S(t) = T(t)$ for all $t \geqslant 0$. The equality of the semigroups implies equality of the generators. □

If A is the generator of a C_0-semigroup T, then sometimes one uses the notation $\mathrm{e}^{tA} := T(t)$ $(t \geqslant 0)$, analogously to the case when $A \in \mathcal{L}(X)$.

1.14 Remark. In Theorems 1.11 and 1.13 we have established important properties of C_0-semigroups and their generators which we summarise as follows, putting them into the context of 'well-posedness'. Let A be the generator of a C_0-semigroup T. Then the initial value problem (also called Cauchy problem)

$$u'(t) = Au(t) \quad (t \geqslant 0), \qquad u(0) = x \in \operatorname{dom}(A)$$

is **well-posed**, i.e., (i) it possesses a unique solution (given by $u(t) = T(t)x$), and (ii) the solution depends continuously on the initial value, which in the present (linear) case is expressed by $\sup_{0\leqslant t\leqslant b} \|u(t)\| \leqslant \sup_{0\leqslant t\leqslant b} \|T(t)\|\|x\|$ for all $b \in (0,\infty)$.

These properties also serve as a motivation to characterise operators that are generators; see Chapter 2. △

For determining the generators of the C_0-semigroups described in Examples 1.7 the following result will be useful.

1.15 Proposition (Nelson's lemma). *Let T be a C_0-semigroup on X, and let A be its generator. Let D be a subspace of* $\operatorname{dom}(A)$ *that is dense in X, and assume that D is invariant under T (i.e. $T(t)(D) \subseteq D$ for all $t \geqslant 0$).*

Then D is a core for A.

Proof. Let $x \in \operatorname{dom}(A)$. We have to show that $(x, Ax) \in \overline{A|_D}$.

Let (x_n) be a sequence in D, $x_n \to x$ in X as $n \to \infty$. Let $n \in \mathbb{N}$, $t > 0$. The function $[0,t] \ni s \mapsto \big(T(s)x_n, AT(s)x_n\big) \in A|_D$ is continuous; therefore (recall Theorem 1.11(a) and Theorem 1.9(b) for the first equality)

$$\begin{aligned}\Big(\int_0^t T(s)x_n \,\mathrm{d}s, T(t)x_n - x_n\Big) &= \Big(\int_0^t T(s)x_n \,\mathrm{d}s, \int_0^t AT(s)x_n \,\mathrm{d}s\Big) \\ &= \int_0^t \big(T(s)x_n, AT(s)x_n\big) \,\mathrm{d}s \in \overline{A|_D}\,.\end{aligned}$$

Letting $n \to \infty$ we conclude that

$$\Big(\int_0^t T(s)x \,\mathrm{d}s, T(t)x - x\Big) \in \overline{A|_D}\,.$$

Dividing by t and taking the limit $t \to 0+$ we obtain $(x, Ax) \in \overline{A|_D}$. □

1.16 Example. Let T be the C_0-group of right translations on $L_p(\mathbb{R})$ (introduced in Example 1.7(a)), and let A be its generator. It is not difficult to show that $D := C_\mathrm{c}^1(\mathbb{R})$ (the continuously differentiable functions with compact support) is a subspace of $\operatorname{dom}(A)$, and $Af = -f'$ for all $f \in C_\mathrm{c}^1(\mathbb{R})$. Also $C_\mathrm{c}^1(\mathbb{R})$ is obviously invariant under T, and is a dense subspace of $L_p(\mathbb{R})$. Thus D is a core for A, by Proposition 1.15. For the reader familiar with Sobolev spaces: this implies that $\operatorname{dom}(A) = W_p^1(\mathbb{R})$. For $p = 2$ we refer to Example 4.18(a).

We refer to Exercise 1.5 for the case of translation semigroups on spaces of continuous functions. △

Notes

The classical treatise on one-parameter semigroups is the monograph by Hille and Phillips [HiPh57]. The notion 'C_0-semigroup' goes back to Phillips [Phi55]; in [HiPh57; Section 10.6] it is put into context with other properties of one-parameter semigroups. In those days other continuity conditions were also considered, for instance

$$\frac{1}{t}\int_0^t T(s)x\,\mathrm{d}s \to x \qquad (t \to 0)$$

for all $x \in X$, which leads to 'C_1-semigroups'.

Many authors use the terminology 'strongly continuous semigroup' instead of C_0-semigroup. However, in Chapter 13 we will encounter semigroups T that are strongly continuous at the origin, while $T(0)$ is not necessarily the identity. For this reason we prefer the terminology 'C_0-semigroup'. Here is a – by far incomplete – list of books on C_0-semigroups: [Dav80], [Paz83], [Gol85], [Nag86], [EnNa00].

The idea of using semigroup invariance to establish a core property, as in Proposition 1.15, is due to Nelson: in [Nel59; proof of Lemma 5.1] he implemented this idea for unitary C_0-groups on Hilbert spaces.

A core can also be seen as a 'space of uniqueness'. In fact, given a subspace D of the domain $\operatorname{dom}(A)$ of the generator A of a C_0-semigroup the following holds. If D is a core for A, then Theorem 1.13(b) implies that A is the only extension of $A|_D$ generating a C_0-semigroup, whereas there are infinitely many such extensions if D is not a core; see [Nag86; Chap. A-II, Theorem 1.33].

The concept of well-posedness that we have introduced in Remark 1.14 is generally attributed to Jacques Hadamard. Indeed, in [Had02] Hadamard discussed problems for the Laplace equation and the wave equation, and he investigated whether they are "parfaitement bien posé" (perfectly well posed), by which he meant that they are "possible" and "déterminé" (i.e. the solution exists and is unique). Only much later, in [Had23; § 19, p. 35], he discussed continuous dependence on the initial data, for the wave equation.

To our knowledge, the three properties of existence, uniqueness and continuous dependence were summed up to one notion for the first time in the book [CoHi37; Chap. III, § 7] by Courant and Hilbert. For a problem from mathematical physics to be "properly posed" they set "the following basic requirements:

(1) The solution must exist.

(2) The solution should be uniquely determined.

(3) The solution should depend continuously on the data (requirement of stability)."

(Our quotation is from the English edition [CoHi62; Chap. III, § 6].) In many cases the converse of (3) is obvious, namely that the data depend continuously on the solution. Then, for linear problems, continuous dependence of the solution on the data follows automatically from existence and uniqueness, by the bounded inverse theorem (see e.g. [Yos68; Section II.5] or [Bre11; Corollary 2.7] for this theorem).

We refer to [Goc16; Section 2.3.3], [MaSh98; Section 15.1] and [MaSc65] for more information on well-posedness and on Hadamard's contributions concerning this topic. In Chapter 19 we will investigate some nonlinear problems that are *not* well-posed.

Equations that involve time are frequently called *evolution equations*; the idea is that the solution evolves in time from a given initial value. In this context, an autonomous driving mechanism and well-posedness together with linearity lead naturally to the concept of a C_0-semigroup. In this book we will see that many physical evolutionary problems are well-posed and yield a semigroup. So, Hille's famous sentence from the foreword of his monograph [Hil48] from 1948 will be confirmed: "I hail a semi-group when I see one and I seem to see them everywhere!"

Exercises

1.1 Let X be a Banach space, and let $A, B \in \mathcal{L}(X)$ satisfy $AB = BA$.

(a) Show that $\mathrm{e}^A := \sum_{j=0}^\infty \frac{1}{j!} A^j$ is absolutely convergent in $\mathcal{L}(X)$, $\|\mathrm{e}^A\| \leqslant \mathrm{e}^{\|A\|}$, and that $\mathrm{e}^{A+B} = \mathrm{e}^A \mathrm{e}^B = \mathrm{e}^B \mathrm{e}^A$.

(b) Show that $t \mapsto \mathrm{e}^{tA}$ is a one-parameter group, that the function is continuously differentiable as an $\mathcal{L}(X)$-valued function, and that $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = A\mathrm{e}^{tA}$ for all $t \in \mathbb{R}$. (In particular, A is the generator of $(\mathrm{e}^{tA})_{t\in\mathbb{R}}$.)

(c) Show that $\mathrm{e}^{zI} = \mathrm{e}^z I$ for all $z \in \mathbb{K}$.

1.2 Let $x_0 \in [0,1]$, and define the operator A from $L_2(0,1)$ to $\mathbb{K}$ by $\mathrm{dom}(A) := C[0,1]$,

$$Af := f(x_0) \qquad (f \in C[0,1]).$$

Show that $\overline{A} = L_2(0,1) \times \mathbb{K}$. (Thus A is not closable.)

1.3 Let X and Y be Banach spaces.

(a) Let A be an operator from X to Y. Show that A is closable if and only if for every null sequence (x_n) in $\mathrm{dom}(A)$ for which (Ax_n) is convergent one has $\lim_{n\to\infty} Ax_n = 0$.

(b) Let (x_n) be a linearly independent null sequence in X, and let (y_n) be a sequence in Y converging to some element $y \neq 0$ in Y. Define an operator A from X to Y by extending the assignments $Ax_n := y_n$ for all $n \in \mathbb{N}$ by linearity to $\mathrm{dom}(A) := \mathrm{lin}\{x_1, x_2, \ldots\}$. Show that the operator A is not closable.

1.4 For a locally compact subset G of $\mathbb{R}^n$ we define $C_\mathrm{c}(G) := \{f \in C(G);\, \mathrm{spt}\, f \text{ compact}\}$. (Here $C(G)$ denotes the space of continuous $\mathbb{K}$-valued functions, and the **support** of f is defined by $\mathrm{spt}\, f := \overline{[f \neq 0]}^G$, where $[f \neq 0] := \{x \in G;\ f(x) \neq 0\}$. We recall that a topological space is called locally compact if every point possesses a compact neighbourhood. Local compactness is not needed for solving this exercise; it is required because the notation C_c and C_0 is only common for locally compact spaces.)

The space $C_0(G)$ is defined as the closure of $C_\mathrm{c}(G)$ in $C_\mathrm{b}(G)$ (the space of bounded continuous functions, provided with the supremum norm).

(a) Show that every function in $C_0(G)$ is uniformly continuous.

(b) Show that

$$C_0(G) = \bigl\{f \in C(G);\, \forall \varepsilon > 0\colon [|f| \geqslant \varepsilon] \text{ compact}\bigr\}$$

(where $[|f| \geqslant \varepsilon] := \{x \in G;\, |f(x)| \geqslant \varepsilon\}$).

Hint: For the inclusion $\supseteq$ show first that it is sufficient to treat the case of real-valued functions. Then approximate real-valued functions f by $\max\{f-\varepsilon,0\}+\min\{f+\varepsilon,0\}$.

(c) Show that

$$C_0(-\infty,0]=\bigl\{f\in C(-\infty,0]\,;\ \lim_{x\to-\infty}f(x)=0\bigr\}$$

and that $C_0(0,1]\cong\{f\in C[0,1]\,;\ f(0)=0\}$.

1.5 (a) Convince yourself that analogously to Example 1.7 one can define the one-parameter semigroup T of right translations on each of the spaces $C_0(\mathbb{R})$, $C_0(-\infty,0]$, $C_0(0,\infty)$, $C_0(0,1]$ (defined in Exercise 1.4), and show that T is a C_0-semigroup.

(b) Show that the generator A of T on $C_0(\mathbb{R})$ is given by

$$\operatorname{dom}(A)=\bigl\{f\in C^1(\mathbb{R})\,;\ f,f'\in C_0(\mathbb{R})\bigr\},\qquad Af=-f'.$$

(c) Show that the generator A of T on $C_0(0,1]$ is given by

$$\operatorname{dom}(A)=\bigl\{f\in C^1(0,1]\,;\ f,f'\in C_0(0,1]\bigr\},\qquad Af=-f'.$$

1.6 Let $(\Omega,\mathcal{A},\mu)$ be a measure space, and let $1\leqslant p<\infty$. Let $a\colon\Omega\to\mathbb{K}$ be a measurable function.

(a) The **maximal multiplication operator** M_a **induced** by the function a is defined by

$$M_a=\bigl\{(f,g)\in L_p(\mu)\times L_p(\mu)\,;\ g=af\bigr\}.$$

Show that M_a is closed and densely defined.

(b) Assume additionally that $\operatorname{Re}a(x)\leqslant 0$ $(x\in\Omega)$. For $t\geqslant 0$ define $T(t)\in\mathcal{L}(L_p(\mu))$ by

$$T(t)f:=\mathrm{e}^{ta}f\qquad(f\in L_p(\mu)).$$

Show that T is a C_0-semigroup on $L_p(\mu)$ and that M_a is the generator of T.

1.7 Let X be a Banach space.

(a) Let $f\colon[0,1]\to X$ be continuous. Show that $\lim_{h\to0+}h^{-1}\int_0^h f(t)\,\mathrm{d}t=f(0)$. (This is the main step for proving Theorem 1.9(a).)

(b) Let T be a C_0-semigroup on X that is continuous with respect to the operator norm. Show that then the generator A of T is an operator $A\in\mathcal{L}(X)$.

Hint: Show that, for small t, the operator $\int_0^t T(s)\,\mathrm{d}s$ is invertible in $\mathcal{L}(X)$; this is a consequence of the Neumann series, which we will recall in Remark 2.3(a). Note that Theorem 1.11(b) implies that the operator $x\mapsto A\int_0^t T(s)x\,\mathrm{d}s$ belongs to $\mathcal{L}(X)$.

1.8 Let X be a Banach space, and let T be a C_0-semigroup on X with generator A. Let $I\subseteq[0,\infty)$ be an interval, let $u\colon I\to X$ be differentiable at some point $t_0\in I$, and suppose that $u(t_0)\in\operatorname{dom}(A)$. Show that the function $I\ni t\mapsto T(t)u(t)$ is differentiable at t_0, with derivative $AT(t_0)u(t_0)+T(t_0)u'(t_0)=T(t_0)(Au(t_0)+u'(t_0))$.

Chapter 2

Characterisation of generators of C_0-semigroups

Generators of C_0-semigroups have special spectral properties. We will study these properties in Section 2.2 and use them to characterise generators of C_0-semigroups in the Hille–Yosida theorem, the main result of this chapter. The exponential formula for C_0-semigroups presented in Section 2.3 will be important for applications. We start with an interlude on spectral theory of operators as well as on more facts about integration.

2.1 Interlude: the resolvent of operators, and some more integration

2.1.1 Resolvent set, spectrum and resolvent

Let X be a Banach space over $\mathbb{K}$, and let A be an operator in X.

We define the **resolvent set** of A,

$$\rho(A) := \bigl\{\lambda \in \mathbb{K};\ \lambda I - A\colon \operatorname{dom}(A) \to X \text{ bijective},\ (\lambda I - A)^{-1} \in \mathcal{L}(X)\bigr\}.$$

The mapping

$$R(\cdot, A)\colon \rho(A) \to \mathcal{L}(X), \quad \lambda \mapsto R(\lambda, A) := (\lambda I - A)^{-1}$$

is called the **resolvent** of A. The **spectrum** of A is the set

$$\sigma(A) := \mathbb{K} \setminus \rho(A).$$

2.1 Remarks. (a) If $\rho(A) \neq \varnothing$ and $\lambda \in \rho(A)$, then $(\lambda I - A)^{-1} \in \mathcal{L}(X)$ is closed – note that every operator belonging to $\mathcal{L}(X)$ is closed. Hence $\lambda I - A$ is closed, and therefore A is closed, by the reasoning presented subsequently in part (b).

(b) If A is a closed operator and $B \in \mathcal{L}(X)$, then the sum $A + B$ is a closed operator. Indeed, if $\bigl((x_n, y_n)\bigr)$ is a sequence in $A + B$, $(x_n, y_n) \to (x, y)$ in $X \times X$ as $n \to \infty$, then $Bx_n \to Bx$, and therefore $Ax_n = (A + B)x_n - Bx_n \to y - Bx$, and the hypothesis that A is closed implies that $(x, y - Bx) \in A$, i.e., $x \in \operatorname{dom}(A)$ and $y = Ax + Bx$.

(c) Let A be a closed operator. Assume that $\lambda \in \mathbb{K}$ is such that $\lambda I - A\colon \operatorname{dom}(A) \to X$ is bijective. Then the inverse $(\lambda I - A)^{-1}$ is a closed operator which is defined on all of X. Therefore the closed graph theorem (see e.g. [Yos68; Sect. II.6, Theorem 1] or [Bre11; Theorem 2.9]) implies that $(\lambda I - A)^{-1} \in \mathcal{L}(X)$. It follows that

$$\rho(A) = \{\lambda \in \mathbb{K};\ \lambda I - A\colon \operatorname{dom}(A) \to X \text{ bijective}\}.$$

(d) Usually, in treatments of operator theory, the above notions are only defined for the case of complex Banach spaces, because many important results of spectral theory depend on complex analysis of one variable. For our purpose it is – for the moment – possible and convenient to include the case of real scalars. △

Before proceeding we include a piece of notation that will be used in different contexts. If (M,d) is a metric space, $x \in M$ and $r \in (0,\infty]$, then

$$B(x,r) := \{y \in M\,;\ d(y,x) < r\} \quad \text{and} \quad B[x,r] := \{y \in M\,;\ d(y,x) \leqslant r\}$$

are the **open ball** and **closed ball** with centre x and radius r, respectively. If necessary, we may also write $B_M(x,r)$ and $B_M[x,r]$ for making it clear in which metric space we are.

The following theorem contains the basic results concerning the resolvent.

2.2 Theorem. *Let A be a closed operator in X.*

(a) *If $\lambda \in \rho(A)$, $x \in \operatorname{dom}(A)$, then $AR(\lambda,A)x = R(\lambda,A)Ax$.*

(b) *For all $\lambda, \mu \in \rho(A)$ one has the* ***resolvent equation***

$$R(\lambda,A) - R(\mu,A) = (\mu - \lambda)R(\mu,A)R(\lambda,A); \tag{2.1}$$

in particular, the resolvents $R(\mu,A)$, $R(\lambda,A)$ commute.

(c) *For $\lambda \in \rho(A)$ one has $B\big(\lambda, \frac{1}{\|R(\lambda,A)\|}\big) \subseteq \rho(A)$, and for $\mu \in B\big(\lambda, \frac{1}{\|R(\lambda,A)\|}\big)$ one has*

$$R(\mu,A) = \sum_{n=0}^{\infty} (\lambda - \mu)^n R(\lambda,A)^{n+1}. \tag{2.2}$$

As a consequence, $\rho(A)$ is an open subset of $\mathbb{K}$, and $R(\cdot,A)\colon \rho(A) \to \mathcal{L}(X)$ is ***analytic*** *(i.e. $R(\cdot,A)$ can be written as a power series about every point of $\rho(A)$).*

2.3 Remarks. (a) For the proof of Theorem 2.2 we recall the **Neumann series**: if $B \in \mathcal{L}(X)$ satisfies $\|B\| < 1$, then $I - B$ is invertible in $\mathcal{L}(X)$, and the inverse is given by $(I-B)^{-1} = \sum_{n=0}^{\infty} B^n$, with absolute convergence of the series.

(b) Let $A \in \mathcal{L}(X)$. Then part (a) implies that $\lambda I - A = \lambda\big(I - \frac{1}{\lambda}A\big)$ is invertible in $\mathcal{L}(X)$ for all $\lambda \in \mathbb{K}$ with $|\lambda| > \|A\|$, with inverse

$$(\lambda I - A)^{-1} = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n.$$

As a consequence, $\{\lambda \in \mathbb{K};\ |\lambda| > \|A\|\} \subseteq \rho(A)$. △

Proof of Theorem 2.2. (a) $AR(\lambda,A)x - \lambda R(\lambda,A)x = -x = R(\lambda,A)Ax - R(\lambda,A)\lambda x$.

(b) Multiplying the identity

$$(\mu I - A) - (\lambda I - A) = (\mu - \lambda)I|_{\mathrm{dom}(A)}$$

from the right by $R(\lambda,A)$ and from the left by $R(\mu,A)$, one obtains the resolvent equation.

(c) Let $\lambda \in \rho(A)$ and $\mu \in B\big(\lambda, \frac{1}{\|R(\lambda,A)\|}\big)$. Then the operator $I - (\lambda-\mu)R(\lambda,A)$ is invertible in $\mathcal{L}(X)$ since $|\lambda-\mu|\|R(\lambda,A)\| < 1$ (Neumann series). Therefore the identity

$$\mu I - A = (\lambda I - A) - (\lambda-\mu)I = \big(I - (\lambda-\mu)R(\lambda,A)\big)(\lambda I - A)$$

shows that the mapping $\mu I - A \colon \mathrm{dom}(A) \to X$ is bijective, with inverse

$$(\mu I - A)^{-1} = R(\lambda,A)\big(I - (\lambda-\mu)R(\lambda,A)\big)^{-1} \in \mathcal{L}(X).$$

Hence $\mu \in \rho(A)$, and the formula (2.2) for the resolvent is a consequence of the Neumann series. It follows that $\rho(A)$ is an open subset of $\mathbb{K}$ and that $R(\cdot,A)$ is analytic. □

2.4 Remarks. (a) Theorem 2.2(c) shows that $\|R(\lambda,A)\| \geqslant \mathrm{dist}(\lambda,\sigma(A))^{-1}$ for all $\lambda \in \rho(A)$. Therefore the norm of the resolvent has to blow up if λ approaches $\sigma(A)$.

(b) As in first year analysis, the analyticity of $R(\cdot,A)$ implies that $R(\cdot,A)$ is infinitely differentiable, and from the power series (2.2) one can read off the derivatives,

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^n R(\lambda,A) = (-1)^n\, n!\, R(\lambda,A)^{n+1} \qquad (\lambda \in \rho(A),\ n \in \mathbb{N}_0).$$ △

2.1.2 Integration of operator-valued functions, and improper integrals

2.5 Proposition. *Let X, Y be Banach spaces, $a, b \in \mathbb{R}$, $a < b$. Let $F \colon [a,b] \to \mathcal{L}(X,Y)$ be strongly continuous. Then the mapping*

$$X \ni x \mapsto \int_a^b F(t)x\,\mathrm{d}t \in Y$$

belongs to $\mathcal{L}(X,Y)$ and has norm less or equal $\int_a^b \|F(t)\|\,\mathrm{d}t$.

Some comments: **strongly continuous** means that $t \mapsto F(t)x$ is continuous for all $x \in X$. (In other words, it means that F is continuous with respect to the **strong operator topology** on $\mathcal{L}(X,Y)$, which is defined as the initial topology with respect to the family of mappings $\big(\mathcal{L}(X,Y) \ni A \mapsto Ax \in Y\big)_{x\in X}$.) It will be part of the proof that the function $\|F(\cdot)\|$ is bounded and measurable.

Proof of Proposition 2.5. (i) First we observe that, in view of the uniform boundedness theorem, the boundedness of $\{\|F(t)x\|;\ t \in [a,b]\}$ for all $x \in X$ implies that $\{\|F(t)\|;\ t \in [a,b]\}$ is bounded.

Next, one easily sees that the set $\{t \in [a,b];\ \|F(t)\| \leqslant \alpha\}$ is closed for all $\alpha \in \mathbb{R}$, and this shows that $\|F(\cdot)\|$ is measurable. (For completeness, we mention that the previous property is called **lower semi-continuity** of $\|F(\cdot)\|$.)

(ii) The linearity of the mapping $x \mapsto \int_a^b F(t)x\,\mathrm{d}t$ is obvious. For $x \in X$ we estimate

$$\left\| \int_a^b F(t)x\,\mathrm{d}t \right\| \leqslant \int_a^b \|F(t)x\|\,\mathrm{d}t \leqslant \int_a^b \|F(t)\|\,\mathrm{d}t\,\|x\|,$$

and this establishes the asserted norm estimate. □

Abbreviating, we will write $\int_a^b F(t)\,\mathrm{d}t$ for the mapping defined in Proposition 2.5. This integral is called the **strong integral**; one has to keep in mind that, in general, it is not the integral of an $\mathcal{L}(X,Y)$-valued function as treated in Subsection 1.3.2.

We will also need 'improper integrals' of continuous Banach space valued functions. For simplicity we restrict our attention to integrals over $[0,\infty)$ (because this is what will be needed next).

2.6 Proposition. *Let X be a Banach space, let $f\colon [0,\infty) \to X$ be continuous, and assume that the function $[0,\infty) \ni t \mapsto \|f(t)\|$ is integrable. Then*

$$\int_0^\infty f(t)\,\mathrm{d}t := \lim_{a\to\infty} \int_0^a f(t)\,\mathrm{d}t$$

exists.

We omit the (easy) proof of this proposition and mention that Proposition 2.5 has its analogue for improper integrals.

2.2 Characterisation of generators of C_0-semigroups

In this section let X be a Banach space.

2.7 Theorem. *Let T be a C_0-semigroup on X, and let A be its generator. Let $M \geqslant 1$, $\omega \in \mathbb{R}$ be such that*

$$\|T(t)\| \leqslant M\mathrm{e}^{\omega t} \qquad (t \geqslant 0)$$

(cf. Proposition 1.4(a)).

Then $\{\lambda \in \mathbb{K};\ \operatorname{Re}\lambda > \omega\} \subseteq \rho(A)$, and for all $\lambda \in \mathbb{K}$ with $\operatorname{Re}\lambda > \omega$ one has

$$R(\lambda, A) = \int_0^\infty \mathrm{e}^{-\lambda t} T(t)\,\mathrm{d}t \quad (\textit{strong improper integral; see Subsection 2.1.2}),$$

$$\|R(\lambda, A)^n\| \leqslant \frac{M}{(\operatorname{Re}\lambda - \omega)^n} \qquad (n \in \mathbb{N}). \tag{2.3}$$

The formula for the resolvent in this theorem means that $\lambda \mapsto R(\lambda, A)$ is the Laplace transform of the semigroup; see [ABHN11; Chapter 2].

In the proof we will use the concept of rescaling. If T is a C_0-semigroup on X with generator A, $\lambda \in \mathbb{K}$, and one defines

$$T_\lambda(t) := \mathrm{e}^{-\lambda t} T(t) \qquad (t \geqslant 0),$$

then T_λ is also a C_0-semigroup, called a **rescaled semigroup**, and the generator of T_λ is given by $A - \lambda I$; see Exercise 2.1.

Proof of Theorem 2.7. Let $\lambda \in \mathbb{K}$, $\operatorname{Re}\lambda > \omega$. Observe that the rescaled semigroup T_λ satisfies the estimate

$$\|T_\lambda(t)\| \leqslant M\mathrm{e}^{(\omega-\operatorname{Re}\lambda)t} \qquad (t \geqslant 0)$$

and that the resolvent of A at λ corresponds to the resolvent of $A - \lambda I$ at 0. Thus for the proof of $\lambda \in \rho(A)$ and of the formula for the resolvent we can assume without loss of generality that $\omega < 0$ and $\lambda = 0$.

The estimate $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ $(t \geqslant 0)$ implies that the strong improper integral

$$R := \int_0^\infty T(t)\,\mathrm{d}t$$

defines an operator $R \in \mathcal{L}(X)$. On the one hand, for all $x \in \operatorname{dom}(A)$ we have

$$RAx = \int_0^\infty T(t)Ax\,\mathrm{d}t = \lim_{a\to\infty} \int_0^a \frac{\mathrm{d}}{\mathrm{d}t}T(t)x\,\mathrm{d}t = \lim_{a\to\infty} (T(a)x - x) = -x.$$

On the other hand, if $x \in X$, then by Theorem 1.11(b) we obtain $A\int_0^a T(t)x\,\mathrm{d}t = T(a)x - x \to -x$ as $a \to \infty$. Since $\int_0^a T(t)x\,\mathrm{d}t \to Rx$ as $a \to \infty$ and A is closed, it follows that $Rx \in \operatorname{dom}(A)$, $ARx = -x$. Thus we have proved the two equalities $RA = -I|_{\operatorname{dom}(A)}$ and $AR = -I$, which imply that $0 \in \rho(A)$ and $R(0, A) = (-A)^{-1} = R$.

For the powers of $R(\lambda, A)$ we now obtain (recall Remark 2.4(b))

$$\begin{aligned} R(\lambda, A)^n &= (-1)^{n-1}\frac{1}{(n-1)!}\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^{n-1} \int_0^\infty \mathrm{e}^{-\lambda t}T(t)\,\mathrm{d}t \\ &= \frac{1}{(n-1)!}\int_0^\infty t^{n-1}\mathrm{e}^{-\lambda t}T(t)\,\mathrm{d}t. \end{aligned} \tag{2.4}$$

(The last equality is obtained by differentiation under the integral; we delegate the details to Exercise 2.2. See also the subsequent Remark 2.8.) By Proposition 2.5 we conclude that

$$\begin{aligned} \|R(\lambda, A)^n\| &\leqslant \frac{1}{(n-1)!}\,M\int_0^\infty t^{n-1}\mathrm{e}^{(\omega-\operatorname{Re}\lambda)t}\,\mathrm{d}t \\ &= \frac{1}{(n-1)!}\,M\Big(\frac{\mathrm{d}}{\mathrm{d}\omega}\Big)^{n-1}\int_0^\infty \mathrm{e}^{(\omega-\operatorname{Re}\lambda)t}\,\mathrm{d}t \\ &= \frac{1}{(n-1)!}\,M\Big(\frac{\mathrm{d}}{\mathrm{d}\omega}\Big)^{n-1}\frac{1}{\operatorname{Re}\lambda-\omega} = \frac{M}{(\operatorname{Re}\lambda-\omega)^n}\,. \end{aligned}$$

□

2.8 Remark. We will mainly be interested in C_0-semigroups satisfying the estimate of Proposition 1.4(a) with $M = 1$, in which case the semigroup is called **quasi-contractive**. For such C_0-semigroups it is sufficient to prove the resolvent estimate (2.3) in Theorem 2.7 for $n = 1$ (because then taking powers one obtains the estimate for all $n \in \mathbb{N}$). Note that for $n = 1$ the second equality in (2.4) is trivial.

The resolvent estimate (2.3) can also be proved by a reduction to the case of a **contractive** C_0-semigroup T, i.e. $\|T(t)\| \leqslant 1$ for all $t \geqslant 0$; cf. Exercise 2.3. △

Next we show that the necessary conditions for the generator we have derived so far are also sufficient. We restrict ourselves to the quasi-contractive case and delegate the proof of the general case to Exercise 2.5.

2.9 Theorem (Hille–Yosida, quasi-contractive case). *Let A be a closed, densely defined operator in X. Assume that there exists $\omega \in \mathbb{R}$ such that $(\omega, \infty) \subseteq \rho(A)$ and*

$$\|R(\lambda, A)\| \leqslant \frac{1}{\lambda - \omega} \qquad (\lambda \in (\omega, \infty)).$$

Then A is the generator of a C_0-semigroup T satisfying the estimate

$$\|T(t)\| \leqslant \mathrm{e}^{\omega t} \qquad (t \geqslant 0).$$

As a preliminary remark we note that it is sufficient to treat the case $\omega = 0$. Indeed, $\tilde{A} := A - \omega I$ satisfies the conditions of Theorem 2.9 with $\omega = 0$. Having obtained the contractive C_0-semigroup $\widetilde{T}$ with generator $\tilde{A}$ one obtains the C_0-semigroup generated by $A = \tilde{A} + \omega I$ as the rescaled semigroup $\widetilde{T}_{-\omega}$.

We now define the **Yosida approximations**

$$A_n := A\Big(I - \frac{1}{n}A\Big)^{-1} = nAR(n, A) = n^2 R(n, A) - nI \in \mathcal{L}(X) \qquad (n \in \mathbb{N})$$

of A. The proof of Theorem 2.9 will consist of three steps:

In the first step we show that the operators A_n generate contractive semigroups.

In the second step we show that these semigroups converge strongly to a C_0-semigroup.

In the third step we show that A is the generator of the limiting semigroup.

Before proceeding with the proof we explain why the operators A_n can be considered as approximations of A; note that the operator occurring in part (b) of the following lemma equals A_n for $\lambda = n \in \mathbb{N}$.

2.10 Lemma. *Let A be a closed, densely defined operator in X. Assume that there exists $\lambda_0 \geqslant 0$ such that $(\lambda_0, \infty) \subseteq \rho(A)$ and $M := \sup_{\lambda > \lambda_0} \|\lambda R(\lambda, A)\| < \infty$. Then:*

(a) *$\lambda R(\lambda, A)x \to x$ ($\lambda \to \infty$) for all $x \in X$.*

(b) *$A\big(\lambda R(\lambda, A)\big)x \to Ax$ ($\lambda \to \infty$) for all $x \in \mathrm{dom}(A)$.*

Proof. (a) If $x \in \mathrm{dom}(A)$, then

$$\lambda R(\lambda, A)x = (\lambda I - A + A)R(\lambda, A)x = x + R(\lambda, A)Ax \to x \qquad (\lambda \to \infty).$$

As $\|\lambda R(\lambda, A)\| \leqslant M$ for all $\lambda > \lambda_0$ and $\mathrm{dom}(A)$ is dense in X, the convergence carries over to all $x \in X$, by Proposition 1.6.

(b) For $x \in \mathrm{dom}(A)$ the convergence proved in part (a) implies

$$A\big(\lambda R(\lambda, A)\big)x = \lambda R(\lambda, A)Ax \to Ax \qquad (\lambda \to \infty). \qquad \square$$

If X is a reflexive Banach space, then the denseness hypothesis for $\mathrm{dom}(A)$ in Lemma 2.10 can be omitted because it follows from the remaining hypotheses; see Exercise 2.7.

Proof of Theorem 2.9. Recall that, without loss of generality, we only treat the case $\omega = 0$.

(i) For $n \in \mathbb{N}$, $t \geqslant 0$ we obtain the estimate

$$\left\|\mathrm{e}^{tA_n}\right\| = \left\|\mathrm{e}^{tn^2 R(n,A)}\mathrm{e}^{-tnI}\right\| \leqslant \mathrm{e}^{tn^2\|R(n,A)\|}\mathrm{e}^{-tn} \leqslant 1.$$

For this computation recall Exercise 1.1(a), and note that $\|R(n,A)\| \leqslant \frac{1}{n}$ because $\omega = 0$.

(ii) For $x \in X$, $t > 0$ and $m, n \in \mathbb{N}$ we compute

$$\begin{aligned}
\mathrm{e}^{tA_m}x - \mathrm{e}^{tA_n}x &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\big(\mathrm{e}^{(t-s)A_n}\mathrm{e}^{sA_m}x\big)\,\mathrm{d}s = \int_0^t \mathrm{e}^{(t-s)A_n}(A_m - A_n)\mathrm{e}^{sA_m}x\,\mathrm{d}s \\
&= \int_0^t \mathrm{e}^{(t-s)A_n}\mathrm{e}^{sA_m}(A_m - A_n)x\,\mathrm{d}s,
\end{aligned}$$

where in the last equality we have used the fact that A_m, A_n as well as the generated semigroups commute. Thus by step (i) we obtain the estimate

$$\left\|\mathrm{e}^{tA_m}x - \mathrm{e}^{tA_n}x\right\| \leqslant t\left\|(A_m - A_n)x\right\|. \tag{2.5}$$

Let $a > 0$. For $n \in \mathbb{N}$ we define the operator $\mathcal{T}_n^a \colon X \to C([0,a];X)$ by

$$\mathcal{T}_n^a x := \big[t \mapsto \mathrm{e}^{tA_n}x\big] \qquad (x \in X)$$

(where $C([0,a];X)$ denotes the Banach space of continuous X-valued functions, provided with the supremum norm). Then step (i) shows that $\mathcal{T}_n^a$ is a contraction, and inequality (2.5) shows that

$$\left\|\mathcal{T}_m^a x - \mathcal{T}_n^a x\right\| \leqslant a\|A_m x - A_n x\| \qquad (x \in X,\ m,n \in \mathbb{N}).$$

For $x \in \mathrm{dom}(A)$ this implies that $(\mathcal{T}_n^a x)_{n\in\mathbb{N}}$ is a Cauchy sequence, because $(A_n x)_{n\in\mathbb{N}}$ is convergent (to Ax), by Lemma 2.10(b). Applying Proposition 1.6 we conclude that there exists $\mathcal{T}^a \in \mathcal{L}\big(X, C([0,a];X)\big)$ such that $\mathcal{T}_n^a \to \mathcal{T}^a$ strongly as $n \to \infty$.

Clearly, if $0 < a < b$, then $\mathcal{T}^b x|_{[0,a]} = \mathcal{T}^a x$ for all $x \in X$, and therefore we can define $T \colon [0,\infty) \to \mathcal{L}(X)$ by

$$T(t)x := \mathcal{T}^a x(t) \qquad (0 \leqslant t \leqslant a,\ x \in X).$$

From $T(\cdot)x|_{[0,a]} = \mathcal{T}^a x$ $(a > 0,\ x \in X)$ we infer that T is strongly continuous. Clearly $T(0) = I$. Taking the strong limit $n \to \infty$ in $\mathrm{e}^{(t+s)A_n} = \mathrm{e}^{tA_n}\mathrm{e}^{sA_n}$ $(t, s \geqslant 0)$ we see that T is a one-parameter semigroup (recall Lemma 1.12). Altogether, T is a contractive C_0-semigroup.

(iii) Let B be the generator of T. Let $x \in \mathrm{dom}(A)$. Using the notation of step (ii), with $a := 1$, we see that $\mathcal{T}_n^1 x \to \mathcal{T}^1 x$ and $(\mathcal{T}_n^1 x)' = \mathcal{T}_n^1 A_n x \to \mathcal{T}^1 Ax$ in $C([0,1];X)$ as $n \to \infty$ (recall Lemma 1.12). This implies that $T(\cdot)x|_{[0,1]} = \mathcal{T}^1 x$ is differentiable with continuous derivative $\mathcal{T}^1 Ax$, and therefore $x \in \mathrm{dom}(B)$ and $Bx = Ax$.

So far we have shown that $A \subseteq B$. We also know that $(0,\infty) \subseteq \rho(B)$, by Theorem 2.7, and that $(0,\infty) \subseteq \rho(A)$, by hypothesis. Now from $I - A \subseteq I - B$, the injectivity of $I - B$ and $\mathrm{ran}(I - A) = X$ we obtain $I - A = I - B$, because a surjective mapping cannot have a proper injective extension, and hence $A = B$. $\square$

2.11 Remarks. (a) The proof of the Hille–Yosida theorem for the general case can be given along the same lines; see Exercise 2.5.

(b) As an interesting feature in the proof of Theorem 2.9 we point out that the approximating semigroups are continuous with respect to the operator norm, whereas in general the limiting semigroup is only strongly continuous. The norm continuity is lost because the approximating semigroups only converge strongly (and generally not with respect to the operator norm). △

2.3 Euler's exponential formula

Given $a \in \mathbb{K}$, there are two well-known ways of approximating e^{ta} besides the exponential series, namely

$$\mathrm{e}^{ta} = \lim_{n\to\infty} \Bigl(1 + \frac{t}{n} a\Bigr)^n \qquad \text{and} \qquad \mathrm{e}^{ta} = \lim_{n\to\infty} \Bigl(1 - \frac{t}{n} a\Bigr)^{-n}.$$

The attempt to replace a by an unbounded generator in the first formula leads to hopeless problems with the domains of the powers of the operators involved, whereas the second formula looks more promising because the occurring inverses are just those whose existence is guaranteed by Theorem 2.7. (In fact, the resulting formulas are those known in numerical analysis as 'backward Euler method'.)

We now show that this approximation idea works for arbitrary C_0-semigroups. The proof is independent of the Hille–Yosida theorem (which was proved by a different kind of approximation).

2.12 Theorem (Euler's exponential formula). *Let T be a C_0-semigroup on a Banach space X, with generator A. Then*

$$T(t)x = \lim_{n\to\infty} \Bigl(I - \frac{t}{n} A\Bigr)^{-n} x$$

for all $x \in X$, with uniform convergence for t in compact subsets of $[0,\infty)$.

2.13 Remarks. (a) Let $M \geqslant 1$ and $\omega \in \mathbb{R}$ be such that $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ for all $t \geqslant 0$. Let $r > 0$. If $\frac{1}{r} > \omega$, then Theorem 2.7 implies $\frac{1}{r} \in \rho(A)$,

$$(I - rA)^{-1} = \frac{1}{r}\Bigl(\frac{1}{r}I - A\Bigr)^{-1} \in \mathcal{L}(X),$$

$$\bigl\|(I - rA)^{-n}\bigr\| \leqslant \Bigl(\frac{1}{r}\Bigr)^n M\Bigl(\frac{1}{r} - \omega\Bigr)^{-n} = M(1 - r\omega)^{-n} \qquad (n \in \mathbb{N}).$$

Now let $a > 0$. Then by the above, the operator $(I - \frac{t}{n}A)^{-n}$ is defined for all $t \in [0,a]$ if $n > a\omega$, and $\|(I - \frac{t}{n}A)^{-n}\| \leqslant M(1 - \frac{t}{n}\omega)^{-n}$. (In fact, if $\omega \leqslant 0$, then these properties hold for all $t \geqslant 0$ and $n \in \mathbb{N}$.)

(b) We note that the expressions $(I - rA)^{-1} = \frac{1}{r}\bigl(\frac{1}{r}I - A\bigr)^{-1}$ for small $r > 0$ correspond to expressions $\lambda(\lambda I - A)^{-1}$ for large $\lambda > 0$. By Lemma 2.10(a) it follows that $(I - rA)^{-1} \to I$ strongly as $r \to 0$. △

Proof of Theorem 2.12. Let $M \geqslant 1$ and $\omega \geqslant 0$ be such that $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ for all $t \geqslant 0$. Let $r \in (0, \frac{1}{\omega})$ (where $\frac{1}{\omega}$ should be read as ∞ if $\omega = 0$). For $s \in (0, \frac{1}{\omega}) \setminus \{r\}$ we compute

$$\begin{aligned}\frac{1}{s-r}\big((I-sA)^{-1}-(I-rA)^{-1}\big) &= \frac{1}{s-r}(I-sA)^{-1}\big((I-rA)-(I-sA)\big)(I-rA)^{-1}\\ &= (I-sA)^{-1}A(I-rA)^{-1} \to A(I-rA)^{-2}\end{aligned}$$

as $s \to r$. This shows that $r \mapsto (I-rA)^{-1}$ is continuously differentiable on $(0, \frac{1}{\omega})$ and that $\frac{\mathrm{d}}{\mathrm{d}r}(I-rA)^{-1} = A(I-rA)^{-2}$.

Now fix $a > 0$ and let $n > a\omega$, as in Remark 2.13(a). If $0 < s \leqslant a$, then $0 < \frac{s}{n} < \frac{1}{\omega}$, and applying the product and chain rules we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(I-\frac{s}{n}A\Big)^{-n} = n\bigg(\Big(I-\frac{s}{n}A\Big)^{-1}\bigg)^{n-1}\frac{1}{n}A\Big(I-\frac{s}{n}A\Big)^{-2} = A\Big(I-\frac{s}{n}A\Big)^{-n-1}.$$

Let $t \in (0,a]$, $x \in \operatorname{dom}(A)$. Then the function $[0,t] \ni s \mapsto T(t-s)(I-\frac{s}{n}A)^{-n}x$ is continuous as well as continuously differentiable on $(0,t]$, with derivative

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\bigg(T(t-s)\Big(I-\frac{s}{n}A\Big)^{-n}x\bigg) &= T(t-s)\bigg(-A+A\Big(I-\frac{s}{n}A\Big)^{-1}\bigg)\Big(I-\frac{s}{n}A\Big)^{-n}x\\ &= T(t-s)\Big(I-\frac{s}{n}A\Big)^{-n}\bigg(\Big(I-\frac{s}{n}A\Big)^{-1}-I\bigg)Ax;\end{aligned}$$

see Exercise 1.8. By the fundamental theorem of calculus (Theorem 1.9) it follows that

$$\begin{aligned}&\big\|(I-\tfrac{t}{n}A)^{-n}x - T(t)x\big\|\\ &\quad= \Big\|\int_0^t T(t-s)(I-\tfrac{s}{n}A)^{-n}\big((I-\tfrac{s}{n}A)^{-1}-I\big)Ax\,\mathrm{d}s\Big\| \qquad (2.6)\\ &\quad\leqslant a\sup_{0\leqslant s\leqslant t\leqslant a}\big\|T(t-s)(I-\tfrac{s}{n}A)^{-n}\big\|\sup_{0\leqslant s\leqslant a}\big\|\big((I-\tfrac{s}{n}A)^{-1}-I\big)Ax\big\|.\end{aligned}$$

Using Remark 2.13(a) we estimate

$$\begin{aligned}\big\|T(t-s)(I-\tfrac{s}{n}A)^{-n}\big\| &\leqslant M\mathrm{e}^{\omega(t-s)}\cdot M(1-\tfrac{s}{n}\omega)^{-n}\\ &\leqslant M^2\mathrm{e}^{\omega a}\sup_{n>a\omega}(1-\tfrac{a}{n}\omega)^{-n} =: M_0 < \infty,\end{aligned}$$

where the supremum is finite because the sequence $\big((1-\frac{a}{n}\omega)^{-n}\big)_{n>a\omega}$ is convergent (to $\mathrm{e}^{a\omega}$). Hence (2.6) yields

$$\sup_{0\leqslant t\leqslant a}\big\|(I-\tfrac{t}{n}A)^{-n}x - T(t)x\big\| \leqslant aM_0\sup_{0\leqslant s\leqslant a}\big\|\big((I-\tfrac{s}{n}A)^{-1}-I\big)Ax\big\|,$$

which tends to 0 as $n \to \infty$ (recall Remark 2.13(b)).

Now the proof is completed pretty much as in step (ii) of the proof of Theorem 2.9. For $n > a\omega$ we define $\mathcal{T}_n^a\colon X \to C([0,a];X)$,

$$\mathcal{T}_n^a x := \big[t \mapsto (I-\tfrac{t}{n}A)^{-n}x\big] \qquad (x \in X).$$

Then $\|\mathcal{T}_n^a\| \leqslant M_0$ for all $n > a\omega$, and $\mathcal{T}_n^a x \to T(\cdot)x|_{[0,a]}$ $(n \to \infty)$ for all $x \in \operatorname{dom}(A)$, as shown above. Using the fact that $\operatorname{dom}(A)$ is dense and applying Proposition 1.6 we obtain $\mathcal{T}_n^a x \to T(\cdot)x|_{[0,a]}$ $(n \to \infty)$ for all $x \in X$. □

2.14 Remark. Euler's exponential formula will be used repeatedly later in the book, for instance in the proof of the invariance of a closed set $C \subseteq X$ under a C_0-semigroup, as follows. If C is invariant under $(I - rA)^{-1}$ for small $r > 0$, then Theorem 2.12 implies that C is invariant under T as well; see also Proposition 9.1.

The exponential formula can also be derived from the Chernoff product formula, for which we refer to Section 13.3; in particular see Example 13.17(b). △

Notes

The Hille–Yosida theorem is basic for the theory of C_0-semigroups and can be found (with varying proofs) in any treatment of C_0-semigroups. Our proof is Yosida's original proof in [Yos48]. A different proof, based on the exponential formula, can be found in §§ 1.2, 1.3, 1.4 of the beautiful Chapter IX in [Kat80], which is a concise introduction to semigroups.

The discovery of the Hille–Yosida theorem in 1948 was a major event in functional analysis and operator theory. Previously known was Stone's theorem from 1932 [Sto32] (announced in [Sto30]): an operator A in a complex Hilbert space generates a one-parameter group of unitary operators if and only if $\mathrm{i}A$ is self-adjoint. Stone was motivated by quantum mechanics, for which John von Neumann developed the mathematical foundations in the years between 1927 and 1932. Von Neumann's work on this subject culminated in his monograph "Mathematische Grundlagen der Quantenmechanik" [Neu32a], which was the first rigorous mathematical treatment of quantum mechanics focusing on unbounded self-adjoint operators. The theory von Neumann developed is still used today whith great success.

Unitary groups describe the dynamics in quantum mechanics. More generally, one-parameter groups describe evolutionary processes that are reversible and in which the states exist for all times. It is natural to consider time evolution for positive times only; this leads to semigroups and the problem of describing their generators. E. Hille and K. Yosida found the solution independently, for the case of contractive semigroups: Yosida published his proof in the first volume of the Journal of the Mathematical Society of Japan [Yos48], where he introduced the Yosida approximations A_n of the unbounded operator A and constructed the semigroup as the limit of e^{tA_n}, as we did in this chapter. Hille, in his book [Hil48; Sections 12.2 and 12.3], employed the inverse Laplace transform and the exponential formula (see Theorem 2.12) to prove the theorem for a slightly more general than the contractive case.

The general case of the Hille–Yosida theorem was proved independently by Miyadera [Miy52; Theorem 2], Phillips [Phi53; Section 2] and Feller [Fel53; Theorem 3]. (The general version of the theorem is sometimes called "Hille–Yosida–Phillips theorem", maybe because Phillips' paper was submitted first.) Hille's book mentioned above has been extended in collaboration with Phillips to [HiPh57]. In this monograph the authors offer two proofs of the general Hille–Yosida theorem: the first is an adaptation of Yosida's proof to the general case, the second relies on the exponential formula; see [HiPh57; Section 12.3].

The proof via the exponential formula can be considered as an application of the backward Euler scheme to the numerical solution of the Cauchy problem

$$u'(t) = Au(t) \quad (t \geqslant 0), \qquad u(0) = x.$$

Indeed, fixing $n \in \mathbb{N}$ and subdividing the interval $[0, t]$ into n equal parts, one can discretise the differential equation $u' = Au$ to derive the difference equation

$$\frac{u_k - u_{k-1}}{t/n} = Au_k \qquad (k = 1, \dots, n),$$

where u_k plays the role of $u(\frac{k}{n}t)$. (For comparison, the *forward* Euler scheme would involve the term Au_{k-1} on the right-hand side instead of Au_k.) Starting from the initial value $u_0 = x$, one easily shows that for large n the solution is given by $u_k = (I - \frac{t}{n}A)^{-k}x$ $(k \in \{0, \dots, n\})$; in particular, for $k = n$ one obtains the approximation $(I - \frac{t}{n}A)^{-n}x$ for $u(t)$ from the exponential formula. We mention that the backward Euler method even works in the theory of nonlinear semigroups, where the central generation theorem is due to Crandall and Liggett; see [CHA&87; Theorem 2.3].

Our entire book demonstrates the power of the Hille–Yosida theorem. We have restricted our proof to the special case of quasi-contractive semigroups; this will be sufficient for our application of the theorem to operators in Hilbert spaces associated with forms, in Section 5.3.

Exercises

2.1 Let T be a C_0-semigroup on a Banach space X, with generator A. Let $\lambda \in \mathbb{K}$. Show that

$$T_\lambda(t) := \mathrm{e}^{-\lambda t}T(t) \qquad (t \geqslant 0)$$

defines a C_0-semigroup (the rescaled semigroup) and that the generator of T_λ is given by $A - \lambda I$.

2.2 Prove the second equality in (2.4).

2.3 Let T be a bounded C_0-semigroup on a Banach space X, with generator A, and let $M := \sup_{t\geqslant 0} \|T(t)\|$.

(a) Show that

$$|||x||| := \sup_{t\geqslant 0} \|T(t)x\| \qquad (x \in X)$$

defines a norm $|||\cdot|||$ on X which is equivalent to $\|\cdot\|$, and that T is a contractive C_0-semigroup on $(X, |||\cdot|||)$.

(b) For any $\alpha_1, \dots, \alpha_n > 0$, show that

$$\big\|(I - \alpha_1 A)^{-1} \cdots (I - \alpha_n A)^{-1}\big\| \leqslant M.$$

2.4 Let T be a C_0-semigroup on a Banach space X, with generator A. Let $M \geqslant 1$, $\omega \in \mathbb{R}$ be such that $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ for all $t \geqslant 0$. For $x \in X$ and $\lambda \in \mathbb{K}$ with $\operatorname{Re}\lambda > \omega$ show that

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

Use this identity to give another proof of the resolvent estimate (2.3).

2.5 Prove the Hille–Yosida theorem for the general case:

Let A be a closed, densely defined operator in a Banach space X. Assume that there exist $M \geqslant 1$ and $\omega \in \mathbb{R}$ such that $(\omega, \infty) \subseteq \rho(A)$ and

$$\|R(\lambda, A)^n\| \leqslant \frac{M}{(\lambda - \omega)^n} \qquad (\lambda \in (\omega, \infty),\ n \in \mathbb{N}).$$

Then A is the generator of a C_0-semigroup T satisfying the estimate

$$\|T(t)\| \leqslant M\mathrm{e}^{\omega t} \qquad (t \geqslant 0).$$

(Hint: Proceed as in the proof of Theorem 2.9, with adapted estimates.)

2.6 Let T be a C_0-semigroup on a Banach space X. For $h > 0$ put $A_h := h^{-1}(T(h) - I)$. Show that $\mathrm{e}^{tA_h}x \to T(t)x$ $(h \to 0)$ for all $x \in X$, uniformly for t in compact subsets of $[0, \infty)$. (Hint: Use a procedure similar to the proof of the exponential formula, Theorem 2.12.)

2.7 Let X be a reflexive Banach space, let A be a closed operator in X, and assume that there exists $\lambda_0 \geqslant 0$ such that $(\lambda_0, \infty) \subseteq \rho(A)$, $M := \sup_{\lambda > \lambda_0} \|\lambda R(\lambda, A)\| < \infty$.

Show that $\operatorname{dom}(A)$ is dense in X.

Hints: 1. Show that $\lambda R(\lambda, A)x \to x$ $(\lambda \to \infty)$ for all $x \in \operatorname{dom}(A)$; cf. Lemma 2.10(a). 2. Now let $x \in X$. Show that there exists a sequence (λ_n) in (λ_0, ∞), $\lambda_n \to \infty$ $(n \to \infty)$ such that $(\lambda_n R(\lambda_n, A)x)$ converges weakly to some $y \in X$. (See e.g. [Yos68; Sect. V.2, Theorem 1] or [Lax02; Sect. 10.2, Theorem 7] for the fact that every bounded sequence in a reflexive Banach space contains a weakly convergent subsequence.) 3. Choose $\mu \in \rho(A)$. Show that $\lambda_n R(\lambda_n, A)R(\mu, A)x \to R(\mu, A)x$, but also $R(\mu, A)\lambda_n R(\lambda_n, A)x \to R(\mu, A)y$ weakly.

Chapter 3

Holomorphic semigroups

The objective of this chapter is to introduce semigroups for which the 'time parameter' t can also be chosen in a complex neighbourhood of the positive real axis. The foundations of these 'holomorphic semigroups' will be given in Section 3.2. Then we characterise which operators generate contractive holomorphic C_0-semigroups. Finally, in Section 3.4 we treat the special case of holomorphic semigroups on Hilbert spaces. We start with an interlude on Banach space valued holomorphy.

3.1 Interlude: vector-valued holomorphic functions

In this section let X, Y be complex Banach spaces. The first issue is to show that for Banach space valued functions several notions of holomorphy coincide.

Let $\Omega \subseteq \mathbb{C}$ be an open set, $f\colon \Omega \to X$. The function f is called **holomorphic** if f is (complex) differentiable at each point of Ω, and f is **weakly holomorphic** if $x' \circ f$ is holomorphic for all $x' \in X'$ ($= \mathcal{L}(X, \mathbb{C})$, the dual space of X). Recall that f is **analytic** if f can be represented as a power series in a neighbourhood of each point of Ω.

3.1 Remarks. (a) It is evident that holomorphy of a function implies weak holomorphy.

(b) Using part (a) one can prove the identity theorem for a holomorphic function $f\colon \Omega \to X$, where $\Omega \subseteq \mathbb{C}$ is open and connected: if $[f = 0]$ has a cluster point in Ω, then $f = 0$. (We use the notation $[f = 0] := \{z \in \Omega;\ f(z) = 0\}$; sets like $[f \geqslant 0]$, $[f > 0]$ etc., occurring later in the book, are defined analogously.)

Indeed, for each $x' \in X'$ the zeros of the function $x' \circ f$ have a cluster point in Ω, and therefore $x' \circ f = 0$, by the identity theorem for $\mathbb{C}$-valued holomorphic functions. From $x' \circ f = 0$ for all $x' \in X'$ one obtains $f = 0$.

The reasoning presented here is an example of how properties of complex-valued holomorphic functions may be transferred to vector-valued functions. An alternative approach is to assure oneself that the proofs in classical complex analysis also work out in the vector-valued case; cf. part (c).

(c) Let $\Omega \subseteq \mathbb{C}$ be an open set, and let $f\colon \Omega \to X$ be holomorphic. The following facts can be proved in the same way as in the case of $\mathbb{C}$-valued functions.

If Ω is convex, and γ is a piecewise continuously differentiable closed path in Ω, then $\int_\gamma f(z)\,\mathrm{d}z = 0$ (Cauchy's integral theorem). Note that the path integral is defined by a parametrisation of the path and therefore reduces to integrals over intervals. As a consequence, path integrals fit into the context explained in Subsection 1.3.2.

If $z_0 \in \Omega$, $r > 0$ are such that $B[z_0, r] \subseteq \Omega$, then f satisfies Cauchy's integral formula

$$f(z) = \frac{1}{2\pi \mathrm{i}} \int_{|\zeta - z_0| = r} \frac{f(\zeta)}{\zeta - z} \, \mathrm{d}\zeta \qquad (z \in B(z_0, r)).$$

The function f is analytic, and one has Cauchy's integral formulas for the derivatives,

$$f^{(n)}(z) = \frac{n!}{2\pi \mathrm{i}} \int_{|\zeta - z_0| = r} \frac{f(\zeta)}{(\zeta - z)^{n+1}} \, \mathrm{d}\zeta \qquad (z \in B(z_0, r), \ n \in \mathbb{N}),$$

with z_0 and r as before. △

A set $E \subseteq X'$ is **separating** (**for** X) if for all $0 \neq x \in X$ there exists $x' \in E$ such that $x'(x) \neq 0$. The set E is called **almost norming** (**for** X) if

$$\|x\|_E := \sup\bigl\{|x'(x)| \,;\, x' \in E, \ \|x'\| \leqslant 1\bigr\} \qquad (x \in X)$$

defines a norm that is equivalent to the norm on X; it is called **norming** if $\|\cdot\|_E = \|\cdot\|$ on X. It is a consequence of the Hahn–Banach theorem that $E = X'$ is norming for X.

3.2 Theorem. *Let $\Omega \subseteq \mathbb{C}$ be an open set, $f \colon \Omega \to X$. Then the following properties are equivalent.*

(i) *f is holomorphic.*

(ii) *f is weakly holomorphic.*

(iii) *There exists an almost norming closed subspace $E \subseteq X'$ such that $x' \circ f$ is holomorphic for all $x' \in E$.*

(iv) *f is locally bounded, and there exists an almost norming set $E \subseteq X'$ such that $x' \circ f$ is holomorphic for all $x' \in E$.*

(v) *f is continuous, and there exists a separating set $E \subseteq X'$ such that $x' \circ f$ is holomorphic for all $x' \in E$.*

Proof. (i) ⇒ (ii) is clear (and was already noted above).

(ii) ⇒ (iii) is clear, with $E = X'$.

(iii) ⇒ (iv). Note that the mapping $\kappa \colon X \to \mathcal{L}(E, \mathbb{C}) = E'$, $\kappa x := [E \ni x' \mapsto x'(x)]$ satisfies $\|\kappa x\| = \|x\|_E$ for all $x \in X$, by the definition of $\|\cdot\|_E$. Since E is almost norming for X, it thus suffices to show that $\kappa \circ f$ is locally bounded. For each $x' \in E$ the function $z \mapsto (\kappa \circ f(z))x' = x'(f(z))$ is holomorphic and hence locally bounded. Therefore the uniform boundedness theorem implies the assertion.

(iv) ⇒ (v). Since almost norming subsets are separating we only have to show that f is continuous.

Let $z_0 \in \Omega$, $r > 0$ be such that $B[z_0, r] \subseteq \Omega$; then $M := \sup\bigl\{\|f(\zeta)\| \,;\, |\zeta - z_0| = r\bigr\} < \infty$. For $x' \in E$, $z \in B[z_0, r/2]$ we then obtain, using Cauchy's integral formula for the derivative,

$$\left|\frac{\mathrm{d}}{\mathrm{d}z} x'(f(z))\right| = \left|\frac{1}{2\pi \mathrm{i}} \int_{|\zeta - z_0| = r} \frac{x'(f(\zeta))}{(\zeta - z)^2} \, \mathrm{d}\zeta\right| \leqslant \frac{1}{2\pi} 2\pi r \|x'\| \frac{M}{(r/2)^2} = \frac{4M}{r} \|x'\|.$$

For $z', z'' \in B[z_0, r/2]$ this implies

$$|x'(f(z') - f(z''))| \leqslant \frac{4M}{r}\|x'\||z' - z''| \qquad (x' \in E),$$

and therefore

$$\|f(z') - f(z'')\|_E \leqslant \frac{4M}{r}|z' - z''|.$$

Since the norm $\|\cdot\|_E$ is equivalent to the norm on X, it follows that f is continuous on $B[z_0, r/2]$.

(v) $\Rightarrow$ (i). Let $z_0 \in \Omega$. We show that f can be expanded into a power series about z_0.

Without loss of generality we assume that $z_0 = 0$. There exists $r > 0$ such that $B[0, r] \subseteq \Omega$. For $n \in \mathbb{N}_0$ we put

$$a_n := \frac{1}{2\pi \mathrm{i}} \int_{|\zeta|=r} \frac{f(\zeta)}{\zeta^{n+1}} \,\mathrm{d}\zeta.$$

With $M := \sup\{\|f(\zeta)\|;\ |\zeta| = r\}$ $(< \infty)$ we obtain $\|a_n\| \leqslant M/r^n$ for all $n \in \mathbb{N}_0$, and therefore the power series $g(z) := \sum_{n=0}^\infty z^n a_n$ converges for all $z \in B(0, r)$. For $x' \in E$, $|z| < r$ we compute, using Cauchy's integral formulas for the derivatives,

$$x'(g(z)) = \sum_{n=0}^\infty z^n \frac{1}{2\pi \mathrm{i}} \int_{|\zeta|=r} \frac{x'(f(\zeta))}{\zeta^{n+1}} \,\mathrm{d}\zeta = \sum_{n=0}^\infty \frac{(x' \circ f)^{(n)}(0)}{n!} z^n = x'(f(z)),$$

where the last equality is just the power series expansion of the holomorphic function $x' \circ f$. Since E is separating we conclude that $f(z) = g(z)$. □

Note that in the proof of the implication '(v) $\Rightarrow$ (i)' it is also shown that f is analytic.

3.3 Remarks. (a) If X is a dual Banach space, then the predual is a norming closed subspace of its bidual X'. (For instance, c_0 is norming for ℓ_1.) This illustrates a possible application of condition (iii) of Theorem 3.2.

(b) If X is reflexive and $E \subseteq X'$ is separating, then $\operatorname{lin} E$ is dense in X' and hence norming; cf. [ArNi00; Remark 1.2 d)]. We refer to [DaLi72] and [ArNi00; Remark 1.2 f)] for examples of separating subspaces that are not almost norming, in the situation of non-reflexive spaces. △

Next we come to the characterisation of holomorphy for $\mathcal{L}(X, Y)$-valued functions. As $\mathcal{L}(X, Y)$ is a Banach space, all the previous criteria apply. However, weak holomorphy is not a useful concept in this case, because typically one has no explicit description of the dual of $\mathcal{L}(X, Y)$.

3.4 Theorem. *Let $\Omega \subseteq \mathbb{C}$ be open, $F\colon \Omega \to \mathcal{L}(X, Y)$. Let B be a dense subset of X, and let $C \subseteq Y'$ be almost norming for Y. Then the following properties are equivalent.*

(i) *F is holomorphic (as an $\mathcal{L}(X, Y)$-valued function).*

(ii) *F is locally bounded, and $F(\cdot)x$ is holomorphic for all $x \in B$.*

(iii) *F is locally bounded, and $y'(F(\cdot)x)$ is holomorphic for all $x \in B$, $y' \in C$.*

Proof. (i) ⇒ (ii) ⇒ (iii) is clear.
(iii) ⇒ (i). It follows from the hypotheses on B and C that the set

$$E := \{A \mapsto y'(Ax);\, x \in B,\ y' \in C\} \subseteq \mathcal{L}(X,Y)'$$

is almost norming for $\mathcal{L}(X,Y)$. Therefore Theorem 3.2, (iv) ⇒ (i), implies the assertion. □

The last issue of this section is the convergence of sequences of holomorphic functions.

3.5 Theorem. *Let $\Omega \subseteq \mathbb{C}$ be open, and let (f_n) be a sequence of holomorphic functions $f_n\colon \Omega \to X$. Assume that (f_n) is locally uniformly bounded (i.e., for each $z_0 \in \Omega$ there exists $r > 0$ such that $B(z_0,r) \subseteq \Omega$ and $\sup\{\|f_n(z)\|;\, z \in B(z_0,r),\ n \in \mathbb{N}\} < \infty$) and that $f(z) := \lim_{n\to\infty} f_n(z)$ exists for all $z \in \Omega$.*
Then (f_n) converges to f locally uniformly, and f is holomorphic.

Proof. In the first step we show that the sequence (f_n) is locally uniformly equicontinuous. Let $z_0 \in \Omega$, $r > 0$ be such that $B[z_0,r] \subseteq \Omega$. Then $M := \sup_{z\in B[z_0,r],\, n\in\mathbb{N}} \|f_n(z)\| < \infty$, by hypothesis. From Cauchy's integral formula for the derivative,

$$f_n'(z) = \frac{1}{2\pi\mathrm{i}} \int_{|\zeta - z_0| = r} \frac{f_n(\zeta)}{(\zeta - z)^2}\, \mathrm{d}\zeta \qquad (z \in B(z_0,r),\ n \in \mathbb{N}),$$

we infer that for all $z \in B[z_0, r/2]$ one has

$$\|f_n'(z)\| \leqslant M\frac{4}{r} \qquad (n \in \mathbb{N}),$$

and we conclude that the sequence (f_n) is uniformly equicontinuous on $B[z_0, r/2]$.
The local uniform equicontinuity of the sequence (f_n) together with the pointwise convergence implies that (f_n) converges to f locally uniformly; see Exercise 3.2(a). Thus $x' \circ f_n \to x' \circ f$ locally uniformly, for all $x' \in X'$. It follows that f is weakly holomorphic, and therefore holomorphic, by Theorem 3.2. □

3.6 Corollary. *Let $\Omega \subseteq \mathbb{C}$ be open, and let (F_n) be a sequence of holomorphic functions $F_n\colon \Omega \to \mathcal{L}(X,Y)$. Assume that (F_n) is locally uniformly bounded and that $F(z) :=$ s-$\lim_{n\to\infty} F_n(z)$ exists for all $z \in \Omega$.*
Then F is holomorphic.

Proof. The hypotheses in combination with Theorem 3.5 imply that $F(\cdot)x$ is holomorphic for all $x \in X$. Therefore F is holomorphic, by Theorem 3.4. □

3.2 Holomorphic semigroups

Let X be a complex Banach space. For $\theta \in (0, \pi/2]$ we define the (open) sector

$$\Sigma_\theta := \{r\mathrm{e}^{\mathrm{i}\alpha};\, r > 0,\ |\alpha| < \theta\} \subseteq \mathbb{C}.$$

We will also use the notation $\Sigma_{\theta,0} := \Sigma_\theta \cup \{0\}$. A **holomorphic semigroup** on X (**of angle** θ) is a function $T\colon \Sigma_{\theta,0} \to \mathcal{L}(X)$, holomorphic on Σ_θ, satisfying

(i) $T(z_1+z_2) = T(z_1)T(z_2)$ for all $z_1, z_2 \in \Sigma_{\theta,0}$.

If additionally

(ii) $\lim_{\Sigma_{\theta'}\ni z\to 0} T(z)x = x$ for all $x \in X$ and all $\theta' \in (0,\theta)$,

then $T(0) = I$, by Remark 1.1(d), and T is called a **holomorphic C_0-semigroup** (**of angle** θ).

Saying that T is a holomorphic semigroup we will always mean that T brings along its domain of definition, in particular, *the angle of* T is defined.

3.7 Remarks. (a) It follows from the definition of a holomorphic C_0-semigroup that for all $\theta' \in (0,\theta)$ there exist $M' \geqslant 1$, $\omega' \in \mathbb{R}$ such that

$$\|T(z)\| \leqslant M' \mathrm{e}^{\omega' \operatorname{Re} z} \qquad (z \in \Sigma_{\theta'});$$

see Exercise 3.4.

(b) The reader might wonder why θ is restricted to the interval $(0,\pi/2]$ in the above definition. The reason is that using the same definition with $\theta \in (\pi/2,\pi]$ one automatically obtains a bounded generator; see Exercise 3.5. As a consequence, the semigroup has a holomorphic extension to all of $\mathbb{C}$. △

The following lemma shows that it suffices to check the semigroup property for real times.

3.8 Lemma. *Let $\theta \in (0,\pi/2]$, and let $T\colon \Sigma_\theta \to \mathcal{L}(X)$ be a holomorphic function satisfying $T(t+s) = T(t)T(s)$ for all $t,s > 0$.*

Then $T(z_1+z_2) = T(z_1)T(z_2)$ for all $z_1, z_2 \in \Sigma_\theta$.

Proof. Fixing $z_1 \in (0,\infty)$, we know that the functions $\Sigma_\theta \ni z_2 \mapsto T(z_1+z_2)$ and $\Sigma_\theta \ni z_2 \mapsto T(z_1)T(z_2)$ are holomorphic and coincide on $(0,\infty)$. The identity theorem (see Remark 3.1(b)) implies that they are equal on Σ_θ. Another application of the identity theorem, with fixed $z_2 \in \Sigma_\theta$, yields the assertion. □

3.9 Proposition. *Let T be a C_0-semigroup on X, and assume that there exist $\theta \in (0,\pi/2]$ and an extension of T to $\Sigma_{\theta,0}$, also called T, holomorphic on Σ_θ and satisfying*

$$\sup_{z\in\Sigma_\theta,\,|z|<1} \|T(z)\| < \infty.$$

Then $\lim_{\Sigma_\theta\ni z\to 0} T(z)x = x$ for all $x \in X$, and T is a holomorphic C_0-semigroup.

Proof. First note that Lemma 3.8 implies property (i) from above.

Let $x \in D := \bigcup_{t>0} \operatorname{ran}(T(t))$, i.e., there exist $y \in X$, $t > 0$ such that $x = T(t)y$. Then $\lim_{\Sigma_\theta\ni z\to 0} T(z)x = x$, by the continuity of the mapping $z \mapsto T(z)x = T(z+t)y$ at 0. Note that D is dense in X, because $T(t) \to I$ strongly as $t \to 0$. Therefore the boundedness assumption implies the assertion (recall Proposition 1.6). □

For a holomorphic C_0-semigroup T we now compute its derivative in terms of the generator A of the C_0-semigroup $T|_{[0,\infty)}$, and we discuss C_0-semigroups that are given by the restriction of T to rays $\mathrm{e}^{\mathrm{i}\alpha}[0,\infty)$.

3.10 Theorem. *Let T be a holomorphic C_0-semigroup of angle $\theta \in (0, \pi/2]$, and let A be the generator of the C_0-semigroup $T|_{[0,\infty)}$. Then:*

(a) *For all $z \in \Sigma_\theta$ one has* $\operatorname{ran}(T(z)) \subseteq \operatorname{dom}(A)$ *and* $T'(z) = AT(z)$. *For all* $x \in \operatorname{dom}(A)$, $z \in \Sigma_\theta$ *one has* $T'(z)x = T(z)Ax$.

(b) *For all* $x \in \operatorname{dom}(A)$, $\theta' \in (0, \theta)$ *one has*

$$\lim_{\Sigma_{\theta'} \ni z \to 0} \frac{1}{z}\big(T(z)x - x\big) = Ax.$$

(c) *For each* $\alpha \in (-\theta, \theta)$ *the mapping* $[0, \infty) \ni t \mapsto T_\alpha(t) := T(\mathrm{e}^{\mathrm{i}\alpha}t)$ *is a C_0-semigroup. The generator A_α of T_α equals* $\mathrm{e}^{\mathrm{i}\alpha}A$.

Proof. (a) Let $x \in X$, $z \in \Sigma_\theta$. Then

$$T'(z)x = \lim_{h\to 0+} \frac{1}{h}\big(T(z+h)x - T(z)x\big) = \lim_{h\to 0+} \frac{1}{h}(T(h) - I)T(z)x.$$

This implies that $T(z)x \in \operatorname{dom}(A)$ and $T'(z)x = AT(z)x$.

If $x \in \operatorname{dom}(A)$, then also $T'(z)x = T(z)\lim_{h\to 0+} \frac{1}{h}(T(h)x - x) = T(z)Ax$.

(b) Let $x \in \operatorname{dom}(A)$, $\theta' \in (0, \theta)$. Then the restriction of $T(\cdot)Ax$ to $\Sigma_{\theta',0}$ is continuous. Using part (a) one finds for $z \in \Sigma_{\theta'}$ that

$$\frac{1}{z}\big(T(z)x - x\big) = \frac{1}{z}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}T(sz)x\,\mathrm{d}s = \int_0^1 T(sz)Ax\,\mathrm{d}s,$$

and the latter tends to Ax as $z \to 0$ in $\Sigma_{\theta'}$.

(c) It is clear that T_α is a C_0-semigroup. Let $x \in \operatorname{dom}(A)$. Then by part (b) one obtains

$$\frac{1}{h}\big(T_\alpha(h)x - x\big) = \mathrm{e}^{\mathrm{i}\alpha}\frac{1}{\mathrm{e}^{\mathrm{i}\alpha}h}\big(T(\mathrm{e}^{\mathrm{i}\alpha}h)x - x\big) \to \mathrm{e}^{\mathrm{i}\alpha}Ax$$

as $h \to 0+$, and hence $x \in \operatorname{dom}(A_\alpha)$, $A_\alpha x = \mathrm{e}^{\mathrm{i}\alpha}Ax$. This shows that $\mathrm{e}^{\mathrm{i}\alpha}(\lambda - A) \subseteq \mathrm{e}^{\mathrm{i}\alpha}\lambda - A_\alpha$ for all $\lambda > 0$. Now choose λ so large that $\lambda \in \rho(A)$ and $\mathrm{e}^{\mathrm{i}\alpha}\lambda \in \rho(A_\alpha)$. Then the above operator inclusion is in fact an equality because a surjective mapping cannot have a proper injective extension, and it follows that $\mathrm{e}^{\mathrm{i}\alpha}A = A_\alpha$. □

3.11 Remarks. (a) Note that holomorphic C_0-semigroups are much more regular on $(0, \infty)$ than general C_0-semigroups. They are always continuous on $(0, \infty)$, even infinitely differentiable with respect to the operator norm.

(b) Theorem 3.10(a) implies that, for all $x \in X$, the function $[0, \infty) \ni t \mapsto T(t)x \in X$ is continuous, continuously differentiable on $(0, \infty)$, and solves the initial value problem

$$u'(t) = Au(t) \quad (t \in (0, \infty)), \qquad u(0) = x.$$

By Theorem 1.13(a) the solution is unique. △

Let T be a holomorphic C_0-semigroup. In the light of Theorem 3.10(b) it is justified to say that the generator of the C_0-semigroup $T|_{[0,\infty)}$ is also the **generator** of the holomorphic C_0-semigroup T. Clearly, the application of Theorem 2.7 yields estimates for the resolvents of the generator. We will not pursue this issue in full generality but restrict our attention to contractive holomorphic semigroups.

3.3 Generation of contractive holomorphic semigroups

As before, let X be a complex Banach space. We call a holomorphic semigroup of angle $\theta \in (0, \pi/2]$ **contractive** if $\|T(z)\| \leqslant 1$ for all $z \in \Sigma_{\theta,0}$.

The following theorem characterises the generation of contractive holomorphic C_0-semigroups.

3.12 Theorem. *For $\theta \in (0, \pi/2]$ and an operator A in X, the following properties are equivalent.*

(i) *A is the generator of a contractive holomorphic C_0-semigroup of angle θ.*

(ii) *For all $\alpha \in (-\theta, \theta)$, $\mathrm{e}^{\mathrm{i}\alpha}A$ is the generator of a contractive C_0-semigroup.*

(iii) *A is closed and densely defined, $\Sigma_\theta \subseteq \rho(A)$, and*

$$\big\|(\lambda I - A)^{-1}\big\| \leqslant \frac{1}{|\lambda|} \qquad (\lambda \in \Sigma_\theta).$$

Proof. (i) $\Rightarrow$ (ii) follows from Theorem 3.10(c).

(ii) $\Rightarrow$ (iii). 'Closed' and 'densely defined' are clear by Theorem 1.11(c). For $\alpha \in (-\theta, \theta)$, we know from Theorem 2.7 that $(0, \infty) \subseteq \rho(\mathrm{e}^{\mathrm{i}\alpha}A)$ and $\|(\mu I - \mathrm{e}^{\mathrm{i}\alpha}A)^{-1}\| \leqslant 1/\mu$ for all $\mu > 0$. The identity $(\mu I - \mathrm{e}^{\mathrm{i}\alpha}A) = \mathrm{e}^{\mathrm{i}\alpha}(\mathrm{e}^{-\mathrm{i}\alpha}\mu I - A)$ then shows that $\{\mathrm{e}^{-\mathrm{i}\alpha}\mu;\ \mu > 0\} \subseteq \rho(A)$ and $\|(\mathrm{e}^{-\mathrm{i}\alpha}\mu I - A)^{-1}\| \leqslant 1/\mu$ for all $\mu > 0$.

(iii) $\Rightarrow$ (i). We employ the exponential formula, Theorem 2.12. For $n \in \mathbb{N}$ we define the holomorphic function $F_n \colon \Sigma_\theta \to \mathcal{L}(X)$,

$$F_n(z) := (I - \tfrac{z}{n}A)^{-n} = \big(\tfrac{n}{z}(\tfrac{n}{z} - A)^{-1}\big)^n.$$

(For the holomorphy of F_n recall Remark 2.4(b).) The hypotheses imply that $\|F_n(z)\| \leqslant 1$ for all $z \in \Sigma_\theta$, $n \in \mathbb{N}$. They further imply that for each $\alpha \in (-\theta, \theta)$ the operator $\mathrm{e}^{\mathrm{i}\alpha}A$ generates a contractive C_0-semigroup T_α, by Theorem 2.9 (Hille–Yosida) and the same computation as in the proof of '(ii) $\Rightarrow$ (iii)'. Let $z \in \Sigma_\theta$, $z = \mathrm{e}^{\mathrm{i}\alpha}t$ with suitable $t > 0$, $\alpha \in (-\theta, \theta)$. Then $F_n(z) = (I - \frac{t}{n}\mathrm{e}^{\mathrm{i}\alpha}A)^{-n} \to T_\alpha(t)$ strongly as $n \to \infty$, by Theorem 2.12, so

$$T(z) := \operatorname{s-lim}_{n\to\infty} F_n(z)$$

exists for all $z \in \Sigma_{\theta,0}$, and $T|_{[0,\infty)} = T_0$ is the C_0-semigroup generated by A. Combining Corollary 3.6 and Proposition 3.9 we conclude that T is a contractive holomorphic C_0-semigroup of angle θ. □

In Theorem 3.20 we will present a further equivalence to the properties stated in Theorem 3.12, for the case of Hilbert spaces.

We add some comments on the generation of holomorphic semigroups that are not necessarily contractive.

3.13 Remarks. (a) The same proof as given for Theorem 3.12 can be used to prove the following equivalence.

Let $\theta \in (0, \pi/2]$, $M \geqslant 1$. Then an operator A is the generator of a **bounded holomorphic C_0-semigroup** of angle θ and with bound M if and only if for each $\alpha \in (-\theta, \theta)$ the operator $\mathrm{e}^{\mathrm{i}\alpha}A$ is the generator of a bounded C_0-semigroup with bound M. (We call the holomorphic semigroup T of angle θ bounded if $\sup_{z\in\Sigma_\theta}\|T(z)\|<\infty$. We point out that the terminology 'bounded holomorphic semigroup' is sometimes used differently.)

(b) If A is the generator of a bounded C_0-semigroup on X with bound $M\geqslant 1$, then $[\operatorname{Re}>0]:=\{\lambda\in\mathbb{C};\,\operatorname{Re}\lambda>0\}\subseteq\rho(A)$ and

$$\big\|(\lambda I-A)^{-1}\big\|\leqslant\frac{M}{\operatorname{Re}\lambda}\qquad(\operatorname{Re}\lambda>0),$$

by Theorem 2.7. If $\theta'\in(0,\pi/2)$ and $\lambda\in\Sigma_{\theta'}$, then $\frac{\operatorname{Re}\lambda}{|\lambda|}\geqslant\cos\theta'$, and this implies

$$\big\|(\lambda I-A)^{-1}\big\|\leqslant\frac{M}{\operatorname{Re}\lambda}\leqslant\frac{M}{\cos\theta'}\frac{1}{|\lambda|}\,.$$

Together with Theorem 3.10(c) the above considerations show that, if A is the generator of a bounded holomorphic semigroup of angle $\theta\in(0,\pi/2]$, then

$$\Sigma_{\theta+\pi/2}\subseteq\rho(A)\quad\text{and}\quad\sup_{\lambda\in\Sigma_{\theta'}}\big\|\lambda(\lambda I-A)^{-1}\big\|<\infty\quad(\theta'\in(0,\theta+\pi/2)).\tag{3.1}$$

This fact has a kind of converse: a closed, densely defined operator A in X is the generator of a holomorphic C_0-semigroup of angle $\theta\in(0,\pi/2]$ that is bounded on all sectors $\Sigma_{\theta'}$ with $\theta'\in(0,\theta)$ if and only if (3.1) holds (see [Kat80; Chap. IX, § 1.6], [EnNa00; Section II.4.a]). Remarkably, one does not need the estimate in (3.1) for powers of $\lambda(\lambda I-A)^{-1}$. △

3.4 The Lumer–Phillips theorem

Let H be a Hilbert space over $\mathbb{K}$. The scalar product of two elements $x,y\in H$ will be denoted by $(x\,|\,y)$, and it is defined to be linear in the first and antilinear in the second argument. An operator A in H is called **accretive** if

$$\operatorname{Re}(Ax\,|\,x)\geqslant 0\qquad(x\in\operatorname{dom}(A)).$$

3.14 Lemma. *Let A be an operator in H.*

(a) *Then A is accretive if and only if*

$$\|(\lambda I+A)x\|\geqslant\lambda\|x\|\qquad(x\in\operatorname{dom}(A))\tag{3.2}$$

for all $\lambda>0$.

(b) *For $\lambda>0$ with $\operatorname{ran}(\lambda I+A)=H$, inequality (3.2) holds if and only if $\lambda\in\rho(-A)$ and $\|(\lambda I+A)^{-1}\|\leqslant\lambda^{-1}$.*

Proof. (a) Assume that A is accretive, and let $\lambda>0$, $x\in\operatorname{dom}(A)$. Then

$$\|(\lambda I+A)x\|\|x\|\geqslant\operatorname{Re}((\lambda I+A)x\,|\,x)\geqslant(\lambda x\,|\,x)=\lambda\|x\|^2.$$

This establishes the asserted inequality (3.2).

On the other hand, assume that (3.2) holds for all $\lambda > 0$, and let $x \in \mathrm{dom}(A)$. Then

$$0 \leqslant \|(\lambda I + A)x\|^2 - \lambda^2\|x\|^2 = 2\lambda \operatorname{Re}(Ax \mid x) + \|Ax\|^2$$

for all $\lambda > 0$, and this implies $\operatorname{Re}(Ax \mid x) \geqslant 0$.

(b) follows from the next remark, which is valid in a more general context. □

3.15 Remark. Let X, Y be Banach spaces, B an operator from X to Y, and $\alpha > 0$. Then

$$\|Bx\|_Y \geqslant \alpha\|x\|_X \qquad (x \in \mathrm{dom}(B))$$

if and only if B is injective and $\|B^{-1}y\| \leqslant \alpha^{-1}\|y\|$ for all $y \in \mathrm{ran}(B)$. If these properties are satisfied and additionally $\mathrm{ran}(B) = Y$, then $B^{-1} \in \mathcal{L}(Y, X)$ and $\|B^{-1}\| \leqslant \alpha^{-1}$. △

An accretive operator A in H satisfying $\mathrm{ran}(I + A) = H$ is called **m-accretive**. A historical note on the prefix 'm': it should be remindful of the word 'maximal'. A *maximal accretive operator* is an accretive operator of which no proper extension is also an accretive operator. With this definition, a densely defined operator is m-accretive if and only if it is a maximal accretive operator; see [Phi59; Corollary of Theorem 1.1.1] (or Exercise 14.7(c)). However, there exist maximal accretive operators that are not m-accretive; see [Phi59; footnote (6)]. (We point out that such operators still have proper m-accretive extensions that are *relations*; see Exercise 14.7(b).)

3.16 Theorem (Lumer–Phillips). *Let A be an operator in H. Then $-A$ is the generator of a contractive C_0-semigroup if and only if A is m-accretive.*

Note the minus sign that makes $-A$ a generator. The background is that one likes the operator A to have the "positivity" property of accretivity; see also Remark 3.21.

An essential part of the proof of Theorem 3.16 is contained in the following lemma.

3.17 Lemma. *Let A be an accretive operator in H, and assume that there exists $\lambda_0 > 0$ such that $\mathrm{ran}(\lambda_0 I + A) = H$. Then $[\mathrm{Re} > 0] = \{\lambda \in \mathbb{K};\ \mathrm{Re}\,\lambda > 0\} \subseteq \rho(-A)$ (in particular, A is m-accretive),*

$$\big\|(\lambda I + A)^{-1}\big\| \leqslant \frac{1}{\lambda} \qquad (\lambda > 0),$$

A is closed, and $\mathrm{dom}(A)$ is dense in H.

Proof. From Lemma 3.14 one obtains $\lambda_0 \in \rho(-A)$. Moreover, if $0 < \lambda \in \rho(-A)$, then $\|(\lambda I + A)^{-1}\| \leqslant 1/\lambda$, and Theorem 2.2(c) implies $B(\lambda, \lambda) \subseteq B\big(\lambda, \|(\lambda I + A)^{-1}\|^{-1}\big) \subseteq \rho(-A)$. Combining these statements one sees that $[\mathrm{Re} > 0] \subseteq \rho(-A)$. In particular it follows that A is closed, by Remark 2.1(a).

Now let $x \in \mathrm{dom}(A)^\perp$. Then $y := (I + A)^{-1}x \in \mathrm{dom}(A)$, and hence

$$0 = \operatorname{Re}(x \mid y) = \operatorname{Re}((I + A)y \mid y) \geqslant \|y\|^2.$$

This implies that $y = 0$, $x = (I + A)y = 0$; therefore $\mathrm{dom}(A)^\perp = \{0\}$, and it follows that $\overline{\mathrm{dom}(A)} = \mathrm{dom}(A)^{\perp\perp} = H$. □

3.18 Remark. Every accretive operator $A \in \mathcal{L}(H)$ is automatically m-accretive. This follows from Lemma 3.17 because $\{\lambda \in \mathbb{R};\ \lambda > \|A\|\} \subseteq \rho(-A)$, by Remark 2.3(b). △

Proof of Theorem 3.16. If $-A$ generates a contractive C_0-semigroup, then Theorem 2.7 and Lemma 3.14 imply that A is m-accretive. The reverse implication follows from Lemma 3.17 and Theorem 2.9. □

For the remainder of this section we assume that H is a complex Hilbert space. In order to formulate a conclusion concerning the generation of holomorphic semigroups we need the following notions. For an operator A in H we define the **numerical range** $\operatorname{num}(A) := \bigl\{(Ax \mid x);\ x \in \operatorname{dom}(A),\ \|x\| = 1\bigr\}$. We call A **sectorial of angle** $\theta \in [0, \pi/2)$ if

$$\operatorname{num}(A) \subseteq \overline{\Sigma_\theta} = \bigl\{r\mathrm{e}^{\mathrm{i}\alpha};\ r \geqslant 0,\ |\alpha| \leqslant \theta\bigr\},$$

and we call A **m-sectorial** (**of angle** θ) if additionally $\operatorname{ran}(I + A) = H$. Here we supplement the notation introduced in Section 3.2 by $\Sigma_0 := (0, \infty)$. (Recall that previously Σ_θ was defined only for $0 < \theta \leqslant \pi/2$.)

We note that our definition of 'sectorial' is slightly more restrictive than the one used in [Kat80; Chap. V, § 3.10]. Unhappily, our terminology also conflicts with a notion introduced in [PrSo93; Section 3] that has become an important concept in the functional calculus for closed operators.

3.19 Remarks. Let A be an operator in H.

(a) Obviously A is accretive if and only if $\operatorname{num}(A) \subseteq [\operatorname{Re} \geqslant 0]$.

(b) We note that for any angle α one has $\operatorname{num}(\mathrm{e}^{\mathrm{i}\alpha}A) = \mathrm{e}^{\mathrm{i}\alpha}\operatorname{num}(A)$. Let $\theta \in (0, \pi/2]$. Then it follows from (a) that $\mathrm{e}^{\mathrm{i}\alpha}A$ is accretive for all $\alpha \in (-\theta, \theta)$ if and only if $\operatorname{num}(A) \subseteq \overline{\Sigma_{\pi/2-\theta}}$. △

We now draw a conclusion of the Lumer–Phillips theorem for generators of contractive holomorphic semigroups.

3.20 Theorem. *Let A be an operator in the complex Hilbert space H, and let $\theta \in (0, \pi/2]$. Then $-A$ generates a contractive holomorphic C_0-semigroup of angle θ if and only if A is m-sectorial of angle $\pi/2 - \theta$.*

Proof. By Theorems 3.12 and 3.16 it suffices to prove the equivalence

$$\mathrm{e}^{\mathrm{i}\alpha}A \text{ is m-accretive for all } \alpha \in (-\theta, \theta) \iff A \text{ is m-sectorial of angle } \pi/2 - \theta.$$

Either property implies that A is m-accretive, and hence $\operatorname{ran}(I + \mathrm{e}^{\mathrm{i}\alpha}A) = \operatorname{ran}(\mathrm{e}^{-\mathrm{i}\alpha}I + A) = H$ for all $\alpha \in (-\theta, \theta)$, by Lemma 3.17. Thus the equivalence follows from Remark 3.19(b). □

3.21 Remark. In the context of generators of contractive C_0-semigroups on Banach spaces, one usually considers *dissipative* instead of *accretive* operators. An operator A is called **dissipative** if $-A$ is accretive. One reason we prefer using the notion of accretive operators is that they will arise naturally in the context of forms. △

Notes

The equivalence of (i), (ii), (iii) in Theorem 3.2 is due to Dunford [Dun38; Theorem 76]. Theorem 3.4 is also due to Dunford; see [Hil39; footnote to Theorem 1], [HiPh57; Theorem 3.10.1].

Surprisingly, the properties (iv) and (v) of Theorem 3.2 can be relaxed jointly to a more general condition: f is holomorphic if f is locally bounded and there exists a separating set $E \subseteq X'$ such that $x' \circ f$ is holomorphic for all $x' \in E$. This generalisation is due to Grosse-Erdmann [GrE92] (published in [GrE04]); a short proof, based on the theorem of Krein–Šmulian, has been given in [ArNi00; Theorem 3.1] (see also [Voi20; Example 12.2]). An even further weakening can be found in [ABK20; Theorem 1.1], where it is shown that instead of local boundedness of f it suffices to assume that there exists a function $g \in L_{1,\mathrm{loc}}(\Omega)$ such that $\|f(z)\| \leqslant g(z)$ for a.e. $z \in \Omega$. An extra assumption of this kind cannot be omitted entirely: in [ArNi00; Theorem 1.5] a function f is constructed such that f has a discontinuity but nevertheless $x' \circ f$ is holomorphic for all x' from a norming (non-closed!) subspace of X'; cf. [ArNi00; Remark 1.4 f)].

The generation theorem for holomorphic semigroups, Theorem 3.12, is classically treated by defining the semigroup as a contour integral. We refer to the literature for this kind of proof. Our proof is a variant of the approaches presented in [AEH97; Section 4] and in [ArEl12a; Section 2]. We note that the characterisation stated at the end of Remark 3.13(b) can also be proved without writing the semigroup as a contour integral; see [AEH97; Theorem 4.3].

For a C_0-semigroup T on a complex Banach space, the asymptotic behaviour of $T(t)$ for $t \to 0+$ determines whether or not T is holomorphic, by which we mean that there exists $0 < \theta \leqslant \pi/2$ such that T extends to a holomorphic semigroup of angle θ. For example, T is holomorphic if

$$\limsup_{t\to 0+} \|T(t) - I\| < 2.$$

This result was obtained independently by Beurling [Beu70] and Kato [Kat70]. Earlier in the same year, Neuberger [Neu70] had proved that under the same hypothesis T is differentiable on $(0,\infty)$ with respect to the operator norm. The converse of the Beurling–Kato theorem is true if T is contractive and X is uniformly convex, in particular if X is an L_p-space, with $1 < p < \infty$; this is due to Pazy [Paz83; Chap. 2, Corollary 5.8]. (We refer to [Ada75; Corollary 2.29] for the uniform convexity of L_p-spaces; see also Exercise 10.4.)

A result of Beurling [Beu70; Theorem III] allows the following characterisation of holomorphy that is valid for arbitrary C_0-semigroups and arbitrary complex Banach spaces. The C_0-semigroup T is holomorphic if and only if there exists a complex polynomial p such that

$$\limsup_{t\to 0+} \|p(T(t))\| < \sup_{|z|\leqslant 1} |p(z)|; \tag{3.3}$$

see Fackler [Fac13; Section 2], where further details and some simplifications of Beurling's proofs can be found. See also Exercise 3.6(b), where the reader is asked to show that for every holomorphic C_0-semigroup there exists a polynomial as in (3.3).

The notion of accretivity and the Lumer–Phillips theorem can be generalised to Banach spaces; see [Paz83; Section 1.4], [EnNa00; Section II.3.b]. The Hilbert space case is all we need in this book, and the restriction to this case makes our treatment significantly simpler.

Exercises

3.1 Define the subspace E of $c_0' = \ell_1$ by

$$E := \Big\{x = (x_n) \in \ell_1\,;\ \sum_{n=1}^{\infty} x_n = 0\Big\}.$$

Show that E is almost norming, but not norming for c_0.

3.2 (a) Let M be a compact metric space, X a Banach space, (f_n) an equicontinuous sequence in $C(M;X)$. Show that the sequence (f_n) is uniformly equicontinuous.

Assume further that $f(s) := \lim_{n\to\infty} f_n(s)$ exists for all $s \in M$. Show that $f \in C(M;X)$ and that $\|f_n - f\|_\infty \to 0$ as $n \to \infty$.

(b) Let X, Y be Banach spaces, (T_n) a sequence in $\mathcal{L}(X,Y)$, $T_n \to T \in \mathcal{L}(X,Y)$ strongly. Let $M \subseteq X$ be compact. Show that $T_n x \to Tx$ $(n \to \infty)$ uniformly for $x \in M$. (Hint: Recall the uniform boundedness theorem.)

3.3 Let $\Omega \subseteq \mathbb{C}$ be an open set. Let (f_n) be a bounded sequence of bounded holomorphic functions $f_n \colon \Omega \to \mathbb{C}$ (i.e. $\sup_{z\in\Omega,\, n\in\mathbb{N}} |f_n(z)| < \infty$). Define the function $f \colon \Omega \to \ell_\infty$ by $f(z) := (f_n(z))_{n\in\mathbb{N}}$.

(a) Show that f is holomorphic. (Hint: Find a suitable norming set for ℓ_∞, and use Theorem 3.2.)

(b) Assume additionally that $f_\infty(z) := \lim_{n\to\infty} f_n(z)$ exists for all $z \in \Omega$. Show that then $f \colon \Omega \to c$ is holomorphic (where c denotes the subspace of ℓ_∞ consisting of the convergent sequences), and that f_∞ is holomorphic.

(c) Show that the convergence $f_n \to f_\infty$ is locally uniform. (Hint: Use the continuity of $f \colon \Omega \to c$.)

(Comment: The whole setup of this exercise could also start with X-valued functions, thereby yielding an alternative proof of Theorem 3.5.)

3.4 Let X be a complex Banach space, and let T be a holomorphic C_0-semigroup on X of angle $\theta \in (0, \pi/2]$.

(a) Show that for each $\theta' \in (0,\theta)$ there exists $\delta > 0$ such that

$$\sup_{z\in\Sigma_{\theta'},\, \mathrm{Re}\, z \leqslant \delta} \|T(z)\| < \infty.$$

(b) Prove the statement in Remark 3.7(a).

(c) Show that the estimate in Remark 3.7(a) can be written equivalently as follows: for each $\theta' \in (0,\theta)$ there exist $M'' \geqslant 1$, $\omega'' \in \mathbb{R}$ such that

$$\|T(z)\| \leqslant M'' \mathrm{e}^{\omega''|z|} \qquad (z \in \Sigma_{\theta'}).$$

3.5 Let X be a complex Banach space, and let $\pi/2 < \theta \leqslant \pi$. Suppose that T is a holomorphic C_0-semigroup of angle θ (with the same definition as in Section 3.2, but now for $\theta \in (\pi/2, \pi]$). Show that the generator of T belongs to $\mathcal{L}(X)$. (Hint: Show that the C_0-group S defined by $S(s) := T(\mathrm{i}s)$ $(s \in \mathbb{R})$ is continuous with respect to the operator norm.)

3.6 Let T be a holomorphic C_0-semigroup, with generator A, on a complex Banach space X.

(a) Assume that T is bounded. Show that there exists $c > 0$ such that $\|AT(t)\| \leqslant c/t$ for all $t > 0$. (Hint: Use Cauchy's integral formula for the derivative.)

(b) For $n \in \mathbb{N}$ let p_n be the polynomial given by $p_n(z) := z^{n+1} - z^n$ $(z \in \mathbb{C})$. Show that there exists $n \in \mathbb{N}$ such that

$$\sup_{0<t<1/n} \|p_n(T(t))\| < \sup_{|z|\leqslant 1} |p_n(z)|.$$

(Hint: Show by rescaling that an estimate as in part (a) holds for all $t \in (0, 2)$. Then use the equality $p_n(T(t)) = T(nt + t) - T(nt) = \int_{nt}^{(n+1)t} AT(s)\,\mathrm{d}s$ (strong integral).)

3.7 Let T be a bounded holomorphic C_0-semigroup of angle $\theta \in (0, \pi/2]$, with generator A.

(a) Show that there exists a strongly continuous extension (also called T) to the closure of $\Sigma_{\theta,0}$. (Hint: Show first that the extension can be defined on $\bigcup_{t>0} \operatorname{ran}(T(t))$.)

(b) Show that $T_{\pm\theta}$, defined by $T_{\pm\theta}(t) := T(\mathrm{e}^{\pm\mathrm{i}\theta}t)$ $(t \geqslant 0)$, are C_0-semigroups (the **boundary semigroups** of T), with generators $\mathrm{e}^{\pm\mathrm{i}\theta}A$. (Hint concerning the generator property: For $x \in \operatorname{dom}(A)$ and $|\alpha| < \theta$ consider the functions $t \mapsto T(\mathrm{e}^{\mathrm{i}\alpha}t)x$ and their derivatives; then take the limits as $\alpha \to \pm\theta$.)

(c) If $\theta = \pi/2$, then show that $T_{\pi/2}(t) := T(\mathrm{i}t)$ $(t \in \mathbb{R})$ defines a C_0-group $T_{\pi/2}$ (the **boundary group** of T), with generator $\mathrm{i}A$.

3.8 Let A be an operator in a complex Hilbert space.

(a) Let $0 < \theta < \pi/2$ and suppose that $\mathrm{e}^{\pm\mathrm{i}\theta}A$ generate contractive C_0-semigroups. Show that A generates a contractive holomorphic C_0-semigroup of angle θ. (Hint: Use Theorem 3.20.)

(b) Show that the assertion in (a) does not hold for $\theta = \pi/2$.

Chapter 4

The Sobolev space H^1, and applications

In Section 4.1 we present the definition and some basic properties of the Sobolev space H^1. The treatment is prepared by several important tools from analysis. The main objective of this chapter is the Hilbert space treatment of the Laplace operator in Section 4.2. In particular, the Dirichlet Laplacian will be presented as our first non-trivial example of a generator of a contractive holomorphic C_0-semigroup.

4.1 The Sobolev space H^1

4.1.1 Convolution

Convolution is an important tool for regularizing locally integrable functions. We recall that the space of **locally integrable functions** on an open set $\Omega \subseteq \mathbb{R}^n$ is given by

$$L_{1,\mathrm{loc}}(\Omega) := \bigl\{f\colon \Omega \to \mathbb{K};\ \forall x \in \Omega\ \exists r > 0\colon B(x,r) \subseteq \Omega,\ f|_{B(x,r)} \in L_1(B(x,r))\bigr\}$$

(where as usual functions are identified if they agree a.e.). Moreover, $C_\mathrm{c}^k(\Omega) := C^k(\Omega) \cap C_\mathrm{c}(\Omega)$ is the space of k times continuously differentiable functions with compact support, for $k \in \mathbb{N}_0$, and similarly $C_\mathrm{c}^\infty(\Omega) := C^\infty(\Omega) \cap C_\mathrm{c}(\Omega)$.

4.1 Lemma. *Let $f \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$, $k \in \mathbb{N}_0 \cup \{\infty\}$, $\rho \in C_\mathrm{c}^k(\mathbb{R}^n)$. We define the* ***convolution*** *of ρ and f,*

$$\rho * f(x) := \int_{\mathbb{R}^n} \rho(x-y)f(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} \rho(y)f(x-y)\,\mathrm{d}y \qquad (x \in \mathbb{R}^n).$$

Then $\rho * f \in C^k(\mathbb{R}^n)$, and for all $\alpha \in \mathbb{N}_0^n$, $|\alpha| \leqslant k$, one has

$$\partial^\alpha(\rho * f) = (\partial^\alpha \rho) * f.$$

Proof. (i) The integrals exist because ρ is bounded and has compact support.

(ii) Continuity of $\rho * f$: There exists $R > 0$ such that $\operatorname{spt}\rho \subseteq B(0,R)$. Let $R' > 0$, $\delta > 0$. For $x, x' \in B(0,R')$, $|x - x'| < \delta$, one obtains

$$\begin{aligned}|\rho * f(x) - \rho * f(x')| &= \left|\int_{B(0,R+R')} \bigl(\rho(x-y) - \rho(x'-y)\bigr)f(y)\,\mathrm{d}y\right| \\ &\leqslant \sup\bigl\{|\rho(z) - \rho(z')|;\ |z - z'| < \delta\bigr\} \int_{B(0,R+R')} |f(y)|\,\mathrm{d}y.\end{aligned}$$

The second factor in the last expression is finite because f is locally integrable, and the first factor becomes small for small δ because ρ is uniformly continuous. This completes the proof if $k = 0$. For the remainder of the proof let $k \geqslant 1$.

(iii) Let $j \in \{1, \dots, n\}$. The existence of the partial derivative of $\rho * f$ with respect to the j-th variable and the equality $\partial_j(\rho * f) = (\partial_j \rho) * f$ are a consequence of the differentiability of integrals with respect to a parameter. The function $(\partial_j \rho) * f$ is continuous, by step (ii) above, and therefore $\rho * f$ is continuously differentiable with respect to the j-th variable.

(iv) Induction establishes the assertion for all $\alpha \in \mathbb{N}_0^n$, $|\alpha| \leqslant k$. □

A sequence $(\rho_k)_{k \in \mathbb{N}}$ in $C_c(\mathbb{R}^n)$ is called a **delta sequence** if $\rho_k \geqslant 0$, $\int \rho_k(x)\,dx = 1$ and $\operatorname{spt} \rho_k \subseteq B[0, 1/k]$ for all $k \in \mathbb{N}$. (The term 'delta sequence' is motivated by the fact that the sequence $\big(C_c^\infty(\mathbb{R}^n) \ni \varphi \mapsto \int \rho_k \varphi\,dx\big)_{k \in \mathbb{N}}$ approximates the 'Dirac delta distribution' $C_c^\infty(\mathbb{R}^n) \ni \varphi \mapsto \varphi(0)$.)

4.2 Remarks. (a) We recall the standard example of a C_c^∞-function φ. The source of this function is the well-known function $\psi \in C^\infty(\mathbb{R})$,

$$\psi(t) := \begin{cases} 0 & \text{if } t \leqslant 0, \\ e^{-1/t} & \text{if } t > 0. \end{cases}$$

Then $\varphi(x) := \psi(1 - |x|^2)$ $(x \in \mathbb{R}^n)$ defines a function $0 \leqslant \varphi \in C_c^\infty(\mathbb{R}^n)$, with the property that $\varphi(x) \neq 0$ if and only if $|x| < 1$.

(b) If $0 \leqslant \rho \in C_c(\mathbb{R}^n)$, $\int \rho(x)\,dx = 1$, $\operatorname{spt} \rho \subseteq B[0, 1]$, and we define

$$\rho_k(x) := k^n \rho(kx) \qquad (x \in \mathbb{R}^n,\ k \in \mathbb{N}),$$

then (ρ_k) is a delta sequence. △

4.3 Proposition. *Let (ρ_k) be a delta sequence in $C_c(\mathbb{R}^n)$.*

(a) *Let $f \in C(\mathbb{R}^n)$. Then $\rho_k * f \to f$ uniformly on compact subsets of $\mathbb{R}^n$ as $k \to \infty$.*

(b) *Let $1 \leqslant p \leqslant \infty$, $f \in L_p(\mathbb{R}^n)$. Then $\rho_k * f \in L_p(\mathbb{R}^n)$,*

$$\|\rho_k * f\|_p \leqslant \|f\|_p \qquad (k \in \mathbb{N}).$$

If $1 \leqslant p < \infty$, then

$$\|\rho_k * f - f\|_p \to 0 \qquad (k \to \infty).$$

Proof. (a) The assertion is an easy consequence of the uniform continuity of f on compact subsets of $\mathbb{R}^n$.

(b) (i) For $p = \infty$ the estimate is straightforward.

If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then we estimate, using Hölder's inequality in the second step,

$$\begin{aligned} |\rho_k * f(x)| &= \left| \int \rho_k(x-y)^{\frac{1}{q}+\frac{1}{p}} f(y)\,dy \right| \\ &\leqslant \left(\int \rho_k(x-y)\,dy \right)^{1/q} \left(\int \rho_k(x-y)|f(y)|^p\,dy \right)^{1/p} \\ &= \left(\int \rho_k(x-y)|f(y)|^p\,dy \right)^{1/p}. \end{aligned}$$

This estimate also holds (trivially) for $p = 1$. Thus, with Fubini's theorem in the second step,

$$\begin{aligned}\int |\rho_k * f(x)|^p \,\mathrm{d}x &\leqslant \iint \rho_k(x-y)|f(y)|^p \,\mathrm{d}y\,\mathrm{d}x \\ &= \iint \rho_k(x-y)\,\mathrm{d}x\,|f(y)|^p \,\mathrm{d}y = \|f\|_p^p .\end{aligned}$$

(ii) Now let $1 \leqslant p < \infty$. For $k \in \mathbb{N}$ we define $T_k \in \mathcal{L}(L_p(\mathbb{R}^n))$ by $T_k g := \rho_k * g$ for $g \in L_p(\mathbb{R}^n)$; then step (i) shows that $\|T_k\| \leqslant 1$.

If $g \in C_{\mathrm{c}}(\mathbb{R}^n)$, then $T_k g \to g$ in $L_p(\mathbb{R}^n)$ as $k \to \infty$. Indeed, there exists $R > 0$ such that $\operatorname{spt} g \subseteq B[0, R]$. It is easy to check that this implies that $\operatorname{spt}(\rho_k * g) \subseteq B[0, R+1]$ for all $k \in \mathbb{N}$; see also Exercise 4.2(a). Moreover $\rho_k * g \to g$ uniformly on $B[0, R+1]$, by part (a). Therefore $\rho_k * g \to g$ in $L_p(\mathbb{R}^n)$ as $k \to \infty$.

Now the denseness of $C_{\mathrm{c}}(\mathbb{R}^n)$ in $L_p(\mathbb{R}^n)$ – see Remark 4.4 below – together with Proposition 1.6 implies that $T_k g \to g$ in $L_p(\mathbb{R}^n)$ as $k \to \infty$, for all $g \in L_p(\mathbb{R}^n)$. □

4.4 Remark. For an open set $\Omega \subseteq \mathbb{R}^n$ and $1 \leqslant p < \infty$, the set $C_{\mathrm{c}}(\Omega)$ is dense in $L_p(\Omega)$. For this property we refer to [Kna05; Corollary 6.4(a)]; the proof given in this reference for $\Omega = \mathbb{R}^n$ and $p \in \{1, 2\}$ carries over to the more general case stated above. For a more elementary proof we refer to Exercise G.3. △

4.5 Corollary. *Let $\Omega \subseteq \mathbb{R}^n$ be open, $1 \leqslant p < \infty$. Then $C_{\mathrm{c}}^\infty(\Omega)$ is dense in $L_p(\Omega)$.*

Proof. Let (ρ_k) be a delta sequence in $C_{\mathrm{c}}^\infty(\mathbb{R}^n)$.

Let $g \in C_{\mathrm{c}}(\Omega)$, and let $\tilde{g}$ be the extension of g to $\mathbb{R}^n$ by zero. Then $\rho_k * \tilde{g} \in C^\infty(\mathbb{R}^n)$ for all $k \in \mathbb{N}$, by Lemma 4.1. If $1/k < \operatorname{dist}(\operatorname{spt} g, \mathbb{R}^n \setminus \Omega)$, then $\operatorname{spt}(\rho_k * \tilde{g}) \subseteq \operatorname{spt} g + B[0, 1/k] \subseteq \Omega$ (see Exercise 4.2(a)), and therefore $(\rho_k * \tilde{g})|_\Omega \in C_{\mathrm{c}}^\infty(\Omega)$. From Proposition 4.3 we know that $\rho_k * \tilde{g} \to \tilde{g}$ in $L_p(\mathbb{R}^n)$ as $k \to \infty$. So, we have shown that $C_{\mathrm{c}}^\infty(\Omega)$ is dense in $C_{\mathrm{c}}(\Omega)$ with respect to the L_p-norm.

Now the denseness of $C_{\mathrm{c}}(\Omega)$ in $L_p(\Omega)$ yields the assertion. □

Note that in the above proof, $\operatorname{dist}(\operatorname{spt} g, \mathbb{R}^n \setminus \Omega) = \infty$ in the case $\Omega = \mathbb{R}^n$.

4.1.2 Distributional derivatives

We start with a simple version of integration by parts for functions on an open set $\Omega \subseteq \mathbb{R}^n$. Let $\psi \in C_{\mathrm{c}}^1(\Omega)$ and $j \in \{1, \ldots, n\}$. Then

$$\int_\Omega \partial_j \psi(x)\,\mathrm{d}x = 0. \tag{4.1}$$

Indeed, let $\tilde{\psi}$ be the extension of ψ to $\mathbb{R}^n$ by zero. Then Fubini's theorem yields

$$\int_\Omega \partial_j \psi(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \partial_j \tilde{\psi}(x)\,\mathrm{d}x = \int_{\check{x}_j \in \mathbb{R}^{n-1}} \Big(\int_{x_j \in \mathbb{R}} \partial_j \tilde{\psi}(x)\,\mathrm{d}x_j \Big)\,\mathrm{d}\check{x}_j = 0,$$

where we use the notation $\check{x}_j = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$, and where the integral over $x_j \in \mathbb{R}$ vanishes by the fundamental theorem of calculus. Now let $f \in C^1(\Omega)$, $\varphi \in C_{\mathrm{c}}^\infty(\Omega)$.

Then, applying (4.1) to the function $f\varphi$ and observing the product rule of differentiation, we obtain

$$\int_\Omega \partial_j f(x)\varphi(x)\,\mathrm{d}x = -\int_\Omega f(x)\partial_j\varphi(x)\,\mathrm{d}x. \tag{4.2}$$

This equality is fundamental for the subsequent definition of distributional derivatives.

Let $P(\partial) = \sum_{\alpha\in\mathbb{N}_0^n,\,|\alpha|\leqslant k} a_\alpha\partial^\alpha$ be a partial differential operator with constant coefficients $a_\alpha \in \mathbb{K}$ ($|\alpha| \leqslant k$), where $k \in \mathbb{N}$. (An important example is $P(\partial) = \sum_{j=1}^n \partial_j^2 = \Delta$, the Laplace operator.) Let $\Omega \subseteq \mathbb{R}^n$ be open, $f \in C^k(\Omega)$. Then for all "test functions" $\varphi \in C_\mathrm{c}^\infty(\Omega)$, repeated application of (4.2) implies

$$\int_\Omega \big(P(\partial)f\big)\varphi\,\mathrm{d}x = \int_\Omega f \sum_{|\alpha|\leqslant k}(-1)^{|\alpha|}a_\alpha\partial^\alpha\varphi\,\mathrm{d}x.$$

Now let $f \in L_{1,\mathrm{loc}}(\Omega)$. We say that $P(\partial)f \in L_{1,\mathrm{loc}}(\Omega)$ if there exists $g \in L_{1,\mathrm{loc}}(\Omega)$ such that

$$\int_\Omega g\varphi\,\mathrm{d}x = \int_\Omega f \sum_{|\alpha|\leqslant k}(-1)^{|\alpha|}a_\alpha\partial^\alpha\varphi\,\mathrm{d}x$$

for all $\varphi \in C_\mathrm{c}^\infty(\Omega)$, and we then say that $P(\partial)f = g$ holds in the **distributional sense**. In particular, if $\partial^\alpha f \in L_{1,\mathrm{loc}}(\Omega)$, then we call $\partial^\alpha f$ a **distributional** (or 'generalised', or 'weak') **derivative** of f.

In order to justify this definition we have to show that $g = P(\partial)f$ is unique as an element of $L_{1,\mathrm{loc}}(\Omega)$. This uniqueness is immediate from the following "fundamental lemma of the calculus of variations".

4.6 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $f \in L_{1,\mathrm{loc}}(\Omega)$,*

$$\int_\Omega f\varphi\,\mathrm{d}x = 0 \qquad (\varphi \in C_\mathrm{c}^\infty(\Omega)).$$

Then $f = 0$.

The statement '$f = 0$' means that f is the zero element of $L_{1,\mathrm{loc}}(\Omega)$, i.e., if f is a representative, then $f = 0$ a.e.

Proof of Lemma 4.6. Let $\psi \in C_\mathrm{c}^\infty(\Omega)$. We show that $\psi f = 0$ a.e. Defining

$$g(x) := \begin{cases} \psi(x)f(x) & \text{if } x \in \Omega, \\ 0 & \text{if } x \in \mathbb{R}^n \setminus \Omega, \end{cases}$$

we have $g \in L_1(\mathbb{R}^n)$.

Let (ρ_k) be a delta sequence in $C_\mathrm{c}^\infty(\mathbb{R}^n)$. From Proposition 4.3 we know that $\rho_k * g \to g$ in $L_1(\mathbb{R}^n)$. For $x \in \mathbb{R}^n$, $k \in \mathbb{N}$ we obtain

$$\rho_k * g(x) = \int_\Omega \rho_k(x-y)\psi(y)f(y)\,\mathrm{d}y = 0,$$

because $\rho_k(x-\cdot)\psi \in C_\mathrm{c}^\infty(\Omega)$. We have shown that $\rho_k * g = 0$, and hence $g = 0$ (as an L_1-function).

From $\psi f = 0$ a.e. for all $\psi \in C_\mathrm{c}^\infty(\Omega)$ we conclude that $f = 0$ a.e.; see Exercise 4.1. □

4.7 Remark. It follows from Lemma 4.6 and the preceding considerations that the distributional definition of $P(\partial)f$ given above is consistent with the 'classical' definition. To be more precise, let $f \in C^k(\Omega)$ and put $g := P(\partial)f$, computed with classical derivatives. Then $P(\partial)f$ exists in $L_{1,\mathrm{loc}}(\Omega)$ in the distributional sense and has the (unique) continuous representative g. △

In the remainder of this subsection we give more information on the one-dimensional case. The aim is to establish the following distributional variant of the fundamental theorem of calculus.

4.8 Proposition. *Let $-\infty \leqslant a < x_0 < b \leqslant \infty$, $f, g \in L_{1,\mathrm{loc}}(a,b)$. Then $f' = g$ in the distributional sense if and only if there exists $c \in \mathbb{K}$ such that*

$$f(x) = c + \int_{x_0}^{x} g(y)\,\mathrm{d}y \qquad (a.e.\ x \in (a,b)).$$

Note that the right-hand side of the previous equality is continuous as a function of x, and that therefore f has a continuous representative.

For the proof of Proposition 4.8 we need a preparatory lemma.

4.9 Lemma. *Let $-\infty \leqslant a < b \leqslant \infty$, $h \in L_{1,\mathrm{loc}}(a,b)$, and assume that $h' = 0$ in the distributional sense. Then there exists $c \in \mathbb{K}$ such that $h = c$ a.e.*

Proof. (i) We start with the observation that a function $\psi \in C_\mathrm{c}^\infty(a,b)$ is the derivative of a function in $C_\mathrm{c}^\infty(a,b)$ if and only if $\int \psi(x)\,\mathrm{d}x = 0$; in this case one obtains $\int \psi(x)h(x)\,\mathrm{d}x = 0$ since $h' = 0$. (We omit the domain (a,b) of integration from our notation.)

(ii) Let $\rho \in C_\mathrm{c}^\infty(a,b)$, $\int \rho(x)\,\mathrm{d}x = 1$, and put $c := \int \rho(x)h(x)\,\mathrm{d}x$. For all $\varphi \in C_\mathrm{c}^\infty(a,b)$ one obtains

$$\begin{aligned}\int \varphi(x)\big(h(x) - c\big)\,\mathrm{d}x &= \int \varphi(x)h(x)\,\mathrm{d}x - \int \varphi(y)\,\mathrm{d}y \int \rho(x)h(x)\,\mathrm{d}x \\ &= \int \Big(\varphi(x) - \int \varphi(y)\,\mathrm{d}y\,\rho(x)\Big)h(x)\,\mathrm{d}x = 0\end{aligned}$$

by step (i) because $\int \big(\varphi(x) - \int \varphi(y)\,\mathrm{d}y\,\rho(x)\big)\,\mathrm{d}x = 0$. Now the assertion is a consequence of Lemma 4.6. □

Proof of Proposition 4.8. (i) We first prove the sufficiency. Let $\varphi \in C_\mathrm{c}^\infty(a,b)$, and choose $x_1 \in (a, \inf \operatorname{spt} \varphi)$. Then

$$f(x) = c_1 + \int_{x_1}^{x} g(y)\,\mathrm{d}y \qquad (\text{a.e. } x \in (a,b)),$$

with $c_1 := c + \int_{x_0}^{x_1} g(y)\,\mathrm{d}y$. Since $\operatorname{spt}\varphi \subseteq (x_1, b)$ and $\int_a^b \varphi'(x)\,\mathrm{d}x = 0$, we obtain

$$\begin{aligned}\int_a^b \varphi'(x)f(x)\,\mathrm{d}x &= \int_{x_1}^b \varphi'(x)\Big(c_1 + \int_{x_1}^x g(y)\,\mathrm{d}y\Big)\,\mathrm{d}x = \iint\limits_{x_1<y<x<b} \varphi'(x)g(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \iint\limits_{x_1<y<x<b} \varphi'(x)g(y)\,\mathrm{d}x\,\mathrm{d}y = -\int_a^b \varphi(y)g(y)\,\mathrm{d}y.\end{aligned}$$

Thus $f' = g$.

(ii) Conversely assume that $f' = g$, and put

$$h(x) := f(x) - \int_{x_0}^{x} g(y)\,\mathrm{d}y \qquad (a < x < b).$$

Then step (i) implies that $h' = f' - g = 0$ in the distributional sense, so by Lemma 4.9 there exists $c \in \mathbb{K}$ such that $h = c$ a.e. □

4.1.3 Definition of $H^1(\Omega)$

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. We define the **Sobolev space**

$$H^1(\Omega) := \bigl\{f \in L_2(\Omega);\, \partial_j f \in L_2(\Omega)\ (j = 1, \ldots, n)\bigr\},$$

with scalar product

$$(f \,|\, g)_{H^1} := (f \,|\, g)_{L_2} + \sum_{j=1}^{n} (\partial_j f \,|\, \partial_j g)_{L_2}$$

(where

$$(f \,|\, g)_{L_2} := \int_\Omega f(x)\overline{g(x)}\,\mathrm{d}x$$

denotes the usual scalar product on $L_2(\Omega)$) and associated norm

$$\|f\|_{H^1} := \Bigl(\|f\|_2^2 + \sum_{j=1}^{n} \|\partial_j f\|_2^2\Bigr)^{1/2}.$$

4.10 Theorem. *The space $H^1(\Omega)$ is a separable Hilbert space.*

For the proof we single out an important property concerning the convergence of distributional derivatives, valid in a more general context.

4.11 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $\alpha \in \mathbb{N}_0^n$. Let (f_k) be a sequence in $L_{1,\mathrm{loc}}(\Omega)$, with $\partial^\alpha f_k \in L_{1,\mathrm{loc}}(\Omega)$ for all $k \in \mathbb{N}$, let $f, g \in L_{1,\mathrm{loc}}(\Omega)$, and assume that $f_k|_K \to f|_K$, $\partial^\alpha f_k|_K \to g|_K$ $(k \to \infty)$ in $L_1(K)$, for all compact sets $K \subseteq \Omega$.*

Then $\partial^\alpha f = g \in L_{1,\mathrm{loc}}(\Omega)$.

Proof. Let $\varphi \in C_\mathrm{c}^\infty(\Omega)$. Then

$$\int \partial^\alpha f_k(x)\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int f_k(x)\partial^\alpha\varphi(x)\,\mathrm{d}x$$

for all $k \in \mathbb{N}$. Taking the limit $k \to \infty$ we obtain

$$\int g(x)\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int f(x)\partial^\alpha\varphi(x)\,\mathrm{d}x.$$

This establishes the assertion. □

Proof of Theorem 4.10. Clearly $H^1(\Omega)$ is a pre-Hilbert space.

Let $J\colon H^1(\Omega) \to \bigoplus_{j=0}^n L_2(\Omega)$ be defined by $Jf := (f, \partial_1 f, \ldots, \partial_n f)$, where $\bigoplus_{j=0}^n L_2(\Omega)$ denotes the orthogonal direct sum. Then clearly J is an isometric operator. Therefore $H^1(\Omega)$ is complete if and only if the range of J is a closed subspace of the Hilbert space $\bigoplus_{j=0}^n L_2(\Omega)$. Let (f_k) be a sequence in $H^1(\Omega)$ such that (Jf_k) is convergent in $\bigoplus_{j=0}^n L_2(\Omega)$ to an element $(f^0, \ldots, f^n)$. This means that $f_k \to f^0$ and $\partial_j f_k \to f^j$ $(j = 1, \ldots, n)$ in $L_2(\Omega)$ as $k \to \infty$. Lemma 4.11 implies $f^j = \partial_j f^0$ $(j = 1, \ldots, n)$, so we have shown that $f := f^0 \in H^1(\Omega)$ and that $(f^0, \ldots, f^n) = Jf \in J(H^1(\Omega))$.

The space $L_2(\Omega)$ is separable, therefore $\bigoplus_{j=0}^n L_2(\Omega)$ is separable, the subspace $J(H^1(\Omega))$ of $\bigoplus_{j=0}^n L_2(\Omega)$ is separable, and thus $H^1(\Omega)$ is separable because J is isometric. □

As in Subsection 4.1.2 we give additional information on the one-dimensional case.

4.12 Theorem. *Let $-\infty < a < b < \infty$. Then every $f \in H^1(a, b)$ possesses a representative in $C[a, b]$, and the embedding $H^1(a, b) \hookrightarrow C[a, b]$ thus defined is continuous.*

Proof. By Proposition 4.8, every function $f \in H^1(a, b)$ has a representative of the form $f(x) = c + \int_{x_0}^x f'(y)\,\mathrm{d}y$, with $x_0 \in (a, b)$ and $c \in \mathbb{K}$. Since $f' \in L_2(a, b) \subseteq L_1(a, b)$, this representative is continuous as a function of $x \in [a, b]$.

Let $f \in H^1(a, b)$, with f chosen as the continuous representative. Then

$$\begin{aligned}\|f\|_{C[a,b]} &\leqslant \inf_{x_0 \in (a,b)} |f(x_0)| + \int_a^b |f'(y)|\,\mathrm{d}y \leqslant \frac{1}{b-a}\int_a^b |f(y)|\,\mathrm{d}y + \int_a^b |f'(y)|\,\mathrm{d}y \\ &\leqslant (b-a)^{-1/2}\|f\|_2 + (b-a)^{1/2}\|f'\|_2 \leqslant \big((b-a)^{-1/2} + (b-a)^{1/2}\big)\|f\|_{H^1}\,.\end{aligned} \tag{4.3}$$

This inequality shows that the embedding is a bounded operator. □

4.13 Remarks. (a) If X, Y are normed spaces, then $X \hookrightarrow Y$ is an abbreviation for saying that X is embedded into Y by an injective operator in $\mathcal{L}(X, Y)$. Usually this means that $X \subseteq Y$ in a natural way and that for some constant $c > 0$ one has

$$\|x\|_Y \leqslant c\|x\|_X \qquad (x \in X).$$

Such a constant is called an **embedding constant**.

(b) As an easy consequence of Theorem 4.12 one also obtains continuous embeddings $H^1(-\infty, d) \hookrightarrow C_0(-\infty, d]$, $H^1(d, \infty) \hookrightarrow C_0[d, \infty)$, for $d \in \mathbb{R}$, and $H^1(-\infty, \infty) \hookrightarrow C_0(-\infty, \infty)$, with embedding constant $c = 2$. To see this, observe that the restriction of an H^1-function to a subinterval (a, b) belongs to $H^1(a, b)$ and apply the estimate 4.3 to subintervals of length $b - a = 1$.

(c) Theorem 4.12 is a simple instance of a Sobolev embedding theorem. △

4.1.4 Denseness properties

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. For $f \in L_{1,\mathrm{loc}}(\Omega)$ we define the **support** of f by

$$\operatorname{spt} f := \Omega \setminus \bigcup\big\{U \subseteq \Omega;\ U \text{ open},\ f|_U = 0 \text{ a.e.}\big\};$$

this definition is consistent with the already defined support for continuous functions. Observe that $f = 0$ a.e. on $\Omega \setminus \operatorname{spt} f$, by Exercise 4.1(b). Furthermore we define

$$H^1_{\mathrm{c}}(\Omega) := \bigl\{ f \in H^1(\Omega);\, \operatorname{spt} f \text{ compact in } \Omega \bigr\},$$
$$H^1_0(\Omega) := \overline{H^1_{\mathrm{c}}(\Omega)}^{H^1(\Omega)}.$$

4.14 Remarks. (a) In a generalised sense, functions in $H^1_0(\Omega)$ 'vanish at the boundary of Ω'. For sets Ω with regular boundary, this will be made more precise in Section 7.2.

(b) For $f \in H^1_0(\Omega)$ the extension to $\mathbb{R}^n$ by zero belongs to $H^1(\mathbb{R}^n)$; see Exercise 4.7. (One first shows this for $f \in H^1_{\mathrm{c}}(\Omega)$ and then uses a denseness argument.) △

4.15 Theorem. (a) $H^1_0(\mathbb{R}^n) = H^1(\mathbb{R}^n)$.

(b) $H^1_0(\Omega) = \overline{C^\infty_{\mathrm{c}}(\Omega)}^{H^1(\Omega)}$.

We need the following auxiliary result.

4.16 Lemma. (a) *Let* $\alpha \in \mathbb{N}_0^n$, $f, \partial^\alpha f \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$, $\rho \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$. *Then*

$$\partial^\alpha(\rho * f) = \rho * (\partial^\alpha f).$$

(b) *Let* $f \in H^1(\mathbb{R}^n)$, *and let* (ρ_k) *be a delta sequence in* $C^\infty_{\mathrm{c}}(\mathbb{R}^n)$. *Then* $\rho_k * f \to f$ *in* $H^1(\mathbb{R}^n)$ *as* $k \to \infty$.

Proof. (a) Using Lemma 4.1 in the first equality and the definition of the distributional derivative in the third, we obtain

$$\begin{aligned}\partial^\alpha(\rho * f)(x) &= \int \partial^\alpha \rho(x-y) f(y)\,\mathrm{d}y = (-1)^{|\alpha|} \int \partial^\alpha[\rho(x-\cdot)](y) f(y)\,\mathrm{d}y \\ &= \int \rho(x-y)\partial^\alpha f(y)\,\mathrm{d}y = \rho * \partial^\alpha f(x).\end{aligned}$$

(b) From Proposition 4.3(b) we know that $\rho_k * f \to f$ in $L_2(\mathbb{R}^n)$ as $k \to \infty$. Using part (a) we further obtain

$$\partial_j(\rho_k * f) = \rho_k * \partial_j f \to \partial_j f \qquad (k \to \infty)$$

in $L_2(\mathbb{R}^n)$, for $j = 1, \dots, n$, and the assertion follows. □

Proof of Theorem 4.15. (a) Let $f \in H^1(\mathbb{R}^n)$, $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$. Then from Exercise 4.5(c) we obtain $\partial_j(\varphi f) = \partial_j \varphi f + \varphi \partial_j f \in L_2(\mathbb{R}^n)$ for all $j \in \{1, \dots, n\}$. Hence $\varphi f \in H^1_{\mathrm{c}}(\mathbb{R}^n)$.

Choose $\psi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$, $\psi|_{B(0,1)} = 1$, and put $\psi_k := \psi(\cdot/k)$ $(k \in \mathbb{N})$. Then $\psi_k f \to f$ and $\partial_j(\psi_k f) = \partial_j \psi_k f + \psi_k \partial_j f \to 0 \cdot f + \partial_j f$ in $L_2(\mathbb{R}^n)$ as $k \to \infty$. Thus the functions $\psi_k f \in H^1_{\mathrm{c}}(\mathbb{R}^n)$ approximate f in $H^1(\mathbb{R}^n)$.

(b) The inclusion '$\supseteq$' is trivial. The proof of '$\subseteq$' is analogous to the proof of Corollary 4.5: it is sufficient to show that each $f \in H^1_{\mathrm{c}}(\Omega)$ can be approximated by elements of $C^\infty_{\mathrm{c}}(\Omega)$. This, however, is a consequence of Remark 4.14(b) and Lemma 4.16(b). □

4.17 Remark. In general, the spaces $H^1(\Omega)$ and $H^1_0(\Omega)$ do not coincide. We prove this for the case of bounded $\Omega \neq \varnothing$, using Theorem 4.15(b).

Note that for all $f \in H^1_0(\Omega)$ one has $\int \partial_1 f(x)\,\mathrm{d}x = 0$. This is clear if $f \in C_c^\infty(\Omega)$ and carries over to $H^1_0(\Omega)$ by continuity. Now let $f \in H^1(\Omega)$ be defined by $f(x) := x_1$ $(x \in \Omega)$. Then $\int \partial_1 f(x)\,\mathrm{d}x \neq 0$ and therefore $f \in H^1(\Omega) \setminus H^1_0(\Omega)$. △

4.18 Examples (Right translation semigroups). We come back to Examples 1.7 and describe the generator A, but only for the case $p = 2$.

(a) On $L_2(\mathbb{R})$: Recall from Example 1.16 that $D := C_c^1(\mathbb{R})$ is a core for A, i.e. $A = \overline{A|_D}$, and that $Af = -f'$ for all $f \in D$. Thus $A|_D$, considered as a subspace of $L_2(\mathbb{R}) \times L_2(\mathbb{R})$, is isomorphic to $C_c^1(\mathbb{R})$ as a subspace of $H^1(\mathbb{R})$ (cf. the proof of Theorem 4.10), and one obtains

$$\operatorname{dom}(A) = \overline{C_c^1(\mathbb{R})}^{H^1(\mathbb{R})} = H^1_0(\mathbb{R}) = H^1(\mathbb{R}), \quad Af = -f' \quad (f \in \operatorname{dom}(A)).$$

(b) On $L_2(0,\infty)$: Similarly to part (a) one obtains

$$\operatorname{dom}(A) = \overline{C_c^1(0,\infty)}^{H^1(0,\infty)} = H^1_0(0,\infty), \quad Af = -f' \quad (f \in \operatorname{dom}(A)).$$

(c) On $L_2(-\infty,0)$: In this case, the set $C_c^1(-\infty,0]$ is a subset of $\operatorname{dom}(A)$, $Af = -f'$ for all $f \in C_c^1(-\infty,0]$, and $C_c^1(-\infty,0]$ is invariant under T; hence by Proposition 1.15, $C_c^1(-\infty,0]$ is a core for A. Then as in part (a) we can conclude that

$$\operatorname{dom}(A) = \overline{C_c^1(-\infty,0]}^{H^1(-\infty,0)} = H^1(-\infty,0), \quad Af = -f' \quad (f \in \operatorname{dom}(A)),$$

once we know that $C_c^1(-\infty,0]$ is dense in $H^1(-\infty,0)$.

In order to obtain this denseness property we first note that the method of the proof of Theorem 4.15(a) shows that $\{f \in H^1(-\infty,0);\, \operatorname{spt} f \text{ bounded}\}$ is dense in $H^1(-\infty,0)$. Now let $a < 0$, $f \in H^1(-\infty,0)$, $\operatorname{spt} f \subseteq (a,0)$. Then there exists a sequence (φ_n) in $C_c(-\infty,0)$ with $\operatorname{spt}\varphi_n \subseteq (a,0)$ for all $n \in \mathbb{N}$ such that $\varphi_n \to f'$ in $L_2(-\infty,0)$ as $n \to \infty$. Putting

$$f_n(x) := \int_{-\infty}^x \varphi_n(y)\,\mathrm{d}y \qquad (x \in (-\infty,0),\ n \in \mathbb{N}),$$

one easily sees that $f_n \in C_c^1(-\infty,0]$ for all $n \in \mathbb{N}$, and $f_n \to f$ in $H^1(-\infty,0)$ as $n \to \infty$.

(d) On $L_2(0,1)$: As in part (c) one first shows that $C_c^1(0,1]$ is a core for A and then obtains

$$\operatorname{dom}(A) = \overline{C_c^1(0,1]}^{H^1(0,1)} = \bigl\{f \in H^1(0,1);\, f(0) = 0\bigr\}, \quad Af = -f' \quad (f \in \operatorname{dom}(A)).$$

The required denseness of $C_c^1(0,1]$ is seen as follows. For a function $f \in H^1(0,1)$ with $f(0) = 0$ one chooses a sequence (φ_n) in $C_c(0,1)$ approximating f' in $L_2(0,1)$. Then

$$f_n(x) := \int_0^x \varphi_n(y)\,\mathrm{d}y \qquad (x \in (0,1),\ n \in \mathbb{N})$$

defines a sequence (f_n) in $C_c^1(0,1]$ approximating f in $H^1(0,1)$. △

4.2 A variant of the Poisson problem, and the Dirichlet Laplacian

An important partial differential equation from mathematical physics is the Poisson equation

$$-\Delta u = f \text{ on } \Omega, \quad u|_{\partial\Omega} = 0,$$

for a given function f on an open set $\Omega \subseteq \mathbb{R}^n$. The first aim of the present section is to solve a variant of this equation. One might look for a solution u that is twice differentiable on Ω and continuous on the closure. We will not treat the problem in this form but rather weaken the requirements.

More explicitly, the boundary condition $u|_{\partial\Omega} = 0$ will be modified to the requirement that u should belong to $H_0^1(\Omega)$, and the equation itself will only be required to hold in the distributional sense.

In the second part of the section we will establish the connection to m-accretivity of the negative Dirichlet-Laplacian in $L_2(\Omega)$ and holomorphic semigroups.

4.2.1 The equation $u - \Delta u = f$

4.19 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $f \in L_2(\Omega)$. Then there exists a unique function $u \in H_0^1(\Omega)$ such that $u - \Delta u = f$ in the distributional sense.*

We insert a lemma expressing the distributional equality in another form.

4.20 Lemma. *Let $u \in H^1(\Omega)$, $g \in L_2(\Omega)$. Then $-\Delta u = g$ in the distributional sense if and only if*

$$(v \mid g)_{L_2} = \sum_{j=1}^{n} (\partial_j v \mid \partial_j u)_{L_2} \qquad (v \in H_0^1(\Omega)). \tag{4.4}$$

Proof. By the definition of distributional derivatives, the equation $-\Delta u = g$ is equivalent to

$$(\varphi \mid g) = -(\Delta\varphi \mid u) = \sum_{j=1}^{n} (\partial_j \varphi \mid \partial_j u) \qquad (\varphi \in C_c^\infty(\Omega)),$$

i.e. to the validity of equation (4.4) for all $v \in C_c^\infty(\Omega)$. As both mappings $H^1(\Omega) \ni v \mapsto (v \mid g) \in \mathbb{K}$ and $H^1(\Omega) \ni v \mapsto \sum_{j=1}^n (\partial_j v \mid \partial_j u) \in \mathbb{K}$ are continuous, the equality of the terms in (4.4) extends to the closure of $C_c^\infty(\Omega)$ in $H^1(\Omega)$, i.e. to $H_0^1(\Omega)$ (recall Theorem 4.15(b)). □

Proof of Theorem 4.19. We define a bounded linear functional $\eta\colon H_0^1(\Omega) \to \mathbb{K}$ by

$$\eta(v) := (v \mid f)_{L_2} \qquad (v \in H_0^1(\Omega)).$$

Applying the representation theorem of Fréchet–Riesz (see e.g. [Bre11; Theorem 5.5]) we obtain $u \in H_0^1(\Omega)$ such that

$$\eta(v) = (v \mid u)_{H_0^1(\Omega)} \qquad (v \in H_0^1(\Omega)).$$

Putting this equation and the definition of η together we obtain

$$(v \,|\, f) = (v \,|\, u)_{H_0^1(\Omega)} = (v \,|\, u) + \sum_{j=1}^{n} (\partial_j v \,|\, \partial_j u) \qquad (v \in H_0^1(\Omega)).$$

Shifting the first term on the right-hand side to the left and applying Lemma 4.20 we conclude that $-\Delta u = f - u$ in the distributional sense.

The uniqueness of u is a consequence of the uniqueness in the Fréchet–Riesz representation theorem. □

4.2.2 The Dirichlet Laplacian

In this subsection we reformulate the result of Subsection 4.2.1 in operator language. As before, let $\Omega \subseteq \mathbb{R}^n$ be an open set.

We define the **Dirichlet Laplacian** Δ_{D} in $L_2(\Omega)$,

$$\Delta_{\mathrm{D}} := \big\{(u, f) \in L_2(\Omega) \times L_2(\Omega)\,;\; u \in H_0^1(\Omega),\; \Delta u = f\big\}.$$

In other words,

$$\begin{aligned} \operatorname{dom}(\Delta_{\mathrm{D}}) &:= \big\{u \in H_0^1(\Omega)\,;\; \Delta u \in L_2(\Omega)\big\}, \\ \Delta_{\mathrm{D}} u &:= \Delta u \qquad (u \in \operatorname{dom}(\Delta_{\mathrm{D}})). \end{aligned}$$

The name 'Dirichlet Laplacian' may be ambiguous, so we give a short explanation. In principle, 'Dirichlet boundary conditions' are of the form $u|_{\partial\Omega} = \varphi$ for a given function φ defined on $\partial\Omega$. Recall that the membership of u in $H_0^1(\Omega)$ is a version of Dirichlet boundary condition zero. So, 'Dirichlet Laplacian' should be regarded as an abbreviation of 'Laplacian with Dirichlet boundary condition zero'.

4.21 Theorem. *The negative Dirichlet Laplacian $-\Delta_{\mathrm{D}}$ is m-accretive. The operator Δ_{D} is the generator of a contractive C_0-semigroup on $L_2(\Omega)$.*

In the complex case, $-\Delta_{\mathrm{D}}$ is m-sectorial of angle 0, and Δ_{D} is the generator of a contractive holomorphic C_0-semigroup of angle $\pi/2$ on $L_2(\Omega;\mathbb{C})$.

Proof. Theorem 4.19 states that $\operatorname{ran}(I - \Delta_{\mathrm{D}}) = L_2(\Omega)$. For $u \in \operatorname{dom}(\Delta_{\mathrm{D}})$ an application of Lemma 4.20 yields

$$(-\Delta_{\mathrm{D}} u \,|\, u) = (-\Delta u \,|\, u) = \sum_{j=1}^{n} (\partial_j u \,|\, \partial_j u) \in [0, \infty).$$

This implies that $-\Delta_{\mathrm{D}}$ is m-accretive, and in the complex case it also shows that $-\Delta_{\mathrm{D}}$ is m-sectorial of angle 0. The remaining assertions follow from Theorem 3.16 (Lumer–Phillips) and Theorem 3.20. □

The statement that $-\Delta_{\mathrm{D}}$ is m-sectorial of angle 0 is equivalent to saying that $-\Delta_{\mathrm{D}}$ is an accretive self-adjoint operator; this will be explained in Chapter 6.

The semigroup T generated by Δ_{D} describes heat propagation in Ω. In fact, let $u_0 \in L_2(\Omega)$, and put $u(t) := T(t)u_0$ for $t \geqslant 0$. Then $u \in C([0,\infty); L_2(\Omega)) \cap C^\infty(0,\infty; L_2(\Omega))$,

$u(t) \in \mathrm{dom}(\Delta_{\mathrm{D}})$ for all $t > 0$, and

$$\begin{aligned} &u'(t) = \Delta u(t), \quad u(t)|_{\partial\Omega} = 0 \qquad (t > 0), \\ &u(0) = u_0\,; \end{aligned}$$

see Remarks 3.11 and 4.14(a) (as well as Exercise 4.9 for the case $\mathbb{K} = \mathbb{R}$). If we consider a homogeneous body Ω (a bounded open subset of $\mathbb{R}^n$) and $u_0(x)$ as the temperature at $x \in \Omega$ at time 0, then $u(t)(x)$ is the temperature at time $t > 0$ at x. The boundary condition means that the temperature is kept at 0 at the boundary. One expects that $\lim_{t\to\infty} u(t) = 0$. That this is indeed the case will be seen in Exercise 5.2(c).

Notes

The theory of partial differential equations started in the 18th century, after the invention of Newton's "fluxions" and Leibnitz' ingenious contribution, the introduction of the symbols of differential and integral calculus. Up to the beginning of the 20th century, solutions of differential equations were always discussed under the basic assumption that all the derivatives of the solution appearing in the equation are classical derivatives.

It was in the early 20th century, in connection with the development of functional analysis and integration theory, that 'generalised solutions' of problems in differential equation were considered. The pioneering idea appears to be contained in the paper of Beppo Levi [Lev06] from 1906: in his discussion of the Dirichlet problem (in two space dimensions) he used an approximation procedure to find a continuous solution whose generalised first partial derivatives belong to L_2. (For more information on the Dirichlet problem we refer to the Notes of Chapter 7.)

The groundbreaking contribution of Sergei Lvovich Sobolev [Sob38] in 1938 was to introduce normed spaces of functions with distributional derivatives belonging to L_p and to study embeddings between certain of these spaces. (Our Theorem 4.12 is a particularly simple instance of a 'Sobolev embedding theorem'.) In the early 1950's, French mathematicians were in search of a name for this kind of spaces. The suggestion to call them 'Beppo Levi spaces' was strongly opposed by Levi himself; as a result, the name 'Sobolev spaces' was coined and now is universally accepted. For a detailed account of the above sketchy description of the development in the 20th century we refer to [WKK09; Section 10.1] and to [Nau02].

In Section 4.1 we have collected basics of Sobolev spaces as far as we have needed them in this chapter. Additional properties will be presented in subsequent chapters, according to what will be needed and used later. In this book we will mainly need the 'Hilbert-Sobolev spaces' $H^1(\Omega)$ and $H^1_0(\Omega)$, with the exception of Chapter 17 and Appendices D and H, where we use $W^1_1(\Omega)$. More generally, for $k \in \mathbb{N}_0$ and $p \in [1, \infty]$, the Sobolev space $W^k_p(\Omega)$ is defined by

$$W^k_p(\Omega) := \bigl\{u \in L_p(\Omega)\,;\ \partial^\alpha u \in L_p(\Omega)\ (\alpha \in \mathbb{N}^n_0,\ |\alpha| \leqslant k)\bigr\},$$

where the derivatives $\partial^\alpha u$ are understood in the distributional sense. For more information we refer to [AdFo05]. We mention that one also defines Sobolev spaces of negative order; for an example we refer to the definition of $H^{-1}(\Omega)$ in Section 16.1.

Exercises

4.1 Let $\Omega \subseteq \mathbb{R}^n$ be open.

(a) Show that there exists a **standard exhaustion** $(\Omega_k)_{k\in\mathbb{N}}$ of Ω, i.e., Ω_k is open, relatively compact in Ω_{k+1} $(k \in \mathbb{N})$, and $\bigcup_{k\in\mathbb{N}} \Omega_k = \Omega$. (Hint: For $\Omega \neq \mathbb{R}^n$ use

$$\Omega_k := \bigl\{x \in \Omega;\ |x| < k,\ \operatorname{dist}(x, \mathbb{R}^n \setminus \Omega) > 1/k\bigr\}.)$$

(b) Let $f \in L_{1,\mathrm{loc}}(\Omega)$, and assume that $f = 0$ locally, i.e., for all $x \in \Omega$ there exists $r > 0$ such that $B(x,r) \subseteq \Omega$ and $f|_{B(x,r)} = 0$. Show that $f = 0$. (All '$= 0$' should be interpreted as a.e.)

4.2 (a) Let $f \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$, $\rho \in C_{\mathrm{c}}(\mathbb{R}^n)$. Show that $\operatorname{spt}(\rho * f) \subseteq \operatorname{spt} f + \operatorname{spt} \rho$. (Hint: Show first that $\operatorname{spt} f + \operatorname{spt} \rho$ is closed.)

(b) Let $K \subseteq U \subseteq \mathbb{R}^n$, K compact, U open. Show that there exists $\psi \in C_{\mathrm{c}}^\infty(\mathbb{R}^n)$ such that $\operatorname{spt} \psi \subseteq U$ and $\mathbf{1}_K \leqslant \psi \leqslant 1$. (Hint: Note that $\operatorname{dist}(K, \mathbb{R}^n \setminus U) > 0$. Find ψ as the convolution of a suitable function $\rho \in C_{\mathrm{c}}^\infty(\mathbb{R}^n)$ with a suitable indicator function.)

4.3 (a) Let K be a compact subset of a metric space M, and let $(U_j)_{j=1,\ldots,m}$ be a covering of K by open subsets of M. Show that there exists a *shrinking* of $(U_j)_{j=1,\ldots,m}$, i.e. an open covering $(V_j)_{j=1,\ldots,m}$ of K such that $\overline{V_j} \subseteq U_j$ for all $j \in \{1, \ldots, m\}$.

Hints: Use the compactness of K to find a finite covering $\bigl(B(x, r_x)\bigr)_{x\in F}$ of K, with a finite set $F \subseteq K$ and $r_x > 0$ for all $x \in F$, such that for each $x \in F$ the closed ball $B[x, r_x]$ is contained in one of the sets U_j. Then put

$$V_j := \bigcup_{x\in F \text{ with } B[x,r_x]\subseteq U_j} B(x, r_x) \qquad (j = 1, \ldots, m).$$

(b) Let $\Omega \subseteq \mathbb{R}^n$ be open, $K \subseteq \Omega$ compact, $(U_j)_{j=1,\ldots,m}$ a covering of K by open subsets of Ω. Show that there exists a family $(\chi_j)_{j=1,\ldots,m}$ in $C_{\mathrm{c}}^\infty(\Omega)$, $\chi_j \geqslant 0$ and $\operatorname{spt} \chi_j \subseteq U_j$ for all $j \in \{1, \ldots, m\}$, such that $\sum_{j=1}^m \chi_j(x) = 1$ for all $x \in K$. The family $(\chi_j)_{j=1,\ldots,m}$ is called a *partition of unity on K subordinate to the covering $(U_j)_{j=1,\ldots,m}$ of K*.

Hints: Use part (a) to find a shrinking $(V_j)_{j=1,\ldots,m}$ of $(U_j)_{j=1,\ldots,m}$ with the additional property that each set V_j is bounded. Construct a family $(W_j)_{j=1,\ldots,m}$ of pairwise disjoint measurable sets such that $\bigcup_{j=1}^m W_j = \bigcup_{j=1}^m V_j$. Then choose $\rho \in C_{\mathrm{c}}^\infty(\mathbb{R}^n)_+$ with $\int \rho(x)\,\mathrm{d}x = 1$, $\operatorname{spt} \rho \subseteq B(0, \varepsilon)$ for suitably small $\varepsilon > 0$ and put $\chi_j := \rho * \mathbf{1}_{W_j}$ to obtain the asserted partition of unity.

(c) Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $P(\partial)$ be a partial differential operator with constant coefficients, $f, g \in L_{1,\mathrm{loc}}(\Omega)$, and assume that $g = P(\partial)f$ in the distributional sense *locally* on Ω, i.e., for every $x \in \Omega$ there exists an open neighbourhood U_x such that $g|_{U_x} = P(\partial)(f|_{U_x})$ in the distributional sense. Show that $g = P(\partial)f$ in the distributional sense.

Hint: For $\varphi \in C_{\mathrm{c}}^\infty(\Omega)$, $K := \operatorname{spt} \varphi$ and a finite covering $(U_x)_{x\in F}$ of K by sets as above use a partition of unity as indicated in part (b).

4.4 Let $H \subseteq \mathbb{R}^2$ be the half-plane $H := \{(x_1, x_2);\ x_1 \geqslant 0\}$, and let $f \in L_{1,\mathrm{loc}}(\mathbb{R}^2)$ be defined by $f := \mathbf{1}_H$.

(a) Show that $\int \partial_1\varphi f \,\mathrm{d}x = -\int_{x_2\in\mathbb{R}} \varphi(0, x_2) \,\mathrm{d}x_2$ for all $\varphi \in C_\mathrm{c}^\infty(\mathbb{R}^2)$ and that there is no $g \in L_{1,\mathrm{loc}}(\mathbb{R}^2)$ such that $\int \partial_1\varphi f \,\mathrm{d}x = \int \varphi g \,\mathrm{d}x$ for all $\varphi \in C_\mathrm{c}^\infty(\mathbb{R}^2)$.

(b) Decide which of the partial derivatives $\partial_1 f, \partial_2 f, \partial^{(1,1)} f$ belong to $L_{1,\mathrm{loc}}(\mathbb{R}^2)$.

4.5 Let $\Omega \subseteq \mathbb{R}^n$ be open.

(a) Let $\alpha, \beta \in \mathbb{N}_0^n$ and $f, \partial^\alpha f \in L_{1,\mathrm{loc}}(\Omega)$. Show that

$$\partial^{\alpha+\beta} f \in L_{1,\mathrm{loc}}(\Omega) \iff \partial^\beta(\partial^\alpha f) \in L_{1,\mathrm{loc}}(\Omega) \implies \partial^{\alpha+\beta} f = \partial^\beta(\partial^\alpha f).$$

(b) Let $f \in L_{1,\mathrm{loc}}(\Omega)$, $\nabla f := (\partial_1 f, \dots, \partial_n f)^\top \in L_{1,\mathrm{loc}}(\Omega; \mathbb{K}^n)$. Show that

$$\Delta f \in L_{1,\mathrm{loc}}(\Omega) \iff \operatorname{div} \nabla f \in L_{1,\mathrm{loc}}(\Omega) \implies \Delta f = \operatorname{div} \nabla f.$$

(It is part of the task to give a definition of the distributional divergence $\operatorname{div} w \in L_{1,\mathrm{loc}}(\Omega)$ of a vector field $w \in L_{1,\mathrm{loc}}(\Omega; \mathbb{K}^n)$, analogously to the definition of distributional derivatives in Subsection 4.1.2.)

(c) Let $j \in \{1, \dots, n\}$, $f, \partial_j f \in L_{1,\mathrm{loc}}(\Omega)$, and let $\varphi \in C^\infty(\Omega)$. Show that $\partial_j(\varphi f) = \partial_j\varphi f + \varphi\partial_j f$.

4.6 Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $\alpha \in \mathbb{N}_0^n$ and $f, \partial^\alpha f \in L_{1,\mathrm{loc}}(\Omega)$.

(a) Show that $\operatorname{spt} \partial^\alpha f \subseteq \operatorname{spt} f$. (Hint: Look at the complements of the supports.)

(b) Assume additionally that $\operatorname{spt} f$ is compact in Ω, and define $\tilde f$ as the extension of f to $\mathbb{R}^n$ by zero. Show that $\partial^\alpha \tilde f \in L_1(\mathbb{R}^n)$ and that $\partial^\alpha \tilde f$ is the extension of $\partial^\alpha f$ to $\mathbb{R}^n$ by zero. (Hint: Using Exercise 4.2, choose a function $\psi \in C_\mathrm{c}^\infty(\mathbb{R}^n)$ such that $\operatorname{spt} \psi \subseteq \Omega$ and $\psi = 1$ in a neighbourhood of $\operatorname{spt} f$.)

4.7 Let $\Omega \subseteq \mathbb{R}^n$ be open.

(a) Let $f \in H_0^1(\Omega)$, and define $\tilde f$ as the extension of f to $\mathbb{R}^n$ by zero. Show that $\tilde f \in H^1(\mathbb{R}^n)$ and that $\partial_j \tilde f$ is the extension of $\partial_j f$ to $\mathbb{R}^n$ by zero, for $j = 1, \dots, n$. (Hint: First consider $f \in H_\mathrm{c}^1(\Omega)$ and apply Exercise 4.6(b).)

(b) Let $\Omega := (-1, 0) \cup (0, 1)$. Find a function $f \in H^1(\Omega) \setminus H_0^1(\Omega)$ with the property that the extension $\tilde f$ to $\mathbb{R}$ by zero belongs to $H^1(\mathbb{R})$.

4.8 Let $n \geqslant 3$. Show that $H^1(\mathbb{R}^n) = H_0^1(\mathbb{R}^n \setminus \{0\})$. For the more ambitious reader: Show this property for $n = 2$ as well.

4.9 Let $\Omega \subseteq \mathbb{R}^n$ be open. Denote by $A_\mathbb{K}$ the Dirichlet-Laplacian Δ_D in $L_2(\Omega; \mathbb{K})$ and by $T_\mathbb{K}$ the C_0-semigroup on $L_2(\Omega; \mathbb{K})$ generated by $A_\mathbb{K}$. Show that $T_\mathbb{C}(t)f = T_\mathbb{R}(t)f$ for all $t \geqslant 0$, $f \in L_2(\Omega; \mathbb{R})$.

Chapter 5

Forms and operators

We now introduce the main object of this book – namely forms in Hilbert spaces. They are very popular in analysis because, in combination with the Lax–Milgram lemma, they are best adapted for establishing existence and uniqueness of weak solutions of elliptic partial differential equations. Moreover, having the Lumer–Phillips machinery at our disposal, we can go much further and associate holomorphic semigroups with forms.

5.1 Forms: algebraic properties

In this section we introduce forms and put together some algebraic properties. As the domain we consider a vector space V over $\mathbb{K}$.

A **sesquilinear form** on V is a mapping $a\colon V \times V \to \mathbb{K}$ such that

$$a(u+v, w) = a(u, w) + a(v, w), \quad a(\lambda u, w) = \lambda a(u, w),$$
$$a(u, v+w) = a(u, v) + a(u, w), \quad a(u, \lambda v) = \overline{\lambda} a(u, v)$$

for all $u, v, w \in V$, $\lambda \in \mathbb{K}$. (It is common usage to call a a form on V, although a is defined on $V \times V$ – rather than on V.)

If $\mathbb{K} = \mathbb{R}$, then a sesquilinear form is the same as a bilinear form. If $\mathbb{K} = \mathbb{C}$, then a is linear in the first and antilinear in the second argument, i.e., only half of the linearity conditions are satisfied for the second argument: the form is $1\frac{1}{2}$-linear, or sesquilinear since the Latin 'sesqui' means 'one and a half'.

For simplicity we will mostly use the term **form** instead of sesquilinear form. A form a is called **symmetric** if

$$a(u, v) = \overline{a(v, u)} \qquad (u, v \in V),$$

and a is **accretive** if

$$\operatorname{Re} a(u, u) \geqslant 0 \qquad (u \in V).$$

In the literature, a symmetric form is also called "positive" if it is accretive, but we will not use this terminology.

We will frequently use the notation

$$a(u) := a(u, u) \qquad (u \in V)$$

for the associated quadratic form.

5.1 Remarks. (a) Each form a on V satisfies the **parallelogram identity**

$$a(u+v) + a(u-v) = 2a(u) + 2a(v) \qquad (u, v \in V).$$

If a is symmetric, one has

$$\operatorname{Re} a(u,v) = \frac{1}{4}\big(a(u+v) - a(u-v)\big) \qquad (u, v \in V).$$

(b) If $\mathbb{K} = \mathbb{C}$, then each form a satisfies the **polarisation identity**

$$a(u,v) = \frac{1}{4}\Big(a(u+v) - a(u-v) + \mathrm{i}\big(a(u+\mathrm{i}v) - a(u-\mathrm{i}v)\big)\Big) \qquad (u, v \in V),$$

as an elementary computation shows. In particular, each form is determined by its quadratic terms. This in turn implies that a is symmetric if and only if $a(u) \in \mathbb{R}$ for all $u \in V$. (Consider the form $b(u,v) := a(u,v) - \overline{a(v,u)}$ and its quadratic terms $b(u)$.) It follows that a is symmetric and accretive if and only if $a(u) \in [0, \infty)$ for all $u \in V$. Clearly, these two characterisations are only true if $\mathbb{K} = \mathbb{C}$.

(c) If $\mathbb{K} = \mathbb{R}$ and a is symmetric, then part (a) implies the **polarisation identity** in the variant

$$a(u,v) = \frac{1}{4}\big(a(u+v) - a(u-v)\big) \qquad (u, v \in V). \qquad \triangle$$

Now let again $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. We present a version of the Cauchy–Schwarz inequality that involves two different symmetric forms, which will be useful later on. Recall that a scalar product is a symmetric form a that is positive definite, i.e. $a(u) > 0$ for all $u \in V \setminus \{0\}$. For the Cauchy–Schwarz inequality to hold we do not need definiteness.

5.2 Proposition (Cauchy–Schwarz inequality). *Let $a, b\colon V \times V \to \mathbb{K}$ be two symmetric forms. Assume that $|a(u)| \leqslant b(u)$ for all $u \in V$. Then*

$$|a(u,v)| \leqslant b(u)^{1/2} b(v)^{1/2} \qquad (u, v \in V). \tag{5.1}$$

Proof. Let $u, v \in V$. In order to prove (5.1) we may assume that $a(u,v) \in \mathbb{R}$ (in the complex case replace u by γu with a suitable $\gamma \in \mathbb{C}$, $|\gamma| = 1$). Then from Remark 5.1(a) and the hypothesis one obtains

$$|a(u,v)| \leqslant \frac{1}{4}\big(|a(u+v)| + |a(u-v)|\big) \leqslant \frac{1}{4}\big(b(u+v) + b(u-v)\big) = \frac{1}{2}\big(b(u) + b(v)\big).$$

Let $s > b(u)^{1/2}$, $t > b(v)^{1/2}$. Then

$$|a(u,v)| = st\left|a\Big(\frac{1}{s}u, \frac{1}{t}v\Big)\right| \leqslant st \cdot \frac{1}{2}\Big(\frac{b(u)}{s^2} + \frac{b(v)}{t^2}\Big) \leqslant st.$$

Taking the infimum over s and t we obtain (5.1). $\square$

Next we introduce the adjoint form. Let $a\colon V\times V\to\mathbb{K}$ be a form. Then

$$a^*(u,v) := \overline{a(v,u)} \qquad (u,v\in V)$$

defines a form $a^*\colon V\times V\to\mathbb{K}$. Note that a is symmetric if and only if $a=a^*$. In the case of complex scalars, the forms

$$\operatorname{Re} a := \frac{1}{2}(a+a^*) \quad \text{and} \quad \operatorname{Im} a := \frac{1}{2\mathrm{i}}(a-a^*)$$

are symmetric, and

$$a = \operatorname{Re} a + \mathrm{i}\operatorname{Im} a.$$

We call $\operatorname{Re} a$ the **real part** and $\operatorname{Im} a$ the **imaginary part** of a. Observe that $(\operatorname{Re} a)(u) = \operatorname{Re} a(u)$ and $(\operatorname{Im} a)(u) = \operatorname{Im} a(u)$ for all $u\in V$.

There is another algebraic notion – only used for the case $\mathbb{K}=\mathbb{C}$ – that will play an important role. A form $a\colon V\times V\to\mathbb{C}$ is **sectorial** if there exists $\theta\in[0,\pi/2)$ such that $a(u)\in\overline{\Sigma_\theta}$ for all $u\in V$. If we want to specify the angle, we say that a is **sectorial of angle** θ. It is obvious that a form $a\colon V\times V\to\mathbb{C}$ is sectorial if and only if a is accretive and there exists a constant $c\geqslant 0$ such that

$$|\operatorname{Im} a(u)| \leqslant c\operatorname{Re} a(u) \qquad (u\in V). \tag{5.2}$$

(The angle θ and the constant c are related by $c=\tan\theta$. Note that without accretivity of a, the estimate (5.2) with $c=0$ would just imply that $a(u)\in\mathbb{R}$ for all $u\in V$.)

5.2 Representation theorems

We now consider forms whose domain is a Hilbert space V over $\mathbb{K}$. We recall the classical representation theorem of Fréchet–Riesz: if η is a bounded linear functional on V, then there exists a unique $v\in V$ such that

$$\eta(u) = (u\,|\,v)_V \qquad (u\in V).$$

The purpose of this section is to generalise this result.

First of all, it will be natural to use the antidual V^* of V instead of the dual space V'. More precisely, if $\mathbb{K}=\mathbb{R}$, then $V^*=V'$ is the dual space of V, and if $\mathbb{K}=\mathbb{C}$, then we denote by V^* the space of all continuous antilinear functionals. Note that η is antilinear if and only if $\bar\eta = \big[u\mapsto\overline{\eta(u)}\,\big]$ is linear. The space V^* is a Banach space for the norm $\|\eta\|_{V^*} = \sup_{\|v\|_V\leqslant 1}|\eta(v)|$. For $\eta\in V^*$ we will mostly write

$$\langle\eta,v\rangle := \eta(v) \qquad (v\in V).$$

The theorem of Fréchet–Riesz can be reformulated by saying that for each $\eta\in V^*$ there exists a unique $u\in V$ such that

$$\eta(v) = (u\,|\,v)_V \qquad (v\in V).$$

The mapping $\Phi\colon V\to V^*,\ u\mapsto(u\,|\,\cdot)_V$ is the **Fréchet–Riesz isomorphism**: it is easy to see that Φ is linear and isometric; the Fréchet–Riesz theorem shows that Φ is surjective.

We now aim for the generalisation of the Fréchet–Riesz theorem, the omnipresent Lax–Milgram lemma.

A form $a\colon V \times V \to \mathbb{K}$ is called **bounded** if there exists $M \geqslant 0$ such that

$$|a(u,v)| \leqslant M\|u\|_V\|v\|_V \qquad (u,v \in V).$$

This condition is equivalent to continuity of a; see Exercise 5.1. If a is a bounded form on V, then $a(u,\cdot) \in V^*$ for all $u \in V$, with $\|a(u,\cdot)\| \leqslant M\|u\|_V$, so we obtain a bounded linear operator $\mathcal{A}\colon V \to V^*$, $u \mapsto a(u,\cdot)$ satisfying

$$\langle \mathcal{A}u, v\rangle = a(u,v) \qquad (u,v \in V) \tag{5.3}$$

and $\|\mathcal{A}\|_{\mathcal{L}(V,V^*)} \leqslant M$. The operator $\mathcal{A}$ is called the **Lax–Milgram operator** associated with the form a. If, in addition, a is **coercive**, i.e. if there exists $\alpha > 0$ such that

$$\operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2 \qquad (u \in V), \tag{5.4}$$

then $\mathcal{A}$ is an isomorphism; this is the famous Lax–Milgram lemma shown below.

We first treat the 'operator version' of the lemma, which is based on the following definition. An operator A in V is called **strictly accretive** if there exists $\alpha > 0$ such that

$$\operatorname{Re}(Au \mid u)_V \geqslant \alpha\|u\|_V^2 \qquad (u \in \operatorname{dom}(A)) \tag{5.5}$$

(expressed differently, $A - \alpha I$ is accretive), and A is **strictly m-accretive** if A is strictly accretive and m-accretive. Lemma 3.17 implies that A is strictly m-accretive if and only if A is strictly accretive and $\operatorname{ran}(A) = V$.

5.3 Lemma. *Let V be a Hilbert space, and let $A \in \mathcal{L}(V)$ be strictly accretive. Then A is invertible in $\mathcal{L}(V)$, and $\|A^{-1}\| \leqslant \frac{1}{\alpha}$, with $\alpha > 0$ as in* (5.5).

Proof. The operator $A - \alpha I$ is accretive and hence m-accretive by Remark 3.18. Thus Lemma 3.17 implies that $A = \alpha I + (A - \alpha I)$ has the asserted properties. □

5.4 Theorem (Lax–Milgram lemma). *Let V be a Hilbert space, and let $a\colon V \times V \to \mathbb{K}$ be a bounded coercive form. Then the operator $\mathcal{A}\colon V \to V^*$ defined in* (5.3) *is an isomorphism, and $\|\mathcal{A}^{-1}\|_{\mathcal{L}(V^*,V)} \leqslant \frac{1}{\alpha}$ with $\alpha > 0$ from* (5.4). *In particular, for every $\eta \in V^*$ there exists $u \in V$ such that $\eta(v) = a(u,v)$ for all $v \in V$.*

Proof. Composing $\mathcal{A}$ with the inverse of the Fréchet–Riesz isomorphism $\Phi\colon V \to V^*$ we obtain an operator $\Phi^{-1}\mathcal{A} \in \mathcal{L}(V)$ satisfying

$$\operatorname{Re}\bigl(\Phi^{-1}\mathcal{A}u \bigm| u\bigr)_V = \operatorname{Re}\langle \mathcal{A}u, u\rangle = \operatorname{Re} a(u,u) \geqslant \alpha\|u\|_V^2 \qquad (u \in V).$$

From Lemma 5.3 we conclude that $\Phi^{-1}\mathcal{A}$ is invertible in $\mathcal{L}(V)$ and that $\|(\Phi^{-1}\mathcal{A})^{-1}\| \leqslant \frac{1}{\alpha}$. As Φ is an isometric isomorphism we obtain the assertions. □

If the form is symmetric, then the Lax–Milgram lemma is the same as the theorem of Fréchet–Riesz. In fact, then a is an **equivalent scalar product**, i.e. $a(u)^{1/2}$ defines an equivalent norm on V.

5.3 Semigroups by forms

Here we come to the heart of this book: we prove the first generation theorem. With a bounded coercive form we associate an operator that is strictly m-accretive and thus yields a contractive C_0-semigroup, by the Lumer–Phillips theorem, Theorem 3.16. In the complex case the associated operator is m-sectorial and thus yields a contractive holomorphic C_0-semigroup, by Theorem 3.20. Recall that m-accretivity and m-sectoriality of an operator A involve the condition $\operatorname{ran}(I + A) = H$. In applications, proving this range condition often amounts to solving a partial differential equation (usually of elliptic type); see Section 4.2, for instance. If the operator is associated with a form, then the Lax–Milgram lemma does this job, so the range condition is automatically satisfied.

The setup we describe below is sometimes referred to as the "complete case" in the literature because the form domain V is a Hilbert space. Once we have presented a series of examples in the further development, we will also introduce – in Chapter 12 – the "incomplete case" in which the form domain is just a vector space.

First we will explain how we associate an operator with a form. Let V, H be Hilbert spaces over $\mathbb{K}$, and let $a\colon V \times V \to \mathbb{K}$ be a bounded form. Let $j \in \mathcal{L}(V, H)$ be an operator with dense range. We define

$$A := \big\{(x, y) \in H \times H\,;\ \exists u \in V\colon j(u) = x,\ a(u, v) = (y \,|\, j(v))\ (v \in V)\big\}; \tag{5.6}$$

here and in what follows, scalar products without index are scalar products in H.

5.5 Proposition. *Assume that*

$$u \in \ker(j),\ a(u) = 0 \quad \textit{implies} \quad u = 0. \tag{5.7}$$

(a) *Then the relation A defined above is a linear operator in H. We call A the* ***operator associated with*** *(a, j) and write $A \sim (a, j)$.*

(b) *If a is accretive, then A is accretive.*

(c) *If $\mathbb{K} = \mathbb{C}$ and a is sectorial, then A is sectorial of the same angle as a.*

Proof. (a) It is easy to see that A is a subspace of $H \times H$. Let $(0, y) \in A$; we have to show that $y = 0$. By definition there exists $u \in V$ such that $j(u) = 0$ and $a(u, v) = (y \,|\, j(v))$ for all $v \in V$. In particular, $a(u) = (y \,|\, j(u)) = 0$. Assumption (5.7) implies that $u = 0$. Hence $(y \,|\, j(v)) = 0$ for all $v \in V$. Since j has dense range, it follows that $y = 0$.

(b), (c) If $x \in \operatorname{dom}(A)$, then there exists $u \in V$ such that $j(u) = x$ and $a(u, v) = (Ax \,|\, j(v))$ for all $v \in V$, in particular $a(u) = (Ax \,|\, j(u)) = (Ax \,|\, x)$.

Thus we have shown that $\{(Ax \,|\, x)\,;\ x \in \operatorname{dom}(A)\} \subseteq \{a(u)\,;\ u \in V\}$, and this proves (b) as well as (c). □

We now prove the first generation theorem for forms. Note that coercivity implies (5.7).

5.6 Theorem (Generation theorem, part 1). *Let $a\colon V \times V \to \mathbb{K}$ be bounded coercive form, and let $j \in \mathcal{L}(V, H)$ have dense range. Let A be the operator associated with (a, j). Then A is strictly m-accretive; in particular, $-A$ generates a contractive C_0-semigroup on H.*

Proof. By hypothesis, there exists $\alpha > 0$ such that $\operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2$ for all $u \in V$. Also, there exists $c > 0$ such that $\|j(u)\|_H \leqslant c\|u\|_V$ for all $u \in V$. Let $x \in \operatorname{dom}(A)$. From the definition of A we know that there exists $u \in V$ such that $j(u) = x$ and $a(u,u) = (Ax \,|\, j(u)) = (Ax \,|\, x)$; hence

$$\operatorname{Re}(Ax \,|\, x) = \operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2 \geqslant \frac{\alpha}{c^2}\|x\|_H^2 .$$

This shows that A is strictly accretive.

For the proof of strict m-accretivity we show that $\operatorname{ran}(A) = H$. Let $y \in H$. Then $\eta(v) := (y \,|\, j(v))$ defines a functional $\eta \in V^*$. By the Lax–Milgram lemma, Theorem 5.4, there exists $u \in V$ such that

$$a(u,v) = \eta(v) = (y \,|\, j(v)) \qquad (v \in V).$$

Therefore $x := j(u) \in \operatorname{dom}(A)$ and $y = Ax \in \operatorname{ran}(A)$.

Now the Lumer–Phillips theorem, Theorem 3.16, implies that $-A$ generates a contractive C_0-semigroup. □

We point out that, in the situation of Theorem 5.6, there exists $\varepsilon > 0$ such that $\|T(t)\| \leqslant \mathrm{e}^{-\varepsilon t}$ for all $t \geqslant 0$; see Exercise 5.2(b).

In an implicit way, the proof of Theorem 5.6 contains a formula for A^{-1}. This formula will be helpful in the proof of Proposition 6.18 and very important in the proofs of Theorems 14.17 and 15.2. For this reason we state it explicitly, as follows.

5.7 Proposition. *Let the hypotheses be as in Theorem 5.6. Let $\mathcal{A}\colon V \to V^*$ be the Lax–Milgram operator associated with the form a as in* (5.3)*, and define $k\colon H \to V^*$, $y \mapsto (y \,|\, j(\cdot))$. Then $k \in \mathcal{L}(H,V^*)$, and*

$$A^{-1} = j\mathcal{A}^{-1}k.$$

(Recall from Theorem 5.4 that $\mathcal{A}$ is invertible with inverse in $\mathcal{L}(V^*,V)$.)

Proof. The boundedness of k is seen from

$$|\langle k(y), v\rangle| \leqslant \|y\|_H\|j(v)\|_H \leqslant \|j\|\|y\|_H\|v\|_V \qquad (y \in H,\ v \in V).$$

We return to the notation used in the proof of Theorem 5.6. Starting with $y \in H$ we obtain $\eta = k(y)$, $u = \mathcal{A}^{-1}\eta$, $x = j(u)$. This results in $x = j\mathcal{A}^{-1}k(y)$, and using $Ax = y$ and the invertibility of A we conclude that $A^{-1} = j\mathcal{A}^{-1}k$. □

In the complex case one also obtains results concerning sectoriality.

5.8 Theorem (Generation theorem, part 2). *Let $\mathbb{K} = \mathbb{C}$, let $a\colon V \times V \to \mathbb{C}$ be a bounded coercive form, and let $j \in \mathcal{L}(V,H)$ have dense range. Let A be the operator associated with (a,j). Then the form a is sectorial, and the operator A is m-sectorial, i.e., $-A$ generates a contractive holomorphic C_0-semigroup on H.*

Proof. By assumption there exist $M \geqslant 0$, $\alpha > 0$ such that

$$|a(u,v)| \leqslant M\|u\|_V\|v\|_V, \quad \operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2$$

for all $u, v \in V$. Thus

$$\frac{|\operatorname{Im} a(u)|}{\operatorname{Re} a(u)} \leqslant \frac{M\|u\|_V^2}{\alpha\|u\|_V^2} = \frac{M}{\alpha}$$

for all $u \in V \setminus \{0\}$. This implies that there exists $\theta \in [0, \pi/2)$ such that $a(u) \in \overline{\Sigma_\theta}$ for all $u \in V$, i.e. a is sectorial. The remaining assertions are immediate consequences of Proposition 5.5(c), Theorem 5.6 and Theorem 3.20. □

We give a first example as an illustration.

5.9 Example (Multiplication operators). Let (Ω, μ) be a measure space, and let $m\colon \Omega \to \mathbb{K}$ be measurable. Assume that there exist $\delta, c > 0$ such that

$$w(x) := \operatorname{Re} m(x) \geqslant \delta, \quad |m(x)| \leqslant c \operatorname{Re} m(x)$$

for all $x \in \Omega$. Let $V := L_2(\Omega, w\mu)$. Then $a(u,v) := \int u\bar{v}m\,\mathrm{d}\mu$ defines a bounded coercive form a on V. Let $H := L_2(\Omega, \mu)$, $j(u) := u$ for all $u \in V$. Then $j \in \mathcal{L}(V, H)$ because $w \geqslant \delta$, and $\operatorname{ran}(j) = V$ is dense in H since V is the domain of the maximal multiplication operator induced by the function $\sqrt{w}$; see Exercise 1.6(a). Let $A \sim (a, j)$. Then A is the maximal multiplication operator induced by m, i.e.

$$\begin{aligned} \operatorname{dom}(A) &= \{u \in L_2(\Omega, \mu)\,;\ mu \in L_2(\Omega, \mu)\}, \\ Au &= mu. \end{aligned}$$

Proof. If $u \in L_2(\Omega, \mu)$ is such that $mu \in L_2(\Omega, \mu)$, then $u \in V$ and

$$a(u,v) = \int u\bar{v}m\,\mathrm{d}\mu = (mu \,|\, v) \qquad (v \in V);$$

hence $u \in \operatorname{dom}(A)$ and $Au = mu$. On the other hand, if $u \in \operatorname{dom}(A)$, then $f := Au \in L_2(\Omega, \mu)$ satisfies

$$\int u\bar{v}m\,\mathrm{d}\mu = a(u,v) = (f \,|\, v) = \int f\bar{v}\,\mathrm{d}\mu$$

for all $v \in V$. In particular, given $n \in \mathbb{N}$, this equality holds for all $v \in L_2(\Omega, \mu)$ with $[v \neq 0] \subseteq [|m| \leqslant n]$, and it follows that $mu = f$ a.e. on $[|m| \leqslant n]$. Hence from $\Omega = \bigcup_{n\in\mathbb{N}} [|m| \leqslant n]$ we obtain $mu = f$ a.e. □

In many applications the coercivity (accretivity, sectoriality) assumption in the above results can only be reached by 'shifting' the form a. From Section 2.2 we recall the concept of rescaling a C_0-semigroup, which corresponds to 'shifting' the generator: if $-A$ is the generator of a C_0-semigroup T and $\omega \in \mathbb{R}$, then $-(A + \omega)$ generates the semigroup $T_\omega = (\mathrm{e}^{-\omega t}T(t))_{t\geqslant 0}$. (The notation '$A + \omega$' is an abbreviation of '$A + \omega I$'; the ω stands for multiplication by the scalar ω, which is just the operator ωI.)

One frequently uses the word "quasi" as prefix if something is true after rescaling. An operator A in a Hilbert space H is **quasi-accretive** or **quasi-m-accretive** if there exists $\omega \in \mathbb{R}$ such that $A+\omega$ is accretive or m-accretive, respectively. By Theorem 3.16 (Lumer–Phillips), an operator A is quasi-m-accretive if and only if $-A$ is the generator of a quasi-contractive C_0-semigroup.

In the case of a complex Hilbert space H, an operator A in H is **quasi-sectorial** or **quasi-m-sectorial** if there exists $\omega \in \mathbb{R}$ such that $A+\omega$ is sectorial or m-sectorial, respectively. A **quasi-contractive holomorphic semigroup** is a holomorphic semigroup T such that $\|\mathrm{e}^{-\omega z}T(z)\| \leqslant 1$ for all $z \in \Sigma_\theta$, for some $\theta \in (0,\pi/2]$ and some $\omega \in \mathbb{R}$. Thus A is quasi-m-sectorial if and only if $-A$ generates a quasi-contractive holomorphic C_0-semigroup.

Next we explain how to shift on the level of forms.

5.10 Remark. Let $a\colon V\times V \to \mathbb{K}$ be a bounded form, and let $j \in \mathcal{L}(V,H)$ have dense range. Let $\omega \in \mathbb{R}$, and define the shifted form a_ω on V by

$$a_\omega(u,v) := a(u,v) + \omega(j(u)\,|\,j(v)) \qquad (u,v \in V). \tag{5.8}$$

Assume that (5.7) is satisfied. Then clearly the form a_ω satisfies (5.7) as well (with a replaced by a_ω). We show that $A+\omega$ is the operator associated with (a_ω, j).

Let A_ω be the operator associated with (a_ω, j). Let $x, y \in H$. Then for all $u, v \in V$ with $j(u) = x$ we have

$$a_\omega(u,v) = (y\,|\,j(v)) \iff a(u,v) = (y-\omega x\,|\,j(v)).$$

This shows that

$$(x,y) \in A_\omega \iff (x, y-\omega x) \in A \iff (x,y) \in A+\omega I. \qquad \triangle$$

Let $a\colon V\times V \to \mathbb{K}$ be a bounded form, and let $j \in \mathcal{L}(V,H)$ have dense range. In analogy to our terminology for operators we could say that a is *quasi-coercive* (*with respect to* j) if one of the forms a_ω is coercive, i.e., if there exist $\omega \in \mathbb{R}$, $\alpha > 0$ such that

$$\operatorname{Re} a(u) + \omega\|j(u)\|_H^2 \geqslant \alpha\|u\|_V^2 \qquad (u \in V); \tag{5.9}$$

for simplicity we prefer to call the form a ***j*-coercive** in this case. (This property has been introduced in [ArEl12b] under the name 'j-elliptic'.) It is obvious that (5.9) implies (5.7); thus Proposition 5.5 is applicable to j-coercive forms.

From Theorems 5.6 and 5.8 we now obtain the following more general generation result; we refer to Exercises 5.3 and 5.4 for a discussion of whether the hypotheses are necessary.

5.11 Corollary. *Let $j \in \mathcal{L}(V,H)$ have dense range, let $a\colon V\times V \to \mathbb{C}$ be a bounded j-coercive form, and let A be the operator associated with (a,j).*

(a) *Then A is quasi-m-accretive. If additionally a is accretive, then A is m-accretive.*

(b) *If $\mathbb{K} = \mathbb{C}$, then A is quasi-m-sectorial. If additionally a is sectorial, then A is m-sectorial of the same angle as a.*

Proof. Define the form a_ω on V by (5.8), with ω as in (5.9). Then a_ω is coercive, and $A+\omega \sim (a_\omega, j)$, by Remark 5.10. Hence the operator $A+\omega$ is m-accretive (by Theorem 5.6), and in the case $\mathbb{K} = \mathbb{C}$ it is m-sectorial (by Theorem 5.8). This proves the first assertions of (a) and (b).

For the second assertion of (a) we recall Proposition 5.5(b), which shows that A is accretive. Applying Lemma 3.17 one concludes that A is m-accretive. For the second assertion of (b) one argues analogously, recalling Proposition 5.5(c). □

In Corollary 5.11, the operator $-A$ generates a holomorphic C_0-semigroup on H if $\mathbb{K} = \mathbb{C}$. (If $\mathbb{K} = \mathbb{R}$, then $-A$ generates a C_0-semigroup T on H that is analytic on $(0,\infty)$. This can be proved by a complexification procedure; see for instance [Ouh05; Remark following Theorem 1.54].)

5.12 Remarks. (a) Note that operators associated with forms as described above are always quasi-m-accretive, and hence the associated C_0-semigroups are always quasi-contractive.

(b) We point out that in the definition of the operator associated with a form the specific scalar product of the Hilbert space enters decisively. Changing it to an equivalent scalar product one obtains a different associated operator, which need not be quasi-accretive in the original scalar product. For an example illustrating this issue we refer to Exercise 5.9.

(c) It is important to keep in mind that not all quasi-m-accrective operators are associated with a form as in Corollary 5.11. We illustrate this with the operator A of differentiation in $H := L_2(\mathbb{R})$, given by $\operatorname{dom}(A) := H^1(\mathbb{R})$, $Af := f'$. Recall from Example 4.18(a) that $-A$ is the generator of the (contractive!) C_0-group T of right translations, hence A is m-accretive.

In the case $\mathbb{K} = \mathbb{C}$, an operator associated with a form as in Corollary 5.11(b) is always minus the generator of a holomorphic semigroup. However, the C_0-semigroup T is not holomorphic. For the case $\mathbb{K} = \mathbb{R}$ we refer to Exercise 5.10(c). △

Later – in Chapter 8 – we will meet interesting situations in which j is not injective. In most applications, however, j is an embedding; then we call (a, j) an **embedded form**, and we will usually suppress the letter j.

We conclude this section with a brief description of the embedded case. Let V and H be Hilbert spaces, $V \overset{d}{\hookrightarrow} H$. This is an abbreviation for saying that V is continuously embedded into H (see Remark 4.13(a)) and that V is dense in H. Now let $a\colon V \times V \to \mathbb{K}$ be a bounded form. We say that a is **quasi-coercive** if

$$\operatorname{Re} a(u) + \omega \|u\|_H^2 \geqslant \alpha \|u\|_V^2$$

for all $u \in V$ and some $\alpha > 0$, $\omega \in \mathbb{R}$. (This means that, in the previous terminology, the form a is j-coercive, where $j\colon V \hookrightarrow H$ is the embedding.) In this case the definition of the operator A associated with a (abbreviated by $A \sim a$, without mentioning the given embedding of V into H) reads as follows. For $x, y \in H$ one has

$$x \in \operatorname{dom}(A),\ Ax = y \iff x \in V,\ a(x, v) = (y \mid v)_H \ (v \in V).$$

The operator A is quasi-m-accretive by Corollary 5.11, and quasi-m-sectorial if $\mathbb{K} = \mathbb{C}$.

In the literature, quasi-coercive forms are sometimes called H-elliptic or just elliptic.

5.4 The classical Dirichlet form

Let Ω be an open subset of $\mathbb{R}^n$. The **classical Dirichlet form** on $H_0^1(\Omega)$ is defined by

$$a(u,v) := \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x = \sum_{j=1}^n \int_\Omega \partial_j u\,\overline{\partial_j v}\,\mathrm{d}x \qquad (u,v \in H_0^1(\Omega)),$$

where $\xi \cdot \eta := \sum_{j=1}^n \xi_j \eta_j$ for $\xi, \eta \in \mathbb{K}^n$ (which for $\xi,\eta \in \mathbb{R}^n$ is the standard scalar product). It is clear that a is bounded; in fact

$$|a(u,v)| \leqslant \|\nabla u\|_2 \|\nabla v\|_2 \leqslant \|u\|_{H^1}\|v\|_{H^1}\,.$$

Here $\nabla u = (\partial_1 u, \ldots, \partial_n u)^\top$ is the **gradient** of u, and $\|\nabla u\|_2 := \bigl(\sum_{j=1}^n \int_\Omega |\partial_j u|^2\,\mathrm{d}x\bigr)^{1/2}$; thus $\|u\|_{H^1}^2 = \|u\|_2^2 + \|\nabla u\|_2^2$.

We will prove that the Dirichlet form is coercive if Ω is bounded, or more generally, if Ω is **contained in a slab**, i.e., there exist $\beta > 0$ and $j_0 \in \{1,\ldots,n\}$ such that $|x_{j_0}| \leqslant \beta$ for all $x \in \Omega$.

5.13 Theorem (Poincaré's inequality). *Assume that Ω is contained in a slab. Then there exists a constant $c_{\mathrm{P}} > 0$ such that*

$$\int_\Omega |u(x)|^2\,\mathrm{d}x \leqslant c_{\mathrm{P}} \int_\Omega |\nabla u(x)|^2\,\mathrm{d}x \qquad (u \in H_0^1(\Omega)).$$

Proof. By Theorem 4.15(b) it suffices to prove the inequality for all $u \in C_c^\infty(\Omega)$. Let $\beta > 0$ and $j_0 \in \{1,\ldots,n\}$ be such that $|x_{j_0}| \leqslant \beta$ for all $x = (x_1,\ldots,x_n) \in \Omega$. We may assume that $j_0 = 1$; otherwise we permute the coordinates. Let $h \in C^1[-\beta,\beta]$, $h(-\beta) = 0$. Then by Hölder's inequality we estimate

$$\begin{aligned}\int_{-\beta}^{\beta} |h(x)|^2\,\mathrm{d}x &= \int_{-\beta}^{\beta} \left|\int_{-\beta}^{x} h'(y)\,\mathrm{d}y\right|^2 \mathrm{d}x \leqslant \int_{-\beta}^{\beta} \left(\int_{-\beta}^{x} |h'(y)|^2\,\mathrm{d}y\right)\left(\int_{-\beta}^{x} 1\,\mathrm{d}y\right)\mathrm{d}x \\ &\leqslant (2\beta)^2 \int_{-\beta}^{\beta} |h'(y)|^2\,\mathrm{d}y.\end{aligned}$$

Now let $u \in C_c^\infty(\mathbb{R}^n)$, $\operatorname{spt} u \subseteq \Omega$. Applying the above estimate to $h(r) := u(r, x_2, \ldots, x_n)$ we obtain

$$\int_\Omega |u(x)|^2\,\mathrm{d}x \leqslant 4\beta^2 \int_\mathbb{R} \cdots \int_\mathbb{R} \int_{-\beta}^{\beta} |\partial_1 u(x_1,\ldots,x_n)|^2\,\mathrm{d}x_1 \cdots \mathrm{d}x_n \leqslant 4\beta^2 \int_\Omega |\nabla u(x)|^2\,\mathrm{d}x. \quad \square$$

In fact, we saw that the constant c_{P} in Theorem 5.13 can be chosen as d^2, where $d := 2\beta$ is an upper estimate for the width of Ω. (For bounded domains the best constant can be determined as $c_{\mathrm{P}} = 1/\lambda_1^{\mathrm{D}}$, where λ_1^{D} is the first eigenvalue of $-\Delta_{\mathrm{D}}$; we will come back to this fact in Example 6.19.) Next we revisit the Dirichlet Laplacian.

5.14 Example (Dirichlet Laplacian). Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $H := L_2(\Omega)$, $V := H_0^1(\Omega)$ and observe that $V \overset{\mathrm{d}}{\hookrightarrow} H$. As before, define $a\colon V \times V \to \mathbb{K}$ by

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

Then a is bounded and quasi-coercive. The operator A in H associated with a is given by

$$\begin{aligned}\operatorname{dom}(A) &= \bigl\{u \in H_0^1(\Omega);\ \Delta u \in L_2(\Omega)\bigr\},\\ Au &= -\Delta u,\end{aligned}$$

i.e. $-A$ is the Dirichlet Laplacian Δ_{D} defined in Subsection 4.2.2.

The form a is symmetric and accretive. If Ω is contained in a slab, then a is coercive.

Proof. The inequality $a(u) + 1\|u\|_H^2 \geqslant \|u\|_V^2$ $(u \in V)$ – in fact an equality – shows that a is quasi-coercive. For $u, f \in L_2(\Omega)$ one has $u \in \operatorname{dom}(A)$, $Au = f$ if and only if $u \in H_0^1(\Omega)$ and

$$\int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x = \int_\Omega f\overline{v}\,\mathrm{d}x \qquad (v \in H_0^1(\Omega)).$$

By Lemma 4.20 the latter is equivalent to $-\Delta u = f$ in the distributional sense.

Symmetry and accretivity of a are obvious. Now assume that Ω is contained in a slab. Let $u \in H_0^1(\Omega)$. Then $a(u) \geqslant \frac{1}{c_{\mathrm{P}}}\int_\Omega |u|^2\,\mathrm{d}x$ by Poincaré's inequality, and thus $a(u) \geqslant \frac12\int_\Omega |\nabla u|^2\,\mathrm{d}x + \frac{1}{2c_{\mathrm{P}}}\int_\Omega |u|^2\,\mathrm{d}x \geqslant \alpha\|u\|_{H^1}^2$, where $\alpha = \min\{\frac12, \frac{1}{2c_{\mathrm{P}}}\}$. Hence a is coercive. □

Finally we give an example in which j is not a canonical embedding.

5.15 Example (Multiplicative perturbation of Δ_{D}). Let $\Omega \subseteq \mathbb{R}^n$ be an open set that is contained in a slab. Let $m\colon \Omega \to \mathbb{C}$ be measurable, $|m(x)| \geqslant \varepsilon > 0$ for all $x \in \Omega$. Define the operator A in $L_2(\Omega)$ by

$$\begin{aligned}\operatorname{dom}(A) &= \bigl\{u \in L_2(\Omega);\ mu \in \operatorname{dom}(\Delta_{\mathrm{D}}),\ \overline{m}\Delta(mu) \in L_2(\Omega)\bigr\},\\ Au &= -\overline{m}\Delta(mu).\end{aligned}$$

Then A is m-sectorial of angle 0.

Proof. Let $H := L_2(\Omega)$, $V := H_0^1(\Omega)$, $a(u,v) := \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x$, and let $j \in \mathcal{L}(V,H)$ be given by $j(v) := \frac{1}{m}v$. Then j has dense range. Indeed, if $g \in j(V)^\perp$, then $\int_\Omega \frac{1}{m}v\overline{g}\,\mathrm{d}x = 0$ for all $v \in C_{\mathrm{c}}^\infty(\Omega)$, and hence $\frac{1}{m}\overline{g} = 0$ by Lemma 4.6. It follows that $j(V)^\perp = \{0\}$, which implies that $j(V)$ is dense in $L_2(\Omega)$.

Clearly a is sectorial of angle 0, so the operator $B \sim (a,j)$ is m-sectorial of angle 0, by Corollary 5.11(b). We show that $A = B$. For $u, f \in L_2(\Omega)$ one has $u \in \operatorname{dom}(B)$ and $Bu = f$ if and only if there exists $w \in H_0^1(\Omega)$ such that $\frac{w}{m} = u$ and $\int_\Omega \nabla w \cdot \overline{\nabla v}\,\mathrm{d}x = \int_\Omega f\,\overline{(\frac{v}{m})}\,\mathrm{d}x$ for all $v \in H_0^1(\Omega)$. This in turn is equivalent to $mu \in \operatorname{dom}(\Delta_{\mathrm{D}})$ and $-\Delta(mu) = \frac{f}{\overline{m}}$, i.e. $u \in \operatorname{dom}(A)$ and $Au = f$. □

Notes

The approach to forms presented here is the "French approach" following Lions [DaLi92]. As a new ingredient we have introduced the mapping j, following the paper [ArEl12b] by Arendt and ter Elst (see also [ArEl12a]). It will bear fruit when we investigate the Dirichlet-to-Neumann operator (in Chapter 8) and also in Chapter 12, where we consider what is called the "incomplete case" in [ArEl12b] and [ArEl12a].

The Lax–Milgram lemma was proved in 1954 [LaMi54; Theorem 2.1] and has been *the* standard tool for establishing weak solutions since then.

As an interesting bit of history, we note that Hilbert used bilinear forms to treat integral equations in his famous papers from the beginning of the 20th century. His ideas led his students to develop the notion of operators in Hilbert spaces. As a result, operators have taken a more central role, and physical and other problems are formulated with the help of operators. In the 1950's, form methods were developed to solve equations defined by operators. Forms are most appropriate for numerical treatments because a form $a\colon V \times V \to \mathbb{C}$ can easily be restricted to a finite-dimensional subspace $V_m \times V_m$, whereas for operators there might be only few invariant subspaces. The method of finite elements is based on such restrictions.

If a form is symmetric, then solving the associated inhomogeneous equation $\mathcal{A}u = L$ (see below for the notation) is equivalent to a minimisation problem. We want to explain this in more detail. Let V be a real Hilbert space and $a\colon V \times V \to \mathbb{R}$ a symmetric bounded coercive form. Consider the associated Lax–Milgram operator $\mathcal{A}\colon V \to V^*$, given by

$$\langle \mathcal{A}u, v\rangle = a(u,v).$$

Given $L \in V^*$, the Lax–Milgram lemma says that there is a unique $u \in V$ such that $\mathcal{A}u = L$. It turns out that u is the unique element of V minimising the function

$$V \ni w \mapsto \tfrac{1}{2}a(w) - \langle L, w\rangle.$$

Indeed, let $v \in V$. Then

$$\begin{aligned}\tfrac{1}{2}a(u+v) - \langle L, u+v\rangle &= \tfrac{1}{2}a(u) + a(u,v) + \tfrac{1}{2}a(v) - \langle L, u\rangle - \langle L, v\rangle \\ &= \tfrac{1}{2}a(u) + \tfrac{1}{2}a(v) - \langle L, u\rangle > \tfrac{1}{2}a(u) - \langle L, u\rangle\end{aligned}$$

unless $v = 0$.

Because of this argument, form methods are also called 'variational methods': we look at the 'variations' $a(u+v)$ of $a(u)$. It was Johann Peter Gustav Lejeune Dirichlet (1805–1859) who proposed such a variational method for the solution of the Dirichlet problem. We refer to the Notes of Chapter 7 for more information.

Exercises

5.1 (a) Let V be a Hilbert space, and let $a\colon V \times V \to \mathbb{K}$ be a sesquilinear form. Show that the properties

(i) a is bounded,
(ii) there exists $A \in \mathcal{L}(V)$ such that $a(u,v) = (Au \,|\, v)$ for all $u, v \in V$,
(iii) a is continuous

are equivalent.

(b) Show that one has the additional equivalence

(iv) there exists $M \geqslant 0$ such that $|a(u)| \leqslant M\|u\|_V^2$ for all $u \in V$

if a is symmetric *or* $\mathbb{K} = \mathbb{C}$.

Hint: Use the Cauchy–Schwarz inequality, Proposition 5.2, and in the complex case use the decomposition $a = \operatorname{Re} a + \mathrm{i} \operatorname{Im} a$.

(c) Find an unbounded form $a\colon V \times V \to \mathbb{R}$ such that $a(u) = 0$ for all $u \in V$, with the real Hilbert space $V := \ell_2(\mathbb{N}; \mathbb{R}^2)$. (This will show that condition (iv) is not equivalent to (i), (ii), (iii) without the additional requirement in part (b).)

Hint: First find a suitable definition of a on the space $\ell_{2,\mathrm{c}}(\mathbb{N}; \mathbb{R}^2) \times \ell_{2,\mathrm{c}}(\mathbb{N}; \mathbb{R}^2)$, where the index 'c' means that in $(u_n) \in \ell_{2,\mathrm{c}}(\mathbb{N}; \mathbb{R}^2)$ only finitely many components are $\neq 0$. In order to achieve the extension of a to $V \times V$, supplement the canonical basis $\check{B}$ of $\ell_{2,\mathrm{c}}(\mathbb{N}; \mathbb{R}^2)$ to an (algebraic) basis B of V and define $a(b_1, b_2) := 0$ for $(b_1, b_2) \in (B \times B) \setminus (\check{B} \times \check{B})$.

5.2 (a) Let A be an operator in a Hilbert space H. Show that A is strictly m-accretive if and only if $-A$ generates a C_0-semigroup T satisfying $\|T(t)\| \leqslant \mathrm{e}^{-\varepsilon t}$ $(t \geqslant 0)$ for some $\varepsilon > 0$.

(b) Let V, H be Hilbert spaces, let $j \in \mathcal{L}(V, H)$ have dense range, and let $a\colon V \times V \to \mathbb{K}$ be a bounded coercive form. Let $A \sim (a, j)$, and let T be the semigroup generated by $-A$. Show that there exists $\varepsilon > 0$ such that $\|T(t)\| \leqslant \mathrm{e}^{-\varepsilon t}$ for all $t \geqslant 0$.

(c) Let $\Omega \subseteq \mathbb{R}^n$ be an open set that is contained in a slab. Show that

$$\big\|\mathrm{e}^{t\Delta_{\mathrm{D}}}\big\|_{\mathcal{L}(L_2(\Omega))} \leqslant \mathrm{e}^{-\varepsilon t} \qquad (t \geqslant 0)$$

for some $\varepsilon > 0$. Express $\varepsilon > 0$ in terms of the width d of Ω; see the remark after the proof of Theorem 5.13.

5.3 Let V and H be Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$, let $a\colon V \times V \to \mathbb{K}$ be a bounded quasi-coercive form, and let $A \sim a$.

(a) Show that $\operatorname{dom}(A)$ is dense in V. (Hints: Without loss of generality assume that a is coercive. Then the Lax–Milgram operator $\mathcal{A}^*\colon V \to V^*$, associated with the adjoint form a^*, is surjective, and A is surjective as well. Use these two facts to show that the orthogonal complement of $\operatorname{dom}(A)$ in V equals $\{0\}$.)

(b) Show that a is accretive if (and only if) A is m-accretive.

(c) In the case $\mathbb{K} = \mathbb{C}$, show that a is sectorial if (and only if) A is m-sectorial.

(d) Let V_1 be a Hilbert space, $V_1 \overset{\mathrm{d}}{\hookrightarrow} H$, and let $a_1\colon V_1 \times V_1 \to \mathbb{K}$ be a bounded quasi-coercive form with the property that also $A \sim a_1$. Show that $V_1 = V$ and $a_1 = a$. (Hint: Observe that $a_1 = a$ on $\operatorname{dom}(A) \times \operatorname{dom}(A)$, and exploit the fact that $\operatorname{dom}(A)$ is dense in V and in V_1.)

5.4 Let $a, b, c, d \in \mathbb{K}$, and define a form γ on $\mathbb{K}^2$ by $\gamma(u, v) := \big(\big(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\big) u \,\big|\, v\big)$.

(a) Under what conditions on a, b, c, d is the form γ accretive? (Hint: The matrix $C := \big(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\big)$ is accretive if and only if the self-adjoint matrix $C + C^*$ is accretive.)

(b) Let $j\colon \mathbb{K}^2 \to \mathbb{K}, u \mapsto u_1$. Show that the operator A associated with (γ, j) is the operator of multiplication by $a - bc/d$ if $d \neq 0$. Under what conditions on d is γ a j-coercive form?

(c) Find a, b, c, d such that the form γ is j-coercive and not accretive, but the operator A associated with (γ, j) is accretive. (The existence of such an example explains why in Exercise 5.3(b) one needs to assume a to be an embedded form.)

5.5 Let $-\infty < a < b < \infty$. In this exercise we will always use the continuous representative for a function in $H^1(a,b)$; recall Theorem 4.12 for the embedding $H^1(a,b) \hookrightarrow C[a,b]$.

(a) Show that each $u \in H^1(a,b)$ is Hölder continuous with exponent $1/2$, i.e. $|u(t) - u(s)| \leqslant c|t-s|^{1/2}$ $(t,s \in [a,b])$ for some $c > 0$.

(b) Show that the embedding $H^1(a,b) \hookrightarrow C[a,b]$ is compact (i.e., the unit ball of $H^1(a,b)$ is relatively compact in $C[a,b]$).

(c) Let $H^2(a,b) := \{u \in L_2(a,b);\ u', u'' \in L_2(a,b)\}$. On $H^2(a,b)$ we define the norm by $\|u\|_{H^2} := \big(\|u\|_2^2 + \|u'\|_2^2 + \|u''\|_2^2\big)^{1/2}$. Show that $H^2(a,b)$ is a Hilbert space and that $H^2(a,b) \hookrightarrow C^1[a,b]$ if $C^1[a,b]$ carries the norm $\|u\|_{C^1} = \|u\|_{C[a,b]} + \|u'\|_{C[a,b]}$. (Hint: Recall Exercise 4.5(a).)

5.6 Let $-\infty < a < b < \infty$.

(a) Show that $C^1[a,b]$ is dense in $H^1(a,b)$. (Hint: Given $f \in H^1(a,b)$, approximate f' in $L_2(a,b)$ by a sequence (g_n) in $C_c(a,b)$. With x_0 and c from Proposition 4.8 define $f_n \in C^1[a,b]$ by $f_n(x) := c + \int_{x_0}^x g_n(y)\,\mathrm{d}y$.)

(b) Let $f, g \in H^1(a,b)$. Show that $fg \in H^1(a,b)$, $(fg)' = f'g + fg'$ and

$$\int_a^b f'g\,\mathrm{d}x = f(b)g(b) - f(a)g(a) - \int_a^b fg'\,\mathrm{d}x.$$

(Hint: Use Theorem 4.12.)

5.7 Let $-\infty < a < b < \infty$ and $\alpha, \beta > 0$. Define the operator A in $L_2(a,b)$ by

$$\begin{aligned} \mathrm{dom}(A) &= \big\{u \in H^2(a,b);\ -u'(a) + \alpha u(a) = 0,\ u'(b) + \beta u(b) = 0\big\}, \\ Au &= -u''. \end{aligned}$$

(See Exercise 5.5(c) for the definition of $H^2(a,b)$ and the existence of $u'(a)$ and $u'(b)$.)

(a) Show that A is m-accretive. (Hint: Consider the form

$$H^1(a,b) \times H^1(a,b) \ni (u,v) \mapsto \int_a^b u'\overline{v'}\,\mathrm{d}x + \alpha u(a)\overline{v(a)} + \beta u(b)\overline{v(b)}$$

and use Exercise 5.6(b).)

(b) Show that $\|\mathrm{e}^{-tA}\|_{\mathcal{L}(L_2(a,b))} \leqslant \mathrm{e}^{-\varepsilon t}$ $(t \geqslant 0)$ for some $\varepsilon > 0$.

5.8 Let V and H be complex Hilbert spaces, $j \in \mathcal{L}(V,H)$ an operator with dense range, $a\colon V \times V \to \mathbb{C}$ a bounded form satisfying (5.7), and let $A \sim (a,j)$. Let H_0 denote the real Hilbert space obtained from H by restricting scalars to $\mathbb{R}$ and defining the modified scalar product on H_0 by

$$(x \,|\, y)_0 = \mathrm{Re}(x \,|\, y) \qquad (x, y \in H_0 = H).$$

(a) Define $a_0\colon V_0 \times V_0 \to \mathbb{R}$ (where V_0 is the real Hilbert space associated with V) by

$$a_0(u,v) := \mathrm{Re}\, a(u,v) \qquad (u, v \in V_0).$$

Show that a_0 is a bounded form satisfying (5.7). For $A_0 \sim (a_0, j)$, show that $A_0 = A$.

(b) Assume additionally that a is j-coercive. Show that then a_0 is j-coercive, and that $T_0 = T$, where T and T_0 are the C_0-semigroups generated by $-A$ and $-A_0$, respectively.

5.9 (a) Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $m \in L_\infty(\Omega;\mathbb{R})$, $\delta > 0$ be such that $m \geqslant \delta$ a.e. Show that the operator $m\Delta_{\rm D}$ generates a bounded holomorphic C_0-semigroup on $H := L_2(\Omega)$ (where $\Delta_{\rm D}$ is the Dirichlet Laplacian; see Example 5.14).

Hint: Consider the (equivalent) scalar product on H given by $(u \,|\, v)_m := \int_\Omega u\bar{v}\,\frac{1}{m}\,{\rm d}x$ and show that $A := -m\Delta_{\rm D}$ is associated with the classical Dirichlet form on $H_0^1(\Omega) \overset{\rm d}{\hookrightarrow} (H, (\cdot\,|\,\cdot)_m)$.

(b) Let $\Omega := \mathbb{R}$ and put $m := 2$ on $(-\infty, 0]$, $m := 1$ on $(0,\infty)$. Show that the C_0-semigroup generated by $m\Delta_D$ is not quasi-contractive on $L_2(\mathbb{R})$, by showing that $A = -m\Delta_D$ is not quasi-accretive with respect to the standard scalar product $(\cdot\,|\,\cdot)$ on $L_2(\mathbb{R})$.

Hints: 1. For $u \in C_{\rm c}^\infty(\mathbb{R};\mathbb{R})$, show that $(Au \,|\, u) = \int_\mathbb{R} |u'|^2 m\,{\rm d}x - u(0)u'(0)$. 2. Choose $u, v \in C_{\rm c}^\infty(\mathbb{R};\mathbb{R})$ with $u(0) = 1$, $v'(0) = 1$. For $k \in \mathbb{N}$ put $u_k := u + \frac{1}{k}v(k^2\cdot)$. Show that $(u_k)_k$ is a bounded sequence in $H_0^1(\mathbb{R})$ and that $(Au_k \,|\, u_k) \to -\infty$ as $k \to \infty$.

5.10 (a) Let A be a quasi-m-sectorial operator in a complex Hilbert space H. Show that there exist $c \geqslant 0$, $\omega \in \mathbb{R}$ such that

$$|(Ax \,|\, y)| \leqslant c\big(\mathrm{Re}(Ax \,|\, x) + \omega\|x\|^2\big)^{1/2}\big(\mathrm{Re}(Ay \,|\, y) + \omega\|y\|^2\big)^{1/2} \qquad (x, y \in \mathrm{dom}(A)). \tag{5.10}$$

(Hint: Choose $\omega \geqslant 0$ such that $A + \omega$ is sectorial.)

(b) Let V, H be Hilbert spaces, let $j \in \mathcal{L}(V, H)$ have dense range, and let $a\colon V \times V \to \mathbb{K}$ be a bounded j-coercive form. Show that the operator $A \sim (a, j)$ satisfies (5.10) for some $c \geqslant 0$, $\omega \in \mathbb{R}$. (By part (a) and Corollary 5.11(b) this is clear if $\mathbb{K} = \mathbb{C}$.) Conclude that A is bounded if $\mathrm{Re}(Ax \,|\, x) = 0$ for all $x \in \mathrm{dom}(A)$.

Hint: Given $x \in \mathrm{dom}(A)$, let $u \in V$ be as in the definition (5.6) of A. Show that $\alpha\|u\|_V^2 \leqslant \mathrm{Re}(Ax \,|\, x) + \omega\|x\|_H^2$, with $\alpha > 0$, $\omega \in \mathbb{R}$ as in (5.9). Then use the boundedness of a.

(c) Let $\mathbb{K} = \mathbb{R}$, and let A be the operator of differentiation in $H := L_2(\mathbb{R})$, given by $\mathrm{dom}(A) := H^1(\mathbb{R})$, $Af := f'$. Show that $(Af \,|\, f) = 0$ for all $f \in \mathrm{dom}(A)$, and conclude that A is not associated with a bounded j-coercive form.

Chapter 6

Adjoint operators, and compactness

The main objective of this chapter is to show that, for a bounded open set $\Omega \subseteq \mathbb{R}^n$, the space $L_2(\Omega)$ has an orthonormal basis consisting of eigenfunctions of the Dirichlet Laplacian Δ_{D} on Ω. The proof is based on two crucial facts: $(I - \Delta_{\mathrm{D}})^{-1}$ is a compact self-adjoint operator, and for such operators there exist sufficiently many eigenfunctions. One of the aims of this chapter is to explain all the notions used in the previous sentence.

6.1 Adjoints of operators, and self-adjoint operators

Throughout this section let G and H be Hilbert spaces over $\mathbb{K}$. We want to prove the following result.

6.1 Theorem. *Let A be an operator in H. Then the following properties are equivalent.*

(i) *A is self-adjoint and accretive.*

(ii) *A is symmetric and m-accretive.*

If H is a complex Hilbert space, then there is the following additional equivalent property.

(iii) *A is m-sectorial of angle* 0.

Evidently, before proceeding to the proof, we first have to explain the notions of self-adjointness and symmetry. The proof will be given at the very end of this section.

Before defining the adjoint of an operator we explain the idea behind this notion. If A is an operator from G to H, then the adjoint A^* should be the maximal operator from H to G such that

$$(Ax \,|\, y)_H = (x \,|\, A^*y)_G \qquad \big(x \in \operatorname{dom}(A),\ y \in \operatorname{dom}(A^*)\big).$$

We recall that the orthogonal direct sum $G \oplus H$ is the product space $G \times H$ provided with the scalar product

$$\big((x,y) \,\big|\, (x_1,y_1)\big)_{G\oplus H} := (x \,|\, x_1)_G + (y \,|\, y_1)_H \qquad \big((x,y),(x_1,y_1) \in G \times H\big),$$

which makes it a Hilbert space.

For an operator A from G to H we define the **adjoint**

$$\begin{aligned} A^* &:= \big\{(y,x) \in H \times G;\ \forall x_1 \in \operatorname{dom}(A)\colon (Ax_1 \,|\, y)_H = (x_1 \,|\, x)_G\big\} \\ &= \big\{(y,x) \in H \oplus G;\ \forall (x_1,y_1) \in A\colon \big((x_1,-y_1) \,\big|\, (x,y)\big)_{G\oplus H} = 0\big\} \\ &= ((-A)^\perp)^{-1}. \end{aligned}$$

It is obvious from the last equality that A^* is a closed subspace of $H \oplus G$. We mention that the operator $-A$, as a linear relation, is given by

$$-A = \{(x, -y)\,;\, (x, y) \in A\}.$$

(It is somewhat unfortunate that, in principle, A^* is already defined as the antidual space to the subspace A of $G \oplus H$. We think that this ambiguity will not be a problem.)

6.2 Remarks. Let A be an operator from G to H.

(a) If B is an operator from H to G such that

$$(Ax \,|\, y)_H = (x \,|\, By)_G \qquad \big(x \in \operatorname{dom}(A),\ y \in \operatorname{dom}(B)\big),$$

then it is immediate from the definition of the adjoint that $B \subseteq A^*$ and $A \subseteq B^*$.

(b) If A^* is an operator, then the definition implies that $(Ax \,|\, y)_H = (x \,|\, A^*y)_G$ for all $x \in \operatorname{dom}(A)$, $y \in \operatorname{dom}(A^*)$.

(c) If B is an operator from G to H, $A \subseteq B$, then clearly $B^* \subseteq A^*$. △

6.3 Theorem. *Let A be an operator from G to H.*

(a) *Then A^* is an operator if and only if* $\operatorname{dom}(A)$ *is dense in G.*

(b) *Assume that* $\operatorname{dom}(A)$ *is dense. Then $A^{**} := (A^*)^* = \overline{A}$, and* $\operatorname{dom}(A^*)$ *is dense if and only if A is closable.*

(c) *If $A \in \mathcal{L}(G, H)$, then $A^* \in \mathcal{L}(H, G)$ and $\|A^*\| = \|A\|$.*

Proof. (a) From the definition of A^* it is easy to see that $\{x \in G\,;\, (0, x) \in A^*\} = \operatorname{dom}(A)^\perp$. This equality implies the assertion.

(b) It is easy to see that in the expression $((-A)^\perp)^{-1}$ for A^* the order of the operations $A \mapsto -A$, $A \mapsto A^\perp$ and $A \mapsto A^{-1}$ does not matter. It follows that $A^{**} = A^{\perp\perp} = \overline{A}$.

By part (a), $\operatorname{dom}(A^*)$ is dense if and only if $(A^*)^*$ is an operator, and because $A^{**} = \overline{A}$, the latter is equivalent to the closability of A.

(c) For each $y \in H$ the mapping $G \ni x \mapsto (Ax \,|\, y)$ is a bounded linear functional; hence by the Fréchet–Riesz representation theorem there exists $z \in G$ such that $(Ax \,|\, y)_H = (x \,|\, z)_G$ for all $x \in G$, and this implies $y \in \operatorname{dom}(A^*)$. Then $|(Ax \,|\, y)_H| = |(x \,|\, A^*y)_G|$ for all $x \in G$, $y \in H$, and taking the supremum over all x, y with $\|x\| \leqslant 1$, $\|y\| \leqslant 1$ one obtains $\|A\| = \|A^*\|$. □

6.4 Example. We define the operator A in $L_2(\mathbb{R})$ by

$$\operatorname{dom}(A) := C_c^\infty(\mathbb{R}), \qquad Af := f' \quad (f \in \operatorname{dom}(A)).$$

Then for $f, g \in L_2(\mathbb{R})$ we obtain:

$$\begin{aligned}(g, f) \in A^* &\iff \forall \varphi \in C_c^\infty(\mathbb{R})\colon (A\varphi \,|\, g) = (\varphi \,|\, f)\\ &\iff \forall \varphi \in C_c^\infty(\mathbb{R})\colon \int g\varphi' = \int f\varphi\\ &\iff f = -g' \text{ in the distributional sense.}\end{aligned}$$

This shows that $\operatorname{dom}(A^*) = H^1(\mathbb{R})$, $A^*g = -g'$ $(g \in \operatorname{dom}(A^*))$; in particular, $A \subseteq -A^*$.

Now it follows that $\overline{A} = -A^*$ because the subspace $-A^*$ of $L_2(\mathbb{R}) \oplus L_2(\mathbb{R})$ is isometrically isomorphic to $H^1(\mathbb{R})$, under the mapping $-A^* \ni (g, g') \mapsto g \in H^1(\mathbb{R})$, and $\operatorname{dom}(A) = C_c^\infty(\mathbb{R})$ is dense in $H^1(\mathbb{R})$, by Theorem 4.15. △

An operator A in H is called **symmetric** if $\operatorname{dom}(A)$ is dense and $A \subseteq A^*$; A is called **self-adjoint** if $A = A^*$ (note that then $\operatorname{dom}(A)$ is dense, by Theorem 6.3(a)).

We collect some properties of symmetric operators.

6.5 Remarks. Let A be an operator in H, $\operatorname{dom}(A)$ dense.

(a) Then A is symmetric if and only if $(Ax \mid y) = (x \mid Ay)$ for all $x, y \in \operatorname{dom}(A)$. This follows immediately from Remark 6.2.

(b) If H is complex, then A is symmetric if and only if $(Ax \mid x) \in \mathbb{R}$ for all $x \in \operatorname{dom}(A)$. This follows from Remark 5.1(b), applied to the form $\operatorname{dom}(A) \times \operatorname{dom}(A) \ni (x, y) \mapsto (Ax \mid y)$, and part (a) above.

(c) If A is symmetric, then A is closable, and $\overline{A}$ is symmetric. Indeed, from $A \subseteq A^*$ and the fact that A^* is a closed operator one concludes that A is closable and that $\overline{A} \subseteq A^* = \overline{A}^*$ (where the last equality is clear from the definition). In this situation one says that A is **essentially self-adjoint** if $\overline{A}$ is self-adjoint. △

Particularly simple examples of self-adjoint operators are those possessing 'sufficiently many' eigenvectors, which we will present next. We note that an eigenvalue of a symmetric operator A is always real: if $0 \neq x \in H$, $Ax = \lambda x$, then $\lambda(x \mid x) = (Ax \mid x) \in \mathbb{R}$.

6.6 Example (Diagonal self-adjoint operators). Let A be a self-adjoint operator in an infinite-dimensional separable Hilbert space H, and assume that there exists an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of H consisting of eigenvectors of A, with corresponding eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$. Then

$$\operatorname{dom}(A) = \Big\{x \in H\,;\ \sum_{n=1}^{\infty} |\lambda_n|^2 |(x \mid e_n)|^2 < \infty\Big\},$$

$$Ax = \sum_{n=1}^{\infty} \lambda_n (x \mid e_n) e_n \qquad (x \in \operatorname{dom}(A)),$$

and $\operatorname{lin}\{e_n\,;\ n \in \mathbb{N}\}$ is a core for A.

We call A the **diagonal operator** associated with the orthonormal basis $(e_n)_{n\in\mathbb{N}}$ and the sequence $\lambda := (\lambda_n)_{n\in\mathbb{N}}$.

Proof. Using the unitary operator $J\colon H \to \ell_2$, $x \mapsto \big((x \mid e_n)\big)_{n\in\mathbb{N}}$, we transform the situation to the case when A is a self-adjoint operator in the Hilbert space ℓ_2, possessing the canonical unit vectors as eigenvectors.

Now we apply Exercise 6.1 to the maximal multiplication operator M_λ in ℓ_2, choosing $\mathbb{N}$ with counting measure as the measure space and the sequence $(\lambda_n)_{n\in\mathbb{N}}$ as the multiplying function m. Observe that $Af = M_\lambda f$ for all $f \in c_{\mathrm{c}} := \operatorname{lin}\{e_n\,;\ n \in \mathbb{N}\}$, i.e., A is an extension of the operator $A_0 := M_\lambda|_{c_{\mathrm{c}}}$ defined in Exercise 6.1. Moreover $\overline{A_0} = M_\lambda$ by part (c) of that exercise, i.e. $c_{\mathrm{c}} = \operatorname{dom}(A_0)$ is a core for M_λ. Because A is closed, we conclude that $M_\lambda \subseteq A$ and hence, by Remark 6.2(c), that $A = A^* \subseteq M_\lambda^* = M_\lambda$. Thus $A = M_\lambda$ is the operator described above. □

6.7 Remark. Assume that in the above example $\lambda_n \geqslant 0$ for all $n \in \mathbb{N}$. Then it follows from Exercise 1.6(b) that $-A$ generates the C_0-semigroup $(M_{\mathrm{e}^{-t\lambda}})_{t\geqslant 0}$, written in the representation of A mentioned at the beginning of the proof given above. Spelling this out in the setting on H, the C_0-semigroup T generated by $-A$ is given by

$$T(t)x = \sum_{n=1}^{\infty} \mathrm{e}^{-t\lambda_n} (x \,|\, e_n) e_n \qquad (t \geqslant 0,\ x \in H). \qquad \triangle$$

In order to connect self-adjointness with m-accretivity we need some preliminary results.

6.8 Lemma. *Let A be an operator from G to H. Then* $\ker(A^*) = \operatorname{ran}(A)^\perp$.

Proof. Let $y \in H$, and recall that $y \in \ker(A^*)$ means $(y, 0) \in A^*$. Hence, by the definition of A^*, $y \in \ker(A^*)$ is equivalent to $(Ax \,|\, y)_H = 0$ for all $x \in \operatorname{dom}(A)$, i.e., to $y \perp \operatorname{ran}(A)$. □

We point out that in the following result, denseness of $\operatorname{dom}(A)$ is not part of the hypotheses.

6.9 Proposition. *Let A be an operator in H, $(Ax \,|\, y) = (x \,|\, Ay)$ for all $x, y \in \operatorname{dom}(A)$, and* $\operatorname{ran}(A) = H$. *Then A is self-adjoint.*

Proof. Remark 6.2(a) shows that $A \subseteq A^*$, and Lemma 6.8 yields $\ker(A^*) = \operatorname{ran}(A)^\perp = \{0\}$. These two facts, together with $\operatorname{ran}(A) = H$, imply that $A^* = A$. (This follows the principle that 'a surjective mapping cannot have a proper injective extension' – which also holds for relations.) □

6.10 Lemma. *Let A be an operator in H,* $\operatorname{dom}(A)$ *dense, and let $\lambda \in \mathbb{K}$. Then $(\lambda I + A)^* = \overline{\lambda} I + A^*$.*

Proof. For $x \in \operatorname{dom}(A)$, $y \in \operatorname{dom}(A^*)$ we compute

$$\big((\lambda + A)x \,\big|\, y\big) = \big(x \,\big|\, \overline{\lambda} y\big) + \big(x \,\big|\, A^* y\big) = \big(x \,\big|\, (\overline{\lambda} + A^*) y\big),$$

and using Remark 6.2(a) we obtain $\overline{\lambda} + A^* \subseteq (\lambda + A)^*$. This inclusion also implies that

$$A^* = (-\lambda + (\lambda + A))^* \supseteq -\overline{\lambda} + (\lambda + A)^*,$$

i.e. $\overline{\lambda} + A^* \supseteq (\lambda + A)^*$. The two inclusions prove the assertion. □

We can now prove the main result of this section.

Proof of Theorem 6.1. (i) ⇒ (ii), (iii). Clearly it suffices to show that $\operatorname{ran}(I + A) = H$. Recall that $\|(I + A)x\| \geqslant \|x\|$ for all $x \in \operatorname{dom}(A)$, by the accretivity of A; therefore $(I + A)^{-1} \colon \operatorname{ran}(I + A) \to X$ is continuous. As A is closed, and hence $(I + A)^{-1}$ is closed, it follows that $\operatorname{ran}(I + A) = \operatorname{dom}((I + A)^{-1})$ is closed, so it remains to show that $\operatorname{ran}(I + A)$ is dense.

By Lemma 6.10 the operator $I + A$ is self-adjoint, and $I + A$ is injective since A is accretive. Now Lemma 6.8 yields $\operatorname{ran}(I + A)^\perp = \ker((I + A)^*) = \ker(I + A) = \{0\}$.

(ii) ⇒ (i). The hypothesis implies that $I + A$ satisfies the conditions in Proposition 6.9, and therefore $I + A$ is self-adjoint. Then Lemma 6.10 shows that A is self-adjoint.

(iii) ⇒ (ii). Being an m-sectorial operator, A is m-accretive. Sectoriality of angle 0 means that $(Ax \,|\, x) \geqslant 0$ for all $x \in \operatorname{dom}(A)$. Applying Remark 6.5(b) one concludes that A is symmetric. □

6.2 Adjoints of forms and operators

In this section we assume that V and H are Hilbert spaces over $\mathbb{K}$ and that $j \in \mathcal{L}(V, H)$ has dense range. Let $a\colon V \times V \to \mathbb{K}$ be a bounded j-coercive form. For the reader's convenience we recall the definition of the operator A associated with (a, j),

$$A = \big\{(x, y) \in H \times H\,;\ \exists u \in V\colon j(u) = x,\ a(u, v) = (y \,|\, j(v))\ (v \in V)\big\};$$

see Section 5.3. We note that the adjoint form a^* is j-coercive as well.

The following result shows the close connection between adjoints of forms and operators.

6.11 Theorem. *Let $A \sim (a, j)$, $B \sim (a^*, j)$. Then $B = A^*$. If a is symmetric, then A is self-adjoint.*

Proof. Without loss of generality we suppose that a is coercive; see Remark 5.10 and Lemma 6.10. Then for $x \in \operatorname{dom}(A)$, $y \in \operatorname{dom}(B)$ there exist $u, v \in V$ such that $j(u) = x$, $j(v) = y$ and

$$(Ax \,|\, y) = (Ax \,|\, j(v)) = a(u, v) = \overline{a^*(v, u)} = \overline{(By \,|\, j(u))} = (x \,|\, By).$$

Hence $B \subseteq A^*$, by Remark 6.2(a).

From Theorem 5.6 we know that $\operatorname{ran}(B) = \operatorname{ran}(A) = H$, and therefore Lemma 6.8 implies that A^* is injective. These properties imply that $B = A^*$ (recall that a surjective mapping cannot have a proper injective extension).

If a is symmetric, then $A = B = A^*$. □

6.3 The spectral theorem for compact self-adjoint operators

We recall that, for Banach spaces X, Y, an operator $A\colon X \to Y$ is called **compact** if $A(B_X(0, 1))$ is a relatively compact subset of Y, where $B_X(0, 1)$ is the open unit ball of X. The set $\mathcal{K}(X, Y)$ of compact operators is a closed subspace of $\mathcal{L}(X, Y)$, and the composition of a compact operator with a bounded operator is compact. The latter is called the **ideal property** of compact operators.

The following result is the spectral theorem for compact self-adjoint operators in a Hilbert space H.

6.12 Theorem (Hilbert). *Let $A \in \mathcal{L}(H)$ be compact and self-adjoint. Then there exist $J \subseteq \mathbb{N}$ and an orthonormal system $(e_n)_{n\in J}$ of eigenvectors of A with corresponding eigenvalues $(\lambda_n)_{n\in J}$ in $\mathbb{R} \setminus \{0\}$, $\lambda_n \to 0$ as $n \to \infty$ if J is infinite, such that*

$$Ax = \sum_{n\in J} \lambda_n (x \,|\, e_n) e_n \qquad (x \in H).$$

6.13 Remark. Note that in the preceding theorem it follows that $\{e_n\,;\ n \in J\}^\perp = \ker(A)$; in particular, $(e_n)_{n\in J}$ is an orthonormal basis if and only if A is injective. △

For the proof of Theorem 6.12 we need the following two propositions.

6.14 Proposition. *Let $A \in \mathcal{L}(H)$ be self-adjoint. Then*

$$\|A\| = \sup\{|(Ax \,|\, x)|;\ \|x\| \leqslant 1\}.$$

Proof. '$\geqslant$' is obvious from $|(Ax \,|\, x)| \leqslant \|A\|\|x\|^2$.

'$\leqslant$'. For $x, y \in H$ we put $a(x,y) := (Ax \,|\, y)$ and $b(x,y) := M(x \,|\, y)$, where $M := \sup\{|a(x)|;\ \|x\| \leqslant 1\}$. Then the Cauchy–Schwarz inequality (Proposition 5.2) implies that $|(Ax \,|\, y)| \leqslant M\|x\|\|y\|$ $(x, y \in H)$; therefore $\|A\| \leqslant M$. □

6.15 Proposition. *Let $0 \neq A \in \mathcal{L}(H)$ be self-adjoint. Then $\|A\| \in \sigma(A)$ or $-\|A\| \in \sigma(A)$. If additionally A is compact, then $\|A\|$ or $-\|A\|$ is an eigenvalue of A.*

Proof. By Proposition 6.14 there exist a sequence (x_n) in H with $\|x_n\| = 1$ $(n \in \mathbb{N})$ and $\lambda \in \mathbb{R}$ with $|\lambda| = \|A\|$ such that $(Ax_n \,|\, x_n) \to \lambda$. Then

$$\|Ax_n - \lambda x_n\|^2 = \|Ax_n\|^2 - 2\lambda(Ax_n \,|\, x_n) + \lambda^2 \leqslant 2\lambda^2 - 2\lambda(Ax_n \,|\, x_n) \to 0;$$

hence $Ax_n - \lambda x_n \to 0$ as $n \to \infty$. This implies that $\lambda \in \sigma(A)$.

If A is compact, then there exists a subsequence (x_{n_k}) of (x_n) such that (Ax_{n_k}) converges to some $y \in H$. It follows that

$$x_{n_k} = \frac{1}{\lambda}(\lambda x_{n_k} - Ax_{n_k}) + \frac{1}{\lambda}Ax_{n_k} \to \frac{1}{\lambda}y =: x,$$

$Ax = \lim_{k\to\infty} Ax_{n_k} = y = \lambda x$ and $\|x\| = 1$. □

Proof of Theorem 6.12. We prove the assertion with $J = \mathbb{N}$ or $J = \{1, \dots, N\}$ for some $N \in \mathbb{N}_0$. The orthonormal system $(e_n)_{n\in J}$ and the corresponding family $(\lambda_n)_{n\in J}$ with the properties described in the theorem will be constructed recursively.

Assume that $k \in \mathbb{N}_0$ and that $e_1, \dots, e_k$ and $\lambda_1, \dots, \lambda_k$ have been constructed. Let $H_{k+1} := \{e_1, \dots, e_k\}^\perp$. Then $A(H_{k+1}) \subseteq H_{k+1}$; indeed, $x \perp e_n$ implies $(Ax \,|\, e_n) = (x \,|\, Ae_n) = \lambda_n(x \,|\, e_n) = 0$ for $n = 1, \dots, k$. Then $A_{k+1} := A|_{H_{k+1}}$, considered as an operator in H_{k+1}, is compact. If $A_{k+1} = 0$, then the construction is complete, with $J = \{1, \dots, k\}$. If $A_{k+1} \neq 0$, then Proposition 6.15 implies that there exist $\lambda_{k+1} \in \mathbb{R} \setminus \{0\}$ with $|\lambda_{k+1}| = \|A_{k+1}\|$ and $e_{k+1} \in H_{k+1}$ with $\|e_{k+1}\| = 1$ such that $A_{k+1}e_{k+1} = \lambda_{k+1}e_{k+1}$.

The construction yields $|\lambda_n| = \|A_n\|$ $(n \in J)$, and since $(\|A_n\|)_{n\in J}$ is decreasing, so is $(|\lambda_n|)_{n\in J}$.

If $J = \mathbb{N}$, then $e_n \to 0$ weakly as $n \to \infty$. It follows that $|\lambda_n| = \|Ae_n\| \to 0$ as $n \to \infty$ because A is compact; see Exercise 6.6(a).

Finally we establish the representation of A. For $x \in H$ we have

$$\Big\|Ax - \sum_{n=1}^{k} \lambda_n(x \,|\, e_n)e_n\Big\| = \Big\|A_{k+1}\Big(x - \sum_{n=1}^{k}(x \,|\, e_n)e_n\Big)\Big\| \leqslant \|A_{k+1}\|\|x\|,$$

which is 0 if $J = \{1, \dots, k\}$ and converges to 0 as $k \to \infty$ if $J = \mathbb{N}$. □

An operator A in a Banach space X is said to have **compact resolvent** if there exists $\lambda \in \rho(A)$ such that $R(\lambda, A)$ is a compact operator. It follows from the resolvent equation and the ideal property of compact operators that then $R(\mu, A)$ is compact for all $\mu \in \rho(A)$.

The following theorem is an application of Hilbert's theorem to accretive self-adjoint operators with compact resolvent. It shows that Example 6.6 – with a sequence (λ_n) in $[0, \infty)$ tending to ∞ – is generic for these operators.

6.16 Theorem. *Let A be an accretive self-adjoint operator with compact resolvent in an infinite-dimensional Hilbert space H. Then there exist an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ and a sequence $(\lambda_n)_{n\in\mathbb{N}}$ in $[0, \infty)$ with $\lim_{n\to\infty} \lambda_n = \infty$ such that A is the associated diagonal operator.*

Proof. It follows from Theorem 6.1 that $(I + A)^{-1}$ exists in $\mathcal{L}(H)$, and by hypothesis this operator is compact. It is easy to see that $(I + A)^{-1}$ is symmetric, hence self-adjoint. Applying Theorem 6.12 to $(I + A)^{-1}$ one obtains an orthonormal system $(e_n)_{n\in J}$ of eigenvectors, with a corresponding family $(\mu_n)_{n\in J}$ in $\mathbb{R} \setminus \{0\}$ of eigenvalues, such that $(I + A)^{-1}x = \sum_{n\in J} \mu_n(x \,|\, e_n)e_n$ for all $x \in H$. This representation together with the injectivity of $(I + A)^{-1}$ implies that $(e_n)_{n\in J}$ is an orthonormal basis; see Remark 6.13. Hence J is countably infinite, without loss of generality $J = \mathbb{N}$.

One easily sees that e_n is an eigenvector of A with eigenvalue $\lambda_n := \mu_n^{-1} - 1$, for all $n \in \mathbb{N}$. The accretivity of A implies

$$\lambda_n = (\lambda_n e_n \,|\, e_n) = (Ae_n \,|\, e_n) \geqslant 0 \qquad (n \in \mathbb{N}).$$

Since $(\mu_n)_{n\in\mathbb{N}}$ is a null sequence, it follows that $\lim_{n\to\infty} \lambda_n = \infty$.

In view of Example 6.6, this completes the proof. □

A sequence (λ_n) in $\mathbb{R}$ with $\lim_{n\to\infty} \lambda_n = \infty$ can always be rearranged to an increasing sequence, and in applications of Theorem 6.16 one usually assumes the sequence of eigenvalues to be increasing. Note that the hypotheses of Theorem 6.16 imply that H is separable.

We now describe how diagonal operators and operators with compact resolvent arise in the context of forms. We start by presenting an example illustrating these issues.

6.17 Example (Diagonal forms). Let V, H be infinite-dimensional separable Hilbert spaces, $j\colon V \overset{\mathrm{d}}{\hookrightarrow} H$ an embedding, let a be a symmetric bounded quasi-coercive form on V, and let $A \sim (a, j)$. Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of H, and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ that is bounded below.

Then A is the diagonal operator associated with $(e_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$ (as defined in Example 6.6) if and only if

$$V = \Big\{u \in H\,;\, \sum_{n=1}^{\infty} |\lambda_n| |(u \,|\, e_n)_H|^2 < \infty\Big\},$$

$$a(u, v) = \sum_{n=1}^{\infty} \lambda_n (u \,|\, e_n)_H (e_n \,|\, v)_H \qquad (u, v \in V).$$

(Then an appropriate scalar product on V is given by $(u\,|\,v)_V := a(u,v) + \omega(u\,|\,v)_H$ $(u, v \in V)$, for large enough $\omega \in \mathbb{R}$.)

Indeed, if the form a is as stated above, then every e_n is an eigenvector of the (self-adjoint) operator A, with corresponding eigenvalue λ_n, and A is as described in Example 6.6. The reverse implication is a consequence of the uniqueness of the form discussed in Exercise 5.3(d).

We add the observation that the embedding j is compact if and only if $\lim_{n\to\infty} \lambda_n = \infty$. Indeed, if the latter property holds, then one easily sees that the finite rank operators $P_n j$, where P_n denotes the orthogonal projection onto $\operatorname{lin}\{e_1, \ldots, e_n\}$, approximate j in the operator norm. On the other hand, if there exists a bounded subsequence (λ_{n_j}) of (λ_n), then the sequence (e_{n_j}) is a bounded sequence in V that does not have a convergent subsequence in H.

By an analogous argument (or by Theorem 6.16 and Proposition 6.18 below) one also sees that A has compact resolvent if and only if $\lim_{n\to\infty} \lambda_n = \infty$. △

If A is an operator associated with a form, then one has the following convenient condition for A to have compact resolvent.

6.18 Proposition. *Let V, H be Hilbert spaces, let $j \in \mathcal{L}(V, H)$ have dense range, let $a\colon V \times V \to \mathbb{K}$ be a bounded j-coercive form, and let $A \sim (a, j)$. Assume additionally that j is compact.*

Then A has compact resolvent.

Proof. Without loss of generality we may suppose that a is coercive. Then by Proposition 5.7, the inverse of A can be expressed explicitly as $A^{-1} = j\mathcal{A}^{-1}k$. As $\mathcal{A}^{-1}$ and k are bounded operators, we see that A^{-1} is compact. □

We finally apply our results to the Dirichlet Laplacian Δ_{D} in $L_2(\Omega)$, for bounded open $\Omega \subseteq \mathbb{R}^n$, as announced in the introduction of this chapter. We will use the compactness of the embedding $H_0^1(\Omega) \hookrightarrow L_2(\Omega)$, which will be proved in the next section.

6.19 Example (Spectral decomposition of the Dirichlet Laplacian). If $\Omega \subseteq \mathbb{R}^n$ is a bounded open set, then Δ_{D} has compact resolvent. There exist an orthonormal basis $(\varphi_k)_{k\in\mathbb{N}}$ of $L_2(\Omega)$ and an increasing sequence $(\lambda_k)_{k\in\mathbb{N}}$ in $(0, \infty)$, with $\lim_{k\to\infty} \lambda_k = \infty$, such that $-\Delta_{\mathrm{D}}$ is the associated diagonal operator. In particular, $\varphi_k \in \operatorname{dom}(\Delta_{\mathrm{D}})$ and

$$-\Delta_{\mathrm{D}}\varphi_k = \lambda_k\varphi_k$$

for all $k \in \mathbb{N}$. One has Poincaré's inequality

$$\int_\Omega |u|^2 \,\mathrm{d}x \leqslant c_{\mathrm{P}} \int_\Omega |\nabla u|^2 \,\mathrm{d}x \qquad (u \in H_0^1(\Omega)),$$

with optimal constant $c_{\mathrm{P}} = \frac{1}{\lambda_1}$.

Proof. From Example 5.14 we know that $-\Delta_{\mathrm{D}}$ is associated with the classical Dirichlet form on $H_0^1(\Omega)$ (which is coercive by Theorem 5.13) and $j\colon H_0^1(\Omega) \hookrightarrow L_2(\Omega)$. Combining Theorems 5.6 and 6.11 we conclude that that $-\Delta_{\mathrm{D}}$ is a strictly accretive self-adjoint operator. As the embedding j is compact by Theorem 6.21 below, Proposition 6.18 implies that $-\Delta_D$ has compact resolvent, and the assertions concerning the sequences (φ_k) and (λ_k) follow from Theorem 6.16. (Note that 0 cannot be an eigenvalue because $-\Delta_D$ is strictly accretive.)

Now for $u \in \operatorname{dom}(-\Delta_D)$ we obtain

$$\int_\Omega |\nabla u|^2 \,\mathrm{d}x = (-\Delta_D u \mid u) = \Big(\sum_{k=1}^\infty \lambda_k (u \mid \varphi_k)\varphi_k \,\Big|\, u\Big) = \sum_{k=1}^\infty \lambda_k |(u \mid \varphi_k)|^2 \geqslant \lambda_1 \|u\|_2^2\,,$$

with equality for $u = \varphi_1$. As $\operatorname{dom}(\Delta_D)$ contains $C_c^\infty(\Omega)$ and hence is dense in $H_0^1(\Omega)$, the inequality $\int_\Omega |\nabla u|^2 \,\mathrm{d}x \geqslant \lambda_1 \|u\|_2^2$ carries over to all $u \in H_0^1(\Omega)$. □

6.20 Remark. The existence of an orthonormal basis of eigenvectors of $-\Delta_D$ is a highlight and triumph of Hilbert space theory as applied to partial differential equations. There is no way to obtain this kind of result by computing eigenfunctions, even if the boundary is nice.

In the case of an interval in one dimension it is not difficult to compute the eigenfunctions (see Exercise 6.9), and this can be generalised to n-dimensional rectangles. There is also a formula for balls, but this is more complicated and involves Bessel functions. △

6.4 Interlude: compactness of the embedding $H_0^1(\Omega) \hookrightarrow L_2(\Omega)$

In this section we prove the following special case of the Rellich–Kondrachov theorem.

6.21 Theorem (Rellich–Kondrachov). *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Then the embedding $j\colon H_0^1(\Omega) \hookrightarrow L_2(\Omega)$ is compact.*

The next result, a key ingredient for the proof of Theorem 6.21, provides the link to a basic fact concerning compactness in function spaces, the Arzelà–Ascoli theorem.

6.22 Proposition. *Let $\rho \in C_c(\mathbb{R})$, let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, and let $1 \leqslant p \leqslant \infty$. Then the mapping*

$$J_\rho\colon L_p(\mathbb{R}^n) \to L_p(\Omega),\ u \mapsto (\rho * u)|_\Omega$$

is compact.

Proof. We show that

$$F := \big\{(\rho * u)|_{\overline{\Omega}};\, u \in L_p(\mathbb{R}^n),\ \|u\|_p \leqslant 1\big\}$$

is relatively compact in $C(\overline{\Omega})$. Clearly, the set is bounded because $|\rho * u(x)| \leqslant \|\rho\|_q \|u\|_p$ for all $x \in \mathbb{R}^n$ (where $\frac{1}{q} + \frac{1}{p} = 1$, as usual). Moreover, the estimate

$$|\rho * u(x) - \rho * u(y)| \leqslant \int \big|\big(\rho(x-z) - \rho(y-z)\big)u(z)\big| \,\mathrm{d}z \leqslant \|\rho(x-\cdot) - \rho(y-\cdot)\|_q \|u\|_p$$

shows that F is equicontinuous. By the Arzelà–Ascoli theorem it follows that F is relatively compact in $C(\overline{\Omega})$. The embedding $C(\overline{\Omega}) \hookrightarrow L_p(\Omega)$ is continuous, and therefore F is also relatively compact in $L_p(\Omega)$, i.e. J_ρ is compact. □

We will also need the following observation concerning approximation.

6.23 Lemma. *Let $u \in H^1(\mathbb{R}^n)$. Then $\|u(\cdot - y) - u\|_2 \leqslant |y|\|u\|_{H^1}$ for all $y \in \mathbb{R}^n$.*

Proof. It suffices to prove the estimate for u from the dense subspace $C_c^1(\mathbb{R}^n)$ of $H^1(\mathbb{R}^n)$. Note that $u(x-y) - u(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} u(x-ty)\,\mathrm{d}t$ for all $x, y \in \mathbb{R}^n$. By the Cauchy–Schwarz inequality it follows that $|u(x-y) - u(x)|^2 \leqslant \int_0^1 \left|\nabla u(x-ty)\cdot y\right|^2 \mathrm{d}t$, and integration over $x \in \mathbb{R}^n$ yields $\|u(\cdot - y) - u\|_2^2 \leqslant \int_0^1 \|\nabla u(\cdot - ty)\|_2^2 |y|^2 \,\mathrm{d}t \leqslant |y|^2 \|u\|_{H^1}^2$. □

Proof of Theorem 6.21. (i) We define bounded operators $E\colon H_0^1(\Omega) \to H^1(\mathbb{R}^n)$, $Eu := \tilde{u}$ (extension of functions to $\mathbb{R}^n$ by zero) and $R\colon L_2(\mathbb{R}^n) \to L_2(\Omega)$, $Rf := f|_\Omega$ (restriction of functions to Ω). Let (ρ_k) be a delta sequence in $C_c(\mathbb{R}^n)$. For $k \in \mathbb{N}$ let $J_k\colon H^1(\mathbb{R}^n) \to L_2(\mathbb{R}^n)$, $J_k f := \rho_k * f$; then Proposition 6.22 implies that RJ_k is compact. We will show that $J_k \to J$ in $\mathcal{L}(H^1(\mathbb{R}^n), L_2(\mathbb{R}^n))$, where $J\colon H^1(\mathbb{R}^n) \hookrightarrow L_2(\mathbb{R}^n)$ is the embedding. Then $j = RJE = \lim_{k\to\infty} RJ_kE$ is compact by the ideal property of compact operators and the closedness of $\mathcal{K}(H^1(\Omega), L_2(\Omega))$ in $\mathcal{L}(H^1(\Omega), L_2(\Omega))$.

(ii) Let $u \in C_c^1(\mathbb{R}^n)$ and $k \in \mathbb{N}$. For $x \in \mathbb{R}^n$ we compute, using the Cauchy–Schwarz inequality in the second step,

$$
\begin{aligned}
|\rho_k * u(x) - u(x)| &= \left| \int \rho_k(y)^{1/2} \rho_k(y)^{1/2} \big(u(x-y) - u(x)\big)\,\mathrm{d}y \right| \\
&\leqslant \Big(\int \rho_k(y)\,\mathrm{d}y \Big)^{1/2} \Big(\int \rho_k(y) |u(x-y) - u(x)|^2 \,\mathrm{d}y \Big)^{1/2}.
\end{aligned}
$$

Integrating over $x \in \mathbb{R}^n$ we obtain, with Fubini's theorem in the second step,

$$
\|\rho_k * u - u\|_2^2 \leqslant \iint \rho_k(y) |u(x-y) - u(x)|^2 \,\mathrm{d}y\,\mathrm{d}x = \int \rho_k(y) \|u(\cdot - y) - u\|_2^2 \,\mathrm{d}y.
$$

By Lemma 6.23 it follows that $\|\rho_k * u - u\|_2 \leqslant \frac{1}{k}\|u\|_{H^1}$. We conclude that $J_k \to J$ in $\mathcal{L}(H^1(\mathbb{R}^n), L_2(\mathbb{R}^n))$ since $C_c^1(\mathbb{R}^n)$ is dense in $H^1(\mathbb{R}^n)$. □

Notes

It is difficult to attribute the development of adjoint operators to a source. One of the first more systematic treatments is given in [Neu32b]. The idea of also including linear relations is contained in [Are61]. In Section 14.1 we will treat the adjoint for linear relations and self-adjoint linear relations.

The first version of the spectral theorem for compact self-adjoint operators appeared in [Hil06]. (This paper is also contained in the collection [Hil12].) The diagonal structure of the Dirichlet Laplacian seems to be classical and difficult to attribute; however, the use of compactness methods for this purpose can be attributed to Rellich [Rel30]. In this paper the first version of the Rellich–Kondrachov theorem appeared as well as the application to the Dirichlet Laplacian. The proof we give for the Rellich–Kondrachov theorem is not the one usually found in textbooks.

Exercises

6.1 Let (Ω, μ) be a measure space, and let $m\colon \Omega \to \mathbb{K}$ be measurable. Let

$$\mathcal{C} := \Big\{ C \subseteq \Omega\,;\ C \text{ measurable},\ \mu(C) < \infty,\ \sup_{x\in C} |m(x)| < \infty \Big\},$$

and define the operator A_0 in $L_2(\Omega, \mu)$ by

$$\begin{aligned} \operatorname{dom}(A_0) &:= \operatorname{lin}\{\mathbf{1}_C\,;\ C \in \mathcal{C}\}, \\ A_0 f &:= mf \qquad (f \in \operatorname{dom}(A_0)). \end{aligned}$$

(a) Show that $\operatorname{dom}(A_0)$ is dense in $L_2(\Omega, \mu)$, and that $A_0^* = M_{\overline{m}}$, where $M_{\overline{m}}$ denotes the maximal multiplication operator induced by $\overline{m}$,

$$\begin{aligned} \operatorname{dom}(M_{\overline{m}}) &:= \{ f \in L_2(\Omega, \mu)\,;\ \overline{m} f \in L_2(\Omega, \mu) \}, \\ M_{\overline{m}} f &:= \overline{m} f \qquad (f \in \operatorname{dom}(M_{\overline{m}})). \end{aligned}$$

(b) Show further that $M_{\overline{m}}^* = M_m$ (maximal multiplication operator induced by m).

(c) Show that $\overline{A_0} = M_m$, i.e. $\operatorname{dom}(A_0)$ is a core for M_m.

6.2 Let A be a self-adjoint operator in a complex Hilbert space H. Show that $\sigma(A) \subseteq \mathbb{R}$. (Hint: Argue as in the proof of Theorem 6.1.)

6.3 Let H be a Hilbert space.

(a) Let $B \in \mathcal{L}(H)$. Show that $\|B^*B\| = \|B\|^2$. (Hint: For the less trivial inequality start with $\|B^*B\| = \sup_{\|x\|,\|y\| \leqslant 1} |(B^*Bx \,|\, y)| \geqslant \sup_{\|x\| \leqslant 1} |(B^*Bx \,|\, x)|$.)

(b) Let T be a bounded C_0-semigroup on H, with generator A. Show that $T(t)$ is self-adjoint for all $t \geqslant 0$ if and only if $-A$ is an accretive self-adjoint operator, and that in this case $\|T(t)\| \leqslant 1$ for all $t \geqslant 0$.

Hint: Use part (a) to show that the boundedness hypothesis together with the self-adjointness of $T(t)$ implies that $\|T(t)\| \leqslant 1$. For the proof of the self-adjointness of A recall Proposition 6.9.

(c) Suppose that H is a complex Hilbert space, and let A be an accretive self-adjoint operator. Show that $-A$ generates a holomorphic C_0-semigroup of angle $\pi/2$, $\|T(z)\| \leqslant 1$ for all $\operatorname{Re} z > 0$.

6.4 (a) Let H and V be Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$, let $a\colon V \times V \to \mathbb{K}$ be a bounded quasi-coercive form, and let $A \sim a$. Show that a is symmetric if (and only if) A is self-adjoint. (Hint: Exercise 5.3(a).)

(b) Use the setup of Exercise 5.4 to find an example showing that in part (a) one needs to assume a to be an embedded form.

6.5 (a) Let X be a vector space, Y an n-dimensional subspace, Z an $(n{-}1)$-codimensional subspace (i.e., there exists an $(n{-}1)$-dimensional subspace Z_0 of X such that $Z \cap Z_0 = \{0\}$ and $Z + Z_0 = X$). Show that $Y \cap Z \neq \{0\}$.

(b) Let V, H be infinite-dimensional Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$, and let a be a symmetric bounded quasi-coercive form on V. Assume that the operator A associated with a is a self-adjoint diagonal operator in H as described in Example 6.6 and that the sequence of eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$ is increasing.

For a finite-dimensional subspace W of V, $W \neq \{0\}$, define

$$M_a(W) := \max\{a(u);\, u \in W,\ \|u\|_H = 1\}.$$

Prove the **min-max principle**:

$$\lambda_n = \min\bigl\{M_a(W);\, W \text{ subspace of } V,\ \dim W = n\bigr\} \qquad (n \in \mathbb{N}).$$

(Hint: Use the description of a from Example 6.17.)

(c) Let $\Omega_1 \subseteq \Omega_2 \subseteq \mathbb{R}^n$ be bounded open sets, and let Δ_j be the Dirichlet Laplacian in $L_2(\Omega_j)$ and $(\lambda_k^j)_{k\in\mathbb{N}}$ the corresponding increasing sequence of eigenvalues, for $j = 1, 2$.

Prove the **domain monotonicity of eigenvalues**: $\lambda_k^1 \geqslant \lambda_k^2$ for all $k \in \mathbb{N}$.

6.6 Let X, Y be Banach spaces, $A \in \mathcal{L}(X, Y)$.

(a) Assume that A is a compact operator. Show that A maps weakly convergent sequences in X to convergent sequences in Y.

(b) Show that the converse of (a) is true if X is reflexive: if A maps weakly convergent sequences in X to convergent sequences in Y, then A is compact.

Hints: 1. Recall that a sequence (x_n) in X is called weakly convergent to $x \in X$ if $\eta(x_n) \to \eta(x)$ for all $\eta \in X'$, and that, if X is reflexive, every bounded sequence in X contains a weakly convergent subsequence.

2. Recall that $A \in \mathcal{L}(X, Y)$ is also continuous with respect to the weak topologies on X and Y; in particular, if a sequence (x_n) in X is weakly convergent to $x \in X$, then (Ax_n) is weakly convergent to Ax in Y.

6.7 Let G, H be Hilbert spaces, $A \in \mathcal{L}(G, H)$. Show that the following properties are equivalent:

(i) A is compact,

(ii) A^* is compact,

(iii) A^*A is compact.

Hint concerning '(iii) $\Rightarrow$ (i)': Given a sequence (x_n) in G converging weakly to 0, show that $\|Ax_n\|^2 \to 0$; then apply Exercise 6.6.

Note. The equivalence of (i) and (ii) is a special case of "Schauder's theorem": For Banach spaces X, Y, an operator $A \in \mathcal{L}(X, Y)$ is compact if and only if the dual operator $A' \in \mathcal{L}(Y', X')$ is compact.

6.8 Let V, H be Hilbert spaces, let $j \in \mathcal{L}(V, H)$ have dense range, let $a\colon V \times V \to \mathbb{K}$ be a bounded j-coercive form, and let $A \sim (a, j)$. Show that j is compact if (and only if) A has compact resolvent. (Note that, in contrast to Exercises 5.3(b) and 6.4(a), j is not supposed to be an embedding.)

Hint: Assume without loss of generality that a is coercive, and let $\mathcal{A}$, k be as in Proposition 5.7. Proceed similarly as in Exercise 6.7 to prove that k is compact, showing that $a(\mathcal{A}^{-1}kx_n) \to 0$ for any weak null sequence (x_n) in H. Then derive the compactness of j by arguing once again with weak null sequences.

6.9 (a) Let $-\infty \leqslant a < x_0 < b \leqslant \infty$, let $f \in C(a,b)$, and assume that $g_1 := \big(f|_{(a,x_0)}\big)' \in L_1(a,x_0)$, $g_2 := \big(f|_{(x_0,b)}\big)' \in L_1(x_0,b)$. Define $g \in L_1(a,b)$ by $g|_{(a,x_0)} := g_1$, $g|_{(x_0,b)} := g_2$. Show that $f' = g$. (Hint: Use Proposition 4.8.)

(b) Let $-\infty < a < b < \infty$. Show that $H_0^1(a,b) = \{f \in H^1(a,b)\,;\ f(a) = f(b) = 0\}$.

Hint: For $f \in H^1(a,b)$ with $f(a) = f(b) = 0$, show that the extension of f to $\mathbb{R}$ by zero belongs to $H^1(\mathbb{R})$. Then decompose f as $f = g + h$ with $g, h \in H^1(\mathbb{R})$, $\operatorname{spt} g \subseteq [a,b)$, $\operatorname{spt} h \subseteq (a,b]$, and show that $g, h \in H_0^1(a,b)$, using Example 1.7(a).

(c) Compute the orthonormal basis of eigenfunctions and the eigenvalues of $-\Delta_{\mathrm{D}}$ for $\Omega = (0,\pi)$. Determine the optimal value of the Poincaré constant for the open set $(0,\pi)$ (see Example 6.19).

Chapter 7

Neumann and Robin boundary conditions

So far our study of the Laplacian was restricted to homogeneous Dirichlet boundary conditions. Our aim in this chapter is to investigate Neumann boundary conditions

$$\partial_\nu u = 0 \quad \text{on } \partial\Omega$$

and more generally Robin boundary conditions

$$\partial_\nu u + \beta u = 0 \quad \text{on } \partial\Omega.$$

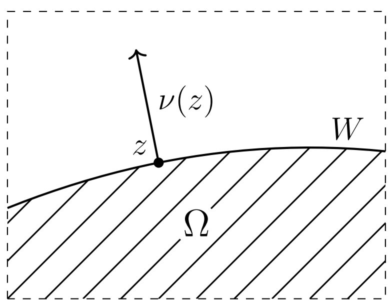
If we think of heat conduction in a body Ω, then Neumann boundary conditions describe an isolated body, whereas Robin boundary conditions describe when part of the heat is absorbed at the boundary at a rate β.

We start with the description of properties of the boundary for an open subset of $\mathbb{R}^n$. The main issue of Section 7.1 will be the discussion of Gauss' theorem and some of its consequences. In an interlude in Section 7.2 we present properties of $H^1(\Omega)$ that will be needed in Sections 7.4 and 7.5 to formulate Neumann and Robin boundary conditions and to derive properties of the corresponding operators.

7.1 Gauss' theorem

Before formulating Gauss' theorem we introduce some notation; in particular we define the notions of open sets with C^1-boundary and of the outer unit normal. Throughout this section, $\Omega \subseteq \mathbb{R}^n$ is a bounded open set.

Let $W \subseteq \partial\Omega$ be an open subset (of the metric space $\partial\Omega$). We say that W is a **standard C^1-graph** (with respect to Ω) if there exist an open set $W' \subseteq \mathbb{R}^{n-1}$, an open interval $(a,b) \subseteq \mathbb{R}$ and a C^1-function $g\colon W' \to (a,b)$ such that $W = \big\{(y, g(y))\,;\ y \in W'\big\}$, i.e. W is the graph of g, and for every $(y,t) \in W' \times (a,b)$ one has



$$(y,t) \in \Omega \quad \text{if and only if} \quad t < g(y).$$

It is easy to see that then $(y,t) \notin \overline{\Omega}$ if and only if $t > g(y)$. The set W is called a $\boldsymbol{C^1}$**-graph** (with respect to Ω) if there exists an orthogonal matrix $B \in \mathbb{R}^{n\times n}$ such that $B(W)$ is a standard C^1-graph with respect to $B(\Omega)$. This means of course that W is a standard C^1-graph with respect to another cartesian coordinate system. We say that Ω has $\boldsymbol{C^1}$**-boundary** if for each $z \in \partial\Omega$ there exists an open neighbourhood $W \subseteq \partial\Omega$ of z such that W is a C^1-graph with respect to Ω.

Similarly we now define other regularity properties of the boundary.

7.1 Remarks. (C^k-, Lipschitz, continuous boundary)

(a) For $k \in \mathbb{N} \cup \{\infty\}$ we call W a $\boldsymbol{C^k}$**-graph** if the function $g\colon W' \to (a,b)$ in the definition given above is a C^k-function. We speak of a **Lipschitz graph** if g is Lipschitz continuous (and not necessarily C^1). We call W a **continuous graph** if we merely require that g is continuous.

(b) We say that Ω has $\boldsymbol{C^k}$**-boundary** (**Lipschitz boundary**, **continuous boundary**) if for each $z \in \partial\Omega$ there exists an open neighbourhood $W \subseteq \partial\Omega$ of z such that W is a C^k-graph (Lipschitz graph, continuous graph). △

In this way we have defined a hierarchy of regularity properties. Continuous boundary is the weakest property we consider – it will suffice for many of our results – and C^∞-boundary is the strongest. If Ω has C^1-boundary, then it also has Lipschitz boundary. Each polygon in $\mathbb{R}^2$ and each convex polyhedron in $\mathbb{R}^3$ has Lipschitz boundary, but not C^1-boundary; so there are good reasons to consider Lipschitz boundary. Nevertheless we will assume C^1-boundary in many results that would only require Lipschitz boundary, because then things become much easier.

Next we introduce the outer (or exterior) normal of an open set with C^1-boundary. It can be characterised intrinsically by basic geometrical properties as follows; see [ArUr23; Theorem 7.4 and its proof].

7.2 Remark. Assume that Ω has C^1-boundary. Then for each $z \in \partial\Omega$ there is a unique vector $\nu(z) \in \mathbb{R}^n$ satisfying

(i) $|\nu(z)| = 1$;

(ii) if $\gamma \in C^1(-1,1;\mathbb{R}^n)$ is such that $\gamma(0) = z$ and $\gamma(t) \in \partial\Omega$ for all $t \in (-1,1)$, then $\nu(z) \perp \gamma'(0)$;

(iii) there exists $\varepsilon > 0$ such that $z + t\nu(z) \notin \overline{\Omega}$ (and $z - t\nu(z) \in \Omega$) for all $0 < t < \varepsilon$.

We call $\nu(z)$ the **outer unit normal** at z. Condition (ii) says that $\nu(z)$ is orthogonal to the boundary and (iii) that $\nu(z)$ points out of Ω.

If W is as in the description of a standard C^1-graph, then for $z = (y, g(y)) \in W$ the outer unit normal is given by

$$\nu(z) = \frac{1}{\sqrt{|\nabla g(y)|^2 + 1}} \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix}.$$

The function ν is continuous on $\partial\Omega$ with values in $\mathbb{R}^n$. △

We assume that the reader is acquainted with the integration of functions on $(n-1)$-dimensional C^1-manifolds in $\mathbb{R}^n$. We point out that the boundary $\partial\Omega$ of a bounded open set Ω with C^1-boundary is an $(n-1)$-dimensional C^1-manifold. The **surface measure** on

$\partial\Omega$ will be denoted by σ. The following formula is basic and can be taken as the definition of σ. If Ω is a bounded open set with C^1-boundary, $W \subseteq \partial\Omega$ is a standard C^1-graph, and W', g are as in the definition, then

$$\int_{\partial\Omega} h(z)\,\mathrm{d}\sigma(z) = \int_{W'} h(y, g(y))\sqrt{1 + |\nabla g(y)|^2}\,\mathrm{d}y$$

for all $h \in C(\partial\Omega)$ with support in W. The weight factor in the integral on the right-hand side is such that the $(n-1)$-dimensional Lebesgue measure on W' is transferred to the appropriate Borel measure σ on the $(n-1)$-dimensional manifold $\partial\Omega$. For more information about the surface measure σ we refer to Section A.1; in particular see formula (A.3).

We insert a few comments concerning Borel measures on a topological space (X, τ). The **Borel σ-algebra** $\mathcal{B}$ of X is defined as the σ-algebra generated by τ, i.e. as the smallest σ-algebra containing all open sets in X. The members of $\mathcal{B}$ are called **Borel sets**. A **Borel measure** μ on X is a (positive) locally finite measure defined on $\mathcal{B}$, where 'locally finite' means that every point of X possesses an open neighbourhood U with $\mu(U) < \infty$. If X is compact, then every Borel measure on X is finite.

We can now formulate Gauss' theorem. By $C^1(\overline{\Omega})$ we denote the space of all functions $u \in C(\overline{\Omega}) \cap C^1(\Omega)$ for which $\partial_j u$ has a continuous extension to $\overline{\Omega}$ for each $j \in \{1, \dots, n\}$; we keep the notation $\partial_j u$ for this extension. (Concerning the notation $C^1(\overline{\Omega})$ we point out Exercise 7.1.)

7.3 Theorem (Gauss). *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Then for all $u \in C^1(\overline{\Omega})$, $j \in \{1, \dots, n\}$ one has*

$$\int_{\Omega} \partial_j u(x)\,\mathrm{d}x = \int_{\partial\Omega} u(z)\nu_j(z)\,\mathrm{d}\sigma(z). \tag{7.1}$$

Here $\nu \in C(\partial\Omega; \mathbb{R}^n)$ is the outer unit normal, $\nu(z) = (\nu_1(z), \dots, \nu_n(z))$.

A proof of Theorem 7.3 is given in Section A.2. We mention that σ is the unique Borel measure on $\partial\Omega$ such that (7.1) holds; see [ArUr23; Section 7.2].

As it should be, the notions introduced above can be interpreted for the case $n = 1$. To get a start, the Lebesgue measure on $\mathbb{R}^0 = \{0\}$ is the counting measure, and all functions $g\colon \mathbb{R}^0 \to \mathbb{R}$ are differentiable with derivative 0. As a consequence, a bounded set $\Omega \subseteq \mathbb{R}$ has C^1-boundary (or continuous boundary, or C^∞-boundary, for that matter) if and only if it is a finite union of bounded open intervals whose closures are pairwise disjoint. The surface measure on $\partial\Omega$, the set of the endpoints of the intervals, is the counting measure.

Specifically, for an interval $\Omega = (a, b)$ we have $\partial\Omega = \{a, b\}$, $\nu(a) = -1$, $\nu(b) = 1$. Applying Gauss' theorem with $u \in C^1[a, b] = C^1(\overline{(a, b)})$ we obtain

$$\int_a^b u'(x)\,\mathrm{d}x = \int_{\{a,b\}} u(z)\nu(z)\,\mathrm{d}\sigma(z) = u(b) - u(a),$$

the second part of the fundamental theorem of calculus.

Next we derive an important consequence of Gauss' theorem. We define $C^2(\overline{\Omega}) := \big\{u \in C^1(\overline{\Omega});\, \partial_j u \in C^1(\overline{\Omega})\ (j = 1, \dots, n)\big\}$. Then for $u \in C^2(\overline{\Omega})$ the functions $\partial_j \partial_k u$ are in $C(\overline{\Omega})$

for all $j, k \in \{1, \ldots, n\}$. For $u \in C^1(\overline{\Omega})$ the function $\partial_\nu u\colon \partial\Omega \to \mathbb{K}$, given by

$$\partial_\nu u(z) := \nu(z) \cdot \nabla u(z) = \sum_{j=1}^{n} \nu_j(z)\partial_j u(z),$$

is called the **normal derivative** of u. Note that $\partial_\nu u \in C(\partial\Omega)$.

7.4 Corollary (Green's formulas). *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary, and let $u \in C^2(\overline{\Omega})$. Then*

$$\int_\Omega (\Delta u)v \,\mathrm{d}x + \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_{\partial\Omega} (\partial_\nu u)v \,\mathrm{d}\sigma \qquad (v \in C^1(\overline{\Omega})), \tag{7.2}$$

$$\int_\Omega (v\Delta u - u\Delta v) \,\mathrm{d}x = \int_{\partial\Omega} (v\partial_\nu u - u\partial_\nu v) \,\mathrm{d}\sigma \qquad (v \in C^2(\overline{\Omega})). \tag{7.3}$$

Proof. By Gauss' theorem one has

$$\int_\Omega (\partial_j^2 u)v \,\mathrm{d}x + \int_\Omega \partial_j u \, \partial_j v \,\mathrm{d}x = \int_\Omega \partial_j\big((\partial_j u)v\big) \,\mathrm{d}x = \int_{\partial\Omega} (\partial_j u)v \,\nu_j \,\mathrm{d}\sigma.$$

Summation over $j = 1, \ldots, n$ yields (7.2).

Exchanging u and v in (7.2) and subtracting the result from (7.2) one obtains (7.3). □

In the literature one finds various definitions of C^1-boundary (all equivalent to our definition). The following equivalence is useful for showing that a given set has C^1-boundary; its proof relies on the implicit function theorem. A bounded open set $\Omega \subseteq \mathbb{R}^n$ has C^1-boundary if and only if for each $z \in \partial\Omega$ there exist an open neighbourhood $U \subseteq \mathbb{R}^n$ of z and a C^1-function $h\colon U \to \mathbb{R}$ such that $\Omega \cap U = [h < 0]$ and $\nabla h(z) \neq 0$ for all $z \in \partial\Omega \cap U$. We mention that then the outer unit normal is given by $\nu(z) = \nabla h(z)/|\nabla h(z)|$ for all $z \in \partial\Omega \cap U$.

For instance, for the open unit ball $\Omega := B(0,1)$ in $\mathbb{R}^n$ the formula $\Omega = \{x \in \mathbb{R}^n;\ |x|^2 - 1 < 0\}$ immediately shows that Ω has C^1-boundary. (In this case the proof by finding local parametrisations of the boundary is not difficult, but carrying out all the details is somewhat tedious.)

7.2 Interlude: more on $H^1(\Omega)$; denseness, trace and compactness

In this section we present some fundamental properties of the Sobolev space $H^1(\Omega)$ that are needed for our treatment of the Neumann and Robin Laplacians in Sections 7.4 and 7.5 below. According to the general philosophy of this book, we give complete proofs of these properties. In a first reading the reader might want to skip some of the more technical proofs and first look at the application of the results.

The first major issue is a denseness property for $H^1(\Omega)$. Denseness results are needed to transfer inequalities or equalities for smooth functions to more general functions. An example for this procedure is the proof of Poincaré's inequality in Chapter 5; further examples follow in this section.

In the proof of the denseness result as well as in later arguments we need the following property of bounded open sets with continuous boundary.

7.5 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with continuous boundary. Then for each $x \in \partial\Omega$ there exist an open neighbourhood $U \subseteq \mathbb{R}^n$ of x and a vector $y \in \mathbb{R}^n \setminus \{0\}$ such that*

$$(\overline{\Omega} \cap U) - \tau y \subseteq \Omega, \quad (\partial\Omega \cap U) + \tau y \subseteq \mathbb{R}^n \setminus \overline{\Omega} \qquad (0 < \tau < 1).$$

Proof. There exists an open neighbourhood $W \subseteq \partial\Omega$ of x such that W is a continuous graph with respect to Ω. Without loss of generality we assume that W is a standard continuous graph. Let $W' \subseteq \mathbb{R}^{n-1}$ and (a, b) be as in the beginning of Section 7.1. Then there exists $\varepsilon > 0$ such that $U := W' \times (a+\varepsilon, b-\varepsilon)$ is an open neighbourhood of x, and one easily sees that the assertions are satisfied for $y = \varepsilon e_n$, with the n-th unit vector e_n. □

The property stated in Lemma 7.5 is in fact equivalent to Ω having continuous boundary. This follows from Exercise 7.2(b), where the reader is asked to show that Ω having continuous boundary is equivalent to the 'segment property' in [Ada75; Chap. III, 3.17] (see property (ii) formulated in Exercise 7.2(b)).

The following lemma contains 'local versions' of Proposition 4.3(a) and Lemma 4.16(b) that will be needed below.

7.6 Lemma. *Let $\Omega, \Omega' \subseteq \mathbb{R}^n$ be open sets, Ω relatively compact in Ω'.*

(a) *Let $\delta > 0$, $\rho \in C_c^\infty(\mathbb{R}^n)$, $\operatorname{spt} \rho \subseteq B(0, \delta)$. Let $f, g \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$, $f = g$ on $\Omega + B(0, \delta)$. Then $\rho * f = \rho * g$ on $\overline{\Omega}$.*

(b) *Let (ρ_k) be a delta sequence in $C_c^\infty(\mathbb{R}^n)$. If $u \in H^1(\Omega')$, then $(\rho_k * \tilde{u})|_\Omega \to u|_\Omega$ in $H^1(\Omega)$ as $k \to \infty$. If $u \in C(\overline{\Omega'})$, then $(\rho_k * \tilde{u})|_{\overline{\Omega}} \to u|_{\overline{\Omega}}$ in $C(\overline{\Omega})$ as $k \to \infty$. (As before, $\tilde{u}$ denotes the extension of u to $\mathbb{R}^n$ by zero.)*

Proof. (a) follows immediately from the definition of convolution.

(b) Either assumption on u implies that $\tilde{u} \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$. There exist $\delta > 0$ and $\varphi_0 \in C_c^\infty(\Omega')$ such that $\varphi_0 = 1$ on $\Omega_0 := \Omega + B(0, \delta)$. Then $u_0 := \varphi_0 u$ satisfies $\tilde{u}_0 = \tilde{u}$ on Ω_0. For $k > 1/\delta$ one has $\operatorname{spt} \rho_k \subseteq B(0, \delta)$ and hence $\rho_k * \tilde{u}_0 = \rho_k * \tilde{u}$ on $\overline{\Omega}$ by part (a).

Now if $u \in H^1(\Omega')$, then $u_0 \in H_c^1(\Omega')$ by Exercise 4.5(c), and thus $\tilde{u}_0 \in H^1(\mathbb{R}^n)$ by Exercise 4.7(a), so $\rho_k * \tilde{u}_0 \to \tilde{u}_0$ in $H^1(\mathbb{R}^n)$ by Lemma 4.16(b). It follows that $(\rho_k * \tilde{u})|_\Omega = (\rho_k * \tilde{u}_0)|_\Omega \to \tilde{u}_0|_\Omega = u|_\Omega$ in $H^1(\Omega)$.

If $u \in C(\overline{\Omega}')$, then $\tilde{u}_0 \in C(\mathbb{R}^n)$, and applying Proposition 4.3(a) one obtains the second assertion of (b). □

We now turn to the announced denseness property for $H^1(\Omega)$.

7.7 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with continuous boundary. Then the set*

$$\check{C}^\infty(\Omega) := \left\{\varphi|_\Omega ;\, \varphi \in C_c^\infty(\mathbb{R}^n)\right\}$$

is dense in $H^1(\Omega)$. In particular, $C^1(\overline{\Omega})$ is dense in $H^1(\Omega)$.

Proof. (i) Let $x \in \partial\Omega$. Choose an open neighbourhood $U \subseteq \mathbb{R}^n$ of x and a vector $y \in \mathbb{R}^n \setminus \{0\}$ as in Lemma 7.5. Let $u \in H^1(\Omega)$ have relatively compact support in U; we show that u can be approximated by functions in $\check{C}^\infty(\Omega)$.

Let $\tilde{u}$ denote the extension of u to $\mathbb{R}^n$ by zero. Then $W_u := \partial\Omega \cap \operatorname{spt} \tilde{u}$ is compact in U, and $W_u + \tau y \subseteq \mathbb{R}^n \setminus \overline{\Omega}$ for all $\tau \in (0,1)$; see Lemma 7.5. By Exercise 7.3 we have $\tilde{u} \in H^1(\mathbb{R}^n \setminus W_u)$ and $\partial_j \tilde{u} = \widetilde{\partial_j u}$ on $\mathbb{R}^n \setminus W_u$ for $j = 1, \dots, n$. We shift $\tilde{u}$ 'outwards', putting $u_\tau := \tilde{u}(\cdot - \tau y)$ for $\tau \in (0,1)$; then

$$u_\tau \in H^1\big(\mathbb{R}^n \setminus (W_u + \tau y)\big), \quad \overline{\Omega} \subseteq \mathbb{R}^n \setminus (W_u + \tau y). \tag{7.4}$$

In particular, $u_\tau|_\Omega \in H^1(\Omega)$ and $\partial_j(u_\tau|_\Omega) = \partial_j \tilde{u}(\cdot - \tau y)|_\Omega$ for $j = 1, \dots, n$. Applying Exercise 7.4(a) we conclude that $u_\tau|_\Omega \to u$ in $H^1(\Omega)$ as $\tau \to 0$. Now fix $\tau \in (0,1)$, and let $(\rho_k)_{k\in\mathbb{N}}$ be a delta sequence in $C_c^\infty(\mathbb{R}^n)$. Then $\rho_k * u_\tau \in C_c^\infty(\mathbb{R}^n)$ for all $k \in \mathbb{N}$, and in view of (7.4), Lemma 7.6(b) shows that $(\rho_k * u_\tau)|_\Omega \to u_\tau|_\Omega$ in $H^1(\Omega)$ as $k \to \infty$. All in all we have established that the functions $(\rho_k * u_\tau)|_\Omega \in \check{C}^\infty(\Omega)$ approximate u in $H^1(\Omega)$.

(ii) A compactness argument shows that $\partial\Omega$ can be covered by open sets $U_1, \dots, U_m \subseteq \mathbb{R}^n$ such that for each $k \in \{1, \dots, m\}$, each function $u \in H^1(\Omega)$ with relatively compact support in U_k can be approximated by functions in $\check{C}^\infty(\Omega)$. Putting $U_0 := \Omega$ we obtain an open covering $(U_k)_{k=0,\dots,m}$ of $\overline{\Omega}$. There exists a partition of unity $(\varphi_k)_{k=0,\dots,m}$ in $C_c^\infty(\mathbb{R}^n)_+$ on $\overline{\Omega}$, subordinate to $(U_k)_{k=0,\dots,m}$; see Exercise 4.3(b).

Let $u \in H^1(\Omega)$. Then $\varphi_k u \in H^1(\Omega)$ for $k = 0, \dots, m$, by Exercise 4.5(c). Now, $\varphi_0 u \in H_c^1(\Omega)$ can be approximated by $C_c^\infty(\Omega)$-functions, by Theorem 4.15(b), and $\varphi_1 u, \dots, \varphi_m u$ can be approximated by $\check{C}^\infty(\Omega)$-functions, by the choice of $U_1, \dots, U_m$. As a consequence, u can be approximated by $\check{C}^\infty(\Omega)$-functions. □

7.8 Remarks. The procedure used in the proof of Theorem 7.7 yields simultaneous approximation with respect to other properties:

(a) If $u \in H^1(\Omega) \cap C(\overline{\Omega})$, then the approximations additionally converge to u in the supremum norm, an observation that will be important in the proof of Theorem 7.11.

Indeed, multiplication of u by C_c^∞-functions does not effect the continuity property of u. Then in step (i) of the proof of Theorem 7.7 one sees that $u_\tau|_{\overline{\Omega}} \to u$ and $(\rho_k * u_\tau)|_{\overline{\Omega}} \to u_\tau|_{\overline{\Omega}}$ uniformly, where the first convergence follows from the uniform continuity of u, and for the second convergence one applies the second assertion of Lemma 7.6(b).

(b) Another instance is the property that positive functions in $H^1(\Omega)$ can be approximated by positive functions in $\check{C}^\infty(\Omega)$. In order to see this, one just has to note that the product as well as the convolution of two positive functions is again a positive function.

(c) For use in the next proof we observe that in step (i) of the above proof one has the estimate $\|u_\tau|_\Omega - u\|_2 \leqslant \tau|y|\|u\|_{H^1}$ for all $\tau \in (0,1)$ (similarly as in Lemma 6.23).

Indeed, by Theorem 7.7 it is sufficient to prove the estimate for $u \in C^1(\overline{\Omega})$. Then $u_\tau(x) - u(x) = \int_0^1 \frac{d}{dt}\tilde{u}(x - t\tau y)\,\mathrm{d}t$ for all $x \in \Omega$. By the Cauchy–Schwarz inequality it follows that $|u_\tau(x) - u(x)|^2 \leqslant \int_0^1 \left|\nabla\tilde{u}(x - t\tau y) \cdot \tau y\right|^2 \mathrm{d}t$, and integration over $x \in \Omega$ yields $\|u_\tau - u\|_2^2 \leqslant \int_0^1 \|\nabla\tilde{u}(\cdot - t\tau y)|_\Omega\|_2^2 \tau^2 |y|^2 \,\mathrm{d}t \leqslant \tau^2|y|^2\|u\|_{H^1}^2$. △

For Ω with continuous boundary we can now transfer the compactness of the embedding $H_0^1(\Omega) \hookrightarrow L_2(\Omega)$ (shown in Theorem 6.21) to $H^1(\Omega)$.

7.9 Theorem (Rellich–Kondrachov). *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with continuous boundary. Then the embedding $j\colon H^1(\Omega) \hookrightarrow L_2(\Omega)$ is compact.*

Proof. (i) We start with the following observation: if $\Omega' \subseteq \mathbb{R}^n$ is open, Ω relatively compact in Ω', then the mapping $R\colon H^1(\Omega') \to L_2(\Omega)$, $Ru := u|_\Omega$ is compact.

Indeed, there exists $\psi \in C_c^\infty(\Omega')$ such that $\psi = 1$ on Ω. We define $B\colon H^1(\Omega') \to L_2(\Omega')$ by $Bu := \psi u$. Since B acts as a bounded operator from $H^1(\Omega')$ to $H_0^1(\Omega')$ (cf. Exercise 4.5(c)), Theorem 6.21 implies that B is compact. Now R is compact since $Ru = (Bu)|_\Omega$ for all $u \in H^1(\Omega')$.

(ii) Let $U_0, \ldots, U_m$ and $\varphi_0, \ldots, \varphi_m$ be as in step (ii) of the proof of Theorem 7.7. There exist vectors $y^1, \ldots, y^m \in \mathbb{R}^n$ (pointing 'out of Ω') such that $(\partial\Omega \cap U_k) + \tau y^k \subseteq \mathbb{R}^n \setminus \overline{\Omega}$ for $k = 1, \ldots, m$ and $\tau \in (0,1)$; see Lemma 7.5. We write $j = \sum_{k=0}^m J_k$, with $J_k \in \mathcal{L}(H^1(\Omega), L_2(\Omega))$, $J_k u := \varphi_k u$ for $k = 0, \ldots, m$. Note that J_0 is compact by the same reasoning as in step (i) above. If $k \in \{1, \ldots, m\}$ and $\tau \in (0,1)$, then for the open set $\Omega' := \mathbb{R}^n \setminus \big((\partial\Omega \cap \operatorname{spt}\varphi_k) + \tau y^k\big)$ the mapping $H^1(\Omega) \ni u \mapsto (\varphi_k \tilde{u})(\cdot - \tau y^k) \in H^1(\Omega')$ is a bounded operator (recall (7.4)), and hence step (i) implies that

$$J_{\tau,k}\colon H^1(\Omega) \to L_2(\Omega), \quad J_{\tau,k} u := (\varphi_k \tilde{u})(\cdot - \tau y^k)|_\Omega$$

is a compact operator. Since $J_{\tau,k} \to J_k$ in $\mathcal{L}(H^1(\Omega), L_2(\Omega))$ as $\tau \to 0$ by Remark 7.8(c), it follows that j is compact. □

We refer to [EdEv87; Theorem 4.17] for a different proof of Theorem 7.9.

The following proposition is an important characterisation of $H_0^1(\Omega)$. Its proof uses a method similar to the proof of Theorem 7.7. For a function $u\colon \Omega \to \mathbb{K}$ we will again denote by $\tilde{u}$ the extension of u to $\mathbb{R}^n$ by zero.

7.10 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with continuous boundary. Then*

$$H_0^1(\Omega) = \big\{u \in H^1(\Omega);\ \tilde{u} \in H^1(\mathbb{R}^n)\big\}.$$

Proof. The inclusion '$\subseteq$' was noted in Remark 4.14(b). (In fact, this inclusion holds for arbitrary open sets Ω.)

In order to prove '$\supseteq$' we assume that $u \in H^1(\Omega)$ is such that $\tilde{u} \in H^1(\mathbb{R}^n)$. Let $\varphi_0, \ldots, \varphi_m$ be as in step (ii) of the proof of Theorem 7.7. There exist vectors $y^1, \ldots, y^m \in \mathbb{R}^n$ (pointing 'out of Ω') such that $\operatorname{spt}(\varphi_k \tilde{u}) - \tau y^k \subseteq \Omega$ and hence $(\varphi_k \tilde{u})(\cdot + \tau y^k)|_\Omega \in H_c^1(\Omega)$ for $k = 1, \ldots, m$, $0 < \tau < 1$ (see Lemma 7.5). Exercise 7.4(b) shows that $(\varphi_k \tilde{u})(\cdot + \tau y^k) \to \varphi_k \tilde{u}$ in $H^1(\mathbb{R}^n)$ as $\tau \to 0$, for $k = 1, \ldots, m$. Since $\varphi_0 \tilde{u} \in H_c^1(\Omega)$ as well, the previous considerations show that u can be approximated by functions belonging to $H_c^1(\Omega)$. This implies that $u \in H_0^1(\Omega)$. □

We point out that in the proof given above the functions $\varphi_k \tilde{u}$ are shifted in the direction opposite to the one in the proof of Theorem 7.7, step (i). In the present case the functions are shifted 'inwards' to produce compact support, in the former proof they are shifted 'outwards' in order to produce smoothness on $\overline{\Omega}$.

Next we show that for a set Ω with C^1-boundary one can define a trace mapping $\operatorname{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$ such that for $u \in C^1(\overline{\Omega})$ one has $\operatorname{tr} u = u|_{\partial\Omega}$. Here and in what follows, $L_2(\partial\Omega)$ denotes the L_2-space with respect to the surface measure σ.

7.11 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Then:*

(a) *There exists $c \geqslant 0$ such that*

$$\|u|_{\partial\Omega}\|^2_{L_2(\partial\Omega)} \leqslant c\|u\|_{L_2(\Omega)}\|u\|_{H^1(\Omega)} \tag{7.5}$$

for all $u \in C^1(\overline{\Omega})$.

(b) *There is a unique bounded linear operator* $\operatorname{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$, *called the* ***trace operator****, such that* $\operatorname{tr} u = u|_{\partial\Omega}$ *for all $u \in C(\overline{\Omega}) \cap H^1(\Omega)$, and then* (7.5) *holds for all $u \in H^1(\Omega)$ (with $u|_{\partial\Omega}$ replaced by* $\operatorname{tr} u$ *on the left-hand side).*

(c) *The mapping* tr *is a compact operator.*

Proof. (a) Let $x \in \partial\Omega$, and let $y := \nu(x)$ be the outer unit normal at x. Since ν is continuous, by Remark 7.2, there exists an open neighbourhood $U \subseteq \mathbb{R}^n$ of x such that $y \cdot \nu(z) \geqslant \frac{1}{2}$ for all $z \in U \cap \partial\Omega$. Let $\varphi \in C_c^\infty(\mathbb{R}^n)$ satisfy $\operatorname{spt}\varphi \subseteq U$, $\varphi \geqslant 0$, and $\varphi = 1$ on a neighbourhood of x. Then $\varphi = 1$ on an open neighbourhood $W_x \subseteq \partial\Omega$ of x.

Let $u \in C^1(\overline{\Omega})$. Then, by Theorem 7.3 (Gauss),

$$\begin{aligned}\frac{1}{2}\int_{W_x} |u|^2 \,\mathrm{d}\sigma &\leqslant \int_{\partial\Omega} (\varphi u)\bar{u} y \cdot \nu \,\mathrm{d}\sigma = \int_\Omega \big(\nabla(\varphi u)\bar{u} + (\varphi u)\nabla\bar{u}\big) \cdot y \,\mathrm{d}x \\ &\leqslant \|\varphi u\|_{H^1}\|u\|_{L_2} + \|\varphi u\|_{L_2}\|u\|_{H^1} \leqslant c_\varphi\|u\|_{L_2}\|u\|_{H^1}\,,\end{aligned}$$

with a constant $c_\varphi > 0$ only depending on φ.

We have shown that for each $x \in \partial\Omega$ there exist an open neighbourhood $W_x \subseteq \partial\Omega$ and a constant $c_x \geqslant 0$ such that

$$\int_{W_x} |u|^2 \,\mathrm{d}\sigma \leqslant c_x\|u\|_{L_2}\|u\|_{H^1}$$

for all $u \in C^1(\overline{\Omega})$. A standard compactness argument completes the proof of (7.5) for $u \in C^1(\overline{\Omega})$.

(b) The inequality (7.5) together with the denseness of $C^1(\overline{\Omega})$ in $H^1(\Omega)$ implies that the mapping $u \mapsto u|_{\partial\Omega}$ has a continuous extension $\operatorname{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$.

So far we only know that $\operatorname{tr} u = u|_{\partial\Omega}$ holds for $u \in C^1(\overline{\Omega})$. In order to prove this equality for $u \in C(\overline{\Omega}) \cap H^1(\Omega)$, we use the remarkable feature of the proof of Theorem 7.7 mentioned in Remark 7.8(a). As explained there, for $u \in C(\overline{\Omega}) \cap H^1(\Omega)$ an approximating sequence (u_k) in $C^1(\overline{\Omega})$ can be chosen that converges to u in $C(\overline{\Omega})$ as well as in $H^1(\Omega)$. For this sequence, $(u_k|_{\partial\Omega})$ converges to $\operatorname{tr} u$ in $L_2(\partial\Omega)$ and uniformly to $u|_{\partial\Omega}$, which implies $\operatorname{tr} u = u|_{\partial\Omega}$.

(c) Let (u_k) be a bounded sequence in $H^1(\Omega)$. By Theorem 7.9 (Rellich–Kondrachov) there exists a subsequence (u_{k_m}) converging in $L_2(\Omega)$, and then (7.5) implies that $(\operatorname{tr} u_{k_m})$ is a Cauchy sequence in $L_2(\partial\Omega)$, hence convergent. □

Sometimes, by abuse of notation, we still write $u|_{\partial\Omega} := \operatorname{tr} u$ for $u \in H^1(\Omega)$. In integrals we will frequently omit the trace notation to make things more readable, writing $\|\operatorname{tr} u\|^2_{L_2(\partial\Omega)} = \int_{\partial\Omega} |u|^2 \,\mathrm{d}\sigma$ for example.

The trace is compatible with our definition of $H_0^1(\Omega)$ as the following result shows.

7.12 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Then*

$$H_0^1(\Omega) = \bigl\{u \in H^1(\Omega)\,;\ \operatorname{tr} u = 0\bigr\}.$$

Proof. The inclusion '$\subseteq$' follows from the continuity of the trace operator and the denseness of $C_c^\infty(\Omega)$ in $H_0^1(\Omega)$.

For the proof of '$\supseteq$' we first note that

$$\int_\Omega \partial_j u\, v\,\mathrm{d}x + \int_\Omega u\,\partial_j v\,\mathrm{d}x = \int_{\partial\Omega} uv\,\nu_j\,\mathrm{d}\sigma \tag{7.6}$$

for all $u, v \in H^1(\Omega)$, $j = 1, \ldots, n$. This is immediate from Theorem 7.3 (Gauss) if $u, v \in \check{C}^\infty(\Omega)$, and (7.6) carries over to general u, v by Theorems 7.7 and 7.11.

As before, the extension to $\mathbb{R}^n$ by zero of a function defined on Ω will be denoted by a tilde. Let $u \in H^1(\Omega)$ be such that $\operatorname{tr} u = 0$. Then (7.6) implies

$$\int_{\mathbb{R}^n} \tilde{u}\,\partial_j\varphi\,\mathrm{d}x = \int_\Omega u\,\partial_j\varphi\,\mathrm{d}x = -\int_\Omega \partial_j u\,\varphi\,\mathrm{d}x$$

for all $\varphi \in C_c^\infty(\mathbb{R}^n)$. This shows that $\tilde{u} \in H^1(\mathbb{R}^n)$ (with distributional derivative $\partial_j\tilde{u} = \widetilde{\partial_j u}$ for $j = 1, \ldots, n$), and hence $u \in H_0^1(\Omega)$, by Proposition 7.10. □

For a different proof of Theorem 7.12 we refer to [Eva10; Sect. 5.5, Theorem 2].

7.3 Weak normal derivative

In this section we define the normal derivative in a weak sense for certain functions in $H^1(\Omega)$, by requiring the validity of Green's formula (7.2). Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary, and let $u \in H^1(\Omega)$, $\Delta u \in L_2(\Omega)$. We say that $\partial_\nu u \in L_2(\partial\Omega)$ if there exists $h \in L_2(\partial\Omega)$ such that

$$\int_\Omega (\Delta u)v\,\mathrm{d}x + \int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}x = \int_{\partial\Omega} hv\,\mathrm{d}\sigma \qquad (v \in H^1(\Omega)).$$

In this case we define the **weak normal derivative** of u by $\partial_\nu u := h$. In order to prove the uniqueness of h we note that $C^1(\overline{\Omega}) \subseteq H^1(\Omega)$. The Stone–Weierstrass theorem (see Appendix B, Theorem B.2) implies that the set $\bigl\{\varphi|_{\partial\Omega}\,;\ \varphi \in C^1(\overline{\Omega})\bigr\}$ is dense in $C(\partial\Omega)$. As $C(\partial\Omega)$ is dense in $L_2(\partial\Omega)$ – see Theorem G.9 – we obtain the uniqueness. (Recall that $L_2(\partial\Omega)$ is understood with respect to the surface measure σ.)

Now let $\Omega \subseteq \mathbb{R}^n$ be any open set, and let $u \in H^1(\Omega)$ be such that $\Delta u \in L_2(\Omega)$. If Ω is bounded and has C^1-boundary, then by the definition given above, $\partial_\nu u = 0$ if and only if

$$\int_\Omega (\Delta u)v\,\mathrm{d}x + \int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}x = 0 \qquad (v \in H^1(\Omega)). \tag{7.7}$$

It is remarkable that this condition makes sense for an arbitrary open set $\Omega \subseteq \mathbb{R}^n$. Therefore, for a function $u \in H^1(\Omega)$ with $\Delta u \in L_2(\Omega)$ we will write '$\partial_\nu u = 0$' (including the quotes!) if (7.7) holds.

7.4 The Neumann Laplacian

Let $\Omega \subseteq \mathbb{R}^n$ be an open set (not necessarily bounded). Using the notation '$\partial_\nu u = 0$', introduced in Section 7.3, we define the **Laplacian with Neumann boundary conditions** or simply **Neumann Laplacian** Δ_{N} in $L_2(\Omega)$ by

$$\begin{aligned} \operatorname{dom}(\Delta_{\mathrm{N}}) &:= \bigl\{u \in H^1(\Omega);\, \Delta u \in L_2(\Omega),\ \text{`}\partial_\nu u = 0\text{'}\bigr\}, \\ \Delta_{\mathrm{N}} u &:= \Delta u \qquad (u \in \operatorname{dom}(\Delta_{\mathrm{N}})). \end{aligned}$$

7.13 Theorem. *The negative Neumann Laplacian $-\Delta_{\mathrm{N}}$ is self-adjoint and accretive; it is associated with the classical Dirichlet form on $H^1(\Omega)$.*

Proof. Define $a\colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{K}$ by $a(u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x$. Then a is bounded. We consider $H^1(\Omega)$ as a subspace of $L_2(\Omega)$. Since $a(u) + \|u\|_{L_2(\Omega)}^2 = \|u\|_{H^1(\Omega)}^2$, the form a is quasi-coercive. Moreover a is symmetric and accretive. Let $A \sim a$. We show that $A = -\Delta_{\mathrm{N}}$; then the assertions follow.

Let $u \in \operatorname{dom}(A)$, $Au = f$. Then by definition $u \in H^1(\Omega)$ and $\int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x = \int_\Omega f\bar{v}\,\mathrm{d}x$ for all $v \in H^1(\Omega)$. Inserting test functions $v \in C_c^\infty(\Omega)$ one obtains $-\Delta u = f$. Thus $\int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x + \int_\Omega (\Delta u)\bar{v}\,\mathrm{d}x = 0$ for all $v \in H^1(\Omega)$, i.e. '$\partial_\nu u = 0$'. We have shown that $A \subseteq -\Delta_{\mathrm{N}}$. Conversely, if $u \in \operatorname{dom}(\Delta_{\mathrm{N}})$ and $-\Delta_{\mathrm{N}} u = f$, then

$$a(u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x + \int_\Omega (\Delta u)\bar{v}\,\mathrm{d}x + \int_\Omega f\bar{v}\,\mathrm{d}x = \int_\Omega f\bar{v}\,\mathrm{d}x \qquad (v \in H^1(\Omega)).$$

Thus $u \in \operatorname{dom}(A)$ and $Au = f$. □

Applying Theorem 7.9 we now conclude that A has compact resolvent if Ω satisfies our weakest regularity property.

7.14 Theorem (Spectral decomposition of the Neumann Laplacian). *If $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with continuous boundary, then Δ_{N} has compact resolvent. There exist an orthonormal basis $(\varphi_k)_{k\in\mathbb{N}}$ of $L_2(\Omega)$ and an increasing sequence $(\lambda_k)_{k\in\mathbb{N}}$ in $[0,\infty)$, with $\lambda_1 = 0$ and $\lim_{k\to\infty} \lambda_k = \infty$, such that $-\Delta_{\mathrm{N}}$ is the associated diagonal operator. In particular, $\varphi_k \in \operatorname{dom}(\Delta_{\mathrm{N}})$ and*

$$-\Delta_{\mathrm{N}} \varphi_k = \lambda_k \varphi_k$$

for all $k \in \mathbb{N}$. If Ω is connected, then $\lambda_2 > 0$.

Proof. From Theorem 7.9 (Rellich–Kondrachov) we know that the embedding $j\colon H^1(\Omega) \hookrightarrow L_2(\Omega)$ is compact. Therefore Proposition 6.18 – in combination with Theorem 7.13 – implies that Δ_{N} has compact resolvent.

The statement concerning the eigenfunctions and eigenvalues now follows from Theorem 6.16, except for the properties of λ_1 and λ_2. However, it is immediate that $\varphi_1 = \operatorname{vol}_n(\Omega)^{-1/2}\mathbf{1}_\Omega$ is an eigenfunction of $-\Delta_{\mathrm{N}}$ with eigenvalue 0.

Now assume that Ω is connected. If $\varphi \in \operatorname{dom}(\Delta_{\mathrm{N}})$ satisfies $-\Delta_{\mathrm{N}}\varphi = 0$, then $\|\nabla\varphi\|_2^2 = (-\Delta_{\mathrm{N}}\varphi \,|\, \varphi) = 0$; hence Lemma 7.15 – proved below – implies that φ is constant. This shows that the eigenspace belonging to the eigenvalue $\lambda_1 = 0$ is one-dimensional. □

The property that for connected Ω one has $\lambda_2 > 0$, together with the formula for the associated semigroup T from Remark 6.7, gives important information concerning the asymptotic behaviour of $T(t)$ as $t \to \infty$.

7.15 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be a connected open set, and let $u \in L_{1,\mathrm{loc}}(\Omega)$, $\nabla u = 0$ in the distributional sense. Then u is constant.*

Proof. For each $x \in \Omega$ we will construct a neighbourhood on which u has a constant representative; then the assertion follows because Ω is connected. Let $\varphi \in C_c^\infty(\Omega)$ be such that $\varphi = 1$ in a neighbourhood of x, and let v denote the extension of φu to $\mathbb{R}^n$ by zero. Then $\partial_j v \in L_1(\mathbb{R}^n)$ by Exercises 4.5(c) and 4.6(b), for $j = 1, \dots, n$, and $\nabla v = 0$ in a neighbourhood of x.

Let (ρ_k) be a delta sequence in $C_c^\infty(\mathbb{R}^n)$; then $\rho_k * v \to v$ in $L_1(\mathbb{R}^n)$. Moreover $\nabla(\rho_k * v) = \rho_k * \nabla v$ by Lemma 4.16(a), and by Lemma 7.6(a) there exists a neighbourhood $U \subseteq \Omega$ of x such that $u = v$ and $\rho_k * \nabla v = 0$ on U for large enough k. We conclude that $u|_U = v|_U$ is the limit of constant functions and thus has a constant representative. $\square$

7.5 The Robin Laplacian

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Given $\beta \in L_\infty(\partial\Omega)$, we define the **Laplacian with Robin boundary conditions** or briefly **Robin Laplacian** Δ_β in $L_2(\Omega)$ by

$$\begin{aligned} \mathrm{dom}(\Delta_\beta) &:= \big\{u \in H^1(\Omega);\ \Delta u \in L_2(\Omega),\ \partial_\nu u + \beta u|_{\partial\Omega} = 0\big\}, \\ \Delta_\beta u &:= \Delta u \qquad (u \in \mathrm{dom}(\Delta_\beta)). \end{aligned}$$

Note that the condition '$\partial_\nu u + \beta u|_{\partial\Omega} = 0$' should be read as '$\partial_\nu u = -\beta \operatorname{tr} u$', with the weak normal derivative $\partial_\nu u$ as defined in Section 7.3.

7.16 Theorem. *Let β be real-valued. Then the operator $-\Delta_\beta$ is self-adjoint and quasi-accretive, with compact resolvent. In particular, Δ_β generates a quasi-contractive C_0-semigroup T_β on $L_2(\Omega)$. If $\beta \geqslant 0$, then $-\Delta_\beta$ is accretive and $\|T_\beta(t)\| \leqslant 1$ for all $t \geqslant 0$.*

Proof. Define the form $a\colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{K}$ by $a(u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x + \int_{\partial\Omega} \beta u \bar{v}\,\mathrm{d}\sigma$. Then $|a(u,v)| \leqslant \|\nabla u\|_2 \|\nabla v\|_2 + \|\beta\|_{L_\infty(\partial\Omega)} \|\operatorname{tr} u\|_{L_2(\partial\Omega)} \|\operatorname{tr} v\|_{L_2(\partial\Omega)}$. Since the trace operator $\operatorname{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$ is bounded, it follows that a is bounded.

We consider $H^1(\Omega)$ as a subspace of $L_2(\Omega)$ and show that a is quasi-coercive. By Theorem 7.11 and the 'Peter–Paul inequality' (i.e. Young's inequality, $ab \leqslant \frac{1}{2}\big(\gamma a^2 + \frac{1}{\gamma} b^2\big)$ for all $a, b \geqslant 0$, $\gamma > 0$) there exists $c > 0$ such that

$$\Big|\int_{\partial\Omega} \beta |u|^2\,\mathrm{d}\sigma\Big| \leqslant \|\beta\|_{L_\infty(\partial\Omega)} \|\operatorname{tr} u\|_{L_2(\partial\Omega)}^2 \leqslant \frac{1}{2}\|u\|_{H^1(\Omega)}^2 + c\|u\|_{L_2(\Omega)}^2$$

for all $u \in H^1(\Omega)$. This implies

$$a(u) \geqslant \int_\Omega |\nabla u|^2\,\mathrm{d}x - \frac{1}{2}\|u\|_{H^1(\Omega)}^2 - c\|u\|_{L_2(\Omega)}^2 = \frac{1}{2}\|u\|_{H^1(\Omega)}^2 - (1+c)\|u\|_{L_2(\Omega)}^2$$

for all $u \in H^1(\Omega)$, and thus a is quasi-coercive.

Let A be the operator associated with a. We show that $A = -\Delta_\beta$. Let $(u, f) \in A$. Then $u \in H^1(\Omega)$ and

$$\int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x + \int_{\partial\Omega} \beta u \bar{v}\,\mathrm{d}\sigma = \int_\Omega f \bar{v}\,\mathrm{d}x \qquad (v \in H^1(\Omega)). \tag{7.8}$$

Taking $v \in C_\mathrm{c}^\infty(\Omega)$ we see that $-\Delta u = f$. Replacing f by $-\Delta u$ in (7.8) we find

$$\int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x + \int_\Omega (\Delta u)\bar{v}\,\mathrm{d}x = -\int_{\partial\Omega} \beta u \bar{v}\,\mathrm{d}\sigma \qquad (v \in H^1(\Omega)), \tag{7.9}$$

i.e. $\partial_\nu u = -\beta u|_{\partial\Omega}$. Thus $(u, f) \in -\Delta_\beta$. Conversely, if $u \in \mathrm{dom}(\Delta_\beta)$, then (7.9) holds. Putting $f = -\Delta u$ we obtain (7.8) and thus $(u, f) \in A$.

Since a is symmetric and quasi-coercive, A is self-adjoint and quasi-accretive. By Theorem 7.9 (Rellich–Kondrachov) the embedding $H^1(\Omega) \hookrightarrow L_2(\Omega)$ is compact, so A has compact resolvent.

Finally, for $\beta \geqslant 0$ the form a is accretive, and this implies the last assertion of the theorem. □

Notes

In Section 7.1 we partially follow [ArUr23]. It is possible to extend Gauss' theorem to open sets with Lipschitz boundary, for which a suitable surface measure on $\partial\Omega$ is defined analogously. The theorem of Gauss is due to Lagrange in 1792 but has been rediscovered by Carl Friedrich Gauss in 1813, by George Green in 1825, and by Mikhail V. Ostrogradsky in 1831. For this reason one finds it in the literature under these different names. Obviously it can also be written as

$$\int_\Omega \operatorname{div} u\,\mathrm{d}x = \int_{\partial\Omega} u \cdot \nu\,\mathrm{d}\sigma,$$

for each vector field $u \in C^1(\overline{\Omega}; \mathbb{R}^n)$. In this form it is frequently called the **divergence theorem**. Physicists and engineers love this version of the theorem because of its immediate interpretation.

Victor Gustave Robin (1855–1897) was a French mathematician. (The reader should correctly pronounce the nasal in the second syllable of "Robin".) He was teaching mathematical physics at the Sorbonne in Paris. Not much is known about him since he burnt his manuscripts. But he worked on thermodynamics, and the Russian school introduced the name Robin boundary conditions. In fact, these boundary conditions had already been introduced by Isaac Newton (1643–1727). We refer to [GuAb98a], [GuAb98b] for the interesting history of Gustave Robin and "his" boundary conditions. Neumann boundary conditions carry their name to honour Carl G. Neumann (1832–1925) who was a professor at Halle, Basel, Tübingen and Leipzig. He introduced the Neumann series for matrices.

Using the tools presented in this chapter we can now continue the considerations started at the end of the Notes of Chapter 5. In his lectures in Berlin, Dirichlet introduced methods for solving partial differential equations via forms. The most famous is the Dirichlet problem, which is the following. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, and let $g \in C(\partial\Omega)$. The classical **Dirichlet problem** consists in finding a function $u \in C(\overline{\Omega}) \cap C^2(\Omega)$ such that $\Delta u = 0$ on Ω and $u|_{\partial\Omega} = g$. The term "Dirichlet problem" was coined by Bernhard Riemann, in honour of his teacher. And in fact Dirichlet promoted a strategy to its solution. The solution should be the function u with $u|_{\partial\Omega} = g$ having the least energy $\int_\Omega |\nabla u|^2 \,\mathrm{d}x$. On the basis of our knowledge obtained so far, we are able to make this precise, albeit in a slightly different context. Let us assume that Ω has C^1-boundary and that $g\colon \partial\Omega \to \mathbb{R}$ is the trace of some $G \in H^1(\Omega)$, i.e. $g = \operatorname{tr} G$. Then the following holds.

Dirichlet's principle. *There is a unique function $u \in H^1(\Omega)$ such that $\Delta u = 0$ and $\operatorname{tr} u = g$. Moreover u is the unique minimiser of the function*

$$G + H_0^1(\Omega) = \{v \in H^1(\Omega)\,;\ \operatorname{tr} v = g\} \ni v \mapsto \int_\Omega |\nabla v|^2 \,\mathrm{d}x. \tag{7.10}$$

(See Theorem 7.12 for the equality in (7.10). Weyl's lemma, see Appendix C, implies that the solution u in fact belongs to $C^\infty(\Omega)$.)

For the proof observe that for $u \in H^1(\Omega)$ one has $\Delta u = 0$ and $\operatorname{tr} u = g$ if and only if for the function $w := G - u$ one has $w \in H_0^1(\Omega)$ and

$$\int_\Omega \nabla w \cdot \nabla v \,\mathrm{d}x = \int_\Omega \nabla G \cdot \nabla v \,\mathrm{d}x \qquad (v \in H_0^1(\Omega)), \tag{7.11}$$

and that by the Fréchet–Riesz theorem there is a unique $w \in H_0^1(\Omega)$ satisfying (7.11). (Recall that, by Poincaré's inequality, the scalar product $(u \,|\, v)_0 := \int \nabla u \cdot \nabla v \,\mathrm{d}x$ is equivalent to the standard scalar product on $H_0^1(\Omega)$.) Moreover, by the 'variational method' formulated in the Notes of Chapter 5, w is the unique minimiser of the function

$$H_0^1(\Omega) \ni v \mapsto \int_\Omega |\nabla v|^2 \,\mathrm{d}x - 2\int_\Omega \nabla G \cdot \nabla v \,\mathrm{d}x = \int |\nabla (G - v)|^2 \,\mathrm{d}x - \int |\nabla G|^2 \,\mathrm{d}x.$$

This shows that $u = G - w \in G + H_0^1(\Omega)$ is the unique minimiser of (7.10).

In 1906, Hadamard [Had06] showed that there are cases in which the Dirichlet problem cannot be treated by the variational method even if the set Ω is 'nice' and a classical solution exists. More precisely, for $\Omega = B_{\mathbb{R}^2}(0,1)$ he constructed a function $u \in C(\overline{\Omega}) \cap C^2(\Omega)$ satisfying $\Delta u = 0$ and $u \notin H^1(\Omega)$; then u is the unique solution corresponding to the boundary function $g = u|_{\partial\Omega} \in C(\partial\Omega)$ (see e.g. [ArUr23; Example 6.67]). This phenomenon does not occur for more regular boundary data g.

The classical Dirichlet problem was finally settled by Oskar Perron [Per23] in 1923. His treatment can be considered as the final point of a development in which Poincaré, Lebesgue, Courant, Lichtenstein, Zaremba and others had contributed. The existence of a solution requires some boundary regularity. In particular, if $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with Lipschitz boundary, then the Dirichlet problem has a solution $u \in C(\overline{\Omega}) \cap C^2(\Omega)$ for any $g \in C(\partial\Omega)$. We refer to [WKK09; Satz 3.3.10, Satz 3.4.3 and Bemerkung 3.4.4] for a proof of this result; see also [Joh82; Section 4.4]. The relation between Perron's and the variational solutions is discussed in [ArDa08a]. For a thorough historical account of the Dirichlet problem and Dirichlet's principle in the 19th and early 20th centuries we refer to [Går79], [WKK09; Section 1.4].

A bounded open set Ω is called **Dirichlet regular** if for each $g \in C(\partial\Omega)$ there exists a (necessarily unique, see Exercise C.3) function $u \in C(\overline{\Omega})$ such that $\Delta u = 0$ and $u|_{\partial\Omega} = g$. In this terminology, Perron's result implies that every bounded open set with Lipschitz-boundary is Dirichlet regular. Lebesgue's cusp provides an example of an open bounded set $\Omega \subseteq \mathbb{R}^3$ that has continuous boundary but is not Dirichlet regular (see e.g. [CoHi62; Chap. IV, §1.4], [ArDa08b; Section 7]).

Exercises

7.1 In this exercise we illustrate that the space $C^1(\overline{\Omega})$ does not only depend on $\overline{\Omega}$. Let $\Gamma \subseteq [0,1]$ denote Cantor's ternary set, and let $\Omega := (0,1) \setminus \Gamma$. Then $\overline{\Omega} = [0,1] = \overline{(0,1)}$. Let $u\colon [0,1] \to [0,1]$ be the "Cantor function", the increasing function generating the usual singular continuous measure on Γ; to wit,

$$\begin{aligned} &u(0) = 0, && u = 1/2 \text{ on } (1/3, 2/3), && u(1) = 1, \\ &u = 1/4 \text{ on } (1/9, 2/9), && u = 3/4 \text{ on } (7/9, 8/9), \\ &u = 1/8 \text{ on } (1/27, 2/27), && \text{etc.} \end{aligned}$$

and u extended to $[0,1]$ by continuity. Show that $u \in C^1(\overline{\Omega}) \setminus C^1(\overline{(0,1)})$.

7.2 (a) Let M be a compact metric space, and let $g\colon M \to \mathbb{R}$ be such that the graph $\{(x, g(x))\,;\, x \in M\}$ of g is compact. Show that g is continuous.

(b) Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Show that the following properties are equivalent:

(i) Ω has continuous boundary;

(ii) Ω has the 'segment property', i.e. for each $x \in \partial\Omega$ there exist an open neighbourhood $U \subseteq \mathbb{R}^n$ of x and a vector $y \in \mathbb{R}^n$ such that $(\overline{\Omega} \cap U) - (0,1)y \subseteq \Omega$.

(The implication '(i) ⇒ (ii)' was shown in Lemma 7.5.)

Hint: Transform property (ii) to the 'standard situation' in which y is a positive multiple of the n-th unit vector e_n. Without loss of generality suppose that $U = W' \times (a,b)$ with an open set $W' \subseteq \mathbb{R}^{n-1}$ and $a, b \in \mathbb{R}$, $a < b$. For $z \in W := \partial\Omega \cap U$ show that $(z - (0,\infty)y) \cap U \subseteq \Omega$ and $(z + (0,\infty)y) \cap U \subseteq \mathbb{R}^n \setminus \overline{\Omega}$. Then apply part (a) to show that W is the graph of a continuous function $g\colon W_0' \to (a,b)$, with an open(!) set $W_0' \subseteq W'$. (The point is that W is the graph of a function g that is not continuous a priori, but will be continuous automatically, by part (a).)

7.3 Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $u \in H^1(\Omega)$, $\tilde{u}$ the extension of u to $\mathbb{R}^n$ by zero.

(a) Show that $\operatorname{spt}\tilde{u} = \overline{\operatorname{spt} u}$.

(b) Define $W_u := \operatorname{spt}\tilde{u} \cap \partial\Omega$. Show that $\tilde{u} \in H^1(\mathbb{R}^n \setminus W_u)$, $\partial_j \tilde{u} = \widetilde{\partial_j u}$ on $\mathbb{R}^n \setminus W_u$ for $j = 1, \dots, n$. (Hint: Use Exercise 4.3(c).)

7.4 (a) Let $1 \leqslant p < \infty$. For $x \in \mathbb{R}^n$ define $T_x \in \mathcal{L}(L_p(\mathbb{R}^n))$ by $T_x f := f(\cdot - x)$ for $f \in L_p(\mathbb{R}^n)$. Show that the function $\mathbb{R}^n \ni x \mapsto T_x \in \mathcal{L}(L_p(\mathbb{R}^n))$ is strongly continuous.

Hint: Show that $x \mapsto T_x\varphi \in L_p(\mathbb{R}^n)$ is continuous for all $\varphi \in C_c(\mathbb{R}^n)$, and use the denseness of $C_c(\mathbb{R}^n)$ in $L_p(\mathbb{R}^n)$.

(b) Let $u \in H^1(\mathbb{R}^n)$. Show that $u(\cdot - x) \to u$ in $H^1(\mathbb{R}^n)$ as $x \to 0$.

7.5 Let $\Omega \subseteq \mathbb{R}^n$ be a connected bounded open set with C^1-boundary. Let $0 < \beta \in L_\infty(\partial\Omega)$ (i.e. $0 \leqslant \beta \in L_\infty(\partial\Omega) \setminus \{0\}$). Denote by T_β the C_0-semigroup on $L_2(\Omega)$ generated by the Robin Laplacian Δ_β. Show that

$$\|T_\beta(t)\| \leqslant \mathrm{e}^{-\varepsilon t} \qquad (t \geqslant 0)$$

for some $\varepsilon > 0$. (Hint: Use Lemma 7.15.)

7.6 Let $\Omega \subseteq \mathbb{R}^n$ be a C^1-domain with a hole; more precisely, assume that there exist bounded open sets $\widetilde{\Omega}$ and ω with C^1-boundary such that $\overline{\omega} \subseteq \widetilde{\Omega}$ and $\Omega = \widetilde{\Omega} \setminus \overline{\omega}$. Then one may write $\partial\Omega = \Gamma_1 \cup \Gamma_2$, where $\Gamma_1 := \partial\widetilde{\Omega}$ and $\Gamma_2 := \partial\omega$. Let $\beta \in L_\infty(\Gamma_2)$ be real-valued. Define the Laplacian with Robin boundary condition $\partial_\nu u + \beta u = 0$ on Γ_2 and Dirichlet boundary condition zero on Γ_1, and show that it is a self-adjoint operator. (It is part of the task to give a proper definition of the indicated boundary conditions and to 'translate' the boundary conditions into a suitable bounded quasi-coercive form.)

7.7 Let $a\colon V \times V \to \mathbb{C}$ be a symmetric bounded form that is j-coercive, where V, H are complex Hilbert spaces and $j \in \mathcal{L}(V, H)$ has dense range. Let $b\colon V \times V \to \mathbb{C}$ be a bounded form and assume that there exists $c \geqslant 0$ such that

$$|b(u)| \leqslant c\|u\|_V \|j(u)\|_H \qquad (u \in V).$$

(a) Show that $a + b\colon V \times V \to \mathbb{C}$ is j-coercive.

(b) Denote by A the (quasi-m-sectorial) operator associated with $(a + b, j)$. Show that the numerical range $\operatorname{num}(A)$ is contained in the region 'surrounded' by a parabola with vertex on the real axis and opened in the direction of the positive real axis.

(c) Show that A is quasi-m-sectorial of any angle $\varphi \in (0, \pi/2)$ and that $-A$ generates a holomorphic C_0-semigroup of angle $\pi/2$. (Hint: Look at the operator $A+\omega$, for arbitrarily large $\omega \in \mathbb{R}$, and remember rescaled semigroups.)

7.8 Let $\mathbb{K} = \mathbb{C}$, and let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Let $\beta \in L_\infty(\partial\Omega)$ (not necessarily real-valued), and let Δ_β be the Robin Laplacian in $L_2(\Omega)$.

(a) Show that $-\Delta_\beta$ has the properties described in parts (b) and (c) of Exercise 7.7.

(b) Suppose that $\operatorname{Re}\beta \geqslant 0$. Show that $-\Delta_\beta$ is accretive and that Δ_β generates a contractive C_0-semigroup on $L_2(\Omega)$.

7.9 Convince yourself that Dirichlet's principle from the above Notes holds for arbitrary bounded open sets $\Omega \subseteq \mathbb{R}^n$ if formulated as follows: for each $G \in H^1(\Omega)$ there exists a unique function $u \in H^1(\Omega)$ such that $\Delta u = 0$, $G - u \in H^1_0(\Omega)$, and u is the unique minimiser of $G + H^1_0(\Omega) \ni v \mapsto \|\nabla v\|_2^2$.

Chapter 8

The Dirichlet-to-Neumann operator

The Dirichlet-to-Neumann operator plays an important role in the theory of inverse problems. For instance, from measurements of electrical currents at the surface of the human body one wishes to determine conductivity inside the body. But the Dirichlet-to-Neumann operator also plays a big role in many other parts of analysis. Here we use form methods to show that it is a self-adjoint operator in $L_2(\partial\Omega)$. It becomes important that our setting allows the mapping j to be non-injective: throughout this chapter j will be the trace operator.

8.1 The Dirichlet-to-Neumann operator for the Laplacian

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. We use the classical Dirichlet form

$$a(u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x \qquad (u,v \in H^1(\Omega)). \tag{8.1}$$

If we choose the canonical injection of $H^1(\Omega)$ into $L_2(\Omega)$, the associated operator is the Neumann Laplacian. Here we will choose as j the trace operator from $H^1(\Omega)$ to $L_2(\partial\Omega)$, introduced in Theorem 7.11; then j has dense range, as noted in Section 7.3. We will show that the form a is j-coercive and investigate the self-adjoint operator in $L_2(\partial\Omega)$ associated with (a,j). It turns out that this is the Dirichlet-to-Neumann operator D_0 which maps the 'Dirichlet data' $u|_{\partial\Omega} \in L_2(\partial\Omega)$ of a harmonic function $u \in H^1(\Omega)$ to the 'Neumann data' $\partial_\nu u \in L_2(\partial\Omega)$; see Theorem 8.3 below. (A function u on Ω is called **harmonic** if it is twice continuously differentiable and $\Delta u = 0$.)

For the proof of the j-coercivity of a we need an auxiliary result, which is a version of what sometimes comes under the heading "Ehrling's lemma". We state it for the general case of Banach spaces; in our context it will only be needed for Hilbert spaces.

8.1 Lemma. *Let X, Y, Z be Banach spaces, X reflexive, $K \in \mathcal{L}(X,Y)$ compact and $S \in \mathcal{L}(X,Z)$ injective. Then for all $\varepsilon > 0$ there exists $c_\varepsilon \geqslant 0$ such that*

$$\|Kx\|_Y \leqslant \varepsilon\|x\|_X + c_\varepsilon\|Sx\|_Z \qquad (x \in X).$$

Proof. For a contradiction, assume that there exists $\varepsilon > 0$ such that for each $n \in \mathbb{N}$ there exists $x_n \in X$ with

$$\|Kx_n\|_Y > \varepsilon\|x_n\|_X + n\|Sx_n\|_Z$$

and $\|x_n\|_X = 1$. Passing to a subsequence we may assume that (x_n) is weakly convergent to some $x \in X$, by the reflexivity of X. Then $Sx_n \to Sx$ weakly; see the hints in Exercise 6.6. Moreover $\|Sx_n\|_Z < \frac{1}{n}\|Kx_n\|_X \to 0$, which implies $Sx = 0$; so $x = 0$ because S is injective. Since K is compact it follows that $Kx_n \to 0$ in norm; see Exercise 6.6(a). But $\|Kx_n\|_Y > \varepsilon$ for all $n \in \mathbb{N}$, a contradiction. □

8.2 Proposition. *With $j = \mathrm{tr}$, the classical Dirichlet form a is j-coercive.*

Proof. The bounded linear operator $S\colon H^1(\Omega) \to L_2(\Omega;\mathbb{K}^n) \oplus L_2(\partial\Omega)$ given by

$$Su := (\nabla u, \mathrm{tr}\, u)$$

is injective. Indeed, if $Su = 0$, then $\nabla u = 0$ and $u \in H_0^1(\Omega)$ by Theorem 7.12, and thus $u = 0$ by Poincaré's inequality (Theorem 5.13).

As the embedding $H^1(\Omega) \hookrightarrow L_2(\Omega)$ is compact, by Theorem 7.9 (Rellich–Kondrachov), the application of Lemma 8.1 yields a constant $c \geqslant 0$ such that for all $u \in H^1(\Omega)$ one has

$$\begin{aligned}\int_\Omega |u|^2\,\mathrm{d}x &\leqslant \frac{1}{2}\|u\|^2_{H^1(\Omega)} + c\|Su\|^2\\ &= \frac{1}{2}\int_\Omega |u|^2\,\mathrm{d}x + \frac{1}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x + c\int_\Omega |\nabla u|^2\,\mathrm{d}x + c\int_{\partial\Omega} |\mathrm{tr}\, u|^2\,\mathrm{d}\sigma.\end{aligned}$$

Adding $\frac{1}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x - \frac{1}{2}\int_\Omega |u|^2\,\mathrm{d}x$ to this inequality we obtain

$$\frac{1}{2}\|u\|^2_{H^1(\Omega)} \leqslant (c+1)a(u) + c\|\mathrm{tr}\, u\|^2_{L_2(\partial\Omega)} \qquad (u \in H^1(\Omega)), \tag{8.2}$$

and it follows that a is j-coercive. □

We now come back to the Dirichlet-to-Neumann operator.

8.3 Theorem. *Let j be the trace operator, and let a be the classical Dirichlet form* (8.1). *Then the operator D_0 in $L_2(\partial\Omega)$ associated with (a, j) is given by*

$$D_0 = \big\{(g,h) \in L_2(\partial\Omega) \times L_2(\partial\Omega);\ \exists u \in H^1(\Omega)\colon \Delta u = 0,\ u|_{\partial\Omega} = g,\ \partial_\nu u = h\big\}.$$

The operator D_0 is self-adjoint and accretive and has compact resolvent. We call D_0 the ***Dirichlet-to-Neumann operator*** *(with respect to Δ).*

The attentive reader may have noticed a discrepancy between the description of the Dirichlet-to-Neumann operater in the first paragraph of the present section, where it is stated that the Dirichlet data of a *harmonic* function u are mapped to the Neumann data, and the description of D_0 given above, where u is merely required to satisfy $\Delta u = 0$ in the distributional sense. This discrepancy is resolved by Weyl's lemma, for which we refer to Appendix C.

The action of the operator D_0 can be described as follows. Given $g \in L_2(\partial\Omega)$ one seeks a solution $u_g \in H^1(\Omega)$ of the Dirichlet problem $\Delta u = 0$, $\mathrm{tr}\, u = g$. As mentioned in the Notes of Chapter 7, a solution u_g exists for all $g \in \mathrm{ran}(\mathrm{tr})$. If this solution satisfies $\partial_\nu u_g \in L_2(\partial\Omega)$, then $g \in \mathrm{dom}(D_0)$ and $D_0 g = \partial_\nu u_g$.

Proof of Theorem 8.3. Let $(g, h) \in D_0$. Since $D_0 \sim (a, j)$, there exists $u \in H^1(\Omega)$ such that $u|_{\partial\Omega} = j(u) = g$ and

$$\int_\Omega \nabla u \cdot \overline{\nabla v} \, \mathrm{d}x = a(u, v) = (h \,|\, j(v))_{L_2(\partial\Omega)} = \int_{\partial\Omega} h\overline{v} \, \mathrm{d}\sigma \tag{8.3}$$

for all $v \in H^1(\Omega)$. Employing this equality with $v \in C_c^\infty(\Omega)$ we obtain $-\Delta u = 0$. Thus, adding $\int (\Delta u)\overline{v} \, \mathrm{d}x = 0$ in (8.3) we find that

$$\int_\Omega (\Delta u)\overline{v} \, \mathrm{d}x + \int_\Omega \nabla u \cdot \overline{\nabla v} \, \mathrm{d}x = \int_{\partial\Omega} h\overline{v} \, \mathrm{d}\sigma \qquad (v \in H^1(\Omega)).$$

Hence $\partial_\nu u = h$ by our definition of the weak normal derivative. This proves the inclusion '$\subseteq$' in the asserted equality for D_0.

In order to prove the reverse inclusion let $u \in H^1(\Omega)$ satisfy $\Delta u = 0$ and $h := \partial_\nu u \in L_2(\partial\Omega)$, and put $g := u|_{\partial\Omega}$. Then

$$a(u, v) = \int_\Omega \nabla u \cdot \overline{\nabla v} \, \mathrm{d}x + \int_\Omega (\Delta u)\overline{v} \, \mathrm{d}x = \int_{\partial\Omega} h\overline{v} \, \mathrm{d}\sigma = (h \,|\, j(v))_{L_2(\partial\Omega)} \qquad (v \in H^1(\Omega));$$

consequently $(g, h) = (j(u), h) \in D_0$.

The symmetry and accretivity of a imply that D_0 is self-adjoint and accretive. Finally, it was shown in Theorem 7.11(c) that $\mathrm{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$ is compact; hence D_0 has compact resolvent, by Proposition 6.18. □

Our next aim is to define Dirichlet-to-Neumann operators with respect to more general Dirichlet problems. The following interlude contains preparatory material for this treatment.

8.2 Interlude: the Fredholm alternative in Hilbert space

We need a detail from the spectral theory of compact operators which we formulate and prove only for operators in Hilbert spaces. It will be used in the proof of Lemma 8.10.

8.4 Theorem (Fredholm alternative). *Let H be a Hilbert space, $K \in \mathcal{L}(H)$ compact. Then the operator $I + K$ is injective if and only if it is surjective, and in this case $I + K$ is invertible in $\mathcal{L}(H)$.*

Proof. (i) In this step we treat the case when $\dim \operatorname{ran}(K) < \infty$. We define $H_1 := \ker(K)$, $H_2 := \ker(K)^\perp$ and denote by P_1, P_2 the orthogonal projections onto H_1, H_2, respectively. Observe that $\dim H_2 < \infty$. (Indeed, from $K = \sum_{j=1}^n (\cdot \,|\, x_j)y_j$ one obtains $\ker(K) \supseteq \{x_1, \ldots, x_n\}^\perp$, and therefore $\ker(K)^\perp \subseteq \operatorname{lin}\{x_1, \ldots, x_n\}$.) On the orthogonal sum $H_1 \oplus H_2$ the operator $I + K$ can be written as the operator matrix

$$I + K = \begin{pmatrix} I_1 & P_1K|_{H_2} \\ 0 & I_2 + P_2K|_{H_2} \end{pmatrix},$$

where I_1, I_2 are the identity operators in H_1, H_2, respectively. It follows from Exercise 8.3 that $I + K$ is injective/surjective/invertible in $\mathcal{L}(H)$ if and only if the corresponding property holds for $I_2 + P_2K|_{H_2}$. Thus the assertions follow from (finite-dimensional) linear algebra.

(ii) Now assume that $\dim \operatorname{ran}(K) = \infty$. Note that $K(B_H(0,1))$ is a relatively compact set in a metric space, and as such is separable; therefore $\overline{\operatorname{ran}(K)}$ is separable. Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\overline{\operatorname{ran}(K)}$, and for $n \in \mathbb{N}$ let P_n be the orthogonal projection from $\overline{\operatorname{ran}(K)}$ onto $\operatorname{lin}\{e_1, \dots, e_n\}$. Then $P_nx \to x$ $(n \to \infty)$ for all $x \in \overline{\operatorname{ran}(K)}$. Hence $P_n \to I$ uniformly on the compact set $\overline{K(B_H(0,1))}$, by Exercise 3.2(b), and this implies that $P_nK \to K$ in the operator norm.

Therefore K can be written as a sum $K = K_1 + K_2$, where $\|K_1\| < 1$ and $\operatorname{ran}(K_2)$ is finite-dimensional. Then $I + K_1$ is invertible in $\mathcal{L}(H)$ (Neumann series), and thus one can write $(I + K_1)^{-1}(I + K) = I + F$, with the finite rank operator $F = (I + K_1)^{-1}K_2$. Hence, $I + K$ is injective/surjective/invertible in $\mathcal{L}(H)$ if and only if the corresponding property holds for $I + F$. Thus the assertions follow from step (i). □

The label 'alternative' in Theorem 8.4 may be explained as follows: the *only* alternative to $I + K$ being invertible is that $I + K$ is neither injective nor surjective.

8.5 Remark. Let G, H be Hilbert spaces, and let $B, K \in \mathcal{L}(G, H)$, B invertible, K compact. Then the operator $B + K$ is injective if and only if it is surjective. This follows from $B + K = B(I + B^{-1}K)$ and Theorem 8.4, applied to the compact operator $B^{-1}K \in \mathcal{L}(G)$. △

8.3 Quasi-m-accretive and self-adjoint operators via essentially coercive forms

Let V, H be Hilbert spaces, and let $a\colon V \times V \to \mathbb{K}$ be a bounded form. Let $j \in \mathcal{L}(V, H)$ have dense range. Throughout this section we assume that

$$u \in \ker(j),\ a(u,v) = 0 \text{ for all } v \in \ker(j) \ \text{ implies } \ u = 0. \tag{8.4}$$

This condition is slightly weaker than (5.7), which was the standard assumption in Section 5.3. The weaker assumption (8.4) still implies that the relation

$$A = \big\{(x,y) \in H \times H\,;\ \exists u \in V\colon j(u) = x,\ a(u,v) = (y \,|\, j(v))\ (v \in V)\big\},$$

defined as in (5.6), is an operator; see Proposition 8.6.

As before we call A the operator associated with (a, j). Given $\omega \in \mathbb{K}$, we denote by a_ω the shifted form defined by $a_\omega(u,v) := a(u,v) + \omega(j(u) \,|\, j(v))$.

8.6 Proposition. *Assume that* (8.4) *is satisfied. Then the relation A given above is an operator. For $\omega \in \mathbb{K}$ the form a_ω satisfies* (8.4)*, and $A + \omega$ is the operator associated with (a_ω, j). If* (8.4) *is also satisfied for a^*, and B is the operator associated with (a^*, j), then $B \subseteq A^*$ and $A \subseteq B^*$. In particular, if a is symmetric and* $\operatorname{dom}(A)$ *is dense, then A is symmetric.*

The proof is delegated to Exercise 8.4. We will be working with the space

$$V_j(a) := \big\{u \in V\,;\ a(u,v) = 0\ (v \in \ker(j))\big\} = \bigcap_{v\in\ker(j)} \ker a(\cdot, v). \tag{8.5}$$

8.7 Remarks. (a) Condition (8.4) is equivalent to $V_j(a) \cap \ker(j) = \{0\}$.

(b) It is easy to see that $V_j(a_\omega) = V_j(a)$ for all $\omega \in \mathbb{K}$.

(c) $V_j(a)$ is a closed subspace of V. One might think of $V_j(a)$ as the 'orthogonal complement' of $\ker(j)$ with respect to a. If a is symmetric and j-coercive, then a_ω is an equivalent scalar product on V for large $\omega \in \mathbb{R}$, and $V_j(a) = V_j(a_\omega)$ is indeed the orthogonal complement of $\ker(j)$ with respect to a_ω. For the case of more general forms we refer to Proposition 8.16.

(d) If $(x, y) \in A$ and $u \in V$ are as in the definition of A, then $u \in V_j(a)$: for all $v \in \ker(j)$ one obtains $a(u, v) = (y \,|\, j(v)) = 0$.

(e) Assume that the Lax–Milgram operator $\mathcal{A}\colon V \to V^*$ given by $\mathcal{A}u := a(u, \cdot)$ is surjective. Then A is surjective. Indeed, for all $y \in H$ one has $(y \,|\, j(\cdot)) \in V^*$, hence $(y \,|\, j(\cdot)) = \mathcal{A}u$ for some $u \in V$; then $y = Aj(u) \in \operatorname{ran}(A)$. △

8.8 Example. We determine the space $V_j(a)$ for the case of the Dirichlet-to-Neumann operator in $L_2(\partial\Omega)$ associated with (a, j), where $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with C^1-boundary, a the classical Dirichlet form on $V = H^1(\Omega)$, and $j = \operatorname{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$. Then $\ker(j) = H^1_0(\Omega)$, by Theorem 7.12, and an element $u \in H^1(\Omega)$ satisfies

$$0 = a(u, v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x \qquad (v \in \ker(j) = H^1_0(\Omega))$$

if and only if $\Delta u = 0$, by Lemma 4.20. Hence $V_j(a) = \{u \in H^1(\Omega)\,;\ \Delta u = 0\}$.

By Remark 8.7(c) and Proposition 8.2 we see that $V_j(a)$ is the orthogonal complement of $H^1_0(\Omega)$ with respect to a_ω, for suitable $\omega \in \mathbb{R}$. Thus every $G \in H^1(\Omega)$ has a unique decomposition $G = u + w$ with $u \in V_j(a)$ and $w \in H^1_0(\Omega)$. Then u satisfies $\Delta u = 0$ and $\operatorname{tr} u = \operatorname{tr} G$, i.e., u is the solution of the Dirichlet problem mentioned in the Notes of Chapter 7. In Exercise 8.8 the reader is asked to show that the decomposition $H^1(\Omega) = V_j(a) \oplus H^1_0(\Omega)$ is not orthogonal in the standard scalar product of $H^1(\Omega)$. △

In the next lemma we draw an important conclusion from parts (d) and (e) of Remarks 8.7. Afterwards we will indicate a condition under which the assumptions of the lemma are satisfied.

8.9 Lemma. *In addition to* (8.4) *assume that the restriction of a to $V_j(a)$ is coercive and that $\mathcal{A}$ is surjective. Then A is strictly m-accretive. If $\mathbb{K} = \mathbb{C}$, then A is m-sectorial.*

Proof. For $x \in \operatorname{dom}(A)$ and $u \in V$ with $j(u) = x$ as in the definition of A one has $u \in V_j(a)$ by Remark 8.7(d). Thus by the coercivity assumption one obtains

$$\operatorname{Re}(Ax \,|\, x) = \operatorname{Re}(Ax \,|\, j(u)) = \operatorname{Re} a(u, u) \geqslant \alpha \|u\|_V^2 \geqslant \frac{\alpha}{c^2}\|x\|_H^2\,,$$

with $c > 0$ such that $c \geqslant \|j\|$ and some $\alpha > 0$ (not depending on x). This shows that A is strictly accretive. (The previous argument is essentially the same as in the proof of Theorem 5.6.) Now the surjectivity of A – see Remark 8.7(e) – implies that A is strictly m-accretive.

Now let $\mathbb{K} = \mathbb{C}$. As seen above one has

$$\{(Ax \,|\, x);\ x \in \operatorname{dom}(A)\} \subseteq \{a(u);\ u \in V_j(a)\}.$$

Since $\check{a} := a|_{V_j(a)\times V_j(a)}$ is coercive, Theorem 5.8 shows that $\check{a}$ is sectorial, so the above inclusion implies that A is sectorial. Then from the m-accretivity of A one concludes that A is m-sectorial. □

We call the form a **essentially coercive** if there exist a Hilbert space $\widetilde{H}$ and a compact operator $\tilde{\jmath}\colon V \to \widetilde{H}$ such that a is $\tilde{\jmath}$-coercive, i.e.

$$\operatorname{Re} a(u) + \widetilde{\omega}\|\tilde{\jmath}(u)\|_{\widetilde{H}}^2 \geqslant \widetilde{\alpha}\|u\|_V^2 \qquad (u \in V), \tag{8.6}$$

with some $\widetilde{\omega}, \widetilde{\alpha} > 0$. Observe that the operator $\mathcal{K} \in \mathcal{L}(V, V^*)$, defined by

$$\langle \mathcal{K}u, v\rangle := (\tilde{\jmath}(u) \,|\, \tilde{\jmath}(v))_{\widetilde{H}} \qquad (u, v \in V), \tag{8.7}$$

is compact. Indeed, the operator

$$k\colon \widetilde{H} \to V^*,\ x \mapsto (x \,|\, \tilde{\jmath}(\cdot))_{\widetilde{H}}$$

belongs to $\mathcal{L}(\widetilde{H}, V^*)$; hence $\mathcal{K} = k \circ \tilde{\jmath}$ is compact, by the ideal property of compact operators.

The property 'essentially coercive' has been introduced in [AEKS14] under the name 'compactly elliptic'. Our terminology is adapted from [ArCh20], where an equivalent property is called 'essentially positive-coercive'; see [ArCh20; Theorem 4.4]. (The background for 'positive' is that in the latter paper more general versions of coercivity are under consideration.) We refer to Exercise 8.5 for further equivalent properties.

Theorem 8.11 below states that for essentially coercive a the operator associated with (a, j) is quasi-m-accretive. To prove this result we will use Lemma 8.9 and the following two properties of essentially coercive forms.

8.10 Lemma. *Assume that* (8.4) *is satisfied and that a is essentially coercive.*

(a) *Then there exist $\omega \geqslant 0$, $\alpha > 0$ such that*

$$\operatorname{Re} a(u) + \omega\|j(u)\|_H^2 \geqslant \alpha\|u\|_V^2 \qquad (u \in V_j(a)). \tag{8.8}$$

(This means that the restriction of a_ω to $V_j(a)$ is coercive.)

(b) *With ω from* (a)*, the Lax–Milgram operator $\mathcal{A}_\omega\colon V \to V^*$ defined by $\mathcal{A}_\omega u := a_\omega(u, \cdot)$ is an isomorphism.*

Proof. (a) By Remark 8.7(a), j is injective on $V_j(a)$. Therefore Lemma 8.1 implies that

$$\widetilde{\omega}\|\tilde{\jmath}(u)\|_{\widetilde{H}}^2 \leqslant \frac{\widetilde{\alpha}}{2}\|u\|_V^2 + \omega\|j(u)\|_H^2 \qquad (u \in V_j(a)),$$

with $\widetilde{\omega}, \widetilde{\alpha} > 0$ and $\tilde{\jmath}$, $\widetilde{H}$ as in (8.6) and some $\omega \geqslant 0$. From (8.6) we then obtain

$$\operatorname{Re} a(u) + \omega\|j(u)\|_H^2 \geqslant \frac{\widetilde{\alpha}}{2}\|u\|_V^2 \qquad (u \in V_j(a)).$$

(b) If $\mathcal{A}_\omega u = 0$, then $a(u,v) = \langle \mathcal{A}_\omega u, v\rangle - \omega(j(u)\,|\,j(v)) = 0$ for all $v \in \ker(j)$, i.e. $u \in V_j(a)$; hence $\alpha\|u\|_V^2 \leqslant \operatorname{Re}\langle \mathcal{A}_\omega u, u\rangle = 0$ by (8.8). Thus $\mathcal{A}_\omega$ is injective.

We define a bounded form $b_\omega\colon V \times V \to \mathbb{K}$ by

$$b_\omega(u,v) := a_\omega(u,v) + \widetilde{\omega}(\tilde{\jmath}(u)\,|\,\tilde{\jmath}(v))_{\widetilde{H}}\,;$$

then b_ω is coercive by (8.6) since $\omega \geqslant 0$. The Lax–Milgram lemma, Theorem 5.4, implies that the operator $\mathcal{B}_\omega\colon V \to V^*$, $u \mapsto b_\omega(u,\cdot)$ is an isomorphism. Moreover $\mathcal{A}_\omega = \mathcal{B}_\omega - \widetilde{\omega}\mathcal{K}$, with the compact operator $\mathcal{K} \in \mathcal{L}(V,V^*)$ from (8.7). As $\mathcal{A}_\omega$ is injective, the 'Fredholm alternative' (in the guise of Remark 8.5) shows that $\mathcal{A}_\omega$ is an isomorphism. □

We now show that the operator associated with (a,j) is quasi-m-accretive.

8.11 Theorem. *Let $a\colon V \times V \to \mathbb{K}$ be a bounded form, and let $j \in \mathcal{L}(V,H)$ have dense range. Assume that* (8.4) *is satisfied and that a is essentially coercive. Let A be the operator associated with (a,j).*

(a) *Then A is quasi-m-accretive.*

(b) *If $\mathbb{K} = \mathbb{C}$, then A is quasi-m-sectorial.*

(c) *If a is symmetric, then A is self-adjoint.*

(d) *If j is compact, then A has compact resolvent.*

Proof. (a) By Lemma 8.10 there exists $\omega \geqslant 0$ such that $a_\omega|_{V_j(a)\times V_j(a)}$ is coercive and $\mathcal{A}_\omega$ is surjective. Recall from Proposition 8.6 that $A + \omega$ is the operator associated with (a_ω, j). Now Lemma 8.9 (combined with Remark 8.7(b)) shows that $A+\omega$ is strictly m-accretive, and hence A is quasi-m-accretive.

(b) In this case Lemma 8.9 implies that $A+\omega$ is m-sectorial, hence A is quasi-m-sectorial.

(c) If a is symmetric, then A is symmetric, by Proposition 8.6. Since A is quasi-m-accretive, Theorem 6.1 implies that A is self-adjoint.

(d) We show that $(A+\omega)^{-1} = j\mathcal{A}_\omega^{-1}k$, with $k\colon H \to V^*$, $y \mapsto (y\,|\,j(\cdot))$; then A has compact resolvent. By definition, $(y,x) \in (A+\omega)^{-1}$ if and only if there exists $u \in V$ such that $j(u) = x$ and $\mathcal{A}_\omega u = (y\,|\,j(\cdot)) = k(y)$, and this is equivalent to $x = j\mathcal{A}_\omega^{-1}k(y)$. (The previous argument is essentially the same as in the proof of Proposition 5.7; see also the proof of Proposition 6.18.) □

8.4 The Dirichlet-to-Neumann operator with respect to $\Delta + m$

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary, and let $m \in L_\infty(\Omega)$ be real-valued. We want to study the Dirichlet-to-Neumann operator D_m with respect to '$(\Delta+m)$-harmonic' functions. More precisely, we define D_m in $L_2(\partial\Omega)$ by requiring that for $g,h \in L_2(\partial\Omega)$ one has $g \in \operatorname{dom}(D_m)$ and $D_m g = h$ if there exists a solution $u \in H^1(\Omega)$ of $\Delta u + mu = 0$, $u|_{\partial\Omega} = g$ such that $\partial_\nu u = h$. We will show that D_m is a self-adjoint operator if

$$0 \notin \sigma(\Delta_{\mathrm{D}} + m), \tag{8.9}$$

which we will suppose throughout. Here $\Delta_{\mathrm{D}} + m$ is the Dirichlet Laplacian perturbed by the bounded operator of multiplication by the function m. We observe that $\Delta_{\mathrm{D}} + m$ is self-adjoint and has compact resolvent; see Exercise 8.6.

As in Section 8.1 we choose $H = L_2(\partial\Omega)$, $V = H^1(\Omega)$ and as $j\colon H^1(\Omega) \to L_2(\partial\Omega)$ the trace operator. According to the intended setup we now define the form $a\colon V \times V \to \mathbb{C}$ by

$$a(u,v) := \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x - \int_\Omega m u \bar{v}\,\mathrm{d}x \qquad (u, v \in H^1(\Omega)).$$

8.12 Remark. In general, the form a is not j-coercive. For example, let $\lambda > \lambda_1$, where λ_1 is the first Dirichlet eigenvalue (cf. Example 6.19), and put $m := \lambda$. Let $u \in H_0^1(\Omega)$ be an eigenfunction of $-\Delta_{\mathrm{D}}$ corresponding to λ_1. Then $\int_\Omega |\nabla u|^2\,\mathrm{d}x = \lambda_1 \int_\Omega |u|^2\,\mathrm{d}x$, so

$$a(u) + \omega\|j(u)\|^2_{L_2(\partial\Omega)} = \int_\Omega |\nabla u|^2\,\mathrm{d}x - \lambda\int_\Omega |u|^2\,\mathrm{d}x = (\lambda_1 - \lambda)\int_\Omega |u|^2\,\mathrm{d}x < 0$$

for all $\omega \in \mathbb{R}$. Thus the form is not j-coercive. (Note that (8.9) is satisfied if λ is not an eigenvalue of Δ_{D}.) △

Since the form a is not necessarily j-coercive, the theory developed in Chapter 5 is not applicable. However, we can apply the results of the previous section.

8.13 Theorem. *Suppose that* (8.9) *holds. Then*

$$D_m := \big\{(g,h) \in L_2(\partial\Omega) \times L_2(\partial\Omega);\, \exists u \in H^1(\Omega)\colon \Delta u + mu = 0,\ u|_{\partial\Omega} = g,\ \partial_\nu u = h\big\}$$

is a quasi-accretive self-adjoint operator with compact resolvent.

Proof. Let a be the form defined above, and let $j \in \mathcal{L}(H^1(\Omega), L_2(\partial\Omega))$ be the trace operator. We first show that condition (8.4) is satisfied. We recall from Theorem 7.12 that $\ker(j) = H_0^1(\Omega)$. Let $u \in \ker(j)$ be such that $a(u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x - \int_\Omega m u\bar{v}\,\mathrm{d}x = 0$ for all $v \in \ker(j) = H_0^1(\Omega)$. Then $u \in \operatorname{dom}(\Delta_{\mathrm{D}} + m)$ and $\Delta_{\mathrm{D}} u + mu = 0$, by the definition of Δ_{D}. This implies $u = 0$ since $0 \notin \sigma(\Delta_{\mathrm{D}} + m)$ by our assumption (8.9).

In order to show that a is essentially coercive, we choose $\widetilde{H} := L_2(\Omega)$ and as $\tilde{\jmath}$ the embedding $H^1(\Omega) \hookrightarrow L_2(\Omega)$. Then

$$a(u) + (\|m\|_\infty + 1)\|\tilde{\jmath}(u)\|_2^2 = \int_\Omega |\nabla u|^2\,\mathrm{d}x - \int_\Omega m|u|^2\,\mathrm{d}x + (\|m\|_\infty + 1)\|u\|_2^2 \geqslant \|u\|_{H^1}^2$$

for all $u \in H^1(\Omega)$, and from Theorem 7.9 we know that $\tilde{\jmath}$ is compact.

Let A be the operator associated with (a, j). By Theorem 8.11, parts (a) and (c), A is self-adjoint and quasi-accretive. As j is compact by Theorem 7.11(c), Theorem 8.11(d) implies that A has compact resolvent.

We show that $A = D_m$. If $(g,h) \in A$, then there exists $u \in H^1(\Omega)$ such that $u|_{\partial\Omega} = g$ and

$$\int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x - \int_\Omega m u\bar{v}\,\mathrm{d}x = \int_{\partial\Omega} h\bar{v}\,\mathrm{d}\sigma \qquad (v \in H^1(\Omega)). \tag{8.10}$$

Inserting test functions $v \in C_c^\infty(\Omega)$ we obtain $-\Delta u - mu = 0$. Plugging $mu = -\Delta u$ into (8.10) we deduce that $\partial_\nu u = h$. Thus $(g,h) \in D_m$. Conversely, if $(g,h) \in D_m$, then there exists $u \in H^1(\Omega)$ such that $\Delta u + mu = 0$, $u|_{\partial\Omega} = g$ and $\partial_\nu u = h$. Thus, by the definition of the weak normal derivative,

$$\int_{\partial\Omega} h\bar{v}\,\mathrm{d}\sigma = \int_\Omega \nabla u\cdot\overline{\nabla v}\,\mathrm{d}x + \int_\Omega (\Delta u)\bar{v}\,\mathrm{d}x = \int_\Omega \nabla u\cdot\overline{\nabla v}\,\mathrm{d}x - \int_\Omega mu\bar{v}\,\mathrm{d}x = a(u,v)$$

for all $v \in H^1(\Omega)$, and hence $(g,h) \in A$. □

8.5 Decomposition of the form domain

In this section we continue the analysis started in Section 8.3. Throughout the section let V, H be Hilbert spaces, let $j \in \mathcal{L}(V,H)$ have dense range, and let $a\colon V\times V \to \mathbb{K}$ be a bounded essentially coercive form satisfying (8.4).

Note that a coercive form is automatically essentially coercive (and satisfies (8.4)). Thus the following results are also valid – and interesting! – for coercive forms. It will become apparent that the properties treated in this section are invariant under shifting the form, and therefore they hold for bounded j-coercive forms as well.

In the applications in this chapter the space $V_j(a)$ is only a 'small' subspace of V because j is not injective; see for instance Example 8.8. We now show that the associated operator does not change if one restricts both a and j to $V_j(a)$.

8.14 Proposition. *Let $V_j(a)$ be defined by* (8.5). *Then $j|_{V_j(a)}$ is injective and has dense range, and the operator A associated with (a,j) is also associated with $(a|_{V_j(a)\times V_j(a)}, j|_{V_j(a)})$.*

Proof. The injectivity of $j|_{V_j(a)}$ is is clear from Remark 8.7(a). Let $\check{a} := a|_{V_j(a)\times V_j(a)}$ and $\check{\jmath} := j|_{V_j(a)}$. From Remark 8.7(d) we obtain $j(V_j(a)) \supseteq \operatorname{dom}(A)$, and as A is densely defined, we conclude that $\check{\jmath}$ has dense range. Lemma 8.10(a) shows that the form $\check{a}$ is $\check{\jmath}$-coercive; hence the operator $\check{A}$ associated with $(\check{a},\check{\jmath})$ is quasi-m-accretive, by the results of Chapter 5.

Let $(x,y) \in A$. Then there exists $u \in V_j(a)$ such that $a(u,v) = (y \mid j(v))$ for all $v \in V$, hence a fortiori for all $v \in V_j(a)$, and it follows that $(x,y) \in \check{A}$. Thus we have shown that $A \subseteq \check{A}$. As both operators A and $\check{A}$ are quasi-m-accretive, we conclude equality (by the well-known reasoning that a surjective mapping cannot have a proper injective extension). □

8.15 Remark. Proposition 8.14 says that the operator A associated with (a,j) can also be obtained from an embedded form (with the embedding $j|_{V_j(a)}$). In the next result we show that there is even more structure. △

Let X be a normed space, and let $X = X_1 \oplus X_2$ be the algebraic direct sum of two subspaces X_1, X_2, i.e. $X = X_1 + X_2$ and $X_1 \cap X_2 = \{0\}$. If the norm $\|\cdot\|$ on X is equivalent to the norm $\|\cdot\|_{\mathrm{s}}$ defined by

$$\|x_1 + x_2\|_{\mathrm{s}} := \|x_1\| + \|x_2\| \qquad (x_1 \in X_1,\ x_2 \in X_2), \tag{8.11}$$

then X is called the **topological direct sum** of X_1, X_2. The equivalence of the norms is equivalent to the property that both X_1 and X_2 are closed; see Exercise 8.7(b).

We emphasise that in the following result the direct sum is not necessarily orthogonal.

8.16 Proposition. *Assume that* (8.4) *is satisfied and that* a *is essentially coercive. Then* $V = V_j(a) \oplus \ker(j)$ *is a topological direct sum.*

Proof. Let $\widehat{V} := \ker(j)$. Clearly the restriction $\hat{a} := a|_{\widehat{V}\times\widehat{V}}$ of a is essentially coercive, and $\hat{\jmath} := j|_{\widehat{V}} \colon \widehat{V} \to \widehat{H} := \{0\}$ has dense range. Moreover $\widehat{V}_{\hat{\jmath}}(\hat{a}) = \{0\}$ since a and j satisfy (8.4). Thus $\hat{a}$ and $\hat{\jmath}$ satisfy (8.4) as well, and since (8.8) is trivially satisfied with $\omega = 0$, we conclude from Lemma 8.10(b) that $\ker(j) \ni u \mapsto \hat{a}(u,\cdot) = a(u,\cdot)|_{\ker(j)} \in \ker(j)^*$ is an isomorphism.

Now let $w \in V$. Then $a(w,\cdot)|_{\ker(j)} \in \ker(j)^*$, and hence there exists $u \in \ker(j)$ such that $a(u,v) = a(w,v)$ for all $v \in \ker(j)$, or equivalently $w - u \in V_j(a)$. Thus we have shown that $w \in V_j(a) + \ker(j)$. This implies the assertion since $V_j(a) \cap \ker(j) = \{0\}$ and both $V_j(a)$ and $\ker(j)$ are closed. □

Notes

A large part of the material in this chapter is adapted from [AEKS14]. The main results of Sections 8.3 and 8.5, Theorem 8.11 and Proposition 8.16, go beyond this paper and are due to H. Vogt. A different proof of Theorem 8.11 can be given by means of results in [Sau13].

In the investigation of the Dirichlet-to-Neumann operator D_m it is possible to avoid assumption (8.9). However, if $0 \in \sigma(\Delta_{\mathrm{D}} + m)$, then D_m is no longer a self-adjoint operator but rather a self-adjoint linear relation; we refer to Section 14.1 for this notion. In the complex case, the resolvent $(\mathrm{i}s - D_m)^{-1}$ is a bounded operator on $L_2(\partial\Omega)$, and the mapping $L_\infty(\Omega) \ni m \mapsto (\mathrm{i}s - D_m)^{-1} \in \mathcal{L}(L_2(\partial\Omega))$ is continuous; see [AEKS14; Theorem 7.3]. This gives valuable information on the stability of the inverse problem which interests engineers and medical doctors likewise.

In fact, one of the most famous inverse problems, the Calderón problem, can be formulated in terms of the Dirichlet-to-Neumann operator with respect to $\Delta - m$. It was Alberto Calderón (1920-1998) who encountered this problem when he worked as an electrical engineer for Yacimientos Petrolíferos Fiscales, the state oil company of Argentina. Calderón worked on *Electrical Impedance Tomography*, a method that can be used to detect oil, but also to find a tumor in the lungs. It consists in determining the electrical conductivity of a medium by making voltage and current measurements at the boundary of a medium. The mathematical problem Calderón formulates is the following.

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with Lipschitz boundary, and let $\gamma \colon \Omega \to (0,\infty)$ be a bounded measurable function with strictly positive lower bound. One should think of Ω as a body with 'electrical conductivity' γ. If one applies a 'voltage' function $g \colon \partial\Omega \to \mathbb{R}$ at the boundary, then the induced 'electric potential' $u \colon \Omega \to \mathbb{R}$ satisfies the Dirichlet problem

$$\begin{aligned} \operatorname{div}(\gamma\nabla u) &= 0 \qquad \text{on } \Omega, \\ u &= g \qquad \text{on } \partial\Omega, \end{aligned}$$

and one measures the 'current' $\gamma\partial_\nu u$ through the surface. The voltage-to-current map $g \mapsto \gamma\partial_\nu u$ is the Dirichlet-to-Neumann operator Λ_γ,

$$\Lambda_\gamma g = \gamma\partial_\nu u.$$

In other words, Λ_γ maps the 'Dirichlet data' $u|_{\partial\Omega}$ of the 'div$(\gamma\nabla)$-harmonic' function u to the 'Neumann data' $\gamma\partial_\nu u$.

The inverse problem Calderón formulates is "to decide whether γ is uniquely determined by Λ_γ and to calculate γ in terms of Λ_γ, if γ is indeed determined by Λ_γ"; see [Cal06] for a reprint of his paper from 1980. (Calderón does not formulate the problem in terms of the Dirichlet-to-Neumann operator, but rather in terms of a quadratic form; we refer to [Uhl09; p. 2] for the equivalence of Calderón's and the above formulation in terms of Λ_γ.)

We add a few comments on the notation used above. For the definition of the distributional divergence of a vector field we refer to Exercise 4.5(b) and to the paragraph preceding Remark 11.3. The weak normal derivative $\gamma\partial_\nu u$, or expressed differently, the weak normal trace of the vector field $\gamma\nabla u$ at $\partial\Omega$, is defined similarly as in Section 7.3.

In the remaining discussion we will only deal with the first of the two issues Calderón states, the uniqueness question. Assuming suitable regularity of γ, one can transform this problem into a uniqueness problem for the Dirichlet-to-Neumann operator D_m treated in Section 8.4, where m is related to the conductivity γ in a one-to-one way, as described in [Uhl09; Section 5]. Let $m_1, m_2 \in L_\infty(\Omega)$ and consider the associated Dirichlet-to-Neumann operators

$$D_{m_j} := \big\{(g,h) \in L_2(\partial\Omega) \times L_2(\partial\Omega);\ \exists u \in H^1(\Omega)\colon \Delta u + m_j u = 0,\ u|_{\partial\Omega} = g,\ \partial_\nu u = h\big\},$$

for $j = 1,2$; see Theorem 8.13. Then, analogously to the above Calderón uniqueness problem, one can ask: if $D_{m_1} = D_{m_2}$, does it follow that $m_1 = m_2$?

In this latter form, the problem has been solved in [KrUh14; Theorem 1.1], where the uniqueness is proved in a very general context. The original uniqueness problem of Calderón has been solved by Sylvester and Uhlmann [SyUh87; Theorem 0.1], for Ω with smooth boundary and conductivities γ in $C^\infty(\overline{\Omega})$, and by Haberman and Tataru [HaTa13; Theorem 1.1], for Ω with Lipschitz boundary and conductivities in $C^1(\overline{\Omega})$. All these results are for the case of dimension $n \geqslant 3$. For more information (also regarding dimension $n = 2$) we refer to [Uhl09] from which this account is adapted and where also further references can be found. We mention that in real-world applications, partial measurements at the boundary are of course more realistic, and much research goes on in this direction; see [KeSa14].

Exercises

8.1 Let $a, b \in \mathbb{R}$, $a < b$.

(a) Compute the Dirichlet-to-Neumann operator D_0 for $\Omega = (a,b)$, and compute the C_0-semigroup generated by $-D_0$.

(b) For $a = -1$, $b = 1$ interpret the result in the light of Exercise 8.2.

8.2 Let $U_n := B_{\mathbb{R}^n}(0,1)$ be the open unit ball in $\mathbb{R}^n$, $S_{n-1} := \partial U_n$ the unit sphere. The following facts can be used for the solution of this exercise: for each $\varphi \in C(S_{n-1})$ there exists a unique solution $u \in C(\overline{U_n})$ of the Dirichlet problem

$$u \text{ harmonic on } U_n, \quad u = \varphi \text{ on } S_{n-1}\,.$$

(We mention that the solution can be written down explicitly with the aid of the Poisson kernel, but this will not be needed for solving the exercise.) The solution satisfies $u|_{U_n} \in C^\infty(U_n)$ and $\|u\|_\infty \leqslant \|\varphi\|_\infty$. Writing $G\varphi := u$ one obtains $G \in \mathcal{L}(C(S_{n-1}), C(\overline{U_n}))$.
Define $T(t) \in \mathcal{L}(C(S_{n-1}))$ by

$$T(t)\varphi(z) := u(\mathrm{e}^{-t}z) \qquad (z \in S_{n-1},\ t \geqslant 0)$$

(with φ and $u = G\varphi$ as above).

(a) Show that T is a contractive C_0-semigroup on $C(S_{n-1})$.

(b) Let A be the generator of T. Show that $D := \bigcup_{t>0} \operatorname{ran}(T(t))$ is a core for A, and that $A\varphi = -\partial_\nu(G\varphi)$ for all $\varphi \in D$.

(c) Put $A_{\min} := A|_D$. Show that $-A_{\min}$ is a restriction of the Dirichlet-to-Neumann operator D_0 in $L_2(S_{n-1})$, and that $D_0 = \overline{-A_{\min}}$ (where $A_{\min}$ is to be regarded as an operator in $L_2(S_{n-1})$). Conclude that T extends to a contractive C_0-semigroup T_2 on $L_2(S_{n-1})$, and that $-D_0$ is the generator of T_2.

(We refer to [Lax02; Section 36.2] for this exercise.)

8.3 Let H_1, H_2 be Hilbert spaces, $H := H_1 \oplus H_2$.

(a) Let $A_1 \in \mathcal{L}(H_1)$, $A_2 \in \mathcal{L}(H_2)$ be invertible in $\mathcal{L}(H_1)$, $\mathcal{L}(H_2)$, respectively, and let $B \in \mathcal{L}(H_2, H_1)$. Show that the operator matrix $\begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$ is invertible in $\mathcal{L}(H)$, and compute its inverse.

(b) Let $A \in \mathcal{L}(H_2, H_1)$, $B \in \mathcal{L}(H_2)$. Show that $\begin{pmatrix} I_1 & A \\ 0 & B \end{pmatrix}\begin{pmatrix} I_1 & -A \\ 0 & I_2 \end{pmatrix} = \begin{pmatrix} I_1 & 0 \\ 0 & B \end{pmatrix}$, where I_1, I_2 are the identity operators in H_1, H_2, respectively. Conclude that $\begin{pmatrix} I_1 & A \\ 0 & B \end{pmatrix}$ is injective/surjective/invertible in $\mathcal{L}(H)$ if and only if the corresponding property holds for B.

8.4 Prove Proposition 8.6. (Hint: Inspect the proof of Proposition 5.5.)

8.5 Let a be a bounded form on a Hilbert space V. Show that the following properties are equivalent:

(i) a is essentially coercive;

(ii) there exists $\alpha > 0$ such that for all sequences (u_n) in V one has

$$\|u_n\| = 1\ (n \in \mathbb{N}),\ u_n \to 0 \text{ weakly} \implies \liminf \operatorname{Re} a(u_n) \geqslant \alpha;$$

(iii) there exist $\alpha > 0$ and a finite-dimensional subspace $V_1 \subseteq V$ such that

$$\operatorname{Re} a(u) \geqslant \alpha \|u\|_V^2 \qquad (u \in V_1^\perp).$$

Hint concerning '(ii) $\Rightarrow$ (iii)': Assume that (iii) does not hold and construct an orthonormal sequence (u_n) in V such that $\limsup \operatorname{Re} a(u_n) \leqslant 0$.

Hint concerning '(iii) $\Rightarrow$ (i)': For $u \in V_1^\perp$, $v \in V_1$ write $a(u+v) = a(u) + a(u,v) + a(v,u) + a(v)$ and use the Peter–Paul inequality to show that there exists $\omega > 0$ (not depending on u, v) such that $\operatorname{Re} a(u+v) \geqslant \frac{\alpha}{2}\|u+v\|_V^2 - \omega\|v\|_V^2$. Then take the orthogonal projection onto V_1 as the compact operator $\tilde{\jmath}$ in the definition of 'essentially coercive'.

8.6 (a) Let G, H be Hilbert spaces. Let A be a densely defined operator from G to H, $B \in \mathcal{L}(G, H)$. Show that $(A+B)^* = A^* + B^*$.

(b) Let $\Omega \subseteq \mathbb{R}^n$ be open, $m \in L_\infty(\Omega)$ real-valued. Show that $-(\Delta_{\mathrm{D}} + m)$ is self-adjoint and quasi-accretive. If Ω is bounded, then $\Delta_{\mathrm{D}} + m$ has compact resolvent.

8.7 Let X be a normed space, and suppose that $X = X_1 \oplus X_2$ is the algebraic direct sum of two subspaces X_1, X_2.

(a) Show that X is the topological direct sum of X_1, X_2 if and only if the projection from X onto X_1 along X_2 is a bounded operator.

(b) Show that if two of the properties

(i) $X = X_1 \oplus X_2$ is a topological direct sum,

(ii) X_1 and X_2 are closed,

(iii) X is a Banach space

hold, then also the third property holds. (Hint concerning '(ii), (iii) $\Rightarrow$ (i)': closed graph theorem.)

8.8 Let $\varnothing \neq \Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Show that the decomposition $H^1(\Omega) = V_j(a) \oplus H_0^1(\Omega)$ from Example 8.8 is not an orthogonal sum. (Hint: Choose $0 \neq u_0 \in C_c^2(\Omega)_+$, and decompose the function $u := 1 + u_0$.)

Chapter 9

Invariance of closed convex sets

In this chapter we investigate criteria for a closed convex set to be invariant under a semigroup. To begin with, we present criteria involving properties of the generator. Applying these criteria to the Dirichlet and Neumann Laplacians one realises that further properies of H^1-functions are needed; these will be provided in an interlude on lattice properties of H^1. In the last section we present criteria involving properties of forms that have a wide range of applications. Examples include the Robin Laplacian and the Dirichlet-to-Neumann operator, for which we refer to Exercises 9.4 and 9.5. In Chapter 11 we will apply the criteria to semigroups generated by elliptic operators.

9.1 Invariance for semigroups

Let T be a C_0-semigroup on a Banach space X over $\mathbb{K}$, with generator A. Our aim is to characterise when a closed convex subset C of X is invariant under the semigroup T, i.e. $T(t)(C) \subseteq C$ for all $t \geqslant 0$. This means that, for an initial value in C, the solution of the corresponding Cauchy problem remains in C for all $t \geqslant 0$. First we show that invariance under T is equivalent to invariance under the resolvent.

9.1 Proposition. *Let T be a C_0-semigroup on a Banach space X, with generator A. Let $C \subseteq X$ be a closed convex set. Then the following properties are equivalent.*

(i) *C is invariant under T.*

(ii) *There exists $\omega \in \mathbb{R}$ such that $(\omega, \infty) \subseteq \rho(A)$ and $\lambda R(\lambda, A)(C) \subseteq C$ for all $\lambda > \omega$.*

Noting that $\lambda R(\lambda, A) = (I - \frac{1}{\lambda}A)^{-1}$ we see that condition (ii) can be expressed equivalently by requiring that there exists $r_0 > 0$ such that $\{1/r;\, 0 < r < r_0\} \subseteq \rho(A)$ and $(I - rA)^{-1}(C) \subseteq C$ for all $0 < r < r_0$. It is this version of condition (ii) that will mostly be used below.

We insert a fact concerning integration that is needed in the proof of the implication '(i) ⇒ (ii)' in Proposition 9.1. It should be understood as a statement on generalised convex combinations.

9.2 Lemma. *Let C be a closed convex subset of a Banach space X, $a, b \in \mathbb{R}$, $a < b$. Let $u\colon [a, b] \to C$ be continuous and $\varphi \in C[a, b]$, $\varphi \geqslant 0$, $\int_a^b \varphi(t)\,\mathrm{d}t = 1$.*

Then $\int_a^b \varphi(t)u(t)\,\mathrm{d}t \in C$.

Proof. For simplicity of notation (and without loss of generality) we assume that $[a,b] = [0,1]$. For $n \in \mathbb{N}$ we put

$$u_n := u(0)\mathbf{1}_{\{0\}} + \sum_{k=1}^{n} u(k/n)\mathbf{1}_{((k-1)/n,k/n]}\,.$$

Then $\|\varphi u_n - \varphi u\|_\infty \to 0$ as $n \to \infty$; hence $\int_0^1 \varphi u_n \,\mathrm{d}t \to \int_0^1 \varphi u \,\mathrm{d}t$. Moreover $\int_0^1 \varphi u_n \,\mathrm{d}t = \sum_{k=1}^{n} \int_{(k-1)/n}^{k/n} \varphi(t)\,\mathrm{d}t\, u(k/n) \in C$, as a convex combination of elements of C. Since C is closed we obtain the assertion. □

Proof of Proposition 9.1. (i) ⇒ (ii). Let $\omega \in \mathbb{R}$, $M \geqslant 0$ be such that $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ for all $t \geqslant 0$, and let $\lambda > \omega$. Then $\lambda R(\lambda, A) = \int_0^\infty \lambda \mathrm{e}^{-\lambda t} T(t)\,\mathrm{d}t$ (strong improper integral). Let $x \in C$. For $\tau > 0$ we obtain $(1-\mathrm{e}^{-\lambda\tau})^{-1}\int_0^\tau \lambda\mathrm{e}^{-\lambda t}T(t)x\,\mathrm{d}t \in C$, by Lemma 9.2. Letting $\tau \to \infty$ we conclude that $\lambda R(\lambda, A)x \in C$.

(ii) ⇒ (i). This follows from the exponential formula (Theorem 2.12):

$$T(t)x = \lim_{n\to\infty}\Bigl(I - \frac{t}{n}A\Bigr)^{-n} x \in C \qquad (t \geqslant 0,\ x \in C). \qquad \square$$

In order to motivate why one is interested in the invariance of closed convex sets, we indicate several examples.

9.3 Remarks. Let (Ω, μ) be a measure space, $H := L_2(\mu; \mathbb{K})$.

(a) Let $C \subseteq L_2(\mu)$ be the **positive cone**, $C := L_2(\mu)_+ := \{u \in L_2(\mu)\,;\ u \geqslant 0\}$. Clearly C is a closed convex subset of $L_2(\mu)$. An operator $S \in \mathcal{L}(H)$ leaves C invariant if and only if S is **positive**, i.e. $Su \geqslant 0$ for all $u \geqslant 0$.

(b) Let $\mathbb{K} = \mathbb{C}$, and let $C := L_2(\mu; \mathbb{R})$ be the closed convex subset of real-valued functions. An operator $S \in \mathcal{L}(H)$ leaves C invariant if and only if S is **real**, i.e. Su is real-valued for all real-valued u.

(c) Let $C := \{u \in L_2(\mu)\,;\ \|u\|_\infty \leqslant 1\}$. Then C is convex and closed, and $S \in \mathcal{L}(H)$ leaves C invariant if and only if S is $\boldsymbol{L_\infty}$**-contractive**, i.e. $\|Su\|_\infty \leqslant \|u\|_\infty$ for all $u \in L_2 \cap L_\infty(\mu)$.

(d) Let $C := \{u \in L_2(\mu)\,;\ u \leqslant 1\}$; then C is convex and closed. We show that $S \in \mathcal{L}(H)$ leaves C invariant if and only if S is **sub-Markovian**, i.e. S is positive and L_∞-contractive.

Indeed, assume that $S(C) \subseteq C$. If $u \in L_2(\mu)_+$, then $-\alpha u \leqslant 1$ and therefore $-\alpha Su \leqslant 1$, for all $\alpha \geqslant 0$, and this implies that $Su \geqslant 0$. Hence S is a positive operator; in particular, S is real. Now let $u \in L_2 \cap L_\infty(\mu)$, $\|u\|_\infty \leqslant 1$. Then for all $\gamma \in \mathbb{K}$ with $|\gamma| = 1$ one obtains

$$\mathrm{Re}(\gamma Su) = \mathrm{Re}\bigl(S(\gamma u)\bigr) = S\bigl(\mathrm{Re}(\gamma u)\bigr) \leqslant 1. \tag{9.1}$$

(Note that in the complex case, for $v \in L_2(\mu)$, one has $\mathrm{Re}\, Sv = \mathrm{Re}\, S(\mathrm{Re}\, v + \mathrm{i}\,\mathrm{Im}\, v) = \mathrm{Re}\bigl(S(\mathrm{Re}\, v) + \mathrm{i}S(\mathrm{Im}\, v)\bigr) = S(\mathrm{Re}\, v)$, because $S(\mathrm{Re}\, v)$ and $S(\mathrm{Im}\, v)$ are real.) Using (9.1) for γ from a countable dense subset of the unit circle, one concludes that $\|Su\|_\infty \leqslant 1$.

Conversely, if S is positive and L_∞-contractive, then $u \leqslant 1$ implies $Su = Su^+ - Su^- \leqslant Su^+ \leqslant 1$. Here, $u^+ := u \vee 0$ is the **positive part** of u, and the **negative part** of u is defined by $u^- := (-u)^+ = u^+ - u$. (Notice that the negative part is positive.) △

Concerning notation, the symbol '$\vee$' used above denotes 'supremum', i.e., for functions $u, v\colon \Omega \to \mathbb{R}$ the supremum is given by $(u \vee v)(x) = \max\{u(x), v(x)\}$ $(x \in \Omega)$. Similarly we will use the symbol '$\wedge$' to denote 'infimum', $(u \wedge v)(x) = \min\{u(x), v(x)\}$ $(x \in \Omega)$.

9.4 Remark. Concerning the notion of 'positive' operators we mention that this property is sometimes called 'positivity preserving'. This is because in the literature 'positive' is often used as a synonym for 'accretive', in the context of symmetric operators in Hilbert spaces. △

We are looking for another characterisation of T-invariance of a closed convex set, involving more directly the generator A instead of its resolvent. This is possible in Hilbert spaces. Let H be a Hilbert space over $\mathbb{K}$, and let $\varnothing \neq C \subseteq H$ be convex and closed. We denote by $P_C\colon H \to C$ the **minimising projection** from H onto C, which maps $x \in H$ to its best approximation $P_C x$ in C, i.e., $P_C x$ is the unique element of C satisfying

$$\|x - P_C x\| = \inf\{\|x - y\|\,;\, y \in C\}.$$

We will often use the fact that $P_C x$ is the unique element of C satisfying

$$\operatorname{Re}(y - P_C x \mid x - P_C x) \leqslant 0 \qquad (y \in C); \tag{9.2}$$

see Exercise 9.2(a) (or [Bre11; Section 5.1]). Geometrically, this means that the vectors in the scalar product in (9.2) form an obtuse angle; if C has a tangent hyperplane at $P_C x$, then $x - P_C x$ is orthogonal to this hyperplane. The mapping P_C is a contraction; see Exercise 9.2(b). In particular, P_C is continuous. Clearly the minimising projection P_C satisfies $P_C \circ P_C = P_C$; so it deserves the name 'projection'. We could not find a commonly accepted name for this mapping in the literature. One should keep in mind that, in general, P_C is not a linear operator.

9.5 Remark. We illustrate the minimising projection for the closed convex subsets of $L_2(\mu)$ treated in Remarks 9.3. All the statements are immediate consequences of Exercise 9.3.

For the positive cone $C = L_2(\mu)_+$ the minimising projection is given by $P_C u = (\operatorname{Re} u)^+$.
For $\mathbb{K} = \mathbb{C}$ and $C = L_2(\mu;\mathbb{R})$ one has $P_C u = \operatorname{Re} u$.
For the set $C = \{u \in L_2(\mu)\,;\, \|u\|_\infty \leqslant 1\}$ one finds that $P_C u = (\operatorname{sgn} u)(|u| \wedge 1)$.
For $C = \{u \in L_2(\mu)\,;\, u \leqslant 1\}$ one obtains $P_C u = (\operatorname{Re} u) \wedge 1$.

In the description given above we have used the **signum function** $\operatorname{sgn}\colon \mathbb{K} \to \mathbb{K}$, which is defined as $\operatorname{sgn}\alpha := \frac{\alpha}{|\alpha|}$ if $0 \neq \alpha \in \mathbb{K}$, and $\operatorname{sgn} 0 := 0$. △

The following result has a geometric appeal. Its assumption (9.3) expresses that the 'driving term' $Au(t)$ in the equation $u'(t) = Au(t)$ always points 'sufficiently' from $u(t)$ towards C. (For $\omega \leqslant 0$ this is quite intuitive. If $\omega > 0$, one can interpret that it is more and more true the closer $u(t)$ is to C.)

9.6 Proposition. *Let T be a C_0-semigroup on a Hilbert space H, with generator A. Let $\varnothing \neq C \subseteq H$ be a closed convex set, and denote by $P := P_C$ the minimising projection. Assume that there exists $\omega \in \mathbb{R}$ such that*

$$\operatorname{Re}(Ax \mid x - Px) \leqslant \omega\|x - Px\|^2 \qquad (x \in \operatorname{dom}(A)). \tag{9.3}$$

Then C is invariant under T.

Proof. In view of Proposition 9.1 it suffices to show that $(I-rA)^{-1}(C)\subseteq C$ for small $r>0$. (Observe that $(I-rA)^{-1}\in\mathcal{L}(H)$ for small $r>0$.) Without loss of generality we assume that $\omega>0$. Let $0<r<1/\omega$, and let $x\in\operatorname{dom}(A)$ be such that $(I-rA)x\in C$. We have to show that $x\in C$. Applying (9.2) with $y=(I-rA)x\in C$ we obtain

$$\operatorname{Re}\big((I-rA)x-Px\,\big|\,x-Px\big)\leqslant 0.$$

Thus

$$\begin{aligned}\|x-Px\|^2&=\operatorname{Re}\big(rAx+(I-rA)x-Px\,\big|\,x-Px\big)\\&\leqslant r\operatorname{Re}(Ax\,|\,x-Px)\leqslant r\omega\|x-Px\|^2.\end{aligned}$$

Using $r\omega<1$ we conclude that $\|x-Px\|=0$, $x=Px\in C$. □

The converse of Proposition 9.6 holds for quasi-contractive semigroups. (The example $C=\{0\}$ shows that the assumption of quasi-contractivity cannot be omitted.)

9.7 Proposition. *Let H,T,A have the same properties as before, and assume that T is quasi-contractive, i.e., there exists $\omega\in\mathbb{R}$ such that $\|T(t)\|\leqslant e^{\omega t}$ for all $t\geqslant 0$. Let $\varnothing\neq C\subseteq H$ be a closed convex set, and assume that C is invariant under T.*

Then (9.3) *holds with the minimising projection $P:=P_C$.*

Proof. Let $x\in\operatorname{dom}(A)$. Then (9.2) implies $\operatorname{Re}\big(T(t)Px-Px\,\big|\,x-Px\big)\leqslant 0$, and this inequality can be rewritten as $0\leqslant\operatorname{Re}\big(-T(t)Px+Px\,\big|\,x-Px\big)$. One then obtains

$$\begin{aligned}\operatorname{Re}\big(T(t)x-x\,\big|\,x-Px\big)&\leqslant\operatorname{Re}\big(T(t)(x-Px)-(x-Px)\,\big|\,x-Px\big)\\&\leqslant(e^{\omega t}-1)\|x-Px\|^2.\end{aligned}$$

Dividing by t and taking the limit $t\to 0+$ one concludes that

$$\operatorname{Re}(Ax\,|\,x-Px)\leqslant\omega\|x-Px\|^2.$$ □

For the case of contractive C_0-semigroups we summarise the results of Propositions 9.6 and 9.7 as an equivalence.

9.8 Corollary. *Let T be a contractive C_0-semigroup on a Hilbert space H, with generator A, and let $C\neq\varnothing$ be a closed convex subset of H. Then C is invariant under T if and only if*

$$\operatorname{Re}(Ax\,|\,x-Px)\leqslant 0\qquad(x\in\operatorname{dom}(A)),\tag{9.4}$$

where $P:=P_C$ is the minimising projection.

We specify Corollary 9.8 for the case when $H=L_2(\mu;\mathbb{R})$ and $C=L_2(\mu)_+$ is the positive cone; see Remark 9.3(a). Then $Pu=u^+$ for all $u\in H$ (see Remark 9.5), and condition (9.4) becomes

$$(Au\,|\,-u^-)=(Au\,|\,u-Pu)\leqslant 0\qquad(u\in\operatorname{dom}(A)).\tag{9.5}$$

Replacing u by $-u$ one transforms (9.5) into

$$(Au\,|\,u^+)=-\big(A(-u)\,\big|\,(-u)^-\big)\leqslant 0\qquad(u\in\operatorname{dom}(A)).\tag{9.6}$$

An operator A in $L_2(\mu;\mathbb{R})$ satisfying (9.6) is called **dispersive**, and A is **m-dispersive** if additionally $\operatorname{ran}(I-A) = L_2(\mu;\mathbb{R})$. A semigroup T on $L_2(\mu)$ is called **positive** if $T(t)$ is positive for all $t \geqslant 0$.

We now obtain the following 'positive version' of the Lumer–Phillips theorem.

9.9 Corollary. *Let A be an operator in $L_2(\mu;\mathbb{R})$. Then A is the generator of a positive contractive C_0-semigroup if and only if A is m-dispersive.*

Proof. If A satisfies (9.6) – hence also (9.5) –, then $(-Au \mid u) = -(Au \mid u^+) - (Au \mid -u^-) \geqslant 0$ for all $u \in \operatorname{dom}(A)$, i.e. $-A$ is accretive. Now the assertion is an immediate consequence of Corollary 9.8 and Theorem 3.16 (Lumer–Phillips). □

9.2 Application to Laplacians

We recall that the Dirichlet Laplacian Δ_{D} in $L_2(\Omega)$ is associated with the classical Dirichlet form on $V = H_0^1(\Omega)$. We also recall that the Neumann Laplacian Δ_{N} in $L_2(\Omega)$ is associated with the classical Dirichlet form on $V = H^1(\Omega)$. (See Example 5.14 and Theorem 7.13.)

9.10 Example. Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Then the C_0-semigroup on $L_2(\Omega)$ generated by Δ_{D} is **sub-Markovian**, i.e. $\mathrm{e}^{t\Delta_{\mathrm{D}}}$ is sub-Markovian for all $t \geqslant 0$. This will be shown below. △

In the complex case, the semigroup being sub-Markovian means in particular that it leaves $L_2(\Omega;\mathbb{R})$ invariant. For the proof of this property one could employ Corollary 9.8. However, it is also instructive to use another reasoning, on the basis of the following proposition.

9.11 Proposition. *Let $X_1 \hookrightarrow X_2$ be Banach spaces. For $j = 1, 2$ let T_j be a C_0-semigroup on X_j with generator A_j, and assume that $A_1 \subseteq A_2$. Then $T_2(t)|_{X_1} = T_1(t)$ for all $t \geqslant 0$; in particular X_1 is invariant under T_2.*

Proof. Let $x \in \operatorname{dom}(A_1)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} T_1(t)x = A_1 T_1(t)x = A_2 T_1(t)x \qquad (t \geqslant 0)$$

shows that $T_1(t)x = T_2(t)x$ for all $t \geqslant 0$, by Theorem 1.13(a). From the denseness of $\operatorname{dom}(A_1)$ in X_1 and extension by continuity one obtains $T_2(t)|_{X_1} = T_1(t)$ for all $t \geqslant 0$. □

Proof of the assertion in Example 9.10. (i) We recall that $\Delta_{\mathrm{D},\mathbb{K}} := \Delta_{\mathrm{D}}$ is a generator in $L_2(\Omega;\mathbb{K})$. We regard $L_2(\Omega;\mathbb{C})$ as a Banach space over $\mathbb{R}$ and use Proposition 9.11 with $X_1 := L_2(\Omega;\mathbb{R})$, $X_2 := L_2(\Omega;\mathbb{C})$. Then $\Delta_{\mathrm{D},\mathbb{R}} \subseteq \Delta_{\mathrm{D},\mathbb{C}}$ implies that $\mathrm{e}^{t\Delta_{\mathrm{D},\mathbb{C}}}|_{L_2(\Omega;\mathbb{R})} = \mathrm{e}^{t\Delta_{\mathrm{D},\mathbb{R}}}$ for all $t \geqslant 0$. (This property was also the subject of Exercise 4.9.)

(ii) We show that the (closed convex) set $C := \{u \in L_2(\Omega;\mathbb{R})\,;\ u \leqslant 1\}$ is invariant under $(\mathrm{e}^{t\Delta_{\mathrm{D}}})_{t\geqslant 0}$; then it follows from Remark 9.3(d) that $\mathrm{e}^{t\Delta_{\mathrm{D}}}$ is sub-Markovian for all $t \geqslant 0$.

By step (i) we may assume without loss of generality that $\mathbb{K} = \mathbb{R}$. Then the minimising projection onto C is given by $Pu = u \wedge 1$; see Remark 9.5. We want to show that

$$(\Delta_{\mathrm{D}} u \,|\, u - u \wedge 1) \leqslant 0$$

for all $u \in \mathrm{dom}(\Delta_{\mathrm{D}})$.

We need to know that $u \wedge 1 \in H_0^1(\Omega)$ and $\nabla(u \wedge 1) = \mathbf{1}_{[u<1]} \nabla u$. These properties will be proved in the next section (and the present example should serve as a motivation for this treatment); see Theorem 9.16. Accepting these properties we obtain

$$(\Delta_{\mathrm{D}} u \,|\, u - u \wedge 1) = -(\nabla u \,|\, \mathbf{1}_{[u \geqslant 1]} \nabla u) = -\int_{[u \geqslant 1]} |\nabla u|^2 \,\mathrm{d}x \leqslant 0.$$

Now the application of Corollary 9.8 yields the invariance of C. □

We note that the same arguments show that the Neumann Laplacian generates a sub-Markovian C_0-semigroup on $L_2(\Omega)$. We refer to Exercise 9.4 for the discussion of invariance properties for the Robin Laplacian.

9.3 Interlude: lattice properties of $H^1(\Omega)$

Throughout this section we will use $\mathbb{K} = \mathbb{R}$, i.e. all the functions will be real-valued. We start with a warm-up.

9.12 Lemma. *Let $-\infty \leqslant a < b \leqslant \infty$, and let $u \in C^1(a, b)$. Then $\partial|u| = (\operatorname{sgn} u)u'$ in the distributional sense; in particular, $\partial|u| = 0$ on $[u = 0]$.*

Proof. We use a sequence (F_k) of functions $F_k \in C^\infty(\mathbb{R};\mathbb{R})$ with $F_k(t) = |t| - \frac{1}{k}$ for $|t| \geqslant \frac{2}{k}$, $F_k(t) = 0$ for $|t| \leqslant \frac{1}{2k}$ and $|F_k'(t)| \leqslant 1$ for all $t \in \mathbb{R}$. (F_1 can be obtained as a convolution of $t \mapsto (|t| - 1)^+$ with a suitable C_{c}^∞-function, and then $F_k(t) := \frac{1}{k} F_1(kt)$ $(t \in \mathbb{R})$.)

By the chain rule we have $(F_k \circ u)' = (F_k' \circ u)u'$. Observe that $F_k \circ u \to |u|$, uniformly on compact subsets of (a, b), and that $(F_k \circ u)' \to (\operatorname{sgn} u)u'$ pointwise on (a, b). Moreover $|(F_k \circ u)'| \leqslant |u'|$ for all $k \in \mathbb{N}$. Therefore $(F_k \circ u)' \to (\operatorname{sgn} u)u'$ in L_1 on compact subsets of (a, b). This implies $\partial|u| = (\operatorname{sgn} u)u'$, by Lemma 4.11.

The last statement is a consequence of $\operatorname{sgn} u = 0$ on $[u = 0]$. □

Our aim is to establish similar properties in more general situations. The first point is that the chain rule also holds for distributional derivatives.

In the remainder of this section let $\Omega \subseteq \mathbb{R}^n$ be an open set.

9.13 Proposition. *Let $F \in C^1(\mathbb{R})$, $\|F'\|_\infty \leqslant 1$, $j \in \{1, \ldots, n\}$, $u, \partial_j u \in L_{1,\mathrm{loc}}(\Omega)$. Then $F \circ u \in L_{1,\mathrm{loc}}(\Omega)$, $\partial_j(F \circ u) = (F' \circ u)\partial_j u$.*

Proof. Without loss of generality assume that $F(0) = 0$; then $|F(t)| \leqslant |t|$ for all $t \in \mathbb{R}$. Being the distributional derivative of a function is a local property (see Exercise 4.3(c)), and therefore (after suitable multiplication by a C_{c}^∞-function) it is sufficient to treat the case when $\Omega = \mathbb{R}^n$ and $u, \partial_j u \in L_1(\mathbb{R}^n)$.

Let $(\rho_k)_{k\in\mathbb{N}}$ be a delta sequence in $C_c^\infty(\mathbb{R}^n)$. Then $u_k := \rho_k * u \to u$, $\partial_j u_k = \rho_k * \partial_j u \to \partial_j u$ in $L_1(\mathbb{R}^n)$, and for a suitable subsequence (u_{k_m}) these convergences also hold a.e. as well as dominated, in the sense that there exists $h \in L_1(\mathbb{R}^n)$ such that $|u_{k_m}|, |\partial_j u_{k_m}| \leqslant h$ for all $m \in \mathbb{N}$. As $u_k \in C^\infty(\mathbb{R}^n)$, one has $F \circ u_{k_m} \in C^1(\mathbb{R}^n)$, and

$$F \circ u_{k_m} \to F \circ u, \quad \partial_j(F \circ u_{k_m}) = (F' \circ u_{k_m})\partial_j u_{k_m} \to (F' \circ u)\partial_j u \quad \text{a.e.}$$

since F and F' are continuous. Furthermore, $|F \circ u_{k_m}| \leqslant |u_{k_m}| \leqslant h$, $|\partial_j(F \circ u_{k_m})| \leqslant |\partial_j u_{k_m}| \leqslant h$, by the hypotheses on F and on the subsequence, and therefore $F \circ u_{k_m} \to F \circ u$, $\partial_j(F \circ u_{k_m}) \to (F' \circ u)\partial_j u$ in $L_1(\mathbb{R}^n)$. This implies that $\partial_j(F \circ u) = (F' \circ u)\partial_j u$, by Lemma 4.11. □

Next we extend the chain rule of Proposition 9.13 to more general composition functions F.

9.14 Proposition. *Let $F \in C(\mathbb{R})$, and assume that there exist a function $G\colon \mathbb{R} \to \mathbb{R}$ and a sequence (F_k) in $C^1(\mathbb{R})$ such that $\|F_k'\|_\infty \leqslant 1$ for all $k \in \mathbb{N}$, and $F_k \to F$, $F_k' \to G$ pointwise as $k \to \infty$. Let $j \in \{1, \ldots, n\}$, $u, \partial_j u \in L_{1,\mathrm{loc}}(\Omega)$.*

Then $F \circ u \in L_{1,\mathrm{loc}}(\Omega)$, $\partial_j(F \circ u) = (G \circ u)\partial_j u$.

Proof. From Proposition 9.13 we know that $F_k \circ u \in L_{1,\mathrm{loc}}(\Omega)$, $\partial_j(F_k \circ u) = (F_k' \circ u)\partial_j u$. Without loss of generality we may assume that $F(0) = 0$ and that $F_k(0) = 0$ for all $k \in \mathbb{N}$, which implies that the functions $|F_k \circ u|$ are dominated by $|u|$. Then, applying the dominated convergence theorem on compact subsets of Ω and using Lemma 4.11 one obtains the assertions. □

9.15 Corollary. *Let $j \in \{1, \ldots, n\}$, $u, \partial_j u \in L_{1,\mathrm{loc}}(\Omega)$. Then $u^+, u \wedge 1 \in L_{1,\mathrm{loc}}(\Omega)$, $\partial_j(u^+) = \mathbf{1}_{[u>0]}\partial_j u$, $\partial_j(u \wedge 1) = \mathbf{1}_{[u<1]}\partial_j u$.*

Proof. Similarly to the construction of the sequence (F_k) in the proof of Lemma 9.12 one can construct a sequence (F_k) in $C^1(\mathbb{R})$ converging to the function $F := [t \mapsto t^+]$ pointwise, with $\|F_k'\|_\infty \leqslant 1$ $(k \in \mathbb{N})$ and such that $F_k' \to \mathbf{1}_{(0,\infty)}$ pointwise. Then Proposition 9.14 implies the assertion for u^+. The reasoning for $u \wedge 1$ is analogous. □

A **vector sublattice** V of $L_2(\Omega)$ is a subspace with the property that $u, v \in V$ implies that $u \vee v, u \wedge v \in V$. A subspace V is a vector sublattice if and only if for all $u \in V$ one has $u^+ \in V$ or $|u| \in V$; this equivalence follows from the relations

$$|u| = u^+ - u^-, \quad u^+ = \tfrac{1}{2}(u + |u|) \qquad (u \in L_2(\Omega)),$$
$$u \vee v = \tfrac{1}{2}(u + v + |u - v|), \quad u \wedge v = \tfrac{1}{2}(u + v - |u - v|) \qquad (u, v \in L_2(\Omega)).$$

The vector sublattice V is called **Stonean** if $u \wedge 1 \in V$ for all $u \in V$.

In this terminology, the following result says in particular that $H^1(\Omega)$ and $H_0^1(\Omega)$ are Stonean vector sublattices of $L_2(\Omega)$.

9.16 Theorem. *Let $u \in H^1(\Omega; \mathbb{R})$. Then $u^+, u \wedge 1 \in H^1(\Omega)$, $\nabla u^+ = \mathbf{1}_{[u>0]}\nabla u$ and $\nabla(u \wedge 1) = \mathbf{1}_{[u<1]}\nabla u$.*

If $u \in H_0^1(\Omega; \mathbb{R})$, then $u^+, u \wedge 1 \in H_0^1(\Omega)$.

Proof. It was shown in Corollary 9.15 that the indicated derivatives for u^+ and $u \wedge 1$ are the distributional derivatives. As they belong to $L_2(\Omega;\mathbb{R}^n)$, the first part of the theorem is proved.

Now let $u \in H_0^1(\Omega)$. There exists a sequence (u_k) in $H_c^1(\Omega)$ such that $u_k \to u$ in $H^1(\Omega)$. Then (u_k^+) is a bounded sequence in $H_c^1(\Omega) \subseteq H_0^1(\Omega)$, by the first part of the theorem, and $u_k^+ \to u^+$ in $L_2(\Omega)$. Thus, Remark 9.17 below shows that $u^+ \in H_0^1(\Omega)$. The argument for $u \wedge 1$ is analogous. □

9.17 Remark. Let V, H be Hilbert spaces, $V \hookrightarrow H$. Let (v_k) be a bounded sequence in V that is weakly convergent in H to $u \in H$. Then $u \in V$, and $v_k \to u$ weakly in V.

Indeed, there exist $v \in V$ and a subsequence (v_{k_m}) such that $v_{k_m} \to v$ weakly in V. Then $v_{k_m} \to v$ weakly in H as well; hence $v = u$. A standard sub-subsequence argument shows that $v_k \to u$ weakly in V; see Exercise 9.7. △

9.4 Invariance described by forms

In this section we transform the invariance criterion obtained in Proposition 9.6 to conditions on forms instead of operators. In the context of forms the associated C_0-semigroups are quasi-contractive, and the condition (9.3) is equivalent to the invariance of C under the semigroup, by Proposition 9.7.

We restrict our treatment to the case of embedded forms, i.e., we assume that V is a Hilbert space that is densely embedded into H and that $a\colon V \times V \to \mathbb{K}$ is a bounded quasi-coercive form. We recall that quasi-coercivity means that there exist $\omega \in \mathbb{R}$, $\alpha > 0$ such that

$$\operatorname{Re} a(u) + \omega\|u\|_H^2 \geqslant \alpha\|u\|_V^2 \qquad (u \in V). \tag{9.7}$$

Let $A \sim a$. It follows from Corollary 5.11 that $-A$ is the generator of a C_0-semigroup T; note the change of the sign in the generator with respect to Section 9.1.

The notation used above will be fixed troughout this section. Coming back to invariance, let $\varnothing \neq C \subseteq H$ be a closed convex set, and let $P := P_C$ be the minimising projection.

9.18 Proposition. *Let C be invariant under T. Then $P(V) \subseteq V$.*

9.19 Remark. At the first glance, this property might seem rather unexpected, because the elements of V have some quality (or 'regularity'), and it is surprising that this quality is preserved under P. To make the point, the elements of the domain of the generator will not be mapped to the domain of the generator, in general (see Exercise 9.4(c)). △

For the proof we single out a technical detail, which will be useful in several of the subsequent proofs.

9.20 Lemma. *Let $(u_n), (v_n)$ be sequences in V, $u_n \to u$ in H, (v_n) bounded in V, and*

$$\operatorname{Re} a(u_n, v_n - u_n) \geqslant 0 \qquad (n \in \mathbb{N}).$$

Then $u \in V$, and $u_n \to u$ weakly in V.

Proof. Without loss of generality we assume that $\omega \geqslant 0$. Using (9.7) we estimate

$$\begin{aligned}\alpha\|u_n\|_V^2 &\leqslant \operatorname{Re} a(u_n,u_n) + \omega\|u_n\|_H^2 \leqslant \operatorname{Re} a(u_n,v_n) + \omega\|u_n\|_H^2 \\ &\leqslant M\|u_n\|_V\|v_n\|_V + \omega c\|u_n\|_V\|u_n\|_H\,,\end{aligned}$$

where M denotes the bound of a and c the norm of the embedding $V \hookrightarrow H$. This inequality implies that (u_n) is bounded in V, and by Remark 9.17 it follows that $u_n \to u$ weakly in V. □

Proof of Proposition 9.18. Let $u \in V$. By (9.2) we obtain

$$\operatorname{Re}(Pu - y \,|\, u - y) = \operatorname{Re}(Pu - y \,|\, u - Pu) + \|Pu - y\|^2 \geqslant 0 \qquad (y \in C). \tag{9.8}$$

For $r > 0$ small enough (say, $0 < r < r_0 < \infty$) we define $R_r := (I + rA)^{-1}$. Then $AR_r = \frac{1}{r}(I + rA - I)R_r = \frac{1}{r}(I - R_r)$, so

$$a(R_r x, v) = (AR_r x \,|\, v) = \frac{1}{r}(x - R_r x \,|\, v) \qquad (x \in H,\ v \in V). \tag{9.9}$$

Now $R_r Pu \in C$, by Proposition 9.1. Using (9.9) and applying (9.8) with $y = R_r Pu$, we obtain

$$\operatorname{Re} a(R_r Pu, u - R_r Pu) = \frac{1}{r}\operatorname{Re}(Pu - R_r Pu \,|\, u - R_r Pu) \geqslant 0.$$

Since $R_r Pu \to Pu$ in H as $r \to 0$ (by Remark 2.13(b)), Lemma 9.20 implies that $Pu \in V$. □

We insert an auxiliary result that will be used in the proof of the next theorem.

9.21 Lemma. (a) *As in the proof of Proposition 9.18 we define* $R_r := (I + rA)^{-1}$ *for* $r \in (0, r_0)$, *with suitable* $r_0 > 0$. *Then* $R_r u \to u$ *weakly in* V *as* $r \to 0$, *for all* $u \in V$.

(b) *The set* dom(A) *is dense in* V.

(c) *Let* (u_n) *be a sequence converging weakly in* V *to* $u \in V$. *Then* $a(u_n, v) \to a(u, v)$ *for all* $v \in V$.

Proof. (a) Let $u \in V$. By (9.9) we obtain

$$a(R_r u, u - R_r u) = \frac{1}{r}(u - R_r u \,|\, u - R_r u) \geqslant 0.$$

Since $R_r u \to u$ in H as $r \to 0$, Lemma 9.20 implies that $R_r u \to u$ weakly in V as $r \to 0$.

(b) Part (a) shows that dom(A) is weakly dense in V. It follows that every bounded linear functional vanishing on dom(A) is zero, and this implies the assertion. (For another proof of (b) we refer to Exercise 5.3(a).)

(c) is clear since $a(\cdot, v) \in V'$. □

We now come to the fundamental result characterising invariance by properties of the form. The inequality (9.11) which appears below has already been commented upon in the paragraph before Proposition 9.6. In order to give a geometrical interpretation to (9.10) we note that, loosely speaking, the expression $a(Pu, u - Pu)$ can be understood as $(A(Pu) \,|\, u - Pu)$ (except that Pu does not necessarily lie in dom(A)). So, the condition gives information on the driving term $-Au(t)$ in the equation $u' = -Au$, whenever $u(t)$ is the image Pu of some $u \in H \setminus C$: in these points, the driving term 'points towards C'.

9.22 Theorem. *Under the previous assumptions the following properties are equivalent.*

(i) *C is invariant under T.*

(ii) *$P(V) \subseteq V$, and*

$$\operatorname{Re} a(Pu, u - Pu) \geqslant 0 \tag{9.10}$$

for all $u \in V$.

(iii) *There exists a dense subset D of V such that $P(D) \subseteq V$, and* (9.10) *holds for all $u \in D$.*

(iv) *$P(V) \subseteq V$, and*

$$\operatorname{Re} a(u, u - Pu) \geqslant -\omega \|u - Pu\|^2 \qquad (u \in V) \tag{9.11}$$

for some $\omega \in \mathbb{R}$.

Proof. (i) $\Rightarrow$ (ii). $P(V) \subseteq V$ was established in Proposition 9.18. Let $u \in V$. Then for $0 < r < r_0$ we have

$$\operatorname{Re} a(R_r Pu, u - Pu) = \frac{1}{r} \operatorname{Re}(Pu - R_r Pu \,|\, u - Pu) \geqslant 0$$

by (9.9) and (9.2), and from Lemma 9.21 we obtain

$$\operatorname{Re} a(Pu, u - Pu) \geqslant 0.$$

(ii) $\Rightarrow$ (iii) is trivial.

(iii) $\Rightarrow$ (ii). Let $u \in V$. There exists a sequence (u_n) in D such that $u_n \to u$ in V as $n \to \infty$. By hypothesis we have

$$\operatorname{Re} a(Pu_n, u_n - Pu_n) \geqslant 0 \qquad (n \in \mathbb{N}). \tag{9.12}$$

From the continuity of $P\colon H \to H$ we obtain $Pu_n \to Pu$ in H, and therefore Lemma 9.20 implies that $Pu \in V$ and that $Pu_n \to Pu$ weakly in V.

In order to prove (9.10) we first observe that an equivalent norm on V is given by

$$\|v\|_a := \big(\operatorname{Re} a(v) + \omega \|v\|_H^2\big)^{1/2} \qquad (v \in V),$$

with ω from (9.7). Hence, the weak convergence $Pu_n \to Pu$ in V implies that $\|Pu\|_a \leqslant \liminf_{n\to\infty} \|Pu_n\|_a$. Using (9.12) we obtain

$$\begin{aligned}\operatorname{Re} a(Pu) + \omega \|Pu\|_H^2 &\leqslant \liminf_{n\to\infty} \big(\operatorname{Re} a(Pu_n) + \omega \|Pu_n\|_H^2\big) \\ &\leqslant \liminf_{n\to\infty} \operatorname{Re} a(Pu_n, u_n) + \omega \|Pu\|_H^2 .\end{aligned}$$

Since $a(Pu_n, u_n) = a(Pu_n, u_n - u) + a(Pu_n, u) \to 0 + a(Pu, u)$ as $n \to \infty$ (recall Lemma 9.21(c)), we conclude that $\operatorname{Re} a(Pu, u - Pu) = \operatorname{Re} a(Pu, u) - \operatorname{Re} a(Pu) \geqslant 0$.

(ii) $\Rightarrow$ (iv) follows from the identity

$$a(u, u - Pu) = a(Pu, u - Pu) + a(u - Pu)$$

and the quasi-coercivity of a.

(iv) $\Rightarrow$ (i). Since $a(u, u - Pu) = (Au \,|\, u - Pu)$ for all $u \in \operatorname{dom}(A)$, condition (9.11) implies (9.3) for the generator $-A$ of T. Then the assertion follows from Proposition 9.6. $\square$

9.23 Example. We come back to Example 9.10. Let us again show step (ii) of its proof, i.e., that the C_0-semigroup on $L_2(\Omega;\mathbb{R})$ generated by the Dirichlet or Neumann Laplacian is sub-Markovian. We check property (ii) of Theorem 9.22 for the convex set $C = \{u \in L_2(\Omega;\mathbb{R});\ u \leqslant 1\}$ and the minimising projection $Pu = u \wedge 1$: Theorem 9.16 implies that P leaves $V = H_0^1(\Omega)$ (and also $V = H^1(\Omega)$) invariant, and

$$a(Pu, u - Pu) = \int \nabla(u \wedge 1) \cdot (\nabla u - \nabla(u \wedge 1))\,\mathrm{d}x = \int \mathbf{1}_{[u<1]}\nabla u \cdot \mathbf{1}_{[u\geqslant 1]}\nabla u\,\mathrm{d}x = 0$$

shows that property (ii) is satisfied. Hence C is invariant by Theorem 9.22. △

Further applications of Theorem 9.22 will be given in Section 10.4.

9.24 Remark. We emphasise that in Theorem 9.22 the associated semigroup is not assumed to be contractive. If the convex set C is a cone, i.e. C is invariant under multiplication by positive scalars, then the quasi-contractive case can easily be reduced to the contractive case by rescaling. This reduction is not possible if C is not a cone. An example for an application of Theorem 9.22 to a non-contractive semigroup in which C is not a cone can be found in Exercise 9.6. △

Notes

Clearly it is of fundamental interest to ask for criteria that describe when certain sets are invariant under the time evolution of a system, and questions of this kind have a long history, in particular in the finite-dimensional case, for linear and nonlinear problems.

In the seminal papers [BeDe58], [BeDe59], Beurling and Deny investigated such questions in infinite-dimensional spaces. Phillips gave a characterisation of positive contraction semigroups in [Phi62]. Invariance of closed convex sets was studied in [BrPa70; Section 2] in the more general context of nonlinear contraction semigroups.

The focus of our presentation is the characterisation of invariance via the form associated with the semigroup. An early (and little cited) result in this direction is a characterisation of sub-Markovian semigroups due to Kunita [Kun70], whose treatment included non-contractive semigroups. The form criteria for invariance of general closed convex sets (under contractive semigroups) are due to Ouhabaz [Ouh96]; we also refer the reader to Ouhabaz' book [Ouh05]. The invariance criterion in Theorem 9.22 is taken from [MVV05; Theorem 2.1]. A generalisation of Ouhabaz' results to a nonlinear setting can be found in [Bar96].

The lattice properties of the Sobolev spaces $H_0^1(\Omega)$ and $H^1(\Omega)$ were developed in the 70's; see the paper of Marcus and Mizel [MaMi79a], which also includes earlier references. Meanwhile, these properties can be found in several books on Sobolev spaces or partial differential equations. We refer to [EdEv87; Section VI.2] for a more general chain rule than presented in this chapter.

In Theorem 9.16 we treat chain rules for the composition of functions in $H^1(\Omega;\mathbb{R})$ with the functions $F\colon \mathbb{R} \to \mathbb{R}$, $F(t) := t^+$ and $F(t) := t \wedge 1$. The corresponding much more general chain rule on $W_{p,0}^1(\Omega;\mathbb{R})$ (where $1 < p < \infty$) for the case of Lipschitz continuous functions $F\colon \mathbb{R} \to \mathbb{R}$ satisfying $F(0) = 0$ is due to Stampacchia [Sta64; Lemme 1.1]. (The notation $W^1_{p,0}(\Omega;\mathbb{R})$ is analogous to $W^1_{1,0}(\Omega)$; see Appendix D.) Stampacchia proved that, for $1 < p < \infty$, the assignment $T_F u := F \circ u$ provides a mapping satisfying

$$T_F(W^1_{p,0}(\Omega)) \subseteq W^1_{p,0}(\Omega). \tag{9.13}$$

For this result we also refer to [EdEv87; Theorem 2.1].

It was shown by Marcus and Mizel [MaMi79b; Sect. 3] that the inclusion (9.13) also holds for $p = 1$ and that the mapping

$$T_F \colon W^1_{p,0}(\Omega) \to W^1_{p,0}(\Omega)$$

is continuous for all $p \in [1,\infty)$. We refer to the References in [MaMi79b] and [MaMi79a] for more information on related topics.

Exercises

9.1 Let (Ω,μ) be a measure space, $1 \leqslant p \leqslant \infty$, and let $A \in \mathcal{L}(L_p(\mu))$ be a positive operator, i.e. $Au \geqslant 0$ for all $u \in L_p(\mu)$ with $u \geqslant 0$.

(a) Show that $|Au| \leqslant A|u|$ for all $u \in L_p(\mu)$.

(b) Show that

$$\|A\| = \sup\bigl\{\|Au\|_p;\ u \in L_p(\mu),\ u \geqslant 0,\ \|u\|_p \leqslant 1\bigr\}.$$

9.2 Let H be a Hilbert space, $\varnothing \neq C \subseteq H$ a closed convex subset, and let P_C be the minimising projection onto C.

(a) Let $x \in H$, $z \in C$. Show that $z = P_C x$ if and only if

$$\operatorname{Re}(y - z \,|\, x - z) \leqslant 0 \qquad (y \in C);$$

see (9.2). (Hints: For '$\Rightarrow$' take $y \in C$ and look at the derivative at 0 of the function $[0,1] \ni t \mapsto \|x - (z + t(y-z))\|^2$. For '$\Leftarrow$' take $y \in C$ and compare $\|x - y\| = \|(x-z) + (z-y)\|$ with $\|x - z\|$.)

(b) Use part (a) to show that P_C is a contraction, i.e. $\|P_C x - P_C y\| \leqslant \|x - y\|$ for all $x, y \in H$.

9.3 Let (Ω,μ) be a measure space, $\check{C} \subseteq \mathbb{K}$ a closed convex set, $0 \in \check{C}$, and let $\check{P} \colon \mathbb{K} \to \check{C}$ be the minimising projection. Show that

$$C := \bigl\{u \in L_2(\mu);\ u(x) \in \check{C} \text{ for } \mu\text{-a.e. } x\bigr\} \neq \varnothing$$

is convex and closed.

Show that the minimising projection $P \colon L_2(\mu) \to C$ is given by $(Pu)(x) = \check{P}(u(x))$ $(x \in \Omega)$.

9.4 Let Δ_β be the Robin Laplacian from Section 7.5.

(a) Let β be real-valued. Show that the C_0-semigroup generated by Δ_β is positive.

(b) Let $\beta \geqslant 0$. Show that the C_0-semigroup generated by Δ_β is sub-Markovian.

(c) Give an example of a closed convex set C and a semigroup generator $-A$ such that $\operatorname{dom}(A)$ is not invariant under the minimising projection P_C. (Hint: Exercises 5.7 and 5.5(c).)

9.5 (a) Let V, H, a, j be as in Proposition 5.5, and assume that minus the operator associated with (a, j) is a generator. Let $\varnothing \neq C \subseteq H$ be convex and closed, P the minimising projection onto C. Let $\hat{P}\colon V \to V$ be a mapping satisfying $Pj = j\hat{P}$, and assume that

$$\operatorname{Re} a(u, u - \hat{P}u) \geqslant 0 \qquad (u \in V).$$

Show that C is invariant under the C_0-semigroup associated with (a, j). (Hint: Use Proposition 9.6.)

(b) Let V, H, a, j be as in Corollary 5.11, and let C, P be as in part (a). Show that C is invariant under the C_0-semigroup associated with (a, j) if and only if there exists a mapping $\hat{P}\colon V \to V$ such that $Pj = j\hat{P}$ and

$$\operatorname{Re} a(\hat{P}u, u - \hat{P}u) \geqslant 0 \qquad (u \in V).$$

Hint concerning 'only if': Apply Theorem 9.22, using Proposition 8.14 and the decomposition of the form domain V given in Proposition 8.16; assume without loss of generality that $V_j(a) \subseteq H$, $j(u) = u$ for all $u \in V_j(a)$, and define $\hat{P} := Pj$.

(c) Show that the C_0-semigroup generated by the Dirichlet-to-Neumann operator from Section 8.1 is sub-Markovian. (Hint: Show that the mapping $u \mapsto u \wedge 1$ on $H^1(\Omega)$ is consistent with the trace operator.)

9.6 Let $\Omega \subseteq \mathbb{R}^n$ be open, let $b \in L_\infty(\Omega; \mathbb{R}^n)$, and define the operator A in $L_2(\Omega)$ by

$$\begin{aligned} \operatorname{dom}(A) &:= \bigl\{u \in H_0^1(\Omega);\ -\Delta u \in L_2(\Omega)\bigr\},\\ Au &:= -\Delta u + b \cdot \nabla u \qquad (u \in \operatorname{dom}(A)). \end{aligned}$$

(a) Show that A is associated with a bounded quasi-coercive form on V, with $V := H_0^1(\Omega) \hookrightarrow H := L_2(\Omega)$ (and therefore $-A$ generates a quasi-contractive C_0-semigroup). Show that the semigroup generated by $-A$ is holomorphic of angle $\pi/2$ if $\mathbb{K} = \mathbb{C}$. (Hint: Use Exercise 7.7.)

(b) Show that the C_0-semigroup generated by $-A$ is sub-Markovian.

9.7 Let Ω be a topological space, let $x \in \Omega$, and let (x_n) be a sequence in Ω. Assume that each subsequence of (x_n) contains a subsequence converging to x. Show that then $x_n \to x$ as $n \to \infty$. (This fact will be referred to as the *(standard) sub-subsequence argument*.)

9.8 As in Theorem 9.22 let V and H be Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$, $a\colon V \times V \to \mathbb{K}$ a bounded quasi-coercive form, but assume additionally that $C = H_1$ is a (closed) subspace of H. Let $P \in \mathcal{L}(H)$ be the orthogonal projection onto H_1. Let $A \sim a$, and let T be the C_0-semigroup generated by $-A$.

(a) Show that H_1 is invariant under T if and only if $P(V) \subseteq V$ and $a(u, v) = 0$ for all $u \in P(V)$, $v \in (I - P)(V)$. (Hint: For the necessity show first that $\operatorname{Re} a(u, v) = 0$ for u, v as specified by considering $u \pm v$.)

(b) Now assume that H_1 is invariant under T, and denote by T_1 the restriction of T to H_1; then T_1 is a C_0-semigroup on H_1. Let $V_1 := V \cap H_1 = P(V)$ and $a_1 := a|_{V_1 \times V_1}$. (Note that a_1 is a bounded quasi-coercive form on V_1.) With $A_1 \sim a_1$, show that $A_1 = A \cap (H_1 \times H_1)$ and that $-A_1$ is the generator of T_1. (Hint: Use Proposition 9.11.)

Chapter 10

Interpolation of holomorphic semigroups

In the first section of this chapter we will present an extremely powerful tool of functional analysis, important in many areas: complex interpolation. It should be looked upon as the surprising fact that (elementary) complex methods are a useful tool for deriving inequalities. The main result is the Stein interpolation theorem; as a particular case one also obtains the famous and important Riesz–Thorin interpolation theorem.

For us, the important consequence will be that a holomorphic semigroup on some L_{p_1} that is bounded on some L_{p_0} for real times can be 'interpolated' holomorphically to other L_p-spaces. In the last section we demonstrate the interplay of invariance, interpolation and duality in applications to C_0-semigroups on L_2.

10.1 Interlude: the Stein interpolation theorem

Throughout this section the scalar field will be $\mathbb{K} = \mathbb{C}$.

10.1.1 The three lines theorem

In this subsection we prove a version of the maximum principle for holomorphic functions on an unbounded set. First we recall the maximum principle for bounded sets.

If $\Omega \subseteq \mathbb{C}$ is a bounded open set, and $h\colon \overline{\Omega} \to \mathbb{C}$ is continuous and holomorphic on Ω, then $\|h\|_{\overline{\Omega}} \leqslant \|h\|_{\partial\Omega}$. This is an easy consequence of Cauchy's integral formula. Here and in what follows we denote by $\|\cdot\|_M$ the supremum norm taken over the set M.

In this and the next section the set $S \subseteq \mathbb{C}$ will be the open strip

$$S := \{z \in \mathbb{C};\, 0 < \operatorname{Re} z < 1\}. \tag{10.1}$$

10.1 Proposition. *Let $h\colon \overline{S} \to \mathbb{C}$ be continuous, bounded, and holomorphic on S. Then*

$$\|h\|_{\overline{S}} \leqslant \|h\|_{\partial S}\,.$$

Proof. Let $n \in \mathbb{N}$. Then the function $\psi_n(z) := \frac{n}{z+n}$ is continuous on $\overline{S}$ and holomorphic on S. With $S_k := \{z \in S;\, |\operatorname{Im} z| < k\}$, for $k \in \mathbb{N}$, we have $\lim_{k\to\infty} \|\psi_n h\|_{S\setminus S_k} = 0$ since h is bounded. Using the maximum principle we conclude that

$$\|\psi_n h\|_{\overline{S}} \leqslant \|\psi_n h\|_{\overline{S}\setminus S_k} \to \|\psi_n h\|_{\partial S} \qquad (k \to \infty);$$

hence $\|\psi_n h\|_{\overline{S}} \leqslant \|\psi_n h\|_{\partial S} \leqslant \|h\|_{\partial S}$. Letting $n \to \infty$ we obtain the assertion. □

10.2 Remark. Taking more astute functions ψ_n one may weaken the assumption that h is bounded; it suffices that $|h(z)| \leqslant e^{ce^{\alpha|\operatorname{Im} z|}}$ with $c > 0$ and $0 < \alpha < \pi$ (see Exercise 10.1). △

The following result is a refinement of Proposition 10.1 in which one distiguishes between the suprema of h at $[\mathrm{Re} = 0]$ and at $[\mathrm{Re} = 1]$.

10.3 Theorem (Three lines theorem). *Let h be as in Proposition 10.1. Then*

$$\|h\|_{[\mathrm{Re}=\tau]} \leqslant \|h\|_{[\mathrm{Re}=0]}^{1-\tau} \|h\|_{[\mathrm{Re}=1]}^{\tau}$$

for all $0 \leqslant \tau \leqslant 1$.

Proof. For $j = 0, 1$ let $b_j > \|h\|_{[\mathrm{Re}=j]}$, and let $\tau \in [0, 1]$. We apply Proposition 10.1 to the function $z \mapsto \big(\frac{b_0}{b_1}\big)^z h(z)$ and obtain

$$\left(\frac{b_0}{b_1}\right)^{\tau} \|h\|_{[\mathrm{Re}=\tau]} \leqslant \left\| z \mapsto \left(\frac{b_0}{b_1}\right)^{z} h(z) \right\|_{\partial S} \leqslant \max\left\{ b_0, \left(\frac{b_0}{b_1}\right)^{1} b_1 \right\} = b_0\,;$$

hence $\|h\|_{[\mathrm{Re}=\tau]} \leqslant b_0^{1-\tau} b_1^{\tau}$. Taking the infima over b_1, b_2 one obtains the assertion. □

10.4 Remark. It follows that $\tau \mapsto \|h\|_{[\mathrm{Re}=\tau]}$ is a log-convex function, i.e. $\tau \mapsto \ln \|h\|_{[\mathrm{Re}=\tau]}$ is convex. △

10.1.2 The Stein interpolation theorem

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $\mathcal{A}_{\mathrm{c}} \subseteq \{A \in \mathcal{A};\, \mu(A) < \infty\}$ be a $\cap$-stable collection of subsets of Ω, with the property that the space of **simple functions** over $\mathcal{A}_{\mathrm{c}}$,

$$S(\mathcal{A}_{\mathrm{c}}) := \operatorname{lin}\{\mathbf{1}_A\,;\, A \in \mathcal{A}_{\mathrm{c}}\},$$

is dense in $L_1(\mu)$. (The index 'c' should be remindful of 'compact': if $\Omega \subseteq \mathbb{R}^n$ is an open set and μ is the Lebesgue measure, then one can choose as $\mathcal{A}_{\mathrm{c}}$ the collection of compact subsets of Ω.)

10.5 Remarks. (a) The $\cap$-stability of $\mathcal{A}_{\mathrm{c}}$ implies that the product uv of two elements $u, v \in S(\mathcal{A}_{\mathrm{c}})$ belongs to $S(\mathcal{A}_{\mathrm{c}})$; hence $S(\mathcal{A}_{\mathrm{c}})$ is an algebra.

Let $\mathcal{R}$ be the ring of subsets generated by $\mathcal{A}_{\mathrm{c}}$; then $S(\mathcal{A}_{\mathrm{c}}) = S(\mathcal{R})$. Indeed, clearly $S(\mathcal{A}_{\mathrm{c}}) \subseteq S(\mathcal{R})$. On the other hand one shows – see Exercise 10.2(a) – that $\mathbf{1}_A \in S(\mathcal{A}_{\mathrm{c}})$ for all $A \in \mathcal{R}$; hence $S(\mathcal{R}) \subseteq S(\mathcal{A}_{\mathrm{c}})$. In Exercise 10.2(b) the reader is asked to show that every function $u \in S(\mathcal{R})$ has a 'disjoint representation' $u = \sum_{A \in \mathcal{F}} c_A \mathbf{1}_A$, with a finite collection $\mathcal{F}$ of pairwise disjoint elements of $\mathcal{R}$ and suitable $c_A \in \mathbb{C}$ $(A \in \mathcal{F})$. As an important consequence one obtains $\varphi \circ u \in S(\mathcal{A}_{\mathrm{c}})$ for any function $\varphi\colon \mathbb{C} \to \mathbb{C}$ and all $u \in S(\mathcal{A}_{\mathrm{c}})$.

(b) A standard situation in measure theory would be to assume that $\mathcal{A}_{\mathrm{c}}$ is a $\cap$-stable generator of $\mathcal{A}$ consisting of sets of finite measure, with the property that each set $A \in \mathcal{A}$ of finite measure can be covered by countably many elements of $\mathcal{A}_{\mathrm{c}}$. Then $S(\mathcal{A}_{\mathrm{c}})$ is dense in $L_1(\mu)$; see Exercise 10.3(d). △

It will be important to know that $S(\mathcal{A}_c)$ is dense not only in $L_1(\mu)$ but also in $L_p(\mu)$, for all $p \in (1,\infty)$.

10.6 Lemma. *With the notation introduced above, the set $S(\mathcal{A}_c)$ is dense in $L_p(\mu)$ for all $p \in [1,\infty)$. More strongly, if $p_1, p_2 \in [1,\infty)$ and $u \in L_{p_1} \cap L_{p_2}(\mu)$, then there exists a sequence (u_n) in $S(\mathcal{A}_c)$ converging to u in $L_{p_1}(\mu)$ as well as in $L_{p_2}(\mu)$.*

Proof. (i) Let $p \in [1,\infty)$, $u \in L_p(\mu)$. Then $u\mathbf{1}_{[1/n \leqslant |u| \leqslant n]} \to u$ in $L_p(\mu)$ as $n \to \infty$, by the dominated convergence theorem. Hence it is sufficient to show that each function $v \in L_1 \cap L_\infty(\mu)$ belongs to the closure of $S(\mathcal{A}_c)$ in $L_p(\mu)$.

(ii) Let $v \in L_1 \cap L_\infty(\mu)$. Since $S(\mathcal{A}_c)$ is dense in $L_1(\mu)$, there exist a sequence $(\tilde v_k)$ in $S(\mathcal{A}_c)$ and a function $\tilde v \in L_1(\mu)$ such that $\tilde v_k \to v$ μ-a.e. as $k \to \infty$, and $|\tilde v_k| \leqslant \tilde v$ for all $k \in \mathbb{N}$. For $k \in \mathbb{N}$ we put $v_k := \tilde v_k \mathbf{1}_{[|\tilde v_k| \leqslant \|v\|_\infty + 1]}$; then $v_k \in S(\mathcal{A}_c)$, by Remark 10.5(a). Moreover, $|v_k| \leqslant \tilde v \wedge (\|v\|_\infty + 1) \in L_p(\mu)$ for all $k \in \mathbb{N}$ and $v_k \to v$ μ-a.e., so the dominated convergence theorem implies that $v_k \to v$ in $L_p(\mu)$ as $k \to \infty$.

(iii) Concerning the second statement we note that for $u \in L_{p_1} \cap L_{p_2}(\mu)$ the sequence defined in step (i) converges to u in $L_{p_1}(\mu)$ and in $L_{p_2}(\mu)$. Similarly, for $v \in L_1 \cap L_\infty(\mu)$, the sequence (v_n) constructed in step (ii) converges in $L_{p_1}(\mu)$ and in $L_{p_2}(\mu)$. $\square$

We will use the notation

$$L_{1,\mathrm{loc}}(\mathcal{A}_c) := \bigl\{u\colon \Omega \to \mathbb{C};\ u \text{ measurable},\ \mathbf{1}_A u \in L_1(\mu)\ (A \in \mathcal{A}_c),\ [u \neq 0]\ \sigma\text{-finite}\bigr\},$$

where we understand the elements of $L_{1,\mathrm{loc}}(\mathcal{A}_c)$ as equivalence classes of a.e. equal functions. With these conventions one has $uv \in L_1(\mu)$ for all $u \in S(\mathcal{A}_c)$, $v \in L_{1,\mathrm{loc}}(\mathcal{A}_c)$. For completeness we recall that a set $A \in \mathcal{A}$ is called σ-finite if there exists a sequence (A_n) in $\mathcal{A}$ with $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$ such that $A \subseteq \bigcup_{n\in\mathbb{N}} A_n$. We note that $L_p(\mu) \subseteq L_{1,\mathrm{loc}}(\mathcal{A}_c)$ for all $p \in [1,\infty)$.

Let the strip $S \subseteq \mathbb{C}$ be defined as in Subsection 10.1.1, and let $p_0, p_1, q_0, q_1 \in [1,\infty]$, $M_0, M_1 \geqslant 0$. For $\tau \in (0,1)$ we denote

$$\frac{1}{p_\tau} := \frac{1-\tau}{p_0} + \frac{\tau}{p_1}, \quad \frac{1}{q_\tau} := \frac{1-\tau}{q_0} + \frac{\tau}{q_1}, \quad M_\tau := M_0^{1-\tau} M_1^\tau.$$

Finally, let $L(S(\mathcal{A}_c), L_{1,\mathrm{loc}}(\mathcal{A}_c))$ denote the set of linear operators $B\colon S(\mathcal{A}_c) \to L_{1,\mathrm{loc}}(\mathcal{A}_c)$ (without continuity requirement), and let $\Phi\colon \overline{S} \to L(S(\mathcal{A}_c), L_{1,\mathrm{loc}}(\mathcal{A}_c))$ be a mapping satisfying the following two conditions.

(St1) $\|\Phi(j + \mathrm{i}s)u\|_{q_j} \leqslant M_j \|u\|_{p_j}$ for all $u \in S(\mathcal{A}_c)$, all $s \in \mathbb{R}$ and $j = 0, 1$. (The estimate means, in particular, that $\Phi(j + \mathrm{i}s)u \in L_{q_j}(\mu)$ for all the indicated terms.)

(St2) For all $A, B \in \mathcal{A}_c$ the function $\overline{S} \ni z \mapsto \int (\Phi(z)\mathbf{1}_A)\mathbf{1}_B \,\mathrm{d}\mu$ is continuous and bounded, and its restriction to S is holomorphic.

We can now state the Stein interpolation theorem, the main result of this section.

10.7 Theorem (Stein). *In the context described above it follows that*

$$\|\Phi(\tau + \mathrm{i}s)u\|_{q_\tau} \leqslant M_\tau \|u\|_{p_\tau}$$

for all $u \in S(\mathcal{A}_c)$, $s \in \mathbb{R}$ and $\tau \in (0,1)$.

Before proceeding to the proof we apply the theorem to the important special case of a constant function Φ.

10.8 Corollary (Riesz–Thorin). *Let the notation be as above, and let $B\colon S(\mathcal{A}_c) \to L_{1,\mathrm{loc}}(\mathcal{A}_c)$ be a linear operator satisfying $\|Bu\|_{q_j} \leqslant M_j\|u\|_{p_j}$ ($u \in S(\mathcal{A}_c)$, $j = 0,1$).*

Then for all $u \in S(\mathcal{A}_c)$ and $\tau \in (0,1)$ one has $\|Bu\|_{q_\tau} \leqslant M_\tau\|u\|_{p_\tau}$.

As an application we recall Example 9.10, where it was shown that the operators $\mathrm{e}^{t\Delta_{\mathrm{D}}}$ are sub-Markovian. Since we know that $\mathrm{e}^{t\Delta_{\mathrm{D}}}$ is contractive in $L_2(\Omega)$, we conclude from Corollary 10.8 that $\mathrm{e}^{t\Delta_{\mathrm{D}}}$ is contractive in $L_p(\Omega)$ for all $p \in [2,\infty]$.

The following fact will be needed in the proof of Theorem 10.7.

10.9 Lemma. *Let the notation $S(\mathcal{A}_c)$, $L_{1,\mathrm{loc}}(\mathcal{A}_c)$ be as above. Let $p, p' \in [1,\infty]$, $\frac{1}{p}+\frac{1}{p'} = 1$, $u \in L_{1,\mathrm{loc}}(\mathcal{A}_c)$, and assume that there exists $c \geqslant 0$ such that*

$$\Bigl|\int uv\,\mathrm{d}\mu\Bigr| \leqslant c\|v\|_{p'} \qquad (v \in S(\mathcal{A}_c)). \tag{10.2}$$

Then $u \in L_p(\mu)$, $\|u\|_p \leqslant c$.

Proof. (i) In the first step we show that (10.2) also holds with $|u|$ in place of u. Let $v \in S(\mathcal{A}_c)$. Then $\overline{\operatorname{sgn} u}\,\mathbf{1}_{[v\neq 0]} \in L_1(\mu)$, so there exists a sequence (w_k) in $S(\mathcal{A}_c)$ such that $w_k \to \overline{\operatorname{sgn} u}\,\mathbf{1}_{[v\neq 0]}$ μ-a.e. Then $v_k := v \operatorname{sgn} w_k \in S(\mathcal{A}_c)$ by Remark 10.5(a), $|uv_k| \leqslant |uv| \in L_1(\mu)$ for all $k \in \mathbb{N}$, and $uv_k = uv\operatorname{sgn} w_k \to |u|v$ μ-a.e., hence

$$\Bigl|\int |u|v\,\mathrm{d}\mu\Bigr| = \lim_{k\to\infty}\Bigl|\int uv_k\,\mathrm{d}\mu\Bigr| \leqslant c\sup_{k\in\mathbb{N}}\|v_k\|_{p'} \leqslant c\|v\|_{p'}\,.$$

(ii) In view of step (i) we now suppose, without loss of generality, that $u \geqslant 0$.

In the case $p = 1$ let $A \in \mathcal{A}$ be a set of finite measure. We choose a sequence (w_k) in $S(\mathcal{A}_c)$ such that $w_k \to \mathbf{1}_A$ μ-a.e. and put $v_k := \mathbf{1}_{[w_k\neq 0]}$. Then (v_k) is a sequence in $S(\mathcal{A}_c)_+$ with $\mathbf{1}_A \leqslant \liminf_{k\to\infty} v_k$, and Fatou's lemma implies

$$\int_A u\,\mathrm{d}\mu \leqslant \int u\liminf_{k\to\infty} v_k\,\mathrm{d}\mu \leqslant \liminf_{k\to\infty}\int uv_k\,\mathrm{d}\mu \leqslant c\sup_{k\in\mathbb{N}}\|v_k\|_{p'} \leqslant c.$$

As $[u\neq 0]$ is σ-finite, the case $p = 1$ is settled.

In the case $p \in (1,\infty]$ let $v \in L_{p'}(\mu)_+$. There exists a sequence $(\tilde v_k)$ in $S(\mathcal{A}_c)$ such that $\tilde v_k \to v$ in $L_{p'}(\mu)$ and μ-a.e. as $k\to\infty$. Putting $v_k := (\operatorname{Re}\tilde v_k)^+$ we obtain a sequence (v_k) in $S(\mathcal{A}_c)_+$, still converging to v in $L_{p'}(\mu)$ and μ-a.e. Hence Fatou's lemma yields

$$\int uv\,\mathrm{d}\mu = \int u\liminf_{k\to\infty} v_k\,\mathrm{d}\mu \leqslant \liminf_{k\to\infty}\int uv_k\,\mathrm{d}\mu \leqslant c\lim_{k\to\infty}\|v_k\|_{p'} = c\|v\|_{p'}\,.$$

In step (iii) we show that this inequality implies the assertion for $p \in (1,\infty]$.

(iii) Let $p \in (1,\infty)$. Since $[u\neq 0]$ is σ-finite, u can be approximated pointwise by an increasing sequence (u_k) in $L_p(\mu)_+$. The inequality proved in step (ii) yields

$$\|u_k\|_p^p \leqslant \int uu_k^{p-1}\,\mathrm{d}\mu \leqslant c\bigl\|u_k^{p-1}\bigr\|_{p'} = c\|u_k\|_p^{p-1}\,;$$

hence $\|u_k\|_p \leqslant c$ for all $k \in \mathbb{N}$, and consequently $\|u\|_p \leqslant c$.

In the case $p = \infty$ we obtain $\int u\mathbf{1}_A\,\mathrm{d}\mu \leqslant c\|\mathbf{1}_A\|_1$ for all $A \in \mathcal{A}$ with $\mu(A) < \infty$, and hence $\int_A (u-c)\,\mathrm{d}\mu \leqslant 0$. It follows that $u \leqslant c$ μ-a.e. because $[u\neq 0]$ is σ-finite. □

Proof of Theorem 10.7. In order to motivate the subsequent computations we first explain why it is sufficient to prove the inequality

$$\left|\int (\Phi(\tau)u)v\,\mathrm{d}\mu\right| \leqslant M_\tau \|u\|_{p_\tau}\|v\|_{q'_\tau} \qquad \big(u,v\in S(\mathcal{A}_{\mathrm{c}}),\ \tau\in(0,1)\big) \tag{10.3}$$

(where $\frac{1}{q_\tau}+\frac{1}{q'_\tau}=1$). In view of Lemma 10.9, this inequality implies the assertion for $s=0$; then the assertion for general $s\in\mathbb{R}$ follows by an application of the previous consideration to the function $z\mapsto\Phi(z+\mathrm{i}s)$.

For the proof of (10.3) let $\tau\in(0,1)$ and $u,v\in S(\mathcal{A}_{\mathrm{c}})$, without loss of generality $\|u\|_{p_\tau}=\|v\|_{q'_\tau}=1$. For $z\in\overline{S}$ put $\alpha(z):=\frac{1-z}{p_0}+\frac{z}{p_1}$, $\beta(z):=\frac{1-z}{q'_0}+\frac{z}{q'_1}$,

$$F(z):=\begin{cases}|u|^{\alpha(z)p_\tau}\operatorname{sgn}u & \text{if } p_\tau<\infty,\\ u & \text{if } p_\tau=\infty,\end{cases}$$

$$G(z):=\begin{cases}|v|^{\beta(z)q'_\tau}\operatorname{sgn}v & \text{if } q'_\tau<\infty,\\ v & \text{if } q'_\tau=\infty;\end{cases}$$

then $F(\tau)=u$, $G(\tau)=v$. Observe that $F(z),G(z)\in S(\mathcal{A}_{\mathrm{c}})$ for all $z\in\overline{S}$; cf. Remark 10.5(a).

Finally we define

$$h(z):=\int\big(\Phi(z)F(z)\big)G(z)\,\mathrm{d}\mu \qquad (z\in\overline{S}).$$

Then h is continuous, bounded, and holomorphic on S. In fact, since u and v have 'disjoint representations' as in Remark 10.5(a), it suffices to prove these properties for the case $u=c\mathbf{1}_A$, $v=d\mathbf{1}_B$, with $c,d\in\mathbb{C}$, $A,B\in\mathcal{A}_{\mathrm{c}}$. If $p_\tau,q'_\tau<\infty$, then

$$h(z)=|c|^{\alpha(z)p_\tau}(\operatorname{sgn}c)|d|^{\beta(z)q'_\tau}(\operatorname{sgn}d)\int(\Phi(z)\mathbf{1}_A)\mathbf{1}_B\,\mathrm{d}\mu \qquad (z\in\overline{S}),$$

and this function has the required properties, by condition (St2); for the boundedness of h note that $\operatorname{Re}\alpha(z),\operatorname{Re}\beta(z)\in[0,1]$ for all $z\in\overline{S}$. An analogous – easier – argument applies if one or both of p_τ and q'_τ are equal to ∞.

The definition of F is such that $\|F(\sigma+\mathrm{i}s)\|_{p_\sigma}=1$ for all $\sigma\in[0,1]$, $s\in\mathbb{R}$. Indeed, note that $\operatorname{Re}\alpha(\sigma+\mathrm{i}s)=\frac{1}{p_\sigma}$. If $p_\sigma<\infty$, then $p_\tau<\infty$, and therefore

$$\|F(\sigma+\mathrm{i}s)\|_{p_\sigma}^{p_\sigma}=\int\big||u|^{\alpha(\sigma+\mathrm{i}s)p_\tau}\big|^{p_\sigma}\,\mathrm{d}\mu=\|u\|_{p_\tau}^{p_\tau}=1.$$

If $p_\sigma=\infty$, then $\operatorname{Re}\alpha(\sigma+\mathrm{i}s)=0$, and therefore

$$\|F(\sigma+\mathrm{i}s)\|_{p_\sigma}=\big\||u|^0\mathbf{1}_{[u\neq0]}\big\|_\infty=1.$$

Analogously one shows that $\|G(\sigma+\mathrm{i}s)\|_{q'_\sigma}=1$. By condition (St1) it follows that

$$\|\Phi(\mathrm{i}s)F(\mathrm{i}s)\|_{q_0}\leqslant M_0\|F(\mathrm{i}s)\|_{p_0}=M_0\,,$$

$$|h(\mathrm{i}s)|=\left|\int\big(\Phi(\mathrm{i}s)F(\mathrm{i}s)\big)G(\mathrm{i}s)\,\mathrm{d}\mu\right|\leqslant\|\Phi(\mathrm{i}s)F(\mathrm{i}s)\|_{q_0}\|G(\mathrm{i}s)\|_{q'_0}\leqslant M_0$$

for all $s\in\mathbb{R}$. In the same way one obtains $|h(1+\mathrm{i}s)|\leqslant M_1$ for all $s\in\mathbb{R}$.

At this point we can apply Theorem 10.3 and conclude that $\left|\int(\Phi(\tau)u)v\,\mathrm{d}\mu\right|=|h(\tau)|\leqslant M_\tau$, thus completing the proof of (10.3). □

10.2 Interpolation of semigroups

As in Section 10.1 the scalar field will be $\mathbb{K} = \mathbb{C}$.

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $p_1 \in [1, \infty)$, $\theta \in (0, \pi/2]$, let T be a bounded holomorphic C_0-semigroup of angle θ on $L_{p_1}(\mu)$, and put $M_1 := \sup_{z\in\Sigma_\theta} \|T(z)\|_{\mathcal{L}(L_{p_1}(\mu))}$. Let $p_0 \in [1, \infty]$, $p_0 \neq p_1$, and assume that $T|_{[0,\infty)}$ is L_{p_0}-bounded; by this we mean that there exists $M_0 \geqslant 0$ such that

$$\|T(t)u\|_{p_0} \leqslant M_0 \|u\|_{p_0} \qquad (u \in L_{p_1} \cap L_{p_0}(\mu),\ t \geqslant 0).$$

The following result on interpolation of holomorphic semigroups is our main application of the Stein interpolation theorem.

10.10 Theorem. *Let the hypotheses be as above, let $\tau \in (0,1)$, and put $\theta_\tau := \tau\theta$, $\frac{1}{p_\tau} := \frac{1-\tau}{p_0} + \frac{\tau}{p_1}$, $M_\tau := M_0^{1-\tau} M_1^\tau$.*

Then for all $z \in \Sigma_{\theta_\tau,0}$ the operator $T(z)|_{L_{p_1}\cap L_{p_\tau}(\mu)}$ extends (uniquely) to an operator $T_\tau(z) \in \mathcal{L}(L_{p_\tau}(\mu))$, and T_τ thus defined is a bounded holomorphic C_0-semigroup of angle θ_τ, $\|T_\tau(z)\| \leqslant M_\tau$ for all $z \in \Sigma_{\theta_\tau,0}$.

Proof. We use the notation $S(\mathcal{A}_c)$, $L_{1,\mathrm{loc}}(\mathcal{A}_c)$ from the beginning of Subsection 10.1.2, with $\mathcal{A}_c := \{A \in \mathcal{A};\ \mu(A) < \infty\}$. Also let S be the strip defined in (10.1).

(i) The essential part of the proof is to establish the boundedness of $T(z)|_{S(\mathcal{A}_c)}$ with respect to the L_{p_τ}-norm, as follows.

Let $\theta' \in (0, \theta)$. With the 'semi-sector' $\Sigma'_{\theta'} := \{re^{i\alpha};\ r > 0,\ 0 \leqslant \alpha \leqslant \theta'\}$, the function $\psi\colon \overline{S} \to \Sigma'_{\theta'}$, $z \mapsto e^{i\theta' z}$ is continuous, bijective, and holomorphic on S. For $\sigma \in [0,1]$, it maps the line $[\mathrm{Re} = \sigma]$ onto the ray $\{re^{i\theta'\sigma};\ r > 0\}$. From these properties it follows that the function $\Phi := T \circ \psi\colon \overline{S} \to L(S(\mathcal{A}_c), L_{1,\mathrm{loc}}(\mathcal{A}_c))$ satisfies the hypotheses of Theorem 10.7 with $q_0 = p_0$, $q_1 = p_1$. We just comment on the hypothesis (St2): for $A \in \mathcal{A}_c$ the function $\overline{S} \ni z \mapsto T(\psi(z))\mathbf{1}_A \in L_{p_1}(\mu)$ is continuous, bounded, and holomorphic on S; hence for all $B \in \mathcal{A}_c$ the function $\overline{S} \ni z \mapsto \int \big(T(\psi(z))\mathbf{1}_A\big)\mathbf{1}_B \,\mathrm{d}\mu$ has the corresponding properties.

For every $s \in \mathbb{R}$, Theorem 10.7 implies that $T(\psi(\tau + is))|_{S(\mathcal{A}_c)}$ is bounded with respect to the L_{p_τ}-norm, with norm $\leqslant M_\tau$. In other words, the L_{p_τ}-norm of $T(z)|_{S(\mathcal{A}_c)}$ is bounded by M_τ, for all z in the ray $\{re^{i\theta'\tau};\ r > 0\}$. Note that the union of these rays, for $\theta' \in (0,\theta)$, is equal to the open semi-sector $\mathring{\Sigma}'_{\theta_\tau}$.

For the complementary open semi-sector $\{\bar{z};\ z \in \mathring{\Sigma}'_{\theta_\tau}\}$ the reasoning is analogous, and for $z \geqslant 0$ the boundedness statement follows from Corollary 10.8.

(ii) Let $z \in \Sigma_{\theta_\tau}$. As $S(\mathcal{A}_c)$ is a dense subspace of $L_{p_\tau}(\mu)$ (note that $p_\tau < \infty$), the operator $T(z)|_{S(\mathcal{A}_c)}$ has a unique extension $T_\tau(z) \in \mathcal{L}(L_{p_\tau}(\mu))$. The operators $T(z)$ and $T_\tau(z)$ coincide on $L_{p_1} \cap L_{p_\tau}(\mu)$ since for all $u \in L_{p_1} \cap L_{p_\tau}(\mu)$ there exists a sequence (u_n) in $S(\mathcal{A}_c)$ such that $u_n \to u$ in $L_{p_1}(\mu)$ as well as in $L_{p_\tau}(\mu)$, by Lemma 10.6.

In order to show that $\Sigma_{\theta_\tau} \ni z \mapsto T_\tau(z)$ is holomorphic we use the results of Section 3.1. For $u, v \in S(\mathcal{A}_c)$ the function $\Sigma_{\theta_\tau} \ni z \mapsto \int (T(z)u)v \,\mathrm{d}\mu$ is holomorphic. As $S(\mathcal{A}_c)$ is dense in $L_{p_\tau}(\mu)$ and in $L_{p'_\tau}(\mu)$ and $T_\tau\colon \Sigma_{\theta_\tau} \to \mathcal{L}(L_{p_\tau}(\mu))$ is bounded, Theorem 3.4 implies that $z \mapsto T_\tau(z) \in \mathcal{L}(L_{p_\tau}(\mu))$ is holomorphic on Σ_{θ_τ}.

It remains to show that T_τ is strongly continuous at 0. We use Hölder's inequality

$$\|v\|_{p_\tau} \leqslant \|v\|_{p_0}^{1-\tau} \|v\|_{p_1}^{\tau} \qquad \big(v \in L_{p_0} \cap L_{p_1}(\mu)\big)$$

and obtain $\|T(t)u - u\|_{p_\tau} \to 0$ as $t \to 0$ for all $u \in S(\mathcal{A}_c)$, because $\{\|T(t)u\|_{p_0}\,;\, t \geqslant 0\}$ is bounded and $T(\cdot)u$ is continuous at 0 in $L_{p_1}(\mu)$. Then the combination of Lemma 1.5 and Proposition 3.9 implies that T_τ is strongly continuous at 0. □

10.3 Adjoint semigroups

In this section we insert some information on adjoint semigroups. Let T be a C_0-semigroup on a (real or complex) Hilbert space H. Then clearly $T^* := (T(t)^*)_{t\geqslant 0}$ is a one-parameter semigroup on H. It turns out that T^* is strongly continuous, but this is not obvious; we will prove it by looking at the generator of T. However, for simplicity we will restrict our treatment to the case of quasi-contractive semigroups.

10.11 Theorem. *Let H be a Hilbert space, let T be a quasi-contractive C_0-semigroup on H, and let A be its generator. Then A^* is the generator of a quasi-contractive C_0-semigroup, and the generated C_0-semigroup is T^* as defined above.*

Proof. By rescaling we can reduce the situation to the case in which T is contractive. Then Theorem 2.7 implies that $(0,\infty) \subseteq \rho(A)$ and that $\|(\lambda - A)^{-1}\| \leqslant \frac{1}{\lambda}$ for all $\lambda > 0$.

As A is closed, A^* is densely defined; see Theorem 6.3(b). Similarly as in the proof of Theorem 6.3(b) one obtains $(\lambda - A^*)^{-1} = ((\lambda - A)^*)^{-1} = ((\lambda - A)^{-1})^*$ for all $\lambda > 0$, and it follows that $\lambda \in \rho(A^*)$ and $\|(\lambda - A^*)^{-1}\| \leqslant \frac{1}{\lambda}$. Therefore the Hille–Yosida theorem, Theorem 2.9, implies that A^* generates a contractive C_0-semigroup.

From the exponential formula, Theorem 2.12, we conclude that the C_0-semigroup generated by A^* is the adjoint semigroup T^*. □

10.12 Remarks. (a) As a particular consequence of Theorem 10.11 we conclude: the semigroup T is **self-adjoint**, i.e. $T(t)$ is self-adjoint for all $t \geqslant 0$, if and only if its generator A is self-adjoint. (See also Exercise 6.3(b).)

(b) If $\mathbb{K} = \mathbb{C}$ and the semigroup T in Theorem 10.11 is holomorphic of some angle $\theta \in (0, \pi/2]$, then T^* (defined as the adjoint of $T|_{[0,\infty)}$) has a holomorphic extension to the sector Σ_θ. This extension is given by

$$T^*(z) := T(\bar{z})^* \qquad (z \in \Sigma_\theta). \tag{10.4}$$

Indeed, it is not difficult to show that T^*, defined by (10.4), is holomorphic. Hence T^* is a holomorphic C_0-semigroup. △

10.13 Remarks. (a) Using the general Hille–Yosida generation theorem (see Exercise 2.5) one also obtains Theorem 10.11 for general C_0-semigroups on H.

(b) If X is a Banach space and T is a C_0-semigroup on X, then it is not generally true that $T'(t) := T(t)'$ $(t \geqslant 0)$ defines a C_0-semigroup on X', where $T(t)' \in \mathcal{L}(X')$ is the dual operator. It is true, however, if X is reflexive. This follows from the fact that any weakly continuous one-parameter semigroup is strongly continuous; see [EnNa00; Chap. I, Theorem 5.8]. △

10.4 Applications of invariance criteria and interpolation

Throughout this section let $(\Omega, \mathcal{A}, \mu)$ be a measure space.

We start our treatment with a criterion for a semigroup on a complex L_2-space to be real. A C_0-semigroup T on $L_2(\mu;\mathbb{C})$ is called **real** if $T(t)$ is a real operator for all $t \geqslant 0$.

10.14 Proposition. *Let V be a complex Hilbert space, $V \overset{\mathrm{d}}{\hookrightarrow} H := L_2(\mu;\mathbb{C})$, and let $a\colon V \times V \to \mathbb{R}$ be a bounded quasi-coercive form. Let A be the operator associated with a, and let T be the C_0-semigroup generated by $-A$.*

(a) *Then T is real if and only if $\operatorname{Re} u \in V$ for all $u \in V$ and $a(u,v) \in \mathbb{R}$ for all real $u, v \in V$.*

(b) *Assume that T is real, and let $H_{\mathrm{r}} := L_2(\mu;\mathbb{R})$, $V_{\mathrm{r}} := V \cap H_{\mathrm{r}}$ and $a_{\mathrm{r}} := a|_{V_{\mathrm{r}} \times V_{\mathrm{r}}}$. Then $V_{\mathrm{r}} \overset{\mathrm{d}}{\hookrightarrow} H_{\mathrm{r}}$, and a_{r} is quasi-coercive; let $A_{\mathrm{r}} \sim a_{\mathrm{r}}$. Then $A_{\mathrm{r}} = A \cap (H_{\mathrm{r}} \times H_{\mathrm{r}})$, and $-A_{\mathrm{r}}$ is the generator of the restriction T_{r} of T to H_{r}.*

Proof. (a) We recall that the minimising projection P onto the closed convex set $L_2(\mu;\mathbb{R})$ is given by $Pu = \operatorname{Re} u$ $(u \in L_2(\mu;\mathbb{C}))$. From the equivalence '(i) $\Leftrightarrow$ (ii)' in Theorem 9.22 we know that $L_2(\mu;\mathbb{R})$ is invariant under T if and only if $\operatorname{Re} u \in V$ for all $u \in V$ and for all real $u, v \in V$ one has

$$0 \leqslant \operatorname{Re} a\big(P(u \pm \mathrm{i}v), (I-P)(u \pm \mathrm{i}v)\big) = \operatorname{Re} a(u, \pm \mathrm{i}v) = \pm \operatorname{Im} a(u,v),$$

i.e. $a(u,v) \in \mathbb{R}$.

(b) Note that $V_{\mathrm{r}} = P(V)$ is dense in H_{r} since V is dense in H and $P \in \mathcal{L}(H)$. Moreover, $P(V) \subseteq V$ implies $V = V_{\mathrm{r}} + \mathrm{i}V_{\mathrm{r}}$. We now show that $A_{\mathrm{r}} \subseteq A$. Indeed, if $(u,f) \in A_{\mathrm{r}}$, then $u \in V_{\mathrm{r}} \subseteq V$, and for all $v \in V_{\mathrm{r}}$ we have $a(u,v) = (f \,|\, v)$. From $V = V_{\mathrm{r}} + \mathrm{i}V_{\mathrm{r}}$ we conclude that $a(u,v) = (f \,|\, v)$ for all $v \in V$, and thus $(u,f) \in A$. Regarding $L_2(\mu;\mathbb{C})$ as a real Banach space and invoking Proposition 9.11 we obtain the assertion. (Incidentally, this argument also proves the sufficiency in part (a).) □

An operator $S \in \mathcal{L}(L_2(\mu))$ is called ***L_1*-contractive** if $\|Su\|_1 \leqslant \|u\|_1$ for all $u \in L_2 \cap L_1(\mu)$, and S is called **substochastic** if S is positive and L_1-contractive. The same terminology will be used for semigroups if all the semigroup operators satisfy the corresponding property.

For simplicity we state the criteria in the next result only for real L_2-spaces. In the complex case one typically uses Proposition 10.14 to show that the semigroup T is real, and then one can apply Theorem 10.15 to the semigroup T_{r}.

10.15 Theorem (Beurling–Deny). *Let V be a real Hilbert space, $V \overset{\mathrm{d}}{\hookrightarrow} H := L_2(\mu;\mathbb{R})$, and let $a\colon V \times V \to \mathbb{R}$ be a bounded quasi-coercive form. Let A be the operator associated with a, and let T be the C_0-semigroup generated by $-A$. Then one has the following properties.*

(a) *T is positive if and only if $u^+ \in V$, $a(u^+, u^-) \leqslant 0$ for all $u \in V$.*

(b) *T is sub-Markovian if and only if $u \wedge 1 \in V$, $a(u \wedge 1, (u-1)^+) \geqslant 0$ for all $u \in V$.*

(c) *T is substochastic if and only if $u \wedge 1 \in V$, $a((u-1)^+, u \wedge 1) \geqslant 0$ for all $u \in V$.*

For the proof of part (c) we need an auxiliary result.

10.16 Lemma. *Let $S \in \mathcal{L}(L_2(\mu))$. Then S is L_∞-contractive if and only if S^* is L_1-contractive, and S is sub-Markovian if and only if S^* is substochastic.*

Proof. Assume that S is L_∞-contractive. Let $v \in L_2 \cap L_1(\mu)$, $\|v\|_1 \leqslant 1$. Then

$$\left|\int u\,\overline{S^*v}\,\mathrm{d}\mu\right| = \left|\int (Su)\bar{v}\,\mathrm{d}\mu\right| \leqslant 1 \qquad \big(u \in L_2 \cap L_\infty(\mu),\ \|u\|_\infty \leqslant 1\big),$$

and from Lemma 10.9, applied with $\mathcal{A}_\mathrm{c} := \{A \in \mathcal{A};\ \mu(A) < \infty\}$, one concludes that $\|S^*v\|_1 \leqslant 1$. The converse statement is proved in the same way.

For the second statement it remains to observe that S is positive if and only if S^* is positive. This equivalence easily follows from the fact that $f \in L_2(\mu)$ is positive if and only if $\int fg\,\mathrm{d}\mu \geqslant 0$ for all $g \in L_2(\mu)_+$, for which we refer to Exercise 10.5. □

Proof of Theorem 10.15. In view of Remarks 9.3 and 9.5, the statements (a) and (b) are easy consequences of '(i) ⇔ (ii)' in Theorem 9.22. For (a) we note that $u - u^+ = -u^-$, and for (b) we note that $u - u \wedge 1 = (u-1)^+$.

For part (c) we recall that $-A^*$ is the generator of the C_0-semigroup T^*, by Theorem 10.11, and that A^* is associated with the form a^*, by Theorem 6.11. Lemma 10.16 implies that T is substochastic if and only if T^* is sub-Markovian, and the latter is equivalent to $u \wedge 1 \in V$ and $a((u-1)^+, u \wedge 1) = a^*(u \wedge 1, (u-1)^+) \geqslant 0$ for all $u \in V$, by part (b). □

10.17 Remarks. (a) The equivalences formulated in Theorem 10.15 are the **Beurling–Deny criteria**. A form a satisfying the conditions in parts (b) and (c) of Theorem 10.15 is called a (**non-symmetric**) **Dirichlet form**.

(b) Theorem 10.15 remains true if one only requires the conditions for u from a dense subset of V; see property (iii) of Theorem 9.22. △

10.18 Theorem. *Let T be a C_0-semigroup on $L_2(\mu)$.*

(a) *Assume that T is sub-Markovian and substochastic. Then for all $p \in [1,\infty)$ the operators $T(t)|_{L_2 \cap L_p(\mu)}$ extend to operators $T_p(t) \in \mathcal{L}(L_p(\mu))$, and T_p thus defined is a contractive C_0-semigroup on $L_p(\mu)$. For $p, q \in [1,\infty)$ the semigroups T_p, T_q are* ***consistent****, i.e. $T_p(t)|_{L_p \cap L_q(\mu)} = T_q(t)|_{L_p \cap L_q(\mu)}$ for all $t \geqslant 0$.*

(b) *Assume that T is self-adjoint and sub-Markovian. Then the assertions of* (a) *hold. If $\mathbb{K} = \mathbb{C}$, then for all $p \in (1,\infty)$ the semigroup T_p extends to a contractive holomorphic C_0-semigroup of angle*

$$\theta_p = \begin{cases} \big(1 - \frac{1}{p}\big)\pi & \text{if } 1 < p < 2, \\ \frac{1}{p}\pi & \text{if } 2 \leqslant p < \infty. \end{cases}$$

Proof. (a) Let $1 \leqslant p < \infty$. For every $t > 0$, Exercise 10.6 (or Corollary 10.8, if $\mathbb{K} = \mathbb{C}$) implies that $T(t)|_{L_1 \cap L_\infty(\mu)}$ extends to a contractive operator $T_p(t)$ on $L_p(\mu)$. It is standard to show that T_p is a one-parameter semigroup. The strong continuity of T_p at 0 is obtained as follows. If $f \in L_p \cap L_2(\mu)$ and (t_n) is a null sequence in $(0,\infty)$, then $T(t_n)f \to f$ in $L_2(\mu)$, so for a subsequence one has $T(t_{n_k})f \to f$ a.e. Now the contractivity of T_p

in combination with Lemma 10.19, proved subsequently, implies that $T_p(t_{n_k})f \to f$ in $L_p(\mu)$. Applying the standard sub-subsequence argument – see Exercise 9.7 – one obtains $T_p(t_n)f \to f$ in $L_p(\mu)$ as $n \to \infty$. Lemma 1.5 concludes the argument.

The consistency is shown as in step (ii) of the proof of Theorem 10.10.

(b) From $T(t)^* = T(t)$ for all $t \geqslant 0$ and Lemma 10.16 it follows that T is also sub-stochastic. Thus part (a) is applicable.

Next, the generator A of T is self-adjoint by Remark 10.12(a), and as T is contractive, $-A$ is accretive. Now if $\mathbb{K} = \mathbb{C}$, then $-A$ is m-sectorial of angle 0, by Theorem 6.1. Hence A is the generator of a contractive holomorphic C_0-semigroup of angle $\pi/2$; see Theorem 3.20. In view of Theorem 10.10 this implies the remaining assertions. □

10.19 Lemma. *Let $1 \leqslant p < \infty$ and $f \in L_p(\mu)$. Let (f_n) be a sequence in $L_p(\mu)$ such that $f_n \to f$ a.e. and $\limsup_{n\to\infty} \|f_n\|_p \leqslant \|f\|_p$. Then $f_n \to f$ in $L_p(\mu)$.*

Proof. For $n \in \mathbb{N}$ put $\tilde{f}_n := \operatorname{sgn} f_n \,(|f| \wedge |f_n|)$. Then $\tilde{f}_n \to f$ in $L_p(\mu)$ by the dominated convergence theorem. Moreover $|f_n| = |\tilde{f}_n| + |f_n - \tilde{f}_n|$; hence $|f_n|^p \geqslant |\tilde{f}_n|^p + |f_n - \tilde{f}_n|^p$ for all $n \in \mathbb{N}$. Therefore $\|f_n - \tilde{f}_n\|_p^p \leqslant \|f_n\|_p^p - \|\tilde{f}_n\|_p^p$, which implies $\limsup \|f_n - \tilde{f}_n\|_p^p \leqslant 0$, and it follows that $f_n = (f_n - \tilde{f}_n) + \tilde{f}_n \to f$ in $L_p(\mu)$. □

Notes

The three lines theorem is generally attributed to Hadamard. In fact, in [Had96] Hadamard announced a variant, the 'three circles theorem'; the well-established version of the three lines theorem was first stated and proved by Doetsch [Doe20]. The Stein interpolation theorem, essentially in the form presented here, is contained in [Ste56]. We refer to this paper for some history of the Riesz–Thorin convexity theorem, finally proved by Thorin by the complex variable method, which initiated a whole new branch of functional analysis. In fact, the paper [Ste56] can be considered as the start of interpolation theory, for which we refer to the seminal paper of Calderón [Cal64] as well as to the monographs [BeLö76], [Lun18].

The application of invariance and interpolation as described in Section 10.4 is well-established in the theory of semigroups for diffusion equations, Schrödinger semigroups and related topics. The proof of Theorem 10.18(b) via interpolation appears natural and elegant; however, it does not yield the optimal angle of holomorphy for the L_p-semigroup. Indeed, Liskevich and Perelmuter showed in [LiPe95] that one obtains the larger angle $\arccos|1 - \frac{2}{p}|$ if one exploits more directly that the semigroup is sub-Markovian. There seems to be no way of obtaining the larger angle via interpolation; an example in [Voi96] shows that the angle from [LiPe95] is optimal. It came as a big surprise when Kriegler [Kri11] showed that extension to a contractive holomorphic C_0-semigroup on L_p of angle $\arccos|1 - \frac{2}{p}|$ is still valid if one merely assumes that T is self-adjoint and L_∞-contractive (and not necessarily positive). We refer to [HKV16] for an elementary proof of Kriegler's result.

Exercises

10.1 For this exercise let S be the strip

$$S := \{z \in \mathbb{C};\ -1/2 < \operatorname{Re} z < 1/2\}$$

(of width 1). Let $h\colon \overline{S} \to \mathbb{C}$ be continuous and holomorphic on S. Assume that h is bounded on ∂S and that there exist $c > 0$ and $\alpha \in (0, \pi)$ such that

$$|h(z)| \leqslant \mathrm{e}^{c\mathrm{e}^{\alpha|\operatorname{Im} z|}} \qquad (z \in S).$$

Show that h is bounded by $\|h\|_{\partial S}$. (Hint: Use $\psi_n(z) := \mathrm{e}^{-\frac{1}{n}(\mathrm{e}^{\mathrm{i}\beta z}+\mathrm{e}^{-\mathrm{i}\beta z})}$, with $\alpha < \beta < \pi$.)

10.2 Let Ω be a set, $\mathcal{A}_\mathrm{c}$ a $\cap$-stable collection of subsets of Ω (i.e. $A \cap B \in \mathcal{A}_\mathrm{c}$ for all $A, B \in \mathcal{A}_\mathrm{c}$), and let $\mathcal{R}$ be the ring of subsets generated by $\mathcal{A}_\mathrm{c}$ (i.e. $\mathcal{R}$ is the smallest ring containing $\mathcal{A}_\mathrm{c}$).

(a) Show that $\mathbf{1}_A \in S(\mathcal{A}_\mathrm{c})$ for all $A \in \mathcal{R}$.

Hint: Show first that the collection $\{A \subseteq \Omega;\ \mathbf{1}_A \in S(\mathcal{A}_\mathrm{c})\}$ is a ring. (Recall from the first paragraph of Remark 10.5(a) that $S(\mathcal{A}_\mathrm{c})$ is an algebra.)

(b) Show that every function $u \in S(\mathcal{R})$ has a 'disjoint representation' as described in Remark 10.5(a).

10.3 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, $\mathcal{A}_\mathrm{c} \subseteq \mathcal{A}$, $1 \leqslant p < \infty$. The issue of this exercise is to determine conditions under which $S(\mathcal{A}_\mathrm{c})$ is a dense subspace of $L_p(\mu)$. Without loss of generality the scalar field is assumed to be $\mathbb{K} = \mathbb{R}$.

A set $X \subseteq L_p(\mu)$ is called a vector sublattice if X is a subspace, and $f \vee g, f \wedge g \in X$ for all $f, g \in X$, or equivalently, $|f| \in X$ for all $f \in X$; see also Section 9.3.

(a) Let X be a vector sublattice of $L_p(\mu)$. Show that the closure of X in $L_p(\mu)$ is again a vector sublattice. (Hint: Show that $L_p(\mu) \ni f \mapsto |f| \in L_p(\mu)$ is Lipschitz continuous.)

(b) Assume that $\mu(\Omega) < \infty$, and let X be a closed vector sublattice of $L_p(\mu)$ containing the function $\mathbf{1}$. Show that $\{A;\ \mathbf{1}_A \in X\}$ is a σ-algebra.

(c) Assume that $\mu(\Omega) < \infty$, and let $\mathcal{A}_\mathrm{c}$ be a $\cap$-stable generator of $\mathcal{A}$ containing the set Ω. Show that $S(\mathcal{A}_\mathrm{c})$ is dense in $L_p(\mu)$. (Hint: Exercise 10.2.)

(d) Let $\mathcal{A}_\mathrm{c} \subseteq \{A \in \mathcal{A};\ \mu(A) < \infty\}$ be a $\cap$-stable generator of $\mathcal{A}$ such that each set $A \in \mathcal{A}$ of finite measure can be covered by countably many elements of $\mathcal{A}_\mathrm{c}$. Show that $S(\mathcal{A}_\mathrm{c})$ is dense in $L_p(\mu)$. (Hint: Use part (c) to show that $S(\mathcal{A}_\mathrm{c} \cap A)$ is dense in $L_p(A, \mu|_{\mathcal{A}\cap A})$, for each $A \in \mathcal{A}_\mathrm{c}$, where $\mathcal{A}_\mathrm{c} \cap A := \{B \cap A;\ B \in \mathcal{A}_\mathrm{c}\}$, and $\mathcal{A} \cap A$ is defined analogously.)

10.4 Let (Ω, μ) be a measure space. Show that

$$\big(\|\tfrac{1}{2}(f+g)\|_p^p + \|\tfrac{1}{2}(f-g)\|_p^p\big)^{1/p} \leqslant 2^{-1/p}\big(\|f\|_p^p + \|g\|_p^p\big)^{1/p} \tag{10.5}$$

for all $f, g \in L_p(\mu)$, $2 \leqslant p \leqslant \infty$.

Hint: Use the mapping

$$T\colon L_p(\mu) \times L_p(\mu) \to L_p(\mu) \times L_p(\mu),\ (f, g) \mapsto \big(\tfrac{1}{2}(f+g), \tfrac{1}{2}(f-g)\big).$$

Compute the norm of T for $p = 2$ and for $p = \infty$ and use the Riesz–Thorin interpolation theorem. (The inequality (10.5) is one of Clarkson's inequalities; see Adams [Ada75; Theorem 2.28]. The other inequalities of Clarkson involve p and the conjugate exponent p' and can also be obtained by interpolation, but this is more complicated. Clarkson's inequalities can be used to show that L_p-spaces, for $1 < p < \infty$, are uniformly convex – see [Ada75; Corollary 2.29] –, which in turn implies that these spaces are reflexive.)

10.5 Let (Ω, μ) be a measure space, and let $f \in L_2(\mu)$. Show that $f \in L_2(\mu)_+$ if and only if $\int fg \,\mathrm{d}\mu \geqslant 0$ for all $g \in L_2(\mu)_+$. (Hint: In the case $\mathbb{K} = \mathbb{R}$ consider $g = f^-$, in the case $\mathbb{K} = \mathbb{C}$ first show that $\operatorname{Im} f = 0$.)

10.6 (a) Let $p \in (1, \infty)$, $r \in [0, \infty)$. Show that

$$r = \inf_{\alpha \in (0,\infty)\cap\mathbb{Q}} \bigl(\tfrac{1}{p}\alpha^{1-p}r^p + (1 - \tfrac{1}{p})\alpha\bigr).$$

(b) Let (Ω, μ) be a measure space, and let $S \in \mathcal{L}(L_2(\mu))$ be sub-Markovian and substochastic. Show that S is L_p-contractive for all $p \in (1, \infty)$. (The case $\mathbb{K} = \mathbb{C}$ is already covered by Corollary 10.8, but not the case $\mathbb{K} = \mathbb{R}$!)

Hint: Using (a) twice show first that $S|u| \leqslant \bigl(S|u|^p\bigr)^{1/p}$ for simple functions.

(c) Let (Ω, μ) be a measure space, let $S \in \mathcal{L}(L_2(\mu))$ be sub-Markovian, and assume that there exists $c > 0$ such that $\frac{1}{c}S$ is substochastic. Show that S "extrapolates" to an operator $S_p \in \mathcal{L}(L_p(\mu))$ with $\|S_p\| \leqslant c^{1/p}$, for all $1 < p < \infty$.

10.7 (Continuation of Exercise 9.6) Let the hypotheses be as in Exercise 9.6. Assume additionally that $b \in C^1(\Omega; \mathbb{R}^n)$ and that $\omega := \sup_{x\in\Omega} \operatorname{div} b(x) < \infty$.

(a) Show that $\|T(t)u\|_1 \leqslant \mathrm{e}^{\omega t}\|u\|_1$ for all $t \geqslant 0$ and $u \in L_2 \cap L_1(\Omega)$, where T is the C_0-semigroup generated by the operator $-A$.

Hint: Use $C_\mathrm{c}^1(\Omega)$ as the dense subset of $V = H_0^1(\Omega)$ for the application of the invariance criterion to the semigroup $\bigl(\mathrm{e}^{-\omega t}T(t)\bigr)_{t\geqslant 0}$. Observe that on $C_\mathrm{c}^1(\Omega)$ one can transform the term $(b \cdot \nabla u \,|\, v)$ – using integration by parts – into an expression in which u appears without a derivative.

(b) Compute estimates for $\|T_p(t)\|$ in terms of ω for $t \geqslant 0$, $1 \leqslant p < \infty$, where T_p is the interpolated semigroup on $L_p(\Omega)$, analogous to Theorem 10.18(b).

10.8 The aim of this exercise is to give an alternative method of proving interpolation of holomorphy. (Unlike Theorem 10.10 this method provides no information about the angle of holomorphy.)

Let (Ω, μ) be a measure space, $p_1 \in (1, \infty)$, and let T be a contractive C_0-semigroup on $L_{p_1}(\mu)$ that is also L_{p_0}-contractive for some $p_0 \in [1, \infty] \setminus \{p_1\}$. Assume that T is holomorphic; recall from the Notes of Chapter 3 that this is equivalent to $\limsup_{t\to 0+} \|T(t) - I\| < 2$. Use this equivalence (even though it is not proved in the book) and the Riesz–Thorin theorem, Corollary 10.8, to prove that T extends to a holomorphic C_0-semigroup on $L_p(\mu)$ for all p strictly between p_0 and p_1.

Chapter 11

Elliptic operators

Elliptic operators with measurable coefficients are a classical topic in partial differential equations. They have realisations under diverse boundary conditions that generate semigroups, which results in well-posedness of parabolic initial boundary value problems. Form methods are very efficient to treat these problems and to derive properties of the solutions of the equations. It will be seen that many of the topics presented so far enter our treatment; in particular the properties of the Sobolev space $H^1(\Omega)$ and the Beurling–Deny criteria will play a decisive role. The latter lead to positivity and to sub-Markovian and substochastic behaviour of the generated semigroups.

In order to achieve these (and further) goals we will need additional lattice properties of $H^1(\Omega)$, treated in an interlude in Section 11.3.

11.1 Perturbation of bounded forms

Versions of the following perturbation result have already appeared in the proof of Theorem 7.16 and in Exercise 7.7(a).

11.1 Lemma. *Let V, H be Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$. Let $a\colon V \times V \to \mathbb{K}$ be a bounded quasi-coercive form. Let $b\colon V \times V \to \mathbb{K}$ be a bounded form, and assume that there exists $M \geqslant 0$ such that*

$$|b(u)| \leqslant M\|u\|_V\|u\|_H \qquad (u \in V).$$

Then $a + b\colon V \times V \to \mathbb{K}$ is quasi-coercive.

Proof. The quasi-coercivity of a means that there exist $\omega \in \mathbb{R}$ and $\alpha > 0$ such that

$$\operatorname{Re} a(u) + \omega\|u\|_H^2 \geqslant \alpha\|u\|_V^2$$

for all $u \in V$.

By the Peter–Paul inequality one has

$$\begin{aligned} \operatorname{Re} a(u) + \operatorname{Re} b(u) + \omega\|u\|_H^2 &\geqslant \alpha\|u\|_V^2 - M\|u\|_V\|u\|_H \\ &\geqslant \alpha\|u\|_V^2 - \frac{1}{2}\Bigl(\alpha\|u\|_V^2 + \frac{1}{\alpha}M^2\|u\|_H^2\Bigr). \end{aligned}$$

This implies

$$\operatorname{Re}\bigl(a(u) + b(u)\bigr) + \Bigl(\omega + \frac{M^2}{2\alpha}\Bigr)\|u\|_H^2 \geqslant \frac{\alpha}{2}\|u\|_V^2 \qquad (u \in V). \qquad \square$$

11.2 Elliptic operators

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $a_{jk} \in L_\infty(\Omega)$ $(j,k = 1,\dots,n)$ be coefficient functions satisfying the **ellipticity condition**

$$\operatorname{Re} \sum_{j,k=1}^{n} a_{jk}(x)\xi_k\overline{\xi_j} \geqslant \alpha|\xi|^2 \qquad (\xi \in \mathbb{K}^n) \tag{11.1}$$

for a.e. $x \in \Omega$, with some $\alpha > 0$, and let $b_j, c_j \in L_\infty(\Omega)$ $(j = 1,\dots,n)$, $d \in L_\infty(\Omega)$. Our aim is to define operators in $L_2(\Omega)$ corresponding to the "elliptic operator in divergence form" $\mathcal{A}$ written formally as

$$\begin{aligned}\mathcal{A}u &= -\sum_{j,k=1}^{n} \partial_j(a_{jk}\partial_k u) + \sum_{j=1}^{n}\big(b_j\partial_j u - \partial_j(c_j u)\big) + du \\ &= -\operatorname{div}\big((a_{jk})\nabla u\big) + b\cdot\nabla u - \operatorname{div}(cu) + du.\end{aligned} \tag{11.2}$$

The first order terms $b \cdot \nabla u$ and $\operatorname{div}(cu)$ in $\mathcal{A}u$ are sometimes called **drift terms**. We consider the form

$$a\colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{K}$$

given by

$$a(u,v) := \int_\Omega \Big(\sum_{j,k=1}^{n} a_{jk}\partial_k u\overline{\partial_j v} + \sum_{j=1}^{n}(b_j\partial_j u\overline{v} + c_j u\overline{\partial_j v}) + du\overline{v}\Big)\,\mathrm{d}x. \tag{11.3}$$

11.2 Proposition. *The form a is bounded and quasi-coercive.*

Proof. Let the form $a_0\colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{K}$ be given by

$$a_0(u,v) := \int_\Omega \sum_{j,k=1}^{n} a_{jk}\partial_k u\overline{\partial_j v}\,\mathrm{d}x.$$

Then the boundedness of the coefficients a_{jk} implies that a_0 is bounded. By the ellipticity condition (11.1) one obtains $\operatorname{Re} a_0(u) + \alpha\|u\|_2^2 \geqslant \alpha\|\nabla u\|_2^2 + \alpha\|u\|_2^2 = \alpha\|u\|_{H^1}^2$ for all $u \in H^1(\Omega)$. Thus a_0 is quasi-coercive.

Define $a_1\colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{K}$ by

$$a_1(u,v) := \int_\Omega \Big(\sum_{j=1}^{n}(b_j\partial_j u\overline{v} + c_j u\overline{\partial_j v}) + du\overline{v}\Big)\,\mathrm{d}x;$$

then $a = a_0 + a_1$. The boundedness of the coefficient functions b_j, c_j and d implies that a_1 is bounded. It also implies that there exists $M \geqslant 0$ such that

$$\begin{aligned}|a_1(u)| &\leqslant M\int_\Omega |\nabla u||u|\,\mathrm{d}x + \|d\|_\infty\|u\|_2^2 \\ &\leqslant M\|\nabla u\|_2\|u\|_2 + \|d\|_\infty\|u\|_2^2 \\ &\leqslant (M + \|d\|_\infty)\|u\|_{H^1}\|u\|_2\end{aligned}$$

for all $u \in H^1(\Omega)$. Now the assertion follows from Lemma 11.1. □

Coming back to (11.2), we want to give this formula a precise meaning, using the distributional divergence of a vector field $w \in L_{1,\mathrm{loc}}(\Omega;\mathbb{K}^n)$. We say that $\operatorname{div} w \in L_{1,\mathrm{loc}}(\Omega)$ if there exists $g \in L_{1,\mathrm{loc}}(\Omega)$ such that

$$\int_\Omega w \cdot \nabla\varphi \,\mathrm{d}x = -\int_\Omega g\varphi \,\mathrm{d}x$$

for all test functions $\varphi \in C_\mathrm{c}^\infty(\Omega)$, and then we say that $\operatorname{div} w = g$ **in the distributional sense**. (This definition was already mentioned in Exercise 4.5(b).) We now define the maximal operator A_max in $L_2(\Omega)$ by

$$A_\mathrm{max} u := -\operatorname{div}\bigl((a_{jk})\nabla u + cu\bigr) + b\cdot\nabla u + du$$

for $u \in H^1(\Omega)$ with $\operatorname{div}\bigl((a_{jk})\nabla u + cu\bigr) \in L_2(\Omega)$. In terms of the form a, the maximal operator can be written as

$$A_\mathrm{max} = \bigl\{(u,f) \in H^1(\Omega)\times L_2(\Omega);\, a(u,v) = (f\,|\,v)_{L_2}\ (v \in C_\mathrm{c}^\infty(\Omega))\bigr\}.$$

Next we define suitable restrictions of A_max. Let V be a closed subspace of $H^1(\Omega)$ containing $H_0^1(\Omega)$. Then the restriction of the form a to $V\times V$ is bounded and quasi-coercive, by Proposition 11.2. Denote by A_V the operator associated with $a|_{V\times V}$. Then $-A_V$ generates a C_0-semigroup T_V on $L_2(\Omega)$, and T_V is holomorphic if $\mathbb{K} = \mathbb{C}$; see Section 5.3. Note that the operator A_V is a restriction of A_max because $C_\mathrm{c}^\infty(\Omega) \subseteq H_0^1(\Omega)$.

11.3 Remark. Let $V \subseteq H_0^1(\Omega)$ be as above. For $(u,f) \in A_\mathrm{max}$ to be in A_V it is needed that

(i) $u \in V$ (and not merely $u \in H^1(\Omega)$),

(ii) $a(u,v) = (f\,|\,v)_{L_2}$ for all $v \in V$ (and not merely for $v \in C_\mathrm{c}^\infty(\Omega)$).

These conditions can be interpreted as boundary conditions. △

11.4 Examples. (a) Let us first consider the case $V = H_0^1(\Omega)$. Then $A_{H_0^1}$ is just the restriction of A_max to $\operatorname{dom}(A_\mathrm{max})\cap H_0^1(\Omega)$. We write $A_\mathrm{D} := A_{H_0^1}$ and call A_D the realisation of the elliptic operator $\mathcal{A}$ with **Dirichlet boundary conditions**. We define $T_\mathrm{D} := T_{H_0^1}$.

(b) Next we consider $V = H^1(\Omega)$. We define $T_\mathrm{N} := T_{H^1}$ and call $A_\mathrm{N} := A_{H^1}$ the realisation of the elliptic operator $\mathcal{A}$ with (generalised) **Neumann boundary conditions**. However, we will see that it is not the *normal* derivative of u that is 0 at the boundary, but rather the **conormal derivative**

$$\nu\cdot\bigl((a_{jk})\nabla u + cu\bigr)\big|_{\partial\Omega},$$

which is defined in terms of the coefficients of $\mathcal{A}$.

Clearly, the vector field $w := (a_{jk})\nabla u + cu \in L_2(\Omega;\mathbb{K}^n)$ cannot simply be restricted to $\partial\Omega$, so we first have to explain what is meant by $\nu\cdot w|_{\partial\Omega} = 0$. If Ω is bounded and has C^1-boundary, and $w \in C^1(\overline{\Omega};\mathbb{K}^n)$, then for all $\varphi \in C^1(\overline{\Omega})$ we obtain

$$\int_\Omega w\cdot\nabla\varphi\,\mathrm{d}x + \int_\Omega (\operatorname{div} w)\varphi\,\mathrm{d}x = \int_\Omega \operatorname{div}(w\varphi)\,\mathrm{d}x = \int_{\partial\Omega} \nu\cdot w\varphi\,\mathrm{d}\sigma,$$

by Gauss' theorem; note that the right-hand side vanishes for all φ if and only if $\nu\cdot w|_{\partial\Omega} = 0$.

Thus, for an arbitrary open set $\Omega \subseteq \mathbb{R}^n$ and $w \in L_2(\Omega; \mathbb{K}^n)$ satisfying $\operatorname{div} w \in L_2(\Omega)$, we will write '$\nu \cdot w|_{\partial\Omega} = 0$' if

$$\int_\Omega w \cdot \nabla v \,\mathrm{d}x + \int_\Omega (\operatorname{div} w) v \,\mathrm{d}x = 0 \qquad (v \in H^1(\Omega)).$$

(With this notation, a function $u \in H^1(\Omega)$ with $\Delta u \in L_2(\Omega)$ satisfies '$\nu \cdot \nabla u|_{\partial\Omega} = 0$' if and only if '$\partial_\nu u = 0$'; see (7.7).)

Now let $(u, f) \in A_{\max}$, and recall that A_{N} is a restriction of $A_{\max}$. For (u, f) to be in A_{N}, only condition (ii) of Remark 11.3 plays a role. Thus, $(u, f) \in A_{\mathrm{N}}$ if and only if

$$\int_\Omega \big((a_{jk})\nabla u + cu\big) \cdot \overline{\nabla v} \,\mathrm{d}x = \int_\Omega (f - b \cdot \nabla u - du)\overline{v} \,\mathrm{d}x \qquad (v \in H^1(\Omega)),$$

and since $f = A_{\max} u$, the latter is equivalent to

$$\int_\Omega \big((a_{jk})\nabla u + cu\big) \cdot \overline{\nabla v} \,\mathrm{d}x = -\int_\Omega \operatorname{div}\big((a_{jk})\nabla u + cu\big)\overline{v} \,\mathrm{d}x \qquad (v \in H^1(\Omega)).$$

Therefore, in terms of the above definition, we obtain

$$A_{\mathrm{N}} = \big\{(u, f) \in A_{\max};\ \text{'}\nu \cdot \big((a_{jk})\nabla u + cu\big)\big|_{\partial\Omega} = 0\text{'}\big\}.$$

(c) There are other possible choices of V. For example, assume that Ω is bounded and has C^1-boundary, and let $\Gamma \subseteq \partial\Omega$ be a Borel set. Then we define

$$V = \big\{u \in H^1(\Omega);\ \operatorname{tr} u = 0 \text{ a.e. on } \Gamma\big\}.$$

We call A_V the realisation of the elliptic operator $\mathcal{A}$ with **mixed boundary conditions** (with the interpretation 'Dirichlet on Γ, Neumann on $\partial\Omega \setminus \Gamma$'). △

11.5 Remarks. Assume that the coefficient matrix $(a_{jk}(x))$ is self-adjoint, i.e. $a_{jk}(x) = \overline{a_{kj}(x)}$ $(j, k = 1, \dots, n)$ for a.e. $x \in \Omega$. Then the ellipticity condition (11.1) says that the smallest eigenvalue of the matrix $(a_{jk}(x))$ is $\geqslant \alpha$ for a.e. $x \in \Omega$. The properties asserted in the following statements hold for any closed subspace V of $H^1(\Omega)$ containing $H_0^1(\Omega)$.

(a) In the complex case, the semigroup T_V generated by $-A_V$ is holomorphic of angle $\pi/2$. This follows from Exercise 7.7(c), applied with the decomposition $a = a_0 + a_1$ from Proposition 11.2; note that $a_0|_{V\times V}$ is symmetric by the assumption on (a_{jk}).

(b) If $b_j = \overline{c_j}$ $(j = 1, \dots, n)$ and d is real-valued, then a is symmetric, and the operator A_V is self-adjoint. △

We conclude this section with comments on the interplay between the real and complex cases.

11.6 Remarks. Assume that all the coefficient functions a_{jk}, b_j, c_j, d are real-valued. Then the ellipticity condition (11.1) needs to be checked for $\xi \in \mathbb{R}^n$ only; see Exercise 11.1.

(a) Let $\mathbb{K} = \mathbb{C}$, and let V be a closed subspace of $H^1(\Omega)$ containing $H_0^1(\Omega)$ with the property that $\operatorname{Re} u \in V$ for all $u \in V$. Then Proposition 10.14 implies that T_V is real. Moreover, with $V_{\mathrm{r}} = V \cap L_2(\Omega; \mathbb{R})$, the restriction of a to $V_{\mathrm{r}} \times V_{\mathrm{r}}$ is bounded and quasi-coercive, and the operator associated with $a|_{V_{\mathrm{r}}\times V_{\mathrm{r}}}$ is minus the generator of the restriction of T_V to $L_2(\Omega; \mathbb{R})$.

(b) Conversely, if V is a closed subspace of $H^1(\Omega;\mathbb{R})$ containing $H^1_0(\Omega;\mathbb{R})$, then $V_{\mathrm{c}} = V + \mathrm{i}V \subseteq H^1(\Omega;\mathbb{C})$ is a space as considered in (a). △

In order to study further properties of the semigroup on $L_2(\Omega;\mathbb{R})$ generated by an elliptic operator with real coefficients we need additional properties of the Sobolev space $H^1(\Omega)$.

11.3 Interlude: Further lattice properties of $H^1(\Omega)$

Throughout this section the scalar field will be $\mathbb{K} = \mathbb{R}$, and $\Omega \subseteq \mathbb{R}^n$ will be an open set. In Section 9.3 we have seen that $H^1(\Omega)$ is a vector sublattice of $L_2(\Omega)$. More precisely, if $u \in H^1(\Omega)$, then Theorem 9.16 implies that $u^+, u^- = (-u)^+, |u| = u^+ + u^- \in H^1(\Omega)$ and $\partial_j u^+ = \mathbf{1}_{[u>0]}\partial_j u,\ \partial_j u^- = -\mathbf{1}_{[u<0]}\partial_j u,\ \partial_j|u| = \partial_j u^+ + \partial_j u^- = (\operatorname{sgn} u)\partial_j u$. It follows that

$$\| |u| \|_{H^1} = \|u\|_{H^1} \qquad (u \in H^1(\Omega)) \tag{11.4}$$

since $\partial_j u = \partial_j u^+ - \partial_j u^- = \mathbf{1}_{[u\neq 0]}\partial_j u$ $(j = 1, \ldots, n)$. Incidentally, this last equality implies that $\mathbf{1}_{[u=0]}\partial_j u = 0$, i.e. $\partial_j u = 0$ a.e. on $[u = 0]$ ("Stampacchia's Lemma").

Next we show that the lattice operations are continuous.

11.7 Proposition. (a) *The mapping $H^1(\Omega) \ni u \mapsto |u| \in H^1(\Omega)$ is continuous.*

(b) *The mappings $(u, v) \mapsto u \wedge v$ and $(u, v) \mapsto u \vee v$ are continuous from $H^1(\Omega) \times H^1(\Omega)$ to $H^1(\Omega)$. In particular, the mappings $H^1(\Omega) \ni u \mapsto u^+, u^- \in H^1(\Omega)$ are continuous.*

Proof. (a) Let (u_k) be a sequence in $H^1(\Omega)$, $u_k \to u$ in $H^1(\Omega)$. Then $|u_k| \to |u|$ in $L_2(\Omega)$, and (11.4) implies that $(|u_k|)$ is bounded in $H^1(\Omega)$. By Remark 9.17 it follows that $|u_k| \to |u|$ weakly in $H^1(\Omega)$.

From (11.4) we also obtain

$$\| |u_k| \|_{H^1} = \|u_k\|_{H^1} \to \|u\|_{H^1} = \| |u| \|_{H^1} \qquad (k \to \infty),$$

from which we conclude that $|u_k| \to |u|$ in $H^1(\Omega)$; see the subsequent Remark 11.8.

(b) follows from $u \wedge v = \frac{1}{2}(u + v - |u - v|)$, $u \vee v = \frac{1}{2}(u + v + |u - v|)$ and part (a). □

11.8 Remark. Let H be a Hilbert space, (u_k) a sequence in H, $u_k \to u$ weakly in H, and $\|u_k\| \to \|u\|$ as $k \to \infty$. Then

$$\|u_k - u\|^2 = \|u_k\|^2 + \|u\|^2 - 2\operatorname{Re}(u_k \mid u) \to 0;$$

so $u_k \to u$ in H as $k \to \infty$. △

A **vector sublattice** V of $H^1(\Omega)$ is a subspace with the property that $u^+ \in V$ for all $u \in V$, analogously to the definition of vector sublattices of $L_2(\Omega)$. We already know that $H^1_0(\Omega)$ is a vector sublattice of $H^1(\Omega)$; see Theorem 9.16.

We now show that $H^1_0(\Omega)$ is even an ideal in $H^1(\Omega)$ (in the sense of vector lattices). An **order ideal** (or 'lattice ideal') V in $H^1(\Omega)$ is a subspace with the property that $u \in V$, $v \in H^1(\Omega)$, $|v| \leqslant |u|$ imply that $v \in V$. Note that then V is a vector sublattice (because $|u| \in V$ for all $u \in V$). On the other hand, if V is a vector sublattice with the property that for all $u \in V$, $v \in H^1(\Omega)$ with $0 \leqslant v \leqslant u$ one has $v \in V$, then V is an order ideal.

11.9 Proposition. *The space $H^1_0(\Omega)$ is an order ideal in $H^1(\Omega)$.*

Proof. Let $u \in H^1_0(\Omega)$, $v \in H^1(\Omega)$, $0 \leqslant v \leqslant u$. There exists a sequence (u_k) in $H^1_\mathrm{c}(\Omega)$ such that $u_k \to u$ in $H^1(\Omega)$. It follows from Proposition 11.7 that $u_k \wedge v \to u \wedge v = v$ in $H^1(\Omega)$ as $k \to \infty$. Since $u_k \wedge v \in H^1_\mathrm{c}(\Omega)$ we conclude that $v \in H^1_0(\Omega)$. □

11.10 Proposition. *Assume that Ω is bounded and has C^1-boundary. Let $\Gamma \subseteq \partial\Omega$ be a Borel set, and let V be as in Example 11.4(c). Then V is a closed order ideal in $H^1(\Omega)$.*

Proof. The linearity and continuity of the trace operator $\mathrm{tr}\colon H^1(\Omega) \to L_2(\partial\Omega)$ imply that V is a closed subspace.

For the proof that V is a vector sublattice of $H^1(\Omega)$, the important additional observation is that the trace operator is also a lattice homomorphism, i.e. $\mathrm{tr}(u \vee v) = \mathrm{tr}\, u \vee \mathrm{tr}\, v$ for all $u, v \in H^1(\Omega)$. This property is clear for $u, v \in H^1(\Omega) \cap C(\overline{\Omega})$ and carries over to $H^1(\Omega)$ by denseness and continuity; see Theorems 7.7 and 7.11 as well as Proposition 11.7(b). Therefore, if $u \in V$, then $\mathrm{tr}(u^+)|_\Gamma = (\mathrm{tr}\, u)^+|_\Gamma = 0$, i.e. $u^+ \in V$.

Finally let $u \in V$, $v \in H^1(\Omega)$, $0 \leqslant v \leqslant u$. Then $0 \leqslant \mathrm{tr}\, v|_\Gamma \leqslant \mathrm{tr}\, u|_\Gamma = 0$. This shows that $v \in V$. □

We will also need the following denseness properties.

11.11 Lemma. *The set $C^\infty_\mathrm{c}(\Omega)_+ := \{\varphi \in C^\infty_\mathrm{c}(\Omega)\,;\ \varphi \geqslant 0\}$ is dense in $H^1_0(\Omega)_+ := \{u \in H^1_0(\Omega)\,;\ u \geqslant 0\}$.*

Proof. (i) First we show that $H^1_\mathrm{c}(\Omega)_+ := \{u \in H^1_\mathrm{c}(\Omega)\,;\ u \geqslant 0\}$ is dense in $H^1_0(\Omega)_+$. Let $u \in H^1_0(\Omega)_+$. There exists a sequence (u_k) in $H^1_\mathrm{c}(\Omega)$ such that $u_k \to u$. Then clearly $u_k^+ \in H^1_\mathrm{c}(\Omega)_+$ for all $k \in \mathbb{N}$, and Proposition 11.7 implies that $u_k^+ \to u^+ = u$.

(ii) Let $u \in H^1_\mathrm{c}(\Omega)_+$, and let $\tilde{u}$ denote the extension of u to $\mathbb{R}^n$ by zero. Recall from Exercise 4.7(a) that $\tilde{u} \in H^1(\mathbb{R}^n)$. Let (ρ_k) be a delta sequence in $C^\infty_\mathrm{c}(\mathbb{R}^n)$. Then as in the proof of Corollary 4.5 one sees that $(\rho_k * \tilde{u})|_\Omega \in C^\infty_\mathrm{c}(\Omega)$ for large k, and $(\rho_k * \tilde{u})|_\Omega \to u$ in $H^1(\Omega)$; for the latter convergence one uses Lemma 4.16(b). Clearly $\rho_k * \tilde{u} \geqslant 0$ for all $k \in \mathbb{N}$, and it follows that $u \in \overline{C^\infty_\mathrm{c}(\Omega)_+}^{H^1(\Omega)}$. □

11.12 Proposition. *Assume that Ω is bounded and has continuous boundary. Then $\check{C}^\infty(\Omega)_+ := \{u \in \check{C}^\infty(\Omega)\,;\ u \geqslant 0\}$ is dense in $H^1(\Omega)_+ := \{u \in H^1(\Omega)\,;\ u \geqslant 0\}$.*

Proof. This property follows from the proof of Theorem 7.7, as explained in Remark 7.8(b). □

11.4 Elliptic operators with real coefficients

Here we continue with the setting of Section 11.2, assuming in addition that $\mathbb{K} = \mathbb{R}$. In particular, $\Omega \subseteq \mathbb{R}^n$ is an open set, the coefficients $a_{jk}, b_j, c_j, d \in L_\infty(\Omega)$ are all real-valued, and we assume that the ellipticity condition (11.1) is satisfied. Furthermore we work with the form a on $H^1(\Omega)$ from (11.2), as well as the operator A_V associated with $a|_{V\times V}$, for a closed subspace V of $H^1(\Omega)$ containing $H^1_0(\Omega)$, the operator $A_\mathrm{D} = A_{H^1_0}$ with Dirichlet boundary conditions, the operator $A_\mathrm{N} = A_{H^1}$ with Neumann boundary conditions, and the associated C_0-semigroups T_V, T_D and T_N.

11.13 Proposition. *Let V be a closed vector sublattice of $H^1(\Omega)$ containing $H^1_0(\Omega)$. Then the semigroup T_V generated by $-A_V$ is positive. In particular, T_{D} and T_{N} are positive.*

Proof. Let $u \in V$; then $u^+ \in V$. Corollary 9.15 implies that $\partial_j u^+ = \mathbf{1}_{[u>0]}\partial_j u$, $\partial_j u^- = -\mathbf{1}_{[u<0]}\partial_j u$, and therefore $\partial_k u^+ \partial_j u^- = 0$, $\partial_j u^+ u^- = 0$, $u^+\partial_j u^- = 0$, $u^+u^- = 0$, for $j,k = 1,\dots,n$. Thus $a(u^+,u^-) = 0$. Now it follows from Theorem 10.15(a) that T_V is positive.

The assertion for T_{D} and T_{N} is a consequence of Theorem 9.16. □

Proposition 11.13 also implies that the semigroup is positive for mixed boundary conditions; see Proposition 11.10.

Next we provide conditions under which the associated semigroup is sub-Markovian or substochastic. In Exercise 11.4 the reader is asked to prove that the required inequalities are in fact necessary.

11.14 Theorem. (a) *Assume that $c = (c_1,\dots,c_n)$ satisfies $\operatorname{div} c \in L_{1,\mathrm{loc}}(\Omega)$ and $\operatorname{div} c \leqslant d$. Then T_{D} is sub-Markovian.*

(b) *Let $b = (b_1,\dots,b_n)$ satisfy $\operatorname{div} b \in L_{1,\mathrm{loc}}(\Omega)$ and $\operatorname{div} b \leqslant d$. Then T_{D} is substochastic.*

Proof. (a) Let $u \in H^1_0(\Omega)$. Then $u \wedge 1 \in H^1_0(\Omega)$ and $\partial_j(u\wedge 1) = \mathbf{1}_{[u<1]}\partial_j u$, by Theorem 9.16. Since $u = u\wedge 1 + (u-1)^+$, it follows that $(u-1)^+ \in H^1_0(\Omega)$ and $\partial_j(u-1)^+ = \mathbf{1}_{[u\geqslant 1]}\partial_j u$. Thus $\partial_k(u\wedge 1)\partial_j(u-1)^+ = 0$ and $\partial_j(u\wedge 1)(u-1)^+ = 0$, for $j,k = 1,\dots,n$. It follows that

$$\begin{aligned} a(u\wedge 1,(u-1)^+) &= \int_\Omega \Big(\sum_{j=1}^n c_j(u\wedge 1)\partial_j(u-1)^+ + d(u\wedge 1)(u-1)^+\Big)\,\mathrm{d}x \\ &= \int_\Omega \Big(\sum_{j=1}^n c_j\partial_j(u-1)^+ + d(u-1)^+\Big)\,\mathrm{d}x \end{aligned}$$

(where the last equality holds because $u\wedge 1 = 1$ on $[u \geqslant 1]$). From the hypotheses on $\operatorname{div} c$ we obtain

$$\int_\Omega \Big(\sum_{j=1}^n c_j\partial_j\varphi + d\varphi\Big)\,\mathrm{d}x = \int_\Omega (-\operatorname{div} c + d)\varphi\,\mathrm{d}x \geqslant 0$$

for all $0 \leqslant \varphi \in C_{\mathrm{c}}^\infty(\Omega)$. Since $(u-1)^+$ can be approximated by positive test functions, by Lemma 11.11, it follows that $a(u\wedge 1,(u-1)^+) \geqslant 0$. Now Theorem 10.15(b) implies that T_{D} is sub-Markovian.

(b) The proof is analogous to (a) and uses Theorem 10.15(c). □

We now prove a similar result for boundary conditions that are defined by more general spaces V. Because of the boundary terms we need stronger regularity assumptions on Ω and on the coefficients c and d than in Theorem 11.14.

11.15 Theorem. *Assume that Ω is bounded and has C^1-boundary, and let V be a Stonean sublattice of $H^1(\Omega)$ containing $H^1_0(\Omega)$.*

(a) *If $c \in C^1(\overline{\Omega};\mathbb{R}^n)$, $\operatorname{div} c \leqslant d$ on Ω and $c\cdot\nu \geqslant 0$ on $\partial\Omega$, then T_V is sub-Markovian. (As before, $\nu(z) = (\nu_1(z),\dots,\nu_n(z))$ denotes the outer unit normal at $z \in \partial\Omega$.)*

(b) *If $b \in C^1(\overline{\Omega};\mathbb{R}^n)$, $\operatorname{div} b \leqslant d$ on Ω and $b\cdot\nu \geqslant 0$ on $\partial\Omega$, then T_V is substochastic.*

Proof. (a) As in the proof of Theorem 11.14 one has

$$a(u \wedge 1, (u-1)^+) = \int_\Omega \Big(\sum_{j=1}^n c_j \partial_j (u-1)^+ + d(u-1)^+ \Big) \,\mathrm{d}x \qquad (u \in V).$$

For $0 \leqslant \varphi \in C^1(\overline{\Omega})$ we now use Gauss' theorem (Theorem 7.3) to obtain

$$\int_\Omega \Big(\sum_{j=1}^n c_j \partial_j \varphi + d\varphi \Big) \,\mathrm{d}x = \int_\Omega \Big(-\sum_{j=1}^n \partial_j c_j + d \Big) \varphi \,\mathrm{d}x + \int_{\partial\Omega} \sum_{j=1}^n \nu_j c_j \varphi \,\mathrm{d}\sigma \geqslant 0.$$

By approximation – applying Proposition 11.12 – we deduce that $a(u \wedge 1, (u-1)^+) \geqslant 0$.
The proof of (b) is analogous. □

11.16 Remarks. Let Ω and V be as in Theorem 11.15, and assume that $\mathbf{1}_\Omega \in V$.
(a) If in Theorem 11.15(a) one has the equalities

$$\operatorname{div} c = d \quad \text{on } \Omega, \qquad c \cdot \nu = 0 \quad \text{on } \partial\Omega,$$

then T_V is not only sub-Markovian but **Markovian**, i.e. $T_V(t)\mathbf{1}_\Omega = \mathbf{1}_\Omega$ for all $t \geqslant 0$.
Indeed, as in the proof of Theorem 11.15 one shows that $a(\mathbf{1}_\Omega, v) = 0$ for all $v \in C^1(\overline{\Omega})$, and hence for all $v \in H^1(\Omega)$, by Theorem 7.7. It follows that $\mathbf{1}_\Omega \in \operatorname{dom}(A_V)$ and $A_V \mathbf{1}_\Omega = 0$, and then Theorem 1.13(a) implies the assertion.
(b) Similarly, if in Theorem 11.15(b) one has the equalities

$$\operatorname{div} b = d \quad \text{on } \Omega, \qquad b \cdot \nu = 0 \quad \text{on } \partial\Omega,$$

then T_V is not only substochastic but **stochastic**, i.e. $\|T_V(t)u\|_1 = \|u\|_1$ for all $t \geqslant 0$, $0 \leqslant u \in L_2 \cap L_1(\Omega)$. See Exercise 11.5. △

11.17 Remarks. (a) In the situation of Theorem 11.14(a), the C_0-semigroup T_{D} gives rise to a consistent family $\big(T_{\mathrm{D},p}\big)_{2 \leqslant p < \infty}$ of C_0-semigroups $T_{\mathrm{D},p}$ on $L_p(\Omega)$. This is seen in the same was as in the proof of Theorem 10.18(a). Similarly, in Theorem 11.14(b) one obtains a consistent family of C_0-semigroups $\big(T_{\mathrm{D},p}\big)_{1 \leqslant p \leqslant 2}$.
An analogous observation applies to Theorem 11.15.
(b) Deviating from the initial announcement of this section that only $\mathbb{K} = \mathbb{R}$ is treated, we include a comment on the complex case. It was mentioned in Section 11.2 that the C_0-semigroup T_V, with a suitable $H_0^1(\Omega) \subseteq V \subseteq H^1(\Omega)$, is holomorphic. Therefore Theorem 10.10 implies that the C_0-semigroups discussed in part (a) extend to holomorphic semigroups on $L_p(\Omega)$ if $p \neq 1$. △

11.5 Domination

We use the assumptions and notation from Section 11.2, as recalled at the beginning of Section 11.4, and throughout we assume that $\mathbb{K} = \mathbb{R}$.
We already know from Theorem 11.13 that the semigroups T_{D}, T_{N} and T_V are positive if V is a closed vector sublattice of $H^1(\Omega)$ containing $H_0^1(\Omega)$. Here we investigate the order relation between these semigroups.

11.18 Theorem. *Let $V \subseteq W$ be closed vector sublattices of $H^1(\Omega)$ containing $H_0^1(\Omega)$, and suppose that V is an order ideal in W. Then $0 \leqslant T_V(t) \leqslant T_W(t)$ for all $t \geqslant 0$, i.e. $0 \leqslant T_V(t)f \leqslant T_W(t)f$ for all $f \in L_2(\Omega)_+$.*

Proof. In view of the exponential formula, Theorem 2.12, it suffices to prove the domination property for the resolvents; i.e., for large $\lambda \in \mathbb{R}$ we have to show that

$$(\lambda + A_V)^{-1}f \leqslant (\lambda + A_W)^{-1}f \qquad (0 \leqslant f \in L_2(\Omega)).$$

Adding λ to the coefficient d we may assume $\lambda = 0$ and also that the form a is coercive.

Let $0 \leqslant f \in L_2(\Omega)$ and put $u_1 = A_V^{-1}f$, $u_2 = A_W^{-1}f$. Then $u_1 \in V_+ := V \cap L_2(\Omega)_+$ and $u_2 \in W_+ := W \cap L_2(\Omega)_+$ by the positivity of T_V, T_W and Proposition 9.1. Moreover

$$a(u_1, v) = (f \,|\, v)_{L_2} \qquad (v \in V),$$
$$a(u_2, v) = (f \,|\, v)_{L_2} \qquad (v \in W),$$

and thus $a(u_1 - u_2, v) = 0$ for all $v \in V$. We want to take $v := (u_1 - u_2)^+$, but at first it is only clear that $(u_1 - u_2)^+ \in W$. Observing that $0 \leqslant (u_1 - u_2)^+ \leqslant u_1$ we obtain $v = (u_1 - u_2)^+ \in V$ by the ideal property of V. Hence we have $a(u_1 - u_2, (u_1 - u_2)^+) = 0$. Now recall from the proof of Proposition 11.13 that $a(w^-, w^+) = a((-w)^+, (-w)^-) = 0$ for all $w \in H^1(\Omega)$. It follows that $a((u_1 - u_2)^+) = 0$, which implies $(u_1 - u_2)^+ = 0$ by our coercivity assumption, and consequently $u_1 \leqslant u_2$. We have shown that $A_V^{-1} \leqslant A_W^{-1}$. □

The following conclusion is immediate from Theorem 11.18 and the ideal property of $H_0^1(\Omega)$, Proposition 11.9.

11.19 Corollary. *Let V be a closed vector sublattice of $H^1(\Omega)$ containing $H_0^1(\Omega)$. Then $T_{\mathrm{D}}(t) \leqslant T_V(t)$ for all $t \geqslant 0$. In particular, $T_{\mathrm{D}}(t) \leqslant T_{\mathrm{N}}(t)$ for all $t \geqslant 0$.*

If additionally V is an order ideal in $H^1(\Omega)$, then $T_V(t) \leqslant T_{\mathrm{N}}(t)$ for all $t \geqslant 0$.

We mention that closed order ideals in $H^1(\Omega)$ were characterised by Stollmann [Sto93]. If A_V is the elliptic operator with mixed boundary conditions – see Example 11.4(c) –, then V is a closed order ideal in $H^1(\Omega)$, by Proposition 11.10, and so $T_V(t) \leqslant T_{\mathrm{N}}(t)$ for all $t \geqslant 0$.

Finally we want to prove domain monotonicity for Dirichlet boundary conditions. We consider the semigroup T_{D} on $L_2(\Omega)$, but we may also restrict the coefficients to an open subset $\Omega_1 \subseteq \Omega$ and consider the corresponding semigroup T_{D}^1 on $L_2(\Omega_1)$. We identify $L_2(\Omega_1)$ with a subspace of $L_2(\Omega)$ by extending functions in $L_2(\Omega_1)$ by zero on $\Omega \setminus \Omega_1$.

11.20 Theorem. *One has $T_{\mathrm{D}}^1(t)f \leqslant T_{\mathrm{D}}(t)f$ for all $t \geqslant 0$, $0 \leqslant f \in L_2(\Omega_1)$.*

Proof. By the exponential formula it suffices to show that $(\lambda + A_{\mathrm{D}}^1)^{-1}f \leqslant (\lambda + A_{\mathrm{D}})^{-1}f$ on Ω_1 for large enough $\lambda \in \mathbb{R}$ and $0 \leqslant f \in L_2(\Omega_1)$. As in the proof of Theorem 11.18 we may assume that the form a is coercive and that $\lambda = 0$.

Let $0 \leqslant f \in L_2(\Omega_1)$ and put $u_1 := (A_{\mathrm{D}}^1)^{-1}f$, $u_2 := (A_{\mathrm{D}})^{-1}f$. Then $u_1 \in H_0^1(\Omega_1)_+$, $u_2 \in H_0^1(\Omega)_+$ and

$$a(u_1, v) = (f \,|\, v)_{L_2(\Omega_1)} \qquad (v \in H_0^1(\Omega_1)),$$
$$a(u_2, v) = (f \,|\, v)_{L_2(\Omega)} \qquad (v \in H_0^1(\Omega)).$$

Observe that for $v \in H_0^1(\Omega_1)$ the extension $\tilde{v}$ to Ω by zero belongs to $H_0^1(\Omega)$ and that $\partial_j \tilde{v} = \widetilde{\partial_j v}$ for all $j \in \{1, \ldots, n\}$ (see Exercise 4.7(a)). In this sense we will consider $H_0^1(\Omega_1)$ as embedded into $H_0^1(\Omega)$ and omit the tilde. The above formulas imply $a(u_1 - u_2, v) = 0$ for all $v \in H_0^1(\Omega_1)$. Clearly $0 \leqslant (u_1 - u_2)^+ \leqslant u_1$, and it follows from Proposition 11.9 that $v := (u_1 - u_2)^+ \in H_0^1(\Omega_1)$. Hence $a(u_1 - u_2, (u_1 - u_2)^+) = 0$. Since $a((u_1 - u_2)^-, (u_1 - u_2)^+) = 0$ it follows that $a((u_1 - u_2)^+) = 0$ and thus $(u_1 - u_2)^+ = 0$ by the coercivity assumption. Therefore $u_1 \leqslant u_2$. □

Notes

The application of the Beurling–Deny criteria to elliptic operators has a long history. In the beginning, mostly symmetric operators were studied, in particular in connection with the heat equation with potential ("Schrödinger semigroups"); see for instance [ReSi78], [Dav89]. The application to non-symmetric operators seems to start with [MaRo92], [Ouh92], [Ouh96]; see [Ouh05] for a more recent presentation. Domination is also considered in these references. A domination criterion for positive semigroups is given in Exercise 11.6; it is interesting to note that the conditions stated there are in fact necessary; see [MVV05; Corollary 4.3]. We refer to [MVV05] and the literature cited there for more general domination results and for a treatment in the context of invariance criteria.

Perturbations as in Lemma 11.1 play an important role for evolution equations that are second order in time, and for their associated cosine functions; see Chapter 7 and Section 3.14 as well as the corresponding notes in [ABHN11].

We note that the treatment of second order elliptic operators by forms is particularly effective for operators in divergence form, as written in (11.2). This terminology concerns the second order part of the operator. Transforming an expression $\sum_{j,k=1}^n a_{jk}\partial_j\partial_k u$ into divergence form would require differentiability properties of the coefficients a_{jk} and produce first order terms.

In the context of Section 11.4 it is a remarkable fact that, even if the assumptions of Theorem 11.14 are not satisfied, the semigroup operators $T_{\mathrm{D}}(t)$ always extend to bounded operators $T_{\mathrm{D},p}(t)$ on $L_p(\Omega)$, and that $T_{\mathrm{D},p}$ thus defined is a C_0-semigroup on $L_p(\Omega)$, for all $p \in [1, \infty)$. The same is true for the semigroup T_{N} if Ω is bounded and has C^1-boundary (or, less restrictively, satisfies an interior cone condition). This follows from the results of [Dan00; Section 6], where it is shown that the semigroup operators have integral kernels satisfying Gaussian estimates. If $\mathbb{K} = \mathbb{C}$, then it follows that the semigroups $T_{\mathrm{D},p}$ and $T_{\mathrm{N},p}$ extend to holomorphic C_0-semigroups on $L_p(\Omega)$, for all $p \in [1, \infty)$, with the same angle of holomorphy as for $p = 2$ (see [Ouh95a; Theorem 2.4] for the case of self-adjoint semigroups, and [Hie96; Theorem 2.3] for the general case). In particular, if the coefficient matrices $(a_{jk}(x))$ are self-adjoint, then all the semigroups are holomorphic of angle $\pi/2$; cf. Remark 11.5(a). This is in contrast to Theorem 10.10, where the angle is obtained by interpolation and depends on p.

Exercises

11.1 Let $(a_{jk}) \in \mathbb{R}^{n\times n}$, $\alpha > 0$ be such that

$$\sum_{j,k=1}^{n} a_{jk}\xi_k\xi_j \geqslant \alpha|\xi|^2$$

for all $\xi \in \mathbb{R}^n$. Show that

$$\operatorname{Re}\sum_{j,k=1}^{n} a_{jk}\xi_k\overline{\xi_j} \geqslant \alpha|\xi|^2$$

for all $\xi \in \mathbb{C}^n$.

11.2 Let $\Omega \subseteq \mathbb{R}^2$ be open, $(a_{jk}) = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$, $b = c = 0$, $d = 0$.

(a) Show that $A_{\mathrm{D}} = -\Delta_{\mathrm{D}}$ (with A_{D} as defined in Example 11.4(a)).

(b) Assume that Ω is bounded and has C^1-boundary. Find the conormal derivative corresponding to A_{N}; cf. Example 11.4(b). Show that $A_{\mathrm{N}} \neq -\Delta_{\mathrm{N}}$ if $\Omega \neq \varnothing$. (Hint: Find $u, v \in H^1(\Omega;\mathbb{R})$ such that $\int_\Omega (a_{jk})\nabla u\cdot\nabla v\,\mathrm{d}x \neq \int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}x$ and apply Exercise 5.3(d).)

11.3 Let a_{jk}, b_j, c_j, d be as in Section 11.2.

(a) Assume additionally that $b \in C_{\mathrm{b}}^1(\Omega;\mathbb{K}^n)$ (bounded derivatives!), $c = -b$, $d = -\operatorname{div} b$. Let the formal elliptic operators $\mathcal{A}_1, \mathcal{A}_2$, in the sense of (11.2), be defined by

$$\mathcal{A}_1 u := -\sum_{j,k=1}^{n} \partial_j(a_{jk}\partial_k u) + b\cdot\nabla u, \quad \mathcal{A}_2 u := -\sum_{j,k=1}^{n} \partial_j(a_{jk}\partial_k u) - \operatorname{div}(cu) + du.$$

Show that $A_{1,\mathrm{D}} = A_{2,\mathrm{D}}$.

(b) Assume additionally that $b, c \in C_{\mathrm{b}}^1(\Omega;\mathbb{K}^n)$, $c = b$, and let $\mathcal{A}$ be defined by

$$\mathcal{A}u := -\sum_{j,k=1}^{n} \partial_j(a_{jk}\partial_k u) + b\cdot\nabla u - \operatorname{div}(cu).$$

Show that A_{D} is associated with a formal elliptic operator without drift terms.

11.4 Let $\mathbb{K} = \mathbb{R}$, and let $\Omega \subseteq \mathbb{R}^n$ be open.

(a) Let $u \in H^1(\Omega)_+$. Show that $(u - \frac{1}{n})^+ \to u$ in $H^1(\Omega)$.

(b) In the situation of Theorem 11.14, show that $\operatorname{div} c \leqslant d$ is a necessary condition for T_{D} to be sub-Markovian if $\operatorname{div} c \in L_{1,\mathrm{loc}}(\Omega)$. (Hint: Let $\varphi \in C_{\mathrm{c}}^\infty(\Omega)_+$, put $u_n := (\varphi - \frac{1}{n})^+$ and show that $\int (c\cdot\nabla u_n + du_n)\,\mathrm{d}x \geqslant 0$, for all $n \in \mathbb{N}$).

11.5 Prove Remark 11.16(b).

11.6 Let $\mathbb{K} = \mathbb{R}$, (Ω,μ) a measure space, $H := L_2(\mu)$, V, W Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$, $W \overset{\mathrm{d}}{\hookrightarrow} H$, and let $a\colon V\times V \to \mathbb{R}$, $b\colon W\times W \to \mathbb{R}$ be bounded quasi-coercive forms. Denote by A the operator associated with a and by B the operator associated with b. Assume that the semigroups T generated by $-A$ and S generated by $-B$ are both positive. (Recall that then V and W are sublattices of $L_2(\mu)$, by Theorem 10.15(a).) Assume that

(i) V is an order ideal in W;

(ii) $a(u,v) \geqslant b(u,v)$ for all $0 \leqslant u, v \in V$.

Show that $T(t) \leqslant S(t)$ for all $t \geqslant 0$. (Hint: Proceed as in the proof of Theorem 11.18.)

11.7 Let $\mathbb{K} = \mathbb{R}$, let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary, $\beta, \beta_1, \beta_2 \in L_\infty(\partial\Omega)$, and let $T_\beta, T_{\beta_1}, T_{\beta_2}$ be the semigroups on $L_2(\Omega)$ generated by the corresponding Robin Laplacians, as in Section 7.5. Use Exercise 11.6 to show:

(a) $T_{\beta_1}(t) \leqslant T_{\beta_2}(t)$ for all $t \geqslant 0$ if $\beta_1 \geqslant \beta_2$.

(b) $T_{\mathrm{D}}(t) \leqslant T_\beta(t) \leqslant T_{\mathrm{N}}(t)$ for all $t \geqslant 0$ if $\beta \geqslant 0$, where T_{D} and T_{N} are the semigroups on $L_2(\Omega)$ generated by the Dirichlet and Neumann Laplacians.

11.8 Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary, and let a_{jk}, b_j, c_j, d be as in Section 11.2. In addition let $\beta \in L_\infty(\partial\Omega)$. Define a suitable form on $V := H^1(\Omega)$ such that the associated operator A_β is the realisation of the elliptic operator $\mathcal{A}$ from (11.2) with (generalised) Robin boundary conditions,

$$A_\beta = \bigl\{(u, f) \in A_{\max}\,;\ \nu \cdot \bigl((a_{jk})\nabla u + cu\bigr)\big|_{\partial\Omega} + \beta u|_{\partial\Omega} = 0\bigr\}.$$

Hint: Recall the proof of Theorem 7.16. It is part of the task to define the weak normal trace $\nu \cdot w|_{\partial\Omega}$ of a vector field $w \in L_2(\Omega; \mathbb{K}^n)$ with $\operatorname{div} w \in L_2(\Omega)$.

Chapter 12

Sectorial forms

In this chapter we study sectorial forms $a\colon V \times V \to \mathbb{C}$ in the general situation when V is just a vector space (and not necessarily a Hilbert space). As in Chapter 5, $j\colon V \to H$ is a linear mapping with dense range in a Hilbert space H. In Section 12.1 we describe how to associate an m-sectorial operator in H with the form (a, j) in this general setting. If the form is closed (see Section 12.1 for the definition), then the approach presented here is equivalent to the approach via j-coercive forms. In Section 12.4 we give an alternative description of the associated operator, in the non-closed case. Two examples are given that illustrate the theory: the Robin Laplacian and the Dirichlet-to-Neumann operator are revisited, but now on rough domains.

12.1 Operators associated with sectorial forms

As defined in Section 5.1, a form a is a sesquilinear mapping $a\colon V \times V \to \mathbb{K}$, where V is a $\mathbb{K}$-vector space. In contrast to previous chapters we do not assume V to carry any additional structure. For our purposes it will be convenient to simply call V the **domain** of a and to put $\operatorname{dom}(a) := V$. Evidently, this is a misuse of the symbol 'dom' since the domain of a, in the usual sense, is the cartesian product $V \times V$. However, this notation is useful, has a long history, and should not lead to confusion.

Let $\mathbb{K} = \mathbb{C}$, and let $a\colon \operatorname{dom}(a) \times \operatorname{dom}(a) \to \mathbb{C}$ be a form. We recall that, by definition, a is sectorial if there exists $\theta \in [0, \pi/2)$ such that $a(u) \in \overline{\Sigma_\theta}$ for all $u \in \operatorname{dom}(a)$, and that this holds if and only if a is accretive and there exists $c \geqslant 0$ such that

$$|\operatorname{Im} a(u)| \leqslant c \operatorname{Re} a(u) \tag{12.1}$$

for all $u \in \operatorname{dom}(a)$. We further recall that $a = \operatorname{Re} a + \mathrm{i} \operatorname{Im} a$ with the symmetric forms $\operatorname{Re} a, \operatorname{Im} a\colon \operatorname{dom}(a) \times \operatorname{dom}(a) \to \mathbb{C}$, and that $\operatorname{Re} a(u) = (\operatorname{Re} a)(u)$ and $\operatorname{Im} a(u) = (\operatorname{Im} a)(u)$ for all $u \in \operatorname{dom}(a)$. Using Proposition 5.2 we see that (12.1) implies the key inequality

$$|a(u, v)| \leqslant (1 + c)(\operatorname{Re} a(u))^{1/2}(\operatorname{Re} a(v))^{1/2}, \tag{12.2}$$

which expresses a sort of intrinsic continuity of a. Observe that an accretive form a is symmetric if and only if (12.1) or (12.2) holds with $c = 0$.

Let H be a complex Hilbert space. If $a\colon \operatorname{dom}(a) \times \operatorname{dom}(a) \to \mathbb{C}$ is a sectorial form and $j\colon \operatorname{dom}(a) \to H$ is linear, then for brevity we call the couple (a, j) a **sectorial form in *H***. In the special case when a is symmetric and accretive, we call (a, j) an **accretive symmetric form in *H***. We say that (a, j) is **densely defined** if j has dense range.

12.1 Remark. In our applications in the last two sections of the present chapter we will consider accretive symmetric forms in *real* Hilbert spaces. In this situation the key inequality (12.2) holds with $c = 0$. A careful inspection of our treatment shows that the results of the chapter depend on (12.2) rather than on (12.1) and that they remain valid in real Hilbert spaces if one replaces 'sectorial form' by 'accretive symmetric form', 'sectorial operator' by 'accretive symmetric operator' and 'm-sectorial operator' by 'accretive self-adjoint operator'. $\triangle$

12.2 Proposition. *Let (a, j) be a densely defined sectorial form in H. Then*

$$A_0 := \big\{(x,y) \in H \times H\,;\ \exists u \in \operatorname{dom}(a)\colon j(u) = x,\ a(u,v) = (y \,|\, j(v))_H\ (v \in \operatorname{dom}(a))\big\} \tag{12.3}$$

defines a sectorial operator in H.

Proof. (i) Let $(0, y) \in A_0$. We show that $y = 0$; then it follows that A_0 is an operator. By definition there exists $u \in \operatorname{dom}(a)$ such that $j(u) = 0$ and $a(u, v) = (y \,|\, j(v))_H$ for all $v \in \operatorname{dom}(a)$. In particular, $a(u) = 0$. By (12.2) this implies that $a(u, v) = 0$ for all $v \in \operatorname{dom}(a)$. Consequently $(y \,|\, j(v))_H = 0$ for all $v \in \operatorname{dom}(a)$. Since j has dense range it follows that $y = 0$.

(ii) Let $x \in \operatorname{dom}(A_0)$. Then there exists $u \in \operatorname{dom}(a)$ such that $j(u) = x$ and $a(u, v) = (A_0x \,|\, j(v))_H$ for all $v \in \operatorname{dom}(a)$. In particular, $(A_0x \,|\, x) = a(u)$. Thus A_0 is sectorial. $\square$

In general, the operator A_0 is not m-sectorial (see Exercise 12.2), but below we will construct an m-sectorial extension. The idea is to use a completion procedure that enables us to apply the generation theorems of Chapter 5. The resulting operator will be discussed in Theorem 12.4.

Let (a, j) be a densely defined sectorial form in H. Then

$$(u \,|\, v)_{a,j} := (\operatorname{Re} a)(u, v) + (j(u) \,|\, j(v))_H \tag{12.4}$$

defines a semi-inner product, i.e. a symmetric, accretive sesquilinear form on $\operatorname{dom}(a)$. Thus

$$\|u\|_{a,j} := \sqrt{(u \,|\, u)_{a,j}} = \big(\operatorname{Re} a(u) + \|j(u)\|_H^2\big)^{1/2}$$

defines a semi-norm on $\operatorname{dom}(a)$. (We will suppress 'j' in the index if $\operatorname{dom}(a) \subseteq H$ and j is the natural embedding.)

12.3 Remark (Completion of a semi-inner product space). Let E be a vector space over $\mathbb{K}$ and $(\cdot \,|\, \cdot)$ a **semi-inner product** on E (i.e. the form $(\cdot \,|\, \cdot)$ is symmetric and accretive), with associated semi-norm $\|\cdot\|$.

(a) A **completion** of E is a pair $(\widetilde{E}, q)$ consisting of a Hilbert space $\widetilde{E}$ and a linear mapping $q\colon E \to \widetilde{E}$ that

(i) is **isometric**, i.e. $(q(u) \,|\, q(v))_{\widetilde{E}} = (u \,|\, v)$ for all $u, v \in E$, and

(ii) has dense range.

For the existence of a completion, note that $F := \{u \in E;\ \|u\| = 0\}$ is a subspace of E, and let $q\colon E \to E/F =: G$ denote the quotient map. Since $|(u \,|\, v)| \leqslant \|u\|\|v\|$, by Proposition 5.2, $\|u\| = 0$ implies $(u \,|\, v) = 0$ for all $v \in E$. As a consequence, putting

$(q(u) \mid q(v))_G = (u \mid v)$ one obtains a well-defined scalar product on G that makes q isometric. Then properties (i) and (ii) are satisfied with the completion $\widetilde{E}$ of the pre-Hilbert space G.

It follows from part (c) below that the completion $(\widetilde{E}, q)$ is unique up to unitary equivalence. Sometimes we drop the letter q and speak of $\widetilde{E}$ as the completion of E.

(b) The mapping q is injective if and only if $(\cdot \mid \cdot)$ is positive definite, and q is surjective if and only if $(E, \|\cdot\|)$ is complete. If $(E, \|\cdot\|)$ is complete, but $(\cdot \mid \cdot)$ is not positive definite, then the completion $(\widetilde{E}, q) = (E/F, q)$ is different from E. (Recall that a semi-normed space is called complete if every Cauchy sequence is convergent.)

(c) If H is a Hilbert space and $j \colon E \to H$ is a bounded linear operator, then there exists a unique bounded linear operator $\tilde{\jmath} \colon \widetilde{E} \to H$ such that $\tilde{\jmath} \circ q = j$. Indeed, there exists a unique operator $\tilde{\jmath}_0 \colon E/F \to H$ such that $\tilde{\jmath}_0 \circ q = j$, and this (bounded) operator has the asserted extension $\tilde{\jmath}$. Similarly, if $a \colon E \times E \to \mathbb{K}$ is a bounded form (i.e. $|a(u,v)| \leqslant M\|u\|\|v\|$ for all $u, v \in E$), then there exists a uniquely determined bounded form $\tilde{a} \colon \widetilde{E} \times \widetilde{E} \to \mathbb{K}$ such that $\tilde{a}(q(u), q(v)) = a(u,v)$ for all $u, v \in E$. △

We now construct the m-sectorial operator associated with a sectorial form. Let H be a complex Hilbert space, and let (a, j) be a densely defined sectorial form in H. Then (12.2) says that a is a bounded form on the semi-inner product space $(\operatorname{dom}(a), (\cdot \mid \cdot)_{a,j})$. Let (V, q) denote the completion of $(\operatorname{dom}(a), (\cdot \mid \cdot)_{a,j})$, and let $\tilde{\jmath} \colon V \to H$ and $\tilde{a} \colon V \times V \to \mathbb{C}$ be as explained in Remark 12.3(c). Observe that $\tilde{a}$ is sectorial and that

$$\operatorname{Re} \tilde{a}(q(u)) + \|\tilde{\jmath}(q(u))\|_H^2 = \operatorname{Re} a(u) + \|j(u)\|_H^2 = \|u\|_{a,j}^2 = \|q(u)\|_V^2$$

for all $u \in \operatorname{dom}(a)$. As $\operatorname{ran}(q)$ is dense in V, and $\tilde{a}$ as well as q are continuous, we conclude that $\operatorname{Re} \tilde{a}(v) + \|\tilde{\jmath}(v)\|_H^2 = \|v\|_V^2$ for all $v \in V$. This equality shows that $\tilde{a}$ is $\tilde{\jmath}$-coercive. Let A be the operator associated with $(\tilde{a}, \tilde{\jmath})$, as described in Section 5.3. We call A the **operator associated with** (a, j), and we write $A \sim (a, j)$. We recall that

$$A = \bigl\{(x,y) \in H \times H\,;\ \exists u \in V \colon \tilde{\jmath}(u) = x,\ \tilde{a}(u,v) = (y \mid \tilde{\jmath}(v))_H\ (v \in V)\bigr\}; \tag{12.5}$$

for another description of A we refer to Theorem 12.11.

We summarise the properties of the operator A in the next theorem. A sectorial form (a, j) in a complex Hilbert space is called **closed** if the semi-inner product space $(\operatorname{dom}(a), (\cdot \mid \cdot)_{a,j})$ is complete.

12.4 Theorem. *Let (a, j) be a densely defined sectorial form in a complex Hilbert space H. Then the operator $A \sim (a, j)$ in* (12.5) *is an m-sectorial extension of the operator A_0 defined in* (12.3). *If the form a is symmetric, then A is self-adjoint. If (a, j) is closed, then $A = A_0$.*

Proof. The operator A is m-sectorial by Corollary 5.11. If $(x, y) \in A_0$, then there exists $u \in \operatorname{dom}(a)$ such that $j(u) = x$ and $a(u,v) = (y \mid j(v))$ for all $v \in \operatorname{dom}(a)$. It follows that $\tilde{\jmath}(q(u)) = x$ and

$$\tilde{a}(q(u), \tilde{v}) = (y \mid \tilde{\jmath}(\tilde{v})) \tag{12.6}$$

for all $\tilde{v} \in \operatorname{ran}(q)$. The denseness of $\operatorname{ran}(q)$ in V and the continuity of $\tilde{a}$ and $\tilde{\jmath}$ imply that (12.6) carries over to all $\tilde{v} \in V$, and this shows that $(x, y) \in A$.

Again by denseness and continuity, the symmetry of a carries over to $\tilde{a}$, and then the self-adjointness of A follows from Theorem 6.11.

If (a, j) is closed, then q is surjective. Let $(x, y) \in A$, and let $\tilde{u} \in V$ be such that $\tilde{\jmath}(\tilde{u}) = x$ and $\tilde{a}(\tilde{u}, \tilde{v}) = (y \,|\, \tilde{\jmath}(\tilde{v}))$ for all $\tilde{v} \in V$. There exists $u \in \operatorname{dom}(a)$ such that $\tilde{u} = q(u)$. Then

$$a(u, v) = \tilde{a}(\tilde{u}, q(v)) = \big(y \,\big|\, \tilde{\jmath}(q(v))\big) = (y \,|\, j(v))$$

for all $v \in \operatorname{dom}(a)$, and since $j(u) = \tilde{\jmath}(q(u)) = x$, it follows that $(x, y) \in A_0$. □

The assertions of Theorem 12.4 remain valid – with appropriate changes – if (a, j) is merely a **quasi-sectorial form** in H, i.e., if instead of a being sectorial one assumes that there exists $\omega \in \mathbb{R}$ such that the shifted form

$$a_\omega(u, v) := a(u, v) + \omega(j(u) \,|\, j(v))_H \qquad (u, v \in \operatorname{dom}(a))$$

is sectorial. The number $-\omega$ is called a **vertex** of (a, j), and (a, j) is **closed** if (a_ω, j) is closed. As before, the operator A associated with (a, j) is given by (12.5), where now V is the completion of $(\operatorname{dom}(a), (\cdot \,|\, \cdot)_{a_\omega, j})$; note that a is a bounded form on this semi-inner product space. The 'extension' $\tilde{a}$ of a from Remark 12.3(c) is easily seen to satisfy $\operatorname{Re} \tilde{a}(v) + (\omega + 1)\|\tilde{\jmath}(v)\|_H^2 = \|v\|_V^2$ for all $v \in V$, and hence $\tilde{a}$ is $\tilde{\jmath}$-coercive. This implies that A is a quasi-m-sectorial operator. Observe that A does not depend on the choice of the vertex $-\omega$ because different values of ω result in equivalent semi-norms $\|\cdot\|_{a_\omega, j}$ on $\operatorname{dom}(a)$.

12.5 Remark. In the context described above, let A_ω denote the (sectorial) operator associated with the form (a_ω, j). One easily checks that $\widetilde{a_\omega} = \tilde{a}_\omega$, so it follows from Remark 5.10 that $A_\omega = A + \omega I$.

In view of Remark 12.1, the same observation applies in the real case if (a, j) is a **quasi-accretive symmetric form** in H, i.e. if (a_ω, j) is an accretive symmetric form in H for some $\omega \in \mathbb{R}$. In the literature, an embedded symmetric form is usually called 'bounded from below' if it is quasi-accretive. △

We introduce one more notion that will be used below. For a quasi-sectorial form (a, j) with vertex $-\omega$, a subspace D of $\operatorname{dom}(a)$ is called a **core for** (a, j) if D is dense in $(\operatorname{dom}(a), \|\cdot\|_{a_\omega, j})$. This notion also makes sense for quasi-accretive symmetric forms in real Hilbert spaces.

12.6 Example. If (a, j) is a quasi-sectorial form and D is a core for a, then it is immediate from the definitions that the operator associated with $(a|_{D \times D}, j|_D)$ is the same as the one associated with (a, j).

For instance, if Ω is an open subset of $\mathbb{R}^n$, then the operator associated with the restriction of the classical Dirchlet form to $C_c^\infty(\Omega)$ is the negative Dirichlet Laplacian. △

12.7 Remark. In Theorems 5.6 and 5.8, Corollary 5.11 and Theorem 12.4 we have formulated generation results for m-accretive operators under different hypotheses. Looking more closely, one can see that Theorem 5.6 is the basic result, which is then applied in different situations. In this context we also mention the generation result Theorem 8.11, which is proved by similar arguments as Theorem 5.6, with the Fredholm alternative as an additional argument. △

12.2 The Friedrichs extension

The main result of this section is that every densely defined sectorial operator has an m-sectorial extension. This extension is associated with a suitably constructed closed form, and an important "by-product" of the construction will be that every m-sectorial operator is associated with a unique embedded closed sectorial form.

Let H be a complex Hilbert space. We consider the situation in which a is a sectorial form whose domain is a subspace of H and $j\colon \mathrm{dom}(a) \hookrightarrow H$ is the embedding. In this case we drop the letter j in our notation and speak of a form a in H. As before we call a an **embedded form**.

If $\mathrm{dom}(a)$ is dense and a is closed, then by Theorem 12.4 the associated m-sectorial operator is given by

$$A = \big\{(u,y) \in \mathrm{dom}(a) \times H\,;\, a(u,v) = (y \,|\, v)_H \ (v \in \mathrm{dom}(a))\big\}. \tag{12.7}$$

Here, a being closed means that $\mathrm{dom}(a)$ is complete for the norm $\|\cdot\|_a$ given by

$$\|u\|_a^2 = \mathrm{Re}\, a(u) + \|u\|_H^2\,.$$

12.8 Theorem (Friedrichs extension). *Let B be a densely defined sectorial operator in H. Then there exists a unique densely defined embedded closed sectorial form a in H such that* $\mathrm{dom}(B) \subseteq \mathrm{dom}(a)$, $\mathrm{dom}(B)$ *is a core for a, and*

$$a(u,v) = (Bu \,|\, v)_H \qquad \big(u \in \mathrm{dom}(B),\ v \in \mathrm{dom}(a)\big).$$

Let $A \sim a$. Then $B \subseteq A$.

The operator A is called the **Friedrichs extension** of B. The proof given below also works for accretive symmetric operators in real Hilbert spaces. In that case one obtains an accretive symmetric form a, and the Friedrichs extension A is an accretive self-adjoint operator; cf. Remark 12.1.

Proof of Theorem 12.8. The uniqueness of a follows from the following general fact: if a_1 and a_2 are two embedded closed sectorial forms in H with a common core $D \subseteq \mathrm{dom}(a_1) \cap \mathrm{dom}(a_2)$, then $a_1|_{D\times D} = a_2|_{D\times D}$ implies $a_1 = a_2$.

For the existence define $b\colon \mathrm{dom}(B) \times \mathrm{dom}(B) \to \mathbb{C}$ by $b(u,v) := (Bu \,|\, v)_H$. Then b is densely defined and sectorial. We use the embedding $j\colon \mathrm{dom}(b) \hookrightarrow H$ and the scalar product $(\cdot \,|\, \cdot)_b$ analogous to (12.4) on $\mathrm{dom}(b) = \mathrm{dom}(B)$. Let (V,q) be the completion of $(\mathrm{dom}(b), (\cdot \,|\, \cdot)_b)$; then q is injective (cf. Remark 12.3(b)). We consider $\mathrm{dom}(b)$ as a subset of V and suppress the letter q in our notation.

We now show that $\tilde{\jmath} \in \mathcal{L}(V,H)$ – from Remark 12.3(c) – is injective. Let $u \in V$, $\tilde{\jmath}(u) = 0$. There exists a sequence (u_n) in $\mathrm{dom}(b)$ such that $u_n \to u$ in V. Then $u_n = \tilde{\jmath}(u_n) \to \tilde{\jmath}(u) = 0$ in H. Using (12.2) one obtains

$$\begin{aligned}\mathrm{Re}\, b(u_n) &= \mathrm{Re}\, b(u_n, u_n - u_k) + \mathrm{Re}\, b(u_n, u_k)\\ &\leqslant (1+c)\|u_n\|_b \|u_n - u_k\|_b + |(Bu_n \,|\, u_k)_H| \qquad (k,n \in \mathbb{N}).\end{aligned}$$

Letting $k \to \infty$ we deduce that $\operatorname{Re} b(u_n) \leqslant (1+c)\|u_n\|_V \|u_n - u\|_V$ for all $n \in \mathbb{N}$. It follows that $\|u_n\|_V^2 = \operatorname{Re} b(u_n) + \|u_n\|_H^2 \to 0$ as $n \to \infty$, $\|u\|_V = \lim_{n\to\infty} \|u_n\|_V = 0$.

Because $\tilde{\jmath}$ is injective we can consider V as a subspace of H and $a := \tilde{b}$ – from Remark 12.3(c) – as an embedded form (with $\operatorname{dom}(a) = V$). Then a is a closed sectorial form in H with the required properties, and it follows from (12.7) that the associated operator A is an extension of B. □

12.9 Corollary. *Let A be an m-sectorial operator in H. Then there exists a unique densely defined embedded closed sectorial form a in H such that A is associated with a.*

Proof. The existence follows from Theorem 12.8 since m-sectorial operators do not have proper m-sectorial extensions.

In order to prove the uniqueness, let a be a form with the asserted properties. Then $\operatorname{dom}(A) \subseteq \operatorname{dom}(a)$ and $a(u,v) = (Au \mid v)$ for all $u, v \in \operatorname{dom}(A)$. Moreover $\operatorname{dom}(A)$ is dense in $(\operatorname{dom}(a), \|\cdot\|_a)$ by Lemma 9.21(b). Therefore the uniqueness follows from the uniqueness in Theorem 12.8. (We also refer to Exercise 5.3(c) for the uniqueness assertion.) □

Just like Theorem 12.8, Corollary 12.9 also holds for accretive symmetric operators in real Hilbert spaces.

12.3 Sectorial versus coercive

We start by giving a short summary of the different types of forms that we have introduced so far. In Section 5.3 we discussed bounded j-coercive forms on V and their associated m-accretive operators in H (which are m-sectorial if $\mathbb{K} = \mathbb{C}$). Here, both V and H are Hilbert spaces. In the special case when $V \subseteq H$ and $j\colon V \hookrightarrow H$ is the embedding, the form was called quasi-coercive. In Section 8.3 we investigated the related concept of essentially coercive forms. We recall from Proposition 8.14 that it is always possible to perform a reduction to the embedded case.

In Section 12.1 we have introduced the notion of a (quasi-)sectorial form (a, j) in H. The new aspect of this notion is that the domain of a is not supposed to be a Hilbert space; the corresponding semi-inner product $(\cdot \mid \cdot)_{a,j}$ need not be positive definite, nor does $(\operatorname{dom}(a), \|\cdot\|_{a,j})$ have to be complete. The construction of the associated operator in H by completion is basically a reduction to the case of j-coercive forms. The point in this approach is that the norm $\|\cdot\|_V$, needed for the generation theorems of Section 5, is defined intrinsically by the form itself and the mapping j. In the special case of closed forms the associated operator is given by the same formula as for j-coercive forms, by the last assertion of Theorem 12.4.

In the following remark we summarise the relations between the notions '(closed) quasi-sectorial form' and 'bounded j-coercive form', for $\mathbb{K} = \mathbb{C}$.

12.10 Remarks. (a) Let V be a Hilbert space and $j \in \mathcal{L}(V, H)$. Let $a\colon V \times V \to \mathbb{C}$ be a bounded j-coercive form, say

$$|a(u,v)| \leqslant M\|u\|_V \|v\|_V, \quad \operatorname{Re} a(u) + \omega\|j(u)\|_H^2 \geqslant \alpha\|u\|_V^2 \qquad (u, v \in \operatorname{dom}(a)).$$

Then a_ω is sectorial (see Theorem 5.8) and the norm $\|\cdot\|_{a_\omega,j}$ is equivalent to the given norm $\|\cdot\|_V$ on V. It follows that (a,j) is a closed quasi-sectorial form in H.

(b) Conversely, let (a,j) be a closed quasi-sectorial form in H, with vertex $-\omega$, and assume that $(\cdot\,|\,\cdot)_{a_\omega,j}$ is positive definite. Then $V := \mathrm{dom}(a)$ is a Hilbert space with scalar product $(\cdot\,|\,\cdot)_{a_\omega,j}$, and a is a bounded j-coercive form on V.

(c) Let (a,j) be a quasi-sectorial form in H, with vertex $-\omega$. Then the completion (V,q) of $(\mathrm{dom}(a),(\cdot\,|\,\cdot)_{a_\omega,j})$ is a Hilbert space for the norm given by $\|u\|_V := \|u\|_{\tilde a_\omega,\tilde\jmath} = (\mathrm{Re}\,\tilde a(u) + (\omega+1)\|\tilde\jmath(u)\|_H^2)^{1/2}$. The form $\tilde a$ is bounded for this norm and $\tilde\jmath$-coercive.

If additionally the form (a,j) is closed, then the construction simplifies, because the completion V of $\mathrm{dom}(a)$ is isomorphic to $\mathrm{dom}(a)/\{u\in\mathrm{dom}(a)\,;\,\|u\|_{a_\omega,j}=0\}$, $q\colon \mathrm{dom}(a)\to V$ is the quotient map, and $\tilde a$ is given by

$$\tilde a(q(u),q(v)) = a(u,v) \qquad (u,v\in\mathrm{dom}(a)). \qquad \triangle$$

12.4 More on the non-closed case

In Section 12.1 we have defined the operator associated with a densely defined sectorial form that is not necessarily closed; see (12.5). In (12.8) we give a more direct description of this operator that avoids the reference to completion; it will be applied to concrete examples in Sections 12.5 and 12.6. It is surprising that, out of the purely algebraic condition of sectoriality, by the approximation formula (12.8) we obtain the generator of a holomorphic C_0-semigroup.

12.11 Theorem (Approximation formula). *Let H be a complex Hilbert space, and let (a,j) be a densely defined sectorial form in H. Then for the operator $A\sim(a,j)$ one has*

$$A = \big\{(x,y)\in H\times H\,;\ \textit{there exists } (u_k) \textit{ in } \mathrm{dom}(a) \textit{ such that} \\ \text{(i) } j(u_k)\to x \textit{ as } k\to\infty, \\ \text{(ii) } \lim_{k,\ell\to\infty}\mathrm{Re}\,a(u_k-u_\ell)=0, \\ \text{(iii) } a(u_k,v)\to (y\,|\,j(v))_H\ (v\in\mathrm{dom}(a))\big\}. \tag{12.8}$$

In this description, property (ii) *can be replaced by*
(ii′) $\sup_{k\in\mathbb{N}}\mathrm{Re}\,a(u_k)<\infty$.

Proof. We recall the notation used in the definition (12.5) of A: (V,q) is the completion of $(\mathrm{dom}(a),(\cdot\,|\,\cdot)_{a,j})$, and $\tilde\jmath\colon V\to H$ and $\tilde a\colon V\times V\to\mathbb{C}$ are as explained in Remark 12.3(c).

Let $(x,y)\in A$. Then there exists $w\in V$ such that $\tilde\jmath(w)=x$ and $\tilde a(w,v)=(y\,|\,\tilde\jmath(v))_H$ for all $v\in V$. Since q has dense range there exists a sequence (u_k) in $\mathrm{dom}(a)$ such that $q(u_k)\to w$ in V. By the continuity of $\tilde\jmath$ it follows that $j(u_k)=\tilde\jmath(q(u_k))\to\tilde\jmath(w)=x$. Since q is isometric, we obtain

$$\mathrm{Re}\,a(u_k-u_\ell)+\|j(u_k)-j(u_\ell)\|_H^2 = \|u_k-u_\ell\|_a^2 = \|q(u_k)-q(u_\ell)\|_V^2\to 0$$

as $k,\ell\to\infty$. Finally, for $v\in\mathrm{dom}(a)$ we find that

$$a(u_k,v)=\tilde a(q(u_k),q(v))\to\tilde a(w,q(v))=(y\,|\,\tilde\jmath(q(v)))_H=(y\,|\,j(v))_H\,.$$

Thus we have found a sequence (u_k) satisfying (i), (ii) and (iii) – and hence also the weaker property (ii′).

Conversely, let $(x, y) \in H \times H$ be such that there exists a sequence (u_k) in $\mathrm{dom}(a)$ satisfying (i), (ii′) and (iii). It follows from (i) together with (ii′) that $\sup_{k\in\mathbb{N}} \|u_k\|_{a,j} < \infty$. Thus, taking a subsequence if necessary, we can assume that $(q(u_k))$ converges weakly to some $w \in V$. Hence $j(u_k) = \tilde{\jmath}(q(u_k)) \to \tilde{\jmath}(w)$ weakly in H, and so $\tilde{\jmath}(w) = x$ by (i). Property (iii) implies

$$\tilde{a}(w, q(v)) = \lim_{k\to\infty} \tilde{a}(q(u_k), q(v)) = \lim_{k\to\infty} a(u_k, v) = (y \,|\, j(v))_H = \big(y \,\big|\, \tilde{\jmath}(q(v))\big)_H$$

for all $v \in \mathrm{dom}(a)$. Since q has dense range and $\tilde{a}, \tilde{\jmath}$ are continuous, we conclude that $\tilde{a}(w, \tilde{v}) = (y \,|\, \tilde{\jmath}(\tilde{v}))_H$ for all $\tilde{v} \in V$, and thus $(x, y) \in A$. □

With the same proof one obtains the following real version of Theorem 12.11; for the second assertion we refer to Remark 12.1 and Theorem 12.4.

12.12 Theorem. *Let (a, j) be a densely defined accretive symmetric form in a real Hilbert space H. Then the operator $A \sim (a, j)$ is given by* (12.8) *(where* (ii) *or* (ii′) *can be used), and A is self-adjoint and accretive.*

12.13 Remarks. Here we assume that a is an embedded densely defined sectorial form in H. Let $j\colon \mathrm{dom}(a) \hookrightarrow H$ denote the embedding.

(a) Let $(\tilde{a}, \tilde{\jmath})$ be the 'extension' of (a, j) as described in Remark 12.3(c). We call the form a **closable** if $\tilde{\jmath}$ is injective. Then we may identify $\mathrm{dom}(\tilde{a})$ with a subspace of H and consider $\tilde{a}$ as an embedded form.

An example of a closable form is the form b from the proof of Theorem 12.8 which leads to the Friedrichs extension.

(b) We have seen that we may associate an m-sectorial operator A with the form a, no matter whether a is closable or not. This means that in our context we can simply forget about the notion of closability.

From Corollary 12.9 we know that there exists a unique embedded closed sectorial form $\bar{a}$ in H that is associated with A. We refer to part (c) for a (somewhat involved) description of $\bar{a}$ in terms of a; we will not use this description in our subsequent results.

(c) Using the results of Section 8.3, we can describe $\bar{a}$ as follows; in particular it will turn out that $\mathrm{dom}(a)$ is a core for $\bar{a}$.

Let (V, q) be the completion of $(\mathrm{dom}(a), (\cdot \,|\, \cdot)_a)$; then q is injective by Remark 12.3(b). The application of Proposition 8.16 implies that $V = V_{\tilde{\jmath}}(\tilde{a}) \oplus \ker(\tilde{\jmath})$ is a topological direct sum, where

$$V_{\tilde{\jmath}}(\tilde{a}) = \big\{u \in V\,;\ \tilde{a}(u, v) = 0 \ (v \in \ker(\tilde{\jmath}))\big\}.$$

If we put $\check{\tilde{a}} := \tilde{a}|_{V_{\tilde{\jmath}}(\tilde{a})\times V_{\tilde{\jmath}}(\tilde{a})}$ and $\check{\tilde{\jmath}} := \tilde{\jmath}|_{V_{\tilde{\jmath}}(\tilde{a})}$, then the operator A associated with (a, j) is also associated with the form $(\check{\tilde{a}}, \check{\tilde{\jmath}})$, by Proposition 8.14, and $\check{\tilde{\jmath}}$ is injective. Therefore the 'transported' form $\bar{a}$,

$$\bar{a}(\check{\tilde{\jmath}}(u), \check{\tilde{\jmath}}(v)) := \tilde{a}(u, v) \qquad (u, v \in V_{\tilde{\jmath}}(\tilde{a}))$$

is the unique embedded closed sectorial form associated with A.

Let P, Q denote the projections corresponding to the direct sum $V = V_{\tilde{\jmath}}(\tilde{a}) \oplus \ker(\tilde{\jmath})$. For $u \in \operatorname{dom}(a)$ one has

$$u = j(u) = \tilde{\jmath}(q(u)) = \tilde{\jmath}((P+Q)q(u)) = \tilde{\jmath}(Pq(u)),$$

and this shows that $\operatorname{dom}(a) = \tilde{\jmath}(Pq(\operatorname{dom}(a)) \subseteq \tilde{\jmath}(V_{\tilde{\jmath}}(\tilde{a})) = \operatorname{dom}(\bar{a})$ and that

$$\bar{a}(u,v) = \tilde{a}(Pq(u), Pq(v)) \qquad (u, v \in \operatorname{dom}(a)).$$

Moreover, the denseness of $q(\operatorname{dom}(a))$ in V implies that $Pq(\operatorname{dom}(a))$ is dense in $V_{\tilde{\jmath}}(\tilde{a})$, i.e. $Pq(\operatorname{dom}(a))$ is a core for $\check{\tilde{a}}$; hence $\operatorname{dom}(a)$ is a core for $\bar{a}$.

If additionally the form a is symmetric, then it turns out that $\bar{a}|_{\operatorname{dom}(a)\times\operatorname{dom}(a)}$ is the 'regular part' of a described in [Sim78; Section 2].

(d) We add a comment on the terminology. The term 'closed form' (in the situation of embedded forms) was forged by Kato – see e.g. [Kat80; Chap. VI, § 1.3] – in a vague analogy to 'closed operator'. However, whereas closed operators *are* closed in the product topology, for forms there is no visible closed set. Note that the definition of 'closable' – see [Kat80; Chap. VI, § 1.4] – requires the associated closed form to be an embedded form. △

12.5 The Robin Laplacian for rough domains

Throughout this section the scalar field will be $\mathbb{K} = \mathbb{R}$. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. On $\partial\Omega$ we consider the $(n-1)$-dimensional Hausdorff measure $\mathcal{H}_{n-1}$; we refer to Section D.3 for the definition of Hausdorff measures and some of their properties. Here we mention that the n-dimensional Hausdorff measure on $\mathbb{R}^n$ coincides with the n-dimensional Lebesgue measure. Moreover, if Ω has C^1-boundary, then $\mathcal{H}_{n-1}$ coincides with the surface measure σ. For both of these properties we refer to Appendix E.

We denote by $L_2(\partial\Omega)$ the space of L_2-functions with respect to the Hausdorff measure $\mathcal{H}_{n-1}$; as explained in the previous paragraph, the notation $L_2(\partial\Omega)$ is consistent with the notation from Chapter 7. In the present section we assume that $\mathcal{H}_{n-1}(\partial\Omega) < \infty$.

We define the **trace** tr as the closure of the operator $u \mapsto u|_{\partial\Omega} \colon C(\overline{\Omega}) \cap H^1(\Omega) \to L_2(\partial\Omega)$ in $H^1(\Omega) \times L_2(\partial\Omega)$. For $u \in H^1(\Omega)$ we write $\operatorname{tr} u := \{\varphi \in L_2(\partial\Omega);\ (u,\varphi) \in \operatorname{tr}\}$, which means that

$$\operatorname{tr} u = \big\{\varphi \in L_2(\partial\Omega);\ \text{there exists } (u_k) \text{ in } C(\overline{\Omega}) \cap H^1(\Omega) \text{ such that} \\ u_k \to u \text{ in } H^1(\Omega),\ u_k|_{\partial\Omega} \to \varphi \text{ in } L_2(\partial\Omega)\big\}.$$

In general, the set $\operatorname{tr} u$ may consist of more than one element (and may also be empty); see Exercise 12.7. But if Ω has Lipschitz boundary, then $\operatorname{tr} u$ is a singleton for each $u \in H^1(\Omega)$; in fact tr is a bounded operator. For the case of C^1-boundary this property has already been proved in Theorem 7.11, for the case of Lipschitz boundary we refer to [Alt16; Theorem A8.6].

Let $u \in H^1(\Omega)$ be such that $\Delta u \in L_2(\Omega)$. We say that $\partial_\nu u \in L_2(\partial\Omega)$ if there exists $h \in L_2(\partial\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_\Omega (\Delta u) v \,\mathrm{d}x = \int_{\partial\Omega} hv \,\mathrm{d}\mathcal{H}_{n-1} \tag{12.9}$$

for all $v \in C(\overline{\Omega}) \cap H^1(\Omega)$; in this case we define the **weak normal derivative** $\partial_\nu u := h$. The uniqueness of $\partial_\nu u$ is obtained from the denseness of $\{v|_{\partial\Omega};\, v \in C^1(\overline{\Omega})\}$ in $L_2(\partial\Omega)$ as in Section 7.3.

12.14 Remarks. (a) The definition of the weak normal derivative given above is an extension of the definition in Section 7.3. Note that in the present situation there may exist points of $\partial\Omega$ for which there is no outer normal ν.

(b) The reader should be warned that the definition of $\partial_\nu u$ given above is not consistent with our definition of '$\partial_\nu u = 0$' (in quotes!) in Section 7.3. In fact, '$\partial_\nu u = 0$' means that (12.9) holds with $h = 0$ for all $v \in H^1(\Omega)$ (and not just for $v \in C(\overline{\Omega}) \cap H^1(\Omega)$); see (7.7). Thus, '$\partial_\nu u = 0$' implies $\partial_\nu u = 0$. The converse is true if Ω is such that $C(\overline{\Omega}) \cap H^1(\Omega)$ is dense in $H^1(\Omega)$. We refer to Exercise 12.5 for an example showing that in general $\partial_\nu u = 0$ does not imply '$\partial_\nu u = 0$'. △

Let $\beta \in L_\infty(\partial\Omega)$, $\inf \beta > 0$. Under our present general hypotheses, the **Robin Laplacian** is defined by

$$\Delta_\beta := \big\{(u,f) \in H^1(\Omega) \times L_2(\Omega);\ \Delta u = f,\ \exists \varphi \in \operatorname{tr} u\colon \partial_\nu u = -\beta\varphi\big\}.$$

12.15 Theorem. *The operator $-\Delta_\beta$ is an accretive self-adjoint operator in $L_2(\Omega)$.*

Proof. Put $\operatorname{dom}(a) := C(\overline{\Omega}) \cap H^1(\Omega)$ and

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta u v \,\mathrm{d}\mathcal{H}_{n-1}\,.$$

Then a is densely defined in $L_2(\Omega)$, symmetric and accretive. Let A be the operator associated with (a,j); see Section 12.1. Then A is self-adjoint and accretive by Theorem 12.12, and we use (12.8) to show that $A = -\Delta_\beta$.

'$A \subseteq -\Delta_\beta$'. Let $(u,f) \in A$. Then there exists a sequence (u_k) in $\operatorname{dom}(a)$ such that

(i) $u_k \to u$ in $L_2(\Omega)$,

(ii) $\lim_{k,\ell\to\infty} a(u_k - u_\ell) = 0$,

(iii) $a(u_k, v) = \int_\Omega \nabla u_k \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta u_k v \,\mathrm{d}\mathcal{H}_{n-1} \to \int_\Omega f v \,\mathrm{d}x$ for all $v \in \operatorname{dom}(a)$.

One concludes from (i) and (ii) that $u \in H^1(\Omega)$ and $u_k \to u$ in $H^1(\Omega)$ and – using the hypothesis $\inf\beta > 0$ – that $\varphi := \lim_{k\to\infty} u_k|_{\partial\Omega}$ exists in $L_2(\partial\Omega)$. Then $\varphi \in \operatorname{tr} u$. Now (iii) implies that

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta\varphi v \,\mathrm{d}\mathcal{H}_{n-1} = \int_\Omega f v \,\mathrm{d}x$$

for all $v \in \operatorname{dom}(a)$. Taking $v \in C_c^\infty(\Omega)$ we obtain $-\Delta u = f$. Thus

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_\Omega (\Delta u) v \,\mathrm{d}x = -\int_{\partial\Omega} \beta\varphi v \,\mathrm{d}\mathcal{H}_{n-1}$$

for all $v \in \operatorname{dom}(a)$. This means that $\partial_\nu u = -\beta\varphi$, by our definition (12.9).

'$-\Delta_\beta \subseteq A$'. Let $(u,f) \in -\Delta_\beta$. Then $u \in H^1(\Omega)$, $f = -\Delta u$, and there exists $\varphi \in \operatorname{tr} u$ such that $\partial_\nu u = -\beta\varphi$. By the definition of $\operatorname{tr} u$, there exists a sequence (u_k) in $\operatorname{dom}(a)$

such that $u_k \to u$ in $H^1(\Omega)$ and $u_k|_{\partial\Omega} \to \varphi$ in $L_2(\partial\Omega)$ as $k \to \infty$. Thus (i) and (ii) hold, and

$$a(u_k, v) = \int_\Omega \nabla u_k \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta u_k v \,\mathrm{d}\mathcal{H}_{n-1} \to \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta\varphi v \,\mathrm{d}\mathcal{H}_{n-1} = \int_\Omega f v \,\mathrm{d}x$$

for all $v \in \operatorname{dom}(a)$ since $-\Delta u = f$ and $\partial_\nu u = -\beta\varphi$, i.e. (iii) holds as well. It follows that $(u, f) \in A$, by Theorem 12.12. □

Actually it is not necessary to call on Theorem 12.11 for the proof of Theorem 12.15. We add a proof that merely relies on the generation results of Section 5.3.

Second proof of Theorem 12.15. Let V be the closure of $\{(u, u|_{\partial\Omega})\,;\, u \in C(\overline{\Omega}) \cap H^1(\Omega)\}$ in $H^1(\Omega) \oplus L_2(\partial\Omega)$. (This just means that V is the linear relation tr introduced above.) Put

$$a\big((u,\varphi),(v,\psi)\big) := \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta\varphi\psi \,\mathrm{d}\mathcal{H}_{n-1} \qquad \big((u,\varphi),(v,\psi) \in V\big),$$

and define $j\colon V \to L_2(\Omega)$ by $j(u,\varphi) := u$. Then a is a bounded j-coercive form on V because $\inf \beta > 0$, and j has dense range. Since a is symmetric and accretive, the associated operator A is self-adjoint and accretive by the results of Chapters 5 and 6.

Let $(u, f) \in A$. Then there exists $\varphi \in L_2(\partial\Omega)$ such that $(u,\varphi) \in V$ (so $\varphi \in \operatorname{tr} u$) and

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \beta\varphi\psi \,\mathrm{d}\mathcal{H}_{n-1} = \int_\Omega f v \,\mathrm{d}x \tag{12.10}$$

for all $(v,\psi) \in V$. Employing (12.10) for all $(v, 0)$ with $v \in C_c^\infty(\Omega)$ we obtain $\Delta u = -f$. Then using (12.10) once more for all $(v, v|_{\partial\Omega})$ with $v \in C(\overline{\Omega}) \cap H^1(\Omega)$, we conclude that $\partial_\nu u = -\beta\varphi \in L_2(\partial\Omega)$.

Conversely, let $(u, f) \in -\Delta_\beta$, i.e., $u \in H^1(\Omega)$, $f = -\Delta u$, and there exists $\varphi \in \operatorname{tr} u$ such that $\partial_\nu u = -\beta\varphi$. Then (12.10) holds for all $v \in C(\overline{\Omega}) \cap H^1(\Omega)$, with $\psi = v|_{\partial\Omega}$. By denseness, (12.10) carries over to all $(v,\psi) \in V$; hence $(u, f) \in A$. □

We point out that the form a from the second proof of Theorem 12.15 is what one obtains by applying the completion procedure from Remark 12.3(c) to the form a from the first proof of Theorem 12.15. Note that one starts from an embedded form, but may lose the injectivity of j in the course of the completion.

The results of this and the next section also hold for $\mathbb{K} = \mathbb{C}$. Of course, in the forms and scalar products used in the proofs one then needs to take the complex conjugates of the second arguments.

12.6 The Dirichlet-to-Neumann operator for rough domains

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. We keep the definitions concerning $\mathcal{H}_{n-1}$, tr and ∂_ν from Section 12.5. Again we use $\mathbb{K} = \mathbb{R}$ and assume that $\mathcal{H}_{n-1}(\partial\Omega) < \infty$. We now define the **Dirichlet-to-Neumann operator** D_0.

12.16 Theorem. *The linear relation*

$$D_0 := \big\{(g,h) \in L_2(\partial\Omega) \times L_2(\partial\Omega)\,;\ \exists u \in H^1(\Omega)\colon \Delta u = 0,\ g \in \operatorname{tr} u,\ h = \partial_\nu u\big\}$$

is an accretive self-adjoint operator in $L_2(\partial\Omega)$.

It is part of the assertion of Theorem 12.16 that D_0 is an operator and not just a relation. This is remarkable because, given $g \in L_2(\partial\Omega)$, there may exist several functions $u \in H^1(\Omega)$ such that $\Delta u = 0$ and $g \in \operatorname{tr} u$, due to roughness of the boundary of Ω (see the Notes of this chapter). That D_0 is an operator means that the weak normal derivatives $\partial_\nu u$ of these functions must coincide if they exist in $L_2(\partial\Omega)$. It turns out that, more strongly, at most one of these functions has a weak normal derivative in $L_2(\partial\Omega)$; see Exercise 12.8(b).

For the proof of Theorem 12.16 we need a striking inequality due to Maz'ya. There exists a constant $c > 0$ such that

$$\|u\|^2_{L_2(\Omega)} \leqslant c\Big(\int_\Omega |\nabla u|^2 \,\mathrm{d}x + \int_{\partial\Omega} |u|^2 \,\mathrm{d}\mathcal{H}_{n-1}\Big) \tag{12.11}$$

for all $u \in C(\overline{\Omega}) \cap H^1(\Omega)$. We refer to Corollary D.10 for the proof of this inequality and to Remarks 12.18 and D.11(a) for further comments.

Proof of Theorem 12.16. Put $\operatorname{dom}(a) := C(\overline{\Omega}) \cap H^1(\Omega)$ and $a(u,v) := \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x$. Let $j\colon \operatorname{dom}(a) \to L_2(\partial\Omega)$ be given by $j(u) := u|_{\partial\Omega}$. Then (a,j) is an accretive symmetric form in $L_2(\partial\Omega)$, and (a,j) is densely defined because $C^1(\overline{\Omega}) \subseteq \operatorname{dom}(a)$; recall Section 7.3 for the denseness of $\big\{v|_{\partial\Omega}\,;\ v \in C^1(\overline{\Omega})\big\}$ in $L_2(\partial\Omega)$. Let A be the operator associated with (a,j); see Section 12.1. By Theorem 12.12, A is self-adjoint and accretive; we use (12.8) to show that $A = D_0$.

Let $(g,h) \in A$. Then there exists a sequence (u_k) in $C(\overline{\Omega}) \cap H^1(\Omega)$ such that $u_k|_{\partial\Omega} \to g$ in $L_2(\partial\Omega)$, $\lim_{k,\ell\to\infty} \int_\Omega |\nabla(u_k - u_\ell)|^2 \,\mathrm{d}x = 0$ and $\lim_{k\to\infty} a(u_k, v) = \int_{\partial\Omega} hv \,\mathrm{d}\mathcal{H}_{n-1}$ for all $v \in C(\overline{\Omega}) \cap H^1(\Omega)$. Now inequality (12.11) implies that (u_k) is a Cauchy sequence in $H^1(\Omega)$; let $u \in H^1(\Omega)$ be its limit. Then $g \in \operatorname{tr} u$ and

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_{\partial\Omega} hv \,\mathrm{d}\mathcal{H}_{n-1} \qquad (v \in C(\overline{\Omega}) \cap H^1(\Omega)).$$

Taking test functions $v \in C_c^\infty(\Omega)$ one obtains $\Delta u = 0$. Thus

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_\Omega (\Delta u) v \,\mathrm{d}x = \int_{\partial\Omega} hv \,\mathrm{d}\mathcal{H}_{n-1}$$

for all $v \in C(\overline{\Omega}) \cap H^1(\Omega)$. Hence $\partial_\nu u = h$ by the definition in Section 12.5, and we have shown that $A \subseteq D_0$.

Conversely, let $(g,h) \in D_0$. Then there exists $u \in H^1(\Omega)$ such that $g \in \operatorname{tr} u$, $\Delta u = 0$, $\partial_\nu u = h$. Hence, there exists a sequence (u_k) in $\operatorname{dom}(a)$ such that $u_k|_{\partial\Omega} \to g$ in $L_2(\partial\Omega)$ and $u_k \to u$ in $H^1(\Omega)$. Then $\lim_{k,\ell\to\infty} a(u_k - u_\ell) = 0$ and

$$a(u_k, v) = \int_\Omega \nabla u_k \cdot \nabla v \,\mathrm{d}x \to \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_\Omega (\Delta u) v \,\mathrm{d}x = \int_{\partial\Omega} hv \,\mathrm{d}\mathcal{H}_{n-1}$$

for all $v \in \operatorname{dom}(a)$. This shows that $(g,h) \in A$. □

12.17 Remark. One can prove Theorem 12.16 without using the results of Sections 12.1 to 12.4, similarly as in our second proof of Theorem 12.15. Again one puts $V := \mathrm{tr}$, and one defines the form a on V by

$$a\big((u,\varphi),(v,\psi)\big) := \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x \qquad \big((u,\varphi),(v,\psi) \in V\big)$$

and $j\colon V \to L_2(\partial\Omega)$ by $j\big((u,\varphi)\big) := \varphi$. We point out that then Maz'ya's inequality (12.11) is responsible for the j-coercivity of the form a: for $(u,\varphi) \in V$ one first obtains

$$\frac{1}{c}\|u\|_{L_2(\Omega)}^2 \leqslant \int_\Omega |\nabla u|^2 \,\mathrm{d}x + \int_{\partial\Omega} |\varphi|^2 \,\mathrm{d}\mathcal{H}_{n-1} = a\big((u,\varphi)\big) + \|j(u,\varphi)\|_{L_2(\partial\Omega)}^2 \tag{12.12}$$

by denseness, and then $a\big((u,\varphi)\big) + \|j(u,\varphi)\|_2^2 \geqslant \frac{1}{2c}\|u\|_2^2 + \frac{1}{2}\|\nabla u\|_2^2$.

The reader is asked to carry out the details in Exercise 12.8. △

12.18 Remark. If Ω has C^1-boundary, then we can easily prove a version of (12.11) at this point. Indeed, in the proof of Proposition 8.2 we have shown that

$$\int_\Omega |u|^2 \,\mathrm{d}x \leqslant \frac{1}{2}\int_\Omega |u|^2 \,\mathrm{d}x + \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x + c\int_\Omega |\nabla u|^2 \,\mathrm{d}x + c\int_{\partial\Omega} |u|^2 \,\mathrm{d}\sigma,$$

and this implies

$$\|u\|_{L_2(\Omega)}^2 \leqslant (2c+1)\int_\Omega |\nabla u|^2 \,\mathrm{d}x + 2c\int_{\partial\Omega} |u|^2 \,\mathrm{d}\sigma.$$

Note, however, that the case of C^1-boundary has already been treated in Section 8.1; the purpose of the present section is to treat open sets with rough boundary. (The attentive reader will have noticed that above we have proved (12.11) with σ in place of $\mathcal{H}_{n-1}$. In Theorem E.3 we show that in fact σ and $\mathcal{H}_{n-1}$ coincide on $\partial\Omega$.)

Inequality (12.11) is a consequence of the following remarkable stronger inequality due to Maz'ya. There exists a constant $c(n)$, only depending on the dimension n, such that

$$\|u\|_{L_q(\Omega)} \leqslant c(n)\Big(\int_\Omega |\nabla u| \,\mathrm{d}x + \int_{\partial\Omega} |u| \,\mathrm{d}\mathcal{H}_{n-1}\Big) \tag{12.13}$$

for all $u \in C(\overline{\Omega}) \cap W_1^1(\Omega)$, where $q = \frac{n}{n-1}$; see [Maz11; Example 5.6.2/1 and Theorem 5.6.3]. We refer to Section D.1 for the definition of the Sobolev space $W_1^1(\Omega)$ and to Remark D.11(a) for comments on (12.13) and its proof. △

Notes

The treatment of sectorial forms in the setting of this chapter is adapted from [ArEl12b]. In that paper, our description of the operator A in Theorem 12.11 is taken as the *definition* of the operator associated with a non-closed sectorial form (a, j). The applications treated in Sections 12.5 and 12.6 are also taken from [ArEl12b]. More information on the Dirichlet-to-Neumann operator is contained in [ArEl11].

For semi-bounded symmetric operators, the Friedrichs extension is contained in [Fri34]; see also Freudenthal [Fre36]. The Friedrichs extension for sectorial operators is due to Kato [Kat80; Chap. VI, § 2.3].

One may ask which operators are associated with forms. By Theorem 12.4 and Corollary 12.9, an operator A in a complex Hilbert space H is associated with a densely defined embedded closed sectorial form in H if and only if A is m-sectorial. (Recall that the latter is equivalent to $-A$ being the generator of a contractive holomorphic C_0-semigroup on H.) As was illustrated in Exercise 5.9, one obtains a larger class of operators if one allows changes of the scalar product on H. In this regard, the following result extends the above equivalence for an operator A in a complex Hilbert space H. There exist an embedded closed quasi-sectorial form a and an equivalent scalar product $[\cdot,\cdot]$ on H such that the operator A is associated with the form a in $(H,[\cdot,\cdot])$ if and only if $-A$ generates a holomorphic C_0-semigroup and $A+\omega$ has bounded imaginary powers for large $\omega \in \mathbb{R}$. For this equivalence we refer to [ABH01; Theorem 3.3] and [Haa06; Corollary 7.3.10], where also the notion of bounded imaginary powers is explained.

We now give some further information concerning the trace, as it was defined in Section 12.5. In the remainder of these Notes we always assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with $\mathcal{H}_{n-1}(\partial\Omega) < \infty$. Recall that the linear relation tr in $H^1(\Omega) \times L_2(\partial\Omega)$ is defined as the closure of the operator $u \mapsto u|_{\partial\Omega} \colon C(\overline{\Omega}) \cap H^1(\Omega) \to L_2(\partial\Omega)$. For $u \in H^1(\Omega)$, the set $\operatorname{tr} u$ is sometimes called the 'approximative trace' of u.

An intriguing question is whether $\operatorname{tr} 0 = \{0\}$ (or equivalently, whether $\operatorname{tr} u$ is a singleton for some/all $u \in \operatorname{dom}(\operatorname{tr})$). In general, the answer is "no", as one sees from Exercise 12.7(b); for an example with a connected set Ω see [ArEl11; Example 4.4]. Recent positive results are due to Sauter, who investigates this question systematically: if Ω has continuous boundary, then $\operatorname{tr} 0 = \{0\}$ (see [Sau20; Theorem 4.11]). Surprisingly, in dimension $n = 2$ the same result holds if Ω is merely connected; see [Sau20; Corollary 5.4]. (In fact, Sauter defines and investigates the approximative trace for general non-empty open subsets of $\mathbb{R}^n$.)

Another interesting question is whether the space $H^1_0(\Omega)$ can be characterised by means of the approximative trace. This is true if Ω has continuous boundary: then it follows from [Sau13; Corollary 7.48] in combination with the result mentioned in the previous paragraph that $H^1_0(\Omega) = \{u \in H^1(\Omega)\,;\ \operatorname{tr} u = \{0\}\}$; this generalises Theorem 7.12.

In the general case, the inclusion $H^1_0(\Omega) \subseteq \{u \in H^1(\Omega)\,;\ 0 \in \operatorname{tr} u\}$ is immediate from the definition of the trace. Below we show that the reverse inclusion holds if and only if the Dirichlet problem $\Delta u = 0$, $\operatorname{tr} u \ni g$ has a unique solution $u \in H^1(\Omega)$ for each $g \in \operatorname{ran}(\operatorname{tr})$; see also [AES23; Theorem 5.2]. An example for non-uniqueness can be found in [AES23; Example 5.3], with the set $\Omega = B(0,2) \setminus (S \times \{0\}) \subseteq \mathbb{R}^2$, where $S \subseteq \mathbb{R}$ is the usual 'middle third Cantor set'.

We give a short proof of the equivalence mentioned in the previous paragraph. Let $H^1_\Delta(\Omega) := \{u \in H^1(\Omega)\,;\ \Delta u = 0\}$. It follows from Exercise 7.9 that the mapping $J \colon H^1_\Delta(\Omega) \to H^1(\Omega)/H^1_0(\Omega)$, $u \mapsto u + H^1_0(\Omega)$ is bijective. The equality of the spaces $H^1_0(\Omega)$ and $\{u \in H^1(\Omega)\,;\ 0 \in \operatorname{tr} u\}$ can be rephrased as the property that the mapping $\widetilde{\operatorname{tr}} \colon H^1(\Omega)/H^1_0(\Omega) \to \operatorname{ran}(\operatorname{tr})$, $u + H^1_0(\Omega) \mapsto \operatorname{tr} u$ is bijective. Thus, equality of the spaces holds if and only if $\operatorname{tr} \colon H^1_\Delta(\Omega) \to \operatorname{ran}(\operatorname{tr})$ is bijective, and the latter property signifies uniqueness of solutions of the Dirichlet problem.

Exercises

12.1 Let V be a $\mathbb{C}$-vector space, H a complex Hilbert space, $0 \neq j\colon V \to H$ linear, $a\colon V \times V \to \mathbb{C}$ a sesquilinear form. Let $\omega \in \mathbb{R}$, $\theta \in [0, \pi/2)$. Show that the following two properties are equivalent:

(i) (a, j) is quasi-sectorial with vertex $-\omega$ and angle θ,

(ii) $\{a(u);\, u \in V,\ \|j(u)\|_H = 1\} \subseteq -\omega + \overline{\Sigma_\theta}$.

Hint: First show that (ii) implies $\{a(u);\, u \in V,\ j(u) = 0\} \subseteq \overline{\Sigma_\theta}$.

12.2 Let Ω be a non-empty bounded open subset of $\mathbb{R}^n$, and let a be the classical Dirichlet form in $L_2(\Omega)$ with $\mathrm{dom}(a) := C_c^\infty(\Omega)$. Recall that the m-sectorial operator A associated with a by (12.5) is the negative Dirichlet Laplacian; see Example 12.6.

(a) Show that the operator A_0 from Proposition 12.2 is given by $A_0 = -\Delta$, with $\mathrm{dom}(A_0) = C_c^\infty(\Omega)$.

(b) Show that $\overline{A_0}$ is not m-sectorial. (This implies, in particular, that $\overline{A_0} \neq A$.)

(c) Show that the Friedrichs extension of the operator $-A_0$ (with A_0 from part (a)) is the Dirichlet Laplacian.

Note. The Neumann Laplacian is another accretive self-adjoint extension of $-A_0$.

12.3 Let H be a complex Hilbert space, a an embedded sectorial form in H. Prove the following criteria for a being closed or closable:

(a) a is closed if and only if for any Cauchy sequence (u_n) in $(\mathrm{dom}(a), \|\cdot\|_a)$ with $u_n \to u$ in H one has $\|u_n - u\|_a \to 0$.

(b) a is closable if and only if for any Cauchy sequence (u_n) in $(\mathrm{dom}(a), \|\cdot\|_a)$ with $u_n \to 0$ in H one has $\|u_n\|_a \to 0$.

12.4 Let $H := L_2(-1,1)$, and let the forms a_1, a_2 in H be defined by $\mathrm{dom}(a_j) = C_c^\infty(-1,1)$,

$$a_1(u,v) := u(0)\overline{v(0)}, \qquad a_2(u,v) := \int_{-1}^{1} u'(x)\overline{v'(x)}\,\mathrm{d}x + u(0)\overline{v(0)}$$

for all $u, v \in C_c^\infty(-1,1)$.

For $j = 1, 2$ determine whether a_j is closable. Find the completion of $(\mathrm{dom}(a_j), (\cdot \mid \cdot)_{a_j})$ and the operator associated with a_j.

12.5 This exercise illustrates Remark 12.14(b). Let $u \in C^2(\overline{(-1,1)})$, $u'(-1) = u'(1) = 0$, $u'(0) \neq 0$, and put $\Omega := (-1,0) \cup (0,1)$.

(a) Show that for the weak normal derivative in the sense of Section 12.5 one has $\partial_\nu u = 0$.

(b) Show that one does *not* have '$\partial_\nu u = 0$' – observe the quotes! – in the sense of Section 7.4. (Hint: Choose $v \in H^1(\Omega)$ with $v(0-) \neq v(0+)$ as a 'test function'.)

12.6 Put $\Omega := (-1,0) \cup (0,1)$.

(a) Determine the relation tr of Section 12.5, and show that dom(tr) is not dense.

(b) Find $\partial_\nu u$ for those $u \in H^1(\Omega)$ with $\Delta u \in L_2(\Omega)$ for which $\partial_\nu u \in L_2(\partial\Omega)$.

(c) Determine the Robin Laplacian for $\beta = 1$.

12.7 Put $S := [0,1] \times \{0\} \subseteq \mathbb{R}^2$, and let (x_n) be a bounded sequence in $\mathbb{R}^2 \setminus S$ having the set S as its cluster points. Let (r_n) be a sequence in $(0,\infty)$ such that $\sum_{n=1}^{\infty} r_n < \infty$, $B[x_n, r_n] \cap S = \varnothing$ $(n \in \mathbb{N})$ and $B[x_n, r_n] \cap B[x_m, r_m] = \varnothing$ $(m, n \in \mathbb{N},\ m \neq n)$. Put $\Omega := \bigcup_{n\in\mathbb{N}} B(x_n, r_n)$.

(a) Determine $\partial\Omega$ and $\mathcal{H}_1(\partial\Omega)$ (1-dimensional Hausdorff measure). (Note that $S \subseteq \partial\Omega$.)

(b) Show that $\mathrm{dom}(\mathrm{tr})$ is dense in $H^1(\Omega)$ and that $\mathrm{tr}\, 0 = L_2(S)$ (where $L_2(S)$ is considered as a subspace of $L_2(\partial\Omega)$ in the natural way).

(c) Let D_0 be the Dirichlet-to-Neumann operator for Ω. Show that $L_2(S) \subseteq \mathrm{dom}(D_0)$ and that $D_0|_{L_2(S)} = 0$.

Note. In view of the reference [Sau20; Corollary 5.4] mentioned in the Notes, a bounded open set $\Omega \subseteq \mathbb{R}^2$ with the property $\mathrm{tr}\, 0 \neq \{0\}$ cannot be connected; in fact, this reference implies that such a set must consist of infinitely many connected components.

12.8 (a) Complete the proof of Theorem 12.16 as suggested in Remark 12.17.

(b) Let $g \in L_2(\partial\Omega)$. Show that there is at most one solution $u \in H^1(\Omega)$ of the Dirichlet problem $\Delta u = 0$, $\mathrm{tr}\, u \ni g$ such that the weak normal derivative $\partial_\nu u$ exists. (Hint: Show that for $u \in H^1(\Omega)$, the properties $\Delta u = 0$, $0 \in \mathrm{tr}\, u$ and $\partial_\nu u = 0$ imply $a\big((u,0)\big) = 0$, with the form a from Remark 12.17; then use (12.12).)

Chapter 13

Approximation of strongly continuous semigroups

In this chapter we present two important classical results on the approximation of semigroups, the Trotter approximation theorem and the Chernoff product formula. In view of the applications in Chapters 14 and 15 we treat these topics in the more general context of ‘degenerate strongly continuous semigroups’. This concept will be introduced in Section 13.1. The chapter closes with an introduction to the spectral theorem for self-adjoint operators.

13.1 Degenerate strongly continuous semigroups and m-accretive linear relations

Let X be a Banach space, and let T be a one-parameter semigroup on X. Recall from Remark 1.1(c) that then $T(0)$ is a projection. We call T a **degenerate strongly continuous semigroup** if

$$T(0) = \text{s-lim}_{t\to 0+} T(t)$$

(not necessarily $T(0) = I$!). Note that a C_0-semigroup is also a degenerate strongly continuous semigroup. We would have preferred to omit ‘degenerate’ in the previous definition; however, in the literature the notions ‘strongly continuous semigroup’ and ‘C_0-semigroup’ are usually synonymous.

A degenerate strongly continuous semigroup T is the direct sum of a C_0-semigroup and a semigroup that is identically zero. Indeed, with $P := T(0)$ one obtains $T(t)P = PT(t)$ for all $t \geqslant 0$, and hence $X_\mathrm{a} := \operatorname{ran}(P)$ is invariant under T. We denote by T_a the restriction of T to X_a, i.e. $T_\mathrm{a}(t) := T(t)|_{X_\mathrm{a}}$ for all $t \geqslant 0$. Then T_a is a C_0-semigroup on X_a, whereas $T(t) = T(t)P$ vanishes on $X_\mathrm{u} := \ker(P)$ for all $t \geqslant 0$. In particular it follows that $[0,\infty) \ni t \mapsto T(t)$ is strongly continuous. (We use the indices a and u to denote the ‘active’ subspace of X, where T acts as a C_0-semigroup, and the ‘unactive’ subspace of X, where T vanishes.)

There exist $M \geqslant 1$, $\omega \in \mathbb{R}$ such that $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ $(t \geqslant 0)$. (This property is known for the C_0-semigroup T_a and carries over to T because $T(t) = T_\mathrm{a}(t)P$ for all $t \geqslant 0$.) Let A_a denote the generator of T_a. Then it follows from Theorem 2.7 that

$$R(\lambda, A_\mathrm{a})P = \int_0^\infty \mathrm{e}^{-\lambda t} T_\mathrm{a}(t)\,\mathrm{d}t\,P = \int_0^\infty \mathrm{e}^{-\lambda t} T(t)\,\mathrm{d}t \qquad \text{(strong integral)}$$

for all $\lambda \in \mathbb{K}$ with $\operatorname{Re}\lambda > \omega$. It is easy to see that the mapping $\rho(A_{\rm a}) \ni \lambda \mapsto R(\lambda, A_{\rm a})P \in \mathcal{L}(X)$ is a **pseudo-resolvent**, i.e. $R(\cdot, A_{\rm a})P$ satisfies the resolvent equation (2.1) (see also Exercise 13.1).

We define the **generator** of T as the linear relation

$$A := \big\{(x, y) \in X \times X\,;\ (x, Py) \in A_{\rm a}\big\}. \tag{13.1}$$

It follows from Exercise 13.2 that the pseudo-resolvent $R(\cdot, A_{\rm a})P$ introduced above is the **resolvent** of the generator A, which is defined by

$$R(\lambda, A) := (\lambda I - A)^{-1}$$

for all $\lambda \in \rho(A) := \{\lambda \in \mathbb{K};\ (\lambda I - A)^{-1} \in \mathcal{L}(X)\} = \rho(A_{\rm a})$, where

$$\lambda I - A := \big\{(x, \lambda x - y)\,;\ (x, y) \in A\big\}; \tag{13.2}$$

recall Subsection 1.3.1 concerning the inverse of a relation. According to the definition (13.1),

$$A = A_{\rm a} \oplus (\{0\} \times X_{\rm u}) \qquad \text{(direct sum in } X \times X\text{)}; \tag{13.3}$$

in particular, $\{0\} \times X_{\rm u}$ is a subset of A. One easily sees that the generator A determines the semigroup T uniquely.

We illustrate these notions by an example.

13.1 Example. Let X be a Banach space, $P \in \mathcal{L}(X)$ a projection, $X_{\rm a} := \operatorname{ran}(P)$. Then $T(t) := P$ ($t \geqslant 0$) defines a degenerate strongly continuous semigroup, the generator of T is given by

$$A = \big(X_{\rm a} \times \{0\}\big) \oplus \big(\{0\} \times \ker(P)\big) = X_{\rm a} \times \ker(P),$$

and the resolvent of A is given by $R(\lambda, A) = \lambda^{-1}P$, for $\lambda \in \rho(A_{\rm a}) = \rho(0_{\mathcal{L}(X_{\rm a})})$ ($= \mathbb{K} \setminus \{0\}$ if $P \neq 0$). △

A degenerate strongly continuous semigroup T is called **contractive** if $\|T(t)\| \leqslant 1$ for all $t \geqslant 0$, and T is **quasi-contractive** if there exists $\omega \in \mathbb{R}$ such that $\|T(t)\| \leqslant {\rm e}^{\omega t}$ for all $t \geqslant 0$. The concept of rescaling works as before: if $\lambda \in \mathbb{K}$ and A is the generator of T, then the generator of the rescaled degenerate strongly continuous semigroup $t \mapsto {\rm e}^{-\lambda t}T(t)$ is the relation $A - \lambda I$.

In the remainder of this section we deal with the special case in which the underlying space is a Hilbert space H. Similarly as in Section 3.4, a linear relation A in H is called **accretive** if $\operatorname{Re}(y \,|\, x) \geqslant 0$ for all $(x, y) \in A$, and A is called **m-accretive** if additionally $\operatorname{ran}(I + A) = \{x + y\,;\ (x, y) \in A\} = H$.

If T is a contractive degenerate strongly continuous semigroup on H, then $P := T(0)$ is a contractive projection onto the 'active' subspace $H_{\rm a} := \operatorname{ran}(P)$; hence P is the orthogonal projection onto $H_{\rm a}$. Let $A_{\rm a}$ be the generator of the C_0-semigroup $T_{\rm a}$. Then the generator A of T is the relation $A := A_{\rm a} \oplus (\{0\} \times H_{\rm u})$ (orthogonal direct sum in $H \oplus H$), where $H_{\rm u} = H_{\rm a}^{\perp}$. For $(x, y) \in A$ one obtains

$$\operatorname{Re}(y \,|\, x) = \operatorname{Re}(Py \,|\, x) = \operatorname{Re}(A_{\rm a}x \,|\, x) \leqslant 0$$

(because $-A_{\rm a}$ is m-accretive), i.e. the linear relation $-A = \{(x,-y);\ (x,y) \in A\}$ is accretive. It is even m-accretive since

$$\operatorname{ran}(I - A) = \operatorname{ran}(I_{\rm a} - A_{\rm a}) \oplus \operatorname{ran}(\{0\} \times H_{\rm u}) = H_{\rm a} \oplus H_{\rm u} = H,$$

where $I_{\rm a}$ is the identity operator in $H_{\rm a}$.

Next we show that all m-accretive linear relations arise as an orthogonal direct sum of an m-accretive operator and a 'trivial' m-accretive linear relation, as above.

13.2 Theorem. *Let A be a linear relation in H.*

(a) *Let A be m-accretive. Put $H_{\rm a} := \overline{\operatorname{dom}(A)}$ and $H_{\rm u} := H_{\rm a}^{\perp}$. Then $A_{\rm a} := A \cap (H_{\rm a} \times H_{\rm a})$ is an m-accretive operator in $H_{\rm a}$, $A_{\rm u} := A \cap (H_{\rm u} \times H_{\rm u}) = \{0\} \times H_{\rm u}$ is an m-accretive linear relation in $H_{\rm u}$, and $A = A_{\rm a} \oplus A_{\rm u}$.*

(b) *The linear relation $-A$ is the generator of a contractive degenerate strongly continuous semigroup on H if and only if A is m-accretive.*

For the proof we single out a technical detail.

13.3 Lemma. *Let A be an accretive linear relation in H with dense domain. Then A is an operator.*

Proof. Let $(0,z) \in A$. Then for all $(x,y) \in A$, $\lambda \in \mathbb{K}$, one obtains $(x, y + \lambda z) \in A$,

$$0 \leqslant \operatorname{Re}(y + \lambda z \,|\, x) = \operatorname{Re}(y \,|\, x) + \operatorname{Re}(\lambda(z \,|\, x)).$$

It follows that $z \perp \operatorname{dom}(A)$, hence $z = 0$. □

Proof of Theorem 13.2. (a) Obviously $A_{\rm a}$ and $A_{\rm u}$ are accretive linear relations, and $\operatorname{dom}(A_{\rm u}) = \{0\}$. Let $z \in H_{\rm u}$. By the m-accretivity of A there exists $(x,y) \in A$ such that $x + y = z$. Then $\operatorname{Re}(y \,|\, x) \geqslant 0$, and

$$0 = (z \,|\, x) = (x + y \,|\, x) = \|x\|^2 + (y \,|\, x) \geqslant \|x\|^2$$

implies $x = 0$, hence $(0,z) = (0,y) = (x,y) \in A$. This shows that $A_{\rm u} = \{0\} \times H_{\rm u}$.

Let $(x,y) \in A$. There exist $y_{\rm a} \in H_{\rm a}$, $y_{\rm u} \in H_{\rm u}$ such that $y = y_{\rm a} + y_{\rm u}$. Then $(0, y_{\rm u}) \in A_{\rm u} \subseteq A$, hence $(x, y_{\rm a}) \in A \cap (H_{\rm a} \times H_{\rm a}) = A_{\rm a}$. It follows that $\operatorname{dom}(A_{\rm a}) = \operatorname{dom}(A)$, and since $\operatorname{dom}(A)$ is dense in $H_{\rm a}$, Lemma 13.3 implies that $A_{\rm a}$ is an operator. It also follows that $A = A_{\rm a} \oplus A_{\rm u}$; thus $H = \operatorname{ran}(I + A) = \operatorname{ran}(I_{\rm a} + A_{\rm a}) \oplus H_{\rm u}$, which yields the range condition $\operatorname{ran}(I_{\rm a} + A_{\rm a}) = H_{\rm a}$.

Summing up, we have shown that A is the direct sum of the m-accretive operator $A_{\rm a}$ and the m-accretive linear relation $A_{\rm u}$.

(b) If $-A$ generates a contractive degenerate strongly continuous semigroup, then the considerations before Theorem 13.2 show that A is m-accretive.

Conversely, if A is m-accretive, then the Lumer–Phillips theorem (Theorem 3.16) shows that $-A_{\rm a}$ generates a contractive C_0-semigroup $T_{\rm a}$ on $H_{\rm a}$, where we use the notation from part (a). Let P denote the orthogonal projection onto $H_{\rm a}$. Then $T(t) := T_{\rm a}(t)P$ $(t \geqslant 0)$ defines a degenerate strongly continuous semigroup T on H with generator $-A$. □

A statement analogous to Theorem 13.2 also holds for **quasi-m-accretive** linear relations, i.e. for relations A such that $A + \omega I$ is m-accretive for some $\omega \in \mathbb{R}$.

13.2 A Trotter approximation theorem

In the first part of this section X will be a Banach space. For simplicity we restrict our treatment to the case of contractive degenerate strongly continuous semigroups; the following theorem is in fact also valid for sequences of semigroups that are uniformly bounded on $[0,1]$.

13.4 Theorem (Trotter approximation theorem for degenerate strongly continuous semigroups). *Let T and T_n $(n \in \mathbb{N})$ be contractive degenerate strongly continuous semigroups on X, with generators A and A_n $(n \in \mathbb{N})$, respectively. Assume that there exists $\lambda > 0$ such that*

$$R(\lambda, A_n)x \to R(\lambda, A)x \qquad (n \to \infty) \tag{13.4}$$

for all $x \in \operatorname{ran}(T(0))$. Then

$$T_n(t)x \to T(t)x \qquad (n \to \infty) \tag{13.5}$$

uniformly on compact subsets of $[0, \infty)$, for all $x \in \operatorname{ran}(T(0))$.

13.5 Remark. It is worth noticing that the converse of Theorem 13.4 is also true. Indeed, if T and T_n $(n \in \mathbb{N})$ are contractive degenerate strongly continuous semigroups, and

$$T(t)x = \lim_{n\to\infty} T_n(t)x \tag{13.6}$$

for all $x \in \operatorname{ran}(T(0))$, $t \geqslant 0$, then

$$R(\lambda, A)x = \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t = \lim_{n\to\infty} \int_0^\infty \mathrm{e}^{-\lambda t} T_n(t)x \,\mathrm{d}t = \lim_{n\to\infty} R(\lambda, A_n)x$$

for all $x \in \operatorname{ran}(T(0))$, $\lambda > 0$, by the dominated convergence theorem.

We point out that, even though only pointwise convergence is required in (13.6), Theorem 13.4 implies that the convergence in (13.6) is in fact uniform for t in compact subsets of $[0, \infty)$. △

13.6 Example. The following example shows that the convergence (13.4) for all $x \in X$ does not imply the convergence (13.5) for all $x \in X$.

Let $X := \mathbb{C}$, $A_n := \mathrm{i}n$, $A := \{0\} \times \mathbb{C}$. Then $R(\lambda, A_n) = (\lambda - \mathrm{i}n)^{-1} \to 0 = R(\lambda, A)$ for all $\lambda > 0$, but $T_n(t) = \mathrm{e}^{\mathrm{i}nt}$ does not converge to $T(t) = 0$. △

In the proof of Theorem 13.4 we will need to express semigroup differences in terms of resolvent differences.

13.7 Lemma. *Let T and S be contractive degenerate strongly continuous semigroups, with generators A and B, respectively. Then for all $\lambda, t > 0$ and $x \in X$ we have*

$$R(\lambda, B)\big(T(t) - S(t)\big)R(\lambda, A)x = \int_0^t S(t-s)\big(R(\lambda, A) - R(\lambda, B)\big)T(s)x \,\mathrm{d}s. \tag{13.7}$$

Proof. We write $P_T := T(0)$, and we denote the generator of T_{a} by A_{a}. (Recall the notation T_{a} and A_{a} from Section 13.1.) In the computation made below we will use the identities $T(t) = P_T T(t) = T(t)P_T$, $T(t)R(\lambda, A) = R(\lambda, A)T(t)$ and $R(\lambda, A) = P_T R(\lambda, A) = R(\lambda, A_{\mathrm{a}})P_T$. The corresponding notation and properties will be used for S.

For $x \in X$, the function $s \mapsto S(t-s)R(\lambda, B)T(s)R(\lambda, A)x$ is differentiable on $[0, t]$, with

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s} S(t-s)R(\lambda, B)T(s)R(\lambda, A)x \\
&\qquad = S(t-s)\big(-B_{\mathrm{a}}R(\lambda, B)R(\lambda, A) + R(\lambda, B)A_{\mathrm{a}}R(\lambda, A)\big)T(s)x \\
&\qquad = S(t-s)\big(P_S R(\lambda, A) - R(\lambda, B)P_T\big)T(s)x \\
&\qquad = S(t-s)\big(R(\lambda, A) - R(\lambda, B)\big)T(s)x,
\end{aligned}$$

where in the middle equality we have used $-B_{\mathrm{a}}R(\lambda, B) = (\lambda - B_{\mathrm{a}} - \lambda)R(\lambda, B) = P_S - \lambda R(\lambda, B)$, and similarly $A_{\mathrm{a}}R(\lambda, A) = -P_T + \lambda R(\lambda, A)$. Integrating with respect to s for $0 \leqslant s \leqslant t$ we obtain (13.7). □

Another ingredient in the proof of Theorem 13.4 is the following important fact.

13.8 Lemma. *Let A be the generator of a C_0-semigroup on X. Then* $\mathrm{dom}(A^2)$ *is dense in X.*

Proof. If $B \in \mathcal{L}(X)$ has dense range, then obviously $\mathrm{ran}(B^2) = B(\mathrm{ran}(B))$ is dense. Applying this fact to a resolvent of A – and recalling that $\mathrm{dom}(A)$ is dense – one obtains the assertion. □

Proof of Theorem 13.4. As before we denote by A_{a} the generator of the C_0-semigroup T_{a} on the 'active' subspace $X_{\mathrm{a}} := \mathrm{ran}(T(0))$.

Let $y \in X_{\mathrm{a}}$, $t_0 > 0$. We estimate

$$\begin{aligned}
\big\|\big(T(t) - T_n(t)\big)R(\lambda, A)y\big\| \leqslant{}& \big\|\big(R(\lambda, A) - R(\lambda, A_n)\big)T(t)y\big\| \\
&+ \big\|R(\lambda, A_n)\big(T(t) - T_n(t)\big)y\big\| \\
&+ \big\|T_n(t)\big(R(\lambda, A_n) - R(\lambda, A)\big)y\big\|.
\end{aligned} \tag{13.8}$$

The third term on the right-hand side of (13.8) converges to 0 uniformly for t in $[0, \infty)$. For the first term we note that $\{T(t)y;\, 0 \leqslant t \leqslant t_0\}$ is a compact subset of X_{a} – by the strong continuity of T – and that strong convergence of a sequence of operators implies uniform convergence on compact sets; see Exercise 3.2(b). So we conclude uniform convergence to 0 for $0 \leqslant t \leqslant t_0$.

For the treatment of the second term on the right-hand side of (13.8) we confine our attention to the case in which y is of the form $y = R(\lambda, A)z$, for some $z \in X_{\mathrm{a}}$, and we apply Lemma 13.7 to obtain

$$\begin{aligned}
\big\|R(\lambda, A_n)\big(T(t) - T_n(t)\big)R(\lambda, A)z\big\| &= \Big\|\int_0^t T_n(t-s)\big(R(\lambda, A) - R(\lambda, A_n)\big)T(s)z \,\mathrm{d}s\Big\| \\
&\leqslant \int_0^t \big\|\big(R(\lambda, A) - R(\lambda, A_n)\big)T(s)z\big\| \,\mathrm{d}s.
\end{aligned}$$

As above one concludes that the integrand converges to 0 uniformly for $0 \leqslant s \leqslant t_0$, and therefore $\big\|R(\lambda, A_n)\big(T(t) - T_n(t)\big)y\big\| \to 0$ uniformly for $0 \leqslant t \leqslant t_0$.

So far we have shown that $T_n(t)x \to T(t)x$ as $n \to \infty$, uniformly for t in compact subsets of $[0, \infty)$, for all x of the form $x = R(\lambda, A)^2 z$ with $z \in X_{\rm a}$, i.e. for all $x \in \operatorname{dom}(A_{\rm a}^2)$. As $\operatorname{dom}(A_{\rm a}^2)$ is dense in $X_{\rm a}$ – by Lemma 13.8 – and T_n ($n \in \mathbb{N}$) as well as T are contractive, one obtains the asserted convergence for all $x \in X_{\rm a}$. □

Example 13.6 illustrates that in the setting of Theorem 13.4 there is no information on the convergence of the sequence of semigroups on the complement of $\operatorname{ran}(T(0))$. The objective of the remaining part of this section is to present a setting in which this deficiency can be remedied; see Theorem 13.10.

We will need the following concept of operator convergence that is important in the context of semigroup approximation. If A and A_n, for $n \in \mathbb{N}$, are generators of contractive degenerate strongly continuous semigroups, then we say that (A_n) converges to A in the **strong resolvent sense** if $R(\lambda, A_n) \to R(\lambda, A)$ strongly as $n \to \infty$, for all $\lambda > 0$. It follows from the subsequent Lemma 13.9 that it suffices to require the convergence for one $\lambda > 0$. Actually, we will also use this terminology for generators $-A_n$ and $-A$, as well as for the case when all operators are shifted by a common real multiple of the identity I. We refer to [Kat80; Chap. VIII, § 1.1], where convergence in the strong resolvent sense is introduced under the name 'convergence in the generalised sense'.

Note that a property similar to the above definition was already used in Theorem 13.4, but there the strong convergence of the resolvents was only required on a subspace of the Banach space.

13.9 Lemma. *Let A and A_n ($n \in \mathbb{N}$) be generators of contractive degenerate strongly continuous semigroups on X. Assume that there exists $\lambda > 0$ such that*

$$R(\lambda, A_n) \to R(\lambda, A) \qquad (n \to \infty)$$

strongly. Then

$$R(\mu, A_n)x \to R(\mu, A)x \qquad (n \to \infty),$$

uniformly for μ in compact subsets of $(0, \infty)$, for all $x \in X$.

Proof. A standard computation with applications of the resolvent equation (2.1) yields

$$\begin{aligned} R(\mu, A) &- R(\mu, A_n) \\ &= \big(I + (\lambda - \mu)R(\mu, A_n)\big)R(\mu, A) - R(\mu, A_n)\big(I + (\lambda - \mu)R(\mu, A)\big) \\ &= \big(I + (\lambda - \mu)R(\mu, A_n)\big)\big(R(\lambda, A) - R(\lambda, A_n)\big)\big(I + (\lambda - \mu)R(\mu, A)\big) \end{aligned}$$

for all $\mu > 0$. Let $x \in X$, and let $0 < \mu_0 < \mu_1$. Then the set

$$\big\{\big(I + (\lambda - \mu)R(\mu, A)\big)x\,;\ \mu_0 \leqslant \mu \leqslant \mu_1\big\}$$

is compact because $R(\cdot, A)$ is continuous, and

$$\sup_{\mu_0 \leqslant \mu \leqslant \mu_1,\, n \in \mathbb{N}} \|I + (\lambda - \mu)R(\mu, A_n)\| < \infty.$$

(For the continuity of $R(\cdot, A)$ recall from Section 13.1 that $R(\cdot, A) = R(\cdot, A_{\rm a})P$; see also Exercise 13.1(a).) Since $R(\lambda, A_n)y \to R(\lambda, A)y$ as $n \to \infty$, uniformly for y in compact subsets of X – by Exercise 3.2(b) –, one obtains the assertion. □

In the following supplement to Theorem 13.4 we present a more special setting in which one also obtains information on the convergence of the approximating semigroups on the complement of $\operatorname{ran}(T(0))$, as announced above.

13.10 Theorem. *Let H be a Hilbert space, and let T and T_n ($n \in \mathbb{N}$) be degenerate strongly continuous semigroups of self-adjoint contractions on H, with corresponding generators A and A_n ($n \in \mathbb{N}$). Assume that (A_n) converges to A in the strong resolvent sense. Then*

$$T_n(t)x \to T(t)x \qquad (n \to \infty)$$

uniformly on compact subsets of $[0, \infty)$, for all $x \in \operatorname{ran}(T(0))$, and

$$T_n(t)x \to 0 = T(t)x \qquad (n \to \infty)$$

uniformly for $t \in [\varepsilon, \infty)$, for all $\varepsilon > 0$ and all $x \in \ker(T(0)) = \operatorname{ran}(T(0))^\perp$. (Note that $T(0)$ is a self-adjoint, therefore orthogonal, projection.)

Combining both assertions of Theorem 13.10 one sees that on compact subsets of $(0, \infty)$ the uniform convergence $T_n(\cdot)x \to T(\cdot)x$ holds for all $x \in X$. We point out that, even if all the semigroups T_n in Theorem 13.10 are C_0-semigroups, it is possible that the limiting semigroup T is degenerate, i.e. $\operatorname{ran}(T(0))$ is a proper subspace of H. In this case $T(0)x = 0 \neq x = T_n(0)x$ for all $x \in \operatorname{ran}(T(0))^\perp \setminus \{0\}$. This explains why on $\operatorname{ran}(T(0))^\perp$ the convergence cannot be expected to be uniform on neighbourhoods of 0.

13.11 Example. A simple example for the phenomenon just mentioned is given by $H := \mathbb{R}$, $T_n(t) := \mathrm{e}^{-nt}$ ($t \geqslant 0$, $n \in \mathbb{N}$), $T(t) := 0$ ($t \geqslant 0$). △

For the proof of Theorem 13.10 we need another auxiliary result, Lemma 13.13 below, in which it will be convenient to use the following notation.

13.12 Remark. In Section 11.5 we have used the inequality sign between operators for inequalities in the sense of order in function spaces. In the theory of operators in a Hilbert space H it is also common to denote the accretivity of a symmetric operator A by writing '$A \geqslant 0$'. For self-adjoint operators $A, B \in \mathcal{L}(H)$ we will write '$A \leqslant B$' if $B - A$ is accretive, i.e.

$$(Ax \,|\, x) \leqslant (Bx \,|\, x) \qquad (x \in H).$$

The reader will have to be aware of this double meaning of the inequality sign; usually it will be clear from the context in which sense it should be interpreted. (In particular, in the context of a Hilbert space that is not a priori an L_2-space, the interpretation in the sense of order in function spaces would be meaningless.) △

13.13 Lemma. *Let H be a Hilbert space, and let T be a degenerate strongly continuous semigroup of self-adjoint contractions on H, with generator A. Then*

$$T(t) \leqslant (I - tA)^{-1} \qquad (t \geqslant 0).$$

Proof. As explained in Section 13.1, $P := T(0)$ is an orthogonal projection, and H decomposes as $H_\mathrm{a} \oplus H_\mathrm{u}$, where $H_\mathrm{a} = \operatorname{ran}(P)$, $H_\mathrm{u} = \ker(P)$, H_a is invariant under T and the restriction T_a of T to H_a is a C_0-semigroup, whereas $T(t)|_{H_\mathrm{u}} = 0$ $(t \geqslant 0)$. Moreover the generator A_a of T_a is self-adjoint, and $-A_\mathrm{a}$ is accretive.

Now we apply the spectral theorem for self-adjoint operators, Theorem 13.21, presented in Section 13.4 below, which states that A_a is unitarily equivalent to a maximal multiplication operator M_α in $L_2(\mu)$ for some semi-finite measure space $(\Omega, \mathcal{A}, \mu)$, with a measurable function $\alpha\colon \Omega \to \mathbb{R}$. Moreover $\alpha \leqslant 0$ a.e. since $-A_\mathrm{a}$ is accretive; see Remark 13.22, property (b).

Let $t \geqslant 0$. Then Exercise 13.4(b) shows that $T_\mathrm{a}(t)$ corresponds to the operator of multiplication by $\mathrm{e}^{t\alpha}$, and it is easy to see that $(I - tA_\mathrm{a})^{-1}$ corresponds to multiplication by $(1 - t\alpha)^{-1}$. From $\mathrm{e}^s \geqslant 1 + s$ $(s \geqslant 0)$ it follows that $\mathrm{e}^{t\alpha} \leqslant (1 - t\alpha)^{-1}$ a.e., and this inequality implies $T_\mathrm{a}(t) \leqslant (I - tA_\mathrm{a})^{-1}$, by property (d) of Remark 13.22. From $T(t) = PT_\mathrm{a}(t)P$ and $(I - tA)^{-1} = P(I - tA_\mathrm{a})^{-1}P$ we then obtain the assertion of the lemma. □

Proof of Theorem 13.10. The first assertion is immediate from Theorem 13.4.

For the second convergence we apply Lemma 13.13 to T_n and A_n and obtain

$$\|T_n(t)x\|^2 = (T_n(2t)x \,|\, x) \leqslant \big((I - 2tA_n)^{-1}x \,\big|\, x\big) = \tfrac{1}{2t}\big((\tfrac{1}{2t} - A_n)^{-1}x \,\big|\, x\big) \tag{13.9}$$

for all $x \in H$, $t > 0$. Now let $\varepsilon > 0$ and $x \in \ker(T(0))$. It follows from Lemma 13.9 that

$$(\lambda - A_n)^{-1}x \to (\lambda - A)^{-1}x = (\lambda - A)^{-1}T(0)x = 0$$

as $n \to \infty$, for all $\lambda > 0$. Therefore (13.9) implies that $\|T_n(\varepsilon)x\|^2 \to 0$, and by the contractivity of the semigroups T_n we conclude that $\sup_{t \geqslant \varepsilon} \|T_n(t)x\|^2 \to 0$. □

13.3 The Chernoff product formula

In the first part of this section X will be a Banach space.

13.14 Theorem (Chernoff product formula for degenerate strongly continuous semigroups). *Let T be a contractive degenerate strongly continuous semigroup on X, with generator A. Let $F\colon [0,\infty) \to \mathcal{L}(X)$ satisfy $F(0)T(0) = T(0)$ and $\|F(t)\| \leqslant 1$ for all $t \geqslant 0$. Define*

$$A(s) := \frac{1}{s}\big(F(s) - I\big) \qquad (s > 0),$$

and assume that there exists $\lambda > 0$ such that

$$R\big(\lambda, A(s)\big)x \to R(\lambda, A)x \qquad (s \to 0)$$

for all $x \in \operatorname{ran}(T(0))$. Then

$$F(t/n)^n x \to T(t)x, \tag{13.10}$$

uniformly for t in compact subsets of $[0,\infty)$, for all $x \in \operatorname{ran}(T(0))$.

13.15 Remarks. (a) The hypothesis '$F(0)T(0) = T(0)$' in Theorem 13.14 is a short way of expressing that $F(0)x = x$ for all $x \in \operatorname{ran}(T(0))$. Note that as in Theorem 13.4 there is no information on convergence on the complement of $\operatorname{ran}(T(0))$; see also Theorem 13.18.

(b) We mention that the hypotheses of Theorem 13.14 imply that

$$\left\|\mathrm{e}^{tA(s)}\right\| = \left\|\mathrm{e}^{(t/s)F(s)}\mathrm{e}^{-t/s}\right\| \leqslant \mathrm{e}^{(t/s)\|F(s)\|}\mathrm{e}^{-t/s} \leqslant 1$$

for all $t \geqslant 0$, $s > 0$; see Exercise 1.1(a). This explains why $\lambda > 0$ belongs to $\rho(A(s))$ and shows that $A(s)$ generates a contractive C_0-semigroup, for all $s > 0$. △

The following technical lemma contains a crucial estimate that will be used in the proof of Theorem 13.14.

13.16 Lemma. *Let $S \in \mathcal{L}(X)$ satisfy $\|S^m\| \leqslant M$ for some $M \geqslant 1$ and all $m \in \mathbb{N}$. Then*

$$\left\|\mathrm{e}^{n(S-I)}x - S^n x\right\| \leqslant \sqrt{n}\, M\|Sx - x\|$$

for all $n \in \mathbb{N}$, $x \in X$.

Proof. We will prove the asserted estimate in the form

$$\left\|\mathrm{e}^{nS}x - \mathrm{e}^n S^n x\right\| \leqslant \mathrm{e}^n\sqrt{n}\, M\|Sx - x\|.$$

First we show that

$$\left\|S^k x - S^m x\right\| \leqslant |k-m|\, M\|Sx - x\|$$

for all $k, m \in \mathbb{N}_0$, $x \in X$. Without loss of generality $k > m$. Then

$$\left\|\left(S^k - S^m\right)x\right\| = \left\|\sum_{j=m}^{k-1}\left(S^{j+1} - S^j\right)x\right\| \leqslant \sum_{j=m}^{k-1}\left\|S^j\right\|\|Sx - x\| \leqslant |k-m|\, M\|Sx - x\|.$$

We now estimate

$$\begin{aligned}\left\|\mathrm{e}^{nS}x - \mathrm{e}^n S^n x\right\| &= \left\|\sum_{k=0}^{\infty}\frac{n^k}{k!}\left(S^k - S^n\right)x\right\| \\ &\leqslant \left(\sum_{k=0}^{\infty}\left(\frac{n^k}{k!}\right)^{1/2}\left(\frac{n^k}{k!}\right)^{1/2}|k-n|\right) M\|Sx - x\| \\ &\leqslant \left(\sum_{k=0}^{\infty}\frac{n^k}{k!}\right)^{1/2}\left(\sum_{k=0}^{\infty}\frac{n^k}{k!}(k-n)^2\right)^{1/2} M\|Sx - x\|.\end{aligned}$$

Expanding $(k-n)^2$ one computes $\sum_{k=0}^{\infty}\frac{n^k}{k!}(k-n)^2 = n\mathrm{e}^n$; hence

$$\left\|\mathrm{e}^{nS}x - \mathrm{e}^n S^n x\right\| \leqslant (\mathrm{e}^n)^{1/2}(n\mathrm{e}^n)^{1/2}M\|Sx - x\| = \mathrm{e}^n\sqrt{n}\, M\|Sx - x\|.$$ □

Proof of Theorem 13.14. As before we write $X_{\mathrm{a}} := \operatorname{ran}(T(0))$ and denote by A_{a} the generator of the C_0-semigroup T_{a}.

Let $t_0 > 0$, $x \in X_{\rm a}$. The convergence (13.10) for $t = 0$ is trivial, so it remains to prove the uniform convergence (13.10) for $t \in (0, t_0]$. In view of Remark 13.15(b) we can apply Theorem 13.4 to obtain $\mathrm{e}^{tA(s)}x \to T(t)x$ as $s \to 0$, uniformly for t in compact subsets of $[0, \infty)$. From this convergence one easily infers that

$$\mathrm{e}^{tA(t/n)}x \to T(t)x \qquad (n \to \infty), \tag{13.11}$$

uniformly for $t \in (0, t_0]$.

Now let $x \in \mathrm{dom}(A_{\rm a})$. For $s > 0$ we put

$$x(s) := \big(\lambda - A(s)\big)^{-1}(\lambda - A_{\rm a})x.$$

Then $x(s) \to x$ and $A(s)x(s) = \lambda x(s) - (\lambda - A(s))x(s) \to \lambda x - (\lambda - A_{\rm a})x = A_{\rm a}x$ as $s \to 0$. Applying Lemma 13.16 with $S = F(t/n)$ and $M = 1$ we obtain

$$\begin{aligned}\big\|\big(\mathrm{e}^{tA(t/n)} - F(t/n)^n\big)x(t/n)\big\| &= \big\|\big(\mathrm{e}^{n(F(t/n)-I)} - F(t/n)^n\big)x(t/n)\big\| \\ &\leqslant \sqrt{n}\,\big\|\big(F(t/n) - I\big)x(t/n)\big\| \\ &= \frac{t}{\sqrt{n}}\|A(t/n)x(t/n)\| \to 0 \qquad (n \to \infty),\end{aligned}$$

uniformly for $t \in (0, t_0]$. Now observe that $\big\|\mathrm{e}^{tA(t/n)} - F(t/n)^n\big\| \leqslant 2$ for all $t > 0$, by Remark 13.15(b). We thus conclude that

$$\begin{aligned}&\big\|\big(\mathrm{e}^{tA(t/n)} - F(t/n)^n\big)x\big\| \\ &\quad\leqslant 2\|x - x(t/n)\| + \big\|\big(\mathrm{e}^{tA(t/n)} - F(t/n)^n\big)x(t/n)\big\| \to 0 \qquad (n \to \infty),\end{aligned} \tag{13.12}$$

uniformly for $t \in (0, t_0]$, and the denseness of $\mathrm{dom}(A_{\rm a})$ in $X_{\rm a}$ implies that the last property carries over to all $x \in X_{\rm a}$.

The convergences (13.11) and (13.12) together imply the assertion. □

The following examples serve to illustrate Theorems 13.4 and 13.14. In Exercise 13.3 the reader is asked to carry out the details.

13.17 Examples. Let T be a contractive C_0-semigroup on X, with generator A.

(a) Put $F(s) := T(s)$ and $A(s) := \frac{1}{s}(T(s) - I)$, for $s > 0$. Then $R\big(\lambda, A(s)\big) \to R(\lambda, A)$ strongly as $s \to 0$, for all $\lambda > 0$. Hence Theorem 13.4 implies that

$$T(t) = \text{s-lim}_{s\to 0}\, \mathrm{e}^{tA(s)} \qquad (t \geqslant 0),$$

and Theorem 13.14 yields the trivial 'convergence' $T(t) = \text{s-lim}_{n\to\infty} T(\frac{t}{n})^n$.

(b) Put $F(s) := (I - sA)^{-1}$ and $A(s) := \frac{1}{s}(F(s) - I) = A(I - sA)^{-1}$, for $s > 0$. Again one shows that $R\big(\lambda, A(s)\big) \to R(\lambda, A)$ strongly as $s \to 0$, for all $\lambda > 0$. Then Theorem 13.4 implies

$$T(t) = \text{s-lim}_{s\to 0}\, \mathrm{e}^{tA(I-sA)^{-1}} \qquad (t \geqslant 0). \tag{13.13}$$

(We point out that $A_n := A(\frac{1}{n})$ are the Yosida approximations of A and that (13.13) is the formula used in the proof of Theorem 2.9 for the definition of the semigroup T.) Remarkably, Theorem 13.14 yields the exponential formula of Theorem 2.12. △

For self-adjoint semigroups on Hilbert spaces one obtains the following stronger version of the Chernoff product formula, which will be important in the proof of the Trotter product formula for forms, the main result of Chapter 15.

13.18 Theorem (Chernoff product formula for self-adjoint degenerate strongly continuous semigroups). *Let T be a degenerate strongly continuous semigroup of self-adjoint contractions on a Hilbert space H, with generator A. Let $F\colon [0,\infty) \to \mathcal{L}(H)$ be a function taking its values in the accretive self-adjoint contractions, $F(0)T(0) = T(0)$. Define*

$$A(s) := \frac{1}{s}\big(F(s) - I\big) \qquad (s > 0),$$

and assume that $(A(s))_{s>0}$ converges to A in the strong resolvent sense as $s \to 0$.

Then

$$F(t/n)^n x \to T(t)x \qquad (n \to \infty)$$

uniformly for t in compact subsets of $[0,\infty)$, for all $x \in \operatorname{ran}(T(0))$, and

$$F(t/n)^n x \to 0 = T(t)x \qquad (n \to \infty)$$

uniformly for t in compact subsets of $(0,\infty)$, for all $x \in \ker(T(0)) = \operatorname{ran}(T(0))^\perp$. (Recall that $T(0)$ is a self-adjoint, therefore orthogonal, projection.)

13.19 Remark. Combining the two assertions of Theorem 13.18 one sees that the uniform convergence $F(t/n)^n x \to T(t)x$ for t in compact subsets of $(0,\infty)$ is valid for all $x \in H$. △

Proof of Theorem 13.18. The first statement is immediate from Theorem 13.14.

In the proof of the second assertion we use the spectral theorem for self-adjoint operators, Theorem 13.21, presented in the subsequent interlude. It implies that, for fixed $t > 0$, the self-adjoint operator $tA(t) = F(t) - I$ is unitarily equivalent to an operator M_α of multiplication by a measurable function α; by Remark 13.22, properties (a) and (b), we may assume that α takes its values in $[-1,0]$. Observe that $(1-s)^{-n} \geqslant (1+s)^n \geqslant 1+ns$ $(0 \leqslant s < 1)$ and thus $0 \leqslant (1+\alpha)^n \leqslant (1-n\alpha)^{-1}$, for all $n \in \mathbb{N}$. Applying Remark 13.22, property (d), we obtain

$$0 \leqslant F(t)^n = \big(I + tA(t)\big)^n \leqslant \big(I - ntA(t)\big)^{-1}.$$

Now let $x \in \ker(T(0))$. Then

$$\big\|F(\tfrac{t}{n})^n x\big\|^2 = \big(F(\tfrac{t}{n})^{2n} x \,\big|\, x\big) \leqslant \Big(\big(I - 2tA(\tfrac{t}{n})\big)^{-1} x \,\Big|\, x\Big) = \tfrac{1}{2t}\Big(\big(\tfrac{1}{2t} - A(\tfrac{t}{n})\big)^{-1} x \,\Big|\, x\Big)$$

for all $t > 0$. Similarly as in the proof of Theorem 13.10, Lemma 13.9 implies that $(\lambda - A(s))^{-1}x \to 0$ as $s \to 0$, uniformly for λ in compact subsets of $(0,\infty)$. From this convergence one easily concludes that $\frac{1}{2t}\big(\frac{1}{2t} - A(\frac{t}{n})\big)^{-1} x \to 0$ and thus $\|F(\frac{t}{n})^n x\|^2 \to 0$, uniformly for t in compact subsets of $(0,\infty)$. □

13.20 Remark. For later use (in Chapter 15) we note that in Theorems 13.14 and 13.18 the assertions remain true if the powers $F(\frac{t}{n})^n$ are replaced by $F(\frac{t}{n+1})^n$.

This is a consequence of the following elementary observation. Let $J = [0, \infty)$ or $J = (0, \infty)$, let $f\colon J \to X$, let (f_n) be a sequence of continuous functions $f_n\colon J \to X$, $f_n \to f$ locally uniformly, and let (r_n) be a sequence in $(0, \infty)$, $r_n \to 1$. Then $f_n(r_n\cdot) \to f$ locally uniformly. △

13.4 Interlude: the spectral theorem for self-adjoint operators

Multiplication operators, as described in Exercise 1.6, are particularly nice and simple operators. The spectral theorem for self-adjoint operators – in the version we present below – states that a self-adjoint operator is always unitarily equivalent to a maximal multiplication operator, as follows.

13.21 Theorem (Spectral theorem for self-adjoint operators). *Let H be a Hilbert space, and let A be a self-adjoint operator in H. Then there exist a semi-finite measure space $(\Omega, \mathcal{A}, \mu)$, a measurable function $\alpha\colon \Omega \to \mathbb{R}$ and a unitary operator $J\colon H \to L_2(\Omega, \mathcal{A}, \mu)$ such that*

$$A = J^{-1} M_\alpha J. \tag{13.14}$$

Concerning terminology, a measure space $(\Omega, \mathcal{A}, \mu)$ is called **semi-finite** if every set $A \in \mathcal{A}$ with $\mu(A) = \infty$ contains a set $B \in \mathcal{A}$ with $0 < \mu(B) < \infty$. In the statement of Theorem 13.21, M_α is the maximal multiplication operator described in Exercise 1.6(a), and (13.14) means in particular that $J(\mathrm{dom}(A)) = \mathrm{dom}(M_\alpha)$. We refer to Appendix F for the proof of this version of the spectral theorem.

The representation of A as a multiplication operator is far from unique. In contrast to this non-uniqueness, the subspaces

$$H_{[a,b]} := J^{-1}\bigl(\{\varphi \in L_2(\mu);\ \varphi = \mathbf{1}_{[a\leqslant\alpha\leqslant b]}\varphi\}\bigr) \qquad (a, b \in \mathbb{R},\ a \leqslant b) \tag{13.15}$$

of H do not depend on the measure space and the unitary operator J in Theorem 13.21; see Exercise F.7.

For the next definition recall that measurability of a mapping $f\colon \Omega \to \Omega'$ between two measurable spaces $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ means that for all $A' \in \mathcal{A}'$ the preimage $f^{-1}(A')$ belongs to $\mathcal{A}$. If Ω, Ω' are topological spaces and $\mathcal{A}, \mathcal{A}'$ are the respective Borel σ-algebras, then f is called **Borel measurable**. The mapping f is Borel measurable if and only if the preimage of each open set is a Borel set; cf. [Bau01; Theorem 7.2].

The spectral theorem can be used to define functions of self-adjoint operators: for any Borel measurable function $f\colon \mathbb{R} \to \mathbb{R}$ one puts

$$f(A) := J^{-1} M_{f\circ\alpha} J. \tag{13.16}$$

Observe that then $\mathbf{1}(A) = I$, $\mathrm{id}_{\mathbb{R}}(A) = A$. The uniqueness of the spaces $H_{[a,b]}$ in (13.15) implies that $f(A)$ does not depend on the measure space and the unitary operator J. The proof of this property is delegated to Exercise F.8.

13.22 Remark. In Exercise 13.5 the reader is asked to show that the semi-finiteness of the measure space $(\Omega, \mathcal{A}, \mu)$ in Theorem 13.21 implies the following properties.

(a) A is bounded if and only if $\alpha \in L_\infty(\mu)$; in this case $\|A\| = \|\alpha\|_\infty$.

(b) A is accretive if and only if $\alpha \geqslant 0$ a.e.

(c) $\lambda \in \mathbb{R}$ is an eigenvalue of A if and only if $\mu([\alpha = \lambda]) > 0$.

(d) If $f, g\colon \mathbb{R} \to \mathbb{R}$ are Borel measurable functions and $f \circ \alpha, g \circ \alpha \in L_\infty(\mu)$, then $f(A) \leqslant g(A)$ if and only if $f \circ \alpha \leqslant g \circ \alpha$ μ-a.e. (where the former '$\leqslant$' is meant as in Remark 13.12). △

If A is an accretive self-adjoint operator, represented as in Theorem 13.21, then – as explained in Remark 13.22, property (b) – we may assume that $\alpha \geqslant 0$, and therefore there exists a unique accretive self-adjoint square root $A^{1/2} := J^{-1} M_{\alpha^{1/2}} J$; cf. Exercise 13.6.

In part (a) of the next proposition we present an important relation between accretive symmetric forms and square roots of accretive self-adjoint operators, Kato's 'second representation theorem'; see [Kat80; Chap. VI, Theorem 2.23]. Part (b) contains a description of the form domain of an accretive self-adjoint operator that will be needed in the proof of the Trotter product formula in Chapter 15; see Lemma 15.8. We deal with an embedded form a in H, which means that $\operatorname{dom}(a)$ is a subset of H; see Section 12.2.

13.23 Proposition. *Let a be a closed accretive symmetric form in H with dense domain, and let A be the associated self-adjoint operator.*

(a) *Then* $\operatorname{dom}(a) = \operatorname{dom}(A^{1/2})$ *and*

$$a(u,v) = \big(A^{1/2}u \,\big|\, A^{1/2}v\big) \qquad (u, v \in \operatorname{dom}(a)).$$

(b) *Let T denote the C_0-semigroup generated by $-A$. Then the function*

$$(0,\infty) \ni s \mapsto \frac{1}{s}\Big(\big(I - T(s)\big)x \,\Big|\, x\Big) \in [0,\infty)$$

is decreasing for all $x \in H$,

$$\operatorname{dom}(a) = \Big\{x \in H\,;\ \sup_{s>0} \frac{1}{s}\Big(\big(I - T(s)\big)x \,\Big|\, x\Big) < \infty\Big\},$$

$$A^{1/2}x = \lim_{s\to 0+} \Big(\frac{1}{s}\big(I - T(s)\big)\Big)^{1/2} x \qquad \big(x \in \operatorname{dom}(a) = \operatorname{dom}(A^{1/2})\big),$$

and

$$a(x,y) = \lim_{s\to 0+} \frac{1}{s}\Big(\big(I - T(s)\big)x \,\Big|\, y\Big) \qquad (x, y \in \operatorname{dom}(a)).$$

Proof. In view of Theorem 13.21 and Remark 13.22 we may assume, without loss of generality, that $H = L_2(\mu)$ and $A = M_\alpha$, with $\alpha \geqslant 0$.

(a) It is easy to see that A is associated with the form b given by

$$b(f,g) = \int \alpha f \overline{g} \,\mathrm{d}\mu = \big(A^{1/2}f \,\big|\, A^{1/2}g\big)$$

on $\operatorname{dom}(b) = \big\{f \in L_2(\mu)\,;\ \alpha^{1/2}f \in L_2(\mu)\big\} = \operatorname{dom}(A^{1/2})$; cf. Example 5.9. The uniqueness stated in Corollary 12.9 shows that $a = b$.

(b) It follows from Exercise 1.6(b) that $T(s)$ acts as multiplication by $\mathrm{e}^{-s\alpha}$; hence

$$\frac{1}{s}\Big(\big(I - T(s)\big)f \,\Big|\, f\Big) = \frac{1}{s}\int_\Omega \big(1 - \mathrm{e}^{-s\alpha(\xi)}\big)|f(\xi)|^2 \,\mathrm{d}\mu(\xi) \qquad (f \in L_2(\mu),\ s > 0).$$

For all $\xi \in \Omega$, the function $(0,\infty) \ni s \mapsto \frac{1}{s}\big(1 - \mathrm{e}^{-s\alpha(\xi)}\big) \in [0,\infty)$ is decreasing – this yields the first assertion –, and $\lim_{s\to 0+} \frac{1}{s}\big(1 - \mathrm{e}^{-s\alpha(\xi)}\big) = \alpha(\xi)$. Therefore the monotone convergence theorem implies that $\sup_{s>0} \frac{1}{s}\big((I - T(s))f \,\big|\, f\big) < \infty$ if and only if $\alpha|f|^2 \in L_1(\mu)$, and the latter is equivalent to $f \in \mathrm{dom}(a)$.

Using the properties of the function $s \mapsto \frac{1}{s}\big(1 - \mathrm{e}^{-s\alpha(\xi)}\big)$ once more, one concludes for $f \in \mathrm{dom}(A^{1/2})$ that

$$\Big(\frac{1}{s}\big(I - T(s)\big)\Big)^{1/2} f = \Big(\frac{1}{s}\big(1 - \mathrm{e}^{-s\alpha}\big)\Big)^{1/2} f \to \alpha^{1/2} f = A^{1/2} f \qquad (s \to 0+),$$

by the dominated convergence theorem. The indicated formula for $a(x, y)$ is then immediate because

$$\Big(\big(I - T(s)\big)x \,\Big|\, y\Big) = \Big(\big(I - T(s)\big)^{1/2} x \,\Big|\, \big(I - T(s)\big)^{1/2} y\Big). \qquad \square$$

Notes

Degenerate strongly continuous semigroups have been introduced and studied in [ArBa93] and [Are01] under the name 'continuous degenerate semigroups'. We were tempted to simply call them 'strongly continuous semigroups' – because they are semigroups (according to our definition in Section 1.2) that are strongly continuous. However, this terminology would collide with the use in the literature.

There are other types of one-parameter semigroups, most notably holomorphic semigroups that are not strongly continuous at zero. This type is used in the theory of parabolic equations in non-reflexive spaces such as L_∞ and spaces of continuous functions; see [Lun95; Chapter 2 and Section 3.1.2], [ABHN11; Section 3.7 and Theorem 6.1.9]. (See also Exercise 13.7(c) for an example in a simple context.) For further types of semigroups that have linear relations (rather than operators) as generators we refer to [Haa06; Appendix A.8] and [FaYa98].

The Trotter approximation theorem, Theorem 13.4, appeared in [Tro58], and the Chernoff product formula, Theorem 13.14, was published in [Che68]; see also [Che74]. We have presented versions of these results for sequences of degenerate strongly continuous semigroups. This case was treated in [Are01] by Laplace transform methods; see [Are01; Theorem 4.2] for a variant of Theorem 13.4. The additional convergence in Theorem 13.10 for the self-adjoint case is generalised to sequences of holomorphic semigroups in [Are01; Theorem 5.2] (where the convergence of $T_n(t)x$ is uniform on compact subsets of $(0,\infty)$). The idea for the proof of Theorem 13.18 is due to Kato [Kat78; Section 3].

The spectral theorem for self-adjoint operators comes in different versions. A common formulation involves a spectral resolution $\mathbb{R} \ni \lambda \mapsto E(\lambda)$, an increasing function of orthogonal projections, which are in fact the orthogonal projections onto the spaces $H_{(-\infty,\lambda]} := \overline{\bigcup_{a<\lambda} H_{[a,\lambda]}}$ (see (13.15) for the definition of $H_{[a,\lambda]}$). The version presented in Section 13.4 is also well-established and very useful. For more information we refer to the Notes of Appendix F.

Exercises

13.1 Let X be a Banach space.

(a) Let $\rho \subseteq \mathbb{K}$, and let $R\colon \rho \to \mathcal{L}(X)$ be a pseudo-resolvent, i.e.

$$R(\lambda) - R(\mu) = (\mu - \lambda)R(\mu)R(\lambda) \qquad (\lambda, \mu \in \rho).$$

Show that R is continuous. Show further that there exists a linear relation $A \subseteq X \times X$ such that $R(\lambda) = (\lambda I - A)^{-1}$ for all $\lambda \in \rho$.

Hint for the first assertion: Neumann series, proof of Theorem 2.2(c). Hints for the second assertion: (1) Let $R_1, R_2 \in \mathcal{L}(X)$, $\kappa \in \mathbb{K}$. Show that

$$R_1 - R_2 = \kappa R_1 R_2 \qquad \Longleftrightarrow \qquad R_2^{-1} \subseteq R_1^{-1} + \kappa I. \tag{13.17}$$

(2) Apply (1) with $\lambda, \mu \in \rho$, $R_1 := R(\lambda)$, $R_2 := R(\mu)$, $\kappa := \mu - \lambda$. (3) Define $A := -R(\lambda)^{-1} + \lambda I$ $(\lambda \in \rho)$.

(b) Let $A \subseteq X \times X$ be a linear relation, and define $\rho(A) := \{\lambda \in \mathbb{K};\, (\lambda - A)^{-1} \in \mathcal{L}(X)\}$. Show that $\rho(A) \ni \lambda \mapsto R(\lambda, A) := (\lambda - A)^{-1} \in \mathcal{L}(X)$ is a pseudo-resolvent. Show further that $\rho(A)$ is open.

Hint for the first assertion: Use (13.17). Hint for the second assertion: See the proof of Theorem 2.2(c); for $\lambda \in \rho(A)$ and $\mu \in \mathbb{K}$ close to λ, apply part (a) with $\rho := \{\lambda, \mu\}$.

13.2 Let X be a Banach space and $X_1 \subseteq X$ a complemented closed subspace, i.e., there exists a bounded projection P from X onto X_1. Let A_1 be a closed operator in X_1, and define

$$A := \bigl\{(x, y) \in X \times X\,;\, (x, Py) \in A_1\bigr\}.$$

Show that A_1 is invertible in $\mathcal{L}(X_1)$ if and only if A is invertible in $\mathcal{L}(X)$, and that then $A^{-1} = A_1^{-1}P \in \mathcal{L}(X)$.

13.3 Prove the assertions stated in Examples 13.17.

13.4 (a) Let X_1, X_2 be Banach spaces, and let $J \in \mathcal{L}(X_1, X_2)$ be an isomorphism. For $j = 1, 2$ let A_j be the generator of a C_0-semigroup T_j on X_j. In addition suppose that $A_2 = JA_1J^{-1}$. Show that $T_2(t) = JT_1(t)J^{-1}$ for all $t \geqslant 0$. (Hint: Determine the generator of the C_0-semigroup $\bigl(JT_1(t)J^{-1}\bigr)_{t\geqslant 0}$ on X_2 and recall Theorem 1.13(b).)

(b) Let A be an accretive self-adjoint operator in a Hilbert space H, and let $(\Omega, \mathcal{A}, \mu)$, α, J be as in Theorem 13.21. Show that the C_0-semigroup on H generated by $-A$ corresponds to the C_0-semigroup $(M_{\mathrm{e}^{-t\alpha}})_{t\geqslant 0}$ on $L_2(\mu)$. (Hint: Recall Exercise 1.6(b), and use part (a) above.)

13.5 Prove the properties stated in Remark 13.22. First treat the special case in which $A = M_\alpha$ for some semi-finite measure space $(\Omega, \mathcal{A}, \mu)$ and a measurable function $\alpha\colon \Omega \to \mathbb{R}$. Then convince yourself that the general case follows from this special case.

13.6 Let A and B be accretive self-adjoint operators in H, $B^2 = A$. Show that $B = A^{1/2}$. (Hint: Start with a representation of B as a multiplication operator.)

13.7 Define the operator A in the Banach space ℓ_∞ by $A(x_n) := (-nx_n)$, for $x = (x_n) \in \operatorname{dom}(A) := \{x \in \ell_\infty;\ (nx_n) \in \ell_\infty\}$. For $t \in \mathbb{K}$, $\operatorname{Re} t \geqslant 0$ define $T(t) \in \mathcal{L}(\ell_\infty)$ by $T(t)(x_n) := (\mathrm{e}^{-tn}x_n)$. Prove the following properties.

(a) $[\operatorname{Re} > 0] \subseteq \rho(A)$, and $\|(\lambda - A)^{-1}\| \leqslant \frac{1}{|\lambda|}$ for all $\operatorname{Re}\lambda > 0$.

(b) $(T(t))_{t\geqslant 0}$ is a contractive one-parameter semigroup on ℓ_∞, strongly continuous on $(0,\infty)$, $T(t)(\ell_\infty) \subseteq c_0$ $(t > 0)$, and $\{x \in \ell_\infty;\ x = \lim_{t\to 0+} T(t)x\} = c_0$ (the space of null sequences). In particular, the restriction T_0 of T to c_0 is a C_0-semigroup on c_0.

(c) In the complex case, T is a contractive holomorphic semigroup of angle $\pi/2$. The restriction of T to c_0 is a holomorphic C_0-semigroup.

(d) $\overline{\operatorname{dom}(A)} = c_0$, and $A_0 := A \cap (c_0 \times c_0)$ is the generator of T_0.

(e) A does not generate a degenerate strongly continuous semigroup on ℓ_∞.

Note. The operator A in this exercise is a special type of Hille–Yosida operator; see Exercise 13.8.

13.8 An operator A in a Banach space X is called a **Hille–Yosida operator** if there exist $\omega \in \mathbb{R}$, $M \geqslant 1$ such that $(\omega,\infty) \subseteq \rho(A)$ and

$$\big\|(\lambda - A)^{-n}\big\| \leqslant \frac{M}{(\lambda-\omega)^n} \qquad (\lambda > \omega,\ n \in \mathbb{N}).$$

For such operators, the part $A_0 := A \cap (X_0 \times X_0)$ of A in $X_0 := \overline{\operatorname{dom}(A)}$ generates a C_0-semigroup T_0 on X_0; see [EnNa00; Chap. II, Corollary 3.21]. The operator A in Exercise 13.7 is a Hille–Yosida operator (with $\omega = 0$, $M = 1$). In that exercise there exists an extension of T_0 to a one-parameter semigroup T on all of X, satisfying

(i) $T(t)|_{X_0} = T_0(t) \quad (t \geqslant 0)$,

(ii) $T(t)(\lambda - A)^{-1}x = (\lambda - A)^{-1}T(t)x \quad (x \in X,\ t \geqslant 0)$.

The present exercise shows that this does not hold for Hille–Yosida operators in general.

On $X := C[0,1]$ we define the operator A by $\operatorname{dom}(A) := \{f \in C^1[0,1];\ f(0) = 0\}$, $Af := -f'$ $(f \in \operatorname{dom}(A))$. Prove the following properties.

(a) $(0,\infty) \subseteq \rho(A)$, $\|(\lambda - A)^{-1}\| \leqslant \frac{1}{\lambda}$ $(\lambda > 0)$ (i.e. A is a Hille–Yosida operator).

(b) $X_0 := \overline{\operatorname{dom}(A)} = \{f \in C[0,1];\ f(0) = 0\} \cong C_0(0,1]$. Note that, by Exercise 1.5(c), the C_0-semigroup T_0 on X_0 generated by the operator A_0 defined above is the semigroup of right translations.

(c) There exists no one-parameter semigroup T on X with the above properties (i), (ii). (Hint: These properties imply that $\operatorname{dom}(A)$ is invariant under T_0.)

(d) There exists a bounded projection $P\colon C[0,1] \to C_0(0,1]$, $Pf := f - f(0)\mathbf{1}$ $(f \in C[0,1])$. Then $\widetilde{T}(t) := T_0(t)P$ $(t \geqslant 0)$ defines a degenerate strongly continuous semigroup on $C[0,1]$; determine its generator.

Note. Combining Exercises 2.7 and 2.5 one concludes that a Hille–Yosida operator in a reflexive Banach space X is always the generator of a C_0-semigroup on X; examples as in Exercise 13.7 and in the present exercise are only possible in non-reflexive Banach spaces.

Chapter 14

Form convergence theorems

The objective of this chapter is to show that convergence of a sequence of forms – suitably defined – implies strong resolvent convergence of the sequence of the associated operators. This supplements the central result of the previous chapter, the Trotter approximation theorem (Theorem 13.4), where strong resolvent convergence of a sequence of semigroup generators was shown to imply the convergence of the corresponding semigroups. These results are of considerable interest for differential operators, but they will also play an important role in the proof of the Trotter product formula for forms, presented in Chapter 15. We start by investigating non-densely defined forms and their relation to degenerate strongly continuous semigroups.

14.1 Non-densely defined forms, m-sectorial and self-adjoint linear relations

In Chapter 13 we discussed degenerate strongly continuous semigroups and their generators. In the present section we describe how such generators can be obtained as linear relations associated with forms, starting from the setting of Section 5.3.

Let V, H be Hilbert spaces over $\mathbb{K}$, $j \in \mathcal{L}(V,H)$, and let $a\colon V \times V \to \mathbb{K}$ be a bounded j-coercive form. Clearly, $j\colon V \to H_{\mathrm{a}} := \overline{\operatorname{ran}(j)}$ has dense range. Let A_{a} denote the quasi-m-accretive operator in H_{a} associated with (a,j); see Corollary 5.11. Then $-A_{\mathrm{a}}$ generates a quasi-contractive C_0-semigroup T_{a} on H_{a}. Let P be the orthogonal projection in H onto H_{a}. Then

$$T(t) := T_{\mathrm{a}}(t)P \qquad (t \geqslant 0)$$

defines a quasi-contractive degenerate strongly continuous semigroup on H (with 'active' subspace H_{a}); we say that T is **associated with** (a,j). If $\mathbb{K} = \mathbb{C}$, then A_{a} is quasi-m-sectorial, and one obtains a holomorphic extension of T by putting $T(z) := T_{\mathrm{a}}(z)P$ for $z \in \Sigma_{\theta,0}$, with suitable $\theta \in (0,\pi/2]$.

Let $-A$ be the generator of T; then $A = A_{\mathrm{a}} \oplus \bigl(\{0\} \times H_{\mathrm{a}}^{\perp}\bigr)$ by (13.3). (Notice the change of the sign in the generator with respect to Section 13.1.) In Exercise 14.1(a) the reader is asked to show that A is quasi-m-accretive and that

$$A = \bigl\{(x,y) \in H \times H\,;\, \exists u \in V\colon j(u) = x,\ a(u,v) = (y \mid j(v))_H\ (v \in V)\bigr\}. \tag{14.1}$$

We call A the **linear relation associated with** (a,j). (The formula for the associated linear relation A in H is the same as in the case of densely defined forms; see (5.6).)

If $\operatorname{dom}(a)$ is a subset of H and $j\colon \operatorname{dom}(a) \hookrightarrow H$ is the embedding, then a is an embedded form, and we will suppress j in the notation, as before.

We illustrate these notions by a simple example.

14.1 Example. Let H_{a} be a closed subspace of H, and define the (embedded) form a by $a(u,v) := 0$ for all $u, v \in \operatorname{dom}(a) := H_{\mathrm{a}}$. Then one obtains

$$A = \bigl\{(u,v) \in H_{\mathrm{a}} \times H\,;\, 0 = (y \,|\, v)_H \ (v \in H_{\mathrm{a}})\bigr\} = H_{\mathrm{a}} \times H_{\mathrm{a}}^{\perp},$$

and the associated degenerate strongly continuous semigroup is given by $T(t) = P$ $(t \geqslant 0)$, where P is the orthogonal projection onto H_{a}. △

Now we turn to the setting of Section 12.1. Let H be a complex Hilbert space, and let (a,j) be a (not necessarily densely defined) quasi-sectorial form in H. Then (a,j) is a densely defined quasi-sectorial form in $H_{\mathrm{a}} := \overline{\operatorname{ran}(j)}$, and hence it is associated with a quasi-m-sectorial operator in H_{a}, which we denote by A_{a}. We recall the construction from Section 12.1: let (V,q) denote the completion of $(\operatorname{dom}(a), (\cdot \,|\, \cdot)_{a_\omega,j})$, where $\omega \in \mathbb{R}$ is chosen such that the shifted form

$$a_\omega(u,v) = a(u,v) + \omega(j(u) \,|\, j(v)) \qquad (u, v \in \operatorname{dom}(a))$$

is sectorial. Let $\tilde{\jmath}, \tilde{a}$ be the 'extensions' of j, a as in Remark 12.3(c). Then $\tilde{a}$ is a bounded $\tilde{\jmath}$-coercive form, we are back in the previous setting, and the formula (14.1) for the associated linear relation reads

$$A = \bigl\{(x,y) \in H \times H\,;\, \exists u \in V\colon \tilde{\jmath}(u) = x,\ \tilde{a}(u,v) = (y \,|\, \tilde{\jmath}(v))_H \ (v \in V)\bigr\}; \tag{14.2}$$

as before we call A the **linear relation associated with** (a,j).

14.2 Remark. Let (a,j) be a quasi-sectorial form with vertex $-\omega \in \mathbb{R}$, and let A be the linear relation defined in (14.2). Then as in Remark 12.5 it follows that the linear relation

$$A + \omega I = \{(x, y + \omega x)\,;\, (x,y) \in A\}$$

is associated with the shifted form (a_ω, j). Recall that $-(A + \omega I)$ generates the rescaled semigroup $t \mapsto \mathrm{e}^{-\omega t} T(t)$. △

A linear relation A in H is called **sectorial** if there exists $\theta \in [0, \pi/2)$ such that $(y \,|\, x) \in \overline{\Sigma_\theta}$ for all $(x,y) \in A$, and A is **m-sectorial** if additionally $\operatorname{ran}(I + A) = H$. A linear relation A is called **quasi-sectorial** (**quasi-m-sectorial**) if there exists $\omega \in \mathbb{R}$ such that $A + \omega I$ is sectorial (m-sectorial). If, as above, (a,j) is a (quasi-)sectorial form in H, then the linear relation A in H associated with (a,j) by (14.2) is (quasi-)m-sectorial; see Exercise 14.1(b).

We now return to the general case of a Hilbert space H over $\mathbb{K}$ and consider the special case of an accretive symmetric form (a,j) in H. For symmetric forms the above definitions of the associated degenerate strongly continuous semigroup T and of the associated linear relation A also apply in real Hilbert spaces. Since (a,j) is symmetric, the associated operator A_{a} in $H_{\mathrm{a}} = \overline{\operatorname{ran}(j)}$ is self-adjoint, by Theorem 12.4, and as a consequence the generator $-A$ of the associated degenerate strongly continuous semigroup will also be self-adjoint, in the sense treated subsequently (see Proposition 14.4(a) below).

If A is a linear relation in H, then the **adjoint** A^* of A is defined in the same way as in Section 6.1 as

$$A^* := \big((-A)^\perp\big)^{-1},$$

and A is called **self-adjoint** if $A^* = A$. (Recall that $-A = \{(x,-y);\ (x,y) \in A\}$; see (13.2).) We emphasise that the results on adjoints and on self-adjoint operators presented in Section 6.1 carry over appropriately to linear relations; cf. Exercises 14.8 and 14.9(c). It is clear from the definition that a linear relation A is self-adjoint if and only if its inverse A^{-1} is self-adjoint.

14.3 Remarks. (a) An "extreme" example of a self-adjoint linear relation is $A := \{0\} \times H$; this is the only self-adjoint linear relation with $\operatorname{dom}(A) = \{0\}$.

(b) Let H_0 and H_1 be Hilbert spaces, and for $j = 0,1$ let A_j be a linear relation in H_j. Then the linear relation

$$A_0 \oplus A_1 := \big\{\big((x_0,x_1),(y_0,y_1)\big)\,;\ (x_j,y_j) \in A_j\ (j=0,1)\big\}$$

in $H_0 \oplus H_1$ is self-adjoint if and only if A_0 and A_1 are self-adjoint. △

Note that a self-adjoint relation A is accretive if and only if $(x \,|\, y) \geqslant 0$ for all $(x,y) \in A$. Clearly, A is accretive if and only if A^{-1} is accretive. We have the following correspondence between accretive symmetric forms and accretive self-adjoint linear relations.

14.4 Proposition. *Let H be a Hilbert space.*

(a) *Let (a,j) be an accretive symmetric form in H. Then the linear relation A associated with (a,j), given by* (14.2)*, is accretive and self-adjoint.*

(b) *Conversely, if A is an accretive self-adjoint linear relation in H, then there exists a unique embedded closed accretive symmetric form a in H such that A is associated with a.*

Proof. (a) follows from the formula $A = A_{\mathrm a} \oplus \big(\{0\} \times H_{\mathrm a}^\perp\big)$ and Remark 14.3.

(b) The linear relation A is m-accretive by Exercise 14.8(c) and hence decomposes as $A = A_{\mathrm a} \oplus A_{\mathrm u}$ as in Theorem 13.2(a). Moreover $A_{\mathrm a}$ is self-adjoint by Remark 14.3(b). Then $A_{\mathrm a}$ is associated with a densely defined embedded closed accretive symmetric form a in $H_{\mathrm a}$, and interpreting a as a form in H one obtains A as the associated operator.

Concerning uniqueness, if A is associated with a form a, then $\overline{\operatorname{dom}(a)} = \overline{\operatorname{dom}(A_{\mathrm a})}$, where $A_{\mathrm a}$ is as above, and then the uniqueness of a follows from Corollary 12.9. □

14.2 Form convergence for increasing sequences

Throughout this section we consider *embedded* closed accretive symmetric forms, which means that we are in the context of Proposition 14.4(b). The main result, Theorem 14.10, deals with convergence of an increasing sequence of symmetric forms. (We refer to the Notes for generalisations to non-symmetric forms.) In order to explain what 'increasing' means, we first define an order relation on the set of accretive symmetric forms, and we characterise it via the inverses of the associated linear relations.

Let H be a Hilbert space. Given two (embedded) accretive symmetric forms a and b in H, we say that $a \leqslant b$ if $\operatorname{dom}(a) \supseteq \operatorname{dom}(b)$ and $a(x) \leqslant b(x)$ for all $x \in \operatorname{dom}(b)$. For the definition of inequalities between bounded self-adjoint operators we refer to Remark 13.12. We recall that a linear relation A in H is self-adjoint if and only if A^{-1} is self-adjoint.

14.5 Proposition. *Let a, b be closed accretive symmetric forms in H. Let A, B be the associated self-adjoint linear relations, and assume that the inverses of A and B are operators belonging to $\mathcal{L}(H)$. Then $a \leqslant b$ if and only if $B^{-1} \leqslant A^{-1}$.*

Our proof of this equivalence is quite different from earlier proofs that can be found in the literature; see e.g. [Kun05; Proposition 2.7], [HSSW06; Lemmas 3.2 and 3.4]. The key observation for our treatment is Proposition 14.7 below, whose proof uses the following elementary lemma.

14.6 Lemma. *Let X be a normed space, H a Hilbert space and $P \in \mathcal{L}(X, H)$. Let $\eta \in X'$, $c \geqslant 0$ be such that*

$$|\eta(x)| \leqslant c\|Px\| \qquad (x \in X). \tag{14.3}$$

Then there exists $z \in H$ such that $\|z\| \leqslant c$ and $\eta(x) = (Px \,|\, z)$ for all $x \in X$.

Proof. Define $\tilde{\eta}\colon \operatorname{ran}(P) \to \mathbb{K}$ by $\tilde{\eta}(Px) := \eta(x)$ for all $x \in X$; observe that by (14.3), $\tilde{\eta}$ is well-defined and continuous on $\overline{\operatorname{ran}(P)}$, $\|\tilde{\eta}\| \leqslant c$. The Fréchet–Riesz representation theorem implies that there exists $z \in \overline{\operatorname{ran}(P)}$ such that $\|z\| \leqslant c$ and $\eta(x) = \tilde{\eta}(Px) = (Px \,|\, z)$ for all $x \in X$. □

Assuming that G and H are Hilbert spaces and that C and D are linear relations in $G \times H$, we will say that D **dominates** C if for all $(x, y) \in D$ there exists $z \in H$ such that $(x, z) \in C$ and $\|z\| \leqslant \|y\|$. If C and D are operators, this simply means that $\operatorname{dom}(D) \subseteq \operatorname{dom}(C)$ and $\|Cx\| \leqslant \|Dx\|$ for all $x \in \operatorname{dom}(D)$. The following fundamental property concerning this notion is a more elaborate version of [Kat80; Chap. VI, Lemma 2.30].

14.7 Proposition. *Let G, H be Hilbert spaces, and let C, D be closed linear relations in $G \times H$. Then D dominates C if and only if $C^{\perp}$ dominates $D^{\perp}$.*

Proof. It clearly suffices to prove '⇒'. Let $(x, y) \in C^{\perp}$.

Let $(f, g) \in D$. By hypothesis there exists $h \in H$ such that $(f, h) \in C$ and $\|h\| \leqslant \|g\|$. Then $(f, h) \perp (x, y)$, hence

$$|(-f \,|\, x)| = |(h \,|\, y)| \leqslant \|h\|\|y\| \leqslant \|g\|\|y\|;$$

note that h has dropped out of this inequality. With $P\colon D \to H$, $(f, g) \mapsto g$ and $\eta\colon D \to \mathbb{K}$, $(f, g) \mapsto (-f \,|\, x)$ it follows that $\big|\eta\big((f, g)\big)\big| \leqslant \|y\|\|P(f, g)\|$. We can now apply Lemma 14.6 to obtain $z \in H$ such that $\|z\| \leqslant \|y\|$ and $(-f \,|\, x) = (P(f, g) \,|\, z) = (g \,|\, z)$ for all $(f, g) \in D$, i.e., $(x, z) \in D^{\perp}$ and $\|z\| \leqslant \|y\|$.

To summarise, we have shown that $C^{\perp}$ dominates $D^{\perp}$. □

Proof of Proposition 14.5. Let A_a be the accretive self-adjoint operator in $\overline{\mathrm{dom}(a)}$ associated with a, and denote by P_a the orthogonal projection onto $\overline{\mathrm{dom}(a)}$. Note that A_a^{-1} is the restriction of A^{-1} to $\overline{\mathrm{dom}(a)}$. The (accretive self-adjoint) square root of A^{-1} will be denoted by $A^{-1/2}$; see Section 13.4 for the definition of the square root of accretive self-adjoint operators. We point out that $A_\mathrm{a}^{-1/2}$, the square root of A_a^{-1}, is the restriction of $A^{-1/2}$ to $\overline{\mathrm{dom}(a)}$, and we define

$$A^{1/2} := (A^{-1/2})^{-1} = \big\{(x,y) \in H \times H\,;\ (x, P_a y) \in A_\mathrm{a}^{1/2}\big\}.$$

The corresponding notation and properties will also be used for b.

By Proposition 13.23(a), $a \leqslant b$ if and only if $B_\mathrm{a}^{1/2}$ dominates $A_\mathrm{a}^{1/2}$. The latter, in turn, holds if and only if $B^{1/2}$ dominates $A^{1/2}$. (Clearly $A_\mathrm{a}^{1/2}$ dominates $A^{1/2}$, and the converse is true as well: if $(x,y) \in A^{1/2}$, then $(x, P_a y) \in A_\mathrm{a}^{1/2}$ and $\|P_a y\| \leqslant \|y\|$. By the same token, $B_\mathrm{a}^{1/2}$ and $B^{1/2}$ dominate each other. As it is easily seen that domination is transitive, one concludes the last assertion above.)

On the other hand, $A^{-1} \geqslant B^{-1}$ if and only if $A^{-1/2}$ dominates $B^{-1/2}$.

Finally we observe that

$$(A^{-1/2})^\perp = \big((-A^{-1/2})^*\big)^{-1} = -(A^{-1/2})^{-1} = -A^{1/2}$$

by the self-adjointness of $A^{-1/2}$, and similarly $(B^{-1/2})^\perp = -B^{1/2}$. Now, applying Proposition 14.7 we conclude that $B^{1/2}$ dominates $A^{1/2}$ if and only if $A^{-1/2}$ dominates $B^{-1/2}$.

This proves the asserted equivalence. □

For the proof of the form convergence theorem announced above we also need the next proposition on decreasing sequences of bounded accretive self-adjoint operators. We refer to Corollary 14.21 for a form version – incomparably more general – of this result.

14.8 Proposition. *Let (B_n) be a decreasing sequence of accretive self-adjoint operators in $\mathcal{L}(H)$. Then $B := \text{s-lim}_{n\to\infty} B_n$ exists, and B is an accretive self-adjoint operator.*

Proof. For $x \in H$, the sequence $\big((B_n x \,|\, x)\big)_n$ is decreasing and bounded below by 0, hence convergent. The polarisation identity – see Remark 5.1 – implies that $\big((B_n x \,|\, y)\big)_n$ is convergent for all $x, y \in H$. It follows from Exercise 5.1 that there exists $B \in \mathcal{L}(H)$ such that $(Bx \,|\, y) = \lim_{n\to\infty}(B_n x \,|\, y)$ for all $x, y \in H$. Then it is straightforward that B is self-adjoint and accretive. Finally, by the subsequent Lemma 14.9 we obtain $\|(B_n - B)x\|^2 \leqslant \|B_1 - B\|\big((B_n - B)x \,\big|\, x\big) \to 0$ as $n \to \infty$, for all $x \in H$. □

14.9 Lemma. *Let $A, B \in \mathcal{L}(H)$ be self-adjoint and accretive, $A \leqslant B$. Then $\|Ax\|^2 \leqslant \|B\|(Ax \,|\, x)$ for all $x \in H$.*

Proof. Let $x \in H$. Then by Lemma 5.2 (Cauchy–Schwarz) we obtain

$$|(Ax \,|\, y)|^2 \leqslant (Ax \,|\, x)(Ay \,|\, y) \leqslant (Ax \,|\, x)(By \,|\, y) \leqslant (Ax \,|\, x)\|B\|\|y\|^2 \qquad (y \in H),$$

which for $y = Ax$ implies the assertion. □

We now state the main result of this section.

14.10 Theorem. *Let $(a_n)_{n\in\mathbb{N}}$ be an increasing sequence of closed accretive symmetric forms in H. Then*

$$\operatorname{dom}(a) := \Bigl\{x \in \bigcap_{n\in\mathbb{N}} \operatorname{dom}(a_n)\,;\ \sup_{n\in\mathbb{N}} a_n(x) < \infty\Bigr\}$$

is a vector space,

$$a(x,y) := \lim_{n\to\infty} a_n(x,y) \tag{14.4}$$

exists for all $x, y \in \operatorname{dom}(a)$, and a thus defined is a closed accretive symmetric form in H.

Let A be the self-adjoint linear relation associated with a, and let A_n be the self-adjoint linear relation associated with a_n, for $n \in \mathbb{N}$. Then (A_n) converges to A in the strong resolvent sense, i.e. $(\lambda + A_n)^{-1} \to (\lambda + A)^{-1}$ strongly as $n \to \infty$, for all $\lambda > 0$.

14.11 Remarks. (a) The assertion in Theorem 14.10 is precisely the hypothesis for the linear relations $-A_n$ and $-A$ in Theorem 13.10. This has the important consequence that the degenerate strongly continuous semigroups generated by $-A_n$ converge to the semigroup generated by $-A$, in the sense formulated in Theorem 13.10.

(b) In Theorem 14.10, the closedness hypothesis for the forms a_n is essential. We refer to [BaEl14; Example 3.6] for a counterexample.

(c) In the context of Theorem 14.10 it is rather natural to work with forms that are not necessarily densely defined: even if all the a_n are densely defined, the limiting form a need not be densely defined. A simple example is given by $H := \mathbb{R}$, $a_n(x,y) := nxy$ ($n \in \mathbb{N}$, $x, y \in \mathbb{R}$), which leads to a having trivial domain $\operatorname{dom}(a) = \{0\}$. △

Proof of Theorem 14.10. If $x, y \in \operatorname{dom}(a)$, then the Cauchy–Schwarz inequality, Proposition 5.2, implies that $\bigl(a_n(x+y)\bigr)_{n\in\mathbb{N}}$ is bounded, hence $x + y \in \operatorname{dom}(a)$. As it is obvious that $\lambda x \in \operatorname{dom}(a)$ for all $x \in \operatorname{dom}(a)$, $\lambda \in \mathbb{K}$, it follows that $\operatorname{dom}(a)$ is a vector space. The polarisation identity (see Remark 5.1) implies that the limit (14.4) exists for all $x, y \in \operatorname{dom}(a)$.

Clearly a is an accretive symmetric form. We show that it is closed. Let $(x_n)_{n\in\mathbb{N}}$ be a $\|\cdot\|_a$-Cauchy sequence. Then $x := \lim x_n$ exists in H. Also, (x_n) is a $\|\cdot\|_{a_m}$-Cauchy sequence for all $m \in \mathbb{N}$, and therefore $x_n \to x$ in $\bigl(\operatorname{dom}(a_m), \|\cdot\|_{a_m}\bigr)$ since a_m is closed. Let $n \in \mathbb{N}$. Then

$$\sup_{m\in\mathbb{N}} a_m(x - x_n) = \sup_{m\in\mathbb{N}} \lim_{k\to\infty} a_m(x_k - x_n) \leqslant \sup_{m\in\mathbb{N},\,k\geqslant n} a_m(x_k - x_n) = \sup_{k\geqslant n} a(x_k - x_n) < \infty.$$

This inequality implies that $x \in \operatorname{dom}(a)$, and

$$a(x - x_n) \leqslant \sup_{k\geqslant n} a(x_k - x_n) \to 0 \qquad (n \to \infty)$$

shows that $x_n \to x$ in the $\|\cdot\|_a$-norm. It follows that a is closed.

In view of Lemma 13.9 it is sufficient to prove the asserted strong convergence for $\lambda = 1$. Recall from Remark 14.2 that the linear relation $I + A_n$ is associated with the shifted form $e + a_n$, for $n \in \mathbb{N}$, where e is the 'unit form', $e(x,y) := (x \,|\, y)$ $(x, y \in H)$. Similarly, $I + A$ is associated with $e + a$. Now Proposition 14.5 implies that $\bigl((I + A_n)^{-1}\bigr)_n$ is decreasing and that $(I + A_n)^{-1} \geqslant (I + A)^{-1}$ for all $n \in \mathbb{N}$. From Proposition 14.8 we infer that $R := \text{s-lim}_{n\to\infty}(I + A_n)^{-1}$ exists and is an accretive self-adjoint operator; clearly $(I + A_n)^{-1} \geqslant R \geqslant (I + A)^{-1}$ for all $n \in \mathbb{N}$. Then $B := R^{-1}$ is an accretive self-adjoint linear relation; let b denote the corresponding closed accretive symmetric form as in Proposition 14.4(b). Using Proposition 14.5 we infer that $e + a_n \leqslant b \leqslant e + a$ for all $n \in \mathbb{N}$. Thus from the definition of a we obtain $b = e + a$. Since B is associated with b and $I + A$ is associated with $e + a$ we conclude that $(I + A)^{-1} = B^{-1} = R = \text{s-lim}_{n\to\infty}(I + A_n)^{-1}$. □

14.12 Example. Let $C \subseteq \mathbb{R}^n$ be a closed set, and let (Ω_k) be a decreasing sequence of open subsets of $\mathbb{R}^n$ such that $C = \bigcap_{k=1}^{\infty} \Omega_k$.

For $k \in \mathbb{N}$ let a_k be the classical Dirichlet form with domain $\operatorname{dom}(a_k) = H_0^1(\Omega_k)$,

$$a_k(u, v) = \int \nabla u \cdot \overline{\nabla v} \, \mathrm{d}x \qquad (u, v \in \operatorname{dom}(a_k)).$$

We consider all the forms a_k as forms in $L_2(\mathbb{R}^n)$. Then (a_k) is an increasing sequence of closed accretive symmetric forms, hence convergent in the sense of Theorem 14.10. It is not difficult to see that the limiting form a is given by

$$\operatorname{dom}(a) = H_0^1(C) := \big\{u \in H^1(\mathbb{R}^n)\,;\, u|_{\mathbb{R}^n\setminus C} = 0\big\} = \bigcap_{k\in\mathbb{N}} H_0^1(\Omega_k),$$
$$a(u, v) = \int \nabla u \cdot \overline{\nabla v} \, \mathrm{d}x \qquad (u, v \in \operatorname{dom}(a)).$$

For the degenerate strongly continuous semigroups T_k and T on $L_2(\mathbb{R}^n)$ associated with the forms a_k and a, respectively, Remark 14.11(a) implies that $T_k(t) \to T(t)$ strongly for all $t > 0$. Moreover, by Theorem 11.20 the sequence $((T_k)(t))_k$ is monotone decreasing (in the sense of the order on $L_2(\mathbb{R}^n)$).

Suppose, more specifically, that $C = \overline{\Omega}$ for some bounded open set $\Omega \subseteq \mathbb{R}^n$. Then clearly $H_0^1(\Omega) \subseteq H_0^1(\overline{\Omega})$. If in addition $\partial\Omega$ is a Lebesgue null set, we conclude that

$$H_0^1(\Omega) \subseteq H_0^1(\overline{\Omega}) \subseteq L_2(\Omega) \ \big(= \big\{u \in L_2(\mathbb{R}^n)\,;\, u|_{\mathbb{R}^n\setminus\Omega} = 0\big\}\big).$$

It follows that a is a densely defined form in $L_2(\Omega)$. In [ArDa08b], the associated self-adjoint operator in $L_2(\Omega)$ is called the "pseudo-Dirichlet Laplacian".

It is natural to ask for conditions implying that $H_0^1(\Omega) = H_0^1(\overline{\Omega})$ (in which case the pseudo-Dirichlet Laplacian is just the Dirichlet Laplacian). This equality holds if Ω has continuous boundary: then Ω and $\overline{\Omega}$ only differ by a Lebesgue null set (see Exercise 14.4), and hence the equality follows from Proposition 7.10. The example $n = 1$, $\Omega := (-1, 0) \cup (0, 1)$ shows that the equality does not hold for all Ω, even if $\partial\Omega$ is a null set. We refer to the Notes for more information. △

14.3 Form convergence 'from above'

A counterpart to increasing sequences of accretive symmetric forms are decreasing sequences. This topic will be treated in Corollary 14.21, which is a special case of our main result on non-monotone sequences of forms that converge 'from above' to a limit form, in a suitable sense; see Theorem 14.17.

As a warm-up we begin with the following important approximation result from numerical analysis.

14.13 Lemma (Céa's lemma). *Let V be a Hilbert space, and let a be a bounded coercive form on V,*

$$|a(u,v)| \leqslant M\|u\|_V\|v\|_V, \quad \operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2 \qquad (u,v \in V)$$

for some $M \geqslant 0$, $\alpha > 0$. Let $\check{V}$ be a closed subspace of V.

Let $\eta \in V^*$, and let $u \in V$, $\check{u} \in \check{V}$ be the unique elements such that

$$a(u,v) = \langle \eta, v\rangle_{V^*,V} \quad (v \in V), \qquad a(\check{u},v) = \langle \eta, v\rangle_{V^*,V} \quad (v \in \check{V}).$$

Then $\|u-\check{u}\|_V \leqslant \frac{M}{\alpha}\operatorname{dist}(u,\check{V})$.

Proof. The existence of u and $\check{u}$ is a consequence of Theorem 5.4, the Lax–Milgram lemma, where for $\check{u}$ the Lax–Milgram lemma is applied to the restriction of a to $\check{V}$.

For $v \in \check{V}$ we compute

$$a(u-\check{u},v) = \langle \eta, v\rangle_{V^*,V} - a(\check{u},v) = 0. \tag{14.5}$$

Hence, for any $\check{v} \in \check{V}$ we obtain

$$\begin{aligned}\alpha\|u-\check{u}\|_V^2 &\leqslant \operatorname{Re} a(u-\check{u},(u-\check{v})+(\check{v}-\check{u})) = \operatorname{Re} a(u-\check{u},u-\check{v})\\ &\leqslant M\|u-\check{u}\|_V\|u-\check{v}\|_V\end{aligned}$$

for all $\check{v} \in \check{V}$. □

14.14 Remarks. (a) Let the hypotheses be as in Lemma 14.13, let (V_n) be a sequence of closed subspaces of V such that $\operatorname{dist}(u,V_n) \to 0$, and let $u_n \in V_n$ $(n \in \mathbb{N})$ be the unique elements such that

$$a(u_n,v) = \langle \eta, v\rangle_{V^*,V} \qquad (v \in V_n).$$

Then Lemma 14.13 implies that $u_n \to u$ in V.

(b) The application of Lemma 14.13 sketched in part (a) is an essential ingredient for the Galerkin method in numerical analysis. An important feature in this application is the circumstance that the sequence of subspaces need not be increasing. The estimate $\|u-u_n\|_V \leqslant \frac{M}{\alpha}\operatorname{dist}(u,V_n)$ from Céa's lemma states that, up to the factor $\frac{M}{\alpha}$, the "Galerkin approximation" $u_n \in V_n$ is as close to the solution u as the best approximation to u in V_n. In this context, equation (14.5) is called "Galerkin orthogonality". △

Next we present a generalisation of Céa's lemma that provides the key estimate needed in the proof of our main result.

14.15 Proposition. *Let V be a complex Hilbert space, and let a be a bounded coercive form on V,*

$$|a(u,v)| \leqslant M\|u\|_V\|v\|_V, \quad \operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2 \qquad (u,v \in V)$$

for some $M \geqslant 0$, $\alpha > 0$.

Let $\check V$ be a complex Hilbert space, $\check a$ a bounded coercive form on $\check V$, $J \in \mathcal{L}(\check V, V)$, and assume that there exists $\theta \in [0, \frac{\pi}{2})$ such that

$$\check a(v) - a(Jv) \in \overline{\Sigma_\theta} \qquad (v \in \check V).$$

Let $\eta \in V^*$, and let $u \in V$, $\check u \in \check V$ be the unique elements such that

$$a(u, v) = \langle \eta, v\rangle_{V^*,V} \quad (v \in V), \qquad \check a(\check u, v) = \langle \eta, Jv\rangle_{V^*,V} \quad (v \in \check V).$$

Then

$$\|u - J\check u\|_V^2 \leqslant \inf_{v\in\check V}\Big(\frac{M^2}{\alpha^2}\|u - Jv\|_V^2 + \frac{c^2}{2\alpha}\big|\check a(v) - a(Jv)\big|\Big), \tag{14.6}$$

where $c := 1 + \tan\theta$.

Proof. We define a form b on $\check V$ by

$$b(w, v) := \check a(w, v) - a(Jw, Jv).$$

The assumptions imply that

$$|b(w, v)| \leqslant c(\operatorname{Re} b(w))^{1/2}(\operatorname{Re} b(v))^{1/2} \qquad (w, v \in \check V); \tag{14.7}$$

see (12.2). For $v \in \check V$ we compute

$$a(u - J\check u, Jv) = \langle \eta, Jv\rangle_{V^*,V} - a(J\check u, Jv) = \check a(\check u, v) - a(J\check u, Jv) = b(\check u, v).$$

Hence, for $v \in \check V$ we obtain

$$\begin{aligned}\alpha\|u - J\check u\|_V^2 &\leqslant \operatorname{Re} a\big(u - J\check u, (u - Jv) + J(v - \check u)\big)\\ &= \operatorname{Re} a(u - J\check u, u - Jv) + \operatorname{Re} b(\check u, v - \check u).\end{aligned}$$

We now use (14.7), the boundedness of a and the Peter–Paul inequality (twice) to estimate

$$\begin{aligned}\alpha\|u - J\check u\|_V^2 &\leqslant M\|u - J\check u\|_V\|u - Jv\|_V + c(\operatorname{Re} b(\check u))^{1/2}(\operatorname{Re} b(v))^{1/2} - \operatorname{Re} b(\check u)\\ &\leqslant \frac{\alpha}{2}\|u - J\check u\|_V^2 + \frac{M^2}{2\alpha}\|u - Jv\|_V^2 + \frac{c^2}{4}\operatorname{Re} b(v).\end{aligned}$$

Reshuffling terms we conclude that

$$\|u - J\check u\|_V^2 \leqslant \frac{M^2}{\alpha^2}\|u - Jv\|_V^2 + \frac{c^2}{2\alpha}\operatorname{Re}\big(\check a(v) - a(Jv)\big). \tag{14.8}$$

□

14.16 Remarks. (a) If the space $\check V$ is a subspace of V, $J\colon \check V \hookrightarrow V$ is the embedding and $\check a = a|_{\check V\times\check V}$, then Proposition 14.15 reduces to Céa's lemma.

(b) The following considerations serve to transform (14.6) into an inequality that will be used in the proof of the subsequent theorem. In the setting of Proposition 14.15, let

$\mathcal{A} \in \mathcal{L}(V, V^*)$ and $\check{\mathcal{A}} \in \mathcal{L}(\check{V}, \check{V}^*)$ be the Lax–Milgram operators associated with a and $\check{a}$, respectively. We define the dual operator $J' \in \mathcal{L}(V^*, \check{V}^*)$ by

$$J'\check{\eta} := \check{\eta} J \qquad (\check{\eta} \in V^*).$$

Then, for $\eta \in V^*$ as specified in Proposition 14.15, the elements $u \in V$, $\check{u} \in \check{V}$ are given by

$$u = \mathcal{A}^{-1}\eta, \quad \check{u} = \check{\mathcal{A}}^{-1} J'\eta.$$

With this notation, (14.6) reads

$$\big\|\mathcal{A}^{-1}\eta - J\check{\mathcal{A}}^{-1}J'\eta\big\|_V^2 \leqslant \inf_{v \in \check{V}} \Big(\frac{M^2}{\alpha^2}\big\|\mathcal{A}^{-1}\eta - Jv\big\|_V^2 + \frac{c^2}{2\alpha}\big|\check{a}(v) - a(Jv)\big|\Big). \tag{14.9}$$

△

We now come to our main result on form convergence 'from above', which is formulated in a rather general setting. As before, the forms are not assumed to be densely defined. In contrast to Theorem 14.10, they need not be embedded, symmetric nor closed.

14.17 Theorem. *Let H be a complex Hilbert space, and let (a, j) be a quasi-sectorial form in H. For $n \in \mathbb{N}$ let a_n be a form with $\operatorname{dom}(a_n) \subseteq \operatorname{dom}(a)$. Let $\theta \in [0, \frac{\pi}{2})$, and assume that*

$$a_n(u) - a(u) \in \overline{\Sigma_\theta} \qquad \big(u \in \operatorname{dom}(a_n),\ n \in \mathbb{N}\big). \tag{14.10}$$

(This implies that the forms a_n are quasi-sectorial with the same vertex as a and with a common angle.) Let D be a core for a, and suppose that for all $u \in D$ there exists a sequence (u_n) in $\operatorname{dom}(a)$, $u_n \in \operatorname{dom}(a_n)$ for all $n \in \mathbb{N}$, such that $u_n \to u$ in $\operatorname{dom}(a)$ and $a_n(u_n) \to a(u)$ as $n \to \infty$.

Let A be the quasi-m-sectorial linear relation associated with (a, j) (see (14.2)), and let A_n be the quasi-m-sectorial linear relation associated with $(a_n, j|_{\operatorname{dom}(a_n)})$, for $n \in \mathbb{N}$. Then (A_n) converges to A in the strong resolvent sense, i.e. $(\lambda + A_n)^{-1} \to (\lambda + A)^{-1}$ $(n \to \infty)$ strongly for all $\lambda > \omega$, where $-\omega$ is a vertex of (a, j).

14.18 Remarks. (a) In the situation of Theorem 14.17 it follows from Theorem 13.4, together with a rescaling argument, that the degenerate strongly continuous semigroups T_n generated by the linear relations $-A_n$ converge to the semigroup T generated by $-A$. In fact, because $\ker(T(0)) = \operatorname{ran}(j)^\perp \subseteq \operatorname{ran}(j|_{\operatorname{dom}(a_n)})^\perp = \ker(T_n(0))$ for all $n \in \mathbb{N}$, it even follows that $T_n(t)x \to T(t)x$ $(n \to \infty)$ uniformly on compact subsets of $[0, \infty)$, for all $x \in H$ (and not just for $x \in \operatorname{ran}(T(0))$).

(b) In the hypotheses of Theorem 14.17, the condition $u_n \to u$ in $\operatorname{dom}(a)$ implies that $a(u_n) \to a(u)$. Therefore '$a_n(u_n) \to a(u)$ as $n \to \infty$' is equivalent to requiring '$a_n(u_n) - a(u_n) \to 0$ as $n \to \infty$'.

(c) In the proof of Theorem 14.17 it is shown that the convergence property required for all $u \in D$ is in fact satisfied for all $u \in \operatorname{dom}(a)$; see the argument connected with the reformulation (14.11) of this property.

(d) The hypotheses of the theorem express a kind of convergence $a_n \to a$. In previous results on form convergence 'from above' (see e.g. [Kat80; Chap. VIII, Theorem 3.6] and

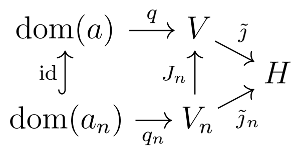
[ArEl12b; Theorem 3.7]) a more restrictive assumption is used, namely that the set

$$D := \Big\{u \in \bigcup_{k\in\mathbb{N}} \bigcap_{n\geqslant k} \operatorname{dom}(a_n)\,;\ a_n(u) \to a(u)\ (n\to\infty)\Big\}$$

is a core for a. Then, in particular, the elements of D belong to almost every $\operatorname{dom}(a_n)$. This property of D is not required in Theorem 14.17. △

Proof of Theorem 14.17. Without loss of generality we assume that $-\omega = 1$ is a vertex of (a, j); cf. Remark 14.2. Then $\operatorname{Re} a$ is a semi-inner product on $\operatorname{dom}(a)$ that is equivalent to the semi-inner product $(\cdot\,|\,\cdot)_{a,j}$ defined in (12.4). Moreover (a, j) is sectorial, and (14.10) implies that (a_n, j) is sectorial for all $n \in \mathbb{N}$.

Let (V, q) denote the completion of $(\operatorname{dom}(a), \operatorname{Re} a)$. Then there exist a unique $\tilde{\jmath} \in \mathcal{L}(V, H)$ and a unique bounded form $\tilde{a}\colon V \times V \to \mathbb{C}$ such that $\tilde{\jmath} \circ q = j$ and $\tilde{a}(q(u), q(v)) = a(u, v)$ for all $u, v \in \operatorname{dom}(a)$ (see Remark 12.3(c)); recall from (14.2) that A is the linear relation associated with $(\tilde{a}, \tilde{\jmath})$. Analogously we define V_n, q_n, $\tilde{\jmath}_n$ and $\tilde{a}_n$ corresponding to a_n, for $n \in \mathbb{N}$; then A_n is the linear relation associated with $(\tilde{a}_n, \tilde{\jmath}_n)$.It follows from (14.10) that the embedding $\operatorname{dom}(a_n) \hookrightarrow \operatorname{dom}(a)$ is continuous for all $n \in \mathbb{N}$, and by Remark 12.3(c) there exists $J_n \in \mathcal{L}(V_n, V)$ such that $J_n \circ q_n = q|_{\operatorname{dom}(a_n)}$. Then $\tilde{\jmath}_n \circ q_n = j|_{\operatorname{dom}(a_n)} = \tilde{\jmath} \circ q|_{\operatorname{dom}(a_n)} = \tilde{\jmath} \circ J_n \circ q_n$; hence $\tilde{\jmath}_n = \tilde{\jmath} \circ J_n$ on $\operatorname{ran}(q_n)$, and by denseness on all of V_n.



Observe that the forms $\tilde{a}$ and $\tilde{a}_n$ are coercive because 1 is a vertex for (a, j). Let $\tilde{\mathcal{A}} \in \mathcal{L}(V, V^*)$ and $\tilde{\mathcal{A}}_n \in \mathcal{L}(V_n, V_n^*)$ $(n \in \mathbb{N})$ be the Lax–Milgram operators associated with $\tilde{a}$ and $\tilde{a}_n$ $(n \in \mathbb{N})$, respectively. By Proposition 14.19 proved subsequently we have $A^{-1} = \tilde{\jmath}\tilde{\mathcal{A}}^{-1}\tilde{k}$ and $A_n^{-1} = \tilde{\jmath}_n\tilde{\mathcal{A}}_n^{-1}\tilde{k}_n$ for all $n \in \mathbb{N}$, where $\tilde{k}(y) = (y\,|\,\tilde{\jmath}(\cdot))_H$ and $\tilde{k}_n(y) = (y\,|\,\tilde{\jmath}_n(\cdot))_H$ for all $y \in H$. We will show that $J_n\mathcal{A}_n^{-1}J_n' \to \mathcal{A}^{-1}$ strongly in $\mathcal{L}(V^*, V)$ as $n \to \infty$. Then, using $\tilde{\jmath}_n = \tilde{\jmath} \circ J_n$ and $\tilde{k}_n = J_n' \circ \tilde{k}$ for all $n \in \mathbb{N}$, we can infer that

$$A_n^{-1} = \tilde{\jmath}(J_n\mathcal{A}_n^{-1}J_n')\tilde{k} \to \tilde{\jmath}\mathcal{A}^{-1}\tilde{k} = A^{-1} \qquad (n \to \infty)$$

strongly in $\mathcal{L}(H)$, and applying Lemma 13.9 we may conclude that $\operatorname{s-lim}_{n\to\infty}(\lambda + A_n)^{-1} = (\lambda + A)^{-1}$ for all $\lambda > \omega = -1$.

Let $n \in \mathbb{N}$. Then for all $u \in \operatorname{dom}(a_n)$ we have

$$\tilde{a}_n(q_n(u)) - \tilde{a}(J_nq_n(u)) = a_n(u) - a(u) \in \overline{\Sigma_\theta}\,.$$

Since $\operatorname{ran}(q_n)$ is dense in V_n, it follows that $\tilde{a}_n(v) - \tilde{a}(J_nv) \in \overline{\Sigma_\theta}$ for all $v \in V_n$. Thus we can apply Proposition 14.15 to the forms $\tilde{a}$ and $\tilde{a}_n$.

For the proof of the strong convergence $\mathcal{A}^{-1} = \operatorname{s-lim}_{n\to\infty} J_n\mathcal{A}_n^{-1}J_n'$ we now evaluate the right-hand side of the inequality (14.9). By the convergence hypothesis of the theorem and Remark 14.18(b) we have

$$\inf_{v\in\operatorname{dom}(a_n)}\big(\|u - v\|_a^2 + |a_n(v) - a(v)|\big) \to 0 \qquad (n \to \infty), \tag{14.11}$$

for all $u \in D$. As D is a core for a, the convergence (14.11) carries over to all $u \in \operatorname{dom}(a)$. In view of the properties of the mappings q, q_n and J_n this implies that

$$\inf_{v\in\operatorname{ran}(q_n)}\big(\|u - J_nv\|_V^2 + |\tilde{a}_n(v) - \tilde{a}(J_nv)|\big) \to 0 \qquad (n \to \infty),$$

for all $u \in \operatorname{ran}(q)$. Then, using the inclusions $\operatorname{ran}(q_n) \subseteq V_n$ as well as the denseness of $\operatorname{ran}(q)$ in V, we also obtain

$$\inf_{v \in V_n} \left(\|u - J_n v\|_V^2 + \left|\tilde{a}_n(v) - \tilde{a}(J_n v)\right| \right) \to 0 \qquad (n \to \infty),$$

for all $u \in V$. Combining this convergence with the inequality (14.9) we conclude that $J_n \mathcal{A}_n^{-1} J_n' \to \mathcal{A}^{-1}$ strongly in $\mathcal{L}(V^*, V)$ as $n \to \infty$. □

In the above proof we have used the following extension of Proposition 5.7 to non-densely defined coercive forms.

14.19 Proposition. *Let V and H be Hilbert spaces, $j \in \mathcal{L}(V, H)$, $a\colon V \times V \to \mathbb{K}$ a bounded coercive form, and let A be the strictly m-accretive linear relation associated with (a, j). Let $\mathcal{A}\colon V \to V^*$ be the Lax–Milgram operator associated with a, and let $k\colon H \to V^*$, $y \mapsto (y \,|\, j(\cdot))$ be as in Proposition 5.7. Then*

$$A^{-1} = j\mathcal{A}^{-1}k.$$

(Recall from Theorem 5.4 that $\mathcal{A}$ is invertible with inverse in $\mathcal{L}(V^*, V)$.)

Proof. For $x, y \in H$ one has $(x, y) \in A$ if and only if there exists $u \in V$ such that $j(u) = x$ and $a(u, v) = (y \,|\, j(v))$ for all $v \in V$. The latter property is equivalent to $\mathcal{A}u = k(y)$, i.e. $u = \mathcal{A}^{-1}k(y)$. Therefore $(x, y) \in A$ if and only if $x = j(u) = (j\mathcal{A}^{-1}k)y$. □

14.20 Remark. Theorem 14.17, with the same proof, also holds for real Hilbert spaces if all the forms a, a_n are symmetric and accretive, and where (14.10) should be interpreted as $a_n(u) - a(u) \geqslant 0$. This is why we do not have to suppose a complex Hilbert space in the subsequent Corollary 14.21. △

In the following result we treat the special case of decreasing sequences of embedded symmetric forms; an interesting feature is that then one does not need to specify the limiting form a in advance. In this context it is natural to work without the assumption of closedness or closability (see Remark 14.23 below).

14.21 Corollary. *Let H be a Hilbert space, and let (a_n) be a decreasing sequence of (embedded) accretive symmetric forms in H. Define* $\operatorname{dom}(a) := \bigcup_{n \in \mathbb{N}} \operatorname{dom}(a_n)$.
Then

$$a(x, y) := \lim_{n \to \infty} a_n(x, y)$$

exists for all $x, y \in \operatorname{dom}(a)$, and a is an accretive symmetric form in H. Let A be the accretive self-adjoint linear relation associated with a, and similarly, let A_n be associated with a_n, for $n \in \mathbb{N}$. Then $A_n \to A$ in the strong resolvent sense, i.e. $(\lambda + A_n)^{-1} \to (\lambda + A)^{-1}$ strongly as $n \to \infty$, for all $\lambda > 0$.

Proof. The existence of $\lim_{n\to\infty} a_n(x, y)$ is obtained as in the proof of Proposition 14.8, and the asserted properties of a are then obvious. We can now apply Theorem 14.17 to obtain the remaining assertions. (Choose $u_n = u$ for large n in the assumptions of the theorem.) □

14.22 Example. Let Ω be an open subset of $\mathbb{R}^n$, and let (Ω_k) be an increasing sequence of open subsets of Ω satisfying $\Omega = \bigcup_{k\in\mathbb{N}} \Omega_k$. For all $k \in \mathbb{N}$ let a_k be the classical Dirichlet form on $C_c^\infty(\Omega_k)$, and let a be the classical Dirichlet form on $C_c^\infty(\Omega)$ (cf. Example 12.6).

Then, for all $k \in \mathbb{N}$, the self-adjoint linear relation in $L_2(\Omega)$ associated with a_k is the negative Dirichlet Laplacian in $L_2(\Omega_k)$, supplemented by $\{0\} \times L_2(\Omega \setminus \Omega_k)$. Corollary 14.21 implies that (A_k) converges in the strong resolvent sense to the negative Dirichlet Laplacian on $L_2(\Omega)$. As in Example 14.12, if T_k and T are the degenerate strongly continuous semigroups on $L_2(\Omega)$ associated with the forms a_k and a, respectively, then $T_k(t) \to T(t)$ strongly for all $t > 0$, and the sequence $((T_k)(t))_k$ is monotone increasing (in the sense of the order on $L_2(\Omega)$).

We refer to Exercise 14.3 for a more sophisticated version of this example. △

14.23 Remark. In contrast to Theorem 14.10, where increasing sequences of forms are treated, in Corollary 14.21 it is not supposed that the forms a_n are closed. Even if the forms a_n are closed, the limiting form a in Corollary 14.21 need not be closable, as is demonstrated by the example in Exercise 14.5. The example also shows how to avoid this problem if one can guess a closed limiting form: then one can apply Theorem 14.17 instead of Corollary 14.21. △

Notes

For an outline of the early history of convergence theorems for sequences of forms – mainly in the 70's – we refer to Reed–Simon [ReSi80; Supplementary Material, Notes to Supplement VIII.7].

The use of non-densely defined forms was advocated in [Sim78; Section 4], for symmetric forms, and further developed, for sectorial forms, in [ArBa93; Section 7], [Kun05], [MVV05], [BaEl14], [VoVo20a]. We refer to [AEKS14] for an application of non-densely defined forms in the context of essential coercivity (or 'compact ellipticity' in the terminology of [AEKS14]).

The form convergence theorem for increasing sequences of forms, Theorem 14.10, is contained in [Sim78; Theorem 4.1]; the full proof can be found in [VoVo20a]. See also Simon [Sim78; Theorem 3.1], Kato [Kat80; Chap. VIII, Theorem 3.13a] for the special case in which the limiting form is densely defined. The proof of Proposition 14.5 is taken from [VoVo20a].

As there is no natural notion of order between sectorial forms, a generalisation of Theorem 14.10 to sequences of sectorial forms is not straightforward. We mention [Ouh95b; Theorem 6], [BaEl14; Theorem 1.2] and [VoVo20a] for results concerning this topic. In the latter two references the input is a sequence (a_n) of sectorial forms for which $a_{n+1} - a_n$, for all $n \in \mathbb{N}$, is sectorial of a fixed angle $\theta \in [0, \frac{\pi}{2})$.

In Example 14.12 the question was raised under what conditions the equality

$$H_0^1(\Omega) = H_0^1(\overline{\Omega}) = \bigl\{u|_{\overline{\Omega}};\ u \in H^1(\mathbb{R}^n),\ u|_{\mathbb{R}^n\setminus\overline{\Omega}} = 0\bigr\}$$

holds for a bounded open set $\Omega \subseteq \mathbb{R}^n$. It turns out that this property is equivalent to an approximation property for harmonic functions. (Recall that $u\colon \Omega \to \mathbb{K}$ is called harmonic if $u \in C^2(\Omega)$ and $\Delta u = 0$.)

The following result is Theorem 11.8 in the survey of Hedberg [Hed93], where methods of abstract potential theory are used for the proof. Versions of the result go back to Keldysh [Kel41], with more direct proofs.

Theorem (Hedberg-Keldysh) *Let $n \geqslant 2$, and let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with $\mathring{\overline{\Omega}} = \Omega$. Then the following properties are equivalent.*

(i) $H_0^1(\Omega) = H_0^1(\overline{\Omega})$.

(ii) *For each function $u \in C(\overline{\Omega})$, $u|_\Omega$ harmonic, there exist functions u_k that are defined and harmonic on an open neighbourhood Ω_k of $\overline{\Omega}$ such that $u_k(x) \to u(x)$ as $k \to \infty$, uniformly for $x \in \overline{\Omega}$.*

We point out that in property (ii) the function u need not belong to $H^1(\Omega)$, even if Ω is a very 'nice' set such as the unit disc $B_{\mathbb{R}^2}(0,1)$; see the discussion of Hadamard's example in the Notes of Chapter 7. Moreover, property (ii) is not related to the solvability of the Dirichlet problem: there is no implication between (ii) and the Dirichlet regularity of Ω. In particular, the set of traces $u|_{\partial\Omega}$ of the functions u occurring in property (ii) may be a proper subset of $C(\partial\Omega)$.

Theorem 14.17 is contained in [VoVo24]; it generalises [Kat80; Chap. VIII, Theorem 3.6], [ArEl12b; Theorem 3.7] and [ChEl18; Theorem 3.7]. The noteworthy new features of our result are that we allow more general domains of the forms a_n – which may in fact have pairwise trivial intersection – and that the limiting form a need not be densely defined. An essential ingredient of the proof is Proposition 14.15, which is an extended version of Céa's lemma, Lemma 14.13. The origin of explicit estimates of the type stated in Lemma 14.13 seems to be [Cea64; Proposition 3.1 on p. 365], where a stronger estimate is stated for the case of symmetric forms. The lemma in the general non-symmetric case is not stated explicitly in [Cea64].

It was mentioned in Remarks 14.14 that Céa's lemma is fundamental for the Galerkin method, which in turn is the basis of the finite-element method for the solution of partial differential equations. We will not expand on this topic here, but we mention the recent paper [ACE22] for the disussion and analysis of various aspects concerning the Galerkin approximation.

Corollary 14.21 generalises [Sim78; Theorems 3.2 and 4.1], where it is assumed that the forms a_n are closed. (It follows from Remark 12.13(c) that the linear relation A in Corollary 14.21 is associated with the form $\overline{a_r}$, where a_r is Simon's 'regular part' of a.)

Exercises

14.1 (a) Let V, H, j and a as well as the remaining notation be as in the second paragraph of Section 14.1, and let A be the linear relation associated with (a, j); see (14.1). Show that A is quasi-m-accretive and that $-A$ is the generator of T.

(b) Now let H be a complex Hilbert space, and let (a, j) be a quasi-sectorial form in H. Show that the linear relation A associated with (a, j) is quasi-m-sectorial. Moreover show that A is m-sectorial if (a, j) is sectorial.

14.2 Verify the assertions stated in Remarks 14.3.

14.3 (a) Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $(\Omega_k)_{k\in\mathbb{N}}$ be a sequence of open subsets of Ω with the property that for each compact set $K \subseteq \Omega$ there exists $k_K \in \mathbb{N}$ such that $K \subseteq \Omega_k$ for all $k \geqslant k_K$. (This property is equivalent to $\bigcup_{n\in\mathbb{N}}\bigcap_{k\geqslant n}\Omega_k = \Omega$.)

Let a be the classical Dirichlet form on $C_c^\infty(\Omega)$,

$$a(f,g) = \int \nabla f \cdot \overline{\nabla g}\,\mathrm{d}x \qquad (f,g \in C_c^\infty(\Omega)),$$

let a_k be the restriction of a to $\operatorname{dom}(a_k) := C_c^\infty(\Omega_k)$, and let A and A_k ($k \in \mathbb{N}$) be the self-adjoint linear relation associated with a and a_k ($k \in \mathbb{N}$), respectively. (In other words, the operator A is the negative Dirichlet Laplacian in $L_2(\Omega)$, whereas for $k \in \mathbb{N}$, A_k is the negative Dirichlet Laplacian in $L_2(\Omega_k)$ supplemented by $\{0\} \times L_2(\Omega \setminus \Omega_k)$ to a self-adjoint linear relation in $L_2(\Omega)$.)

Show that $A_k \to A$ in the strong resolvent sense.

(b) Show that the sets $\Omega := \mathbb{R}^n$, $\Omega_k := B(0,k) \cup (\mathbb{R}^n \setminus B[0,k+1])$ ($k \in \mathbb{N}$) satisfy the conditions posed in (a) above. Show that $k,l \in \mathbb{N}$, $\Omega_k \subseteq \Omega_l$ implies that $k = l$.

This example shows that, in general, in the context of part (a) one cannot expect to find a decreasing subsequence $(a_{k_j})_j$ of $(a_k)_k$.

14.4 Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with continuous boundary. Show that $\partial\Omega$ has Lebesgue measure 0 and that $\mathring{\overline{\Omega}} = \Omega$.

14.5 (From [Kat80; Chap. VIII, Example 3.10]) Let $H := L_2(0,1;\mathbb{R})$, and for $n \in \mathbb{N}$ define the form a_n in H by $\operatorname{dom}(a_n) := H^1(0,1)$,

$$a_n(u,v) := \frac{1}{n}\int_0^1 u'(\xi)v'(\xi)\,\mathrm{d}\xi + u(0)v(0) + u(1)v(1).$$

(a) Show that (a_n) is a decreasing sequence of (embedded) closed accretive symmetric forms. Determine the form a given by Corollary 14.21 and show that a is not closable.

(b) Choose $a_0 := 0$ with domain $\operatorname{dom}(a_0) := H$, and use Theorem 14.17 with $D := H_0^1(0,1)$ to show that the sequence (A_n) converges to $A = 0$ in the strong resolvent sense, where A_n is associated with a_n.

(c) Conclude from (b) that the operator $A = 0$ is associated with the form a from part (a).

14.6 Let H be a Hilbert space, and let A be an accretive linear relation in H. Prove the following properties:

(a) $\overline{A}$ is accretive.

(b) A is closed if and only if $\operatorname{ran}(I + A)$ is closed.

(c) If $z \in \operatorname{ran}(I + A)^\perp$, then

$$A_z := A + \operatorname{lin}\{(z,z)\} = \{(x + \lambda z, y + \lambda z);\, (x,y) \in A\}$$

is accretive, and $z \in \operatorname{ran}(I + A_z)$.

14.7 Let H be a Hilbert space, and let A be an accretive linear relation in H.

(a) Show that A has a **maximal accretive** extension, i.e., there exists an accretive linear relation $B \supseteq A$ such that B has no proper accretive extension. (Hint: Zorn's lemma.)

(b) Assume that A is maximal accretive. Show that A is m-accretive. (Hint: Use Exercise 14.6.)

(c) Assume that A is a densely defined operator. Conclude from part (b) and Lemma 13.3 that A is m-accretive if and only if A is a maximal accretive *operator* (i.e. no proper accretive extension of A is an operator).

14.8 Let H be a Hilbert space, and let A be a linear relation in H. Prove the following properties.

(a) $\ker(A^*) = \operatorname{ran}(A)^\perp$ and $\overline{\operatorname{ran}(A^*)} = \ker(\overline{A})^\perp$.

(b) If A is accretive, then $\ker(A) \subseteq \ker(A^*)$. (Hint: Proceed similarly as in the proof of Lemma 13.3; then use part (a).)

(c) If A is self-adjoint and accretive, then A is m-accretive. (Hint: Mimic the proof of Theorem 6.1.)

14.9 Let H be a complex Hilbert space, and let A be a linear relation in H. Prove the following properties.

(a) $A \subseteq A^*$ if and only if $(x \,|\, y) \in \mathbb{R}$ for all $(x, y) \in A$. (Hint concerning '$\Leftarrow$': Assume that $(x \,|\, y) \in \mathbb{R}$ for all $(x, y) \in A$. Then, given $(x, y), (f, g) \in A$, show that $(x \,|\, \lambda g) + (\lambda f \,|\, y) \in \mathbb{R}$ for all $\lambda \in \mathbb{C}$, and conclude that $(x \,|\, g) = (y \,|\, f)$.)

(b) A is self-adjoint if and only if $\pm \mathrm{i}A$ are both m-accretive. (Hint: If A is self-adjoint, then proceed as in the proof of Theorem 6.1 to show that $\operatorname{ran}(I + \mathrm{i}A) = H$.)

(c) A is self-adjoint and accretive if and only if A is m-sectorial of angle 0.

Chapter 15

The Trotter product formula for forms

The central result of this chapter is the Trotter product formula for sectorial forms, which is the content of Section 15.3. The proof relies on the material presented in Chapters 13 and 14. In addition we will need Vitali's convergence theorem for sequences of holomorphic functions and holomorphic dependence of forms and operators; see Sections 15.1 and 15.2. In Section 15.4 we discuss applications of the Trotter product formula.

15.1 Interlude: The Vitali convergence theorem

Convergence of a sequence of holomorphic functions on a domain $\Omega \subseteq \mathbb{C}$ propagates from very small subsets to the entire domain. This surprising phenomenon has been discovered by Vitali. In the functional analytic setting used in our proof we obtain Vitali's theorem as a consequence of the unique extension theorem.

15.1 Theorem (Vitali). *Let X be a complex Banach space. Let $\Omega \subseteq \mathbb{C}$ be a connected open set, and let (f_n) be a sequence of holomorphic functions $f_n \colon \Omega \to X$. Assume that (f_n) is locally uniformly bounded and that the set*

$$\Omega_0 := \bigl\{ z \in \Omega;\ (f_n(z))_{n\in\mathbb{N}} \text{ convergent} \bigr\}$$

has a cluster point in Ω.

Then there exists a holomorphic function $f \colon \Omega \to X$ such that $f_n \to f$ locally uniformly.

Proof. Denote by $\ell_\infty(\mathbb{N};X)$ the space of bounded sequences in X, provided with the supremum norm, and by $c(\mathbb{N};X)$ the closed subspace of convergent sequences.

We define the function $g\colon \Omega \to \ell_\infty(\mathbb{N};X)$ by $g(z) := (f_n(z))_{n\in\mathbb{N}}$ $(z \in \Omega)$. Then g is holomorphic, by the equivalence of (i) and (iv) in Theorem 3.2; take $\bigl\{\ell_\infty(\mathbb{N};X) \ni (x_n) \mapsto x'(x_k);\ x' \in X',\ k \in \mathbb{N}\bigr\}$ as a norming subset of the dual of $\ell_\infty(\mathbb{N};X)$. (See also Exercise 15.1.) On the set Ω_0, the function g takes its values in $c(\mathbb{N};X)$. Applying the identity theorem for holomorphic functions together with the Hahn–Banach theorem we conclude that $g(z) \in c(\mathbb{N};X)$ for all $z \in \Omega$, i.e. (f_n) converges pointwise.

Now the assertion is a consequence of Theorem 3.5. □

15.2 Holomorphic families of operators defined by forms

The following result shows that the holomorphic dependence – on a complex variable – of a family of sectorial forms implies holomorphic dependence of the associated sectorial operators. A rather striking application, due to B. Simon, will be given in Corollary 15.9.

15.2 Theorem. *Let H be a complex Hilbert space, $V \subseteq H$ a subspace, and let $\Omega \subseteq \mathbb{C}$ be open. For each $z \in \Omega$ let a_z be a closed sectorial form in H with domain* $\operatorname{dom}(a_z) = V$, *and let A_z denote the m-sectorial linear relation associated with a_z (as defined in Section 14.1). Assume that for all $x, y \in V$ the mapping $\Omega \ni z \mapsto a_z(x, y)$ is holomorphic.*

Then the function $\Omega \ni z \mapsto (I + A_z)^{-1} \in \mathcal{L}(H)$ is holomorphic.

Note that the subspace V is not assumed to be dense in H, so that the semigroups generated by $-A_z$ may be degenerate. However, the corresponding 'active' subspaces are all equal to the closure of V in H, in particular they do not depend on z. In Exercise 15.5(a) it is explained why it would be useless to try finding a version of Theorem 15.2 with z-dependent domains $\operatorname{dom}(a_z)$.

15.3 Lemma. *Let X, Y be complex Banach spaces. Let $r > 0$, and let $A\colon B_{\mathbb{C}}(0, r) \to \mathcal{L}(X, Y)$ be holomorphic. Assume that $A(0)$ is invertible with $A(0)^{-1} \in \mathcal{L}(Y, X)$.*

Then there exists $\delta \in (0, r)$ such that $A(z)$ has an inverse in $\mathcal{L}(Y, X)$ for all $z \in B(0, \delta)$, and the function $B(0, \delta) \ni z \mapsto A(z)^{-1} \in \mathcal{L}(Y, X)$ is holomorphic.

Proof. Replacing $A(z)$ by $A(0)^{-1}A(z) \in \mathcal{L}(X)$ $(z \in B(0, r))$ we see that without loss of generality we may assume that $Y = X$.

It is an easy consequence of Remark 2.3(a) (Neumann series) that

$$\mathcal{I}(X) := \bigl\{A \in \mathcal{L}(X);\ A \text{ invertible in } \mathcal{L}(X)\bigr\}$$

is open in $\mathcal{L}(X)$ and that the mapping $\mathcal{I}(X) \ni A \mapsto A^{-1} \in \mathcal{I}(X)$ is continuous. This implies that there exists $\delta \in (0, r)$ such that $A(z) \in \mathcal{I}(X)$ for all $z \in B(0, \delta)$ and $\sup_{|z|<\delta} \|A(z)^{-1}\| < \infty$. Dividing the equality

$$A(w)^{-1} - A(z)^{-1} = A(w)^{-1}\bigl(A(z) - A(w)\bigr)A(z)^{-1} \qquad (w, z \in B(0, \delta))$$

by $w - z$ and taking the limit $w \to z$, we conclude that $z \mapsto A(z)^{-1}$ is holomorphic and that

$$\frac{\mathrm{d}}{\mathrm{d}z}A(z)^{-1} = -A(z)^{-1}A'(z)A(z)^{-1}. \qquad \square$$

Proof of Theorem 15.2. The closedness of a_z means that the space $(V, \|\cdot\|_{a_z})$, with the norm

$$\|u\|_{a_z} = \bigl(\operatorname{Re} a_z(u) + \|u\|_H^2\bigr)^{1/2} \qquad (u \in V),$$

is complete. Using the continuity of the embeddings $(V, \|\cdot\|_{a_z}) \hookrightarrow (H, \|\cdot\|_H)$ and applying the closed graph theorem we conclude that the norms $\|\cdot\|_{a_z}$ are pairwise equivalent. For notational convenience we can therefore make V a Hilbert space with a norm $\|\cdot\|_V$ equivalent to all the norms $\|\cdot\|_{a_z}$. We define the form $e\colon V \times V \to \mathbb{C}$ by $e(u, v) := (u \,|\, v)_H$. Then for $z \in \Omega$, $a_z + e$ is a bounded coercive form on V, by the equivalence of the norms $\|\cdot\|_{a_z}$ and $\|\cdot\|_V$.

Recall that the linear relation $I + A_z$ is associated with the form $a_z + e$; see Remark 14.2. Now we use the description of the inverse of $I + A_z$ given in Proposition 14.19: if we define $\mathcal{B}_z \in \mathcal{L}(V, V^*)$ by

$$\langle \mathcal{B}_z u, v\rangle := a_z(u,v) + (u \,|\, v)_H \qquad (u, v \in V), \tag{15.1}$$

then $\mathcal{B}_z^{-1} \in \mathcal{L}(V^*, V)$ and

$$(I + A_z)^{-1} = j\mathcal{B}_z^{-1}k \in \mathcal{L}(H), \tag{15.2}$$

where $j\colon V \hookrightarrow H$ is the embedding, and $k\colon H \hookrightarrow V^*$ is as in Proposition 14.19.

The definition (15.1) together with the hypotheses and Theorem 3.2 yield the holomorphy of the function $\Omega \ni z \mapsto \mathcal{B}(z) \in \mathcal{L}(V, V^*)$. Then Lemma 15.3 shows that $\Omega \ni z \mapsto \mathcal{B}_z^{-1} \in \mathcal{L}(V^*, V)$ is holomorphic. By (15.2) this implies the holomorphy of $\Omega \ni z \mapsto (I + A_z)^{-1} \in \mathcal{L}(H)$. □

Next we show that the holomorphic depencence on z of a family of generators implies that the corresponding degenerate strongly continuous semigroups also depend holomorphically on z. By $C_{\mathrm{b}}([0,\infty); X)$ we denote the space of bounded continuous functions from $[0,\infty)$ to X.

15.4 Theorem. *Let X be a complex Banach space. Let $\Omega \subseteq \mathbb{C}$ be open, $(A_z)_{z\in\Omega}$ a family of generators of degenerate strongly continuous semigroups T_z on X. Assume that $M := \sup\bigl\{\|T_z(t)\|;\ t \geqslant 0,\ z \in \Omega\bigr\} < \infty$ and that $\Omega \ni z \mapsto (I - A_z)^{-1} \in \mathcal{L}(X)$ is holomorphic. Then*

(a) *the function $\Omega \ni z \mapsto T_z(t) \in \mathcal{L}(X)$ is holomorphic for all $t \geqslant 0$,*

(b) *the function $\Omega \ni z \mapsto T_z(\cdot)x \in C_{\mathrm{b}}([0,\infty); X)$ is holomorphic for all $x \in X$.*

Proof. (a) It follows from the boundedness hypothesis that

$$\bigl\|(\lambda - A_z)^{-n}\bigr\| \leqslant \frac{M}{\lambda^n} \qquad (n \in \mathbb{N},\ \lambda > 0); \tag{15.3}$$

see Theorem 2.7. We show that the holomorphy hypothesis implies that $\Omega \ni z \mapsto (\lambda - A_z)^{-1}$ is holomorphic for all $\lambda > 0$. Put

$$U := \bigl\{\lambda > 0;\ \Omega \ni z \mapsto (\lambda - A_z)^{-1} \text{ holomorphic}\bigr\}.$$

It follows from (15.3) and Theorem 3.5 that U is a closed subset of $(0,\infty)$. Also, $U \neq \varnothing$ because $1 \in U$. If $\mu \in U$, then, using the resolvent equation, one concludes that

$$(\lambda - A_z)^{-1} = (\mu - A_z)^{-1}\bigl(I - (\lambda - \mu)(\mu - A_z)^{-1}\bigr)^{-1} \tag{15.4}$$

for all $z \in \Omega$ and all $\lambda > 0$ such that $|\lambda - \mu|\frac{M}{\mu} < 1$, and by Lemma 15.3 one sees that the right-hand side of (15.4) is holomorphic in z. This shows that U is open, and as $(0,\infty)$ is connected one concludes that $U = (0,\infty)$.

For $z \in \Omega$ let $P_z := T_z(0)$ denote the projection onto the 'active' subspace of T_z. Let $t > 0$. Then it follows from the exponential formula, Theorem 2.12, that

$$T_z(t)x = T_z(t)P_z x = \lim_{n\to\infty}\bigl(I - \tfrac{t}{n}A_z\bigr)^{-n}P_z x = \lim_{n\to\infty}\bigl(I - \tfrac{t}{n}A_z\bigr)^{-n}x \qquad (z \in \Omega,\ x \in X).$$

Inequality (15.3) implies that $\left\|(I-\frac{t}{n}A_z)^{-n}\right\| \leqslant M$ for all $z\in\Omega$. Now Corollary 3.6 yields the holomorphy of $z\mapsto T_z(t)$ for $t>0$. Another application of Corollary 3.6 shows that $z\mapsto T_z(0)=\text{s-lim}_{t\to 0}T_z(t)$ is holomorphic as well.

(b) Let $x\in X$ and define $F\colon \Omega\to C_{\mathrm{b}}([0,\infty);X)$ by $F(z):=[t\mapsto T_z(t)x]$. Then by part (a), the mapping $\Omega\ni z\mapsto F(z)(t)\in X$ is holomorphic for all $t\geqslant 0$. Since F is bounded (by M), it follows that F is holomorphic; see Exercise 15.1. □

15.3 The Trotter product formula for forms

We now come to the main result of this chapter. Not only the results of the present chapter treated so far, but also the topics presented in the previous two chapters will play a decisive role in the proof of this result. We first treat the case of symmetric forms; the more general case of sectorial forms will then be obtained as a corollary. For simplicity we only treat the case in which the associated semigroups are contractive; by a rescaling argument one easily sees that the results remain true in the quasi-contractive case.

Throughout this section H will be a Hilbert space, and the forms will be embedded forms in H.

15.5 Theorem (Trotter product formula for symmetric forms). *Let a,b be closed accretive symmetric forms in H, let $c:=a+b$, with domain* $\operatorname{dom}(c):=\operatorname{dom}(a)\cap\operatorname{dom}(b)$, *and let T_a,T_b,T_c be the degenerate strongly continuous semigroups associated with a,b,c, respectively. Then*

$$\bigl(T_a(\tfrac{t}{n})T_b(\tfrac{t}{n})\bigr)^n x\to T_c(t)x \qquad (n\to\infty),$$

uniformly for t in compact subsets of $[0,\infty)$, for all $x\in\overline{\operatorname{dom}(c)}$, and

$$\bigl(T_a(\tfrac{t}{n})T_b(\tfrac{t}{n})\bigr)^n x\to 0=T_c(t)x \qquad (n\to\infty),$$

uniformly for t in compact subsets of $(0,\infty)$, for all $x\in\operatorname{dom}(c)^\perp$.

In the proof of the above theorem we need some auxiliary results; the first lemma is of a general nature. We recall from Section 13.4 that every accretive self-adjoint operator A has a unique accretive self-adjoint square root $A^{1/2}$.

15.6 Lemma. (a) *Let $A,B\in\mathcal{L}(H)$, A self-adjoint and accretive, $(Ax\,|\,x)\leqslant \operatorname{Re}(Bx\,|\,x)$ for all $x\in H$. Then $\|(I+B)^{-1}A^{1/2}\|\leqslant 1$.*

(b) *Let $A,B\in\mathcal{L}(H)$ be self-adjoint and accretive. Then $|(ABx\,|\,x)|\leqslant\frac12\|A\|(Ax\,|\,x)+\frac12\|B\|(Bx\,|\,x)$ for all $x\in H$.*

Proof. (a) The hypotheses imply that B and B^* are accretive. We have $\|(I+B)^{-1}A^{1/2}\| = \left\|(A^{1/2})^*\bigl((I+B)^{-1}\bigr)^*\right\| = \|A^{1/2}(I+B^*)^{-1}\|$, and for the last norm we estimate

$$\begin{aligned}\left\|A^{1/2}(I+B^*)^{-1}x\right\|^2 &= \bigl(A(I+B^*)^{-1}x\,\big|\,(I+B^*)^{-1}x\bigr)\\ &\leqslant \operatorname{Re}\bigl((I+B)(I+B^*)^{-1}x\,\big|\,(I+B^*)^{-1}x\bigr)\\ &\leqslant \bigl|\bigl((I+B^*)^{-1}x\,\big|\,x\bigr)\bigr|\leqslant\|x\|^2\end{aligned}$$

for all $x\in H$. This proves the asserted inequality.

(b) Recall from Lemma 14.9 that $\|Ax\|^2\leqslant\|A\|(Ax\,|\,x)$, and analogously for B. This implies the assertion since $|(ABx\,|\,x)|=|(Ax\,|\,Bx)|\leqslant\frac12\|Ax\|^2+\frac12\|Bx\|^2$. □

For the proof of Theorem 15.5 as well as for two further auxiliary lemmas we fix the following additional notation. Let A be the accretive self-adjoint linear relation associated with a; then $-A$ is the generator of the degenerate strongly continuous semigroup T_a. We denote the orthogonal projection onto the 'active' subspace of T_a by $P_a := T_a(0)$, and by A_0 we denote the accretive self-adjoint operator associated with a as a form in $\overline{\mathrm{dom}(a)} = \mathrm{ran}(P_a)$. (In Section 14.1, the operator A_0 had been denoted by A_{a}, as the restriction of A to the 'active' subspace H_{a} of H. In the present context we avoid this notation because simultaneously using indices a and a might lead to confusion.) For the forms b and c we will use the corresponding notation; in particular, the accretive self-adjoint linear relations associated with b, c will be denoted by B, C, respectively.

For $t \geqslant 0$, $s > 0$ we define

$$F(t) := T_a(t)T_b(t), \quad R(s) := \tfrac{1}{s}\big(I - F(s)\big)$$

and

$$A(s) := \tfrac{1}{s}\big(I - T_a(s)\big), \quad B(s) := \tfrac{1}{s}\big(I - T_b(s)\big), \quad C(s) := A(s) + B(s).$$

It is easy to check that with these definitions one gets

$$R(s) = C(s) - sA(s)B(s). \tag{15.5}$$

The overall plan for the proof of Theorem 15.5 is to show that

$$\text{s-lim}_{s\to 0}\big(I + R(s)\big)^{-1} = \text{s-lim}_{s\to 0}\big(I + C(s)\big)^{-1} = (I + C)^{-1},$$

and then to apply the Chernoff product formula, Theorems 13.14 and 13.18. For the application of the latter theorem we will also need a symmetrisation procedure (in step (ii) of the proof of Theorem 15.5).

15.7 Lemma. *For all $s > 0$ one has*

$$\mathrm{Re}\big(R(s)x \,\big|\, x\big) \geqslant \tfrac{1}{2}\big(C(s)x \,\big|\, x\big) \geqslant \tfrac{1}{2}\big(A(s)x \,\big|\, x\big) \geqslant 0 \qquad (x \in H);$$

in particular, $R(s)$ is accretive.

Proof. Note that $0 \leqslant A(s) = \frac{1}{s}(I - T_a(s)) \leqslant \frac{1}{s}I$ because $0 \leqslant T_a(s) \leqslant I$, and similarly for $B(s)$. Thus, for $x \in H$ it follows from Lemma 15.6(b) that

$$\begin{aligned}\big|\big(sA(s)B(s)x \,\big|\, x\big)\big| &\leqslant \tfrac{s}{2}\|A(s)\|\big(A(s)x \,\big|\, x\big) + \tfrac{s}{2}\|B(s)\|\big(B(s)x \,\big|\, x\big) \\ &\leqslant \tfrac{1}{2}\big(A(s)x \,\big|\, x\big) + \tfrac{1}{2}\big(B(s)x \,\big|\, x\big) = \tfrac{1}{2}\big(C(s)x \,\big|\, x\big).\end{aligned}$$

This implies

$$\mathrm{Re}\big(R(s)x \,\big|\, x\big) = \big(C(s)x \,\big|\, x\big) - \mathrm{Re}\big(sA(s)B(s)x \,\big|\, x\big) \geqslant \tfrac{1}{2}\big(C(s)x \,\big|\, x\big).$$

Since $C(s) = A(s) + B(s) \geqslant A(s) \geqslant 0$, we obtain the assertion. □

15.8 Lemma. *One has*

(a) $\text{s-lim}_{s\to 0}\big(I + C(s)\big)^{-1} = (I + C)^{-1}$,

(b) $\text{s-lim}_{s\to 0}\, s^{1/2}B(s)\big(I + C(s)\big)^{-1} = 0$.

Proof. (a) Let $s > 0$. As $T_a(s)$ is self-adjoint and $\|T_a(s)\| \leqslant 1$, the form $H \times H \ni (x, y) \mapsto \frac{1}{s}\big((I - T_a(s))x \,\big|\, y\big)$ is bounded, symmetric and accretive, and similarly for T_b. The form associated with the bounded operator $C(s)$ is given by the sum of these two forms,

$$c_s(x, y) := (C(s)x \,|\, y) = \frac{1}{s}\big((I - T_a(s))x \,\big|\, y\big) + \frac{1}{s}\big((I - T_b(s))x \,\big|\, y\big).$$

We observe that $(0, \infty) \ni s \mapsto c_s(x) \in [0, \infty)$ is decreasing for all $x \in H$. Indeed, if $x \in \overline{\mathrm{dom}(a)}$, then Proposition 13.23(b) implies that $s \mapsto \frac{1}{s}\big((I - T_a(s))x \,\big|\, x\big)$ is decreasing, and therefore

$$s \mapsto \frac{1}{s}\big((I - T_a(s))(x + y) \,\big|\, x + y\big) = \frac{1}{s}\big((I - T_a(s))x \,\big|\, x\big) + \frac{1}{s}\|y\|^2$$

is decreasing for all $x \in \overline{\mathrm{dom}(a)}$, $y \in \mathrm{dom}(a)^\perp$. The analogous property holds for b.

If $0 \neq x \in \mathrm{dom}(a)^\perp$, then $T_a(s)x = 0$, therefore $c_s(x) \to \infty$ as $s \to 0$, and similarly for $0 \neq x \in \mathrm{dom}(b)^\perp$. For $x \in \overline{\mathrm{dom}(a)} \cap \overline{\mathrm{dom}(b)}$, Proposition 13.23(b) shows that $\sup_{s>0} c_s(x) < \infty$ if and only if $x \in \mathrm{dom}(a) \cap \mathrm{dom}(b)$, and that

$$\begin{aligned}\lim_{s\to 0} c_s(x, y) &= \lim_{s\to 0} \frac{1}{s}\big((I - T_a(s))x \,\big|\, y\big) + \lim_{s\to 0} \frac{1}{s}\big((I - T_b(s))x \,\big|\, y\big)\\ &= a(x, y) + b(x, y)\end{aligned}$$

for all $x, y \in \mathrm{dom}(a) \cap \mathrm{dom}(b)$. Using Theorem 14.10 we conclude that $\big(I + C(s)\big)^{-1} \to (I + C)^{-1}$ strongly as $s \to 0$.

(b) The main part of the proof consists in showing that

$$\begin{aligned}\text{s-lim}_{s\to 0} A(s)^{1/2}\big(I + C(s)\big)^{-1} &= A_0^{1/2}(I + C)^{-1},\\ \text{s-lim}_{s\to 0} B(s)^{1/2}\big(I + C(s)\big)^{-1} &= B_0^{1/2}(I + C)^{-1}.\end{aligned} \tag{15.6}$$

We fix $x \in H$ and put

$$y(s) := \big(I + C(s)\big)^{-1}x, \quad y_a(s) := A(s)^{1/2}y(s), \quad y_b(s) := B(s)^{1/2}y(s).$$

From part (a) we know that $y := \lim_{s\to 0} y(s) = (I + C)^{-1}x$.

We show that $P_a y_a(s) \to A_0^{1/2}y$ and $P_b y_b(s) \to B_0^{1/2}y$ weakly as $s \to 0$. We have $\|y(s)\| \leqslant \|x\|$, because $C(s)$ is accretive. It follows that

$$\begin{aligned}\|y(s)\|^2 + \|y_a(s)\|^2 + \|y_b(s)\|^2 &= \big(y(s) + A(s)y(s) + B(s)y(s) \,\big|\, y(s)\big)\\ &= (x \,|\, y(s)) \leqslant \|x\|\,\|y(s)\| \leqslant \|x\|^2.\end{aligned} \tag{15.7}$$

This inequality implies that the set $\{y_a(s);\ s > 0\}$ is bounded. Let $z \in \mathrm{dom}(a) = \mathrm{dom}(A_0^{1/2})$. Then Proposition 13.23(b) shows that $A(s)^{1/2}z \to A_0^{1/2}z$, and we obtain

$$\begin{aligned}(P_a y_a(s) \,|\, z) &= \big(P_a A(s)^{1/2}y(s) \,\big|\, z\big) = \big(y(s) \,\big|\, A(s)^{1/2}z\big)\\ &\to \big(y \,\big|\, A_0^{1/2}z\big) = \big(A_0^{1/2}y \,\big|\, z\big) \qquad (s \to 0).\end{aligned}$$

Taking into account $\operatorname{ran}(P_a) = \overline{\operatorname{dom}(a)}$ and $A_0^{1/2}y \in \operatorname{ran}(P_a)$ as well as the boundedness shown above, we obtain the weak convergence $P_a y_a(s) \to A_0^{1/2}y$ asserted above. The corresponding weak convergence $P_b y_b(s) \to B_0^{1/2}y$ follows by the symmetry of the hypotheses with respect to a and b.

For the last part of the proof of (15.6) we note that the equation $y = (I + C)^{-1}x$ ($\in \operatorname{dom}(C_0) \subseteq \operatorname{dom}(c)$) can be reformulated as $(I + C_0)y = P_c x$. From (15.7) we infer, recalling $y(s) \to y = P_c y$, that

$$\begin{aligned}&\|y_a(s)\|^2 + \|y_b(s)\|^2 = (x \,|\, y(s)) - \|y(s)\|^2\\ &\quad\to (x \,|\, y) - \|y\|^2 = (P_c x - y \,|\, y) = (C_0 y \,|\, y) = a(y) + b(y) = \big\|A_0^{1/2}y\big\|^2 + \big\|B_0^{1/2}y\big\|^2\end{aligned}$$

as $s \to 0$. Since $P_a y_a(s) \to A_0^{1/2}y$ and $P_b y_b(s) \to B_0^{1/2}y$ weakly, we conclude that

$$\begin{aligned}\big\|y_a(s) - A_0^{1/2}y\big\|^2 + \big\|y_b(s) - B_0^{1/2}y\big\|^2 &= \|y_a(s)\|^2 - 2\operatorname{Re}\big(y_a(s) \,\big|\, P_a A_0^{1/2}y\big) + \big\|A_0^{1/2}y\big\|^2\\ &\quad + \|y_b(s)\|^2 - 2\operatorname{Re}\big(y_b(s) \,\big|\, P_b B_0^{1/2}y\big) + \big\|B_0^{1/2}y\big\|^2\end{aligned}$$

converges to 0 as $s \to 0$. This completes the proof of (15.6).

Now for the asserted convergence we note – with the notation from above – that

$$\big\|s^{1/2}B(s)\big(I + C(s)\big)^{-1}x\big\|^2 = \big\|s^{1/2}B(s)^{1/2}y_b(s)\big\|^2 = \big(sB(s)y_b(s) \,\big|\, y_b(s)\big),$$

and the latter converges to $\big((I - P_b)B_0^{1/2}y \,\big|\, B_0^{1/2}y\big)$ as $s \to 0$, which in fact equals 0 because $\operatorname{ran}(B_0^{1/2}) \subseteq \overline{\operatorname{dom}(b)} = \operatorname{ran}(P_b)$. □

Proof of Theorem 15.5. (i) In the first step we show that $\big(I + R(s)\big)^{-1} - \big(I + C(s)\big)^{-1} \to 0$ strongly as $s \to 0$. In view of Lemma 15.8(a) it will follow that $\text{s-lim}_{s\to 0}\big(I + R(s)\big)^{-1} = (I + C)^{-1}$, and applying Theorem 13.14 we then obtain the first assertion of the theorem. (Note that $F(0)x = P_a P_b x = x$ for all $x \in \operatorname{ran}(T_c(0)) = \operatorname{ran}(P_c)$.)

Using (15.5) we see that

$$\begin{aligned}\big(I + R(s)\big)^{-1} - \big(I + C(s)\big)^{-1} &= \big(I + R(s)\big)^{-1} sA(s)B(s)\big(I + C(s)\big)^{-1}\\ &= \Big(\big(I + R(s)\big)^{-1}A(s)^{1/2}\Big)\big(sA(s)\big)^{1/2}\Big(s^{1/2}B(s)\big(I + C(s)\big)^{-1}\Big).\end{aligned}$$

For the first factor on the right-hand side, the application of Lemmas 15.6 and 15.7 yields

$$\big\|\big(I + R(s)\big)^{-1}A(s)^{1/2}\big\| \leqslant \sqrt{2} \qquad (s > 0),$$

and $\big\|\big(sA(s)\big)^{1/2}\big\| \leqslant 1$ since $sA(s) \leqslant I$. Therefore Lemma 15.8(b) implies the asserted strong convergence $\big(I + R(s)\big)^{-1} - \big(I + C(s)\big)^{-1} \to 0$.

(ii) In order to obtain the second assertion of the theorem we symmetrise the operators $F(t)$ and $R(s)$. For $t \geqslant 0$, $s > 0$ we define

$$G(t) := T_a(\tfrac{t}{2})T_b(t)T_a(\tfrac{t}{2}), \quad S(s) := \tfrac{1}{s}\big(I - G(s)\big);$$

note that $G(t)$ is an accretive self-adjoint contraction. Moreover

$$sI = s\big(I + S(s)\big)\big(I + S(s)\big)^{-1} = \big((1 + s)I - G(s)\big)\big(I + S(s)\big)^{-1},$$

so we obtain

$$\begin{aligned}(1+s)\bigl(I+S(s)\bigr)^{-1} &= sI + T_a(\tfrac{s}{2})T_b(s)T_a(\tfrac{s}{2})\bigl(I+S(s)\bigr)^{-1}\\ &= sI + T_a(\tfrac{s}{2})T_b(s)\bigl(I+R(s)\bigr)^{-1}T_a(\tfrac{s}{2}),\end{aligned}$$

where the second equality is a consequence of

$$\bigl(I+R(s)\bigr)T_a(\tfrac{s}{2}) = T_a(\tfrac{s}{2})\bigl(I+S(s)\bigr).$$

It follows that

$$\begin{aligned}\text{s-lim}_{s\to 0}\bigl(I+S(s)\bigr)^{-1} &= \text{s-lim}_{s\to 0}\,\tfrac{1}{1+s}\Bigl(sI + T_a(\tfrac{s}{2})T_b(s)\bigl(I+R(s)\bigr)^{-1}T_a(\tfrac{s}{2})\Bigr)\\ &= 0 + P_aP_b(I+C)^{-1}P_a = (I+C)^{-1},\end{aligned}$$

where in the last equality we have used the identities $(I+C)^{-1} = P_c(I+C)^{-1}P_c$ and $P_a(P_bP_c) = P_c = P_cP_a$ (the latter being true because $\operatorname{ran}(P_c) \subseteq \operatorname{ran}(P_a)\cap\operatorname{ran}(P_b)$).

Now Theorem 13.18 implies that $G(\frac{t}{n})^n x \to T_c(t)x$ as $n\to\infty$, uniformly for t in compact subsets of $(0,\infty)$, for all $x\in H$; recall also Remark 13.19. Remark 13.20 shows that the same holds for the convergence $G(\frac{t}{n})^{n-1}x \to T_c(t)x$, and we conclude that

$$F(\tfrac{t}{n})^n x = T_a(\tfrac{t}{2n})G(\tfrac{t}{n})^{n-1}T_a(\tfrac{t}{2n})T_b(\tfrac{t}{n})x \to P_aT_c(t)P_aP_bx = T_c(t)x \qquad (n\to\infty),$$

uniformly for t in compact subsets of $(0,\infty)$, for all $x\in H$. □

We now come to the striking application of Theorem 15.2 announced above.

15.9 Corollary (Trotter product formula for sectorial forms). *Let a and b be closed sectorial forms in a complex Hilbert space H, put $c := a+b$, and let T_a, T_b, T_c be the associated degenerate strongly continuous semigroups on H.*

Then one obtains the same conclusions as in Theorem 15.5.

Proof. There exists a constant $C>0$ such that $|\operatorname{Im} a(u)| \leqslant C\operatorname{Re} a(u)$ for all $u\in\operatorname{dom}(a)$ and $|\operatorname{Im} b(u)| \leqslant C\operatorname{Re} b(u)$ for all $u\in\operatorname{dom}(b)$; see the last paragraph of Section 5.1. Then for $z\in\Omega := \bigl\{z\in\mathbb{C};\ |\operatorname{Re} z| < \frac{1}{C}\bigr\}$ the forms

$$a_z := \operatorname{Re} a + z\operatorname{Im} a,\quad b_z := \operatorname{Re} b + z\operatorname{Im} b,\quad c_z := \operatorname{Re} c + z\operatorname{Im} c$$

are closed sectorial forms in H. Indeed, for $z\in\Omega$ and $u\in\operatorname{dom}(a)$ we have

$$\begin{gathered}\operatorname{Re} a_z(u) = \operatorname{Re} a(u) + (\operatorname{Re} z)\operatorname{Im} a(u) \geqslant (1-|\operatorname{Re} z|C)\operatorname{Re} a(u) \geqslant 0,\\ |\operatorname{Im} a_z(u)| = |\operatorname{Im} z||\operatorname{Im} a(u)| \leqslant |\operatorname{Im} z|\,C\operatorname{Re} a(u) \leqslant \frac{|\operatorname{Im} z|C}{1-|\operatorname{Re} z|C}\operatorname{Re} a_z(u)\end{gathered}$$

and similarly for b and c. Clearly, the mappings $z\mapsto a_z$, $z\mapsto b_z$, $z\mapsto c_z$ are holomorphic in the sense formulated in Theorem 15.2.

Applying first Theorem 15.2 and then Theorem 15.4 we obtain the holomorphic dependence $\Omega\ni z\mapsto T_{a_z}(t)\in\mathcal{L}(H)$ for all $t\geqslant 0$, and the same for T_{b_z} and T_{c_z}. (Note that the boundedness condition of Theorem 15.4 is satisfied since $\|T_{a_z}(t)\| \leqslant 1$ for all $t \geqslant 0$, and similarly for b_z and c_z.)

Fix $t_0 > 0$, $x \in \overline{\operatorname{dom}(c)}$, and let $n \in \mathbb{N}$. We define the function $F_n \colon \Omega \to C([0,t_0];H)$ by

$$F_n(z) := \Bigl[t \mapsto \bigl(T_{a_z}(\tfrac{t}{n})T_{b_z}(\tfrac{t}{n})\bigr)^n x\Bigr];$$

then $\|F_n(z)\| \leqslant \|x\|$ for all $z \in \Omega$. Furthermore, the function $\Omega \ni z \mapsto F_n(z)(t)$ is holomorphic for all $0 \leqslant t \leqslant t_0$. This implies that F_n is holomorphic; see Exercise 15.1. Finally, for $z \in \Omega \cap \mathbb{R}$ we already know from Theorem 15.5 that

$$F_n(z) \to T_{c_z}(\cdot)x|_{[0,t_0]} \qquad (n \to \infty).$$

Now Vitali's theorem, Theorem 15.1, shows that $F_n(z) \to T_{c_z}(\cdot)x|_{[0,t_0]}$ in $C([0,t_0];H)$ as $n \to \infty$, for all $z \in \Omega$. Observing that $a = a_{\mathrm{i}}$, $b = b_{\mathrm{i}}$ and $c = c_{\mathrm{i}}$, we obtain the first convergence asserted in Theorem 15.5.

The second assertion is proved analogously. □

15.4 Applications of the Trotter product formula

The following applications of the Trotter product formula, Theorem 15.5, illustrate the interaction between diffusion semigroups and domain restrictions or multiplication operators.

15.10 Example (Dirichlet boundary conditions via absorption). Let

$$a(u,v) := \int_{\mathbb{R}^n} \nabla u(x) \cdot \overline{\nabla v(x)}\,\mathrm{d}x$$

be the classical Dirichlet form, defined on $\operatorname{dom}(a) := H^1(\mathbb{R}^n)$. Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and define the form b by $\operatorname{dom}(b) := L_2(\Omega) = \{f \in L_2(\mathbb{R}^n);\ f = 0 \text{ a.e. on } \mathbb{R}^n \setminus \Omega\}$,

$$b(u,v) := 0 \qquad (u,v \in \operatorname{dom}(b)).$$

Then $\operatorname{dom}(a)$ is dense in $L_2(\mathbb{R}^n)$, and the C_0-semigroup T_a associated with a is governed by the heat equation. The degenerate strongly continuous semigroup T_b associated with b is given by $T_b(t) = P_\Omega$ $(t \geqslant 0)$, where P_Ω is the orthogonal projection onto $L_2(\Omega)$, $P_\Omega f = \mathbf{1}_\Omega f$ for all $f \in L_2(\mathbb{R}^n)$.

The form $c = a + b$ is then given by

$$\operatorname{dom}(c) = H^1(\mathbb{R}^n) \cap L_2(\Omega), \qquad c(u,v) = a(u,v) \quad (u,v \in \operatorname{dom}(c)),$$

and Theorem 15.5 implies that T_c is given by the interesting formula

$$T_c(t)f = \lim_{k\to\infty} \bigl(T_a(\tfrac{t}{k})P_\Omega\bigr)^k f,$$

where the convergence is uniform for t in compact subsets of $[0,\infty)$ if $f \in L_2(\Omega)$, and uniform for t in compact subsets of $(0,\infty)$ if $f \in L_2(\mathbb{R}^n)$.

Assuming additionally that Ω is bounded and has continuous boundary, one obtains a particularly interesting result. In this case the restricted semigroup $\big(T_c(t)|_{L_2(\Omega)}\big)_{t\geqslant 0}$ is the C_0-semigroup generated by the Dirichlet Laplacian in $L_2(\Omega)$. Indeed, it was shown in Proposition 7.10 that under these hypotheses one has $\operatorname{dom}(c) = H^1(\mathbb{R}^n) \cap L_2(\Omega) = H_0^1(\Omega)$.

We warn the reader that the previous statement does not hold without the hypothesis of continuous boundary, as the example $n = 1$, $\Omega = (-1, 0) \cup (0, 1)$ shows. Indeed, in this situation one has $\operatorname{dom}(c) = H_0^1(-1, 1)$, but $H_0^1(-1, 1)$ is not a subset of $H_0^1(\Omega)$. Clearly one can construct similar examples in any dimension.

We will see in Example 15.11(b) that there is another way of approximating the semigroup T_c. Placing ourselves into the context of Example 15.11 and defining $q\colon \mathbb{R} \to [0, \infty]$ by $q := \infty\mathbf{1}_{\mathbb{R}^n\setminus\Omega}$, we find that the semigroup T_c described above is the same as the one obtained in Example 15.11. △

15.11 Example. Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let

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

be the classical Dirichlet form, defined on $\operatorname{dom}(a) = H_0^1(\Omega)$ (corresponding to the Dirichlet Laplacian) or on $\operatorname{dom}(a) = H^1(\Omega)$ (corresponding to the Neumann Laplacian).

Let $q\colon \Omega \to [0, \infty]$ be measurable, and define

$$b(u, v) := \int_\Omega q(x)\,u(x)\,\overline{v(x)}\,\mathrm{d}x$$

for $u, v \in \operatorname{dom}(b) := \{u \in L_2(\Omega);\ q|u|^2 \in L_1(\Omega)\}$.

Then $c := a + b$ is a closed accretive symmetric form.

(a) One way of approximating T_c is by applying Theorem 15.5. Observe that the degenerate strongly continuous semigroup T_b associated with the form b is given by $T_b(t)f = \mathrm{e}^{-tq}f$ (with $\mathrm{e}^{-t\infty} := 0$); cf. Exercise 1.6(b). The application of Theorem 15.5 yields

$$\big(T_a(\tfrac{t}{k})\mathrm{e}^{-\frac{t}{k}q}\big)^k f \to T_c(t)f, \tag{15.8}$$

with the specifications concerning the sets of convergence according to the assertion in Theorem 15.5. It follows from (15.8) that all the operators $T_c(t)$, for $t > 0$, are sub-Markovian on $L_2(\Omega)$. We now show that there exists a measurable set $\Omega_0 \subseteq \Omega$ such that

$$H_0 := \overline{\operatorname{dom}(c)} = L_2(\Omega_0) = \{f \in L_2(\Omega);\ f = 0 \text{ a.e. on } \Omega \setminus \Omega_0\}; \tag{15.9}$$

then the restriction of the degenerate strongly continuous semigroup T_c associated with c to H_0 is a sub-Markovian C_0-semigroup.

Indeed, the strong continuity of T_c implies that $T_c(0) = \text{s-lim}_{t\to 0+} T_c(t)$ is also sub-Markovian; it is in fact a sub-Markovian projection. More strongly, the convergence (15.8) implies that $0 \leqslant T_c(t) \leqslant T_a(t)$ for all $t > 0$. (Here and in what follows, the inequality $A \leqslant B$ between two operators $A, B \in \mathcal{L}(L_2(\mu))$ denotes inequality in the order sense, i.e. $Af \leqslant Bf$ for all $f \in L_2(\Omega)_+$.) Therefore

$$T_c(0) = \text{s-lim}_{t\to 0} T_c(t) \leqslant \text{s-lim}_{t\to 0} T_a(t) = T_a(0) = I,$$

and consequently $T_c(0)$ is a projection satisfying $0 \leqslant T_c(0) \leqslant I$. By Proposition 15.12 below it follows that

$$H_0 := \overline{\mathrm{dom}(c)} = \mathrm{ran}(T_c(0))$$

is a closed **order ideal** in $L_2(\Omega)$, i.e., H_0 is a closed subspace, and $f \in H_0$, $g \in L_2(\mu)$, $|g| \leqslant |f|$ implies $g \in H_0$; therefore Exercise 15.2(c) shows that there exists a measurable subset $\Omega_0 \subseteq \Omega$ such that (15.9) holds.

(b) Another way of approximating the degenerate strongly continuous semigroup T_c is by using the forms b_k, defined by $\mathrm{dom}(b_k) := L_2(\Omega)$,

$$b_k(u, v) := \int_\Omega (q \wedge k) u \bar{v} \,\mathrm{d}x.$$

Then (b_k) is an increasing sequence of closed accretive symmetric forms, and $(a + b_k)_k$ converges to $c = a + b$ in the sense of Theorem 14.10. Let T_k be the C_0-semigroup associated with $a + b_k$. Then, using Theorem 14.10 together with Theorem 13.10 one concludes that $T_k(t)f \to T_c(t)f$ for all $f \in L_2(\Omega_0)$, uniformly for t in compact subsets of $[0, \infty)$, and that $T_k(t)f \to 0$ for all $f \in L_2(\Omega \setminus \Omega_0)$, uniformly for t in compact subsets of $(0, \infty)$. △

15.12 Proposition. *Let* (Ω, μ) *be a measure space. Let* $P \in \mathcal{L}(L_2(\mu))$ *be a projection satisfying* $0 \leqslant P \leqslant I$ *(in the order sense). Then* $\mathrm{ran}(P)$ *is a closed order ideal in* $L_2(\mu)$.

Proof. The closedness of $\mathrm{ran}(P) = \ker(I - P)$ is clear. Let $f \in \mathrm{ran}(P)$, $g \in L_2(\mu)$, $|g| \leqslant |f|$. Then $|f| = |Pf| \leqslant P|f|$ since $P \geqslant 0$; see Exercise 9.1(a). Note that $I - P \geqslant 0$ as well, so it follows that $|(I - P)g| \leqslant (I - P)|g| \leqslant (I - P)|f| \leqslant 0$. Therefore $(I - P)g = 0$, $g = Pg \in \mathrm{ran}(P)$. □

Notes

The elegant proof of Vitali's theorem, Theorem 15.1, is due to Arendt and Nikolski; see [ArNi00; Theorem 2.1]. Note that even for the case of a sequence of $\mathbb{C}$-valued functions the proof uses the concept of Banach space valued holomorphic functions.

Theorem 15.2 is due to Kato [Kat80; Chap. VII, Theorem 4.2]. Our proof is taken from [VoVo18]; the new feature in this proof is the use of the Lax–Milgram lemma in the form of Proposition 14.19. Theorem 15.4 is taken from [VoVo18] as well. The Trotter product formula has first been proved by Trotter [Tro59] for the case of contractive semigroups, with suitable hypotheses on the domains of the generators. The Trotter product formula for forms, Theorem 15.5, is due to Kato. Our proof follows [Kat78] but is somewhat simpler because we do not prove the theorem in the full generality stated there. Corollary 15.9, due to Simon, is contained in the Addendum in [Kat78].

There has been considerable interest in investigating variants of the Trotter product formula. One natural question is: when do the Trotter products converge in the operator norm or even stronger norms? This question is discussed in [CaZa99], [Tam00], [CaZa01a], [CaZa01b], [NSZ18], [Zag19; Chapter 5]. Results on numerical methods in connection with Trotter products can be found in [JaLu00], [Tha08].

Another issue is the "Lapidus problem", which arises in the context of Theorem 15.5. If $\mathbb{K} = \mathbb{C}$, then the semigroups T_a, T_b, T_c are bounded holomorphic semigroups of angle $\pi/2$, and as such they have strongly continuous extensions to the closed right half-plane (cf. Exercise 3.7(a)). In Exercise 15.3 the reader is asked to show that the Trotter approximation also holds on the open right half-plane. The Lapidus problem is the question of whether the Trotter approximation, for $x \in \overline{\operatorname{dom}(c)}$, holds on the imaginary axis as well. (It follows from [MaSh03; Subsection 3.1] that one cannot expect convergence for $x \notin \overline{\operatorname{dom}(c)}$.) We refer to [JoLa00; Problem 11.3.9 and Section 11.7], [Cac05], [ENZ11] for more information and partial results concerning this (open!) problem.

An interesting special case of the Trotter product formula is when one of the semigroups is given by $[0,\infty) \ni t \mapsto P$ (as in Example 14.1), with an orthogonal projection P; see Exercise 15.6(a). This kind of Trotter product has been studied by Matolcsi in the context of C_0-semigroups on a complex Hilbert space H. For a C_0-semigroup T on H with generator A it is shown in [Mat03; Theorem 3] that $\text{s-lim}_{t\to\infty}(T(\frac{t}{n})P)^n$ exists for all $t \geqslant 0$ and all orthogonal projections P in H if and only if $-A$ is associated with a densely defined closed quasi-sectorial form in H. Moreover, $\text{s-lim}_{t\to\infty}(T(\frac{t}{n})P)^n$ exists for all $t \geqslant 0$ and all *bounded* projections P in H if and only if A is bounded. (In both equivalences the sufficiency of the condition follows from Corollary 15.9; see Exercises 15.6 and 15.7.)

Now suppose that A is the unbounded generator of a C_0-semigroup on H. Then, by Matolcsi's result described above, there exists a bounded projection P in H such that $\text{s-lim}_{t\to\infty}(T(\frac{t}{n})P)^n$ does not exist. One easily finds an equivalent scalar product $[\cdot,\cdot]$ on H such that P is an orthogonal projection in $(H,[\cdot,\cdot])$. As a consequence, $-A$ is not associated with a closed quasi-sectorial form in $(H,[\cdot,\cdot])$ (cf. Exercise 15.6(a)). To summarise, for an unbounded generator A of a C_0-semigroup on H there always exists an equivalent scalar product $[\cdot,\cdot]$ on H such that $-A$ is not associated with a closed quasi-sectorial form a in $(H,[\cdot,\cdot])$. For a concrete example of this phenomenon we refer to Exercise 5.9(b).

In a different context, if T is a positive C_0-semigroup on $L_p(\Omega,\mu)$, where (Ω,μ) is a σ-finite measure space and $p \in [1,\infty)$, it is shown in [ArBa93; Theorem 5.3] that $S(t) := \text{s-lim}_{n\to\infty}(PT(\frac{t}{n}))^n = \text{s-lim}_{n\to\infty}(T(\frac{t}{n})P)^n$ exists for all $t \geqslant 0$, for all projections P in $L_p(\Omega,\mu)$ of the form $Pf = \mathbf{1}_\omega f$ with measurable $\omega \subseteq \Omega$, and S thus defined is a degenerate strongly continuous semigroup. (A very simple case of this result is treated in Example 15.10.) In [MaSh03; Section 3.2] a counterexample is given showing that the above property does not hold for arbitrary positive contractive projections P.

Exercises

15.1 Let M be a topological space, X a complex Banach space, and define $C_{\mathrm{b}}(M;X)$ as the Banach space of bounded continuous functions $f\colon M \to X$, provided with the supremum norm. Let $\Omega \subseteq \mathbb{C}$ be an open set, $F\colon \Omega \to C_{\mathrm{b}}(M;X)$ a bounded function, and assume that there exists a dense subset $D \subseteq M$ such that $\Omega \ni z \mapsto F(z)(t) \in X$ is holomorphic for all $t \in D$.

Show that $F\colon \Omega \to C_{\mathrm{b}}(M;X)$ is holomorphic.

Hint: Show that $E := \{C_{\mathrm{b}}(M;X) \ni f \mapsto \langle f(t), x'\rangle;\ t \in D,\ x' \in X'\} \subseteq C_{\mathrm{b}}(M;X)'$ is a norming subset and apply Theorem 3.2.

15.2 Let (Ω,μ) be a measure space, and let $H_0 \subseteq L_2(\mu)$ be a closed order ideal in $L_2(\mu)$; see Example 15.11 for the definition. Let P denote the orthogonal projection onto H_0.

(a) Let $g \in L_2(\mu)$. Show that $g \perp H_0$ if and only if $|g| \wedge |f| = 0$ for all $f \in H_0$. Show that $|Pf| \wedge |(I-P)f| = 0$ for all $f \in L_2(\mu)$. Hint for '⇒' in the first part: For $f \in H_0$ one has $f \cdot \overline{\operatorname{sgn} f} \cdot \operatorname{sgn} g \in H_0$.

(b) Assume additionally that $\mu(\Omega) < \infty$. Show that there exists a measurable set $\Omega_0 \subseteq \Omega$ such that $P\mathbf{1}_\Omega = \mathbf{1}_{\Omega_0}$, and that $Pf = \mathbf{1}_{\Omega_0} f$ for all $f \in L_2(\Omega)$. (Hint: For the proof of the latter property first consider $0 \leqslant f \leqslant \mathbf{1}_\Omega$.)

(c) Assume that μ is σ-finite. Show that there exists a measurable set $\Omega_0 \subseteq \Omega$ such that $Pf = \mathbf{1}_{\Omega_0} f$ for all $f \in L_2(\Omega)$.

15.3 Let H be a complex Hilbert space, let a, b be closed accretive symmetric forms in H, and put $c := a + b$. Let T_a, T_b, T_c be the holomorphic degenerate strongly continuous semigroups of angle $\pi/2$ associated with a, b, c, respectively. (If A_0 denotes the self-adjoint operator associated with the form a in $\overline{\operatorname{dom}(a)}$ and P_a the orthogonal projection onto $\overline{\operatorname{dom}(a)}$, then $T_a(z) = \mathrm{e}^{-zA_0}P_a$ for all $z \in \mathbb{C}$ with $\operatorname{Re} z \geqslant 0$, and analogously for T_b, T_c.)

(a) Show that $\big(T_a(\frac{z}{n})T_b(\frac{z}{n})\big)^n x \to T_c(z)x$ as $n \to \infty$, uniformly for z in compact subset of the open right half-plane $[\operatorname{Re} > 0] \subseteq \mathbb{C}$, for all $x \in H$. (Hint: Use Vitali's theorem, Theorem 15.1.)

(b) Let $\theta \in (-\pi/2, \pi/2)$. Show that $\big(T_a(\frac{z}{n})T_b(\frac{z}{n})\big)^n x \to T_c(z)x$ as $n \to \infty$, uniformly for z in compact subsets of the ray $\mathrm{e}^{\mathrm{i}\theta}[0,\infty)$, for all $x \in \overline{\operatorname{dom}(c)}$. (Hint: Use Corollary 15.9.)

(c) Let $R > 0$, $\theta \in (0, \pi/2)$, and put $D := \big\{r\mathrm{e}^{\mathrm{i}\alpha};\ 0 \leqslant r \leqslant R,\ |\alpha| \leqslant \theta\big\}$. Show that $\big(T_a(\frac{z}{n})T_b(\frac{z}{n})\big)^n x \to T_c(z)x$ as $n \to \infty$, uniformly for $z \in D$, for all $x \in \overline{\operatorname{dom}(c)}$. (Hint: Combine (a) and (b), and apply Exercise 15.4.)

15.4 Prove the maximum principle for vector-valued holomorphic functions: if $\Omega \subseteq \mathbb{C}$ is a bounded open set, X is a complex Banach space, $f\colon \overline{\Omega} \to X$ is continuous, and $f|_\Omega$ is holomorphic, then $\|f(\zeta)\| \leqslant \sup_{z\in\partial\Omega} \|f(z)\|$ for all $\zeta \in \Omega$. (Hint: Use the fact that X' is norming for X.)

15.5 (a) Let $\Omega \subseteq \mathbb{C}$ be open and connected, and let H be a complex Hilbert space. Let $P\colon \Omega \to \mathcal{L}(H)$ be a holomorphic function taking its values in the set of bounded self-adjoint operators. Show that P is constant. (Note that this exercise applies, in particular, to holomorphic functions taking their values in the orthogonal projections.)

(b) Find a holomorphic function $P\colon \mathbb{C} \to \mathcal{L}(\mathbb{C}^2)$ taking its values in the projections that is not constant, where $\mathbb{C}^2$ is provided with the Euclidean norm.

(c) Find a holomorphic function $P\colon B(0,1) \to \mathcal{L}(\mathbb{C}^2)$ taking its values in the contractive projections that is not constant, where $\mathbb{C}^2$ is provided with the norm given by $\|(x,y)\|_1 = |x| + |y|$. (Hint: Look for a function satisfying $\operatorname{ran} P(z) = \mathbb{C} \times \{0\}$ for all $z \in B(0,1)$.)

15.6 (a) (**Trotter product formula for projections**) Let H be a Hilbert space, and let $P \in \mathcal{L}(H)$ be an orthogonal projection. Let a be a closed quasi-accretive symmetric form in H if $\mathbb{K} = \mathbb{R}$, or a closed quasi-sectorial form in H if $\mathbb{K} = \mathbb{C}$. Show that

$$S(t) := \text{s-lim}_{n\to\infty}(T_a(\tfrac{t}{n})P)^n = \text{s-lim}_{n\to\infty}(PT_a(\tfrac{t}{n}))^n$$

exists for all $t \geqslant 0$, and that $(S(t))_{t\geqslant 0}$ thus defined is a degenerate strongly continuous semigroup. (Hint: Apply Theorem 15.5 or Corollary 15.9, respectively, using a rescaling argument.)

(b) Let H be a Hilbert space, and let $P_1, P_2 \in \mathcal{L}(H)$ be orthogonal projections. Show that $Q := \text{s-lim}_{n\to\infty}(P_1P_2)^n$ exists and is the orthogonal projection onto $\operatorname{ran}(P_1) \cap \operatorname{ran}(P_2)$.

15.7 Let H be a Hilbert space, and let $P \in \mathcal{L}(H)$ be a projection.

(a) Show that there exists an equivalent scalar product $[\cdot,\cdot]$ on H such that P is an orthogonal projection in $(H, [\cdot,\cdot])$.

(b) Assume that $\mathbb{K} = \mathbb{C}$, and let T be a C_0-semigroup on H, with generator $A \in \mathcal{L}(H)$. Show that

$$S(t) := \text{s-lim}_{n\to\infty}(T(\tfrac{t}{n})P)^n = \text{s-lim}_{n\to\infty}(PT(\tfrac{t}{n}))^n$$

exists for all $t \geqslant 0$, and that $(S(t))_{t\geqslant 0}$ thus defined is a degenerate strongly continuous semigroup. (Hint: Show that every bounded operator A is associated with a closed quasi-sectorial form in $(H, [\cdot,\cdot])$ and apply Exercise 15.6(a).)

Note. Using a complexification procedure one can show that the assertion is also true for real Hilbert spaces.

Chapter 16

The Stokes operator

The Stokes operator arises in the context of the (nonlinear!) Navier–Stokes equation and acts in a subspace of a $\mathbb{K}^n$-valued L_2-space. In our context we define it using a variant of the classical Dirichlet form. One of the features appearing in the description of the Stokes operator is the use of a Sobolev space of negative order, which is introduced at the beginning. Another important feature is the appearance of divergence free vector fields. This extra condition of vanishing divergence has interesting implications for the theory of the related Sobolev spaces, and a large part of the chapter is devoted to the investigation of these properties.

16.1 Interlude: the Sobolev space $H^{-1}(\Omega)$

Let $\Omega \subseteq \mathbb{R}^n$ be open. The space $H_0^1(\Omega)$ is a Hilbert space, and by the Fréchet–Riesz theorem, each continuous antilinear functional on $H_0^1(\Omega)$ is represented by an element of $H_0^1(\Omega)$. For some purposes, however, it is more convenient to use a different representation of the antidual space of $H_0^1(\Omega)$ that contains $L_2(\Omega)$ als a dense subspace.

The basic observation is that each element $f \in L_2(\Omega)$ acts in a natural way as a continuous antilinear functional on $H_0^1(\Omega)$, by

$$H_0^1(\Omega) \ni u \mapsto (f \mid u)_{L_2(\Omega)} =: \langle f, u \rangle_{H^{-1}, H_0^1} .$$

The operator $L_2(\Omega) \ni f \mapsto \langle f, \cdot \rangle_{H^{-1}, H_0^1} \in H_0^1(\Omega)^*$ is injective, and it has dense range because $u \in H_0^1(\Omega)$, $\langle f, u \rangle_{H^{-1}, H_0^1} = 0$ for all $f \in L_2(\Omega)$ implies $u = 0$; see Exercise 17.1. In this context the antidual $H_0^1(\Omega)^*$ is denoted by $H^{-1}(\Omega)$.

16.1 Remark. We note that in the situation described above one has

$$H_0^1(\Omega) \hookrightarrow L_2(\Omega) \hookrightarrow H^{-1}(\Omega) = H_0^1(\Omega)^*,$$

with dense embeddings, and the embeddings are dual to each other. In such a situation one calls $\big(H_0^1(\Omega), L_2(\Omega), H^{-1}(\Omega)\big)$ a **Gelfand triple**; this concept will be treated in greater generality in Section 17.1. △

If $f \in L_2(\Omega)$ and $j \in \{1, \ldots, n\}$, then the mapping

$$H_0^1(\Omega) \ni u \mapsto \langle \partial_j f, u \rangle_{H^{-1}, H_0^1} := -(f \mid \partial_j u)_{L_2(\Omega)}$$

belongs to $H_0^1(\Omega)^*$. This definition of $\partial_j f$ is consistent with the definition of the distributional derivative in Section 4.1 because $C_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, and the mapping $\partial_j \colon L_2(\Omega) \to H^{-1}(\Omega)$ is linear and bounded. As a consequence, the differential operator $\Delta = \sum_{j=1}^n \partial_j\partial_j$ acts as a bounded operator $\Delta\colon H_0^1(\Omega) \to H^{-1}(\Omega)$. With this interpretation of the Laplace operator we obtain

$$\langle \Delta u, \varphi \rangle_{H^{-1}, H_0^1} = (u \mid \Delta\varphi)_{L_2(\Omega)}$$

for all $u \in H_0^1(\Omega)$, $\varphi \in C_c^\infty(\Omega)$.

It turns out that the mapping $I - \Delta\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isometric isomorphism. If Ω is bounded and one provides $H_0^1(\Omega)$ with the scalar product $(u, v) \mapsto \int \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x$, then $-\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isometric isomorphism. See Exercise 16.1 for these properties. (Strictly speaking, the notation '$I - \Delta$' is not quite correct: the identity I is meant to be the embedding $H_0^1(\Omega) \hookrightarrow H^{-1}(\Omega)$.)

16.2 The Stokes operator

Let $\Omega \subseteq \mathbb{R}^n$ be open. The Stokes operator is an operator in a subspace H of $L_2(\Omega; \mathbb{K}^n)$. We will define it as the operator associated with a variant of the classical Dirichlet form on

$$V := H_{0,\sigma}^1(\Omega; \mathbb{K}^n) := \big\{u \in H_0^1(\Omega; \mathbb{K}^n) = H_0^1(\Omega)^n\,;\ \operatorname{div} u = 0\big\}.$$

The index 'σ' should be remindful of 'solenoidal', which is the classical term for divergence free vectors fields. Note that V is a closed subspace of $H_0^1(\Omega; \mathbb{K}^n)$. The Hilbert space H is defined as the closure of V in $L_2(\Omega; \mathbb{K}^n)$,

$$H := L_{2,\sigma,0}(\Omega; \mathbb{K}^n) := \overline{V}^{L_2(\Omega;\mathbb{K}^n)}.$$

16.2 Remarks. (a) The space $L_{2,\sigma}(\Omega; \mathbb{K}^n) := \{f \in L_2(\Omega; \mathbb{K}^n)\,;\ \operatorname{div} f = 0\}$ is a closed subspace of $L_2(\Omega; \mathbb{K}^n)$; this is because $\operatorname{div}\colon L_2(\Omega; \mathbb{K}^n) \to H^{-1}(\Omega)$ is continuous. Therefore H is a (closed) subspace of $L_{2,\sigma}(\Omega; \mathbb{K}^n)$.

We point out that $C_c^\infty(\Omega; \mathbb{K}^n) \cap H$ need not be dense in H, so that the notation $L_{2,\sigma,0}$ is not entirely consistent with previous notation. (There is no such reservation for sets Ω that are bounded and have Lipschitz boundary; cf. Theorem 16.14.)

(b) One could consider the form a defined below as a non-densely defined form in the Hilbert space $L_2(\Omega, \mathbb{K}^n)$; cf. Chapter 14. In the present chapter we are not interested in this point of view, instead we work with the Hilbert space H adapted to the form domain V. △

We define the form $a\colon V \times V \to \mathbb{R}$ by

$$a(u, v) := \sum_{j=1}^n \int_\Omega \nabla u_j \cdot \overline{\nabla v_j}\,\mathrm{d}x.$$

In each component of $H_0^1(\Omega)^n$, the form a is the classical Dirichlet form. Therefore we conclude from Section 5.4 that a is symmetric, accretive and quasi-coercive; for bounded Ω it is coercive. It follows that the operator

$$A := \Big\{(u,f) \in V \times H\,;\ \sum_{j=1}^n \int_\Omega \nabla u_j \cdot \overline{\nabla v_j}\,\mathrm{d}x = (f \mid v)_H\ (v \in V)\Big\} \tag{16.1}$$

associated with a is self-adjoint and accretive. In view of Section 16.1, the equality appearing in the description of A can be rewritten as

$$0 = -\sum_{j=1}^n \int_\Omega \nabla u_j \cdot \overline{\nabla v_j}\,\mathrm{d}x + (f \mid v)_H = \sum_{j=1}^n \langle \Delta u_j + f_j, v_j\rangle_{H^{-1},H_0^1} \qquad (v \in V),$$

where f_j is considered as an element of $H^{-1}(\Omega)$ via the injection $L_2(\Omega) \hookrightarrow H^{-1}(\Omega)$, and Δ is the operator $\Delta\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ from Section 16.1.

Note that the antidual of $H_0^1(\Omega)^n$ is given by $H^{-1}(\Omega)^n$, with the dual pairing

$$\langle \eta, u\rangle_{H^{-1},H_0^1} := \sum_{j=1}^n \langle \eta_j, u_j\rangle_{H^{-1},H_0^1} \qquad \big(\eta \in H^{-1}(\Omega)^n,\ u \in H_0^1(\Omega)^n\big).$$

Then the condition in (16.1) becomes

$$\langle \Delta u + f, v\rangle_{H^{-1},H_0^1} = 0 \qquad (v \in V),$$

with the abbreviating notation $\Delta u = (\Delta u_1, \dots, \Delta u_n)$. Defining the polar of $V \subseteq H_0^1(\Omega)^n$,

$$V^\circ := \big\{\eta \in H^{-1}(\Omega)^n\,;\ \langle \eta, v\rangle_{H^{-1},H_0^1} = 0\ (v \in V)\big\},$$

as the 'orthogonal complement' of V in $H^{-1}(\Omega)^n$ we can write the condition in (16.1) in the concise form

$$\eta := \Delta u + f \in V^\circ.$$

Summarising, we obtain the following description of the **Stokes operator**; as mentioned above, it is self-adjoint and accretive.

16.3 Theorem. *The operator A in H associated with the form a is given by*

$$A = \big\{(u,f) \in V \times H\,;\ \exists \eta \in V^\circ\colon -\Delta u + \eta = f\big\}$$

(where the equality '$-\Delta u + \eta = f$' is an equality in $H^{-1}(\Omega)^n$, with $f \in H \hookrightarrow H^{-1}(\Omega)^n$). Written differently,

$$\begin{aligned} \operatorname{dom}(A) &= \big\{u \in V\,;\ \exists \eta \in V^\circ\colon -\Delta u + \eta \in H\big\},\\ Au &= -\Delta u + \eta \quad \textit{(with } \eta \textit{ as in } \operatorname{dom}(A)\textit{)}. \end{aligned}$$

Note that for $f \in L_2(\Omega)$ one has $\langle \nabla f, v\rangle_{H^{-1},H_0^1} = (f \,|\, \operatorname{div} v)_H = 0$ for all $v \in V$, so the gradient $\nabla f \in H^{-1}(\Omega)^n$ belongs to V°. For the physical interpretation of the Stokes operator one requires that *all* elements $\eta \in V^\circ$ appearing in the description of A can be expressed in this way:

$$\text{for all } \eta \in V^\circ \text{ there exists } p \in L_2(\Omega) \text{ with } \eta = \nabla p. \tag{H$_0$}$$

(The index 0 in the label (H_0) refers to the index 0 in $V = H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$.)

We will comment on hypothesis (H_0) in Remark 16.6. In Section 16.4 we show that it holds if Ω is bounded and has Lipschitz boundary. With this hypothesis we get another description of the Stokes operator A.

16.4 Theorem. *Assume that Ω satisfies hypothesis* (H_0)*. Then*

$$A = \bigl\{(u,f) \in V \times H\,;\ \exists p \in L_2(\Omega)\colon -\Delta u + \nabla p = f\bigr\}$$

is the operator associated with the form a. Expressed differently,

$$\operatorname{dom}(A) = \bigl\{u \in V\,;\ \exists p \in L_2(\Omega)\colon -\Delta u + \nabla p \in H\bigr\},$$
$$Au = -\Delta u + \nabla p \quad (\textit{with } p \textit{ as in } \operatorname{dom}(A)).$$

In fluid dynamics one interprets p in the statement of Theorem 16.4 as a pressure, which is the reason for the notation. However, we will not endeavour to enter the physical interpretation of the Stokes operator.

In the following remarks we describe some conditions and properties related to (H_0) in the 'classical' context.

16.5 Remarks. Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $u\colon \Omega \to \mathbb{K}^n$ be a continuous vector field.

(a) Assume that u satisfies the condition

$$\int u \cdot \varphi \,\mathrm{d}x = 0 \quad \text{for all } \varphi \in C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n) \text{ with } \operatorname{div}\varphi = 0. \tag{16.2}$$

(This condition should be considered as a weakened version of the hypothesis '$\eta \in V^\circ$' in (H_0); note that $\{\varphi \in C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n)\,;\ \operatorname{div}\varphi = 0\}$ need not be dense in V.) Then one can show that there exists a potential $p \in C^1(\Omega)$ for u, i.e. $\nabla p = u$. Clearly it is sufficient to treat the case when Ω is connected. Then, fixing an 'initial point' $x^0 \in \Omega$, one defines $p(x) := \int_0^1 u(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t$, where $\gamma\colon [0,1] \to \Omega$ is a piecewise continuously differentiable path connecting x^0 with x. Using (16.2) one can show that p is well-defined – this is the main issue – and that $u = \nabla p$. Conversely, the existence of a potential for the vector field implies the validity of (16.2).

The proof of these statements is delegated to Exercise 16.2(a).

(b) If u is continuously differentiable and satisfies (16.2), then it satisfies the 'compatibility condition' $\partial_j u_k = \partial_k u_j$ for all $j,k \in \{1,\dots,n\}$. Indeed, let $\psi \in C_\mathrm{c}^\infty(\Omega)$, and put $\varphi_j := \partial_k\psi$, $\varphi_k := -\partial_j\psi$, $\varphi_\ell := 0$ for all other components (where without loss of generality $j \neq k$). Then $\operatorname{div}\varphi = 0$, and therefore

$$\int (\partial_j u_k - \partial_k u_j)\psi \,\mathrm{d}x = \int (u_j\partial_k\psi - u_k\partial_j\psi)\,\mathrm{d}x = \int u \cdot \varphi \,\mathrm{d}x = 0.$$

If one requires from u merely the compatibility condition (and not the validity of (16.2)), then one can show that every point of Ω possesses a neighbourhood in which a potential exists, given by the explicit formula (16.4) in Section 16.3 below. An example of a vector field u without a global potential can be found in Exercise 16.2(b). However, if one adds the hypothesis that Ω is simply connected, then one can show that a global potential for u exists. This is a version of the 'Poincaré lemma'.

We will not prove these statements, because in this generality they will not be of importance in our context. However, it is noteworthy that the smoothed version (16.5) of the explicit formula mentioned above will be basic for the further development.

(c) We mention that both of the phenomena described in parts (a) and (b) can also be treated in the more general context of distributions. If u is a vector of distributions on Ω satisfying a condition analogous to (16.2), then one can show that there exists a distribution p such that $\nabla p = u$. (This is a special case of 'de Rham's theorem'.) Likewise, if Ω is simply connected, and a vector $u = (u_1, \dots, u_n)$ of distributions satisfies the compatibility condition $\partial_j u_k = \partial_k u_j$ for all $j, k \in \{1, \dots, n\}$, then one can find a distribution p on Ω with $\nabla p = u$. (This is a generalised version of Poincaré's lemma.) For a recent treatment of these topics we refer to [Voi23a],[Voi23b]. △

The topics described in the preceding remark do not explicitly enter our further treatment, but serve to motivate the investigation of corresponding properties in the context of Sobolev spaces, as sketched in the following remark.

16.6 Remark. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, and let $f \in L_2(\Omega)$. Then clearly $\eta := \nabla f \in H^{-1}(\Omega)^n$ has the following property corresponding to (16.2):

$$\langle \eta, \varphi \rangle = 0 \text{ for all } \varphi \in C_c^\infty(\Omega)^n \text{ satisfying } \operatorname{div} \varphi = 0. \tag{16.3}$$

In view of Remarks 16.5(a) and (c) it is natural to ask whether for any $\eta \in H^{-1}(\Omega)^n$ satisfying (16.3) there exists $f \in L_2(\Omega)$ such that $\nabla f = \eta$. (Note that the condition on η in (16.3) is weaker than '$\eta \in V^\circ$'; so the requirement that η satisfying (16.3) can be represented as $\eta = \nabla f$ – later this will be defined as property $(\mathrm{H_c})$ – is stronger than $(\mathrm{H_0})$.) A positive answer will be given in Theorem 16.14 for the case when Ω has Lipschitz boundary.

The remaining sections of the present chapter are mainly motivated by this topic. △

16.3 Interlude: the Bogovskiĭ formula

This section could also run under the heading "some functions are divergences". For $u \in C_c^\infty(\mathbb{R}^n; \mathbb{K}^n)$ one has $\int \operatorname{div} u \, \mathrm{d}x = 0$, and in fact the converse holds as well: for each $\varphi \in C_c^\infty(\mathbb{R}^n)$ with $\int \varphi \, \mathrm{d}x = 0$ there exists $u \in C_c^\infty(\mathbb{R}^n; \mathbb{K}^n)$ with $\operatorname{div} u = \varphi$. Here we show that for suitable bounded open $\Omega \subseteq \mathbb{R}^n$ one has a Sobolev space version of this statement: for any $f \in L_2(\Omega)$ with $\int f \, \mathrm{d}x = 0$ there exists $u \in H_0^1(\Omega)^n$ such that $\operatorname{div} u = f$; see Theorem 16.11.

Let $u \in C^\infty(\mathbb{R}^n; \mathbb{K}^n)$ be a vector field satisfying the compatibility condition $\partial_j u_k = \partial_k u_j$ for all $j, k \in \{1, \dots, n\}$. Then it is well-known and easy to show that, for any $y \in \mathbb{R}^n$, a potential for u is given by

$$p(x) := \int_0^1 u(y + t(x-y)) \cdot (x-y)\,\mathrm{d}t, \tag{16.4}$$

i.e. $u = \nabla p$. We smooth this formula out with the help of a function $\rho \in C_c^\infty(\mathbb{R}^n)_+$ satisfying $\int \rho(x)\,\mathrm{d}x = 1$: we define

$$Au(x) := \int \rho(y) \int_0^1 u(sy + (1-s)x) \cdot (x-y)\,\mathrm{d}s\,\mathrm{d}y \tag{16.5}$$

and obtain $Au \in C^\infty(\mathbb{R}^n)$, $\nabla(Au) = u$. In order to write A as an integral operator, we substitute $z = sy + (1-s)x$ and $r = \frac{1}{s}$ to obtain

$$\begin{aligned} Au(x) &= \int_0^1 \int \rho\bigl(\tfrac{1}{s}(z-(1-s)x)\bigr)u(z) \cdot \frac{x-z}{s}\, s^{-n}\,\mathrm{d}z\,\mathrm{d}s \\ &= \int_1^\infty \int \rho(x + r(z-x))u(z) \cdot (x-z)\,\mathrm{d}z\, r^{n-1}\,\mathrm{d}r. \end{aligned}$$

This means that one can write $Au(x) = \int k(x,y) \cdot u(y)\,\mathrm{d}y$, with

$$k(x,y) = \int_1^\infty \rho(x + r(y-x))r^{n-1}\,\mathrm{d}r\,(x-y).$$

Let ℓ be the negative transposed kernel of k, i.e.

$$\ell(x,y) := -k(y,x) = (x-y)\int_1^\infty \rho(y + r(x-y))r^{n-1}\,\mathrm{d}r \qquad (x,y \in \mathbb{R}^n).$$

It will be shown in the next theorem that the definition

$$\begin{aligned} Bf(x) &:= \int \ell(x,y)f(y)\,\mathrm{d}y \\ &= \int f(y)(x-y)\int_1^\infty \rho(y + r(x-y))r^{n-1}\,\mathrm{d}r\,\mathrm{d}y, \end{aligned} \tag{16.6}$$

for $x \in \mathbb{R}^n$ and $f \in C_c^\infty(\mathbb{R}^n)$, yields a mapping $B\colon C_c^\infty(\mathbb{R}^n) \to C_c^\infty(\mathbb{R}^n;\mathbb{K}^n)$. This definition is such that for all $u \in C_c^\infty(\mathbb{R}^n;\mathbb{K}^n)$, $f \in C_c^\infty(\mathbb{R}^n)$ one has $\int (Au)f\,\mathrm{d}x = -\int u \cdot Bf\,\mathrm{d}x$.

16.7 Theorem. *For all $f \in C_c^\infty(\mathbb{R}^n)$ one has $Bf \in C_c^\infty(\mathbb{R}^n;\mathbb{K}^n)$,*

$$\operatorname{spt}(Bf) \subseteq \bigl\{\lambda z_1 + (1-\lambda)z_2\,;\ z_1 \in \operatorname{spt} f,\ z_2 \in \operatorname{spt}\rho,\ 0 \leqslant \lambda \leqslant 1\bigr\} =: E. \tag{16.7}$$

If $\int f(x)\,\mathrm{d}x = 0$, then $\operatorname{div} Bf = f$.

Proof. Let $f \in C_c^\infty(\mathbb{R}^n)$.

(i) First we show that $Bf = 0$ on $\mathbb{R}^n \setminus E$. Note that E is a compact set (since $\operatorname{spt} f$, $\operatorname{spt}\rho$ and $[0,1]$ are compact). Let $x \in \mathbb{R}^n \setminus E$. If $y \in \operatorname{spt} f$ and $r \geqslant 1$ then $y + r(x-y) \notin \operatorname{spt}\rho$ (because $z = y + r(x-y) \in \operatorname{spt}\rho$ would lead to $x = \frac{1}{r}z + (1-\frac{1}{r})y \in E$ – a contradiction), and therefore $\rho(y + r(x-y)) = 0$. Hence (16.6) implies $Bf(x) = 0$.

(ii) By the variable transformation $z = x - y$ and then $r = 1 + \frac{s}{|z|}$ in the inner integral we can write Bf in the form

$$Bf(x) = \int f(x-z)\frac{z}{|z|^n}\int_0^\infty \rho\Big(x + s\frac{z}{|z|}\Big)(s+|z|)^{n-1}\,\mathrm{d}s\,\mathrm{d}z.$$

In this form one can differentiate under the integral to obtain $Bf \in C^\infty(\mathbb{R}^n;\mathbb{K}^n)$. (If $R > 0$ is such that ρ, f have their supports in $B(0,R)$, then we know from step (i) that $Bf = 0$ on $\mathbb{R}^n \setminus B(0,R)$, and for $x \in B(0,R)$ we can use a multiple of $(s,z) \mapsto \mathbf{1}_{(0,2R)}(s)\mathbf{1}_{B(0,2R)}(z)|z|^{1-n}$ as a dominating function.)

(iii) Let $f,\varphi \in C_\mathrm{c}^\infty(\mathbb{R}^n)$, $\int f\,\mathrm{d}x = 0$. Note that $\nabla\varphi$ satisfies the compatibility condition, therefore $\nabla A\nabla\varphi = \nabla\varphi$, $\nabla(A\nabla\varphi - \varphi) = 0$, i.e. $A\nabla\varphi - \varphi$ is constant. It follows that $\int f(A\nabla\varphi - \varphi)\,\mathrm{d}x = 0$, and we obtain

$$\int(\operatorname{div} Bf)\varphi\,\mathrm{d}x = -\int Bf\cdot\nabla\varphi\,\mathrm{d}x = \int fA\nabla\varphi\,\mathrm{d}x = \int f\varphi\,\mathrm{d}x.$$

As this holds for all $\varphi \in C_\mathrm{c}^\infty(\mathbb{R}^n)$ we conclude that $\operatorname{div} Bf = f$. □

16.8 Remarks. (a) The formula (16.6) is the **Bogovskiĭ formula**. The linear mapping B in (16.6) as well as its extension obtained in Theorem 16.10 below is the **Bogovskiĭ operator**. (The reader should keep in mind the fact that the Bogovskiĭ operator depends on the function ρ; so the use of the definite article 'the' might be somewhat misleading.)

(b) If $n \geqslant 2$, then there exist vector fields $0 \neq u \in C_\mathrm{c}^\infty(\mathbb{R}^n;\mathbb{K}^n)$ satisfying $\operatorname{div} u = 0$. Therefore the solution of $\operatorname{div} u = f$ obtained by the Bogovskiĭ formula is not unique. For the Bogovskiĭ formula in dimension $n = 1$ we refer to Exercise 16.3.

(c) Let $\Omega \subseteq \mathbb{R}^n$ be an open set containing a ball $B(x^0,r)$ with the property that Ω is star-shaped with respect to each point of $B(x^0,r)$. Let $\rho \in C_\mathrm{c}^\infty(\mathbb{R}^n)_+$ satisfy $\operatorname{spt}\rho \subseteq B(x^0,r)$ and $\int \rho(x)\,\mathrm{d}x = 1$. Then (16.7) implies that $B(C_\mathrm{c}^\infty(\Omega)) \subseteq C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n)$, with the identification $C_\mathrm{c}^\infty(\Omega) = \{\varphi \in C_\mathrm{c}^\infty(\mathbb{R}^n)\,;\, \operatorname{spt}\varphi \subseteq \Omega\}$. In this sense, B is a linear mapping from $C_\mathrm{c}^\infty(\Omega)$ to $C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n)$. Thus, an important consequence of Theorem 16.7 is that for each $\varphi \in C_\mathrm{c}^\infty(\Omega)$ with $\int \varphi(x)\,\mathrm{d}x = 0$ there exists a vector field Φ $(= B\varphi) \in C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n)$ such that $\operatorname{div}\Phi = \varphi$. △

We now show that the last property mentioned in Remark 16.8(c) carries over to more general open sets.

16.9 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be a connected open set, and let $\varphi \in C_\mathrm{c}^\infty(\Omega)$ satisfy $\int \varphi(x)\,\mathrm{d}x = 0$. Then there exists a vector field $\Phi \in C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n)$ such that* $\operatorname{div}\Phi = \varphi$.

Proof. (i) In the first step we prove the assertion for functions of the special form $\varphi = \hat\varphi - \check\varphi$, with $\check\varphi, \hat\varphi \in C_\mathrm{c}^\infty(\Omega)$, $\int \check\varphi(x)\,\mathrm{d}x = \int \hat\varphi(x)\,\mathrm{d}x$, and $\operatorname{spt}\check\varphi \subseteq \check B$, $\operatorname{spt}\hat\varphi \subseteq \hat B$ for open balls $\check B, \hat B \subseteq \Omega$.

There exist open balls $\check B = B_0, B_1, \dots, B_m = \hat B \subseteq \Omega$ such that $B_{j-1} \cap B_j \neq \varnothing$ for $j = 1,\dots,m$. Choose $\varphi_1,\dots,\varphi_m \in C_\mathrm{c}^\infty(\Omega)$ such that $\int \varphi_j(x)\,\mathrm{d}x = \int \check\varphi(x)\,\mathrm{d}x$ and $\operatorname{spt}\varphi_j \subseteq B_{j-1}\cap B_j$ for $j = 1,\dots,m$, and put $\varphi_0 := \check\varphi$, $\varphi_{m+1} := \hat\varphi$. By Remark 16.8(c) there exist $\Phi_0,\dots,\Phi_m \in C_\mathrm{c}^\infty(\Omega;\mathbb{K}^n)$ satisfying $\operatorname{spt}\Phi_j \subseteq B_j$ and $\operatorname{div}\Phi_j = \varphi_{j+1} - \varphi_j$ for $j = 0,\dots,m$. Then the vector field $\Phi := \sum_{j=0}^m \Phi_j$ is as asserted.

(ii) For the general case we find a covering $(B_j)_{j=1,\dots,m}$ of $\operatorname{spt}\varphi$ by open balls $B_j \subseteq \Omega$. There exists a subordinate partition of unity $(\chi_j)_{j=1,\dots,m}$ in $C_c^\infty(\Omega)_+$ on $\operatorname{spt}\varphi$, i.e. a family of functions with $\operatorname{spt}\chi_j \subseteq B_j$ for all $j \in \{1,\dots,m\}$ and $\sum_{j=1}^m \chi_j(x) = 1$ for all $x \in \operatorname{spt}\varphi$; see Exercise 4.3(b). Let $B_0 \subseteq \Omega$ be an open ball, and choose $\varphi_0 \in C_c^\infty(\Omega)$ with $\operatorname{spt}\varphi_0 \subseteq B_0$ and $\int \varphi_0(x)\,dx = 1$. For $j = 1,\dots,m$ we put $\varphi_j := \chi_j\varphi$ and apply step (i) with $\check\varphi := (\int \varphi_j)\varphi_0$ and $\hat\varphi := \varphi_j$ to obtain $\Phi_j \in C_c^\infty(\Omega;\mathbb{K}^n)$ satisfying $\operatorname{div}\Phi_j = \varphi_j - (\int \varphi_j)\varphi_0$. Then $\Phi := \sum_{j=1}^m \Phi_j$ satisfies $\operatorname{div}\Phi = \big(\sum_{j=1}^m \chi_j\big)\varphi - (\int\varphi)\varphi_0 = \varphi$. □

The next issue is extending the Bogovskiĭ operator to a bounded linear operator $B\colon L_2(\Omega) \to H_0^1(\Omega;\mathbb{K}^n)$, for suitable bounded open Ω. Then the equality $\operatorname{div} Bf = f$ carries over to all $f \in L_2^0(\Omega) := \big\{f \in L_2(\Omega);\ \int f(x)\,dx = 0\big\}$ because $C_c^\infty \cap L_2^0(\Omega)$ is dense in $L_2^0(\Omega)$; see Exercise 16.4(a). We confine ourselves to a special form of the 'smoothing function' ρ.

16.10 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set containing a ball $B(x^0, r_0)$ with the property that Ω is star-shaped with respect to every point of $B(x^0, r_0)$. Let $\rho_0, \tilde\rho \in C_c^\infty(\mathbb{R}^n)$ be such that $\operatorname{spt}\rho_0 \subseteq B(0, r_0/2)$, $\operatorname{spt}\tilde\rho \subseteq B(x^0, r_0/2)$, $\int \rho_0(x)\,dx = \int \tilde\rho(x)\,dx = 1$, and let $\rho := \rho_0 * \tilde\rho$. (Note that $\rho \in C_c^\infty(\mathbb{R}^n)$, $\operatorname{spt}\rho \subseteq B(x^0, r_0)$ and $\int \rho(x)\,dx = 1$.)*

Then $B\colon C_c^\infty(\Omega) \to C_c^\infty(\Omega;\mathbb{K}^n)$, defined in Remark 16.8(c), has a continuous (linear) extension $B\colon L_2(\Omega) \to H_0^1(\Omega;\mathbb{K}^n)$. For all $f \in L_2^0(\Omega)$ one has $\operatorname{div} Bf = f$.

The proof of Theorem 16.10 is delegated to Appendix H. Showing that there exists $c > 0$ such that $\|Bf\|_2 \leqslant c\|f\|_2$ is not the problem; see Exercise 16.5. The hard part is the corresponding estimate for the derivatives of Bf. In order to stay with our philosophy to provide complete information on the treated topics we present the proof of the 'hard part' mentioned above in Appendix H, Section H.3.

Traditionally, for proving Theorem 16.10 one uses the Calderón–Zygmund theory of singular integral operators; see Section H.4 and the Notes of Appendix H. Our hypothesis of the special form of $\rho = \rho_0 * \tilde\rho$ helps avoiding harder parts of the Calderón–Zygmund theory and thereby facilitates a much easier access. The restriction to this special form of the Bogovskiĭ operator has no restricting effect on the remaining results because we really need only *one* operator $B\colon C_c^\infty(\Omega) \to C_c^\infty(\Omega;\mathbb{K}^n)$ with the continuity property and the property that $\operatorname{div} Bf = f$ for all $f \in C_c^\infty(\Omega)$ with $\int f(x)\,dx = 0$.

An important consequence of Theorem 16.10 is the surjectivity of the mapping $\operatorname{div}\colon H_0^1(\Omega) \to L_2^0(\Omega)$. (Note that $\operatorname{div}(C_c^\infty(\Omega;\mathbb{K}^n)) \subseteq L_2^0(\Omega)$ and thus $\operatorname{div}(H_0^1(\Omega)) \subseteq L_2^0(\Omega)$ for *any* bounded open $\Omega \subseteq \mathbb{R}^n$.) We now show that Theorem 16.10 implies the surjectivity of $\operatorname{div}\colon H_0^1(\Omega) \to L_2^0(\Omega)$ for more general Ω.

16.11 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a connected bounded open set with Lipschitz boundary. Let $f \in L_2^0(\Omega)$. Then there exists $u \in H_0^1(\Omega;\mathbb{K}^n)$ such that $\operatorname{div} u = f$.*

Proof. (i) It is not difficult to see that for all $x \in \overline\Omega$ there exists an open neighbourhood U_x such that $U_x \cap \Omega$ is star-shaped with respect to the points of a ball in $U_x \cap \Omega$. This is obvious for $x \in \Omega$, and for $x \in \partial\Omega$ it results from the Lipschitz property of $\partial\Omega$. Compactness of $\overline\Omega$ implies that there exists a finite open covering $(\Omega_j)_{j=1,\dots,m}$ of Ω by sets to which Thereom 16.10 can be applied.

(ii) It is not too difficult to show that there exist functions $f_1, \dots, f_m \in L_2^0(\Omega)$ such that $[f_j \neq 0] \subseteq \Omega_j$ for all $j \in \{1, \dots, m\}$ and $f = \sum_{j=1}^m f_j$. (We explain the idea for $m = 2$: in that case $\Omega_1 \cap \Omega_2$ is non-empty, so we can choose $g \in L_2(\Omega)$ with $[g \neq 0] \subseteq \Omega_1 \cap \Omega_2$, $\int g \, \mathrm{d}x = \int \mathbf{1}_{\Omega_1} f \, \mathrm{d}x$, and put $f_1 := \mathbf{1}_{\Omega_1} f - g$, $f_2 := \mathbf{1}_{\Omega_2 \setminus \Omega_1} f + g$. The general case is delegated to Exercise 16.4(b).) Then for all $j \in \{1, \dots, m\}$ there exists $u^j \in H_0^1(\Omega_j; \mathbb{K}^n)$ with $\operatorname{div} u^j = f_j$, and $u := \sum_{j=1}^m u^j$ has the required properties. □

16.4 The hypothesis (H$_0$)

In this section we show that the hypothesis (H$_0$) is satisfied for bounded open sets $\Omega \subseteq \mathbb{R}^n$ with Lipschitz boundary.

16.12 Theorem. *Let Ω be as stated above. Then for any $\eta \in H_{0,\sigma}^1(\Omega; \mathbb{K}^n)^\circ$, i.e. $\eta \in H^{-1}(\Omega)^n$ with $\langle \eta, u \rangle_{H^{-1}, H_0^1} = 0$ for all $u \in H_{0,\sigma}^1(\Omega; \mathbb{K}^n)$, there exists $p \in L_2(\Omega)$ such that $\eta = \nabla p$. In other words, Ω satisfies hypothesis* (H$_0$).

For the proof we need the following special case of the 'closed range theorem'.

16.13 Proposition. *Let G, H be Hilbert spaces, and let $A \in \mathcal{L}(G, H)$ have closed range. Then* $\operatorname{ran}(A^*)$ *is closed,* $\operatorname{ran}(A^*) = \ker(A)^\perp$.

Proof. The bounded inverse theorem implies that the operator $A_1 := A|_{\ker(A)^\perp} \colon \ker(A)^\perp \to \operatorname{ran}(A)$ is boundedly invertible. It follows that $A_1^* \colon \operatorname{ran}(A) \to \ker(A)^\perp$ is boundedly invertible (with $(A_1^*)^{-1} = (A_1^{-1})^*$).

Let $J \colon \ker(A)^\perp \hookrightarrow G$ be the embedding, and let P be the orthogonal projection from H onto $\operatorname{ran}(A) = \ker(A^*)^\perp$. Then J^* is the orthogonal projection from G onto $\ker(A)^\perp = \overline{\operatorname{ran}(A^*)}$, and $P^* \colon \ker(A^*)^\perp \hookrightarrow H$ is the embedding; see Exercise 16.6. Since $A_1^* = (PAJ)^* = J^* A^* P^*$, we conclude that $\operatorname{ran}(A^*) = \operatorname{ran}(A_1^*) = \ker(A)^\perp$, and the latter set is closed. □

Proof of Theorem 16.12. Without loss of generality we assume that Ω is connected.

The fundamental observation for the proof is that the bounded linear operators $\operatorname{div} \colon H_0^1(\Omega; \mathbb{K}^n) \to L_2(\Omega)$ and $\nabla \colon L_2(\Omega) \to H^{-1}(\Omega)^n$ are negative adjoints of each other. Indeed, for $f \in L_2(\Omega)$, $u \in H_0^1(\Omega; \mathbb{K}^n)$ one has

$$(f \mid \operatorname{div} u) = \sum_{j=1}^n \int f \, \overline{\partial_j u_j} \, \mathrm{d}x = -\sum_{j=1}^n \langle \partial_j f, u_j \rangle_{H^{-1}, H_0^1} = -\langle \nabla f, u \rangle_{H^{-1}, H_0^1} .$$

We know from Lemma 6.8 that $\operatorname{ran}(\nabla)^\circ = \ker(\operatorname{div})$, and this implies $\overline{\operatorname{ran}(\nabla)} = \ker(\operatorname{div})^\circ$. (Note that for this argument no special properties of Ω are required.) It is shown in Theorem 16.11 that $\operatorname{ran}(\operatorname{div}) = L_2^0(\Omega)$, which is a closed subspace of $L_2(\Omega)$. Therefore Proposition 16.13 yields $\operatorname{ran}(\nabla) = \ker(\operatorname{div})^\circ$. As $\ker(\operatorname{div}) = H_{0,\sigma}^1(\Omega; \mathbb{K}^n)$ by definition, we obtain the assertion of the theorem. □

16.5 Supplement: the space $H^1_{\sigma,0}(\Omega;\mathbb{K}^n)$ and the hypothesis ($\mathrm{H_c}$)

The final issue of the chapter will be to investigate a slightly stronger version of hypothesis ($\mathrm{H_0}$). Using the space $H^1_{\sigma,0}(\Omega;\mathbb{K}^n) := \overline{C^\infty_{\mathrm{c},\sigma}(\Omega;\mathbb{K}^n)}^{H^1_0}$, where

$$C^\infty_{\mathrm{c},\sigma}(\Omega;\mathbb{K}^n) := \{u \in C^\infty_\mathrm{c}(\Omega;\mathbb{K}^n);\ \operatorname{div} u = 0\} = C^\infty_\mathrm{c}(\Omega;\mathbb{K}^n) \cap H^1_{0,\sigma}(\Omega;\mathbb{K}^n),$$

we can formulate it as follows:

$$\text{for all } \eta \in H^1_{\sigma,0}(\Omega;\mathbb{K}^n)^\circ \text{ there exists } p \in L_2(\Omega) \text{ with } \eta = \nabla p. \tag{$\mathrm{H_c}$}$$

The space $H^1_{\sigma,0}(\Omega;\mathbb{K}^n)$ is a closed subspace of $H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$, and ($\mathrm{H_c}$) can be rephrased as the property that $H^1_{\sigma,0}(\Omega;\mathbb{K}^n)^\circ = \operatorname{ran}(\nabla)$, where $\nabla\colon L_2(\Omega) \to H^{-1}(\Omega)^n,\ f \mapsto \nabla f$.

We insert the important observation that for any open set $\Omega \subseteq \mathbb{R}^n$ one has the inclusions $\operatorname{ran}(\nabla) \subseteq H^1_{0,\sigma}(\Omega;\mathbb{K}^n)^\circ \subseteq H^1_{\sigma,0}(\Omega;\mathbb{K}^n)^\circ$ $(= C^\infty_{\mathrm{c},\sigma}(\Omega;\mathbb{K}^n)^\circ)$. Recall that ($\mathrm{H_0}$) is equivalent to $H^1_{0,\sigma}(\Omega;\mathbb{K}^n)^\circ = \operatorname{ran}(\nabla)$. To summarise, property ($\mathrm{H_c}$) is equivalent to ($\mathrm{H_0}$) together with the equality $H^1_{\sigma,0}(\Omega;\mathbb{K}^n) = H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$, and this equality holds if and only if $C^\infty_{\mathrm{c},\sigma}(\Omega;\mathbb{K}^n)$ is dense in $H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$.

The following theorem strengthens Theorem 16.12.

16.14 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Then* ($\mathrm{H_c}$) *is satisfied, and $H^1_{\sigma,0}(\Omega;\mathbb{K}^n) = H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$, i.e. $C^\infty_{\mathrm{c},\sigma}(\Omega;\mathbb{K}^n)$ is dense in $H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$.*

Proof. (i) Assume additionally that Ω is star-shaped with respect to some ball $B(x^0,r) \subseteq \Omega$. Without loss of generality assume that $x^0 = 0$. Then one can easily see that $\lambda\overline{\Omega} \subseteq \Omega$ for all $\lambda \in (0,1)$. Let $u \in H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$. Extend u to $\mathbb{R}^n$ by zero; then it follows from Exercise 4.7 that $\operatorname{div} u = 0$ for the extended function as well. For $\lambda \in (0,1)$ put $u_\lambda := u(\lambda^{-1}\cdot)$; then $\operatorname{spt} u_\lambda \subseteq \lambda\overline{\Omega}$ is a compact subset of Ω. If $(\rho_k)_{k\in\mathbb{N}}$ is a delta sequence in $C^\infty_\mathrm{c}(\mathbb{R}^n)$, then one concludes that $\operatorname{div}(\rho_k * u_\lambda) = 0$ for all $k \in \mathbb{N}$, $\rho_k * u_\lambda \in C^\infty_\mathrm{c}(\Omega;\mathbb{K}^n)$ for large k, and $\rho_k * u_\lambda \to u_\lambda$ in $H^1_0(\Omega;\mathbb{K}^n)$ as $k \to \infty$. As a consequence, $u_\lambda \in H^1_{\sigma,0}(\Omega;\mathbb{K}^n)$ for all $\lambda \in (0,1)$. Now $u_\lambda \to u$ in $H^1_0(\Omega;\mathbb{K}^n)$ as $\lambda \to 1$; hence $u \in H^1_{\sigma,0}(\Omega;\mathbb{K}^n)$.

So we have shown that $H^1_{\sigma,0}(\Omega;\mathbb{K}^n) = H^1_{0,\sigma}(\Omega;\mathbb{K}^n)$. This implies ($\mathrm{H_c}$) since ($\mathrm{H_0}$) is satisfied by Theorem 16.12.

(ii) Without loss of generality we assume that Ω is connected. From the proof of Theorem 16.11 we recall that there exists an open covering $(\Omega_j)_{j=1,\dots,m}$ of Ω for which each Ω_j is star-shaped with respect to a ball.

Let $\eta \in H^1_{\sigma,0}(\Omega;\mathbb{K}^n)^\circ$. Then clearly $\eta^j := \eta|_{H^1_0(\Omega_j;\mathbb{K}^n)} \in H^1_{\sigma,0}(\Omega_j;\mathbb{K}^n)^\circ$ for $j = 1,\dots,m$, and from step (i) we conclude that there exists $f_j \in L_2(\Omega_j)$ (unique up to a constant) such that $\nabla f_j = \eta^j$. We will show that one can choose versions of the f_j such that the family $(f_j)_{j=1,\dots,m}$ is consistent, in the sense that $f_j = f_k$ on $\Omega_j \cap \Omega_k$ for $j,k = 1,\dots,m$.

For all $j \in \{1,\dots,m\}$ we choose a function $\varphi_j \in C^\infty_\mathrm{c}(\Omega_j)_+$ with $\int_{\Omega_j} \varphi_j(x)\,\mathrm{d}x = 1$. For each $j \in \{1,\dots,m\}$ we use Proposition 16.9 to obtain a vector field $\Phi_j \in C^\infty_\mathrm{c}(\Omega;\mathbb{K}^n)$ such that $\operatorname{div}\Phi_j = \varphi_1 - \varphi_j$, and we choose the version of f_j that satisfies $(f_j \,|\, \varphi_j) = \langle \eta, \Phi_j\rangle$. Given $j,k \in \{1,\dots,m\}$ we now show that $f_j = f_k$ on $\Omega_j \cap \Omega_k$. Let $\varphi \in C^\infty_\mathrm{c}(\Omega_j \cap \Omega_k)$, $\int \varphi(x)\,\mathrm{d}x = 1$;

then by Proposition 16.9 we can find $\Phi_{\varphi,j} \in C_c^\infty(\Omega_j; \mathbb{K}^n)$ such that $\operatorname{div} \Phi_{\varphi,j} = \varphi - \varphi_j$, and similarly $\Phi_{\varphi,k} \in C_c^\infty(\Omega_k; \mathbb{K}^n)$ with $\operatorname{div} \Phi_{\varphi,k} = \varphi - \varphi_k$. Then

$$(f_j \,|\, \varphi) = (f_j \,|\, \operatorname{div} \Phi_{\varphi,j}) + (f_j \,|\, \varphi_j) = -\langle \nabla f_j, \Phi_{\varphi,j} \rangle + \langle \eta, \Phi_j \rangle = \langle \eta, \Phi_j - \Phi_{\varphi,j} \rangle.$$

Note that η belonging to $H^1_{\sigma,0}(\Omega; \mathbb{K}^n)^\circ$ means that $\langle \eta, \Phi \rangle$ depends only on the divergence of $\Phi \in C_c^\infty(\Omega; \mathbb{K}^n)$. Since $\operatorname{div}(\Phi_j - \Phi_{\varphi,j}) = (\varphi_1 - \varphi_j) - (\varphi - \varphi_j) = \varphi_1 - \varphi = \operatorname{div}(\Phi_k - \Phi_{\varphi,k})$ we thus obtain $(f_j \,|\, \varphi) = (f_k \,|\, \varphi)$. It follows that $(f_j \,|\, \varphi) = (f_k \,|\, \varphi)$ for all $\varphi \in C_c^\infty(\Omega_j \cap \Omega_k)$, the consistency of the family $(f_j)_{j=1,\ldots,n}$ is established, and the function f, $f|_{\Omega_j} := f_j$ for $j = 1, \ldots, m$, is a well-defined function in $L_2(\Omega)$ satisfying $\nabla f = \eta$. □

The fundamental problem in the proof of Theorem 16.14 is that one cannot simply approximate elements of $H^1_{0,\sigma}(\Omega; \mathbb{K}^n)$ by convolution with a delta sequence. We refer to [ACM15; proof of Theorem 4.1] for another approach to dealing with this problem.

16.15 Remarks. In these remarks we sketch how Theorems 16.12 and 16.14 are obtained in [Tem77; Chap. I, §1].

(a) The basis is Nečas' inequality, stating that there exists $c > 0$ such that

$$\|f\|_{L_2} \leqslant c\Big(\sum_{j=1}^n \|\partial_j f\|_{H^{-1}} + \|f\|_{H^{-1}}\Big) \qquad (f \in L_2(\Omega)),$$

if Ω has Lipschitz boundary, asserted in [Neč12; Lemma 7.1, p. 186]. Using the compactness of the embedding $H^1_0(\Omega) \hookrightarrow L_2(\Omega)$ one can show that this inequality implies

$$\|f\|_2 \leqslant c \sum_{j=1}^n \|\partial_j f\|_{H^{-1}} \qquad \Big(f \in L_2(\Omega) \text{ with } \int_\Omega f(x)\,\mathrm{d}x = 0\Big) \tag{16.8}$$

if Ω is bounded and has Lipschitz boundary. From (16.8) one concludes that $\operatorname{ran}(\nabla)$ is closed (where $\nabla\colon L_2(\Omega) \to H^{-1}(\Omega)^n$); see Exercise 16.7. As $\operatorname{ran}(\nabla)$ is also dense in $H^1_{0,\sigma}(\Omega; \mathbb{K}^n)^\circ$, Theorem 16.12 is proved.

(b) Theorem 16.14 is also derived in [Tem77; Chap. I, Remark 1.4]. The argument there is based on a stronger version of (16.8), in which f is not a priori in $L_2(\Omega)$. The source [Neč66] for this version, cited in [Tem77], was not available to us. △

Notes

Our introduction and presentation of the Stokes operator follows [Mon06] and [ArEl12a]. As in several of the previous chapters, properties of Sobolev spaces play an important role – for the Stokes operator we need Sobolev spaces of divergence free vector fields. These spaces are also of importance for the treatment of the Navier–Stokes equation.

The treatment given for the space $H^1_{0,\sigma}(\Omega; \mathbb{K}^n)$ in Section 16.3 can be found to a large part in [Gal11; Section III.3]. The *derivation* of the Bogovskiĭ formula presented at the beginning of Section 16.3 was found by the authors, and the same holds for the proof of Theorem 16.14. We mention that the Bogovskiĭ operator is also treated in the

L_p-context in [Gal11] and in Sobolev spaces of negative order in [GHH06]. The treatment in Section 16.4 uses ideas contained in [Tem77; Chap. I, § 1].

In [ACM15; Theorem 4.1] one can find an account of various important relations between spaces arising in connection with the Stokes operator and the Navier–Stokes equation.

Exercises

16.1 Let $\Omega \subseteq \mathbb{R}^n$ be open.

(a) Show that the mapping $I - \Delta\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isometric isomorphism.

(b) Assume additionally that Ω is bounded, and provide $H_0^1(\Omega)$ with the scalar product $(u, v) \mapsto \int \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x$. Show that $-\Delta\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isometric isomorphism.

16.2 (a) Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $u\colon \Omega \to \mathbb{K}^n$ be a continuous vector field. Show that the following properties are equivalent:

(i) u is a gradient field, $u = \nabla p$ for some $p \in C^1(\Omega)$;

(ii) u satisfies (16.2), i.e., $\int u \cdot \varphi \,\mathrm{d}x = 0$ for all $\varphi \in C_{\mathrm{c}}^1(\Omega; \mathbb{K}^n)$ with $\operatorname{div} \varphi = 0$;

(iii) u is conservative, i.e., $\int_\gamma u = \int_0^1 u(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t = 0$ for all piecewise continuously differentiable closed paths $\gamma\colon [0, 1] \to \Omega$.

Hint concerning '(ii) $\Rightarrow$ (iii)': Let (ρ_k) be a delta sequence in $C_{\mathrm{c}}^\infty(\mathbb{R}^n)$. Show that $\varphi_k(y) := \int_0^1 \rho_k(\gamma(t) - y)\gamma'(t)\,\mathrm{d}t$ defines a divergence free element of $C_{\mathrm{c}}^\infty(\Omega; \mathbb{K}^n)$, for k large enough, and that $\int u \cdot \varphi_k \,\mathrm{d}x \to \int_\gamma u$ for all $u \in C(\Omega; \mathbb{K}^n)$.

Hint concerning '(iii) $\Rightarrow$ (i)': Define p as in Remark 16.5(a).

Supplemental comment: The sequence (φ_k) converges in a suitable sense to a divergence free distributional vector field supported on $\operatorname{im} \gamma = \{\gamma(t);\ 0 \leqslant t \leqslant 1\}$. The limiting distributional vector field is given by $C_{\mathrm{c}}^\infty(\Omega; \mathbb{K}^n) \ni \psi \mapsto \int_0^1 \psi(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t \in \mathbb{K}$.

(b) Show that the vector field $u\colon \mathbb{R}^2 \times \{0\} \to \mathbb{R}^2$, $u(x_1, x_2) := \left(\dfrac{-x_2}{x_1^2 + x_2^2}, \dfrac{x_1}{x_1^2 + x_2^2}\right)$, satisfies the compatibility condition $\partial_1 u_2 = \partial_2 u_1$. Show that u has a potential on $(0, \infty) \times \mathbb{R}$, given by $p(x_1, x_2) := \arctan \frac{x_2}{x_1}$. Extend this potential to $\mathbb{R}^2 \setminus (-\infty, 0] \times \{0\}$, and show that it cannot be extended to $\mathbb{R}^2 \setminus \{0\}$.

16.3 Find a simple form of the Bogovskiĭ formula in space dimension $n = 1$. In this form it should be immediately visible that Bf does not depend on ρ for f with $\int f(x)\,\mathrm{d}x = 0$, whereas Bf depends on ρ if $\int f(x)\,\mathrm{d}x \neq 0$. (Hint: Use the expression for the Bogovskiĭ operator indicated in the proof of Theorem 16.7, step (ii).)

16.4 Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set.

(a) Show that $L_2^0(\Omega)$ is a closed subspace of $L_2(\Omega)$ and that $C_{\mathrm{c}}^\infty \cap L_2^0(\Omega)$ is dense in $L_2^0(\Omega)$. (Hint: For the denseness show first that $L_{2,\mathrm{c}}^0(\Omega) := \{f \in L_2^0(\Omega);\ \operatorname{spt} f \text{ compact}\}$ is dense in $L_2^0(\Omega)$.)

(b) Assume additionally that Ω is connected. Let $(\Omega_j)_{j=1,\dots,m}$ be a finite open covering of Ω. Let $f \in L_2^0(\Omega)$. Show that there exist functions $f_1, \dots, f_m \in L_2^0(\Omega)$, $[f_j \neq 0] \subseteq \Omega_j$ for $j = 1, \dots, m$, such that $f = \sum_{j=1}^m f_j$.

Hint: Proceed by induction on m, replacing two members of the covering with non-empty intersection by their union in the induction step. Use the following fact established in step (ii) of the proof of Theorem 16.11: if $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ are open, $\Omega_1 \cap \Omega_2 \neq \varnothing$ and $f \in L_2^0(\Omega_1 \cup \Omega_2)$, then there exist $f_j \in L_2^0(\Omega_j)$ $(j = 1, 2)$ such that $f = f_1 + f_2$.

16.5 Let Ω be as in Theorem 16.10. Let $\rho \in C_c^\infty(\mathbb{R}^n)$ be such that $\operatorname{spt} \rho \subseteq B(x^0, r_0)$, $\int \rho(x) \, \mathrm{d}x = 1$, and let the Bogovskiĭ operator B be as in (16.6). Show that there exists $c > 0$ such that $\|Bf\|_2 \leqslant c\|f\|_2$ for all $f \in C_c^\infty \cap L_2^0(\Omega)$.

Hint: Recall the dominating function indicated in the proof of Theorem 16.7. Then either use Proposition 4.3(b), or else prove an L_1-L_1 bound and an L_∞-L_∞ bound for B, and then use Riesz–Thorin.

16.6 Let H be a Hilbert space. Let $H_0 \subseteq H$ be a closed subspace, $J \colon H_0 \hookrightarrow H$ the embedding. Show that J^* is the orthogonal projection from H onto H_0.

16.7 Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Show that the inequality (16.8) implies that $\operatorname{ran}(\nabla)$ is closed, for the mapping $\nabla \colon L_2(\Omega) \to H^{-1}(\Omega)^n$.

Chapter 17

Non-autonomous equations

So far we have considered forms that do not depend on time. The aim of this chapter is to establish well-posedness of the non-autonomous problem

$$u'(t) + \mathcal{A}(t)u(t) = f(t), \quad u(0) = u_0\,, \tag{17.1}$$

where each operator $\mathcal{A}(t)$ is associated with a form $a(t,\cdot,\cdot)$, defined on a t-independent Hilbert space $V \overset{\mathrm{d}}{\hookrightarrow} H$. In contrast to previous results, the existence of solutions will be investigated in the antidual V^*, which will be considered as a superspace of H. In fact, the situation is slightly more delicate and interesting: the solution will take its values in V, but the equation will be solved only in V^*.

We begin this chapter by describing the setup for V^*. Sections 17.2 and 17.3 are devoted to the Bochner integral for Hilbert space valued functions and to Hilbert space valued Sobolev spaces. Of special interest are the 'maximal regularity spaces', certain mixed Sobolev spaces, and their properties. In Section 17.4 we present the important representation theorem of J.-L. Lions, which extends the Lax–Milgram lemma. Lions' representation theorem is the key for the elegant treatment of (17.1), in the last section.

17.1 Gelfand triples

Let V, H be Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$. We have frequently encountered this situation in previous chapters; typical examples are $H_0^1(\Omega) \hookrightarrow L_2(\Omega)$ or $H^1(\Omega) \hookrightarrow L_2(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ is open. In Section 16.1 we have introduced the Gelfand triple corresponding to the embedding $H_0^1(\Omega) \hookrightarrow L_2(\Omega)$; here we treat the general setup.

For $x \in H$ we define $\langle x, \cdot\rangle \in V^*$ by

$$\langle x, u\rangle = \langle x, u\rangle_{V^*,V} := (x \,|\, u)_H \qquad (u \in V).$$

It is not difficult to see that the mapping $H \ni x \mapsto \langle x, \cdot\rangle \in V^*$ belongs to $\mathcal{L}(H, V^*)$ and is injective. It also has dense range because $u \in V$, $\langle x, u\rangle = 0$ for all $x \in H$ implies $u = 0$; see Exercise 17.1. Incidentally, the mapping $x \mapsto \langle x, \cdot\rangle$ coincides with the mapping $k \in \mathcal{L}(H, V^*)$ used in Proposition 5.7. We will identify $\langle x, \cdot\rangle$ and x and thereby consider H as a subspace of V^*. Thus we now have a triple of injected spaces

$$V \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} V^*,$$

called a **Gelfand triple**. We illustrate this situation in the following remark.

17.1 Remarks. (a) Let (Ω, μ) be a measure space, $m\colon \Omega \to [\delta, \infty)$ a measurable function, where $\delta > 0$, and let $H := L_2(\Omega, \mu)$, $V := L_2(\Omega, m\mu)$. Then $V \overset{\mathrm{d}}{\hookrightarrow} H$; for the denseness of the embedding we refer to Example 5.9. We can identify V^* with $L_2(\Omega, \frac{1}{m}\mu)$, where the duality is given by

$$\langle f, u\rangle = \int_\Omega f\overline{u}\,\mathrm{d}\mu \qquad \big(f \in L_2(\Omega, \tfrac{1}{m}\mu),\ u \in V\big).$$

One can easily see this by using the fact that $\Phi u := mu$ defines an isometric isomorphism $\Phi\colon L_2(\Omega; m\mu) \to L_2(\Omega; \frac{1}{m}\mu)$.

(b) In fact, the situation described in part (a) is generic. This is a consequence of the spectral theorem for self-adjoint operators, Theorem 13.21; we refer to Exercise 17.2(a) for more information. △

Now let $a\colon V \times V \to \mathbb{K}$ be a bounded form. Recall from (5.3) the corresponding Lax–Milgram operator $\mathcal{A} \in \mathcal{L}(V, V^*)$, given by $\langle \mathcal{A}u, v\rangle = a(u, v)$. Let A be the operator in H associated with the form a; see Proposition 5.5. From the definition it is immediate that A is the part of $\mathcal{A}$ in H, i.e.

$$A = \mathcal{A} \cap (V \times H),$$

with $V \subseteq H \subseteq V^*$ considered as subspaces. If a is quasi-coercive, then we know that $-A$ generates a holomorphic C_0-semigroup on H. One can show that $-\mathcal{A}$, considered as an operator in V^*, generates a holomorphic C_0-semigroup on V^*. We will not pursue this issue, but we recall from Remark 3.11(b) that as a consequence one can solve the initial value problem

$$u' + \mathcal{A}u = 0, \quad u(0) = u_0 \in V^*$$

in V^*. Our aim in the present chapter is to study the *inhomogeneous* Cauchy problem

$$u' + \mathcal{A}u = f, \quad u(0) = u_0\,,$$

where $\mathcal{A}$ *depends on time*; see Section 17.5.

17.2 Interlude: The Bochner integral for Hilbert space valued functions

There is a general theory extending the Lebesgue integral to Banach space valued functions (the Bochner integral, see [ABHN11; Section 1.1]). On separable Hilbert spaces one may use a more elementary approach, which we will present here (cf. [ArUr23; Section 8.5]).

Let H be a separable Hilbert space, and let $-\infty \leqslant a < b \leqslant \infty$. A function $f\colon (a, b) \to H$ is called **measurable** if $(f(\cdot) \mid x)$ is measurable for all $x \in H$. If (x_n) is a dense sequence in the unit ball $B_H(0, 1)$, then $\|f(t)\| = \sup_{n\in\mathbb{N}} |(f(t) \mid x_n)|$ for all $t \in (a, b)$, and this implies that $\|f(\cdot)\|\colon (a, b) \to \mathbb{R}$ is measurable. We define

$$L_1(a, b; H) := \Big\{f\colon (a, b) \to H;\ f \text{ measurable}, \ \int_a^b \|f(t)\|\,\mathrm{d}t < \infty\Big\},$$

where the elements of $L_1(a, b; H)$ are to be understood as equivalence classes of a.e. equal functions.

17.2 Lemma. *Let $f \in L_1(a,b;H)$. Then there exists a unique $x \in H$ such that*

$$\int_a^b (f(t)\,|\,y)_H \,\mathrm{d}t = (x\,|\,y)_H \qquad (y \in H),$$

and we define $\int_a^b f(t)\,\mathrm{d}t := x$. The mapping $L_1(a,b;H) \ni f \mapsto \int_a^b f(t)\,\mathrm{d}t \in H$ is a bounded linear operator, $\big\|\int_a^b f(t)\,\mathrm{d}t\big\| \leqslant \int_a^b \|f(t)\|\,\mathrm{d}t$ for all $f \in L_1(a,b;H)$.

Lemma 17.2 is an easy consequence of the theorem of Fréchet–Riesz; the proof is delegated to Exercise 17.3(a). The above definition of the integral is consistent with the definition given in Subsection 1.3.2, by Theorem 1.8(a). As in that theorem, bounded linear operators commute with integration; see Exercise 17.3(b).

We also introduce the space

$$L_2(a,b;H) := \Big\{ f\colon (a,b) \to H\,;\ f \text{ measurable},\ \int_a^b \|f(t)\|^2 \,\mathrm{d}t < \infty \Big\},$$

again identifying a.e. equal functions. Note that $L_2(a,b;H) \subseteq L_1(a,b;H)$ if (a,b) is a bounded interval.

17.3 Proposition. *The space $L_2(a,b;H)$ is a Hilbert space for the scalar product*

$$(f\,|\,g)_{L_2(a,b;H)} := \int_a^b (f(t)\,|\,g(t))_H \,\mathrm{d}t.$$

In order to see that the function $t \mapsto (f(t)\,|\,g(t))_H$ in Proposition 17.3 is measurable, we first prove the following denseness property.

17.4 Lemma. *For each $f \in L_2(a,b;H)$ there exists a sequence (f_n) in*

$$L_2(a,b) \otimes H := \operatorname{lin}\big\{\varphi(\cdot)x\,;\ \varphi \in L_2(a,b),\ x \in H\big\}$$

such that $\|f_n(\cdot)\| \leqslant \|f(\cdot)\|$ for all $n \in \mathbb{N}$ and $f_n \to f$ a.e.

In particular, $L_2(a,b) \otimes H$ is dense in $L_2(a,b;H)$, and $L_2(a,b;H)$ is separable.

Proof. The assertions are obvious if $\dim H < \infty$, so we assume that H is infinite-dimensional. Let $f \in L_2(a,b;H)$. Let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of H. Then, with $f_n := \sum_{k=1}^n (f(\cdot)\,|\,e_k)\,e_k$, the sequence (f_n) has the required properties. The dominated convergence theorem implies that $f_n \to f$ in $L_2(a,b;H)$. □

Proof of Proposition 17.3. We first show that $t \mapsto (f(t)\,|\,g(t))_H$ is measurable. If $h \in L_2(a,b) \otimes H$, $h = \sum_{j=1}^n \varphi_j(\cdot)x_j$, then $(h(\cdot)\,|\,g(\cdot))_H = \sum_{j=1}^n \varphi_j(\cdot)(x_j\,|\,g(\cdot))_H$ is measurable. Every $f \in L_2(a,b;H)$ can be approximated pointwise a.e. by a sequence in $L_2(a,b) \otimes H$, by Lemma 17.4, and the measurability of $(f(\cdot)\,|\,g(\cdot))_H$ follows.

The completeness is proved in the same way as the completeness of the scalar-valued $L_2(a,b)$. (Show that every absolutely convergent series is convergent.) □

A nice and important application of Lemma 17.4 is the following fact which will be needed in the next section.

17.5 Lemma. *Let $u \in L_2(\mathbb{R}; H)$. Then the function $\mathbb{R} \ni \tau \mapsto u(\cdot - \tau) \in L_2(\mathbb{R}; H)$ is continuous.*

Proof. The assertion is obvious if $u \in C_\mathrm{c}(\mathbb{R}) \otimes H = \mathrm{lin}\big\{\varphi(\cdot)x\,;\ \varphi \in C_\mathrm{c}(\mathbb{R}),\ x \in H\big\}$. By Lemma 17.4 and the denseness of $C_\mathrm{c}(\mathbb{R})$ in $L_2(\mathbb{R})$ there exists a sequence (u_k) in $C_\mathrm{c}(\mathbb{R}) \otimes H$ such that $u_k \to u$ in $L_2(\mathbb{R}; H)$ as $k \to \infty$. Then $u_k(\cdot - \tau) \to u(\cdot - \tau)$ in $L_2(\mathbb{R}; H)$ as $k \to \infty$, uniformly for $\tau \in \mathbb{R}$. This implies the assertion. □

17.3 Vector-valued Sobolev spaces

We now define Hilbert space valued Sobolev spaces. As before, let H be a separable Hilbert space.

Let $-\infty \leqslant a < b \leqslant \infty$. As in the scalar case, given $u \in L_2(a, b; H)$, a function $u' \in L_2(a, b; H)$ is called a **weak derivative** of u if

$$-\int_a^b u(t)\varphi'(t)\,\mathrm{d}t = \int_a^b u'(t)\varphi(t)\,\mathrm{d}t \qquad (\varphi \in C_\mathrm{c}^\infty(a, b)).$$

Such a weak derivative is unique whenever it exists; see Exercise 17.4(a). We define the **Sobolev space**

$$H^1(a, b; H) := \big\{u \in L_2(a, b; H)\,;\ u \text{ has a weak derivative } u' \text{ in } L_2(a, b; H)\big\}.$$

It is easy to see that $H^1(a, b; H)$ is a separable Hilbert space for the scalar product

$$(u \,|\, v)_{H^1} = \int_a^b \big((u(t) \,|\, v(t))_H + (u'(t) \,|\, v'(t))_H\big)\,\mathrm{d}t$$

(cf. Theorem 4.10 and Proposition 17.3).

For the remainder of this section we will assume throughout that $-\infty < a < b < \infty$. We state the following basic result concerning $H^1(a, b; H)$ without proof. Parts (a) and (b) can be proved essentially as Proposition 4.8, and part (c) is analogous to Theorem 4.12.

17.6 Proposition. (a) *Let $v \in L_2(a, b; H)$, $u_0 \in H$, and put $u(t) := u_0 + \int_a^t v(s)\,\mathrm{d}s$ $(t \in (a, b))$. Then $u \in H^1(a, b; H)$ and $u' = v$.*

(b) *Conversely, let $u \in H^1(a, b; H)$. Then there exists $u_0 \in H$ such that*

$$u(t) = u_0 + \int_a^t u'(s)\,\mathrm{d}s \qquad (a.e.\ t \in (a, b)).$$

In particular, each function $u \in H^1(a, b; H)$ has a representative in $C([a, b]; H)$, and with this representative one has $\int_a^b u'(s)\,\mathrm{d}s = u(b) - u(a)$.

(c) *The embedding $H^1(a, b; H) \hookrightarrow C([a, b]; H)$ thus defined is continuous.*

We will always use the continuous representative for functions $u \in H^1(a, b; H)$.

17.7 Remark. We note the following product rule for differentiation: if $u \in H^1(a,b;H)$ and $\varphi \in C^\infty[a,b]$, then $\varphi u \in H^1(a,b;H)$, and $(\varphi u)' = \varphi' u + \varphi u'$.

Indeed, for $\psi \in C_c^\infty(a,b)$ one has

$$\begin{aligned}\int_a^b (\varphi u)\psi' \,\mathrm{d}t &= \int_a^b (\varphi\psi)' u \,\mathrm{d}t - \int_a^b \varphi'\psi u \,\mathrm{d}t \\ &= -\int_a^b \varphi\psi u' \,\mathrm{d}t - \int_a^b \varphi' u\psi \,\mathrm{d}t = -\int_a^b (\varphi u' + \varphi' u)\psi \,\mathrm{d}t. \qquad \triangle\end{aligned}$$

Next we suppose that V is a separable Hilbert space, $V \overset{\mathrm{d}}{\hookrightarrow} H$, and we identify H with a dense subspace of V^* as in Section 17.1. The mixed Sobolev space

$$MR(a,b;V,V^*) := L_2(a,b;V) \cap H^1(a,b;V^*),$$

with $-\infty < a < b < \infty$ as before, plays an important role for evolutionary problems. Here the symbol "*MR*" stands for "maximal regularity". It is easy to see that $MR(a,b;V,V^*)$ is a separable Hilbert space for the norm

$$\|u\|_{MR(a,b;V,V^*)} := \big(\|u\|^2_{L_2(a,b;V)} + \|u'\|^2_{L_2(a,b;V^*)}\big)^{1/2}.$$

By Proposition 17.6(b) each $u \in MR(a,b;V,V^*)$ has a representative $u \in C([a,b];V^*)$. We will see in Theorem 17.9 below that the representative is even continuous with values in the smaller space H. In the statement of this result we will use the scalar **Sobolev space**

$$W_1^1(a,b) := \big\{u \in L_1(a,b)\,;\; u' = \partial u \in L_1(a,b)\big\},$$

with norm

$$\|u\|_{W_1^1(a,b)} := \|u\|_{L_1(a,b)} + \|u'\|_{L_1(a,b)} \qquad (u \in W_1^1(a,b)).$$

17.8 Proposition. *Every $u \in W_1^1(a,b)$ has a representative in $C[a,b]$, and the embedding $W_1^1(a,b) \hookrightarrow C[a,b]$ thus defined is continuous; explicitly,*

$$\|u\|_{C[a,b]} \leqslant \frac{1}{b-a}\|u\|_{L_1(a,b)} + \|u'\|_{L_1(a,b)}\,.$$

The proof is the same as for Theorem 4.12; the inequality is on the first line of (4.3). Proposition 4.8 implies that, choosing the continuous representative of $u \in W_1^1(a,b)$, one has the formula

$$u(t) = u(a) + \int_a^t u'(s)\,\mathrm{d}s \qquad (t \in [a,b]). \tag{17.2}$$

For the following theorem, concerning the regularity of elements in $MR(a,b;V,V^*)$, recall the embeddings $V \hookrightarrow H \hookrightarrow V^*$ and the dual pairing $\langle\cdot,\cdot\rangle$ from Section 17.1.

17.9 Theorem. (a) *If $u \in MR(a,b;V,V^*)$, then the function $\|u(\cdot)\|_H^2$ belongs to $W_1^1(a,b)$, and*

$$\big(\|u(\cdot)\|_H^2\big)' = 2\,\mathrm{Re}\langle u'(\cdot), u(\cdot)\rangle. \tag{17.3}$$

(b) *One has $MR(a,b;V,V^*) \hookrightarrow C([a,b];H)$.*

Observe that the function $t \mapsto f(t) := \langle u'(t), u(t)\rangle$ in part (a) is measurable, by the same reasoning as in the proof of Proposition 17.3. Since $|f(t)| \leqslant \|u'(t)\|_{V^*}\|u(t)\|_V$ for all $t \in (a,b)$, the Cauchy–Schwarz inequality then implies that $f \in L_1(a,b)$.

For the proof of Theorem 17.9 we need the following denseness property. The result is analogous to Theorem 7.7.

17.10 Proposition. *The space $C^\infty([a,b];V)$ is dense in $MR(a,b;V,V^*)$.*

Proof. (i) Let $\alpha \in C^\infty(\mathbb{R})$ be such that $\alpha = 1$ in a neighbourhood of $(-\infty, a]$, $\alpha = 0$ in a neighbourhood of $[b,\infty)$. For $\tau > 0$ put $u_\tau := (\alpha u)(\cdot + \tau) + \big((1-\alpha)u\big)(\cdot - \tau)$. Then $u_\tau \to u$ in $MR(a,b;V,V^*)$ as $\tau \to 0$, by Lemma 17.5. (For the derivative of u_τ, we note that $(u_\tau)' = (\alpha u)'(\cdot+\tau) + \big((1-\alpha)u\big)'(\cdot-\tau)$, by the choice of α.) For fixed $\tau > 0$ we choose $\psi \in C_c^\infty(\mathbb{R})$ with $\operatorname{spt}\psi \subseteq (a-\tau, b+\tau)$ and $\psi = 1$ on (a,b); then $\psi u_\tau \in MR(\mathbb{R};V,V^*)$. (For convenience we use the notation $MR(\mathbb{R};V,V^*) := L_2(\mathbb{R};V) \cap H^1(\mathbb{R};V^*)$, although MR was only defined for bounded intervals (a,b).)

(ii) Now it suffices to approximate $v \in MR(\mathbb{R};V,V^*)$ by a sequence in $C^\infty(\mathbb{R};V) \cap MR(\mathbb{R};V,V^*)$. Clearly $MR(\mathbb{R};V,V^*)$ is invariant under translations, and Lemma 17.5 implies that $\mathbb{R} \ni t \mapsto v(\cdot - t) \in MR(\mathbb{R};V,V^*)$ is continuous. Let ρ_k be a delta sequence in $C_c^\infty(\mathbb{R})$. For $k \in \mathbb{N}$ we define $v_k := \int_{\mathbb{R}} \rho_k(s)v(\cdot - s)\,\mathrm{d}s \in MR(\mathbb{R};V,V^*)$. Then

$$\begin{aligned}\|v_k - v\|_{MR(\mathbb{R};V,V^*)} &\leqslant \int_{\mathbb{R}} \rho_k(s)\|v(\cdot - s) - v\|_{MR(\mathbb{R};V,V^*)}\,\mathrm{d}s\\ &\leqslant \sup\big\{\|v(\cdot - s) - v\|_{MR(\mathbb{R};V,V^*)};\ |s| \leqslant 1/k\big\} \to 0 \qquad (k\to\infty).\end{aligned}$$

For $t \in \mathbb{R}$, Proposition 17.6(c) implies that point evaluation $MR(\mathbb{R};V,V^*) \ni u \mapsto u(t) \in V^*$ is a bounded linear operator. Hence, using Exercise 17.3(b), we obtain

$$v_k(t) = \int_{\mathbb{R}} \rho_k(s)v(t-s)\,\mathrm{d}s =: \rho_k * v(t),$$

where the integral is taken in V^*. Applying Exercise 17.3(b) again (with the embedding $V \hookrightarrow V^*$ as the operator $A \in \mathcal{L}(V,V^*)$) we see that the integral can also be taken in V; then with a proof as for Lemma 4.1 one shows that $\rho_k * v \in C^\infty(\mathbb{R};V)$. ☐

Proof of Theorem 17.9. (a) If $u \in C^1([a,b];V)$, then it is immediate that $\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)\|_H^2 = (u'(t)\,|\,u(t))_H + (u(t)\,|\,u'(t))_H = 2\operatorname{Re}\langle u'(t), u(t)\rangle$. Now let $u \in MR(a,b;V,V^*)$. By Proposition 17.10 there exists a sequence (u_n) in $C^1([a,b];V)$ converging to u in $MR(a,b;V,V^*)$. Then

$$\big(\|u_n(\cdot)\|_H^2\big)' = 2\operatorname{Re}\langle u_n'(\cdot), u_n(\cdot)\rangle \to 2\operatorname{Re}\langle u'(\cdot), u(\cdot)\rangle$$

in $L_1(a,b)$. Moreover $u_n \to u$ in $L_2(a,b;H)$ since $V \hookrightarrow H$, and therefore $\|u_n(\cdot)\|_H^2 \to \|u(\cdot)\|_H^2$ in $L_1(a,b)$. Hence, $2\operatorname{Re}\langle u'(\cdot), u(\cdot)\rangle$ is the distributional derivative of $\|u(\cdot)\|_H^2$, by Lemma 4.11.

(b) For $u \in C^1([a,b];V)$, applying Proposition 17.8 to the function $t \mapsto \|u(t)\|_H^2$ we obtain

$$\begin{aligned}\|u\|^2_{C([a,b];H)} &\leqslant \frac{1}{b-a}\int_a^b \|u(t)\|_H^2 \,\mathrm{d}t + \int_a^b \left|\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)\|_H^2\right| \mathrm{d}t \\ &\leqslant \frac{1}{b-a}\int_a^b \|u(t)\|_H^2 \,\mathrm{d}t + \int_a^b 2\|u'(t)\|_{V^*}\|u(t)\|_V \,\mathrm{d}t \\ &\leqslant \frac{c^2}{b-a}\|u\|^2_{L_2(a,b;V)} + 2\|u'\|_{L_2(a,b;V^*)}\|u\|_{L_2(a,b;V)},\end{aligned}$$

with the embedding constant $c > 0$ of $V \hookrightarrow H$. As $C^1([a,b];V)$ is dense in $MR(a,b;V,V^*)$, by Proposition 17.10, this inequality shows that $MR(a,b;V,V^*) \hookrightarrow C([a,b];H)$. □

17.4 Lions' representation theorem

Of great importance in the theory of Hilbert spaces is the representation theorem of Fréchet–Riesz, which often gives us weak solutions of partial differential equations. One of the consequences of this theorem is the reflexivity of Hilbert spaces. Reflexivity is a key ingredient in the proof of the following much more general representation theorem.

17.11 Theorem (Lions' representation theorem). *Let $\mathcal{V}$ be a Hilbert space, $\mathcal{W}$ a normed space, $\mathcal{W} \hookrightarrow \mathcal{V}$. Let $E\colon \mathcal{V} \times \mathcal{W} \to \mathbb{K}$ be sesquilinear, and assume that*

(i) *$E(\cdot, w) \in \mathcal{V}'$ for all $w \in \mathcal{W}$,*

(ii) *$|E(w,w)| \geqslant \varepsilon\|w\|^2_{\mathcal{W}}$ for all $w \in \mathcal{W}$, with some $\varepsilon > 0$.*

Let $L \in \mathcal{W}^*$. Then there exists $u \in \mathcal{V}$ such that $L(w) = E(u,w)$ for all $w \in \mathcal{W}$ and $\|u\|_{\mathcal{V}} \leqslant \frac{c}{\varepsilon}\|L\|_{\mathcal{W}^*}$, where $c > 0$ is an embedding constant of the embedding $\mathcal{W} \hookrightarrow \mathcal{V}$.

The fact that $\mathcal{W}$ is not supposed to be complete makes Lions' theorem more widely applicable than the Lax–Milgram lemma. The larger the norm on $\mathcal{W}$, the less restrictive is the assumption on L to be continuous with respect to the norm of $\mathcal{W}$. On the other hand, the 'coercivity hypothesis' (ii) becomes more restrictive if we take larger norms. Even if we choose as norm on $\mathcal{W}$ the norm of $\mathcal{V}$, it is an advantage that E need not be defined on all of $\mathcal{V}$ in the second variable; note that there is no continuity requirement on E with respect to the second variable.

In Exercise 17.7(a) we present a slightly more general version of Theorem 17.11.

Proof of Theorem 17.11. We define an antilinear operator $T\colon \mathcal{W} \to \mathcal{V}'$ by $Tw := E(\cdot, w)$ for all $w \in \mathcal{W}$. It follows from property (ii) that

$$\varepsilon\|w\|^2_{\mathcal{W}} \leqslant |E(w,w)| \leqslant \|E(\cdot,w)\|_{\mathcal{V}'}\|w\|_{\mathcal{V}} \leqslant c\|Tw\|_{\mathcal{V}'}\|w\|_{\mathcal{W}};$$

thus $\frac{\varepsilon}{c}\|w\|_{\mathcal{W}} \leqslant \|Tw\|_{\mathcal{V}'}$ for all $w \in \mathcal{W}$. This implies that T is injective and that $T^{-1}\colon T(\mathcal{W}) \to \mathcal{W}$ is antilinear and bounded, where $T(\mathcal{W})$ is provided with the norm of $\mathcal{V}'$. Since $L \in \mathcal{W}^*$, the functional $L \circ T^{-1}$ extends to a bounded linear functional $\ell \in (\mathcal{V}')'$, with $\|\ell\| \leqslant \frac{c}{\varepsilon}\|L\|_{\mathcal{W}^*}$. By the reflexivity of $\mathcal{V}$ there exists $u \in \mathcal{V}$ such that $\ell(v') = v'(u)$ for all $v' \in \mathcal{V}'$, and $\|u\| = \|\ell\|$. In particular, for $w \in \mathcal{W}$ one has $L(w) = L(T^{-1}Tw) = \ell(Tw) = Tw(u) = E(u,w)$. □

17.12 Remark (Uniqueness in Theorem 17.11). The vector $u \in \mathcal{V}$ is unique if and only if

$$v \in \mathcal{V},\ E(v,w) = 0 \text{ for all } w \in \mathcal{W} \text{ implies } v = 0.$$

This is the same as saying that $T(\mathcal{W})$ is dense in $\mathcal{V}'$, with the operator T from the proof. △

17.5 The non-autonomous equation

We now come to the main result of this chapter. We study the non-autonomous inhomogeneous evolution equation

$$u'(t) + \mathcal{A}(t)u(t) = f(t), \quad u(0) = u_0\,. \tag{17.4}$$

Our assumptions are as follows. Let V, H be separable Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$. With these spaces we form the Gelfand triple $V \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} V^*$ and use the notation introduced in Section 17.1.

Let $\tau \in (0,\infty)$, and let a be a **non-autonomous form** on V, i.e., $a\colon [0,\tau] \times V \times V \to \mathbb{K}$ is a mapping such that

(Li1) $a(t,\cdot,\cdot)\colon V \times V \to \mathbb{K}$ is sesquilinear for all $t \in [0,\tau]$;
(Li2) $a(\cdot,x,y)$ is measurable for all $x, y \in V$.

We also suppose that there exist $M \geqslant \alpha > 0$ such that

(Li3) $|a(t,x,y)| \leqslant M\|x\|_V\|y\|_V$ for all $t \in [0,\tau]$, $x, y \in V$ (**boundedness**);
(Li4) $\operatorname{Re} a(t,x,x) \geqslant \alpha\|x\|_V^2$ for all $t \in [0,\tau]$, $x \in V$ (**coercivity**).

We point out that α and M are independent of t. It follows from Lemma 17.4 and the conditions (Li2) and (Li3) that $t \mapsto a(t,u(t),v(t))$ is integrable for all $u, v \in L_2(0,\tau;V)$. This property will be used below without further notice.

For each $t \in [0,\tau]$ we denote by $\mathcal{A}(t) \in \mathcal{L}(V,V^*)$ the operator given by

$$\langle \mathcal{A}(t)x, y\rangle = a(t,x,y) \qquad (x, y \in V).$$

First we want to interpret $\mathcal{A}$ as a ‘multiplication operator’ in the following way.

17.13 Proposition. *Defining*

$$(\mathcal{A}u)(t) := \mathcal{A}(t)u(t) \qquad (t \in (0,\tau))$$

for $u \in L_2(0,\tau;V)$, one obtains a bounded linear operator $\mathcal{A}\colon L_2(0,\tau;V) \to L_2(0,\tau;V^*)$.

Proof. Let $u \in L_2(0,\tau;V)$. Recalling that the function $t \mapsto \langle \mathcal{A}u(t), y\rangle = a(t,u(t),y)$ is measurable for all $y \in V$, one deduces that $\mathcal{A}u$ is measurable. The estimate

$$\|\mathcal{A}(t)u(t)\|_{V^*} \leqslant M\|u(t)\|_V \qquad (t \in (0,\tau)) \tag{17.5}$$

shows that $\mathcal{A}u = \mathcal{A}(\cdot)u(\cdot) \in L_2(0,\tau;V^*)$. Obviously the operator $\mathcal{A}$ is linear, and by (17.5) it is bounded. □

In the proof of Theorem 17.15 below we will apply Theorem 17.11 with a suitable sesquilinear mapping E; the definition of E is motivated by the following proposition. We will use the notation $C_{\mathrm{c}}^\infty[0,\tau) := \{\varphi \in C^\infty[0,\tau)\,;\ \operatorname{spt}\varphi \text{ compact}\}$.

17.14 Proposition. *Let $u_0 \in H$, $f \in L_2(0,\tau;V^*)$, $u \in L_2(0,\tau;V)$. Then the following properties are equivalent.*

(i) *$u \in MR(0,\tau;V,V^*)$ and $u' + \mathcal{A}u = f$, $u(0) = u_0$.*

(ii) *For all $\varphi \in C_{\mathrm{c}}^\infty[0,\tau)$, $x \in V$ one has*

$$-\int_0^\tau (u(t)\,|\,\varphi'(t)x)_H\,\mathrm{d}t + \int_0^\tau a(t,u(t),\varphi(t)x)\,\mathrm{d}t = \int_0^\tau \langle f(t),\varphi(t)x\rangle\,\mathrm{d}t + (u_0\,|\,\varphi(0)x)_H\,.$$

Proof. Note that property (ii) is equivalent to

$$-\int_0^\tau \varphi'(t)u(t)\,\mathrm{d}t + \int_0^\tau \varphi(t)(\mathcal{A}u)(t)\,\mathrm{d}t = \int_0^\tau \varphi(t)f(t)\,\mathrm{d}t + \varphi(0)u_0 \tag{17.6}$$

for all $\varphi \in C_{\mathrm{c}}^\infty[0,\tau)$. This is an equation in V^*; recall that $\mathcal{A}u = \mathcal{A}(\cdot)u(\cdot) \in L_2(0,\tau;V^*)$.

(i) $\Rightarrow$ (ii). Let $\varphi \in C_{\mathrm{c}}^\infty[0,\tau)$. Using $\mathcal{A}u = f - u'$ and Remark 17.7 we obtain

$$-\int_0^\tau \varphi'(t)u(t)\,\mathrm{d}t + \int_0^\tau \varphi(t)(\mathcal{A}u)(t)\,\mathrm{d}t = -\int_0^\tau (\varphi u)'(t)\,\mathrm{d}t + \int_0^\tau \varphi(t)f(t)\,\mathrm{d}t.$$

Since $\varphi(\tau) = 0$ and $u(0) = u_0$, it follows from Proposition 17.6(b) that (17.6) holds.

(ii) $\Rightarrow$ (i). Using (17.6) with $\varphi \in C_{\mathrm{c}}^\infty(0,\tau)$ we obtain

$$-\int_0^\tau \varphi'(t)u(t)\,\mathrm{d}t = \int_0^\tau \varphi(t)\big(f(t) - (\mathcal{A}u)(t)\big)\,\mathrm{d}t.$$

This implies that $u \in H^1(0,\tau;V^*)$ and $u' = f - \mathcal{A}u$.

From (i) $\Rightarrow$ (ii) it follows that (17.6) also holds with $u(0)$ in place of u_0. Choosing $\varphi \in C_{\mathrm{c}}^\infty[0,\tau)$ with $\varphi(0) = 1$ we conclude that $u(0) = u_0$. □

We can now formulate and prove the main result of this chapter.

17.15 Theorem (Lions). *Let $u_0 \in H$, $f \in L_2(0,\tau;V^*)$. Then there exists a unique $u \in MR(0,\tau;V,V^*)$ such that*

$$u' + \mathcal{A}u = f, \quad u(0) = u_0\,. \tag{17.7}$$

For the solution u one has the estimate

$$\|u\|_{L_2(0,\tau;V)} \leqslant \tfrac{1}{\alpha}\|f\|_{L_2(0,\tau;V^*)} + \sqrt{\tfrac{2}{\alpha}}\|u_0\|_H\,. \tag{17.8}$$

Note that both terms u', $\mathcal{A}u$ belong to the same space $L_2(0,\tau;V^*)$ as f. For this reason we say that the problem has **maximal regularity in $\boldsymbol{V^*}$**. Note also that, in view of Theorem 17.9(b), a solution $u \in MR(0,\tau;V,V^*)$ of (17.7) can only exist if the initial value u_0 belongs to H. In Exercise 17.7(b) we indicate how to obtain a slightly sharper estimate than (17.8) for the solution u.

17.16 Remark. Theorem 17.15 implies that the linear operator

$$H \times L_2(0,\tau;V^*) \ni (u_0, f) \mapsto u \in MR(0,\tau;V,V^*)$$

(mapping the data u_0, f to the solution u) is bounded. Concerning the norm of u in $L_2(0,\tau;V)$ we refer to (17.8). For the derivative of u we apply (17.5) to estimate

$$\|u'\|_{L_2(0,\tau;V^*)} = \|f - \mathcal{A}u\|_{L_2(0,\tau;V^*)} \leqslant \|f\|_{L_2(0,\tau;V^*)} + M\|u\|_{L_2(0,\tau;V)}$$

and then use (17.8) once more. △

Proof of Theorem 17.15. To prove the existence we apply Lions' representation theorem, Theorem 17.11, with $\mathcal{V} := L_2(0,\tau;V)$ and the pre-Hilbert space

$$\mathcal{W} := C_c^\infty[0,\tau) \otimes V = \operatorname{lin}\bigl\{\varphi(\cdot)x\,;\ \varphi \in C_c^\infty[0,\tau),\ x \in V\bigr\}$$

with norm $\|w\|_{\mathcal{W}} := \bigl(\alpha\|w\|_{\mathcal{V}}^2 + \frac{1}{2}\|w(0)\|_H^2\bigr)^{1/2}$. We define $E\colon \mathcal{V}\times\mathcal{W} \to \mathbb{K}$ by

$$E(v,w) = -\int_0^\tau (v(t)\,|\,w'(t))_H\,\mathrm{d}t + \int_0^\tau a(t,v(t),w(t))\,\mathrm{d}t.$$

For $w \in \mathcal{W}$ one has

$$|E(v,w)| \leqslant \|v\|_{\mathcal{V}}\|w'\|_{L_2(0,\tau;V^*)} + M\|v\|_{\mathcal{V}}\|w\|_{\mathcal{V}} \qquad (v \in \mathcal{V}).$$

Thus condition (i) of Theorem 17.11 is satisfied.

In order to verify (ii) let $w \in \mathcal{W}$. Then $\|w(\cdot)\|_H^2 \in C_c^\infty[0,\tau)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\|w(t)\|_H^2 = 2\operatorname{Re}(w'(t)\,|\,w(t))_H$; thus

$$-\int_0^\tau \operatorname{Re}(w'(t)\,|\,w(t))_H\,\mathrm{d}t = \frac{1}{2}\|w(0)\|_H^2\,.$$

By the coercivity condition (Li4) it follows that

$$\operatorname{Re}E(w,w) \geqslant \frac{1}{2}\|w(0)\|_H^2 + \alpha\|w\|_{\mathcal{V}}^2 = \|w\|_{\mathcal{W}}^2\,; \tag{17.9}$$

thus condition (ii) of Theorem 17.11 is satisfied.

Define $L \in \mathcal{W}^*$ by

$$L(w) := \int_0^\tau \langle f(t), w(t)\rangle\,\mathrm{d}t + (u_0\,|\,w(0))_H \qquad (w \in \mathcal{W}).$$

By Lions' representation theorem, there exists $u \in \mathcal{V}$ such that $E(u,w) = L(w)$ for all $w \in \mathcal{W}$, which means that u satisfies property (ii) of Proposition 17.14. It follows that $u \in MR(0,\tau;V,V^*)$ and that u is a solution of (17.7).

In order to prove the uniqueness, let $u \in MR(0,\tau;V,V^*)$ be such that $u' + \mathcal{A}u = 0$ and $u(0) = 0$. Then by Theorem 17.9(a) we obtain

$$\bigl(\|u(\cdot)\|_H^2\bigr)'(t) = 2\operatorname{Re}\langle u'(t), u(t)\rangle = -2\operatorname{Re}\langle(\mathcal{A}u)(t), u(t)\rangle = -2\operatorname{Re}a(t,u(t),u(t)) \leqslant 0$$

for a.e. $t \in [0,\tau]$. Using formula (17.2) we conclude that $\|u(\cdot)\|_H$ is decreasing, so $u(t) = 0$ for all $t \in [0,\tau]$.

Obviously $\|w\|_{\mathcal{V}} \leqslant \frac{1}{\sqrt{\alpha}}\|w\|_{\mathcal{W}}$ for all $w \in \mathcal{W}$, and by the Cauchy–Schwarz inequality one obtains $\|L\|_{\mathcal{W}^*} \leqslant \bigl(\frac{1}{\alpha}\|f\|_{L_2(0,\tau;V^*)}^2 + 2\|u_0\|_H^2\bigr)^{1/2} \leqslant \frac{1}{\sqrt{\alpha}}\|f\|_{L_2(0,\tau;V^*)} + \sqrt{2}\|u_0\|_H$. In view of these inequalities and (17.9), the estimate in Theorem 17.11 implies (17.8). □

17.17 Remark. The notion of solution we use here is rather weak. This is because our space of test functions $\mathcal{W}$ is small and consists of very regular functions. Of course, if we have a weak notion of solution, the existence of solutions might become easy to prove (an extreme version would be to call every function a solution). But the uniqueness may become hard to prove. Here for existence we need the key equality (17.3) only for functions in $\mathcal{W}$, for which it is in fact trivial. With the help of Lions' representation theorem the proof of existence is rather elegant and easy. In contrast, uniqueness is more technical since (17.3) is needed for a rather general class of functions. △

Notes

The results of Sections 17.4 and 17.5 go back to J.-L. Lions and his school in the 1950s and 1960s. Lions' representation theorem, Theorem 17.11, was first proved in [Lio57]. The existence and uniqueness result Theorem 17.15 is contained in Lions' book ([Lio61; Chap. IV, Théorème 1.1]), with a slightly different formulation. Lions cites [Vis56] and [Lio59] for the first proofs of existence. Uniqueness is shown in [Lio59]; this had been an open problem for a while. It is also possible to prove Theorem 17.15 by the Galerkin method; see [DaLi92]. This is of interest for the numerical treatment but less elegant than via the representation theorem. For more recent results on these and related topics we refer to [ACE23], [ACE24].

Exercises

17.1 Let V be a Hilbert space and $M \subseteq V^*$ a subset with the property that

$$u \in V,\ \langle \eta, u\rangle = 0 \text{ for all } \eta \in M \quad \text{implies} \quad u = 0.$$

Show that M is a dense subset of V^*.

Hint: Let $V^* \ni \eta \mapsto \hat\eta \in V$ be the inverse of the Fréchet–Riesz isomorphism, i.e. $(\hat\eta \,|\, \cdot)_V = \eta$ for all $\eta \in V^*$. Show that $\widehat{M} = \{\hat\eta;\ \eta \in M\}$ is dense in V. (Note that, in view of the reflexivity of V, the assertion can also be obtained from the Hahn–Banach theorem.)

17.2 Let V, H be Hilbert spaces, $V \overset{\mathrm{d}}{\hookrightarrow} H$.

(a) Show that there exist a measure space (Ω, μ), a unitary operator $J\colon H \to L_2(\Omega, \mu)$ and a measurable function $m\colon \Omega \to [\delta, \infty)$ (with $\delta > 0$) such that $J|_V$ is a unitary operator from V onto $L_2(\Omega, m\mu)$; see Remarks 17.1. (Hint: Show that $a(u, v) := (u \,|\, v)_V$ $(u, v \in V)$ defines a symmetric bounded coercive form, and let A be the associated strictly accretive self-adjoint operator. Use the spectral theorem for self-adjoint operators, Theorem 13.21.)

(b) In the context of part (a), let the form $a\colon V \times V \to \mathbb{K}$ be given by

$$a(u, v) = \int_\Omega m u \bar v \,\mathrm{d}\mu.$$

Show that the operator $\mathcal{A}$ in V^*, described in Section 17.1, is given by

$$\mathcal{A}u = mu \qquad (u \in \operatorname{dom}(\mathcal{A}) = V).$$

(c) Let $\mathbb{K} = \mathbb{C}$. Show that the operator $-\mathcal{A}$ from part (b) generates a holomorphic C_0-semigroup T on V^*. Show that $T(t)(V^*) \subseteq V$ for all $t > 0$, and that $(0,\infty) \ni t \mapsto T(t)f \in V$ is infinitely differentiable, for all $f \in V^*$.

17.3 (a) Prove Lemma 17.2.

(b) Let G be a Hilbert space, $A \in \mathcal{L}(H,G)$. Show that $A\int_a^b f(t)\,\mathrm{d}t = \int_a^b Af(t)\,\mathrm{d}t$ for all $f \in L_1(a,b;H)$.

17.4 Let $-\infty < a < b < \infty$.

(a) Let H be a separable Hilbert space, $u \in L_1(a,b;H)$. Assume that

$$\int_a^b \varphi(t)u(t)\,\mathrm{d}t = 0$$

for all $\varphi \in C_c^\infty(a,b)$. Show that $u = 0$ a.e.

Deduce the uniqueness of the weak derivative defined in Section 17.3.

(b) Find a (non-separable) Hilbert space H and a function $u\colon (a,b) \to H$ with the properties that $\|u(t)\| = 1$ for all $t \in (a,b)$ and $(u(\cdot)\,|\,v) = 0$ a.e. for all $v \in H$. (Note that then $\int_a^b \varphi(t)u(t)\,\mathrm{d}t = 0$ for all $\varphi \in C_c^\infty(a,b)$, if the integral were defined as in Lemma 17.2.)

17.5 Let $V \overset{\mathrm{d}}{\hookrightarrow} H$ be Hilbert spaces, let a be a bounded coercive form on V, with associated operator A, and let S be the C_0-semigroup generated by $-A$. Let $\tau > 0$, and regard a as a non-autonomous form (not depending on t).

Let $u_0 \in H$. Show that $u := S(\cdot)u_0$ is the solution of

$$u' + \mathcal{A}u = 0, \quad u(0) = u_0$$

obtained from Theorem 17.15. (Hint: First treat the case $u_0 \in \operatorname{dom}(A)$. For the general case use (17.8).)

17.6 Prove the existence and uniqueness parts of Theorem 17.15 when the form a is not necessarily coercive but quasi-coercive, i.e.,

$$\operatorname{Re} a(t,x,x) + \omega\|x\|_H^2 \geqslant \alpha\|x\|_V^2 \qquad (t \in [0,\tau],\ x \in V)$$

holds for some $\omega \geqslant 0$, $\alpha > 0$. Show that the solution satisfies the estimate

$$\|u\|_{L_2(0,\tau;V)} \leqslant \mathrm{e}^{\omega\tau}\big(\tfrac{1}{\alpha}\|f\|_{L_2(0,\tau;V^*)} + \sqrt{\tfrac{2}{\alpha}}\|u_0\|_H\big).$$

Hint: Solve $v' + (\mathcal{A}+\omega)v = \mathrm{e}^{-\omega\cdot}f$.

17.7 (a) Prove a generalised version of Lions' representation theorem, Theorem 17.11, in which hypothesis (ii) is replaced by the weaker condition

(ii′) $|E(w,w)| \geqslant \varepsilon\|w\|_{\mathcal{W}}\|w\|_{\mathcal{V}}$ for all $w \in \mathcal{W}$, with some $\varepsilon > 0$.

Show that then one obtains the estimate $\|u\|_{\mathcal{V}} \leqslant \frac{1}{\varepsilon}\|L\|_{\mathcal{W}^*}$ for the solution u. (Hint: Inspect the proof of Theorem 17.11.)

(b) Using part (a), give a proof of Theorem 17.15, with the improved estimate

$$\|u\|_{L_2(0,\tau;V)} \leqslant \tfrac{1}{\alpha}\|f\|_{L_2(0,\tau;V^*)} + \tfrac{1}{\sqrt{2\alpha}}\|u_0\|_H$$

for the solution u.

Hint: $\mathcal{V}$ and $\mathcal{W}$ as before, but $\|w\|_{\mathcal{W}} := \max\bigl\{\|w\|_{\mathcal{V}}, \sqrt{\tfrac{2}{\alpha}}\,\|w(0)\|_H\bigr\}$; then

$$\|w\|_{\mathcal{W}}\|w\|_{\mathcal{V}} \leqslant \|w\|_{\mathcal{V}}^2 + \tfrac{1}{2\alpha}\|w(0)\|_H^2 \leqslant \tfrac{1}{\alpha}\operatorname{Re} E(w,w) \qquad (w \in \mathcal{W})$$

(with the aid of Peter–Paul in the first inequality), and $\|L\|_{\mathcal{W}^*} \leqslant \|f\|_{L_2(0,\tau;V^*)} + \sqrt{\tfrac{\alpha}{2}}\,\|u_0\|_H$.

(c) Consider $V = H = \mathbb{R}$ and $\mathcal{A} = \alpha$ in order to show that the coefficient $\frac{1}{\sqrt{2\alpha}}$ at $\|u_0\|_H$, in the estimate in (b), is optimal (if one wants a τ-independent estimate).

17.8 (a) Let $x, a, b \in [0,\infty)$ satisfy the 'quadratic inequality' $x^2 \leqslant ax + b$. Show that this implies $x \leqslant a + \sqrt{b}$.

(b) Under the hypotheses of Theorem 17.15, let $u \in MR(0,\tau;V,V^*)$ be a solution of (17.7).

Insert $u' = -\mathcal{A}u + f$ into (17.3), integrate the resulting equation, and apply the coercivity of a, to obtain

$$\alpha\|u\|_{L_2(0,\tau;V)}^2 \leqslant \|f\|_{L_2(0,\tau;V^*)}\|u\|_{L_2(0,\tau;V)} + \tfrac{1}{2}\|u_0\|_H^2 .$$

Apply part (a) to reproduce the estimate from Exercise 17.7(b).

Conclude the uniqueness for solutions of (17.7).

Chapter 18

Maximal regularity for non-autonomous equations

The last chapter was devoted to the non-autonomous Cauchy problem

$$u'(t) + \mathcal{A}(t)u(t) = f(t), \quad u(0) = u_0\,,$$

where $\mathcal{A}$ is associated with a non-autonomous form on V. The problem was treated in the setup of a Gelfand triple $V \overset{d}{\hookrightarrow} H \overset{d}{\hookrightarrow} V^*$, and in Theorem 17.15, well-posedness was obtained in V^*. In the present chapter we will establish well-posedness and maximal regularity in H. The price to pay is that the non-autonomous form has to be Lipschitz continuous in time. The importance of maximal regularity in H will be demonstrated in Chapter 19, by applications to nonlinear problems. We refer to the Notes for more information on what is true and not true concerning maximal regularity in H.

18.1 Well-posedness in H

As in Chapter 17 let V, H be separable Hilbert spaces, $V \overset{d}{\hookrightarrow} H$, giving rise to the Gelfand triple $V \overset{d}{\hookrightarrow} H \overset{d}{\hookrightarrow} V^*$. Let $\tau \in (0,\infty)$, and assume that a is a bounded non-autonomous form on V, i.e., a mapping $a\colon [0,\tau] \times V \times V \to \mathbb{K}$ satisfying the properties (Li1), (Li2) and (Li3) stated in Section 17.5. For $t \in [0,\tau]$ we recall the definition of $\mathcal{A}(t) \in \mathcal{L}(V,V^*)$,

$$\langle \mathcal{A}(t)x, y\rangle = a(t,x,y) \qquad (x,y \in V),$$

and we define the operator $A(t)$ as the part of $\mathcal{A}(t)$ in H, $A(t) = \mathcal{A}(t) \cap (V \times H)$; see the discussion following Remarks 17.1. In analogy to the notation $MR(a,b;V,V^*)$, introduced in Section 17.3, we define

$$MR(0,\tau;V,H) := L_2(0,\tau;V) \cap H^1(0,\tau;H),$$

with norm

$$\|u\|_{MR(0,\tau;V,H)} := \big(\|u\|^2_{L_2(0,\tau;V)} + \|u'\|^2_{L_2(0,\tau;H)}\big)^{1/2} \qquad (u \in MR(0,\tau;V,H)).$$

We say that the form a satisfies **maximal regularity in H** if, given $u_0 \in V$ and $f \in L_2(0,\tau;H)$, there exists a unique function $u \in MR(0,\tau;V,H)$ such that

$$u' + \mathcal{A}u = f, \quad u(0) = u_0\,. \tag{18.1}$$

(Concerning the notation $\mathcal{A}u$ we refer to Proposition 17.13.) If u satisfies the above properties, then

$$u(t) \in \operatorname{dom}(A(t)) \quad \text{and} \quad u'(t) + A(t)u(t) = f(t) \qquad \text{for a.e. } t.$$

Thus both functions u', $A(\cdot)u(\cdot)$ belong to the same space $L_2(0,\tau;H)$ as f, which is the reason for the term "maximal regularity in H". Maximal regularity in H is not always valid for bounded non-autonomous forms; additional assumptions are needed. Here we will make the following assumptions.

We suppose that the form $a\colon [0,\tau]\times V\times V\to\mathbb{K}$ can be written as $a = a_1 + b$, where a_1 and b are bounded non-autonomous forms on V, with the following additional requirements.

The form a_1 is **symmetric**, i.e.

$$a_1(t,x,y) = \overline{a_1(t,y,x)} \qquad (t\in[0,\tau],\ x,y\in V),$$

coercive, i.e., there exists $\alpha > 0$ such that

$$a_1(t,x,x) \geqslant \alpha\|x\|_V^2 \qquad (t\in[0,\tau],\ x\in V),$$

and **Lipschitz continuous**, i.e., there exists a constant $M_1' \geqslant 0$ such that

$$|a_1(t,x,y) - a_1(s,x,y)| \leqslant M_1'|t-s|\,\|x\|_V\|y\|_V \qquad (s,t\in[0,\tau],\ x,y\in V).$$

In this context we will call M_1' a **Lipschitz bound** of a_1.

For the form b we do not require any regularity in time (besides measurability) but impose a stronger boundedness property, namely

$$|b(t,x,y)| \leqslant M_b\|x\|_V\|y\|_H \qquad (t\in[0,\tau],\ x,y\in V), \tag{18.2}$$

for some $M_b \geqslant 0$.

18.1 Remarks. (a) The above hypotheses imply that the form a is **quasi-coercive**, i.e.

$$\operatorname{Re} a(t,x,x) + \omega\|x\|_H^2 \geqslant \alpha\|x\|_V^2 \qquad (t\in[0,\tau],\ x\in V),$$

for some $\alpha > 0$, $\omega\in\mathbb{R}$.

Indeed, using the Peter–Paul inequality we see that

$$|b(t,x,x)| \leqslant M_b\|x\|_V\|x\|_H \leqslant \frac{\alpha}{2}\|x\|_V^2 + \frac{M_b^2}{2\alpha}\|x\|_H^2\,,$$

and hence by the coercivity of a_1 we obtain

$$\operatorname{Re} a(t,x,x) = a_1(t,x,x) + \operatorname{Re} b(t,x,x) \geqslant \frac{\alpha}{2}\|x\|_V^2 - \frac{M_b^2}{2\alpha}\|x\|_H^2\,,$$

for all $t\in[0,\tau]$, $x\in V$.

(b) Let a be a Lipschitz continuous symmetric bounded non-autonomous form on V which is quasi-coercive. Then a satisfies the above requirements, with

$$a_1(t,x,y) := a(t,x,y) + \omega(x\,|\,y)_H\,, \qquad b(t,x,y) := -\omega(x\,|\,y)_H\,.$$

(c) Let $a = a_1 + b$ be a non-autonomous form satisfying the hypotheses formulated above. Then, as shown in (a), the form a is quasi-coercive. Let $u_0 \in V$, $f \in L_2(0,\tau;H)$. Then, by Theorem 17.15 and Exercise 17.6, there exists a unique solution $u \in MR(0,\tau;V,V^*)$ of (18.1). On the other hand, a solution $u \in MR(0,\tau;V,H)$ of (18.1) will belong to $MR(0,\tau;V,V^*)$; hence, the uniqueness statement in Theorem 18.2 below is a consequence of the uniqueness recalled above. △

Now we formulate the main result of this chapter. We assume that $a = a_1 + b$ is a non-autonomous form satisfying the hypotheses formulated above. The operators $\mathcal{A}(t) \in \mathcal{L}(V,V^*)$ are given by $\langle \mathcal{A}(t)x, y\rangle = a(t,x,y)$, as before.

18.2 Theorem. *Let $u_0 \in V$, $f \in L_2(0,\tau;H)$. Then there exists a unique solution $u \in MR(0,\tau;V,H)$ of*

$$u' + \mathcal{A}u = f, \quad u(0) = u_0\,.$$

In short, the form a satisfies maximal regularity in H.

For the solution u one has the estimate

$$\|u\|_{MR(0,\tau;V,H)} \leqslant 2\mathrm{e}^{\gamma\tau}\bigl(\|f\|_{L_2(0,\tau;H)} + M_{1,0}^{1/2}\|u_0\|_V\bigr), \tag{18.3}$$

where $\gamma := \frac{1+M_1'+M_b^2}{\alpha}$, and $M_{1,0}$ is a bound for the form $a_1(0,\cdot,\cdot)$.

The proof of this theorem will be given in Section 18.5. First we will need to study Lipschitz continuous functions and product rules; see Sections 18.3 and 18.4.

The solution u in Theorem 18.2 depends continuously on the part b of the form; this is the topic of Exercise 18.1.

18.3 Remarks. (a) With the form $b(t,\cdot,\cdot)$ we associate the Lax–Milgram operator $\mathcal{B}(t) \in \mathcal{L}(V,V^*)$, for all $t \in [0,\tau]$. Because of the H-norm at the element y in the estimate (18.2), the element $\mathcal{B}(t)x \in V^*$ belongs to H, for all $x \in V$, $t \in [0,\tau]$. This means that in fact the range of $\mathcal{B}(t)$ is contained in H, and that $\mathcal{B}(t)$ also acts as a bounded operator from V to H, with norm $\leqslant M_b$.

The properties of $\mathcal{B}$ imply that $\mathcal{B}u \in L_2(0,\tau;H)$ and $\|\mathcal{B}u\|_{L_2(0,\tau;H)} \leqslant M_b\|u\|_{L_2(0,\tau;V)}$ for all $u \in L_2(0,\tau;V)$. (The measurability of $\mathcal{B}u$ is proved in the same way as in the proof of Proposition 17.13.)

(b) With the notation from part (a) and with $\mathcal{A}_1(t)$ denoting the Lax–Milgram operator associated with $a_1(t,\cdot,\cdot)$, the equation in Theorem 18.2 can be written as

$$u' + \mathcal{A}_1u + \mathcal{B}u = f, \quad u(0) = u_0\,.$$

The solution $u \in MR(0,\tau;V,H)$ also has maximal regularity in H in the sense that both terms $\mathcal{A}_1u$ and $\mathcal{B}u$ are in $L_2(0,\tau;H)$. Indeed, $\mathcal{A}u \in L_2(0,\tau;H)$ together with $\mathcal{B}u \in L_2(0,\tau;H)$ implies that $\mathcal{A}_1u = \mathcal{A}u - \mathcal{B}u \in L_2(0,\tau;H)$. From Theorem 19.7 in the next chapter one can conclude a further regularity property: the solution u even belongs to $C([0,\tau];V)$.

(c) Assume that, in Theorem 18.2, the form a_1 does not depend on t, and let A_1 be the self-adjoint operator in H associated with a_1. Then the solution u of $u' + \mathcal{A}u = f$ has the stronger regularity property $u \in MR(0,\tau;\mathrm{dom}_{A_1},H)$, where dom_{A_1} denotes $\mathrm{dom}(A_1)$, provided with the 'graph norm' $\|\cdot\|_{A_1}$,

$$\|x\|_{A_1} = \big(\|x\|_H^2 + \|A_1x\|_H^2\big)^{1/2} \qquad (x \in \mathrm{dom}(A_1)).$$

Indeed, A_1 is the part of $\mathcal{A}_1$ in H, and $A_1^{-1} = \mathcal{A}_1^{-1}|_H$ maps H continuously to dom_{A_1}. This shows that $u = A_1^{-1}(\mathcal{A}_1u) \in L_2(0,\tau;\mathrm{dom}_{A_1})$. △

18.2 Examples for non-autonomous problems

In this section we give examples illustrating the existence theorems for non-autonomous problems presented so far. The first issue is an application of Theorem 17.15.

18.4 Example. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Let $\tau \in (0,\infty)$, and for $j,k = 1,\dots,n$ let $a_{jk}\colon [0,\tau] \to L_\infty(\Omega)$ be bounded and measurable as an $L_2(\Omega)$-valued function. Assume that there exists $\alpha > 0$ such that

$$\mathrm{Re}\sum_{j,k=1}^n a_{jk}(t,x)\xi_k\overline{\xi_j} \geqslant \alpha|\xi|^2$$

for all $t \in [0,\tau]$, $x \in \Omega$, $\xi \in \mathbb{K}^n$.

Let $V := H_0^1(\Omega)$ and $H := L_2(\Omega)$; then $V^* = H^{-1}(\Omega)$ (cf. Section 16.1). Define $a\colon [0,\tau]\times V\times V \to \mathbb{K}$ by

$$a(t,u,v) := \int_\Omega \sum_{j,k=1}^n a_{jk}(t,x)\partial_k u(x)\overline{\partial_j v(x)}\,\mathrm{d}x.$$

Obviously a satisfies the properties (Li1), (Li2) and (Li4) in Section 17.5. In order to see the measurability property (Li3), note that, by hypothesis,

$$t \mapsto \int_\Omega a_{jk}(t,x)f(x)\,\mathrm{d}x$$

is measurable for all $f \in L_2(\Omega)$, and hence, by approximation, for $f = \partial_k u\cdot\overline{\partial_j v} \in L_1(\Omega)$ as well.

Let $u_0 \in L_2(\Omega)$, $f \in L_2(0,\tau;H^{-1}(\Omega))$. Then Theorem 17.15 implies that there exists a unique solution $u \in MR(0,\tau;H_0^1(\Omega),H^{-1}(\Omega))$ of the initial value problem

$$u' + \mathcal{A}u = f, \quad u(0) = u_0$$

(where $\mathcal{A}$ is defined as in Section 17.5). △

Our second example illustrates the application of Theorem 18.2.

18.5 Example. Let the hypotheses be as in Example 18.4, but rename the non-autonomous form a to a_1. Additionally to these hypotheses assume that, for $j,k = 1,\dots,n$, the functions $a_{jk}\colon [0,\tau] \to L_\infty(\Omega)$ are Lipschitz continuous, and that $a_{jk}(t) = \overline{a_{kj}(t)}$ for all $t \in [0,\tau]$, $j,k \in \{1,\dots,n\}$.

In addition, for $j = 1,\dots,n$ let $b_j\colon [0,\tau] \to L_\infty(\Omega)$ be bounded and measurable as an $L_2(\Omega)$-valued function, and define $b\colon [0,\tau] \times V \times V \to \mathbb{K}$ by

$$b(t,u,v) := \int_\Omega \Bigl(\sum_{j=1}^n b_j(t,x)\partial_j u(x)\Bigr)\overline{v(x)}\,\mathrm{d}x.$$

One easily sees that a_1 and b satisfy the requirements stated in Section 18.1. Define $\mathcal{A}$ as in Section 18.1.

Let $u_0 \in H_0^1(\Omega)$, $f \in L_2(0,\tau; L_2(\Omega))$. Then by Theorem 18.2 there exists a unique $u \in MR(0,\tau; H_0^1(\Omega), L_2(\Omega))$ such that

$$u' + \mathcal{A}u = f, \quad u(0) = u_0\,.$$

Formulated in detail, the equation $u' + \mathcal{A}u = f$ reads

$$\partial_t u(t,x) = \sum_{j,k=1}^n \partial_j(a_{jk}(t,x)\partial_k u)(t,x) - \sum_{j=1}^n b_j(t,x)\partial_j u(t,x) + f(t,x);$$

this is a non-autonomous diffusion equation with a drift term. △

18.3 Interlude: Lipschitz continuous functions, scalar product rule

Let $-\infty \leqslant a < b \leqslant \infty$. It follows from Proposition 4.8 that each function $u \in C(a,b)$ with distributional derivative $u' \in L_\infty(a,b)$ is Lipschitz continuous. This property has a converse, as follows.

18.6 Proposition. *Let $u\colon (a,b) \to \mathbb{K}$ be Lipschitz continuous, i.e.*

$$|u(t) - u(s)| \leqslant L|t-s| \qquad (s,t \in (a,b)),$$

for some $L \geqslant 0$. Then $u' = \partial u \in L_\infty(a,b)$, with $\|u'\|_\infty \leqslant L$.

Proof. It is easy to see that one can extend u to a Lipschitz continuous function $u\colon \mathbb{R} \to \mathbb{K}$ with the same Lipschitz constant L. Therefore, without loss of generality, we may assume that $(a,b) = \mathbb{R}$.

Let $(\rho_k)_k$ be a delta sequence in $C_c^\infty(\mathbb{R})$. Then $\rho_k * u \in C^1(\mathbb{R})$ for all $k \in \mathbb{N}$, by Lemma 4.1. For $k \in \mathbb{N}$, $s,t \in \mathbb{R}$ we estimate

$$|\rho_k * u(t) - \rho_k * u(s)| \leqslant \int_\mathbb{R} \rho_k(r)\bigl|u(t-r) - u(s-r)\bigr|\,\mathrm{d}r \leqslant L|t-s|.$$

It follows that $\|(\rho_k * u)'\|_\infty \leqslant L$, for all $k \in \mathbb{N}$.

Let $R > 0$. Then $\bigl((\rho_k * u)'|_{(-R,R)}\bigr)_k$ is a bounded sequence in $L_2(-R,R)$ and therefore contains a weakly convergent subsequence $\bigl((\rho_{k_j} * u)'|_{(-R,R)}\bigr)_j$, with weak limit $v_R \in$

$L_2(-R,R)$. Proposition 4.3(a) shows that $\rho_k * u \to u$ uniformly on $(-R,R)$. Arguing as in the proof of Lemma 4.11, one concludes from these convergences that $\partial\big(u|_{(-R,R)}\big) = v_R$. Moreover, for all $w \in L_2(-R,R)$ one obtains

$$\left|\int_{-R}^{R} v_R(t)w(t)\,\mathrm{d}t\right| = \lim_{j\to\infty}\left|\int_{-R}^{R} (\rho_{k_j} * u)'(t)w(t)\,\mathrm{d}t\right| \leqslant L\|w\|_1\,,$$

and this implies that $\|v_R\|_\infty \leqslant L$.

Using a standard diagonal procedure one can choose the subsequence $\big((\rho_{k_j} * u)'\big)_j$ such that $\big((\rho_{k_j} * u)'|_{(-R,R)}\big)_j$ is weakly convergent simultaneously for all $R > 0$. Then the definition $v|_{(-R,R)} := v_R$, for all $R > 0$, yields the distributional derivative v of u, with $\|v\|_\infty \leqslant L$. □

The second topic of this section is the following product rule for scalar functions with distributional derivatives in $L_{1,\mathrm{loc}}$; it serves as a warm-up for the next interlude.

18.7 Proposition. *Let $-\infty \leqslant a < b \leqslant \infty$, and let $u, v \in C(a,b)$ be such that $u', v' \in L_{1,\mathrm{loc}}(a,b)$. Then $(uv)' = u'v + uv' \in L_{1,\mathrm{loc}}(a,b)$.*

Proof. We only need to prove the formula on (c,b), for $c \in (a,b)$. For $c \leqslant t < b$ we compute, applying Fubini's theorem in the last step,

$$\begin{aligned} u(t)v(t) - u(c)v(c) &= \int_c^t u'(s)\,\mathrm{d}s\, v(t) + u(c)\int_c^t v'(r)\,\mathrm{d}r \\ &= \int_c^t u'(s)\Big(v(s) + \int_s^t v'(r)\,\mathrm{d}r\Big)\,\mathrm{d}s + \int_c^t \Big(u(r) - \int_c^r u'(s)\,\mathrm{d}s\Big)v'(r)\,\mathrm{d}r \\ &= \int_c^t \big(u'(s)v(s) + u(s)v'(s)\big)\,\mathrm{d}s + 0. \end{aligned}$$

This establishes the asserted equality, by Proposition 4.8. □

18.4 Interlude: the product rule – vector-valued

In this section let G and H be separable Hilbert spaces of infinite dimension, and let $\tau \in (0,\infty)$ be fixed. In the arguments below we will need the space

$$L_\infty(0,\tau;H) := \big\{u\colon (0,\tau) \to H\,;\ u \text{ measurable, } \|u(\cdot)\|_H \text{ essentially bounded}\big\},$$

where a.e. equal functions are identified, and with norm

$$\|u\|_\infty := \big\|\|u(\cdot)\|_H\big\|_{L_\infty(0,\tau)}\,.$$

18.8 Proposition. *Let $T \in \mathcal{L}(G, L_\infty(0,\tau;H))$. Then there exists a bounded function $k\colon (0,\tau) \to \mathcal{L}(G,H)$, $\|k(t)\| \leqslant \|T\|$ for all $t \in (0,\tau)$, such that $Tx = k(\cdot)x$ for all $x \in G$.*

Proof. We choose an orthonormal basis $(e_j)_{j\in\mathbb{N}}$ of G and define $G_0 := \mathrm{lin}_{\mathbb{Q}_\mathbb{K}}\{e_j\,;\, j\in\mathbb{N}\}$ as the set of rational linear combinations of $\{e_j\,;\, j\in\mathbb{N}\}$, where $\mathbb{Q}_\mathbb{R} := \mathbb{Q}$ and $\mathbb{Q}_\mathbb{C} := \mathbb{Q}+\mathrm{i}\mathbb{Q}$. For each $x\in G_0$ we choose a representative $k_x\colon (0,\tau)\to H$ of Tx. Then there exists a null set $N\subseteq(0,\tau)$ such that for all $t\in(0,\tau)\setminus N$, $x,y\in G_0$, $\lambda\in\mathbb{Q}_\mathbb{K}$ one has

$$\begin{aligned}\|k_x(t)\|_H &\leqslant \|T\|\|x\|_G\,, \\ k_{\lambda x+y}(t) &= \lambda k_x(t)+k_y(t).\end{aligned} \tag{18.4}$$

(This is because of the linearity of T and the countability of $\mathbb{Q}_\mathbb{K}$ and G_0.)

For $t\in(0,\tau)\setminus N$, the continuous $\mathbb{Q}_\mathbb{K}$-linear mapping $G_0\ni x\mapsto k_x(t)\in H$ has a continuous $\mathbb{K}$-linear extension to G, denoted by $k(t)$. Putting $k(t):=0$ $(t\in N)$, we thus obtain a function $k\colon(0,\tau)\to\mathcal{L}(G,H)$, and (18.4) implies $\|k(t)\|_{\mathcal{L}(G,H)}\leqslant\|T\|$ for all $t\in(0,\tau)$.

It remains to show that $k(\cdot)x=k_x(\cdot)$, which by definition holds for all $x\in G_0$, carries over to all $x\in G$. From the measurability of $k(\cdot)x$ for all $x\in G_0$ we deduce by (18.4) that $k(\cdot)x$ is measurable for all $x\in G$. We define the operator $T_0\in\mathcal{L}(G,L_\infty(0,\tau;H))$ by $T_0x:=k(\cdot)x$ $(x\in G)$. Since T and T_0 coincide on G_0, the denseness of G_0 implies that $T=T_0$, i.e. $Tx=k(\cdot)x$ for all $x\in G$. □

18.9 Remark. The following fact concerning weak derivatives will be useful. If $u,v\in L_2(0,\tau;H)$, then $u\in H^1(0,\tau;H)$ and $u'=v$ if and only if $(u(\cdot)\,|\,y)'=(v(\cdot)\,|\,y)$ in the distributional sense, for all $y\in H$.

This equivalence is an easy consequence of the definitions. △

18.10 Proposition. (a) *Let $f\colon[0,\tau]\to H$ be Lipschitz continuous, with Lipschitz constant L. Then $f\in H^1(0,\tau;H)$, $\|f'(t)\|\leqslant L$ for a.e. $t\in(0,\tau)$.*

(b) *Let $\mathcal{D}\colon[0,\tau]\to\mathcal{L}(G,H)$ be Lipschitz continuous, with Lipschitz constant L. Then there exists a bounded function $\mathcal{D}'\colon(0,\tau)\to\mathcal{L}(G,H)$, $\|\mathcal{D}'(t)\|_{\mathcal{L}(G,H)}\leqslant L$ for all $t\in(0,\tau)$, such that $(\mathcal{D}(\cdot)x)'=\mathcal{D}'(\cdot)x$ for all $x\in G$.*

Proof. (a) Using Proposition 18.6 we can define $T\in\mathcal{L}(H,L_\infty(0,\tau))$ by $Tx:=(x\,|\,f(\cdot))'$ $(x\in H)$; then $\|T\|\leqslant L$. We now apply Proposition 18.8 and use the Fréchet–Riesz representation theorem to identify H with $\mathcal{L}(H;\mathbb{K})$ by the anti-isomorphism $y\mapsto(\cdot\,|\,y)$. Thus we obtain a bounded function $k\colon(0,\tau)\to H$, $\|k(t)\|_H\leqslant L$ for all $t\in(0,\tau)$, such that

$$(x\,|\,f(\cdot))'=Tx=(x\,|\,k(\cdot))\qquad(x\in H).$$

This property yields the measurability of k, and Remark 18.9 implies that $f\in H^1(0,\tau;H)$, $f'=k$.

(b) For each $x\in G$, part (a) above shows that $Tx:=(\mathcal{D}(\cdot)x)'\in L_\infty(0,\tau;H)$, $\|Tx\|_\infty\leqslant L\|x\|$. Applying Proposition 18.8 to the operator $T\in\mathcal{L}(G,L_\infty(0,\tau;H))$ we obtain a function $\mathcal{D}'\colon(0,\tau)\to\mathcal{L}(G,H)$ as asserted. □

For the proof of the following 'product rule' we refer to Exercise 18.2.

18.11 Lemma. *Let $u,v\in H^1(0,\tau;H)$. Then $(u(\cdot)\,|\,v(\cdot))\in H^1(0,\tau)$,*

$$(u(\cdot)\,|\,v(\cdot))'=(u'(\cdot)\,|\,v(\cdot))+(u(\cdot)\,|\,v'(\cdot)).$$

18.12 Proposition. *Let $\mathcal{D}\colon [0,\tau] \to \mathcal{L}(G,H)$ be Lipschitz continuous, and let $\mathcal{D}'$ be as in Proposition 18.10(b). Let $u \in H^1(0,\tau;G)$. Then $\mathcal{D}u \in H^1(0,\tau;H)$,*

$$(\mathcal{D}u)' = \mathcal{D}'u + \mathcal{D}u'.$$

(Here we have used the notation $\mathcal{D}v := \mathcal{D}(\cdot)v(\cdot)$, for $v \in L_2(0,\tau;G)$, and similarly for $\mathcal{D}'$.)
The mapping $u \mapsto \mathcal{D}u$ belongs to $\mathcal{L}(H^1(0,\tau;G), H^1(0,\tau;H))$.

Proof. Let $y \in H$; then the function $\mathcal{D}^*(\cdot)y\colon [0,\tau] \to G$ is Lipschitz continuous. For all $x \in G$ we get – observing Remark 18.9 –

$$(x \,|\, \mathcal{D}(\cdot)^*y)'_G = (\mathcal{D}(\cdot)x \,|\, y)'_H = (\mathcal{D}'(\cdot)x \,|\, y)_H = (x \,|\, \mathcal{D}'(\cdot)^*y)_G .$$

Applying Remark 18.9 again we conclude that $(\mathcal{D}(\cdot)^*y)' = \mathcal{D}'(\cdot)^*y$. Thus, by Lemma 18.11 we obtain

$$\begin{aligned}\big(\mathcal{D}(\cdot)u(\cdot) \,\big|\, y\big)'_H &= \big(u(\cdot) \,\big|\, \mathcal{D}^*(\cdot)y\big)'_G = \big(u'(\cdot) \,\big|\, \mathcal{D}^*(\cdot)y\big)_G + \big(u(\cdot) \,\big|\, (\mathcal{D}^*(\cdot)y)'\big)_G \\ &= \big(\mathcal{D}(\cdot)u'(\cdot) + \mathcal{D}'(\cdot)u(\cdot) \,\big|\, y\big)_H .\end{aligned}$$

As this equality holds for all $y \in H$, we can use Remark 18.9 once more to conclude that $\mathcal{D}u \in H^1(0,\tau;H)$ and $(\mathcal{D}u)' = \mathcal{D}'u + \mathcal{D}u'$.

In view of the formula for $(\mathcal{D}u)'$, the continuity of the mapping $H^1(0,\tau;G) \ni u \mapsto \mathcal{D}u \in H^1(0,\tau;H)$ is obvious. □

We employ Proposition 18.12 for the proof of a product rule for non-autonomous forms, as follows. Note that in this result there is no coercivity requirement on the non-autonomous form a.

18.13 Corollary. *Let V be a separable Hilbert space. Let $a\colon [0,\tau] \times V \times V \to \mathbb{K}$ be a Lipschitz continuous bounded non-autonomous form with Lipschitz bound M'. Then $a(\cdot, u(\cdot), v(\cdot)) \in H^1(0,\tau)$ for all $u, v \in H^1(0,\tau;V)$. There exists a bounded non-autonomous form $a'\colon (0,\tau) \times V \times V \to \mathbb{K}$, $|a'(t,x,y)| \leqslant M'\|x\|_V\|y\|_V$ for all $x, y \in V$, $t \in (0,\tau)$, such that*

$$a(\cdot, u(\cdot), v(\cdot))' = a'(\cdot, u(\cdot), v(\cdot)) + a(\cdot, u'(\cdot), v(\cdot)) + a(\cdot, u(\cdot), v'(\cdot)),$$

for all $u, v \in H^1(0,\tau;V)$.

Proof. We start by recalling that with the form a we can associate the Lax–Milgram operaters $\mathcal{A}(t) \in \mathcal{L}(V, V^*)$, defined by

$$\mathcal{A}(t)x = a(t, x, \cdot) \qquad (x \in V,\ t \in (0,\tau)).$$

It is easy to see that the Lipschitz continuity of the form a is equivalent to the Lipschitz continuity of $\mathcal{A}\colon [0,\tau] \to \mathcal{L}(V, V^*)$ with Lipschitz bound M'. Now Proposition 18.12 and Lemma 18.11 imply that $a(\cdot, u(\cdot), v(\cdot)) \in H^1(0,\tau)$,

$$\begin{aligned}a(\cdot, u(\cdot), v(\cdot))' &= \langle \mathcal{A}u, v\rangle'_{V^*,V} = \langle \mathcal{A}'u, v\rangle + \langle \mathcal{A}u', v\rangle + \langle \mathcal{A}u, v'\rangle \\ &= \langle \mathcal{A}'u, v\rangle + a(\cdot, u'(\cdot), v(\cdot)) + a(\cdot, u(\cdot), v'(\cdot)),\end{aligned} \qquad (18.5)$$

where we have used the dual pairing $\langle\cdot,\cdot\rangle_{V^*,V}$ instead of the scalar product in V. It remains to find the form a' as in the assertion. From Proposition 18.10(b) we know that $\mathcal{A}'\colon (0,\tau)\to\mathcal{L}(V,V^*)$ is a bounded function, $\|\mathcal{A}'(t)\|\leqslant M'$ $(t\in(0,\tau))$. Putting

$$a'(t,x,y) := \langle\mathcal{A}'(t)x,y\rangle \qquad (x,y\in V,\ t\in(0,\tau))$$

we obtain a bounded non-autonomous form $a'\colon (0,\tau)\times V\times V\to\mathbb{K}$ with bound M'. (The measurability property (Li2) from Section 17.5 follows from the measurability of $\mathcal{A}'(\cdot)x = (\mathcal{A}(\cdot)x)'$ $(x\in V)$, in Proposition 18.10(b).) □

18.5 Proof of maximal regularity in H

Throughout this section we assume that $a = a_1 + b$ is a non-autonomous form as in Section 18.1.

Given $u_0\in V\subseteq H$ and $f\in L_2(0,\tau;H)\subseteq L_2(0,\tau;V^*)$, we already know the existence of a solution $u\in MR(0,\tau;V,V^*)$ of (18.1), from Theorem 17.15 and Exercise 17.6. Our method of proof will *not* be to show that this solution, in the present case, has better qualities than stated in Theorem 17.15, but we give a new existence proof. The idea of the new proof is to again use Lions' representation theorem, but with a different sesquilinear form E that results from a characterisation of solutions which is different from Proposition 17.14 – see the third and fourth paragraphs of the proof given below.

Proof of Theorem 18.2. As observed in Remark 18.1(c), the uniqueness of the solution follows from Theorem 17.15 and Exercise 17.6.

To prove the existence, let

$$\mathcal{V} := \bigl\{u\in MR(0,\tau;V,H)\,;\ u(0)\in V\bigr\}.$$

(Recall that $MR(0,\tau;H,V)\hookrightarrow H^1(0,\tau;H)\hookrightarrow C([0,\tau];H)$; hence the requirement that $u(0)$ should belong to V makes sense.) Then $\mathcal{V}$ is a Hilbert space for the norm $\|\cdot\|_{\mathcal{V}}$,

$$\|u\|_{\mathcal{V}}^2 := \int_0^\tau \|u'(t)\|_H^2\,\mathrm{d}t + \int_0^\tau \|u(t)\|_V^2\,\mathrm{d}t + a_1(0,u(0),u(0)).$$

Let $\mathcal{W} := H^1(0,\tau;V)$ with the same norm; note that $\mathcal{W}$ is not complete.

If $u\in\mathcal{V}$ is a solution of (18.1) (with the given initial value $u_0\in V$ and the given inhomogeneity f), then one easily sees that $E(u,w) = L(w)$ for all $w\in\mathcal{W}$, where the sesquilinear form $E\colon\mathcal{V}\times\mathcal{W}\to\mathbb{K}$ and $L\in\mathcal{W}^*$ are defined by

$$E(v,w) := \int_0^\tau (v'(t)\,|\,w'(t))_H \mathrm{e}^{-\gamma t}\,\mathrm{d}t + \int_0^\tau a(t,v(t),w'(t))\mathrm{e}^{-\gamma t}\,\mathrm{d}t + a_1(0,v(0),w(0)),$$
$$L(w) := \int_0^\tau (f(t)\,|\,w'(t))_H \mathrm{e}^{-\gamma t}\,\mathrm{d}t + a_1(0,u_0,w(0)) \qquad (v\in\mathcal{V},\ w\in\mathcal{W});$$

here, the constant $\gamma\in\mathbb{R}$ is arbitrary – for the moment – and will be fixed later.

Conversely, assume that $u\in\mathcal{V}$ is such that

$$E(u,w) = L(w) \qquad (w\in\mathcal{W}). \tag{18.6}$$

We show that then u is a solution of (18.1). Let $\psi \in C_c^\infty(0,\tau)$ and $x \in V$. Then $w(t) := \int_0^t \psi(s)\,ds\; x$ defines an element $w \in \mathcal{W}$; hence by (18.6) we obtain

$$\int_0^\tau (u'(t)\,|\,x)_H \psi(t) e^{-\gamma t}\,dt + \int_0^\tau a(t,u(t),x)\psi(t)e^{-\gamma t}\,dt = \int_0^\tau (f(t)\,|\,x)_H\psi(t)e^{-\gamma t}\,dt.$$

Since $\psi \in C_c^\infty(0,\tau)$ is arbitrary, it follows that

$$(u'(t)\,|\,x)_H + a(t,u(t),x) = (f(t)\,|\,x)_H$$

for a.e. $t \in (0,\tau)$. This being valid for all $x \in V$ implies that $u'(t) + \mathcal{A}(t)u(t) = f(t)$ for a.e. $t \in (0,\tau)$, i.e. u solves the differential equation. Applying (18.6) with the constant function $w(t) = x$ we obtain $a_1(0,u(0),x) = a_1(0,u_0,x)$ for all $x \in V$, which implies $u(0) = u_0$. Thus, u solves (18.1).

Clearly $E(\cdot,w) \in \mathcal{V}'$ for all $w \in \mathcal{W}$, i.e., E satisfies condition (i) of Lions' representation theorem, Theorem 17.11. We show that condition (ii) is satisfied as well if γ is chosen appropriately. This will achieve the proof of existence: then by Theorem 17.11 there exists $u \in \mathcal{V}$ satisfying (18.6), which means that u is a solution of (18.1), as shown above.

Let $w \in \mathcal{W} = H^1(0,\tau;V)$. By Corollary 18.13 there exists a non-autonomous form $a_1'\colon (0,\tau)\times V \times V \to \mathbb{K}$ which is bounded by M_1' such that $a_1(\cdot,w(\cdot),w(\cdot)) \in H^1(0,\tau)$ and

$$a_1(\cdot,w(\cdot),w(\cdot))' = a_1'(\cdot,w(\cdot),w(\cdot)) + 2\operatorname{Re} a_1(\cdot,w(\cdot),w'(\cdot)); \tag{18.7}$$

for the expression on the right-hand side we have applied the symmetry of a_1. We multiply by $e^{-\gamma\cdot}$; then by the product rule and the accretivity of a_1 it follows that

$$\begin{aligned}
&\big(a_1(\cdot,w(\cdot),w(\cdot))e^{-\gamma\cdot}\big)' \\
&\quad= -\gamma a_1(\cdot,w(\cdot),w(\cdot))e^{-\gamma\cdot} + a_1'(\cdot,w(\cdot),w(\cdot))e^{-\gamma\cdot} + 2\operatorname{Re} a_1(\cdot,w(\cdot),w'(\cdot))e^{-\gamma\cdot} \\
&\quad\leqslant -\gamma\alpha\|w(\cdot)\|_V^2 e^{-\gamma\cdot} + M_1'\|w(\cdot)\|_V^2 e^{-\gamma\cdot} + 2\operatorname{Re} a_1(\cdot,w(\cdot),w'(\cdot))e^{-\gamma\cdot},
\end{aligned}$$

which we rewrite as

$$\operatorname{Re} a_1(\cdot,w(\cdot),w'(\cdot))e^{-\gamma\cdot} \geqslant \frac{1}{2}\big(a_1(\cdot,w(\cdot),w(\cdot))e^{-\gamma\cdot}\big)' + \frac{1}{2}(\gamma\alpha - M_1')\|w(\cdot)\|_V^2 e^{-\gamma\cdot}.$$

By the Peter–Paul inequality we have

$$|b(t,w(t),w'(t))| \leqslant M_b\|w(t)\|_V\|w'(t)\|_H \leqslant \frac{1}{2}\|w'(t)\|_H^2 + \frac{M_b^2}{2}\|w(t)\|_V^2$$

for all $t \in [0,\tau]$. Altogether, recalling $a = a_1 + b$, we can now estimate

$$\begin{aligned}
\operatorname{Re} E(w,w) &= \int_0^\tau \|w'(t)\|_H^2 e^{-\gamma t}\,dt + \operatorname{Re}\int_0^\tau a(t,w(t),w'(t))e^{-\gamma t}\,dt + a_1(0,w(0),w(0)) \\
&\geqslant \frac{1}{2}\int_0^\tau \|w'(t)\|_H^2 e^{-\gamma t}\,dt + \frac{1}{2}\int_0^\tau \big(a_1(\cdot,w(\cdot),w(\cdot))e^{-\gamma\cdot}\big)'(t)\,dt \\
&\qquad + \frac{\gamma\alpha - M_1' - M_b^2}{2}\int_0^\tau \|w(t)\|_V^2 e^{-\gamma t}\,dt + a_1(0,w(0),w(0)).
\end{aligned}$$

Choosing $\gamma > 0$ such that $\gamma\alpha - M_1' - M_b^2 = 1$ and applying the fundamental theorem of calculus in W_1^1 (see (17.2)), we conclude that

$$\begin{aligned}\operatorname{Re} E(w,w) &\geqslant \frac{1}{2}\int_0^\tau \|w'(t)\|_H^2 \mathrm{e}^{-\gamma t}\,\mathrm{d}t + \frac{1}{2}a_1(\tau, w(\tau), w(\tau))\mathrm{e}^{-\gamma\tau} - \frac{1}{2}a_1(0,w(0),w(0)) \\ &\quad + \frac{1}{2}\int_0^\tau \|w(t)\|_V^2 \mathrm{e}^{-\gamma t}\,\mathrm{d}t + a_1(0,w(0),w(0)) \\ &\geqslant \frac{1}{2}\mathrm{e}^{-\gamma\tau}\|w\|_{\mathcal{V}}^2 .\end{aligned} \tag{18.8}$$

Thus E satisfies condition (ii) of Theorem 17.11.

In order to obtain the estimate (18.3) we first note that the embedding constant of $\mathcal{W} \hookrightarrow \mathcal{V}$ is $c = 1$. Since $\gamma > 0$, the norm of L can be estimated by

$$\|L\|_{\mathcal{W}^*} \leqslant \big(\|f\|_{L_2(0,\tau;H)}^2 + a_1(0,u_0,u_0)\big)^{1/2} \leqslant \|f\|_{L_2(0,\tau;H)} + M_{1,0}^{1/2}\|u_0\|_V ,$$

and (18.8) together with the estimate in Theorem 17.11 yields (18.3). (Note that the $\mathcal{V}$-norm of u on the left-hand side of the resulting inequality dominates the $MR(0,\tau;H,V)$-norm of u.) □

Notes

Let us consider a non-autonomous form as described in Section 18.1. In the main result of this chapter, Theorem 18.2, it is proved that the form satisfies maximal regularity in H if it is symmetric and Lipschitz continuous (up to some perturbation). This result was proved in [ADLO14; Theorem 4.2] in greater generality and with a different approach. The special case in which the form a is symmetric and C^1 (instead of Lipschitz continuous) is due to J.-L. Lions [Lio61; Chap. IV, Théorème 6.1]. Lions [Lio61; p. 68] asked whether maximal regularity in H still holds if the form is merely continuous (in time) or even measurable. Today this question is settled and we want to tell first what is known if we restrict ourselves to the initial value $u_0 = 0$.

We say that the problem

$$u' + \mathcal{A}u = f, \quad u(0) = 0 \tag{18.9}$$

satisfies **maximal regularity in *H*** if for all $f \in L_2(0,\tau;H)$ there exists a unique solution $u \in MR(0,\tau;V,H)$ of (18.9). It was proved by Ouhabaz and Spina [OuSp10] that (18.9) has maximal regularity in H if a is Hölder continuous with exponent $\beta > 1/2$, i.e.

$$|a(t,u,v) - a(s,u,v)| \leqslant M|t-s|^\beta \|u\|_V\|v\|_V$$

for all $s,t \in [0,\tau]$, $u,v \in V$. Fackler [Fac17] showed that the exponent $\beta > 1/2$ is optimal even in the symmetric case. More precisely, he constructed a symmetric non-autonomous form a satisfying

$$|a(t,u,v) - a(s,u,v)| \leqslant M|t-s|^{1/2} \|u\|_V\|v\|_V$$

for all $s, t \in [0, \tau]$, $u, v \in V$ such that (18.9) does not have maximal regularity in H. We refer to [ADF17] for more information. It remained an open question whether one has maximal regularity in H for Hölder exponents $\beta \leqslant 1/2$ in the special case of a non-autonomous form associated with a family of differential operators as in Example 18.4. In the case of complex coefficients a negative answer was given by Bechtel, Mooney and Veraar [BMV24] in dimension $n \geqslant 2$. The problem remains open for real coefficients and in dimension $n = 1$.

Another question is for which initial values the solution has maximal regularity in H. This is even a problem in the autonomous case. Let $V \overset{\mathrm{d}}{\hookrightarrow} H$, and let $a\colon V \times V \to \mathbb{K}$ be a bounded coercive form. Denote by A the operator in H associated with a and by S the C_0-semigroup generated by $-A$. Let $u_0 \in H$. Then $u := S(\cdot)u_0$ is the unique solution $u \in MR(0, \tau; V, V^*)$ of $u'(t) + \mathcal{A}u(t) = 0$, $u(0) = u_0$ (where $\mathcal{A}$ is given by $\langle \mathcal{A}u, v\rangle = a(u, v)$); see Exercise 17.5. It is known from semigroup theory that $u \in H^1(0, \tau; H)$ if and only if $u_0 \in \operatorname{dom}(A^{1/2})$, the domain of the square root of A (which we do not want to define here). If a is symmetric, then $\operatorname{dom}(A^{1/2}) = V$ (see Proposition 13.23), but in general it is not easy to determine $\operatorname{dom}(A^{1/2})$. McIntosh gave an example of a form a for which $\operatorname{dom}(A^{1/2}) \neq V$ (see [Are04; Section 5.5] for references and more details). The following was a big problem, open for many years.

Kato's square root problem. *Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $a\colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{K}$ be the form defined in* (11.3). *Does it then follow that* $\operatorname{dom}(A^{1/2}) = H^1(\Omega)$*?*

The problem was finally solved by Auscher et al. [AHL&02], who showed that the answer is positive for $\Omega = \mathbb{R}^n$ (see [EHT16] for bounded Lipschitz domains Ω).

Exercises

18.1 The aim of this exercise is to establish continuous dependence of the solution u on the non-autonomous form b, in Theorem 18.2.

For a bounded non-autonomous form $b\colon [0, \tau] \times V \times H \to \mathbb{K}$, as in Section 18.1, we denote by $B(t) \in \mathcal{L}(V, H)$ the 'refined' Lax–Milgram operator defined by

$$(B(t)x \,|\, y)_H = b(t, x, y) \qquad (x \in V,\ y \in H),$$

for $t \in [0, \tau]$. Note that b is bounded with bound M_b if and only if $\sup_{0 \leqslant t \leqslant \tau} \|B(t)\|_{\mathcal{L}(V,H)} \leqslant M_b$.

(a) Let (B_n) be a sequence of operator functions as above, associated with a sequence of forms (b_n) satisfying $\sup_n M_{b_n} < \infty$. For $n \in \mathbb{N}$ and $f \in L_2(0, \tau; H)$ let v_n be the solution of

$$v_n' + \mathcal{A}_1 v_n + B_n v_n = f, \quad v_n(0) = 0.$$

Show that there exists a constant $c \geqslant 0$, not depending on n and f, such that

$$\|v_n\|_{MR(0,\tau;V,H)} \leqslant c\|f\|_{L_2(0,\tau;H)}\,.$$

(Hint: Evaluate the constant in (18.3).)

(b) Let (B_n) be a sequence as in (a), $B_n(t) \to B(t)$ strongly in $\mathcal{L}(V,H)$ for a.e. $t \in [0,\tau]$, with B as above. Let $u_0 \in V$, $f \in L_2(0,\tau;H)$, and let $u_n, u \in MR(0,\tau;V,H)$ be the unique solutions of

$$\begin{aligned} u_n' + \mathcal{A}_1 u_n + B_n u_n = f, \quad u_n(0) = u_0\,, \\ u' + \mathcal{A}_1 u + Bu = f, \quad\ \ u(0) = u_0\,, \end{aligned}$$

respectively. Show that $u_n \to u$ in $MR(0,\tau;V,H)$.

Hints: For $v_n := u_n - u$ one obtains, subtracting the two equations,

$$v_n' + \mathcal{A}_1 v_n + B_n v_n = (B - B_n)u, \quad v_n(0) = 0.$$

Show that $(B - B_n)u \to 0$ in $L_2(0,\tau;H)$, using the dominated convergence theorem. Then apply part (a).

18.2 Prove Lemma 18.11. (Hint: Imitate the proof of Proposition 18.7.)

18.3 Let $H := L_2(0,\pi)$, and let $a\colon V \times V \to \mathbb{K}$ be the classical Dirichlet form,

$$a(u,v) := \int_0^\pi u'\overline{v'}\,\mathrm{d}x$$

on V as indicated below. Denote by $\mathcal{A} \in \mathcal{L}(V,V^*)$ the associated Lax–Milgram operator in the Gelfand triple $V \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} V^*$. Let $u := \sin$.

(a) Show that $\mathcal{A}u = u$ if $V = H_0^1(0,\pi)$.

(b) Show that $\mathcal{A}u = u - \delta_0 - \delta_\pi$ if $V = H^1(0,\pi)$, where δ_0, δ_π denote the Dirac measures at $0,\pi$. Conclude that u does not belong to the domain of the operator associated with a on $H^1(0,\pi)$.

18.4 Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. In parts (a) and (b) let $a(u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x$ be the classical Dirichlet form on the indicated space V. Also, denote by $\mathcal{A} \in \mathcal{L}(V,V^*)$ the associated Lax–Milgram operator in the Gelfand triple $V \overset{\mathrm{d}}{\hookrightarrow} L_2(\Omega) \overset{\mathrm{d}}{\hookrightarrow} V^*$.

(a) Let $V := H_0^1(\Omega)$, $u \in H_0^1(\Omega)$. Show that $\mathcal{A}u \in L_2(\Omega)$ if and only if $\Delta u \in L_2(\Omega)$ (in the distributional sense), and that then

$$\langle \mathcal{A}u, v\rangle = -\int \Delta u \overline{v}\,\mathrm{d}x \qquad (v \in H_0^1(\Omega)).$$

(b) Assume that Ω has C^1-boundary, and let $V := H^1(\Omega)$, $u \in C^2(\overline{\Omega})$. Show that

$$\langle \mathcal{A}u, v\rangle = -\int_\Omega \Delta u \overline{v}\,\mathrm{d}x + \int_{\partial\Omega} \overline{v}\nabla u \cdot \nu\,\mathrm{d}\sigma \qquad (v \in H^1(\Omega)).$$

Determine those $u \in C^2(\overline{\Omega})$ for which $\mathcal{A}u \in L_2(\Omega)$. (Hint: Prove the formula first for $v \in C^1(\overline{\Omega})$; then use Theorem 7.11.)

(c) As in (b), assume that Ω has C^1-boundary, and let $V := H^1(\Omega)$. Let $\beta \in L_\infty(\partial\Omega)$, and define $a\colon V \times V \to \mathbb{K}$ by

$$a(u,v) := \int_\Omega \nabla u \cdot \overline{\nabla v}\,\mathrm{d}x + \int_{\partial\Omega} \beta u \overline{v}\,\mathrm{d}\sigma \qquad (u,v \in V).$$

(Recall from Section 7.5 that a is bounded and quasi-coercive.)

Let $u \in C^2(\overline{\Omega})$. Show that

$$\langle \mathcal{A}u, v\rangle = -\int_\Omega \Delta u \overline{v}\,\mathrm{d}x + \int_{\partial\Omega} \overline{v}(\beta u + \nabla u \cdot \nu)\,\mathrm{d}\sigma \qquad (v \in H^1(\Omega)).$$

Determine those $u \in C^2(\overline{\Omega})$ for which $\mathcal{A}u \in L_2(\Omega)$.

18.5 Under the hypotheses of Theorem 18.2, show that there exists a constant $c > 0$, not depending on τ, such that

$$\|u\|_{MR(0,\tau;V,H)} \leqslant c\mathrm{e}^{\tilde{\gamma}\tau}\big(\|f\|_{L_2(0,\tau;H)} + \|u_0\|_V\big)$$

for all solutions $u \in MR(0,\tau;V,H)$ of (18.1), where $\tilde{\gamma} := \frac{M_b^2}{2\alpha}$. (The aim of this exercise is improving the constant γ in (18.3). In particular, if $b = 0$, then the estimate does not depend on τ.)

Hints: Let $u \in MR(0,\tau;V,H)$ be a solution of (18.1).

1. Using Exercise 17.6 and Remark 18.1(a), show that

$$\|u\|_{L_2(0,\tau;V)} \leqslant c_0\mathrm{e}^{\tilde{\gamma}\tau}\big(\tfrac{2}{\alpha}\|f\|_{L_2(0,\tau;H)} + \tfrac{2}{\sqrt{\alpha}}\|u_0\|_V\big),$$

where c_0 is an embedding constant of $V \hookrightarrow H$.

2. Using

$$\|u'\|^2_{L_2(0,\tau;H)} = (f - \mathcal{B}u \,|\, u') - (\mathcal{A}_1 u \,|\, u') = (f - \mathcal{B}u \,|\, u') - a_1(\cdot, u(\cdot), u'(\cdot))$$

and (18.7), show that

$$\|u'\|^2_{L_2(0,\tau;H)} \leqslant \|f - \mathcal{B}u\|_{L_2(0,\tau;H)}\|u'\|_{L_2(0,\tau;H)} + \tfrac{1}{2}\big(M_1'\|u\|^2_{L_2(0,\tau;V)} + a_1(0,u_0,u_0)\big).$$

From Exercise 17.8(a) conclude that

$$\|u'\|_{L_2(0,\tau;H)} \leqslant \|f\|_{L_2(0,\tau;H)} + \Big(M_b + \sqrt{\tfrac{M_1'}{2}}\Big)\|u\|_{L_2(0,\tau;V)} + \sqrt{\tfrac{M_{1,0}}{2}}\|u_0\|_V\,.$$

Chapter 19

Nonlinear non-autonomous equations

The aim of this chapter is to treat nonlinear non-autonomous Cauchy problems of the form

$$u'(t) + \mathcal{A}(t)u(t) = Fu(t), \quad u(0) = u_0\,, \tag{19.1}$$

where F is a suitable nonlinear mapping between certain function spaces. A sketch of the method for obtaining a solution is as follows. Given w in the function space on which F is defined, find a solution u of the linear problem

$$u'(t) + \mathcal{A}(t)u(t) = Fw(t), \quad u(0) = u_0\,, \tag{19.2}$$

thus obtaining a (nonlinear) mapping $T\colon w \mapsto u$. Then a fixed point of T will be a solution of (19.1).

The linear problem (19.2) has been treated in the previous two chapters, and it will become clear that the regularity of solutions obtained in Chapter 18 is essential for the method sketched above. Another essential issue is the existence of a fixed point of the mapping T. The corresponding fixed point theorems are the subject of Section 19.1; a compactness property that is needed for their application will be treated in Section 19.2.

19.1 Schauder's and Schaefer's fixed point theorems

Fixed point theorems belong to the principal tools for treating nonlinear problems. In the well-known Banach fixed point theorem the mapping is assumed to be a strict contraction; then there exists a unique fixed point. In contrast, in Schauder's fixed point theorem – which is what we want to apply – the essential hypothesis is compactness, and the fixed point cannot be expected to be unique.

19.1 Theorem (Schauder). *Let X be a Banach space, $C \subseteq X$ a closed convex subset, and let T be a continuous self-map of C, with the property that $T(C)$ is relatively compact in X. Then T has a fixed point, i.e. there exists $x \in C$ such that $Tx = x$.*

We refer to Appendix I for the proof; see Theorem I.8.

As a first application, we show how one can apply Schauder's fixed point theorem to a nonlinear elliptic problem.

19.2 Example (A semilinear elliptic equation). Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, and let $f\colon \mathbb{K} \to \mathbb{K}$ be bounded and continuous. Then there exists $u \in H^1_0(\Omega)$ such that

$$-\Delta u = f \circ u.$$

Proof. Note that for all $u \in L_2(\Omega)$, $f \circ u$ belongs to $L_\infty(\Omega) \subseteq L_2(\Omega)$. We show that the mapping $F\colon L_2(\Omega) \to L_2(\Omega)$, given by $F(u) := f \circ u$, is continuous. Indeed, let (u_k) be a sequence in $L_2(\Omega)$, $u_k \to u$ in $L_2(\Omega)$. There exists a subsequence (u_{k_m}) converging to u almost everywhere. The dominated convergence theorem implies that $f \circ u_{k_m} \to f \circ u$ in $L_2(\Omega)$. Now the standard sub-subsequence argument shows that $f \circ u_k \to f \circ u$ as $k \to \infty$.

Let $R := (-\Delta_D)^{-1} \in \mathcal{L}(L_2(\Omega))$ be the inverse of the negative Dirichlet Laplacian; see Example 6.19. Then $T := R \circ F\colon L_2(\Omega) \to L_2(\Omega)$ is continuous. From Example 6.19 we know that R is a compact operator. The range of F is a bounded subset of $L_2(\Omega)$, and therefore the range of $R \circ F$ is relatively compact in $L_2(\Omega)$. Now Schauder's fixed point theorem, Theorem 19.1, implies the existence of a fixed point $u = Tu = RF(u)$; then $u \in \mathrm{dom}(\Delta_{\mathrm{D}}) \subseteq H_0^1(\Omega)$ and $-\Delta u = f \circ u$. □

In the previous example, the function f in the nonlinearity is assumed to be bounded, which is a rather strong hypothesis. The following fixed point theorem, due to Schaefer, will enable us to treat more general nonlinearities. In this theorem, the relative compactness of the range of the mapping is no longer required; instead, an a priori estimate is needed, which is in fact a very natural assumption for problems in partial differential equations.

A (nonlinear) mapping $T\colon X \to Y$, where X, Y are Banach spaces, is called **compact** if $T(B_X(0,r))$ is relatively compact for all $r > 0$.

19.3 Theorem (Schaefer). *Let X be a Banach space, and let $T\colon X \to X$ be continuous and compact. Assume that the* ***Schaefer set***

$$S := \big\{x \in X\,;\, \exists \lambda \in (0,1)\colon \lambda Tx = x\big\}$$

is bounded. Then T has a fixed point.

Proof. Let $r > 0$ be such that $S \subseteq B(0,r)$, and let $P\colon X \to B[0,r]$ be the projection given by $Px := r\frac{x}{\|x\|}$ if $\|x\| > r$, $P|_{B[0,r]} = \mathrm{id}_{B[0,r]}$. Then $T \circ P$ is continuous and has relatively compact range; therefore Schauder's fixed point theorem, Theorem 19.1, implies that $T \circ P$ has a fixed point $x \in X$, $TPx = x$.

We show that $x \in B[0,r]$; then it follows that $x = Px$ is a fixed point of T. If $\|x\|$ were greater than r one would obtain $Px = \lambda x = \lambda TPx$, with $\lambda = \frac{r}{\|x\|} \in (0,1)$. This would imply that $Px \in S$, contradicting the assumption on r, because $\|Px\| = r$. □

As an application of Schaefer's fixed point theorem we show how the nonlinearity in Example 19.2 can be generalised.

19.4 Example (A semilinear elliptic equation, revisited). Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Denote by $\lambda_1 = \min \sigma(-\Delta_D) > 0$ the first Dirichlet eigenvalue, and let $0 \leqslant \alpha < \lambda_1$, $\beta \geqslant 0$. Let $f\colon \mathbb{K} \to \mathbb{K}$ be continuous, satisfying

$$|f(z)| \leqslant \alpha|z| + \beta \qquad (z \in \mathbb{K}).$$

Let $g \in L_2(\Omega)$. Then there exists $u \in H_0^1(\Omega)$ such that

$$-\Delta u = f \circ u + g.$$

Proof. We show that the mapping $F\colon L_2(\Omega) \to L_2(\Omega)$, defined by $F(u) := f \circ u + g$, is continuous. Let (u_k) be a sequence in $L_2(\Omega)$, $u_k \to u$ in $L_2(\Omega)$. There exist a subsequence (u_{k_m}) and a function $h \in L_2(\Omega)$ such that $u_{k_m} \to u$ a.e. and $|u_{k_m}| \leqslant h$ a.e. for all $m \in \mathbb{N}$. Then $|f \circ u_{k_m}| \leqslant \alpha h + \beta$ a.e. for all $m \in \mathbb{N}$, and the dominated convergence theorem implies that $F(u_{k_m}) \to F(u)$ in $L_2(\Omega)$. Now the standard sub-subsequence argument (Exercise 9.7) yields the asserted convergence $F(u_k) \to F(u)$ as $k \to \infty$.

As in Example 19.2 we define $R := (-\Delta_D)^{-1} \in \mathcal{L}(L_2(\Omega))$ and recall that R is a compact operator. Then $T := R \circ F\colon L_2(\Omega) \to L_2(\Omega)$ is continuous. In order to show that T is compact, let $r > 0$, $u \in B_{L_2(\Omega)}(0, r)$. Since $|f \circ u| \leqslant \alpha|u| + \beta$, it follows that $\|F(u)\|_{L_2(\Omega)} \leqslant \alpha r + \beta\|\mathbf{1}_\Omega\|_{L_2(\Omega)} + \|g\|_{L_2(\Omega)}$. Hence, $F(B(0,r))$ is bounded in $L_2(\Omega)$, and $T(B(0,r)) = R\bigl(F(B(0,r))\bigr)$ is relatively compact in $L_2(\Omega)$.

We have shown that T is a compact continuous mapping. In order to apply Schaefer's fixed point theorem, Theorem 19.3, we need to prove that the Schaefer set

$$S := \bigl\{v \in L_2(\Omega);\, \exists \lambda \in (0,1)\colon \lambda T v = v\bigr\}$$

is bounded. If this is achieved, the theorem implies that there exists $u \in L_2(\Omega)$ with $Tu = u$; then $u \in H_0^1(\Omega)$ and $-\Delta u = f \circ u + g$.

Let $v \in S$, i.e., $v \in H_0^1(\Omega)$ and there exists $\lambda \in (0,1)$ such that

$$-\Delta v = \lambda(f \circ v + g).$$

Then

$$\int_\Omega |\nabla v|^2 \,\mathrm{d}x = \lambda \int_\Omega (f \circ v + g)\overline{v} \,\mathrm{d}x \leqslant \int_\Omega \alpha |v|^2 \,\mathrm{d}x + \int_\Omega \beta|v| \,\mathrm{d}x + \int_\Omega |g||v| \,\mathrm{d}x.$$

Using the Peter–Paul inequality we obtain

$$\int_\Omega |\nabla v|^2 \,\mathrm{d}x \leqslant \alpha \int_\Omega |v|^2 \,\mathrm{d}x + \varepsilon \int_\Omega |v|^2 \,\mathrm{d}x + \frac{1}{4\varepsilon}\int_\Omega (\beta + |g|)^2 \,\mathrm{d}x,$$

for all $\varepsilon > 0$. By Poincaré's inequality (see Example 6.19) one has $\lambda_1 \int_\Omega |v|^2 \leqslant \int_\Omega |\nabla v|^2$; hence it follows that

$$(\lambda_1 - \alpha - \varepsilon)\int_\Omega |v|^2 \,\mathrm{d}x \leqslant \frac{1}{4\varepsilon}\int_\Omega (\beta + |g|)^2 \,\mathrm{d}x.$$

Choosing $0 < \varepsilon < \lambda_1 - \alpha$, we conclude that the Schaefer set is bounded, and the proof is complete. □

19.5 Remark. The assertion of Example 19.4 no longer holds for $\alpha = \lambda_1$. Indeed, choose $f(z) := \lambda_1 z$ and $g := \varphi_1$, where $\varphi_1 \in H_0^1(\Omega)$, $-\Delta\varphi_1 = \lambda_1\varphi_1$, $\|\varphi_1\|_{L_2(\Omega)} = 1$. Suppose that $u \in H_0^1(\Omega)$ satisfies $-\Delta u = \lambda_1 u + \varphi_1$. Then $\int \nabla u \cdot \overline{\nabla v} = \lambda_1 \int u\overline{v} + \int \varphi_1 \overline{v}$ for all $v \in H_0^1(\Omega)$. In particular, for $v = \varphi_1$ one obtains

$$\lambda_1 \int u\overline{\varphi_1} \,\mathrm{d}x = \int \nabla u \cdot \overline{\nabla\varphi_1} \,\mathrm{d}x = \lambda_1 \int u\overline{\varphi_1} \,\mathrm{d}x + \int |\varphi_1|^2 \,\mathrm{d}x,$$

which is absurd. △

19.2 Interlude: Compact embeddings of mixed spaces

Let $\tau \in (0,\infty)$. We have seen that the embedding $H^1(0,\tau) \hookrightarrow L_2(0,\tau)$ is compact; see Theorem 7.9. Now, if H is an infinite dimensional separable Hilbert space, then the embedding $H^1(0,\tau;H) \hookrightarrow L_2(0,\tau;H)$ is no longer compact. (Indeed, let (e_n) be an orthonormal sequence in H, and consider the functions $u_n(t) := e_n$ $(0 \leqslant t \leqslant \tau)$ in $H^1(0,\tau;H)$. The sequence $(u_n)_{n\in\mathbb{N}}$ has no convergent subsequence in $L_2(0,\tau;H)$.) However, things change if we consider the vector-valued spaces taking values in two different Hilbert spaces.

19.6 Theorem (Aubin–Lions). *Let V, H be separable Hilbert spaces, with compact embedding $V \hookrightarrow H$. Then the embedding*

$$MR(0,\tau;V,H) \hookrightarrow L_2(0,\tau;H)$$

is compact.

Proof. As $MR(0,\tau;V,H)$ is a Hilbert space, hence reflexive, it is sufficient to show that every weakly convergent sequence (u_n) in $MR(0,\tau;V,H)$ is convergent in $L_2(0,\tau;H)$; see Exercise 6.6(b). Without loss of generality $u_n \to 0$ weakly in $MR(0,\tau;V,H)$, i.e., $u_n \to 0$ weakly in $L_2(0,\tau;V)$ and $u_n' \to 0$ weakly in $L_2(0,\tau;H)$, and also without loss of generality $\|u_n'\|_{L_2(0,\tau;H)} \leqslant 1$ for all $n \in \mathbb{N}$. We have to prove that $u_n \to 0$ in $L_2(0,\tau;H)$.

First we fix $t \in (0,\tau)$ and show that $u_n(t) \to 0$ in H. Let $\varepsilon > 0$. For the following computations recall Proposition 17.6(b) (as well as Remark 17.7, for the product rule). For $0 < s_0 < \tau - t$ we have

$$\begin{aligned} u_n(t) &= \frac{1}{s_0}\int_0^{s_0} \frac{\mathrm{d}}{\mathrm{d}s}\big((s-s_0)u_n(t+s)\big)\,\mathrm{d}s \\ &= \frac{1}{s_0}\int_0^{s_0} u_n(t+s)\,\mathrm{d}s + \frac{1}{s_0}\int_0^{s_0}(s-s_0)u_n'(t+s)\,\mathrm{d}s, \end{aligned}$$

where the integrals should be taken in H. Applying the Cauchy–Schwarz inequality we find that the H-norm of the second term in the last expression is $\leqslant \sqrt{s_0}$; we choose $s_0 \leqslant \varepsilon^2$ to make this norm $\leqslant \varepsilon$. Note that the integral in the first term can also be considered as an integral in V, by Exercise 17.3(b) (applied with the embedding $V \hookrightarrow H$ as the operator $A \in \mathcal{L}(V,H)$). Moreover the mapping $L_2(0,\tau;V) \ni u \mapsto \frac{1}{s_0}\int_0^{s_0} u(t+s)\,\mathrm{d}s \in V$ is linear and continuous, hence continuous with respect to the weak topologies. It follows that the sequence $\big(\frac{1}{s_0}\int_0^{s_0} u_n(t+s)\,\mathrm{d}s\big)_n$ converges weakly to 0 in V, hence in norm to 0 in H, by the compactness of the embedding $V \hookrightarrow H$. Thus we have shown that $\limsup_{n\to\infty}\|u_n(t)\|_H \leqslant \varepsilon$.

From Proposition 17.6(c) we recall that $H^1(0,\tau;H)$ is continuously embedded into $C([0,\tau];H)$; hence there exists a constant $c \geqslant 0$ such that $\|u_n(t)\|_H \leqslant c$ for all $t \in (0,\tau)$, $n \in \mathbb{N}$. Now the dominated convergence theorem implies that $u_n \to 0$ in $L_2(0,\tau;H)$ as $n \to \infty$. □

The compactness property in Theorem 19.6 can be reinforced to compactness of the embedding $MR(0,\tau;V,H) \hookrightarrow C([0,\tau];H)$; see Exercise 19.3.

19.3 More on maximal regularity

In this section we return to the non-autonomous problem (18.1) and prove a further regularity property of the solutions obtained in Theorem 18.2. As in Section 18.1, let V, H be separable Hilbert spaces, $V \overset{d}{\hookrightarrow} H$. Let $\tau \in (0,\infty)$. We assume that $a\colon [0,\tau]\times V\times V \to \mathbb{K}$ is a Lipschitz continuous symmetric bounded coercive non-autonomous form. The bound of a will be denoted by M, the Lipschitz bound by M' and the coercivity constant by α. (The present hypotheses on the form a correspond to the hypotheses on a_1 in Section 18.1; the form b from that section will not be used here.)

For the discussion below as well as for later use we define the space

$$MR_a(0,\tau;V,H) := \{u \in MR(0,\tau;V,H);\ \mathcal{A}u \in L_2(0,\tau;H)\},$$

with norm

$$\|u\|_{MR_a(0,\tau;V,H)} := \big(\|u\|^2_{MR(0,\tau;V,H)} + \|\mathcal{A}u\|^2_{L_2(0,\tau;H)}\big)^{1/2}.$$

We point out that a solution $u \in MR(0,\tau;V,H)$ of (18.1) automatically lies in $MR_a(0,\tau;V,H)$; using $\mathcal{A}u = f - u'$ and the estimate (18.3) one obtains

$$\begin{aligned}\|u\|_{MR_a(0,\tau;V,H)} &\leqslant \|u\|_{MR(0,\tau;V,H)} + \|\mathcal{A}u\|_{L_2(0,\tau;H)} \leqslant 2\|u\|_{MR(0,\tau;V,H)} + \|f\|_{L_2(0,\tau;H)}\\ &\leqslant 5\mathrm{e}^{\gamma\tau}\big(\|f\|_{L_2(0,\tau;H)} + M_0^{1/2}\|u_0\|_V\big),\end{aligned} \tag{19.3}$$

where $\gamma = \frac{1+M'}{\alpha}$, and M_0 is a bound for the form $a(0,\cdot,\cdot)$.

19.7 Theorem. *With the previous notation one has*

$$MR_a(0,\tau;V,H) \hookrightarrow C([0,\tau];V).$$

More explicitly, for each $u \in MR_a(0,\tau;V,H)$ one has $u \in C([0,\tau];V)$ and

$$\|u\|_{C([0,\tau];V)} \leqslant c_0\|u\|_{MR_a(0,\tau;V,H)}\,, \tag{19.4}$$

with a constant $c_0 \geqslant 0$ depending only on τ, α, M, M'. One also has

$$\|u\|_{C([0,\tau];V)} \leqslant c_1\big(\|u\|_{MR_a(0,\tau;V,H)} + \|u(0)\|_V\big), \tag{19.5}$$

with a constant $c_1 \geqslant 0$ depending only on α, M, M', but not on τ.

The proof requires a few preliminary results. The first of these deals with forms without t-dependence.

19.8 Proposition. *Let a be a bounded coercive form on V, $\operatorname{Re} a(u) \geqslant \alpha\|u\|_V^2$ for all $u \in V$, where $\alpha > 0$. Let $\mathcal{A}$ be the associated Lax–Milgram operator, considered as an operator in V^* with domain* $\operatorname{dom}(\mathcal{A}) = V$.

(a) *Then $\|(I+\mathcal{A})^{-1}\|_{\mathcal{L}(V^*,H)} \leqslant 1/(2\sqrt{\alpha})$ and $\|(I+\mathcal{A})^{-1}\|_{\mathcal{L}(H,V)} \leqslant 1/(2\sqrt{\alpha})$.*

(b) *Assume additionally that a is symmetric, and let $M \geqslant 0$ be a bound of a (i.e. $a(u) \leqslant M\|u\|_V^2$ for all $u \in V$). Then $\|(I+\mathcal{A})^{-1}\|_{\mathcal{L}(V)} \leqslant \sqrt{M/\alpha}$.*

Proof. (a) For $u \in V$ we estimate

$$2\|u\|_H\sqrt{\alpha}\|u\|_V \leqslant \|u\|_H^2 + \alpha\|u\|_V^2 \leqslant \operatorname{Re}\langle (I+\mathcal{A})u, u\rangle \leqslant \|(I+\mathcal{A})u\|_{V^*}\|u\|_V\,.$$

This implies $\|(I+\mathcal{A})^{-1}\eta\|_H \leqslant 1/(2\sqrt{\alpha})\|\eta\|_{V^*}$ for all $\eta \in V^*$.

Similarly, for $u \in \operatorname{dom}(A)$ one has

$$2\sqrt{\alpha}\|u\|_V\|u\|_H \leqslant \|u\|_H^2 + \alpha\|u\|_V^2 \leqslant \operatorname{Re}((I+\mathcal{A})u\,|\,u)_H \leqslant \|(I+\mathcal{A})u\|_H\|u\|_H\,,$$

and it follows that $\|(I+\mathcal{A})^{-1}x\|_V \leqslant 1/(2\sqrt{\alpha})\|x\|_H$ for all $x \in H$.

(b) As a is symmetric, the associated operator A in H is an accretive self-adjoint operator. Let $u \in V$, and put $v := (I+A)^{-1}u$. Then $Av = u - v \in V$,

$$\begin{aligned} a(u) = a(v+Av) &= a(v) + 2a(v,Av) + a(Av) \\ &= a(v) + 2\|Av\|_H^2 + a(Av) \geqslant a(v) = a((I+A)^{-1}u), \end{aligned}$$

hence

$$\alpha\|(I+A)^{-1}u\|_V^2 \leqslant a((I+A)^{-1}u) \leqslant a(u) \leqslant M\|u\|_V^2\,.$$

Since $(I+\mathcal{A})^{-1}|_V = (I+A)^{-1}|_V$ we obtain the assertion. □

The following proposition deals with approximation of the identity on $MR_a(0,\tau;V,H)$ and of the embedding $MR_a(0,\tau;V,H) \hookrightarrow L_2(0,\tau;V)$; it is the final technical tool for the proof of Theorem 19.7, and we will use it to prove the compactness property stated in Proposition 19.10 below.

19.9 Proposition. *Under the hypotheses as in Theorem 19.7, for $n \in \mathbb{N}$ we define*

$$\begin{aligned} \mathcal{D}_n(t) &:= (I + \tfrac{1}{n}\mathcal{A}(t))^{-1} \in \mathcal{L}(V^*,V) \qquad (t \in [0,\tau]), \\ \mathcal{D}_n u(t) &:= \mathcal{D}_n(t)u(t) \qquad (t \in (0,\tau),\ H^1(0,\tau;V^*)) \end{aligned}$$

(where we consider $\mathcal{A}(t)$ as an unbounded operator in V^*, with $\operatorname{dom}(\mathcal{A}(t)) = V$). Then

(a) *$\mathcal{D}_n \in \mathcal{L}(H^1(0,\tau;V^*), H^1(0,\tau;V))$, and $\mathcal{D}_n \in \mathcal{L}(MR_a(0,\tau;V,H))$;*

(b) *$\mathcal{D}_n \to I$ $(n \to \infty)$ strongly in $\mathcal{L}(MR_a(0,\tau;V,H))$;*

(c) *$\mathcal{D}_n \to j$ in $\mathcal{L}(MR_a(0,\tau;V,H), L_2(0,\tau;V))$ as $n \to \infty$, where j denotes the embedding $MR_a(0,\tau;V,H) \hookrightarrow L_2(0,\tau;V)$.*

Proof. (a) Note that $I + \frac{1}{n}\mathcal{A}(t)$ is the Lax–Milgram operator associated with the coercive form $\frac{1}{n}a(t,u,v) + (u\,|\,v)_H$ on V. Thus the Lax–Milgram lemma, Theorem 5.4, implies that $\|\mathcal{D}_n(t)\|_{\mathcal{L}(V^*,V)} \leqslant n/\alpha$ $(t \in [0,\tau])$. For all $s,t \in [0,\tau]$ we have

$$\mathcal{D}_n(t) - \mathcal{D}_n(s) = \tfrac{1}{n}\mathcal{D}_n(t)(\mathcal{A}(s) - \mathcal{A}(t))\mathcal{D}_n(s), \tag{19.6}$$

which shows that $\mathcal{D}_n\colon [0,\tau] \to \mathcal{L}(V^*,V)$ is Lipschitz continuous with Lipschitz constant $\frac{n}{\alpha^2}M'$ (where M' is the Lipschitz constant of a). By Proposition 18.12 it follows that $\mathcal{D}_n \in \mathcal{L}(H^1(0,\tau;V^*), H^1(0,\tau;V))$.

Using the continuous embeddings $MR(0,\tau;V,H) \hookrightarrow H^1(0,\tau;V^*)$ and $H^1(0,\tau;V) \hookrightarrow MR(0,\tau;V,H)$ we conclude that $\mathcal{D}_n \in \mathcal{L}(MR(0,\tau;V,H))$. Now the accretivity of $\frac{1}{n}a(t)$ implies that

$$\|\mathcal{D}_n(t)\|_{\mathcal{L}(H)} \leqslant 1 \qquad (t \in [0,\tau],\ n \in \mathbb{N}). \tag{19.7}$$

Since $\mathcal{A}\mathcal{D}_n = \mathcal{A}(I + \frac{1}{n}\mathcal{A})^{-1} = (I + \frac{1}{n}\mathcal{A})^{-1}\mathcal{A} = \mathcal{D}_n\mathcal{A}$ in $\mathcal{L}(V, V^*)$, we obtain

$$\|\mathcal{A}\mathcal{D}_n u\|_{L_2(0,\tau;H)} = \|\mathcal{D}_n\mathcal{A}u\|_{L_2(0,\tau;H)} \leqslant \|\mathcal{A}u\|_{L_2(0,\tau;H)}$$

for all $u \in MR_a(0,\tau;V,H)$. Combining this estimate with the property obtained before we conclude that $\mathcal{D}_n \in \mathcal{L}(MR_a(0,\tau;V,H))$.

(b) Let $u \in MR_a(0,\tau;V,H)$, and put $u_n := \mathcal{D}_n u$ $(n \in \mathbb{N})$. The fundamental observation for the asserted convergence is that $\mathcal{D}_n(t) \to I$ $(n \to \infty)$ strongly in $\mathcal{L}(H)$ for all $t \in [0,\tau]$, by Lemma 2.10. Using (19.7) and the dominated convergence theorem, we conclude that $u_n \to u$ and $\mathcal{A}u_n = \mathcal{A}\mathcal{D}_n u = \mathcal{D}_n\mathcal{A}u \to \mathcal{A}u$ in $L_2(0,\tau;H)$ as $n \to \infty$.

The convergence $\mathcal{A}u_n \to \mathcal{A}u$ in $L_2(0,\tau;H)$ implies $\mathcal{A}u_n \to \mathcal{A}u$ in $L_2(0,\tau;V^*)$. Because of the accretivity of a we have $\|\mathcal{A}(t)^{-1}\|_{\mathcal{L}(V^*,V)} \leqslant 1/\alpha$ for all $t \in [0,\tau]$; hence $u_n = \mathcal{A}^{-1}\mathcal{A}u_n \to \mathcal{A}^{-1}\mathcal{A}u = u$ in $L_2(0,\tau;V)$.

It remains to prove the convergence of the derivatives $u_n' = (\mathcal{D}_n u)'$ to u' in $L_2(0,\tau;H)$. The Lipschitz continuity of $\mathcal{D}_n\colon [0,\tau] \to \mathcal{L}(V^*,V)$, shown in part (a), implies that $\mathcal{D}_n$ is Lipschitz continuous considered as a function from $[0,\tau]$ to $\mathcal{L}(H)$. Hence by Propositions 18.10(b) and 18.12, $\mathcal{D}_n'(\cdot)$ exists in $\mathcal{L}(H)$ and $u_n' = \mathcal{D}_n' u + \mathcal{D}_n u'$. As we know from the previous considerations that $\mathcal{D}_n u' \to u'$ $(n \to \infty)$ in $L_2(0,\tau;H)$, we still have to show that $\mathcal{D}_n' u \to 0$ $(n \to \infty)$ in $L_2(0,\tau;H)$.

First, Proposition 19.8(a) implies that

$$\|\mathcal{D}_n(t)\|_{\mathcal{L}(V^*,H)} \leqslant \sqrt{n/(4\alpha)}, \quad \|\mathcal{D}_n(t)\|_{\mathcal{L}(H,V)} \leqslant \sqrt{n/(4\alpha)} \qquad (t \in [0,\tau],\ n \in \mathbb{N}). \tag{19.8}$$

From these inequalities and (19.6) we conclude that the functions $\mathcal{D}_n\colon [0,\tau] \to \mathcal{L}(H)$ are Lipschitz continuous, with a Lipschitz constant $M'/(4\alpha)$ independent of $n \in \mathbb{N}$. Thus by Proposition 18.10(b) the derivatives of $\mathcal{D}_n\colon [0,\tau] \to \mathcal{L}(H)$ can in fact be chosen such that

$$\|\mathcal{D}_n'(t)\|_{\mathcal{L}(H)} \leqslant M'/(4\alpha) \qquad (t \in (0,\tau),\ n \in \mathbb{N}).$$

Interpreting $\mathcal{D}_n'$ as the 'multiplication operator' $\mathcal{D}_n'\colon L_2(0,\tau;H) \to L_2(0,\tau;H)$, $v \mapsto \mathcal{D}_n'(\cdot)v(\cdot)$ we conclude that $(\mathcal{D}_n')$ is a bounded sequence in $\mathcal{L}(L_2(0,\tau;H))$.

From Proposition 19.8(b) we obtain the estimate $\|\mathcal{D}_n(t)\|_{\mathcal{L}(V)} \leqslant \sqrt{M/\alpha}$ for all $t \in [0,\tau]$, $n \in \mathbb{N}$. Combining this estimate with (19.6) and (19.8) we conclude that, for $x \in V$, the function $\mathcal{D}_n(\cdot)x\colon (0,\tau) \to H$ is Lipschitz continuous with Lipschitz constant $c\|x\|_V/\sqrt{n}$, where $c := \sqrt{M}M'/(2\alpha)$. Hence, Proposition 18.10(a) implies that

$$\|(\mathcal{D}_n(\cdot)x)'\|_{L_\infty(0,\tau;H)} \leqslant c\|x\|_V/\sqrt{n} \to 0 \qquad (n \to \infty).$$

It follows that $\mathcal{D}_n'(f(\cdot)x) = f(\mathcal{D}_n(\cdot)x)' \to 0$ in $L_2(0,\tau;H)$ for all $x \in V$, $f \in L_2(0,\tau)$. Since $\operatorname{lin}\{f(\cdot)x;\ x \in V,\ f \in L_2(0,\tau)\}$ is dense in $L_2(0,\tau;H)$ and the sequence $(\mathcal{D}_n')$ is bounded in $\mathcal{L}(L_2(0,\tau;H))$, we conclude that $\mathcal{D}_n' \to 0$ strongly in $\mathcal{L}(L_2(0,\tau;H))$; in particular, $\mathcal{D}_n' u \to 0$ in $L_2(0,\tau;H)$.

(c) It is easy to see that $j - \mathcal{D}_n = \frac{1}{n}\mathcal{D}_n\mathcal{A}$. Thus from $\|\mathcal{A}\|_{\mathcal{L}(MR_a(0,\tau;V,H),L_2(0,\tau;H))} \leqslant 1$ and the second estimate in (19.8) we obtain

$$\|j - \mathcal{D}_n\|_{\mathcal{L}(MR_a(0,\tau;V,H),L_2(0,\tau;V))} \leqslant \tfrac{1}{n}\|\mathcal{D}_n\|_{\mathcal{L}(L_2(0,\tau;H),L_2(0,\tau;V))} \leqslant \tfrac{1}{2\sqrt{n\alpha}} \to 0 \qquad (n \to \infty). \ \square$$

Proof of Theorem 19.7. Let $u \in MR_a(0,\tau;V,H)$.

(i) For the proof of the asserted inequalities we first assume in addition that $u \in H^1(0,\tau;V)$; then $u \in C([0,\tau];V)$ by Proposition 17.6(b). By Corollary 18.13 and the symmetry of a we obtain $a(\cdot,u(\cdot),u(\cdot)) \in H^1(0,\tau) \subseteq W_1^1(0,\tau)$ and

$$a(\cdot,u(\cdot),u(\cdot))' = a'(\cdot,u(\cdot),u(\cdot)) + 2\,\mathrm{Re}(\mathcal{A}u(\cdot)\,|\,u'(\cdot))_H\,.$$

Applying Proposition 17.8 we obtain

$$\begin{aligned}\alpha\|u\|^2_{C([0,\tau];V)} &\leqslant \|a(\cdot,u(\cdot),u(\cdot))\|_{C[0,\tau]} \qquad (19.9)\\ &\leqslant \frac{1}{\tau}\|a(\cdot,u(\cdot),u(\cdot))\|_{L_1(0,\tau)} + \|a(\cdot,u(\cdot),u(\cdot))'\|_{L_1(0,\tau)}\\ &\leqslant \frac{1}{\tau}\int_0^\tau M\|u(t)\|_V^2\,\mathrm{d}t + \int_0^\tau \big(|a'(t,u(t),u(t))| + 2\|\mathcal{A}u(t)\|_H\|u'(t)\|_H\big)\,\mathrm{d}t\\ &\leqslant \Big(\frac{M}{\tau}+M'\Big)\|u\|^2_{L_2(0,\tau;V)} + \|\mathcal{A}u\|^2_{L_2(0,\tau;H)} + \|u'\|^2_{L_2(0,\tau;H)}\,.\end{aligned}$$

In view of the 'fundamental theorem of calculus formula' (17.2), the estimate (19.9) can also be continued as

$$\begin{aligned}\alpha\|u\|^2_{C([0,\tau];V)} &\leqslant a(0,u(0),u(0)) + \int_0^\tau \big|a(\cdot,u(\cdot),u(\cdot))'(t)\big|\,\mathrm{d}t\\ &\leqslant M\|u(0)\|_V^2 + M'\|u\|^2_{L_2(0,\tau;V)} + \|\mathcal{A}u\|^2_{L_2(0,\tau;H)} + \|u'\|^2_{L_2(0,\tau;H)}\,.\end{aligned}$$

The above two estimates for $\|u\|_{C([0,\tau];V)}$ establish (19.4) and (19.5) under the additional assumption $u \in H^1(0,\tau;H)$. The constants in the estimates have the dependencies stated in the theorem.

(ii) For general u we use the mappings $\mathcal{D}_n$ defined in Proposition 19.9 and put $u_n := \mathcal{D}_n u$ ($n \in \mathbb{N}$). Then Proposition 19.9(a) shows that (u_n) is a sequence in $MR_a(0,\tau;V,H) \cap H^1(0,\tau;V)$, and $u_n \to u$ in $MR_a(0,\tau;V,H)$ by Proposition 19.9(b). The inequality (19.4) implies that (u_n) is a Cauchy sequence in $C([0,\tau];V)$, and since $u_n \to u$ in $L_2(0,\tau;V)$, it follows that u has a representative in $C([0,\tau];V)$ satisfying (19.4). Once the convergence $u_n \to u$ in $C([0,\tau];V)$ is established, it is clear that (19.5) also carries over from u_n to u. □

The following compactness property will be needed below in an application of Schaefer's fixed point theorem.

19.10 Proposition. *Suppose that the embedding $V \hookrightarrow H$ is compact. Then the space $MR_a(0,\tau;V,H)$ is compactly embedded into $L_2(0,\tau;V)$.*

Proof. Let $n \in \mathbb{N}$, and let $\mathcal{D}_n$ be defined as in Proposition 19.9. We show that $\mathcal{D}_n\colon MR_a(0,\tau;V,H) \to L_2(0,\tau;V)$ is compact. The Aubin–Lions lemma, Theorem 19.6, says that the embedding $MR(0,\tau;V,H) \hookrightarrow L_2(0,\tau;H)$ is compact, and the second estimate in (19.8) implies that $\mathcal{D}_n\colon L_2(0,\tau;H) \to L_2(0,\tau;V)$ is bounded. This shows, more strongly, that $\mathcal{D}_n\colon MR(0,\tau;V,H) \to L_2(0,\tau;V)$ is compact.

Then Proposition 19.9(c) implies that the embedding $MR_a(0,\tau;H,V) \hookrightarrow L_2(0,\tau;V)$ is approximated in norm by compact mappings, hence is compact. □

19.4 An abstract semilinear problem

Let V and H be Hilbert spaces, with compact embedding $V \overset{\mathrm{d}}{\hookrightarrow} H$. Let $\tau \in (0, \infty)$, and let $a\colon [0, \tau] \times V \times V \to \mathbb{K}$ be a Lipschitz continuous symmetric bounded non-autonomous form which is quasi-coercive, i.e., there exist $\omega \in \mathbb{R}$, $\alpha > 0$ such that

$$a(t, u, u) + \omega \|u\|_H^2 \geqslant \alpha \|u\|_V^2 \qquad (t \in [0, \tau],\ u \in V).$$

Denote by $\mathcal{A}(t)\colon V \to V^*$ the Lax–Milgram operator associated with $a(t, \cdot, \cdot)$.

Next we introduce the nonlinear term. Let $F\colon L_2(0, \tau; V) \to L_2(0, \tau; H)$ be continuous, and suppose that there exists a constant $c_F \geqslant 0$ such that

$$\|Fw\|_{L_2(0,t;H)} \leqslant c_F(1 + \|w\|_{L_2(0,t;V)}) \qquad (t \in (0, \tau],\ w \in L_2(0, \tau; V)). \tag{19.10}$$

The inequality in (19.10) expresses that the values of Fw on the interval $(0, t)$ should not depend too much on the 'future' values of w (on the interval (t, τ)). In Section 19.6 we will describe concrete examples of mappings F with these properties.

Given these hypotheses, we formulate the main result of the present chapter, concerning the existence of solutions of non-autonomous semilinear parabolic equations.

19.11 Theorem. *Let $u_0 \in V$. Then there exists $u \in MR(0, \tau; V, H)$ such that $u(0) = u_0$ and*

$$u'(t) + \mathcal{A}(t)u(t) = Fu(t) \tag{19.11}$$

for a.e. $t \in [0, \tau]$.

Note that there is no statement concerning uniqueness; we refer to Exercise 19.5 for an example with non-uniqueness.

Proof of Theorem 19.11. By the quasi-coercivity of a there exists $\omega \geqslant 0$ such that the form a_ω defined by

$$a_\omega(t, x, y) := a(t, x, y) + \omega(x \,|\, y)_H \qquad (x, y \in V,\ t \in [0, \tau])$$

is coercive. Note that a function $u \in MR(0, \tau; V, H)$ satisfies the properties stated in Theorem 19.11 if and only if u is a solution of

$$u'(t) + (\mathcal{A}(t) + \omega)u(t) = F_\omega u(t), \quad u(0) = u_0\,,$$

where

$$F_\omega w := Fw + \omega w \qquad (w \in L_2(0, \tau; V)).$$

This reduces the proof to the case in which a is coercive.

Applying Theorem 18.2, we obtain for each $w \in L_2(0, \tau; V)$ a unique solution $u =: Tw \in MR(0, \tau; V, H)$ of

$$u' + \mathcal{A}u = Fw, \quad u(0) = u_0\,.$$

We will show that $T\colon L_2(0, \tau; V) \to L_2(0, \tau; V)$ satisfies the assumptions of Schaefer's fixed point theorem.

Since F is continuous, Theorem 18.2 implies that T is continuous from $L_2(0,\tau;V)$ to $MR(0,\tau;V,H)$ and hence to $L_2(0,\tau;V)$. Combining (19.3) and (19.10) we estimate

$$\begin{aligned}\|Tw\|_{MR_a(0,t;V,H)} &\leqslant 5\mathrm{e}^{\gamma t}\big(\|Fw\|_{L_2(0,t;H)} + M_0^{1/2}\|u_0\|_V\big)\\ &\leqslant 5\mathrm{e}^{\gamma\tau}\big(c_F(\|w\|_{L_2(0,t;V)}+1) + M_0^{1/2}\|u_0\|_V\big)\end{aligned} \tag{19.12}$$

for all $w \in L_2(0,\tau;V)$, $0 < t \leqslant \tau$. It follows that T maps bounded sets in $L_2(0,\tau;V)$ to bounded sets in $MR_a(0,\tau;V,H)$. Since the embedding $MR_a(0,\tau;V,H) \hookrightarrow L_2(0,\tau;V)$ is compact, by Proposition 19.10, we conclude that $T\colon L_2(0,\tau;V) \to L_2(0,\tau;V)$ is compact.

We want to show that the Schaefer set

$$S = \{w \in L_2(0,\tau;V);\ \exists\lambda \in (0,1)\colon \lambda Tw = w\}$$

is bounded in $L_2(0,\tau;V)$. Let $w \in S$, so $w = \lambda Tw$ with $\lambda \in (0,1)$. Then (19.5) yields

$$\|w\|_{C([0,t];V)} = \lambda\|Tw\|_{C([0,t];V)} \leqslant c_1\big(\|Tw\|_{MR_a(0,t;V,H)} + \|Tw(0)\|_V\big)$$

for all $t \in (0,\tau]$. Since $Tw(0) = u_0$, inequality (19.12) implies that there exists a constant $c > 0$ (not depending on w) such that

$$\|w\|_{C([0,t];V)} \leqslant c\|w\|_{L_2(0,t;V)} + c$$

and hence

$$\|w(t)\|_V^2 \leqslant 2c^2\|w\|_{L_2(0,t;V)}^2 + 2c^2,$$

for all $t \in (0,\tau]$. Then Gronwall's inequality, proved below, shows that $\|w(t)\|_V^2 \leqslant 2c^2\mathrm{e}^{2c^2t}$ for all $t \in [0,\tau]$, and the boundedness of S in $L_2(0,\tau;V)$ follows.

Now we can apply Schaefer's fixed point theorem, Theorem 19.3, and find $u \in L_2(0,\tau;V)$ such that $Tu = u$. By the definition of T, any function u with this property is a solution as asserted in the theorem. □

In the proof above we have used Gronwall's inequality in its simplest form, as follows.

19.12 Lemma (Gronwall's inequality). *Let $f\colon [0,\tau] \to [0,\infty)$ be continuous, and assume that*

$$f(t) \leqslant \alpha + \beta\int_0^t f(s)\,\mathrm{d}s \qquad (0 < t < \tau),$$

with $\alpha, \beta \geqslant 0$. Then $f(t) \leqslant \alpha\mathrm{e}^{\beta t}$ for all $t \in [0,\tau]$.

Proof. Let $g(t) := \alpha + \beta\int_0^t f(s)\,\mathrm{d}s$ for all $t \in [0,\tau]$. Then $g(0) = \alpha$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{e}^{-\beta t}g(t)\big) = \mathrm{e}^{-\beta t}(g'(t) - \beta g(t)) = \beta\mathrm{e}^{-\beta t}(f(t) - g(t)) \leqslant 0$$

for all $t \in (0,\tau)$. This shows that $\mathrm{e}^{-\beta t}g(t) \leqslant \alpha$ for all $t \in [0,\tau]$, and the assertion follows. □

19.13 Remark. Recall that in the general context introduced in Section 18.1, the operator $\mathcal{A}$ consisted of two terms, the second coming from a non-autonomous form b that satisfies (18.2). Here we explain how such a term can be included in Theorem 19.11 as well.

Indeed, let b be as in Section 18.1. Then, defining $\mathcal{B}(t) \in \mathcal{L}(V, H)$ $(0 < t < \tau)$ as in Remark 18.3(a), one obtains a mapping $\mathcal{B} \in \mathcal{L}(L_2(0,\tau;V), L_2(0,\tau;H))$ satisfying (19.10) with $c_{\mathcal{B}} = M_b$. The equation

$$u'(t) + \big(\mathcal{A}(t) + \mathcal{B}(t)\big)u(t) = Fu(t),$$

where $\mathcal{A}$ and F are as in Theorem 19.11, can be rewritten as

$$u'(t) + \mathcal{A}(t)u(t) = Fu(t) - \mathcal{B}(t)u(t),$$

and the new nonlinear term, given by $Fw - \mathcal{B}w$ $(w \in L_2(0,\tau;V))$, satisfies (19.10) with constant $c_F + M_b$. △

19.5 Interlude: $L_2((0,\tau)\times\Omega) = L_2(0,\tau;L_2(\Omega))$

With the following proposition we prepare the application of Theorem 19.11 to situations in which $H = L_2(\Omega)$ and F is a Nemytskii type operator.

19.14 Proposition. *Let $\tau > 0$, and let $\Omega \subseteq \mathbb{R}^n$ be open. Then for each $w \in L_2((0,\tau)\times\Omega)$ there exists a unique $\widetilde{w} \in L_2(0,\tau;L_2(\Omega))$ such that $w(t,\cdot) = \widetilde{w}(t)$ in $L_2(\Omega)$ for a.e. $t \in (0,\tau)$. The mapping $w \mapsto \widetilde{w}$ thus defined is an isometric isomorphism.*

The property that $w(t,\cdot) = \widetilde{w}(t)$ for a.e. $t \in (0,\tau)$ should be read with the understanding that it holds for all representatives of the elements w and $\widetilde{w}$.

Proof of Proposition 19.14. We will work with the version of L_2-spaces in which all representatives of the elements are Borel measurable. We consider w as a representative of the given element in $L_2((0,\tau)\times\Omega)$; then by Fubini's theorem

$$\|w\|^2_{L_2((0,\tau)\times\Omega)} = \int_0^\tau \int_\Omega |w(t,x)|^2 \,\mathrm{d}x\,\mathrm{d}t.$$

Thus, the set $N := \big\{t \in (0,\tau);\ \int_\Omega |w(t,x)|^2 \,\mathrm{d}x = \infty\big\}$ is a Borel null set. Put

$$w_1(t,\cdot) := \begin{cases} w(t,\cdot) & \text{if } t \in (0,\tau)\setminus N,\\ 0 & \text{if } t \in N.\end{cases}$$

Then $w_1 = w$ a.e. on $(0,\tau)\times\Omega$, and $w_1(t,\cdot) \in L_2(\Omega)$ for all $t \in (0,\tau)$. The function $(0,\tau) \ni t \mapsto w_1(t,\cdot) \in L_2(\Omega)$ is measurable since for $g \in L_2(\Omega)$, the function $(0,\tau)\times\Omega \ni (t,x) \mapsto w_1(t,x)\bar{g}(x)$ is integrable (as product of the L_2-functions w_1 and $(t,x) \mapsto \bar{g}(x)$), and thus

$$t \mapsto \int_\Omega w_1(t,x)\bar{g}(x)\,\mathrm{d}x = (w_1(t,\cdot)\,|\,g)$$

is measurable. Hence, $\widetilde{w} := [t \mapsto w_1(t,\cdot)] \in L_2(0,\tau;L_2(\Omega))$, and $\|\widetilde{w}\|_{L_2(0,\tau;L_2(\Omega))} = \|w_1\|_{L_2((0,\tau)\times\Omega)}$. For the uniqueness of $\widetilde{w}$ it suffices to note that, if w_2 is another representative of w, then

$$\int_0^\tau \int_\Omega |w_2(t,x) - w(t,x)|^2 \,\mathrm{d}x \,\mathrm{d}t = 0,$$

hence $w_2(t,\cdot) = w(t,\cdot) = w_1(t,\cdot)$ for a.e. $t \in (0,\tau)$.

The above considerations show that

$$J\colon L_2((0,\tau)\times\Omega) \to L_2(0,\tau;L_2(\Omega)),\ w \mapsto \widetilde{w}$$

is an isometric linear mapping. It remains to show that J has dense range. For $\varphi \in L_2(0,\tau)$ and $g \in L_2(\Omega)$ one easily sees that $w := [(t,x) \mapsto \varphi(t)g(x)] \in L_2((0,\tau)\times\Omega)$ and $Jw = \varphi(\cdot)g$. Thus, $\operatorname{ran}(J)$ contains the set $L_2(0,\tau) \otimes L_2(\Omega)$, which by Lemma 17.4 is dense in $L_2(0,\tau;L_2(\Omega))$. □

19.6 Non-autonomous semilinear parabolic equations

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, let $H := L_2(\Omega)$, and let V be a closed subspace of $H^1(\Omega)$ containing $H_0^1(\Omega)$. We suppose that the embedding $V \hookrightarrow H$ is compact. This is the case if $V = H_0^1(\Omega)$, by Theorem 6.21, but it also holds under the previous more general hypothesis if Ω has continuous boundary, by Theorem 7.9. Let $\tau \in (0,\infty)$, and let the non-autonomous form a and the notation be as in Section 19.4. In particular, we ask the reader to think of a non-autononous form a as in Example 18.4; this will justify the heading of the present section.

Let $f\colon (0,\tau)\times\Omega\times\mathbb{K}\times\mathbb{K}^n \to \mathbb{K}$ be a Borel measurable function with the property that $f(t,x,\cdot,\cdot)\colon \mathbb{K}\times\mathbb{K}^n \to \mathbb{K}$ is continuous for all $t \in (0,\tau)$, $x \in \Omega$. Assume that there exist $L \geqslant 0$ and $0 \leqslant g \in L_2(0,\tau;L_2(\Omega))$ such that

$$|f(t,x,q,p)| \leqslant g(t)(x) + L(|q| + |p|) \qquad (t \in (0,\tau),\ x \in \Omega,\ q \in \mathbb{K},\ p \in \mathbb{K}^n). \tag{19.13}$$

Given these hypotheses, we can formulate the following consequence of Theorem 19.11; the proof will consist in showing that the right-hand side of (19.14) corresponds to a mapping F as in Theorem 19.11.

19.15 Theorem. *Let $u_0 \in V$. Then there exists $u \in MR(0,\tau;V,H)$ such that $u(0) = u_0$ and*

$$u'(t) + \mathcal{A}(t)u(t) = f(t,\cdot,u(t,\cdot),\nabla u(t,\cdot)) \tag{19.14}$$

for a.e. $t \in [0,\tau]$.

Note that, as in Theorem 19.11, there is no statement concerning uniqueness; the example of Exercise 19.5 applies also for the kind of equation treated here. The right-hand side of (19.14) is quite general; in particular it can include nonlinear first order terms. Before we prove Theorem 19.15, we present an application, which is a generalisation of Example 18.5, with nonlinear drift terms.

19.16 Example. Let $b\colon [0,\tau] \times \Omega \times \mathbb{K} \to \mathbb{K}^n$ be a bounded Borel measurable function, $b(t,x,\cdot)$ continuous for all $t \in [0,\tau]$, $x \in \Omega$. Then $b(t,\cdot,v(\cdot)) \in L_\infty(\Omega;\mathbb{K}^n)$ for all $v \in L_2(\Omega)$, $t \in [0,\tau]$.

Let $g \in L_2(0,\tau;L_2(\Omega))$ and $u_0 \in V$. Applying Theorem 19.15 with $f(t,x,q,p) := g(t)(x) - b(t,x,q)\cdot p$ we obtain: there exists $u \in MR(0,\tau;V,L_2(\Omega))$ such that $u(0) = u_0$ and

$$u'(t) + \mathcal{A}(t)u(t) + b(t,\cdot,u(t,\cdot))\cdot\nabla u(t,\cdot) = g(t) \tag{19.15}$$

for a.e. $t \in [0,\tau]$.

We mention that for this kind of nonlinearity the solution can already be obtained from Schauder's fixed point theorem (whereas in the proof of Theorem 19.11 we have applied Schaefer's fixed point theorem). The method is to map each function $w \in L_2(0,\tau;L_2(\Omega))$ to the solution $u =: Tw \in MR(0,\tau;V,L_2(\Omega))$ of

$$u'(t) + \mathcal{A}(t)u(t) + b(t,\cdot,w(t,\cdot))\cdot\nabla u(t,\cdot) = g(t), \quad u(0) = u_0\,,$$

which exists by Theorem 18.2. Then Exercise 18.1 can be used to show that $T\colon L_2(0,\tau;L_2(\Omega)) \to MR(0,\tau;V,L_2(\Omega))$ is continuous. (In fact, as $L_2(0,\tau;L_2(\Omega))$ is a metric space, it is sufficient to prove sequential continuity.) Moreover, it follows from the estimate (18.3) that the range of T is a bounded subset of $MR(0,\tau;V,L_2(\Omega))$; hence by the Aubin–Lions lemma, Theorem 19.6, the range of T is relatively compact in $L_2(0,\tau;H)$. Applying Schauder's fixed point theorem, Theorem 19.3, one obtains $u \in MR(0,\tau;V,L_2(\Omega)) \subseteq L_2(0,\tau;L_2(\Omega))$ satisfying $Tu = u$, i.e. u solves (19.15). In Exercise 19.2 the reader is asked to provide the details of these arguments. △

The next lemma is needed for the application of Theorem 19.11 in the proof of Theorem 19.15.

19.17 Lemma. *The mapping* $F\colon L_2(0,\tau;V) \to L_2(0,\tau;H)$ *given by*

$$Fw(t) := f(t,\cdot,w(t,\cdot),\nabla w(t,\cdot)) \qquad \big(w \in L_2(0,\tau;V),\ t \in (0,\tau)\big) \tag{19.16}$$

is defined and continuous, and satisfies (19.10) *with* $c_F := 2L + \|g\|_{L_2(0,\tau;H)}$.

Proof. Using Proposition 19.14 and the Borel measurability of f one easily shows that $Fw \in L_2(0,\tau;H)$ for all $w \in L_2(0,\tau;V)$. Moreover the assumed estimate (19.13) implies

$$\|Fw\|_{L_2(0,t;H)} \leqslant \|g\|_{L_2(0,\tau;H)} + 2L\|w\|_{L_2(0,t;V)} \qquad (t \in (0,\tau),\ w \in L_2(0,\tau;V)). \tag{19.17}$$

This proves the last assertion of the lemma.

Let $w_k \to w$ in $L_2(0,\tau;V)$. Then $w_k \to w$ and $\partial_j w_k \to \partial_j w$ $(j = 1,\dots,n)$ in $L_2(0,\tau;H)$. Using Proposition 19.14 again, we find a subsequence (w_{k_m}) and a function $h \in L_2((0,\tau)\times\Omega)$ such that $w_{k_m} \to w$, $\nabla w_{k_m} \to \nabla w$ a.e. on $(0,\tau)\times\Omega$ and $|w_{k_m}| \leqslant h$, $|\nabla w_{k_m}| \leqslant h$ a.e. for all $m \in \mathbb{N}$. Then $v_m(t,x) := f(t,x,w_{k_m}(t,x),\nabla w_{k_m}(t,x)) \to f(t,x,w(t,x),\nabla w(t,x))$ for a.e. $(t,x) \in (0,\tau)\times\Omega$, and $|v_m| \leqslant g + 2Lh$ a.e. Now the dominated convergence theorem together with Proposition 19.14 and the standard sub-subsequence argument yields the asserted continuity of F. □

Proof of Theorem 19.15. From Lemma 19.17 we know that the mapping F defined in (19.16) satisfies the assumption needed in Theorem 19.11, and with this choice of F, equation (19.14) becomes a particular version of (19.11).

Hence Theorem 19.15 follows immediately from Theorem 19.11. □

Finally we give an example of a mapping F exhibiting actions that are 'non-local' in time.

19.18 Example. Let $0 < t_0 < \tau$, and define

$$Fw(t) := (|w| \wedge 1)(t + t_0) + |\nabla w|(t - t_0) \qquad (0 < t < \tau,\ w \in L_2(0, \tau; V)), \tag{19.18}$$

where we suppose w to be extended to $\mathbb{R}$ by zero on $\mathbb{R} \setminus (0, \tau)$.

In Exercise 19.4 the reader is asked to show that F is continuous and satisfies (19.10). The first term in this example takes values of w 'from the future' – appropriately damped – whereas the second is a delay term. △

Notes

Mathematicians working in partial differential equations believe in the following principle: if there is an a priori estimate for solutions, then a solution should exist. Schaefer's fixed point theorem gives us at hand a precise framework in which this statement is correct. This fixed point theorem is part of the Leray–Schauder theory (which started by a paper of Leray and Schauder [LeSc34] and is presented in many textbooks, e.g., in [Sch69], [Ber77], [Dei85]). It was Schaefer [Sch55] who found the short proof we present here and which works even in the framework of locally convex spaces. See also [Eva10; Sect. 9.2.2, Theorem 4] for Schaefer's fixed point theorem.

The idea of using maximal regularity in order to gain compactness via the Aubin–Lions lemma, as we present it in Section 19.4, stems from [ArCh10]. (In that paper, the locally convex version of Schaefer's fixed point theorem allows the authors to treat unbounded domains.) What is new in our approach is that the linear part is chosen non-autonomous throughout. Moreover, we first treat the problem in a more abstract context with a rather general class of nonlinear terms, which includes delay terms. An important tool in our proof is the embedding result Theorem 19.7, which is based on the approximation obtained in Proposition 19.9. For related results in a more general context we refer to [ADLO14; Sections 3 and 4].

Exercises

19.1 Show that under the hypotheses of Theorem 19.3 the operator λT has a fixed point, for all $\lambda \in (0, 1)$; in particular, the Schaefer set of T is non-empty. (Hint: Look at the Schaefer set of λT.)

19.2 Carry out the following details of the procedure sketched in Example 19.16.

(a) Show that for all $w \in L_2(0,\tau; L_2(\Omega))$ an operator function $B_w\colon (0,\tau) \to \mathcal{L}(V, L_2(\Omega))$ is defined by

$$B_w v(t) := b(t,\cdot,w(t,\cdot)) \cdot \nabla v(\cdot) \qquad (v \in V,\ t \in [0,\tau]).$$

(b) Let (w_j) be a sequence in $L_2(0,\tau; L_2(\Omega))$, $w_j \to w$ in $L_2(0,\tau; L_2(\Omega))$. Show that there exists a subsequence (w_{j_k}) such that $B_{w_{j_k}}(t) \to B_w(t)$ strongly in $\mathcal{L}(V, L_2(\Omega))$ for a.e. $t \in [0,\tau]$, and that then $Tw_{j_k} \to Tw$ in $MR(0,\tau; V, L_2(\Omega))$. Apply the sub-subsequence argument to show that $Tw_j \to Tw$ in $MR(0,\tau; V, L_2(\Omega))$.

(c) Show that $T\big(L_2(0,\tau; L_2(\Omega))\big)$ is bounded in $MR(0,\tau; V, L_2(\Omega))$.

19.3 Let $\tau > 0$, and let H be a Hilbert space.

(a) Show that the unit ball of $H^1(0,\tau; H)$ is a uniformly equicontinuous subset of $C([0,\tau]; H)$. (Hint: Recall Proposition 17.6.)

(b) Additionally let V be a Hilbert space, with compact embedding $V \hookrightarrow H$. Show that the embedding $MR(0,\tau; V, H) \hookrightarrow C([0,\tau]; H)$ is compact. (Hint: Use the proof of Theorem 19.6 and part (a) above to show that every weakly convergent sequence in $MR(0,\tau; V, H)$ is convergent in $C([0,\tau]; H)$.)

19.4 Show that the mapping F defined in Example 19.18 is continuous and satisfies (19.10).

19.5 (a) Find infinitely many solutions of the initial value problem for the ordinary differential equation $y' = \sqrt{|y|}$, $y(0) = 0$ on $[0,1]$.

(b) As in Section 19.6 let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set and $H := L_2(\Omega)$. Let $V := H^1(\Omega)$, $\tau := 1$, and define the form a by $a(u,v) := \int \nabla u \cdot \nabla v \,\mathrm{d}x$ $(u, v \in V)$. Define $f\colon (0,1) \times \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ by

$$f(t,x,q,p) := \sqrt{|q|}\,.$$

Show that f satisfies (19.13), and find infinitely many solutions $u \in MR(0,1; V, H)$ of the initial value problem

$$u'(t) + \mathcal{A}u(t) = f(t,\cdot,u(t,\cdot),\nabla u(t,\cdot)), \quad u(0) = 0.$$

Appendices

Appendix A

The divergence theorem

In this appendix we provide a proof of Gauss' theorem, also known as the divergence theorem. In the first section we will sketch, without proofs, basic facts concerning embedded manifolds in $\mathbb{R}^n$ and integration on manifolds. On the basis of these facts we will give a complete proof of the divergence theorem in the second section.

A.1 Integration on submanifolds of $\mathbb{R}^n$

In this section we sketch the definition of the integral for functions on submanifolds of $\mathbb{R}^n$. We start with some facts concerning submanifolds, by which we always mean C^1-submanifolds.

Let $k, n \in \mathbb{N}_0$, $k < n$. A set $M \subseteq \mathbb{R}^n$ is a **k-dimensional submanifold** if for all $z \in M$ there exist an open neighbourhood U of z in $\mathbb{R}^n$, an open set $V \subseteq \mathbb{R}^n$ and a diffeomorphism $h\colon U \to V$ such that $h(M \cap U) = \big(\mathbb{R}^k \times \{0_{n-k}\}\big) \cap V$ (where 0_{n-k} denotes the zero element in $\mathbb{R}^{n-k}$). There are several other equivalent descriptions of submanifolds. The following is the one we will need below.

A set $M \subseteq \mathbb{R}^n$ is a k-dimensional submanifold if and only if M has local parametrisations, i.e., for all $z \in M$ there exist an open neighbourhood $W \subseteq M$, an open set $T \subseteq \mathbb{R}^k$ (the parameter domain) and a regular mapping $\Phi\colon T \to \mathbb{R}^n$ (the local parametrisation) such that $\Phi(T) = W$. Here, 'regular' means: Φ is continuously differentiable, $\Phi'(t)$ has full rank k for all $t \in T$, Φ is injective, and $\Phi^{-1}\colon W \to T$ is continuous.

In particular, it follows that the boundary $\partial\Omega$ of a bounded open set Ω with C^1-boundary is an $(n-1)$-dimensional manifold: looking at the definition of a C^1-graph W one easily finds a local parametrisation $\Phi\colon T \to \partial\Omega$ with $\Phi(T) = W$.

The definition of the integral on submanifolds is based on the local parametrisations and includes a weight factor that involves the quantity $\gamma(A) := \sqrt{\det(A^\top A)}$, for matrices $A \in \mathbb{R}^{n\times k}$. (Note that $A^\top A \in \mathbb{R}^{k\times k}$ is positive definite, hence $\det(A^\top A) \geqslant 0$. As a motivation for this weight factor we mention that the k-dimensional volume of $A([0,1]^k)$ – the image of the cube $[0,1]^k \subseteq \mathbb{R}^k$ under A – is given by $\gamma(A)$.)

A.1 Remarks (The surface measure σ on M). Let $M \subseteq \mathbb{R}^n$ be a k-dimensional submanifold.

(a) Let $\Phi\colon T \to M$ be a local parametrisation, $W := \Phi(T)$. Then $\Phi\colon T \to W$ is a homeomorphism, and we define the surface measure σ_W on the Borel sets $A \subseteq W$ by

$$\sigma_W(A) := \int_T \mathbf{1}_A(\Phi(t))\,\gamma(\Phi'(t))\,\mathrm{d}t \in [0,\infty].$$

Note that the integral exists because the function $\mathbf{1}_A \circ \Phi$ is Borel measurable by Exercise A.1(b). Moreover $\mathbf{1}_A \circ \Phi$ is the indicator function on T of the Borel set $\Phi^{-1}(A)$, and thus σ_W is the image on W of the weighted Lebesgue-Borel measure on T, with weight $\gamma(\Phi'(\cdot))$. This definition has the consequence that a Borel measurable function $f\colon W \to \mathbb{K}$ is integrable with respect to σ_W if and only if $f(\Phi(\cdot))\gamma(\Phi'(\cdot))$ is integrable on T, and then

$$\int_W f(z)\,\mathrm{d}\sigma_W(z) = \int_T f(\Phi(t))\gamma(\Phi'(t))\,\mathrm{d}t. \tag{A.1}$$

(b) Now let $\Phi_j\colon T_j \to M$ be local parametrisations, $W_j := \Phi_j(T_j)$ for $j = 1, 2$. It is a major step in the theory to show that the measures σ_{W_1} and σ_{W_2} are consistent, i.e. they coincide on the Borel subsets of $W_1 \cap W_2$. For the proof one establishes that $\int_{W_1} f(z)\,\mathrm{d}\sigma_{W_1}(z) = \int_{W_2} f(z)\,\mathrm{d}\sigma_{W_2}(z)$ for all functions $f \in C_\mathrm{c}(M)$ with $\operatorname{spt} f \subseteq W_1 \cap W_2$, using the theorem of local invertibility and the change of variable formula for k-dimensional integrals.

(c) In order to define the surface measure σ on the Borel subsets of M we assume for simplicity that M is compact. (This is sufficient for our purposes because our aim is to define σ on $\partial\Omega$ for bounded open sets Ω with C^1-boundary.)

The compactness of M implies that there exists a finite set of local parametrisations $\Phi_j\colon T_j \to M$, $j = 1, \dots, m$, such that $M = \bigcup_{j=1}^m W_j$, with $W_j := \Phi_j(T_j)$. By part (b) above, the Borel measures σ_{W_j} on the Borel subsets of W_j are mutually consistent. It follows that there exists a unique Borel measure σ on M such that the restriction of σ to the Borel subsets of W_j equals σ_{W_j} for all $j \in \{1, \dots, m\}$; see Exercise A.2.

As a consequence one obtains: if $f \in L_1(M, \sigma)$ is written as a sum $f = \sum_{j=1}^m f_j$, where $f_j \in L_1(M, \sigma)$ and $f_j = 0$ on $M \setminus W_j$ for all $j \in \{1, \dots, m\}$, then

$$\int_M f(z)\,\mathrm{d}\sigma(z) = \sum_{j=1}^m \int_{W_j} f_j(z)\,\mathrm{d}\sigma_{W_j}(z) = \sum_{j=1}^m \int_{T_j} f_j(\Phi_j(t))\gamma(\Phi_j'(t))\,\mathrm{d}t.$$

(d) An important property of the surface measure σ is ‘orthogonal invariance’. More precisely, if $B \in \mathbb{R}^{n\times n}$ is an orthogonal matrix, then

$$\int_{B(M)} f(z)\,\mathrm{d}\sigma(z) = \int_M f(Bz)\,\mathrm{d}\sigma(z) \tag{A.2}$$

for all $f \in L_1(B(M), \sigma)$; in particular, inserting $f = \mathbf{1}_{B(A)}$ for a Borel set $A \subseteq M$ one obtains $\sigma(B(A)) = \sigma(A)$. Here we suppose as in part (c) that M is compact; then one easily sees that $B(M)$ is a compact k-dimensional submanifold of $\mathbb{R}^n$ as well. We point out that on the left-hand side of (A.2) σ denotes the surface measure on $B(M)$, and on the right-hand side the surface measure on M. The basic observation for the proof of (A.2) is the identity

$$\gamma(BA) = \det\bigl((BA)^\top BA\bigr) = \det\bigl(A^\top B^\top BA\bigr) = \gamma(A) \qquad (A \in \mathbb{R}^{n\times k}).$$

If $\Phi\colon T \to W \subseteq M$ is a local parametrisation of M, then obviously $B \circ \Phi\colon T \to B(W) \subseteq B(M)$ is a local parametrisation of $B(M)$, and the above identity implies $\gamma\bigl((B\circ\Phi)'(t)\bigr) = \gamma(\Phi'(t))$. Then it is straightforward to derive (A.2). $\triangle$

In the proof of Gauss' theorem we will apply the formula (A.1) to the special case when $k = n-1$ and $W \subseteq M$ is a standard C^1-graph. With the notation introduced in Section 7.1, a parametrisation is given by $\Phi\colon W' \to W,\ y \mapsto \begin{pmatrix} y \\ g(y) \end{pmatrix}$, where $W' \subseteq \mathbb{R}^{n-1}$ is an open set and $g \in C^1(W')$. In order to evaluate (A.1) we have to determine $\gamma(\Phi'(y))$. Now $\Phi'(y) = \begin{pmatrix} E_{n-1} \\ g'(y) \end{pmatrix}$, hence

$$\Phi'(y)^\top \Phi'(y) = \begin{pmatrix} E_{n-1} & \nabla g(y) \end{pmatrix} \begin{pmatrix} E_{n-1} \\ g'(y) \end{pmatrix} = E_{n-1} + \nabla g(y) g'(y),$$

and the subsequent Lemma A.2 implies that $\gamma(\Phi'(y))^2 = 1 + |\nabla g(y)|^2$. Thus from (A.1) one obtains

$$\int_W f(z)\,\mathrm{d}\sigma(z) = \int_{W'} f(\Phi(y))\sqrt{1 + |\nabla g(y)|^2}\,\mathrm{d}y. \tag{A.3}$$

A.2 Lemma. *Let $k \in \mathbb{N}$, $v \in \mathbb{R}^k$, and let $E_k \in \mathbb{R}^{k\times k}$ denote the unit matrix. Then* $\det(E_k + vv^\top) = 1 + |v|^2$.

Proof. Without loss of generality $v \neq 0$. Then v is an eigenvector of the matrix $E_k + vv^\top$ with eigenvalue $1 + |v|^2$, whereas the $(k-1)$-dimensional orthogonal complement $v^\perp$ is the eigenspace of $E_k + vv^\top$ associated with the eigenvalue 1. This implies the assertion. □

A.2 Proof of the divergence theorem

Throughout this section let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. We will prove Gauss' theorem, Theorem 7.3, in the equivalent form of the divergence theorem. (Concerning the equivalence we observe that the choice $u(x) = v(x)e_j$ in (A.4) yields (7.1) for the scalar function v; conversely, applying (7.1) to the components of the vector field one obtains (A.4).)

A.3 Theorem (Divergence theorem). *Let $u \in C^1(\overline{\Omega}; \mathbb{K}^n)$. Then*

$$\int_\Omega \operatorname{div} u(x)\,\mathrm{d}x = \int_{\partial\Omega} u(z)\cdot\nu(z)\,\mathrm{d}\sigma(z), \tag{A.4}$$

where $\nu(z)$ is the outer unit normal at $z \in \partial\Omega$.

The following local part of the divergence theorem contains the main technical difficulty of the proof.

A.4 Proposition. *Let $W \subseteq \partial\Omega$ be a C^1-graph. Then there exists an open set $U \subseteq \mathbb{R}^n$ with the following properties: $W = U \cap \partial\Omega$, and if $u \in C^1(\overline{\Omega}; \mathbb{K}^n)$ is such that* $\operatorname{spt} u$ *(formed in $\overline{\Omega}$) is a compact subset of U, then*

$$\int_\Omega \operatorname{div} u(x)\,\mathrm{d}x = \int_{\partial\Omega} u(z)\cdot\nu(z)\,\mathrm{d}\sigma(z), \tag{A.5}$$

where $\nu(z)$ is the outer unit normal at $z \in \partial\Omega$.

Proof. (i) In this step we suppose that W is a standard C^1-graph. Let $W' \subseteq \mathbb{R}^{n-1}$, (a,b) and $g\colon W' \to (a,b)$ be as in Section 7.1. Put $U := W' \times (a,b)$, and let $u \in C^1(\overline{\Omega};\mathbb{K}^n)$, $\operatorname{spt} u \subseteq U$.

Recall from Remark 7.2 that the outer unit normal is given by

$$\nu(z) = \frac{1}{\sqrt{|\nabla g(y)|^2+1}} \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix}$$

for $z = (y, g(y)) \in W$, with $y \in W'$. Thus, applying (A.3) we can rewrite the right-hand side of (A.5) as

$$\begin{aligned} \int_{\partial\Omega} u(z)\cdot\nu(z)\,\mathrm{d}\sigma(z) &= \int_{W'} u(y,g(y))\cdot \frac{1}{\sqrt{|\nabla g(y)|^2+1}} \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix} \sqrt{1+|\nabla g(y)|^2}\,\mathrm{d}y \\ &= \int_{W'} u(y,g(y))\cdot \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix} \mathrm{d}y. \end{aligned}$$

In order to transform the left-hand side of (A.5) into the latter expression we extend u by $0 \in \mathbb{K}^n$ to a function on $\mathbb{R}^n$ and define $v\colon W' \times (-\infty, 0] \to \mathbb{K}^n$,

$$v(y,t) := u(y, t+g(y)).$$

From $u(y,t) = v(y, t-g(y))$ we then obtain

$$\operatorname{div} u(y,t) = \sum_{j=1}^{n-1} \partial_j v_j(y, t-g(y)) + \partial_n v(y, t-g(y)) \cdot \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix}.$$

Therefore

$$\begin{aligned} \int_\Omega \operatorname{div} u(x)\,\mathrm{d}x &= \int_{W'} \int_{-\infty}^{g(y)} \operatorname{div} u(y,t)\,\mathrm{d}t\,\mathrm{d}y \\ &= \int_{W'} \int_{-\infty}^{0} \operatorname{div} u(y,t+g(y))\,\mathrm{d}t\,\mathrm{d}y \\ &= \sum_{j=1}^{n-1} \int_{-\infty}^{0} \int_{W'} \partial_j v_j(y,t)\,\mathrm{d}y\,\mathrm{d}t + \int_{W'} \int_{-\infty}^{0} \partial_n v(y,t) \cdot \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix} \mathrm{d}t\,\mathrm{d}y. \end{aligned} \tag{A.6}$$

Since $\int_{W'} \partial_j \varphi(y)\,\mathrm{d}y = 0$ for all $\varphi \in C_\mathrm{c}^1(W')$ (see (4.1)), we conclude that the left-hand term in the last expression vanishes. In the right-hand term we apply the fundamental theorem of calculus; then we can continue (A.6) by

$$\int_\Omega \operatorname{div} u(x)\,\mathrm{d}x = \int_{W'} v(y,0)\cdot \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix} \mathrm{d}y = \int_{W'} u(y,g(y))\cdot \begin{pmatrix} -\nabla g(y) \\ 1 \end{pmatrix} \mathrm{d}y.$$

(ii) Now we treat the case of a general C^1-graph W. By definition, there exists an orthogonal matrix $B \in \mathbb{R}^{n\times n}$ such that $\widetilde{W} := B(W)$ is a standard C^1-graph with respect to $\widetilde{\Omega} := B(\Omega)$. By step (i) there exists an open set $\widetilde{U} \subseteq \mathbb{R}^n$ with the following properties: $\widetilde{W} = \widetilde{U} \cap \partial\widetilde{\Omega}$, and

$$\int_{\widetilde{\Omega}} \operatorname{div} \tilde{u}(x) \, \mathrm{d}x = \int_{\partial\widetilde{\Omega}} \tilde{u}(z) \cdot \tilde{\nu}(z) \, \mathrm{d}\sigma(z) \tag{A.7}$$

for all $\tilde{u} \in C^1\big(\overline{\widetilde{\Omega}}; \mathbb{K}^n\big)$ with $\operatorname{spt} u \subseteq \widetilde{U}$, where $\tilde{\nu}$ denotes the outer unit normal for the set $\widetilde{\Omega}$.

Put $U := B^{-1}(\widetilde{U})$ and note that $W = U \cap \partial\Omega$. Let $u \in C^1(\overline{\Omega}; \mathbb{K}^n)$ be such that $\operatorname{spt} u \subseteq U$, and put $\tilde{u} := B \circ u \circ B^{-1}$. Then $\tilde{u} \in C^1\big(\overline{\widetilde{\Omega}}; \mathbb{K}^n\big)$ and $\operatorname{spt} u \subseteq \widetilde{U}$, so $\tilde{u}$ satisfies (A.7). From Exercise A.3 and the invariance of the Lebesgue measure under orthogonal transformations we obtain

$$\int_{\widetilde{\Omega}} \operatorname{div} \tilde{u}(x) \, \mathrm{d}x = \int_{\widetilde{\Omega}} \operatorname{div} u(B^{-1}x) \, \mathrm{d}x = \int_{\Omega} \operatorname{div} u(x) \, \mathrm{d}x.$$

On the other hand $\tilde{\nu}(Bz) = B\nu(z)$ for all $z \in W$, and the invariance of the measure σ under orthogonal transformations (see Remark A.1(d)) implies that

$$\begin{aligned} \int_{\partial\widetilde{\Omega}} \tilde{u}(z) \cdot \tilde{\nu}(z) \, \mathrm{d}\sigma(z) &= \int_{\partial\Omega} \tilde{u}(Bz) \cdot \tilde{\nu}(Bz) \, \mathrm{d}\sigma(z) = \int_{\partial\Omega} Bu(z) \cdot B\nu(z) \, \mathrm{d}\sigma(z) \\ &= \int_{\partial\Omega} u(z) \cdot \nu(z) \, \mathrm{d}\sigma(z). \end{aligned}$$

Plugging the above equalities into (A.7) we obtain (A.5). □

Proof of Theorem A.3. Let $y \in \partial\Omega$. Since Ω has C^1-boundary we can find a C^1-graph $W_y \subseteq \partial\Omega$ containing y. We choose an open set $U_y \subseteq \mathbb{R}^n$ corresponding to W_y with the properties described in Proposition A.4.

For the open covering $\{U_y;\ y \in \partial\Omega\}$ of the compact set $\partial\Omega$ there exists a finite subcovering $\{U_y;\ y \in F\}$. Supplementing this covering by the set Ω, we obtain a finite open covering of $\overline{\Omega}$ and can choose a subordinate partition of unity on $\overline{\Omega}$, i.e., functions $0 \leqslant \varphi \in C_\mathrm{c}^\infty(\mathbb{R}^n)$ with $\operatorname{spt} \varphi \subseteq \Omega$, $0 \leqslant \varphi_y \in C_\mathrm{c}^\infty(\mathbb{R}^n)$ with $\operatorname{spt} \varphi_y \subseteq U_y$ for $y \in F$, and such that $\varphi + \sum_{y \in F} \varphi_y = 1$ on $\overline{\Omega}$; see Exercise 4.3(b).

Now let $u \in C^1(\overline{\Omega}; \mathbb{K}^n)$. Then $\int_\Omega \operatorname{div}(\varphi u)(x) \, \mathrm{d}x = 0$ by (4.1), and

$$\begin{aligned} \int_\Omega \operatorname{div} u(x) \, \mathrm{d}x &= \int_\Omega \operatorname{div}\Big(\big(\varphi + \sum_{y \in F} \varphi_y\big) u\Big)(x) \, \mathrm{d}x \\ &= \int_\Omega \operatorname{div}(\varphi u)(x) \, \mathrm{d}x + \sum_{y \in F} \int_\Omega \operatorname{div}(\varphi_y u)(x) \, \mathrm{d}x \\ &= \sum_{y \in F} \int_{\partial\Omega} \varphi_y(z) u(z) \cdot \nu(z) \, \mathrm{d}\sigma(z), \end{aligned}$$

where in the last step we have applied Proposition A.4. This establishes (A.4) since $\sum_{y \in F} \varphi_y = 1$ on $\partial\Omega$, and the proof of the divergence theorem is complete. □

A.5 Remark. In Appendix E, Section E.2, it is shown that in fact the surface measure σ coincides with the $(n-1)$-dimensional Hausdorff measure on $\partial\Omega$. We avoided using this property because we wanted to stick to the more elementary definition of the surface measure by parametrisations. △

Notes

The divergence theorem plays an important role for applications of the theory of forms. In this book we rely on the theorem whenever traces and weak normal derivatives are involved; see Chapters 7, 8, 11 and 12. More generally, it is central for the theory of partial differential equations.

We apply the divergence theorem only in the elementary form of Theorem A.3, for bounded open sets with C^1-boundary. It was hard to find an accessible source for this version in the English literature. So, according to our principle to provide complete information for topics going beyond a certain level, we decided to include a complete proof.

We mention that in German textbooks on analysis one can find proofs; see [For83; §15], [Kab99; Theorem 20.3]. The main inspiration for our proof was the proof presented in [Tre75; Lemma 10.1]. A different access can be found in [ArUr23; Chapter 7]. For a more general version of the divergence theorem in the context of geometric measure theory we refer to [EvGa92; Sect. 5.8, Theorem 1].

Exercises

A.1 Let (X, τ) and (Y, ρ) be topological spaces, $\Phi\colon X \to Y$ a homeomorphism. Let $\mathcal{B}_\tau$ and $\mathcal{B}_\rho$ denote the Borel σ-algebras of X and Y, respectively.

(a) Show that $\mathcal{B}_\rho = \{\Phi(A);\, A \in \mathcal{B}_\tau\}$.

(b) Let $f\colon Y \to \mathbb{K}$. Show that f is Borel measurable on Y if and only if $f \circ \Phi$ is Borel measurable on X.

A.2 Let Ω be a set, $\mathcal{A}$ a σ-algebra on Ω. Let $\Omega_1, \dots, \Omega_m \in \mathcal{A}$ be such that $\bigcup_{j=1}^m \Omega_j = \Omega$. For each $j \in \{1, \dots, m\}$ let μ_j be a measure on $\mathcal{A}_j := \{A \in \mathcal{A};\, A \subseteq \Omega_j\}$ with $\mu_j(\Omega_j) < \infty$. Suppose that the measures μ_j are mutually consistent, i.e. for all $j, k \in \{1, \dots, m\}$, $A \in \mathcal{A}_j \cap \mathcal{A}_k$ one has $\mu_j(A) = \mu_k(A)$.

Show that there exists a unique measure μ on $\mathcal{A}$ such that $\mu(A) = \mu_j(A)$ for all $j \in \{1, \dots, m\}$, $A \in \mathcal{A}_j$.

Hint: Choose a family $(B_j)_{j=1,\dots,m}$ in $\mathcal{A}$ such that $B_j \in \mathcal{A}_j$ for all $j \in \{1, \dots, m\}$, $\bigcup_{j=1}^m B_j = \Omega$ and $B_j \cap B_k = \varnothing$ if $j \neq k$. Define $\mu(A) := \sum_{j=1}^m \mu_j(A \cap B_j)$.

A.3 Let $\Omega \subseteq \mathbb{R}^n$ be open, $u \in C^1(\Omega; \mathbb{K}^n)$, and let $B \in \mathbb{R}^{n\times n}$ be an invertible matrix, $\widetilde{\Omega} := B(\Omega)$. Put $\tilde{u} := B \circ u \circ B^{-1} \in C^1(\widetilde{\Omega}; \mathbb{K}^n)$, spelled out

$$\tilde{u}(y) := Bu(B^{-1}y) \qquad (y \in \widetilde{\Omega}).$$

Show that $\operatorname{div} \tilde{u} = (\operatorname{div} u) \circ B^{-1}$. (This fact is used in the proof of Proposition A.4 for orthogonal matrices B.)

Hints: 1. For $u \in C^1(\Omega; \mathbb{K}^n)$ the divergence can be written as $\operatorname{div} u(x) = \operatorname{tr} u'(x)$, where the latter expression denotes the trace of the derivative (Jacobi matrix) of $u'(x)$. 2. For matrices $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times m}$ one has $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Appendix B

The Stone–Weierstrass theorem

The aim of this appendix is to prove the Stone–Weierstrass theorem, Theorem B.2. The proof is based on Stone's theorem, Theorem B.1, which is interesting and important in its own right. In a supplementary remark we mention that our treatment also provides a proof of the Weierstrass approximation theorem.

Throughout this appendix let K be a compact topological space. For real-valued functions $f, g \in C(K)$ we define $f \vee g, f \wedge g \in C(K)$ by

$$f \vee g(x) := \max\{f(x), g(x)\}, \quad f \wedge g(x) := \min\{f(x), g(x)\} \qquad (x \in K).$$

(With these operations, the space $C(K;\mathbb{R})$ becomes a vector lattice.) If $F \subseteq C(K;\mathbb{R})$ is a finite subset, then $\bigvee F \in C(K)$ is defined by $\bigl(\bigvee F\bigr)(x) := \max\{f(x);\, f \in F\}$ $(x \in K)$, and similarly for '$\bigwedge$'.

A **sublattice** of $C(K;\mathbb{R})$ is a subset L with the property that $f \vee g, f \wedge g \in L$ for all $f, g \in L$. A **vector sublattice** of $C(K;\mathbb{R})$ is a subspace that is also a sublattice.

If L is a subset of $C(K)$, then we say that L **separates the points** of K if for all $x \neq y$ in K there exists a function $f \in L$ such that $f(x) \neq f(y)$.

B.1 Theorem (Stone). *Let L be a vector sublattice of $C(K;\mathbb{R})$ such that*

(i) *L separates the points of K,*

(ii) *L contains the constant function* $\mathbf{1}$.

Then L is dense in $C(K;\mathbb{R})$.

Proof. First we note that, given points $x \neq y$ in K and scalars $\alpha, \beta \in \mathbb{R}$, there exists $f \in L$ such that $f(x) = \alpha$, $f(y) = \beta$. Indeed, there exists $g \in L$ such that $g(x) \neq g(y)$, and then f can be obtained as a linear combination of g and $\mathbf{1}$.

Let $f \in C(K;\mathbb{R})$, and let $\varepsilon > 0$. We need to show that there exists $h \in L$ with $\|f - h\|_\infty \leqslant \varepsilon$. First, fix $x \in K$. For all $y \in K$ there exists $g_y \in L$ such that $g_y(x) = f(x)$, $g_y(y) = f(y)$. Then $U_y := [g_y > f - \varepsilon]$ is an open neighbourhood of y, and from the open covering $\{U_y;\, y \in K\}$ of K we can choose a finite subcovering $\{U_y;\, y \in F\}$. Putting $h_x := \bigvee_{y \in F} g_y$ we have found a function $h_x \in L$ satisfying $h_x(x) = f(x)$, $h_x \geqslant f - \varepsilon$.

For all $x \in K$, the set $V_x := [h_x < f + \varepsilon]$ is an open neighbourhood of x. The open covering $\{V_x;\, x \in K\}$ of K contains a finite subcovering $\{V_x;\, x \in G\}$, and then the function $h := \bigwedge_{x \in G} h_x \in L$ satisfies $f - \varepsilon \leqslant h \leqslant f + \varepsilon$, i.e. $\|f - h\|_\infty \leqslant \varepsilon$. □

B.2 Theorem (Stone–Weierstrass). *Let L be a subalgebra of $C(K;\mathbb{K})$ such that*

(i) *L separates the points of K,*

(ii) *L contains the constant function* $\mathbf{1}$.

If $\mathbb{K} = \mathbb{C}$, we suppose additionally that
(iii) *for all $f \in L$ one has $\overline{f} \in L$.*
Then L is dense in $C(K; \mathbb{K})$.

Proof. Without loss of generality we can assume that L is closed. Indeed, it is easy to see that the closure of L is again a subalgebra of $C(K; \mathbb{K})$ satisfying the conditions (i), (ii) (and (iii)).

First we treat the case $\mathbb{K} = \mathbb{R}$. We show that L is a sublattice of $C(K; \mathbb{R})$; then Theorem B.1 implies the assertion. Let $f \in L$. By the Weierstrass approximation theorem there exists a sequence (p_n) of real polynomials such that $\sup_{|t| \leqslant \|f\|_\infty} \big| |t| - p_n(t) \big| \to 0$ as $n \to \infty$. Then $\big\| |f| - p_n \circ f \big\|_\infty \to 0$ as $n \to \infty$. Since $p_n \circ f \in L$ for all $n \in \mathbb{N}$, we conclude that $|f| \in L$. This has the consequence that for all $f, g \in L$ one has

$$f \vee g = \tfrac{1}{2}(f + g + |f - g|) \in L, \quad f \wedge g = \tfrac{1}{2}(f + g - |f - g|) \in L,$$

i.e. L is a sublattice of $C(K; \mathbb{R})$.

In the complex case one immediately deduces that $L_{\mathrm{r}} := L \cap C(K; \mathbb{R})$ is a subalgebra of $C(K; \mathbb{R})$ satisfying conditions (i) and (ii); hence the case treated above implies that L_{r} is dense in $C(K; \mathbb{R})$. Thus $L = L_{\mathrm{r}} + \mathrm{i} L_{\mathrm{r}}$ is dense in $C(K; \mathbb{C})$. □

B.3 Remarks. (a) In the proof of Theorem B.2 we quoted the Weierstrass approximation theorem, but in fact we did not need the full extent of the theorem since we only had to approximate the absolute value function on bounded intervals. Here we show how this can be done.

From basic analysis we recall that the power series

$$(1 + s)^{1/2} = \sum_{n=0}^{\infty} \binom{1/2}{n} s^n$$

is uniformly absolutely convergent on $[-1, 1]$. Therefore the expression

$$|t| = \big(1 + (t^2 - 1)\big)^{1/2} = \sum_{n=0}^{\infty} \binom{1/2}{n} (t^2 - 1)^n$$

yields a series of polynomials converging to $|t|$, uniformly for $-1 \leqslant t \leqslant 1$. Clearly this implies that $|t|$ can be approximated uniformly on arbitrary bounded intervals.

(b) The Weierstrass approximation theorem can be obtained as a special case of Theorem B.2 if in the proof one uses the argument from part (a) to approximate the absolute value function. △

Notes

The foundations for the use of order and lattice properties in the approximation of continuous functions as in Theorems B.1 and B.2 go back to Stone [Sto37]; see in particular the discussion on the Weierstrass approximation theorem in [Sto37; p. 467]. For Theorem B.1 we also refer to Kakutani [Kak41; Theorem 4], and [Sto48] contains further discussion of the topic.

Appendix C

Weyl's lemma

Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $u \in L_{1,\mathrm{loc}}(\Omega)$ satisfy $\Delta u = 0$ in the distributional sense. The issue of this appendix is to show that then u is a.e. equal to a harmonic function $\tilde{u}$, i.e. $\tilde{u} \in C^2(\Omega)$, $\Delta \tilde{u} = 0$.

C.1 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $u \in C^2(\Omega)$, $\Delta u = 0$. Then u has the **mean value property**: if $x \in \Omega$, $r > 0$ are such that $B[x,r] \subseteq \Omega$, then*

$$u(x) = \frac{1}{r^{n-1}\sigma_{n-1}} \int_{\partial B(x,r)} u(y)\,\mathrm{d}\sigma(y) = \frac{1}{\sigma_{n-1}} \int_{S_{n-1}} u(x + r\xi)\,\mathrm{d}\sigma(\xi), \tag{C.1}$$

where σ_{n-1} is the $(n-1)$-dimensional volume of the unit sphere $S_{n-1} \subseteq \mathbb{R}^n$.

Proof. For $x \in \mathbb{R}^n$, $r > 0$ such that $B[x,r] \subseteq \Omega$ we compute

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}r} \int_{S_{n-1}} u(x + r\xi)\,\mathrm{d}\sigma(\xi) &= \int_{S_{n-1}} \nabla u(x + r\xi) \cdot \xi\,\mathrm{d}\sigma(\xi) \\
&= \frac{1}{r^{n-1}} \int_{\partial B(x,r)} \nabla u(y) \cdot \nu(y)\,\mathrm{d}\sigma(y) \\
&= \frac{1}{r^{n-1}} \int_{B(x,r)} \Delta u(y)\,\mathrm{d}y = 0,
\end{aligned} \tag{C.2}$$

where in the last step we have applied the divergence theorem, Theorem A.3. It follows that the expression on the right-hand side of (C.1) does not depend on r. This expression tends to $u(x)$ as $r \to 0$, which implies the assertion. □

C.2 Theorem (Weyl's lemma). *Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $u \in L_{1,\mathrm{loc}}(\Omega)$, $\Delta u = 0$ (in the distributional sense). Then there exists an infinitely differentiable harmonic representative $\tilde{u}$ of u.*

Proof. Below we show that for every $x \in \Omega$ there exist an open neighbourhood $U_x \subseteq \Omega$ and a representative $u_x \in C^\infty(U_x)$ of $u|_{U_x}$. Applying Exercise C.1 one then obtains an infinitely differentiable representative $\tilde{u}$ of u, and Remark 4.7 implies that $\tilde{u}$ is harmonic.

Let $x \in \Omega$, and let $r > 0$ be such that $B[x, 3r] \subseteq \Omega$. We will establish the property asserted above for $U_x := B(x,r)$. Let v denote the extension of $u\mathbf{1}_{B(x,3r)}$ by zero to a function in $L_1(\mathbb{R}^n)$, and let $(\rho_k)_{k\in\mathbb{N}}$ be a delta sequence in $C_\mathrm{c}^\infty(\mathbb{R}^n)$. Then $v_k := \rho_k * v \to v$

in $L_1(\mathbb{R}^n)$ by Proposition 4.3(b). Moreover, Lemma 4.1 shows that $v_k \in C_c^\infty(\mathbb{R}^n)$ and, for $z \in B(x, 2r)$, $k > \frac{1}{r}$, that

$$\Delta v_k(z) = \int_\Omega \Delta\rho_k(z-y)u(y)\,\mathrm{d}y = \int_\Omega \rho_k(z-y)\Delta u(y)\,\mathrm{d}y = 0,$$

where the middle equality holds because $\rho_k(z-\cdot) \in C_c^\infty(\Omega)$. Hence Proposition C.1 implies that $v_k|_{B(x,2r)}$ satisfies the mean value property, for all $k > \frac{1}{r}$.

Let $\rho \in C_c^\infty(\mathbb{R}^n)_+$ with $\operatorname{spt}\rho \subseteq B(0,r)$, $\int \rho(x)\,\mathrm{d}x = 1$, be a radial function, i.e. $\rho(y) = \rho(|y|e_1)$ for all $y \in B(0,r)$, where e_1 is the first unit vector in $\mathbb{R}^n$. Then, for $z \in B(x,r)$, one uses the mean value property of $v_k|_{B(z,r)}$ to see that

$$\begin{aligned}\rho * v_k(z) &= \int_{B(0,r)} \rho(y)v_k(z-y)\,\mathrm{d}y = \int_{s=0}^{r} \rho(se_1)\int_{\partial B(0,s)} v_k(z-y)\,\mathrm{d}\sigma(y)\,\mathrm{d}s \\ &= \Big(\sigma_{n-1}\int_0^r \rho(se_1)s^{n-1}\,\mathrm{d}s\Big)v_k(z) = \Big(\int_{B(0,r)} \rho(y)\,\mathrm{d}y\Big)v_k(z) = v_k(z).\end{aligned}$$

Now the convergence $v_k \to v$ in $L_1(\mathbb{R}^n)$ implies that $v_k = \rho * v_k \to \rho * v$ uniformly on $B(x,r)$ as $k \to \infty$; hence the C^∞-function $(\rho * v)|_{B(x,r)}$ is a representative of $v|_{B(x,r)} = u|_{B(x,r)}$. □

Notes

Theorem C.2 goes back to Hermann Weyl [Wey40; Lemma 2]. In modern language of partial differential operators one could state it by saying that the Laplace operator is a *hypoelliptic* differential operator. We refer to [Rud91; Corollary of Theorem 8.12], where Weyl's lemma is obtained as a consequence of an important result concerning the regularity of distributional solutions of general elliptic equations. In contrast, the proof we provide is rather elementary, being based only on the mean value property of harmonic functions and properties of convolutions. The mean value property of harmonic functions is a standard fact in potential theory and can be found in the standard literature on partial differential equations; see, for instance, [Eva10; Section 2.2.2].

We mention that Weyl's lemma also holds for the case when u is a distribution instead of a locally integrable function; with vitually the same proof as given for Theorem C.2.

Exercises

C.1 Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $u \in L_{1,\mathrm{loc}}(\Omega)$. Suppose that for all $x \in \Omega$ there exist an open neighbourhood $U_x \subseteq \Omega$ and a function $u_x \in C(U_x)$ such that $u_x = u$ a.e. on U_x. Show that there exists a function $\tilde{u} \in C(\Omega)$ such that $\tilde{u}|_{U_x} = u_x$ for all $x \in \Omega$. Apply Exercise 4.1 to show that $\tilde{u}$ is a representative of u.

Conclude that $\tilde{u}$ is infinitely differentiable if every u_x is infinitely differentiable.

C.2 Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $u \in C(\Omega)$ be a function having the mean value property (C.1). Show that u is an infinitely differentiable harmonic function. (Hint: For the property that $u \in C^\infty(\Omega)$ look at the proof of Theorem C.2; for 'harmonic' use (C.2) and argue by contradiction.)

C.3 Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set.

(a) Let $u \in C(\overline{\Omega}) \cap C^2(\Omega)$ be harmonic on Ω. Prove the **maximum principle**

$$\max_{x\in\overline{\Omega}} |u(x)| = \max_{x\in\partial\Omega} |u(x)|.$$

(Hint: Use the mean value property (C.1).)

(b) Show that

$$C_\Delta(\overline{\Omega}) := \bigl\{u \in C(\overline{\Omega});\ u|_\Omega \text{ harmonic}\bigr\}$$

is a closed subspace of $C(\overline{\Omega})$, and conclude that $\bigl\{u|_{\partial\Omega};\ u \in C_\Delta(\overline{\Omega})\bigr\}$ is a closed subspace of $C(\partial\Omega)$. (Hint concerning the last statement: Part (a) shows that the trace mapping $T\colon C_\Delta(\overline{\Omega}) \to C(\partial\Omega),\ u \mapsto u|_{\partial\Omega}$ is a linear isometry.)

Note. With the trace mapping T from above, the image $T(C_\Delta(\overline{\Omega}))$ is the set of 'admissible boundary values for the classical Dirichlet problem'. For $g \in T(C_\Delta(\overline{\Omega}))$ the Dirichlet problem $\Delta u = 0$, $u|_{\partial\Omega} = g$ has the unique solution $u = T^{-1}g$, and the mapping T^{-1} is an isometry, in particular continuous.

Appendix D

Hausdorff measure and an inequality due to Maz'ya

D.1 The Sobolev space $W_1^1(\Omega)$

We will need the Sobolev space $W_1^1(\Omega)$ and some of its properties. Let $\Omega \subseteq \mathbb{R}^n$ be an open set. We define

$$W_1^1(\Omega) := \bigl\{u \in L_1(\Omega);\, \partial_j u \in L_1(\Omega)\ (j = 1, \dots, n)\bigr\},$$

with norm

$$\|u\|_{W_1^1} := \|u\|_1 + \sum_{j=1}^n \|\partial_j u\|_1 .$$

Then the space $W_1^1(\Omega)$ is a separable Banach space. This is proved in the same way as the corresponding property for $H^1(\Omega)$; see Theorem 4.10 and its proof.

Analogously to H_c^1 and H_0^1 we define

$$\begin{aligned} W_{1,c}^1(\Omega) &:= \bigl\{u \in W_1^1(\Omega);\, \operatorname{spt} u \text{ compact in } \Omega\bigr\},\\ W_{1,0}^1(\Omega) &:= \overline{W_{1,c}^1(\Omega)}^{W_1^1(\Omega)}. \end{aligned}$$

The property $W_{1,0}^1(\Omega) = \overline{C_c^\infty(\Omega)}^{W_1^1(\Omega)}$ is proved in the same way as the corresponding property for H^1; see Theorem 4.15 and its proof.

We will need some lattice properties of W_1^1-functions. These are consequences of the properties treated in Section 9.3.

D.1 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $u, v \in L_{1,\mathrm{loc}}(\Omega; \mathbb{R})$, $\nabla u, \nabla v \in L_{1,\mathrm{loc}}(\Omega; \mathbb{R}^n)$. Then $\nabla(u \wedge v) \in L_{1,\mathrm{loc}}(\Omega; \mathbb{R}^n)$, $|\nabla(u \wedge v)| \leqslant |\nabla u| + |\nabla v|$.*

Proof. Let $j \in \{1, \dots, n\}$. From Corollary 9.15 we recall $\partial_j u^+ = \mathbf{1}_{[u>0]} \partial_j u$. This implies

$$\partial_j u^- = \partial_j(u^+ - u) = (\mathbf{1}_{[u>0]} - 1)\partial_j u = -\mathbf{1}_{[u\leqslant 0]}\partial_j u;$$

hence

$$\partial_j |u| = \partial_j u^+ - \partial_j u^- = \bigl(\mathbf{1}_{[u>0]} - \mathbf{1}_{[u\leqslant 0]}\bigr)\partial_j u.$$

Applying this equality to $u \wedge v = \frac{1}{2}(u + v - |u - v|)$ we obtain

$$\partial_j(u \wedge v) = \frac{1}{2}\big(\partial_j u + \partial_j v - (\mathbf{1}_{[u>v]} - \mathbf{1}_{[u\leqslant v]})(\partial_j u - \partial_j v)\big) = \mathbf{1}_{[u\leqslant v]}\partial_j u + \mathbf{1}_{[u>v]}\partial_j v.$$

We conclude that

$$\nabla(u \wedge v) = \mathbf{1}_{[u\leqslant v]}\nabla u + \mathbf{1}_{[u>v]}\nabla v,$$

which implies the assertion. □

Lemma D.1 will be applied in the proof of Theorem D.9 in the following more specific situation.

D.2 Lemma. *Let $\Omega, \Omega_1 \subseteq \mathbb{R}^n$ be open sets, Ω_1 bounded, let $u \in C(\overline{\Omega}) \cap W_1^1(\Omega)$, $v \in C(\overline{\Omega_1}) \cap W_1^1(\Omega_1)$, $u, v \geqslant 0$, and assume that $u(x) < v(x)$ for all $x \in \partial\Omega_1 \cap \overline{\Omega}$. Put $v := \infty$ on $\mathbb{R}^n \setminus \overline{\Omega_1}$, and define $w := u \wedge v$ on $\overline{\Omega}$. Then $w \in C(\overline{\Omega}) \cap W_1^1(\Omega)$,*

$$\int_\Omega |\nabla w(x)|\,\mathrm{d}x \leqslant \int_\Omega |\nabla u(x)|\,\mathrm{d}x + \int_{\Omega_1} |\nabla v(x)|\,\mathrm{d}x. \tag{D.1}$$

Proof. The continuity of w follows from the continuity of w on $\overline{\Omega} \cap \overline{\Omega_1}$ and the continuity of $w = u$ on $\overline{\Omega} \setminus \Omega_1$. (Note that $\{\overline{\Omega} \cap \overline{\Omega_1}, \overline{\Omega} \setminus \Omega_1\}$ is a finite covering of $\overline{\Omega}$ by closed sets.)

Lemma D.1 shows that $w|_{\Omega\cap\Omega_1} \in W_1^1(\Omega \cap \Omega_1)$,

$$\int_{\Omega\cap\Omega_1} |\nabla w(x)|\,\mathrm{d}x \leqslant \int_{\Omega\cap\Omega_1} |\nabla u(x)|\,\mathrm{d}x + \int_{\Omega_1} |\nabla v(x)|\,\mathrm{d}x. \tag{D.2}$$

The hypotheses imply that there exists an open set $U \subseteq \Omega$ with $\partial\Omega_1 \cap \Omega \subseteq U$ such that $u(x) < v(x)$ for all $x \in U \cap \overline{\Omega_1}$. Then $w = u$ on $(\Omega \setminus \Omega_1) \cup U$, and therefore $w \in W_1^1(\Omega)$,

$$\int_\Omega |\nabla w(x)|\,\mathrm{d}x = \int_{\Omega\setminus\Omega_1} |\nabla u(x)|\,\mathrm{d}x + \int_{\Omega\cap\Omega_1} |\nabla w(x)|\,\mathrm{d}x. \tag{D.3}$$

Combining (D.2) and (D.3) one obtains (D.1). □

D.3 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, $u \in W_1^1(\Omega) \cap C_0(\Omega)$. Then $u \in W_{1,0}^1(\Omega)$.*

Proof. It is sufficient to treat the case when u is real-valued. Splitting $u = u^+ - u^-$ we reduce the problem to the case $u \geqslant 0$.

For $s > 0$ one has $u - s \in W_1^1(\Omega)$, and Corollary 9.15 implies that $\nabla(u - s)^+ = \mathbf{1}_{[u>s]}\nabla(u - s) = \mathbf{1}_{[u>s]}\nabla u$. Note that $u \leqslant s$ on a neighbourhood of $\partial\Omega$ because $u \in C_0(\Omega)$, and therefore $\operatorname{spt}(u - s)^+$ is compact, which shows that $(u - s)^+ \in W_{1,\mathrm{c}}^1(\Omega)$. For $s \to 0$ one has $(u - s)^+ \to u$, $\nabla(u - s)^+ \to \mathbf{1}_{[u>0]}\nabla u$ pointwise and dominated, hence in $L_1(\Omega)$, $L_1(\Omega; \mathbb{R}^n)$, respectively. Because $\mathbf{1}_{[u>0]}\nabla u = \nabla u^+ = \nabla u$, these two convergences mean that $(u - s)^+ \to u$ in $W_1^1(\Omega)$, hence $u \in W_{1,0}^1(\Omega)$. □

D.4 Remark. Proposition D.3 remains valid with W_1^1 replaced by H^1. This version is remindful of the inclusion '$\supseteq$' in Proposition 7.10, but note the differences in the regularity of $\partial\Omega$ and in the assumptions on u at the boundary. △

So far we have treated lattice properties concerning differentiation only for real-valued functions. In order to include complex-valued functions in our treatment, we need the following formula concerning the differentiation of absolute values.

D.5 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $u \in L_{1,\mathrm{loc}}(\Omega)$, $j \in \{1, \dots, n\}$, $\partial_j u \in L_{1,\mathrm{loc}}(\Omega)$. Then*

$$\partial_j |u| = \mathrm{Re}(\overline{\operatorname{sgn} u}\, \partial_j u) \in L_{1,\mathrm{loc}}(\Omega).$$

Proof. Given $\delta > 0$, the equality

$$\partial_j (|u|^2 + \delta)^{1/2} = \partial_j (\bar{u} u + \delta)^{1/2} = (|u|^2 + \delta)^{-1/2}\, \mathrm{Re}(\bar{u}\, \partial_j u)$$

is straightforward if u is continuously differentiable. Using the same arguments as in the proof of Proposition 9.13 one can easily show that this equality remains valid under the present hypotheses.

Letting $\delta \to 0$ and applying Lemma 4.11 we conclude that

$$\partial_j |u| = \mathrm{Re}\Big(\frac{\bar{u}}{|u|} \mathbf{1}_{[u \neq 0]}\, \partial_j u \Big) = \mathrm{Re}(\overline{\operatorname{sgn} u}\, \partial_j u). \qquad \square$$

The last issue in this section is Poincaré's inequality for the L_1-context. Here we say that a set $\Omega \subseteq \mathbb{R}^n$ is **contained in a slab** if there exist $\beta > 0$ such that after a suitable translation and rotation applied to Ω, one has $\Omega \subseteq (-\beta, \beta) \times \mathbb{R}^{n-1}$; in Section 5.4 we used a more restricted version of this notion. We note that Theorem 5.13 is also valid with the present more general version, and that the proof given below is analogous to the proof of Theorem 5.13.

D.6 Theorem (Poincaré's inequality). *Assume that Ω is contained in a slab. Then*

$$\int_\Omega |u(x)|\, \mathrm{d}x \leqslant c_{\mathrm{P}} \int_\Omega |\nabla u(x)|\, \mathrm{d}x \qquad (u \in W^1_{1,0}(\Omega)),$$

where c_{P} is the width of the slab.

Proof. It was mentioned above that $W^1_{1,0}(\Omega)$ is the closure of $C_{\mathrm{c}}^\infty(\Omega)$ in $W^1_1(\Omega)$; therefore it suffices to prove the inequality for all $u \in C_{\mathrm{c}}^\infty(\Omega)$. Under translations and orthogonal transformations of the independent variable, the gradient of a function is transformed along with the function. (This is obvious for translations. Concerning orthogonal transformations for a function $u \in C^1(\Omega)$ and a matrix $B \in \mathbb{R}^{n \times n}$, the chain rule implies that $\nabla(u \circ B) = \big((u \circ B)'\big)^\top = \big((u' \circ B)B\big)^\top = B^\top(\nabla u \circ B)$.) As a consequence we may assume that $\Omega \subseteq (-\beta, \beta) \times \mathbb{R}^{n-1}$, where $2\beta > 0$ is the width of the slab. Let $h \in C^1[-\beta, \beta]$, $h(-\beta) = 0$. Then we estimate

$$\int_{-\beta}^{\beta} |h(x)|\, \mathrm{d}x = \int_{-\beta}^{\beta} \Big| \int_{-\beta}^{x} h'(y)\, \mathrm{d}y \Big|\, \mathrm{d}x \leqslant \int_{-\beta}^{\beta} \int_{-\beta}^{\beta} |h'(y)|\, \mathrm{d}y\, \mathrm{d}x = 2\beta \int_{-\beta}^{\beta} |h'(y)|\, \mathrm{d}y.$$

Let $u \in C_{\mathrm{c}}^\infty(\Omega)$. Applying the above estimate with $h(r) = u(r, x_2, \dots, x_n)$ we obtain

$$\int_\Omega |u(x)|\, \mathrm{d}x \leqslant 2\beta \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{-\beta}^{\beta} |\partial_1 u(x_1, \dots, x_n)|\, \mathrm{d}x_1 \cdots \mathrm{d}x_n \leqslant 2\beta \int_\Omega |\nabla u(x)|\, \mathrm{d}x. \qquad \square$$

D.2 On the Hausdorff measure $\mathcal{H}_d$

In this section we present an auxiliary result concerning the approximation of the Hausdorff measure of dimension $d \in [0, \infty)$ on a metric space M. We start by introducing some notation and giving a short introduction to Hausdorff measures.

For a set $C \subseteq M$ we define $\operatorname{rd}(C) := \frac{1}{2} \operatorname{diam}(C)$, and for a countable collection $\mathcal{C}$ of subsets of M we define $\operatorname{rd}(\mathcal{C}) := \sup_{C \in \mathcal{C}} \operatorname{rd}(C)$ and

$$S_d(\mathcal{C}) := \omega_d \sum_{C \in \mathcal{C}} \operatorname{rd}(C)^d,$$

where

$$\omega_d := \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} + 1)},$$

which is the volume of the unit ball in $\mathbb{R}^d$ if $d \in \mathbb{N}_0$. (Even though the notation 'rd' should be remindful of 'radius', the reader should be aware that a set C will not necessarily be contained in a ball with radius $\operatorname{rd}(C)$.)

Let $B \subseteq M$. For $\delta > 0$ we put

$$\mathcal{H}_{d,\delta}(B) := \inf\bigl\{S_d(\mathcal{C})\,;\, \mathcal{C} \text{ countable covering of } B,\ \operatorname{rd}(\mathcal{C}) \leqslant \delta\bigr\}.$$

Then

$$\mathcal{H}_d^*(B) := \lim_{\delta \to 0} \mathcal{H}_{d,\delta}(B) = \sup_{\delta > 0} \mathcal{H}_{d,\delta}(B)$$

is the **outer d-dimensional Hausdorff measure** of B; note that the limit exists and is equal to the supremum because $\mathcal{H}_{d,\delta}(B)$ is decreasing in δ. Carathéodory's construction of measurable sets yields a measure $\mathcal{H}_d$, the **d-dimensional Hausdorff measure**, and it turns out that Borel sets are measurable. For all of these properties (and more) we refer to [EvGa92; Chapter 2] and [Fed69; Section 2.10.2]. If $d \in \mathbb{N}$, and $E = \mathbb{R}^d \times \{0\} \subseteq \mathbb{R}^n$, then $\mathcal{H}_d$ is the Lebesgue measure on $\mathbb{R}^d \cong E$; see Appendix E. We mention that for $d = 0$ the Hausdorff measure $\mathcal{H}_d$ is the counting measure. This follows from the fact that $\omega_0 = 1$ and $\operatorname{rd}(C)^0 = 1$ for all $C \subseteq M$.

Observe that, in the definition of $\mathcal{H}_{d,\delta}(B)$, one can also take the infimum over all countable *open* coverings of B and still obtain the same resulting value for $\mathcal{H}_d^*(B)$. Indeed, for $\varepsilon > 0$ there exists a countable covering $\mathcal{C}$ with $\operatorname{rd}(\mathcal{C}) \leqslant \delta/2$ and $S_d(\mathcal{C}) \leqslant \mathcal{H}_d^*(B) + \varepsilon/2$. Choose $(\varepsilon_C)_{C \in \mathcal{C}} \in (0, \infty)^{\mathcal{C}}$ such that $\sum_{C \in \mathcal{C}} \varepsilon_C \leqslant \varepsilon/2$. Then for all $C \in \mathcal{C}$ there exists an open set $U_C \supseteq C$ such that

$$\operatorname{rd}(U_C) \leqslant \delta \quad \text{and} \quad \omega_d \operatorname{rd}(U_C)^d \leqslant \omega_d \operatorname{rd}(C)^d + \varepsilon_C\,.$$

Then $\mathcal{U} := \{U_C\,;\, C \in \mathcal{C}\}$ is a countable open covering of B with $\operatorname{rd}(\mathcal{U}) \leqslant \delta$ and $S_d(\mathcal{U}) \leqslant \mathcal{H}_d^*(B) + \varepsilon$.

D.7 Proposition. *Let M be a metric space, $d \in [0, \infty)$, and assume that $\mathcal{H}_d(M) < \infty$. Let $\varepsilon > 0$.*

Then for all $\delta > 0$ there exists a countable partition $\mathcal{A}$ of M with $\operatorname{rd}(\mathcal{A}) \leqslant \delta$*, consisting of Borel subsets of M and such that*

$$\sum_{C \in \mathcal{A}} \left| \mathcal{H}_d(C) - \omega_d \operatorname{rd}(C)^d \right| \leqslant \varepsilon. \tag{D.4}$$

If M is compact, then the partition $\mathcal{A}$ can be chosen finite.

Proof. (i) The definition of $\mathcal{H}_d$ implies that there exists $\delta_\varepsilon > 0$ such that for all countable coverings $\mathcal{C}$ of M with $\operatorname{rd}(\mathcal{C}) \leqslant \delta_\varepsilon$ one has

$$\mathcal{H}_d(M) - \varepsilon \leqslant S_d(\mathcal{C}). \tag{D.5}$$

(ii) Next we show that for all $\delta > 0$ there exists a countable partition $\mathcal{A}$ of M with $\operatorname{rd}(\mathcal{A}) \leqslant \delta$, consisting of Borel subsets of M and such that

$$S_d(\mathcal{A}) \leqslant \mathcal{H}_d(M) + \varepsilon;$$

if M is compact, then the partition $\mathcal{A}$ can be chosen finite.

As pointed out above, there exists a countable open covering $\mathcal{U}$ of M with $\operatorname{rd}(\mathcal{U}) \leqslant \delta$ and $S_d(\mathcal{U}) \leqslant \mathcal{H}_d(M) + \varepsilon$. If M is compact, then there exists a finite subcovering of $\mathcal{U}$. A standard procedure to produce a pairwise disjoint covering yields the desired partition $\mathcal{A}$.

(iii) Let $0 < \delta \leqslant \delta_\varepsilon$ (from step (i)), and let $\mathcal{A}$ be a partition of M as in step (ii). Let $\mathcal{A}_1 \subseteq \mathcal{A}$; then $\mathcal{A}_1$ is a partition of $M_1 := \bigcup \mathcal{A}_1$. We show that

$$\mathcal{H}_d(M_1) - \varepsilon \leqslant S_d(\mathcal{A}_1) \leqslant \mathcal{H}_d(M_1) + 2\varepsilon. \tag{D.6}$$

Let $\mathcal{C}_2$ be a countable covering of $M_2 := M \setminus M_1$ with $\operatorname{rd}(\mathcal{C}_2) \leqslant \delta_\varepsilon$. Then the covering $\mathcal{A}_1 \cup \mathcal{C}_2$ of M satisfies $\operatorname{rd}(\mathcal{A}_1 \cup \mathcal{C}_2) \leqslant \delta_\varepsilon$, so from (D.5) we obtain

$$\mathcal{H}_d(M_1) + \mathcal{H}_d(M_2) - \varepsilon = \mathcal{H}_d(M) - \varepsilon \leqslant S_d(\mathcal{A}_1 \cup \mathcal{C}_2) \leqslant S_d(\mathcal{A}_1) + S_d(\mathcal{C}_2).$$

Since one can approximate $\mathcal{H}_d(M_2)$ arbitrarily well by $S_d(\mathcal{C}_2)$, choosing $\mathcal{C}_2$ suitably, this inequality implies

$$\mathcal{H}_d(M_1) - \varepsilon \leqslant S_d(\mathcal{A}_1),$$

the left-hand inequality of (D.6).

The application of the result obtained so far to the partition $\mathcal{A}_2 := \mathcal{A} \setminus \mathcal{A}_1$ of M_2 yields $\mathcal{H}_d(M_2) - \varepsilon \leqslant S_d(\mathcal{A}_2)$. Putting this inequality together with the inequality stated in step (ii) we obtain

$$S_d(\mathcal{A}_1) = S_d(\mathcal{A}) - S_d(\mathcal{A}_2) \leqslant \mathcal{H}_d(M) + \varepsilon - (\mathcal{H}_d(M_2) - \varepsilon) = \mathcal{H}_d(M_1) + 2\varepsilon,$$

the right-hand inequality of (D.6).

We now choose $\mathcal{A}_1 := \big\{ C \in \mathcal{A};\, \omega_d \operatorname{rd}(C)^d \leqslant \mathcal{H}_d(C) \big\}$ and apply (D.6) with $\mathcal{A}_1$, M_1 and with $\mathcal{A}_2$, M_2 (as defined above):

$$\begin{aligned} &\sum_{C \in \mathcal{A}} \left| \mathcal{H}_d(C) - \omega_d \operatorname{rd}(C)^d \right| \\ &\qquad = \sum_{C \in \mathcal{A}_1} \big(\mathcal{H}_d(C) - \omega_d \operatorname{rd}(C)^d \big) + \sum_{C \in \mathcal{A}_2} \big(\omega_d \operatorname{rd}(C)^d - \mathcal{H}_d(C) \big) \\ &\qquad = \big(\mathcal{H}_d(M_1) - S_d(\mathcal{A}_1) \big) + \big(S_d(\mathcal{A}_2) - \mathcal{H}_d(M_2) \big) \leqslant 3\varepsilon. \end{aligned}$$

□

D.8 Remark. The crucial point of the inequality in Proposition D.7 is that not only is the sum $S_d(\mathcal{A})$ close to $\mathcal{H}_d(M)$, but the individual terms $\omega_d \operatorname{rd}(C)^d$ of the sum also approximate the corresponding terms $\mathcal{H}_d(C)$, with a small total error. △

D.3 Maz'ya's inequality

The following theorem contains the central result of this appendix. The inequality we prove is not quite Maz'ya's inequality (12.13), but it is strong enough to allow the derivation of the inequality (12.11) in Section 12.6.

D.9 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with $\mathcal{H}_{n-1}(\partial\Omega) < \infty$. Then there exists a constant $c(n,\Omega)$ (depending only on n and the width of Ω) such that*

$$\|u\|_{L_1(\Omega)} \leqslant c(n,\Omega)\Big(\int_\Omega |\nabla u(x)|\,\mathrm{d}x + \int_{\partial\Omega} |u|\,\mathrm{d}\mathcal{H}_{n-1}\Big) \tag{D.7}$$

for all $u \in C(\overline{\Omega}) \cap W_1^1(\Omega)$.

Proof. Let $u \in C(\overline{\Omega}) \cap W_1^1(\Omega)$. It follows from Proposition D.5 that $|u| \in C(\overline{\Omega}) \cap W_1^1(\Omega)$ and that $\big|\nabla|u|\big| \leqslant |\nabla u|$. This shows that it is sufficient to treat the case $u \geqslant 0$.

Let $u \in C(\overline{\Omega}) \cap W_1^1(\Omega)$, $u \geqslant 0$, and let $\varepsilon > 0$. Then, by the (uniform) continuity of u, there exists $\delta \in (0,\varepsilon]$ such that $|u(x) - u(y)| < \varepsilon$ whenever $x, y \in \overline{\Omega}$, $|x - y| < \delta$. By Proposition D.7 there exists a finite partition $\mathcal{A}$ of $\partial\Omega$ with $\operatorname{rd}(\mathcal{A}) \leqslant \delta/4$, consisting of Borel sets and such that (D.4) holds. We choose a family $(x_C)_{C\in\mathcal{A}}$ with $x_C \in C$ for all $C \in \mathcal{A}$. (Recall that, by definition, the sets in a partition are supposed to be non-empty.)

Clearly $\{B_{\mathbb{R}^n}[x_C, \operatorname{diam}(C)];\ C \in \mathcal{A}\}$ is a covering of $\partial\Omega$, where we use the notation $B[x,r]$ for the closed ball with centre x and radius r. For each $C \in \mathcal{A}$, $s \in (0, \delta/2)$ we define a function

$$\psi_{C,s}(x) := \begin{cases} (u(x_C) + \varepsilon)\,\frac{1}{s}\operatorname{dist}\big(x, B(x_C, \operatorname{diam}(C))\big) & \text{if } x \in B[x_C, \operatorname{diam}(C) + s], \\ \infty & \text{otherwise.} \end{cases}$$

Note that $\psi_{C,s} \in W_1^1(B(x_C, \operatorname{diam}(C) + s))$. Note also that $\operatorname{diam}(C) + s \leqslant \delta$, and hence $\psi_{C,s}(x) = u(x_C) + \varepsilon > u(x)$ for all $x \in \partial B(x_C, \operatorname{diam}(C) + s) \cap \overline{\Omega}$, $C \in \mathcal{A}$. These properties and Lemma D.2 – applied repeatedly – imply that

$$u_{\varepsilon,s} := u \wedge \inf_{C\in\mathcal{A}} \psi_{C,s} \quad \text{on } \overline{\Omega}$$

belongs to $C(\overline{\Omega}) \cap W_1^1(\Omega)$. Since $\delta \leqslant \varepsilon$, the function $u_{\varepsilon,s}$ coincides with u on $\Omega_\varepsilon := \{x \in \Omega;\ B(x,\varepsilon) \subseteq \Omega\}$, and as $u_{\varepsilon,s}$ vanishes on $\partial\Omega$, Proposition D.3 shows that $u_{\varepsilon,s} \in W_{1,0}^1(\Omega)$.

The following computations prepare the application of Poincaré's inequality to $u_{\varepsilon,s}$. Lemma D.2 implies

$$\int_\Omega |\nabla u_{\varepsilon,s}(x)|\,\mathrm{d}x \leqslant \int_\Omega |\nabla u(x)|\,\mathrm{d}x + \sum_{C\in\mathcal{A}} \int_{B(x_C, \operatorname{diam}(C)+s)} |\nabla\psi_{C,s}(x)|\,\mathrm{d}x. \tag{D.8}$$

Observing that $\psi_{C,s} = 0$ on $B(x_C, \operatorname{diam}(C))$ and $|\nabla\psi_{C,s}| = (u(x_C)+\varepsilon)/s$ on the spherical shell $B(x_C, \operatorname{diam}(C)+s) \setminus B(x_C, \operatorname{diam}(C))$, we obtain

$$\int_{B(x_C,\operatorname{diam}(C)+s)} |\nabla\psi_{C,s}(x)|\,\mathrm{d}x = (u(x_C)+\varepsilon)\frac{1}{s}\omega_n\big((\operatorname{diam}(C)+s)^n - \operatorname{diam}(C)^n\big). \quad \text{(D.9)}$$

We note that, for $s \to 0$, the latter expression tends to

$$(u(x_C)+\varepsilon)n\omega_n \operatorname{diam}(C)^{n-1} = 2^{n-1}\frac{n\omega_n}{\omega_{n-1}}(u(x_C)+\varepsilon)\omega_{n-1}\operatorname{rd}(C)^{n-1}.$$

Recalling that $u|_{\Omega_\varepsilon} = u_{\varepsilon,s}|_{\Omega_\varepsilon}$, we conclude from Theorem D.6 that

$$\|u\|_{L_1(\Omega_\varepsilon)} \leqslant \|u_{\varepsilon,s}\|_{L_1(\Omega)} \leqslant c_{\mathrm{P}}\int_\Omega |\nabla u_{\varepsilon,s}(x)|\,\mathrm{d}x. \quad \text{(D.10)}$$

Inserting (D.8) and (D.9) into (D.10) and taking the limit $s \to 0$ we obtain

$$\|u\|_{L_1(\Omega_\varepsilon)} \leqslant c_{\mathrm{P}}\Big(\int_\Omega |\nabla u(x)|\,\mathrm{d}x + 2^{n-1}\frac{n\omega_n}{\omega_{n-1}}\sum_{C\in\mathcal{A}}(u(x_C)+\varepsilon)\omega_{n-1}\operatorname{rd}(C)^{n-1}\Big). \quad \text{(D.11)}$$

Exploiting (D.4) we can estimate the sum on the right-hand side of (D.11) by

$$\begin{aligned}
&\sum_{C\in\mathcal{A}}(u(x_C)+\varepsilon)\big(\big|\omega_{n-1}\operatorname{rd}(C)^{n-1} - \mathcal{H}_{n-1}(C)\big| + \mathcal{H}_{n-1}(C)\big)\\
&\qquad\leqslant (\|u\|_\infty+\varepsilon)\varepsilon + \sum_{C\in\mathcal{A}}\int_C (u(x_C)+\varepsilon)\,\mathrm{d}\mathcal{H}_{n-1}\\
&\qquad\leqslant (\|u\|_\infty+\varepsilon)\varepsilon + \int_{\partial\Omega}(u+2\varepsilon)\,\mathrm{d}\mathcal{H}_{n-1}\,.
\end{aligned}$$

Because of this inequality, the estimate (D.11) implies

$$\begin{aligned}
\|u\|_{L_1(\Omega_\varepsilon)} \leqslant c_{\mathrm{P}}\bigg(&\int_\Omega |\nabla u(x)|\,\mathrm{d}x\\
&+ 2^{n-1}\frac{n\omega_n}{\omega_{n-1}}\Big(\int_{\partial\Omega} u\,\mathrm{d}\mathcal{H}_{n-1} + \varepsilon(2\mathcal{H}_{n-1}(\partial\Omega) + \|u\|_\infty + \varepsilon)\Big)\bigg).
\end{aligned}$$

Letting $\varepsilon \to 0$ we finally obtain

$$\|u\|_{L_1(\Omega)} \leqslant c_{\mathrm{P}}\Big(\int_\Omega |\nabla u(x)|\,\mathrm{d}x + 2^{n-1}\frac{n\omega_n}{\omega_{n-1}}\int_{\partial\Omega}|u|\,\mathrm{d}\mathcal{H}_{n-1}\Big). \quad \text{(D.12)}$$

□

From Theorem D.9 we now derive the inequality (12.11) presented in Section 12.6.

D.10 Corollary. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with $\mathcal{H}_{n-1}(\partial\Omega) < \infty$. Then*

$$\|u\|^2_{L_2(\Omega)} \leqslant 4c(n,\Omega)^2\int_\Omega |\nabla u(x)|^2\,\mathrm{d}x + 2c(n,\Omega)\int_{\partial\Omega}|u|^2\,\mathrm{d}\mathcal{H}_{n-1}$$

for all $u \in C(\overline{\Omega}) \cap H^1(\Omega)$, with $c(n,\Omega)$ from Theorem D.9.

Proof. Let $u \in H^1(\Omega)$, $j \in \{1, \ldots, n\}$. Then $\partial_j|u| = \mathrm{Re}(\overline{\operatorname{sgn} u}\, \partial_j u) \in L_2(\Omega)$ by Proposition D.5. Using Corollary 9.15 we conclude that $\partial_j(|u| \wedge k) = \mathbf{1}_{[|u|<k]}\, \partial_j|u| \in L_2(\Omega)$ for all $k \in \mathbb{N}$, and $\partial_j(|u| \wedge k) \to \partial_j|u|$ in $L_2(\Omega)$ as $k \to \infty$. Then $(|u| \wedge k)^2 \to |u|^2$ in $L_1(\Omega)$, and applying Corollary 9.13 we obtain $\partial_j(|u| \wedge k)^2 = 2(|u| \wedge k)\, \partial_j(|u| \wedge k) \to 2|u|\, \partial_j|u|$ in $L_1(\Omega)$ as $k \to \infty$. This shows that $|u|^2 \in W_1^1(\Omega)$, $\nabla|u|^2 = 2|u|\, \nabla|u|$.

Assume additionally that $u \in C(\overline{\Omega})$. We apply Theorem D.9 to $|u|$ and obtain

$$\|u\|_2^2 \leqslant c(n, \Omega)\Big(2\int_\Omega |u(x)|\big|\nabla|u|(x)\big|\,\mathrm{d}x + \int_{\partial\Omega} |u|^2\,\mathrm{d}\mathcal{H}_{n-1}\Big). \tag{D.13}$$

Applying the Cauchy–Schwarz inequality, Proposition D.5 and the Peter–Paul inequality we get

$$\int_\Omega |u(x)|\big|\nabla|u|(x)\big|\,\mathrm{d}x \leqslant \|u\|_2\Big(\int_\Omega |\nabla u(x)|^2\,\mathrm{d}x\Big)^{1/2} \leqslant \gamma\|u\|_2^2 + \frac{1}{4\gamma}\int_\Omega |\nabla u(x)|^2\,\mathrm{d}x$$

for all $\gamma > 0$. Inserting this inequality into (D.13), with $\gamma := \frac{1}{4c(n,\Omega)}$, and reshuffling terms we finally get

$$\|u\|_2^2 \leqslant 2c(n, \Omega)\Big(2c(n, \Omega)\int_\Omega |\nabla u(x)|^2\,\mathrm{d}x + \int_{\partial\Omega} |u|^2\,\mathrm{d}\mathcal{H}_{n-1}\Big). \qquad \square$$

D.11 Remarks. (a) In the last step of the proof of Theorem D.9, in the estimate (D.10), we have applied the Poincaré inequality. There is another important inequality, estimating the L_q-norm, for $q = \frac{n}{n-1}$, in terms of the L_1-norm of the gradient, namely the **Sobolev inequality**

$$\|u\|_{L_q(\mathbb{R}^n)} \leqslant c(n)\int_{\mathbb{R}^n} |\nabla u(x)|\,\mathrm{d}x, \tag{D.14}$$

valid for all $u \in W_1^1(\mathbb{R}^n)$, with $c(n)$ only dependent on the dimension n. We refer to [Bre11; Theorem 9.9] or [Ada75; Remark 5.11] for this inequality. Applying (D.14) instead of the Poincaré inequality in the proof of Theorem D.9, one obtains as a result the 'better' inequality (12.13) (with a different value of $c(n)$). We put quotes around 'better' because clearly one can derive (D.7) from an inequality of the type (12.13), but then the constant in (D.7) will necessarily depend on the volume of Ω. In our derivation, however, that constant depends only on the dimension and on the width of Ω.

(b) Here we comment on the close relationship between (12.13), (D.14) and the **isoperimetric inequality**

$$\mathrm{vol}_n(A)^{(n-1)/n} \leqslant c(n)\mathcal{H}_{n-1}(\partial A) \tag{D.15}$$

for any measurable set $A \subseteq \mathbb{R}^n$ of finite n-dimensional volume, where

$$c(n) := \frac{1}{n\omega_n^{1/n}} = \frac{\Gamma(n/2+1)^{1/n}}{n\sqrt{\pi}}$$

is the **isoperimetric constant**. For a proof of (D.15) we refer to [Fed69; Theorem 3.2.43 and Remark 3.2.44] or [Maz11; Theorem 9.1.5 and Remark 9.2.2]. In fact, the isoperimetric constant $c(n)$ is also the optimal constant in (12.13) and (D.14).

We point out that (12.13) implies (D.15) for bounded open sets $A = \Omega$ (put $u = \mathbf{1}_\Omega$). Moreover, (12.13) implies (D.14) for functions $u \in W^1_{1,\mathrm{c}}(\mathbb{R}^n)$. On the other hand, assuming that one knows (D.15) for 'nice' sets, one can prove (D.14) for C_c^∞-functions, using the coarea formula of integration theory. Finally, one can extend (D.14) to L_1-functions 'of bounded variation', i.e. with distributional derivatives that are finite measures, and from this inequality one can derive (12.13). (This last step is rather non-trivial, and in a way contains the technical refinements one always encounters if one wants to prove (D.15) or (12.13) for the general case.)

In view of the above comments one says that in some sense the inequalities (D.15), (12.13) and (D.14) are equivalent. △

Notes

Inequalities like (12.13), due to Maz'ya, or the general isoperimetric inequality are notorious for their challenging technical level. The inequalities become easier if one does not insist on the optimal constant. For the isoperimetric inequality, at one point one has to estimate the 'perimeter' of an open set Ω by $\mathcal{H}_{n-1}(\partial\Omega)$; see for instance [AFP00; Theorem 3.62]. The proof of that theorem inspired us to approximate integrals with respect to a Hausdorff measure by integrals of suitable 'simple functions' on partitions arising in connection with the definition of the Hausdorff measure (see Proposition D.7) and to express the integral over the resulting simple functions by integrals over the gradient of suitable auxiliary functions (see the proof of Theorem D.9).

We refer to [VoVo20b] for this treatment of Maz'ya's inequality as well as for further comments on the relation between Mazya's inequality, the isoperimetric inequality and the Sobolev inequality.

Appendix E

The n- and $(n-1)$-dimensional Hausdorff measures on $\mathbb{R}^n$

E.1 $\mathcal{H}_n$ on $\mathbb{R}^n$

In this section we show that the n-dimensional Hausdorff measure on $\mathbb{R}^n$ is the Lebesgue measure λ^n. For the definition of the Hausdorff measure and the notation used below we refer to Section D.3.

E.1 Theorem. *For all $n \in \mathbb{N}_0$ one has $\mathcal{H}_n = \lambda^n$.*

The assertion includes the statement that the measurable subsets coincide. This will be established by showing that the outer measures $\mathcal{H}_n^*$ and $\lambda^{n,*}$ (on arbitrary subsets of $\mathbb{R}^n$) are identical.

We recall that the outer n-dimensional Lebesgue measure is defined – for instance – by

$$\lambda^{n,*}(A) := \inf\Big\{\sum_{R\in\mathcal{R}} \lambda^n(R)\,;\ \mathcal{R} \text{ countable collection of closed } n\text{-dimensional rectangles, } A \subseteq \bigcup \mathcal{R}\Big\},$$

for $A \subseteq \mathbb{R}^n$; see [Str94; Section 2.1]. The properties we will need below are that $\lambda^{n,*}$ is countably subadditive and that

$$\lambda^{n,*}(A) = \inf\{\lambda^n(U)\,;\ U \subseteq \mathbb{R}^n \text{ open, } A \subseteq U\} \qquad (A \subseteq \mathbb{R}^n);$$

see [Str94; Lemmas 2.1.2 and 2.1.4].

The most important tool for the aim formulated above is the **isodiametric inequality**

$$\lambda^{n,*}(A) \leqslant \omega_n \operatorname{rd}(A)^n, \tag{E.1}$$

valid for all bounded subsets $A \subseteq \mathbb{R}$, where $\operatorname{rd}(A) := \frac{1}{2}\operatorname{diam}(A)$. This inequality will be proved in Section E.3.

We will also need the following nice property concerning the Lebesgue measure.

E.2 Lemma. *Let $U \subseteq \mathbb{R}^n$ be an open set, and let $\delta > 0$. Then there exists a pairwise disjoint countable collection $\mathcal{B}$ of open balls $B \subseteq U$ with* $\operatorname{rd}(\mathcal{B}) \leqslant \delta$ *and* $\lambda^n(U \setminus \bigcup \mathcal{B}) = 0$.

Proof. We take it for granted from measure theory that there exists a pairwise disjoint countable collection $\mathcal{Q}$ of open cubes $Q \subseteq U$ of side lengths $\leqslant 2\delta$ and such that $\lambda^n(U \setminus \bigcup \mathcal{Q}) = 0$.

In view of this property it clearly suffices to prove the assertion for the case when U is an open cube U_0 with side length $s \leqslant 2\delta$. Put $\mathcal{B}_0 := \{B[c, s/2]\}$, where c is the centre of the cube U_0; then $U_1 := U_0 \setminus \bigcup \mathcal{B}_0 = U_0 \setminus B[c, s/2]$ is an open set, with

$$q := \lambda^n(U_1)/\lambda^n(U_0) = 1 - 2^{-n}\omega_n < 1.$$

The open set U_1 can be covered, up to a null set, by a pairwise disjoint countable collection $\mathcal{U}_1$ of open cubes $Q \subseteq U_1$. Applying the previous procedure to each of these cubes one obtains a countable collection $\mathcal{B}_1$ of closed balls and a new open set $U_2 = U_1 \setminus \bigcup \mathcal{B}_1$ with $\lambda^n(U_2)/\lambda^n(U_1) = q$. Repeating the procedure we obtain a decreasing sequence (U_k) of open sets with

$$\lambda^n(U_k) \leqslant q\lambda^n(U_{k-1}) \leqslant \cdots \leqslant q^k\lambda^n(U_0) \to 0 \qquad (k \to \infty),$$

and then

$$\lambda^n\Bigl(U_0 \setminus \bigcup_{k\in\mathbb{N}_0} \bigcup \mathcal{B}_k\Bigr) = 0,$$

where $\mathcal{B}_k$ is the countable collection of balls corresponding to U_k. Replacing each of the (closed) balls in $\bigcup_{k\in\mathbb{N}} \mathcal{B}_k$ by its interior we obtain the desired countable collection $\mathcal{B}$ of open balls with the property $\lambda^n(U_0 \setminus \bigcup \mathcal{B}) = 0$. □

Proof of Theorem E.1. Let $A \subseteq \mathbb{R}^n$.

(i) First we show that $\mathcal{H}_n^*(A) \leqslant \lambda^{n,*}(A)$. There is nothing to prove if $\lambda^{n,*}(A) = \infty$; so assume that $\lambda^{n,*}(A) < \infty$. Let $\delta > 0$, $\varepsilon > 0$. There exists an open set $U \supseteq A$ with $\lambda^n(U) \leqslant \lambda^{n,*}(A) + \varepsilon$. By Lemma E.2 there exists a pairwise disjoint countable collection $\mathcal{B}$ of open balls contained in U, with $\mathrm{rd}(\mathcal{B}) \leqslant \delta$ and

$$\sum_{B\in\mathcal{B}} \lambda^n(B) = \lambda^n(U) \leqslant \lambda^{n,*}(A) + \varepsilon.$$

As $U \setminus \bigcup \mathcal{B}$ is a Lebesgue null set, there exists a countable covering $\mathcal{Q}$ of $U \setminus \bigcup \mathcal{B}$ with $\mathrm{rd}(\mathcal{Q}) \leqslant \delta$, consisting of cubes, and such that $\sum_{Q\in\mathcal{Q}} \lambda^n(Q) \leqslant \frac{1}{\omega_n}\bigl(\frac{2}{\sqrt{n}}\bigr)^n\varepsilon$. Then $\mathcal{B} \cup \mathcal{Q}$ is a countable covering of A, $\mathrm{rd}(\mathcal{B} \cup \mathcal{Q}) \leqslant \delta$, and

$$\begin{aligned} S_n(\mathcal{B} \cup \mathcal{Q}) &= \omega_n \sum_{B\in\mathcal{B}} \mathrm{rd}(B)^n + \omega_n \sum_{Q\in\mathcal{Q}} \mathrm{rd}(Q)^n \\ &= \lambda^n(U) + \omega_n\Bigl(\frac{\sqrt{n}}{2}\Bigr)^n \sum_{Q\in\mathcal{Q}} \lambda^n(Q) \leqslant \lambda^{n,*}(A) + 2\varepsilon, \end{aligned}$$

which shows that $\mathcal{H}_{n,\delta}(A) \leqslant \lambda^{n,*}(A) + 2\varepsilon$. Letting first $\varepsilon \to 0$ and then $\delta \to 0$ we obtain the assertion of the first step.

(ii) We now prove that $\lambda^{n,*}(A) \leqslant \mathcal{H}_n^*(A)$. Let $\mathcal{C}$ be a countable covering of A with $\mathrm{rd}(\mathcal{C}) \leqslant 1$. Then the isodiametric inequality (E.1) shows that

$$\lambda^{n,*}(A) \leqslant \sum_{C\in\mathcal{C}} \lambda^{n,*}(C) \leqslant \omega_n \sum_{C\in\mathcal{C}} \mathrm{rd}(C)^n,$$

by the countable subadditivity of $\lambda^{n,*}$. Taking the infimum on the right-hand side we obtain

$$\lambda^{n,*}(A) \leqslant \mathcal{H}_{n,1}(A) \leqslant \mathcal{H}_n^*(A). \qquad \square$$

E.2 $\mathcal{H}_{n-1}$ on $\mathbb{R}^n$

In this section we show that that the $(n-1)$-dimensional Hausdorff measure on $\partial\Omega$ is the surface measure σ described in Section 7.1.

E.3 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1-boundary. Then $\sigma(A) = \mathcal{H}_{n-1}(A)$ for all Borel sets $A \subseteq \partial\Omega$.*

The following lemma will be needed in the proof of Theorem E.3.

E.4 Lemma. *Let (M_1, ρ_1) and (M_2, ρ_2) be metric spaces, $\varphi\colon M_1 \to M_2$ Lipschitz continuous, with Lipschitz constant $L \geqslant 0$, and let $d \in [0, \infty)$. Then*

$$\mathcal{H}_d^*(\varphi(A)) \leqslant L^d \mathcal{H}_d^*(A) \tag{E.2}$$

for all $A \subseteq M_1$.

Proof. Let $\delta > 0$, and let $\mathcal{C}$ a covering of $A \subseteq M_1$ with $\mathrm{rd}(\mathcal{C}) \leqslant \delta$. Then

$$\varphi(\mathcal{C}) := \{\varphi(C)\,;\ C \in \mathcal{C}\}$$

is a covering of $\varphi(A)$, with $\mathrm{rd}(\varphi(\mathcal{C})) \leqslant L\delta$. We estimate

$$\mathcal{H}_{d,L\delta}(\varphi(A)) \leqslant S_d(\varphi(\mathcal{C})) = \omega_d \sum_{C\in\mathcal{C}} \mathrm{rd}(\varphi(C))^d \leqslant \omega_d L^d \sum_{C\in\mathcal{C}} \mathrm{rd}(C)^d = L^d S_d(\mathcal{C}).$$

Taking the infimum on the right-hand side we obtain $\mathcal{H}_{d,L\delta}(\varphi(A)) \leqslant L^d \mathcal{H}_{d,\delta}(A)$, which for $\delta \to 0$ yields (E.2). $\square$

Proof of Theorem E.3. In the case $n = 1$ both measures σ and $\mathcal{H}_{n-1}$ are the counting measure, so we only need to treat the case $n \geqslant 2$.

(i) Let $z \in \partial\Omega$, and let $\nu(z)$ be the outer unit normal at z. In the first step we construct a special local parametrisation of $\partial\Omega$ around z. There exists an orthogonal matrix B such that $B\nu(z) = e_n$, the n-th unit vector. Applying B to $\partial\Omega$ one can see that there exists an open neighbourhood W of z in $\partial\Omega$ such that $B(W)$ is a standard C^1-graph, i.e., there exist an open set $W' \subseteq \mathbb{R}^{n-1}$ and a C^1-function $g\colon W' \to \mathbb{R}$ such that $B(W) = \{(y, g(y))\,;\ y \in W'\}$. There exists $z' \in W'$ such that $Bz = (z', g(z'))$. Then $\nabla g(z') = 0$ because $B\nu(z) = e_n$; one could say that $B(W)$ is *flat at Bz*.

Next, let $\alpha > 1$. Reduce W' to a smaller convex open neighbourhood of z' by requiring $\sqrt{1+|\nabla g(y')|^2} \leqslant \alpha$ for all $y' \in W'$, and reduce the set W accordingly. It is easy to see that then the mapping $\varphi\colon W' \to B(W),\ y' \mapsto (y', g(y'))$ is Lipschitz continuous with Lipschitz constant α. Let $A \subseteq W$ be a Borel set, $A' := \varphi^{-1}(B(A))$. Then

$$\sigma(A) = \sigma(B(A)) = \int_{A'} \sqrt{1+|\nabla g(y')|^2}\,\mathrm{d}y',$$

and by estimating the integrand below and above we obtain

$$\lambda^{n-1}(A') \leqslant \sigma(A) \leqslant \alpha\lambda^{n-1}(A'). \tag{E.3}$$

Let $\mathrm{pr}_{n-1}\colon \mathbb{R}^n \to \mathbb{R}^{n-1}$ denote the canonical projection, $\mathrm{pr}_{n-1}(x_1,\ldots,x_n) := (x_1,\ldots,x_{n-1})$, and note that $A' = \mathrm{pr}_{n-1}(B(A))$. From Lemma E.4 we conclude that

$$\mathcal{H}_{n-1}(A') = \mathcal{H}_{n-1}\big(\mathrm{pr}_{n-1}(B(A))\big) \leqslant \mathcal{H}_{n-1}(B(A)) = \mathcal{H}_{n-1}(\varphi(A')) \leqslant \alpha^{n-1}\mathcal{H}_{n-1}(A'),$$

which in view of Theorem E.1 and the obvious equality $\mathcal{H}_{n-1}(B(A)) = \mathcal{H}_{n-1}(A)$ can be rewritten as

$$\lambda^{n-1}(A') \leqslant \mathcal{H}_{n-1}(A) \leqslant \alpha^{n-1}\lambda^{n-1}(A'). \tag{E.4}$$

Combining the estimates (E.3) and (E.4) we obtain

$$\big|\sigma(A) - \mathcal{H}_{n-1}(A)\big| \leqslant (\alpha^n - 1)\lambda^{n-1}(A') \leqslant (\alpha^n - 1)\sigma(A).$$

(ii) In order to continue we rephrase the result of step (i) as follows. Let $\alpha > 1$. Then for every $z \in \partial\Omega$ there exists an open neighbourhood W_z in $\partial\Omega$ such that for every Borel set $A_z \subseteq W_z$ one has $\big|\sigma(A_z) - \mathcal{H}_{n-1}(A_z)\big| \leqslant (\alpha^n - 1)\sigma(A_z)$. From the open covering $\{W_z\,;\ z \in \partial\Omega\}$ of $\partial\Omega$ we choose a finite subcovering $\{W_z\,;\ z \in F\}$.

Let $A \subseteq \partial\Omega$ be a Borel set. Then there exists a partition $\{A_z\,;\ z \in F\}$ of A, with a Borel set $A_z \subseteq W_z$ for all $z \in F$, and we obtain

$$|\sigma(A) - \mathcal{H}_{n-1}(A)| \leqslant \sum_{z\in F}\big|\sigma(A_z) - \mathcal{H}_{n-1}(A_z)\big| \leqslant \sum_{z\in F}(\alpha^n - 1)\sigma(A_z) \leqslant (\alpha^n - 1)\sigma(\partial\Omega).$$

As the inequality between the outer terms of this chain of inequalities holds for all $\alpha > 1$, we conclude that $\sigma(A) = \mathcal{H}_{n-1}(A)$. □

E.5 Remarks. (a) The proof of Theorem E.3 can be modified to show that more generally one has equality of the d-dimensional Hausdorff measure and the "usual" d-dimensional surface measure on embedded d-dimensional manifolds, for $d = 1,\ldots,n-1$.

(b) Much more generally, it is shown in the literature that the image of an injective Lipschitz continuous map $\varphi\colon U \to \mathbb{R}^n$, where $U \subseteq \mathbb{R}^d$ (with $d \in \{1,\ldots,n-1\}$) is a bounded open set, has d-dimensional Hausdorff measure given by $\mathcal{H}_d(\varphi(U)) = \int_U \sqrt{\det(\varphi'(y)^\top\varphi'(y))}\,\mathrm{d}y$. This fact and the formula carry the name of "area formula" in geometric measure theory. (It deserves to be mentioned that, by Rademacher's theorem, the derivative occurring in the formula exists almost everywhere.) We refer to [EvGa92; Sect. 3.3, Theorem 1], [AFP00; Theorem 2.71], for the full formulation and the proof of the area formula. △

E.3 The isodiametric inequality

The issue of this section is to give a proof of the isodiametric inequality (E.1).

E.6 Theorem. *Let $A \subseteq \mathbb{R}^n$ be a bounded set. Then*

$$\lambda^{n,*}(A) \leqslant \omega_n \operatorname{rd}(A)^n. \tag{E.5}$$

E.7 Remarks. (a) Note that a set $A \subseteq \mathbb{R}^n$ need not be contained in a ball of radius $\operatorname{rd}(A)$; an example in $\mathbb{R}^2$ is the isolateral triangle.

(b) If A is symmetric with respect to the origin, i.e. $A = -A$, then $2|x| = |x - (-x)| \leqslant \operatorname{diam}(A)$ for all $x \in A$, hence $A \subseteq B[0, \operatorname{rd}(A)]$, and (E.5) holds. △

The idea of the proof of Theorem E.6 is the reduction to the case of Remark E.7(b), by successive symmetrisation of A. The one-dimensional symmetrisation of a bounded set $B \subseteq \mathbb{R}$ is the open interval

$$S(B) := \big(-\tfrac{1}{2}\lambda^{1,*}(B), \tfrac{1}{2}\lambda^{1,*}(B)\big)$$

(which is empty if B is a null set). Now let $A \subseteq \mathbb{R}^n$ be bounded, and let $(e_1, \dots, e_n)$ denote the standard orthonormal basis of $\mathbb{R}^n$. We define the **Steiner symmetrisation** $S_n(A)$ of A with respect to the coordinate hyperplane $e_n^\perp$ by

$$S_n(A) := \bigcup_{\check{x} \in \mathbb{R}^{n-1}} \{\check{x}\} \times S(A_{\check{x}}),$$

where, for $\check{x} \in \mathbb{R}^{n-1}$, we use the notation

$$A_{\check{x}} := \{x_n \in \mathbb{R};\ (\check{x}, x_n) \in A\}.$$

(Usually, the Steiner symmetrisation is defined with closed intervals; we found it more convenient to work with open intervals.) The Steiner symmetrisation consists in subdividing $\mathbb{R}^n$ into one-dimensional fibers and then applying the one-dimensional symmetrisation in each fiber. Clearly $S_n(A)$ is symmetric with respect to $e_n^\perp$.

The symmetrisations $S_k(A)$ of A with respect to the coordinate hyperplanes $e_k^\perp$, for $k = 1, \dots, n-1$, are defined analogously.

E.8 Lemma. *Let $A \subseteq \mathbb{R}^n$ be a bounded set.*

(a) *Then* $\operatorname{diam}(S_n(A)) \leqslant \operatorname{diam}(A)$.

(b) *If $k \in \{1, \dots, n-1\}$, and A is symmetric with respect to $e_k^\perp$, then $S_n(A)$ is symmetric with respect to $e_k^\perp$ as well.*

(c) *If A is a Borel set, then $S_n(A)$ is a Borel set, and $\lambda^n(S_n(A)) = \lambda^n(A)$.*

Proof. (a) Let $x = (\check{x}, x_n)$, $y = (\check{y}, y_n) \in S_n(A)$, and put $r_1 := \inf A_{\check{x}}$, $r_2 := \sup A_{\check{x}}$ and $s_1 := \inf A_{\check{y}}$, $s_2 := \sup A_{\check{y}}$. Then

$$\begin{aligned}
|x_n - y_n| &\leqslant |x_n| + |y_n| < \tfrac{1}{2}\lambda^{1,*}(A_{\check{x}}) + \tfrac{1}{2}\lambda^{1,*}(A_{\check{y}}) \\
&\leqslant \tfrac{1}{2}(r_2 - r_1) + \tfrac{1}{2}(s_2 - s_1) = \tfrac{1}{2}(r_2 - s_1) + \tfrac{1}{2}(s_2 - r_1) \\
&\leqslant \max\{r_2 - s_1, s_2 - r_1\}.
\end{aligned}$$

Therefore

$$\begin{aligned}|x-y|^2 &= |\check{x}-\check{y}|^2 + (x_n-y_n)^2 \\ &< |\check{x}-\check{y}|^2 + \max\big\{(r_2-s_1)^2, (s_1-r_2)^2\big\} \\ &= \max\big\{|(\check{x},r_2)-(\check{y},s_1)|^2, |(\check{y},s_2)-(\check{x},r_1)|^2\big\} \\ &\leqslant \operatorname{diam}(A)^2,\end{aligned}$$

where the last inequality holds because the points $(\check{x}, r_1), (\check{x}, r_2), (\check{y}, s_1), (\check{y}, s_2)$ belong to the closure of A. This proves the assertion.

(b) The symmetry of A with respect to $e_k^{\perp}$ implies that $A_{R_k\check{x}} = A_{\check{x}}$ for all $\check{x} \in \mathbb{R}^{n-1}$, where R_k denotes the reflection of $\mathbb{R}^{n-1}$ in the $(n-2)$-dimensional coordinate hyperplane $e_k^{\perp}$. Then it follows from the definition of $S_n(A)$ that $S_n(A)$ is symmetric, too.

(c) From Fubini's theorem we know that the function

$$f\colon \mathbb{R}^{n-1} \to \mathbb{R}, \quad \check{x} \mapsto \int_{\mathbb{R}} \mathbf{1}_A(\check{x}, x_n)\,\mathrm{d}x_n = \lambda^1(A_{\check{x}})$$

is Borel measurable. It follows that the function $g\colon \mathbb{R}^n \to \mathbb{R}$, $g(\check{x}, x_n) := \frac{1}{2}f(\check{x}) - |x_n|$ ($\check{x} \in \mathbb{R}^{n-1},\ x_n \in \mathbb{R}$), is Borel measurable, and therefore $S_n(A) = [g > 0]$ is a Borel set. Using Fubini's theorem again we conclude that

$$\lambda^n(A) = \int_{\mathbb{R}^n} \mathbf{1}_A(x)\,\mathrm{d}x = \int_{\check{x}\in\mathbb{R}^{n-1}} \lambda^1(A_{\check{x}})\,\mathrm{d}\check{x} = \lambda^n(S_n(A)). \qquad \square$$

Proof of Theorem E.6. Without loss of generality we suppose that A is closed (and hence a Borel set). Put $A_0 := S_1(\dots S_{n-1}(S_n(A))\dots)$; then Lemma E.8(c) implies $\lambda^n(A) = \lambda^n(A_0)$. Moreover, by Lemma E.8(b), A_0 is symmetric whith respect to all the coordinate hyperplanes $e_1^{\perp}, \dots, e_n^{\perp}$, hence with respect to the origin, and therefore $\lambda^n(A_0) \leqslant \omega_n \operatorname{rd}(A_0)^n$ by Remark E.7(b). Finally, Lemma E.8(a) shows that $\operatorname{rd}(A_0) \leqslant \operatorname{rd}(A)$, and this concludes the proof. $\square$

Notes

The contents of this appendix are standard subjects of geometric measure theory and can be found, for instance, in [Fed69], [EvGa92], [AFP00]. In Section E.2 we treat the equality of the surface measure and the $(n-1)$-dimensional Hausdorff measure in less generality than in the indicated references, thereby allowing an easier access. For comments on the general case we refer to Remark E.5(b). Our proof of the isodiametric inequality in Section E.3 is a reformulation of the proof given in [EvGa92; Section 2.2].

Exercises

E.1 Let (M_1, ρ_1), (M_2, ρ_2) be metric spaces, $d \in [0, \infty)$, $\varphi\colon M_1 \to M_2$ Hölder continuous with exponent $\alpha \in (0, \infty)$, Hölder constant $H \geqslant 0$, i.e. $\rho_2(\varphi(x), \varphi(y)) \leqslant H\rho_1(x, y)^{\alpha}$ for all $x, y \in M_1$. Show that $\mathcal{H}^*_{d/\alpha}(\varphi(A)) \leqslant H^{d/\alpha}\mathcal{H}^*_d(A)$, for all $A \subseteq M_1$. (See Lemma E.4 for a special case of this assertion.)

Appendix F

The spectral theorem for self-adjoint operators

In this appendix we prove the spectral theorem, which we have already stated in Chapter 13.

F.1 Theorem. *Let H be a Hilbert space, and let A be a self-adjoint operator in H. Then there exist a semi-finite measure space $(\Omega, \mathcal{A}, \mu)$, a measurable function $\alpha\colon \Omega \to \mathbb{R}$ and a unitary operator $J\colon H \to L_2(\Omega, \mathcal{A}, \mu)$ such that*

$$A = J^{-1} M_\alpha J. \tag{F.1}$$

The identity (F.1) means that

$$\operatorname{dom}(A) = \{x \in H\,;\, \alpha Jx \in L_2(\mu)\}, \qquad Ax = J^{-1}(\alpha Jx).$$

Clearly, the spectral theorem is only of interest in non-trivial Hilbert spaces. For this reason we exclude the Hilbert space $H = \{0\}$ in the present appendix.

F.1 The spectral theorem for bounded self-adjoint operators

Throughout this section let H be a (real or complex) Hilbert space and $A \in \mathcal{L}(H)$ a self-adjoint operator. It will be important to remember that $\sigma(A) \subseteq \mathbb{R}$; in the complex case this holds because the operators $\pm \mathrm{i}A$ are m-accretive, by Remark 3.18. (See also Exercise 6.2.)

The proof of the spectral theorem relies on some basic properties of the functional calculus for self-adjoint operators. The function spaces needed below will consist of real- or complex-valued functions, according to whether H is a real or complex Hilbert space.

We start by associating with each polynomial $p = a_0 + a_1 t + \cdots + a_n t^n$ the operator $p(A) := a_0 I + a_1 A + \cdots + a_n A^n$. Let $\mathcal{P}$ denote the set of all polynomials. Then it is straightforward that the mapping $\mathcal{P} \ni p \mapsto p(A) \in \mathcal{L}(H)$ is an algebra homomorphism. The following theorem states further properties and extends this homomorphism to the continuous functions on the spectrum of A. A polynomial p can be considered as a function on $\mathbb{K}$; we denote its restriction to $\sigma(A)$ by $p|_{\sigma(A)}$.

F.2 Theorem. *There exists an algebra homomorphism* $\Phi\colon C(\sigma(A)) \to \mathcal{L}(H)$ *satisfying*

$\Phi(p|_{\sigma(A)}) = p(A)$ *for all polynomials* p, *in particular* $\Phi(\mathbf{1}_{\sigma(A)}) = I$, $\Phi(\mathrm{id}_{\sigma(A)}) = A$,

$\|\Phi(f)\| \leqslant \|f\|_\infty$ *for all* $f \in C(\sigma(A))$,

$\Phi(\bar f) = \Phi(f)^*$ *for all* $f \in C(\sigma(A))$,

$\Phi(f) \geqslant 0$ *for all* $f \in C(\sigma(A))$, $f \geqslant 0$.

Moreover, Φ *is uniquely determined by the first two properties.*

Proof. We first assume that p is a real polynomial, i.e. all the coefficients of p are real. Then $p(A)$ is self-adjoint. Let $\lambda \in \mathbb{R}$ and suppose that $p(t) \neq \lambda$ for all $t \in \sigma(A)$. We will show that $\lambda \notin \sigma(p(A))$; then it follows that $\sigma(p(A)) \subseteq \mathrm{ran}(p|_{\sigma(A)})$, and hence $\|p(A)\| \leqslant \|p|_{\sigma(A)}\|_\infty$ by Proposition 6.15.

Replacing the polynomial p by $p - \lambda$ we see that we may assume without loss of generality that $\lambda = 0$. Then we can decompose p as

$$p(t) = c \prod_{j\in J}(t-\alpha_j) \prod_{k\in K}\big((t-\beta_k)^2 + \gamma_k^2\big),$$

with finite sets J, K and $c \in \mathbb{R}\setminus\{0\}$, $\alpha_j \in \mathbb{R}\setminus\sigma(A)$ $(j \in J)$, $\beta_k, \gamma_k \in \mathbb{R}$, $\gamma_k \neq 0$ $(k \in K)$. The latter, quadratic factors correspond to non-real zeros of p – occurring in pairs of complex conjugate zeros $\beta_k \pm \mathrm{i}\gamma_k$. It follows that $p(A)$ is invertible in $\mathcal{L}(H)$, i.e. $0 \notin \sigma(p(A))$, because all the factors $A - \alpha_j$ $(j \in J)$ and $(A-\beta_k)^2 + \gamma_k^2$ $(k \in K)$ are invertible. (Concerning the latter factors, note that $(A-\beta_k)^2$ is m-accretive.)

For general $p \in \mathcal{P}$ one readily sees that $p(A)^* = \bar p(A)$, where $\bar p$ is the polynomial obtained from p by replacing the coefficients of p by their complex conjugates. Then, using the assertion for real polynomials shown above, one estimates

$$\|p(A)x\|^2 = \big(p(A)^*p(A)x \,\big|\, x\big) = \big((\bar p p)(A)x \,\big|\, x\big) \leqslant \big\|(\bar p p)|_{\sigma(A)}\big\|_\infty \|x\|^2$$

for all $x \in H$, and this implies that $\|p(A)\| \leqslant \|p|_{\sigma(A)}\|_\infty$. As a consequence, if $p, q \in \mathcal{P}$ are such that $p|_{\sigma(A)} = q|_{\sigma(A)}$, then $p(A) = q(A)$. Therefore, putting

$$\Phi(p|_{\sigma(A)}) := p(A) \qquad (p \in \mathcal{P})$$

one obtains a well-defined mapping $\Phi\colon \mathcal{P}_{\sigma(A)} \to \mathcal{L}(H)$, where $\mathcal{P}_{\sigma(A)} := \big\{p|_{\sigma(A)}\,;\, p \in \mathcal{P}\big\}$. Moreover, Φ thus defined is an algebra homomorphism.

We have proved the asserted properties (except the final one) for Φ, defined on $\mathcal{P}_{\sigma(A)}$. The set $\mathcal{P}_{\sigma(A)}$ is dense in $C(\sigma(A))$ by the Stone–Weierstrass theorem, Theorem B.2 (or by the Weierstrass approximation theorem, together with Exercise F.1). Since Φ is a bounded operator, it has a unique continuous extension to $C(\sigma(A))$, and the properties shown so far carry over. Finally $\Phi(f) = \Phi(f^{1/2})^*\Phi(f^{1/2}) \geqslant 0$ for all $0 \leqslant f \in C(\sigma(A))$. □

We will use the notation $f(A) := \Phi(f)$ for all $f \in C(\sigma(A))$. For $x \in H \setminus \{0\}$ we define

$$H_x := \overline{\mathrm{lin}\{A^n x\,;\, n \in \mathbb{N}_0\}} = \overline{\{p(A)x\,;\, p \in \mathcal{P}\}} = \overline{\{f(A)x\,;\, f \in C(\sigma(A))\}},$$

the **cyclic subspace** generated by x. Obviously H_x is invariant under A, and therefore under $f(A)$ for all $f \in C(\sigma(A))$. If $H_x = H$, then x is called a **cyclic vector** (for A); the following theorem contains the spectral theorem in this special case.

F.3 Theorem. *Let $x \in H \setminus \{0\}$. Then there exist a unique Borel measure μ on $\sigma(A)$ such that*

$$(f(A)x \,|\, x) = \int_{\sigma(A)} f \,\mathrm{d}\mu \qquad \bigl(f \in C(\sigma(A))\bigr) \tag{F.2}$$

and a unitary operator $J\colon H_x \to L_2(\sigma(A), \mu)$ such that $f(A)|_{H_x} = J^{-1}M_f J$ for all $f \in C(\sigma(A))$.

The measure μ from Theorem F.3 is the **spectral measure** associated with x. It is a finite measure, in fact

$$\mu(\sigma(A)) = \int_{\sigma(A)} 1 \,\mathrm{d}\mu = \|x\|^2 \tag{F.3}$$

by (F.2).

Proof of Theorem F.3. We start with the crucial observation that

$$C(\sigma(A)) \ni f \mapsto (f(A)x \,|\, x) \in \mathbb{K}$$

defines a positive linear functional on $C(\sigma(A))$, by Theorem F.2. Thus from the Riesz–Markov theorem – see Appendix G – we obtain a unique Borel measure μ on $\sigma(A)$ satisfying (F.2). For $f, g \in C(\sigma(A))$ one has $g(A)^* f(A) = (f\overline{g})(A)$ by Theorem F.2, and it follows that

$$\bigl(f(A)x \,\big|\, g(A)x\bigr) = \bigl((f\overline{g})(A)x \,\big|\, x\bigr) = \int_{\sigma(A)} f\overline{g} \,\mathrm{d}\mu. \tag{F.4}$$

Now we consider $C(\sigma(A))$ as a subspace of $L_2(\sigma(A), \mu)$, with μ-a.e. equal functions identified. With this understanding, equality (F.4) shows that the mapping

$$\bigl(C(\sigma(A)), \|\cdot\|_{L_2(\sigma(A),\mu)}\bigr) \ni f \mapsto f(A)x \in H_0 := \bigl\{f(A)x;\ f \in C(\sigma(A))\bigr\}$$

is isometric, with isometric inverse

$$J_0\colon H_0 \to L_2(\sigma(A), \mu), \qquad J_0(f(A)x) := f \quad \bigl(f \in C(\sigma(A))\bigr).$$

As H_0 is dense in H_x and $C(\sigma(A))$ is dense in $L_2(\sigma(A), \mu)$, by Theorem G.9, the continuous extension J of J_0 is a unitary operator $J\colon H_x \to L_2(\sigma(A), \mu)$.

Now let $f \in C(\sigma(A))$. If $g, h \in C(\sigma(A))$, $y := g(A)x$, $z := h(A)x$ ($\in H_0$), then

$$(f(A)y \,|\, z) = \bigl(f(A)g(A)x \,\big|\, h(A)x\bigr) = \int_{\sigma(A)} fg\overline{h} \,\mathrm{d}\mu = (fJy \,|\, Jz)_{L_2(\sigma(A),\mu)}\,.$$

This extends to

$$(f(A)y \,|\, z) = (fJy \,|\, Jz)_{L_2(\sigma(A),\mu)} = (J^* M_f J y \,|\, z)$$

for all $y, z \in H_x$, which shows that $f(A)|_{H_x} = J^{-1}M_f J$. □

If $x, y \in H$ and y is orthogonal to H_x, then

$$\bigl(f(A)y \,\big|\, g(A)x\bigr) = \bigl(y \,\big|\, (\overline{f}g)(A)x\bigr) = 0$$

for all $f, g \in C(\sigma(A))$, and this implies that $H_y \perp H_x$.

F.4 Proposition. *There exists a set $X \subseteq H \setminus \{0\}$ such that $H_x \perp H_y$ for all $x, y \in X$ with $x \neq y$ and such that $H = \bigoplus_{x\in X} H_x$.*

We note that the (possibly uncountably infinite) orthogonal direct sum $\bigoplus_{x\in X} H_x$ is the closed linear hull of $\bigcup_{x\in X} H_x$ and consists of all elements $y \in H$ that can be written as $y = \sum_{x\in X} y_x$ with suitable $y_x \in H_x$, where the set $\{x \in X\,;\, y_x \neq 0\}$ is countable; see Exercises F.3(c) and F.4(b). For simplicity we will also write the element y as a family $(y_x)_{x\in X}$.

Proof of Proposition F.4. The set

$$\mathcal{X} := \bigl\{X \subseteq H \setminus \{0\}\,;\; H_x \perp H_y \text{ for all } x, y \in X \text{ with } x \neq y\bigr\}$$

is ordered by inclusion, and any chain $\mathcal{C}$ in $\mathcal{X}$ is dominated, by $\bigcup_{X\in\mathcal{C}} X$. Zorn's lemma implies that the $\mathcal{X}$ has a maximal element X. For this set X one has $H = \bigoplus_{x\in X} H_x$, because otherwise there would exist $0 \neq y \in \bigl(\bigoplus_{x\in X} H_x\bigr)^{\perp}$, and then $X \cup \{y\} \in \mathcal{X}$, in contradiction to the maximality of X. □

F.5 Remark. It is not hard to see that the set X in Proposition F.4 is countable if and only if H is separable. If H is separable one does not need Zorn's lemma for the proof, because one can construct the set X, starting with a countable dense subset of H. △

Proof of Theorem F.1 for the case of bounded self-adjoint operators.
Let $X \subseteq H \setminus \{0\}$ be a set with the properties stated in Proposition F.4. For each $x \in X$ let μ_x be the spectral measure associated with x, and let A_x be the restriction of A to H_x. Let $M_{\mathrm{id},x}$ denote the operator of multiplication by id on $L_2(\sigma(A), \mu_x)$. Then there exists a unitary operator $J_x \colon H_x \to L_2(\sigma(A), \mu_x)$ such that $A_x = J_x^{-1} M_{\mathrm{id},x} J_x$; see Theorem F.3 and the comments after its proof. We compose the family $(J_x)_{x\in X}$ to a unitary operator

$$\check{J} := \bigoplus_{x\in X} J_x \colon \bigoplus_{x\in X} H_x \,(= H) \to \bigoplus_{x\in X} L_2(\sigma(A), \mu_x)$$

in the canonical way, i.e. $\check{J}(y_x)_{x\in X} = (J_x y_x)_{x\in X}$. Similarly we compose the family $(M_{\mathrm{id},x})_{x\in X}$ to a bounded operator M on the Hilbert space $\bigoplus_{x\in X} L_2(\sigma(A), \mu_x)$, putting $M(f_x)_{x\in X} = (M_{\mathrm{id},x} f_x)_{x\in X}$. Then $A = \check{J}^{-1} M \check{J}$.

In order to transform M into a multiplication operator we define $\Omega := \sigma(A) \times X$. For $B \subseteq \Omega$ and $x \in X$ we put $B_x := \{\lambda \in \sigma(A)\,;\, (\lambda, x) \in B\}$; then $B = \bigcup_{x\in X} (B_x \times \{x\})$. We further define

$$\mathcal{A} := \bigl\{B \subseteq \Omega\,;\; B_x \text{ measurable } (x \in X)\bigr\},$$
$$\mu(B) := \sum_{x\in X} \mu_x(B_x) \qquad (B \in \mathcal{A}).$$

Then $(\Omega, \mathcal{A}, \mu)$ is a measure space and

$$\hat{J} \colon \bigoplus_{x\in X} L_2(\sigma(A), \mu_x) \to L_2(\Omega, \mu), \quad \hat{J}(f_x)_{x\in X} := [(\lambda, x) \mapsto f_x(\lambda)]$$

is a unitary operator; see Exercise F.5. Moreover, defining $\alpha\colon \Omega \to \mathbb{R}$ by $\alpha(\lambda, x) := \lambda$ one easily sees that α is measurable and that $\hat{J}^{-1} M_\alpha \hat{J} = M$. We conclude that

$$A = \big(\hat{J}\check{J}\big)^{-1} M_\alpha \hat{J}\check{J}.$$

Finally we show that the measure space $(\Omega, \mathcal{A}, \mu)$ constructed above is semi-finite. We first mention that for every $x \in H$ the spectral measure μ_x is a finite measure; see (F.3). Now, if $B \in \mathcal{A}$ is such that $\mu(B) = \infty$, then by the definition of μ there exists $x \in X$ such that $\mu_x(B_x) > 0$. As we already know that $\mu_x(B_x) < \infty$ we have found the required subset $B_x \times \{x\}$ of B for the property of semi-finiteness of μ. □

F.6 Remarks. (a) If H is separable, then the set X can be chosen such that μ is finite. Indeed, as $\mu(\Omega) = \sum_{x\in X} \mu_x(\sigma(A)) = \sum_{x\in X} \|x\|^2$ (recall (F.3)), one only has to choose X such that $\sum_{x\in X} \|x\|^2 < \infty$, and this can be arranged if X is countable.

(b) In Exercise F.9 the reader is asked to show that any measure μ can be modified to a semi-finite measure without changing the space $L_2(\mu)$. △

F.2 The spectral theorem for self-adjoint operators; the general case

In this section we use the spectral theorem for bounded self-adjoint operators to give a proof for the general case. As before, let H be a (real or complex) Hilbert space.

We begin with some preliminary material. The main tool is the following result of von Neumann.

F.7 Theorem. *Let G be a Hilbert space, and let A be a closed operator from H to G,* $\operatorname{dom}(A)$ *dense. Then A^*A is an accretive self-adjoint operator in H.*

Proof. Let $z \in H$. From the equality $A^\perp = -(A^*)^{-1}$ in $H \oplus G$ (see Section 6.1) we conclude that $(z, 0) \in H \oplus G$ can be written as $(z, 0) = (x, Ax) + (-A^*y, y)$, with suitable $x \in \operatorname{dom}(A)$, $y \in \operatorname{dom}(A^*)$. It follows that $z = x - A^*y = x + A^*Ax$, so we have shown that $\operatorname{ran}(I + A^*A) = H$. Clearly

$$(A^*Ax \,|\, y) = (Ax \,|\, Ay) = (x \,|\, A^*Ay) \qquad \big(x, y \in \operatorname{dom}(A^*A) \subseteq \operatorname{dom}(A)\big).$$

In particular $(A^*Ax \,|\, x) \geqslant 0$ for all $x \in \operatorname{dom}(A^*A)$, i.e. A^*A is accretive, and then Proposition 6.9 implies that $I + A^*A$ is self-adjoint. □

F.8 Remark. Note that the assertion of Theorem F.7 also contains the information that $\operatorname{dom}(A^*A)$ is dense in H – a property that is not a priori obvious. △

Given an accretive self-adjoint operator C, we will use the notation $(I + C)^{-1/2}$ as a shorthand for the (accretive) square root of the bounded accretive self-adjoint operator $(I + C)^{-1}$. This square root exists by Theorem F.2.

F.9 Lemma. *Let A be a self-adjoint operator in H. Then A^2 is an accretive self-adjoint operator, $B := A(I + A^2)^{-1/2}$ belongs to $\mathcal{L}(H)$ and is a self-adjoint contraction, and -1 and 1 are not eigenvalues of B.*

Proof. From $A^* = A$ and Theorem F.7 we conclude that A^2 is an accretive self-adjoint operator; hence $(I + A^2)^{-1}$ is an accretive self-adjoint operator.

Let $x \in \mathrm{dom}(A)$, $y := (I + A^2)^{-1}x$. Then $y \in \mathrm{dom}(A^3)$,

$$A(I + A^2)^{-1}x = Ay = (I + A^2)^{-1}A(I + A^2)y = (I + A^2)^{-1}Ax,$$

and therefore $A\,p\big((I+A^2)^{-1}\big)x = p\big((I+A^2)^{-1}\big)Ax$ for all polynomials p. Since $(I+A^2)^{-1/2}$ can be approximated in $\mathcal{L}(H)$ by polynomials in $(I + A^2)^{-1}$ (see Theorem F.2) and since A is closed, it follows that $(I + A^2)^{-1/2}x \in \mathrm{dom}(A)$ and

$$A(I + A^2)^{-1/2}x = (I + A^2)^{-1/2}Ax. \tag{F.5}$$

This implies that

$$\begin{aligned}\big\|A(I + A^2)^{-1/2}x\big\|^2 &= \big((I + A^2)^{-1}Ax \,\big|\, Ax\big) = \big(A^2(I + A^2)^{-1}x \,\big|\, x\big) \\ &\leqslant \big((I + A^2)(I + A^2)^{-1}x \,\big|\, x\big) = \|x\|^2.\end{aligned}$$

As $\mathrm{dom}(A)$ is dense in H and A is closed, we conclude that $\mathrm{ran}\big((I + A^2)^{-1/2}\big) \subseteq \mathrm{dom}(A)$ and $B \in \mathcal{L}(H)$, $\|B\| \leqslant 1$.

For all $x, y \in \mathrm{dom}(A)$ one has

$$\big(A(I + A^2)^{-1/2}x \,\big|\, y\big) = \big((I + A^2)^{-1/2}Ax \,\big|\, y\big) = \big(x \,\big|\, A(I + A^2)^{-1/2}y\big),$$

which, again by the denseness of $\mathrm{dom}(A)$, shows that B is symmetric, hence self-adjoint.

Using (F.5) we obtain

$$I - B^2 = I - A^2(I + A^2)^{-1} = (I + A^2)^{-1}. \tag{F.6}$$

Hence, if $x \in H$ satisfies $Bx = \pm x$, then $0 = x - B^2x = (I + A^2)^{-1}x$ and thus $x = 0$. □

Proof of Theorem F.1, general case.
By Lemma F.9, the operator $B := A(I + A^2)^{-1/2}$ is self-adjoint and contractive. Moreover (F.6) implies that

$$B = A(I + A^2)^{-1/2} = A(I - B^2)^{1/2}.$$

We know from Theorem F.1 for the case of a bounded self-adjoint operator – see Section F.1 for its proof – that B is unitarily equivalent to a multiplication operator M_β on $L_2(\mu)$ for some semi-finite measure space $(\Omega, \mathcal{A}, \mu)$, with measurable $\beta\colon \Omega \to (-1, 1)$. (For the property that β takes its values in $(-1, 1)$ we apply properties (a) and (c) of Remark 13.22, taking into account the contractivity of B and the injectivity of $\pm I - B$.)

In order to simplify notation we now assume – without loss of generality – that $H = L_2(\mu)$ and $B = M_\beta$. Then we have $M_\beta = B = A(I - B^2)^{1/2} = AM_{(1-\beta^2)^{1/2}}$, and composing on the right with $M_{(1-\beta^2)^{-1/2}}$ we obtain

$$M_\beta M_{(1-\beta^2)^{-1/2}} = A|_{\mathrm{ran}(I-B^2)^{1/2}} \subseteq A.$$

One easily sees that $M_\beta M_{(1-\beta^2)^{-1/2}} = M_\alpha$, where $\alpha := \beta(1 - \beta^2)^{-1/2}$; indeed, this is an immediate consequence of $\mathrm{dom}(M_\alpha) = \mathrm{dom}(M_{(1-\beta^2)^{-1/2}})$, which in turn follows from the equality $\alpha^2 + 1 = 1/(1 - \beta^2)$.

The conclusion is that $M_\alpha \subseteq A$, and since both M_α and A are self-adjoint one obtains $M_\alpha = A$. □

F.10 Remark. The idea of the proof of Theorem F.1 for the general case is to find an injective function $f\colon \sigma(A) \to \mathbb{R}$ such that $f(A)$ becomes a bounded self-adjoint operator (to which the result for the bounded case can be applied). The sophisticated choice $f(t) = t(1+t^2)^{-1/2}$ above is needed when $\sigma(A) = \mathbb{R}$. If there exists a point $\lambda \in \mathbb{R} \setminus \sigma(A)$, then one can use $f(t) := (\lambda - t)^{-1}$ and argue as above, thereby reducing the proof to the bounded self-adjoint operator $(\lambda - A)^{-1}$. △

Notes

The version of the spectral theorem we have presented in Theorem 13.21 is well-known and well-established. Nevertheless we have been at a loss finding a complete and simple proof in the literature, without the restriction to separable or complex Hilbert spaces. This is why we decided to include a complete proof.

The proof we have given draws on various sources. The proof of the functional calculus for bounded operators, Theorem F.2, is inspired by Riesz–Sz.-Nagy [RiNa68; Kap. VII, Nr. 106], with a shortcut made possible by our Proposition 6.15. The procedure for finding the spectral measure for elements in the Hilbert space (Theorem F.3), the decomposition of the Hilbert space into cyclic subspaces (Proposition F.4) and the construction of the measure space for the representation of the operator as a multiplication operator in Section F.1 are rather standard and have no special sources. The proof of Theorem F.7 is taken from Kato [Kat80; Chap. V, Theorem 3.24]. To our knowledge, the method for deriving the spectral theorem for unbounded operators from the case of bounded operators using the operator $A(I+A^2)^{-1/2}$, as presented in Section F.2, appeared for the first time in Schmüdgen's book [Sch12].

Exercises

F.1 Let $K \subseteq \mathbb{R}$ be a compact set, and let $a, b \in \mathbb{R}$, $a < b$, $K \subseteq [a,b]$. Show that for any $f \in C(K)$ there exists a function $\tilde{f} \in C[a,b]$ such that $\tilde{f}|_K = f$, $\|\tilde{f}\|_\infty = \|f\|_\infty$. Conclude that every function $f \in C(K)$ can be approximated by polynomials, uniformly on K. (Hint: There exists a pairwise disjoint family $(U_j)_{j\in N}$, with N countable, of open intervals such that $K = \mathbb{R} \setminus \bigcup_{j\in N} U_j$.)

Note. The existence of $\tilde{f}$ stated above is a special case of the Tietze extension theorem; see for instance [Haa14; Theorem 15.15].

F.2 Let H be a Hilbert space, $A \in \mathcal{L}(H)$ self-adjoint.

(a) Show that the mapping $\Phi\colon C(\sigma(A)) \to \mathcal{L}(H)$, $f \mapsto f(A)$ from Theorem F.2 is injective.

Hint: Assume that there exists $f \in C(\sigma(A))$, $f \neq 0$ such that $f(A) = 0$. Choose $\lambda \in \sigma(A)$ such that $f(\lambda) \neq 0$. Show that $(\lambda - A)^2 = ((\lambda - \mathrm{id})^2 + |f|^2)(A)$ is invertible in $\mathcal{L}(H)$, and conclude that $\lambda \in \rho(A)$.

(b) Let $x \in H$ be a cyclic vector for A. Show that the mapping $C(\sigma(A)) \ni f \mapsto f(A)x \in H$ is injective.

F.3 Let Y be a Banach space, X an index set and $\mathcal{F} := \{F \subseteq X\,;\, F \text{ finite}\}$. Let $(y_x)_{x\in X}$ be a summable family of vectors $y_x \in Y$. Here, the family is called *summable*, with *sum* $y =: \sum_{x\in X} y_x \in Y$, if

$$\forall \varepsilon > 0\ \exists E \in \mathcal{F}\ \forall F \in \mathcal{F},\ F \supseteq E\colon \Big\| \sum_{x\in F} y_x - y \Big\| \leqslant \varepsilon. \tag{F.7}$$

(a) Show that y is uniquely determined by (F.7).

(b) Show that, with the above notation, $\big\|\sum_{x\in F\setminus E} y_x\big\| \leqslant 2\varepsilon$, and conclude that $\|y_x\| \leqslant 2\varepsilon$ for all $x \in X \setminus E$.

(c) Show that $X_0 := \{x \in X\,;\, y_x \neq 0\}$ is countable.

(d) Assume that the set X_0 from part (c) is infinite. Show that $\sum_{x\in X} y_x = \sum_{n=1}^\infty y_{x_n}$ for every enumeration (x_n) of the (countably infinite) set X_0.

(e) Assume that $Y = \mathbb{R}$ and that $y_x \geqslant 0$ for all $x \in X$. Show that $\sum_{x\in X} y_x = \sup\big\{\sum_{x\in F} y_x\,;\, F \in \mathcal{F}\big\}$.

F.4 Let H be a Hilbert space, X an index set and $(H_x)_{x\in X}$ a family of closed subspaces $H_x \subseteq H$ such that $H_x \perp H_y$ for all $x, y \in X$ with $x \neq y$. For $x \in X$ let P_x denote the orthogonal projection onto H_x.

(a) Let $F \subseteq X$ be finite. Show that $\sum_{x\in F} P_x$ is the orthogonal projection onto $\bigoplus_{x\in F} H_x$.

(b) Assume that $\operatorname{lin}\big(\bigcup_{x\in X} H_x\big)$ is dense in H. Show that every element $y \in H$ can be written as $y = \sum_{x\in X} P_x y$.

F.5 Let X be an index set, $K \subseteq \mathbb{R}$ a Borel set and $\Omega := K \times X$. For each $x \in X$ let μ_x be a Borel measure on K. For $B \subseteq \Omega$ and $x \in X$ denote $B_x := \{\lambda \in K\,;\, (\lambda, x) \in B\}$; then $B = \bigcup_{x\in X} (B_x \times \{x\})$. Define

$$\mathcal{A} = \{B \subseteq \Omega;\ B_x \text{ measurable } (x \in X)\}, \qquad \mu(B) := \sum_{x\in X} \mu_x(B_x) \quad (B \in \mathcal{A}).$$

(a) Show that $(\Omega, \mathcal{A}, \mu)$ thus defined is a measure space. (If one takes $K = \sigma(A)$, then $(\Omega, \mathcal{A}, \mu)$ is the measure space constructed in the proof of Theorem F.1.)

(b) Let $f \in \mathscr{L}_1(\Omega, \mathcal{A}, \mu)$. Show that $f(\cdot, x) \in \mathscr{L}_1(K, \mu_x)$ for all $x \in X$ and

$$\int_\Omega f \,\mathrm{d}\mu = \sum_{x\in X} \int_K f(\lambda, x)\,\mathrm{d}\mu_x(\lambda).$$

Hint: Why can one assume without loss of generality that $f \geqslant 0$? It may be regarded as obvious that $\int_\Omega f\,\mathrm{d}\mu = \int_K f(\lambda, x)\,\mathrm{d}\mu_x(\lambda)$ if there exists $x \in X$ such that $f(\cdot, y) = 0$ for all $y \in X \setminus \{x\}$. Then consider the function $f_F := \mathbf{1}_{K\times F} f \in \mathscr{L}_1(\Omega, \mathcal{A}, \mu)$ for finite sets $F \subseteq X$ and use the property $0 \leqslant f_F \leqslant f$ to conclude that $\int_K f(\lambda, x)\,\mathrm{d}\mu_x(\lambda) = 0$ for all but countably many $x \in X$. Finally use the monotone convergence theorem.

(c) Use part (b) to show that

$$\hat{J}\colon \bigoplus_{x\in X} L_2(K, \mu_x) \to L_2(\Omega, \mu), \quad (f_x)_{x\in X} \mapsto \big((\lambda, x) \mapsto f_x(\lambda)\big)$$

is a unitary operator. (It is part of the exercise to show that the function $(\lambda, x) \mapsto f_x(\lambda)$ is measurable.)

F.6 Prove the statements in Remark F.5.

F.7 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, $H := L_2(\mu)$. Let $\alpha\colon \Omega \to \mathbb{R}$ be measurable, $A := M_\alpha$, and define $\operatorname{dom}^\infty(A) := \bigcap_{n\in\mathbb{N}} \operatorname{dom}(A^n)$.

For $a \leqslant b$ define

$$H_{[a,b]} := \big\{f \in L_2(\mu)\,;\ f = \mathbf{1}_{[a\leqslant\alpha\leqslant b]} f\big\}.$$

Put $\lambda := \frac{a+b}{2}$, $r := \frac{b-a}{2}$, and show that

$$H_{[a,b]} = \big\{f \in \operatorname{dom}^\infty(A)\,;\ \|(A-\lambda)^n f\| \leqslant r^n \|f\|\ (n \in \mathbb{N})\big\}. \tag{F.8}$$

Convince yourself that this implies that the spaces $H_{[a,b]}$ in (13.15) do not depend on the representation of A as a multiplication operator in Theorem 13.21; note that in the description (F.8) the function α does not appear.

F.8 Let A be a self-adjoint operator in a Hilbert space H. Let $(\Omega_j, \mathcal{A}_j, \mu_j)$, α_j, J_j be as in Theorem 13.21, for $j = 1, 2$. Show that

$$J_1^{-1} M_{f\circ\alpha_1} J_1 = J_2^{-1} M_{f\circ\alpha_2} J_2 \tag{F.9}$$

for all Borel measurable functions $f\colon \mathbb{R} \to \mathbb{R}$, by implementing the following steps.

(i) It is sufficient to prove the assertion for all bounded Borel measurable functions.

(ii) Let F denote the set of all bounded Borel measurable functions $f\colon \mathbb{R} \to \mathbb{R}$ for which (F.9) holds. Show that F is an algebra (under pointwise addition and multiplication of functions), that F contains the indicator functions of compact intervals (use the independence of the spaces $H_{[a,b]}$, Exercise F.7) and the function $\mathbf{1}$, and that for all bounded, pointwise convergent sequences in F the limit belongs to F (a consequence of the dominated convergence theorem).

(iii) Use (ii) to show that the set $\mathcal{B} := \{B \subseteq \mathbb{R}\,;\ \mathbf{1}_B \in F\}$ is the Borel σ-algebra of $\mathbb{R}$. Conclude that F is the set of all bounded Borel measurable functions (by using the fact that every bounded Borel measurable function can be obtained as the pointwise limit of a bounded sequence in $S(\mathcal{B}) = \operatorname{lin}\{\mathbf{1}_B\,;\ B \in \mathcal{B}\}$).

Note. Stripping hypotheses to a minimum, we include the following two observations. Let Ω be a set, S an algebra of functions $f\colon \Omega \to \mathbb{R}$. Then $\mathcal{B} := \{B \subseteq \Omega\,;\ \mathbf{1}_B \in S\}$ is a ring of subsets of Ω (see the hint to Exercise 10.5(a)). If additionally $\mathbf{1} \in S$, then $\Omega \in \mathcal{B}$, and $\mathcal{B}$ is an algebra of subsets (see step (iii) above).

F.9 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. We define

$$\mathcal{A}_{\text{fin}} := \{A \in \mathcal{A}\,;\ \mu(A) < \infty\},$$

and for $A \in \mathcal{A}$ we put

$$\mu_{\text{sf}}(A) := \sup\{\mu(B)\,;\ B \in \mathcal{A}_{\text{fin}},\ B \subseteq A\}.$$

Prove the following properties.

(a) μ_{sf} is a semi-finite measure, and $\mu_{\text{sf}}(A) = \mu(A)$ for all $A \in \mathcal{A}_{\text{fin}}$. Every μ-null set is a μ_{sf}-null set.

(b) Every μ-integrable function $f\colon \Omega \to [0,\infty)$ is μ_{sf}-integrable, and $\int f\,\mathrm{d}\mu_{\mathrm{sf}} = \int f\,\mathrm{d}\mu$. If $f\colon \Omega \to [0,\infty)$ is μ_{sf}-integrable, then there exists a set $A \in \mathcal{A}$ with $\mu_{\mathrm{sf}}(\Omega \setminus A) = 0$ such that $\mathbf{1}_A f$ is μ-integrable.

(c) If $f \in L_2(\mu)$, then every μ-representative of f belongs to $L_2(\mu_{\mathrm{sf}})$, and the mapping $L_2(\mu) \to L_2(\mu_{\mathrm{sf}})$ thus defined is an isometric isomorphism. (Hint: Given $f \in L_2(\mu_{\mathrm{sf}})$, use part (b) to find a μ-square-integrable representative of f.)

Appendix G

The Riesz–Markov representation theorem

The aim of this appendix is to prove the Riesz–Markov theorem in the special context of compact metric spaces. We denote by $C(K)$ the space of continuous functions on a compact space K. A linear functional $\eta\colon C(K) \to \mathbb{K}$ is called **positive** if $\eta(f) \geqslant 0$ for all $0 \leqslant f \in C(K)$.

G.1 Theorem (Riesz–Markov). *Let K be a compact metric space, and let $\eta\colon C(K) \to \mathbb{K}$ be a positive linear functional. Then there exists a unique Borel measure μ on K such that*

$$\eta(f) = \int f \,\mathrm{d}\mu$$

for all $f \in C(K)$.

We only need to prove the theorem for $\mathbb{K} = \mathbb{R}$. Indeed, in the case $\mathbb{K} = \mathbb{C}$ observe that $\eta_0 := \eta|_{C(K;\mathbb{R})}\colon C(K;\mathbb{R}) \to \mathbb{R}$ is a positive linear functional. Thus Theorem G.1 yields a measure μ representing η_0 as above, and then

$$\eta(f) = \eta_0(\operatorname{Re} f) + \mathrm{i}\eta_0(\operatorname{Im} f) = \int \operatorname{Re} f \,\mathrm{d}\mu + \mathrm{i} \int \operatorname{Im} f \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu,$$

for all $f \in C(K;\mathbb{C})$.

From now on the scalar field will be $\mathbb{K} = \mathbb{R}$. Throughout this appendix let K be a compact topological space, and let η be a positive linear functional on $C(K)$. Only at the very end we will assume additionally that K is a metric space.

G.1 Elementary properties of positive linear functionals on $C(K)$

We will work with the lattice operations $f \vee g$ and $f \wedge g$ for $f, g \in C(K) = C(K;\mathbb{R})$ (and in fact for any functions $f, g\colon K \to \mathbb{R}$),

$$f \vee g(x) := \max\{f(x), g(x)\}, \quad f \wedge g(x) := \min\{f(x), g(x)\} \qquad (x \in K).$$

We denote the positive cone of $C(K)$ by $C(K)_+ := \{f \in C(K);\ f \geqslant 0\}$.

It is important that positive functionals automatically have the following ‘continuity property’.

G.2 Proposition. *The functional η is σ-continuous, i.e., whenever (f_n) is a decreasing sequence in $C(K)_+$ with $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in K$, then $\lim_{n\to\infty} \eta(f_n) = 0$.*

The essential step in the proof is Dini's lemma, as follows.

G.3 Lemma. *Let (f_n) be a decreasing sequence in $C(K)_+$ with $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in K$. Then $\lim_{n\to\infty} \|f_n\|_\infty = 0$.*

Proof. Let $\varepsilon > 0$. For all $x \in K$ there exists $n_x \in \mathbb{N}$ such that $f_{n_x}(x) < \varepsilon$, and therefore $U_x := [f_{n_x} < \varepsilon]$ is an open neighbourhood of x. From the open covering $(U_x)_{x\in K}$ of K we can choose a finite subcovering $(U_x)_{x\in F}$. Put $m := \max_{x\in F} n_x$. Since $f_m \leqslant f_{n_x}$ for all $x \in F$, we obtain $0 \leqslant f_m \leqslant \varepsilon$ and thus $\|f_n\|_\infty \leqslant \varepsilon$ for all $n \geqslant m$. □

Proof of Proposition G.2. From Lemma G.3 and the positivity of η we conclude that $0 \leqslant \eta(f_n) \leqslant \|f_n\|_\infty \eta(\mathbf{1}_K) \to 0$ as $n \to \infty$. □

G.2 The upper integral and $\mathscr{L}_1(K, \eta)$

Recall that η is supposed to be a positive linear functional on $C(K)$. Given a function $f\colon K \to [0,\infty)$, we call a sequence (φ_n) a **hull sequence for** f if (φ_n) is a monotone increasing sequence in $C(K)_+$ and $\lim_{n\to\infty} \varphi_n(x) \geqslant f(x)$ for all $x \in K$. We define the upper integral of f by

$$\eta^*(f) := \inf\Big\{\lim_{n\to\infty} \eta(\varphi_n);\ (\varphi_n) \text{ a hull sequence for } f\Big\} \in [0,\infty].$$

We mention that in these definitions the limits of the increasing sequences $(\eta(\varphi_n))_n$ and $(\varphi_n(x))_n$ $(x \in K)$ are allowed to be ∞.

G.4 Proposition. (a) *If $f, g\colon K \to [0,\infty)$ satisfy $f \leqslant g$, then*

$$\eta^*(f) \leqslant \eta^*(g).$$

(b) *For all $f\colon K \to [0,\infty)$, $\lambda \geqslant 0$ one has*

$$\eta^*(\lambda f) = \lambda \eta^*(f).$$

(c) *For all $f, g\colon K \to [0,\infty)$ one has*

$$\eta^*(f+g) \leqslant \eta^*(f) + \eta^*(g).$$

Proof. (a) follows directly from the definition.

(b) is obvious for $\lambda = 0$. In the case $\lambda > 0$, (φ_n) is a hull sequence for f if and only if $(\lambda\varphi_n)$ is a hull sequence for λf. Taking the infimum over all hull sequences for f on both sides of the equality $\lim_{n\to\infty} \eta(\lambda\varphi_n) = \lambda \lim_{n\to\infty} \eta(\varphi_n)$ one obtains the assertion.

(c) Let (φ_n) and (ψ_n) be hull sequences for f and g, respectively. Then $(\varphi_n + \psi_n)$ is a hull sequence for $f + g$, and hence

$$\eta^*(f+g) \leqslant \lim_{n\to\infty} \eta(\varphi_n + \psi_n) = \lim_{n\to\infty} \eta(\varphi_n) + \lim_{n\to\infty} \eta(\psi_n).$$

Using the definition of $\eta^*(f)$ and $\eta^*(g)$ one obtains the asserted inequality. □

Next we show that the subadditivity of η^* from Proposition G.4(c) can be reinforced to the following version of countable subadditivity.

G.5 Proposition. *Let (f_n) be a sequence of functions $f_n\colon K \to [0,\infty)$ that satisfies $\sum_{n=1}^\infty f_n(x) < \infty$ for all $x \in K$. Then*

$$\eta^*\Big(\sum_{n=1}^\infty f_n\Big) \leqslant \sum_{n=1}^\infty \eta^*(f_n).$$

Proof. The inequality is trivial if the right-hand side is ∞. Assume that the right-hand side is less than ∞, and let $\varepsilon > 0$. Then for all $n \in \mathbb{N}$ there exists a hull sequence $(\varphi_{nk})_k$ for f_n such that $\lim_{k\to\infty} \eta(\varphi_{nk}) \leqslant \eta^*(f_n) + 2^{-n}\varepsilon$. We put

$$\psi_k := \sum_{n=1}^k \varphi_{nk} \qquad (k \in \mathbb{N}).$$

One easily checks that (ψ_k) is a hull sequence for $\sum_{n=1}^N f_n$ for each $N \in \mathbb{N}$, and hence for $\sum_{n=1}^\infty f_n$. Moreover

$$\eta(\psi_k) = \sum_{n=1}^k \eta(\varphi_{nk}) \leqslant \sum_{n=1}^k \big(\eta^*(f_n) + 2^{-n}\varepsilon\big) \leqslant \sum_{n=1}^\infty \eta^*(f_n) + \varepsilon$$

for all $k \in \mathbb{N}$. Thus

$$\eta^*\Big(\sum_{n=1}^\infty f_n\Big) \leqslant \lim_{k\to\infty} \eta(\psi_k) \leqslant \sum_{n=1}^\infty \eta^*(f_n) + \varepsilon. \qquad \square$$

G.6 Proposition. *For all $\varphi \in C(K)_+$ one has $\eta^*(\varphi) = \eta(\varphi)$.*

Proof. The constant sequence $(\varphi)_{n\in\mathbb{N}}$ is a hull sequence for φ, and this implies the inequality $\eta^*(\varphi) \leqslant \eta(\varphi)$.

If (φ_n) is a hull sequence for φ, then $(\varphi_n \wedge \varphi)$ is a hull sequence for φ as well. Moreover $(\varphi - \varphi_n \wedge \varphi)_{n\in\mathbb{N}}$ is decreasing, with $\varphi - \varphi_n \wedge \varphi \to 0$ pointwise as $n \to \infty$; hence by Proposition G.2 we obtain

$$\eta(\varphi) = \lim_{n\to\infty} \eta(\varphi_n \wedge \varphi) \leqslant \lim_{n\to\infty} \eta(\varphi_n).$$

This shows that $\eta(\varphi) \leqslant \eta^*(\varphi)$. $\square$

We now define $q\colon \mathbb{R}^K \to [0,\infty]$ by

$$q(f) := \eta^*(|f|) \qquad (f \in \mathbb{R}^K),$$

where $\mathbb{R}^K$ is the vector space of all function $f\colon K \to \mathbb{R}$. Then Proposition G.4 implies that $\mathscr{L} := \big\{f \in \mathbb{R}^K;\ q(f) < \infty\big\}$ is a subspace of $\mathbb{R}^K$ and that the restriction of q to $\mathscr{L}$ is a semi-norm. We define the space of ***η*-integrable functions** $\mathscr{L}_1(K,\eta)$ as the closure of $C(K)$ in the semi-normed space $(\mathscr{L}, q)$. It follows from Proposition G.6 that $|\eta(f)| = |\eta(f^+) - \eta(f^-)| \leqslant \eta(|f|) = \eta^*(|f|) = q(f)$ for all $f \in C(K)$. Therefore η has a unique continuous extension to $\mathscr{L}_1(K,\eta)$, which we will still denote by η.

In the following theorem we collect properties that will be needed in order to define the measure μ on a suitable σ-algebra. In particular, the version of the monotone convergence theorem stated in part (c) will be responsible for the σ-additivity of the measure.

G.7 Theorem. (a) *The functional* $\eta\colon \mathscr{L}_1(K,\eta) \to \mathbb{R}$ *is positive, and* $\eta(f) = \eta^*(f)$ *for all* $f \in \mathscr{L}_1(K,\eta)_+$.

If $f, g \in \mathscr{L}_1(K,\eta)$, *then* $|f|, f \vee g, f \wedge g \in \mathscr{L}_1(K,\eta)$.

(b) *Let* (f_n) *be a monotone increasing sequence in* $\mathscr{L}_1(K,\eta)_+$ *and* $f\colon K \to [0,\infty)$ *such that* $f_n \to f$ *pointwise,* $\sup_{n\in\mathbb{N}} \eta(f_n) < \infty$. *Then* $f \in \mathscr{L}_1(K,\eta)$ *and* $\eta(f) = \lim_{n\to\infty} \eta(f_n)$.

Proof. (a) Let $f \in \mathscr{L}_1(K,\eta)$. There exists a sequence (φ_n) in $C(K)$ converging to f in $\mathscr{L}_1(K,\eta)$. Then by Proposition G.4(a) one sees that

$$q(|f| - |\varphi_n|) = \eta^*\big(\big||f| - |\varphi_n|\big|\big) \leqslant \eta^*(|f - \varphi_n|) = q(f - \varphi_n) \to 0,$$

which shows that $|f| \in \mathscr{L}_1(K,\eta)$ and $|\varphi_n| \to |f|$ in $\mathscr{L}_1(K,\eta)$ as $n \to \infty$. Proposition G.6 implies that $\eta(|\varphi_n|) = q(\varphi_n)$ for all $n \in \mathbb{N}$, and letting $n \to \infty$ one obtains $\eta(|f|) = q(f)$, by the continuity of η and q. In particular, if $f \geqslant 0$, then $\eta(f) = q(f) = \eta^*(f) \geqslant 0$.

The last two assertions of part (a) then follow from $f \vee g = \frac{1}{2}(f + g + |f - g|)$ and $f \wedge g = \frac{1}{2}(f + g - |f - g|)$.

(b) We supplement the sequence (f_n) by $f_0 := 0$. Then by part (a) we obtain

$$\sum_{j=0}^{N} \eta^*(f_{j+1} - f_j) = \sum_{j=0}^{N} \eta(f_{j+1} - f_j) = \eta(f_{N+1}) \leqslant \sup_{n\in\mathbb{N}} \eta(f_n) < \infty \qquad (N \in \mathbb{N}),$$

which implies that the series $\sum_{j=0}^{\infty} \eta^*(f_{j+1} - f_j)$ is convergent. Hence by the countable subadditivity of η^*, see Proposition G.5, we infer that

$$\eta^*(f - f_n) = \eta^*\Big(\sum_{j=n}^{\infty} (f_{j+1} - f_j)\Big) \leqslant \sum_{j=n}^{\infty} \eta^*(f_{j+1} - f_j) \to 0 \qquad (n \to \infty).$$

This establishes that $f \in \mathscr{L}_1(K,\eta)$ and $\eta(f_n) \to \eta(f)$ $(n \to \infty)$, as asserted. □

G.3 The measure μ and the proof of the Riesz–Markov theorem

We define the collection $\mathcal{A}$ of ***η*-measurable sets**,

$$\mathcal{A} := \{A \subseteq K\,;\, \mathbf{1}_A \in \mathscr{L}_1(K,\eta)\},$$

and for $A \in \mathcal{A}$ we define $\mu(A) := \eta(\mathbf{1}_A)$.

G.8 Proposition. *The collection* $\mathcal{A}$ *is a* σ*-algebra, and* μ *is a measure on* $\mathcal{A}$.

Proof. From $\mathbf{1}_\varnothing = 0 \in \mathscr{L}_1(K,\eta)$ we obtain $\varnothing \in \mathcal{A}$ and $\mu(\varnothing) = \eta(0) = 0$. If $A \in \mathcal{A}$, then $\mathbf{1}_{K\setminus A} = \mathbf{1}_K - \mathbf{1}_A \in \mathscr{L}_1(K,\eta)$, hence $K \setminus A \in \mathcal{A}$.

If $A, B \in \mathcal{A}$, then $\mathbf{1}_{A\cup B} = \mathbf{1}_A \vee \mathbf{1}_B \in \mathscr{L}_1(K,\eta)$ by Theorem G.7(a), and hence $A \cup B \in \mathcal{A}$. If additionally $A \cap B = \varnothing$, then $\mu(A \cup B) = \eta(\mathbf{1}_{A\cup B}) = \eta(\mathbf{1}_A + \mathbf{1}_B) = \mu(A) + \mu(B)$.

Finally, if (A_n) is an increasing sequence in $\mathcal{A}$, $A := \bigcup_n A_n$, then $\mathbf{1}_{A_n} \to \mathbf{1}_A$ pointwise and $\sup_n \eta(\mathbf{1}_{A_n}) \leqslant \eta(\mathbf{1}_K) < \infty$. Therefore, Theorem G.7(b) implies that $\mathbf{1}_A \in \mathscr{L}_1(K,\eta)$ and $\mu(A) = \eta(\mathbf{1}_A) = \lim_{n\to\infty} \eta(\mathbf{1}_{A_n}) = \lim_{n\to\infty} \mu(A_n)$.

These properties show that $\mathcal{A}$ is a σ-algebra and that μ is a measure on $\mathcal{A}$. □

It turns out that every $f \in C(K)$ is $\mathcal{A}$-measurable and that $\eta(f) = \int f \,\mathrm{d}\mu$ for all $f \in C(K)$; in fact this equality is even true for all $f \in \mathscr{L}_1(K,\eta)$, and $\mathscr{L}_1(K,\eta) = \mathscr{L}_1(\mu)$; see Exercise G.1(a). But these properties do not imply the Riesz–Markov theorem since μ is not a Borel measure in general.

Proof of Theorem G.1. For the existence part let $\mathcal{A}$ and μ be as defined above, and let $\mathcal{B}$ be the Borel σ-algebra of K. We first prove that $\mathcal{B} \subseteq \mathcal{A}$. (Only here we need K to be a *metric* space.) Since $\mathcal{A}$ is a σ-algebra, if suffices to show that $U \in \mathcal{A}$ for all open sets $U \subsetneq K$. Putting

$$f(x) := \operatorname{dist}(x, K \setminus U) \qquad (x \in K),$$

the distance of x to $K \setminus U$, we have a function $f \in C(K)$ satisfying $U = [f > 0]$. Then $f_n := (nf) \wedge 1$ defines a monotone increasing sequence (f_n) in $C(K)_+$ converging to $\mathbf{1}_U$ pointwise. Since $\sup_n \eta(f_n) \leqslant \eta(\mathbf{1}_K) < \infty$, Theorem G.7(b) implies that $\mathbf{1}_U \in \mathscr{L}_1(K,\eta)$, $U \in \mathcal{A}$.

Clearly $\mu_0 := \mu|_{\mathcal{B}}$ is a Borel measure, and for $B \in \mathcal{B} \subseteq \mathcal{A}$ we have

$$\eta(\mathbf{1}_B) = \mu_0(B) = \int \mathbf{1}_B \,\mathrm{d}\mu_0 \,.$$

By linearity of η and the integral, the equality $\eta(f) = \int f \,\mathrm{d}\mu_0$ extends to all functions $f \in \mathcal{S}(\mathcal{B}) := \operatorname{lin}\{\mathbf{1}_B ;\, B \in \mathcal{B}\}$ (the simple functions over $\mathcal{B}$). Now let $f \in C(K)$. We put $f_n(x) := \frac{1}{n}\lfloor nf(x)\rfloor$ for all $n \in \mathbb{N}$, $x \in K$, where $\lfloor t\rfloor$ denotes the integer part of $t \in \mathbb{R}$. Then $f_n \in \mathcal{S}(\mathcal{B})$ since f_n is a composition of Borel measurable functions and the range of f_n is finite, for each $n \in \mathbb{N}$. Moreover $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Using Theorem G.7(a) we conclude that $|\eta(f) - \eta(f_n)| \leqslant \|f - f_n\|_\infty \eta(\mathbf{1}_K) \to 0$ as $n \to \infty$, and also $|\int f \,\mathrm{d}\mu_0 - \int f_n \,\mathrm{d}\mu_0| \leqslant \|f - f_n\|_\infty \mu_0(K) \to 0$ as $n \to \infty$. Thus, taking the limit $n \to \infty$ in $\eta(f_n) = \int f_n \,\mathrm{d}\mu_0$ we obtain $\eta(f) = \int f \,\mathrm{d}\mu_0$.

For the uniqueness part let μ and ν be two Borel measures representing η, i.e. $\eta(f) = \int f \,\mathrm{d}\mu = \int f \,\mathrm{d}\nu$ for all $f \in C(K)$. Then $\mu + \nu$ is a Borel measure. Let $B \in \mathcal{B}$; we need to show that $\mu(B) = \nu(B)$. Note that $\mathbf{1}_B \in L_1(\mu + \nu)$ since $\mu + \nu$ is a finite measure. By Theorem G.9 below there exists a sequence (f_n) in $C(K)$ converging to $\mathbf{1}_B$ in $L_1(\mu + \nu)$. Then also $\|f_n - \mathbf{1}_B\|_{L_1(\mu)} \leqslant \|f_n - \mathbf{1}_B\|_{L_1(\mu+\nu)} \to 0$, and in the same way $\|f_n - \mathbf{1}_B\|_{L_1(\nu)} \to 0$ as $n \to \infty$. Thus

$$\mu(B) = \int \mathbf{1}_B \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \,\mathrm{d}\nu = \int \mathbf{1}_B \,\mathrm{d}\nu = \nu(B). \qquad \square$$

G.9 Theorem. *Let K be a compact metric space, and let μ be a Borel measure on K. Let $1 \leqslant p < \infty$. Then $C(K)$ is dense in $L_p(\mu)$.*

Proof. Exercise 10.3(a) implies that the closure X of $C(K)$ in $L_p(\mu)$ is a vector sublattice of $L_p(\mu)$. Thus, by Exercise 10.3(b), $\mathcal{A} := \{A;\, \mathbf{1}_A \in X\}$ is a σ-algebra. If $U \subsetneq K$ is an open set, then as in the above proof of Theorem G.1 we find a monotone increasing sequence (f_n) in $C(K)_+$ converging to $\mathbf{1}_U$ pointwise, and hence in $L_p(\mu)$ by the dominated convergence theorem, which implies $\mathbf{1}_U \in X$, $U \in \mathcal{A}$. It follows that $\mathcal{A}$ contains the Borel σ-algebra $\mathcal{B}$, i.e. $\mathbf{1}_B \in X$ for all $B \in \mathcal{B}$. Since $\mathcal{S}(\mathcal{B}) = \operatorname{lin}\{\mathbf{1}_B ;\, B \in \mathcal{B}\}$ is dense in $L_p(\mu)$, we conclude that $X = L_p(\mu)$. $\square$

G.10 Remarks. (a) A more standard proof of the uniqueness in Theorem G.1 is as follows. One easily sees that $\mu(U) = \nu(U)$ for all open sets U, using the approximation of $\mathbf{1}_U$ by the sequence (f_n) from the first paragraph of the proof of Theorem G.1. Since the open sets form a $\cap$-stable generator of the Borel σ-algebra and μ, ν are finite, the uniqueness theorem from measure theory, see e.g. [Bau01; Theorem 5.4], implies that $\mu = \nu$. Our argument above avoids the use of the uniqueness theorem.

(b) If K is just a compact topological space (and not a metric space), then the assertions of Theorems G.1 and G.9 remain valid, with one modification: one has to replace the Borel σ-algebra by the smaller **Baire $\boldsymbol{\sigma}$-algebra** $\mathcal{B}_0$, the σ-algebra generated by the collection

$$\big\{[f > 0];\ f \in C(K)\big\}.$$

Then in Theorem G.1 one obtains a unique **Baire measure** on K, i.e. a measure on $\mathcal{B}_0$ satisfying $\mu(K) < \infty$, and in Theorem G.9 one has to assume that μ is a Baire measure.

We note that $\mathcal{B}_0$ is always contained in $\mathcal{A}$, and that $\mathcal{B}_0 = \mathcal{B}$ if K is a metric space. The reader is asked to carry out the details in Exercises G.1(b) and G.2. △

Notes

The Riesz–Markov representation theorem, often simply called the 'Riesz representation theorem', is a classical result going back to F. Riesz [Rie09], Banach 1937 (see footnote 1 in [Sak38]), Saks [Sak38], Markov [Mar38] and Kakutani [Kak41; Theorem 9] in versions of increasing generality. For the case of compact metric spaces we refer to [Sak38] for a particularly nice and simple proof based on Carathéodory's extension theorem. We have been at a loss finding a concise proof just containing the bare facts. This is why we decided to include a proof for the mininal version we need in the proof of the spectral theorem; see Appendix F. Our proof is a reduced version of the Daniell integral: we first extend the given positive linear functional η to a larger class of functions, but then develop only part of the integration theory along this line.

We point out that our proof – when applied to the case of general compact topological spaces – only yields a Baire measure, which for compact *metric* spaces is automatically a Borel measure. If one modifies the extension of the functional η by admitting arbitrary directed sets as 'increasing hull nets' instead of increasing sequences in the space of continuous functions, then one obtains a Borel measure with the desired properties for the case of non-metric Hausdorff spaces as well. This is the procedure presented by Bourbaki [Bou65; Chap. IV, §1], for Hausdorff locally compact spaces. A discussion comparing the methods and the respective results can be found in [Flo81; Anhang]. A discussion concerning the set of different Borel measures leading to the desired representation for continuous functions can be found in [Bau01; §§28 and 29].

Exercises

G.1 Let K be a compact topological space, let $\eta\colon C(K) \to \mathbb{K}$ be a positive linear functional, and let $\mathcal{A}$ and μ be defined as in Section G.3.

(a) Show that $\mathscr{L}_1(K,\eta) = \mathscr{L}_1(\mu)$ (in particular each $f \in C(K)$ is $\mathcal{A}$-measurable) and that $\eta(f) = \int f \,\mathrm{d}\mu$ for all $f \in \mathscr{L}_1(K,\eta)$.

Hint: Note that from the definition of $\mathcal{A}$ it follows that $\mathcal{A} = \{A \subseteq K\,;\, \mathbf{1}_A \in \mathscr{L}_1(K,\eta)\} = \{A \subseteq K\,;\, \mathbf{1}_A \in \mathscr{L}_1(\mu)\}$. Show that both $\mathscr{L}_1(K,\eta)$ and $\mathscr{L}_1(\mu)$ are vector lattices with the property that each positive function can be approximated pointwise by a monotone increasing sequence in $S(\mathcal{A}) = \mathrm{lin}\{\mathbf{1}_A\,;\, A \in \mathcal{A}\}$; use an approximation procedure similar to the one in the proof of Theorem G.1.

(b) Let $\mathcal{B}_0$ be the Baire σ-algebra of K. Show that $\mathcal{B}_0 \subseteq \mathcal{A}$ and that $\mu|_{\mathcal{B}_0}$ is the unique Baire measure on K representing η. Moreover, show that $\mathcal{B}_0$ equals the Borel σ-algebra of K if K is a metric space.

Hint: In order to show that $[f > 0] \in \mathcal{A}$ for all $f \in C(K)$, consider the sequence (f_n) defined by $f_n := \big(nf \wedge 1\big) \vee 0$ and proceed similarly as in the proof of Theorem G.1.

G.2 Let K be a compact topological space, let μ be a Baire measure on K, and let $1 \leqslant p < \infty$. Show that $C(K)$ is dense in $L_p(\mu)$. (Hint: Mimic the proof of Theorem G.9, using the sequence (f_n) given in the hint to Exercise G.1(b).)

G.3 Let Ω be a locally compact metric space, let μ be a Borel measure on Ω, and let $1 \leqslant p < \infty$. Assume that each Borel set $B \subseteq \Omega$ of finite measure can be covered by countably many compact sets. Show that $C_{\mathrm{c}}(\Omega)$ is dense in $L_p(\mu)$. (Hint: Exercise 10.3(d) and Theorem G.9.)

Note. The above hypothesis holds for any open subset Ω of $\mathbb{R}^n$; see Exercise 4.1(a).

Appendix H

Singular integrals and the Bogovskiĭ operator

The issue of this appendix is the proof of Theorem 16.10. The most important tool and key result is Theorem H.6, concerning convolution operators with singular kernels.

H.1 The Fourier transformation

For $f \in L_1(\mathbb{R}^n)$ we define the **Fourier transform** of f,

$$Ff(\xi) = \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x \qquad (\xi \in \mathbb{R}^n).$$

Then $\hat{f}\colon \mathbb{R}^n \to \mathbb{C}$ is continuous (by the dominated convergence theorem) and bounded, $\|\hat{f}\|_\infty \leqslant \frac{1}{(2\pi)^{n/2}}\|f\|_1$, and the **Fourier transformation** $F\colon L_1(\mathbb{R}^n) \to C_b(\mathbb{R}^n)$ is a bounded linear operator. More strongly, the range of F is a subset of $C_0(\mathbb{R}^n)$.

We will need the following important facts concerning the Fourier transformation; see [Duo01; Chapter 1] or [Yos68; Chapter VI], for instance.

H.1 Remarks. (a) For $f \in L_1 \cap L_2(\mathbb{R}^n)$ one has $\hat{f} \in L_2(\mathbb{R}^n)$, $\|\hat{f}\|_2 = \|f\|_2$, and F extends from $L_1 \cap L_2(\mathbb{R}^n)$ to a unitary operator $F_2\colon L_2(\mathbb{R}^n) \to L_2(\mathbb{R}^n)$.

(b) If $f, g \in L_1(\mathbb{R}^n)$, then $F(f * g) = (2\pi)^{n/2}\hat{f}\hat{g}$, and the same same formula holds if $f \in L_1(\mathbb{R}^n)$, $g \in L_2(\mathbb{R}^n)$. (For the convolution $f * g$ we refer to Proposition H.2 below.)

(c) For all $f \in C_\mathrm{c}^\infty(\mathbb{R}^n)$, $\alpha \in \mathbb{N}_0^n$ one has

$$\widehat{\partial^\alpha f}(\xi) = (-\mathrm{i}\xi)^\alpha \hat{f}(\xi) \qquad (\xi \in \mathbb{R}^n).$$

(This property makes the Fourier transformation a fundamental tool in the theory of partial differential equations.)

(d) If $f \in L_2(\mathbb{R}^n)$, $j \in \{1, ..., n\}$, and the function $\xi \mapsto \xi_j \hat{f}(\xi)$ belongs to $L_2(\mathbb{R}^n)$, then f has a distributional derivative $\partial_j f$ belonging to $L_2(\mathbb{R}^n)$, and $\|\partial_j f\|_2 = \|\xi \mapsto \xi_j \hat{f}(\xi)\|_2$.

Indeed, for all $\varphi \in C_\mathrm{c}^\infty(\mathbb{R}^n)$ one has

$$\begin{aligned}(f \mid \partial_j\varphi) &= \big(\hat{f} \mid \widehat{\partial_j\varphi}\big) = \big(\hat{f} \mid \xi \mapsto (-\mathrm{i}\xi_j\hat{\varphi}(\xi)\big) = \big(\xi \mapsto \mathrm{i}\xi_j\hat{f}(\xi) \mid \hat{\varphi}\big)\\ &= \big(F_2^{-1}(\xi \mapsto \mathrm{i}\xi_j\hat{f}(\xi)) \mid \varphi\big).\end{aligned}$$

This shows that $-F_2^{-1}(\xi \mapsto \mathrm{i}\xi_j\hat{f}(\xi)) \in L_2(\mathbb{R}^n)$ is the distributional derivative $\partial_j f$. △

H.2 Proposition. *Let $\rho \in L_1(\mathbb{R}^n)$, and let $1 \leqslant p \leqslant \infty$, $f \in L_p(\mathbb{R}^n)$. Then*

$$\rho * f(x) := \int_{\mathbb{R}^n} \rho(x-y) f(y) \, \mathrm{d}y = \int_{\mathbb{R}^n} \rho(y) f(x-y) \, \mathrm{d}y$$

exists for a.e. $x \in \mathbb{R}^n$, and

$$\|\rho * f\|_p \leqslant \|\rho\|_1 \|f\|_p \, .$$

Proof. The function $\mathbb{R}^{2n} \ni (x,y) \mapsto \rho(x-y)u(y)$ is measurable because the mapping $(x,y) \mapsto (x-y,y)$ is a diffeomorphism. Let $1 \leqslant q \leqslant \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. For $g \in L_q(\mathbb{R}^n)$ we estimate, using Fubini–Tonelli and Hölder,

$$\begin{aligned} \iint |\rho(y) f(x-y)| \, \mathrm{d}y \, |g(x)| \, \mathrm{d}x &= \int |\rho(y)| \int |f(x-y)||g(x)| \, \mathrm{d}x \, \mathrm{d}y \\ &\leqslant \int |\rho(y)| \|f(\cdot - y)\|_p \|g\|_q \, \mathrm{d}y = \|\rho\|_1 \|f\|_p \|g\|_q \, . \end{aligned}$$

This inequality implies the assertions since $L_q(\mathbb{R}^n)$ is norming for $L_p(\mathbb{R}^n)$. □

H.2 Singular integrals of convolution type

In this section we discuss convolution operators mapping $L_2(\Omega)$ to $H^1(\Omega)$, for suitable open sets $\Omega \subseteq \mathbb{R}^n$. We will be working with the following version of the convolution of functions.

H.3 Remark. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set, and put $R := \operatorname{diam}(\Omega)$. Considering $L_2(\Omega)$ in the canonical way as a subspace of $L_2(\mathbb{R}^n)$, we define the convolution $k * f$ of $k \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$ and $f \in L_2(\Omega)$ by

$$k * f(x) := \int_\Omega k(x-y) f(y) \, \mathrm{d}y = \int (\mathbf{1}_{B(0,R)} k)(x-y) f(y) \, \mathrm{d}y \qquad (x \in \Omega).$$

In other words, $k * f = (\mathbf{1}_{B(0,R)} k) * f|_\Omega$. Proposition H.2 implies that $k * f$ is defined a.e. on Ω and that $\|k * f\|_{L_2(\Omega)} \leqslant \|k\|_{L_1(B(0,R))} \|f\|_{L_2(\Omega)}$. △

A function $k \colon \mathbb{R}^n \setminus \{0\} \to \mathbb{K}$ is **positively homogeneous of degree** $\alpha \in \mathbb{R}$ if $k(\lambda x) = \lambda^\alpha k(x)$ for all $x \in \mathbb{R}^n \setminus \{0\}$, $\lambda > 0$. Equivalently, one could formulate that, with $k_0 := k|_{S_{n-1}} \colon S_{n-1} \to \mathbb{K}$, one has $k(x) = |x|^\alpha k_0\big(\frac{x}{|x|}\big)$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Our first result, Theorem H.5 below, is concerned with a non-singular case that can be treated by using classical convolution inequalities. For the proof we will need the following observation.

H.4 Remark. Let $k \in C^1(\mathbb{R}^n \setminus \{0\})$ be positively homogeneous of degree $\alpha > 1 - n$. Then for all $R > 0$ the function k belongs to $W_1^1(B(0,R))$. The proof of this statement is delegated to Exercise H.3. For the definition of the Sobolev space $W_1^1(\Omega)$ we refer to Section D.1. △

H.5 Theorem. *Let $k \in C^1(\mathbb{R}^n \setminus \{0\})$ be positively homogeneous of degree $\alpha > 1 - n$. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Then the convolution $f \mapsto k * f$ defines a bounded linear operator $T\colon L_2(\Omega) \to H^1(\Omega)$, and there exists a constant $c_{\alpha,R} > 0$, only depending on α and $R := \operatorname{diam}(\Omega)$ (and the dimension n), such that*

$$\|T\| \leqslant c_{\alpha,R} \max\bigl\{\bigl\|k|_{S_{n-1}}\bigr\|_\infty, \bigl\|(\nabla k)|_{S_{n-1}}\bigr\|_\infty\bigr\}. \tag{H.1}$$

Proof. Recall Remark H.3 for the definition of the convolution $k * f$ and the estimate $\|k * f\|_{L_2(\Omega)} \leqslant \|k\|_{L_1(B(0,R))}\|f\|_{L_2(\Omega)}$, for $f \in L_2(\Omega)$.

Now let $f \in C_c^\infty(\Omega)$. The gradient $\nabla(k * f)$ on Ω is given by

$$\nabla(k * f)(x) = k * \nabla f(x) = \int_\Omega k(x-y)\nabla f(y)\,\mathrm{d}y = \int (\mathbf{1}_{B(0,R)}\nabla k)(x-y) f(y)\,\mathrm{d}y,$$

where in the last step we have used Remark H.4. From

$$\Bigl|\int \mathbf{1}_{B(0,R)}\nabla k(x-y)f(y)\,\mathrm{d}y\Bigr| \leqslant \int \bigl(\mathbf{1}_{B(0,R)}|\nabla k|\bigr)(x-y)|f(y)|\,\mathrm{d}y$$

it follows that

$$\|\nabla k * f\|_{L_2(\Omega;\mathbb{K}^n)} \leqslant \bigl\|\bigl(\mathbf{1}_{B(0,R)}|\nabla k|\bigr) * |f|\bigr\|_{L_2(\Omega)} \leqslant \bigl\||\nabla k|\bigr\|_{L_1(B(0,R))}\|f\|_{L_2(\Omega)}.$$

Now, $\|k\|_{L_1(B(0,R))}$ can be estimated by

$$\begin{aligned}\|k\|_{L_1(B(0,R))} &\leqslant \int_{B(0,R)} |x|^\alpha\,\mathrm{d}x\,\|k|_{S_{n-1}}\|_\infty = \sigma_{n-1}\int_0^R r^{n+\alpha-1}\,\mathrm{d}r\,\|k|_{S_{n-1}}\|_\infty \\ &= \sigma_{n-1}\frac{1}{n+\alpha}R^{n+\alpha}\|k|_{S_{n-1}}\|_\infty\,,\end{aligned}$$

where σ_{n-1} denotes the $(n-1)$-dimensional volume of S_{n-1}. From Exercise H.2 we know that ∇k is positively homogeneous of degree $\alpha - 1$; hence as above we obtain

$$\bigl\||\nabla k|\bigr\|_{L_1(B(0,R))} \leqslant \sigma_{n-1}\frac{1}{n+\alpha-1}R^{n+\alpha-1}\bigl\|(\nabla k)|_{S_{n-1}}\bigr\|_\infty\,.$$

Combining the above estimates we conclude that

$$\|k * f\|_{H^1(\Omega)} \leqslant c_{\alpha,R}\max\bigl\{\bigl\|k|_{S_{n-1}}\bigr\|_\infty, \bigl\|(\nabla k)|_{S_{n-1}}\bigr\|_\infty\bigr\}\|f\|_{L_2(\Omega)}\,,$$

with

$$c_{\alpha,R} := \sigma_{n-1}\Bigl(\frac{1}{n+\alpha}R^{n+\alpha} + \frac{1}{n+\alpha-1}R^{n+\alpha-1}\Bigr).$$

As $C_c^\infty(\Omega)$ is dense in $L_2(\Omega)$ we obtain the asserted properties; see Exercise H.4. □

The following main result of the present section is the version of Theorem H.5 for $\alpha = 1 - n$. The touchy issue in this result is that the derivatives of the function k have a non-integrable singularity at the point 0. As a consequence, the convolution with $\partial_j k$ is not easily interpreted as a mapping on L_2.

H.6 Theorem. *Let $k \in C^1(\mathbb{R}^n \setminus \{0\})$ be positively homogeneous of degree $1-n$, and let $R > 0$.*

Then $k_R := k\mathbf{1}_{B(0,R)} \in L_1(\mathbb{R}^n)$, $k_R * f \in H^1(\mathbb{R}^n)$ for all $f \in L_2(\mathbb{R}^n)$, and there exists $M_R \geqslant 0$ such that

$$\|k_R * f\|_{H^1(\mathbb{R}^n)} \leqslant M_R \|f\|_{L_2(\mathbb{R}^n)} \qquad (f \in L_2(\mathbb{R}^n)). \tag{H.2}$$

More explicitly, there exist constants b and b', only depending on n, such that the estimate holds with

$$M_R := (\sigma_{n-1}R + b)\|k|_{S_{n-1}}\|_\infty + b'\|(\nabla k)|_{S_{n-1}}\|_\infty .$$

If $\Omega \subseteq \mathbb{R}^n$ is a bounded open set, $R := \operatorname{diam}(\Omega)$, then

$$\|k * f\|_{H^1(\Omega)} \leqslant M_R \|f\|_{L_2(\Omega)} \qquad (f \in L_2(\Omega)). \tag{H.3}$$

In the proof we will need the inequality

$$\Big|\int_a^b \frac{\mathrm{e}^{-\mathrm{i}r}}{r}\,\mathrm{d}r\Big| \leqslant \ln\frac{1}{a} + 2 \qquad (0 < a < 1,\ b > a). \tag{H.4}$$

Indeed, if $b \leqslant 1$ then $\big|\int_a^b \frac{\mathrm{e}^{-\mathrm{i}r}}{r}\,\mathrm{d}r\big| \leqslant \int_a^1 \frac{1}{r}\,\mathrm{d}r = \ln\frac{1}{a}$. If $b > 1$, then we split the integral, estimate $\big|\int_a^1 \frac{\mathrm{e}^{-\mathrm{i}r}}{r}\,\mathrm{d}r\big| \leqslant \ln\frac{1}{a}$ and use integration by parts to obtain

$$\Big|\int_1^b \frac{\mathrm{e}^{-\mathrm{i}r}}{r}\,\mathrm{d}r\Big| \leqslant \Big|\frac{\mathrm{i}\mathrm{e}^{-\mathrm{i}r}}{r}\Big|_1^b - \int_1^b \frac{-\mathrm{i}\mathrm{e}^{-\mathrm{i}r}}{r^2}\,\mathrm{d}r\Big| \leqslant \frac{1}{b} + 1 + 1 - \frac{1}{b} = 2.$$

Proof of Theorem H.6. (i) By the hypotheses there exists $c \geqslant 0$ such that $|k(x)| \leqslant c|x|^{1-n}$ for all $x \in \mathbb{R}^n \setminus \{0\}$. This implies that $k_R \in L_1(\mathbb{R}^n)$, $\|k_R\|_1 \leqslant \sigma_{n-1}R\|k|_{S_{n-1}}\|_\infty$ (see the proof of Theorem H.5), and from Proposition H.2 it follows that $k_R * f \in L_2(\mathbb{R}^n)$, $\|k_R * f\|_2 \leqslant \sigma_{n-1}R\|k|_{S_{n-1}}\|_\infty\|f\|_2$ for all $f \in L_2(\mathbb{R}^n)$.

(ii) In order to estimate the derivatives of $k_R * f$, we show that there exist constants b, $b' > 0$, depending only on n, such that

$$(2\pi)^{n/2}|\xi\,\widehat{k_R}(\xi)| \leqslant b\|k|_{S_{n-1}}\|_\infty + b'\|(\nabla k)|_{S_{n-1}}\|_\infty \qquad (\xi \in \mathbb{R}^n). \tag{H.5}$$

Observe that the estimate (H.5) is invariant under orthogonal transformations of $\mathbb{R}^n$. (Note that $\hat{f}(A\cdot) = \widehat{f(A\cdot)}$ for all $f \in L_1(\mathbb{R}^n)$ and all orthogonal matrices.) Thus, it suffices to prove (H.5) for $\xi = te_1$, with $t > 0$ and the first unit vector e_1.

By the positive $(1-n)$-homogeneity of k we obtain

$$(2\pi)^{n/2}t\,\widehat{k_R}(te_1) = \int_{B(0,R)} t\mathrm{e}^{-\mathrm{i}te_1\cdot x}k(x)\,\mathrm{d}x = \int_{B(0,tR)} \mathrm{e}^{-\mathrm{i}x_1}k(x)\,\mathrm{d}x = \lim_{\varepsilon\to 0} h(\varepsilon, tR), \tag{H.6}$$

where

$$h(\varepsilon, s) := \int_{\Omega_{\varepsilon,s}} \mathrm{e}^{-\mathrm{i}x_1}k(x)\,\mathrm{d}x, \quad \Omega_{\varepsilon,s} := B(0,s) \setminus B[0,\varepsilon] \qquad (0 < \varepsilon < s).$$

Let $\alpha \in \{0,1\}$ and $0 < \varepsilon < s$. Applying Gauss' theorem, Theorem 7.3, we compute

$$\begin{aligned} -\mathrm{i}h(\varepsilon,s) &= \int_{\Omega_{\varepsilon,s}} \partial_1(\mathrm{e}^{-\mathrm{i}x_1} - \alpha)\,k(x)\,\mathrm{d}x \\ &= \int_{\partial\Omega_{\varepsilon,s}} \nu_1(x)(\mathrm{e}^{-\mathrm{i}x_1} - \alpha)\,k(x)\,\mathrm{d}\sigma(x) - \int_{\Omega_{\varepsilon,s}} (\mathrm{e}^{-\mathrm{i}x_1} - \alpha)\,\partial_1 k(x)\,\mathrm{d}x. \end{aligned} \tag{H.7}$$

Since $\nu(x) = \pm x/|x|$ and ρ is positively $(1-n)$-homogeneous, the first integral on the right-hand side equals $\int_{S_{n-1}} y_1\big((\mathrm{e}^{-\mathrm{i}sy_1} - \alpha) - (\mathrm{e}^{-\mathrm{i}\varepsilon y_1} - \alpha)\big)k(y)\,\mathrm{d}\sigma(y)$. Using generalised polar coordinates in $\mathbb{R}^n$ we see that the second integral equals $\int_{S_{n-1}}\int_\varepsilon^s(\mathrm{e}^{-\mathrm{i}ry_1} - \alpha)\,\partial_1 k(ry)r^{n-1}\,\mathrm{d}r\,\mathrm{d}\sigma(y)$. Observing that $\partial_1 k$ is positively homogeneous of degree $-n$ (see Exercise H.2(a)), we conclude that

$$-\mathrm{i}h(\varepsilon,s) = \int_{S_{n-1}} y_1(\mathrm{e}^{-\mathrm{i}sy_1} - \mathrm{e}^{-\mathrm{i}\varepsilon y_1})\,k(y)\,\mathrm{d}\sigma(y) - \int_{S_{n-1}}\int_\varepsilon^s \frac{\mathrm{e}^{-\mathrm{i}ry_1} - \alpha}{r}\,\mathrm{d}r\,\partial_1 k(y)\,\mathrm{d}\sigma(y),$$

and therefore

$$|h(\varepsilon,s)| \leqslant \int_{S_{n-1}} 2|y_1|\,\mathrm{d}\sigma(y)\,\|k|_{S_{n-1}}\|_\infty + \int_{S_{n-1}}\left|\int_\varepsilon^s \frac{\mathrm{e}^{-\mathrm{i}ry_1} - \alpha}{r}\,\mathrm{d}r\right|\mathrm{d}\sigma(y)\,\|(\nabla k)|_{S_{n-1}}\|_\infty\,.$$

In the case $s \leqslant 1$ we choose $\alpha = 1$ and use the estimate $\left|\frac{\mathrm{e}^{-\mathrm{i}ry_1}-1}{r}\right| \leqslant |y_1|$ to obtain

$$|h(\varepsilon,s)| \leqslant \int_{S_{n-1}} 2|y_1|\,\mathrm{d}\sigma(y)\,\|k|_{S_{n-1}}\|_\infty + \int_{S_{n-1}} |y_1|\,\mathrm{d}\sigma(y)\,\|(\nabla k)|_{S_{n-1}}\|_\infty\,. \tag{H.8}$$

In the case $0 < \varepsilon < 1 < s$ we use the decomposition $h(\varepsilon,s) = h(\varepsilon,1) + h(1,s)$. The first term $h(\varepsilon,1)$ is estimated as above, whereas for the second term $h(1,s)$ we choose $\alpha = 0$ and compute

$$\left|\int_1^s \frac{\mathrm{e}^{-\mathrm{i}ry_1}}{r}\,\mathrm{d}r\right| = \left|\int_1^s \frac{\mathrm{e}^{-\mathrm{i}r|y_1|}}{r}\,\mathrm{d}r\right| = \left|\int_{|y_1|}^{s|y_1|} \frac{\mathrm{e}^{-\mathrm{i}r}}{r}\,\mathrm{d}r\right| \leqslant \ln\frac{1}{|y_1|} + 2,$$

where in the last estimate we have used (H.4). Therefore

$$|h(1,s)| \leqslant \int_{S_{n-1}} 2|y_1|\,\mathrm{d}\sigma(y)\,\|k|_{S_{n-1}}\|_\infty + \int_{S_{n-1}}\left(\ln\frac{1}{|y_1|} + 2\right)\mathrm{d}\sigma(y)\,\|(\nabla k)|_{S_{n-1}}\|_\infty\,.$$

(We refer to Exercise H.6 for the finiteness $\int_{S_{n-1}} \ln\frac{1}{|y_1|}\,\mathrm{d}\sigma(y) < \infty$.) Combining this estimate with (H.8) for $s = 1$ or using (H.8) by itself we conclude that

$$|h(\varepsilon,s)| \leqslant b\|k|_{S_{n-1}}\|_\infty + b'\|(\nabla k)|_{S_{n-1}}\|_\infty$$

for $\varepsilon < \min\{s,1\}$, with $b := \int_{S_{n-1}} 4|y_1|\,\mathrm{d}\sigma(y)$ and $b' := \int_{S_{n-1}}\big(|y_1| + \ln\frac{1}{|y_1|} + 2\big)\,\mathrm{d}\sigma(y)$. Now applying (H.6) we obtain (H.5).

(iii) We now prove the estimate on the derivatives of $k_R * f$. Using Remarks H.1(d), (b), (a) and inequality (H.5) we obtain

$$\begin{aligned}\|\nabla(k_R * f)\|_2 &= \big\|\xi \mapsto \xi\widehat{k_R * f}(\xi)\big\|_2 = \big\|\xi \mapsto \xi(2\pi)^{n/2}\widehat{k_R}(\xi)\hat{f}(\xi)\big\|_2 \\ &\leqslant \big\|\xi \mapsto (2\pi)^{n/2}\xi\widehat{k_R}(\xi)\big\|_\infty \|\hat{f}\|_2 \\ &\leqslant \big(b\|k|_{S_{n-1}}\|_\infty + b'\|(\nabla k)|_{S_{n-1}}\|_\infty\big)\|f\|_2\end{aligned}$$

for all $f \in L_2(\mathbb{R}^n)$.

Combining this estimate with the estimate for $\|k_R * f\|_2$ presented in step (i) of the proof we obtain (H.2).

(iv) For the last statement of the theorem we recall from Remark H.3 that $k * f = k_R * f$ on Ω, so (H.3) is immediate from (H.2). □

H.7 Remark. In Exercise H.2(b) the reader is asked to show that in the situation of Theorem H.6 one has $\int_{S_{n-1}} \partial_j k(y)\,\mathrm{d}\sigma(y) = 0$ for all $j \in \{1,\dots,n\}$. Thus it turns out that in formula (H.7) of the proof of Theorem H.6 there is no need to introduce the parameter α. Proving (H.7) without α is sufficient, because in the last term of the formulas one can always add it, and this is what is needed in a later part of the proof. △

H.3 Proof of Sobolev estimates for the Bogovskiĭ operator

For easy reference we recall the definition of the Bogovskiĭ operator,

$$\begin{aligned}Bf(x) &:= \int f(y)(x-y)\int_1^\infty \rho(y + r(x-y))r^{n-1}\,\mathrm{d}r\,\mathrm{d}y \\ &= \int z\int_0^\infty \rho(x+tz)(t+1)^{n-1}\,\mathrm{d}t\, f(x-z)\,\mathrm{d}z,\end{aligned}$$

for $x \in \mathbb{R}^n$, $f \in C_c^\infty(\mathbb{R}^n)$. We also recall from the hypotheses of Theorem 16.10 that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set, star-shaped with respect to a ball $B(x^0, r_0) \subseteq \Omega$, and that $\rho = \rho_0 * \tilde{\rho}$, where $\operatorname{spt}\rho_0 \subseteq B(0, r_0/2)$, $\operatorname{spt}\tilde{\rho} \subseteq B(x^0, r_0/2)$, $\int \rho_0(x)\,\mathrm{d}x = \int \tilde{\rho}(x)\,\mathrm{d}x = 1$.

Proof of Theorem 16.10. Recall that the operator B maps $C_c^\infty(\Omega)$ to $C_c^\infty(\Omega;\mathbb{K}^n)$. In order to prove the theorem it is sufficient to find a constant $M \geqslant 0$ such that $\|Bf\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant M\|f\|_{L_2(\Omega)}$ for all $f \in C_c^\infty(\Omega)$. We now fix $f \in C_c^\infty(\Omega)$ and proceed in three steps to establish this inequality.

(i) Using the definition of ρ we rewrite Bf, for $x \in \mathbb{R}^n$, as

$$\begin{aligned}Bf(x) &= \int_z z\int_0^\infty \int_y \rho_0(y)\tilde{\rho}(x+tz-y)\,\mathrm{d}y\,(t+1)^{n-1}\,\mathrm{d}t\, f(x-z)\,\mathrm{d}z \\ &= \int_w \rho_0(x-w)\int_z z\int_0^\infty \tilde{\rho}(w+tz)(t+1)^{n-1}\,\mathrm{d}t\, f(x-z)\,\mathrm{d}z\,\mathrm{d}w \\ &= \int_w \rho_0(x-w)\,k_w * f(x)\,\mathrm{d}w,\end{aligned}$$

where ρ_0, $\tilde\rho$ and f are considered as functions on $\mathbb{R}^n$ with value zero outside their supports, and where

$$k_w(z) := z\int_0^\infty \tilde\rho(w+tz)(t+1)^{n-1}\,\mathrm{d}t;$$

recall Remark H.3 for the definition of $k_w * f$. In fact, we are only interested in the values of Bf on Ω. Observing that $\rho_0(x-w)=0$ for all $x\in\Omega$, $w\notin\Omega_0 := \Omega + B(0,r_0/2)$ we obtain

$$Bf(x) = \int_{\Omega_0} \rho_0(x-w)k_w * f(x)\,\mathrm{d}w \qquad (x\in\Omega). \tag{H.9}$$

(ii) In this step we prove an estimate for the H^1-norm of Bf in terms of the H^1-norms of the functions $k_w * f$ in the integral in (H.9).

Put $R := \operatorname{diam}(\Omega)$. Defining

$$k_{w,R} := \mathbf{1}_{B(0,R)}k_w \qquad (w\in\mathbb{R}^n),$$

we obtain

$$k_w * f(x) = \int_\Omega k_w(x-y)f(y)\,\mathrm{d}y = k_{w,R} * f(x) \qquad (w\in\mathbb{R}^n,\ x\in\Omega). \tag{H.10}$$

As in step (ii) of the proof of Theorem 16.7 we rewrite k_w as

$$k_w(z) = \frac{z}{|z|^n}\int_0^\infty \tilde\rho\Big(w + s\frac{z}{|z|}\Big)(s+|z|)^{n-1}\,\mathrm{d}s \tag{H.11}$$

to see that there exists $c\geqslant 0$ such that

$$|k_{w,R}(z)| \leqslant \frac{c}{|z|^{n-1}} \qquad (z\in\mathbb{R}^n,\ w\in\Omega_0);$$

hence $k_{w,R}\in L_1(\mathbb{R}^n;\mathbb{K}^n)$. It is easy to see that, for $|z|\neq 0$, the function $\Omega_0\ni w\mapsto k_{w,R}(z)$ is continuous. (Observe that the domain of the s-integration in (H.11) is contained in the bounded interval $[0,R]$ because $\operatorname{spt}\tilde\rho(w+\cdot) = \operatorname{spt}\tilde\rho - w \subseteq B(x^0-w,r_0/2)\subseteq B(0,R)$.) By the dominated convergence theorem it follows that the mapping $\Omega_0\ni w\mapsto k_{w,R}\in L_1(\mathbb{R}^n;\mathbb{K}^n)$ is bounded and continuous.

From Lemma 4.1 we know that $k_{w,R}*f$ is continuously differentiable and $\partial_j(k_{w,R}*f) = k_{w,R}*\partial_j f$. Proposition H.2 implies that the linear operator ${}^*\!f\colon L_1(\mathbb{R}^n)\to C_\mathrm{b}(\mathbb{R}^n)$, $h\mapsto h*f$ is bounded, and $\|{}^*\!f\|_{\mathcal{L}(L_1,C_\mathrm{b})}\leqslant \|f\|_\infty$. Combining these facts we infer that the function $\Omega_0\ni w\mapsto k_{w,R}*f\in C_\mathrm{b}^1(\mathbb{R}^n;\mathbb{K}^n)$ is bounded and continuous. Taking into account that the function $\Omega_0\ni w\mapsto \rho_0(\cdot - w)\in C_\mathrm{b}^1(\mathbb{R}^n)$ is bounded and continuous, we conclude that the function

$$\Omega_0\ni w\mapsto \rho_0(\cdot-w)(k_{w,R}*f)\in C_\mathrm{b}^1(\mathbb{R}^n;\mathbb{K}^n)$$

is bounded and continuous as well; see Exercise H.7(b).

In order to obtain the estimate we are aiming for, we define the function $F\colon\Omega_0\to H^1(\Omega;\mathbb{K}^n)$,

$$F(w) := \rho_0(\cdot-w)\,k_w * f|_\Omega \qquad (w\in\Omega_0).$$

Then (H.10) implies that $F(w) = \rho_0(\cdot - w)k_{w,R} * f|_\Omega$, and in view of the properties mentioned above we can apply Exercise H.8 to conclude that

$$\begin{aligned}\|Bf\|_{H^1(\Omega;\mathbb{K}^n)} &= \Big\| x \mapsto \int_{\Omega_0} F(w)(x)\,\mathrm{d}w \Big\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant \mathrm{vol}_n(\Omega_0) \sup_{w\in\Omega_0} \|F(w)\|_{H^1(\Omega;\mathbb{K}^n)} \\ &= \mathrm{vol}_n(\Omega_0) \sup_{w\in\Omega_0} \|\rho_0(\cdot - w)k_w * f\|_{H^1(\Omega;\mathbb{K}^n)}\,.\end{aligned}$$

By Exercise H.7(a) we can estimate

$$\|\rho_0(\cdot - w)k_w * f\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant \sqrt{2}\Big(\|\rho_0\|_\infty^2 + \sum_{j=1}^n \|\partial_j\rho_0\|_\infty^2\Big)^{1/2} \|k_w * f\|_{H^1(\Omega;\mathbb{K}^n)}\,.$$

As a result we obtain

$$\|Bf\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant C_R \sup_{w\in\Omega_0} \|k_w * f\|_{H^1(\Omega;\mathbb{K}^n)}\,, \tag{H.12}$$

where $C_R := \mathrm{vol}_n(\Omega_0)\sqrt{2}\big(\|\rho_0\|_\infty^2 + \sum_{j=1}^n \|\partial_j\rho_0\|_\infty^2\big)^{1/2}$.

Establishing a suitable estimate for the supremum in (H.12) will be the issue of the following step (iii) of the proof.

(iii) In order to establish the connection to Section H.2 on singular integrals we write $k_w = \sum_{m=0}^{n-1} k_w^m$, where

$$k_w^m(z) := \binom{n-1}{m} z \int_0^\infty \tilde\rho(w+tz)t^m\,\mathrm{d}t.$$

It is easy to see that k_w^m is positively homogeneous of degree m, for all $m \in \{0,\dots,n-1\}$.

We will use the elementary inequality

$$\|k_w^m * f\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant \sqrt{n} \sup_{l=1,\dots,n} \|(k_w^m)_l * f\|_{H^1(\Omega)}\,, \tag{H.13}$$

where $(k_w^m)_l$ are the components of the vector-valued kernels k_w^m. In order to apply Theorems H.5 and H.6 to the right-hand side of (H.13), we need estimates for the components $(k_w^m)_l$ on S_{n-1}.

Let $l \in \{1,\dots,n\}$. For $m \in \{0,\dots,n-1\}$, $w \in \Omega_0$ we have

$$\big\|(k_w^m)_l|_{S_{n-1}}\big\|_\infty \leqslant c_m := \tbinom{n-1}{m}\|\tilde\rho\|_\infty \tfrac{1}{m+1}R^{m+1}$$

and

$$\begin{aligned}\big\|\nabla(k_w^m)_l|_{S_{n-1}}\big\|_\infty &= \tbinom{n-1}{m} \sup_{z\in S_{n-1}} \Big| \int_0^\infty \tilde\rho(w+tz)t^m\,\mathrm{d}t\, e_l + z_l \int_0^\infty \nabla\tilde\rho(w+tz)t^{m+1}\,\mathrm{d}t \Big| \\ &\leqslant c_m' := \tbinom{n-1}{m}\big(\|\tilde\rho\|_\infty \tfrac{1}{m+1}R^{m+1} + \|\nabla\tilde\rho\|_\infty \tfrac{1}{m+2}R^{m+2}\big).\end{aligned}$$

For $m = 0,\dots,n-2$ we can now apply Theorem H.5 to obtain

$$\|(k_w^m)_l * f\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant c_{m,R}\max\{c_m, c_m'\}\|f\|_{L_2(\Omega)} \qquad (w \in \Omega_0).$$

For $m = n-1$ we apply (H.3) of Theorem H.6 and obtain

$$\big\|(k_w^{n-1})_l * f\big\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant \big((\sigma_{n-1} + b)c_{n-1} + b'c_{n-1}'\big)\|f\|_{L_2(\Omega)} \qquad (w \in \Omega_0).$$

Summing up over $m = 0, ..., n-1$, using (H.13), and applying (H.12) we conclude that

$$\|Bf\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant C_R\sqrt{n}\biggl(\sum_{m=0}^{n-2} c_{m,R}\max\{c_m, c'_m\} + (\sigma_{n-1}+b)c_{n-1} + b'c'_{n-1}\biggr)\|f\|_{L_2(\Omega)}\,,$$

and this completes the proof of the theorem. □

H.4 Comparison with the 'traditional' proof

The 'traditional' way of proving the continuity of the Bogovskiĭ operator – see Theorem 16.10 – is to invoke the theory developed in [CaZy56]; this approach goes back already to Bogovskiĭ's original paper [Bog79]. In our attempt to provide complete information on the topics treated in the present book we realised that we could avoid part of the treatment of singular integrals by considering only the case of the Bogovskiĭ operator in which the function ρ in (16.6) is of the form $\rho = \rho_0 * \tilde{\rho}$ described in Theorem 16.10. In order explain the difference between the traditional way and our proof we briefly sketch part of the theory of convolutions with singular kernels.

Let $k\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{K}$ be measurable and positively homogeneous of degree $-n$, and assume that $k|_{S_{n-1}} \in L_1(S_{n-1})$. Then

$$K\varphi(x) := k * \varphi(x) := \lim_{\varepsilon\to 0+}\int_{|y|>\varepsilon} k(y)\varphi(x-y)\,\mathrm{d}y$$

can exist for all $x \in \mathbb{R}^n$, $\varphi \in C_c^\infty(\mathbb{R}^n)$ only if

$$\int_{S_{n-1}} k(\xi)\,\mathrm{d}\sigma(\xi) = 0; \tag{H.14}$$

see [Duo01; Proposition 4.1]. If k is an odd function (which implies that (H.14) holds), then K acts as a bounded operator on $L_p(\mathbb{R}^n)$, for all $p \in (1,\infty)$; see [Duo01; Corollary 4.8]. If k satisfies (H.14), the odd part of $k|_{S_{n-1}}$ belongs to $L_1(S_{n-1})$ and the even part of $k|_{S_{n-1}}$ belongs to $L_q(S_{n-1})$ for some $q > 1$, then K is bounded on $L_p(\mathbb{R}^n)$ for all $p \in (1,\infty)$; see [Duo01; Theorem 4.12]. (We note that the proof of our Theorem H.6 is closely related to methods of the proof of the results mentioned above. Our restriction to the case $p = 2$ facilitates the use of the Fourier transformation.)

For the application to the Bogovskiĭ operator one needs the more general concept of 'extended convolution operators'. In this case one has a measurable function $k\colon \mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}) \to \mathbb{K}$, $k(x,\cdot)$ positively homogeneous of degree $-n$ for all $x \in \mathbb{R}^n$, and one wants to define

$$K\varphi(x) := \lim_{\varepsilon\to 0+}\int_{|y|>\varepsilon} k(x,y)\varphi(x-y)\,\mathrm{d}y$$

for $x \in \mathbb{R}^n$, $\varphi \in C_c^\infty(\mathbb{R}^n)$. Assume additionally that

$$\sup_{x\in\mathbb{R}^n} |k(x,\cdot)|_{S_{n-1}}| \in L_1(S_{n-1})$$

and that $k(x,\cdot)$ is an odd function for all $x \in \mathbb{R}^n$. Then K as defined above is bounded on $L_p(\mathbb{R}^n)$ for all $p \in (1,\infty)$; see [Duo01; Theorem 4.16]. Alternatively, assume that there exists $q > 1$ such that

$$\sup_{x\in\mathbb{R}^n} \|k(x,\cdot)\|_q < \infty$$

and that $k(x,\cdot)$ satisfies (H.14) for all $x \in \mathbb{R}^n$. Then K is bounded on $L_p(\mathbb{R}^n)$ for all $p \in [q/(q-1),\infty)$; see [CaZy56; Theorem 2].

In the proof of the continuity of the Bogovskiĭ operator (in the case when ρ is not necessarily of the form $\rho = \rho_0 * \tilde{\rho}$) one uses a decomposition as in step (iii) of our proof in Section H.3. Then the critical terms are of the form

$$\begin{aligned} k(x,y) &= \frac{\partial}{\partial y_j}\left(y_l \int_0^\infty \rho(x+ty)t^{n-1}\,\mathrm{d}t \right) \\ &= \delta_{jl}\int_0^\infty \rho(x+ty)t^{n-1}\,\mathrm{d}t + y_l \int_0^\infty \partial_j\rho(x+ty)t^n\,\mathrm{d}t \qquad (j,l = 1,\dots,n), \end{aligned}$$

responsible for the $L_2(\Omega)$-boundedness of the operator mapping a function f to the ∂_j-derivative of the l-th component of Bf. In this kernel, the function $k(x,\cdot)$ is positively homogeneous of degree $-n$, and $k(x,\cdot)$ satisfies (H.14) for all $x \in \mathbb{R}^n$, by Exercise H.2(b). But $k(x,\cdot)$ is not odd, so one has to apply [CaZy56; Theorem 2]. In our more special case in which ρ in the Bogovskiĭ operator is of the form $\rho = \rho_0 * \tilde{\rho}$ we avoid the application of [CaZy56; Theorem 2] and only have to use the version of [Duo01; Theorem 4.12] presented in Theorem H.6.

Notes

The information regarding the Fourier transformation is standard; we only include it for easy reference. The Bogovskiĭ operator appeared first in [Bog79], in connection with the treatment of the Stokes operator, and Bogovskiĭ refers to [CaZy56] for the property that B maps $L_p^0(\Omega)$ to $W^1_{p,0}(\Omega;\mathbb{K})$. A more detailed description of this application of [CaZy56] can be found in [Gal11; Section III.3]. For more comments on singular integrals and the difference between the treatments of the general and our more special Bogovskiĭ operator we refer to Section H.4.

Exercises

H.1 Let $f \in L_2(\mathbb{R}^n)$, $j \in \{1,\dots,n\}$, $\partial_j f \in L_2(\mathbb{R}^n)$. Show that $[\xi \mapsto \xi_j \hat{f}(\xi)] \in L_2(\mathbb{R}^n)$.

Note. Combined with Remark H.1(d) this shows that $\partial_j f \in L_2(\mathbb{R}^n)$ if and only if $[\xi \mapsto \xi_j \hat{f}(\xi)] \in L_2(\mathbb{R}^n)$.

H.2 Let $\alpha \in \mathbb{R}$, and let the function $k \in C^1(\mathbb{R}^n \setminus \{0\})$ be positively homogeneous of degree α. Let $j \in \{1,\dots,n\}$.

(a) Show that $\partial_j k$ is positively homogeneous of degree $\alpha - 1$.

(b) If $\alpha = 1-n$, show that $\int_{S_{n-1}} \partial_j k(y)\,\mathrm{d}\sigma(y) = 0$. Hint: Choose $\zeta \in C_\mathrm{c}^\infty(0,\infty)_+$, and show that $\int_0^\infty \zeta(r) r^{-1}\,\mathrm{d}r \int_{S_{n-1}} \partial_j k(y)\,\mathrm{d}\sigma(y) = \int_0^\infty \zeta(r) \int_{S_{n-1}} \partial_j k(ry) r^{n-1}\,\mathrm{d}\sigma(y)\,\mathrm{d}r = \cdots = -\int_0^\infty \zeta'(r)\,\mathrm{d}r \int_{S_{n-1}} y_j k(y)\,\mathrm{d}\sigma(y) = 0$.

H.3 Give a proof of Remark H.4. Hint: Choose $\psi \in C_\mathrm{c}^\infty(\mathbb{R}^n)$, $\psi = 1$ in a neighboourhood of 0. Show that $k_m := (1-\psi(m\cdot))k \in W_1^1(B(0,R))$ $(m \in \mathbb{N})$, $k_m \to k$ in $W_1^1(B(0,R))$ $(m \to \infty)$. (Observe Exercise H.2.)

H.4 Let X, Y, Z be Banach spaces, $Y \hookrightarrow Z$, $T \in \mathcal{L}(X,Z)$, and suppose that there exist a dense subspace $D \subseteq X$ and a constant $c \geqslant 0$ such that $\|Tx\|_Y \leqslant c\|x\|_X$ for all $x \in D$. Show that $\operatorname{ran}(T) \subseteq Y$, $T \in \mathcal{L}(X,Y)$, $\|T\|_{\mathcal{L}(X,Y)} \leqslant c$.

H.5 Let $k \in W_1^1(\mathbb{R}^n)$. Show that $k * f \in H^1(\mathbb{R}^n)$ for all $f \in L_2(\mathbb{R}^n)$, and $Tf := k * f$ $(f \in L_2(\mathbb{R}^n))$ defines a bounded operator $T\colon L_2(\mathbb{R}^n) \to H^1(\mathbb{R}^n)$, with $\|T\| \leqslant \big(\|k\|_1^2 + \sum_{j=1}^n \|\partial_j k\|_1^2\big)^{1/2} \|f\|_2$. (Hint: Use the denseness of $C_\mathrm{c}^\infty(\mathbb{R}^n)$ in $W_1^1(\mathbb{R}^n)$ to show that $\partial_j(k*f) = (\partial_j k) * f$.)

Note. This property is similar to Theorem H.5.

H.6 (a) Let $g\colon (0,1] \to [0,\infty)$ be a decreasing function, $\int_0^1 g(r)\,\mathrm{d}r < \infty$. Show that $\int_{S_{n-1}} g(|y_1|)\,\mathrm{d}\sigma(y) < \infty$.

Hint: Estimate $\int_{S_{n-1}} g(|y_1|)\,\mathrm{d}\sigma(y) \leqslant n\int_0^1 r^{n-1} \int_{S_{n-1}} g(|ry_1|)\,\mathrm{d}\sigma(y)\,\mathrm{d}r$, use generalised polar coordinates, and estimate the resulting integral over $B(0,1)$ by the integral over $[-1,1]^n$.

(b) Show that $\int_{S_{n-1}} \ln \frac{1}{|y_1|}\,\mathrm{d}\sigma(y) < \infty$.

H.7 Let $m, n \in \mathbb{N}$, and let $\Omega \subseteq \mathbb{R}^n$ be an open set.

(a) Show that $C_\mathrm{b}^1(\Omega) \times H^1(\Omega;\mathbb{K}^m) \ni (\varphi, g) \mapsto \varphi g \in H^1(\Omega;\mathbb{K}^m)$ defines a continuous bilinear mapping,

$$\|\varphi g\|_{H^1(\Omega;\mathbb{K}^n)} \leqslant \sqrt{2}\Big(\|\varphi\|_\infty^2 + \sum_{j=1}^n \|\partial_j\varphi\|_\infty^2\Big)^{1/2} \|g\|_{H^1(\Omega;\mathbb{K}^n)}\,.$$

(b) Show that $C_\mathrm{b}^1(\Omega) \times C_\mathrm{b}^1(\Omega;\mathbb{K}^m) \ni (\varphi, g) \mapsto \varphi g \in C_\mathrm{b}^1(\Omega;\mathbb{K}^m)$ defines a continuous bilinear mapping.

H.8 Let $\Omega \subseteq \mathbb{R}^n$, $\Omega' \subseteq \mathbb{R}^m$ be bounded open sets. Let $F\colon \Omega' \to C_\mathrm{b}^1(\Omega)$ be bounded and continuous.

(a) Show that $f(x) := \int_{\Omega'} F(w)(x)\,\mathrm{d}w$ defines a function $f \in C_\mathrm{b}^1(\Omega)$, with $\partial_j f(x) = \int_{\Omega'} \partial_j(F(w))(x)\,\mathrm{d}w$ for all $x \in \Omega$, $j = 1, \dots, n$.

(b) Show that

$$\|f\|_{H^1(\Omega)} \leqslant \int_{\Omega'} \|F(w)\|_{H^1(\Omega)}\,\mathrm{d}w.$$

Hint: Exploit the equality $\|f\|_{H^1(\Omega)} = \sup\big\{|(f\,|\,g)_{H^1(\Omega)}|;\ g \in H^1(\Omega),\ \|g\|_{H^1(\Omega)} \leqslant 1\big\}$.

Note. If we had the Bochner integral for the general context at our disposal, the inequality would be elementary.

(c) Convince yourself that the corresponding property holds if $C_\mathrm{b}^1(\Omega)$ is replaced by $C_\mathrm{b}^1(\Omega;\mathbb{K}^n)$, and the norm in $H^1(\Omega)$ by the norm in $H^1(\Omega;\mathbb{K}^n)$.

Appendix I

The fixed point theorems of Brouwer and Schauder

The fixed point theorems proved in this appendix are important for the existence of solutions for nonlinear problems. In contrast to the Banach fixed point theorem, which also provides uniqueness of solutions, the fixed point theorems treated here only provide existence. The Schauder fixed point theorem is used in Chapter 19. The Brouwer fixed point theorem, interesting and important in its own right, is needed for the proof of Schauder's theorem.

I.1 Brouwer's fixed point theorem

Throughout this section we fix $n \in \mathbb{N}$, and we denote the closed unit ball $B_{\mathbb{R}^n}[0,1]$ in $\mathbb{R}^n$ by B. As a consequence, $\mathring{B}$ will be the open unit ball, the interior of B. Slightly abusing the usual terminology, we will call a function continuously differentiable on B if it belongs to $C^1(\overline{B(0,1)})$.

I.1 Theorem (Brouwer). *Let f be a continuous mapping from B to itself. Then f has a fixed point, i.e., there exists $x \in B$ such that $f(x) = x$.*

In a first step we will show that it is sufficient to prove the theorem under the stronger hypothesis that f is continuously differentiable on B. As the whole procedure of the proof is based on contradiction, we will formulate the following proposition 'negatively', i.e., by starting with an assumption which – in the end – we want to disprove.

I.2 Proposition. *Assume that f is a continuous self-map of B without a fixed point. Then there exists a continuously differentiable mapping $\tilde{f}\colon B \to \mathring{B}$ without a fixed point.*

Proof. The compactness of B implies that $\varepsilon := \inf\{|f(x) - x|;\ x \in B\} > 0$. Clearly one can find $r \in (0,1)$ such that $|rf(x) - f(x)| \leqslant \varepsilon/3$ for all $x \in B$. It follows from the Stone–Weierstrass theorem (Theorem B.2) that rf can be approximated by a mapping $\tilde{f}$ that is a polynomial in each component, $\bigl|\tilde{f}(x) - rf(x)\bigr| \leqslant \min\{\varepsilon/3, (1-r)/2\}$. (The set of polynomials on B is an algebra that separates the points of B and contains the function $\mathbf{1}$; hence, the polynomials are dense in $C(B)$.) As a consequence, $\bigl|\tilde{f}(x)\bigr| \leqslant 1 - (1-r)/2$ and $\bigl|\tilde{f}(x) - x\bigr| \geqslant \varepsilon/3$ for all $x \in B$, and $\tilde{f}$ is as asserted. $\square$

We will see below that the existence of a mapping $\tilde{f}$ as in Proposition I.2 would entail the existence of a continuously differentiable retraction of g of B to ∂B, i.e. a continuously differentiable mapping g from B to $\mathbb{R}^n$ with $g(B) \subseteq \partial B$ and $g|_{\partial B} = \mathrm{id}_{\partial B}$. The following result shows that such a retraction does not exist. In fact, this result is the key point in the proof of Brouwer's fixed point theorem and contains the main technical difficulty.

I.3 Proposition. *There is no continuously differentiable retraction of B to ∂B.*

Proof. The proof proceeds via contradiction. We assume that there exists a continuously differentiable retraction g of B to ∂B. We define a homotopy $g\colon [0,1] \times B \to B$ between id_B and g by

$$g_t := (1-t)\,\mathrm{id}_B + tg \qquad (t \in [0,1]).$$

Then $g_0 = \mathrm{id}_B$, $g_1 = g$ and $g_t|_{\partial B} = \mathrm{id}_{\partial B}$ for all $t \in [0,1]$.

We will prove the properties

(i) $t \to p(t) := \int_{\mathring{B}} \det g_t'(x)\,\mathrm{d}x$ is a polynomial,

(ii) $p(0) = \omega_n > 0$, and there exists $\varepsilon \in (0,1]$ such that $p|_{[0,\varepsilon]}$ is constant,

(iii) $p(1) = 0$.

Obviously these properties cannot hold simultaneously, which means that the contradiction will be established.

(i) The entries of the matrix $g_t'(x)$ are given by $(1-t)\delta_{jk} + t\partial_k g_j(x)$ $(j,k = 1,\dots,n)$; hence $\det g_t'(x)$ is a polynomial in t,

$$a_n(x)t^n + a_{n-1}(x)t^{n-1} + \cdots + a_0(x),$$

with continuous coefficients $a_j\colon B \to \mathbb{R}$ $(j = 0,\dots,n)$. Integration over $\mathring{B}$ yields the assertion.

(ii) Put $c := \|g'\|_\infty$ $(= \sup_{|x|\leqslant 1} |g'(x)|$, where $|g'(x)|$ denotes the Euclidean matrix norm). Then for all $t \in [0,1]$ and $x,y \in B$ one obtains

$$|g_t(x) - g_t(y)| \geqslant (1-t)|x-y| - tc|x-y| = (1-t(c+1))|x-y|;$$

hence g_t is injective for $0 \leqslant t < 1/(c+1)$. For all $x \in \mathring{B}$ one has $\det g_0'(x) = 1$, hence the uniform continuity of $(t,x) \mapsto \det g_t'(x)$ on $[0,1] \times B$ implies that there exists $\varepsilon \in \big(0, 1/(c+1)\big)$ such that $\det g_t'(x) > 0$ for all $t \in [0,\varepsilon]$, $x \in \mathring{B}$. This shows that, for $t \in [0,\varepsilon]$, the mapping $g_t|_{\mathring{B}}$ is a diffeomorphism from $\mathring{B}$ to the open subset $g_t(\mathring{B})$ of $\mathring{B}$. From

$$\partial\big(g_t(\mathring{B})\big) = \overline{g_t\big(\mathring{B}\big)} \setminus g_t\big(\mathring{B}\big) \subseteq g_t(B) \setminus g_t\big(\mathring{B}\big) = g_t\big(\partial\mathring{B}\big) = \partial\mathring{B}$$

we conclude that in fact $g_t\big(\mathring{B}\big) = \mathring{B}$.

We now apply the transformation formula: for all $t \in [0,\varepsilon]$ we have

$$p(t) = \int_{\mathring{B}} \det g_t'(x)\,\mathrm{d}x = \int_{g_t(\mathring{B})} 1\,\mathrm{d}x = \lambda^n\big(g_t(\mathring{B})\big) = \lambda^n(\mathring{B}).$$

(iii) $g(\mathring{B}) \subseteq \partial B$ and the theorem of local invertibility imply that $\det g'(x) = 0$ for all $x \in \mathring{B}$; hence $p(1) = 0$. □

In order to construct a retraction as in Proposition I.3 from a (hypothetical) self-map of B without a fixed point, we need the following lemma.

I.4 Lemma. *For $y \in \mathring{B}$ and $z \in \mathbb{R}^n \setminus \{0\}$, the half-ray $y + (0,\infty)z$ meets the unit sphere $S_{n-1} = \partial B$ in precisely one point $\varphi(y,z) \in S_{n-1}$. The function $\varphi\colon \mathring{B} \times (\mathbb{R}^n \setminus \{0\}) \to \partial B$ thus defined is continuously differentiable.*

Proof. For $y \in \mathring{B}$, $z \in \mathbb{R}^n \setminus \{0\}$, the function $s \mapsto |y+sz|^2 = |y|^2 + 2(y \,|\, z)s + |z|^2 s^2$ is quadratic, equals $|y|^2 < 1$ for $s = 0$, and tends to ∞ as $s \to \infty$; hence there is a unique $s > 0$ such that $|y+sz|^2 = 1$. This solution is given by

$$s(y,z) = |z|^{-2}\Big(-(y \,|\, z) + \sqrt{(y \,|\, z)^2 + (1-|y|^2)|z|^2}\Big),$$

which depends continuously differentiably on (y,z). We conclude that $\varphi(y,z) = y + s(y,z)z$ satisfies the assertion. □

Proof of Theorem I.1. Assume that there exists a continuous self-map f of B without a fixed point. By Proposition I.2 we may assume without loss of generality that f is a continuously differentiable mapping from B to $\mathring{B}$. We define $g\colon B \to \partial B$ by $g(x) := \varphi(f(x), x - f(x))$, with φ from Lemma I.4. Then g is a continuously differentiable retraction of B to ∂B. This is a contradiction to Proposition I.3.

Hence, an f as assumed above cannot exist. □

I.5 Corollary. (a) *Let $K \subseteq \mathbb{R}^n$ be homeomorphic to B. Then every continuous self-map of K has a fixed point.*

(b) *Let $K \neq \varnothing$ be a compact convex subset of $\mathbb{R}^n$. Then every continuous self-map of K has a fixed point.*

Proof. (a) Let $h\colon B \to K$ be a homeomorphism, and let f be a continuous self-map of K. Then $h^{-1} \circ f \circ h$ is a continuous self-map of B, and a fixed point x of $h^{-1} \circ f \circ h$ is mapped to a fixed point $h(x)$ of f.

(b) Provided with the Euclidean scalar product, $\mathbb{R}^n$ is a Hilbert space. Let $P\colon \mathbb{R}^n \to K$ be the minimising projection onto K. Let $R > 0$ be such that $K \subseteq B[0,R]$, and let f be a continuous self-map of K. Then $f \circ P$ is a continuous self-map of $B[0,R]$, and a fixed point $x \in B[0,R]$ of $f \circ P$ will automatically belong to $f(K) \subseteq K$, hence be a fixed point of f. □

I.6 Remark. Having Theorem I.1 at our disposal we can reinforce Propoposition I.3 to the statement that there is no continuous retraction of B to ∂B. Indeed, if g were a continuous retraction of B to ∂B, then $-g$ would be a continuous self-map of B without a fixed point. △

I.2 Schauder's fixed point theorem

I.7 Theorem (Schauder). *Let X be a Banach space, let $K \neq \varnothing$ be a compact convex subset of X, and let f be a continuous self-map of K. Then f has a fixed point.*

Proof. We may assume that the scalar field is $\mathbb{R}$. Assume, for a contradiction, that f has no fixed point. Then

$$\varepsilon := \inf\{\|f(x) - x\|\,;\ x \in K\} > 0. \tag{I.1}$$

For convenience, and without loss of generality, we assume that $\varepsilon = 1$. (This can be achieved by replacing the norm $\|\cdot\|$ on X by $\frac{1}{\varepsilon}\|\cdot\|$.) Put

$$\alpha(x) := (1 - \|x\|)^+ \qquad (x \in X).$$

There exist $x_1, \dots, x_n \in K$ such that $\{B(x_j, 1)\,;\ j = 1, \dots, n\}$ is a covering of K. For $x \in K$, $j \in \{1, \dots, n\}$ put $\alpha_j(x) := \alpha(x - x_j)$,

$$\beta_j(x) := \frac{\alpha_j(x)}{\alpha_1(x) + \cdots + \alpha_n(x)} \ (\geqslant 0).$$

Note that the denominator in the expression defining β_j is strictly positive, and observe that $\sum_{j=1}^n \beta_j(x) = 1$ $(x \in K)$.

Let C be the convex hull of $\{x_1, \dots, x_n\}$ and observe that $C \subseteq K$ is a compact convex subset of the finite-dimensional space $\operatorname{lin}\{x_1, \dots, x_n\}$. Then

$$g\colon K \to C, \quad x \mapsto \sum_{j=1}^n \beta_j(x) x_j$$

is continuous. Note that $\beta_j(x) = 0$ if $\|x - x_j\| \geqslant 1$, hence

$$\|x - g(x)\| = \left\| \sum_{j=1}^n \beta_j(x)(x - x_j) \right\| < 1 \qquad (x \in K). \tag{I.2}$$

Brouwer's fixed point theorem, in the form of Corollary I.5(b), implies that the continuous self-map $g \circ f|_C$ of C has a fixed point $x_f \in C$. Using (I.2) we obtain

$$\|f(x_f) - x_f\| = \|f(x_f) - g(f(x_f))\| < 1,$$

a contradiction to (I.1). □

Instead of supposing compactness of the set on which the mapping f is defined, one can also suppose that the mapping has relatively compact range, as follows.

I.8 Theorem (Schauder). *Let X be a Banach space, $C \subseteq X$ a closed convex subset, and let f be a continuous self-map of C, with the property that $f(C)$ is relatively compact in X. Then f has a fixed point.*

Proof. The closed convex hull of a relatively compact set in a Banach space is compact; see Exercise I.1. Thus, the closed convex hull $K := \overline{\operatorname{co}(f(C))}$ is compact, and it is also a subset of C. Hence, the restriction of f to K is a continuous self-map of K, and Theorem I.7 implies that f has a fixed point. □

Notes

The Brouwer fixed point theorem originally appeared in [Bro11; Satz 4]. A combinatorial proof, on the basis of “Sperner’s lemma”, [Spe28], was given in [KKM29]. Meanwhile multitudinous proofs can be found in the literature. Most of them belong to the area of homology theory and use the degree of mappings in $\mathbb{R}^n$. Our proof, which only relies on ‘elementary analysis’, is along the lines of the proof in [DuSc58; Section V.12]; see also [Heu98; Kap. XXVII, 228]. Another interesting proof is given in [Mil78], where also other sources are indicated.

The Schauder fixed point theorem appeared in [Sch30]. Our proof follows [Rud91; Theorem 5.28], where in fact Tikhonov’s fixed point theorem [Tyc35] is proved, the generalisation of Schauder’s fixed point theorem to locally convex topological vector spaces.

Exercises

I.1 Let X be a Banach space, $C \subseteq X$ a compact subset. Show that the closed convex hull of C is compact. (Hint: First prove the claim for finite sets C. Then use the precompactness of C to show that $\operatorname{co}(C)$ is precompact as well.)

Bibliography

[Ada75] R. A. Adams: *Sobolev Spaces*. Academic Press, New York, 1975.

[AdFo05] R. A. Adams and J. J. F. Fournier: *Sobolev Spaces*. 2nd, reprinted edition. Elsevier, Amsterdam, 2005.

[Alt16] H. W. Alt: *Linear Functional Analysis. An Application-Oriented Introduction.* Springer, London, 2016.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara: *Functions of Bounded Variation and Free Discontinuity Problems.* Clarandon Press, Oxford, 2000.

[ACM15] C. Amrouche, P. G. Ciarlet, and C. Mardare: On a lemma of Jaques-Louis Lions and its relation to other fundamental results. *J. Math. Pures Appl.* **104**, 207–226 (2015).

[ABHN11] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander: *Vector-valued Laplace transforms and Cauchy problems.* Second edition. Birkhäuser, Basel, 2011.

[ArCh10] W. Arendt and R. Chill: Global existence for quasilinear diffusion equations in isotropic nondivergence form. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **IX**, 523–539 (2010).

[ArDa08a] W. Arendt and D. Daners: The Dirichlet problem by variational methods. *Bull. London Math. Soc.* **40**, 51–56 (2008).

[ArDa08b] W. Arendt and D. Daners: Varying domains: stability of the Dirichlet and the Poisson problem. *Discrete Contin. Dyn. Syst.* **21**, 21–39 (2008).

[ADF17] W. Arendt, D. Dier, and S. Fackler: J. L. Lions' problem on maximal regularity. *Arch. Math. (Basel)* **109**, 59–72 (2017).

[ADLO14] W. Arendt, D. Dier, H. Laasri, and E. M. Ouhabaz: Maximal regularity for evolution equations governed by non-autonomous forms. *Adv. Diff. Equ.* **19**, 1043–1066 (2014).

[AEH97] W. Arendt, O. El Mennaoui, and M. Hieber: Boundary values of holomorphic semigroups. *Proc. Amer. Math. Soc.* **125**, no. 3, 635–647 (1997).

[ArEl11] W. Arendt and A. F. M. ter Elst: The Dirichlet-to-Neumann operator on rough domains. *J. Differential Equations* **251**, no. 8, 2100–2124 (2011).

[ArEl12a] W. Arendt and A. F. M. ter Elst: From forms to semigroups. *Spectral theory, mathematical system theory, evolution equations, differential and difference equations.* Oper. Theory: Adv. Appl. 221. Birkhäuser/Springer, Basel, 2012, 47–69.

[ArEl12b] W. Arendt and A. F. M. ter Elst: Sectorial forms and degenerate differential operators. *J. Operator Theory* **67**, 33–72 (2012).

[AEKS14] W. Arendt, A. F. M. ter Elst, J. B. Kennedy, and M. Sauter: The Dirichlet-to-Neumann operator via hidden compactness. *J. Funct. Anal.* **266**, 1757–1786 (2014).

[AES23] W. Arendt, A. F. M. ter Elst, and M. Sauter: The Perron solution for elliptic equations without the maximum principle. *Math. Ann.* (2023). `https://doi.org/10.1007/s00208-023-02761-0`.

[ArNi00] W. Arendt and N. Nikolski: Vector-valued holomorphic functions revisited. *Math. Z.* **234**, 777–805 (2000).

[ArUr23] W. Arendt and K. Urban: *Partial Differential Equations*. Springer, New York, 2023.

[Are01] W. Arendt: Approximation of degenerate semigroups. *Taiwanese J. Math.* **5**, 279–295 (2001).

[Are04] W. Arendt: Semigroups and evolution equations: functional calculus, regularity and kernel estimates. *Handbook of Differential Equations. Vol. I.* Ed. by C. M. Dafermos and E. Feireisl. Elsevier, 2004.

[ArBa93] W. Arendt and C. J. K. Batty: Absorption semigroups and Dirichlet boundary conditions. *Math. Ann.* **295**, 427–448 (1993).

[ABK20] W. Arendt, M. Bernhard, and M. Kreuter: Elliptic problems and holomorphic functions in Banach spaces. *Illinois J. Math.* **64**, no. 3, 331–347 (2020).

[ABH01] W. Arendt, S. Bu, and M. Haase: Functional calculus, variational methods and Liapunov's theorem. *Arch. Math. (Basel)* **77**, 65–75 (2001).

[ArCh20] W. Arendt and I. Chalendar: Essentially coercive forms and asymptotically compact semigroups. *J. Math. Anal. Appl.* **491**, no. 2, 124318, 29 pp. (2020).

[ACE22] W. Arendt, I. Chalendar, and R. Eymard: Galerkin approximation for linear problems in Banach and Hilbert spaces. *IMA J. Numer. Anal.* **42**, 165–198 (2022).

[ACE23] W. Arendt, I. Chalendar, and R. Eymard: Lions' representation theorem and applications. *J. Math. Anal. Appl.* **522**, 126946, 23 p. (2023).

[ACE24] W. Arendt, I. Chalendar, and R. Eymard: Space-time error estimates for approximations of linear parabolic problems with generalized time boundary conditions. *IMA J. Numer. Anal.* **00**, 1–53 (2024). URL: `https://doi.org/10.1093/imanum/drae%20028`.

[Are61] R. Arens: Operational calculus of linear relations. *Pacific J. Math.* **11**, 9–23 (1961).

[AHL&02] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, and P. Tchamitchian: The solution of the square root problem for second order elliptic operators on $\mathbb{R}^n$. *Ann. of Math. (2)* **156**, 633–654 (2002).

[Bar96] L. Barthélemy: Invariance d'un convexe fermé par un semi-groupe associé à une forme non-linéaire. *Abstr. Appl. Anal.* **1**, 237–262 (1996).

[BaEl14] C. J. K. Batty and A. F. M. ter Elst: On series of sectorial forms. *J. Evol. Equ.* **14**, 29–47 (2014).

[Bau01] H. Bauer: *Measure and Integration Theory.* Walter de Gruyter, Berlin, New York, 2001.

[BMV24] S. Bechtel, C. Mooney, and M. Veraar: Counterexamples to maximal regularity for operators in divergence form. *Arch. Math. (Basel)* **123**, 199–209 (2024).

[Ber77] M. S. Berger: *Nonlinearity and Functional Analysis.* Academic Press, New York, 1977.

[BeLö76] J. Bergh and J. Löfström: *Interpolation Spaces. An Introduction.* Springer, Berlin, 1976.

[BeDe58] A. Beurling and J. Deny: Espaces de Dirichlet. *Acta Math.* **99**, 203–229 (1958).

[BeDe59] A. Beurling and J. Deny: Dirichlet spaces. *Proc. Acad. Sci. U.S.A.* **45**, 208–215 (1959).

[Beu70] A. Beurling: On analytic extensions of semigroups of operators. *J. Funct. Anal.* **6**, 387–400 (1970).

[Bog79] M. E. Bogovskiĭ: Solution of the first boundary value problem for the equation of continuity of an imcompressible medium. *Soviet Math. Dokl.* **20**, 1094–1098 (1979).

[Bou65] N. Bourbaki: *Intégration, Chap. 1 à 4. Deuxième éd. revue et augmentée.* Hermann, Paris, 1965.

[Bre11] H. Brezis: *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, New York, 2011.

[BrPa70] H. Brezis and A. Pazy: Semigroups of nonlinear contractions on convex sets. *J. Functional Analysis* **6**, 237–281 (1970).

[Bro11] L. E. J. Brouwer: Über Abbildung von Mannigfaltigkeiten. *Math. Ann.* **71**, 97–115 (1911).

[Cac05] V. Cachia: On a product formula for unitary groups. *Bull. London Math. Soc.* **37**, 621–626 (2005).

[CaZa99] V. Cachia and V. A. Zagrebnov: Operator-norm convergence of the Trotter product formula for sectorial generators. *Lett. Math. Phys.* **50**, 203–211 (1999).

[CaZa01a] V. Cachia and V. A. Zagrebnov: Operator-norm convergence of the Trotter product formula for holomorphic semigroups. *J. Operator Theory* **46**, 199–213 (2001).

[CaZa01b] V. Cachia and V. A. Zagrebnov: Trotter product formula for nonself-adjoint Gibbs semigroups. *J. London Math. Soc. (2)* **64**, 436–444 (2001).

[Cal64] A. P. Calderón: Intermediate spaces and interpolation, the complex method. *Studia Math.* **24**, 113–190 (1964).

[Cal06] A. P. Calderón: On an inverse boundary value problem. Reprint of the original paper published by the Brazilian Mathematical Society in ATAS of SBM (Rio de Janeiro), 65–73 (1980). *Comput. Appl. Math.* **25**, 133–138 (2006).

[CaZy56] A. P. Calderón and A. Zygmund: On singular integrals. *Amer. J. Math.* **78**, 289–309 (1956).

[Cea64] J. Céa: Approximation variationelle des problèmes aux limites. *Ann. Inst. Fourier (Grenoble)* **14**, 345–444 (1964).

[Che68] P. R. Chernoff: Note on product formulas for operator semigroups. *J. Funct. Anal.* **2**, 238–242 (1968).

[Che74] P. R. Chernoff: *Product formulas, nonlinear semigroups, and addition of unbounded operators.* Mem. Amer. Math. Soc. 140. American Mathematical Society, 1974.

[ChEl18] R. Chill and A. F. M. ter Elst: Weak and strong approximation of semigroups on Hilbert spaces. *Integral Equations Operator Theory* **90**, no. 1, Paper No. 9, 22 pp. (2018).

[CHA&87] P. Clément, H. J. A. M. Heijmans, S. Angenent, C. J. van Duijn, and B. de Pagter: *One-Parameter Semigroups.* North-Holland, Amsterdam, 1987.

[CoHi37] R. Courant and D. Hilbert: *Methoden der mathematischen Physik II.* Springer, Berlin, 1937.

[CoHi62] R. Courant and D. Hilbert: *Methods of Mathematical Physics, Vol. II, Partial Differential Equations.* Wiley–Interscience, New York, 1962.

[Dan00] D. Daners: Heat kernel estimates for operators with boundary conditions. *Math. Nachr.* **217**, 13–41 (2000).

[DaLi92] R. Dautray and J.-L. Lions: *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 5.* Springer, Berlin, 1992.

[Dav80] E. B. Davies: *One-Parameter Semigroups.* Academic Press, London, 1980.

[Dav89] E. B. Davies: *Heat kernels and spectral theory.* Cambridge University Press, Cambridge, 1989.

[DaLi72] W. J. Davis and J. Lindenstrauss: On total nonnorming subspaces. *Proc. Amer. Math. Soc.* **31**, 109–111 (1972).

[Dei85] K. Deimling: *Nonlinear Functional Analysis.* Springer, Berlin, 1985.

[Doe20] G. Doetsch: Über die obere Grenze des absoluten Betrages einer analytischen Funktion auf Geraden. *Math. Z.* **8**, 237–240 (1920).

[Dun38] N. Dunford: Uniformity in linear spaces. *Trans. Amer. Math. Soc.* **44**, 305–356 (1938).

[DuSc58] N. Dunford and J. Schwartz: *Linear Operators, Part I: General Theory.* John Wiley & Sons, New York, 1958.

[Duo01] J. Duoandikoetxea: *Fourier Analysis.* American Mathematical Society, Providence, RI, 2001.

[EdEv87] D. E. Edmunds and W. D. Evans: *Spectral Theory and Differential Operators.* Clarendon Press, Oxford, 1987.

[EHT16] M. Egert, R. Haller-Dintelmann, and P. Tolksdorf: The Kato square root problem follows from an extrapolation property of the Laplacian. *Publ. Math.* **60**, 451–483 (2016).

[EnNa00] K.-J. Engel and R. Nagel: *One-Parameter Semigroups for Linear Evolution Equations.* Springer, New York, 2000.

[Eva10] L. C. Evans: *Partial Differential Equations.* Second edition. American Mathematical Society, Providence, RI, 2010.

[EvGa92] L. C. Evans and R. F. Gariepy: *Measure Theory and Fine Properties of Functions.* CRC Press, Boca Raton, FL, 1992.

[ENZ11] P. Exner, H. Neidhardt, and V. A. Zagrebnov: Remarks on the Trotter–Kato product formula for unitary groups. *Integral Equations Operator Theory* **69**, 451–478 (2011).

[Fac13] S. Fackler: Regularity of semigroups via the asymptotic behaviour at zero. *Semigroup Forum* **87**, 1–17 (2013).

[Fac17] S. Fackler: J.-L. Lions' problem concerning maximal regularity of equations governed by non-autonomous forms. *Ann. Inst. H. Poincaré (C) Non Linear Analysis* **34**, 699–709 (2017).

[FaYa98] A. Favini and A. Yagi: *Degenerate Differential Equations in Banach Spaces.* Marcel Dekker, Inc., New York, 1998.

[Fed69] H. Federer: *Geometric Measure Theory.* Springer, Berlin, 1969.

[Fel53] W. Feller: On the generation of unbounded semi-groups of bounded linear operators. *Ann. of Math. (2)* **58**, 166–174 (1953).

[Flo81] K. Floret: *Maß- und Integrationstheorie.* Teubner, Stuttgart, 1981.

[For83] O. Forster: *Analysis 3, Integralrechnung im* $\mathbb{R}^n$ *mit Anwendungen.* 2. edition. Springer, Wiesbaden, 1983.

[Fre36] H. Freudenthal: Ueber die Friedrichssche Fortsetzung halbbeschränkter Hermitescher Operatoren. *Proc. Acad. Amsterdam* **39**, 832–833 (1936).

[Fri34] K. Friedrichs: Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren. *Math. Ann.* **109**, 465–487 (1934).

[Gal11] G. P. Galdi: *An Introduction to the Mathematical Theory of the Navier-Stokes Equations. Steady-State Problems.* Second edition. Springer, New York, 2011.

[Går79] L. Gårding: The Dirichlet Problem. *Math. Intelligencer* **2**, no. 1, 43–53 (1979).

[GHH06] M. Geißert, H. Heck, and M. Hieber: On the equation $\operatorname{div} u = g$ and Bogovskiĭ's operator in Sobolev spaces of negative order. *Partial differential equations and functional analysis.* Oper. Theory Adv. Appl. 168. Birkhäuser, Basel, 2006, 113–121.

[Goc16] M. Gockenbach: *Linear Inverse Problems and Tikhonov Regularization.* Mathematical Association of America, Inc., 2016.

[Gol85] J. A. Goldstein: *Semigroups of Linear Operators and Applications.* Oxford University Press, New York, 1985.

[GrE92] K.-G. Grosse-Erdmann: *The Borel–Okada theorem revisited.* Habilitationsschrift. Hagen: Fernuniversität Hagen, 1992.

[GrE04] K.-G. Grosse-Erdmann: A weak criterion for vector-valued holomorphy. *Math. Proc. Cambridge Philos. Soc.* **136**, 399–411 (2004).

[GuAb98a] K. Gustafson and T. Abe: The third boundary condition – was it Robin's? *Math. Intelligencer* **20**, no. 1, 63–71 (1998).

[GuAb98b] K. Gustafson and T. Abe: (Victor) Gustave Robin: 1855–1897. *Math. Intelligencer* **20**, no. 2, 47–53 (1998).

[HKV16] M. Haase, P. C. Kunstmann, and H. Vogt: On the numerical range of generators of symmetric L_∞-contractive semigroups. *Arch. Math. (Basel)* **107**, 553–559 (2016).

[Haa06] M. Haase: *The Functional Calculus of Sectorial Operators.* Birkhäuser, Basel, 2006.

[Haa14] M. Haase: *Functional Analysis. An Elementary Introduction.* American Mathematical Society, Providence, R.I., 2014.

[HaTa13] B. Haberman and D. Tataru: Uniqueness in Calderón's problem with Lipschitz conductivities. *Duke Math. J.* **162**, 497–516 (2013).

[Had96] J. Hadamard: Sur les fonctions entières. *Bull. Soc. Math. France* **24**, 186–187 (1896).

[Had02] J. Hadamard: Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin* **XIII**, 49–52 (1902).

[Had06] J. Hadamard: Sur le principe de Dirichlet. *Bull. Soc. Math. France* **34**, 135–138 (1906).

[Had23] J. Hadamard: *Lectures on Cauchy's Problem in Linear Partial Differential Equations.* Yale University Press, New Haven, 1923.

[HSSW06] S. Hassi, A. Sandovici, H. S. V. de Snoo, and H. Winkler: Form sums of nonnegative selfadjoint operators. *Acta Math. Hungar.* **111**, 81–105 (2006).

[Hed93] L. I. Hedberg: Approximation by harmonic functions, and stability of the Dirichlet problem. *Expo. Math.* **11**, 193–259 (1993).

[Heu98] H. Heuser: *Lehrbuch der Analysis, Teil 2.* 10., durchges. Aufl. Springer Fachmedien, Wiesbaden, 1998.

[Hie96] M. Hieber: Gaussian estimates and holomorphy of semigroups on L^p spaces. *J. London Math. Soc. (2)* **54**, 148–160 (1996).

[Hil06] D. Hilbert: Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen IV. *Nachr. Gesellsch. Wiss. Göttingen (often cited as "Nachr. Akad. Wiss. Göttingen"). Math.-Phys. Kl.,* 49–91 (1906).

[Hil12] D. Hilbert: *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen.* B. G. Teubner, Leipzig und Berlin, 1912. (Wiederabdruck von sechs Mitteilungen aus den Nachrichten der K. Gesellschaft der Wissenschaften zu Göttingen 1904–1910.)

[Hil48] E. Hille: *Functional Analysis and Semi-Groups.* Amer. Math. Soc., New York, 1948.

[Hil39] E. Hille: Notes on linear transformations. II. Analyticity of semi-groups. *Ann. of Math. (2)* **40**, 1–47 (1939).

[HiPh57] E. Hille and R. S. Phillips: *Functional Analysis and Semi-Groups.* Revised edition. Amer. Math. Soc., Providence, Rh. I., 1957.

[JaLu00] T. Jahnke and C. Lubich: Error bounds for exponential operator splittings. *BIT* **40**, 735–744 (2000).

[Joh82] F. John: *Partial Differential Equations.* 4th edition. Springer-Verlag, New York, 1982.

[JoLa00] G. W. Johnson and M. L. Lapidus: *The Feynman Integral and Feynman's Operational Calculus.* Clarendon Press, Oxford, 2000.

[Kab99] W. Kaballo: *Einführung in die Analysis III.* Spektrum Akademischer Verlag, Heidelberg, 1999.

[Kak41] S. Kakutani: Concrete representation of abstract (M)-spaces (A characterization of the space of continuous functions). *Ann. of Math. (2)* **42**, 994–1024 (1941).

[Kat61a] T. Kato: Abstract evolution equations of parabolic type in Banach and Hilbert spaces. *Nagoya Math. J.* **19**, 93–125 (1961).

[Kat61b] T. Kato: Fractional powers of dissipative operators. *J. Math. Soc. Japan* **13**, 246–274 (1961).

[Kat70] T. Kato: A characterisation of holomorphic semigroups. *Proc. Amer. Math. Soc.* **25**, 495–498 (1970).

[Kat78] T. Kato: *Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups.* Topics in functional analysis (essays dedicated to M. G. Kreĭn on the occasion of his 70th birthday), pp. 185–195, Adv. in Math. Suppl. Stud., vol. 3. Academic Press, New York, 1978.

[Kat80] T. Kato: *Perturbation Theory for Linear Operators.* Springer, Berlin, 1980.

[Kel41] M. V. Keldysh: On the solubility and the stability of Dirichlet's problem. *Amer. Math. Soc. Translations (2)* **51**, 1–73 (1966). *Uspekhi Mat. Nauk* **8**, 171–231 (1941).

[KeSa14] C. Kenig and M. Salo: Recent progress in the Calderón problem with partial data. *Contemp. Math.* **615**, 193–222 (2014).

[Kna05] A. W. Knapp: *Basic Real Analysis*. Birkhäuser, Boston, 2005.

[KKM29] B. Knaster, C. Kuratowski, and S. Mazurkiewics: Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe. *Fund. Math.* **14**, 132–137 (1929).

[Kri11] C. Kriegler: Analyticity angles for non-commutative diffusion semigroups. *J. London Math. Soc.* **83**, 168–186 (2011).

[KrUh14] K. Krupchyk and G. Uhlmann: Uniqueness in an inverse boundary problem for a magnetic Schrödinger operator with a bounded magnetic potential. *Comm. Math. Phys.* **327**, 993–1009 (2014).

[Kun70] H. Kunita. Sub-Markov semi-groups in Banach lattices. *Proceedings of the International Conference on Functional Analysis and Related Topics (Tokyo, 1969)*. Univ. of Tokyo Press, Tokyo, 1970, 332–343.

[Kun05] M. Kunze: *Form Methods for Linear Evolution Problems on Hilbert Spaces*. Diplomarbeit, Fakultät für Mathematik und Wirtschaftswissenschaften. Universität Ulm, 2005.

[LaMi54] P. D. Lax and A. N. Milgram: Parabolic equations. *Contributions to the theory of partial differential equations*. Ed. by L. Bers, S. Bochner, and F. John. Ann. of Math. Stud., no. 33. Princeton University Press, Princeton, NJ, 1954, 167–190.

[Lax02] P. D. Lax: *Functional Analysis*. Wiley-Interscience, New York, 2002.

[LeSc34] J. Leray and J. Schauder: Topologie et équations fonctionelles. *Ann. Sci. Éc. Norm. Supér. (3)* **51**, 45–78 (1934).

[Lev06] B. Levi: Sul principio de Dirichlet. *Rend. Circ. Mat. Palerma* **22**, 293–359 (1906).

[Lio57] J. L. Lions: Sur les problèmes mixtes pour certains systèmes paraboliques dans des ouverts non-cylindriques. *Ann. Inst. Fourier* **7**, 143–182 (1957).

[Lio59] J. L. Lions: Problemi misti nel senso di Hadamard classici e generalizzati. *Milan J. Math. (Rend. Semin. Mat. Fis. Milano)* **29**, 235–239 (1959).

[Lio61] J. L. Lions: *Equations Différentielles Opérationnelles et Problèmes aux Limites*. Springer, Berlin, 1961.

[LiPe95] V. A. Liskevich and M. A. Perelmuter: Analyticity of submarkovian semigroups. *Proc. Amer. Math. Soc.* **123**, 1097–1104 (1995).

[Lun95] A. Lunardi: *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Birkhäuser Verlag, Basel, 1995.

[Lun18] A. Lunardi: *Interpolation theory*. 3nd edition. Appunti, Sc. Norm. Super. Pisa (N.S.) 9, 2018.

[MaRo92] Z.-M. Ma and M. Röckner: *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*. Springer, Berlin, 1992.

[MVV05] A. Manavi, H. Vogt, and J. Voigt: Domination of semigroups associated with sectorial forms. *J. Operator Theory* **54**, 9–25 (2005).

[MaSc65] S. Mandelbrojt and L. Schwartz: Jacques Hadamard. *Bull. Amer. Math. Soc.* **71**, 107–129 (1965).

[MaMi79a] M. Marcus and V. J. Mizel: Complete characterization of functions which act, via superposition, on Sobolev spaces. *Trans. Amer. Math. Soc.* **251**, 187–218 (1979).

[MaMi79b] M. Marcus and V. J. Mizel: Every superposition operators mapping one Sobolev space into another is continuous. *J. Funct. Anal.* **33**, 217–229 (1979).

[Mar38] A. Markoff: On mean values and exterior densities. *Mat. Sbornik* **4 (46)**, 165–191 (1938).

[Mat03] M. Matolcsi: On the relation of closed forms and Trotter's product formula for projections. *J. Funct. Anal.* **205**, 401–413 (2003).

[MaSh03] M. Matolcsi and R. Shvidkoy: Trotter's product formula for projections. *Arch. Math. (Basel)* **81**, 309–317 (2003).

[Maz11] V. G. Maz'ya: *Sobolev Spaces, with Applications to Elliptic Partial Differential Equations.* 2nd, revised and augmented edition. Springer, Berlin, 2011.

[MaSh98] V. Maz'ya and T. Shaposhnikova: *Jacques Hadamard, A Universal Mathematician.* American Mathematical Society, Providence, RI; London Mathematical Society, London, 1998. URL: `https://doi.org/10.1090/hmath/014`.

[Mil78] J. Milnor: "Hairy ball theorem" and Brouwer fixed point theorem. *Amer. Math. Monthly* **85**, 521–524 (1978).

[Miy52] I. Miyadera: Generation of a strongly continuous semi-group operators. *Tôhoku Math. J. (2)* **4**, 109–114 (1952).

[Mon06] S. Monniaux: Navier-Stokes equations in arbitrary domains: the Fujita-Kato scheme. *Math. Res. Lett.* **13**, 455–461 (2006).

[Nag86] R. Nagel (ed.): *One-Parameter Semigroups of Positive Operators.* Springer, Berlin, 1986.

[Nau02] J. Naumann: *Remarks on the Prehistory of Sobolev Spaces.* Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Mathematik, 2002. DOI: `https://dx.doi.org/10.18452/2615`.

[Neč66] J. Nečas: Sur les normes équivalentes dans W_p^k et sur la coercivité des formes formellement positives. *Séminaire Équations aux Dérivées Partielles, Sém. math. sup. Univ. Montréal, 1965, Les Presses de l'Université de Montréal*, 102–128 (1966).

[Neč12] J. Nečas: *Direct Methods in the Theory of Elliptic Equations.* (original French version 1967). Springer, Heidelberg, 2012.

[NSZ18] H. Neidhardt, A. Stephan, and V. A. Zagrebnov: Remarks on the operator-norm convergence of the Trotter product formula. *Integral Equations Operator Theory* **90**, no. 2, Paper No. 15, 14 pp. (2018).

[Nel59] E. Nelson: Analytic vectors. *Ann. of Math. (2)* **70**, 572–615 (1959).

[Neu70] J. W. Neuberger: Analyticity and quasi-analyticity for one-parameter semi-groups. *Proc. Amer. Math. Soc.* **25**, 488–494 (1970).

[Neu32a] J. von Neumann: *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin, 1932.

[Neu32b] J. von Neumann: Über adjungierte Funktionaloperatoren. *Ann. of Math. (2)* **33**, 294–310 (1932).

[Ouh92] E.-M. Ouhabaz: L^∞-contractivity of semigroups generated by sectorial forms. *J. London Math. Soc.* **(2) 46**, 529–542 (1992).

[Ouh95a] E.-M. Ouhabaz: Gaussian estimates and holomorphy of semigroups. *Proc. Amer. Math. Soc.* **123**, no. 5, 1465–1474 (1995).

[Ouh95b] E.-M. Ouhabaz: Second order elliptic operators with essential spectrum $[0, \infty)$ on L^p. *Comm. Partial Differential Equations* **20**, 763–773 (1995).

[Ouh96] E.-M. Ouhabaz: Invariance of closed convex sets and domination criteria for semigroups. *Potential Anal.* **5**, 611–625 (1996).

[Ouh05] E. M. Ouhabaz: *Analysis of Heat Equations on Domains*. Princeton University Press, Princeton, 2005.

[OuSp10] E. M. Ouhabaz and C. Spina: Maximal regularity for non-autonomous Schrödinger type equations. *J. Differential Equations* **248**, 1668–1683 (2010).

[Paz83] A. Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, 1983.

[Per23] O. Perron: Eine neue Bahandlung der Randwertaufgabe für $\Delta u = 0$. *Math. Z.* **18**, 42–54 (1923).

[Phi53] R. S. Phillips: Perturbation theory for semi-groups of linear operators. *Trans. Amer. Math. Soc.* **74**, 199–221 (1953).

[Phi55] R. S. Phillips: Semi-groups of operators. *Bull. Amer. Math. Soc.* **61**, 16–33 (1955).

[Phi59] R. S. Phillips: Dissipative operators and hyperbolic systems of partial differential equations. *Trans. Amer. Math. Soc.* **91**, 193–254 (1959).

[Phi62] R. S. Phillips: Semi-groups of positive contraction operators. *Czechoslovak Math. J.* **12**, 294–313 (1962).

[PrSo93] J. Prüss and H. Sohr: Imaginary powers of elliptic second order differential operators in L^p-spaces. *Hiroshima Math. J.* **23**, 161–192 (1993).

[ReSi78] M. Reed and B. Simon: *Methods of Modern Mathematical Physics IV: Analyis of Operators*. Academic Press, New York, 1978.

[ReSi80] M. Reed and B. Simon: *Methods of Modern Mathematical Physics I: Functional Analysis*. Revised and enlarged edition. Academic Press, San Diego, 1980.

[Rel30] F. Rellich: Ein Satz über mittlere Konvergenz. *Nachr. Gesellsch. Wiss. Göttingen (often cited as "Nachr. Akad. Wiss. Göttingen"). Math.-Phys. Kl.,* 30–35 (1930).

[Rie09] F. Riesz: Sur les opérations fonctionelles linéaires. *Comptes Rendus Acad. Sci. Paris* **149**, 974–977 (1909).

[RiNa68] F. Riesz and B. Sz.-Nagy: *Vorlesungen über Funktionalanalysis*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1968.

[Rud91] W. Rudin: *Functional Analysis*. 2nd edition. McGraw-Hill, New York, 1991.

[Sak38] S. Saks: Integration in abstract metric spaces. *Duke Math. J.* **4**, 408–411 (1938).

[Sau13] M. Sauter: *Degenerate Elliptic Operators with Boundary Conditions via Form Methods*. PhD thesis. The University of Auckland, 2013. URL: `https://hdl.handle.net/2292/21318`.

[Sau20] M. Sauter: Uniqueness of the approximative trace. *Indiana Univ. Math. J.* **69**, no. 1, 171–204 (2020).

[Sch55] H. H. Schaefer: Über die Methode der a priori-Schranken. *Math. Ann.* **129**, 415–416 (1955).

[Sch30] J. Schauder: Der Fixpunktsatz in Funktionalräumen. *Studia Math.* **2**, 171–180 (1930).

[Sch12] K. Schmüdgen: *Unbounded Self-adjoint Operators on Hilbert Space*. Springer, Dordrecht, 2012.

[Sch69] J. T. Schwartz: *Nonlinear Functional Analysis*. Gordon and Breach, New York, 1969.

[Sim78] B. Simon: A canonical decomposition for quadratic forms with applications to monotone convergence theorems. *J. Funct. Anal.* **28**, 377–385 (1978).

[Sob38] S. L. Sobolev: On a theorem of functional analysis. Russian original with French synopsis: *Math. Sb. (N.S.)* **4**, 471–491 (1938). *Amer. Math. Soc. Transl. (2)* **34**, 39–68 (1963).

[Spe28] E. Sperner: Neuer Beweis für die Invarianz der Dimensionszahl und des Gebietes. *Abh. Math. Semin. Univ. Hamb.* **6**, 265–272 (1928).

[Sta64] G. Stampacchia: Équations elliptiques du second ordre à coefficients discontinus. *Séminaire Jean Leray*, no. 3, 1–77 (1963-1964).

[Ste56] E. M. Stein: Interpolation of linear operators. *Trans. Amer. Math. Soc.* **83**, 482–492 (1956).

[Sto93] P. Stollmann: Closed ideals in Dirichlet spaces. *Potential Anal.* **2**, 263–268 (1993).

[Sto30] M. H. Stone: Linear Transformations in Hilbert Space III. Operational Methods and Group Theory. *Proc. Natl. Acad. Sci. USA* **16**, 172–175 (1930).

[Sto32] M. H. Stone: On one-parameter unitary groups in Hilbert space. *Ann. of Math. (2)* **33**, 643–648 (1932).

[Sto37] M. H. Stone: Applications of the theory of Boolean rings to general topology. *Trans. Amer. Math. Soc.* **41**, 375–481 (1937).

[Sto48] M. H. Stone: The generalized Weierstrass approximation theorem. *Math. Mag.* **21**, 167–184 (1948).

[Str94] D. W. Stroock: *A Concise Introduction to the Theory of Integration*. 2nd edition. Birkhäuser, Boston, 1994.

[SyUh87] J. Sylvester and G. Uhlmann: A global uniqueness theorem for an inverse boundary value problem. *Ann. Math.* **125**, 153–169 (1987).

[Tam00] H. Tamura: A remark on operator-norm convergence of Trotter–Kato product formula. *Integral Equations Operator Theory* **37**, 350–356 (2000).

[Tem77] R. Temam: *Navier-Stokes Equations. Theory and Numerical Analysis*. North-Holland, Amsterdam, 1977.

[Tha08] M. Thalhammer: High-order exponential operator splitting methods for time-dependent Schrödinger equations. *SIAM J. Numer. Anal.* **26**, 2022–2038 (2008).

[Tre75] F. Treves: *Basic linear partial differential equations*. Academic Press, Orlando, 1975.

[Tro58] H. F. Trotter: Approximation of semi-groups of operators. *Pacific J. Math.* **8**, 887–919 (1958).

[Tro59] H. F. Trotter: On the product of semi-groups of operators. *Proc. Amer. Math. Soc.* **10**, 545–551 (1959).

[Tyc35] A. Tychonoff: Ein Fixpunktsatz. *Math. Ann.* **111**, 767–776 (1935).

[Uhl09] G. Uhlmann: Electrical impedance tomography and Calderón's problem. *Inverse Problems* **25**, 123011 (39pp) (2009).

[Vis56] M. I. Vishik: The problem of Cauchy with operators as coefficients, the mixed boundary problem for systems of differential equations and an approximate method of their solution. in Russian. *Mat. Sb. (N.S.)* **39 (81)**, 51–148 (1956).

[VoVo18] H. Vogt and J. Voigt: Holomorphic families of forms, operators and C_0-semigroups. *Monatsh. Math.* **187**, 375–380 (2018).

[VoVo20a] H. Vogt and J. Voigt: Increasing sequences of sectorial forms. *Czechoslovak Math. J.* **70 (145)**, 1033–1046 (2020).

[VoVo20b] H. Vogt and J. Voigt: On Hausdorff measure and an inequality due to Maz'ya. *Arch. Math. (Basel)* **114**, 573–583 (2020).

[VoVo24] H. Vogt and J. Voigt: On sequences of sectorial forms converging 'from above'. *Discrete Contin. Dyn. Syst. Ser. S* **17**, no. 5&6, 2021–2029 (2024).

[Voi96] J. Voigt: The sector of holomorphy for symmetric submarkovian semigroups. *Functional Analysis: Proceedings of the First International Workshop held at Trier University, Sept. 26–Oct. 1, 1994*. Ed. by S. Dierolf, S. Dineen, and P. Domański. de Gruyter, Berlin, 1996, 449–453.

[Voi20] J. Voigt: *A Course on Topological Vector Spaces*. Birkhäuser, Cham (Switzerland), 2020.

[Voi23a] J. Voigt: On the existence of distributional potentials. *Math. Nachr.* **296**, 424–433 (2023).

[Voi23b] J. Voigt: Couplings and primitives on topological spaces. *Arch. Math. (Basel)* **121**, 183–188 (2023).

[Wey40] H. Weyl: The method of orthogonal projection in potential theory. *Duke Math. J.* **7**, 411–444 (1940).

[WKK09] E. Wienholtz, H. Kalf, and T. Kriecherbauer: *Elliptische Differentialgleichungen zweiter Ordnung*. Springer-Verlag, Berlin, 2009.

[Yos48] K. Yosida: On the differentiability and the representation of one-parameter semi-groups of linear operators. *J. Math. Soc. Japan* **1**, 15–21 (1948).

[Yos68] K. Yosida: *Functional Analysis*. 2nd edition. Springer, New York, 1968.

[Zag19] V. A. Zagrebnov: *Gibbs Semigroups*. Birkhäuser, Cham, Switzerland, 2019.

Index of notation

$\mathbb{N}$	the set of natural numbers $\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$
$\mathbb{R}$	the field of real numbers
$\mathbb{C}$	the field of complex numbers
$\mathbb{K}$	$\mathbb{R}$ or $\mathbb{C}$
$A^\top$	transpose of a matrix $A \in \mathbb{K}^{m\times n}$
$\operatorname{lin} B$	linear span of a subset B of a vector space
$B(x, r)$	open ball with centre x, radius r, page 16
$B[x, r]$	closed ball with centre x, radius r, page 16
Σ_θ	open sector of (semi-)angle $\theta \in (0, \pi/2]$, page 30
$\Sigma_{\theta,0}$	$\Sigma_\theta \cup \{0\}$, page 30
Σ_0	$(0, \infty)$, page 36
V'	dual space of a Hilbert space V, page 57
V^*	antidual space of a Hilbert space V, page 57
$L_2(\mu)_+$	cone of functions $0 \leqslant u \in L_2(\mu)$, page 116
$\operatorname{diam}(C)$	diameter of a set C in a metric space
$\operatorname{rd}(C)$	$\frac{1}{2}\operatorname{diam}(C)$ for a set C in a metric space, should be remindful of 'radius', page 290
$\operatorname{rd}(\mathcal{C})$	$\sup_{C\in\mathcal{C}} \operatorname{rd}(C)$, for a countable collection $\mathcal{C}$ of subsets of a metric space, page 290
$\mathcal{L}(X, Y)$	space of bounded linear operators from X to Y, with Banach spaces X, Y, page 3
s-lim	limit in the strong operator topology, page 3
$\operatorname{dom}(A)$	domain of A, page 5

$\operatorname{ran}(A)$	range of A, page 5
$\ker(A)$	kernel of A, page 5
$\operatorname{num}(A)$	numerical range of A, page 36
e^{tA}	the C_0-semigroup generated by a generator A, page 10
A^*	adjoint of the operator A, page 71
A^*	adjoint of the linear relation A, page 187
$A \sim (a, j)$	operator A associated with (a, j), page 155
tr	trace operator, assigning boundary values to functions, page 92
Δ_D	Dirichlet Laplacian, page 51
Δ_N	Neumann Laplacian, page 94
Δ_β	Robin Laplacian, page 95
Δ_β	Robin Laplacian for rough domains, page 162
$C(G)$	space of continuous functions with values in $\mathbb{K}$, for a locally compact set $G \subseteq \mathbb{R}^n$, page 12
$C_\mathrm{b}(G)$	space of bounded continuous functions, page 12
$C_\mathrm{c}(G)$	space of continuous functions with compact support, page 12
$C_0(G)$	closure of $C_\mathrm{c}(G)$ in $C_\mathrm{b}(G)$, page 12
$C^k(\Omega)$	space of k times continuously differentiable functions on an open set $\Omega \subseteq \mathbb{R}^n$, for $k \in \mathbb{N}_0$
$C^\infty(\Omega)$	space of infinitely differentiable functions
$C^k(\overline{\Omega})$	space of functions $u \in C(\overline{\Omega}) \cap C^k(\Omega)$, with derivatives up to order k continuously extendable to $\overline{\Omega}$, for bounded open $\Omega \subseteq \mathbb{R}^n$, page 87
$L_{1,\mathrm{loc}}(\Omega)$	space of locally integrable functions, page 41
$H^1(\Omega)$	Sobolev space of L_2-functions with distributional first order derivatives in L_2, page 46
$H^1_\mathrm{c}(\Omega)$	functions in $H^1(\Omega)$ with compact support, page 48
$H^1_0(\Omega)$	closure of $H^1_\mathrm{c}(\Omega)$ in $H^1(\Omega)$, page 48
$H^{-1}(\Omega)$	antidual of $H^1_0(\Omega)$ in the Gelfand triple $\big(H^1_0(\Omega), L_2(\Omega), H^{-1}(\Omega)\big)$, page 215

$W_1^1(\Omega)$	Sobolev space of L_1-functions, with distributional first order derivatives in L_1, page 287
$W_{1,\mathrm{c}}^1(\Omega),\ W_{1,0}^1(\Omega)$	analogous to $H_\mathrm{c}^1(\Omega),\ H_0^1(\Omega)$, page 287
$H^1(a,b;H)$	vector-valued Sobolev space, page 232
$\mathbf{1}_A$	indicator function of a subset A of a set Ω, $\mathbf{1}_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in \Omega \setminus A. \end{cases}$
$[f \leqslant g]$	$\{x \in \Omega;\, f(x) \leqslant g(x)\}$, for a set Ω and functions $f, g\colon \Omega \to \mathbb{R}$. Similarly for $[f < g]$, $[f = g]$ etc.
$f \vee g,\ f \wedge g$	$f \vee g(x) := \max\{f(x), g(x)\},\ f \wedge g(x) := \min\{f(x), g(x)\}$ $(x \in \Omega)$, for a set Ω and functions $f, g\colon \Omega \to \mathbb{R}$
$f^+,\ f^-$	positive and negative part of a real-valued function f, $f^+ := f \vee 0$, $f^- := (-f)^+$
sgn	signum function, page 117
$\rho * u$	convolution of ρ and f, page 41
$\operatorname{spt} f$	support of a continuous function, page 12
$\operatorname{spt} f$	support of an $L_{1,\mathrm{loc}}$-function, page 47
$\lvert\alpha\rvert$	$\sum_{j=1}^n \alpha_j$, order of the multiindex $\alpha \in \mathbb{N}_0^n$
$\partial_j u$	partial derivative with respect to the j'th variable, for a differentiable function u of n variables
$\partial^\alpha u$	$\partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} u$, for a multiindex $\alpha \in \mathbb{N}_0^n$
$\partial^\alpha f$	distributional derivative of $f \in L_{1,\mathrm{loc}}(\Omega)$, page 44
∇u	gradient of a function $u \in C^1(\Omega)$ or $u \in H^1(\Omega)$, $\nabla u := (\partial_1 u, \ldots, \partial_n u)^\top$, page 64
$\partial_\nu u$	normal derivative, page 88
$\partial_\nu u$	weak normal derivative, page 93
$\partial_\nu u$	weak normal derivative, page 162

Index