Spectral Theory

Lecture Notes of Summer Semester 2023

Roland Schnaubelt

These lecture notes are based on my course from summer semester 2023. I kept the numbering and the contents of the results presented in the lectures (except for minor corrections and improvements). Some equation numbers have been shifted a bit. Typically, the proofs and calculations in the notes are somewhat shorter than those given in the lecture. Moreover, the drawings and many additional, mostly oral remarks from the lectures are omitted here. On the other hand, I have added several, often lengthy proofs not shown in the lectures (in particular about Sobolev spaces). These are marked by footnotes or collected in appendices. Occasionally I use the notation, definitions and basic results of my lecture notes Analysis 1–4 and Functional Analysis without further notice.

I like to thank Heiko Hoffmann, Bernhard Konrad and Martin Meyries for their support in preparing earlier versions of these lecture notes and for several corrections and comments.

Karlsruhe, July 27, 2023 Roland Schnaubelt

Contents

CHAPTER 1

General spectral theory

The spectrum is the infinite-dimensional analogue of the set of matrix eigenvalues. Several core results of matrix theory can be extended (usually in modified form) to linear operators T on a Banach or Hilbert space, where the proofs are typically quite different and one often needs additional assumptions on T. Besides bounded T, we also treat a class of discontinuous operators, the 'closed' ones, which is relevant for differential operators. We first discuss this class, and then establish the basic results of spectral theory.

General notation. $X \neq \{0\}$, $Y \neq \{0\}$, and $Z \neq \{0\}$ are Banach spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ with norms $\|\cdot\|$ (or $\|\cdot\|_X$ etc.). A linear map $T : X \to Y$ is continuous if and only if the *operator norm* $\|T\| = \sup_{\|x\| \leqslant 1} \|Tx\|$ is finite. Endowed with this norm, the set

$$\mathcal{B}(X,Y) = \{T : X \to Y \,|\, T \text{ is linear and continuous}\}$$

is a Banach space, where we put $\mathcal{B}(X) := \mathcal{B}(X,X)$. We also set $X^\star = \mathcal{B}(X,\mathbb{F})$ and write $x^\star(x) = \langle x, x^\star \rangle_{X \times X^\star} = \langle x, x^\star \rangle$ for $x \in X$ and $x^\star \in X^\star$.

Let $\mathrm{D}(A)$ be a linear subspace of X and $A : \mathrm{D}(A) \to Y$ be linear. Then A, or $(A, \mathrm{D}(A))$, is called *linear operator from X to Y* (and *on* X if $X = Y$) with *domain* $\mathrm{D}(A)$. Its *kernel* ar *range* are denoted by

$$\mathrm{N}(A) = \{x \in \mathrm{D}(A) \,|\, Ax = 0\} \quad \text{resp.} \quad \mathrm{R}(A) = \{y \in Y \,|\, \exists\, x \in \mathrm{D}(A) \colon y = Ax\}.$$

1.1. Closed operators

We recall one of the basic examples of an unbounded operator: Let $X = C([0,1])$ be endowed with $\|\cdot\|_\infty$ and consider $Af = f'$ with domain $\mathrm{D}(A) = C^1([0,1])$. Then A is linear, but not bounded. Indeed, the functions $u_n \in \mathrm{D}(A)$ given by $u_n(t) = (1/\sqrt{n}) \sin(nt)$ for $n \in \mathbb{N}$ satisfy $\|u_n\|_\infty \to 0$ and

$$\|Au_n\|_\infty \geqslant |u_n'(0)| = \sqrt{n} \to \infty \quad \text{as } n \to \infty.$$

However, if $f_n \in \mathrm{D}(A) = C^1([0,1])$ fulfill $f_n \to f$ and $Af_n = f_n' \to g$ in $C([0,1])$ as $n \to \infty$, then $f \in \mathrm{D}(A)$ and $Af = g$ (see Analysis 1). This observation leads us to the following basic definition.

Definition 1.1. *Let A be a linear operator from X to Y. The operator A is called* closed *if for all $x_n \in \mathrm{D}(A)$, $n \in \mathbb{N}$, possessing limits $x = \lim_{n\to\infty} x_n$ in X and $y = \lim_{n\to\infty} Ax_n$ in Y, we have $x \in \mathrm{D}(A)$ and $Ax = y$.*

For closed A, we thus have $\lim_{n\to\infty} Ax_n = A(\lim_{n\to\infty} x_n)$ if (x_n) **and** (Ax_n) converge. We discuss some basic examples, where we are a bit sloppy when working in L^p-spaces. In Examples 2.6 b) and 2.12 a) of [**FA**] one can find a more precise treatment of related issues. Differential operators on L^p-spaces are studied in Section 3.4.

EXAMPLE 1.2. a) Clearly, every operator $A \in \mathcal{B}(X, Y)$ is closed (with $\mathrm{D}(A) = X$). On $X = C([0,1])$ the operator $Af = f'$ with $\mathrm{D}(A) = C^1([0,1])$ is closed, as seen above. Below we equip ${}^{\mathrm{d}}\!/\!{}_{\mathrm{d}x}$ with boundary conditions.

b) Let $X = C([0,1])$. The operator $Af = f'$ with

$$\mathrm{D}(A) = \{f \in C^1([0,1]) \,|\, f(0) = 0\}.$$

is closed in X. Indeed, let $f_n \in \mathrm{D}(A)$ and $f, g \in X$ be such that $f_n \to f$ and $Af_n = f_n' \to g$ in X as $n \to \infty$. Again by Analysis 1, the function f belongs to $C^1([0,1])$ and $f' = g$. Since $0 = f_n(0) \to f(0)$ as $n \to \infty$, we obtain $f \in \mathrm{D}(A)$ and thus $Af = f' = g$. This means that A is closed on X. In the same way we see that $A_1 f = f'$ with

$$\mathrm{D}(A_1) = \{f \in C^1([0,1]) \,|\, f'(0) = 0,\ f(1) = 0\}$$

is closed in X. There are many more variants.

c) Let $X = C([0,1])$ and $Af = f'$ with

$$\mathrm{D}(A) = C_c^1((0,1]) = \{f \in C^1([0,1]) \,|\, \operatorname{supp} f \subseteq (0,1]\},$$

where the support $\operatorname{supp} f$ of f is the closure of $\{t \in [0,1] \,|\, f(t) \neq 0\}$ in $[0,1]$. This operator is not closed. In fact, consider the functions $f_n \in \mathrm{D}(A)$ and $f \in C^1([0,1])$ given by

$$f(t) = t^2, \qquad f_n(t) = \begin{cases} 0, & 0 \leqslant t < 1/n, \\ (t - 1/n)^2, & 1/n \leqslant t \leqslant 1, \end{cases}$$

for every $n \in \mathbb{N}$. We then have the limits $f_n \to f$ and $f_n' \to f'$ in X as $n \to \infty$. However, since $\operatorname{supp} f = [0,1]$ the map f does not belong to $\mathrm{D}(A)$.

d) Let $X = L^p(\mathbb{R}^d)$, $1 \leqslant p \leqslant \infty$, and $m : \mathbb{R}^d \to \mathbb{C}$ be measurable. Define $Af = mf$ with

$$\mathrm{D}(A) = \{f \in X \,|\, mf \in X\}.$$

This is the *maximal domain*. Then A is closed. Indeed, let $f_n \to f$ and $Af_n = mf_n \to g$ in X as $n \to \infty$. Then there is a subsequence such that $f_{n_j}(x) \to f(x)$ and $m(x)f_{n_j}(x) \to g(x)$ for a.e. $x \in \mathbb{R}^d$, as $j \to \infty$. Hence, $mf = g$ in $L^p(\mathbb{R}^d)$ and we thus obtain $f \in \mathrm{D}(A)$ and $Af = g$.

e) Let $X = L^1([0,1])$, $Y = \mathbb{C}$, and $Af = f(0)$ with $\mathrm{D}(A) = C([0,1])$. Then A is not closed from X to Y. In fact, look at $f_n \in \mathrm{D}(A)$ given by

$$f_n(t) = \begin{cases} 1 - nt, & 0 \leqslant t \leqslant 1/n, \\ 0, & 1/n < t \leqslant 1, \end{cases}$$

for $n \in \mathbb{N}$. Here $\|f_n\|_1 = \frac{1}{2n}$ tends to 0 as $n \to \infty$, but $Af_n = f_n(0) = 1$. $\Diamond$

To study closed operators, one uses the following concepts.

DEFINITION 1.3. *Let A be a linear operator from X to Y. The* graph *of A is given by*

$$\mathrm{G}(A) = \{(x, Ax) \in X \times Y \,|\, x \in \mathrm{D}(A)\}.$$

The graph norm *of A is defined by $\|x\|_A = \|x\|_X + \|Ax\|_Y$. We write* $[\mathrm{D}(A)]$ *if we equip* $\mathrm{D}(A)$ *with* $\|\cdot\|_A$.

Note that $\|\cdot\|_A$ is equivalent to $\|\cdot\|_X$ if $A \in \mathcal{B}(X,Y)$. We endow $X \times Y$ with the norm $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$. Recall that a sequence in $X \times Y$ converges if and only if its components in X and in Y converge. We collect basic properties of the above notions, where part c) connects them closely.

Lemma 1.4. *Every linear operator A from X to Y satisfies the following assertions.*

a) $\mathrm{G}(A) \subseteq X \times Y$ *is a linear subspace.*

b) $[\mathrm{D}(A)]$ *is a normed vector space and* $A \in \mathcal{B}([\mathrm{D}(A)], Y)$.

c) A is closed if and only if $\mathrm{G}(A)$ *is closed in $X \times Y$ if and only if* $[\mathrm{D}(A)]$ *is a Banach space.*

d) Let A be injective and put $\mathrm{D}(A^{-1}) := \mathrm{R}(A)$. *Then, A is closed from X to Y if and only if A^{-1} is closed from Y to X.*

Proof. Statements a) and b) follow from the definitions, and assertion c) implies d) since

$$\mathrm{G}(A^{-1}) = \big\{(y, A^{-1}y) \,\big|\, y \in \mathrm{R}(A)\big\} = \big\{(Ax, x) \,\big|\, x \in \mathrm{D}(A)\big\}$$

is closed in $Y \times X$ if and only if $\mathrm{G}(A)$ is closed in $X \times Y$. We next show c).

The operator A is closed if and only if for all $x_n \in \mathrm{D}(A)$, $n \in \mathbb{N}$, and $(x,y) \in X \times Y$ with $(x_n, Ax_n) \to (x,y)$ in $X \times Y$ as $n \to \infty$, we have $x \in \mathrm{D}(A)$ and $Ax = y$; i.e., $(x,y) \in \mathrm{G}(A)$. This property is equivalent to the closedness of $\mathrm{G}(A)$. Since $\|(x,Ax)\|_{X\times Y} = \|x\|_X + \|Ax\|_Y$, a Cauchy sequence or a converging sequence in $\mathrm{G}(A)$ corresponds to a Cauchy or a converging sequence in $[\mathrm{D}(A)]$, respectively. So $[\mathrm{D}(A)]$ is complete if and only if $(\mathrm{G}(A), \|\cdot\|_{X\times Y})$ is complete. By Corollary 1.13 of [**FA**], the latter is equivalent of the closedness of $\mathrm{G}(A)$ in $X \times Y$. □

By part c), a closed operator A on X can also be viewed as a bounded one acting from $[\mathrm{D}(A)]$ to X. However, in spectral theory one has to treat A as a map on X. The following *closed graph theorem* is a variant of the Open Mapping Theorem 4.28 in [**FA**].

Theorem 1.5. *Let X and Y be Banach spaces and A be a closed operator from X to Y. Then A is* bounded *(i.e., $\|Ax\|_Y \leqslant c\|x\|_X$ for some $c \geqslant 0$ and all $x \in \mathrm{D}(A)$) if and only if* $\mathrm{D}(A)$ *is closed in X. In particular, a closed operator with* $\mathrm{D}(A) = X$ *belongs to* $\mathcal{B}(X,Y)$.

Proof. Let $\mathrm{D}(A)$ be closed in X. Then $\mathrm{D}(A)$ is a Banach space for $\|\cdot\|_X$ (by Analysis 2) and $\|\cdot\|_A$ (by Lemma 1.4). Since $\|x\|_X \leqslant \|x\|_A$ for all $x \in \mathrm{D}(A)$, a corollary to the open mapping theorem (see Corollary 4.29 in [**FA**]) shows that there is a constant $c > 0$ such that $\|Ax\|_Y \leqslant \|x\|_A \leqslant c\|x\|_X$ for all $x \in \mathrm{D}(A)$.

Conversely, let A be bounded and let $x_n \in \mathrm{D}(A)$ converge to $x \in X$ with respect to $\|\cdot\|_X$. Then $\|Ax_n - Ax_m\|_Y \leqslant c\|x_n - x_m\|_X$, and so the sequence $(Ax_n)_n$ is Cauchy in Y. There thus exists $y := \lim_{n\to\infty} Ax_n$ in Y. The closedness of A shows that x belongs to $\mathrm{D}(A)$; i.e., $\mathrm{D}(A)$ is closed in X. □

We next show that Theorem 1.5 is wrong without completeness and give an example of a non-closed, everywhere defined operator on each infinite-dimensional Banach space.

Remark 1.6. a) Let M be given by $(Mf)(t) = tf(t)$, $t \in \mathbb{R}$, on $C_c(\mathbb{R})$ with $\|\cdot\|_\infty$. This linear operator is everywhere defined, unbounded and closed. In fact, take $f_n, f, g \in C_c(\mathbb{R})$ such that $f_n(t) \to f(t)$ and $(Mf_n)(t) = tf_n(t) \to g(t)$ uniformly for $t \in \mathbb{R}$ as $n \to \infty$. Then $g(t) = tf(t)$ for all $t \in \mathbb{R}$; i.e., $g = Mf$ and M is closed. Further, pick $\varphi_n \in C_c(\mathbb{R})$ with $\|\varphi_n\|_\infty = 1$ and $\varphi_n(n) = 1$. Since $\|M\varphi_n\|_\infty \geqslant |M\varphi_n(n)| = n$, the operator M is unbounded.

b) Let X be an infinite-dimensional Banach space and let $\mathcal{B}$ be an algebraic basis of X, see Theorem III.5.1 in [**La**]. (Hence, for each $x \in X$ there are unique $\alpha_b(x) = \alpha_b$ for $b \in \mathcal{B}$ with $x = \sum_{b \in \mathcal{B}} \alpha_b b$, where only finitely many of the coefficients $\alpha_b(x)$ are non-zero.)[1] We may assume that $\|b\| = 1$ for all $b \in \mathcal{B}$. Choose a countable subset $\mathcal{B}_0 = \{b_k \mid k \in \mathbb{N}\}$ of $\mathcal{B}$ and set

$$M_{\mathcal{B}} b_k = k b_k \text{ for each } b_k \in \mathcal{B}_0, \text{ and } M_{\mathcal{B}} b = 0 \text{ for each } b \in \mathcal{B} \backslash \mathcal{B}_0.$$

Then $M_{\mathcal{B}}$ can be extended to a linear operator on X which is unbounded, since $\|M_{\mathcal{B}} b_k\| = k$ and $\|b_k\| = 1$. Thus $M_{\mathcal{B}}$ is not closed by Theorem 1.5. $\Diamond$

We discuss permanence properties of closed operators, which are more delicate than for bounded ones. In the proof, one just checks the definition.

Proposition 1.7. *Let A be closed from X to Y, $T \in \mathcal{B}(X,Y)$, and $S \in \mathcal{B}(Z,X)$. Then the following operators are closed.*

a) $B = A + T$ with $\mathrm{D}(B) = \mathrm{D}(A)$.

b) $C = AS$ with $\mathrm{D}(C) = \{z \in Z \,|\, Sz \in \mathrm{D}(A)\}$.

Proof. a) Let $x_n \in \mathrm{D}(B)$, $n \in \mathbb{N}$, $x \in X$, and $y \in Y$ such that $x_n \to x$ in X and $Bx_n = Ax_n + Tx_n \to y$ in Y as $n \to \infty$. Since T is bounded, there exists $Tx = \lim_{n\to\infty} Tx_n$ and so $Ax_n \to y - Tx$ as $n \to \infty$. The closedness of A then yields $x \in \mathrm{D}(A) = \mathrm{D}(B)$ and $Ax = y - Tx$; i.e., $Bx = Ax + Tx = y$.

b) Let $z_n \in \mathrm{D}(C)$, $n \in \mathbb{N}$, $z \in Z$, and $y \in Y$ such that $z_n \to z$ in Z and $ASz_n \to y$ in Y as $n \to \infty$. By the boundedness of S, the vectors $x_n := Sz_n$ converge to Sz. Since $Ax_n \to y$ and A is closed, we obtain $Sz \in \mathrm{D}(A)$ and $ASz = y$; i.e., $z \in \mathrm{D}(C)$ and $Cz = y$. $\square$

We state simple consequences which are needed in the next section.

Corollary 1.8. *Let A be linear on X and $\lambda \in \mathbb{F}$. Then the following assertions hold.*

a) The operator A is closed on X if and only if $\lambda I - A$ is closed on X.

b) Let $\lambda I - A$ be bijective with $(\lambda I - A)^{-1} \in \mathcal{B}(X)$. Then A is closed.

Proof. Assertion a) follows from Proposition 1.7 since $A = -((\lambda I - A) - \lambda I)$. For the second part, Lemma 1.4 proves that $\lambda I - A$ is closed, and then assertion a) yields b). $\square$

The following examples show that closedness can be lost when taking sums or products of closed operators. See the exercises for further related results.

Example 1.9. a) Let $E = C_b(\mathbb{R}^2)$ and $A_k = \partial_k$ with

$$\mathrm{D}(A_k) = \{f \in E \,|\, \text{the partial derivative } \partial_k f \text{ exists and belongs to } E\},$$

[1] This statement was given in the lectures in a somewhat sloppy form.

for $k \in \{1, 2\}$. Set $B = \partial_1 + \partial_2$ on

$$\mathrm{D}(B) = \mathrm{D}(A_1) \cap \mathrm{D}(A_2) = C_b^1(\mathbb{R}^2) := \{f \in C^1(\mathbb{R}^2) \,|\, f, \partial_1 f, \partial_2 f \in E\}.$$

By an exercise, A_1 and A_2 are closed. However, B is not closed.

Indeed, take $\phi_n \in C_b^1(\mathbb{R})$ converging uniformly to some $\phi \in C_b(\mathbb{R})\backslash C^1(\mathbb{R})$. Set $f_n(x, y) = \phi_n(x - y)$ and $f(x, y) = \phi(x - y)$ for $(x, y) \in \mathbb{R}^2$ and $n \in \mathbb{N}$. We then obtain $f \in E$, $f_n \in \mathrm{D}(B)$, $\|f_n - f\|_\infty = \|\phi_n - \phi\|_\infty \to 0$ and $Bf_n = \phi_n' - \phi_n' = 0 \to 0$ as $n \to \infty$, but $f \notin \mathrm{D}(B)$.

b) Let $X = C([0, 1])$, $Af = f'$ with $\mathrm{D}(A) = C^1([0, 1])$ and $m \in C([0, 1])$ with $m = 0$ on $[0, 1/2]$. Define $T \in \mathcal{B}(X)$ by $Tf = mf$ for all $f \in X$. Then the operator TA with $\mathrm{D}(TA) = \mathrm{D}(A)$ is not closed.

To see this, take maps $f_n \in \mathrm{D}(A)$ with $f_n = 1$ on $[1/2, 1]$ and $f_n \to f$ in X with $f \notin C^1([0, 1])$. Then, $TAf_n = mf_n' = 0$ tends to 0, but $f \notin \mathrm{D}(A)$. $\Diamond$

1.2. The spectrum

We start with the basic definitions of spectral theory. For deeper investigations of spectra one has to take complex numbers $\mathbb{F} = \mathbb{C}$. However, several results are also true for the real case $\mathbb{F} = \mathbb{R}$. Since this case is needed sometimes, we develop the theory for $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ as long as it makes sense.

Definition 1.10. *Let A be a closed operator on X. The* resolvent set *of A is given by*

$$\rho(A) = \{\lambda \in \mathbb{F} \,|\, \lambda I - A : \mathrm{D}(A) \to X \text{ is bijective}\}, \tag{1.1}$$

and its spectrum *by*

$$\sigma(A) = \mathbb{F}\backslash\rho(A).$$

We further define the point spectrum *of A by*

$$\sigma_{\mathrm{p}}(A) =, \{\lambda \in \mathbb{F} \,|\, \exists\, v \in \mathrm{D}(A)\backslash\{0\} \text{ with } \lambda v = Av\} \subseteq \sigma(A),$$

where we call $\lambda \in \sigma_{\mathrm{p}}(A)$ an eigenvalue *of A and the corresponding v an* eigenvector *or* eigenfunction *of A. For $\lambda \in \rho(A)$ the operator*

$$R(\lambda, A) := (\lambda I - A)^{-1} : X \to X$$

and the set $\{R(\lambda, A) \,|\, \lambda \in \rho(A)\}$ are called the resolvent*.*

Eigenvalues are usually much easier to compute than general $\lambda \in \sigma(A)$. So they may help a lot to determine the spectrum. However, Examples 1.21 and 1.25 yield unbounded and bounded operators with empty point spectrum and non-void (even 'large') spectrum, where $\dim X = \infty$. Observe that computing the resolvent amounts to solve the equation $\lambda u - Au = f$ for each given $f \in X$ and a unique $u \in \mathrm{D}(A)$. In the simple examples below this can be done explicitely, which is one way to calculate the spectrum. Before we note that resolvent operators are automatically bounded.

Remark 1.11. a) Let A be closed on X and $\lambda \in \rho(A)$. Note that the resolvent $R(\lambda, A)$ has the range $\mathrm{D}(A)$. Corollary 1.8 and Lemma 1.4 further show that $R(\lambda, A)$ is closed on X, and thus it belongs to $\mathcal{B}(X)$ by Theorem 1.5. In fact, even $R(\lambda, A) : X \to [\mathrm{D}(A)]$ is bounded, see Theorem 1.13.

b) Let A be a linear operator such that $\lambda I - A : \mathrm{D}(A) \to X$ has a bounded inverse for some $\lambda \in \mathbb{F}$. Then A is closed by Corollary 1.8. In this case, the closedness assumption in Definition 1.10 is redundant. $\Diamond$

In the literature, the spectrum is sometimes defined for general linear operators, assuming in addition the boundedness of $(\lambda I - A)^{-1}$ in (1.1).

For unbounded A, the spectrum can be empty or equal to $\mathbb{F}$, as we see in the next examples which also demonstrate the influence of boundary conditions. We set $e_\lambda(t) = \mathrm{e}^{\lambda t}$ for $\lambda \in \mathbb{C}$, $t \in J$, and any interval $J \subseteq \mathbb{R}$.

EXAMPLE 1.12. a) Let $X = \mathbb{C}^m$ and $T \in \mathcal{B}(X)$. Then $\sigma(T)$ only consists of the eigenvalues $\lambda_1, \dots, \lambda_k$ of T, where $1 \leqslant k \leqslant m$. (See linear algebra.)

b) Let $X = C([0,1])$ and $Au = u'$ with $\mathrm{D}(A) = C^1([0,1])$. Then $\sigma(A) = \sigma_{\mathrm{p}}(A) = \mathbb{F}$. Indeed, e_λ belongs to $\mathrm{D}(A)$ and $Ae_\lambda = \lambda e_\lambda$ for each $\lambda \in \mathbb{F}$.

c) Let $X = C([0,1])$ and $Au = u'$ with $\mathrm{D}(A) = \{u \in C^1([0,1]) \,|\, u(0) = 0\}$. Then A is closed by Example 1.2. Moreover, $\sigma(A)$ is empty. In fact, let $\lambda \in \mathbb{C}$ and $f \in X$. We then have $u \in \mathrm{D}(A)$ and $(\lambda I - A)u = f$ if and only if $u \in C^1([0,1])$, $u'(t) = \lambda u(t) - f(t)$ for $t \in [0,1]$, and $u(0) = 0$, which is equivalent to

$$u(t) = -\int_0^t \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s =: (R_\lambda f)(t),$$

for all $0 \leqslant t \leqslant 1$. Hence, $\sigma(A) = \emptyset$ and $R(\lambda, A) = R_\lambda$. $\Diamond$

Let $U \subseteq \mathbb{F}$ be open. The *derivative* of $f : U \to Y$ at $\lambda \in U$ is given by

$$f'(\lambda) = \lim_{\mu \to \lambda} \frac{1}{\mu - \lambda}(f(\mu) - f(\lambda)) \ \in Y,$$

if the limit exists in Y. In the next theorem we collect fundamental properties of the spectrum and the resolvent of closed operators.

THEOREM 1.13. *Let A be a closed operator on X and let $\lambda \in \rho(A)$. Then the following assertions hold.*

a) $AR(\lambda, A) = \lambda R(\lambda, A) - I$, $AR(\lambda, A)x = R(\lambda, A)Ax$ for $x \in \mathrm{D}(A)$, and

$$\frac{1}{\mu - \lambda}(R(\lambda, A) - R(\mu, A)) = R(\lambda, A)R(\mu, A) = R(\mu, A)R(\lambda, A) \tag{1.2}$$

if $\mu \in \rho(A)\backslash\{\lambda\}$. The formula in display is called the resolvent equation.

b) The spectrum $\sigma(A)$ is closed, where $B(\lambda, 1/\|R(\lambda,A)\|) \subseteq \rho(A)$ and

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

if $|\lambda - \mu| < 1/\|R(\lambda,A)\| =: r_\lambda$. This series converges in $\mathcal{B}(X, [\mathrm{D}(A)])$, absolutely and uniformly on $B(\lambda, \delta r_\lambda)$ for each $\delta \in (0,1)$. Moreover, we have

$$\|R(\mu, A)\|_{\mathcal{B}(X,[\mathrm{D}(A)])} \leqslant \frac{c(\lambda)}{1 - \delta}$$

for all $\mu \in B(\lambda, \delta r_\lambda)$ and a constant $c(\lambda)$ given by (1.3).

c) The function $\rho(A) \to \mathcal{B}(X, [\mathrm{D}(A)])$; $\lambda \mapsto R(\lambda, A)$, is infinitely often differentiable with

$$\left(\tfrac{\mathrm{d}}{\mathrm{d}\lambda}\right)^n R(\lambda, A) = (-1)^n n!\, R(\lambda, A)^{n+1} \qquad \textit{for every } \ n \in \mathbb{N}.$$

d) $\|R(\lambda, A)\| \geqslant \frac{1}{\mathrm{d}(\lambda, \sigma(A))}$.

PROOF. a) The first claims follow from the identities

$$x = (\lambda I - A)R(\lambda, A)x = R(\lambda, A)(\lambda I - A)x,$$

where $x \in X$ in the first and $x \in \mathrm{D}(A)$ in the second one. For $\mu \in \rho(A)$, we further have

$$\begin{aligned}(\lambda R(\lambda, A) - AR(\lambda, A))R(\mu, A) &= R(\mu, A),\\ R(\lambda, A)(\mu R(\mu, A) - AR(\mu, A)) &= R(\lambda, A).\end{aligned}$$

Formula (1.2) follows by subtracting the above equations and also interchanging λ and μ.

b) Let $|\mu - \lambda| \leqslant \delta/\|R(\lambda,A)\|$ for some $\delta \in (0, 1)$ and $x \in X$ with $\|x\| \leqslant 1$. Also using statement a), we compute

$$\begin{aligned}&\|(\lambda - \mu)^n R(\lambda, A)^{n+1}x\|_A\\ &\qquad \leqslant \frac{\delta^n}{\|R(\lambda, A)\|^n}\left(\|AR(\lambda, A)R(\lambda, A)^n x\| + \|R(\lambda, A)^{n+1}x\|\right)\\ &\qquad \leqslant \delta^n\left(\|\lambda R(\lambda, A)\| + 1 + \|R(\lambda, A)\|\right) =: \delta^n c(\lambda).\end{aligned} \tag{1.3}$$

Due to this inequality and Lemma 4.23 in [**FA**], the series R_μ converges and can be estimated as asserted. Part a) yields

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

Employing this fact and that A belongs to $\mathcal{B}([\mathrm{D}(A)], X)$, we infer

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

and similarly $R_\mu(\mu I - A)x = x$ for all $x \in \mathrm{D}(A)$. Hence, μ is contained in $\rho(A)$ and $R_\mu = R(\mu, A)$. This means that $\rho(A)$ is open and $\sigma(A)$ is closed.

Assertion c) is a consequence of the power series expansion, as in the scalar case. Statement b) also implies d). $\square$

By the next result, the spectrum of a multiplication operator is directly given via the multiplier. As a by-product we see that each closed set $S \subseteq \mathbb{F}$ occurs as the spectrum of a closed operator, complementing Theorem 1.13 b).

PROPOSITION 1.14. *Let $\Omega \subseteq \mathbb{R}^d$ be non-empty, $m \in C(\Omega)$, $E = C_b(\Omega)$, and $Af = mf$ with $\mathrm{D}(A) = \{f \in E \mid mf \in E\}$. Then A is closed,*

$$\sigma(A) = \overline{m(\Omega)},$$

and $R(\lambda, A)g = \frac{1}{\lambda - m}g$ for all $\lambda \in \rho(A)$ and $g \in E$.

For every closed (resp., non-empty and compact) subset $S \subseteq \mathbb{F}$ there is a closed (resp., bounded) operator B on a Banach space with $\sigma(B) = S$.

PROOF. The closedness of A can be shown as in Remark 1.6. Let $g \in E$. If $u \in \mathrm{D}(A)$ satisfies $\lambda u - Au = (\lambda - m)u = g$, we obtain $u(x) = (\lambda - m(x))^{-1}g(x)$ for all $x \in \Omega$ with $\lambda \neq m(x)$. So we first take $\lambda \notin \overline{m(\Omega)}$. Then the function $f := \frac{1}{\lambda - m}g$ belongs to E and satisfies $\lambda f - mf = g$ so that $mf = \lambda f - g \in E$. As a result, f is an element of $\mathrm{D}(A)$ and it is the unique solution in $\mathrm{D}(A)$ of the equation $\lambda u - Au = g$. This means that $\lambda \in \rho(A)$, $R(\lambda, A)g = \frac{1}{\lambda - m}g$, and $\sigma(A) \subseteq \overline{m(\Omega)}$.

In the case that $\lambda = m(x)$ for some $x \in \Omega$, we compute
$$((\lambda I - A)f)(x) = \lambda f(x) - m(x)f(x) = 0$$
for every $f \in \mathrm{D}(A)$. Consequently, $\lambda I - A$ is not surjective and so $\lambda \in \sigma(A)$; i.e., $m(\Omega) \subseteq \sigma(A)$. The closedness of the spectrum now yields $\sigma(A) = \overline{m(\Omega)}$.

The final assertion follows from Example 1.12 c) if $S = \emptyset$. Otherwise, consider $\Omega = S$ and $m(x) = x$ where one identifies $\mathbb{R}^2$ with $\mathbb{C}$ if $\mathbb{F} = \mathbb{C}$. Define A and E as above. Then $\sigma(A)$ is equal to S, and A is bounded if S is compact (where $C_b(S) = C(S)$). □

A similar result is valid in L^p-spaces, cf. Example IX.2.6 in [**Co2**]. We next study a variant of the first derivative with a non-trivial spectrum. Here we use the closedness of the spectrum since we can compute eigenvalues only for a (dense) subset of $\sigma(A)$.

Example 1.15. Let $X = C_0(\mathbb{R}_{\geqslant 0}) = \{f \in C(\mathbb{R}_{\geqslant 0}) \mid \lim_{t\to\infty} f(t) = 0\}$ with $\mathbb{F} = \mathbb{C}$ be endowed with $\|\cdot\|_\infty$. On X we consider $Af = f'$ with
$$\mathrm{D}(A) = C_0^1(\mathbb{R}_{\geqslant 0}) = \big\{f \in C^1(\mathbb{R}_{\geqslant 0}) \,\big|\, f, f' \in X\big\}.$$
As in Example 1.2 one sees that A is closed. Moreover, we have $\sigma(A) = \{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda \leqslant 0\}$ and $\sigma_{\mathrm{p}}(A) = \{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda < 0\} =: \mathbb{C}_-$.

Proof. First note that for $\lambda \in \mathbb{C}_-$ the function e_λ belongs to $\mathrm{D}(A)$ and $Ae_\lambda = e_\lambda' = \lambda e_\lambda$. This means that $\mathbb{C}_- \subseteq \sigma_{\mathrm{p}}(A) \subseteq \sigma(A)$, and hence $\overline{\mathbb{C}_-} \subseteq \sigma(A)$ by the closedness of the spectrum.

Next, let $\operatorname{Re}\lambda > 0$ and $f \in X$. We then have $u \in \mathrm{D}(A)$ and $\lambda u - Au = f$ if and only if $u \in X \cap C^1(\mathbb{R}_{\geqslant 0})$ and $u'(t) = \lambda u(t) - f(t)$ for all $t \geqslant 0$. This equation is solved by
$$u(t) = \int_t^\infty \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s =: (R_\lambda f)(t), \quad t \geqslant 0.$$
We still have to check $R_\lambda f \in X$. Let $\varepsilon > 0$. There is a number $t_\varepsilon \geqslant 0$ such that $|f(s)| \leqslant \varepsilon$ for all $s \geqslant t_\varepsilon$. We can now estimate
$$|R_\lambda f(t)| \leqslant \int_t^\infty \mathrm{e}^{(\operatorname{Re}\lambda)(t-s)} |f(s)|\,\mathrm{d}s \leqslant \varepsilon \int_0^\infty \mathrm{e}^{-\operatorname{Re}\lambda r}\,\mathrm{d}r = \frac{\varepsilon}{\operatorname{Re}\lambda},$$
for all $t \geqslant t_\varepsilon$, where we substituted $r = s - t$. As a result, u is contained in $\mathrm{D}(A)$ and solves $\lambda u - Au = f$.

Let $v \in \mathrm{D}(A)$ be another solution. Then $w := u - v \in \mathrm{D}(A)$ satisfies $w' = \lambda w$ and hence $w = ce_\lambda$ for some $c \in \mathbb{C}$. Because of $\operatorname{Re}\lambda > 0$ the function e_λ does not belong to X, implying $w = 0$ and the uniqueness of solutions in $\mathrm{D}(A)$. We have shown that $\lambda \in \rho(A)$ with $R_\lambda = R(\lambda, A)$, and hence $\sigma(A) = \overline{\mathbb{C}_-}$.

Finally, assume there is a number $\lambda \in \mathrm{i}\mathbb{R}$ and a function $v \in C_0^1(\mathbb{R}_{\geqslant 0})$ with $v' = \lambda v$. It follows $v(t) = \mathrm{e}^{\lambda(t)} v(0)$ and so $|v(t)| = |v(0)|$ for all $t \geqslant 0$. Letting $t \to \infty$, we infer that $|v(0)| = 0$ and thus $v = 0$. Therefore A has no eigenvalues on $\mathrm{i}\mathbb{R}$. □

Complementing Theorem 1.13, we state additional properties of the spectrum if the operator is bounded, e.g., it is compact. Using this fact, for $T \in \mathcal{B}(X)$ we define the *spectral radius*
$$\mathrm{r}(T) = \max\big\{|\lambda| \,\big|\, \lambda \in \sigma(T)\big\}.$$

THEOREM 1.16. *Let $T \in \mathcal{B}(X)$. Then $\sigma(T)$ is a compact set and for $\lambda \in \mathbb{F}$ with $|\lambda| > \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$ we have*

$$R(\lambda, T) = \sum_{n=0}^{\infty} \lambda^{-n-1} T^n =: R_\lambda.$$

Let also $\mathbb{F} = \mathbb{C}$. Then $\sigma(T)$ is non-empty and the spectral radius is given by

$$\mathrm{r}(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|T^n\|^{1/n} \leqslant \|T\|.$$

PROOF. 1) Since $\|T^{n+m}\| \leqslant \|T^n\| \, \|T^m\|$ for all $n, m \in \mathbb{N}$, by an elementary lemma (see Lemma VI.1.4 in [**We**]) there exists the limit

$$r := \lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|T^n\|^{1/n} \leqslant \|T\|.$$

Let $|\lambda| > r$. We estimate

$$\limsup_{n\to\infty} \|\lambda^{-n} T^n\|^{1/n} = \frac{1}{|\lambda|} \lim_{n\to\infty} \|T^n\|^{1/n} = \frac{r}{|\lambda|} < 1.$$

Lemma 4.23 in [**FA**] now yields the convergence in $\mathcal{B}(X)$ of the series R_λ. Moreover, we have

$$(\lambda I - T) R_\lambda = \sum_{n=0}^{\infty} \lambda^{-n} T^n - \sum_{n=0}^{\infty} \lambda^{-n-1} T^{n+1} = I,$$

and similarly $R_\lambda(\lambda I - T) = I$. Hence, λ belongs to $\rho(T)$ and $R_\lambda = R(\lambda, T)$. Due to its closedness, the spectrum $\sigma(T) \subseteq \overline{B}(0, r)$ is compact. Therefore $\mathrm{r}(T)$ exists as the maximum of a compact subset of $\mathbb{R}$, and $\mathrm{r}(T) \leqslant r$.

2) Let $\mathbb{F} = \mathbb{C}$. Take $\Phi \in \mathcal{B}(X)^\star$ and define $f_\Phi(\lambda) := \Phi(R(\lambda, T))$ for $\lambda \in D = \mathbb{C}\backslash\overline{B}(0, \mathrm{r}(T))$. Note that $f_\Phi : D \to \mathbb{C}$ is complex differentiable and

$$f_\Phi(\lambda) = \sum_{n=0}^{\infty} \lambda^{-n-1} \Phi(T^n) =: S_\lambda \qquad \text{if } \; |\lambda| > r.$$

By Theorem V.1.11 in [**Co1**], there are unique coeffcients $a_m \in \mathbb{C}$ with

$$f_\Phi(\lambda) = \sum_{m=-\infty}^{\infty} a_m \lambda^m \qquad \text{for } \; \lambda \in D.$$

The series S_λ thus converges for all $\lambda \in D$, and so

$$\forall\, \lambda \in D, \; \Phi \in \mathcal{B}(X)^\star : \qquad \sup_{n\in\mathbb{N}} |\Phi(\lambda^{-n-1} T^n)| < \infty.$$

A corollary to the uniform boundedness principle (see Corollary 5.12 in [**FA**]) thus yields that

$$c(\lambda) := \sup_{n\in\mathbb{N}} \|\lambda^{-n-1} T^n\| < \infty$$

for each $\lambda \in D$. This fact leads to

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \lim_{n\to\infty} |\lambda| \, (|\lambda| \, \|\lambda^{-n-1} T^n\|)^{1/n} \leqslant |\lambda| \lim_{n\to\infty} (|\lambda| \, c(\lambda))^{1/n} = |\lambda|$$

for all $|\lambda| > \mathrm{r}(T)$. Together with step 1), we arrive at $r = \mathrm{r}(T)$.

3) Suppose that $\sigma(T) = \emptyset$. The functions f_Φ from part 2) are now holomorphic on $\mathbb{C}$ for every $\Phi \in \mathcal{B}(X)^\star$. Step 1) implies that

$$|f_\Phi(\lambda)| \leqslant \|\Phi\|\,|\lambda|^{-1} \sum_{n=0}^{\infty} \frac{\|T\|^n}{|\lambda|^n} \leqslant \frac{2\|\Phi\|}{|\lambda|},$$

for all $\lambda \in \mathbb{C}$ with $|\lambda| \geqslant 2\|T\|$. Therefore, f_Φ is bounded and hence constant by Liouville's theorem from complex analysis. The above estimate then shows that $\Phi(R(\lambda, T)) = 0$ for all $\lambda \in \mathbb{C}$ and $\Phi \in \mathcal{B}(X)^\star$. Employing the Hahn–Banach theorem (see Corollary 5.10 in [**FA**]), we obtain $R(\lambda, T) = 0$, which is impossible since $R(\lambda, T)$ is injective and $X \neq \{0\}$. $\square$

We note that already on $X = \mathbb{R}^2$ the matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has empty real spectrum, but the complex eigenvalues $\pm\mathrm{i}$. The next example first shows that the spectral radius can be much smaller than the norm, and then uses $\mathrm{r}(T) \leqslant \|T\|$ for a basic operator. (We let $\mathbb{F} = \mathbb{C}$ in the example.)

Example 1.17. a) We define the Volterra operator V on $X = C([0,1])$ by

$$Vf(t) = \int_0^t f(s)\,\mathrm{d}s$$

for $t \in [0,1]$ and $f \in X$. Then V belongs to $\mathcal{B}(X)$ with $\|V^n\| \leqslant 1/(n!)$ since

$$|V^n f(t)| \leqslant \int_0^t \int_0^{s_1} \dots \int_0^{s_{n-1}} \|f\|_\infty \,\mathrm{d}s_n \dots \mathrm{d}s_1 \leqslant \frac{1}{n!}\|f\|_\infty$$

for all $n \in \mathbb{N}$, $t \in [0,1]$, and $f \in X$. Moreover, taking $f = \mathbb{1}$ we obtain $\|V^n\| \geqslant \|V^n \mathbb{1}\|_\infty = 1/(n!)$ and so $\|V^n\| = 1/(n!)$. Theorem 1.16 thus yields

$$\mathrm{r}(V) = \lim_{n\to\infty} (n!)^{-1/n} = 0 < 1 = \|V\| \quad \text{and} \quad \sigma(V) = \{0\}.$$

Observe that $\sigma_{\mathrm{p}}(V) = \emptyset$ since $Vf = 0$ implies that $f = (Vf)' = 0$.

b) Let left shift L given by $Lx = (x_{n+1})$ on $X \in \{c_0, \ell^p \,|\, 1 \leqslant p \leqslant \infty\}$ has the spectrum $\sigma(L) = \overline{B}(0,1)$. We further obtain $\sigma_{\mathrm{p}}(L) = B(0,1)$ if $X \neq \ell^\infty$ and $\sigma_p(L) = \overline{B}(0,1)$ if $X = \ell^\infty$.

Proof. The operator $L \in \mathcal{B}(X)$ has norm 1 (see Example 2.9 in [**FA**]), and so $\sigma(L) \subseteq \overline{B}(0,1)$. Clearly, $L(1,0,\dots) = 0$. Let $0 < |\lambda| \leqslant 1$. Observe that $Lv = \lambda v$ is equivalent to $v_{n+1} = \lambda v_n$ for all $n \in \mathbb{N}$ and hence to $v_n = \lambda^{n-1} v_1$. Choosing $v_1 = 1$, we obtain the eigensequences $v = (\lambda^n)_{n\in\mathbb{N}}$ in ℓ^∞ for λ, and thus $\sigma(L) = \sigma_p(L) = \overline{B}(0,1)$ for $X = \ell^\infty$. Now, let$X \neq \ell^\infty$. Here we have $v \in X$ if and only if $0 < |\lambda| < 1$. It follows $B(0,1) \subseteq \sigma_{\mathrm{p}}(L) \subseteq \sigma(L)$ and $\sigma(L) = \overline{B}(0,1)$ by the closedness of the spectrum. $\square$

We decompose the spectrum into parts related to eigenvalues, cf. Theorem 1.24. In this context also other definitions are used in the literature.

Definition 1.18. *Let A be a linear operator on X. Then*

$$\sigma_{\mathrm{ap}}(A) = \big\{\lambda \in \mathbb{F} \,\big|\, \exists\, x_n \in \mathrm{D}(A) \textit{ with } \|x_n\| = 1 \textit{ for all } n \in \mathbb{N} \textit{ and } \lambda x_n - Ax_n \to 0 \textit{ as } n \to \infty\big\}$$

is the approximate point spectrum *of A and*

$$\sigma_{\mathrm{r}}(A) = \big\{\lambda \in \mathbb{F} \,\big|\, (\lambda I - A)\,\mathrm{D}(A) \textit{ is not dense in } X\big\}$$

the residual spectrum *of A.*

One calls $\lambda \in \sigma_{\mathrm{ap}}(A)$ an *approximate eigenvalue* and the corresponding x_n *approximate eigenvectors*. If one has $\lambda \in \mathbb{C}$ and $x_n \in \mathrm{D}(A)$ with $\|x_n\| \geqslant \delta > 0$ for all $n \in \mathbb{N}$ and $\lambda x_n - Ax_n \to 0$ as $n \to \infty$, then λ belongs to $\sigma_{\mathrm{ap}}(A)$ with approximate eigenvectors $\|x_n\|^{-1}x_n$, since $\|x_n\|^{-1} \leqslant 1/\delta$.

In the next result we characterize $\sigma_{\mathrm{ap}}(A)$ and decompose the spectrum into $\sigma_{\mathrm{ap}}(A)$ and $\sigma_{\mathrm{r}}(A)$. The last statement implies that $\sigma_{\mathrm{ap}}(A)$ is non-empty if $\sigma(A) \notin \{\emptyset, \mathbb{F}\}$.

Proposition 1.19. *Let A be closed on X. The following assertions are true (with possibly non-disjoint unions).*

a) $\sigma_{\mathrm{ap}}(A) = \sigma_{\mathrm{p}}(A) \cup \{\lambda \in \mathbb{F} \,|\, (\lambda I - A)\,\mathrm{D}(A)$ *is not closed in* $X\}$.

b) $\sigma(A) = \sigma_{\mathrm{ap}}(A) \cup \sigma_{\mathrm{r}}(A)$.

c) $\partial\sigma(A) \subseteq \sigma_{\mathrm{ap}}(A)$.

Proof. 1) Let $\lambda \notin \sigma_{\mathrm{ap}}(A)$. Note that this fact holds if and only if there is a constant $c > 0$ with $\|\lambda x - Ax\| \geqslant c\|x\|$ for all $x \in \mathrm{D}(A)$. This lower estimate implies that $\lambda \notin \sigma_{\mathrm{p}}(A)$. Moreover, let $y_n := \lambda x_n - Ax_n \to y$ in X as $n \to \infty$ for some $x_n \in \mathrm{D}(A)$. Then the lower estimate shows that (x_n) is Cauchy in X, and so x_n tends to some x in X. Hence, $Ax_n = \lambda x_n - y_n$ converges to $\lambda x - y$, so that x belongs to $\mathrm{D}(A)$ and $\lambda x - Ax = y$ by the closedness of A. Consequently, $(\lambda I - A)\,\mathrm{D}(A)$ is closed.

Conversely, let $(\lambda I - A)\,\mathrm{D}(A)$ be closed and $\lambda \notin \sigma_{\mathrm{p}}(A)$. Then the inverse $(\lambda I - A)^{-1}$ exists and is closed on its closed domain $(\lambda I - A)\,\mathrm{D}(A)$ due to Lemma 1.4. The closed graph theorem 1.5 then yields the boundedness of $(\lambda I - A)^{-1}$. It follows

$$\|x\| = \|(\lambda I - A)^{-1}(\lambda I - A)x\| \leqslant C\|(\lambda I - A)x\|$$

for all $x \in \mathrm{D}(A)$ and a constant $C > 0$. This means that $\lambda \notin \sigma_{\mathrm{ap}}(A)$. We thus have shown assertion a), which implies b).

2) Let $\lambda \in \partial\sigma(A)$. Then there are points λ_n in $\rho(A)$ with $\lambda_n \to \lambda$ as $n \to \infty$. By Theorem 1.13 d), the norms $\|R(\lambda_n, A)\|$ tend to ∞ as $n \to \infty$, and there thus exist $y_n \in X$ with $\|y_n\| = 1$ for all $n \in \mathbb{N}$ and $0 \neq a_n := \|R(\lambda_n, A)y_n\| \to \infty$ as $n \to \infty$. Set $x_n = \frac{1}{a_n}R(\lambda_n, A)y_n \in \mathrm{D}(A)$. We then have $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\lambda x_n - Ax_n = (\lambda - \lambda_n)x_n + \frac{1}{a_n}y_n$ converges to 0 as $n \to \infty$. As a result, λ is an element of $\sigma_{\mathrm{ap}}(A)$. $\square$

In the next result we determine the spectra of certain operators which (formally) arise as functions $f(A)$ of A, namely the resolvent of A, where $f(\mu) = (\lambda - \mu)^{-1}$ for $\lambda \in \rho(A)$ and $\mu \in \sigma(A)$, as well as an affine transformation of A, where $f(\mu) = \alpha\mu + \beta$ for $\alpha, \beta \in \mathbb{F}$. This often useful, as seen below, and will be generalized in Section 4.2 and Chapter 5.

Proposition 1.20. *Let A be closed on X, $\lambda \in \rho(A)$, $\alpha \in \mathbb{F}\backslash\{0\}$ and $\beta \in \mathbb{F}$. Then the following assertions hold.*

a) $\sigma(R(\lambda, A))\backslash\{0\} = (\lambda - \sigma(A))^{-1} = \left\{\frac{1}{\lambda - \nu} \,\middle|\, \nu \in \sigma(A)\right\}$.

b) $\sigma_j(R(\lambda, A))\backslash\{0\} = (\lambda - \sigma_j(A))^{-1}$ *for* $j \in \{\mathrm{p}, \mathrm{ap}, \mathrm{r}\}$.

c) If x is an eigenvector for the eigenvalue $\mu \neq 0$ of $R(\lambda, A)$, then $y = \mu R(\lambda, A)x$ is an eigenvector for the eigenvalue $\nu = \lambda - 1/\mu$ of A. If $y \in \mathrm{D}(A)$

is an eigenvector for the eigenvalue $\nu = \lambda - 1/\mu$ of A with $\mu \in \mathbb{F}\backslash\{0\}$, then $x = \mu^{-1}(\lambda y - Ay)$ is an eigenvector for the eigenvalue μ of $R(\lambda, A)$.

d) $\mathrm{r}(R(\lambda, A)) = 1/\mathrm{d}(\lambda,\sigma(A))$.

e) If A is unbounded, then $0 \in \sigma(R(\lambda, A))$.

f) $\sigma(\alpha A+\beta I) = \alpha\sigma(A)+\beta$ and $\sigma_j(\alpha A+\beta I) = \alpha\sigma_j(A)+\beta$, $j \in \{\mathrm{p}, \mathrm{ap}, \mathrm{r}\}$.

Proof. Let $\mu \in \mathbb{F}\backslash\{0\}$. Taking out the bijective map $\mu R(\lambda, A) : X \to \mathrm{D}(A)$, we obtain

$$\mu I - R(\lambda, A) = \Big(\big(\lambda - \tfrac{1}{\mu}\big)I - A\Big)\mu R(\lambda, A). \tag{1.4}$$

Hence, the bijectivity of $\mu I - R(\lambda, A) : X \to X$ is equivalent to that of $(\lambda - \frac{1}{\mu})I - A : \mathrm{D}(A) \to X$. As a result, μ belongs to $\rho(R(\lambda, A))$ if and only if $\lambda - \frac{1}{\mu}$ belongs to $\rho(A)$ if and only if $\mu = (\lambda - \nu)^{-1}$ for some $\nu \in \rho(A)$. We have shown part a).

In the same way, one derives assertion b) for $j = \mathrm{p}$, assertion c) and that $\mu I - R(\lambda, A)$ and $(\lambda - \frac{1}{\mu})I - A$ have the same range. Using also Proposition 1.19, we then deduce statement b) also for $j = \mathrm{ap}$ and $j = \mathrm{r}$.

Assertion d) is a consequence of a). In part e), the inverse $R(\lambda, A)^{-1} = \lambda I - A$ is unbounded so that $0 \in \sigma(R(\lambda, A))$. Similar as a) and b), the last statement follows from the equality

$$\lambda I - (\alpha A + \beta I) = \alpha\big(\tfrac{\lambda-\beta}{\alpha}I - A\big). \qquad \square$$

Approximate eigenvectors are often 'close' to an eigenvector of the operator acting on a 'larger' space. Such a fact can be used to construct them, as in the following basic examples.

Example 1.21. a) Let $X = L^p(\mathbb{R})$, $1 \leqslant p \leqslant \infty$, with $\mathbb{F} = \mathbb{C}$ and the (left) translation $T(t)$ be given by $(T(t)f)(s) = f(s + t)$ for $s \in \mathbb{R}$, $f \in X$, and $t \in \mathbb{R}$. Then $\sigma(T(t)) = \partial B(0, 1)$ for $t \neq 0$.

Proof. Recall from Example 4.12 in [**FA**] that $T(t)$ is an isometry on X with inverse $(T(t))^{-1} = T(-t)$ for every $t \in \mathbb{R}$. Using Theorem 1.16, we deduce $\sigma(T(t)) \subseteq \overline{B}(0, 1)$. Proposition 1.20 further yields $\sigma(T(t))^{-1} = \sigma(T(t)^{-1}) = \sigma(T(-t)) \subseteq \overline{B}(0, 1)$ so that $\sigma(T(t)) \subseteq \partial B(0, 1)$ for all $t \in \mathbb{R}$. Fix $t \neq 0$ and take $\lambda \in \mathrm{i}\mathbb{R}$. Then e_λ belongs to $C_b(\mathbb{R}) \subseteq L^\infty(\mathbb{R})$ and

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

for all $s \in \mathbb{R}$. We infer $\sigma(T(t)) = \sigma_{\mathrm{p}}(T(t)) = \partial B(0, 1)$ for $p = \infty$.

If $p \in [1, \infty)$, we use e_λ to construct approximate eigenfunctions if $p < \infty$. For $n \in \mathbb{N}$ set $f_n = n^{-1/p}\mathbb{1}_{[0,n]}e_\lambda$. We compute $\|f_n\|_p = n^{-1/p}\|\mathbb{1}_{[0,n]}\|_p = 1$ and, employing the above formula in display,

$$\|T(t)f_n - \mathrm{e}^{\lambda t}f_n\|_p = n^{-1/p}\|\mathrm{e}^{\lambda t}(\mathbb{1}_{[-t,n-t]} - \mathbb{1}_{[0,n]})e_\lambda\|_p = n^{-1/p}|2t|^{-1/p} \longrightarrow 0,$$

as $n \to \infty$. It follows $\sigma(T(t)) = \partial B(0, 1)$ if $t \neq 0$. $\square$

b) Let $X = C_0(\mathbb{R})$ with $\mathbb{F} = \mathbb{C}$ and $Au = u'$ with $\mathrm{D}(A) = C_0^1(\mathbb{R}) := \{u \in C^1(\mathbb{R}) \,|\, u, u' \in C_0(\mathbb{R})\}$. Then $\sigma(A) = \mathrm{i}\mathbb{R}$ and $\sigma_{\mathrm{p}}(A) = \emptyset$.

Proof. As in Example 1.15 one sees that $\lambda \in \rho(A)$ if $\operatorname{Re}\lambda \neq 0$ with

$$R(\lambda, A)f(t) = \int_t^\infty \mathrm{e}^{\lambda(t-s)}f(s)\,\mathrm{d}s \qquad \text{if} \quad \operatorname{Re}\lambda > 0 \qquad \text{and}$$

$$R(\lambda, A)f(t) = -\int_{-\infty}^{t} \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s \qquad \text{if} \quad \operatorname{Re}\lambda < 0$$

for all $t \in \mathbb{R}$ and $f \in X$. Let $\operatorname{Re}\lambda = 0$. Then λ is not an eigenvalue, cf. Example 1.15. Choose $\varphi_n \in C_c^1(\mathbb{R})$ with $\|\varphi_n'\|_\infty \leqslant 1/n$ and $\|\varphi_n\|_\infty = 1$, and set $u_n = \varphi_n e_\lambda$ for all $n \in \mathbb{N}$. Note that $\|u_n\|_\infty = 1$, $u_n \in \mathrm{D}(A)$, and

$$Au_n = \varphi_n' e_\lambda + \varphi_n e_\lambda' = \varphi_n' e_\lambda + \lambda u_n.$$

Since $\|\varphi_n' e_\lambda\|_\infty \leqslant 1/n$, we obtain $\lambda \in \sigma_{\mathrm{ap}}(\mathbb{R})$. □

We now introduce the adjoint of a densely defined linear operator in order to obtain a convenient description of the residual spectrum, for instance.

Definition 1.22. *Let A be a linear operator from X to Y with dense domain. We define its* adjoint *$A^\star$ from $Y^\star$ to $X^\star$ by setting*

$$\begin{aligned} &\mathrm{D}(A^\star) = \{y^\star \in Y^\star \,|\, \exists\, z^\star \in X^\star \;\; \forall\, x \in \mathrm{D}(A) : \langle Ax, y^\star\rangle = \langle x, z^\star\rangle\}, \\ &A^\star y^\star = z^\star. \end{aligned} \tag{1.5}$$

Observe that for all $x \in \mathrm{D}(A)$ and $y^\star \in \mathrm{D}(A^\star)$ we obtain

$$\langle Ax, y^\star\rangle = \langle x, A^\star y^\star\rangle.$$

We note that the operator $Af = f'$ with $\mathrm{D}(A) = \{f \in C^1([0,1]) \,|\, f(0) = 0\}$ is not densely defined on $X = C([0,1])$ since $\overline{\mathrm{D}(A)} = \{f \in X \,|\, f(0) = 0\}$. We first collect basic properties of the adjoint that follow rather directly from the definition.

Remark 1.23. Let A be linear from X to Y with $\overline{\mathrm{D}(A)} = X$.

a) Since $\mathrm{D}(A)$ is dense, there is at most one vector $z^\star = A^\star y^\star$ as in (1.5), so that $A^\star : \mathrm{D}(A^\star) \to X^\star$ is a map. It is clear that $A^\star$ is linear. For $A \in \mathcal{B}(X,Y)$, Definition 1.22 coincides with the definition of $A^\star$ in §5.4 of [**FA**], where $\mathrm{D}(A^\star) = Y^\star$.

b) The operator $A^\star$ is closed from $Y^\star$ to $X^\star$.

Proof. Let $y_n^\star \in \mathrm{D}(A^\star)$, $y^\star \in Y^\star$, and $z^\star \in X^\star$ such that $y_n^\star \to y^\star$ in $Y^\star$ and $z_n^\star := A^\star y_n^\star \to z^\star$ in $X^\star$ as $n \to \infty$. To check that $y^\star \in \mathrm{D}(A^\star)$, take $x \in \mathrm{D}(A)$. We derive

$$\langle x, z^\star\rangle = \lim_{n\to\infty} \langle x, z_n^\star\rangle = \lim_{n\to\infty} \langle Ax, y_n^\star\rangle = \langle Ax, y^\star\rangle,$$

and thus $y^\star$ belongs to $\mathrm{D}(A^\star)$ and $A^\star y^\star = z^\star$. □

c) Let $T \in \mathcal{B}(X,Y)$ and $\alpha \in \mathbb{F}\backslash\{0\}$. Then $\alpha A + T$ with $\mathrm{D}(\alpha A + T) = \mathrm{D}(A)$ has the adjoint $(\alpha A + T)^\star = \alpha A^\star + T^\star$ with $\mathrm{D}((\alpha A + T)^\star) = \mathrm{D}(A^\star)$.

Proof. Let $x \in \mathrm{D}(A)$ and $y^\star \in Y^\star$. We obtain

$$\langle(\alpha A + T)x, y^\star\rangle = \alpha\langle Ax, y^\star\rangle + \langle x, T^\star y^\star\rangle.$$

Hence, $y^\star$ is contained in $\mathrm{D}((\alpha A + T)^\star)$ if and only if $y^\star$ belongs to $\mathrm{D}(A^\star)$, and then $(\alpha A + T)^\star y^\star = \alpha A^\star y^\star + T^\star y^\star$. □

It is often difficult to treat adjoints of unbounded operators in examples. This topic will be discussed later on. Here we focus on the connection to spectral theory, where we can characterize $\sigma_{\mathrm{r}}(A)$ by the point spectrum of $A^\star$ and show that taking adjoints does not change the spectrum.

THEOREM 1.24. *Let A be a closed operator on X with dense domain. Then the following assertions hold.*

a) $\sigma_{\mathrm{r}}(A) = \sigma_{\mathrm{p}}(A^\star)$.

b) $\sigma(A) = \sigma(A^\star)$ *and* $R(\lambda, A)^\star = R(\lambda, A^\star)$ *for every* $\lambda \in \rho(A)$.

PROOF. a) Due to a corollary of the Hahn–Banach theorem (see Corollary 5.13 in [**FA**]), the set $(\lambda I - A)\,\mathrm{D}(A)$ is not dense in X if and only if there is a vector $y^\star \in X^\star \backslash \{0\}$ such that $\langle \lambda x - Ax, y^\star \rangle = 0$ for every $x \in \mathrm{D}(A)$. This equation is equivalent to $\langle Ax, y^\star \rangle = \langle x, \lambda y^\star \rangle$, which in turn means that $y^\star \in \mathrm{D}(A^\star)\backslash\{0\}$ and $A^\star y^\star = \lambda y^\star$; i.e., $\lambda \in \sigma_{\mathrm{p}}(A^\star)$.

b) Let $\lambda \in \rho(A)$. To show that $R(\lambda, A)^\star$ is the resolvent of $A^\star$, take $y^\star \in \mathrm{D}(A^\star)$ and $x \in X$. We compute

$$\begin{aligned}\langle x, R(\lambda, A)^\star (\lambda I - A^\star) y^\star \rangle &= \langle R(\lambda, A)x, (\lambda I - A^\star) y^\star \rangle \\ &= \langle (\lambda I - A) R(\lambda, A) x, y^\star \rangle = \langle x, y^\star \rangle,\end{aligned}$$

using Definition 1.22 and that $R(\lambda, A)x$ belongs to $\mathrm{D}(A)$. It follows $R(\lambda, A)^\star (\lambda I - A^\star) y^\star = y^\star$ so that $\lambda I - A^\star$ is injective. Next, pick $x^\star \in X^\star$. Set $y^\star = R(\lambda, A)^\star x^\star$. For $x \in \mathrm{D}(A)$, we obtain

$$\langle (\lambda I - A)x, y^\star \rangle = \langle R(\lambda, A)(\lambda I - A)x, x^\star \rangle = \langle x, x^\star \rangle.$$

Therefore $y^\star$ is an element of $\mathrm{D}(A^\star)$ and $x^\star = (\lambda I - A)^\star y^\star = (\lambda I - A^\star) y^\star$, where we use Remark 1.23 c). Consequently, the operator $\lambda I - A^\star$ is surjective, and thus bijective with inverse $R(\lambda, A^\star) = R(\lambda, A)^\star$.

Conversely, let $\lambda \in \rho(A^\star)$. Then λ does not belong to $\sigma_{\mathrm{p}}(A^\star) = \sigma_{\mathrm{r}}(A)$ by part a). Take $x \in \mathrm{D}(A)$. Due to a corollary of the Hahn–Banach theorem (see Corollary 5.10 in [**FA**]), there is a functional $y^\star \in X^\star$ such that $\|y^\star\| = 1$ and $\langle x, y^\star \rangle = \|x\|$. As above, we calculate

$$\begin{aligned}\|x\| = \langle x, y^\star \rangle &= \langle x, (\lambda I - A^\star) R(\lambda, A^\star) y^\star \rangle = \langle (\lambda I - A)x, R(\lambda, A^\star) y^\star \rangle \\ &\leqslant \|R(\lambda, A^\star)\| \, \|\lambda x - Ax\|;\end{aligned}$$

i.e., λ does not belong to $\sigma_{\mathrm{ap}}(A)$. Proposition 1.19 now yields $\lambda \notin \sigma(A)$. □

We give a typical application of the above results in a case where we know the adjoint explicitely, so that its eigenvalues can be computed.

EXAMPLE 1.25. Let $X \in \{c_0, \ell^p \,|\, 1 \leqslant p \leqslant \infty\}$ with $\mathbb{F} = \mathbb{C}$. Let $Rx = (0, x_1, x_2, \dots)$ be the right shift on X. We have $\sigma(R) = \overline{B}(0,1)$ and $\sigma_{\mathrm{p}}(R) = \emptyset$ for all X, $\sigma_{\mathrm{r}}(R) = \overline{B}(0,1)$ for $X = \ell^1$, and $\sigma_{\mathrm{r}}(R) = B(0,1)$ for $X \in \{c_0, \ell^p \,|\, 1 < p < \infty\}$.

PROOF. First, let $X \neq \ell^\infty$. From Example 5.44 of [**FA**] we know that $R^\star = L$, where the left shift L acts on ℓ^1 if $X = c_0$ and on $\ell^{p'}$ otherwise. Since $\sigma(L) = \overline{B}(0,1)$ by Example 1.17, Theorem 1.24 yields $\sigma(R) = \sigma(R^\star) = \sigma(L) = \overline{B}(0,1)$. Similarly, $\sigma_{\mathrm{r}}(R) = \sigma_{\mathrm{p}}(L) = B(0,1)$ if $X = c_0$ or $X = \ell^p$ with $1 < p < \infty$, and $\sigma_{\mathrm{r}}(R) = \sigma_{\mathrm{p}}(L) = \overline{B}(0,1)$ if $X = \ell^1$.

Let $X = \ell^\infty$. Here we use $R = L^\star$ for L on ℓ^1 so that again $\sigma(R) = \overline{B}(0,1)$.

Clearly, $Rx = 0$ yields $x = 0$. If $\lambda x = Rx = (0, x_1, x_2, \dots)$ and $\lambda \neq 0$, then $0 = \lambda x_1$ and so $x_1 = 0$. Iteratively one sees that $x = 0$. Hence, R has no eigenvalues. □

We show below that the spectrum is stable under 'small' perturbations. (A classic treatment of such questions is given in [**Ka**].) To this end, we first introduce an important notion that allows us to compare the 'size' of closed operators, and we discuss it a bit.

Let A be a linear operator from X to Y. Then a linear operator B from X to Y is called A*–bounded* (or if *relatively bounded* with respect to A) if $\mathrm{D}(A) \subseteq \mathrm{D}(B)$ and $B \in \mathcal{B}([\mathrm{D}(A)], Y)$.

Remark 1.26. Let A and B be linear from X to Y with $\mathrm{D}(A) \subseteq \mathrm{D}(B)$.

a) The operator B is A–bounded if and only if there are constants $a, b \geqslant 0$ such that

$$\|Bx\| \leqslant a\|Ax\| + b\|x\| \tag{1.6}$$

for all $x \in \mathrm{D}(A)$. Let $X = Y$ and A be closed with $\lambda \in \rho(A)$. Then the A–boundedness of B is also equivalent to the boundedness of $BR(\lambda, A)$. For instance, we then have (1.6) with $a := \|BR(\lambda, A)\|$ and $b := |\lambda|\, a$.

b) Let A be closed and let (1.6) be satisfied with $a < 1$. Then $A + B$ with $\mathrm{D}(A + B) = \mathrm{D}(A)$ is also closed, by an exercise. In view of a), the next result also requires that b in (1.6) is sufficiently small. ◊

Theorem 1.27. *Let A be a closed operator on X and $\lambda \in \rho(A)$. Further, let B be linear on X with $\mathrm{D}(A) \subseteq \mathrm{D}(B)$. Assume that $\|BR(\lambda, A)\| < 1$. Then $A + B$ with $\mathrm{D}(A + B) = \mathrm{D}(A)$ is closed, λ belongs to $\rho(A + B)$, and*

$$R(\lambda, A + B) = R(\lambda, A) \sum_{n=0}^{\infty} (BR(\lambda, A))^n = R(\lambda, A)(I - BR(\lambda, A))^{-1},$$

$$\|R(\lambda, A + B)\| \leqslant \frac{\|R(\lambda, A)\|}{1 - \|BR(\lambda, A)\|}\,.$$

Proof. By Proposition 4.24 of [**FA**], the operator $I - BR(\lambda, A)$ has the inverse

$$S_\lambda = \sum_{n=0}^{\infty} (BR(\lambda, A))^n$$

in $\mathcal{B}(X)$. Hence, $\lambda I - A - B = (I - BR(\lambda, A))(\lambda I - A) : \mathrm{D}(A) \to X$ is bijective with the bounded inverse $R(\lambda, A)S_\lambda$. Remark 1.11 thus yields the closedness of $A + B$ on $\mathrm{D}(A)$, and so $\lambda \in \rho(A + B)$. The asserted estimate also follows from Proposition 4.24 of [**FA**]. □

The smallness condition in the above theorem is sharp in general: Let $X = \mathbb{C}$, $a \in \mathbb{C} \cong \mathcal{B}(\mathbb{C})$, $a \neq 0$, and $b = a$. Then a is invertible, but $a - a = 0$ is not. Here we have $\lambda = 0$ and $|bR(0, a)| = |\frac{a}{a}| = 1$.

CHAPTER 2

Spectral theory of compact operators

Compact operators often occur in applications since integral operators are compact in many situations. So it is a crucial fact that their spectral theory is still close to the matrix case in several respects. We first discuss the relevant properties of compact operators and then establish the core spectral results. Subsequently, we extend the theory to closed operators having a compact resolvent. We also treat a more flexible class of operators and a related subset of the spectrum that it is invariant under compact perturbations. Finally, we sketch a typical application to the dynamics of reaction-diffusion equations.

2.1. Compact operators

We first recall a few facts from, e.g., Section 1.3 of [**FA**]. A non-empty subset $S \subseteq X$ is *compact* if each sequence in S has a subsequence with limit in S. Equivalently, S is compact if every open covering of S has a finite subcovering. We call $S \subseteq X$ *relatively compact* if $\overline{S}$ is compact, which means that each sequence in S has a converging subsequence (with limit in $\overline{S}$). Finally, $S \subseteq X$ is relatively compact if and only if it is *totally bounded*; i.e., for each $\varepsilon > 0$ there are finitely many balls in X with radius ε covering it, where one may chose the centers in S. Compact sets are bounded and closed. The converse is true if and only if X has finite dimension. A closed subset of compact set is also compact.

We start with the basic definition for this chapter and state simple facts.

Definition 2.1. *A linear map $T : X \to Y$ is called* compact *if $T\overline{B}(0,1)$ is relatively compact in Y. The set of all compact linear operators is denoted by $\mathcal{B}_0(X,Y)$.*

Remark 2.2. Let $T : X \to Y$ be linear.

a) If T is compact, then $T\overline{B}(0,1)$ is bounded and thus T is bounded; i.e., $\mathcal{B}_0(X,Y) \subseteq \mathcal{B}(X,Y)$.

b) The following assertions are equivalent.

i) T is compact.
ii) T maps bounded sets of X into relatively compact sets of Y.
iii) For every bounded sequence $(x_n)_n$ in X there exists a convergent subsequence $(Tx_{n_j})_j$ in Y.

Proof. Let T be compact. Take a bounded set $B \subseteq X$. Then B is contained in $\overline{B}(0,r)$ for some $r > 0$ and thus $\overline{TB}$ in $\overline{T\overline{B}(0,r)} = r\overline{T\overline{B}(0,1)}$, which is compact. Hence, $\overline{TB}$ is compact. The other implications ii) $\Rightarrow$ iii) $\Rightarrow$ i) are clear. $\square$

c) The space of *operators of finite rank* is defined by

$$\mathcal{B}_{00}(X,Y) = \{T \in \mathcal{B}(X,Y) \,\big|\, \dim TX < \infty\},$$

cf. Example 5.16 of [**FA**]. For $T \in \mathcal{B}_{00}(X,Y)$, the set $T\overline{B}(0,1)$ is relatively compact by the Bolzano–Weierstraß theorem; i.e., $\mathcal{B}_{00}(X,Y) \subseteq \mathcal{B}_0(X,Y)$.

d) The identity $I : X \to X$ is compact if and only if $\overline{B}(0,1)$ is compact if and only if $\dim X < \infty$. See Theorem 1.42 of [**FA**]. ◊

The next result says that $\mathcal{B}_0(X,Y)$ is a closed two-sided ideal in $\mathcal{B}(X,Y)$. The proofs are typical for the area.

PROPOSITION 2.3. *The set $\mathcal{B}_0(X,Y)$ is a closed linear subspace of $\mathcal{B}(X,Y)$. Let $T \in \mathcal{B}(X,Y)$ and $S \in \mathcal{B}(Y,Z)$. If one of the operators T or S is compact, then ST is compact.*

PROOF. Let $x_k \in X$ with $k \in \mathbb{N}$ satisfy $c := \sup_{k\in\mathbb{N}} \|x_k\| < \infty$.

1) Let $T, R \in \mathcal{B}_0(X,Y)$ be compact. If $\alpha \in \mathbb{F}$, then αT is also compact. There further exists a converging subsequence $(Tx_{k_j})_j$. Since $(x_{k_j})_j$ is still bounded, there is another converging subsequence $(Rx_{k_{j_l}})_l$. So $((T+R)x_{k_{j_l}})_l$ has a limit and $T+R$ belongs to $\mathcal{B}_0(X,Y)$; which thus is a linear subspace.

2) Let $T_n \in \mathcal{B}_0(X,Y)$ tend in $\mathcal{B}(X,Y)$ to some $T \in \mathcal{B}(X,Y)$ as $n \to \infty$. The compactness of T_1 yields a subsequence $(T_1 x_{\nu_1(j)})_j$ with limit y_1. Because of $\|x_{\nu_1(j)}\| \leqslant c$ for all j, there is a subsubsequence ν_2 of ν_1 such that $(T_2 x_{\nu_2(j)})_j$ converges. Note that $(T_1 x_{\nu_2(j)})_j$ still tends to y_1. Iteratively, we obtain subsequences ν_l of ν_{l-1} such that $(T_n x_{\nu_l(j)})_j$ converges for all $n \leqslant l$.

We use the diagonal sequence given by $u_m = x_{\nu_m(m)}$ for $m \in \mathbb{N}$. Then $(T_n u_m)_m$ converges as $m \to \infty$ for each $n \in \mathbb{N}$. Let $\varepsilon > 0$. Fix an index $N = N_\varepsilon \in \mathbb{N}$ such that $\|T_N - T\| \leqslant \varepsilon$. Then fix $M \in \mathbb{N}$ with $\|T_N(u_m - u_k)\| \leqslant \varepsilon$ for all $m \geqslant k \geqslant M$. For these indices we obtain

$$\begin{aligned} \|Tu_m - Tu_k\| &\leqslant \|(T-T_N)u_m\| + \|T_N(u_m - u_k)\| + \|(T_N - T)u_k\| \\ &\leqslant c\varepsilon + \varepsilon + c\varepsilon. \end{aligned}$$

Therefore $(Tu_m)_m$ is a Cauchy sequence, and we have shown that T is compact. Hence, $\mathcal{B}_0(X,Y)$ is closed in $\mathcal{B}(X,Y)$.

3) Let $S \in \mathcal{B}_0(X,Y)$. Since $(Tx_k)_k$ is bounded, there is a converging subsequence $(STx_{k_j})_j$, so that ST is compact. Instead, let $T \in \mathcal{B}_0(X,Y)$. We then find a subsequence $(Tx_{k_l})_l$ with a limit y, and thus STx_{k_l} tends to Sy. Again, ST is compact. □

REMARK 2.4. Strong limits of compact operators may fail to be compact. Consider, e.g., $X = \ell^2$ and $T_n x = (x_1, \dots, x_n, 0, 0, \dots)$ for all $x \in X$ and $n \in \mathbb{N}$. Then T_n belongs to $\mathcal{B}_{00}(X) \subseteq \mathcal{B}_0(X)$, but $T_n x$ tends to $x = Ix$ as $n \to \infty$ for every $x \in X$ and $I \notin \mathcal{B}_0(X)$. ◊

We next discuss compactness of several integral operators.

EXAMPLE 2.5. a) Let $X \in \{C([0,1]), L^p([0,1]) \,|\, 1 \leqslant p \leqslant \infty\}$, $Y = C([0,1])$, and $k \in C([0,1]^2)$. Setting

$$Tf(t) = \int_0^1 k(t,\tau) f(\tau)\,\mathrm{d}\tau$$

for $f \in X$ and $t \in [0,1]$, we define the integral operator $T : X \to Y$ for the kernel k. Then T belongs to $\mathcal{B}_0(X,Y)$.

Proof. By Analysis 2 or 3, the function Tf is continuous for all $f \in X$ and $T : X \to Y$ is linear. Since $\|Tf\|_\infty \leqslant \|k\|_\infty \|f\|_1 \leqslant \|k\|_\infty \|f\|_p$ (using that $\lambda([0,1]) = 1$), the map T is contained in $\mathcal{B}(X,Y)$. In particular, TB is bounded in Y where $B := \overline{B}_X(0,1)$. To show compactness, we use the Arzela–Ascoli Theorem 1.47 from [**FA**]. For $t, s \in [0,1]$ and $f \in B$ we have

$$\begin{aligned} |Tf(t) - Tf(s)| &\leqslant \int_0^1 |k(t,\tau) - k(s,\tau)|\, |f(\tau)| \,\mathrm{d}\tau \\ &\leqslant \sup_{\tau \in [0,1]} |k(t,\tau) - k(s,\tau)|\, \|f\|_1 \leqslant \sup_{\tau \in [0,1]} |k(t,\tau) - k(s,\tau)|. \end{aligned}$$

The right-hand side tends to 0 as $|t-s| \to 0$ uniformly in $f \in B$, because k is uniformly continuous. Therefore TB is equicontinuous. Theorem 1.47 in [**FA**] then implies that TB is relatively compact; i.e., $T \in \mathcal{B}_0(X,Y)$. □

b) Let $X = C([0,1])$ and $Vf(t) = \int_0^t f(s)\,\mathrm{d}s$ for $t \in (0,1)$ and $f \in X$. This defines a bounded operator V on X with norm 1, see Example 1.17. Let $f \in \overline{B}(0,1)$. Then $\|Vf\|_\infty \leqslant 1$ and $\|(Vf)'\|_\infty = \|f\|_\infty \leqslant 1$. The Arzela–Ascoli theorem (see Corollary 1.48 in [**FA**]) now yields the compactness of V.

c)[1] Let $X = L^2(\mathbb{R})$. For $f \in X$, we define

$$Tf(t) = \int_{\mathbb{R}} \mathrm{e}^{-|t-s|} f(s)\,\mathrm{d}s, \qquad t \in \mathbb{R}.$$

By Theorem 2.14 of [**FA**], the operator $T : X \to X$ is linear and bounded. We claim that T is not compact.

Proof. Take $f_n = \mathbb{1}_{[n,n+1]}$. For $n > m$ in $\mathbb{N}$, we compute $\|f_n\|_2 = 1$ and

$$\begin{aligned} \|Tf_n - Tf_m\|_2^2 &\geqslant \int_{n+1}^{n+2} \left| \int_n^{n+1} \mathrm{e}^{s-t}\,\mathrm{d}s - \int_m^{m+1} \mathrm{e}^{s-t}\,\mathrm{d}s \right|^2 \mathrm{d}t \\ &= \int_{n+1}^{n+2} \mathrm{e}^{-2t} \big(\mathrm{e}^{n+1} - \mathrm{e}^n - e^{m+1} + \mathrm{e}^m\big)^2 \,\mathrm{d}t \\ &\geqslant \tfrac{1}{2}(\mathrm{e}^{-2n-2} - \mathrm{e}^{-2n-4})(\mathrm{e}^{n+1} - 2\mathrm{e}^n)^2 \\ &= \tfrac{1}{2}(\mathrm{e}^{-2} - \mathrm{e}^{-4})(\mathrm{e} - 2)^2 > 0. \end{aligned}$$

Hence, (Tf_n) has no converging subsequence. □

d) Let $E = L^2(\mathbb{R}^m)$ and $k \in L^2(\mathbb{R}^{2m})$. For $f \in E$, we set

$$Tf(x) = \int_{\mathbb{R}^m} k(x,y) f(y)\,\mathrm{d}y, \qquad x \in \mathbb{R}^m.$$

As seen in Example 5.44 of [**FA**], this defines an operator $T \in \mathcal{B}(E)$. We claim that T is compact.

Proof. There are maps $k_n \in C_c(\mathbb{R}^{2m})$ that converge to k in E. (See Analysis 3.) Let T_n be the corresponding integral operators in $\mathcal{B}(E)$. There is a closed ball $B_n \subseteq \mathbb{R}^m$ such that $\operatorname{supp} k_n \subseteq B_n \times B_n$. We then have

$$T_n f(x) = \begin{cases} 0, & x \in \mathbb{R}^m \backslash B_n, \\ \int_{B_n} k_n(x,y) f(y)\,\mathrm{d}y, & x \in B_n. \end{cases}$$

[1]Parts c) and d) were partly sketched in the lectures.

for $f \in E$ and $n \in \mathbb{N}$. Let $R_n f = f\restriction_{B_n}$. Fix $n \in \mathbb{N}$ and take a bounded sequence (f_k) in E. Arguing as in part a), one finds a subsequence such that $(R_n T_n f_{k_j})_j$ has a limit g_0 in $C(B_n)$. Since B_n has finite measure, this sequence converges also in $L^2(B_n)$. The maps $T_n f_{k_j}$ then tend in E to the 0–extension g of g_0 as $j \to \infty$, and so T_n is compact. Let $f \in \overline{B}_E(0,1)$. Hölder's inequality in the inner integral yields

$$\begin{aligned}\|Tf - T_n f\|_2^2 &= \int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} (k(x,y) - k_n(x,y)) f(y) \, \mathrm{d}y \right|^2 \mathrm{d}x \\ &\leqslant \int_{\mathbb{R}^m} \Big(\int_{\mathbb{R}^m} |k(x,y) - k_n(x,y)|^2 \, \mathrm{d}y \int_{\mathbb{R}^m} |f(y)|^2 \, \mathrm{d}y \Big) \, \mathrm{d}x \\ &\leqslant \|k - k_n\|_2^2\end{aligned}$$

for all $n \in \mathbb{N}$. The operators T_n thus converge to T in $\mathcal{B}(E)$ so that T is compact by Proposition 2.3. □

Summarizing, integral operators are usually compact if the base space is compact or has finite measure, or if the kernel decays fast enough at infinity. (In part c) we have $k(t,s) = \mathrm{e}^{-|t-s|}$ without decay on strips $\{(t,s) \in \mathbb{R}^2 \mid |t-s| \leqslant c\}$. The next result due to Schauder will allow us to use duality in the context of compact operators.

THEOREM 2.6. *An operator $T \in \mathcal{B}(X,Y)$ is compact if and only if its adjoint $T^\star \in \mathcal{B}(Y^\star, X^\star)$ is compact.*

PROOF. 1) Let T be compact. Take $y_n^\star \in Y^\star$ with $\sup_{n\in\mathbb{N}} \|y_n^\star\| =: c < \infty$. The set $K := \overline{T\overline{B}_X(0,1)}$ is a compact metric space for the restriction of the norm of Y. Set $f_n := y_n^\star\restriction_K \in C(K)$ for each $n \in \mathbb{N}$. Putting $\kappa := \max_{y\in K} \|y\| < \infty$, we obtain

$$\|f_n\|_\infty = \max_{y\in K} |\langle y, y_n^\star \rangle| \leqslant c\kappa$$

for every $n \in \mathbb{N}$. Moreover, $(f_n)_{n\in\mathbb{N}}$ is equicontinuous since

$$|f_n(y) - f_n(z)| = |\langle y - z, y_n^\star \rangle| \leqslant \|y_n^\star\| \, \|y - z\| \leqslant c\|y - z\|$$

for all $n \in \mathbb{N}$ and $y, z \in K$. The Arzela–Ascoli theorem then yields a subsequence $(f_{n_j})_j$ converging in $C(K)$. We further compute

$$\begin{aligned}\|T^\star y_{n_j}^\star - T^\star y_{n_l}^\star\|_{X^\star} &= \sup_{\|x\|\leqslant 1} |\langle x, T^\star(y_{n_j}^\star - y_{n_l}^\star \rangle| = \sup_{\|x\|\leqslant 1} |\langle Tx, y_{n_j}^\star - y_{n_l}^\star \rangle| \\ &\leqslant \|f_{n_j} - f_{n_l}\|_{C(K)} \longrightarrow 0\end{aligned}$$

as $j, l \to \infty$. This means that $(T^\star y_{n_j}^\star)_j$ converges and so $T^\star$ is compact.

2) Let $T^\star$ be compact. By step 1), the bi-adjoint T^{**} is compact. Let $J_X : X \to X^{**}$ be the canonical isometric embedding. Proposition 5.45 in [**FA**] says that $T^{**} J_X = J_Y T$, and hence $J_Y T$ is compact by Proposition 2.3. Let (x_n) be bounded in X. We then obtain a converging subsequence $(J_Y T x_{n_j})_j$ which is Cauchy. Since J_Y is isometric, also $(Tx_{n_j})_j$ is Cauchy and thus has a limit; i.e., T is compact. □

2.2. The Fredholm alternative

In this section we establish the main spectral properties of compact operators K. These follow from the deep Theorem 2.7 due to Riesz (1918) and Schauder (1930) which describes the mapping properties of $I - K$ in detail.

We need some facts from functional analysis to study kernels and ranges by means of duality. To this end, for non-empty sets $M \subseteq X$ and $N_* \subseteq X^\star$ we define the *annihilators*

$$M^\perp = \{x^\star \in X^\star \,|\, \forall y \in M : \langle y, x^\star \rangle = \},$$
$$^\perp N_* = \{x \in X \,|\, \forall y^\star \in N_* : \langle x, y^\star \rangle = 0\}.$$

These sets are equal to $X^\star$ or X if and only if $M = \{0\}$ or $N_* = \{0\}$, respectively, see Remark 5.21 in [**FA**]. Let $T \in \mathcal{B}(X)$. Proposition 5.46 in [**FA**] says that

$$\begin{aligned} \mathrm{R}(T)^\perp &= \mathrm{N}(T^\star), \qquad & \overline{\mathrm{R}(T)} &= {}^\perp\mathrm{N}(T^\star), \\ \mathrm{N}(T) &= {}^\perp\mathrm{R}(T^\star), \qquad & \overline{\mathrm{R}(T^\star)} &\subseteq \mathrm{N}(T)^\perp. \end{aligned} \tag{2.1}$$

In particular, $\mathrm{R}(T)$ is dense if and only if $T^\star$ is injective; and if $\mathrm{R}(T^\star)$ is dense, then T is injective.

The following *Riesz–Schauder theorem* extends fundamental results for matrices known from linear algebra. The core equivalence of injectivity and surjectivity of $I - K$ fails for non-compact K. (Take for instance $K = I - R$ for the right shift R on ℓ^p, which is injective but not surjective, and $K = I - L$ for the left shift L on ℓ^p, which is surjective but not injective.)

THEOREM 2.7. *Let $K \in \mathcal{B}_0(X)$ and set $T = I - K$. Then the following assertions hold.*

a) $\mathrm{R}(T)$ is closed.

b) $\dim \mathrm{N}(T) < \infty$ and $\operatorname{codim} \mathrm{R}(T) := \dim X/\mathrm{R}(T) < \infty$.

c) T is bijective $\iff$ T is surjective $\iff$ T is injective $\iff$ $T^\star$ is bijective $\iff$ $T^\star$ is surjective $\iff$ $T^\star$ is injective. More precisely, we have

$$\dim \mathrm{N}(T) = \operatorname{codim} \mathrm{R}(T) = \dim \mathrm{N}(T^\star) = \operatorname{codim} \mathrm{R}(T^\star). \tag{2.2}$$

Before proving the theorem, we first reformulate it as the *Fredholm alternative* for the solvability of the equation $\lambda x - Lx = y$ for compact L. For integral operators this result goes back to Fredholm (1900).

COROLLARY 2.8. *Let $L \in \mathcal{B}_0(X)$, $\lambda \in \mathbb{F}\backslash\{0\}$, and $x, y \in X$. Then one of the following alternatives holds.*

A) The homogeneous problem $\lambda x = Lx$ has only the trivial solution $x = 0$. Then for every $y \in X$ there is a unique solution $x \in X$ of $\lambda x - Lx = y$ given by $x = R(\lambda, L)y$.

B) The equation $\lambda x = Lx$ has the n-dimensional solution space $\mathrm{N}(\lambda I - L)$ for some $n \in \mathbb{N}$. Then there are n linearly independent solutions $x_1^\star, \dots, x_n^\star \in X^\star$ of $\lambda x^\star = L^\star x^\star$. The equation $\lambda x - Lx = y$ has a solution $x \in X$ if and only if $\langle y, x_k^\star \rangle = 0$ for all $k \in \{1, \dots, n\}$. Every $z \in X$ satisfying $\lambda z - Lz = y$ is of the form $z = x + x_0$, where $\lambda x - Lx = y$ and $x_0 \in \mathrm{N}(\lambda I - L)$.

Proof of Corollary 2.8. We set $K = \frac{1}{\lambda}L \in \mathcal{B}_0(X)$ and note that $\lambda x - Lx = y$ is equivalent to $(I-K)x = \frac{1}{\lambda}y$. By Theorem 2.7 b), we have either $\dim \mathrm{N}(I-K) = 0$ (case A) or $\dim \mathrm{N}(I-K) = n \in \mathbb{N}$ (case B).

In the first case, $I-K$ is bijective due to Theorem 2.7 c) which yields A).

In the second case, Theorem 2.7 a) shows that $\mathrm{R}(I-K)$ is closed so that $\mathrm{R}(I-K) = {}^{\perp}\mathrm{N}(I-K^{\star})$ by (2.1). We thus deduce the solvability condition from case B) noting that $\dim \mathrm{N}(I-K^{\star}) = n$ due to (2.2). If $x - Kx = y$ and $z - Kz = y$, then $z-x$ belongs to $\mathrm{N}(I-K)$, as required in case B). □

We also note that the Fredholm alternative fails for $\lambda = 0$, and give simple, but typical application to differential equations.

Example 2.9. a) Let $X = C([0,1])$ and $Vf(t) = \int_0^t f(s)ds$ for $t \in [0,1]$ and $f \in X$. We then have

$$\mathrm{R}(V) = \{g \in C^1([0,1]) \,|\, g(0) = 0\},$$

which neither closed nor dense in X. In particular, $Vf = g$ can not be solved for all $g \in X$. Nevertheless, V is injective and compact by Examples 1.17 and 2.5, respectively.

b) Let $X = C([0,1])$ with $\mathbb{F} = \mathbb{R}$ and $q, f \in X$ with $q \geqslant 0$. Then there is a unique function $u \in C^2([0,1])$ solving the boundary value problem

$$u''(t) - q(t)u(t) = f(t), \quad t \in [0,1], \qquad u(0) = 0 = u(1). \tag{2.3}$$

It is given by the integral equation

$$u(t) - \int_0^1 k(t,s)q(s)u(s)\,\mathrm{d}s = \int_0^1 k(t,s)f(s)\,\mathrm{d}s, \qquad t \in [0,1], \tag{2.4}$$

with the kernel

$$k(t,s) = \begin{cases} (t-1)s, & 0 \leqslant s \leqslant t \leqslant 1, \\ t(s-1), & 0 \leqslant t < s \leqslant 1. \end{cases}$$

Proof. 1) Since we want to deduce solvability from uniqueness by means of the Fredholm alternative, we first take two solutions $u, v \in C^2([0,1])$ of (2.3). Then $w = u - v \in C^2([0,1])$ solves (2.3) with $f = 0$. We multiply the latter problem by w, take the integral over $[0,1]$, and integrate by parts. The boundary conditions then imply

$$0 \leqslant \int_0^1 qw^2\,\mathrm{d}t = \int_0^1 w''w\,\mathrm{d}t = -\int_0^1 (w')^2\,\mathrm{d}t,$$

so that $w' = 0$ and $w(t) = w(0) = 0$ for all t. Problem (2.3) has thus at most one solution.

2) The operator K given by $Kg(t) = \int_0^1 k(t,s)q(s)g(s)\,\mathrm{d}s$ for $t \in [0,1]$ and $g \in X$ is compact on X by Example 2.5. It is straightforward to check that a solution $u \in C^2([0,1])$ to (2.4) belongs to $C^2([0,1])$ and solves (2.3). So step 1) yields uniqueness of (2.4) with $f = 0$. Case A) in Corollary 2.8 now leads to the assertion. □

Proof of Theorem 2.7. In the first two steps we show $\dim \mathrm{N}(T) < \infty$ and use this fact to establish part a). In a third step, a duality argument completes the proof of assertion b). The lenghty two final steps then prove statement c), using properties of kernels and ranges of the powers T^k.

1) The space $N := \mathrm{N}(T) = T^{-1}(\{0\})$ is closed in X. For $x \in N$ we have $Kx = x \in N$, so that K leaves N invariant and its restriction K_N to N coincides with the identity on N. On the other hand, K_N is still compact so that $\dim N < \infty$ by Remark 2.2 d).

2) Since $\dim N < \infty$, there is a closed subspace $C \subseteq X$ such that $N \cap C = \{0\}$ and $N + C = X$; i.e., $X = N \oplus C$. See Proposition 5.17 in [**FA**].

Let $\tilde{T} : C \to \mathrm{R}(T)$ be the restriction of T to C. We endow C and $\mathrm{R}(T)$ with the norm of X, so that C is a Banach space by its closedness. To show that $\mathrm{R}(T)$ is closed, we want to invert $\tilde{T}$.

Let $\tilde{T}x = 0$ for some $x \in C$. Then x also belongs to N and so $x = 0$. Let $y \in \mathrm{R}(T)$. There is a vector $x \in X$ with $Tx = y$. We can write $x = x_0 + x_1$ with $x_0 \in N$ and $x_1 \in C$. Hence, $\tilde{T}x_1 = Tx_1 + Tx_0 = y$, and so $\tilde{T}$ is bijective.

A corollary to the open mapping theorem (see Corollary 4.31 in [**FA**]) now yields that $\mathrm{R}(T) = \mathrm{R}(\tilde{T})$ is closed if and only if $\tilde{T}^{-1} : \mathrm{R}(T) \to C$ is bounded. Suppose that $\tilde{T}^{-1}$ was unbounded. Then there would exist elements $\bar{y}_n = \tilde{T}\bar{x}_n$ of $\mathrm{R}(T)$ with $\bar{x}_n \in C$ such that $\bar{y}_n \to 0$ as $n \to \infty$ and $\|\bar{x}_n\| = \|\tilde{T}^{-1}\bar{y}_n\| \geqslant \delta$ for some $\delta > 0$ and all $n \in \mathbb{N}$. We set $x_n = \|\bar{x}_n\|^{-1}\bar{x}_n$ and note that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and that

$$y_n := x_n - Kx_n = \tilde{T}x_n = \frac{1}{\|\bar{x}_n\|}\bar{y}_n \;\longrightarrow\; 0$$

as $n \to \infty$. The compactness of K yields a subsequence $(x_{n_j})_j$ and a vector $z \in X$ such that $Kx_{n_j} \to z$ as $j \to \infty$. We obtain the limit $x_{n_j} = y_{n_j} + Kx_{n_j} \to z$ and so $\|z\| = 1$. Observe that z belongs to the closed set C. On the other hand, z is contained in N because of

$$Tz = z - Kz = \lim_{j\to\infty} Kx_{n_j} - K \lim_{j\to\infty} x_{n_j} = 0$$

implying that $z \in C \cap N = \{0\}$. This fact contradicts $\|z\| = 1$, and hence assertion a) is true.

3) Theorem 2.6 provides the compactness of $K^\star$ so that $\dim \mathrm{N}(I - K^\star) < \infty$ by step 1). Using (2.1) and Proposition 5.23 in [**FA**], we further obtain

$$\mathrm{N}(T^\star) = \mathrm{R}(T)^\perp \cong (X/\mathrm{R}(T))^\star.$$

Since $\mathrm{N}(T^\star)$ is finite-dimensional, linear algebra yields that

$$\infty > \dim \mathrm{N}(T^\star) = \dim (X/\mathrm{R}(T))^\star = \dim X/\mathrm{R}(T) = \operatorname{codim} \mathrm{R}(T), \qquad (2.5)$$

showing statement b). We next prove in two steps the remaining equalities in (2.2) which then imply the first part of assertion c).[2]

4) *Claim A: There is a closed linear subspace $\hat{N}$ with $\dim \hat{N} < \infty$ and a closed linear subspace $\hat{R}$ of X such that*

$$X = \hat{N} \oplus \hat{R}, \quad T\hat{N} \subseteq \hat{N}, \quad T\hat{R} \subseteq \hat{R} \quad \text{and} \quad T_2 := T{\restriction}_{\hat{R}} : \hat{R} \to \hat{R} \text{ is bijective.}$$

Assume that Claim A has been shown. Setting $T_1 := T{\restriction}_{\hat{N}} \in \mathcal{B}(\hat{N})$, we obtain the following properties.

(i) $\dim \hat{N}/\mathrm{R}(T_1) = \dim \mathrm{N}(T_1)$ (by the dimension formula in $\mathbb{C}^n$).

[2] In the lectures (2.2) was not shown. Instead of steps 4) and 5) below, only the (much easier) proof of the first part of c) was given.

(ii) $\mathrm{N}(T) = \mathrm{N}(T_1)$. In fact, writing $x = x_1 + x_2$ for $x \in X$, $x_1 \in \hat{N}$ and $x_2 \in \hat{R}$, we deduce that $Tx = 0$ if and only if $T_2 x_2 = -T_1 x_1 \in \hat{N} \cap \hat{R} = \{0\}$. As T_2 is injective, the latter statement is equivalent to $x_2 = 0 = T_1 x_1$. Hence, x belongs to $\mathrm{N}(T)$ if and only if $x \in \hat{N}$ and $T_1 x = 0$; i.e., $\mathrm{N}(T) = \mathrm{N}(T_1)$.

(iii) We define the map

$$\Phi : \hat{N}/\mathrm{R}(T_1) \to X/\mathrm{R}(T); \quad x + \mathrm{R}(T_1) \mapsto x + \mathrm{R}(T),$$

for $x \in \hat{N} \subseteq X$. Because of $\mathrm{R}(T_1) \subseteq \mathrm{R}(T)$, the map Φ is well defined. Of course, it is linear. We want to show that Φ is bijective, which leads to

$$\dim \hat{N}/\mathrm{R}(T_1) = \dim X/\mathrm{R}(T).$$

Proof of (iii). Let $\Phi(x + \mathrm{R}(T_1)) = 0$ for some $x \in \hat{N}$ which yields $x = Ty$ for a vector $y \in X$. By Claim A, we have the decomposition $y = y_1 + y_2$ for $y_1 \in \hat{N}$ and $y_2 \in \hat{R}$. Hence, $T_2 y_2 = x - T_1 y_1$ is contained in $\hat{R} \cap \hat{N} = \{0\}$ so that $y_2 = 0$ by the injectivity of T_2. So x belongs to $\mathrm{R}(T_1)$ and Φ is injective.

Take $x \in X$. Again there are elements $x_1 \in \hat{N}$ and $x_2 \in \hat{R} = T\hat{R}$ with $x = x_1 + x_2$. We now conclude that $x - x_1 = x_2 \in \mathrm{R}(T)$ and thus

$$\Phi(x_1 + \mathrm{R}(T_1)) = x_1 + \mathrm{R}(T) = x + \mathrm{R}(T).$$

Hence, Φ is bijective. $\Diamond$

Properties (i)–(iii) lead to

$$\dim \mathrm{N}(T) = \dim \mathrm{N}(T_1) = \dim \hat{N}/\mathrm{R}(T_1) = \dim X/\mathrm{R}(T) = \operatorname{codim} \mathrm{R}(T). \quad (2.6)$$

Since also $K^\star$ is compact by Theorem 2.6, we further obtain

$$\dim \mathrm{N}(T^\star) = \operatorname{codim} \mathrm{R}(T^\star). \quad (2.7)$$

In view of (2.5)–(2.7), part c) follows from Claim A.

5) *Proof of Claim A.* We set $N_k = \mathrm{N}(T^k)$ and $R_k = \mathrm{R}(T^k)$ for $k \in \mathbb{N}_0$. Observe that

$$\begin{aligned} \{0\} = N_0 \subseteq N_1 \subseteq N_2 \subseteq \dots, \qquad & X = R_0 \supseteq R_1 \supseteq R_2 \supseteq \dots, \\ TN_k \subseteq N_{k-1} \subseteq N_k, \quad \text{and} \quad & TR_k = R_{k+1} \subseteq R_k \end{aligned} \quad (2.8)$$

for all $k \in \mathbb{N}_0$. We also have

$$T^k = (I - K)^k = I - \sum_{j=1}^{k} \binom{k}{j} (-1)^{j+1} K^j =: I - C_k,$$

where C_k is compact for each $k \in \mathbb{N}$ due to Proposition 2.3. Assertions a) and b) now imply that

$$N_k,\ R_k \ \text{ are closed} \quad \text{and} \quad \dim N_k < \infty \quad (2.9)$$

for every $k \in \mathbb{N}$. We need four more claims to establish Claim A.

Claim 1: There is a minimal $n \in \mathbb{N}_0$ such that $N_n = N_{n+j}$ for all $j \in \mathbb{N}_0$.

Indeed, suppose that $N_j \subsetneqq N_{j+1}$ for all $j \in \mathbb{N}_0$. Then Riesz' Lemma 1.44 in [**FA**] would give $x_j \in N_j$ with $\|x_j\| = 1$ and $\mathrm{d}(x_j, N_{j-1}) \geqslant 1/2$ for every $j \in \mathbb{N}_0$. (Here we use that N_{j-1} is closed.) Take $l > k \geqslant 0$. Since $Tx_l + x_k - Tx_k$ is contained in N_{l-1} by (2.8), we deduce that

$$\|Kx_l - Kx_k\| = \|x_l - (Tx_l + x_k - Tx_k)\| \geqslant 1/2.$$

As a result, $(Kx_k)_k$ has no converging subsequence, which contradicts the compactness of K. So there is a minimal $n \in \mathbb{N}_0$ with $N_n = N_{n+1}$. Let $x \in N_{n+2}$. Then, Tx belongs to $N_{n+1} = N_n$ so that x is an element of N_{n+1}. This means that $N_{n+1} = N_{n+2}$, and Claim 1 follows by induction.

Claim 2: There is a minimal $m \in \mathbb{N}_0$ such that $R_m = R_{m+j}$ for all $j \in \mathbb{N}_0$.

In fact, suppose that $R_{j+1} \subsetneqq R_j$ for all $j \in \mathbb{N}_0$. Again from Riesz' lemma we obtain vectors $x_j \in R_j$ with $\|x_j\| = 1$ and $\mathrm{d}(x_j, R_{j+1}) \geqslant 1/2$ for every $j \in \mathbb{N}_0$. (Here we use that R_{j+1} is closed.) Take $l > k \geqslant 0$. Since $Tx_k + x_l - Tx_l \in R_{k+1}$ by (2.8), we deduce that

$$\|Kx_k - Kx_l\| = \|x_k - (Tx_k + x_l - Tx_l)\| \geqslant 1/2.$$

This lower bound contradicts the compactness of K. So there exists a minimal $m \in \mathbb{N}_0$ with $R_m = R_{m+1}$. Let $y \in R_{m+1}$. Then there is a vector $x \in X$ with $y = T^{m+1}x = TT^m x$. Hence, y is contained in $TR_m = TR_{m+1} = R_{m+2}$ by (2.8), implying $R_{m+1} = R_{m+2}$. Inductively, we obtain Claim 2.

Claim 3: $N_n \cap R_n = \{0\}$ and $N_m + R_m = X$.

Indeed, let $x \in N_n \cap R_n$ for the first part. Then $T^n x = 0$ and we have a pre-image $y \in X$ with $T^n y = x$. Hence, $T^{2n} y = 0$ and so y belongs to $N_{2n} = N_n$ by Claim 1, which yields $x = T^n y = 0$.

For the second part, let $x \in X$. By Claim 2, the vector $T^m x$ is contained in $R_m = R_{2m}$; i.e., $T^m x = T^{2m} y$ for some $y \in X$. Therefore, $x = (x - T^m y) + T^m y$ is an element of $N_m + R_m$.

Claim 4: $n = m$.

In fact, suppose that $n > m$. Due to Claim 1 and Claim 2, there is a vector $x \in N_n \backslash N_m$ and we have $R_n = R_m$. Claim 3 further gives $y \in N_m \subseteq N_n$ and $z \in R_m = R_n$ with $x = y + z$. Therefore, $z = x - y$ also belongs to N_n so that $z = 0$ by Claim 3. We obtain the contradiction $x = y \in N_m$.

Second, suppose that $n < m$. Claim 1 and Claim 2 yield $N_n = N_m$ and an element $x \in R_n \backslash R_m$. Owing to Claim 3, we have $x = y + z$ for vectors $y \in N_m = N_n$ and $z \in R_m \subseteq R_n$. Therefore, $y = x - z$ is contained in R_n so that $y = 0$ by Claim 3. It follows $x = z \in R_m$, which is impossible.

We can now finish the proof of Claim A, setting $\hat{N} := N_n$ and $\hat{R} := R_n$. By (2.9), the spaces $\hat{N}$ and $\hat{R}$ are closed and $\dim \hat{N} < \infty$. From Claims 3 and 4 we then infer that $X = \hat{N} \oplus \hat{R}$. Moreover, (2.8) and Claim 2 yield $T\hat{N} \subseteq \hat{N}$ and $T\hat{R} = \hat{R}$. If $Tx = 0$ for some $x = T^n y \in \hat{R}$ and $y \in X$, then $y \in N_{n+1} = N_n$ by Claim 1. Therefore, $x = 0$ and $T\restriction_{\hat{R}}$ is bijective. □

Reformulating the Riesz–Schauder theorem, we next describe the spectrum of a compact operator if $\dim X = \infty$. It contains 0 and at most countably many eigenvalues that tend to 0 if they are infinitely many. Moreover, the eigenspaces $\mathrm{N}(\lambda I - K)$ are finite-dimensional. Recall from Example 2.9 that the Voltera operator V is compact with $\sigma(V) = \{0\}$ and $\sigma_{\mathrm{p}}(V) = \emptyset$.

THEOREM 2.10. *Let $\dim X = \infty$ and $K \in \mathcal{B}(X)$ be compact. Then the following assertions hold.*

a) $\sigma(K) = \{0\} \,\dot\cup\, \{\lambda_j \,|\, j \in J\}$, where $J \in \{\emptyset, \mathbb{N}, \{1, \ldots, n\} \,|\, n \in \mathbb{N}\}$.

b) $\sigma(K)\backslash\{0\} = \sigma_{\mathrm{p}}(K)\backslash\{0\}$. *For all* $\lambda \in \sigma(K)\backslash\{0\}$ *the range of* $\lambda I - K$ *is closed and*

$$\dim \mathrm{N}(\lambda I - K) = \operatorname{codim} \mathrm{R}(\lambda I - K) < \infty.$$

c) For each $\varepsilon > 0$ *the set* $\sigma(K)\backslash B(0,\varepsilon)$ *is finite, so that* $\lambda_j \to 0$ *as* $j \to \infty$ *if* $J = \mathbb{N}$.

Proof. Suppose $0 \notin \sigma(K)$; i.e., K is invertible. By Proposition 2.3 the identity $I = K^{-1}K$ would be compact, which contradicts $\dim X = \infty$. Observe that assertion a) now follows from c) by taking $\varepsilon = 1/n$ for $n \in \mathbb{N}$.

For $\lambda \in \mathbb{F}\backslash\{0\}$ we have $\lambda I - K = \lambda(I - \frac{1}{\lambda}K)$. Since $\frac{1}{\lambda}K \in \mathcal{B}_0(X)$, Theorem 2.7 implies that either $\lambda \in \rho(K)$ or $\lambda \in \sigma_{\mathrm{p}}(K)$ with

$$\dim \mathrm{N}(\lambda I - K) = \dim \mathrm{N}(I - \tfrac{1}{\lambda}K) = \operatorname{codim} \mathrm{R}(I - \tfrac{1}{\lambda}K) = \operatorname{codim} \mathrm{R}(\lambda I - K) < \infty.$$

So we have established part b).

To prove statement c), we suppose that for some $\varepsilon_0 > 0$ we have points λ_n in $\sigma(K)\backslash B(0,\varepsilon_0)$ with $\lambda_n \neq \lambda_m$ for all $n \neq m$ in $\mathbb{N}$ and vectors x_n in $X\backslash\{0\}$ with $Kx_n = \lambda_n x_n$. In linear algebra it is shown that eigenvectors to different eigenvalues are linearly independent. Hence, the subspaces

$$X_n := \operatorname{lin}\{x_1, \dots, x_n\}$$

satisfy $X_n \subsetneqq X_{n+1}$ for every $n \in \mathbb{N}$. Moreover, $KX_n \subseteq X_n$ and X_n is closed for each $n \in \mathbb{N}$ (since $\dim X_n < \infty$). Riesz' Lemma 1.44 in [**FA**] gives vectors y_n in X_n such that $\|y_n\| = 1$ and $\mathrm{d}(y_n, X_{n-1}) \geqslant 1/2$ for each $n \in \mathbb{N}$. There are coefficients $\alpha_{n,j} \in \mathbb{F}$ with $y_n = \alpha_{n,1}x_1 + \cdots + \alpha_{n,n}x_n$, and hence

$$\lambda_n y_n - K y_n = \sum_{j=1}^{n} (\lambda_n - \lambda_j)\alpha_{n,j} x_j = \sum_{j=1}^{n-1} (\lambda_n - \lambda_j)\alpha_{n,j} x_j$$

belongs to X_{n-1}. For $n > m$, the vector $\lambda_n y_n - Ky_n + Ky_m$ is thus contained in X_{n-1} so that

$$\|Ky_n - Ky_m\| = |\lambda_n| \left\| y_n - \frac{1}{\lambda_n}(\lambda_n y_n - Ky_n + Ky_m) \right\| \geqslant \frac{|\lambda_n|}{2} \geqslant \frac{\varepsilon_0}{2}.$$

This fact contradicts the compactness of K. □

Using this result, we next shown non-compactness of our basic operators. A true application of the Riesz–Schauder theory is sketched in Example 2.25.

Example 2.11. The following bounded operators are not compact since their spectra are not finite or a null sequence, where we let $\mathbb{F} = \mathbb{C}$:

a) the left and right shifts L and R on c_0 or ℓ^p, $1 \leqslant p \leqslant \infty$, see Examples 1.17 and 1.25;

b) the translation $T(t)f = f(\cdot + t)$ for $t \in \mathbb{R}\backslash\{0\}$, on $L^p(\mathbb{R})$, $1 \leqslant p \leqslant \infty$, see Example 1.21;

c) multiplication operators $Tf = mf$ on $C_b(S)$ for a pathwise connected $S \subseteq \mathbb{R}^d$ and a given map $m \in C_b(S)\backslash\{0\}$, see Proposition 1.14. (Observe that $m(S)$ is pathwise connected in $\mathbb{C}$.) ◊

2.3. Closed operators with compact resolvent

We now transfer the results of the previous section to a class of closed operators introduced in the next definition.

Definition 2.12. *A closed operator A on X has a* compact resolvent *if there exists a number $\lambda \in \rho(A)$ such that $R(\lambda, A) \in \mathcal{B}(X)$ is compact.*

Besides other properties, we first note that an operator with $\rho(A) \neq \emptyset$ has compact resolvent if and only if its domain is compactly embedded into X.

Remark 2.13. Let A be closed on X and $\lambda \in \rho(A)$. Then the following assertions are true.

a) Let $R(\lambda, A)$ be compact. The resolvent equation (1.2) yields

$$R(\mu, A) = R(\lambda, A) + (\lambda - \mu)R(\lambda, A)R(\mu, A),$$

for $\mu \in \rho(A)$, so that also $R(\mu, A)$ is compact due to Proposition 2.3.

b) Recall $[\mathrm{D}(A)] = (\mathrm{D}(A), \|\cdot\|_A)$. The following assertions are equivalent.

i) A has a compact resolvent.
ii) Each bounded sequence in $[\mathrm{D}(A)]$ has a subsequence with limit in X.
iii) The inclusion map $J : [\mathrm{D}(A)] \to X$ is compact.

Proof. Let statement i) hold. Take $x_n \in \mathrm{D}(A)$ with $\|x_n\|_A \leqslant c$ for $n \in \mathbb{N}$. Set $y_n = \lambda x_n - Ax_n$. Then $\|y_n\| \leqslant (|\lambda| + 1)c$ for every $n \in \mathbb{N}$ so that $x_n = R(\lambda, A)y_n$ has a subsequence which converges in X by i), and so claim ii) is true. The implication 'ii) $\Rightarrow$ iii)' follows from Remark 2.2. Let part iii) be valid. Define $R_\lambda \in \mathcal{B}(X, [\mathrm{D}(A)])$ by $R_\lambda x = R(\lambda, A)x$ for $x \in X$. Then $R(\lambda, A) = JR_\lambda : X \to X$ is compact due to Proposition 2.3. $\square$

c) Let $T \in \mathcal{B}(X)$ have a compact resolvent and $\lambda \in \rho(T)$. We then have $\dim X < \infty$ since $I = (\lambda I - T)R(\lambda, T)$ is compact by Proposition 2.3.

d) Let $\mathrm{D}(A)$ be dense in X. Then, A has a compact resolvent if and only if $A^\star$ has a compact resolvent. Indeed, first recall from Theorem 1.24 that $\lambda \in \rho(A^\star)$ and $R(\lambda, A^\star) = R(\lambda, A)^\star$. Theorem 2.6 then yields that the compactness of $R(\lambda, A)$ and of $R(\lambda, A)^\star$ are equivalent. $\Diamond$

Results like the Arzela–Ascoli theorem provide compactly embedded function spaces on bounded spatial domains. Still one has to show that the resolvent is non-empty to apply the above characterization. This is indicated in the following example, where we use the Hölder space

$$C^\alpha(S) = \Big\{ f \in C_b(S) \,\Big|\, [f]_\alpha = \sup_{x \neq y} \frac{|f(y) - f(x)|}{|x - y|^\alpha} < \infty \Big\}$$

for $\alpha \in (0, 1)$ and $S \subseteq \mathbb{R}^m$. It is a Banach space with norm $\|f\|_{C^\alpha} = \|f\|_\infty + [f]_\alpha$. For $\alpha = 1$ the corresponding (Lipschitz) space is denoted by $C^{1-}(S)$. We have $C^{1-}(S) \hookrightarrow C^\beta(S) \hookrightarrow C^\alpha(S) \hookrightarrow C_b(S)$ if $0 < \alpha \leqslant \beta < 1$, where the embeddings are given by the inclusion map.

Example 2.14. a) Let $K \subseteq \mathbb{R}^m$ be compact, E be a closed subspace of $C(K)$, and A be closed on E with $\lambda \in \rho(A)$. Assume that $[\mathrm{D}(A)] \hookrightarrow C^\alpha(K)$ for some $\alpha \in (0, 1)$. A bounded sequence (f_n) in $[\mathrm{D}(A)]$ is thus bounded in $C^\alpha(K)$. The theorem of Arzela–Ascoli (see Corollary 1.48 in

[**FA**]) yields a subsequence with limit in $C(K)$, and hence in E. According to Remark 2.13 b), the operator A has a compact resolvent in E.

b) Let $X = C([0,1])$, $Au = u'$, and $\mathrm{D}(A) = \{u \in C^1([0,1]) \,|\, u(0) = 0\}$. The spectrum of A is empty by Example 1.12. So part a) implies that A has compact resolvent (since $C^1([0,1]) \hookrightarrow C^{1-}([0,1])$).

c) Let $X = C([0,1])$, $Au = u'$, and $\mathrm{D}(A) = C^1([0,1])$. The resolvent set of A is empty by Example 1.12. Therefore A has no (compact) resolvent although $[\mathrm{D}(A)]$ is compactly embedded in X by a). ◊

The next result easily follows from Theorem 2.10. It says that the spectrum of an operator with compact resolvent only contains a discrete set of eigenvalues, having finite-dimensional eigenspaces. Here the spectrum may be empty in view of the above example.

Theorem 2.15. *Let $\dim X = \infty$ and A be a closed operator with compact resolvent. Then the following assertions are true.*

a) The spectrum $\sigma(A)$ is either empty or $\sigma(A) = \sigma_{\mathrm{p}}(A)$ contains at most countably many eigenvalues λ_j.

b) If $\sigma(A)$ is infinite, then $|\lambda_j| \to \infty$ as $j \to \infty$.

c) For all $\lambda_j \in \sigma(A)$, the range of $\lambda_j I - A$ is closed and $\dim N(\lambda_j I - A) = \operatorname{codim} \mathrm{R}(\lambda_j I - A) < \infty$.

Proof. Fix $\mu \in \rho(A)$. By Theorem 2.10, the spectrum $\sigma(R(\mu, A))$ only contains 0 and either no or finitely many eigenvalues $\mu_j \neq 0$ or a nullsequence of eigenvalues $\mu_j \neq 0$. Moreover, the range of $\mu_j I - R(\mu, A)$ is closed and $\dim N(\mu_j I - R(\mu, A)) = \operatorname{codim} \mathrm{R}(\mu_j I - R(\mu, A)) < \infty$ for all j. Proposition 1.20 yields $(\mu - \sigma(A))^{-1} = \sigma(R(\mu, A))\backslash\{0\}$ and $(\mu - \sigma_{\mathrm{p}}(A))^{-1} = \sigma_{\mathrm{p}}(R(\mu, A))\backslash\{0\}$. These facts imply assertions a) and b), where $\lambda_j = \mu - \mu_j^{-1}$. Observe that

$$\lambda_j I - A = \mu_j^{-1}(\mu_j I - R(\mu, A))(\mu I - A)$$

on $\mathrm{D}(A)$. Since $\mu I - A : \mathrm{D}(A) \to X$ is bijective, also part c) follows from the results of Theorem 2.10 stated above. □

We use the above theorem to compute the spectra of two basic operators with compact resolvent.

Example 2.16. a) Let $X = C([0,1])$ with $\mathbb{F} = \mathbb{C}$ and $Au = u'$ with $\mathrm{D}(A) = \{u \in C^1([0,1]) \,|\, u(0) = u(1)\}$. Then A is closed, has a compact resolvent, and $\sigma(A) = \sigma_{\mathrm{p}}(A) = 2\pi \mathrm{i}\mathbb{Z}$.

Proof. The closedness is shown as in Example 1.2. Let $f \in X$. A function u belongs to $\mathrm{D}(A)$ and satisfies $u - Au = f$ if and only if $u \in C^1([0,1])$, $u(0) = u(1)$, and $u' = u - f$. These properties are equivalent to

$$u(t) = c\mathrm{e}^t - \int_0^t \mathrm{e}^{t-s} f(s)\,\mathrm{d}s =: R_c f(t), \qquad t \in [0,1], \qquad \text{and} \qquad u(0) = u(1),$$

for some $c = c(f)$. Here c has to satisfy

$$c = R_c f(0) = R_c f(1) = c\mathrm{e} - \mathrm{e}\int_0^1 \mathrm{e}^{-s} f(s)\,\mathrm{d}s,$$

$$c = \frac{\mathrm{e}}{\mathrm{e}-1}\int_0^1 \mathrm{e}^{-s} f(s)\,\mathrm{d}s.$$

We derive $1 \in \rho(A)$ and

$$R(1,A)f(t) = \frac{\mathrm{e}^{t+1}}{\mathrm{e}-1}\int_0^1 \mathrm{e}^{-s} f(s)\,\mathrm{d}s - \int_0^t \mathrm{e}^{t-s} f(s)\,\mathrm{d}s, \qquad t \in [0,1].$$

Due to Example 2.14 a), the operator A thus has a compact resolvent, and so $\sigma(A) = \sigma_{\mathrm{p}}(A)$ by Theorem 2.15. Finally, $\lambda \in \mathbb{C}$ belongs to $\sigma_{\mathrm{p}}(A)$ if and only if there is $u \in C^1([0,1])\backslash\{0\}$ with $u(0) = u(1)$ and $u' = \lambda u$, which is equivalent to $u = e_\lambda u(0) \neq 0$ and $1 = e_\lambda(0) = e_\lambda(1) = \mathrm{e}^\lambda$; i.e., $\lambda \in 2\pi\mathrm{i}\mathbb{Z}$. $\square$

b) Let $X = C([0,1])$ with $\mathbb{F} = \mathbb{C}$ and $Au = u''$ with $\mathrm{D}(A) = \{u \in C^2([0,1]) \,|\, u(0) = u(1) = 0\}$. Then A is closed, has a compact resolvent, and $\sigma(A) = \sigma_{\mathrm{p}}(A) = \{-\pi^2 k^2 \,|\, k \in \mathbb{N}\}$.

PROOF. 1) Example 2.9 b) with $q = 0$ provides an inverse for A; i.e., 0 belongs to $\rho(A)$. So A is closed, and it has a compact resolvent by Example 2.14 a). To compute the eigenvalues, note that $v_k(t) = \sin(\pi k t)$ is an eigenfunction for A and the eigenvalue $\lambda = -\pi^2 k^2$, where $k \in \mathbb{N}$. Conversely, let $\lambda \in \sigma_{\mathrm{p}}(A)$. Then we have a map $u \in C^2([0,1])$ with $u(0) = u(1) = 0$ and $u'' = \lambda u$. There thus exist $a, b, \mu \in \mathbb{C}$ with $\mu^2 = \lambda$ and $u = ae_\mu + be_{-\mu} \neq 0$. The conditions $u(0) = 0$ and $u(1) = 0$ then yield $a+b = 0$ and $a\mathrm{e}^\mu + b\mathrm{e}^{-\mu} = 0$, respectively. Hence, $e^{2\mu} = 1$ and $\mu \neq 0$; i.e., $\mu = \mathrm{i}\pi k$ and $\lambda = -\pi^2 k^2$ for some $k \in \mathbb{Z}\backslash\{0\}$. Theorem 2.15 now yields the assertion.

2) We also compute the resolvent operators for later use (and to present an important technique). We take $\lambda \in \rho(A)\backslash\{0\} = \mathbb{C}\backslash\{-\pi^2 k^2 \,|\, k \in \mathbb{N}_0\}$, as the case $\lambda = 0$ was treated in Example 2.9. Then there is a number $\mu \in \mathbb{C}\backslash\{0\}$ with $\lambda = \mu^2$. Let $f \in X$. Set

$$u_0(t) = \frac{1}{2\mu}\int_0^1 \mathrm{e}^{-\mu|t-s|} f(s)\,\mathrm{d}s = \frac{1}{2\mu}\int_0^t \mathrm{e}^{\mu(s-t)} f(s)\,\mathrm{d}s + \frac{1}{2\mu}\int_t^1 \mathrm{e}^{\mu(t-s)} f(s)\,\mathrm{d}s$$

for $t \in [0,1]$. Then $u_0 \in C^2([0,1])$ satisfies $\mu^2 u_0 - u_0'' = f$. (We see in Example 3.47 how to guess this formula.) We now want to add a function $u_1 \in C^2([0,1])$ with $u_1'' = \mu^2 u_1$ so that $u = u_0 + u_1$ fulfills the boundary conditions $u(0) = 0 = u(1)$. Then u will belong to $\mathrm{D}(A)$ and solve $\lambda u - Au = f$. Uniqueness of solutions was already shown in the first step.

From step 1) we know that $u_1 = ae_\mu + be_{-\mu}$ for some numbers $a = a(f,\mu)$ and $b = b(f,\mu)$ in $\mathbb{C}$. We have to fulfill the boundary conditions

$$u(0) = a + b + \frac{1}{2\mu}\int_0^1 \mathrm{e}^{-\mu s} f(s)\,\mathrm{d}s = 0,$$
$$u(1) = a\mathrm{e}^\mu + b\mathrm{e}^{-\mu} + \frac{\mathrm{e}^{-\mu}}{2\mu}\int_0^1 \mathrm{e}^{\mu s} f(s)\,\mathrm{d}s = 0,$$

These two equations are equivalent to

$$\begin{pmatrix} a(f,\mu) \\ b(f,\mu) \end{pmatrix} = \frac{1}{\mathrm{e}^\mu - \mathrm{e}^{-\mu}} \begin{pmatrix} \mathrm{e}^{-\mu} & -1 \\ -e^\mu & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2\mu}\int_0^1 \mathrm{e}^{-\mu s} f(s)\,\mathrm{d}s \\ \frac{\mathrm{e}^{-\mu}}{2\mu}\int_0^1 \mathrm{e}^{\mu s} f(s)\,\mathrm{d}s \end{pmatrix}$$
$$= \frac{1}{2\mu(\mathrm{e}^\mu - \mathrm{e}^{-\mu})} \begin{pmatrix} \mathrm{e}^{-\mu}\int_0^1 (\mathrm{e}^{-\mu s} - \mathrm{e}^{\mu s}) f(s)\,\mathrm{d}s \\ \int_0^1 (\mathrm{e}^{-\mu}\mathrm{e}^{\mu s} - \mathrm{e}^{\mu}\mathrm{e}^{-\mu s}) f(s)\,\mathrm{d}s \end{pmatrix}.$$

Note that $e^{\mu} \neq e^{-\mu}$. We obtain

$$R(\lambda, A)f(t) = a(f,\mu)e^{\mu t} + b(f,\mu)e^{-\mu t} + \frac{1}{2\mu}\int_0^1 e^{-\mu|t-s|} f(s)\,ds$$

for $\lambda \in \rho(A)\backslash\{0\}$, $f \in X$ and $t \in [0,1]$. □

We note that one can also compute the resolvents in part a). So the power of Theorem 2.15 is not really needed here, but it can simplify the reasoning and gives extra information.

2.4. Fredholm operators and the essential spectrum

In this section we briefly study a class of operators which satisfy most of the assertions of the Riesz–Schauder Theorems 2.7 or 2.15, except for the restrictive equality $\dim \mathrm{N}(T) = \operatorname{codim} \mathrm{R}(T)$. This class turns out to be stable under compact perturbations, and it arises in many applications as we indicate in Example 2.25. We further introduce part of the spectrum which also does not change under compact perturbations.

Definition 2.17. *A map $T \in \mathcal{B}(X,Y)$ is called a* Fredholm operator *if*

a) its range $\mathrm{R}(T)$ *is closed in* Y,

b) $\dim \mathrm{N}(T) < \infty$,

c) $\operatorname{codim} \mathrm{R}(T) = \dim Y/\mathrm{R}(T) < \infty$.

In this case the index *of T is the integer* $\operatorname{ind}(T) = \dim \mathrm{N}(T) - \operatorname{codim} \mathrm{R}(T)$.

For a closed operator A in Y we use the above definition with $X = [\mathrm{D}(A)]$. One sometimes calls $\dim \mathrm{N}(T)$ the nullity and $\operatorname{codim} \mathrm{R}(T)$ the defect of T. We first discuss the above concept.

Remark 2.18. a) An invertible operator $T \in \mathcal{B}(X,Y)$ is Fredholm with index 0, of course. Loosely speaking, Fredholmity means 'invertibility except for finite-dimensional spaces', cf. Proposition 2.19.

b) Theorem 2.7 says that $\lambda I - K : X \to X$ is Fredholm with index 0 if $K \in \mathcal{B}(X)$ is compact and $\lambda \in \mathbb{F}\backslash\{0\}$. The same is true for $\lambda I - A : [\mathrm{D}(A)] \to X$ if $\lambda \in \mathbb{F}$ and A is closed in X with compact resolvent, by Theorem 2.15.

c) Each integer can occur as an index of a Fredholm operator. For instance, let $T = L^n$ for the left shift L on ℓ^p, $1 \leqslant p \leqslant \infty$, and some $n \in \mathbb{N}$; i.e., $Tx = (x_{n+k})_k$. Because of $\mathrm{R}(T) = \ell^p$ and $\mathrm{N}(T) = \operatorname{lin}\{e_1, \dots, e_n\}$, the map T is Fredholm with index n. Moreover, $S = R^n$ has index $-n$ for the right shift R on ℓ^p and $n \in \mathbb{N}$ since $Sx = (0, \dots, 0, x_1, x_2, \dots)$ with n zeros, so that $\mathrm{N}(S) = \{0\}$ and $\ell^p/\mathrm{R}(S) \cong \operatorname{lin}\{e_1, \dots, e_n\}$, cf. Example 2.20 of [**FA**].

d) The Fredholm operators do not form a linear subspace of $\mathcal{B}(X)$ if $\dim X = \infty$. For instance, the identity I is Fredholm, but $I - I = 0$ not.

e) One can omit condition a) in Definition 2.17, as shown by Kato (1958).

Proof.[3] Let $T \in \mathcal{B}(X,Y)$ satisfy $\operatorname{codim} \mathrm{R}(T) = n \in \mathbb{N}_0$. If $n = 0$, then $\mathrm{R}(T) = Y$ and we are done. Hence, take $n \in \mathbb{N}$. The operator $Q : X \to X/\mathrm{N}(T)$; $Qx = x + \mathrm{N}(T) = \hat{x}$, is a surjective contraction with kernel $\mathrm{N}(T)$, see Proposition 2.19 in [**FA**]. By $\hat{T}(x + \mathrm{N}(T)) = \hat{T}\hat{x} := Tx$, we define a bijective operator $\hat{T} \in \mathcal{B}(X/\mathrm{N}(T), \mathrm{R}(T))$ such that $T = \hat{T}Q$, cf.

[3] This proof and that of the following comment were omitted in the lectures.

Proposition A.1.3 of [**Co2**]. Further, there are $y_1, \dots, y_n \in Y$ such that $Y = \mathrm{R}(T) + \mathrm{lin}\{y_1, \dots, y_n\}$ and $\mathrm{R}(T) \cap \mathrm{lin}\{y_1, \dots, y_n\} = \{0\}$. On the Banach space $E = (X/\mathrm{N}(T)) \times \mathbb{F}^n$ we define the operator

$$S : E \to Y; \quad S(\hat{x}, (\lambda_1, \dots, \lambda_n)) = \hat{T}\hat{x} + \sum\nolimits_{j=1}^{n} \lambda_j y_j = Tx + \sum\nolimits_{j=1}^{n} \lambda_j y_j.$$

It straightforward to check that S is linear, bounded and bijective. The Open Mapping Theorem 4.28 in [**FA**]) then says that S is an isomorphism. Consequently, $\mathrm{R}(T) = S((X/\mathrm{N}(T)) \times \{0\})$ is closed. $\square$

In context of part e) of the above remark, we stress that for each infinite-dimensional Banach space X there are non-closed subspaces Z of X with codimension 1. To see this, take a countable subset $\mathcal{B}_0 = \{b_k \,|\, k \in \mathbb{N}\}$ of an algebraic basis $\mathcal{B}$ of X as in Remark 1.6. Set $\varphi(b_k) = k$ and $\varphi(b) = 0$ for $b \in \mathcal{B}\backslash\mathcal{B}_0$. Then φ extends to an unbounded linear map $\varphi : X \to \mathbb{F}$. Define $\hat{\varphi}$ as in the proof of Remark 2.18 e). It is a linear bijection from $X/\mathrm{N}(\varphi)$ to $\mathrm{R}(\varphi) = \mathbb{F}$; i.e., $\operatorname{codim} \mathrm{N}(\varphi) = 1$. However, $Z = \mathrm{N}(\varphi)$ is not closed by Proposition III.5.3 of [**Co2**].

The analysis in this section is based on a characterization of Fredholm operators with index 0 established in the next result. (See Theorem 3.15.8 in [**Si**] for a related characterization of Fredholm operators with any index.)

Proposition 2.19. *Let $T \in \mathcal{B}(X, Y)$. Then T is Fredholm with index $\mathrm{ind}(T) = 0$ if and only if there exists an invertible operator $J \in \mathcal{B}(Y, X)$ and a finite rank operator $K \in \mathcal{B}_{00}(X)$ such that $JT = I_X - K$.*

Proof. 1) Let T be Fredholm with index 0. We first modify T and X to obtain an invertible map 'close' to T. To this end, by means of Fredholmity we can choose closed subspaces X_1 of X and Y_0 of Y such that $X = \mathrm{N}(T) \oplus X_1$, $Y = Y_0 \oplus \mathrm{R}(T)$ and $\dim Y_0 = \operatorname{codim} \mathrm{R}(T) = \dim \mathrm{N}(T)$. (Use Proposition 5.17 of [**FA**].) We set

$$T_1 : X_1 \times Y_0 \to Y; \quad T_1(x_1, y_0) = Tx_1 + y_0,$$

replacing $\mathrm{N}(T)$ by the complement of $\mathrm{R}(T)$. Clearly, T_1 is linear and continuous. If $T_1(x_1, y_0) = 0$ for some (x_1, y_0) in $X_1 \times Y_0$, then $y_0 = -Tx_1 \in \mathrm{R}(T)$ so that $y_0 = 0$. Hence, x_1 belongs to $\mathrm{N}(T) \cap X_1 = \{0\}$; i.e., T_1 is injective. Let $y \in Y$. Then $y = y_0 + y_1$ for some $y_0 \in Y_0$ and $y_1 = Tx_1$ with $x_1 \in X$. As a result, $y = y_0 + Tx_1 = T_1(x_1, y_0)$ and T_1 is bijective. The inverse $T_1^{-1} : Y \to X_1 \times Y_0$ is then bounded by the Open Mapping Theorem 4.28 in [**FA**]. Observe that $T_1^{-1} y_1 = (x_1, 0)$ if $Tx_1 = y_1$ for some $x_1 \in X_1$.

There exists an isomorphism $S : Y_0 \to \mathrm{N}(T)$ since these spaces have the same finite dimension. To relate $X_1 \times Y_0$ with X, we define

$$S_1 : X_1 \times Y_0 \to X; \quad S_1(x_1, y_0) = x_1 + Sy_0.$$

As above, one checks that S_1 is invertible. We now introduce the invertible operator $J = S_1 T_1^{-1} : Y \to X$ and the map $K := I_X - JT \in \mathcal{B}(X)$. For $x_1 \in X_1$ we compute

$$Kx_1 = x_1 - S_1 T_1^{-1} T x_1 = x_1 - S_1(x_1, 0) = x_1 - x_1 = 0.$$

Since $X = \mathrm{N}(T) \oplus X_1$, we derive $KX \subseteq K\,\mathrm{N}(T)$, and hence $\dim \mathrm{R}(K) \leqslant \dim \mathrm{N}(T) < \infty$ as asserted.

2) Conversely, let $JT = I_X - K$ for some $K \in \mathcal{B}_{00}(X)$ and invertible $J \in \mathcal{B}(Y,X)$. Then JT is Fredholm with index 0 by Theorem 2.7. Since J is invertible, also TX is closed and we have $\dim \mathrm{N}(T) = \dim \mathrm{N}(JT) =: n < \infty$. Moreover, as above we obtain a closed subspace X_0 of X such that $X = JTX \oplus X_0$ and $\dim X_0 = n$. It follows $Y = TX \oplus J^{-1}X_0$ and thus $\operatorname{codim} TX = \dim J^{-1}X_0 = n$, using Example 2.20 of [**FA**]. Hence, T is Fredholm with index 0. □

We can now show an important perturbation result for Fredholmity and the index. The quantitative smallness condition from Theorem 1.27 (on invertibility) is replaced by compactness; i.e., by 'topological' smallness.

THEOREM 2.20. *Let $T \in \mathcal{B}(X,Y)$ be Fredholm and $K \in \mathcal{B}(X,Y)$ be compact. Then the sum $T+K \in \mathcal{B}(X,Y)$ is Fredholm with* $\operatorname{ind}(T+K) = \operatorname{ind}(T)$.

PROOF. Set $n = \operatorname{ind}(T) \in \mathbb{Z}$.

1) Let $n = 0$. Proposition 2.19 yields an invertible operator $J \in \mathcal{B}(Y,X)$ and a map $K_1 \in \mathcal{B}_{00}(X)$ with $JT = I_X - K_1$. The product JK is compact by Proposition 2.3. We thus deduce from Theorem 2.7 that $J(T+K) = I_X - (K_1 - JK)$ is Fredholm with index 0. As in step 2) of the proof of Proposition 2.19, it follows that also $T+K$ is Fredholm with index 0.

2) Let $n > 0$. Set $\tilde{Y} = Y \times \mathbb{F}^n$, and define

$$\tilde{T} : X \to \tilde{Y}; \quad \tilde{T}x = (Tx, 0), \qquad \tilde{K} : X \to \tilde{Y}; \quad \tilde{K}x = (Kx, 0).$$

It is straightforward to check that $\mathrm{R}(\tilde{T})$ is closed, $\mathrm{N}(T) = \mathrm{N}(\tilde{T})$, $\tilde{Y}/\mathrm{R}(\tilde{T}) \cong (Y/\mathrm{R}(T)) \times \mathbb{F}^n$. In particular, $\operatorname{codim} \mathrm{R}(\tilde{T}) = \operatorname{codim} \mathrm{R}(T) + n$ and so $\tilde{T}$ is Fredholm with index 0. Since $\tilde{K}$ is still compact, by step 1) the sum $\tilde{T} + \tilde{K}$ is also Fredholm with index 0. Noting that $(\tilde{T} + \tilde{K})x = (Tx + Kx, 0)$, we infer that $T + K$ is Fredholm with index n.

3) Let $n < 0$. Set $\hat{X} = X \times \mathbb{F}^{|n|}$, and define the maps

$$\hat{T} : \hat{X} \to Y; \quad \hat{T}(x,\xi) = Tx, \qquad \hat{K} : \hat{X} \to Y; \quad \hat{K}(x,\xi) = Kx.$$

Starting from $\dim \mathrm{N}(\hat{T}) = \dim \mathrm{N}(T) + |n|$ and $\mathrm{R}(\hat{T}) = \mathrm{R}(T)$, one derives the assertion as in part 2). □

To exploit the above result in spectral theory, we need another definition.

DEFINITION 2.21. *For $T \in \mathcal{B}(X)$ we define the* essential spectrum *by*

$$\sigma_{\mathrm{ess}}(T) = \{\lambda \in \mathbb{F} \mid \lambda I - T \text{ is not Fredholm}\}.$$

We also set

$$\sigma^0_{\mathrm{ess}}(T) = \{\lambda \in \mathbb{F} \mid \lambda I - T \text{ is not Fredholm of index 0}\}.$$

For a closed operator A on X we analogously introduce

$$\sigma_{\mathrm{ess}}(A) = \{\lambda \in \mathbb{F} \mid \lambda I - A : [\mathrm{D}(A)] \to X \text{ is not Fredholm}\},$$
$$\sigma^0_{\mathrm{ess}}(A) = \{\lambda \in \mathbb{F} \mid \lambda I - A : [\mathrm{D}(A)] \to X \text{ is not Fredholm of index 0}\}.$$

Observe that $\sigma_{\mathrm{ess}}(B) \subseteq \sigma^0_{\mathrm{ess}}(B) \subseteq \sigma(B)$ and that $\lambda \in \sigma(B) \backslash \sigma^{(0)}_{\mathrm{ess}}(B)$ is an eigenvalue with finite-dimensional eigenspace since $\lambda I - B$ is then Fredholm. (Here $B \in \mathcal{B}(X)$ or B is closed on X.)

There are various differing concepts of the essential spectrum in the literature. Typically, they lead to the same *essential spectral radius*

$$\mathrm{r}_{\mathrm{ess}}(T) = \sup\{|\lambda| \,\big|\, \lambda \in \sigma_{\mathrm{ess}}(T)\},$$

if $T \in \mathcal{B}(X)$. The next concept and the following lemma are used below in our perturbation result Theorem 2.24 for $\sigma_{\mathrm{ess}}(A)$.

Definition 2.22. *Let A and B be linear from X to Y with $\mathrm{D}(A) \subseteq \mathrm{D}(B)$. Then B is called* A–compact *(or,* relatively compact with respect to A*) if $B : [\mathrm{D}(A)] \to Y$ is compact.*

Note that an A–compact operator is automatically A–bounded (but it does not need to be closed from X to Y). Moreover, A–compactness is just compactness if $A \in \mathcal{B}(X, Y)$. Relative compactness is further discussed in the excercises.

Lemma 2.23. *Let A be closed from X to Y and B be A–compact. Then $A + B$ with $\mathrm{D}(A + B) = \mathrm{D}(A)$ is closed and B is relatively compact with respect to $A + B$.*

Proof. 1) We first check that B is $(A+B)$–compact. Let $(x_n) \subseteq \mathrm{D}(A)$ be bounded for $\|\cdot\|_{A+B}$. In particular, (x_n) is bounded in X.

We want to use the A-boundedness of B. To this end, suppose that $\alpha_n := \|Ax_n\|$ tends to infinity as $n \to \infty$. Set $\tilde{x}_n = \alpha_n^{-1} x_n$ for $n \in \mathbb{N}$. (We may assume that $Ax_n \neq 0$ for all n.) We then have the limits $\tilde{x}_n \to 0$ in X and $(A + B)\tilde{x}_n = \alpha_n^{-1}(A + B)x_n \to 0$ in Y as $n \to \infty$, whereas $\|A\tilde{x}_n\| = 1$ for all n. Since B is A–compact, there is a subsequence $(B\tilde{x}_{n_j})_j$ converging to some z in Y. Hence, $A\tilde{x}_{n_j}$ tends to $-z$. The closedness of A then yields $z = 0$, which is impossible since $1 = \|A\tilde{x}_{n_j}\| \to \|z\|$ as $j \to \infty$.

We conclude that there exists a subsequence such that $(Ax_{n_k})_k$ is bounded in Y. Employing again the A–compactness of B, we obtain another subsequence $(Bx_{n_{k_l}})_l$ with a limit in Y; i.e., B is $(A + B)$–compact.

2) Let $x_n \in \mathrm{D}(A)$ tend to some x in X and $(A + B)x_n$ to some y in Y as $n \to \infty$. By part 1), there is subsequence and a vector $z \in Y$ such that $Bx_{n_j} \to z$ in Y as $j \to \infty$. As a result, Ax_{n_j} tends to $y - z$. Since A is closed, we infer that $x \in \mathrm{D}(A) = \mathrm{D}(A + B)$ and $Ax = y - z$. This means that $x_{n_j} \to x$ in $[\mathrm{D}(A)]$, and hence $Bx_{n_j} \to Bx = z$ by continuity. It follows $(A + B)x = y$ and so $A + B$ is closed. □

By means of Theorem 2.20 we show that essential spectra are not changed by compact perturbations. They can thus affect only eigenvalues with finite-dimensional eigenspaces.

Theorem 2.24. *Let A be closed on X, B be A–compact, and $\mathrm{D}(A+B) = \mathrm{D}(A)$. Then*

$$\sigma_{\mathrm{ess}}(A + B) = \sigma_{\mathrm{ess}}(A) \qquad \textit{and} \qquad \sigma^0_{\mathrm{ess}}(A + B) = \sigma^0_{\mathrm{ess}}(A).$$

Proof. Let $\lambda I - A : [\mathrm{D}(A)] \to$ be Fredholm (with index 0). Since $-B$ is A–compact, Theorem 2.20 shows that also $\lambda I - A - B$ is Fredholm (with index 0). Conversely, let $\lambda I - A - B : [\mathrm{D}(A + B)] \to$ be Fredholm (with index 0). By Lemma 2.23, the operator B is $(A+B)$–compact, so that again Theorem 2.20 yields the Fredholmity (with index 0) of $\lambda I - A$. □

We apply the above result to a typical situation arising in partial differential equations. However, we can only give a rough sketch.

EXAMPLE 2.25. We study the asymptotic stability of stationary solutions to the reaction-convection-diffusion equation

$$\begin{aligned} \partial_t u(t,s) &= a\partial_{ss}u(t,s) + b\partial_s u(t,s) + f(u(t,s)), \qquad t \geqslant 0,\ s \in \mathbb{R}, \\ u(0,s) &= u_0(s), \qquad s \in \mathbb{R}, \end{aligned} \tag{2.10}$$

for diffusion and convection constants $a > 0$ and $b \in \mathbb{R}$, a given initial state $u_0 \in X = C_{ub}(\mathbb{R}) = \{v \in C_b(\mathbb{R}) \,|\, v \text{ is uniformly continuous}\}$, and a function $f \in C^1(\mathbb{C})$ with $f(\mathbb{R}) \subseteq \mathbb{R}$ describing (auto-)reaction (where we mean real differentiability). We interpret $u(t,s) \geqslant 0$ as the density of a species.

It is known that there is a maximal existence time $\bar{t} = \bar{t}(u_0) \in (0,\infty]$ and a unique solution u of (2.10) in the space $C([0,\bar{t}),X) \cap C^1((0,\bar{t}),X) \cap C((0,\bar{t}),C^2_{ub}(\mathbb{R}))$. Moreover, if $u_0 \geqslant 0$ and $f(0) \geqslant 0$, we have $u \geqslant 0$. (See Proposition 7.3.1 in [**Lu**] and Theorem 3.8 in [**nEE**].)

Let $u_* \in C^2_{ub}(\mathbb{R},\mathbb{R}) = \{v \in C^2(\mathbb{R},\mathbb{R}) \,|\, v, v', v'' \in X\}$ be a stationary solution of (2.10); i.e., $u(t,s) = u_*(s)$ solves (2.10), which is equivalent to

$$0 = au_*'' + bu_*' + f(u_*) \qquad \text{on } \ \mathbb{R}.$$

One now asks whether such special solutions describe well the behavior of (2.10), at least locally near u_*. One possible answer is the principle of linearized stability. Here one proceeds similar as for ordinary differential equations in Analysis 4. In $X = C_b(\mathbb{R})$ define the maps A and F by

$$Av = av'' + bv' \quad \text{with} \quad \mathrm{D}(A) = C^2_{ub}(\mathbb{R}), \qquad F(v) = f \circ v.$$

One can then check that $F \in C^1(X,X)$ with derivative $F'(v) \in \mathcal{B}(X)$ at $v \in X$ given by $F'(v)w = f'(v)w$ for $w \in X$. (One defines differentiability in Banach spaces as in $\mathbb{R}^n$, and the formula for $F'(v)$ has to be modified a bit for $\mathbb{C}$-valued v or w.) We introduce the linearized operator at u_* by setting

$$A_*v = Av + F'(u_*)v = av'' + bv' + f'(u_*)v \quad \text{with} \quad \mathrm{D}(A_*) = C^2_{ub}(\mathbb{R}).$$

The principle of linearized stability now says the following. If $\mathrm{s}(A_*) := \sup\{\operatorname{Re}\lambda \,|\, \lambda \in \sigma(A)\} < -\delta < 0$, then there are constants $c, r > 0$ such that

$$\forall\, u_0 \in \overline{B}_X(u_*,r): \quad \bar{t}(u_0) = \infty \quad \text{and} \quad \|u(t) - u_*\|_\infty \leqslant c\mathrm{e}^{-\delta t}\|u_0 - u_*\|_\infty$$

for all $t \geqslant 0$, where u solves (2.10). (See Theorem 3.13 in [**nEE**].) We note that such results fail for certain partial differential equations. Here it works since (2.10) is of 'parabolic type'.

Of course, one now has to compute the sign of the *spectral bound* $\mathrm{s}(A_*)$ (or different properties of $\sigma(A_*)$ for more refined versions of the above result). We sketch a partial answer for the important special case that $u_*(s)$ has limits $\xi_\pm$ in $\mathbb{R}$ as $s \to \pm\infty$. Then the limit operators

$$A_\pm = A + F'(\xi_\pm \mathbb{1}) \quad \text{with} \quad \mathrm{D}(A_\pm) = C^2_{ub}(\mathbb{R})$$

have constant coefficients which simplifies the computation of their spectral properties. We now follow the survey article [**Sa**].

Let $\lambda \in \mathbb{C}$. We rewrite $A_\pm u - \lambda u = g$ as the first order system by

$$L(\lambda)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} := \begin{pmatrix} v_1' \\ v_2' \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ \frac{1}{a}(\lambda - f'(\xi_\pm)) & -\frac{b}{a} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{a}g \end{pmatrix},$$

where $(v_1, v_2) \mathrel{\widehat{=}} (u, u')$. We set

$$M_\pm(\lambda) = \begin{pmatrix} 0 & 1 \\ \frac{1}{a}(\lambda - f'(\xi_\pm)) & -\frac{b}{a} \end{pmatrix}.$$

We denote by $X_\pm^u(\lambda)$ the linear span of all (generalized) eigenvectors of $M_\pm(\lambda)$ for eigenvalues μ with $\operatorname{Re}\mu > 0$. Theorems 3.2 and 3.3 and Remark 3.3 in [**Sa**] then yield

$$\lambda I - A_* \text{ is Fredholm} \iff \lambda \notin \sigma_{\mathrm{ess}}(A_*) \iff \sigma(M_\pm(\lambda)) \cap \mathrm{i}\mathbb{R} = \emptyset,$$
$$\operatorname{ind}(\lambda I - A_*) = \dim X_-^u(\lambda) - \dim X_+^u(\lambda),$$
$$\lambda \notin \sigma_{\mathrm{ess}}^0(A_*) \iff \sigma(M_\pm(\lambda)) \cap \mathrm{i}\mathbb{R} = \emptyset \quad \text{and} \quad \dim X_-^u(\lambda) = \dim X_+^u(\lambda).$$

Note that here non-zero indices naturally occur. The proofs of these results use Theorems 2.20 and 2.24 and properties of the ordinary differential equation governed by the matrices

$$M_\lambda(s) = \begin{pmatrix} 0 & 1 \\ \frac{1}{a}(\lambda - f'(u_*(s))) & -\frac{b}{a} \end{pmatrix}, \qquad s \in \mathbb{R}.$$

One thus has to study the eigenvalues of $M_\pm(\lambda)$ (which is easy) to determine the location of the essential spectrum. It then remains the (difficult) task to locate the eigenvalues of A_* to verify $\mathrm{s}(A_*) < 0$. In particular for one spatial dimension, corresponding tools are discussed in [**Sa**]. ♢

2.5. Appendix: The Dirichlet problem and boundary integrals

In[4] this section, we discuss a principal application of Theorem 2.7 to partial differential equations. Here we work with real-valued functions for simplicity. Let $D \subseteq \mathbb{R}^3$ be open, bounded and connected with $\partial D \in C^2$ and outer unit normal ν at ∂D (see part 3) below). Let $\varphi \in C(\partial D)$ be given.

Claim. There is a unique solution u in $C^2(D) \cap C(\overline{D}) := \{u \in C(\overline{D}) \mid u{\restriction}_D \in C^2(D)\}$ of the Dirichlet problem

$$\begin{aligned} \Delta u(x) &= 0, \qquad x \in D, \\ u(x) &= \varphi(x), \qquad x \in \partial D. \end{aligned} \tag{2.11}$$

1) Tools from partial differential equations and uniqueness. We first state the strong maximum principle for the Laplacian, see Theorem 2.2.4 in [**Ev**].

(MP) Let $u \in C^2(D) \cap C(\overline{D})$ satisfy $\Delta u = 0$ on D. Then $\max_{\overline{D}} u = \max_{\partial D} u$. If there is a point $x_0 \in D$ such that $u(x_0) = \max_{\overline{D}} u$, then u is constant.

Hence, if $u, v \in C^2(D) \cap C(\overline{D})$ solve (2.11), then $w = u - v \in C^2(D) \cap C(\overline{D})$ satisfies $\Delta w = 0$ on D and $w = 0$ on ∂D. The maximum principle (MP) thus yields that $\max_{\overline{D}} w = 0$. Similarly, the maximum of $-w$ is 0, so that $w = 0$. This means that the problem (2.11) has at most one solution.

Theorem 2.7 of [**PW**] implies the following version of Hopf's lemma.

(HL) Let $u \in C^2(D) \cap C(\overline{D})$ satisfy $\Delta u \geqslant 0$ on D. Assume that there is a point $x_0 \in \partial D$ such that $u(x_0) = \max_{\overline{D}} u$, $\partial_\nu u(x_0)$ exists, and $\partial_\nu u$ is continuous at x_0. Then either u is constant or $\partial_\nu u(x_0) > 0$.

[4] This section was not part of the lectures.

2) The double layer potential. We want to reformulate (2.11) using an integral operator. To this aim, we first consider the *Newton potential* on $\mathbb{R}^3$ given by $\gamma(x) = \frac{1}{4\pi|x|_2}$ for $x \in \mathbb{R}^3\backslash\{0\}$. It satisfies $\Delta\gamma = 0$ on $\mathbb{R}^3\backslash\{0\}$. We define

$$k(x,y) = \frac{\partial}{\partial\nu(y)}\gamma(x-y) = -(\nabla\gamma)(x-y)\cdot\nu(y) = \frac{(x-y)\cdot\nu(y)}{4\pi|x-y|_2^3} \qquad (2.12)$$

for all $x \in \mathbb{R}^3$ and $y \in \partial D$ with $x \neq y$, where the dot denotes the Euclidean scalar product in $\mathbb{R}^3$. One introduces the *double layer potential* by setting

$$Sg(x) = \int_{\partial D} k(x,y)g(y)\,\mathrm{d}\sigma(y) \qquad (2.13)$$

for all $x \in \mathbb{R}^3\backslash\partial D$ and $g \in C(\partial D)$, where the surface integral is recalled below in part 3). Standard results from Analysis 3 then imply that $Sg \in C^\infty(\mathbb{R}^3\backslash\partial D)$ and $\Delta Sg = 0$ on $\mathbb{R}^3\backslash\partial D$. For each $\varphi \in C(\partial D)$, one thus obtains the solution $u = (Sg)\restriction_D$ of (2.11) if one can find a map $g \in C(\partial D)$ satisfying

$$\lim_{\substack{x\to z\\ x\in D}} Sg(x) = \varphi(z) \qquad \text{for all} \;\; z \in \partial D. \qquad (2.14)$$

3) The surface integral. A compact boundary $\partial\Omega$ of an open subset $\Omega \subseteq \mathbb{R}^m$ belongs to C^k, $k \in \mathbb{N}$, if there are open subsets $\tilde{U}_j$ and $\tilde{V}_j$ of $\mathbb{R}^m$ and C^k–diffeomorphisms $\Psi_j : \tilde{V}_j \to \tilde{U}_j$, $j \in \{1,\dots,l\}$, such that the functions Ψ_j and Ψ_j^{-1} and their derivatives up to order k have continuous extensions to $\partial\tilde{V}_j$ and $\partial\tilde{U}_j$, respectively, $\partial\Omega \subseteq \tilde{V}_1 \cup \cdots \cup \tilde{V}_l$, and Ψ_j maps $V_j := \tilde{V}_j \cap \partial\Omega$ onto $U_j := \tilde{U}_j \cap (\mathbb{R}^{m-1}\times\{0\})$. We set $F_j = \Psi_j^{-1}\restriction_{U_j}$. Below we use these notions for $k = 2$ and $D = \Omega$. We also identify U_j with a subset of $\mathbb{R}^2$ writing $t \in \mathbb{R}^2$ instead of $(t,0) \in \mathbb{R}^3$.

We recall that the surface integral for a (Borel) measurable function $h : \partial D \to \mathbb{R}$ is given by

$$\int_{\partial D} h(y)\,\mathrm{d}\sigma(y) = \sum_{j=1}^{m}\int_{U_j}\varphi_j(F_j(t))h(F_j(t))\sqrt{\det F'(t)^T F'(t)}\,\mathrm{d}t,$$

if the right hand side exists. Here, $0 \leqslant \varphi_j \in C_c^\infty(\mathbb{R}^3)$ satisfy $\operatorname{supp}\varphi_j \subseteq \tilde{V}_j$ and $\sum_{j=1}^m \varphi_j = 1$. This definition does not depend on the choice of $\Psi_j : \tilde{V}_j \to \tilde{U}_j$ and φ_j. Moreover, $\sigma(B) = \int_{\partial D}\mathbb{1}_B\,\mathrm{d}\sigma$ defines a (finite) measure on the Borel sets of ∂D. In particular, the above integral has the usual properties of integrals. We mostly omit the index $j \in \{1,\dots,l\}$.

Recall that for $y = F(t) \in V$ and $t \in U \subseteq \mathbb{R}^2$, the tangent plane of ∂D at y is spanned by $\partial_1 F(t)$ and $\partial_2 F(t)$, where $t \in U \subseteq \mathbb{R}^2$. Taylor's formula applied to $\Psi^{-1} \in C_b^2(\tilde{U})$ at $(t,0)$ yields that

$$\begin{aligned} x := \Psi^{-1}(s,0) &= y + (\Psi^{-1})'(t,0)\begin{pmatrix} s-t \\ 0 \end{pmatrix} + \mathcal{O}(|s-t|_2^2) \\ &= y + F'(t)(s-t) + \mathcal{O}(|s-t|_2^2). \end{aligned}$$

for $s \in U$. Using that $\nu(y)$ is orthogonal to $\partial_j F(t)$, we deduce that

$$(x-y)\cdot\nu(y) = \nu(y)^T F'(t)(s-t) + \mathcal{O}(|s-t|_2^2) = \mathcal{O}(|s-t|_2^2). \qquad (2.15)$$

On the other hand, Ψ^{-1} and Ψ are globally Lipschitz so that

$$c\,|s-t|_2 \leqslant |x-y|_2 \leqslant C\,|s-t|_2 \tag{2.16}$$

for all $x = F(s) \in V$, $y = F(t) \in V$ with $s, t \in U$ and some constants $C, c > 0$. In the following we denote by c various, possibly differing constants.

4) Compactness of a version of S on ∂D. Using (2.12), (2.15) and (2.16), we obtain

$$|k(x,y)| = \frac{|(x-y)\cdot\nu(y)|}{4\pi\,|x-y|_2^3} \leqslant \frac{c}{|x-y|_2} \leqslant \frac{c}{|s-t|_2} \tag{2.17}$$

for all $x = F(s)$ and $y = F(t)$ in V with $x \neq y$. As a result, the integrands

$$\varphi(F(t))k(F(s),F(t))g(F(t))\sqrt{\det F'(t)^T F'(t)}$$

of Sg are bounded by a constant times $|s-t|_2^{-1}\|g\|_\infty$ for all $x = F(s)$ and $y = F(t)$ in ∂D with $x \neq y$. We next set $k(x,x) = 0$ for $x \in \partial D$ and

$$k_n(x,y) = \begin{cases} k(x,y), & |x-y|_2 > 1/n, \\ n^3(4\pi)^{-1}(x-y)\cdot\nu(y), & |x-y|_2 \leqslant 1/n, \end{cases}$$

for $n \in \mathbb{N}$. By means of (2.15) and (2.16), we estimate

$$|k_n(x,y)| \leqslant cn^3|s-t|_2^2 \leqslant c|s-t|_2^{-1} \tag{2.18}$$

if $|x-y|_2 \leqslant 1/n$ because then $|s-t|_2 \leqslant c/n$.

Since k_n is continuous on $\partial D \times \partial D$, we can define an operator $T_n \in \mathcal{B}(C(\partial D))$ by

$$T_n g(x) = \int_{\partial D} k_n(x,y)g(y)\,\mathrm{d}\sigma$$

for $x \in \partial D$ and $g \in C(\partial D)$. As in Example 2.5 a), one shows that T_n is compact thanks to the Arzela-Ascoli theorem. For $g \in C(\partial D)$ and $x \in \partial D$, we set $D(x,n) = D \cap B(x, \frac{1}{n})$ and calculate

$$\begin{aligned}
\int_{\partial D} |(k(x,y)-k_n(x,y))g(y)|\,\mathrm{d}\sigma(y) &= \int_{D(x,n)} |k(x,y)-k_n(x,y)|\,|g(y)|\,\mathrm{d}\sigma(y) \\
&\leqslant \|g\|_\infty \int_{D(x,n)} (|k(x,y)| + |k_n(x,y)|)\,\mathrm{d}\sigma(y) \\
&\leqslant c\,\|g\|_\infty \sum_{j=1}^{l} \int_{U_j \cap B(s,\frac{c}{n})} \frac{\mathrm{d}t}{|s-t|_2} \\
&\leqslant c\,\|g\|_\infty \int_{B(0,\frac{c}{n})} \frac{\mathrm{d}v}{|v|_2} \\
&\leqslant c\,\|g\|_\infty \int_0^{\frac{c}{n}} \frac{r\mathrm{d}r}{r} \leqslant \frac{c\,\|g\|_\infty}{n},
\end{aligned} \tag{2.19}$$

employing (2.16), (2.17), (2.18), and polar coordinates in $\mathbb{R}^2$.

Hence, for each $x \in \partial D$ the function $y \mapsto k(x,y)g(y)$ is integrable for the surface measure σ on ∂D. So we can define

$$Tg(x) := \int_{\partial D} k(x,y)g(y)\,\mathrm{d}\sigma(y),$$

for $x \in \partial D$ and $g \in C(\partial D)$. By (2.19), the functions $T_n g$ converge uniformly on ∂D to Tg as $n \to \infty$ so that $Tg \in C(\partial D)$. Estimate (2.19) actually implies that the differences $T_n - T$ belong to $\mathcal{B}(C(\partial D))$ and converge to 0 in this space. Hence, T is contained in $\mathcal{B}(C(\partial D))$ and it is compact by Proposition 2.3 since all T_n are compact.

5) Facts from potential theory. In Theorems VIII and IX in Chapter VI of [**Ke**] it is shown that

$$\lim_{\substack{x \to z\\ x \in D}} Sg(x) = Tg(z) - \tfrac{1}{2}g(z) \quad \text{and} \quad \lim_{\substack{x \to z\\ x \in \mathbb{R}^3\setminus D}} Sg(x) = Tg(z) + \tfrac{1}{2}g(z), \quad (2.20)$$

for all $z \in \partial D$ and $g \in C(\partial D)$. We set

$$v(x) = \begin{cases} Sg(x), & x \in \mathbb{R}^3\setminus\overline{D}, \\ g(x), & x \in \partial D. \end{cases} \quad (2.21)$$

If $v = 0$ on D, then there exists $\partial_\nu v(y) = 0$ for $y \in \partial D$ due to Theorem X in Chapter VI of [**Ke**].

6) Conclusion. Let $\varphi \in (\partial D)$. In view of (2.14) and (2.20), the function $Sg \in C^2(D)$ has an extension $u \in C^2(D) \cap C(\overline{D})$ solving (2.11) provided that $g \in C(\partial D)$ satisfies $\frac{1}{2}g - Tg = -\varphi$.

Since T is compact, thanks to the Fredholm alternative Corollary 2.8 it remains to establish the injectivity of $\frac{1}{2}I - T$. So let $g_0 \in C(\partial D)$ satisfy $\frac{1}{2}g_0 = Tg_0$. By the previous paragraph, the extension of Sg_0 to $\overline{D}$ then solves (2.11) with $\varphi = 0$. This problem has also the trivial solution $u = 0$. The uniqueness of (2.11) thus yields $Sg_0 = 0$ on D.

Define v_0 by (2.21) with g_0 instead of g. Then $\partial_\nu v_0 = 0$ on ∂D due to the result mentioned after (2.21). We are now looking for a contradiction with Hopf's lemma (HL), employing Sg_0 on $\mathbb{R}^3\setminus\overline{D}$. Fix $r_0 > 0$ such that $\overline{D} \subseteq B(0, r_0)$. For $r \geqslant r_0 + 1$, $x \in \partial B(0,r)$ and $y \in \partial D$, from (2.13) and (2.12) we deduce that

$$|k(x,y)| \leqslant \frac{c}{|x-y|_2^2} \leqslant \frac{c}{(r-r_0)^2} \leqslant \frac{c}{r^2},$$

$$|Sg_0(x)| \leqslant \int_{\partial D} \frac{c}{r^2}\|g_0\|_\infty d\sigma \leqslant \frac{c}{r^2}.$$

Suppose that $g_0 \neq 0$. We can thus fix a radius $r \geqslant r_0 + 1$ such that

$$\|g_0\|_\infty > \max_{x \in \partial B(r)} |Sg_0(x)|.$$

In particular, v_0 is not constant on $\overline{B(r)}\setminus D$. Since $\Delta v_0 = 0$ on $B(r)\setminus\overline{D}$, the strong maximum principle (MP) says that v_0 does not attain its maximum on $B(r)\setminus\overline{D}$. Since the maximum exists on $\overline{B(r)}\setminus D$, it must be attained at a point $y_0 \in \partial D$, and hence $v_0(y_0) > v_0(x)$ for all $x \in B(r)\setminus\overline{D}$. Noting that $-\nu(y_0)$ is the outer unit normal of $B(r)\setminus\overline{D}$ at y_0, we infer from (HL) that $\partial_\nu v_0(y_0) < 0$ contradicting $\partial_\nu v_0 = 0$ on ∂D. As a result, $\frac{1}{2}I - T$ is injective and we have established the claim on (2.11).

CHAPTER 3

Fourier transform, Sobolev spaces, and weak derivatives

The Fourier transform is a fundamental tool in many branches of mathematics and its applications. In the first section of this chapter we study its basic properties in an L^2-context. If one wants to treat partial differential equations in L^2- (or L^p-) spaces, one needs weak derivatives and the Sobolev spaces $W^{k,p}$. The second section gives a brief introduction to these topics and it establishes important links between the Fourier transform and the spaces $W^{k,2}(\mathbb{R}^m)$. We then discuss deeper properties of Sobolev spaces, mostly without proof. The last section is devoted to differential operators using Sobolev spaces and the Fourier transform.

3.1. The Fourier transform

In this section we let $\mathbb{F} = \mathbb{C}$. We start with the definition of the Fourier transform for integrable $f : \mathbb{R}^m \to \mathbb{C}$, where we write

$$\xi \cdot x = \sum_{k=1}^{m} \xi_k x_k$$

for the (real) scalar product of $\xi = (\xi_k)_k \in \mathbb{C}^m$ and $x = (x_k)_k \in \mathbb{C}^m$. Note that $|x|_2^2 = x \cdot x$ for $x \in \mathbb{R}^m$.

Definition 3.1. *Let $f \in L^1(\mathbb{R}^m)$. The* Fourier transform *of f is*

$$\hat{f} : \mathbb{R}^m \to \mathbb{C}; \quad \hat{f}(\xi) = (\mathcal{F}f)(\xi) := (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x. \tag{3.1}$$

In the literature several variants of the constants in (3.1) are used. These choices affect the constants of many results of the theory, so that one has to be careful when using different sources. In Theorem 3.11 b) we see that f can be expressed as the superposition

$$f(x) = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{\mathrm{i}x\cdot\xi} \hat{f}(\xi)\,\mathrm{d}\xi, \qquad x \in \mathbb{R}^m,$$

of 'plane waves' $x \mapsto e_{\mathrm{i}\xi}(x) = \mathrm{e}^{\mathrm{i}x\cdot\xi}$ with coefficients $\hat{f}(\xi)$, provided that $f \in L^2(\mathbb{R}^m)$ and $\hat{f} \in L^1(\mathbb{R}^m)$, see also (3.7). So $\hat{f}(\xi)$ can be view as the component of f at 'frequency' ξ, noting that $e_{\mathrm{i}\xi}(x) = \cos(x\cdot\xi) + \mathrm{i}\sin(x\cdot\xi)$.

Set $\varphi(x,\xi) = \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)$ for $f \in L^1(\mathbb{R}^m)$. Observe that $|\varphi(x,\xi)| = |f(x)|$ is integrable in $x \in \mathbb{R}^m$ for every $\xi \in \mathbb{R}^m$ and that $\mathbb{R}^m \ni \xi \mapsto \varphi(\xi, x)$ is continuous for a.e. $x \in \mathbb{R}^m$. By means of a corollary to the theorem of dominated convergence, we thus conclude

$$\hat{f} \text{ is continuous on } \mathbb{R}^m \quad \text{and} \quad \|\hat{f}\|_\infty \leqslant (2\pi)^{-\frac{m}{2}}\,\|f\|_1 \ \text{ for } f \in L^1(\mathbb{R}^m). \tag{3.2}$$

We discuss some basic and instructive examples.

EXAMPLE 3.2. a) Let $m = 1$ and $f = \mathbb{1}_{[a,b]}$. We then have $\hat{f}(0) = (b-a)/\sqrt{2\pi}$ and

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}x = \frac{\mathrm{i}(\mathrm{e}^{-\mathrm{i}b\xi} - \mathrm{e}^{-\mathrm{i}a\xi})}{\sqrt{2\pi}\,\xi}, \qquad \xi \neq 0.$$

b) Let $m = 1$ and $f(x) = (1+x^2)^{-1}$. Using complex curve integrals, one can show $\hat{f}(\xi) = \sqrt{\pi/2}\,\mathrm{e}^{-|\xi|}$ for $\xi \in \mathbb{R}$, see Analysis 4 and the next example.

In these two examples (non-)rapid decay and (non-)smoothness on f correspond to (non-)smoothness and (non-)rapid decay of $\hat{f}$, respectively, cf. Lemma 3.7. In the next part, both f and $\hat{f}$ are smooth and decay rapidly.

c) Let $\gamma(x) = \exp(-\frac{1}{2}|x|_2^2)$ for $x \in \mathbb{R}^m$ be the standard Gaussian. We show that γ is a fixed vector of the Fourier transform; i.e., $\widehat{\gamma} = \gamma$. Let $\xi \in \mathbb{R}^m$. Observe that $\frac{1}{2}(x+\mathrm{i}\xi)\cdot(x+\mathrm{i}\xi) = \frac{1}{2}|x|_2^2 + \mathrm{i}\xi\cdot x - \frac{1}{2}|\xi|_2^2$. We then obtain

$$\begin{aligned}
\hat{\gamma}(\xi) &= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-(\mathrm{i}\xi\cdot x + \frac{1}{2}|x|_2^2)}\,\mathrm{d}x = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\frac{1}{2}|\xi|_2^2}\,\mathrm{e}^{-\frac{1}{2}(x+\mathrm{i}\xi)\cdot(x+\mathrm{i}\xi)}\,\mathrm{d}x \\
&= \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^m \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}(x_k+\mathrm{i}\xi_k)^2}\,\mathrm{d}x_k = \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^m \frac{1}{\sqrt{2\pi}} \int_{\mathrm{i}\xi_k+\mathbb{R}} \mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z \\
&= \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^m \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}t^2}\,\mathrm{d}t = \gamma(\xi),
\end{aligned}$$

employing the formula $\int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}t^2}\,\mathrm{d}t = \sqrt{2\pi}$ from Analysis 3. In the penultimate equality we shifted the path of integration within $\mathbb{C}$. To justify this shift, we fix $\eta \in \mathbb{R}\backslash\{0\}$ and use the rectangular path Γ_n with vertices $-n$, n, $n+\mathrm{i}\eta$, and $-n+\mathrm{i}\eta$. Cauchy's integral theorem yields $\int_{\Gamma_n} \mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z = 0$. The two vertical lines $S_n^\pm$ in Γ_n have length $|\eta|$, and on $S_n^\pm$ it holds

$$\left|\mathrm{e}^{-\frac{1}{2}z^2}\right| = \mathrm{e}^{-\frac{1}{2}\operatorname{Re}(\pm n+\mathrm{i}\tau)^2} \leqslant \mathrm{e}^{-\frac{1}{2}n^2}\,\mathrm{e}^{\frac{1}{2}|\eta|^2}$$

for $0 \leqslant |\tau| \leqslant |\eta|$. Hence, $\int_{S_n^\pm} \mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z$ tends to 0 as $n \to \infty$, and the above shift is justified. $\Diamond$

Let $f \in L^p(\mathbb{R}^m)$, $1 \leqslant p \leqslant \infty$, and $t, x \in \mathbb{R}^m$. To describe important mapping properties of $\mathcal{F}$, we set

$$e_{\mathrm{i}t}(x) = \mathrm{e}^{\mathrm{i}t\cdot x},$$

and introduce the *translation* operator T_t by

$$(T_t f)(x) = f(x+t).$$

As in Example 4.12 of [**FA**] one sees that $T_t : L^p(\mathbb{R}^m) \to L^p(\mathbb{R}^m)$ is an isometric isomorphism with inverse T_{-t}. For $a > 0$ we further define the *dilation* operator D_a by

$$(D_a f)(x) = f(ax).$$

Observe that $D_{1/a}D_a = D_a D_{1/a} = I$, and that the substitution $y = ax$ yields

$$\|D_a f\|_p^p = \int_{\mathbb{R}^m} |f(ax)|^p\,\mathrm{d}x = \int_{\mathbb{R}^m} a^{-m}\,|f(y)|^p\,\mathrm{d}y = a^{-m}\,\|f\|_p^p \tag{3.3}$$

for $p < \infty$ (and analogously for $p = \infty$). As a result, $a^{m/p} D_a : L^p(\mathbb{R}^m) \to L^p(\mathbb{R}^m)$ is an isometric isomorphism. Finally, also the *reflection* operator

$$R : L^p(\mathbb{R}^m) \to L^p(\mathbb{R}^m); \quad Rf(x) = f(-x),$$

is an isometric isomorphism with inverse $R^{-1} = R$, since we have $R^2 = I$.

Recall that Hölder's inequality implies the continuity of the bilinear map

$$L^p(B) \times L^{p'}(B) \to \mathbb{F}; \quad (f, g) \mapsto \int_B fg \,\mathrm{d}x, \tag{3.4}$$

for all $1 \leqslant p \leqslant \infty$ and Borel sets $B \subseteq \mathbb{R}^m$, where $\frac{1}{p} + \frac{1}{p'} = 1$. Moreover, let $1 \leqslant p, q, r \leqslant \infty$ with $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, $f \in L^p(\mathbb{R}^m)$, and $g \in L^q(\mathbb{R}^m)$. Theorem 2.14 of [**FA**] shows that the *convolution*

$$f * g(x) = (f * g)(x) = \int_{\mathbb{R}^m} f(x - y) g(y) \,\mathrm{d}y, \qquad \text{a.e. } x \in \mathbb{R}^m,$$

belongs to $L^r(\mathbb{R}^m)$ and satisfies *Young's inequality*

$$\|f * g\|_r \leqslant \|f\|_p \, \|g\|_q. \tag{3.5}$$

The bilinear map $(f, g) \mapsto f * g$ is thus continuous from $L^p(\mathbb{R}^m) \times L^q(\mathbb{R}^m)$ to $L^r(\mathbb{R}^m)$. We only need the case $q = 1$, where one has $r = p \in [1, \infty]$. We now prove basic operational properties of the Fourier transform on $L^1(\mathbb{R}^m)$.

PROPOSITION 3.3. *Let $f, g \in L^1(\mathbb{R}^m)$, $t \in \mathbb{R}^m$, and $a > 0$. The following formulas hold.*

a) $\mathcal{F}(T_t f) = e_{\mathrm{i}t} \hat{f}$.

b) $\mathcal{F}(e_{\mathrm{i}t} f) = T_{-t} \hat{f}$.

c) $\mathcal{F}(D_a f) = a^{-m} \, D_{1/a} \hat{f}$.

d) $\mathcal{F}(f * g) = (2\pi)^{\frac{m}{2}} \, \hat{f} \hat{g}$.

PROOF. Let $f, g \in L^1(\mathbb{R}^m)$, $t, \xi \in \mathbb{R}^m$, and $a > 0$. Using the substitutions $y = x + t$ and $z = ax$, we check assertions a), b) and c) by calculating

$$\begin{aligned}
\mathcal{F}(T_t f)(\xi) &= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi \cdot x} f(x + t) \,\mathrm{d}x = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi \cdot (y - t)} f(y) \,\mathrm{d}y \\
&= \mathrm{e}^{\mathrm{i}\xi \cdot t} \hat{f}(\xi), \\
\mathcal{F}(e_{\mathrm{i}t} f)(\xi) &= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi \cdot x} \mathrm{e}^{\mathrm{i}t \cdot x} f(x) \,\mathrm{d}x = \hat{f}(\xi - t), \\
\mathcal{F}(D_a f)(\xi) &= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi \cdot x} f(ax) \,\mathrm{d}x = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} a^{-m} \, \mathrm{e}^{-\mathrm{i}\frac{1}{a}\xi \cdot z} f(z) \,\mathrm{d}z \\
&= a^{-m} \hat{f}(\tfrac{1}{a}\xi).
\end{aligned}$$

To prove part d), we first recall from the proof of Theorem 2.14 in [**FA**] that the map $\mathbb{R}^{2m} \ni (x, y) \mapsto f(y - x) g(x)$ is integrable. Hence, also the function $(x, y) \mapsto \mathrm{e}^{-\mathrm{i}\xi \cdot y} f(y - x) g(x)$ belongs to $L^1(\mathbb{R}^{2m})$. Fubini's theorem then yields

$$\begin{aligned}
\mathcal{F}(f * g)(\xi) &= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi \cdot y} f(y - x) g(x) \,\mathrm{d}x \,\mathrm{d}y \\
&= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi \cdot (y - x)} f(y - x) \mathrm{e}^{-\mathrm{i}\xi \cdot x} g(x) \,\mathrm{d}y \,\mathrm{d}x
\end{aligned}$$

$$= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}\xi\cdot z} f(z)\,\mathrm{d}z\, \mathrm{e}^{-\mathrm{i}\xi\cdot x} g(x)\,\mathrm{d}x = (2\pi)^{\frac{m}{2}} \hat{f}(\xi)\hat{g}(\xi),$$

where we also employed the subtitution $z = y - x$ in one of y–integrals. □

We illustrate the above properties by computing the Fourier transform of a general Gaussian.

Example 3.4. We set $f(x) = \exp(-\frac{a}{2}|x-v|_2^2)$ for all $x \in \mathbb{R}^m$ and some $a > 0$ and $v \in \mathbb{R}^m$. The Fourier transform of this Gaussian function is given by $\hat{f}(\xi) = a^{-m/2}\exp(-\mathrm{i}v\cdot\xi)\exp(-\frac{1}{2a}|\xi|_2^2)$ for $\xi \in \mathbb{R}^m$. In fact, we have $f = T_{-v}D_{\sqrt{a}}\gamma$. Proposition 3.3 and Example 3.2 thus yield

$$\hat{f} = e_{-\mathrm{i}v}\mathcal{F}(D_{\sqrt{a}}\gamma) = e_{-\mathrm{i}v}a^{-\frac{m}{2}}D_{1/\sqrt{a}}\hat{\gamma} = a^{-\frac{m}{2}}e_{-\mathrm{i}v}D_{1/\sqrt{a}}\gamma,$$

as asserted. ◊

As one of its main properties, the Fourier transform maps derivatives into multiplication by polynomials, and vice versa. To state this fact concisely, we use the *multi-index notation*: For $\alpha = (\alpha_1,\dots,\alpha_m) \in \mathbb{N}_0^m$ and $x = (x_1,\dots,x_m) \in \mathbb{R}^m$, we set

$$|\alpha| = \alpha_1+\dots+\alpha_m, \quad x^\alpha = x_1^{\alpha_1}\cdot\ldots\cdot x_m^{\alpha_m}, \quad \partial^\alpha = \partial_1^{\alpha_1}\dots\partial_m^{\alpha_m} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_m^{\alpha_m}}.$$

We further write $x^\alpha f$ for the function $\mathbb{R}^m \ni x \mapsto x^\alpha f(x)$, etc. Observe that

$$|x^\alpha| = |x_1|^{\alpha_1}\cdot\ldots\cdot|x_m|^{\alpha_m} \leqslant |x|_2^{|\alpha|} \leqslant 1 + |x|_2^k \tag{3.6}$$

for $x \in \mathbb{R}^m$ and $|\alpha| \leqslant k$.

To relate the Fourier transform with derivatives, we need a space of smooth functions. Unfortunately, $C_c^\infty(\mathbb{R}^m)$ is not invariant under the Fourier transform. Instead one uses the (somewhat less convenient) 'Schwartz space' on which $\mathcal{F}$ becomes a bijection, as seen below.

Definition 3.5. *For $f \in C^\infty(\mathbb{R}^m)$, $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^m$, we set*

$$p_{k,\alpha}(f) = \sup_{x\in\mathbb{R}^m} |x|_2^k\,|\partial^\alpha f(x)|.$$

We define the Schwartz space $\mathcal{S}_m$ *by*

$$\mathcal{S}_m = \big\{f \in C^\infty(\mathbb{R}^m)\,\big|\, p_{k,\alpha}(f) < \infty \;\; \textit{for all} \;\; k \in \mathbb{N}_0,\ \alpha \in \mathbb{N}_0^m\big\}.$$

Notice that $\mathcal{S}_m$ is a vector space and that all derivatives of f in $\mathcal{S}_m$ decay faster than $|x|_2^{-k}$ for every $k \in \mathbb{N}$, as $|x|_2 \to \infty$. One thus calls $f \in \mathcal{S}_m$ *rapidly decreasing*. Clearly, the map $\gamma(x) = \mathrm{e}^{-|x|_2^2/2}$ belongs to $\mathcal{S}_m$. Moreover, one can replace $p_{k,\alpha}(f)$ by $p_{2k,\alpha}(f)$ in the definition of $\mathcal{S}_m$ without changing $\mathcal{S}_m$. We discuss further basic properties of this space.

Remark 3.6. a) Let $f \in \mathcal{S}_m$, $k \in \mathbb{N}_0$, and $\alpha \in \mathbb{N}_0^m$. We estimate

$$\begin{aligned} |x|_2^k\,|\partial^\alpha f(x)| &= (1+|x|_2^{m+1})^{-1}\,(|x|_2^k + |x|_2^{k+m+1})\,|\partial^\alpha f(x)| \\ &\leqslant (1+|x|_2^{m+1})^{-1}\,(p_{k,\alpha}(f) + p_{k+m+1,\alpha}(f)) \end{aligned}$$

for all $x \in \mathbb{R}^m$. Since the function $x \mapsto (1+|x|_2^{m+1})^{-1}$ is integrable on $\mathbb{R}^m$ (see Analysis 3), we deduce that $g := |x|_2^k\partial^\alpha f$ is in $L^1(\mathbb{R}^m) \cap C_0(\mathbb{R}^m)$, and hence $|x|_2^k\partial^\alpha f$ belongs to $L^p(\mathbb{R}^m)$ for all $p \in [1,\infty]$ in view of $|g|^p \leqslant |g|\,\|g\|_\infty^{p-1}$.

b) Because of $C_c^\infty(\mathbb{R}^m) \subseteq \mathcal{S}_m \subseteq L^p(\mathbb{R}^m)$, Proposition 4.13 in [**FA**] yields that $\mathcal{S}_m$ is dense in $L^p(\mathbb{R}^m)$ for every $p \in [1, \infty)$.

c) Observe that $p_{k,\alpha}$ is a seminorm on $\mathcal{S}_m$ for all $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^m$, where p_{00} is the supnorm. We order these seminorms as a sequence $(p_j)_{j\in\mathbb{N}}$. Due to Proposition 1.8 in [**FA**], the Schwartz space $\mathcal{S}_m$ has the metric

$$\mathrm{d}(f,g) = \sum_{j=1}^{\infty} 2^{-j} \frac{p_j(f-g)}{1+p_j(f-g)},$$

and $\mathrm{d}(f_n, f) \to 0$ as $n \to \infty$ if and only if $p_{k,\alpha}(f - f_n) \to 0$ as $n \to \infty$ for all $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^m$.

Moreover, $\mathcal{S}_m$ is complete for this metric.[1] Indeed, let (f_n) be Cauchy in $(\mathcal{S}_m, \mathrm{d})$. Then $(|x|_2^k \partial^\alpha f_n)_n$ is Cauchy in $C_b(\mathbb{R}^m)$ and thus has a limit $f_{k,\alpha}$ in $C_b(\mathbb{R}^m)$ for each $(k,\alpha) \in \mathbb{N}_0 \times \mathbb{N}_0^n$. Letting $k = 0$, we deduce that $f := f_{0,0}$ belongs to $C^\infty(\mathbb{R}^m)$ and $\partial^\alpha f = f_{0,\alpha}$. So the products $|x|_2^k \partial^\alpha f_n$ also tend to $|x|_2^k \partial^\alpha f$ pointwise, and hence this function coincides with $f_{k,\alpha}$. This means that f is contained in $\mathcal{S}_m$ and $\mathrm{d}(f_n, f) \to 0$ as $n \to \infty$. $\Diamond$

The next lemma deals with the announced relation between Fourier transform and derivatives. We use the *Laplace operator* given by $\Delta = \partial_1^2 + \cdots + \partial_m^2$ and the space of smooth, polynomially bounded functions

$$\mathcal{E}_m = \big\{ f \in C^\infty(\mathbb{R}^m) \,\big|\, \forall\, \alpha \in \mathbb{N}_0^m \;\; \exists\, n_\alpha \in \mathbb{N}_0 : \sup_{|x|_2 \geqslant 1} |x|_2^{-n_\alpha} \, |\partial^\alpha f(x)| < \infty \big\}.$$

Note that Schwartz functions and polynomials belong to $\mathcal{E}_m$.

Lemma 3.7. *Let $f \in \mathcal{S}_m$, $g \in \mathcal{E}_m$, and $\alpha \in \mathbb{N}_0^m$. Then the following assertions hold.*

a) $\widehat{f} \in C^\infty(\mathbb{R}^m)$, $\partial^\alpha \widehat{f} = (-\mathrm{i})^{|\alpha|} \mathcal{F}(x^\alpha f)$, $\mathcal{F}(\partial^\alpha f) = \mathrm{i}^{|\alpha|} \xi^\alpha \widehat{f}$.

b) $\mathcal{F}\Delta f = \mathcal{F}\partial_1^2 f + \cdots + \mathcal{F}\partial_m^2 f = \mathrm{i}^2(\xi_1^2 + \cdots + \xi_m^2)\mathcal{F}f = -|\xi|_2^2 \mathcal{F}f$.

c) The maps $f \mapsto gf$, $f \mapsto \partial^\alpha f$, and R are continuous from $\mathcal{S}_m$ to $\mathcal{S}_m$.

d) The Fourier transform is continuous from $\mathcal{S}_m$ to $\mathcal{S}_m$.

Proof. Let $\xi, x \in \mathbb{R}^m$, $f, f_n \in \mathcal{S}_m$ for $n \in \mathbb{N}$, $g \in \mathcal{E}_m$, $\alpha \in \mathbb{N}_0^m$, and $k \in \mathbb{N}_0$. We show a) for $\alpha = e_j$, the assertion then follows by induction. There exists

$$\frac{\partial}{\partial \xi_j} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) = -\mathrm{i}x_j \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) =: \varphi_j(\xi, x),$$

and $\mathbb{R}^m \ni x \mapsto |\varphi_j(\xi, x)| = |x_j f(x)|$ is integrable by Remark 3.6. A corollary to the theorem of dominated convergence thus shows that

$$\exists\; \frac{\partial}{\partial \xi_j} \hat{f}(\xi) = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} -\mathrm{i}x_j \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) \,\mathrm{d}x = -\mathrm{i}\mathcal{F}(x_j f)(\xi).$$

For the second part of a) we write $[-n, n]^l = C_n^l$ for $l, n \in \mathbb{N}$ and $x = (x', x_j)$ with $x' = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m) \in \mathbb{R}^{m-1}$. Using that $\partial_j f$ and f are integrable and integrating by parts in x_j, we compute

$$\mathcal{F}(\partial_j f)(\xi) = (2\pi)^{-\frac{m}{2}} \lim_{n\to\infty} \int_{C_n^{m-1}} \int_{-n}^{n} \mathrm{e}^{-\mathrm{i}\xi\cdot(x', x_j)} \partial_j f(x', x_j) \,\mathrm{d}x_j \,\mathrm{d}x'$$

[1] The proof was omitted in the lectures.

$$= (2\pi)^{-\frac{m}{2}} \lim_{n\to\infty} \Big[\int_{C_n^m} \mathrm{i}\xi_j \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x + \int_{C_n^{m-1}} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x', x_j)\Big|_{x_j=-n}^{n} \mathrm{d}x' \Big]$$
$$= \mathrm{i}\xi_j\, \widehat{f}(\xi).$$

Here the second integral J_n in the second line tends to 0 as $n \to \infty$ since

$$|J_n| \leqslant \sum_{N=\pm n} \int_{C_n^{m-1}} |(x',N)|_2^{-m}\, |(x',N)|_2^m\, |f(x',N)|\,\mathrm{d}x' \leqslant 2(2n)^{m-1} n^{-m} p_{m0}(f).$$

Assertion b) is a consequence of a). For part c), note that the function $|x|_2^k \partial^\alpha(fg)$ is a linear combination of terms $|x|_2^{k+n_\gamma} \partial^\beta f\, |x|_2^{-n_\gamma} \partial^\gamma g$ for $\beta, \gamma \in \mathbb{N}_0^m$ with $\beta + \gamma = \alpha$. Since $g \in \mathcal{E}_m$, also employing (3.6) we obtain

$$p_{k,\alpha}(fg) \leqslant c \sum_{\beta \leqslant \alpha} (p_{k+l,\beta}(f) + p_{k,\beta}(f)),$$

where $l = \max_{|\gamma|\leqslant|\alpha|} n_\gamma$ and c only depends on k, α, m and g. Hence, fg belongs to $\mathcal{S}_m$. The asserted continuity of $f \mapsto fg$ follows by replacing f with $f - f_n$. Similarly, one checks the second and third part of c).

By means of claims a) and b), we further compute

$$|\xi|_2^{2k} \partial^\alpha \widehat{f} = (-\mathrm{i})^{|\alpha|} |\xi|_2^{2k} \mathcal{F}(x^\alpha f) = (-\mathrm{i})^{|\alpha|} (-1)^k \mathcal{F}(\Delta^k(x^\alpha f)).$$

Due to part c) and Remark 3.6, the function $\Delta^k(x^\alpha f)$ belongs to $\mathcal{S}_m \subseteq L^1(\mathbb{R}^m)$ so that its Fourier transform can be estimated by means of (3.2). This means that $\widehat{f}$ is contained in $\mathcal{S}_m$, and we also obtain

$$p_{2k,\alpha}(\mathcal{F}(f - f_n)) \leqslant c \|\Delta^k(x^\alpha(f - f_n))\|_1.$$

Using again Leibniz' rule and Remark 3.6, the term on the right-hand side can be bounded by a linear combination of certain seminorms $p_{i,\beta}(f - f_n)$; i.e., $\mathcal{F}_m : \mathcal{S}_m \to \mathcal{S}_m$ is continuous. □

We infer the *Riemann–Lebesgue Lemma*, which improves on (3.2).

Corollary 3.8. *If $f \in L^1(\mathbb{R}^m)$, then $\hat{f} \in C_0(\mathbb{R}^m)$. Hence, $\mathcal{F}$ belongs to $\mathcal{B}(L^1(\mathbb{R}^m), C_0(\mathbb{R}^m))$.*

Proof. Let $f \in L^1(\mathbb{R}^m)$. Remark 3.6 provides functions $f_n \in \mathcal{S}_m$ converging to f in $L^1(\mathbb{R}^m)$. Lemma 3.7 shows that $\hat{f}_n$ belongs to $\mathcal{S}_m \subseteq C_0(\mathbb{R}^m)$. By (3.2), the functions $\hat{f}_n$ tend to $\hat{f}$ in supnorm so that $\hat{f}$ is contained in $C_0(\mathbb{R}^m)$. The second assertion then follows from (3.2). □

The next lemma is the crucial step towards the main results of this section. Observe that in its second part a double integral disappears due to cancellations of the highly oscillating integrands.

Lemma 3.9. *The following assertions hold.*

a) $\int_{\mathbb{R}^m} \hat{f} g\,\mathrm{d}x = \int_{\mathbb{R}^m} f \hat{g}\,\mathrm{d}x$ for all $f, g \in \mathcal{S}_m$.

b) $\mathcal{F}^2 = R$; i.e., $(\mathcal{F}\mathcal{F}f)(x) = f(-x)$ for all $f \in \mathcal{S}_m$ and $x \in \mathbb{R}^m$.

Proof. Let $f, g \in \mathcal{S}_m$. Since $(x, y) \mapsto \mathrm{e}^{-\mathrm{i}y\cdot x} f(x)g(y)$ is integrable on $\mathbb{R}^{2m}$, Fubini's theorem yields

$$\int_{\mathbb{R}^m} \hat{f}(y) g(y)\,\mathrm{d}y = \int_{\mathbb{R}^m} (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}y\cdot x} f(x) g(y)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_{\mathbb{R}^m} f(x)\,(2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}y\cdot x} g(y)\,\mathrm{d}y\,\mathrm{d}x = \int_{\mathbb{R}^m} f(x)\hat{g}(x)\,\mathrm{d}x.$$

In the second assertion one is led to the integrand $\mathrm{e}^{-\mathrm{i}y\cdot\xi}\mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)$ which is not integrable for $(\xi, x) \in \mathbb{R}^{2m}$. So Fubini's theorem does not apply directly, and one has to use a regularization. To that purpose, fix $\xi \in \mathbb{R}^m$ and $a > 0$. Set $h_a = e_{-\mathrm{i}\xi} D_a \gamma \in \mathcal{S}_m$; i.e., $h_a(y) = \mathrm{e}^{-\mathrm{i}\xi\cdot y} \exp(-\frac{a^2}{2}|y|_2^2)$ for $y \in \mathbb{R}^m$. Due to the theorem of dominated convergence with the majorant $|\hat{f}|$, the integral

$$J_a := \int_{\mathbb{R}^m} \hat{f}(y) h_a(y)\,\mathrm{d}y = \int_{\mathbb{R}^m} \hat{f}(y)\mathrm{e}^{-\mathrm{i}\xi\cdot y}\gamma(ay)\,\mathrm{d}y$$

converges to $(2\pi)^{\frac{m}{2}}(\mathcal{F}\hat{f})(\xi)$ as $a \to 0$. On the other hand, part a), Proposition 3.3 and Example 3.2 imply that

$$\begin{aligned} J_a &= \int_{\mathbb{R}^m} f(x)\,\mathcal{F}(e_{-\mathrm{i}\xi} D_a \gamma)(x)\,\mathrm{d}x = \int_{\mathbb{R}^m} f(x)\,a^{-m}(T_\xi D_{1/a}\gamma)(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^m} f(x)\,a^{-m}\gamma(\tfrac{1}{a}(x+\xi))\,\mathrm{d}x = \int_{\mathbb{R}^m} f(az-\xi)\gamma(z)\,\mathrm{d}z, \end{aligned}$$

where we also substitute $z = \frac{1}{a}(x+\xi)$. By means of the theorem of dominated convergence with the majorant $\|f\|_\infty \gamma$ we conclude that J_a tends to $f(-\xi)\|\gamma\|_1 = (2\pi)^{m/2} f(-\xi)$ as $a \to 0$, which shows assertion b). □

We now establish the asserted bijectity of $\mathcal{F}$ on $\mathcal{S}_m$ and compute its inverse. We further complement Proposition 3.3 d) on convolutions. Equation (3.8) will be crucial to extend our results to $L^2(\mathbb{R}^m)$.

Proposition 3.10. *The Fourier transform $\mathcal{F} : \mathcal{S}_m \to \mathcal{S}_m$ is a homeomorphism with $\mathcal{F}^4 = I$ and $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$. Moreover, for all $f, g \in \mathcal{S}_m$ and $x \in \mathbb{R}^m$ we have*

$$\mathcal{F}^{-1}g(x) = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{\mathrm{i}x\cdot\xi} g(\xi)\,\mathrm{d}\xi, \tag{3.7}$$

$$(\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2} := \int_{\mathbb{R}^m} f(x)\overline{g(x)}\,\mathrm{d}x, \tag{3.8}$$

$$f * g \in \mathcal{S}_m, \tag{3.9}$$

$$\mathcal{F}(fg) = (2\pi)^{-\frac{m}{2}} \hat{f} * \hat{g}. \tag{3.10}$$

Proof. Lemma 3.9 shows that $I = R^2 = \mathcal{F}^4 = \mathcal{F}\mathcal{F}^3 = \mathcal{F}^3\mathcal{F}$ on $\mathcal{S}_m$ so that $\mathcal{F} : \mathcal{S}_m \to \mathcal{S}_m$ has the continuous inverse $\mathcal{F}^3 = R\mathcal{F}$. This fact already gives (3.7). Let $f, g \in \mathcal{S}_m$ and $x, \xi \in \mathbb{R}^m$. Equation (3.7) then yields

$$\begin{aligned} \mathcal{F}(\overline{\mathcal{F}g})(\xi) &= (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i}x\cdot\xi}\, \overline{\hat{g}(\xi)}\,\mathrm{d}\xi = (2\pi)^{-\frac{m}{2}} \overline{\int_{\mathbb{R}^m} \mathrm{e}^{\mathrm{i}x\cdot\xi} \hat{g}(\xi)\,\mathrm{d}\xi} \\ &= \overline{(\mathcal{F}^{-1}\mathcal{F}g)(x)} = \overline{g(x)}. \end{aligned}$$

So we can deduce from Lemma 3.9 a) that

$$(\hat{f}|\hat{g})_{L^2} = \int_{\mathbb{R}^m} \hat{f}\,\overline{\hat{g}}\,\mathrm{d}\xi = \int_{\mathbb{R}^m} f\,\mathcal{F}(\overline{\mathcal{F}g})\,\mathrm{d}x = \int_{\mathbb{R}^m} f\overline{g}\,\mathrm{d}x = (f|g)_{L^2}.$$

For the final two assertions, Proposition 3.3 and Lemma 3.7 imply that $\mathcal{F}(f * g) = (2\pi)^{m/2}\hat{f}\hat{g} =: \varphi$ belongs to $\mathcal{S}_m$. Hence, the convolution $f * g =$

$\mathcal{F}^{-1}\varphi$ is an element $\mathcal{S}_m$. Replacing f and g by $\hat{f}$ and $\hat{g}$, we further infer

$$\mathcal{F}(\hat{f} * \hat{g}) = (2\pi)^{\frac{m}{2}} \mathcal{F}^2(f)\mathcal{F}^2(g) = (2\pi)^{\frac{m}{2}} R(fg) = (2\pi)^{\frac{m}{2}} \mathcal{F}^2(fg)$$

since $R = \mathcal{F}^2$. We apply $\mathcal{F}^{-1}$ and arrive at (3.10). $\square$

The equality (3.8) yields $\|\mathcal{F}f\|_2 = \|f\|_2$ for all $f \in \mathcal{S}_m$. Since $\mathcal{S}_m$ is dense in $L^2(\mathbb{R}^m)$ by Remark 3.6, we can extend $\mathcal{F}$ to a linear isometry $\mathcal{F}_2 : L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^m)$ which is also called *Fourier transform* (use Lemma 2.13 of [**FA**]). Let $f \in L^2(\mathbb{R}^m) \cap L^1(\mathbb{R}^m)$. By Theorem 4.21 in [**FA**], we have functions $f_n \in C_c^\infty(\mathbb{R}^m) \subseteq \mathcal{S}_m$ which tend to f in $L^2(\mathbb{R}^m)$ and in $L^1(\mathbb{R}^m)$. Since $\mathcal{F}f_n \to \mathcal{F}_2 f$ in $L^2(\mathbb{R}^m)$, there is a subsequence $\mathcal{F}f_{n_j}$ converging to $\mathcal{F}_2 f$ a.e. as $j \to \infty$ due to Riesz–Fischer. On the other hand, $\mathcal{F}f_{n_j}$ converges uniformly to $\mathcal{F}f$ by (3.2). Thus, $\mathcal{F}_2 f = \mathcal{F}f$ a.e.. We now write $\mathcal{F} : L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^m)$ instead of $\mathcal{F}_2$, and also $\mathcal{F}_2 f = \hat{f}$.

Warning: $\mathcal{F}f$ is **not** given by the formula (3.1) if $f \in L^2(\mathbb{R}^m)\backslash L^1(\mathbb{R}^m)$.

In the next theorem we collect the main properties of $\mathcal{F}$ on $L^2(\mathbb{R}^m)$, except for its behavior under derivatives which will be dealt with in Theorem 3.25. Unitary operators are introduced in Definition 5.43 in [**FA**] or in Section 4.1.

THEOREM 3.11. *The Fourier transform on $\mathcal{S}_m$ extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^m)$ which is given by* (3.1) *on $L^2(\mathbb{R}^m) \cap L^1(\mathbb{R}^m)$. Let $f, g \in L^2(\mathbb{R}^m)$, $h \in L^1(\mathbb{R}^m) \cap L^2(\mathbb{R}^m)$, $t \in \mathbb{R}^m$, and $a > 0$. Then the following assertions hold.*

a) $\mathcal{F}^2 = R,\ \ \mathcal{F}^4 = I,\ \ \mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$.

b) $\mathcal{F}^{-1}h(x) = (2\pi)^{-m/2}\int_{\mathbb{R}^m} e^{ix\cdot\xi} h(\xi)\,d\xi$ *for* $x \in \mathbb{R}^m$ (inversion formula).

c) $(\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2}$ (Plancherel identity).

d) $\int_{\mathbb{R}^m} \hat{f} g\,dx = \int_{\mathbb{R}^m} f\hat{g}\,dx$.

e) $\mathcal{F}(T_t f) = e_{it}\hat{f},\ \ \mathcal{F}(e_{it}f) = T_{-t}\hat{f},\ \ \mathcal{F}(D_a f) = a^{-m} D_{1/a}\hat{f}$.

f) Let $\chi, \psi \in L^1(\mathbb{R}^m)$ with $\varphi := R\mathcal{F}\psi \in C_0(\mathbb{R}^m)$ and write $\hat{\varphi} = \psi$. Then $\mathcal{F}(\chi * f) = (2\pi)^{\frac{m}{2}}\,\hat{\chi}\hat{f},\quad \mathcal{F}(\varphi f) = (2\pi)^{-\frac{m}{2}}\,\hat{\varphi} * \hat{f}$ (convolution theorem).

PROOF. As seen above, $\mathcal{F}$ is an isometry on $L^2(\mathbb{R}^m)$. The equations $\mathcal{F}^2 = R$, $\mathcal{F}^4 = I$, and those in assertions c)–e) hold on the dense subspace $\mathcal{S}_m$ as shown in Proposition 3.3, Lemma 3.9 and Proposition 3.10. Since the maps $\mathcal{F}$, R, T_t, D_a, $f \mapsto e_{it}f$ and the scalar product (cf. (3.4)) are continuous from $L^2(\mathbb{R}^m)$, resp. from $L^2(\mathbb{R}^m) \times L^2(\mathbb{R}^m)$, to $L^2(\mathbb{R}^m)$, these identities can be extended to $L^2(\mathbb{R}^m)$ by approximation.

From $\mathcal{F}^4 = I$ we infer $I = \mathcal{F}\mathcal{F}^3 = \mathcal{F}^3\mathcal{F}$ so that $\mathcal{F} : L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^m)$ has the inverse $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$, and parts a) and b) are shown. As a bijective isometry on a Hilbert space, $\mathcal{F}$ is unitary by Proposition 5.52 in [**FA**].

The first part of claim f) follows from Proposition 3.3 by approximation, using (3.5). For the second part, take $f_n, \psi_n \in \mathcal{S}_m$ with $f_n \to f$ in $L^2(\mathbb{R}^m)$ and $\psi_n \to \psi$ in $L^1(\mathbb{R}^m)$ as $n \to \infty$. Then $\varphi_n := \mathcal{F}^{-1}\psi_n = R\mathcal{F}\psi$ tends to $\varphi = R\mathcal{F}\psi$ in $L^\infty(\mathbb{R}^m)$ by (3.2). Hence, $\varphi_n f_n$ converges to φf and $\widehat{\varphi_n} * \widehat{f_n} = \psi_n * \widehat{f_n}$ to $\psi * \hat{f} = \hat{\varphi} * \hat{f}$ in $L^2(\mathbb{R}^m)$ due to (3.4) and (3.5), respectively. Equation (3.10) and the continuity of $\mathcal{F}$ now imply the second part of f). $\square$

In Theorem 3.25 we complement the above result. In the appendix Section 3.6 we treat the Fourier transform in a more general framework.

3.2. Basic properties of Sobolev spaces

In the remainder of the chapter $U \subseteq \mathbb{R}^m$ is open and non-empty. We are looking for properties of C^1-functions which can be generalized to a theory of derivatives suited to L^p-spaces. Looking at the theorem of dominated convergence, for instance, one sees that here the basic concepts should not be based on pointwise limits. It turns out that integration by parts is an excellent starting point for such a theory.

Let $f \in C^1(U)$ and $\varphi \in C_c^\infty(U)$. Extend $\varphi f \in C_c^1(U)$ by 0 to a function $g \in C_c^1(\mathbb{R}^m)$. Then $\partial_1 g = \partial_1\varphi\, f + \varphi \partial_1 f$ on U. Take a number $a > 0$ such that $\operatorname{supp} g \subseteq (-a,a)^m =: C^m$ and write $x = (x_1, x')$. We then derive

$$\begin{aligned}\int_U \partial_1 f\,\varphi\,\mathrm{d}x &= -\int_U f\partial_1\varphi\,\mathrm{d}x + \int_U \partial_1 g\,\mathrm{d}x\\ &= -\int_U f\partial_1\varphi\,\mathrm{d}x + \int_{C^{m-1}}\int_{-a}^{a} \partial_1 g(x_1,x')\,\mathrm{d}x_1\,\mathrm{d}x'\\ &= -\int_U f\partial_1\varphi\,\mathrm{d}x + \int_{C^{m-1}} (g(a,x') - g(-a,x'))\,\mathrm{d}x'\\ &= -\int_U f\partial_1\varphi\,\mathrm{d}x.\end{aligned}$$

Inductively one shows that

$$\int_U \partial^\alpha f\,\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\int_U f\partial^\alpha\varphi\,\mathrm{d}x, \tag{3.11}$$

for all $f \in C^k(U)$, $\varphi \in C_c^\infty(U)$, and $\alpha \in \mathbb{N}_0^m$ with $|\alpha| \leqslant k$. Throughout derivatives like ∂^α only act on the following map (if there are no parentheses).

To imitate (3.11) in a definition, we set

$$L^p_{\mathrm{loc}}(U) = \big\{ f : U \to \mathbb{F} \,\big|\, f \text{ measurable, } f\restriction_K \in L^p(K) \text{ for all compact } K \subseteq U\big\}$$

for $p \in [1,\infty]$. Note $L^p(U) \subseteq L^p_{\mathrm{loc}}(U) \subseteq L^1_{\mathrm{loc}}(U)$. We extend $f \in L^p_{\mathrm{loc}}(U)$ by 0 to a measurable function $f : \mathbb{R}^m \to \mathbb{F}$ without further notice. Convergence in $L^p_{\mathrm{loc}}(U)$ means that the restrictions to K converge in $L^p(K)$ for all compact $K \subseteq U$. This limit concept can be described by a complete metric as in Example 1.9 of [**FA**]. We introduce a new notion of derivative.

Definition 3.12. *Let $f \in L^1_{\mathrm{loc}}(U)$ and $\alpha \in \mathbb{N}_0^m$. Let $g \in L^1_{\mathrm{loc}}(U)$ satisfy*

$$\int_U g\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\int_U f\partial^\alpha\varphi\,\mathrm{d}x \tag{3.12}$$

for all $\varphi \in C_c^\infty(U)$. Then $g =: \partial^\alpha f$ is called weak derivative *of f. We set*

$$W_\alpha(U) = \big\{ f \in L^1_{\mathrm{loc}}(U) \,\big|\, \exists\, \partial^\alpha f \in L^1_{\mathrm{loc}}(U)\big\}.$$

For $k \in \mathbb{N}$ and $p \in [1,\infty]$, one defines the Sobolev spaces *by*

$$W^{k,p}(U) = \big\{ f \in L^p(U) \,\big|\, f \in W_\alpha(U),\ \partial^\alpha f \in L^p(U) \ \ \textit{for all } |\alpha| \leqslant k\big\}$$

and endows them with

$$\|f\|_{k,p} = \begin{cases} \Big(\sum\limits_{0\leqslant|\alpha|\leqslant k} \|\partial^\alpha f\|_p^p \Big)^{1/p}, & 1 \leqslant p < \infty, \\ \max\limits_{0\leqslant|\alpha|\leqslant k} \|\partial^\alpha f\|_\infty, & p = \infty, \end{cases}$$

where $\partial^0 f := f$. *We set* $W^{0,p}_{(\mathrm{loc})}(U) = L^p_{(\mathrm{loc})}(U)$, $\partial^{e_j} = \partial_j$, $\partial = \partial_1$ *if* $m = 1$, *and*

$$W^{k,p}_{\mathrm{loc}} = \big\{ f \in L^p_{\mathrm{loc}}(U) \,\big|\, f \in W_\alpha(U),\ \partial^\alpha f \in L^p_{\mathrm{loc}}(U) \text{ for all } |\alpha| \leqslant k \big\}.$$

As usual, $L^p_{\mathrm{loc}}(U)$, $W_\alpha(U)$, $W^{k,p}_{\mathrm{loc}}(U)$, and $W^{k,p}(U)$ are spaces of equivalence classes modulo $\mathcal{N} = \{f : U \to \mathbb{F} \,|\, f \text{ measurable},\ f = 0 \text{ a.e.}\}$. After recalling an important fact from Lemma 4.15 in [**FA**] (se also Lemma 3.50 in the appendix), we list basic properties of weak derivatives.

LEMMA 3.13. *Let* $g \in L^1_{\mathrm{loc}}(U)$ *satisfy* $\int_U g\varphi \,\mathrm{d}x = 0$ *for all* $\varphi \in C_c^\infty(U)$. *Then* $g = 0$ *a.e..*

REMARK 3.14. Let $\alpha, \beta \in \mathbb{N}_0^m$, $p \in [1, \infty]$, and $k \in \mathbb{N}$.

a) Lemma 3.13 implies that $\partial^\alpha f$ is uniquely determined for a.e. $x \in U$. From (3.11) we then infer that $C^k(U) + \mathcal{N}$ is contained in $W_\alpha(U)$ for $|\alpha| \leqslant k$ and that weak and classical derivatives coincide for $f \in C^k(U)$.

b) $W_\alpha(U)$ is a vector space and the map $\partial^\alpha : W_\alpha(U) \to L^1_{\mathrm{loc}}(U)$ is linear.

c) Let $f \in W_\alpha(U) \cap W_{\alpha+\beta}(U)$. Then $\partial^\alpha f$ belongs to $W_\beta(U)$ and $\partial^\beta\partial^\alpha f = \partial^{\alpha+\beta} f$. If also $f \in W_\beta(U)$, then $\partial^\beta f \in W_\alpha(U)$ and $\partial^\alpha\partial^\beta f = \partial^{\alpha+\beta} f = \partial^\beta\partial^\alpha f$. For $f \in W^{|\alpha|,1}_{\mathrm{loc}}(U)$ we obtain $\partial^\alpha f = \partial_1^{\alpha_1} \dots \partial_m^{\alpha_m} f$ and may change the order.

PROOF. Let $\varphi \in C_c^\infty(U)$ and $f \in W_\alpha(U) \cap W_{\alpha+\beta}(U)$. We just check the definition of $\partial^\beta(\partial^\alpha f)$, where put ∂^α on φ by means of (3.12) and then use Schwarz' theorem from Analysis 2. So we compute

$$\begin{aligned}(-1)^{|\beta|} \int_U \partial^\alpha f \,\partial^\beta \varphi \,\mathrm{d}x &= (-1)^{|\alpha|+|\beta|} \int_U f \,\partial^\alpha\partial^\beta\varphi \,\mathrm{d}x = (-1)^{|\alpha+\beta|} \int_U f \,\partial^{\alpha+\beta}\varphi \,\mathrm{d}x \\ &= \int_U \partial^{\alpha+\beta} f \,\varphi \,\mathrm{d}x;\end{aligned}$$

i.e., $\partial^\alpha f \in W_\beta(U)$ and $\partial^\beta\partial^\alpha f = \partial^{\alpha+\beta} f$. The second claim follows by interchanging α and β; the last one is clear. □

d) $(W^{k,p}(U), \|\cdot\|_{k,p})$ is a normed vector space. A sequence $(f_n)_n$ converges in $W^{k,p}(U)$ if and only if $(\partial^\alpha f_n)_n$ converges in $L^p(U)$ for each α with $|\alpha| \leqslant k$. Note that $\|f\|_{1,p}^p = \|f\|_p^p + \|\,|\nabla f|_p\,\|_p^p$ for $f \in W^{1,p}(U)$ and $p < \infty$.

e) The map $J : W^{k,p}(U) \to L^p(U)^N;\ f \mapsto (\partial^\alpha f)_{|\alpha|\leqslant k}$, is a linear isometry, where N is the number of α in $\mathbb{N}_0^m$ with $|\alpha| \leqslant k$ and $L^p(U)^N$ has the norm $\|(f_j)_j\| = |(\|f_j\|_p)_j|_p$. Since the p-norm and the 1-norm on $\mathbb{R}^N$ are equivalent, there is a constant $c_N > 0$ such that

$$c_N \sum_{0\leqslant|\alpha|\leqslant k} \|\partial^\alpha f\|_p \leqslant \|f\|_{k,p} \leqslant \sum_{0\leqslant|\alpha|\leqslant k} \|\partial^\alpha f\|_p$$

for all $f \in W^{k,p}(U)$. ◊

We continue with simple, but instructive one-dimensional examples.

EXAMPLE 3.15. a) Let $f \in C(\mathbb{R})$ be such that $f_\pm := f{\restriction}_{\mathbb{R}_{\gtrless 0}}$ belong to $C^1(\mathbb{R}_{\gtrless 0})$. We then have $f \in W^{1,1}_{\mathrm{loc}}(\mathbb{R})$ with

$$\partial f = \left\{ \begin{array}{ll} f_+' & \text{on } \mathbb{R}_{\geqslant 0}, \\ f_-' & \text{on } \mathbb{R}_- = (-\infty, 0) \end{array} \right\} =: g.$$

For $f(x) = |x|$, we thus obtain $\partial f = \mathbb{1}_{\mathbb{R}_{\geqslant 0}} - \mathbb{1}_{\mathbb{R}_-}$.

PROOF. For every $\varphi \in C_c^\infty(\mathbb{R})$, we compute

$$\begin{aligned} \int_{\mathbb{R}} f\varphi' \,\mathrm{d}t &= \int_{-\infty}^0 f_-\varphi' \,\mathrm{d}t + \int_0^\infty f_+\varphi' \,\mathrm{d}t \\ &= -\int_{-\infty}^0 f_-'\varphi \,\mathrm{d}t + f_-\varphi\big|_{-\infty}^0 - \int_0^\infty f_+'\varphi dt + f_+\varphi\big|_0^\infty \\ &= -\int_{\mathbb{R}} g\varphi \,\mathrm{d}t, \end{aligned}$$

since $f_+(0) = f_-(0)$ by the continuity of f. □

b) The function $f = \mathbb{1}_{\mathbb{R}_{\geqslant 0}}$ does not belong to $W^{1,1}(\mathbb{R})$.

PROOF. Assume there would exist $g = \partial f \in L^1_{\mathrm{loc}}(\mathbb{R})$. Then we obtain

$$\int_{\mathbb{R}} g\varphi \,\mathrm{d}t = -\int_{\mathbb{R}} \mathbb{1}_{\mathbb{R}_{\geqslant 0}}\varphi' \,\mathrm{d}t = -\int_0^\infty \varphi'(t) \,\mathrm{d}t = \varphi(0)$$

for all $\varphi \in C_c^\infty(\mathbb{R})$. Taking φ with $\operatorname{supp}\varphi \subseteq \mathbb{R}_+ = (0,\infty)$, from Lemma 3.13 we deduce $g = 0$ on $\mathbb{R}_+$. Similarly, it follows that $g = 0$ on $\mathbb{R}_-$. The equation in display then yields $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$, which is wrong. □

c) Set $f(x, y) = \mathbb{1}_{\mathbb{R}_{\geqslant 0}}(x)$ for $(x, y) \in \mathbb{R}^2$. Observe that $\int_{\mathbb{R}} \partial^\alpha \varphi(x, y) \,\mathrm{d}y = 0$ for all $\varphi \in C_c^\infty(\mathbb{R}^2)$ and $\alpha \in \mathbb{N}_0^2$ with $\alpha_2 \neq 0$. As a result, the weak derivatives $\partial^\alpha f = 0$ exist for such α, e.g., $\partial_2 f = 0 = \partial^{(1,1)} f = 0$. However, as in part b) one sees that the weak derivative $\partial_1 f$ does not exist. ◊

So far we have just used the definition of weak derivatives by duality. For further examples and deeper results one needs mollifiers, which we recall and discuss next.

Fix a function $0 \leqslant \chi \in C^\infty(\mathbb{R}^m)$ with support $\overline{B}(0, 1)$ and $\chi > 0$ on $B(0, 1)$. For $x \in \mathbb{R}^m$ and $\varepsilon > 0$, we set

$$k(x) = \frac{1}{\|\chi\|_1}\chi(x) \qquad \text{and} \qquad k_\varepsilon(x) = \varepsilon^{-m} k(\tfrac{1}{\varepsilon}x).$$

Note that $0 \leqslant k_\varepsilon \in C^\infty(\mathbb{R}^m)$, $k_\varepsilon(x) > 0$ if and only if $|x|_2 < \varepsilon$, and $\|k_\varepsilon\|_1 = 1$. Let $f \in L^1_{\mathrm{loc}}(U)$ and $\varepsilon > 0$. We now introduce the *mollifier* G_ε by

$$G_\varepsilon f(x) = \int_{B(x,\varepsilon)} k_\varepsilon(x - y) f(y) \,\mathrm{d}y = \int_{B(0,\varepsilon)} k_\varepsilon(z) f(x - z) \,\mathrm{d}z, \qquad (3.13)$$

for $x \in U^\varepsilon := \{x \in U \,|\, 0 < \varepsilon < \mathrm{d}(x, \partial U)\}$. If the 0-extension of f belongs to $L^1_{\mathrm{loc}}(\mathbb{R}^m)$, the above definition works for $x \in \mathbb{R}^m$ or $x \in U$ and all $\varepsilon > 0$, and we have $G_\varepsilon f = k_\varepsilon * f$ for this extension. For a subset S of a Banach space and $\varepsilon > 0$, we define

$$S_\varepsilon = S + \overline{B}(0, \varepsilon).$$

From Proposition 4.13 in [**FA**] and its proof we recall that

$$G_\varepsilon f \in C^\infty(U^\varepsilon) \quad \text{and} \quad G_\varepsilon f \in C^\infty(\mathbb{R}^m) \quad \text{if } f \in L^1_{\mathrm{loc}}(\mathbb{R}^m), \tag{3.14}$$

$$\operatorname{supp} G_\varepsilon f \subseteq S_\varepsilon \quad \text{for } S := \operatorname{supp} f, \quad S_\varepsilon \text{ is compact if } S \text{ is compact}, \tag{3.15}$$

$$\|G_\varepsilon f\|_{L^p(U)} \leqslant \|G_\varepsilon f\|_{L^p(\mathbb{R}^m)} \leqslant \|f\|_p \quad \text{if } f \in L^p(U) \text{ and } 1 \leqslant p \leqslant \infty, \tag{3.16}$$

$$G_\varepsilon f \to f \text{ in } L^p(U) \quad \text{as } \varepsilon \to 0 \quad \text{if } f \in L^p(U) \text{ and } 1 \leqslant p < \infty, \quad \text{or if } p = \infty \text{ and } f \in C_{ub}(U). \tag{3.17}$$

The next lemma is the key to many properties of weak derivatives. Abusing notation, we write $G_\varepsilon f \to f$ in $L^p_{\mathrm{loc}}(U)$ as $\varepsilon \to 0$ if we have $(G_\varepsilon f)\upharpoonright_K \to f\upharpoonright_K$ in $L^p(K)$ as $\varepsilon \to 0$ for all compact $K \subseteq U$ and $0 < \varepsilon < \mathrm{d}(K, \partial U)$.

Lemma 3.16. *Let $\alpha \in \mathbb{N}_0^m$, $p \in [1, \infty]$, and $\varepsilon > 0$.*

a) Let $f \in W_\alpha(U)$. If $U = \mathbb{R}^m$, we have $\partial^\alpha G_\varepsilon f = G_\varepsilon \partial^\alpha f$, and otherwise

$$(\partial^\alpha G_\varepsilon f)(x) = (G_\varepsilon \partial^\alpha f)(x) \quad \text{for } x \in U \text{ with } \varepsilon < \mathrm{d}(x, \partial U).$$

b) Let $f \in W_\alpha(U)$. Then the functions $G_\varepsilon f$ converge to f and $\partial^\alpha G_\varepsilon f$ tend to $\partial^\alpha f$ in $L^1_{\mathrm{loc}}(U)$ as $\varepsilon \to 0$. If, in addition, f and $\partial^\alpha f$ belong to $L^p_{\mathrm{loc}}(U)$ for some $p \in (1, \infty)$, then we have convergence in $L^p_{\mathrm{loc}}(U)$.

c) Let $f \in W_\beta(U)$ for some multi-indices $\beta \in \mathbb{N}_0^m$. For each null sequence (ε_j) in $\mathbb{R}_+$, we obtain a subsequence $\varepsilon_n := \varepsilon_{j_n} \to 0$ such that $G_{\varepsilon_n} f \to f$ and $\partial^\beta G_{\varepsilon_n} f \to \partial^\beta f$ a.e. on U for all these β as $n \to \infty$, where $n \geqslant N_K$ for $x \in K$, some $N_K \in \mathbb{N}$ and any compact $K \subseteq U$.

d) Let $f, g \in L^1_{\mathrm{loc}}(U)$ and $f_n \in W_\alpha(U)$ such that $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^1_{\mathrm{loc}}(U)$ as $n \to \infty$. Then f is contained in $W_\alpha(U)$ and $\partial^\alpha f = g$. If these limits exist in $L^p(U)$ (or $L^p_{\mathrm{loc}}(U)$) and for all α with $|\alpha| \leqslant k$, then f is an element of $W^{k,p}(U)$ (or $W^{k,p}_{\mathrm{loc}}(U)$). Moreover, ∂^α with domain

$$\mathrm{D}(\partial^\alpha) = \{f \in W_\alpha(U) \cap L^p(U) \,|\, \partial^\alpha f \in L^p(U)\}$$

is closed in $L^p(U)$.

Proof. a) Let $f \in W_\alpha(U)$, $\varepsilon > 0$, and $x \in U^\varepsilon$. Then the map $y \mapsto \varphi_{\varepsilon,x}(y) := k_\varepsilon(x - y)$ belongs to $C_c^\infty(U)$ since $\operatorname{supp} \varphi_{\varepsilon,x} = \overline{B}(x, \varepsilon)$. Using a corollary to Lebesgue's theorem and (3.12), we can thus deduce

$$\begin{aligned}\partial^\alpha G_\varepsilon f(x) &= \int_U \partial_x^\alpha k_\varepsilon(x - y) f(y)\,\mathrm{d}y = (-1)^{|\alpha|} \int_U (\partial^\alpha \varphi_{\varepsilon,x})(y) f(y)\,\mathrm{d}y \\ &= \int_U \varphi_{\varepsilon,x}(y) \partial^\alpha f(y)\,\mathrm{d}y = G_\varepsilon \partial^\alpha f(x).\end{aligned}$$

If $U = \mathbb{R}^m$ this argument works for all x.

b) Let $f \in W_\alpha(U)$ with $f, \partial^\alpha f \in L^p_{\mathrm{loc}}(U)$ for some $p < \infty$. Choose a compact subset $K \subseteq U$ and fix $\delta > 0$ with $K_\delta \subseteq U$. Take $\varepsilon \in (0, \delta]$. Note that the integrand of $G_\varepsilon g(x)$ is then supported in K_δ for all $x \in K$ and $g \in L^1_{\mathrm{loc}}(U)$, see (3.13). Hence, part a) and (3.17) imply the limit

$$\mathbb{1}_K \partial^\alpha G_\varepsilon f = \mathbb{1}_K G_\varepsilon \partial^\alpha f = \mathbb{1}_K G_\varepsilon (\mathbb{1}_{K_\delta} \partial^\alpha f) \longrightarrow \mathbb{1}_K \mathbb{1}_{K_\delta} \partial^\alpha f = \mathbb{1}_K \partial^\alpha f$$

in $L^p(K)$ as $\varepsilon \to 0$. So the asserted convergence in $L^p_{\mathrm{loc}}(U)$ is true.

c) Let $f \in W_\alpha(U)$. For $k \in \mathbb{N}$, we define

$$K_k = \{x \in U \,|\, \mathrm{d}(x, \partial U) \geqslant \tfrac{1}{k} \text{ and } |x|_2 \leqslant k\}$$

These sets are compact and $\bigcup_{k\in\mathbb{N}} K_k = U$. Let $\varepsilon_j \to 0$. For each $k \in \mathbb{N}$ there is a null set $N_k \subseteq K_k$ and a subsequence ν_k in ν_{k-1} (with $\nu_0(i) = i$) such that $G_{\varepsilon_{\nu_k(i)}} f(x)$ tends to $f(x)$ and $\partial^\alpha G_{\varepsilon_{\nu_k(i)}} f(x)$ to $\partial^\alpha f(x)$ for all $x \in K_k \backslash N_k$ as $i \to \infty$. We find a 'diagonal' subsequence $\varepsilon_n = \varepsilon_{j_n} \to 0$ such that $G_{\varepsilon_n} f(x) \to f(x)$ and $\partial^\alpha G_{\varepsilon_n} f(x) \to \partial^\alpha f(x)$ for $x \in U\backslash(\bigcup_{k\in\mathbb{N}} N_k)$ as $n \to \infty$ where $\varepsilon_n < \frac{1}{k}$ for $x \in K_k$ and $\bigcup_{k\in\mathbb{N}} N_k$ is a null set. By another diagonal sequence, we can achieve this for countably many $\partial^\beta f$ at the same time.

d) Let $\varphi \in C_c^\infty(U)$ and $S = \operatorname{supp}\varphi$. Since S is compact, the assumptions of assertion d) yield $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^1(S)$ as $n \to \infty$. From (3.4) on S and (3.12) we then infer

$$\begin{aligned}\int_U f\partial^\alpha\varphi\,\mathrm{d}x &= \lim_{n\to\infty}\int_U f_n\partial^\alpha\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\lim_{n\to\infty}\int_U \partial^\alpha f_n\,\varphi\,\mathrm{d}x\\ &= (-1)^{|\alpha|}\int_U g\varphi\,\mathrm{d}x;\end{aligned}$$

i.e., f belongs to $W_\alpha(U)$ and $\partial^\alpha f = g$. The remaining claims follow easily. □

In the next examples we also argue by approximation, but using a different, more explicit regularisation method.

Example 3.17. Let $U = B(0,1)$.

a) Let $m \geqslant 2$, $1 \leqslant p < m$, and $f(x) = \ln|x|_2$ for $x \in U\backslash\{0\}$. Then f belongs to $W^{1,p}(U)$ with

$$\partial_j f(x) = \frac{x_j}{|x|_2^2} =: g_j(x),$$

for $x \neq 0$ and $j \in \{1,\dots,m\}$. Moreover, $f \in L^q(U)\backslash L^\infty(U)$ for all $q \in [1,\infty)$.

Proof. Using polar coordinates and $|x_j| \leqslant r$, we obtain

$$\begin{aligned}\|f\|_q^q &= c\int_0^1 |\ln r|^q\, r^{m-1}\,\mathrm{d}r < \infty,\\ \|g_j\|_p^p &\leqslant c\int_0^1 \frac{r^p}{r^{2p}} r^{m-1}\,\mathrm{d}r = c\int_0^1 r^{m-p-1}\,\mathrm{d}r < \infty,\end{aligned}$$

since $p < m$. Hence, f is contained in $L^q(U)$ and g_j in $L^p(U)$. Define $u_n \in C^\infty(\overline{U}) \subseteq W^{1,p}(U)$ by $u_n(x) = \ln(n^{-2} + |x|_2^2)^{1/2}$ for $n \in \mathbb{N}$. Observe that $\partial_j u_n(x) = (n^{-2} + |x|_2^2)^{-1}x_j$, $u_n(x) \to f(x)$, and $\partial_j u_n(x) \to g_j(x)$ as $n \to \infty$ for all $x \in U\backslash\{0\}$. We have the pointwise bounds

$$|u_n(x)| \leqslant \begin{cases} |f(x)|, & n^{-2} + |x|_2^2 \leqslant 1,\\ \ln\sqrt{2}, & n^{-2} + |x|_2^2 > 1,\end{cases} \qquad |\partial_j u_n(x)| \leqslant g_j(x), \quad x \in U.$$

Lebesgue's theorem thus yields that $u_n \to f$ and $\partial_j u_n \to f_j$ in $L^p(U)$ as $n \to \infty$. The assertion then follows from Lemma 3.16 d). □

b) Let $p \in [1,\infty)$ and $\beta \in (1-\frac{m}{p}, 1]$. Set $u(x) = |x|_2^\beta$ and $f_j(x) = \beta x_j |x|_2^{\beta-2}$ for $x \in U\backslash\{0\}$ and $j \in \{1,\dots,m\}$. As in part a) one shows that $u \in W^{1,p}(U)$ and $\partial_j u = f_j$. (See Example 4.18 in [**FA**].) ◊

We now obtain the basic functional analytic properties of Sobolev spaces.

Proposition 3.18. *Let $1 \leqslant p \leqslant \infty$ and $k \in \mathbb{N}$. Then $W^{k,p}(U)$ is a Banach space which is isometrically isomorphic to a closed subspace of a $L^p(U)^N$ for some $N \in \mathbb{N}$. It is separable if $p < \infty$ and reflexive if $1 < p < \infty$. Moreover, $W^{k,2}(U)$ is a Hilbert space endowed with the scalar product*

$$(f|g)_{k,2} = \sum_{|\alpha|\leqslant k} \int_U \partial^\alpha f \, \overline{\partial^\alpha g} \, \mathrm{d}x.$$

Proof. Let $(f_n)_n$ be a Cauchy sequence in $W^{k,p}(U)$. Then $(\partial^\alpha f_n)_n$ is a Cauchy sequence in $L^p(U)$ for every $\alpha \in \mathbb{N}_0^m$ with $|\alpha| \leqslant k$ and thus $\partial^\alpha f_n \to g_\alpha$ in $L^p(U)$ for some $g_\alpha \in L^p(U)$ as $n \to \infty$, where we set $f := g_0$. Lemma 3.16 d) implies that f belongs to $W^{k,p}(U)$ with $\partial^\alpha f = g_\alpha$ for $|\alpha| \leqslant k$. Hence, f_n tends to f for $\|\cdot\|_{k,p}$ and $W^{k,p}(U)$ is a Banach space. We then deduce from Remark 3.14 e) that $W^{k,p}(U)$ is isometrically isomorphic to a subspace of $L^p(U)^N$ which is closed by Remark 2.11 in [**FA**]. The remaining assertion now follow by isomorphy from known results of functional analysis, cf. Proposition 4.19 of [**FA**]. □

We next establish product and chain rules by approximation, using core arguments of the area. One can derive many variants by modifications of the proofs. Typically such identities can be shown by approximation if the functions involved belong to the desired L^p-spaces because of Hölder's inequality, for instance, though the case $p = \infty$ may cause additional trouble.

Proposition 3.19. *a) Let $f, g \in W^{1,1}_{\mathrm{loc}}(U) \cap L^\infty(U)$. Then, fg belongs to $W^{1,1}_{\mathrm{loc}}(U) \cap L^\infty(U)$ and has the derivatives*

$$\partial_j(fg) = \partial_j f \, g + f \partial_j g, \qquad j \in \{1, \ldots, m\}. \tag{3.18}$$

b) Let $1 \leqslant p \leqslant \infty$, $f \in W^{1,p}(U)$ and $g \in W^{1,p'}(U)$. Then, fg is contained in $W^{1,1}(U)$ and satisfies (3.18).

Proof. We approximate first g and then f to make the arguments more transparent.

1) Let $f, g \in W^{1,1}_{\mathrm{loc}}(U)$. Set $f_n = G_{\varepsilon_n} f \in C^\infty(U)$ and $g_n = G_{\varepsilon_n} g \in C^\infty(U)$ with $\varepsilon_n \to 0$ as in Lemma 3.16 c). Fix $k \in \mathbb{N}$ and take $\varphi \in C_c^\infty(U)$ and $j \in \{1, \ldots, m\}$. Choose an open and bounded set V such that $\operatorname{supp} \varphi \subseteq V \subseteq \overline{V} \subseteq U$. Since $f_n \to f$ and $\partial_j f_n \to \partial_j f$ on $L^1(\overline{V})$ by Lemma 3.16 a), formulas (3.4) and (3.11) yield

$$\begin{aligned}\int_U f g_k \partial_j \varphi \, \mathrm{d}x &= \lim_{n\to\infty} \int_V f_n g_k \partial_j \varphi \, \mathrm{d}x = -\lim_{n\to\infty} \int_V \big(\partial_j f_n \, g_k + f_n \partial_j g_k\big) \varphi \, \mathrm{d}x \\ &= -\int_U \big(\partial_j f \, g_k + f \partial_j g_k\big) \varphi \, \mathrm{d}x,\end{aligned}$$

so that $f g_k \in W^{1,1}_{\mathrm{loc}}(U)$ and $\partial_j(f g_k) = \partial_j f \, g_k + f \partial_j g_k$.

2) Let $f, g \in W^{1,1}_{\mathrm{loc}}(U) \cap L^\infty(U)$ and g_k as in 1). Note that $g_k \to g$ and $\partial_j g_k \to \partial_j g$ in $L^1_{\mathrm{loc}}(U)$ as $k \to \infty$. Since f is bounded, we obtain

$$\int_U f g \partial_j \varphi \, \mathrm{d}x = \lim_{k\to\infty} \int_U f g_k \partial_j \varphi \, \mathrm{d}x = \lim_{k\to\infty} -\Big[\int_U \partial_j f \, g_k \varphi \, \mathrm{d}x + \int_U f \partial_j g_k \, \varphi \, \mathrm{d}x\Big]$$

using step 1). The last integral converges to $\int_U f\partial_j g\,\varphi\,\mathrm{d}x$, again because of $f \in L^\infty(U)$. For the penultimate integral we use that $g_k \to g$ a.e. on $\operatorname{supp}\varphi$ by Lemma 3.16 c) and that $\|g_k\|_\infty \leqslant \|g\|_\infty$ by (3.16). Lebesgue's theorem (with the majorant $|\partial_j f|\,\|g\|_\infty\|\varphi\|_\infty\mathbb{1}_{\operatorname{supp}\varphi}$) then implies

$$\int_U fg\partial_j\varphi\,\mathrm{d}x = -\int_U \big(\partial_j f\,g + f\partial_j g\big)\varphi\,\mathrm{d}x.$$

The map $\partial_j f\,g + f\partial_j g$ is contained in $L^1_{\mathrm{loc}}(U)$ by our assumptions. Assertion a) is shown.

3) Let $f \in W^{1,p}(U)$ and $g \in W^{1,p'}(U)$. If $p \in (1,\infty]$ we show (3.18) as in step 2), using (3.4) and that g_k, $\partial_j g_k$ converge in $L^{p'}_{\mathrm{loc}}(U)$ by Lemma 3.16 b). If $p = 1$, we replace the roles of f and g. Hölder's inequality and (3.18) finally yield that fg and $\partial_j(fg)$ are contained in $L^1(U)$. $\square$

Proposition 3.20. *Let $1 \leqslant p \leqslant \infty$, $j \in \{1,\dots,m\}$, and $f \in W^{1,p}_{\mathrm{loc}}(U)$.*

a) Let f be real-valued and $h \in C^1(\mathbb{R})$ with $h' \in C_b(\mathbb{R})$. Then $h\circ f$ belongs to $W^{1,p}_{\mathrm{loc}}(U)$ with derivatives

$$\partial_j(h\circ f) = (h'\circ f)\partial_j f.$$

b) Let $V \subseteq \mathbb{R}^m$ be open and $\Phi = (\Phi_1,\dots,\Phi_m) : V \to U$ be a diffeomorphism such that Φ' and $(\Phi^{-1})'$ are bounded. Then $f\circ\Phi$ belongs to $W^{1,p}_{\mathrm{loc}}(V)$ with derivatives

$$\partial_j(f\circ\Phi) = \sum_{k=1}^m ((\partial_k f)\circ\Phi)\partial_j\Phi_k.$$

In both results we can replace $W^{1,p}_{\mathrm{loc}}(U)$ by $W^{1,p}(U)$, where in part a) we then also assume $h(0) = 0$ if $\lambda(U) = \infty$ and $p < \infty$.

Proof.[2] By Lemma 3.16, there are maps$f_n \in C^\infty(U)$ such that $f_n \to f$ and $\partial_j f_n \to \partial_j f$ in $L^1_{\mathrm{loc}}(U)$ and a.e. as $n \to \infty$.

a) The function $h\circ f$ belongs to $L^p_{\mathrm{loc}}(U)$ since

$$|h(f(x))| \leqslant |h(f(x)) - h(0)| + |h(0)| \leqslant \|h'\|_\infty |f(x)| + |h(0)|$$

for all $x \in U$. It is contained in $L^p(U)$ if $f \in L^p(U)$ and if $h(0) = 0$ in the case that $\lambda(U) = \infty$ and $p \neq \infty$. Note that $h\circ f_n$ belongs to $C^1(U)$ and $\partial_j(h\circ f_n) = (h'\circ f_n)\partial_j f_n$. Let $K \subseteq U$ be compact. We compute

$$\int_K |h(f_n(x)) - h(f(x))|\,\mathrm{d}x \leqslant \|h'\|_\infty \int_K |f_n(x) - f(x)|\,\mathrm{d}x \longrightarrow 0,$$

$$\int_K |h'(f_n(x))\partial_j f_n(x) - h'(f(x))\partial_j f(x)|\,\mathrm{d}x$$

$$\leqslant \|h'\|_\infty \int_K |\partial_j f_n(x) - \partial_j f(x)|\,\mathrm{d}x + \int_K |h'(f_n(x)) - h'(f(x))||\partial_j f(x)|\,\mathrm{d}x \to 0$$

as $n \to \infty$ where we also used Lebesgue's theorem and the majorant $2\|h'\|_\infty\mathbb{1}_K|\partial_j f|$ in the last integral. Since $(h'\circ f)\partial_j f$ belongs to $L^p_{\mathrm{loc}}(U)$, Lemma 3.16 d) yields assertion a). If $f \in W^{1,p}(U)$, then $(h'\circ f)\partial_j f \in L^p(U)$ and so $h\circ f$ is an element $W^{1,p}(U)$.

[2] Not shown in the lectures.

b) Let $B = \Phi^{-1}(A)$ for an open set $A \subseteq U$ and $g \in L^p(A)$. For $p < \infty$, from the transformation rule we deduce

$$\int_B |g(\Phi(x))|^p \,\mathrm{d}x = \int_A |g(y)|^p \,|\det[(\Phi^{-1})'(y)]|\,\mathrm{d}y \leqslant c\|g\|^p_{L^p(A)}.$$

An analogous estimate is also true for $p = \infty$. Hence, $f \circ \Phi$ is contained in $L^p_{\mathrm{loc}}(V)$ and $f_n \circ \Phi$ converges to $f \circ \Phi$ in $L^1_{\mathrm{loc}}(V)$. We further have

$$\partial_j(f_n \circ \Phi) = \sum_{k=1}^m ((\partial_k f_n) \circ \Phi)\, \partial_j \Phi_k.$$

The above estimate also implies that $(\partial_k f_n) \circ \Phi$ tends to $(\partial_k f) \circ \Phi$ in $L^1_{\mathrm{loc}}(V)$ as $n \to \infty$, where this map is an element of $L^p_{\mathrm{loc}}(V)$. Since $\partial_j \Phi_k$ is bounded, Lemma 3.16 d) now yields that $f \circ \Phi$ belongs to $W^{1,p}_{\mathrm{loc}}(U)$ and has the asserted derivative. If $f \in L^p(U)$, we can replace throughout $L^p_{\mathrm{loc}}(V)$ by $L^p(V)$. □

We extend the chain rule for $h \circ f$ to certain Lipschitz functions, using an adapted regularization of h.

Corollary 3.21. *Let $f \in W^{1,1}_{\mathrm{loc}}(U)$ be real-valued. Then the maps f_+, f_-, and $|f|$ belong to $W^{1,1}_{\mathrm{loc}}(U)$ with*

$$\partial_j f_\pm = \pm \mathbb{1}_{\{f \gtrless 0\}} \partial_j f \quad \text{and} \quad \partial_j |f| = \big(\mathbb{1}_{\{f>0\}} - \mathbb{1}_{\{f<0\}}\big)\, \partial_j f$$

for all $j \in \{1, \dots, m\}$. Here one can replace $W^{1,1}_{\mathrm{loc}}$ by $W^{1,p}$ for all $1 \leqslant p \leqslant \infty$.

Proof.[3] We employ the map $h_\varepsilon \in C^1(\mathbb{R})$ given by $h_\varepsilon(t) := \sqrt{t^2 + \varepsilon^2} - \varepsilon \leqslant t$ for $t \geqslant 0$ and $h_\varepsilon(t) := 0$ for $t < 0$, where $\varepsilon > 0$. Observe that $\|h'_\varepsilon\|_\infty = 1$ for $\varepsilon > 0$ and that $h_\varepsilon(t) \to \mathbb{1}_{\mathbb{R}_+}(t)t$ for $t \in \mathbb{R}$ as $\varepsilon \to 0$. Proposition 3.20 shows that $h_\varepsilon \circ f \in W^{1,1}_{\mathrm{loc}}(U)$ and

$$\int_U h_\varepsilon(f) \partial_j \varphi \,\mathrm{d}x = -\int_U h'_\varepsilon(f) \partial_j f\, \varphi \,\mathrm{d}x = -\int_{\{f>0\}} \frac{f}{\sqrt{f^2 + \varepsilon^2}} \partial_j f\, \varphi \,\mathrm{d}x$$

for each $\varphi \in C^\infty_c(U)$. Using the majorants $\|\partial_j \varphi\|_\infty \mathbb{1}_S |f|$ and $\|\varphi\|_\infty \mathbb{1}_S |\partial_j f|$ with $S = \operatorname{supp} \varphi$, we deduce from Lebesgue's convergence theorem that

$$\int_U f_+ \partial_j \varphi \,\mathrm{d}x = -\int_{\{f>0\}} \frac{f}{|f|} \partial_j f\, \varphi \,\mathrm{d}x = -\int_U \mathbb{1}_{\{f>0\}} \partial_j f\, \varphi \,\mathrm{d}x.$$

There thus exists $\partial_j f_+ = \mathbb{1}_{\{f>0\}} \partial_j f \in L^1_{\mathrm{loc}}(U)$. Clearly, $\partial_j f_+$ belongs to $L^p(U)$ if $f \in W^{1,p}(U)$. The other claims follow from $f_- = (-f)_+$ and $|f| = f_+ + f_-$. □

We discuss three special cases, namely $m = 1$, $p = \infty$, and $p = 2$ for $U = \mathbb{R}^m$.

Theorem 3.22. *Let $J \subseteq \mathbb{R}$ be an open interval, $1 \leqslant p < \infty$, and $f \in L^p_{\mathrm{loc}}(J)$. Then f belongs to $W^{1,p}_{\mathrm{loc}}(J)$ if and only if there is a map $g \in L^p_{\mathrm{loc}}(J)$ and a continuous representative of f which satisfy*

$$f(t) = f(s) + \int_s^t g(\tau)\,\mathrm{d}\tau \tag{3.19}$$

for all $s, t \in J$. In this case, we have $g = \partial f$ a.e.. Let also $\partial f \in L^1(J)$. Then f extends to $\overline{J}$ continuously and (3.19) is true for $t, s \in \overline{J}$.

[3] Not shown in the lectures.

PROOF. 1) Let $f \in W^{1,p}_{\mathrm{loc}}(J)$. Take the functions $f_n = G_{\varepsilon_n} f \in C^\infty(J)$ from Lemma 3.16 c). Then for a.e. $t \in J$ and for a.e. $t_0 \in J$ we have

$$f(t) - f(t_0) = \lim_{n\to\infty} (f_n(t) - f_n(t_0)) = \lim_{n\to\infty} \int_{t_0}^t f_n'(\tau)\,\mathrm{d}\tau = \int_{t_0}^t \partial f(\tau)\,\mathrm{d}\tau.$$

Fix one t_0. Since $t \mapsto \int_{t_0}^t \partial f(\tau)\,\mathrm{d}\tau$ is continuous, we obtain a continuous representative of f which fulfills (3.19) for $t \in J$, $s = t_0$ and $g = \partial f$. Subtracting the above equation with t_0 and $s \in J$ instead of t, we deduce (3.19) with $g = \partial f$ for all $t, s \in J$. The addendum with $\partial f \in L^1(J)$ follows easily.

2) Let (3.19) be satisfied by some $f, g \in L^p_{\mathrm{loc}}(J)$. As in the proof of Lemma 3.16 b) we find maps $g_n \in C^\infty(J)$ with $g_n \to g$ in $L^p_{\mathrm{loc}}(J)$ as $n \to \infty$. For every $s \in J$ and $n \in \mathbb{N}$, the function $J \ni t \mapsto f_n(t) := f(s) + \int_s^t g_n(\tau)\,\mathrm{d}\tau$ belongs to $C^\infty(J)$ with $f_n' = g_n$. For $[a,b] \subseteq J$ with $s \in [a,b]$, we estimate

$$\begin{aligned}
\|f_n - f\|^p_{L^p([a,b])} &= \int_a^b \left| \int_s^t (g_n(\tau) - g(\tau))\,\mathrm{d}\tau \right|^p \mathrm{d}t \\
&\leqslant \int_a^b |t-s|^{p/p'} \left(\int_a^b |g_n(\tau) - g(\tau)|^p \,\mathrm{d}\tau \right)^{p/p} \mathrm{d}t \\
&\leqslant (b-a)^{1+p/p'} \|g_n - g\|^p_{L^p([a,b])},
\end{aligned}$$

using (3.19) and Hölder's inequality. Hence, f_n tends to f in $L^p_{\mathrm{loc}}(J)$ as $n \to \infty$. Lemma 3.16 d) then yields $f \in W^{1,p}_{\mathrm{loc}}(J)$ and $\partial f = g$. □

We discuss the relationship of the weak and pointwise derivatives if $m = 1$.

REMARK 3.23. a) Let $J = (a,b)$ for some $a < b$ in $\mathbb{R}$ and $f : J \to \mathbb{F}$. We then have $f \in W^{1,1}(J)$ if and only if f is *absolutely continuous*; i.e., for all $\varepsilon > 0$ there is a number $\delta > 0$ such that for all points $a < \alpha_1 < \beta_1 < \alpha_2 < \cdots < \alpha_n < \beta_n < b$ with $n \in \mathbb{N}$ and $\sum_{j=1}^n (\beta_j - \alpha_j) \leqslant \delta$ we obtain

$$\sum\nolimits_{j=1}^n |f(\beta_j) - f(\alpha_j)| \leqslant \varepsilon.$$

(Note that a Lipschitz continuous function is absolutely continuous and that an absolutely continuous function is uniformly continuous.) In this case, f is differentiable for a.e. $t \in J$ and the pointwise derivative f' is equal a.e. to the weak derivative $\partial f \in L^1(J)$.

PROOF.[4] Let $f \in W^{1,1}(J)$. Then formula (3.19) yields

$$\sum_{j=1}^n |f(\beta_j) - f(\alpha_j)| = \sum_{j=1}^n \left| \int_{\alpha_j}^{\beta_j} \partial f(\tau)\,\mathrm{d}\tau \right| \leqslant \int_{\bigcup_{j=1}^n (\alpha_j,\beta_j)} |\partial f(\tau)|\,\mathrm{d}\tau =: S,$$

where $S \to 0$ as $\lambda(\bigcup_{j=1}^n (\alpha_j, \beta_j)) \to 0$.

The converse implication and the last assertion is shown in Theorem 7.20 of [**Ru1**] (combined with our Theorem 3.22). □

b) There is a continuous increasing map $f : [0,1] \to \mathbb{R}$ with $f(0) = 0$ and $f(1) = 1$ such that $f'(t) = 0$ exists for a.e. $t \in [0,1]$. This function satisfies

$$1 = f(1) \neq f(0) + \int_0^1 f'(\tau)\,\mathrm{d}\tau = 0,$$

[4] The proof was omitted in the lectures.

and so f violates (3.19), does not belong to $W^{1,1}((0,1))$, and is not absolutely continuous. (See §7.16 in [**Ru1**].)

c) Let $f \in W^{1,1}_{\mathrm{loc}}(J)$ and ∂f be continuous. Then f is continuously differentiable, since then

$$\frac{f(t)-f(s)}{t-s} = \frac{1}{t-s}\int_s^t \partial f(\tau)\,\mathrm{d}\tau \;\longrightarrow\; \partial f(s)$$

as $t \to s$ in J, thanks to Theorem 3.22. $\Diamond$

We now characterize $W^{1,\infty}(U)$ at least for convex U, other domains are treated in the Corollary 3.29.

Proposition 3.24. *a) The space $C^{1-}(U) = \{f \in C_b(U) \,|\, f \text{ Lipschitz}\}$ is embedded into $W^{1,\infty}(U)$, and one has*

$$\|f\|_{1,\infty} \leqslant \|f\|_\infty + [f]_{\mathrm{Lip}} = \|f\|_{C^{1-}}$$

for $f \in C^{1-}(U)$, where $[f]_{\mathrm{Lip}}$ is the Lipschitz constant of f.

b) Let U be convex. Then $W^{1,\infty}(U)$ is embedded into $C^{1-}(U)$ and one has

$$\|f\|_{C^{1-}} \leqslant \sqrt{m}\,\|f\|_{1,\infty}, \qquad f \in W^{1,\infty}(U).$$

Proof.[5] a) Let $f \in C^{1-}(U)$. Take $\varphi \in C_c^\infty(U)$ with support S, $j \in \{1,\dots,m\}$, and $\delta > 0$ with $S_\delta \subseteq U$. For $\varepsilon \in (0,\delta]$ the difference quotient $\frac{1}{\varepsilon}(\varphi(x+\varepsilon e_j) - \varphi(x))$ converges uniformly on $\operatorname{supp}\varphi$ as $\varepsilon \to 0$, and hence

$$\begin{aligned}\left|\int_U f\partial_j\varphi\,\mathrm{d}x\right| &= \lim_{\varepsilon\to 0}\left|\int_S f(x)\frac{1}{\varepsilon}(\varphi(x+\varepsilon e_j) - \varphi(x))\,\mathrm{d}x\right| \\ &\leqslant \lim_{\varepsilon\to 0}\int_S \frac{1}{\varepsilon}|f(y-\varepsilon e_j) - f(y)|\,|\varphi(y)|\,\mathrm{d}y \\ &\leqslant [f]_{\mathrm{Lip}}\|\varphi\|_1.\end{aligned}$$

Since $C_c^\infty(U)$ is dense in $L^1(U)$, the map $\varphi \mapsto -\int_U f\partial_j\varphi\,\mathrm{d}x$ has a continuous linear extension $F_j : L^1(U) \to \mathbb{F}$. By Theorem 5.4 of [**FA**] there thus exists a function g_j in $L^\infty(U) = L^1(U)^\star$ with $\|g_j\|_\infty = \|F_j\|_{(L^1)^\star} \leqslant [f]_{\mathrm{Lip}}$ such that

$$-\int_U f\partial_j\varphi\,\mathrm{d}x = F_j(\varphi) = \int_U g_j\varphi\,\mathrm{d}x$$

for all $\varphi \in C_c^\infty(U)$. This means that f has the weak derivative $\partial_j f = g_j \in L^\infty(U)$. As a result, f belongs to $W^{1,\infty}(U)$ and $\|f\|_{1,\infty} \leqslant \|f\|_\infty + [f]_{\mathrm{Lip}}$.

b) Let $f \in W^{1,\infty}(U)$ and U be convex. Take $\varepsilon_n \to 0$ from Lemma 3.16 c) and fix a compact $K \subseteq U$. For sufficiently large $n \in \mathbb{N}$, Lemma 3.16 and (3.16) yield

$$|\partial_j G_{\varepsilon_n} f(z)| = |G_{\varepsilon_n}\partial_j f(z)| \leqslant \|\partial_j f\|_\infty \leqslant \|f\|_{1,\infty},$$

for all $j \in \{1,\dots,m\}$ and $z \in K$. Using that $G_{\varepsilon_n} f(x) \to f(x)$ as $n \to \infty$ for all $x \in U\backslash N$ and a null set N, for all $x, y \in U\backslash N$ we thus estimate

$$\begin{aligned}|f(x) - f(y)| &= \lim_{n\to\infty}|G_{\varepsilon_n} f(x) - G_{\varepsilon_n} f(y)| \\ &= \lim_{n\to\infty}\left|\int_0^1 \nabla G_{\varepsilon_n} f(y+\tau(x-y))\cdot(x-y)\,\mathrm{d}\tau\right|\end{aligned}$$

[5] Not shown in the lectures.

$$\leqslant \sqrt{m}\, \max_j \|\partial_j f\|_\infty \, |x-y|_2 \, .$$

Hence, f has a representative with Lipschitz constant $\sqrt{m}\, \max_j \|\partial_j f\|_\infty$. □

In the spirit of Remark 3.23, we mention Rademacher's theorem which says that a Lipschitz function f is differentiable for a.e. $x \in U$ and that the weak derivative $\partial_j f$ coincides with the pointwise one. Actually, it is enough to assume $f \in W^{1,p}_{\mathrm{loc}}(U)$ for some $p \in (m, \infty]$, see Theorem 5.8.5 in [**Ev**].

The next important result describes the space $W^{k,2}(\mathbb{R}^m)$ via the Fourier transform $\mathcal{F}$ in a very convenient way and complements Theorem 3.11. Recall that $\mathcal{F}$ is a bijective isometry on $L^2(\mathbb{R}^m)$ by this theorem.

THEOREM 3.25. *Let $\mathbb{F} = \mathbb{C}$, $k \in \mathbb{N}$, and $\alpha \in \mathbb{N}_0^m$ with $|\alpha| \leqslant k$. We then have*

$$W^{k,2}(\mathbb{R}^m) = \big\{u \in L^2(\mathbb{R}^m) \,\big|\, |\xi|_2^k \hat{u} \in L^2(\mathbb{R}^m)\big\} =: H^k,$$

and the norm of $W^{k,2}(\mathbb{R}^m)$ is equivalent to $(\|u\|_2^2 + \||\xi|_2^k \hat{u}\|_2^2)^{\frac{1}{2}}$. For $u \in W^{k,2}(\mathbb{R}^m)$ it further holds

$$\mathcal{F}(\partial^\alpha u) = \mathrm{i}^{|\alpha|}\, \xi^\alpha \hat{u}. \tag{3.20}$$

PROOF. 1) We first show the asserted norm equivalence for Schwartz functions $u \in \mathcal{S}_m$. They satisfy (3.20) by Lemma 3.7. We thus obtain

$$\|u\|_{k,2}^2 = \sum_{|\alpha|\leqslant k} \|\mathcal{F}\partial^\alpha u\|_2^2 = \sum_{|\alpha|\leqslant k} \|\xi^\alpha \hat{u}\|_2^2 = \int_{\mathbb{R}^m} \sum_{|\alpha|\leqslant k} |\xi^\alpha|^2 |\hat{u}|_2^2 \,\mathrm{d}\xi$$

$$\begin{cases} \leqslant c_k(\|u\|_2^2 + \||\xi|_2^k \hat{u}\|_2^2), \\ \geqslant c_k'(\|u\|_2^2 + \||\xi|_2^k \hat{u}\|_2^2) \end{cases} \tag{3.21}$$

for some $c_k, c_k' > 0$ and all $u \in \mathcal{S}_m$, using (3.6) for '$\leqslant$' and $\alpha = k e_j$ for '$\geqslant$'.

2) Let $u \in W^{k,2}(\mathbb{R}^m)$. By an extension of Proposition 4.13 of [**FA**] (see Theorem 3.27 below), there are $u_n \in \mathcal{S}_m$ which converge to u in $W^{k,2}(\mathbb{R}^m)$ as $n \to \infty$. Since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^m)$, the functions $\widehat{u_n}$ tend to $\hat{u}$ in $L^2(\mathbb{R}^m)$ and (possibly after passing to a subsequence) pointwise a.e., as $n \to \infty$. Hence, the products $|\xi|_2^k \widehat{u_n}$ converge pointwise a.e. to $|\xi|_2^k \hat{u}$. On the other hand, equation (3.21) yields that the sequence $(|\xi|_2^k \widehat{u_n})_n$ is Cauchy in $L^2(\mathbb{R}^m)$, and thus it has the limit $|\xi|_2^k \hat{u}$ in $L^2(\mathbb{R}^m)$. We conclude that $W^{k,2}(\mathbb{R}^m)$ is contained in H^k and that (3.21) is true for $u \in W^{k,2}(\mathbb{R}^m)$.

3) Conversely, take $u \in H^k$. Let $\varphi \in C_c^\infty(\mathbb{R}^m)$ and $|\alpha| \leqslant k$. Then $\xi^\alpha \hat{u}$ is an element of $L^2(\mathbb{R}^m)$. From Theorem 3.11 and Lemma 3.7 we deduce

$$\int_{\mathbb{R}^m} u\, \partial^\alpha \varphi \,\mathrm{d}x = (u|\partial^\alpha \overline{\varphi}) = (\mathcal{F}u|\mathcal{F}\partial^\alpha \overline{\varphi}) = (\hat{u}|\mathrm{i}^{|\alpha|}\xi^\alpha \widehat{\overline{\varphi}}) = ((-\mathrm{i})^{|\alpha|}\xi^\alpha \hat{u}|\mathcal{F}\overline{\varphi})$$

$$= (\mathcal{F}^{-1}((-\mathrm{i})^{|\alpha|}\xi^\alpha \hat{u})|\overline{\varphi}) = (-1)^{|\alpha|} \int_{\mathbb{R}^m} \varphi\, \mathcal{F}^{-1}(\mathrm{i}^{|\alpha|}\xi^\alpha \hat{u}) \,\mathrm{d}x.$$

Therefore u belongs to $W_\alpha(\mathbb{R}^m)$ and $\partial^\alpha u = \mathcal{F}^{-1}(\mathrm{i}^{|\alpha|}\xi^\alpha \hat{u}) \in L^2(\mathbb{R}^m)$; i.e., u is contained in $W^{k,2}(\mathbb{R}^m)$. Hence, $W^{k,2}(\mathbb{R}^m) = H^k$ and their norms are equivalent by (3.21). Applying $\mathcal{F}$ to the above equation for $\partial^\alpha u$, we also derive (3.20) for all $u \in W^{k,2}(\mathbb{R}^m)$. □

We use the above characterization and $\mathcal{F}$ to solve a basic partial differential equation on $\mathbb{R}^m$.

Example 3.26. We consider the diffusion equation

$$\begin{aligned} \partial_t u(t,x) &= \Delta u(t,x), \qquad t>0,\ x\in\mathbb{R}^m,\\ u(0,x) &= u_0(x), \qquad x\in\mathbb{R}^m, \end{aligned} \tag{3.22}$$

for a given initial value $u_0 \in L^2(\mathbb{R}^m)$. We find a unique *solution* of (3.22); i.e., a map u in $C(\mathbb{R}_{\geqslant 0}, L^2(\mathbb{R}^m)) \cap C^1(\mathbb{R}_+, L^2(\mathbb{R}^m)) \cap C(\mathbb{R}_+, W^{2,2}(\mathbb{R}^m))$ which satisfies (3.22) as equations in $L^2(\mathbb{R}^m)$.

To that purpose, we first assume that we have such a solution u. We set $\hat{u}(t) = \mathcal{F}u(t)$ for all $t \geqslant 0$ which defines a function $\hat{u} \in C(\mathbb{R}_{\geqslant 0}, L^2(\mathbb{R}^m))$ since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^m)$. We further compute

$$\mathcal{F}u'(t) = \mathcal{F}\lim_{h\to 0}\tfrac{1}{h}(u(t+h)-u(t)) = \lim_{h\to 0}\tfrac{1}{h}(\widehat{u}(t+h)-\widehat{u}(t))$$

for all $t>0$, so that $\widehat{u} \in C^1(\mathbb{R}_+, L^2(\mathbb{R}^m))$ and $\partial_t\widehat{u} = \mathcal{F}\partial_t u$. Applying $\mathcal{F}$ to (3.22), we then deduce from (3.20) the equations

$$\partial_t \hat{u}(t) = \mathcal{F}\partial_t u(t) = \mathcal{F}\Delta u(t) = -|\xi|_2^2\hat{u}(t), \qquad \hat{u}(0) = \widehat{u_0}. \tag{3.23}$$

If we insert into $\hat{u}(t)$ the arguments $\xi \in \mathbb{R}^m$, for each ξ we obtain the ordinary differential equation $\varphi_\xi'(t) = -|\xi|_2^2\varphi_\xi(t)$, $t \geqslant 0$, with initial value $\widehat{u_0}(\xi)$, which is solved by $\varphi_\xi(t) = \mathrm{e}^{-t|\xi|_2^2}\widehat{u_0}(\xi)$. We thus define

$$u(t) = \mathcal{F}^{-1}(m_t\widehat{u_0}) \quad \text{ with } \quad m_t(\xi) = \mathrm{e}^{-t|\xi|_2^2} \tag{3.24}$$

for $t>0$ and $\xi \in \mathbb{R}^m$. Theorem 3.11 and Example 3.4 yield $\|u(t)\|_2 \leqslant \|u_0\|_2$ for all $t \geqslant 0$ and

$$\begin{aligned} u(t) &= \mathcal{F}^{-1}(D_{\sqrt{2t}}\gamma\,\widehat{u_0}) = (2\pi)^{-\frac{m}{2}}(\mathcal{F}^{-1}(D_{\sqrt{2t}}\gamma)) * u_0\\ &= (2\pi)^{-\frac{m}{2}}(2t)^{-\frac{m}{2}}(D_{1/\sqrt{2t}}\gamma) * u_0, \end{aligned} \tag{3.25}$$
$$u(t,x) = (u(t))(x) = (4\pi t)^{-\frac{m}{2}}\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|x-y|_2^2}{4t}}u_0(y)\,\mathrm{d}y, \quad t>0,\ x\in\mathbb{R}^m.$$

Since $|\xi|_2^k\hat{u}(t) = |\xi|_2^k m_t\widehat{u_0} \in L^2(\mathbb{R}^m)$, Theorem 3.25 implies that $u(t)$ belongs to $W^{k,2}(\mathbb{R}^m)$ for all $k \in \mathbb{N}$ and $t>0$. From (3.20) we then infer

$$\mathcal{F}\Delta u(t) = -|\xi|_2^2\hat{u}(t) = -|\xi|_2^2 m_t\widehat{u_0}, \qquad \Delta u(t) = -\mathcal{F}^{-1}(|\xi|_2^2 m_t\widehat{u_0}).$$

Let $v(t) = \mathcal{F}u(t) = m_t\widehat{u_0}$ for $t>0$. Clearly, $\frac{1}{h}(v(t+h)-v(t))$ tends pointwise to $-|\xi|_2^2 m_t\widehat{u_0}$ as $h\to 0$. Moreover, $|\frac{1}{h}(v(t+h)-v(t))|$ is bounded by $|\xi|_2^2 m_{t/2}\widehat{u_0} \in L^2(\mathbb{R}^m)$ if $|h| \leqslant t/2$. Dominated convergence then implies that v has the derivative $v'(t) = -|\xi|_2^2 m_t\widehat{u_0}$ in $L^2(\mathbb{R}^m)$ for $t>0$. Similarly one sees the continuity of $t \mapsto |\xi|_2^2 m_t\widehat{u_0}$ from $\mathbb{R}_+$ to $L^2(\mathbb{R}^m)$. Theorem 3.25 and the continuity of $\mathcal{F}^{-1}$ on $L^2(\mathbb{R}^m)$ thus yield that u belongs to $C^1(\mathbb{R}_+, L^2(\mathbb{R}^m)) \cap C(\mathbb{R}_+, W^{2,2}(\mathbb{R}^m))$ and satisfies

$$u'(t) = -\mathcal{F}^{-1}(|\xi|_2^2 m_t\widehat{u_0}) = \Delta u(t)$$

for $t>0$. Finally, $m_t\widehat{u_0}$ tends to $\widehat{u_0}$ in $L^2(\mathbb{R}^m)$ as $t\to 0$ by Lebesgue's theorem with majorant $|\widehat{u_0}|$. Hence, u is also contained in $C(\mathbb{R}_{\geqslant 0}, L^2(\mathbb{R}^m))$ with $u(0) = u_0$. In view of (3.23) one sees that every solution of (3.22) is given by (3.24) so that solutions are unique. $\Diamond$

3.3. Main results on Sobolev spaces

In this section we discuss main theorems on Sobolev spaces, which are used throughout analysis and other parts of mathematics. Many of the proofs are rather long and technical. They are mostly omitted since they do not fit to this lecture. (The proofs are presented in the appendix Section 3.5, though often in special cases only.)

These arguments often involve density and extensions arguments, which mostly require assumptions on the boundary ∂U. To this end, we define *locally Lipschitz,* resp. C^k*, boundaries* (or domains), where $k \in \mathbb{N}$: For each $y \in \partial U$, we assume that there are open sets $U'_y \subseteq \mathbb{R}^{m-1}$, an open interval J, and a bounded Lipschitz function $h : U'_y \to J$ (resp. $h \in C_b^k(U'_y, J)$) such that $y \in U_y := U'_y \times J$ and we have, possibly after rotation,

$$\begin{aligned} U \cap U_y &= \{x = (x', x_m) \in U_y \,|\, x_m > h(x')\}, \\ \partial U \cap U_y &= \{x = (x', x_m) \in U_y \,|\, x_m = h(x')\}, \end{aligned}$$

If one can cover ∂U by finitely many such U_y, we call ∂U a *Lipschitz- (or* C^{1-}*-)*, resp. C^k*-, boundary*, and we write $\partial U \in C^{1-}$, resp. $\partial U \in C^k$.[6]

A compact locally Lipschitz boundary is Lipschitz. Unbounded examples are the 'upper' halfspace $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m \,|\, x_m > 0\}$ or, more generally, the bent upper halfspace $U = \{(x', x_m) \in \mathbb{R}^m \,|\, x_m > h(x')\}$ for a map $h \in C^{1-}(\mathbb{R}^{m-1}, \mathbb{R})$. Bounded examples are polyhedra, cylinders, cones, or balls. We note that a local Lipschitz domain is locally given by the chart $\psi(x) = (x', x_m - h(x'))$ with $\psi(U \cap U_y) \subseteq \mathbb{R}^m_+$ and $\psi(\partial U \cap U_y) \subseteq \mathbb{R}^{m-1} \times \{0\}$. For C^1-boundaries this description is equivalent to our definition, but not for Lipschitz domains, in general. (Compare Analysis 3.)

We can now state a crucial density theorem, which can be modified to cover $p = \infty$, too, cf. the exercises. We write $C_c^\infty(\overline{U})$ for the set of restrictions $f\restriction_{\overline{U}}$ of $f \in C_c^\infty(\mathbb{R}^m)$. These functions have bounded support and they and their derivatives extend continuously to ∂U.

THEOREM 3.27. *Let $k \in \mathbb{N}$ and $p \in [1, \infty)$. Then the set $C^\infty(U) \cap W^{k,p}(U)$ is dense in $W^{k,p}(U)$. Moreover, $C_c^\infty(\mathbb{R}^m)$ is dense in $W^{k,p}(\mathbb{R}^m)$. Let also $\partial U \in C^{1-}$. Then $C_c^\infty(\overline{U})$ is dense in $W^{k,p}(U)$.*

The second result is shown in Theorem 4.21 of [**FA**] for $k = 1$. General k can be handled similarly, cf. Corollary 3.23 in [**AF**]. A refinement of this argument yields the first result, see Theorem 5.3.2 in [**Ev**]. The last part is more difficult and proven in Theorem 3.22 of [**AF**] under weaker assumptions on ∂U (and in Theorem 5.3.3 in [**Ev**] for $\partial U \in C^1$ and bounded U).

Many theorems are much easier to tackle on $U = \mathbb{R}^m$. One thus wants to reduce results for (more or less) general U to the full space case. This can be done via so-called extension operators. Let $k \in \mathbb{N}_0$ and $p \in [1, \infty)$. The easiest case are maps $f \in C_c^k(U)$. These can simply be extended by 0 to the function $E_0 f = \tilde{f} \in C_c^k(\mathbb{R}^m)$, where we often omit the tilde. Clearly, one has $E_0(\partial^\alpha f) = \partial^\alpha E_0 f$ for the classical derivative and $|\alpha| \leqslant k$, so that E_0 is

[6] This is a bit stronger than the standard definition if ∂U is unbounded; see [**AF**] for a thorough discussion of boundary regularity in the context of Sobolev spaces.

an isometry for $\|\cdot\|_{k,p}$. To extend E_0, we set

$$W_0^{k,p}(U) = \overline{C_c^\infty(U)}^{k,p} \tag{3.26}$$

with closure in $W^{k,p}(U)$. Theorem 3.27 says that $W_0^{k,p}(\mathbb{R}^m) = W^{k,p}(\mathbb{R}^m)$. By Remark 3.40 below, the space $W_0^{k,p}(U)$ is strictly smaller than $W^{k,p}(U)$ if $\partial U \in C^1$ (and $U \neq \mathbb{R}^m$). Further note that $L^p(U) = W_0^{0,p}(U)$ since $C_c^\infty(U)$ is dense in $L^p(U)$ for $p < \infty$ and every open U.

We can thus extend E_0 to an isometry $E_0 : W_0^{k,p}(U) \to W^{k,p}(\mathbb{R}^m)$. Since also $\partial^\alpha : W_0^{k,p}(U) \to L^p(U)$ and $\partial^\alpha : W_0^{k,p}(\mathbb{R}^m) \to L^p(\mathbb{R}^m)$ are continuous, we have $E_0\partial^\alpha = \partial^\alpha E_0$ on $W_0^{k,p}(U)$ by approximation, where $|\alpha| \leqslant k$.

In this context also restriction operators occur. Let $V \subseteq U$ be open and define $R_V f = f\restriction_V$ for maps $f : U \to \mathbb{F}$. Let $f \in W_\alpha(U)$ and $\varphi \in C_c^\infty(V)$ with 0-extension $\tilde\varphi \in C_c^\infty(U)$. We then compute

$$(-1)^{|\alpha|}\int_V R_V f\, \partial^\alpha\varphi \,\mathrm{d}x = (-1)^{|\alpha|}\int_U f\,\partial^\alpha\tilde\varphi \,\mathrm{d}x = \int_U \partial^\alpha f\,\tilde\varphi \,\mathrm{d}x = \int_V R_V(\partial^\alpha f)\varphi \,\mathrm{d}x,$$

so that $R_V f \in W_\alpha(V)$ and $\partial^\alpha R_V f = R_V\partial^\alpha f$. Hence, R_V induces a contraction $R_V : W^{k,p}(U) \to W^{k,p}(V)$, and we have $R_U E_0 = I$ on $W_0^{k,p}(U)$.

The extension result for $W^{k,p}(U)$ requires assumptions on ∂U. We state a core extension theorem for Lipschitz domains due to *Stein*. A more general version is proven in Theorem VI.5 of [**St**], and a sketch is given in §5.25 of [**AF**]. The proof uses also harmonic analysis.

THEOREM 3.28. *Let $\partial U \in C^{1-}$, $k \in \mathbb{N}$, and $p \in [1,\infty]$. Then there exists an operator $E_{k,p}$ in $\mathcal{B}(W^{k,p}(U), W^{k,p}(\mathbb{R}^m))$ with $E_{k,p}u = u$ on U. These operators coincide on intersections of the respective spaces. We thus write E_U for all of them.*

Note that $R_U E_U$ is the identity on $W^{k,p}(U)$. If $\partial U \in C^1$, U is bounded, $k = 1$ and $p < \infty$, we indicate how construct an extension operator for $W^{1,p}(U)$ in a simpler way (see Theorem 5.4.1 in [**Ev**]). By the density result from Theorem 3.27, we may restrict ourselves to $f \in C_c^\infty(\overline{U})$. One then reduces the problem to the halfspace $\mathbb{R}^m_+$. Using the definition of a C^1 domain, we cover $\overline{U}$ by open sets $U_j = U_{y^j}$ with $j \in \{0, 1, \dots, N\}$ such that $\overline{U_0} \subseteq U$, $U_1, \dots, U_N$ cover ∂U, and there are charts $\psi_j : U_j \to V_j$ for $j \in \{1, \dots, N\}$. Let $\{\varphi_j \,|\, j \in \{0, 1, \dots, N\}\}$ be a smooth partition of unity subject to $\{U_j\}$. The part $R_{U_0}(\varphi_0 f)$ can be treated by means of E_0. The other parts are studied via the functions $g_j := (R_{U_j\cap U}(\varphi_j f)) \circ \psi_j^{-1}$ on $V_j \cap \mathbb{R}^m_+$. We extend these functions to maps Eg_j on $\mathbb{R}^m$ by first setting them 0 on $\mathbb{R}^m_+\backslash V_j$ and then extending them to $x_m = 0$ by continuity and to $x \in \mathbb{R}^m_- = \{x \in \mathbb{R}^m \,|\, x_m < 0\}$ by defining

$$Eg_j(x) = -3g_j(x', -x_m) + 4g_j(x', -x_m/2). \tag{3.27}$$

Observe that Eg_j belongs to $C^1(\mathbb{R}^m)$. Using Propositions 3.19 and 3.20 one can check

$$\|Eg_j\|_{W^{1,p}(\mathbb{R}^m)} \leqslant c\|g_j\|_{W^{1,p}(\mathbb{R}^m_+)} \leqslant c\|f\|_{W^{1,p}(U)}.$$

Using further cut-off functions and composing with ψ_j, one can then glue together the pieces Eg_j to the desired operator $E_U g$.

One can generalize this procedure to the spaces $W^{k,p}(U)$ assuming that $\partial U \in C^k$ and modifying the formula in (3.27). The resulting operator only works for $W^{j,p}(U)$ for $0 \leqslant j \leqslant k$, in contrast to the Stein extension operator.

As an illustration, we use E_U to extend the characterization of $W^{1,\infty}(U)$ from Proposition 3.24 to Lipschitz domains.

Corollary 3.29. *Let $\partial U \in C^{1-}$. Then $W^{1,\infty}(U)$ is isomorphic to $C^{1-}(U)$.*

Proof. Proposition 3.24 yields the embedding $C^{1-}(U) \hookrightarrow W^{1,\infty}(U)$, and the converse one for convex U, e.g., for $U = \mathbb{R}^m$. To show $W^{1,\infty}(U) \hookrightarrow C^{1-}(U)$, take $f \in W^{1,\infty}(U)$. By Theorem 3.27 and Proposition 3.24, the extension $E_U f$ belongs to $C^{1-}(\mathbb{R}^m)$ with norm bounded by $c\|f\|_{W^{1,\infty}(U)}$. Hence, $f = R_U E_U f$ is contained in $C^{1-}(U)$ with the same norm bound. □

Another important topic are embeddings of Sobolev spaces. We first note the easy ones

$$W^{k,p}(U) \hookrightarrow W^{j,p}(U) \quad \text{and} \quad W^{k,p}(U) \hookrightarrow W^{j,q}(U) \;\text{ if } \lambda(U) < \infty, \tag{3.28}$$

for $k \geqslant j \geqslant 0$ and $1 \leqslant q \leqslant p \leqslant \infty$. By a 'scaling argument', we next see which integrability exponents may occur in such embeddings for $U = \mathbb{R}^m$, where we restrict ourselves to $k = 1$.

Remark 3.30. Assume that $\|f\|_q \leqslant c\|f\|_{1,p}$ for some $p, q \in [1,\infty]$, a constant c, and all $f \in C_c^1(\mathbb{R}^m)$. We take $f \in C_c^1(\mathbb{R}^m)\backslash\{0\}$ and $a > 0$. Standard substitutions yield $\|D_a f\|_q = a^{-m/q}\|f\|_q$ and $\|\partial_j D_a f\|_p = a^{1-m/p}\|\partial_j f\|_p$, see (3.3). So the assumed estimate implies

$$a^{-m/q}\|f\|_q = \|D_a f\|_q \leqslant c\|D_a f\|_{1,p} = c\big(a^{-m/p}\|f\|_p + a^{1-m/p}\||\nabla f|_p\|_p\big)$$

We then obtain $p \leqslant q$ in the limit $a \to 0$, and $1 - \frac{m}{p} \geqslant -\frac{m}{q}$ as $a \to \infty$. If we only assume that $\|f\|_q \leqslant c\,\||\nabla f|_p\|_p$, it even follows $1 - \frac{m}{p} = -\frac{m}{q}$. ◊

For $j \in \mathbb{N}_0$ and $\beta \in (0,1)$, we use the spaces

$$C_0^j(\overline{U}) = \big\{f\restriction_{\overline{U}} \,\big|\, f \in C^j(\mathbb{R}^m),\ \forall\, 0 \leqslant |\alpha| \leqslant j : \partial^\alpha f(x) \to 0 \text{ as } |x|_2 \to \infty\big\},$$
$$C_0^{j+\beta}(\overline{U}) = \big\{f \in C_0^j(\overline{U}) \,\big|\, \forall\, 0 \leqslant |\alpha| \leqslant j : \partial^\alpha u \in C^\beta(\overline{U})\big\}.$$

Note that functions in $C_0^j(\overline{U})$ or $C_0^{j+\beta}(\overline{U})$ and their derivatives up to order j have bounded and continuous extensions to ∂U.

We now state the *Sobolev embedding theorem*, where the C^β-part is called *Morrey embedding*. In cases a) and b) the injection is given by inclusion, in c) by picking the continuous representative. So case a), for instance, just means that there is constant c_S with $\|f\|_{j,q} \leqslant c_S\|f\|_{k,p}$ for all $f \in W^{k,p}(U)$.

Theorem 3.31. *Let $k \in \mathbb{N}$, $j \in \mathbb{N}_0$, $p \in [1,\infty)$, und either $U = \mathbb{R}^m$ or $\partial U \in C^{1-}$. We have the following embeddings.*

a) If $q \in [p,\infty)$ and $k - \frac{m}{p} \geqslant j - \frac{m}{q}$, then

$$W^{k,p}(U) \hookrightarrow W^{j,q}(U).$$

b) If $q \in [p,\infty)$ and $k - \frac{m}{p} = j$, then

$$W^{k,p}(U) \hookrightarrow W^{j,q}(U).$$

c) If $\beta \in (0,1)$ and $k - \frac{m}{p} = j + \beta$, then

$$W^{k,p}(U) \hookrightarrow C_0^{j+\beta}(\overline{U}).$$

These embeddings are true for every open U if one replaces W by W_0.

A more general version of this result is shown in Theorem 4.12 of [**AF**], see also Section 5.6 of [**Ev**]. We note that it is enough to treat $U = \mathbb{R}^m$ by extension (Theorem 3.28) and to consider $f \in C_c^\infty(\mathbb{R}^m)$ due to density (Theorem 3.27). Iteratively one can reduce to the case $k = 1$. On this level the desired estimates can be shown by elementary, but tricky arguments involving the fundamental theorem of calculus, Hölder's inequality, and lengthy calculations. The addendum follows by the definition (3.26) of $W_0^{k,p}(U)$.

For $p = 2$ we can show the core estimates also using the Fourier transform. We present the easiest case. Let $k > m/2$ and $f \in C_c^\infty(\mathbb{R}^m)$ so that $\mathcal{F}f \in \mathcal{S}_m$. Using the Fourier inversion formula from Theorem 3.11, estimate (3.2), Hölder's inequality, Theorem 3.25 and polar coordinates, we compute

$$\begin{aligned}\|f\|_\infty &= \|\mathcal{F}^{-1}\mathcal{F}f\|_\infty \leqslant (2\pi)^{-m/2}\,\|(1+|\xi|_2^k)^{-1}(1+|\xi|_2^k)\widehat{f}\|_1\\ &\leqslant (2\pi)^{-m/2}\,\|(1+|\xi|_2^k)^{-1}\|_2\,\|(1+|\xi|_2^k)\widehat{f}\|_2\\ &\leqslant c\|f\|_{k,2}\Big(\int_0^\infty \frac{r^{m-1}}{(1+r^k)^2}\,\mathrm{d}r\Big)^{\frac12} \leqslant c\|f\|_{k,2}\end{aligned}$$

for constants $c > 0$, since $2k > m$.

We reformulate Theorem 3.31 modifying the statement for the borderline cases $k - m/p \in \mathbb{N}$ a bit.

Corollary 3.32. *Let $p \in [1,\infty)$, $k \in \mathbb{N}$, and either $U = \mathbb{R}^m$ or $\partial U \in C^{1-}$.*

a) Let $kp < m$. Then $p^\star := \frac{pm}{m-kp} > p$ and $W^{k,p}(U) \hookrightarrow L^q(U)$ for $q \in [p, p^\star]$.

b) Let $kp = m$. Then $W^{k,p}(U) \hookrightarrow L^q(U)$ for all $q \in [p,\infty)$.

c) Let $kp > m$. Then either $k - \frac{m}{p} \in \mathbb{N}$ or $k - \frac{m}{p} = j + \beta$ for some $j \in \mathbb{N}_0$ and $\beta \in (0,1)$. In the first case, set $j := k - \frac{m}{p} - 1 \in \mathbb{N}_0$ and take any $\beta \in (0,1)$. We obtain $W^{k,p}(U) \hookrightarrow C_0^{j+\beta}(\overline{U})$.

For $k - \frac{m}{p} \in \mathbb{N}_0$ in b) and c) the supnorm embeddings fail in general, but not always; as we discuss next.

Remark 3.33. a) Let $m \geqslant 2$, $\alpha \in (0, 1 - \frac{1}{m})$, and $U = B(0, \frac12)$. Set $f(x) = (-\ln|x|_2)^\alpha$ for $x \in U\backslash\{0\}$ and $f(0) = 0$. Then $f \in W^{1,m}(U)\backslash L^\infty(U)$.

Proof. Arguing as in Example 3.17, one sees that $f \in L^p(U)$ for all $p < \infty$, $f \notin L^\infty(U)$, and

$$\partial_j f(x) = -\alpha(-\ln|x|_2)^{\alpha-1}\frac{x_j}{|x|_2^2}$$

for $x \neq 0$ and $j \in \{1,\dots,m\}$. Using polar coordinates, we further estimate

$$\|\partial_j f\|_m \leqslant c\Big(\int_0^{\frac12} \frac{|\ln r|^{(\alpha-1)m}}{r^m} r^{m-1}\,\mathrm{d}r\Big)^{\frac1m} \leqslant c\Big(\int_0^{\frac12}\frac{\mathrm{d}r}{r(\ln r)^{(1-\alpha)m}}\Big)^{\frac1m} < \infty$$

for some constants $c > 0$, since $(1-\alpha)m > 1$. $\square$

b) However, we have the embedding $W^{m,1}(U) \hookrightarrow C_0(\overline{U})$ if $\partial U \in C^{1-}$. It is enough to prove it for $U = \mathbb{R}^m$ (with a constant c_S) since then

$$\|f\|_{C_0(\overline{U})} \leqslant \|E_U f\|_{C_0(\mathbb{R}^m)} \leqslant c_S \|E_U f\|_{W^{m,1}(\mathbb{R}^m)} \leqslant c_S \|E_U\| \|f\|_{W^{m,1}(U)}$$

for all $f \in W^{m,1}(U)$. By density, we only have to treat $f \in C_c^\infty(\mathbb{R}^m)$. We restrict to $m = 2$, as the other cases are similar. Then the fundamental theorem of calculus yield

$$f(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} \partial_x \partial_y f(s, t) \,\mathrm{d}t \,\mathrm{d}s.$$

for $(x, y) \in \mathbb{R}^2$, so that $\|f\|_\infty \leqslant \|f\|_{2,1}$. ♢

For a bounded Lipschitz domain U, the Morrey embedding and the Arzela–Ascoli theorem imply that the embedding $W^{k,p}(U) \hookrightarrow C(\overline{U})$ is compact if $kp > m$. (One says that $W^{k,p}(U)$ *compactly embedded* into $C(\overline{U})$.) The Rellich–Kondrachov theorem extends this fact to the case $kp < m$. Recall that a compact embedding $J : Y \hookrightarrow X$ means that any bounded sequence (y_n) in Y has a subsequence $(Jy_{n_j})_j$ with limit in X.

Theorem 3.34. *Let U be bounded with $\partial U \in C^{1-}$, $k \in \mathbb{N}$ and $1 \leqslant p < \infty$. Then the following assertions hold.*

a) Let $kp \leqslant m$ and $1 \leqslant q < p^\star = \frac{mp}{m-kp} \in (p, \infty]$. Then $W^{k,p}(U)$ is compactly embedded in $L^q(U)$. (For instance, let $q = p$.)

b) Let $k - \frac{m}{p} > j \in \mathbb{N}_0$. Then $W^{k,p}(U)$ is compactly embedded in $C^j(\overline{U})$.

The second part can be shown as indicated above, for the first one we refer to Theorem 5.7.1 of [**Ev**]. We discuss the sharpness of claim a) above.

Remark 3.35. a) Theorem 3.34 is wrong for unbounded domains, in general. In fact, let $k \in \mathbb{N}$, $p \in [1, \infty)$, and define $f_n = f(\cdot - n)$ in $W^{k,p}(\mathbb{R})$ for any function $0 \neq f \in C^\infty(\mathbb{R})$ with $\operatorname{supp} f \subseteq (-1/2, 1/2)$. Then $\|f_n\|_{k,p}$ and $\|f_n - f_m\|_q > 0$ do not depend on $n \neq m$ in $\mathbb{N}$ so that (f_n) is bounded in $W^{k,p}(\mathbb{R})$ and has no subsequence with limit in L^q for $1 \leqslant q < p^\star$.

b) The embedding $W^{1,p}(U) \hookrightarrow L^{p^\star}(U)$ is never compact, see Example 6.12 in [**AF**]. ♢

As we see in Example 5.11, it is often important to control a function by its derivative. Such results are provided by the *Poincaré inequalities* in the next theorem. By its part a), we can omit the term $\|f\|_p$ in the norm of $W_0^{1,p}(U)$ if, e.g., U is bounded. Here we say that U has *finite width* if is located between two parallel hyperplanes. Moreover, we write $a_U(f) = \lambda(U)^{-1} \int_U f \,\mathrm{d}x$ for the *average* of $f \in L^1(U)$.

Theorem 3.36. *a) Let U have finite width and $p \in [1, \infty)$. We then have $\|f\|_p \leqslant c \||\nabla f|_p\|_p$ for all $f \in W_0^{1,p}(U)$ and some $c > 0$.*

b) Let U be bounded and pathwise connected with $\partial U \in C^{1-}$ and $p \in [1, \infty)$. We then have $\|f - a_U(f)\mathbb{1}\|_p \leqslant c \||\nabla f|_p\|_p$ for all $f \in W^{1,p}(U)$ and some $c > 0$.

The first part is shown in the exercises by a direct calculation and the second in Theorem 5.8.1 in [**Ev**] using Rellich's theorem in a contradiction argument.

By the next result, first-order derivatives are small perturbations of second-order ones (cf. Remark 1.26). We can even omit $\|\partial_j f\|_p$ in the definition of $W^{2,p}(U)$. The following estimate can be extended to derivatives of higher order and is the starting point for a large family of 'interpolative inequalities', see e.g. Sections 5.1–5.3 of [**AF**].

Proposition 3.37. *Let $1 \leqslant p < \infty$ and $\partial U \in C^{1-}$. Let $f \in L^p(U) \cap \bigcap_{|\alpha|=2} W_\alpha$ with $\partial^\alpha f \in L^p(U)$ for $|\alpha| = 2$. We then have $f \in W^{2,p}(U)$ and there are constants $C, \varepsilon_0 > 0$ (depending only on $\|E_U\|$ and m) such that*

$$\Big(\sum_{j=1}^m \|\partial_j f\|_p^p\Big)^{1/p} \leqslant \varepsilon \Big(\sum_{i,j=1}^m \|\partial_{ij} f\|_p^p\Big)^{1/p} + \frac{C}{\varepsilon}\,\|f\|_p, \qquad (3.29)$$

for $\varepsilon \in (0, \varepsilon_0]$. This estimate is true for all open U and $\varepsilon > 0$ if $f \in W_0^{2,p}(U)$.

For $f \in W^{2,p}(U)$ or $f \in W_0^{2,p}(U)$ one can reduce the result to $f \in C_c^\infty(\mathbb{R}^m)$ by extension and density. In this case (3.29) is shown using the fundamental theorem of calculus, Hölder's inequality and further computations, see Theorem 5.2 of [**AF**]. The first part of the theorem can then be proven by approximation at least for certain domains, see Corollary 1.1.11 in [**Ma**] for the general case.

For $p > m$ and Lipschitz domains, we have $W^{1,p}(U) \hookrightarrow C(\overline{U})$ by the Sobolev Theorem 3.31 so that the 'trace map' $f \mapsto f{\restriction}_{\partial U}$ is well defined from $W^{1,p}(U)$ to $C(\partial U)$. Also, Remark 3.33 says that $W^{1,1}(a,b) \hookrightarrow C([a,b])$ for $m = 1$. In other cases it is not clear at all how to give a meaning to the mapping $f \mapsto f{\restriction}_{\partial U}$ on $W^{1,p}(U)$ as for reasonable open sets ∂U is a null set. The following *trace theorem* solves this problem, and it very conveniently describes $W_0^{1,p}(U)$ as the space of maps in $W^{1,p}(U)$ with trace 0.

Theorem 3.38. *Let $p \in [1, \infty)$ and $\partial U \in C^1$ (or U be bounded and $\partial U \in C^{1-}$). Then the trace map $f \mapsto f{\restriction}_{\partial U}$ from $W^{1,p}(U) \cap C(\overline{U})$ to $L^p(\partial U, \sigma)$ has a bounded linear extension $\mathrm{tr} : W^{1,p}(U) \to L^p(\partial U, \sigma)$ whose kernel is $W_0^{1,p}(U)$, where σ is the surface measure on ∂U.*

For C^1-boundaries these results are proved in Theorems 5.36 and 5.37 of [**AF**]; the case of bounded Lipschitz domains is treated in Sections 2.4.2 and 2.4.3 of [**Ne**]. We add a proof for a stronger statement if $p = 2$ and $U = \mathbb{R}_+^m$ using the Fourier transform.

Remark 3.39. We write $\mathbb{R}_+^m = \{(x, y) \mid x \in \mathbb{R}^{m-1}, y > 0\}$ with $m \geqslant 2$.

1) Let $u \in C_c^1(\overline{\mathbb{R}_+^m})$ and $\mathcal{F}_x$ be the Fourier transform with respect to $x \in \mathbb{R}^{m-1}$. The map $y \mapsto (\mathcal{F}_x u)(\xi, y)$ then belongs to $C_c^1(\mathbb{R}_{\geqslant 0})$ for each $\xi \in \mathbb{R}^{m-1}$. Since we deal with C_c^1-functions, we can calculate

$$\begin{aligned}
\int_{\mathbb{R}^{m-1}} |\xi|_2\, |(\mathcal{F}_x u)(\xi, 0)|^2\, \mathrm{d}\xi &= -\int_{\mathbb{R}^{m-1}} |\xi|_2 \int_0^\infty \partial_y\, |(\mathcal{F}_x u)(\xi, y)|^2\, \mathrm{d}y\, \mathrm{d}\xi \\
&= -2\,\mathrm{Re} \int_0^\infty \int_{\mathbb{R}^{m-1}} |\xi|_2\, \overline{\mathcal{F}_x u}\, \partial_y(\mathcal{F}_x u)\, \mathrm{d}\xi\, \mathrm{d}y \\
&\leqslant 2 \Big[\int_0^\infty \int_{\mathbb{R}^{m-1}} |\xi \mathcal{F}_x u|_2^2\, \mathrm{d}\xi\, \mathrm{d}y\Big]^{\frac12} \Big[\int_0^\infty \int_{\mathbb{R}^{m-1}} |\partial_y(\mathcal{F}_x u)|^2\, \mathrm{d}\xi\, \mathrm{d}y\Big]^{\frac12}
\end{aligned}$$

$$= 2\left[\int_0^\infty \int_{\mathbb{R}^{m-1}} |\mathcal{F}_x(\nabla_x u)|_2^2 \,\mathrm{d}\xi\,\mathrm{d}y\right]^{\frac{1}{2}} \left[\int_0^\infty \int_{\mathbb{R}^{m-1}} |\mathcal{F}_x(\partial_y u)|^2 \,\mathrm{d}\xi\,\mathrm{d}y\right]^{\frac{1}{2}}$$
$$= 2\left[\int_0^\infty \int_{\mathbb{R}^{m-1}} |\nabla_x u|_2^2 \,\mathrm{d}x\,\mathrm{d}y\right]^{\frac{1}{2}} \left[\int_0^\infty \int_{\mathbb{R}^{m-1}} |\partial_y u|^2 \,\mathrm{d}x\,\mathrm{d}y\right]^{\frac{1}{2}}$$
$$\leqslant \int_{\mathbb{R}^m_+} |\nabla_x u|_2^2 \,\mathrm{d}z + \int_{\mathbb{R}^m_+} |\partial_y u|_2^2 \,\mathrm{d}z \leqslant \|u\|^2_{W^{1,2}(\mathbb{R}^m_+)},$$

also using Hölder's inequality and Theorems 3.25 and 3.11.

2) For $s > 0$ we define the *Bessel potential space*

$$H^s(\mathbb{R}^n) = \left\{v \in L^2(\mathbb{R}^n) \,\middle|\, |\xi|_2^s\,\hat{v} \in L^2(\mathbb{R}^n)\right\}$$

endowed with the (Hilbertian) norm given by

$$\|v\|^2_{H^s} = \int_{\mathbb{R}^n} (1 + |\xi|_2^2)^s\, |\hat{v}(\xi)|^2 \,\mathrm{d}\xi.$$

(Theorem 3.25 yields $H^k(\mathbb{R}^n) = W^{k,2}(\mathbb{R}^n)$ for $k \in \mathbb{N}$.) We have thus shown that the trace map is continuous from $(C_c^1(\overline{\mathbb{R}^m_+}), \|\cdot\|_{1,2})$ to $H^{\frac{1}{2}}(\mathbb{R}^{m-1})$. Since $C_c^1(\overline{\mathbb{R}^m_+})$ is dense in $W^{1,2}(\mathbb{R}^m_+)$ Theorem 3.27, the trace map is continuous from $W^{1,2}(\mathbb{R}^m_+)$ to $H^{1/2}(\mathbb{R}^{m-1})$.

3) It is possible to show that $\mathrm{tr} : W^{1,2}(\mathbb{R}^m_+) \to H^{\frac{1}{2}}(\mathbb{R}^{m-1})$ is surjective. One can show analogous boundedness and surjectivity results on bounded U with $\partial U \in C^{1-}$ and for $p \in (1,\infty)$ – employing somewhat different spaces of functions on ∂U. See e.g. Theorem 5.7 in [**Ne**]. ♢

We also not an intersting consequence of the characterization of $W_0^{1,p}(U)$.

Remark 3.40. The above theorem implies that $W_0^{1,p}(U)$ differs from $W^{1,p}(U)$ if $U \neq \mathbb{R}^m$ has a C^1-boundary (where $p < \infty$). Indeed, the restriction to U of a map $\varphi \in C_c^\infty(\mathbb{R}^m)$ with $\varphi(x) \neq 0$ for some $x \in \partial U$ belongs to $W^{1,p}(U)$, but has non-zero trace. For less regular ∂U it my happen that $W_0^{1,p}(U) = W^{1,p}(U)$, see Theorem 3.33 in [**AF**]. ♢

Having the continuous trace map, we can now prove *Gauß' divergence theorem* and *Green's formulas* in Sobolev spaces. They are crucial tools in many applications to partial differential operators.

Theorem 3.41. *Let $U = \mathbb{R}^m$ or let $\partial U \in C^{1-}$ be bounded with the outer unit normal ν, $p \in [1,\infty]$, $F \in W^{1,p}(U)^m$, and $\varphi \in W^{1,p'}(U)$. We obtain*

$$\int_U \operatorname{div}(F)\varphi \,\mathrm{d}x = -\int_U F \cdot \nabla\varphi \,\mathrm{d}x + \int_{\partial U} \varphi\, \nu \cdot F \,\mathrm{d}\sigma. \tag{3.30}$$

If $u \in W^{2,p}(U)$ and $v \in W^{2,p'}(U)$, we obtain

$$\begin{aligned}\int_U (\Delta u)v \,\mathrm{d}x &= -\int_U \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial U} (\partial_\nu u)v \,\mathrm{d}\sigma \\ &= \int_U u\Delta v \,\mathrm{d}x + \int_{\partial U} ((\partial_\nu u)v - u\partial_\nu v) \,\mathrm{d}\sigma.\end{aligned} \tag{3.31}$$

If $U = \mathbb{R}^m$, these formulas hold without the boundary integrals. (We omit tr *in the boundary integrals and set $\partial_\nu u = \sum_{j=1}^m \nu_j \,\mathrm{tr}\,\partial_j u$.)*

Proof.[7] We first observe that Green's formulas (3.31) are a straightforward consequence of (3.30) with $F = \nabla u$.

1) Let U be bounded. Gauß' formula (3.30) is shown in analysis courses for $F \in C^1(\overline{U})^m$ and $\varphi \in C^1(\overline{U})$.

a) Let $p \in (1,\infty)$. For $F \in W^{1,p}(U)^m$ and $\varphi \in W^{1,p'}(U)$, Theorem 3.27 provides functions $F_n \in C^1(\overline{U})^m$ and $\varphi_n \in C^1(\overline{U})$ that converge to F and φ in $W^{1,p}(U)^m$ and $W^{1,p'}(U)$ as $n \to \infty$, respectively, since $p, p' < \infty$. Theorem 3.38 then yields that $F_n\restriction_{\partial U} \to \operatorname{tr} F$ in $L^p(U,\sigma)^m$ and $\varphi_n\restriction_{\partial U} \to \operatorname{tr}\varphi$ in $L^{p'}(U,\sigma)$ as $n \to \infty$. Since $\partial_j : W^{1,q}(U) \to L^q(U)$ is continuous for $q \in \{p, p'\}$, we further obtain that the terms with derivatives converge in L^p, resp. in $L^{p'}$. Formula (3.30) now follows by approximation, using (3.4).

b) Next, let $p = 1$ and thus $p' = \infty$. As above, $\operatorname{div} F_n$ and F_n converge to $\operatorname{div} F$ and F in $L^1(U)$ and $L^1(U)^m$, respectively, as well as $F_n\restriction_{\partial U} \to \operatorname{tr} F$ in $L^1(\partial U,\sigma)^m$. By Corollary 3.29, the function φ belongs to $C(\overline{U})$ and one can thus be extended to a map $\varphi \in C_c(\mathbb{R}^m)$, e.g., by $G_1E_0\varphi$. Set $\varphi_n = G_{\frac{1}{n}}\varphi \in C_c^\infty(\mathbb{R}^m)$. Properties (3.17) and (3.16) as well as Lemma 3.16 imply that $\varphi_n \to \varphi$ in $C(\overline{U})$, $\|\nabla\varphi_n\|_\infty \leqslant \|\nabla\varphi\|_\infty$ and $\nabla\varphi_n \to \nabla\varphi$ pointwise a.e., as $n \to \infty$, where we possibly pass to a subsequence also denoted by $(\varphi_n)_n$. We can now take the limit $n \to \infty$ in the first and the third integral of equation (3.30) for F_n and φ_n. For the second integral we use the estimate

$$\Big|\int_U F_n \cdot \nabla\varphi_n \,\mathrm{d}x - \int_U F \cdot \nabla\varphi \,\mathrm{d}x\Big| \leqslant \|F_n - F\|_1 \,\|\nabla\varphi_n\|_\infty + \int_U |F|\,|\nabla\varphi_n - \nabla\varphi| \,\mathrm{d}x$$

and Lebesgue's theorem for the last integral above. So (3.30) holds for F and φ. The case $p = \infty$ is treated analogously.

2) Let U be unbounded. There is a radius $r > 0$ with $\partial U \subseteq B(0,r)$. Let $k \in \mathbb{N}$ with $k > r$. We define cut-off functions $\chi_k \in C_c^\infty(\mathbb{R}^m)$ by $\chi_k(x) = \phi(|x|_2/k)$ for some $\phi \in C^\infty(\mathbb{R})$ with $0 \leqslant \phi \leqslant 1$, $\phi = 1$ on $[0,1]$ and $\phi = 0$ on $[2,\infty)$. Note that $0 \leqslant \chi_k \leqslant 1$, $\operatorname{supp}\chi_k \subseteq \overline{B}(0,2k)$, $\chi_k = 1$ on $B(0,k) \supseteq \partial U$, $\|\chi_k'\|_\infty \leqslant c/k$, and $\chi_k \to \mathbb{1}$ pointwise as $k \to \infty$.

For $F_k := \chi_k F$ and $\varphi_k := \chi_k\varphi$, formula (3.30) follows from step 1) replacing U by $V_k := U \cap B(0,2k)$. Note that $\partial V_k = \partial U \cup \partial B(0,2k)$. The properties of χ_k then yield that $F_k = F$ and $\varphi_k = \varphi$ on ∂U and that they vanish on $\mathbb{R}^m \backslash B(0,2k)$, and so we have

$$\int_U \operatorname{div}(F_k)\varphi_k \,\mathrm{d}x = -\int_U F_k \cdot \nabla\varphi_k \,\mathrm{d}x + \int_{\partial U} \varphi\,\nu \cdot F \,\mathrm{d}\sigma.$$

Next, observe that

$$\operatorname{div}(F_k)\varphi_k = \chi_k^2 \operatorname{div}(F)\varphi + \chi_k\varphi F \cdot \nabla\chi_k, \quad F_k \cdot \nabla\varphi_k = \chi_k^2 F \cdot \nabla\varphi + \chi_k\varphi F \cdot \nabla\chi_k.$$

The terms with $\nabla\chi_k$ are bounded by $\frac{1}{k}|\varphi F|$, so that the corresponding integrals vanish as $k \to \infty$ since φF is integrable on U by Hölder's inequality. The terms with χ_k^2 tend pointwise to $\operatorname{div}(F)\varphi$ and $F \cdot \nabla\varphi$, respectively, and are majorized by these L^1-functions. The theorem of dominated convergence thus implies (3.30) for F and φ. This proof also works for $U = \mathbb{R}^m$. □

[7]The proof was omitted in the lectures.

3.4. Differential operators in L^p spaces

We now use some of above results to study differential operators in L^p spaces and discuss their closedness, spectra and the compactness of their resolvents. The study of these (and related) examples will be continued in Example 4.8 and in Chapter 5.

We first recall that $A = \partial^\alpha$ with $\mathrm{D}(A) = \{u \in W_\alpha(U) \mid u, \partial^\alpha u \in L^p(U)\}$ is closed in $L^p(U)$ by Lemma 3.16, where $1 \leqslant p \leqslant \infty$ and $\alpha \in \mathbb{N}_0^m$. Observe that $\mathrm{D}(A) = W^{1,p}(U)$ if $m = 1 = |\alpha|$. We next extend this result to ∂^2.

EXAMPLE 3.42. Let $J \subseteq \mathbb{R}$ be an open interval, $1 \leqslant p < \infty$, and $X = L^p(J)$. Then $A = \partial^2$ with $\mathrm{D}(A) = W^{2,p}(J)$ is closed.

PROOF. As noted above, A is closed on $D := \{u \in W_2(J) \mid u, \partial^2 u \in L^p(J)\}$. Proposition 3.37 then implies the remaining inclusion $\mathrm{D}(A) \subseteq D$. □

We note that in the one-dimensional case ∂^k is closed on $W^{k,p}(J)$ for all $k \in \mathbb{N}$ by similar proofs based on higher-order versions of Proposition 3.37. This property relies on the fact that the domain $W^{k,p}(J)$ does not contain derivatives in other directions. In contrast, ∂_1 on $W^{1,p}((0,1)^2)$ is not closed in $L^p((0,1)^2)$. This can be shown by functions $u_n(x,y) = \varphi_n(y)$, where $\varphi_n \in C^1([0,1])$ converges in $L^p(0,1)$ to a map $\varphi \notin W^{1,p}(0,1)$.

We next study the first derivative in more detail, first on $\mathbb{R}$. Since $W^{1,p}(\mathbb{R})$ is embedded in $C_0(\mathbb{R})$ by Theorem 3.31, the following domain $\mathrm{D}(A)$ exhibits 'boundary conditions' at $\pm\infty$.

EXAMPLE 3.43. Let $1 \leqslant p < \infty$, $X = L^p(\mathbb{R})$, and $A = \partial$ with $\mathrm{D}(A) = W^{1,p}(\mathbb{R})$. Then $\sigma(A) = \mathrm{i}\mathbb{R}$, $\sigma_{\mathrm{p}}(A) = \emptyset$ and

$$(R(\lambda, A)f)(t) = \begin{cases} \int_t^\infty \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s, & \operatorname{Re}\lambda > 0, \\ -\int_{-\infty}^t \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s, & \operatorname{Re}\lambda < 0, \end{cases}$$

for $t \in \mathbb{R}$ and $f \in X$, cf. Example 1.21 for $X = C_0(\mathbb{R})$.

PROOF. 1) Let $\operatorname{Re}\lambda > 0$ and $\varphi_\lambda = e_{\operatorname{Re}\lambda}\mathbb{1}_{\mathbb{R}_{\leqslant 0}}$. We denote the integral in the assertion by $R_\lambda f(t)$. Since $|R_\lambda(f(t))| \leqslant (\varphi_\lambda * |f|)(t)$ for all $t \in \mathbb{R}$, Young's inequality (3.5) yields that

$$\|R_\lambda f\|_p \leqslant \|\varphi_\lambda\|_1 \, \|f\|_p = \frac{1}{\operatorname{Re}\lambda}\|f\|_p\,,$$

and hence R_λ belongs to $\mathcal{B}(X)$. Let $f_n \in C_c^\infty(\mathbb{R})$ converge to f in X. We then compute $\frac{\mathrm{d}}{\mathrm{d}t}R_\lambda f_n = \lambda R_\lambda f_n - f_n \in X$, which tends to $\lambda R_\lambda f - f$ in X by the above estimate. Moreover, $R_\lambda f_n$ is an element of $\mathrm{D}(A)$. Since A is closed, also $R_\lambda f$ is contained in $\mathrm{D}(A)$ and $AR_\lambda f = \lambda R_\lambda f - f$; i.e., $\lambda I - A$ is surjective.

2) Let $Au = \mu u$ for some $u \in \mathrm{D}(A)$ and $\mu \in \mathbb{C}$. Theorem 3.22 says that u is continuous so that Remark 3.23 c) implies the continuous differentiability of u and thus $u' = \mu u$. Consequently, u is equal to a multiple of e_μ. As $e_\mu \notin X$, we obtain $u = 0$ and the injectivity of $\mu I - A$. We have shown that $\sigma_{\mathrm{p}}(A) = \emptyset$ and that $\lambda \in \rho(A)$ with $R(\lambda, A) = R_\lambda$ if $\operatorname{Re}\lambda > 0$. In the same way, one sees that λ belongs to $\rho(A)$ if $\operatorname{Re}\lambda < 0$, with the asserted resolvent.

3) Let $\lambda \in \mathrm{i}\mathbb{R}$. Take $\varphi_n \in C_c^1(\mathbb{R})$ with $\varphi_n = n^{-1/p}$ on $[-n, n]$, $\varphi_n(t) = 0$ for $|t| \geqslant n+1$ and $\|\varphi_n'\|_\infty \leqslant 2n^{-1/p}$ for every $n \in \mathbb{N}$. Then the map $u_n := \varphi_n e_\lambda$

is contained in $W^{1,p}(\mathbb{R})$ and satisfies $Au_n = \lambda u_n + \varphi_n e_\lambda$,

$$\|u_n\|_p \geqslant \Big(\int_{-n}^{n} |\varphi_n(t)e_\lambda(t)|^p \,\mathrm{d}t\Big)^{1/p} = n^{-1/p}(2n)^{1/p} = 2^{1/p},$$

$$\|\lambda u_n - Au_n\|_p = \|\varphi_n' e_\lambda\|_p = \Big(\int_{\{n\leqslant|t|\leqslant n+1\}} |\varphi_n'(t)e_\lambda(t)|^p \,\mathrm{d}t\Big)^{1/p} \leqslant 2n^{-1/p}2^{1/p}.$$

Consequently, λ is an element of $\sigma_{\mathrm{ap}}(A)$ and so $\sigma(A) = \mathrm{i}\mathbb{R}$. □

We next study the first derivative on subintervals of $\mathbb{R}$. These examples indicate that boundary conditions in $\mathrm{D}(A)$ can remove point spectrum (i.e., enhance injectivity), but may be an obstacle for surjectivity. Also their specific form heavily influences the solvability of the equation $\lambda u - Au = f$ (for given $f \in X$) and thus the spectrum.

Example 3.44. Let $1 \leqslant p < \infty$.

a) Let $X = L^p(0,1)$ and $A = \partial$ with $\mathrm{D}(A) = W^{1,p}(0,1)$. Then $\sigma(A) = \sigma_{\mathrm{p}}(A) = \mathbb{C}$ since e_λ belongs to $\mathrm{D}(A)$ with $Ae_\lambda = \lambda e_\lambda$ for all $\lambda \in \mathbb{C}$.

b) Let $X = L^p(0,1)$ and $A = \partial$ with $\mathrm{D}(A) = \{u \in W^{1,p}(0,1) \,|\, u(0) = 0\}$. Here we use the continuous representative of $u \in W^{1,p}(0,1)$ from Theorem 3.22. Then A is closed, $\sigma(A) = \emptyset$, $R(\lambda, A)f(t) = -\int_0^t \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s$ for $t \in (0,1)$, $f \in X$ and $\lambda \in \mathbb{C}$, and A has a compact resolvent.

Proof. As in Example 3.43 one sees that the above integral defines a bounded inverse of $\lambda I - A$ for all $\lambda \in \mathbb{C}$. In particular, A is closed by Remark 1.11. (Alternatively, take u_n in $\mathrm{D}(A)$ such that $u_n \to u$ and $Au_n \to g$ in X as $n \to \infty$. Lemma 3.16 yields that $u \in W^{1,p}(0,1)$ and $\partial u = g$. Using Theorem 3.22, we further infer that $0 = u_n(0) \to u(0)$. Hence, $u \in \mathrm{D}(A)$ and $Au = \partial u = g$.) Finally, Remark 2.13 and Theorem 3.34 imply that A has a compact resolvent. □

c) Let $X = L^p(\mathbb{R}_+)$ and $A = \partial$ with $\mathrm{D}(A) = W^{1,p}(\mathbb{R}_+)$. Then $\sigma(A) = \overline{\mathbb{C}_-}$, $\sigma_{\mathrm{p}}(A) = \mathbb{C}_-$, and $R(\lambda, A)f(t) = \int_t^\infty \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s$ for $t > 0$, $f \in X$ and $\operatorname{Re}\lambda > 0$.

Proof. The operator is closed due to Lemma 3.16. As in Example 3.43, one computes the resolvent for $\operatorname{Re}\lambda > 0$ and checks that $\mathrm{i}\mathbb{R}$ does not contain eigenvalues. If $\operatorname{Re}\lambda < 0$, then e_λ is an eigenfunction. Because $\sigma(A)$ is closed, it is equal to $\overline{\mathbb{C}_-}$. □

d) Let $X = L^p(\mathbb{R}_+)$ and $A = -\partial$ with $\mathrm{D}(A) = W_0^{1,p}(\mathbb{R}_+)$. Then A is closed, $\sigma(A) = \overline{\mathbb{C}_-}$, $\sigma_{\mathrm{p}}(A) = \emptyset$, and $R(\lambda, A)f(t) = \int_0^t \mathrm{e}^{-\lambda(t-s)} f(s)\,\mathrm{d}s$ for $t > 0$, $f \in X$ and $\operatorname{Re}\lambda > 0$. (Observe that $\tilde{A} = -\partial$ with $\mathrm{D}(\tilde{A}) = W^{1,p}(\mathbb{R}_+)$ has $\sigma(\tilde{A}) = \overline{\mathbb{C}_+}$ and $\sigma_{\mathrm{p}}(\tilde{A}) = \mathbb{C}_+ := -\mathbb{C}_-$ by part c) and Proposition 1.20.)

Proof. If $\operatorname{Re}\lambda > 0$, the formula for the resolvent is verified as in Example 3.43, so that A is closed. The point spectrum is empty since the only possible eigenfunctions e_λ do not fulfill the boundary condition in $\mathrm{D}(A)$.

Take $\operatorname{Re}\lambda < 0$ and $f = \mathbb{1}_{[0,1)}$. Let $u \in \mathrm{D}(A)$ satisfy $\lambda u - Au = f$. Then $u(0) = 0$ and $\partial u = -\lambda u + f$ is continuous except at $t = 1$. Since u is piecewise C^1 by Remark 3.23 c), we obtain $u(1) = \int_0^1 \mathrm{e}^{-\lambda(1-s)}\,\mathrm{d}s$ and

$$u(t) = \mathrm{e}^{-\lambda(t-1)}u(1) = \mathrm{e}^{-\lambda t}\int_0^1 \mathrm{e}^{\lambda s}\,\mathrm{d}s, \qquad t \geqslant 1.$$

So u does not belong to X, and thus f not to the range of $\lambda I - A$; i.e., $\lambda \in \sigma(A)$. The result then follows from the closedness of the spectrum. □

e) Let $X = L^p(0,1)$ and $A = \partial$ with $\mathrm{D}(A) = \{u \in W^{1,p}(0,1) \mid u(0) = u(1)\}$. Then A is closed, $\sigma(A) = \sigma_{\mathrm{p}}(A) = 2\pi\mathrm{i}\mathbb{Z}$, and A has a compact resolvent. These facts can be proved as in Example 2.16 for $X = C([0,1])$, using now Theorem 3.34 for the compactness. $\Diamond$

We now turn our attention to the second derivative.

Example 3.45. Let $X = L^p(\mathbb{R})$, $1 \leqslant p < \infty$, and $A = \partial^2$ with $\mathrm{D}(A) = W^{2,p}(\mathbb{R})$. Then $\sigma(A) = \mathbb{R}_{\leqslant 0}$.

Proof. 1) Set $A_1 = \partial$ and $\mathrm{D}(A_1) = W^{1,p}(\mathbb{R})$. Let $\mu \in \mathbb{C}\backslash\mathbb{R}_{\leqslant 0}$. There exists a number $\lambda \in \mathbb{C}_+$ such that $\mu = \lambda^2$. Observe that

$$(\mu I - A)u = (\lambda I - A_1)(\lambda I + A_1)u = -(\lambda I - A_1)(-\lambda I - A_1)u \qquad (3.32)$$

for all $u \in \mathrm{D}(A)$. Example 3.43 implies that $\lambda I \pm A_1$ are invertible. Hence, $\mu I - A$ is injective. Next, for $v \in W^{1,p}(\mathbb{R})$ the function

$$\partial(\lambda I + A_1)^{-1}v = A_1(\lambda I + A_1)^{-1}v = -\lambda(\lambda I + A_1)^{-1}v + v$$

belongs to $W^{1,p}(\mathbb{R})$. This means that $(\lambda I + A_1)^{-1}$ maps $W^{1,p}(\mathbb{R})$ into $\mathrm{D}(A)$. Given $f \in X$, the map $u := (\lambda I + A_1)^{-1}(\lambda I - A_1)^{-1}f$ thus is an element of $\mathrm{D}(A)$ and $\mu u - Au = f$ in view of the factorization (3.32). We have shown that $\mu \in \rho(A)$ and $R(\mu, A) = (\lambda I + A_1)^{-1}(\lambda I - A_1)^{-1}$.

2) For $\mu \leqslant 0$, we have $\mu = \lambda^2$ for some $\lambda \in \mathrm{i}\mathbb{R}$ and (3.32) is still true. The operator $\lambda I - A_1$ is not surjective since its range is not closed by Example 3.43 and Proposition 1.19. Equation (3.32) thus implies that $\mu I - A$ is not surjective, and hence $\sigma(A) = \mathbb{R}_{\geqslant 0}$. $\square$

Example 3.46. Let $X = L^p(0,1)$, $1 \leqslant p < \infty$, and $A = \partial^2$ with $\mathrm{D}(A) = W^{2,p}(0,1) \cap W_0^{1,p}(0,1)$. Then A is closed, $\sigma(A) = \sigma_{\mathrm{p}}(A) = \{-\pi^2 k^2 \mid k \in \mathbb{N}\}$, and A has a compact resolvent. These facts can be proved as in Example 2.16 for $X = C([0,1])$, using now Theorem 3.34 for the compactness. $\Diamond$

The situation is much more complicated for $m \geqslant 2$. Here the core example is the Laplacian $\Delta = \partial_{11} + \ldots + \partial_{mm}$. It is not clear at all that Δ is closed on $W^{2,p}(U)$ since the derivatives in Δu may exhibit cancellations and $\partial_{jk}u$ for $j \neq k$ do not appear. For $p = 2$ and $U = \mathbb{R}^m$ we can prove closedness using the Fourier transform.

Example 3.47. Let $X = L^2(\mathbb{R}^m)$ and $A = \Delta$ with $\mathrm{D}(A) = W^{2,2}(\mathbb{R}^m)$. Then A is closed and $\sigma(A) = \mathbb{R}_{\leqslant 0}$.

Proof. We employ the Fourier transform $\mathcal{F}$ which is unitary on X by Theorem 3.11. Theorem 3.25 yields

$$\Delta u = \mathcal{F}^{-1}(-|\xi|_2^2\widehat{u}) \quad \text{for} \quad u \in \mathrm{D}(A) = \big\{u \in X \;\big|\; |\xi|_2^2\widehat{u} \in X\big\}.$$

1) Let $\lambda \in \mathbb{C}\backslash\mathbb{R}_{\leqslant 0}$. Set $m_\lambda(\xi) = (\lambda + |\xi|_2^2)^{-1}$ for $\xi \in \mathbb{R}^m$. Observe that $\|m_\lambda\|_\infty \leqslant c_\lambda$ where $c_\lambda = 1/|\operatorname{Im}\lambda|$ if $\operatorname{Re}\lambda \leqslant 0$ and $c_\lambda = 1/|\lambda|$ if $\operatorname{Re}\lambda > 0$. Let $f \in X$. Then $m_\lambda\widehat{f}$ belongs to X so that $R_\lambda f := \mathcal{F}^{-1}(m_\lambda\widehat{f}) \in X$ and

$$\|R_\lambda f\|_2 = \|m_\lambda\widehat{f}\|_2 \leqslant c_\lambda\|\widehat{f}\|_2 = c_\lambda\|f\|_2\,.$$

Since $|\xi|_2^2 m_\lambda$ is bounded on $\mathbb{R}^m$, also the function $|\xi|_2^2\mathcal{F}R_\lambda f = |\xi|_2^2 m_\lambda\widehat{f}$ is an element of X, and so R_λ maps into $\mathrm{D}(A)$. Similarly we see that

$$(\lambda I - \Delta)R_\lambda f = \mathcal{F}^{-1}(\lambda + |\xi|_2^2)\mathcal{F}\mathcal{F}^{-1}m_\lambda\widehat{f} = f,$$

$$R_\lambda(\lambda I - \Delta)u = \mathcal{F}^{-1}m_\lambda \mathcal{F}\mathcal{F}^{-1}(\lambda + |\xi|_2^2)\widehat{u} = u, \quad u \in \mathrm{D}(A).$$

Thus, $\lambda \in \rho(A)$ and $R(\lambda, A) = R_\lambda$. In particular, A is closed and $\sigma(A) \subseteq \mathbb{R}_{\leqslant 0}$.

It can be seen that m_λ is the Fourier transform of an integrable function k_λ, see Proposition 6.1.5 in [**Gr**], so that $R_\lambda f = (2\pi)^{-m/2}k_\lambda * f$ by Theorem 3.11. For $m = 1$, the kernel k_λ is given as in Example 3.2.

2) Let $\lambda \leqslant 0$. Define $m_\lambda(\xi) = (\lambda + |\xi|_2^2)^{-1}$ for $\xi \in \mathbb{R}^m$ with $|\xi|_2 \neq \sqrt{-\lambda} =: \ell$ and $m_\lambda(\xi) = 0$ if $|\xi|_2 = \ell$. Set $g = \mathbb{1}_{B(0,\ell+1)} \in X$. Then $h = \mathcal{F}^{-1}g$ belongs to X and $m_\lambda \widehat{h}$ not. If there was an element u of $\mathrm{D}(A)$ with $\lambda u - \Delta u = h$, we would obtain as above $\lambda\widehat{u} + |\xi|_2^2\widehat{u} = \widehat{h}$ and the contradiction $m_\lambda\widehat{h} = \widehat{u} \in X$. As a result, λ belongs to $\sigma(A)$ and $\sigma(A) = \mathbb{R}_{\leqslant 0}$. □

The above result can be extended to exponents $p \in (1, \infty)$ and, imposing boundary conditions, to domains $U \neq \mathbb{R}^m$, but core parts of the corresponding proofs are beyond the scope of this lecture. We briefly discuss two examples and come back to this issue in Example 5.11.

Example 3.48. Let $1 < p < \infty$, $X = L^p(\mathbb{R}^m)$, and $A = \Delta$ with $\mathrm{D}(A) = W^{2,p}(\mathbb{R}^m)$. Then A is closed.

Proof. The Calderón–Zygmund estimate says that the graph norm of A is equivalent to $\|\cdot\|_{2,p}$ on $C_c^\infty(\mathbb{R}^m)$, see Corollary 9.10 in [**GT**]. Let $u \in W^{2,p}(\mathbb{R}^m)$. By Theorem 3.27, there are $u_n \in C_c^\infty(\mathbb{R}^m)$ converging to u in $W^{2,p}(\mathbb{R}^m)$ as $n \to \infty$, and hence $u_n \to u$ and $\Delta u_n \to \Delta u$ in X. We derive

$$\|u\|_{2,p} = \lim_{n\to\infty} \|u_n\|_{2,p} \leqslant \lim_{n\to\infty} c(\|u_n\|_p + \|\Delta u_n\|_p) = c(\|u\|_p + \|\Delta u\|_p) \leqslant c'\|u\|_{2,p},$$

so that $\|\cdot\|_A$ is equivalent to a complete one and thus A is closed. □

Example 3.49. Let $1 < p < \infty$, $U \subseteq \mathbb{R}^m$ be bounded and open with $\partial U \in C^2$, $X = L^p(U)$, and $A = \Delta$ with $\mathrm{D}(A) = W^{2,p}(U) \cap W_0^{1,p}(U)$. Then the *Dirichlet–Laplacian* A is closed, invertible and has a compact resolvent. We thus have $\sigma(A) = \sigma_{\mathrm{p}}(A)$.

Proof. The closedness of A follows from Theorem 9.14 in [**GT**]. Its bijectivity is shown in Theorem 9.15 of [**GT**]. Remark 2.13, Theorem 3.34 and Theorem 2.15 then imply the other assertions. □

There are variants of Examples 3.48 and 3.49 for $X = L^1(U)$, $X = L^\infty(U)$ and in other sup-norm spaces (with $m \geqslant 2$), see Theorem 5.8 in [**Ta**] as well as Sections 3.1.2 and 3.1.5 in [**Lu**]. Here the descriptions of the domains are much more complicated, and they are not just (subspaces of) Sobolev (or C^2-) spaces. To indicate the difficulties, we note that there is a function $u \notin W^{2,\infty}(B(0,1))$ with $B(0,1)$ in $\mathbb{R}^2$ such that $\Delta u \in L^\infty(B(0,1))$, namely

$$u(x,y) = (x^2 - y^2)\ln(x^2 + y^2), \quad (x,y) \neq (0,0),$$

and $u(0,0) = 0$. Then the second derivative

$$\partial_{xx}u(x,y) = 2\ln(x^2 + y^2) + \frac{4x^2}{x^2 + y^2} + \frac{(6x^2 - 2y^2)(x^2 + y^2) - 4x^2(x^2 - y^2)}{(x^2 + y^2)^2}$$

is unbounded around $(0,0)$, but $\Delta u(x,y) = 8\frac{x^2 - y^2}{x^2 + y^2}$ is bounded.

3.5. Appendix: Density, embedding and trace theorems

We[8] provide proofs for the results from Section 3.3, though often in a simplified setting and partly sketched. The material is presented in a somewhat different way as in Section 3.3. We start with a technical result that we use a few times.

LEMMA 3.50. *Let $K \subseteq U$ be compact. Then there is a function $\psi \in C_c^\infty(U)$ such that $0 \leqslant \psi \leqslant 1$ on U and $\psi = 1$ on K. Let $g \in L^1_{\mathrm{loc}}(U)$ satisfy*

$$\int_U g\varphi \,\mathrm{d}x = 0$$

for all $\varphi \in C_c^\infty(U)$. Then $g = 0$ a.e..

PROOF. 1) Let $0 < \delta < \frac{1}{2}\mathrm{d}(\partial K, \partial U)$. Then $K_{2\delta} = K + \overline{B}(0, 2\delta)$ is compact and $K_{2\delta} \subseteq U$. The function $\psi := G_\delta \mathbb{1}_{K_\delta}$ thus belongs to $C_c^\infty(U)$ by (3.14) and (3.15). Moreover, (3.13) and (3.16) imply that $0 \leqslant \psi(x) \leqslant \|\psi\|_\infty \leqslant \|\mathbb{1}_{K_\delta}\|_\infty = 1$ for all $x \in U$ and

$$\psi(x) = \int_{B(x,\delta)} k_\delta(x-y)\mathbb{1}_{K_\delta}(y)\,\mathrm{d}y = \|k_\delta\|_1 = 1$$

for all $x \in K$. The first claim is shown.

2) Assume that $g \neq 0$ on a Borel set $B \subseteq U$ with $\lambda(B) > 0$. Theorem 2.20 of [**Ru1**] yields a compact set $K \subseteq B \subseteq U$ with $\lambda(K) > 0$. Since $\psi g \in L^1(U)$, the functions $G_\varepsilon(\psi g)$ converge to ψg in $L^1(U)$ as $\varepsilon \to 0$ due to (3.17). Hence, there is a nullset N and a sequence $\varepsilon_j \to 0$ with $\varepsilon_j \leqslant \delta$ such that $(G_{\varepsilon_j}(\psi g))(x) \to g(x) \neq 0$ as $j \to \infty$ for each $x \in K\backslash N$. For every $x \in K\backslash N$ and $j \in \mathbb{N}$, we also deduce

$$(G_{\varepsilon_j}(\psi g))(x) = \int_U k_{\varepsilon_j}(x-y)\psi(y)\,g(y)\,\mathrm{d}y = 0$$

from the assumption, since the function $y \mapsto k_{\varepsilon_j}(x-y)\psi(y)$ belongs to $C_c^\infty(U)$. This is a contradiction. $\square$

In the first density result, we do not have smoothness up to ∂U. The first part of the result is essentially taken from Theorem 4.21 from [**FA**].

THEOREM 3.51. *Let $k \in \mathbb{N}$ and $p \in [1, \infty)$. We then have $W_0^{k,p}(\mathbb{R}^m) = W^{k,p}(\mathbb{R}^m)$. Moreover, the set $C^\infty(U) \cap W^{k,p}(U)$ is dense in $W^{k,p}(U)$.*

PROOF. We prove the theorem only for $k = 1$, the general case can be treated similarly.

1) Let $f \in W^{1,p}(\mathbb{R}^m)$. Take any $\phi \in C^\infty(\mathbb{R})$ with $0 \leqslant \phi \leqslant 1$, $\phi = 1$ on $[0,1]$, and $\phi = 0$ on $[2, \infty)$. Set $\varphi_n(x) = \phi(\frac{1}{n}|x|_2)$ for $n \in \mathbb{N}$ and $x \in \mathbb{R}^m$. We then have $\varphi_n \in C_c^\infty(\mathbb{R}^m)$, $0 \leqslant \varphi_n \leqslant 1$ and $\|\partial_j \varphi_n\|_\infty \leqslant \|\phi'\|_\infty \frac{1}{n}$ for all $n \in \mathbb{N}$, as well as $\varphi_n(x) \to 1$ for all $x \in \mathbb{R}^m$ as $n \to \infty$. Thus $\|\varphi_n f - f\|_p \to 0$ as $n \to \infty$ by Lebesgue's convergence theorem. Further, Proposition 3.19 implies that

$$\begin{aligned}\|\partial_j(\varphi_n f - f)\|_p &= \|(\varphi_n \partial_j f - \partial_j f) + (\partial_j \varphi_n) f\|_p \\ &\leqslant \|\varphi_n \partial_j f - \partial_j f\|_p + \tfrac{1}{n}\|\phi'\|_\infty \|f\|_p\,,\end{aligned}$$

[8] This section was not part of the lectures.

and the right hand side tends to 0 as $n \to \infty$ for each $j \in \{1, \ldots, m\}$. Given $\varepsilon > 0$, we can thus fix an index $N \in \mathbb{N}$ such that $\|\varphi_N f - f\|_{1,p} \leqslant \varepsilon$. Due to (3.14) and (3.15), the functions $G_{\frac{1}{n}}(\varphi_N f)$ belong to $C_c^\infty(\mathbb{R}^m)$ for all $n \in \mathbb{N}$. Equation (3.17) and Lemma 3.16 further yield that

$$G_{\frac{1}{n}}(\varphi_N f) \to \varphi_N f \quad \text{and} \quad \partial_j G_{\frac{1}{n}}(\varphi_N f) = G_{\frac{1}{n}} \partial_j(\varphi_N f) \to \partial_j(\varphi_N f)$$

in $L^p(\mathbb{R}^m)$ as $n \to \infty$, for $j \in \{1, \ldots, m\}$. So there is an index $n \in \mathbb{N}$ with

$$\|G_{\frac{1}{n}}(\varphi_N f) - \varphi_N f\|_{1,p} \leqslant \varepsilon,$$

and thus

$$\|G_{\frac{1}{n}}(\varphi_N f) - f\|_{1,p} \leqslant 2\varepsilon.$$

2) For the second assertion, we only have to consider the case $\partial U \neq \emptyset$. Let $f \in W^{1,p}(U)$. Set

$$U_n = \big\{x \in U \,\big|\, |x|_2 < n \ \text{ and } \ \mathrm{d}(x, \partial U) > \tfrac{1}{n}\big\}$$

for all $n \in \mathbb{N}$. We obtain $U_n \subseteq \overline{U_n} \subseteq U_{n+1} \subseteq U$, $\overline{U_n}$ is compact, and $\bigcup_{n=1}^\infty U_n = U$. Observe that $U = \bigcup_{n=1}^\infty U_{n+1} \backslash \overline{U_{n-1}}$, where U_0, $U_{-1} := \emptyset$. There are functions φ_n in $C_c^\infty(U)$ such that $\operatorname{supp} \varphi_n \subseteq U_{n+1} \backslash \overline{U_{n-1}}$, $\varphi_n \geqslant 0$, and $\sum_{n=1}^\infty \varphi_n(x) = 1$ for all $x \in U$. (Compare Theorem 3.15 in [**AF**].)

Fix $\varepsilon > 0$. As in step 1), for each $n \in \mathbb{N}$ there is a number $\delta_n > 0$ such that $g_n := G_{\delta_n}(\varphi_n f) \in C_c^\infty(U)$, $\operatorname{supp} g_n \subseteq (\operatorname{supp} \varphi_n f)_{\delta_n} \subseteq U_{n+1} \backslash \overline{U_{n-1}}$ and $\|g_n - \varphi_n f\|_{1,p} \leqslant 2^{-n}\varepsilon$. Define $g(x) = \sum_{n=1}^\infty g_n(x)$ for all $x \in U$. Observe that on each ball $\overline{B}(x,r) \subseteq U$ this sum is finite, so that g belongs to $C^\infty(U)$. Since $f = \sum_{n=1}^\infty \varphi_n f$, we further have

$$g(x) - f(x) = \sum_{n=1}^\infty (g_n(x) - \varphi_n(x) f(x)),$$

for all $x \in U$ and $n \in \mathbb{N}$. Due to $\|g_n - \varphi_n f\|_{1,p} \leqslant 2^{-n}\varepsilon$, this series converges absolutely in $W^{1,p}(U)$, and

$$\|f - g\|_{1,p} \leqslant \sum_{n=1}^\infty \|g_n - \varphi_n f\|_{1,p} \leqslant \varepsilon. \qquad \square$$

For 'not too bad' ∂U one can replace in $C^\infty(U)$ by $C^\infty(\overline{U})$ in Theorem 3.51, see Theorem 3.60 below.

We now want to study embeddings of Sobolev spaces. We clearly have

$$W^{k,p}(U) \hookrightarrow W^{j,p}(U), \tag{3.33}$$

$$W^{k,p}(U) \hookrightarrow W^{j,q}(U) \qquad \text{if } \ \lambda(U) < \infty, \tag{3.34}$$

for $k \geqslant j \geqslant 0$ and $1 \leqslant q \leqslant p \leqslant \infty$. (Recall that we set $W^{0,p}(U) = L^p(U)$ for $1 \leqslant p \leqslant \infty$.) The embedding $X \hookrightarrow Y$ means that there is an injective map $J \in \mathcal{B}(X,Y)$. Writing $c = \|J\|$, one obtains $\|f\|_Y \leqslant c\|f\|_X$ if one identifies Jf with f. We next prove the *Sobolev–Morrey embeddings* on $\mathbb{R}^m$.

THEOREM 3.52. *Let $k \in \mathbb{N}$ and $p \in [1,\infty)$. The following embeddings hold.*

a) If $kp < m$, then

$$p^\star := \frac{pm}{m - kp} \in (p, \infty) \quad \text{and} \quad W^{k,p}(\mathbb{R}^m) \hookrightarrow L^q(\mathbb{R}^m) \quad \text{for all} \ \ q \in [p, p^\star].$$

b) If $kp = m$, then

$$W^{k,p}(\mathbb{R}^m) \hookrightarrow L^q(\mathbb{R}^m) \qquad \textit{for all} \ \ q \in [p, \infty).$$

c) If $kp > m$, then there are either $j \in \mathbb{N}_0$ and $\beta \in (0, 1)$ such that $k - \frac{m}{p} = j + \beta$ or $k - \frac{m}{p} \in \mathbb{N}$. In the latter case we set $j := k - \frac{m}{p} - 1 \in \mathbb{N}_0$ and take any $\beta \in (0, 1)$. Then

$$W^{k,p}(\mathbb{R}^m) \hookrightarrow C_0^{j+\beta}(\mathbb{R}^m).$$

In parts a) and b), the embedding J is just the inclusion map, and in part c) the function Jf is the continuous representative of f.

Beore proving them, we rephrase the above results in a slightly modified way using the 'effective regularity index' $k - \frac{m}{p}$ of $W^{k,p}$.

Corollary 3.53. *Let $k \in \mathbb{N}$, $j \in \mathbb{N}_0$, and $p \in [1, \infty)$. We have the following embeddings.*

a) If $q \in [p, \infty)$ and $k - \frac{m}{p} \geqslant j - \frac{m}{q}$, then

$$W^{k,p}(\mathbb{R}^m) \hookrightarrow W^{j,q}(\mathbb{R}^m).$$

b) If $q \in [p, \infty)$ and $k - \frac{m}{p} = j$, then

$$W^{k,p}(\mathbb{R}^m) \hookrightarrow W^{j,q}(\mathbb{R}^m).$$

c) If $\beta \in (0, 1)$ and $k - \frac{m}{p} = j + \beta$, then

$$W^{k,p}(\mathbb{R}^m) \hookrightarrow C_0^{j+\beta}(\mathbb{R}^m).$$

Proof of Corollary 3.53. a) Let $(k - j)p = m$. The embedding

$$W^{k-j,p}(\mathbb{R}^m) \hookrightarrow L^q(\mathbb{R}^m) \tag{3.35}$$

then follows from Theorem 3.52 b) for all $q \in [p, \infty)$. Let $(k - j)p > m$. We then have $W^{k-j,p}(\mathbb{R}^m) \hookrightarrow L^p(\mathbb{R}^m)$ by (3.33) and $W^{k-j,p}(\mathbb{R}^m) \hookrightarrow L^\infty(\mathbb{R}^m)$ by Theorem 3.52 c). Hence the interpolation inequality (3.36) below implies (3.35) for all $q \in [p, \infty]$ in this case. Let $(k - j)p < m$. By assumption, we have $p \leqslant q \leqslant pm(m - (k - j)p)^{-1}$ and thus for these q the embedding (3.35) is a consequence of Theorem 3.52 a). Applying (3.35) to $\partial^\alpha f \in W^{k-j,p}(\mathbb{R}^m)$ for $|\alpha| \leqslant j$ and $f \in W^{k,p}(\mathbb{R}^m)$, we deduce

$$\|\partial^\alpha f\|_q \leqslant c\|\partial^\alpha f\|_{k-j,p} \leqslant c\|f\|_{k,p}.$$

So claim a) is true. Part b) follows from a), and c) from Theorem 3.52 c). □

For the proof of Theorem 3.52 we set $\hat{x}^j = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_m) \in \mathbb{R}^{m-1}$ for all $x \in \mathbb{R}^m$, $j \in \{1, \dots, m\}$, and $m \geqslant 2$. We start with a lemma.

Lemma 3.54. *Let $m \geqslant 2$ and $f_1, \dots, f_m \in L^{m-1}(\mathbb{R}^{m-1}) \cap C(\mathbb{R}^{m-1})$. Set $f(x) = f_1(\hat{x}^1) \cdot \ldots \cdot f_m(\hat{x}^m)$ for $x \in \mathbb{R}^m$. We then have $f \in L^1(\mathbb{R}^m)$ and*

$$\|f\|_{L^1(\mathbb{R}^m)} \leqslant \|f_1\|_{L^{m-1}(\mathbb{R}^{m-1})} \cdot \ldots \cdot \|f_m\|_{L^{m-1}(\mathbb{R}^{m-1})}.$$

Proof. If $m = 2$, then Fubini's theorem shows that

$$\int_{\mathbb{R}^2} |f(x)| \,\mathrm{d}x = \int_{\mathbb{R}} \int_{\mathbb{R}} |f_1(x_2)||f_2(x_1)| \,\mathrm{d}x_1 \,\mathrm{d}x_2 = \|f_1\|_1 \, \|f_2\|_1,$$

as asserted. Assume that the assertion holds for some $m \in \mathbb{N}$ with $m \geqslant 2$.

Take $f_1, \dots, f_{m+1} \in L^m(\mathbb{R}^m) \cap C(\mathbb{R}^m)$. Write $y = (x_1, \dots, x_m) \in \mathbb{R}^m$ and $x = (y, x_{m+1}) \in \mathbb{R}^{m+1}$. For a.e. $x_{m+1} \in \mathbb{R}$, the maps $\hat{y}^j \mapsto |f_j(\hat{y}^j, x_{m+1})|^m$ are integrable on $\mathbb{R}^{m-1}$ for each $j \in \{1, \dots, m\}$ due to Fubini's theorem. Fix such a $x_{m+1} \in \mathbb{R}$ and write

$$\tilde{f}(y, x_{m+1}) := \prod_{j=1}^{m} f_j(\hat{x}^j).$$

Using Hölder's inequality, we obtain

$$\begin{aligned} \int_{\mathbb{R}^m} |f(y, x_{m+1})| \,\mathrm{d}y &= \int_{\mathbb{R}^m} |\tilde{f}(y, x_{m+1})| \, |f_{m+1}(y)| \,\mathrm{d}y \\ &\leqslant \|f_{m+1}\|_{L^m(\mathbb{R}^m)} \left(\int_{\mathbb{R}^m} |\tilde{f}(y, x_{m+1})|^{m'} \,\mathrm{d}y \right)^{\frac{1}{m'}}. \end{aligned}$$

We set $g_j(\hat{y}^j) = |f_j(\hat{y}^j, x_{m+1})|^{m'}$ for $j \in \{1, \dots, m\}$ and $x \in \mathbb{R}^{m+1}$. Since $m'(m-1) = m$, the maps g_j belong to $L^{m-1}(\mathbb{R}^{m-1})$ and the induction hypothesis yields

$$\begin{aligned} \int_{\mathbb{R}^m} |\tilde{f}(y, x_{m+1})|^{m'} \,\mathrm{d}y &= \int_{\mathbb{R}^m} g_1(\hat{y}^1) \cdot \ldots \cdot g_m(\hat{y}^m) \,\mathrm{d}y \leqslant \|g_1\|_{m-1} \cdot \ldots \cdot \|g_m\|_{m-1} \\ &= \prod_{j=1}^{m} \left(\int_{\mathbb{R}^{m-1}} |f_j(\hat{y}^j, x_{m+1})|^m \,\mathrm{d}y \right)^{\frac{1}{m-1}}. \end{aligned}$$

Integrating over $x_{m+1} \in \mathbb{R}$, we thus arrive at

$$\int_{\mathbb{R}^{m+1}} |f| \,\mathrm{d}x \leqslant \|f_{m+1}\|_m \int_{\mathbb{R}} \prod_{j=1}^{m} \left(\int_{\mathbb{R}^{m-1}} |f_j(\hat{x}^j)|^m \,\mathrm{d}y \right)^{\frac{1}{m-1} \frac{m-1}{m}} \mathrm{d}x_{m+1}.$$

Applying the m-fold Hölder inequality to the x_{m+1}-integral, we conclude

$$\begin{aligned} \int_{\mathbb{R}^{m+1}} |f| \,\mathrm{d}x &\leqslant \|f_{m+1}\|_m \prod_{j=1}^{m} \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^{m-1}} |f_j(\hat{x}^j)|^m \,\mathrm{d}y \right)^{\frac{1}{m} \cdot m} \mathrm{d}x_{m+1} \right)^{\frac{1}{m}} \\ &= \|f_1\|_m \cdot \ldots \cdot \|f_{m+1}\|_m. \end{aligned}$$ □

Recall from Analysis 3 that for $f \in L^p(U) \cap L^q(U)$ and $r \in [p, q]$ with $1 \leqslant p < q \leqslant \infty$, we have

$$\|f\|_r \leqslant \|f\|_p^{\theta} \, \|f\|_q^{1-\theta} \leqslant \theta \|f\|_p + (1-\theta) \|f\|_q, \tag{3.36}$$

where $\theta \in [0, 1]$ is given by $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$ and we also used Young's inequality from Analysis 1.

Proof of Theorem 3.52. We only prove the case $k = 1$, the rest can be done by induction, see e.g. §5.6.3 in [**Ev**]. Since $W^{1,p}(\mathbb{R}^m) \hookrightarrow L^p(\mathbb{R}^m)$, in view of (3.36) for assertion a) it suffices to show

$$W^{1,p}(\mathbb{R}^m) \hookrightarrow L^{p^\star}(\mathbb{R}^m).$$

1) Let $f \in C_c^1(\mathbb{R}^m)$. Let first $p = 1 < m$, whence $p^\star = \frac{m}{m-1}$. For $x \in \mathbb{R}^m$ and $j \in \{1, \dots, m\}$, we then obtain

$$|f(x)| = \left| \int_{-\infty}^{x_j} \partial_j f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_m) \,\mathrm{d}t \right| \leqslant \int_{\mathbb{R}} |\partial_j f(x)| \,\mathrm{d}x_j,$$

$$|f(x)|^m \leqslant \prod_{j=1}^m \int_{\mathbb{R}} |\partial_j f(x)| \, \mathrm{d}x_j.$$

Setting $g_j(\hat{x}^j) = (\int_{\mathbb{R}} |\partial_j f(x)| \, \mathrm{d}x_j)^{\frac{1}{m-1}}$, we deduce

$$|f(x)|^{\frac{m}{m-1}} \leqslant \prod_{j=1}^m g_j(\hat{x}^j).$$

After integration over $x \in \mathbb{R}^m$, Lemma 3.54 yields

$$\begin{aligned}
\|f\|_{L^{\frac{m}{m-1}}(\mathbb{R}^m)}^{\frac{m}{m-1}} &\leqslant \int_{\mathbb{R}^m} g_1(\hat{x}^1) \cdot \ldots \cdot g_d(\hat{x}^m) \, \mathrm{d}x \leqslant \prod_{j=1}^m \|g_j\|_{L^{m-1}(\mathbb{R}^{m-1})} \\
&= \prod_{j=1}^m \left(\int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}} |\partial_j f(x)| \, \mathrm{d}x_j \, \mathrm{d}\hat{x}^j \right)^{\frac{1}{m-1}}, \\
\|f\|_{L^{\frac{m}{m-1}}(\mathbb{R}^m)} &\leqslant \prod_{j=1}^m \|\partial_j f\|_{L^1(\mathbb{R}^m)}^{\frac{1}{m}} \leqslant \||\nabla f|_1\|_1 \leqslant \|f\|_{1,1}. \qquad (3.37)
\end{aligned}$$

2) Next, let $p \in (1, m)$ and $p^\star = \frac{pm}{m-p}$. Set $t_\star = \frac{m-1}{m} p^\star = \frac{m-1}{m-p} p > 1$. An elementary calculation shows that $(t_\star - 1)p' = t_\star \frac{m}{m-1} = p^\star$. Set

$$g = f|f|^{t-1} = f(f\overline{f})^{\frac{t-1}{2}}$$

for $t > 1$. We compute

$$\begin{aligned}
\partial_j g &= \partial_j f |f|^{t-1} + f \frac{t-1}{2} (f\overline{f})^{\frac{t-1}{2}-1} \big((\partial_j f)\overline{f} + f(\partial_j \overline{f}) \big) \\
&= \partial_j f |f|^{t-1} + (t-1) f |f|^{t-3} \operatorname{Re}(f \partial_j \overline{f}), \\
|g| &= |f|^t, \quad |\partial_j g| \leqslant t |\partial_j f| |f|^{t-1}.
\end{aligned}$$

Applying (3.37) to g, we estimate

$$\begin{aligned}
\|f\|_{\frac{tm}{m-1}}^t &= \left(\int_{\mathbb{R}^m} |f|^{t\frac{m}{m-1}} \, \mathrm{d}x \right)^{\frac{m-1}{m}} = \left(\int_{\mathbb{R}^m} |g|^{\frac{m}{m-1}} \, \mathrm{d}x \right)^{\frac{m-1}{m}} \\
&\leqslant \prod_{j=1}^m \|\partial_j g\|_1^{\frac{1}{m}} \leqslant \prod_{j=1}^m t^{\frac{1}{m}} \left(\int_{\mathbb{R}^m} |\partial_j f| \, |f|^{t-1} \, \mathrm{d}x \right)^{\frac{1}{m}} \\
&\leqslant t \prod_{j=1}^m \Big(\int_{\mathbb{R}^m} |\partial_j f|^p \, \mathrm{d}x \Big)^{\frac{1}{pm}} \left(\int_{\mathbb{R}^m} |f|^{(t-1)p'} \, \mathrm{d}x \right)^{\frac{1}{p'm}} \\
&\leqslant t \prod_{j=1}^m \||\nabla f|_p\|_p^{\frac{1}{m}} \ \|f\|_{(t-1)p'}^{\frac{t-1}{m}} = t \, \||\nabla f|_p\|_p \ \|f\|_{(t-1)p'}^{t-1},
\end{aligned}$$

where we used Hölder's inequality. For $t = t_\star$, we use the properties of t stated above and obtain

$$\|f\|_{p^\star} \leqslant p \frac{m-1}{m-p} \||\nabla f|_p\|_p \leqslant p \frac{m-1}{m-p} \|f\|_{1,p}. \qquad (3.38)$$

This estimate can be extended to all $f \in W^{1,p}(\mathbb{R}^m)$ by density (see Theorem 3.51). Hence, the inclusion map is the required embedding in part a).

3) Let $p = m$, $f \in C_c^1(\mathbb{R}^m)$, and $t > 1$. Then $p' = \frac{m}{m-1}$, and step 2) yields

$$\|f\|_{t\frac{m}{m-1}} \leqslant t^{\frac{1}{t}} \|f\|_{(t-1)\frac{m}{m-1}}^{1-\frac{1}{t}} \||\nabla f|_p\|_m^{\frac{1}{t}} \leqslant c\left(\|f\|_{(t-1)\frac{m}{m-1}} + \||\nabla f|_p\|_m\right) \quad (3.39)$$

using Young's inequality. For $t = m$, this estimate gives $f \in L^{\frac{m^2}{m-1}}(\mathbb{R}^m)$ and

$$\|f\|_{\frac{m^2}{m-1}} \leqslant c\,\|f\|_{1,m}.$$

Here and below the constants $c > 0$ do not depend on f. For $q \in (m, m\frac{m}{m-1})$, inequality (3.36) further yields

$$\|f\|_q \leqslant c\,(\|f\|_m + \|f\|_{\frac{m^2}{m-1}}) \leqslant c\,\|f\|_{1,m}.$$

Now, we can apply (3.39) with $t = m+1$ and obtain

$$\|f\|_{\frac{m^2+m}{m-1}} \leqslant c\,(\|f\|_{\frac{m^2}{m-1}} + \||\nabla f|_p\|_m) \leqslant c\,\|f\|_{1,m}.$$

As above, we see that $f \in L^q(\mathbb{R}^m)$ for $m \leqslant q \leqslant m\frac{m+1}{m-1}$. We can then iterate this procedure with $t_n = m + n$ and obtain

$$\|f\|_q \leqslant c(q)\|f\|_{1,p}$$

for all $q < \infty$. As above, assertion b) follows by approximation.

4) Let $p > m$, $f \in C_c^1(\mathbb{R}^m)$, $Q(r) = [-\frac{r}{2}, \frac{r}{2}]^m$ for some $r > 0$, and $x_0 \in Q(r)$. We set $M(r) = r^{-m}\int_{Q(r)} f\,\mathrm{d}x$ and $\beta = 1 - \frac{m}{p} \in (0,1)$. Using $|x - x_0|_\infty \leqslant r$ for $x \in Q(r)$, the transformation $y = t(x - x_0)$ and Hölder's inequality, we compute

$$\begin{aligned}
|f(x_0) - M(r)| &= \left| r^{-m} \int_{Q(r)} (f(x_0) - f(x))\,\mathrm{d}x \right| \\
&= r^{-m} \left| \int_{Q(r)} \int_1^0 \frac{\mathrm{d}}{\mathrm{d}t} f(x_0 + t(x - x_0))\,\mathrm{d}t\,\mathrm{d}x \right| \\
&\leqslant r^{-m} \int_{Q(r)} \int_0^1 |\nabla f(x_0 + t(x - x_0)) \cdot (x - x_0)|\,\mathrm{d}t\,\mathrm{d}x \\
&\leqslant r^{1-m} \int_0^1 \int_{Q(r)} |\nabla f(x_0 + t(x - x_0))|_1\,\mathrm{d}x\,\mathrm{d}t \\
&= r^{1-m} \int_0^1 \int_{t(Q(r)-x_0)} |\nabla f(x_0 + y)|_1\,\mathrm{d}y\; t^{-m}\,\mathrm{d}t \\
&\leqslant r^{1-m} \int_0^1 \left[\int_{\mathbb{R}^m} |\nabla f(x_0 + y)|_1^p\,\mathrm{d}y \right]^{\frac{1}{p}} \lambda(t(Q(r) - x_0))^{\frac{1}{p'}} t^{-m}\,\mathrm{d}t \\
&\leqslant c r^{1-m} \||\nabla f|_p\|_p \int_0^1 r^{\frac{m}{p'}} t^{\frac{m}{p'}-m}\,\mathrm{d}t \\
&= C r^{1-\frac{m}{p}} \||\nabla f|_p\|_p
\end{aligned}$$

for constants $C, c > 0$ only depending on m and p, using also that $\frac{m}{p'} - m > -1$ due to $p > m$. A translation then gives

$$\left| f(x_0 + z) - r^{-m} \int_{z+Q(r)} f(y)\,\mathrm{d}y \right| \leqslant C r^{\beta} \||\nabla f|_p\|_p$$

for all $z \in \mathbb{R}^m$. Taking $x = z$, $x_0 = 0$ and $r = 1$, by means of Hölder's inequality we thus obtain

$$\begin{aligned} |f(x)| &\leqslant \Big| f(x) - \int_{x+Q(1)} f \,\mathrm{d}y \Big| + \Big| \int_{x+Q(1)} f \,\mathrm{d}y \Big| \\ &\leqslant C \| |\nabla f|_p \|_p + \|f\|_p \leqslant c \|f\|_{1,p} \end{aligned} \tag{3.40}$$

for all $x \in \mathbb{R}^m$, where c only depends on m and p. Given $x, y \in \mathbb{R}^m$, we find a cube Q of side length $|x - y|_\infty =: r$ such that $x, y \in Q$ and Q is parallel to the axes. Hence,

$$\begin{aligned} |f(x) - f(y)| &\leqslant \Big| f(x) - r^{-m} \int_Q f \,\mathrm{d}y \Big| + \Big| r^{-m} \int_Q f \,\mathrm{d}y - f(y) \Big| \\ &\leqslant 2C \| |\nabla f|_p \|_p \, |x - y|_\infty^\beta \leqslant 2C \| |\nabla f|_p \|_p \, |x - y|_2^\beta. \end{aligned}$$

Let $f \in W^{1,p}(\mathbb{R}^m)$. Then there are $f_n \in C_c^1(\mathbb{R}^m)$ converging to f in $W^{1,p}(\mathbb{R}^m)$. By (3.40), f_n is a Cauchy sequence in $C_0(\mathbb{R}^m)$. Hence, f has a representative $\tilde{f} \in C_0(\mathbb{R}^m)$ such that $f_n \to \tilde{f}$ uniformly as $n \to \infty$. So the above estimates imply that

$$\|\tilde{f}\|_\infty + \sup_{x \neq y} \frac{|\tilde{f}(x) - \tilde{f}(y)|}{|x - y|_2^\beta} \leqslant c \|f\|_{1,p}.$$

The map $f \mapsto \tilde{f}$ is the required embedding. □

Remark 3.55. Theorem 3.52 remains valied on any open set U instead of $\mathbb{R}^m$ if we replace $W^{k,p}$ by $W_0^{k,p}$. In fact, we obtain the desired estimate for $f \in C_c^\infty(U)$ if we apply Theorem 3.52 to the 0–extension of f. The result for $f \in W_0^{k,p}(U)$ then follows by density. ◊

Corollary 3.56. *Let $U \subseteq \mathbb{R}^m$ be open and bounded. We then have* Poincare's inequality

$$\int_U |\nabla u|_p^p \,\mathrm{d}x \geqslant \delta \int_U |u|^p \,\mathrm{d}x \tag{3.41}$$

for some $\delta > 0$ and all $u \in W_0^{1,p}(U)$ and $p \in [1, \infty)$.

Proof. For $p \in [1, m)$, the estimate (3.41) follows from (3.38) since $L^{p^*}(U) \hookrightarrow L^p(U)$. Let $p \in [m, \infty)$. The case $p = m = 1$ easily follows from (3.19). For the other cases, fix $r \in (p, \infty)$ and $u \in W_0^{1,p}(U)$. Then (3.36), (3.37) and $W_0^{1,p}(U) \hookrightarrow L^r(U)$ imply

$$\begin{aligned} \|u\|_p &\leqslant c_\varepsilon \|u\|_1 + \varepsilon \|u\|_r \leqslant c_\varepsilon \| |\nabla u|_1 \|_1 + c\varepsilon \|u\|_{1,p} \\ &\leqslant c_\varepsilon \| |\nabla u|_p \|_p + c\varepsilon \|u\|_p + c\varepsilon \| |\nabla u|_p \|_p \end{aligned}$$

for all $\varepsilon > 0$ and some constants $c_\varepsilon, c > 0$ independent of u, where c does not depend on $\varepsilon > 0$. Choosing a small ε, we derive (3.41). □

Remark 3.57. We show a part of Theorem 3.52 a) for the case $p = 2$ by means of the Fourier transform. We use the Hausdorff–Young inequality

$$\|\mathcal{F} f\|_q \leqslant c \, \|f\|_{q'} \qquad \text{for } \; q \in [2, \infty], \; f \in L^{q'}(\mathbb{R}^m), \tag{3.42}$$

see e.g. Satz V.2.10 in [**We**]. Let $f \in C_c^\infty(\mathbb{R}^m)$ and $k \in \mathbb{N}$. Then $\widehat{f}$ belongs to $\mathcal{S}_m \subseteq L^1(\mathbb{R}^m)$. The case $k > m/2$ was already treated in Remark 3.33.

Let $k < m/2$ and $2 \leqslant q < 2^\star = 2m(m-2k)^{-1}$. The latter is equivalent to

$$\frac{1}{q} > \frac{1}{2} - \frac{k}{m}.$$

To apply Hölder's inequality, we define the number $s \in (2, \infty]$ by

$$\frac{1}{s} = \frac{1}{q'} - \frac{1}{2} = \frac{1}{2} - \frac{1}{q} < \frac{k}{m}.$$

As in Remark 3.33, by means of (3.42) and Theorem 3.25 we estimate

$$\begin{aligned}\|f\|_q &\leqslant c\|(1+|\xi|_2^k)^{-1}(1+|\xi|_2^k)\widehat{f}\|_{q'} \leqslant c\|(1+|\xi|_2^k)^{-1}\|_s \,\|(1+|\xi|_2^k)\widehat{f}\|_2 \\ &\leqslant c\|f\|_{k,2}\Big(\int_0^\infty \frac{r^{m-1}}{(1+r^k)^s}\,\mathrm{d}r\Big)^{\frac{1}{s}} \leqslant c\|f\|_{k,2},\end{aligned}$$

where we have used that $sk > m$ by the above relations between the exponents. One can again conclude that $W^{k,2}(\mathbb{R}^m) \hookrightarrow L^q(\mathbb{R}^m)$, which is Theorem 3.52 a) for $p = 2$ and $q < 2^\star$. $\Diamond$

Most of the following results in this section are based on Stein's extension Theorem 3.28. We prove here a somewhat weaker result.

Remark 3.58. We show Theorem 3.28 for $W^{1,p}(U)$ with $p \in [1, \infty]$ if U is bounded and $\partial U \in C^1$.

1) We write elements in $\mathbb{R}^m_\pm$ as (y, t), where $\mathbb{R}^m_- = -\mathbb{R}^m_+$. For $f \in W^{1,p}(\mathbb{R}^m_-) \cap C^1(\overline{\mathbb{R}^m_-})$, we define

$$E_- f(y, r) = \begin{cases} f(y,t), & (y,t) \in \overline{\mathbb{R}^m_-}, \\ 4f(y, -\frac{t}{2}) - 3f(y, -t), & (y,t) \in \mathbb{R}^m_+. \end{cases}$$

Note that $E_- f$ belongs to $C^1(\mathbb{R}^m)$ and fulfills $\|E_- f\|_{W^{1,p}(\mathbb{R}^m)} \leqslant c\|f\|_{W^{1,p}(\mathbb{R}^m_-)}$ for a constant $c > 0$.

2) We show that $W^{1,p}(\mathbb{R}^m_-) \cap C^1(\overline{\mathbb{R}^m_-})$ is dense in $W^{1,p}(\mathbb{R}^m_-)$, so that E_- can be extended to an extension operator on $W^{1,p}(\mathbb{R}^m_-)$. In fact, let $f \in W^{1,p}(\mathbb{R}^m_-)$ and $\varepsilon > 0$. Theorem 3.51 yields a function g in $C^\infty(\mathbb{R}^m_-) \cap W^{1,p}(\mathbb{R}^m_-)$ with $\|f - g\|_{1,p} \leqslant \varepsilon$. Setting $g_n(y,t) = g(y, t - \frac{1}{n})$ for $t \leqslant 0$, $y \in \mathbb{R}^{m-1}$ and $n \in \mathbb{N}$, we define maps g_n in $C^1(\overline{\mathbb{R}^m_-}) \cap W^{1,p}(\mathbb{R}^m_-)$. Note that

$$\partial^\alpha g_n = R_{\mathbb{R}^m_-} S_n E_0 \partial^\alpha g$$

for $0 \leqslant |\alpha| \leqslant 1$, where $S_n \in \mathcal{B}(L^p(\mathbb{R}^m))$ is given by $S_n h(y,t) = h(y, t - \frac{1}{n})$ for $h \in L^p(\mathbb{R}^m)$. One can see that $S_n h \to h$ in $L^p(\mathbb{R}^m)$ as in Example 4.12 of [**FA**]. Hence, g_n converges to g in $W^{1,p}(\mathbb{R}^m_-)$ implying the claim.

3) Since $\partial U \in C^1$ and U is bounded, there are bounded open subsets $U_0, U_1, \dots, U_N$ of $\mathbb{R}^m$ such that $U \subseteq U_0 \cup \dots \cup U_N$, $\overline{U}_0 \subseteq U$, and $\partial U \subseteq U_1 \cup \dots \cup U_N$, as well as diffeomorphisms $\Psi_j : U_j \to V_j$ such that Ψ_j' and $(\Psi_j^{-1})'$ are bounded, $\Psi_j(U_j \cap U) \subseteq \mathbb{R}^m_-$, and $\Psi_j(U_j \cap \partial U) \subseteq \mathbb{R}^{m-1} \times \{0\}$, for each $j \in \{1, \dots, N\}$. By Analysis 3 we have functions $0 \leqslant \varphi_j \in C_c^\infty(\mathbb{R}^m)$ with $\operatorname{supp} \varphi_j \subseteq U_j$ for all $j \in \{0, 1, \dots, N\}$ and $\sum_{j=0}^m \varphi_j(x) = 1$ for all $x \in \overline{U}$.

Let $j \in \{1, \dots, N\}$. Set $S_j g(y) = g(\Psi_j^{-1}(y))$ for $y \in \mathbb{R}^m_- \cap V_j$ and $S_j g(y) = 0$ for $y \in \mathbb{R}^m_- \backslash V_j$, where $g \in W^{1,p}(U_j \cap U)$. For $h \in W^{1,p}(\mathbb{R}^m)$, set $\hat{S}_j h(x) = h(\Psi_j(x))$ for $x \in U_j$ and $\hat{S}_j h(x) = 0$ for $x \in \mathbb{R}^m \backslash U_j$. Take any $\tilde{\varphi}_j \in C_c^\infty(\mathbb{R}^m)$

with $\operatorname{supp}\tilde{\varphi}_j \subseteq U_j$ and $\tilde{\varphi}_j = 1$ on $\operatorname{supp}\varphi_j$ (see Lemma 3.50). Let $f \in W^{1,p}(U)$. We now define

$$E_1 f = E_0\varphi_0 f + \sum_{j=1}^{m} \tilde{\varphi}_j \hat{S}_j E_- S_j\big(R_{(U_j\cap U)}(\varphi_j f)\big).$$

Using part 2) and Propositions 3.19 and 3.20, we see that E_1 belongs to $\mathcal{B}(W^{1,p}(U), W^{1,p}(\mathbb{R}^m))$. Let $x \in U$. If $x \in U_k$ for some $k \in \{1, \dots, N\}$, we have $\Psi_k(x) \in \mathbb{R}^m_-$. If $x \notin U_j$, then $\tilde{\varphi}_j(x) = 0$. It follows

$$\begin{aligned} E_1 f(x) &= \varphi_0(x)f(x) + \sum_{1\leqslant j\leqslant N,\, x\in U_j} \tilde{\varphi}_j(x)(\varphi_j f)(\Psi_j^{-1}(\Psi_j(x))) \\ &= \sum_{j=0}^{N} \varphi_j(x)f(x) = f(x). \end{aligned}$$

If $x \in U_0\backslash(U_1 \cup \dots \cup U_N)$, we also have $E_1 f(x) = \varphi_0 f(x) = f(x)$. $\Diamond$

Using Theorem 3.28, we can easily extend the above embedding and density results to $U \neq \mathbb{R}^m$.

Theorem 3.59. *Let $\partial U \in C^{1-}$. Theorem 3.52 and Corollary 3.53 then remain true if we replace $\mathbb{R}^m$ by U.*

Proof. Consider e.g. Theorem 3.52 a). We have the embedding

$$J : W^{k,p}(\mathbb{R}^m) \hookrightarrow L^{p^\star}(\mathbb{R}^m)$$

given by the inclusion. Thanks to Theorem 3.28, the map

$$R_U J E_U : W^{k,p}(U) \to L^{p^\star}(U)$$

is continuous and injective. The other assertions are proved similarly. $\square$

Theorem 3.60. *Let $\partial U \in C^{1-}$, $k \in \mathbb{N}$, and $p \in [1, \infty)$. Then $C_c^\infty(\overline{U})$ is dense in $W^{k,p}(U)$.*

Proof. Let $f \in W^{k,p}(U)$. Then $E_U f$ belongs to $W^{k,p}(\mathbb{R}^m)$ by Theorem 3.28. Theorem 3.51 yields functions g_n in $C_c^\infty(\mathbb{R}^m)$ that converge to $E_U f$ in $W^{k,p}(\mathbb{R}^m)$. Hence, $R_U g_n$ is contained in $C_c^\infty(\overline{U}) \subseteq W^{k,p}(U)$ and tends to $f = R_U E_U f$ in $W^{k,p}(U)$ as $n \to \infty$. $\square$

We continue with one of the main compactness results in analysis, due to *Rellich* and *Kondrachov.*

Theorem 3.61. *Let $U \subseteq \mathbb{R}^m$ be bounded and $\partial U \in C^{1-}$, $k \in \mathbb{N}$, and $1 \leqslant p < \infty$. Then the following assertions hold.*

a) Let $kp \leqslant m$ and $1 \leqslant q < p^\star = \frac{mp}{m-kp} \in (p, \infty]$. Then the inclusion map

$$J : W^{k,p}(U) \hookrightarrow L^q(U)$$

is compact. (For instance, let $q = p$.)

b) Let $k - \frac{m}{p} > j \in \mathbb{N}_0$. Then the embedding

$$J : W^{k,p}(U) \hookrightarrow C^j(\overline{U})$$

is compact, where Jf is the continuous representative.

PROOF. We prove the result only for $k = 1$ (and thus $j = 0$), see Theorem 6.3 of [**AF**] for the other cases. Part b) follows from the Arzela–Ascoli theorem since Theorem 3.59 gives constants $\beta, c > 0$ such that $|f(x) - f(y)| \leqslant c|x-y|^\beta$ and $|f(x)| \leqslant c$ for all $x, y \in U$ and $f \in W^{1,p}(U)$ with $\|f\|_{1,p} \leqslant 1$, where $p > m$.

In the case $p < m$, take $f_n \in W^{1,p}(U)$ with $\|f_n\|_{1,p} \leqslant 1$ for all $n \in \mathbb{N}$. Fix an open bounded set $V \subseteq \mathbb{R}^m$ containing $\overline{U}$. Lemma 3.50 yields a function $\varphi \in C_c^\infty(V)$ which is equal to 1 on $\overline{U}$. Let E_U be given by Theorem 3.28. Set $g_n = \varphi E_U f_n \in W^{1,p}(\mathbb{R}^m)$. These functions have support V and $\|g_n\|_{1,p} \leqslant c\,\|\varphi\|_{1,\infty} \|E_U\| =: M$ for all $n \in \mathbb{N}$. Fix $q \in [1, p^\star)$ and take $\theta \in (0,1]$ with $\frac{1}{q} = \frac{\theta}{1} + \frac{1-\theta}{p^\star}$. Inequality (3.36) and Theorem 3.52 yield that

$$\begin{aligned}
\|f_n - f_m\|_{L^q(U)} &\leqslant \|g_n - g_m\|_{L^q(V)} \leqslant \|g_n - g_m\|_{L^1(V)}^\theta \|g_n - g_m\|_{L^{p^\star}(V)}^{1-\theta} \\
&\leqslant c\|g_n - g_m\|_{L^1(V)}^\theta (\|g_n\|_{1,p}^{1-\theta} + \|g_m\|_{1,p}^{1-\theta}) \\
&\leqslant 2cM^{1-\theta} \|g_n - g_m\|_{L^1(V)}^\theta
\end{aligned}$$

for all $n, m \in \mathbb{N}$. So it suffices to construct a subsequence of g_n which converges in $L^1(V)$. For $x \in V$, $n \in \mathbb{N}$ and $\varepsilon > 0$, we compute

$$\begin{aligned}
|g_n(x) - G_\varepsilon g_n(x)| &= \left| \int_{\mathbb{R}^m} k_\varepsilon(x-y)(g_n(x) - g_n(y))\,\mathrm{d}y \right| \\
&\leqslant \varepsilon^{-m} \int_{B(x,\varepsilon)} k(\tfrac{1}{\varepsilon}(x-y))\,|g_n(x) - g_n(y)|\,\mathrm{d}y \\
&= \int_{B(0,1)} k(z)|g_n(x) - g_n(x - \varepsilon z)|\,\mathrm{d}z \\
&= \int_{B(0,1)} k(z) \left| \int_0^\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} g_n(x - tz)\,\mathrm{d}t \right| \mathrm{d}z \\
&\leqslant \int_{B(0,1)} k(z) \int_0^\varepsilon |\nabla g_n(x - tz) \cdot z|\,\mathrm{d}t\,\mathrm{d}z \\
&\leqslant \int_0^\varepsilon \int_{B(0,1)} k(z)\,|\nabla g_n(x - tz)|_2\,\mathrm{d}z\,\mathrm{d}t \\
&= \int_0^\varepsilon \int_{B(x,t)} t^{-m} k(\tfrac{1}{t}(x-y))\,|\nabla g_n(y)|_2\,\mathrm{d}y\,\mathrm{d}t, \\
\|g_n - G_\varepsilon g_n\|_{L^1(V)} &\leqslant \int_0^\varepsilon \|k_t * |\nabla g_n|_2\|_{L^1(\mathbb{R}^m)}\,\mathrm{d}t \leqslant \varepsilon \sup_{0 \leqslant t \leqslant \varepsilon} \|k_t\|_1\,\||\nabla g_n|_2\|_{L^1(\mathbb{R}^m)} \\
&\leqslant c\varepsilon \||\nabla g_n|_p\|_p \leqslant cM\varepsilon =: C\varepsilon, \qquad (3.43)
\end{aligned}$$

where we have used the transformations $z = \frac{1}{\varepsilon}(x-y)$ and $y = x - tz$, as well as Fubini's theorem, Young's inequality (3.5), and $L^p(V) \hookrightarrow L^1(V)$.

On the other hand, the definition of $G_\varepsilon g_n$ yields

$$|G_\varepsilon g_n(x)| \leqslant \|k_\varepsilon\|_\infty \|g_n\|_{L^1(V)} \qquad \text{and} \qquad |\nabla G_\varepsilon g_n(x)| \leqslant \|\nabla k_\varepsilon\|_\infty \|g_n\|_{L^1(V)}$$

for all $x \in V$, $n \in \mathbb{N}$ and each fixed $\varepsilon > 0$. The Arzela-Ascoli theorem now implies that the set $F_\varepsilon := \{G_\varepsilon g_n \,|\, n \in \mathbb{N}\}$ is relatively compact in $C(\overline{V})$ for each $\varepsilon > 0$, and thus in $L^1(V)$ since $C(\overline{V}) \hookrightarrow L^1(V)$. Let $\delta > 0$ be given and fix $\varepsilon = \frac{\delta}{2C}$. Then there are indeces $n_1, \dots, n_l \in \mathbb{N}$ such that

$$F_\varepsilon \subseteq \bigcup_{j=1}^{l} B_{L^1(V)}(G_\varepsilon g_{n_j}, \tfrac{\delta}{2}) =: \bigcup_{j=1}^{l} B_j.$$

Hence, given $n \in \mathbb{N}$, there is an index n_j such that $G_\varepsilon g_n \in B_j$. The estimates (3.43) and (3.16) then yield

$$\|g_n - G_\varepsilon g_{n_j}\|_{L^1(V)} \leqslant \|g_n - G_\varepsilon g_n\|_{L^1(V)} + \|G_\varepsilon(g_n - g_{n_j})\|_{L^1(V)} \leqslant C\varepsilon + \delta/2 = \delta.$$

We have shown that, for each $\delta > 0$, the set $G := \{g_n \,|\, n \in \mathbb{N}\}$ is covered by finitely many open balls B_j of radius δ; i.e., G is totally bounded in $L^1(V)$. Thus G contains a subsequence converging in $L^1(V)$ (see Corollary 1.39 in [**FA**]). In the case $p = m$ one replaces $p^\star$ by any $r \in (q, \infty)$. □

We can now give a proof of the second Poincaré inequality, which we repeat for convenience.

Theorem 3.62. *Let U be bounded and pathwise connected with $\partial U \in C^{1-}$ and $p \in [1, \infty)$. We then have $\|f - a_U(f)\mathbb{1}\|_p \leqslant c\| |\nabla f|_p\|_p$ for all $f \in W^{1,p}(U)$ and some $c > 0$.*

Proof. We show the estimate via contradiction. So assume that for each $n \in \mathbb{N}$ the are functions $u_n \in W^{1,p}(U)$ with $a_n := \|u_n - a_U(u_n)\mathbb{1}\|_p > n\||\nabla u_n|_p\|_p$. We normalize u_n to $v_n := a_n^{-1}(u_n - a_U(u_n)\mathbb{1}) \in W^{1,p}(U)$. Observe that

$$a_U(v_n) = 0, \qquad \|v_n\|_p = 1, \qquad \||\nabla v_n|_p\|_p < \tfrac{1}{n}.$$

In particular, the sequence (v_n) is bounded in $W^{1,p}(U)$ and so a subsequence $(v_{n_k})_k$ converges to some v in $L^p(U)$ by Theorem 3.34. This function then satisfies $a_U(v) = 0$ and $\|v\|_p = 1$. Since $\partial_j v_{n_k} \to 0$ in $L^p(U)$ as $k \to \infty$ for each $j \in \{1, \dots, m\}$, we also obtain $\partial_j v = 0$ by Lemma 3.16. Because of the pathwise connectedness, an exercise shows that $v = \kappa\mathbb{1}$ for some $\kappa \in \mathbb{F}$ and hence $0 = a_U(v) = \kappa$. But the conclusion $v = 0$ contradicts $\|v\|_p = 1$. □

We next present an interpolation estimate for first derivatives. Again there are plenty of variants. It is not so easy to generalize the last part to U with $\partial U \in C^{1-}$ since our extension and density results involve first derivatives, but see Corollary 1.1.11 in [**Ma**].

Proposition 3.63. *Let $1 \leqslant p < \infty$. Let either $f \in W_0^{2,p}(U)$ or $\partial U \in C^1{-}1$ and $f \in W^{2,p}(U)$. Then there are constants $C, \varepsilon_0 > 0$ such that*

$$\Big(\sum_{j=1}^{m} \|\partial_j f\|_p^p\Big)^{1/p} \leqslant \varepsilon \Big(\sum_{i,j=1}^{m} \|\partial_{ij} f\|_p^p\Big)^{1/p} + \frac{C}{\varepsilon}\,\|f\|_p, \tag{3.44}$$

for all $\varepsilon > 0$ if $f \in W_0^{2,p}(U)$ and for all $0 < \varepsilon \leqslant \varepsilon_0$ if $f \in W^{2,p}(U)$. Moreover, a function f in $L^p(\mathbb{R}^m) \cap \bigcap_{|\alpha|=2} W_\alpha(\mathbb{R}^m)$ with $\partial^\alpha f \in L^p(\mathbb{R}^m)$ for all $|\alpha| = 2$ already belongs to $W^{2,p}(\mathbb{R}^m)$.

Proof. 1) Let $f \in C_c^2(U)$ and extend it to $\mathbb{R}^m$ by 0. Take $j = 1$. Write $x = (t, y) \in \mathbb{R} \times \mathbb{R}^{m-1}$ for $x \in \mathbb{R}^m$. Fix $y \in \mathbb{R}^{m-1}$ and set $g(t) = f(t, y)$ for

$t \in \mathbb{R}$. Let $\varepsilon > 0$ and $a, b \in \mathbb{R}$ with $b - a = \varepsilon$. Take any $r \in (a, a + \frac{\varepsilon}{3})$ and $t \in (b - \frac{\varepsilon}{3}, b)$. There there is a number $\overline{s} = \overline{s}(r,t) \in (a,b)$ such that

$$|g'(\overline{s})| = \Big|\frac{g(t) - g(r)}{t - r}\Big| \leqslant \frac{3}{\varepsilon}(|g(t)| + |g(r)|).$$

For every $s \in (a,b)$ we thus obtain

$$|g'(s)| = \Big|g'(\overline{s}) + \int_{\overline{s}}^{s} g''(\tau)\,\mathrm{d}\tau\Big| \leqslant \frac{3}{\varepsilon}(|g(r)| + |g(t)|) + \int_a^b |g''(\tau)|\,\mathrm{d}\tau.$$

Integrating first over r and then over t, we conclude

$$\begin{aligned}
\frac{\varepsilon}{3}|g'(s)| &\leqslant \frac{3}{\varepsilon}\int_a^{a+\frac{\varepsilon}{3}} |g(r)|\,\mathrm{d}r + |g(t)| + \frac{\varepsilon}{3}\int_a^b |g''(\tau)|\,\mathrm{d}\tau,\\
\frac{\varepsilon^2}{9}|g'(s)| &\leqslant \int_a^{a+\frac{\varepsilon}{3}} |g(r)|\,\mathrm{d}r + \int_{b-\frac{\varepsilon}{3}}^{b} |g(t)|\,\mathrm{d}t + \frac{\varepsilon^2}{9}\int_a^b |g''(\tau)|\,\mathrm{d}\tau,\\
|g'(s)| &\leqslant \frac{9}{\varepsilon^2}\int_a^b |g(\tau)|\,\mathrm{d}\tau + \int_a^b |g''(\tau)|\,\mathrm{d}\tau\\
&\leqslant \varepsilon^{\frac{1}{p'}}\frac{9}{\varepsilon^2}\Big(\int_a^b |g(\tau)|^p\,\mathrm{d}\tau\Big)^{\frac{1}{p}} + \varepsilon^{\frac{1}{p'}}\Big(\int_a^b |g''(\tau)|^p\,\mathrm{d}\tau\Big)^{\frac{1}{p}}\\
&\leqslant \varepsilon^{\frac{p-1}{p}}2^{\frac{p-1}{p}}\left(\Big(\frac{9}{\varepsilon^2}\Big)^p\int_a^b |g(\tau)|^p\,\mathrm{d}\tau + \int_a^b |g''(\tau)|\,\mathrm{d}\tau\right)^{\frac{1}{p}},
\end{aligned}$$

where we used Hölder's inequality first for the integrals and then in $\mathbb{R}^2$. We take now the p-th power and then integrate over s arriving at

$$\int_a^b |g'(s)|^p\,\mathrm{d}s \leqslant \varepsilon\varepsilon^{p-1}2^{p-1}\left(\frac{9^p}{\varepsilon^{2p}}\int_a^b |g(\tau)|^p\,\mathrm{d}\tau + \int_a^b |g''(\tau)|^p\,\mathrm{d}\tau\right).$$

Now choose $a = a_k = k\varepsilon$ and $b = b_k = (k+1)\varepsilon$ for $k \in \mathbb{Z}$. Summing the integrals on $[k\varepsilon, (k+1)\varepsilon)$ for $k \in \mathbb{Z}$ and then integrating over $y \in \mathbb{R}^{m-1}$, it follows that

$$\begin{aligned}
\int_{\mathbb{R}} |g'(\tau)|^p\,\mathrm{d}\tau &\leqslant \varepsilon^p 2^{p-1}\left(\frac{9^p}{\varepsilon^{2p}}\int_{\mathbb{R}} |g(\tau)|^p\,\mathrm{d}\tau + \int_{\mathbb{R}} |g''(\tau)|^p\,\mathrm{d}\tau\right),\\
\int_U |\partial_1 f|^p\,\mathrm{d}x &\leqslant (2\varepsilon)^p\int_U |\partial_{11} f|^p\,\mathrm{d}x + \frac{36^p}{(2\varepsilon)^p}\int_U |f|^p\,\mathrm{d}x. \qquad (3.45)
\end{aligned}$$

2) By approximation, (3.45) can be established for all $f \in W_0^{2,p}(U)$. The same result holds for $\partial_j f$ and $\partial_{jj} f$ with $j \in \{2, \ldots, m\}$. We now replace 2ε by ε, sum over j and take the p-th root to arrive at

$$\begin{aligned}
\Big(\sum_{j=1}^m \|\partial_j f\|_p^p\Big)^{\frac{1}{p}} &\leqslant \Big(\varepsilon^p\sum_{j=1}^m \|\partial_{jj} f\|_p^p + \frac{36^p}{\varepsilon^p}\|f\|_p^p\Big)^{\frac{1}{p}}\\
&\leqslant \varepsilon\Big(\sum_{j=1}^m \|\partial_{jj} f\|_p^p\Big)^{\frac{1}{p}} + \frac{36}{\varepsilon}\|f\|_p, \qquad (3.46)
\end{aligned}$$

for all $f \in W_0^{2,p}(U)$, as asserted.

3) Let $u \in W^{2,p}(U)$ and $\partial U \in C^{1-}$. The extension operator E_U in $\mathcal{B}(W^{2,p}(U), W^{2,p}(\mathbb{R}^m))$ from Theorem 3.28 and (3.46) with $U = \mathbb{R}^m$ imply

$$\begin{aligned}
\Big(\sum_{j=1}^m \|\partial_j f\|_{L^p(U)}^p\Big)^{\frac{1}{p}} &\leqslant \Big(\sum_{j=1}^m \|\partial_j E_U f\|_{L^p(\mathbb{R}^m)}^p\Big)^{\frac{1}{p}} \\
&\leqslant \varepsilon \Big(\sum_{j=1}^m \|\partial_{jj} E_U f\|_{L^p(\mathbb{R}^m)}^p\Big)^{\frac{1}{p}} + \frac{36}{\varepsilon}\,\|E_U f\|_{L^p(\mathbb{R}^m)} \\
&\leqslant \varepsilon \|E_U f\|_{W^{2,p}(\mathbb{R}^m)} + \frac{36}{\varepsilon}\,\|E_U f\|_{L^p(U)} \\
&\leqslant c\varepsilon \|f\|_{W^{2,p}(U)} + \frac{c}{\varepsilon}\|f\|_{L^p(U)} \\
&\leqslant c_0\varepsilon \Big(\sum_{i,j=1}^m \|\partial_{ij} f\|_p^p\Big)^{\frac{1}{p}} + c_1\varepsilon \Big(\sum_{j=1}^m \|\partial_j f\|_p^p\Big)^{\frac{1}{p}} + \frac{c}{\varepsilon}\,\|f\|_p
\end{aligned}$$

where we assume that $\varepsilon \in (0,1]$ and the constants c, c_0, c_1 only depend on $\|E_U\|$ and m. Choosing $\varepsilon_1 = \min\{\frac{1}{2c_1}, 1\}$ we arrive at

$$\frac{1}{2}\Big(\sum_{j=1}^m \|\partial_j f\|_p^p\Big)^{\frac{1}{p}} \leqslant c_0\varepsilon \Big(\sum_{i,j=1}^m \|\partial_{ij} f\|_p^p\Big)^{\frac{1}{p}} + \frac{c}{\varepsilon}\|f\|_p$$

if $0 < \varepsilon \leqslant \varepsilon_1$. This inequality implies (3.44) if $\partial U \in C^{1-}$, after replacing ε by $\varepsilon/(2c_0)$ and ε_1 by $\varepsilon_0 = \min\{c_0/c_1, 2c_0\}$.

4) Let $f, \partial^\alpha f \in L^p(\mathbb{R}^m)$ for $|\alpha| = 2$. Set $f_n = G_{1/n} f \in C^\infty(\mathbb{R}^m)$ for $n \in \mathbb{N}$. Then f_n and $\partial^\alpha f_n$ tend to f and $\partial^\alpha f$ in $L^p(\mathbb{R}^m)$. Repeating steps 1) and 2), for f_n we first see as in (3.45) that $\partial_j f_n$ belongs to $L^p(\mathbb{R}^m)$ for all j (since the right-hand side is finite) and then derive (3.46) for f_n and $U = \mathbb{R}^m$. Hence, the sequence $(\partial_j f_n)_n$ is Cauchy in $L^p(\mathbb{R}^m)$ so that $\partial_j f$ exists as its limit in $L^p(\mathbb{R}^m)$ by Lemma 3.16. Estimate (3.45) also follows for f. □

We next show the *trace theorem* which extends the trace map $f \mapsto f{\restriction}_{\partial U}$ from $W^{1,p}(U) \cap C(\overline{U})$ to $W^{1,p}(U)$ and shows the important fact that $W_0^{1,p}(U)$ is the space of functions in $W^{1,p}(U)$ with trace 0.

Theorem 3.64. *Let $p \in [1,\infty)$ and $U \subseteq \mathbb{R}^m$ be bounded with $\partial U \in C^1$. Then the trace map $f \mapsto f{\restriction}_{\partial U}$ from $W^{1,p}(U) \cap C(\overline{U})$ to $L^p(\partial U, \sigma)$ has a bounded linear extension* $\operatorname{tr} : W^{1,p}(U) \to L^p(\partial U, \sigma)$ *whose kernel is $W_0^{1,p}(U)$, where σ is the surface measure on ∂U.*

Proof. 1) Let $u \in C^1(\overline{U})$. By the definition of the surface integral, see Section 2.5 with a slightly different notation, there are finitely many diffeomorphisms $\Psi_j : U_j \to V_j$ and $\varphi_j \in C_c^1(U_j)$ with $0 \leqslant \varphi_j \leqslant 1$ such that $\|u\|_{L^p(\partial U,\sigma)}^p$ is dominated by

$$c \sum_{j=1}^m \int_{V_{j0}} \varphi_j \circ \Psi_j^{-1}\, |u \circ \Psi_j^{-1}|^p \, \mathrm{d}y'$$

where U_j and V_j are open subsets of $\mathbb{R}^m$, the sets U_j cover ∂U, the maps φ_j form a partition of unity subordinated to U_j, $V_{j0} := \{(y', y_m) \in V_j \,|\, y_m = 0\}$, $V_{j+} := \{(y', y_m) \in V_j \,|\, y_m > 0\}$, $\Psi_j(U_j \cap \partial U) = V_{j0}$, and $\Psi_j(U_j \cap U) = V_{j+}$.

We set $v = u \circ \Psi_j^{-1}$ and $\psi = \varphi_j \circ \Psi_j^{-1} \in C_c^1(V_j)$ and drop the indices j below. By means of Fubini's theorem and the fundamental theorem of calculus, we compute

$$\begin{aligned}\int_{V_0} \psi\, |v(y')|^p \,\mathrm{d}y' &= -\int_{V_+} \partial_m(\psi\, |v|^p)\,\mathrm{d}y \\ &= -\int_{V_+} [(\partial_m\psi)\, |v|^p + p\psi|v|^{p-2}\operatorname{Re}(\overline{v}\partial_m v)]\,\mathrm{d}y \\ &\leqslant c\int_{V_+} [|v|^p + |v|^{p-1}\,|\partial_m v|]\,\mathrm{d}y \leqslant c\|v\|_p^p + c\,\|v\|_p^{p-1}\,\|\partial_m v\|_p \\ &\leqslant c(\|v\|_p^p + \|\partial_m v\|_p^p) \leqslant c\|v\|^p_{W^{1,p}(V_+)} \leqslant c\|u\|^p_{W^{1,p}(U)}.\end{aligned}$$

Here we also used Hölder's and Young's inequality and Proposition 3.20. As a result, the map $\operatorname{tr} : (C^1(\overline{U}), \|\cdot\|_{1,p}) \to L^p(\partial U, \sigma)$; $\operatorname{tr} u = u|_{\partial U}$, is continuous. Theorem 3.60 allows us to extend tr to an operator in $\mathcal{L}(W^{1,p}(U), L^p(\partial U, \sigma))$. If we start with a function $u \in W^{1,p}(U) \cap C(\overline{U})$, then we can construct approximations $u_n \in C^1(\overline{U})$ which converge to u in $W^{1,p}(U)$ and in $C(\overline{U})$, see the proof of Theorem 5.3.3 in [**Ev**]. Hence, $\operatorname{tr} u_n = u_n\restriction_{\partial U}$ tends to $u\restriction_{\partial U}$ uniformly on ∂U and to $\operatorname{tr} u$ in $L^p(\partial U, \sigma)$, so that $\operatorname{tr} u = u\restriction_{\partial U}$.

2a) We next observe that the inclusion $W_0^{1,p}(U) \subseteq \mathrm{N}(\mathrm{tr})$ is a consequence of the continuity of tr since tr vanishes on $C_c^\infty(U)$ and this space is dense in $W_0^{1,p}(U)$ by definition. To prove the converse, we start with the model case that $v \in W^{1,p}(V_+)$ has a compact support in V_+ and $\operatorname{tr} v = 0$. Theorem 3.60 yields functions $v_n \in C^1(\overline{V_+})$ converging to v in $W^{1,p}(V_+)$, and hence $\operatorname{tr} v_n = v_n\restriction_{V_0} \to 0$ in $L^p(V_0)$, as $n \to \infty$. Observe that

$$\begin{aligned}|v_n(y', y_m)| &\leqslant |v_n(y', 0)| + \int_0^{y_m} |\partial_m v_n(y', s)|\,\mathrm{d}s, \\ |v_n(y', y_m)|^p &\leqslant 2|v_n(y', 0)|^p + 2\Big(\int_0^{y_m} |\partial_m v_n(y', s)|\,\mathrm{d}s\Big)^p\end{aligned}$$

for $y' \in V_0$ and $y_m \geqslant 0$. Integrating over y' und employing Hölder's inequality, we obtain

$$\int_{V_0} |v_n(y', y_m)|^p\,\mathrm{d}y' \leqslant 2\int_{V_0} |v_n(y', 0)|^p\,\mathrm{d}y' + 2y_m^{p-1}\int_{V_0}\int_0^{y_m} |\partial_m v_n(y', s)|^p\,\mathrm{d}s\,\mathrm{d}y'.$$

By Fubini's theorem, $v_n(\cdot, x_m)$ tends to $v(\cdot, x_m)$ in $L^p(V_0)$ for a.e. $y_m > 0$ and thus for pointwise a.e., after possibly passing to subsequence. Using Fatou's lemma, we can now let $n \to \infty$ and arrive at

$$\int_{V_0} |v(y', y_m)|^p\,\mathrm{d}y' \leqslant 2y_m^{p-1}\int_{V_0}\int_0^{y_m} |\partial_m v(y', s)|^p\,\mathrm{d}s\,\mathrm{d}y' \tag{3.47}$$

for a.e. $y_m > 0$.

We next use a cut-off argument to obtain a support in the interior of V_+. Choose a function $\chi \in C^\infty(\mathbb{R}_{\geqslant 0})$ such that $\chi = 0$ on $[0, 1]$ and $\chi = 1$ on $[2, \infty)$. Set $\chi_n(s) = \chi(ns)$ for $s \geqslant 0$ and $n \in \mathbb{N}$, and define $w_n = \chi_n v$ on V_+. Note that $w_n \to v$ in $L^p(V_+)$ as $n \to \infty$, $\partial_j w_n = \chi_n \partial_j v$ for $j \in \{1, \ldots, m-1\}$

and $\partial_m w_n = \chi_n \partial_m v + n\chi'(n\cdot)v$. Estimate (3.47) then implies

$$\begin{aligned}
&\int_{V_+} |\nabla v - \nabla w_n|_p^p \,\mathrm{d}y \\
&\quad\leqslant c\int_0^{2/n}\int_{V_0} |1-\chi_n|^p\, |\nabla v|_p^p \,\mathrm{d}y'\,\mathrm{d}s + cn^p \int_0^{2/n}\int_{V_0} |v(y',s)|^p \,\mathrm{d}y'\,\mathrm{d}s \\
&\quad\leqslant c\int_0^{2/n}\int_{V_0} |\nabla v|_p^p \,\mathrm{d}y'\,\mathrm{d}s + cn^p \int_0^{2/n} s^{p-1}\int_0^s\int_{V_0} |\partial_m v(y',\tau)|^p \,\mathrm{d}y'\,\mathrm{d}\tau\,\mathrm{d}s \\
&\quad\leqslant c\int_0^{2/n}\int_{V_0} |\nabla v|_p^p \,\mathrm{d}y'\,\mathrm{d}s + c\int_0^{2/n}\int_{V_0} |\partial_m v(y',\tau)|^p \,\mathrm{d}y'\,\mathrm{d}\tau
\end{aligned}$$

for some constants $c > 0$. Because of $v \in W^{1,p}(V_+)$, the above integrals tend to 0 as $n \to \infty$, and so $w_n \to v$ in $W^{1,p}(V_+)$ as $n \to \infty$. Since $w_n = 0$ for $y_m \in (0, 1/n]$, we can mollify w_n to obtain a function $\widehat{w}_n \in C_c^\infty(V_+)$ such that $\|\widehat{w}_n - w_n\|_{1,p} \leqslant 1/n$. This means that $\widehat{w}_n \to v$ in $W^{1,p}(V_+)$ as $n \to \infty$.

2b) We come back to $u \in W^{1,p}(U)$ and consider the sets U_j and V_j and the functions Ψ_j and φ_j from step 1). Let $v_j = (\varphi_j u) \circ \Psi_j^{-1}$. First, observe that the trace of v_j to the set V_{j0} is given by $(\operatorname{tr}\varphi_j) \circ \Psi_j^{-1}\,(\operatorname{tr} u) \circ \Psi_j^{-1}$ if $u \in C(\overline{U})$, in addition. By continuity one can extend this identity to all $u \in W^{1,p}(U)$. Let $\operatorname{tr} u = 0$. Then we can apply part 2a) to v_j and obtain $\widehat{w}_n^j \in C_c^1(V_{j+})$ converging to v_j in $W^{1,p}(V_{j+})$. The function

$$\widehat{u}_n = \sum_{j=1}^m \widehat{w}_n^j \circ (\Psi_j{\restriction}_{U\cap U_j})$$

thus belongs to $C_c^1(U)$ and converges to u in $W^{1,p}(U)$ as $n \to \infty$. Since $\widehat{u}_n$ has compact support, we can mollify $\widehat{u}_n$ to a function $u_n \in C_c^\infty(U)$ with $\|\widehat{u}_n - u_n\|_{1,p} \leqslant 1/n$. This means that $u_n \to u$ in $W^{1,p}(U)$ as $n \to \infty$, and hence $u \in W_0^{1,p}(U)$. □

3.6. Appendix: Tempered distributions

The[9] theory of distributions allows to define derivatives of any order for rather general objects such as locally integrable functions or measures on open subsets of $\mathbb{R}^m$, see e.g. [**Ru2**]. Here we only discuss the subclass of tempered distributions on $\mathbb{R}^m$ to which one can extend the Fourier transform in a very convenient way. We let $\mathbb{F} = \mathbb{C}$.

Definition 3.65. Tempered distributions *are continuous linear functionals on $\mathcal{S}_m$. We write $\mathcal{S}_m^\star := \{u : \mathcal{S}_m \to \mathbb{C} \mid u$ linear, continuous$\}$ for the space of tempered distributions and $\langle \varphi, u\rangle_{\mathcal{S}_m} = u(\varphi)$ for $u \in \mathcal{S}_m^\star$ and $\varphi \in \mathcal{S}_m$.*

Recall that $\varphi_n \to \varphi$ in $\mathcal{S}_m$ means that $p_{k,\alpha}(\varphi_n - \varphi) = \||x|^k \partial^\alpha(\varphi_n - \varphi)\|_\infty \to 0$ for all $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^m$, as $n \to \infty$. It is possible to define weak* convergence in $\mathcal{S}_m^\star$, but we will not deal with such questions, see [**Ru2**]. We collect first instructive examples for tempered distributions, namely functions and measures with some growth restrictions as well as *Dirac distributions.*

[9] This section was not part of the lectures.

Example 3.66. a) Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^m)$ satisfy

$$a_l(f) := \int_{l \leqslant |x|_2 < l+1} |f(x)| \, \mathrm{d}x \leqslant c l^\kappa$$

for all $l \in \mathbb{N}$ and some $\kappa, c \geqslant 0$. This condition is satisfied by polynomially bounded f and by $f \in L^p(\mathbb{R}^m)$ with $p \in [1,\infty]$ (because then $a_l(f) \leqslant c\|f\|_p \, l^{(m-1)/p'}$ by Hölder's inequality). For $\varphi \in \mathcal{S}_m$, we define

$$u_f(\varphi) = \int_{\mathbb{R}^m} \varphi f \, \mathrm{d}x.$$

To show $u_f \in \mathcal{S}_m^\star$, let φ_n tend to φ in $\mathcal{S}_m$. Take $k \in \mathbb{N}$ with $k \geqslant \kappa + 2$. Inserting $|x|^{-k}|x|^k$ in the integrands for $l \geqslant 1$, we estimate

$$\begin{aligned} |u_f(\varphi - \varphi_n)| &\leqslant \sum_{l=0}^{\infty} \int_{l \leqslant |x|_2 < l+1} |\varphi - \varphi_n| \, |f(x)| \, \mathrm{d}x \\ &\leqslant \|f\|_{L^1(B(0,1))} p_{0,0}(\varphi - \varphi_n) + p_{k,0}(\varphi - \varphi_n) \sum_{l=1}^{\infty} c l^{\kappa - k} \\ &\leqslant c' (p_{0,0}(\varphi - \varphi_n) + p_{k,0}(\varphi - \varphi_n)) \end{aligned}$$

for a constant $c' > 0$. Hence, $u_f : \mathcal{S}_m \to \mathbb{C}$ is continuous. The linearity of u_f is clear, and so it belongs to $\mathcal{S}_m^\star$. One often writes f instead of u_f.

b) Let μ be a measure on $\mathcal{B}_m$ with $\mu(B(0,1)) < \infty$ and $\mu(B(0,l+1)\backslash B(0,l)) \leqslant c l^\kappa$ for all $l \in \mathbb{N}$ and some $\kappa, c \geqslant 0$. Then one sees as in part a) that

$$u_\mu(\varphi) = \int_{\mathbb{R}^m} \varphi \, \mathrm{d}\mu, \qquad \varphi \in \mathcal{S}_m,$$

defines a tempered distribution u_μ, which is often simply denoted by μ.

c) Let $y \in \mathbb{R}^m$ and $\alpha \in \mathbb{N}_0^m$. We set $\delta_y^\alpha(\varphi) = \partial^\alpha \varphi(y)$ for $\varphi \in \mathcal{S}_m$. Let φ_n converge to φ in $\mathcal{S}_m$. Then $|\delta_y^\alpha(\varphi) - \delta_y^\alpha(\varphi_n)| \leqslant p_{0,\alpha}(\varphi - \varphi_n)$ tends to 0, so that δ_y^α is contained in $\mathcal{S}_m^\star$. $\Diamond$

We now extend the operators from Section 3.1 defined on $\mathcal{S}_m$ to $\mathcal{S}_m^\star$ by duality.

Definition 3.67. *Let $u \in \mathcal{S}_m^\star$, $g \in \mathcal{E}_m$, and $\alpha \in \mathbb{N}_0^m$. For $\varphi \in \mathcal{S}_m$ we define*

a) $(gu)(\varphi) = \langle \varphi, gu \rangle_{\mathcal{S}_m} := \langle g\varphi, u \rangle_{\mathcal{S}_m} = u(g\varphi)$,

b) $(\partial^\alpha u)(\varphi) = \langle \varphi, \partial^\alpha u \rangle_{\mathcal{S}_m} := (-1)^{|\alpha|} \langle \partial^\alpha \varphi, u \rangle_{\mathcal{S}_m} = (-1)^{|\alpha|} u(\partial^\alpha \varphi)$,

c) $\widehat{u}(\varphi) := (\mathcal{F}u)(\varphi) = \langle \varphi, \mathcal{F}u \rangle_{\mathcal{S}_m} := \langle \mathcal{F}\varphi, u \rangle_{\mathcal{S}_m} = u(\mathcal{F}\varphi)$,

d) $(Ru)(\varphi) = \langle \varphi, Ru \rangle_{\mathcal{S}_m} := \langle R\varphi, u \rangle_{\mathcal{S}_m} = u(R\varphi)$,

e) $(\varphi * u)(x) := \langle T_{-x} R\varphi, u \rangle_{\mathcal{S}_m} = u(T_{-x} R\varphi)$ for every $x \in \mathbb{R}^m$.

By Lemma 3.7, the maps gu, $\partial^\alpha u$, $\mathcal{F}u$, and Ru are continuous and linear from $\mathcal{S}_m$ to $\mathbb{C}$, and hence they belong to $\mathcal{S}_m^\star$. Similarly one sees $T_{-x} R\varphi \in \mathcal{S}_m$.

Observe that we multiply and convolve tempered distributions only with the (very regular) functions in $\mathcal{E}_m$ and $\mathcal{S}_m$, respectively. The following examples and the theorem below indicate that the above definitions extend the known concepts in a natural way and that they allow to generalize several main properties of the Fourier transform to the space $\mathcal{S}_m^\star$.

EXAMPLE 3.68. Let $\varphi \in \mathcal{S}_m$, $g \in \mathcal{E}_d$, $\alpha \in \mathbb{N}_0^m$, and $x, y \in \mathbb{R}^m$.

a) Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^m)$ be as in Example 3.66 a). Then $gu_f = u_{gf}$ since

$$(gu_f)(\varphi) = \int_{\mathbb{R}^m} \varphi g f \,\mathrm{d}x = u_{gf}(\varphi).$$

b) Let $f \in W^{k,p}(\mathbb{R}^m)$ for some $p \in [1, \infty]$ and $|\alpha| \leqslant k \in \mathbb{N}$. Then $\partial^\alpha u_f = u_{\partial^\alpha f}$ since the definitions and Gauß' Theorem 3.41 yield

$$\begin{aligned}\langle \varphi, \partial^\alpha u_f \rangle_{\mathcal{S}_m} &= (-1)^{|\alpha|} \langle \partial^\alpha \varphi, u_f \rangle_{\mathcal{S}_m} = (-1)^{|\alpha|} \int_{\mathbb{R}^m} \partial^\alpha \varphi \, f \,\mathrm{d}x = \int_{\mathbb{R}^m} \varphi \, \partial^\alpha f \,\mathrm{d}x \\ &= \langle \varphi, u_{\partial^\alpha f} \rangle_{\mathcal{S}_m} .\end{aligned}$$

c) Let $f \in L^2(\mathbb{R}^m)$. Then $\mathcal{F} u_f = u_{\mathcal{F} f}$ since Theorem 3.11 implies

$$\langle \varphi, \mathcal{F} u_f \rangle_{\mathcal{S}_m} = \langle \mathcal{F} \varphi, u_f \rangle_{\mathcal{S}_m} = \int_{\mathbb{R}^m} \widehat{\varphi} f \,\mathrm{d}x = \int_{\mathbb{R}^m} \varphi \widehat{f} \,\mathrm{d}x = \langle \varphi, u_{\mathcal{F} f} \rangle_{\mathcal{S}_m} .$$

d) We have $\partial^\alpha \delta_y = (-1)^{|\alpha|} \delta_y^\alpha$ because of

$$\langle \varphi, \partial^\alpha \delta_y \rangle_{\mathcal{S}_m} = (-1)^{|\alpha|} \langle \partial^\alpha \varphi, \delta_y \rangle_{\mathcal{S}_m} = (-1)^{|\alpha|} \partial^\alpha \varphi(y) = (-1)^{|\alpha|} \delta_y^\alpha(\varphi).$$

e) We have $\mathcal{F} \delta_y = (2\pi)^{-m/2} e_{-\mathrm{i}y}$ because of

$$\langle \varphi, \mathcal{F} \delta_y \rangle_{\mathcal{S}_m} = \langle \mathcal{F} \varphi, \delta_y \rangle_{\mathcal{S}_m} = (2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\mathrm{i} y \cdot x} \varphi(x) \,\mathrm{d}x = \big\langle \varphi, (2\pi)^{-\frac{m}{2}} e_{-\mathrm{i}y} \big\rangle_{\mathcal{S}_m} .$$

f) We have $\mathcal{F} e_{\mathrm{i}y} = (2\pi)^{m/2} \delta_y$ since Proposition 3.10 implies

$$\langle \varphi, \mathcal{F} e_{\mathrm{i}y} \rangle_{\mathcal{S}_m} = \langle \mathcal{F} \varphi, e_{\mathrm{i}y} \rangle_{\mathcal{S}_m} = \int_{\mathbb{R}^m} \widehat{\varphi}(\xi) \mathrm{e}^{\mathrm{i} y \cdot \xi} \,\mathrm{d}\xi = (2\pi)^{\frac{m}{2}} (\mathcal{F}^{-1} \widehat{\varphi})(y) = (2\pi)^{\frac{m}{2}} \varphi(y).$$

Assertion f) can also be deduced from e) since $\mathcal{F}^2$ is equal to R in $\mathcal{S}_m^\star$, too, as shown in the next theorem (with a similar proof as above).

g) Let $f \in L^1(\mathbb{R}^m)$. Then $\varphi * u_f = \varphi * f$, since

$$\varphi * u_f(x) = \langle T_{-x} R \varphi, u_f \rangle_{\mathcal{S}_m} = \int_{\mathbb{R}^m} \varphi(-(z - x)) f(z) \,\mathrm{d}z = \varphi * f(x). \qquad \Diamond$$

We now collect the main properties of the above objects on $\mathcal{S}_m^\star$. Observe that the second part of assertion b) does not work on $W^{k,2}(\mathbb{R}^m)$.

THEOREM 3.69. *Let $u \in \mathcal{S}_m^\star$, $\varphi, \psi \in \mathcal{S}_m$, and $\alpha \in \mathbb{N}_0^m$. The following assertions hold.*

a) $\mathcal{F} : \mathcal{S}_m^\star \to \mathcal{S}_m^\star$ is bijective with $\mathcal{F}^4 = I$ and $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$.

b) $\mathcal{F}(\partial^\alpha u) = \mathrm{i}^{|\alpha|} \xi^\alpha \mathcal{F} u$ and $\partial^\alpha (\mathcal{F} u) = (-\mathrm{i})^{|\alpha|} \mathcal{F}(x^\alpha u)$.

c) $\varphi * u \in \mathcal{E}_m$, and hence $\varphi * u$ induces a tempered distribution.

d) $\partial^\alpha (\varphi * u) = (\partial^\alpha \varphi) * u = \varphi * \partial^\alpha u$.

e) $\mathcal{F}(\varphi * u) = (2\pi)^{m/2} \widehat{\varphi} \widehat{u}$ and $\mathcal{F}(\varphi u) = (2\pi)^{-m/2} \widehat{\varphi} * \widehat{u}$.

PROOF. Let $u \in \mathcal{S}_m^\star$, $\varphi \in \mathcal{S}_m$, and $\alpha \in \mathbb{N}_0^m$. For a), Proposition 3.10 yields

$$\langle \varphi, \mathcal{F}^4 u \rangle_{\mathcal{S}_m} = \langle \mathcal{F} \varphi, \mathcal{F}^3 u \rangle_{\mathcal{S}_m} = \cdots = \langle \mathcal{F}^4 \varphi, u \rangle_{\mathcal{S}_m} = \langle \varphi, u \rangle_{\mathcal{S}_m} ,$$

so that $\mathcal{F}^4 = I$ on $\mathcal{S}_m^\star$ and $\mathcal{F} : \mathcal{S}_m^\star \to \mathcal{S}_m^\star$ is bijective with inverse $\mathcal{F}^{-1} = \mathcal{F}^3$. Similarly, we show the remaining equality $\mathcal{F}^2 = R$ by computing

$$\langle \varphi, \mathcal{F}^2 u \rangle_{\mathcal{S}_m} = \langle \mathcal{F}^2 \varphi, u \rangle_{\mathcal{S}_m} = \langle R \varphi, u \rangle_{\mathcal{S}_m} .$$

The first equality in statement b) follows from Lemma 3.7 and

$$\begin{aligned}\langle \varphi, \mathcal{F}\partial^\alpha u\rangle_{\mathcal{S}_m} &= \langle \mathcal{F}\varphi, \partial^\alpha u\rangle_{\mathcal{S}_m} = (-1)^{|\alpha|}\langle \partial^\alpha \mathcal{F}\varphi, u\rangle_{\mathcal{S}_m} = \mathrm{i}^{|\alpha|}\langle \mathcal{F}(x^\alpha\varphi), u\rangle_{\mathcal{S}_m}\\ &= \langle \varphi, \mathrm{i}^{|\alpha|}\xi^\alpha \mathcal{F}u\rangle_{\mathcal{S}_m}.\end{aligned}$$

The second part of b) is established in the same way.

For the proof of assertions c) and d) we refer to Theorem 7.19 in [**Ru2**]. To show e), first take $\psi \in C_c^\infty(\mathbb{R}^m)$. There is a closed interval $I \subset \mathbb{R}^m$ such that $\operatorname{supp}\psi \subset I$. Using the definitions, part a) and Proposition 3.3 in the last step, we compute

$$\begin{aligned}\langle \widehat{\psi}, \mathcal{F}(\varphi * u)\rangle_{\mathcal{S}_m} &= \langle R\psi, \varphi * u\rangle_{\mathcal{S}_m} = \int_{\mathbb{R}^m} \psi(-x)(\varphi * u)(x)\,\mathrm{d}x\\ &= \int_{-I} \psi(-x)u(T_{-x}R\varphi)\,\mathrm{d}x = \int_I u(\psi(z)T_zR\varphi)\,\mathrm{d}z\\ &= u\Big(\int_I \psi(z)T_zR\varphi\,\mathrm{d}z\Big) = \langle R(\psi*\varphi), u\rangle_{\mathcal{S}_m}\\ &= \langle \mathcal{F}(\psi*\varphi), \mathcal{F}u\rangle_{\mathcal{S}_m} = (2\pi)^{\frac{m}{2}}\langle \widehat{\psi}\widehat{\varphi}, \mathcal{F}u\rangle_{\mathcal{S}_m}.\end{aligned} \tag{3.48}$$

Here the second integral in (3.48) is understood as an $\mathcal{S}_m$-valued Riemann integral on I; i.e., as the limit in $\mathcal{S}_m$ of Riemann sums such as

$$S_n(y) = \sum_{j=1}^{N_n} \psi(z_{j,n})(R\varphi)(y + z_{j,n})\operatorname{vol}(Q_{j,n}), \qquad y \in \mathbb{R}^m,$$

where $z_{j,n} \in Q_{j,n}$, the rectangles $Q_{j,n}$ with $j \in \{1, \dots, N_n\}$ subdivide I, and $\max_j \operatorname{vol}(Q_{j,n})$ tends to 0 as $n \to \infty$. Clearly, S_n belongs to $\mathcal{S}_m$. We omit the somewhat tedious, but elementary proof that S_n indeed converges in $\mathcal{S}_m$. Hence, u can be taken out of the approximating Riemann sums by its linearity and out of the limit by its continuity. This fact justifies that we have interchanged u and the integral in (3.48). So far we have shown

$$\langle \widehat{\psi}, \mathcal{F}(\varphi * u)\rangle_{\mathcal{S}_m} = (2\pi)^{\frac{m}{2}}\langle \widehat{\psi}, \widehat{\varphi}\widehat{u}\rangle_{\mathcal{S}_m} \tag{3.49}$$

for all $\psi \in C_c^\infty(\mathbb{R}^m)$. Arguing as in the first part of the proof of Theorem 4.21 in [**FA**], one can show that $C_c^\infty(\mathbb{R}^m)$ is dense in $\mathcal{S}_m$, see Theorem 7.10 in [**Ru2**]. Since the Fourier transform is continuous on $\mathcal{S}_m$ by Lemma 3.7, the identity (3.49) is thus valid for all $\widehat{\psi}$ with $\psi \in \mathcal{S}_m$ due to an approximation argument. We can now replace here $\widehat{\psi}$ by $\psi \in \mathcal{S}_m$ using that $\mathcal{F}$ is bijective on $\mathcal{S}_m$ thanks to Proposition 3.10. So the first part of assertion e) is shown. For the second part, observe that

$$\begin{aligned}\langle \psi, R(\varphi)R(u)\rangle_{\mathcal{S}_m} &= \langle \psi R\varphi, Ru\rangle_{\mathcal{S}_m} = \langle R(\psi)R^2\varphi, u\rangle_{\mathcal{S}_m} = \langle R\psi, \varphi u\rangle_{\mathcal{S}_m}\\ &= \langle \psi, R(\varphi u)\rangle_{\mathcal{S}_m},\end{aligned}$$

for all $\psi \in \mathcal{S}_m$; i.e., $R(\varphi)R(u) = R(\varphi u)$. Employing also a), we then calculate

$$\mathcal{F}(\widehat{\varphi} * \widehat{u}) = (2\pi)^{\frac{m}{2}}\mathcal{F}^2(\varphi)\mathcal{F}^2(u) = (2\pi)^{\frac{m}{2}}R(\varphi u) = (2\pi)^{\frac{m}{2}}\mathcal{F}^2(\varphi u).$$

Applying $\mathcal{F}^{-1}$, the second part of assertion e) follows. $\square$

CHAPTER 4

Self-adjoint operators

As on $\mathbb{C}^m$, self-adjoint operators on a Hilbert space possess a very powerful spectral theory. It has important applications all over mathematics and its applications; the mathematical foundation of quantum mechanics is a prime example (see [**RS**]). For these applications we also have to study unbounded self-adjoint operators. In the first section we focus on the often difficult and most basic problem how to determine whether a given (partial) differential operator is self-adjoint. In the second section we then establish a core result of spectral theory: the spectral theorem for self-adjoint operators, also in the unbounded case.

In this chapter X and Y are Hilbert spaces with scalar product $(\cdot|\cdot)$ and we let $\mathbb{F} = \mathbb{C}$, unless something else is said.

4.1. Basic properties

We start with the under-lying concepts of this chapter. Let A be a densely defined linear operator from X to Y. We define its *Hilbert space adjoint* A' as in the Banach space case by

$$\begin{aligned} \mathrm{D}(A') &:= \{ y \in Y \,|\, \exists\, z \in X \ \forall\, x \in \mathrm{D}(A) : \ (Ax|y) = (x|z) \}, \\ A'y &:= z. \end{aligned} \tag{4.1}$$

As in Remark 1.23 one sees that $A' : \mathrm{D}(A') \to X$ is a linear map, which is closed from Y to X. Let $T \in \mathcal{B}(X,Y)$. Then $\mathrm{D}(T') = Y$ and T' is given by

$$\forall\, x \in X,\, y \in Y : \qquad \big(x\big|T'y\big) = (Tx|y)$$

as in (5.9) in [**FA**]. We recall from Proposition 5.42 of [**FA**] that

$$\|T\| = \|T'\|, \quad T'' = T, \quad (\alpha T + \beta S)' = \overline{\alpha} T' + \overline{\beta} S', \quad (UT)' = T'U' \tag{4.2}$$

for $T, S \in \mathcal{B}(X,Y)$, $U \in \mathcal{B}(Y,Z)$, a Hilbert space Z, and $\alpha, \beta \in \mathbb{C}$.

Definition 4.1. *A densely defined linear operator A on X is called* self-adjoint *if $A = A'$ (in particular,* $\mathrm{D}(A) = \mathrm{D}(A')$ *and A must be closed),* skew-adjoint *if $A = -A'$, and* normal *if $AA' = A'A$. We say that $T \in \mathcal{B}(X,Y)$ is* unitary *if T is invertible with $T^{-1} = T'$.*

Let $\Phi : X \to X^\star$ be the (antilinear) Riesz isomorphism given by $(\Phi(x))(y) = (y|x)$ for all $x, y \in X$. For $\lambda \in \mathbb{C}$ and $X = Y$, we obtain

$$\overline{\lambda} I_X - T' = \Phi^{-1}(\lambda I_{X^\star} - T^\star)\Phi, \qquad \overline{\lambda} I_X - A' = \Phi^{-1}(\lambda I_{X^\star} - A^\star)\Phi$$

with $\mathrm{D}(A') = \Phi^{-1}\,\mathrm{D}(A^\star)$ as in p.109 of [**FA**]. Theorem 1.24 thus implies

$$\sigma(A) = \sigma(A^\star) = \overline{\sigma}(A'), \qquad \sigma_{\mathrm{r}}(A) = \sigma_{\mathrm{p}}(A^\star) = \overline{\sigma}_p(A'), \tag{4.3}$$
$$R(\overline{\lambda}, A') = \Phi^{-1} R(\lambda, A^\star)\Phi = \Phi^{-1} R(\lambda, A)^\star \Phi = R(\lambda, A)' \qquad \text{for} \ \ \lambda \in \rho(A),$$

where the bars mean complex conjugation of each element. For $\mathbb{F} = \mathbb{R}$, the above definitions also make sense and the stated results remain valid (without the extra conjugation). The following relations between the above concepts are straightforward to check.

Remark 4.2. a) Self-adjoint, skew-adjoint or unitary operators are normal (where $X = Y$). If $A = A'$ and $\lambda \in \mathbb{C}$, then $\lambda I - A$ is normal.

b) A densely defined linear operator A on X is skew-adjoint if and only if $\mathrm{i}A$ is self-adjoint. (As in (4.2) one sees that $(\mathrm{i}A)' = -\mathrm{i}A$.) $\Diamond$

Theorem 1.16 says that the spectral radius $\mathrm{r}(T)$ is less or equal than $\|T\|$ for each bounded operator T on a Banach space. Already for the matrix $T = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ one has the strict inequality $\mathrm{r}(T) = 0 < 1 = \|T\|$. We next show $\|T\| = \mathrm{r}(T)$ for normal operators, which is a key to their deeper properties.

Proposition 4.3. *Each operator $T \in \mathcal{B}(X, Y)$ satisfies $\|T'T\| = \|TT'\| = \|T\|^2$. (This is also true if $\mathbb{F} = \mathbb{R}$.) Let $T \in \mathcal{B}(X)$ be normal. We then have $\|T\| = \mathrm{r}(T)$, and thus $T = 0$ if $\sigma(T) = \{0\}$.*

Proof. For $x \in X$, using (4.2) we compute

$$\begin{aligned} \|Tx\|^2 &= (Tx|Tx) = \big(T'Tx|x\big) \leqslant \|T'T\|\|x\|^2, \\ \|T\|^2 &= \sup_{\|x\| \leqslant 1} \|Tx\|^2 \leqslant \|T'T\| \leqslant \|T'\|\|T\| = \|T\|^2; \end{aligned}$$

i.e., $\|T\|^2 = \|T'T\|$. We infer that $\|T\|^2 = \|T'\|^2 = \|T''T'\| = \|TT'\|$.

Next, let T be normal. From the first part we then deduce

$$\|T^2\|^2 = \|T^2(T^2)'\| = \|TTT'T'\| = \|TT'TT'\| = \|TT'(TT')'\| = \|TT'\|^2 = \|T\|^4,$$

so that $\|T^2\| = \|T\|^2$. Iteratively it follows that $\|T^{2^n}\| = \|T\|^{2^n}$ for all $n \in \mathbb{N}$. By means of Theorem 1.16 we conclude

$$\mathrm{r}(T) = \lim_{j \to \infty} \|T^j\|^{1/j} = \lim_{n \to \infty} \|T^{2^n}\|^{2^{-n}} = \|T\|. \qquad \square$$

The following concepts turn out to be very useful to compute adjoints, for instance. In the next definition and remark we also allow for $\mathbb{F} = \mathbb{R}$.

Definition 4.4. *Let A and B be linear operators from a Banach space X to a Banach space Y. We say that B* extends *A (and write $A \subseteq B$) if $\mathrm{D}(A) \subseteq \mathrm{D}(B)$ and $Ax = Bx$ for all $x \in \mathrm{D}(A)$.*

Next, let X be a Hilbert space. A linear operator A on X is called symmetric *if we have $(Ax|y) = (x|Ay)$ for all $x, y \in \mathrm{D}(A)$.*

By (4.1), a self-adjoint operator is symmetric. In the unbounded case the converse is not true, in general, see Example 4.8. We collect direct consequences of these definitions.

Remark 4.5. Let A and B be linear operators from a Banach space X to a Banach space Y.

a) The operator B extends A if and only if its graph $\mathrm{G}(B)$ contains $\mathrm{G}(A)$. Let $A \subseteq B$. Then $A = B$ is equivalent to $\mathrm{D}(B) \subseteq \mathrm{D}(A)$.

b) Let $A \subseteq B$, A be surjective, and B be injective. Then A and B are equal. Hence, if $X = Y$ and there is $\lambda \in \mathbb{F}$ such that $\lambda I - A$ is surjective and $\lambda I - B$ is injective, then we have $A = B$.

PROOF. Let $x \in \mathrm{D}(B)$ and set $y = Bx$. The surjectivity of A yields a vector z in $\mathrm{D}(A) \subseteq \mathrm{D}(B)$ with $y = Az = Bz$. Since B is injective, we obtain $x = z \in \mathrm{D}(A)$ so that $A = B$. $\square$

c) Let X and Y be Hilbert spaces, A and B be densely defined, and $A \subseteq B$. We then have $B' \subseteq A'$.

PROOF. Let $y \in \mathrm{D}(B')$ and $x \in \mathrm{D}(A)$. The assumption implies

$$(Ax|y) = (Bx|y) = (x|B'y),$$

so that y belongs to $\mathrm{D}(A')$ and $B'y = A'y$ by (4.1). $\square$

d) Let A be densely defined and symmetric on a Hilbert space X. Definition (4.1) implies that $A \subseteq A'$. In particular, A is self-adjoint if and only if $\mathrm{D}(A') \subseteq \mathrm{D}(A)$. $\Diamond$

In the next theorem we give very useful spectral conditions for the self-adjointness of a symmetric operator. We start with a crucial lemma.

LEMMA 4.6. *Let A be symmetric, $x \in \mathrm{D}(A)$, and $\alpha, \beta \in \mathbb{R}$. Set $\lambda = \alpha + \mathrm{i}\beta$. We have $(Ax|x) \in \mathbb{R}$ and*

$$\|\lambda x - Ax\|^2 = \|\alpha x - Ax\|^2 + |\beta|^2 \|x\|^2 \geqslant |\beta|^2 \|x\|^2.$$

If A is also closed, then $\sigma_{\mathrm{ap}}(A) \subseteq \mathbb{R}$ and $\|R(\lambda, A)\| \leqslant \frac{1}{|\mathrm{Im}\,\lambda|}$ for all $\lambda \in \rho(A) \setminus \mathbb{R}$.

PROOF. For $x \in \mathrm{D}(A)$ we have $(Ax|x) = (x|Ax) = \overline{(Ax|x)}$ so that $(Ax|x) = (x|Ax)$ is real. From this fact we deduce that

$$\begin{aligned}
\|\lambda x - Ax\|^2 &= (\alpha x - Ax + \mathrm{i}\beta x|\alpha x - Ax + \mathrm{i}\beta x) \\
&= \|\alpha x - Ax\|^2 + 2\,\mathrm{Re}\,(\mathrm{i}\beta x|\alpha x - Ax) + \|\mathrm{i}\beta x\|^2 \\
&= \|\alpha x - Ax\|^2 + 2\,\mathrm{Re}(\mathrm{i}\beta\alpha\|x\|^2 - \mathrm{i}\beta\,(x|Ax)\,) + |\beta|^2\|x\|^2 \\
&= \|\alpha x - Ax\|^2 + |\beta|^2\|x\|^2 \geqslant |\beta|^2\|x\|^2.
\end{aligned}$$

In particular, λ does not belong to $\sigma_{\mathrm{ap}}(A)$ if $\mathrm{Im}\,\lambda = \beta \neq 0$.

Let $\lambda \in \rho(A) \setminus \mathbb{R}$ and $y \in X$. Set $x = R(\lambda, A)y \in \mathrm{D}(A)$. We then calculate

$$\|y\|^2 = \|\lambda x - Ax\|^2 \geqslant |\mathrm{Im}\,\lambda|^2\|x\|^2 = |\mathrm{Im}\,\lambda|^2\|R(\lambda, A)y\|^2. \qquad \square$$

THEOREM 4.7. *Let X be a Hilbert space and A be densely defined, closed and symmetric on X. Then the following assertions are true.*

a) The spectrum $\sigma(A)$ is either a subset of $\mathbb{R}$ or $\sigma(A) = \mathbb{C}$ or $\sigma(A) = \{\lambda \in \mathbb{C} \mid \mathrm{Im}\,\lambda \geqslant 0\}$ or $\sigma(A) = \{\lambda \in \mathbb{C} \mid \mathrm{Im}\,\lambda \leqslant 0\}$.

b) The following assertions are equivalent.

i) $A = A'$.

ii) $\sigma(A) \subseteq \mathbb{R}$.

iii) $\mathrm{i}I - A'$ and $\mathrm{i}I + A'$ are injective.

iv) $(\mathrm{i}I - A)\,\mathrm{D}(A)$ and $(\mathrm{i}I + A)\,\mathrm{D}(A)$ are dense in X.

c) Let $\rho(A) \cap \mathbb{R} \neq \emptyset$. Then A is self-adjoint.

d) Let A be self-adjoint. Then we have

$$\|R(\lambda, A)\| \leqslant \frac{1}{|\mathrm{Im}\,\lambda|} \tag{4.4}$$

for $\lambda \notin \mathbb{R}$. Further, $\sigma(A) = \sigma_{\mathrm{ap}}(A)$ is non-empty and A has no symmetric extension $B \neq A$.

PROOF. a) Let $\lambda \in \rho(A)$ and $\mu \in \sigma(A)$. Suppose that $\operatorname{Im}\lambda > 0$ and $\operatorname{Im}\mu > 0$. The line segment from λ to μ then must contain a point $\nu \in \partial\sigma(A)$. This point belongs to $\sigma_{\mathrm{ap}}(A)$ by Proposition 1.19 and satisfies $\operatorname{Im}\nu > 0$, which contradicts Lemma 4.6 since A is symmetric. Similarly we exclude that $\operatorname{Im}\lambda < 0$ and $\operatorname{Im}\mu < 0$. Since the spectrum is closed, only the four cases in assertion a) remain.

b) Let A be self-adjoint. Lemma 4.6 yields the inclusion $\sigma_{\mathrm{ap}}(A) \subseteq \mathbb{R}$. Due to (4.3), we also have the equalities $\sigma_{\mathrm{r}}(A) = \overline{\sigma}_p(A') = \overline{\sigma}_p(A)$ so that $\sigma_{\mathrm{r}}(A) = \sigma_{\mathrm{p}}(A) \subseteq \mathbb{R}$. From Proposition 1.19 we thus deduce $\sigma(A) \subseteq \mathbb{R}$; i.e., i) implies ii). The implication 'ii)⇒iii)' is obvious. Equation (4.3) also shows that $\pm\mathrm{i}$ belongs to $\sigma_{\mathrm{p}}(A')$ if and only if $\mp\mathrm{i}$ to $\sigma_{\mathrm{r}}(A)$; i.e., claims iii) and iv) are equivalent. Finally, let statement iv) (and thus iii)) be true. The range of the operator $\mathrm{i}I - A$ is closed by Lemma 4.6 and Proposition 1.19. In view of iv), the map $\mathrm{i}I - A$ is then surjective. On the other hand, $\mathrm{i}I - A'$ is injective because of iii), and hence A is equal to A' thanks to Remark 4.5.

c) Assume there is a point λ in $\rho(A) \cap \mathbb{R}$. Then $\rho(A)$ contains a ball around λ by its openness. By part a), the spectrum of A is thus contained in $\mathbb{R}$, and so A is self-adjoint by b).

d) Let $A = A'$. If its spectrum was empty, then A would be invertible with a self-adjoint inverse (see (4.3)). Proposition 1.20 thus yields that $\sigma(A^{-1})$ is equal to $\{0\}$ so that $A^{-1} = 0$ by Proposition 4.3, which is impossible. Hence, $\sigma(A)$ is non-empty. Let $A \subseteq B$ for a symmetric operator B on X. Remark 4.5 then yields $A \subseteq B \subseteq B' \subseteq A' = A$ and so $A = B$. Since $\sigma(A) \subseteq \mathbb{R}$ by b), we have $\sigma(A) = \partial\sigma(A) \subseteq \sigma_{\mathrm{ap}}(A)$ due to Proposition 1.19, and (4.4) follows from Lemma 4.6. □

We discuss several examples with (unbounded) differential operators, complementing the results from Section 3.4. (See Example 5.44 in [**FA**] for the bounded case.) Typically it is straightforward to check symmetry integrating by parts. We then use spectral properties and Theorem 4.7 to establish self-adjointness. The examples also indicate that boundary conditions (possible 'at $\pm\infty$') are often necessary for symmetry. However 'too many' boundary conditions can be an obstacle to self-adjointness.

EXAMPLE 4.8. a) Let $X = L^2(\mathbb{R})$ and $A = \mathrm{i}\partial$ with $\mathrm{D}(A) = W^{1,2}(\mathbb{R})$. Then A is self-adjoint with $\sigma(A) = \mathbb{R}$.

PROOF. For $u, v \in \mathrm{D}(A)$, integrating by parts we deduce

$$(Au|v) = \mathrm{i}\int_{\mathbb{R}} \partial u\, \overline{v}\, \mathrm{d}s = -\mathrm{i}\int_{\mathbb{R}} u\, \partial\overline{v}\, \mathrm{d}s = \int_{\mathbb{R}} u\, \overline{\mathrm{i}\partial v}\, \mathrm{d}s = (u|Av)\,,$$

see Theorem 3.41; i.e., A is symmetric. Proposition 1.20 and Example 3.43 further imply that $\sigma(A) = \mathrm{i}\sigma(-\mathrm{i}A) = \mathrm{i}^2\,\mathbb{R} = \mathbb{R}$. Hence, A is self-adjoint. □

b) Let $X = L^2(\mathbb{R}_+)$ and $A = \mathrm{i}\partial$ on $\mathrm{D}(A) = W^{1,2}(\mathbb{R}_+)$. Then A is not symmetric.

PROOF. For $u, v \in \mathrm{D}(A)$ with $u(0) = v(0) = 1$, as above an integration by parts implies

$$(Au|v) = \mathrm{i}\int_0^\infty \partial u\, \overline{v}\, \mathrm{d}s = -\mathrm{i}\int_0^\infty u\, \partial\overline{v}\, \mathrm{d}s - \mathrm{i}\overline{u(0)}v(0) = (u|Av) - \mathrm{i}. \qquad \square$$

c) Let $X = L^2(\mathbb{R}_+)$ and $A = \mathrm{i}\partial$ on $\mathrm{D}(A) = W_0^{1,2}(\mathbb{R}_+)$. Then A is symmetric, but not self-adjoint, and $\sigma(A) = \{\lambda \in \mathbb{C} \,|\, \operatorname{Im}\lambda \geqslant 0\}$.

Proof. Symmetry is shown as in a) using Theorem 3.41 and that now $u, v \in \mathrm{D}(A)$ have trace 0 at $s = 0$. From Proposition 1.20 and Example 3.44 we deduce that $\sigma(A) = -\mathrm{i}\sigma(\mathrm{i}A) = -\mathrm{i}\{\lambda \in \mathbb{C} \,|\, \operatorname{Re}\lambda \leqslant 0\} = \{\lambda \in \mathbb{C} \,|\, \operatorname{Im}\lambda \geqslant 0\}$. Consequently, A is not self-adjoint. □

d) Let $X = L^2(\mathbb{R}^m)$ and $A = \Delta$ with $\mathrm{D}(A) = W^{2,2}(\mathbb{R}^m)$. Then A is self-adjoint with $\sigma(A) = \mathbb{R}_{\leqslant 0}$.

Proof. For $u, v \in \mathrm{D}(A)$, Theorem 3.41 yields

$$(Au|v) = \int_{\mathbb{R}^m} \Delta u \, \overline{v} \, \mathrm{d}y = \int_{\mathbb{R}^m} u \, \Delta\overline{v} \, \mathrm{d}y = (u|Av) \, ,$$

so that A is symmetric. In Example 3.47 we have seen that $\sigma(A) = \mathbb{R}_{\leqslant 0}$, and hence A is self-adjoint. □

e) Let $U \subseteq \mathbb{R}^m$ be open and bounded with C^2-boundary and $A = \Delta$ with $\mathrm{D}(A) = W^{2,2}(U) \cap W_0^{1,2}(U)$. Then A is self-adjoint (and has compact resolvent by Example 3.49). In fact, the symmetry of A can be shown as in part d) because the traces of $u, v \in D(A)$ vanish by Theorem 3.38. Then A is self-adjoint since it is invertible by Example 3.49.

f) Let $X = L^2(0,1)$, $A_0 = \partial^2$ with $\mathrm{D}(A_0) = W_0^{2,2}(0,1)$, and A be as in assertion e) with $U = (0,1)$. As in e) we see that A_0 is symmetric. But A_0 is not self-adjoint, since $A_0 \subsetneqq A$ and $A = A'$ (see Theorem 4.7 d)). We further claim that $A_0' = \partial^2$ with $\mathrm{D}(A_0') = W^{2,2}(0,1)$.

Proof. For $v \in W^{2,2}(0,1)$ and $u \in \mathrm{D}(A)$, we deduce from Theorem 3.41

$$(A_0 u|v) = \int_0^1 \partial^2 u \, \overline{v} \, \mathrm{d}s = \int_0^1 u \, \partial^2\overline{v} \, \mathrm{d}s + \big[\overline{v}\partial u - u\partial\overline{v}\big]_0^1 = (u|\partial^2 v);$$

i.e., $(\partial^2, W^{2,2}(0,1)) \subseteq A_0'$. Conversely, take $v \in D(A_0')$. For $u \in C_c^\infty(0,1)$ we have $\overline{u} \in C_c^\infty(0,1) \subseteq \mathrm{D}(A)$ and hence obtain

$$\int_0^1 \partial^2\overline{u} \, \overline{v} \, \mathrm{d}s = (A_0\overline{u}|v) = (\overline{u}|A_0' v) = \int_0^1 \overline{u} \, \overline{A_0' v} \, \mathrm{d}s.$$

After complex conjugation, we see that $v \in W_2(0,1) \cap X$ and $\partial^2 v = A_0' v \in X$. The function v thus belongs to $W^{2,2}(0,1)$ by Example 3.42. □

We note that $\sigma_{\mathrm{p}}(A_0') = \mathbb{C}$ in Example 4.8 f) since e_μ is an eigenfunction for the eigenvalue μ^2 for each $\mu \in \mathbb{C}$. As in part b) one can also show that A_0' is not symmetric.

We next prove that small symmetric perturbations preserve self-adjointness. Compare Theorem 1.27 and the exercises for similar results on injectivity and Fredholm properties, respectively.

Theorem 4.9. *Let A be densely defined and self-adjoint on X and let B be symmetric with $\mathrm{D}(A) \subseteq \mathrm{D}(B)$. Assume there are constants $c > 0$ and $\delta \in [0, 1/2)$ such that $\|Bx\| \leqslant c\|x\| + \delta\|Ax\|$ for all $x \in \mathrm{D}(A)$. Then the operator $A + B$ with $\mathrm{D}(A+B) = \mathrm{D}(A)$ is self-adjoint.*

Proof. By Theorem 4.7, the number $\mathrm{i}t$ belongs to $\rho(A)$ for all $t \in \mathbb{R}\backslash\{0\}$. Take $\varepsilon \in (0, 1-2\delta) \subseteq (0,1)$ and $x \in X$. Using (4.4), we estimate

$$\|BR(\mathrm{i}t, A)x\| \leqslant \delta\|AR(\mathrm{i}t, A)x\| + c\|R(\mathrm{i}t, A)^{-1}x\|$$

$$= \delta\|\mathrm{i}tR(\mathrm{i}t, A)x - x\| + c\|R(\mathrm{i}t, A)x\|$$
$$\leqslant \delta\Big(\frac{|t|}{|t|} + 1\Big)normx + \frac{c}{|t|}\|x\| \leqslant (1-\varepsilon)\|x\|,$$

whenever $|t| \geqslant \frac{c}{1-2\delta-\varepsilon}$. Theorem 1.27 now implies that $\pm \mathrm{i}t \in \rho(A+B)$ for such t. Moreover, $A+B$ is symmetric since

$$((A+B)x|y) = (Ax|y) + (Bx|y) = (x|Ay) + (x|By) = (x|(A+B)y)$$

for all $x, y \in \mathrm{D}(A)$. So, $A+B$ is self-adjoint due to Theorem 4.7. □

Actually, in the above theorem it suffices to assume that $\delta < 1$, see Theorem X.13 in [**RS**]. We present a typical application of the theorem which is very important for applications.

Example 4.10. On $L^2(\mathbb{R}^3)$ consider $A = \Delta$ with $\mathrm{D}(A) = W^{2,2}(\mathbb{R}^3)$. Set $Vu = \frac{b}{|x|_2}u$ for $u \in \mathrm{D}(A)$ and some $b \in \mathbb{R}$. Then $A+V$ with domain $\mathrm{D}(A)$ is self-adjoint.

Proof. Recall from Examples 4.8 and 3.47 that A is self-adjoint and its graph norm is equivalent to $\|\cdot\|_{2,2}$. Since $k - \frac{m}{p} = 2 - \frac{3}{2} > 0$, Theorem 3.31 yields $\mathrm{D}(A) \hookrightarrow C_0(\mathbb{R}^3)$. Let $0 < \varepsilon \leqslant 1$. Using polar coordinates, we compute

$$\int_{\mathbb{R}^3} |Vu|^2 \,\mathrm{d}x = b^2 \int_{B(0,\varepsilon)} \frac{|u(x)|^2}{|x|_2^2}\,\mathrm{d}x + b^2 \int_{\mathbb{R}^3\setminus B(0,\varepsilon)} \frac{|u(x)|^2}{|x|_2^2}\,\mathrm{d}x$$
$$\leqslant c\|u\|_\infty^2 \int_0^\varepsilon \frac{r^2}{r^2}\,\mathrm{d}r + \frac{b^2}{\varepsilon^2}\int_{\mathbb{R}^3\setminus B(0,\varepsilon)} |u|^2\,\mathrm{d}x$$
$$\leqslant c\varepsilon\|u\|_{2,2}^2 + \frac{b^2}{\varepsilon^2}\|u\|_2^2 \leqslant c\varepsilon\|Au\|_2^2 + c\varepsilon\|u\|_2^2 + \frac{b^2}{\varepsilon^2}\|u\|_2^2,$$

for constants $c > 0$ independent of $u \in \mathrm{D}(A)$ and ε. Moreover, V is symmetric on $\mathrm{D}(A)$ since

$$(Vu|v) = \int_{\mathbb{R}^3} \frac{b}{|x|_2} u(x)\overline{v(x)}\,\mathrm{d}x = (u|Vv)$$

for all $u, v \in \mathrm{D}(A)$, using that $b|x|_2^{-1}$ is real. For small $\varepsilon > 0$, Theorem 4.9 implies that $A+V$ is self-adjoint. □

The spectra $\sigma(A+V)$ and $\sigma_{\mathrm{p}}(A+V)$ and the eigenfunctions of $A+V$ are computed in §7.3.4 of [**Tr**], where $b > 0$. The above operator $A = \Delta + V$ is used in physics to describe the hydrogen atom. We come back to this point.

4.2. The spectral theorems

Hermitian matrices are unitarily equivalent to diagonal matrices and thus very easy to treat. In this section, for self-adjoint operators we establish infinite-dimensional analogues of this basic result from linear algebra. These 'spectral theorems' can be extended to normal operators, and the separability assumption partly made below can be removed. See Corollaries X.2.8 and X.5.4 in [**DS**] or Theorems 13.24, 13.30 and 13.33 in [**Ru2**].

There are three versions of the spectral theorem. We start with the 'functional calculus' variant for bounded self-adjoint T. To this end, we first introduce the most simple functional calculus for general $T \in \mathcal{B}(Z)$ on a

Banach space Z. Let $p(z) = a_0 + a_1 z + \ldots + a_n z^n$ be a complex polynomial. We then define the operator polynomial

$$p(T) = a_0 I + a_1 T + \ldots + a_N T^N \in \mathcal{B}(Z). \tag{4.5}$$

This gives a map $p \mapsto p(T)$ from the space of polynomials to $\mathcal{B}(Z)$.

For self-adjoint T on a Hilbert space one can extend this map to all $f \in C(\sigma(T))$, obtaining the *continuous functional calculus* in the next theorem. We set $p_1(z) = z$. Recall the $\sigma(T)$ is a compact non-empty subset of $\mathbb{R}$ by Theorems 1.16 and 4.7.

Theorem 4.11. *Let $T \in \mathcal{B}(X)$ be self-adjoint on a Hilbert space X. There exists exactly one map $\Phi_T : C(\sigma(T)) \to \mathcal{B}(X)$; $f \mapsto f(T)$, satisfying*

(C1) $(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T)$,

(C2) $\|f(T)\| = \|f\|_\infty$ *(hence, Φ_T is injective),*

(C3) $\mathbb{1}(T) = I$ *and* $p_1(T) = T$,

(C4) $(fg)(T) = f(T)g(T) = g(T)f(T)$,

(C5) $f(T)' = \overline{f}(T)$

for all $f, g \in C(\sigma(T))$ and $\alpha, \beta \in \mathbb{C}$. In particular, we have $\Phi_T(p) = p(T)$ for each polynomial p, where $p(T)$ is given by (4.5)*.*

Proof. 1) We first show the properties (C1)–(C5) for polynomials $p(t) = a_0 + a_1 t + \ldots + a_N t^N$ and $q(t) = b_0 + b_1 t + \ldots + b_N t^N$ with $t \in \mathbb{R}$ and the map $p \mapsto p(T)$ defined by (4.5), where any $a_j, b_j \in \mathbb{C}$ may be equal to 0. Clearly, (C1) and (C3) are true in this case, and $p(T)' = \sum_{j=0}^N \overline{a}_j (T^j)' = \overline{p}(T)$ due to (4.2) and $T = T'$. By means of (4.5), we further obtain

$$(pq)(T) = \sum_{l=0}^{2N} \Big(\sum_{\substack{0 \leqslant j,k \leqslant N \\ j+k=l}} a_j b_k \Big) T^l = \sum_{j=0}^{N} a_j T^j \sum_{k=0}^{N} b_k T^k = p(T)q(T),$$

and so $(pq)(T) = (qp)(T) = q(T)p(T)$; i.e., (C4) is shown for polynomials. Properties (C4) and (C5) yield the normality of $p(T)$. Hence, Proposition 4.3 and Lemma 4.12 below imply the core identity

$$\|p(T)\| = \mathrm{r}(p(T)) = \max\big\{|\lambda| \,\big|\, \lambda \in \sigma(p(T))\big\} = \max\big\{|\lambda| \,\big|\, \lambda \in p(\sigma(T))\big\} = \|p\|_\infty.$$

2) Let $f \in C(\sigma(T))$. Since $\sigma(T) \subseteq \mathbb{R}$ is compact, Weierstraß' approximation theorem yields real polynomials such that $p_n \to \operatorname{Re} f$ and $q_n \to \operatorname{Im} f$ in $C(\sigma(T))$ as $n \to \infty$, and thus $p_n + \mathrm{i} q_n \to f$. (Note that step 1) applies to $p_n + \mathrm{i} q_n$.) We can thus extend the map $p \mapsto p(T)$ to a linear isometry $\Phi_T : f \mapsto f(T)$ from $C(\sigma(T))$ to $\mathcal{B}(X)$. By continuity, Φ_T also satisfies (C4) and (C5) on $C(\sigma(T))$.

3) Let there be another map $\Psi : C(\sigma(T)) \to \mathcal{B}(X)$ satisfying (C1)–(C5). From (C1), (C3) and (C4) we then infer $\Psi(p) = p(T) = \Phi_T(p)$ for all polynomials, so that $\Psi = \Phi_T$ by continuity and density. □

We observe that we have actually shown uniqueness in the class of linear and continuous maps $\Psi : C(\sigma(T)) \to \mathcal{B}(X)$ fulfilling (C3) and (C4). The next result was used in the above proof. It is a special case of Theorem 5.3 (which is independent of Theorem 4.11, of course). One can show the lemma also in an elementary (but tedious) way, extending Proposition 1.20 f).

Lemma 4.12. *Let $T \in \mathcal{B}(Z)$ for a Banach space Z and let p be a polynomial. Then $\sigma(p(T)) = p(\sigma(T))$.*

We derive some rather direct, but important consquences of the theorem.

Corollary 4.13. *Let $T \in \mathcal{B}(X)$ be self-adjoint and $f \in C(\sigma(T))$. Then the following asssertions hold.*

(C6) Let $Tx = \lambda x$ for some $x \in X$ and $\lambda \in \mathbb{R}$. Then $f(T)x = f(\lambda)x$.

(C7) $f(T)$ is normal.

(C8) $\sigma(f(T)) = f(\sigma(T))$. (spectral mapping theorem)

(C9) $f(T)$ is self-adjoint if and only if f is real-valued.

Proof. Take a sequence of polynomials p_n converging to f uniformly. Let $Tx = \lambda x$. Property (C6) holds for a polynomial p since

$$p(T)x = \sum_{j=0}^{N} a_j T^j x = \sum_{j=0}^{N} a_j \lambda^j x = p(\lambda)x.$$

Using (C2), we then obtain

$$f(T)x = \lim_{n\to\infty} p_n(T)x = \lim_{n\to\infty} p_n(\lambda)x = f(\lambda)x.$$

From properties (C5) and (C4) we infer $f(T)f(T)' = f(T)\overline{f}(T) = \overline{f}(T)f(T) = f(T)'f(T)$ so that $f(T)$ is normal.

We next show (C8). Let $\mu \notin f(\sigma(T))$. The function $g := \frac{1}{\mu - f}$ then is an element of $C(\sigma(T))$. Thus (C3) and (C4) yield

$$(\mu I - f(T))g(T) = g(T)(\mu I - f(T)) = (g(\mu\mathbb{1} - f))(T) = \mathbb{1}(T) = I;$$

i.e., μ is an element of $\rho(f(T))$. Conversely, let $\mu = f(\lambda)$ for some $\lambda \in \sigma(T)$. Then $\mu_n := p_n(\lambda)$ belongs to $\sigma(p_n(T))$ for all $n \in \mathbb{N}$ by Lemma 4.12. As above, the operators $\mu_n I - p_n(T)$ tend to $\mu I - f(T)$ in $\mathcal{B}(X)$. Suppose that $\mu I - f(T)$ was invertible. Then also $\mu_n I - p_n(T)$ would be invertible for large n by Theorem 1.27. This is impossible, and so μ is contained in $\sigma(f(T))$.

For the last assertion, observe that $f(T) = f(T)'$ if and only if $(f - \overline{f})(T) = 0$ if and only if $f - \overline{f} = 0$, because Φ_T is injective. ☐

We use the functional calculus to solve the equation $W^n = T$ within the class self-adjoint 'non-negative' operators, where T is given. Below and in the exercises one finds more applications of this kind.

Corollary 4.14. *Let $n \in \mathbb{N}$ and $T = T' \in \mathcal{B}(X)$ with $\sigma(T) \subseteq \mathbb{R}_{\geqslant 0}$. (In this case one writes $T = T' \geqslant 0$ and calls T* non-negative*.) Then there is a unique self-adjoint operator $W \in \mathcal{B}(X)$ with $\sigma(W) \subseteq \mathbb{R}_{\geqslant 0}$ and $W^n = T$.*

Proof. Consider $w(t) = t^{1/n}$ for $t \in \sigma(T) \subseteq \mathbb{R}_{\geqslant 0}$ and define $W = w(T)$. Then $W^n = w^n(T) = p_1(T) = T$ by (C4) and (C3). Properties (C8) and (C9) imply that $W = W' \geqslant 0$. For the proof of the uniqueness of W, we refer to Korollar VII.1.16 in [**We**] or the exercises. ☐

We next study the special case of a compact self-adjoint operator T (or a self-adjoint A with compact resolvent). This *compact spectral theorem* provides a very convenient decomposition and representation of the operator and very strong spectral-theoretic results close to the finite-dimensional setting. The compact case is of great importance for many areas of mathematics and its applications. We use basic properties of orthogonal projections and orthonormal bases which are discussed in Chapter 3 of [**FA**].

THEOREM 4.15. *Let X be a Hilbert space with $\dim X = \infty$, $T \in \mathcal{B}(X)$ be compact and self-adjoint, and A be densely defined, closed and self-adjoint on X having a compact resolvent. Then the following assertions hold.*

a) i) There is an index set $J \in \{\emptyset, \mathbb{N}, \{1, \dots, N\} \,|\, N \in \mathbb{N}\}$ and eigenvalues $\lambda_j \neq 0$, $j \in J$, such that $\sigma(T) = \{0\} \dot\cup \{\lambda_j \,|\, j \in J\} \subseteq \mathbb{R}$, where $\lambda_j \to 0$ as $j \to \infty$ if $J = \mathbb{N}$.

ii) There is an orthonormal basis of $\mathrm{N}(T)^\perp = \overline{\mathrm{R}(T)}$ consisting of eigenvectors of T for the eigenvalues λ_j.

iii) The eigenspace $E_j(T) := \mathrm{N}(\lambda_j I - T)$ is finite-dimensional and the orthogonal projection P_j onto $E_j(T)$ commutes with T, for each $j \in J$, where $TP_jx = \lambda_j P_j x$.

iv) The sum $T = \sum_{j\in J} \lambda_j P_j$ converges in $\mathcal{B}(X)$.

b) i) We have $\sigma(A) = \sigma_{\mathrm{p}}(A) = \{\mu_n | n \in \mathbb{N}\} \subseteq \mathbb{R}$ with $|\mu_n| \to \infty$ as $n \to \infty$.

ii) There is an orthonormal basis of X consisting of eigenvectors of A.

iii) The eigenspaces $E_n(A) = \mathrm{N}(\mu_n I - A)$ are finite-dimensional and the orthogonal projections Q_n onto $E_n(A)$ satisfy $Q_n X \subseteq \mathrm{D}(A)$ and $Q_n Ax = AQ_n = \mu_n Q_n x$ for all $x \in \mathrm{D}(A)$ and $n \in \mathbb{N}$.

iv) The sum $Ax = \sum_{n=1}^\infty \mu_n Q_n x$ converges in X for all $x \in \mathrm{D}(A)$.

PROOF.[1] 1) Theorem 2.10, 2.15 and 4.7 show the parts i) in assertions a) and b), except for the infiniteness of $\sigma(A)$, as well as the finiteness of $\dim E_j(T)$ and $\dim E_n(A)$ for all j and n. Let $x, y \in \mathrm{D}(A)$ be eigenvectors of A for eigenvalues $\mu_n \neq \mu_k$. Then

$$\mu_n\,(x|y) = (Ax|y) = (x|Ay) = \mu_k(x|y),$$

so that $(x|y) = 0$. Similarly, one sees that $E_j(T) \perp E_k(T)$ if $j \neq k$. By the Gram–Schmidt procedure, each eigenspace $E_j(T)$ and $E_n(A)$ has an orthonormal basis of eigenvectors for $\lambda_j \neq 0$ and μ_n, respectively. The union of these bases gives orthonormal sets $\mathcal{B}_T$ and $\mathcal{B}_A$.

2) Let $J = \mathbb{N}$ in statement a), the other cases are treated similarly. Observe that $\mathcal{B}_T \subseteq \mathrm{R}(T)$. Set $\mathbb{1}_j = \mathbb{1}_{\{\lambda_j\}} \in C(\sigma(T))$ and $\varphi_n = \mathbb{1}_1 + \ldots + \mathbb{1}_n$ for every $j, n \in \mathbb{N}$. We then have $\mathbb{1}_j(T)^2 = \mathbb{1}_j^2(T) = \mathbb{1}_j(T)$ by (C4). Moreover, (C4), (C9) and (C3) imply that $T\mathbb{1}_j(T) = \mathbb{1}_j(T)T$, $\mathbb{1}_j(T)' = \mathbb{1}_j(T)$, and

$$(\lambda_j I - T)\mathbb{1}_j(T) = ((\lambda_j \mathbb{1} - p_1)\mathbb{1}_j)(T) = ((\lambda_j - \lambda_j)\mathbb{1}_j)(T) = 0.$$

If $\lambda_j v = Tv$ for some $v \in X$, we further deduce $\mathbb{1}_j(T)v = \mathbb{1}_j(\lambda_j)v = v$ from (C6). As a result, $\mathbb{1}_j(T)$ is a self-adjoint projection onto $E_j(T)$. For $x \in X$ and $y \in \mathrm{N}(\mathbb{1}_j(T))$, we then obtain $(\mathbb{1}_j(T)x|y) = (x|\mathbb{1}_j(T)y) = 0$ so that $\mathbb{1}_j(T)$ is orthogonal; i.e., $\mathbb{1}_j(T) = P_j$ and $\varphi_n(T) = P_1 + \ldots + P_n$ for all $j, n \in \mathbb{N}$. We have shown claim iii) in a).

[1]Part a) was not shown in the lectures. A more direct proof of it can be found in Theorem 6.7 of [**FA**].

Since $TP_j = \lambda_j P_j$, the operator $\varphi_n(T)T$ is a partial sum of the series in part iv). Employing also (C2), we thus derive iv) from

$$\|T - \varphi_n(T)T\| = \|(p_1 - \varphi_n p_1)(T)\| = \|p_1 - \varphi_n p_1\|_\infty = \sup_{j \geqslant n+1} |\lambda_j| \longrightarrow 0,$$

as $n \to \infty$. It also follows that $\varphi_n(T)y \in \operatorname{lin} \mathcal{B}_T$ converges to y as $n \to \infty$ for all $y \in \mathrm{R}(T)$. Therefore, $\mathcal{B}_T$ is an orthonormal basis of $\overline{\mathrm{R}(T)}$ due to Theorem 3.15 in [**FA**]. Finally, (2.1) shows that $\overline{\mathrm{R}(T)} = {}^\perp \mathrm{N}(T') = \mathrm{N}(T)^\perp$ because $T = T'$ and X is reflexive.

3) Fix $t \in \rho(A) \cap \mathbb{R}$. Then $R(t,A)' = R(t,A') = R(t,A)$ by (4.3), and this operator is compact and has a trivial kernel. By step 2), $X = \mathrm{N}(R(t,A))^\perp$ possesses an orthonormal basis of eigenvectors w of $R(t,A)$ for the eigenvalues $\lambda \neq 0$. Proposition 1.20 yields the eigenvector $v = \lambda R(t,A)w$ of A for the eigenvalue $\mu = t - \lambda^{-1}$. Because $\dim X = \infty$ and the eigenspaces of $R(t,A)$ are finite-dimensional, A has infinitely many distinct eigenvalues; i.e., part i) in assertion b) is shown.

Let $x \in X$ be orthogonal to all eigenvectors of A. For the above v, we obtain $0 = (x|v) = (x|\lambda R(t,A)w) = \lambda\,(R(t,A)x|w)$. Since the eigenvectors w span X, we infer that $R(t,A)x = 0$ and hence $x = 0$. Consequently, $\mathcal{B}_A$ is a basis of X by Theorem 3.15 in [**FA**], and claim ii) of b) is true.

4) Let $\{v_{n,1}, \ldots, v_{n,l_n}\}$ be eigenvectors of A forming an orthonormal basis of $E_n(A)$ for $n \in \mathbb{N}$. From step 3) we then deduce

$$Q_n x = \sum_{j=1}^{l_n} (x|v_{n,j})\, v_{n,j} \in \mathrm{D}(A) \qquad \text{and} \qquad x = \sum_{n=1}^{\infty} Q_n x$$

for all $x \in X$. For $x \in \mathrm{D}(A)$ it follows

$$\begin{aligned} Q_n A x &= \sum_{j=1}^{l_n} (Ax|v_{n,j})\, v_{n,j} = \sum_{j=1}^{l_n} (x|Av_{n,j})\, v_{n,j} = \sum_{j=1}^{l_n} (x|\mu_n v_{n,j})\, v_{n,j} \\ &= \sum_{j=1}^{l_n} (x|v_{n,j})\, \mu_n v_{n,j} = \sum_{j=1}^{l_n} (x|v_{n,j})\, A v_{n,j} = A Q_n X = \mu_n Q_n x. \end{aligned}$$

We thus conclude

$$A \sum_{k=1}^{n} Q_k x = \sum_{k=1}^{n} Q_k A x \longrightarrow \sum_{k=1}^{\infty} Q_k A x = \sum_{k=1}^{\infty} A Q_k x = \sum_{k=1}^{\infty} \mu_k Q_k x,$$

as $n \to \infty$, so that the closedness of A yields the last assertion. □

Remark 4.16. a) In the above proof we also obtain that

$$Ax = \sum_{n=1}^{\infty} \mu_n \sum_{j=1}^{l_n} (x|v_{n,j})\, v_{n,j}$$

for all $x \in \mathrm{D}(A)$. An analogous result holds for T, see Theorem 6.7 in [**FA**].

b) Let T be self-adjoint and compact such that $\mathrm{N}(T)$ is separable. Theorem 3.15 yields an orthonormal basis $\{z_k \,|\, k \in J_0\}$ of $\mathrm{N}(T)$, where $J_0 \subseteq \mathbb{Z}_{\leqslant 0}$ could be empty. (J_0 must be infinite if J from Theorem 4.15 is finite, as $\dim X = \infty$.) Denote by λ_l the non-zero eigenvalues of T (repeated according to their multiplicity) with corresponding orthonormal basis of eigenvectors

$\{w_l \,|\, l \in J_1\}$. The union $\{b_j \,|\, j \in J'\}$ of $\{z_k \,|\, k \in J_0\}$ and $\{w_l \,|\, l \in J_1\}$ is an orthonormal basis of X. By Theorem 3.18 in [**FA**], the map

$$\Phi : X \to \ell^2(J'); \quad \Phi x = \big((x|b_j)\big)_{j\in J'},$$

is unitary with $\Phi^{-1}((\xi_j)_{j\in J'}) = \sum_{j\in J'} \xi_j b_j$. Moreover, the transformed operator $\Phi T\Phi^{-1}$ acts on $\ell^2(J')$ as the multiplication operator

$$\Phi T\Phi^{-1}(\xi_j) = \Phi T \sum_{j\in J'} \xi_j b_j = \Phi \sum_{j\in J'} \lambda_j \xi_j b_j = (\lambda_j\xi_j)_{j\in J'},$$

where $\lambda_j := 0$ if $j \in J_0$. Hence, $\Phi T\Phi^{-1}$ is represented as an infinite diagonal matrix with diagonal elements λ_j. Analogous results hold for A from Theorem 4.15, see e.g. Theorems 4.5.1-3 in [**Tr**]. $\Diamond$

In part a) of the above remark, A is written as a sum over projections. In the non-compact case one can show an analogous result using an integral over projections, see Remark 4.19. In view of part b), T is unitarily equivalent to the multiplication operator $M : (\xi_j)_j \mapsto (\lambda_j\xi_j)_j$ on $\ell^2(J')$ which can be viewed as an L^2-space on $\sigma(T)$. This is a strong statement since M is a rather simple object. We next extend this *multiplicator representation* to the general *bounded self-adjoint case*. Besides its inherent interest, this fact will allow us to pass to unbounded $A = A'$. In the second part of the proof one sees that Ω is a 'disjoint sum' of $\sigma(T)$ and that $h = p_1$, roughly speaking. Thus the information on T is mainly encoded in the measure μ which is essentially given by (4.6) using the functional calculus.

THEOREM 4.17. *Let $T \in \mathcal{B}(X)$ be self-adjoint on a separable Hilbert space X. Then there is a σ–finite measure space $(\Omega, \mathcal{A}, \mu)$, a measurable function $h : \Omega \to \sigma(T)$ and a unitary operator $U : X \to L^2(\mu)$ such that*

$$Tx = U^{-1}hUx \qquad \text{for all } \; x \in X.$$

PROOF. 1) Let $v_1 \in X\backslash\{0\}$. We define the linear subspaces

$$Y_1 = \{f(T)v_1 \,|\, f \in C(\sigma(T))\} \quad \text{and} \quad X_1 = \overline{Y}_1$$

of X. Since $Tf(T)v_1 = (p_1 f)(T)v_1 \in Y_1$ for every $f \in C(\sigma(T))$, we obtain $TY_1 \subseteq Y_1$ and so $TX_1 \subseteq X_1$. We introduce the map

$$\varphi_1 : C(\sigma(T)) \to \mathbb{C}; \quad \varphi_1(f) = (f(T)v_1|v_1),$$

which is linear and bounded because $|\varphi_1(f)| \leqslant \|f(T)\|\|v_1\|^2 = \|f\|_\infty\|v_1\|^2$ due to (C2). If $f \geqslant 0$, then $\sigma(f(T)) \subseteq \mathbb{R}_{\geqslant 0}$ by (C8). So we can deduce from Corollary 4.14 that

$$(f(T)v_1|v_1) = \big(f(T)^{1/2}v_1\big|f(T)^{1/2}v_1\big) = \|f(T)^{1/2}v_1\|^2 \geqslant 0.$$

The Riesz representation theorem of $C(\sigma(T))^\star$ now gives a bounded regular[2] measure μ_1 on $\mathcal{B}(\sigma(T))$ such that

$$(f(T)v_1|v_1) = \varphi_1(f) = \int_{\sigma(T)} f \,\mathrm{d}\mu_1, \tag{4.6}$$

[2]A positive measure μ on a Borel σ-algebra $\mathcal{B}$ is called *regular*, if it satisfies

$$\mu(B) = \inf\{\mu(\mathcal{O}) \,|\, B \subseteq \mathcal{O},\ \mathcal{O} \text{ is open}\} = \sup\{\mu(K) \,|\, K \subseteq B, K \text{ is compact}\}, \quad B \in \mathcal{B}.$$

for all $f \in C(\sigma(T)) \subseteq L^2(\mu_1)$, see Theorem 2.14 of [**Ru1**]. For $x = f(T)v_1 \in Y_1$, we define $V_1 x := f \in L^2(\mu_1)$. We compute

$$\|V_1x\|^2_{L^2(\mu_1)} = \int_{\sigma(T)} |f|^2 \,\mathrm{d}\mu_1 = \varphi_1(\overline{f}f) = \big((\overline{f}f)(T)v_1|v_1\big)$$
$$= \big(f(T)'f(T)v_1|v_1\big) = (f(T)v_1|f(T)v_1) = \|x\|^2_X.$$

In particular, if $x = f(T)v_1 = g(T)v_1$ for some $g \in C(\sigma(T))$, then

$$\|f-g\|_2^2 = \|(f(T)-g(T))v_1\|_X^2 = 0,$$

and so $f = g$ in $L^2(\mu_1)$. As a result, $V_1 : Y_1 \to L^2(\mu_1)$ is a linear isometric map and can be extended to a linear isometry $U_1 : X_1 \to L^2(\mu_1)$.

Observe that $C(\sigma(T)) \subseteq \mathrm{R}(U_1)$. Since $C(\sigma(T))$ is dense in $L^2(\mu_1)$ by Theorem 3.14 in [**Ru1**], the isometry U_1 has dense range. Hence, U_1 is bijective and thus unitary by Proposition 5.52 in [**FA**]. Finally, we compute

$$U_1Tf(T)v_1 = V_1(p_1f)(T)v_1 = p_1f = p_1U_1f(T)v_1,$$

for $f \in C(\sigma(T))$. By density, it follows $Tx = U_1^{-1}p_1U_1x$ for all $x \in X_1$.

2)[3] We are done if there is a vector $v_1 \in X$ with $X_1 = X$. In general this is not true. Using Zorn's Lemma (see Corollary 2.5 in Appendix 2 of [**La**]), we instead find orthogonal spaces X_n as in step 1) which span X.

a) To that aim, we introduce the collection $\mathcal{E}$ of all sets E having as elements at most countably many closed subspaces $X_j \subset X$ of the type constructed in step 1) such that $X_i \perp X_j$ for all $X_i \neq X_j$ in E. The system $\mathcal{E}$ is ordered via inclusion of sets. Let $\mathcal{C}$ be a chain in $\mathcal{E}$; i.e., a subset of $\mathcal{E}$ such that $E \subseteq F$ or $F \subseteq E$ for all $E, F \in \mathcal{C}$. We put $C = \bigcup_{E\in\mathcal{C}} E$. Clearly, E is contained in C for all $E \in \mathcal{C}$. Let $Y, Z \in C$. Then Y and Z are closed subspaces of X as constructed in part 1) and there are sets $E, F \in \mathcal{C}$ with $Y \in E$ and $Z \in F$. We may assume that $E \subseteq F$ and so $X, Y \in F$. The subspaces Y and Z are thus orthogonal (if $Y \neq Z$). As a result, C contains pairwise orthogonal subspaces of X. If $x \perp y$ have norm 1, then $\|x-y\|^2 = \|x\|^2 + \|y\|^2 = 2$. The separability of X then implies that at most countably many subspaces belong to C, so that C is an element of $\mathcal{E}$ and hence an upper bound of $\mathcal{C}$. Zorn's Lemma now gives a maximal element $M = \{X_j \,|\, j \in J\}$ in $\mathcal{E}$, where $J \subseteq \mathbb{N}$ and X_j are pairwise orthogonal subspaces as constructed in 1).

b) Assume that there was a vector $z \in X$ being orthogonal to all X_j. Let Z be the closed span of all vectors $f(T)z$ with $f \in C(\sigma(T))$. Let $g \in C(\sigma(T))$ and $x = g(T)v_i \in X_i$ for some $i \in J$, where v_i generates X_i as in step 1). We then obtain

$$(x|f(T)z) = (f(T)'g(T)v_i|z) = ((\overline{f}g)(T)v_i|z) = 0$$

since $(\overline{f}g)(T)v_i \in X_i$. By density, it follows that Z is orthogonal to all X_j and thus $M \cup \{Z\} \in \mathcal{E}$. The maximality on M now implies that Z belongs to M, implying $z = 0$. Consequently, X is the closed linear span of the elements in the orthogonal subspaces X_j. Each $x \in X$ can thus be uniquely written as $x = \sum_j x_j$ for some $x_j \in X_j$, and we have $\|x\|^2 = \sum_j \|x_j\|^2$ by Pythagoras.

[3]Not shown in the lectures.

c) We now define

$$\Omega = \bigcup_{j \in J} \sigma(T) \times \{j\} \subseteq \mathbb{R}^2, \quad \mathcal{A} = \mathcal{B}(\Omega), \quad \mu(A) = \sum_{j \in J} \mu_j(A_j)$$

with $A_j \times \{j\} = A \cap (\sigma(T) \times \{j\})$, and

$$h(\lambda, j) = \lambda, \qquad Ux = \sum_{j \in J} \tilde{U}_j x_j,$$

where $(\lambda, j) \in \Omega$, U_j and μ_j are given as in step 1), and we set $(\tilde{U}_j x_j)(\lambda, j) = (U_j x_j)(\lambda)$.

It is straightforward to check that $(\Omega, \mathcal{A}, \mu)$ is a σ–finite measure space, and the map h is even continuous for the Euclidean metric. The definition of μ and step 1) show the isometry

$$\|Ux\|^2_{L^2(\mu)} = \sum_j \|U_j x_j\|^2_{L^2(\mu_j)} = \sum_j \|x_j\|^2 = \|x\|^2.$$

To show the surjectivity of U, take $g \in L^2(\mu)$. For the restictions $g_j(\lambda) = g(\lambda, j)$ part 1) gives vectors $x_j \in X_j$ with $g_j = U_j x_j$ for $j \in J$, and hence $g = U \sum_j x_j$. As a result, U is unitary by Proposition 5.52 in [**FA**]. Let $T_j : X_j \to X_j$ be the restriction of T. Since $Tx = \sum_j T_j x_j$, we derive

$$\begin{aligned}(hUx)(\lambda, j) &= \sum_j \lambda(\tilde{U}_j x_j)(\lambda, j) = \sum_j (p_1 U_j x_j)(\lambda) = \sum_j (U_j T_j x_j)(\lambda) \\ &= U\Big(\sum_j T_j x_j\Big)(\lambda, j) = UTx(\lambda, j). \qquad \square\end{aligned}$$

We add an observation to the above proof. Let $\lambda \in \sigma(T) \backslash \sigma_{\mathrm{p}}(T)$. Then $\Lambda = \{\lambda\} \times J \subseteq \Omega$ is a μ–null set. In fact, otherwise the characteristic function f of any subset of Λ with measure in $\mathbb{R}_+$ would be a non-zero element of $L^2(\mu)$. Hence, $x = U^{-1} f \neq 0$ would be an eigenvector of T for the eigenvalue λ, since $Tx = U^{-1}\lambda f = \lambda x$. We further note that in the proof of Theorem 4.17 one could also take $\Omega = \bigcup_{j \in J} \sigma(T_j) \times \{j\}$, where $T_j = T{\restriction}_{X_j}$. It can be shown that $\sigma(T)$ is the closure of $\bigcup_{j \in J} \sigma(T_j)$.

The above representation of bounded self-adjoint operators as multiplication operators now leads to a *multiplication representation* and to a M_b–*functional calculus* for (possibly) *unbounded self-adjoint* operators A. Here $M_b(\sigma(A))$ is the Banach space of bounded Borel-measurable functions on $\sigma(A)$ endowed with the supremum norm. We use this space instead of $L^\infty(\sigma(A))$ to avoid certain technical problems. We set $r_\lambda(z) = (\lambda - z)^{-1}$ for $z \in \mathbb{C}\backslash\{\lambda\}$ and write (C3') replacing in (C3) the map p_1 by r_λ for $\lambda \in \rho(A)$.

Theorem 4.18. *Let A be a self-adjoint operator on a separable Hilbert space X. Then the following assertions hold.*

a) There is a σ–finite measure space $(\Omega, \mathcal{A}, \mu)$, a measurable function $h : \Omega \to \sigma(A)$ and a unitary operator $U : X \to L^2(\mu)$ such that

$$\mathrm{D}(A) = \{x \in X \,|\, hUx \in L^2(\mu)\} \quad \text{and} \quad Ax = U^{-1} hUx.$$

b) There is a contractive map $\Psi_A : M_b(\sigma(A)) \to \mathcal{B}(X)$; $\Psi_A(f) = f(A)$, satisfying (C1) and (C3')–(C5). Moreover, if $f_n \in M_b(\sigma(A))$ are uniformly bounded and converge to $f \in M_b(\sigma(T))$ pointwise, then $f_n(A)x \to f(A)x$ as $n \to \infty$ for all $x \in X$. Finally, for $x \in \mathrm{D}(A)$ and $f \in M_b(\sigma(A))$ the vector $f(A)x$ belongs to $\mathrm{D}(A)$ and $Af(A)x = f(A)Ax$.

Proof. a) We additionally assume that $\sigma(A) \neq \mathbb{R}$ and fix $t \in \rho(A) \cap \mathbb{R}$.[4] Then $R(t,A) \in \mathcal{B}(X)$ is self-adjoint and can be represented as $R(t,A) = U^{-1}mU$ on a space $L^2(\Omega,\mu)$ as in Theorem 4.17. Recall that Proposition 1.20 yields $\sigma(A) = t - [\sigma(R(t,A))\backslash\{0\}]^{-1}$. Set

$$h(\lambda, j) = t - \frac{1}{m(\lambda, j)} = t - \frac{1}{\lambda} \in \sigma(A),$$

for $j \in J$ and $\lambda \in \sigma(R(t,A))\backslash\{0\}$. The sets $\{0\} \times J$ have μ–measure 0 in view of the obervation before the theorem, due to the injectivity of $R(t,A)$. We can thus extend h to a measurable function $h : \Omega \to \sigma(A)$, by setting $h(0,j) = \mu$ for some $\mu \in \sigma(A)$.

Let $x \in \mathrm{D}(A)$. We put $y = tx - Ax \in X$. Using $x = R(t,A)y = U^{-1}mUy$, we compute

$$hUx = hmUy = (tm - \mathbb{1})Uy \in L^2(\mu),$$
$$U^{-1}hUx = tU^{-1}mUy - y = tx - y = Ax.$$

If $x \in X$ satisfies $hUx \in L^2(\mu)$, then we put $y = U^{-1}(t\mathbb{1} - h)Ux \in X$ and obtain $mUy = (tm - mh)Ux = Ux$. Therefore, $x = U^{-1}mUy = R(t,A)y$ belongs to $\mathrm{D}(A)$, and part a) is proved.

b) We define $\Psi_A : f \mapsto f(A)$ by

$$f(A)x = U^{-1}(f \circ h)Ux \tag{4.7}$$

for $f \in M_b(\sigma(A))$ and $x \in X$. We further set $M_f\varphi = (f \circ h)\varphi$ for $\varphi \in L^2(\mu)$. It is straightforward to check that $f(A) \in \mathcal{B}(X)$, Ψ_A is linear, $\mathbb{1}(A) = I$ and (C5) is true. Let $\lambda \in \rho(A)$. We have

$$hUr_\lambda(A)x = h(r_\lambda \circ h)Ux = h(\lambda\mathbb{1} - h)^{-1}Ux \in L^2(\mu)$$

for all $x \in X$. So step a) yields that $r_\lambda(A)X \subseteq \mathrm{D}(A)$ and $(\lambda I - A)r_\lambda(A) = I$. Similarly, one sees that $r_\lambda(A)(\lambda x - Ax) = x$ for all $x \in \mathrm{D}(A)$, and thus (C3') is shown. The contractivity follows from $\|f(A)\| = \|M_f\| \leqslant \|f\|_\infty$. For property (C4) we observe that

$$(fg)(A)x = U^{-1}(f \circ h)(g \circ h)Ux = U^{-1}(f \circ h)UU^{-1}(g \circ h)Ux = f(A)g(A)x,$$

where $f, g \in M_b(\sigma(A))$ and $x \in X$.

Let $f, f_n \in M_b(\sigma(A))$ be uniformly bounded by c such that $f_n \to f$ pointwise as $n \to \infty$. For every $x \in X$, we have $f_n(A)x - f(A)x = U^{-1}((f_n - f) \circ h)Ux$. Since $(f_n - f) \circ h \to 0$ pointwise and $|((f_n - f) \circ h)Ux| \leqslant 2c|Ux|$, Lebesgue's convergence theorem shows that $((f_n - f) \circ h)Ux$ tends to 0 in $L^2(\mu)$ and so $f_n(A)x \to f(A)x$ in X as $n \to \infty$.

Let $x \in \mathrm{D}(A)$. The above results yield that

$$g := h(f \circ h)Ux = (f \circ h)UU^{-1}hUx = (f \circ h)UAx \in L^2(\mu).$$

On the other hand, $g = hUU^{-1}(f \circ h)Ux = hUf(A)x$. Part (a) thus implies that $f(A)x \in \mathrm{D}(A)$ and

$$Af(A)x = U^{-1}hUf(A)x = U^{-1}g = U^{-1}(f \circ h)UAx = f(A)Ax. \qquad \square$$

[4]If $\sigma(A) = \mathbb{R}$, we instead take $t = \mathrm{i}$ and use below the version of Theorem 4.17 for the normal operator $R(\mathrm{i}, A)$ given in Satz VII.1.25 in [**We**]. Part b) was not shown in the lectures.

The mapping Ψ_A is uniquely determined, see Theorem VIII.5 in [**RS**]. There are versions of Theorem 4.18 b) for unbounded f given in Theorems 13.24 and 13.30 of [**Ru2**]. If we have $A = T \in \mathcal{B}(X)$ in the above theorem, then the definition of Ψ_A in (4.7) yields $p_1(T) = T$.

We briefly sketch a third version of the spectral theorem using the *spectral measure* of $A = A'$.

Remark 4.19. Let A be a self-adjoint operator on a separable Hilbert space X. Set $P_S = \mathbb{1}_S(A)$ for a Borel set $S \subseteq \sigma(A)$. Since $\mathbb{1}_S^2 = \mathbb{1}_S = \overline{\mathbb{1}_S}$, the contractive operator P_S is a self-adjoint projection by the M_b-calculus. Take $x \in \mathrm{N}(P_s)$ and $y = P_S z \in \mathrm{R}(P_S)$. We then obtain $(x|y) = (x|P_S z) = (P_S x|z) = 0$ so that the *spectral projection* P_S is orthogonal. Let $x, y \in X$. Using Theorem 4.18 b), one sees that $S \mapsto (P_S x|y)$ is a '$\mathbb{C}$-valued' measure $\mu_{x,y}$ (see Appendix C in [**Co2**]). Moreover, for $f \in M_b(\sigma(A))$ one can show

$$(f(A)x|y) = \int_{\sigma(A)} f(\lambda)\,\mathrm{d}\mu_{x,y}(\lambda)$$

and moreover

$$(Ax|y) = \int_{\sigma(A)} \lambda\,\mathrm{d}\mu_{x,y}(\lambda)$$

for $x \in \mathrm{D}(A) = \{x \in X \mid \int_{\sigma(A)} |\lambda|^2\,\mathrm{d}\mu_{x,x} < \infty\}$ and $y \in X$. One can further define an operator-valued integral $\int f\,\mathrm{d}P \in \mathcal{B}(X)$ such that

$$\int_{\sigma(A)} f(\lambda)\,\mathrm{d}\mu_{x,y}(\lambda) = \Big(\Big(\int_{\sigma(A)} f\,\mathrm{d}P\Big)x\,\Big|\,y\Big)$$

for $x, y \in X$. See Theorem VIII.6 in [**RS**] and Sections VII.1+3 of [**We**]. $\Diamond$

We conclude with one of the most important applications of the above theorem.

Example 4.20. Let H be a self-adjoint operator on a (separable) Hilbert space X. For a given $u_0 \in \mathrm{D}(H)$ we claim that there is exactly one function $u \in C^1(\mathbb{R}, X) \cap C(\mathbb{R}, [\mathrm{D}(H)])$ solving the *Schrödinger equation*

$$\tfrac{\mathrm{d}}{\mathrm{d}t} u(t) = -\mathrm{i}Hu(t), \quad t \in \mathbb{R}, \qquad u(0) = u_0. \tag{4.8}$$

(H is called Hamiltonian.) The solution is given by $u(t) = T(t)u_0$ for unitary operators $T(t)$ on X satisfying $T(0) = I$, $T(t+s) = T(t)T(s) = T(s)T(t)$ and $T(t)^{-1} = T(-t)$ for $t, s \in \mathbb{R}$. Moreover, $t \mapsto T(t)x \in X$ continuous for $t \in \mathbb{R}$ and all $x \in X$. An example for this setting is $X = L^2(\mathbb{R}^3)$ and $H = -(\Delta + \frac{b}{|x|_2})$ with $\mathrm{D}(H) = W^{2,2}(\mathbb{R}^3)$, see Example 4.10.

Proof. For $t \in \mathbb{R}$, we consider the bounded function $f_t : \mathbb{R} \to \mathbb{C}$; $f_t(\xi) = \mathrm{e}^{-\mathrm{i}t\xi}$. Theorem 4.18 allows us to define $T(t) = f_t(H) \in \mathcal{B}(X)$. Since $f_0 = \mathbb{1}$ and $f_t f_s = f_{t+s}$, we obtain $T(0) = I$ and $T(t)T(s) = T(s)T(t)$ for $t, s \in \mathbb{R}$. With $s = -t$ it follows that $T(t)$ has the inverse $T(-t)$. Moreover, $T(t)$ is unitary since $T(t)' = \overline{f_t}(H) = f_{-t}(H) = T(-t)$. Because of $\|f_t\|_\infty = 1$ and the continuity of $t \mapsto f_t(\xi)$ for fixed $\xi \in \mathbb{R}$, the map $\mathbb{R} \ni t \mapsto T(t)z \in X$ is continuous for each $x \in X$ by Theorem 4.18.

Let $u_0 \in \mathrm{D}(H)$. We set $y = \tau u_0 - Hu_0$ for some $\tau \in \rho(H)$, so that $u_0 = R(\tau, H)y = r_\tau(H)y$. We then obtain

$$\tfrac{1}{t-s}(T(t)u_0 - T(s)u_0) = \tfrac{1}{t-s}(f_t(H) - f_s(H))r_\tau(H)y$$

$$= \left(\tfrac{1}{t-s}(f_t - f_s)r_\tau\right)(H)y =: g_{t,s}(H)y$$

for all $t \neq s$. Observe that $g_{t,s}(\xi) \to \frac{-\mathrm{i}\xi}{\tau-\xi} f_s(\xi) =: m(\xi)f_s(\xi)$ as $t \to s$ for all $\xi \in \sigma(H)$ and $\|g_{t,s}\|_\infty \leqslant \|m\|_\infty = \sup_{\xi\in\sigma(H)}|\frac{\xi}{\tau-\xi}| < \infty$ for all $t \neq s$. So Theorem 4.18 shows that there exists $\frac{\mathrm{d}}{\mathrm{d}t}T(t)u_0 = m(H)T(t)y = T(t)m(H)y$. We further compute

$$\begin{aligned} m(H)y &= U(m\circ h)U^{-1}y = U((-\mathrm{i}p_1 r_\tau)\circ h)U^{-1}y \\ &= -U\mathrm{i}hU^{-1}U(r_\tau\circ h)U^{-1}y = -\mathrm{i}HR(\tau,H)y = -\mathrm{i}Hu_0 \end{aligned}$$

by means of Theorem 4.18. Hence, we arrive at

$$\tfrac{\mathrm{d}}{\mathrm{d}t}T(t)u_0 = -\mathrm{i}T(t)Hu_0 = -\mathrm{i}HT(t)u_0,$$

using Theorem 4.18 once more. Due to these equations, $u = T(\cdot)u_0$ belongs to $C^1(\mathbb{R},X) \cap C(\mathbb{R},[\mathrm{D}(H)])$ and solves (4.8).

Let $v \in C^1(\mathbb{R},X) \cap C(\mathbb{R}_{\geqslant 0},[\mathrm{D}(H)])$ be another solution of (4.8). For $t,s \in \mathbb{R}$ and $h \neq 0$ we compute

$$\begin{aligned} &\tfrac{1}{h}(T(t-s-h)v(s+h) - T(t-s)v(s)) - T(t-s)(v'(s) + \mathrm{i}Hv(s)) \\ &= T(t-s-h)\left(\tfrac{1}{h}(v(s+h)-v(s)) - v'(s)\right) + (T(t-s-h) - T(t-s))v'(s) \\ &\quad - \tfrac{1}{-h}(T(t-s-h) - T(t-s))v(s) - T(t-s)\mathrm{i}Hv(s) \longrightarrow 0 \end{aligned}$$

as $h \to 0$. For any $y \in X$ we thus obtain $\frac{\mathrm{d}}{\mathrm{d}s}\left(T(t-s)v(s)|y\right) = 0$, since $v' = -\mathrm{i}Hv$. Consequently,

$$(T(t)x|y) = (T(t)v(0)|y) = (T(0)v(t)|y) = (v(t)|y)\,,$$

which gives $u(t) = v(t)$ for all $t \geqslant 0$. Thus the 'strongly continuous unitary group' $(T(t))_{t\in\mathbb{R}}$ solves (4.8) uniquely. $\square$

CHAPTER 5

Holomorphic functional calculi

We come back to the case of Banach spaces X and Y, but keep $\mathbb{F} = \mathbb{C}$. We want to introduce functional calculi for non self-adjoint operators on X, now using complex curve integrals.

5.1. The bounded case

Let $U \subseteq \mathbb{C}$ be open, $g : U \to Y$ be holomorphic (i.e., complex differentiable), and $\gamma \in C([a,b],U)$ be a piecewise C^1-curve (in U) with range $\Gamma = \gamma([a,b]) \subseteq U$. This means that there are $a = a_0 < a_1 < \cdots < a_N = b$ such that the restrictions of γ to $[a_{k-1}, a_k]$ are C^1. We define the *curve integral*

$$\int_\gamma g \,\mathrm{d}z = \int_a^b g(\gamma(t))\gamma'(t)\,\mathrm{d}t := \sum_{k=1}^N \int_{a_{k-1}}^{a_k} g(\gamma(t))\gamma'(t)\,\mathrm{d}t$$

as a Banach space-valued Riemann integral (having the same definition, results and proofs as for $Y = \mathbb{R}$ in Analysis 2). Using Riemann sums, one checks the usual properties of curve integrals and also that $T\int_\gamma g\,\mathrm{d}z = \int_\gamma Tg\,\mathrm{d}z$ for all $T \in \mathcal{B}(Y,Z)$ and Banach spaces Z.

Let $\gamma_j : [a_j, b_j] \to U$ be piecewise C^1-curves for $j \in \{1,2\}$ such that either $b_1 < a_2$ or $b_1 = a_2$ and $\gamma_1(b_1) = \gamma_2(a_2)$. On $[a_1,b_1] \cup [a_2,b_2]$ we define the 'sum curve' $\gamma_1 \cup \gamma_2(t) = \gamma_j(t)$ for $t \in [a_j,b_j]$. If $b_1 < a_2$, the ranges Γ_1 and Γ_2 can be disjoint, and we call also such curves piecewise C^1.

The index of a closed curve (i.e., $\gamma(a) = \gamma(b)$) at $z \in \mathbb{C}\backslash\Gamma$ is given by

$$\mathrm{n}(\gamma, z) = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{\mathrm{d}w}{w - z}.$$

The index is the number of times that γ winds around z, counted with orientation $\pm$. (See Analysis 4 for basic properties of the index.)

Let γ be closed and piecewise C^1 in U such that $\mathrm{n}(\gamma, z) = 0$ for all $z \notin U$. Then Cauchy's integral theorem and formula

$$\int_\gamma g\,\mathrm{d}z = 0, \tag{5.1}$$

$$\mathrm{n}(\gamma,z)g(z) = \frac{1}{2\pi\mathrm{i}}\int_\gamma \frac{g(w)}{w-z}\,\mathrm{d}w, \tag{5.2}$$

are valid for all $z \in U\backslash\Gamma$. This fact is shown for $Y = \mathbb{C}$ in Theorems IV.5.4 and IV.5.7 of [**Co1**]. For a general Banach space Y, the formulas (5.1) and (5.2) are thus satisfied by functions $z \mapsto \langle g(z), y^\star\rangle$ for every $y^\star \in Y^\star$. Hence, $\langle \int_\gamma g\,\mathrm{d}z, y^\star\rangle = 0$ for all $y^\star \in Y^\star$ so that a corollary of the Hahn–Banach theorem yields (5.1) in Y. Similarly one deduces (5.2).

For compact non-empty subsets $K \subseteq \mathbb{C}$, we introduce the space

$$H(K) = \big\{ f : \mathrm{D}(f) \to \mathbb{C} \,\big|\, K \subseteq \mathrm{D}(f) \subseteq \mathbb{C},\ \mathrm{D}(f) \text{ is open, } f \text{ is holomorphic} \big\}.$$

Let $K \subseteq U \subseteq \mathbb{C}$, K be compact, and U be open. By Proposition VIII.1.1 in [**Co1**] and its proof there exists an *admissible curve* γ *for* K *and* U (or, *in* $U \backslash K$) which means that $\gamma : [a, b] \to U \backslash K$ is piecewise C^1, $\mathrm{n}(\gamma, z) = 1$ for all $z \in K$, and $\mathrm{n}(\gamma, z) = 0$ for all $z \in \mathbb{C} \backslash U$.

Let $T \in \mathcal{B}(X)$, $f \in H(\sigma(T))$, and γ be admissible for $\sigma(T)$ and $\mathrm{D}(f)$. We then define

$$f(T) = \frac{1}{2\pi \mathrm{i}} \int_\gamma f(\lambda) R(\lambda, T) \,\mathrm{d}\lambda \ \in \ \mathcal{B}(X). \tag{5.3}$$

This integral exists in $\mathcal{B}(X)$ since the integrand $\lambda \mapsto f(\lambda)R(\lambda, T)$ is holomorphic on $\rho(T) \cap \mathrm{D}(f) \supseteq \Gamma$ and Γ is compact. Writing $R(\lambda, T)$ as '$\frac{1}{\lambda - T}$', one sees the similarity of (5.3) and (5.2), but $R(\lambda, T)$ does not exist on the possibly 'large' set $\sigma(T)$, whereas the map $w \mapsto \frac{1}{w - z}$ is defined on $\mathbb{C} \backslash \{z\}$.

Let γ' be another admissible curve for $\sigma(T)$ and $\mathrm{D}(f)$. We set $\gamma'' = \gamma \cup (-\gamma')$, where "$-$" denotes the inversion of the orientation. We then have

$$\mathrm{n}(\gamma'', z) = \mathrm{n}(\gamma, z) - \mathrm{n}(\gamma', z) = \begin{cases} 1 - 1 = 0, & z \in \sigma(T), \\ 0 - 0 = 0, & z \in \mathbb{C} \backslash \mathrm{D}(f). \end{cases}$$

So we can apply (5.1) on $U = \mathrm{D}(f) \backslash \sigma(T)$ obtaining

$$0 = \int_{\gamma''} f(\lambda) R(\lambda, T) \,\mathrm{d}\lambda = \int_\gamma f(\lambda) R(\lambda, T) \,\mathrm{d}\lambda - \int_{\gamma'} f(\lambda) R(\lambda, T) \,\mathrm{d}\lambda.$$

Consequently, (5.3) does not depend on the choice of the admissible curve.

We recall that $p_1(z) = z$ for $z \in \mathbb{C}$. For $f \in H(\sigma(T))$ and $T \in \mathcal{B}(X)$ we next establish the *holomorphic functional calculus* which has very similar properties as the continuous calculus from Theorem 4.11 for $T = T'$. However, now the functions have to be defined on a neighborhood of $\sigma(T)$, and they have to be holomorphic and not just continuous.

THEOREM 5.1. *Let* $T \in \mathcal{B}(X)$ *and* $f, g, f_n \in H(\sigma(T))$ *with* $\mathrm{D}(f_n) = \mathrm{D}(f)$ *for* $n \in \mathbb{N}$. *Then the map*

$$\Phi_T : H(\sigma(T)) \to \mathcal{B}(X); \quad f \mapsto f(T),$$

defined by (5.3) *is linear and satisfies*

(H1) $\|f(T)\| \leqslant c \sup_{\lambda \in \Gamma} |f(\lambda)|$ *for a constant* $c = c(\gamma, T) > 0$,

(H2) $\mathbb{1}(T) = I$, $p_1(T) = T$,

(H3) $f(T)g(T) = g(T)f(T) = (fg)(T)$,

(H4) $f(T)^\star = f(T^\star)$,

(H5) if $f_n \to f$ *uniformly on compact subsets of* $\mathrm{D}(f)$, *then* $f_n(T) \to f(T)$ *in* $\mathcal{B}(X)$ *as* $n \to \infty$,

(H6) if $f(\lambda) \neq 0$ *for all* $\lambda \in \sigma(T)$, *then* $\frac{1}{f} \in H(\sigma(T))$ *and* $f(T)$ *has the inverse* $\frac{1}{f}(T)$.

Moreover, Φ_T *is the only linear map from* $H(\sigma(T))$ *to* $\mathcal{B}(X)$ *satisfying (H1)–(H3). For a polynomial* p, *the operators* $p(T)$ *in* (5.3) *and in* (4.5) *coincide. If* X *is a Hilbert space and* $T = T'$, *the above* Φ_T *is the restriction of the map* Φ_T *from Theorem 4.11.*

Proof. It is clear that $f \mapsto f(T)$ is linear. Property (H1) follows from

$$\|f(T)\| \leqslant \frac{1}{2\pi}\ell(\gamma)\sup_{\lambda\in\Gamma}\|R(\lambda,T)\| \sup_{\lambda\in\Gamma}|f(\lambda)| =: c(\gamma,T) \sup_{\lambda\in\Gamma}|f(\lambda)|.$$

Replacing here $f(T)$ by $f(T) - f_n(T) = (f - f_n)(T)$ we also deduce (H5). To check (H4), we recall that $\sigma(T) = \sigma(T^\star)$ and $R(\lambda,T)^\star = R(\lambda,T^\star)$ from Theorem 1.24. Hence,

$$f(T)^\star = \frac{1}{2\pi\mathrm{i}}\int_\gamma f(\lambda)R(\lambda,T)^\star\,\mathrm{d}\lambda = f(T^\star).$$

We next show (H3). We choose a bounded open set $U \subseteq \mathbb{C}$ with $\sigma(T) \subseteq U \subseteq \overline{U} \subseteq \mathrm{D}(f) \cap \mathrm{D}(g)$ and admissible curves γ_f in $U\backslash\sigma(T)$ and γ_g in $(\mathrm{D}(f) \cap \mathrm{D}(g))\backslash\overline{U}$. We then have $\mathrm{n}(\gamma_f,\mu) = 0$ for all $\mu \in \Gamma_g \subseteq \mathbb{C}\backslash U$ and $\mathrm{n}(\gamma_g,\lambda) = 1$ for all $\lambda \in \Gamma_f \subseteq U$. Using the resolvent equation, Fubini's theorem in $\mathcal{B}(X)$ (see Theorem X.6.16 in [**AE**]) and (5.2) in $\mathbb{C}$, we compute

$$\begin{aligned}
f(T)g(T) &= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_f} f(\lambda)R(\lambda,T)\frac{1}{2\pi\mathrm{i}}\int_{\gamma_g} g(\mu)R(\mu,T)\,\mathrm{d}\mu\,\mathrm{d}\lambda\\
&= \left(\frac{1}{2\pi\mathrm{i}}\right)^2\int_{\gamma_f}\int_{\gamma_g} f(\lambda)g(\mu)\frac{1}{\mu-\lambda}(R(\lambda,T) - R(\mu,T))\,\mathrm{d}\mu\,\mathrm{d}\lambda\\
&= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_f} f(\lambda)R(\lambda,T)\frac{1}{2\pi\mathrm{i}}\int_{\gamma_g}\frac{g(\mu)}{\mu-\lambda}\,\mathrm{d}\mu\,\mathrm{d}\lambda\\
&\quad + \frac{1}{2\pi\mathrm{i}}\int_{\gamma_g} g(\mu)R(\mu,T)\frac{1}{2\pi\mathrm{i}}\int_{\gamma_f}\frac{f(\lambda)}{\lambda-\mu}\,\mathrm{d}\lambda\,\mathrm{d}\mu\\
&= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_f} f(\lambda)g(\lambda)R(\lambda,T)\,\mathrm{d}\lambda = (fg)(T).
\end{aligned}$$

This identity also yields $(fg)(T) = (gf)(T) = g(T)f(T)$.

To check (H2), we take $f = \mathbb{1}$ with $D(f) = \mathbb{C}$. We choose the circle $\gamma_0(t) = 2\|T\|\mathrm{e}^{it}$ for $t \in [0,2\pi]$. Theorem 1.16 then leads to

$$\begin{aligned}
\mathbb{1}(T) &= \frac{1}{2\pi\mathrm{i}}\int_{\gamma_0} R(\lambda,T)\,\mathrm{d}\lambda = \frac{1}{2\pi\mathrm{i}}\int_{\gamma_0}\sum_{n=0}^\infty T^n\lambda^{-n-1}\,\mathrm{d}\lambda\\
&= \sum_{n=0}^\infty T^n\frac{1}{2\pi\mathrm{i}}\int_{\gamma_0}\lambda^{-n-1}\,\mathrm{d}\lambda = I,
\end{aligned}$$

since the series converges in $\mathcal{B}(X)$ uniformly on Γ_0 and $\int_{\gamma_0}\lambda^{-k}\,\mathrm{d}\lambda$ is equal to $2\pi\mathrm{i}$ if $k = 1$ and equal to 0 for $k \in \mathbb{Z}\backslash\{1\}$. The property $p_1(T) = T$ is shown similarly.

If $f(\lambda) \neq 0$ for $\lambda \in \sigma(T)$, by continuity f is non-zero on some open set $\mathrm{D}(\frac{1}{f})$ containing $\sigma(T)$. Therefore, $\frac{1}{f}$ belongs to $H(\sigma(T))$ and (H6) follows from (H2) and (H3) since $\mathbb{1} = f\frac{1}{f}$.

Let $\Psi : H(\sigma(T)) \to \mathcal{B}(X)$ be linear and satisfy (H1)–(H3). The linearity, (H2) and (H3) imply that $\Psi(p) = p(T) = \Phi_T(p)$ for every polynomial p. Moreover, Ψ also fulfills (H6) and hence $\Psi(r) = p(T)q(T)^{-1} = \Phi_T(r)$ for rational $r = \frac{p}{q}$ in $H(\sigma(T)$. Let $f \in H(\sigma(T))$. Runge's Theorem VIII.1.8 in [**Co1**] yields a bounded open set U with $\sigma(T) \subseteq U \subseteq \overline{U} \subseteq \mathrm{D}(f)$ and rational

$r_n \in H(\overline{U})$ tending uniformly to f on $\overline{U}$ as $n \to \infty$. Taking an admissible γ in $U \backslash \sigma(T)$, we see that $r_n(T)$ tends to $\Psi(f)$ and $\Phi_T(f)$ in $\mathcal{B}(X)$ by (H1), and thus $\Psi = \Phi_T$. The last claim is shown similarly, using just polynomials. $\square$

We first compute $f(T)$ for power series f and then for the simplest class of T, namely multiplication operators $T = M$.

EXAMPLE 5.2. a) Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ with convergence radius $\rho > 0$ and $T \in \mathcal{B}(X)$ with $\mathrm{r}(T) < \rho$. Since $p_N(z) := \sum_{n=0}^{N} a_n z^n$ tends to f locally uniformly, Theorem 5.1 shows $f(T) = \sum_{n=0}^{\infty} a_n T^n$ with convergence in $\mathcal{B}(X)$. This representation as a power series does not work in the situation of Theorem 5.5 (where f only has local power series); but even if it works, it may be better to use the calculus (see Example 5.4).

b) Let $E = C(K)$ for a compact set $K \subseteq \mathbb{R}^d$ and let $m \in C(K)$. We define $M\varphi = m\varphi$ for $\varphi \in E$. Proposition 1.14 shows that $M \in \mathcal{B}(E)$, $\sigma(M) = m(K)$, and $R(\lambda, M)\varphi = \frac{1}{\lambda - m}\varphi$ for all $\lambda \in \rho(M)$.

Let $f \in H(m(K))$, γ be an admissible curve in $D(f)\backslash m(K)$, $\varphi \in E$, and $x \in K$. Using that the map $\psi \mapsto \psi(x)$ is continuous and linear from E to $\mathbb{C}$ and Cauchy's formula (5.2) for $z = m(x)$, we compute

$$[f(M)\varphi](x) = \Big[\frac{1}{2\pi \mathrm{i}} \int_\gamma f(\lambda) R(\lambda, M)\varphi \, \mathrm{d}\lambda\Big](x) = \frac{1}{2\pi \mathrm{i}} \int_\gamma f(\lambda)(R(\lambda, M)\varphi)(x) \, \mathrm{d}\lambda$$
$$= \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(\lambda)}{\lambda - m(x)} \, \mathrm{d}\lambda \, \varphi(x) = f(m(x))\varphi(x).$$

As a result, $f(M)\varphi = (f \circ m)\varphi$ is also a multiplication operator. $\Diamond$

Also for the holomorphic calculus we show the *spectral mapping theorem.*

THEOREM 5.3. *Let $T \in \mathcal{B}(X)$ and $f \in H(\sigma(T))$. We then have*

$$\sigma(f(T)) = f(\sigma(T)).$$

PROOF. Let $\mu \notin f(\sigma(T))$. Then $g = \mu\mathbb{1} - f$ is nowhere on $\sigma(T)$ and so $g(T)$ is the inverse of $\mu I - f(T)$ by (H6). Hence, μ belongs to $\rho(f(T))$.

Conversely, let $\mu = f(\lambda)$ for some $\lambda \in \sigma(T)$. We set $h(z) = \frac{f(z) - \mu}{z - \lambda}$ for $z \in \mathrm{D}(f)\backslash\{\lambda\}$ and $h(\lambda) = f'(\lambda)$. Since h is bounded, it is holomorphic on $\mathrm{D}(f)$ by Riemann's theorem on removable singularities. We have $h(z)(z - \lambda) = f(z) - \mu$ for all $z \in \mathrm{D}(f)$, and so the calculus yields

$$(\lambda I - T)h(T) = h(T)(\lambda I - T) = (h(\lambda\mathbb{1} - p_1))(T) = (\mu\mathbb{1} - f)(T) = \mu I - f(T).$$

As the operator $(\lambda I - T)$ is not surjective or not injective, $\mu I - f(T)$ is not bijective; i.e, μ is contained in $\sigma(f(T))$. $\square$

We now use the spectral mapping theorem to study the long-term behavior of differential equations. For $X = \mathbb{C}^m$ and matrices A, we reprove results from Analysis 4 on linear ordinary differential equations.

EXAMPLE 5.4. Let $A \in \mathcal{B}(X)$. For $t \in \mathbb{R}$ we set $f_t : \mathbb{C} \to \mathbb{C}$; $f_t(z) = \mathrm{e}^{tz}$, and define $\mathrm{e}^{tA} = f_t(A) \in \mathcal{B}(X)$. As in Example 4.20, one sees that

$$\mathrm{e}^{(t+s)A} = \mathrm{e}^{tA}\mathrm{e}^{sA} = \mathrm{e}^{sA}\mathrm{e}^{tA}, \quad \mathrm{e}^{0A} = I, \quad \exists \, [\mathrm{e}^{tA}]^{-1} = \mathrm{e}^{-tA}$$

for all $t,s\in\mathbb{R}$. Moreover, $t\mapsto \mathrm{e}^{tA}$ belongs to $C^1(\mathbb{R},\mathcal{B}(X))$ with

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = A\mathrm{e}^{tA} = \mathrm{e}^{tA}A.$$

Hence, the map $u(t)=\mathrm{e}^{tA}u_0$ is the unique solution in $C^1(\mathbb{R},X)$ of

$$\tfrac{\mathrm{d}}{\mathrm{d}t}u(t) = Au(t), \quad t\in\mathbb{R}, \qquad u(0)=u_0,$$

where $u_0\in X$ is given. (Uniqueness is shown as in Example 4.20.) Theorem 5.3 further yields

$$\mathrm{r}(\mathrm{e}^{tA}) = \max\big\{|\mu| \,\big|\, \mu\in\sigma(\mathrm{e}^{tA}) = \mathrm{e}^{t\sigma(A)}\big\} = \max\big\{\mathrm{e}^{t\,\mathrm{Re}\,\lambda} \,\big|\, \lambda\in\sigma(A)\big\} = \mathrm{e}^{t\mathrm{s}(A)}$$

for the *spectral bound* $\mathrm{s}(A) := \max\{\mathrm{Re}\,\lambda \mid \lambda\in\sigma(A)\}$. Therefore, if $\mathrm{s}(A)<0$ (i.e., $\sigma(A)\subseteq\mathbb{C}_-$), then we deduce from Theorem 1.16 that

$$1 > \mathrm{r}(e^A) = \lim_{n\to\infty}\|(e^A)^n\|^{1/n} = \lim_{n\to\infty}\|\mathrm{e}^{nA}\|^{1/n}.$$

So we can fix an index $N\in\mathbb{N}$ with $\|\mathrm{e}^{NA}\| =: q < 1$. Writing any given $t\geqslant 0$ as $t = kN+\tau$ for some $k\in\mathbb{N}_0$ and $0\leqslant\tau<N$, we estimate

$$\|\mathrm{e}^{tA}\| = \|(\mathrm{e}^{NA})^k\mathrm{e}^{\tau A}\| \leqslant q^k\,\|\mathrm{e}^{\tau A}\| \leqslant \max_{0\leqslant\tau\leqslant N}\|\mathrm{e}^{\tau A}\|\exp\big(Nk\tfrac{\ln q}{N}\big) \leqslant M\mathrm{e}^{-wt}$$

where $\omega := -\frac{\ln q}{N} > 0$ and $M := \max_{0\leqslant\tau\leqslant N}\|\mathrm{e}^{\tau A}\|\,\mathrm{e}^{|\ln q|}$. So spectral information on the given operator A implies the exponential decay $\|u(t)\|\leqslant M\mathrm{e}^{-\omega t}\|u_0\|$, $t\geqslant 0$, of the solutions u. ◊

In the next theorem we discuss *spectral projections* in the present setting. We first need some preparations. Let $S\in\mathcal{B}(X)$ and $P=P^2\in\mathcal{B}(X)$ be a projection with $SP=PS$. Set $X_1=\mathrm{R}(P)$ and $X_2=\mathrm{N}(P)$. Lemma 2.16 of [**FA**] then yields the direct sum $X=X_1\oplus X_2$. Moreover, if $y=Px\in\mathrm{R}(P)$, then $Sy=SPx=PSx$ also belongs to $\mathrm{R}(P)$. If $x\in\mathrm{N}(P)$, then $PSx = SPx = 0$ so that also Sx is an element of $\mathrm{N}(P)$. As a result, S leaves invariant X_1 and X_2, and the restrictions $S\restriction_{X_j}\in\mathcal{B}(X_j)$ are well defined.

Theorem 5.5. *Let $T\in\mathcal{B}(X)$ and $\sigma(T)=\sigma_1\dot\cup\sigma_2$ for two disjoint closed sets $\sigma_j\neq\emptyset$ in $\mathbb{C}$. Then there is a projection $P\in\mathcal{B}(X)$ such that $f(T)P = Pf(T)$ for $f\in H(\sigma(T))$ and $\sigma(T_j)=\sigma_j$ for $j\in\{1,2\}$, where $T_j = T\restriction_{X_j}\in\mathcal{B}(X_j)$, $X_1=\mathrm{R}(P)$ and $X_2=\mathrm{N}(P)$. We further have $X=X_1\oplus X_2$ and $R(\lambda,T_j)=R(\lambda,T)\restriction_{X_j}$ for $\lambda\in\rho(T)=\rho(T_1)\cap\rho(T_2)$. The projection is given by*

$$P = \frac{1}{2\pi\mathrm{i}}\int_{\gamma_1} R(\lambda,T)\,\mathrm{d}\lambda, \tag{5.4}$$

where γ_1 is an admissible curve for σ_1 and any open $U_1\supseteq\sigma_1$ with $\overline{U_1}\cap\sigma_2=\emptyset$.

Proof. There are open sets U_j with $\overline{U}_1\cap\overline{U}_2=\emptyset$ and $\sigma_j\subseteq U_j$ for $j\in\{1,2\}$. Define $h\in H(\sigma(T))$ by $h=1$ on U_1 and $h=0$ on U_2. We set $P=h(T)\in\mathcal{B}(X)$. Ptroperties (H3) and (H2) then yield $P^2=h^2(T)=h(T)=P$ and $f(T)P=Pf(T)$ for all $f\in H(\sigma(T))$. As seen above, $X=X_1\oplus X_2$ and the operators $T_j=T\restriction_{X_j}\in\mathcal{B}(X_j)$ are well defined.

Formula (5.4) follows by choosing $\gamma=\gamma_1\cup\gamma_2$, where γ_j are admissible curves for σ_j and U_j for $j\in\{1,2\}$. Let $\lambda\notin\sigma_1$. We may shrink U_1 so that $\lambda\notin U_1$ since P does not depend on the choice of γ. We define $g(z)=\frac{1}{\lambda-z}$ for $z\in U_1$ and $g(z)=0$ for $z\in U_2$. Then g belongs to $H(\sigma(T))$ and satisfies

$$g(T)(\lambda I - T) = (\lambda I - T)g(T) = ((\lambda\mathbb{1} - p_1)g)(T) = h(T) = P.$$

Setting $R = g(T)\restriction_{X_1} \in \mathcal{B}(X_1)$, we thus obtain

$$R(\lambda I_{X_1} - T_1) = (\lambda I_{X_1} - T_1)R = I_{X_1}.$$

This means that $\lambda \in \rho(T_1)$, and so $\sigma(T_1) \subseteq \sigma_1$. Similarly, one sees that $\sigma(T_2) \subseteq \sigma_2$. In particular, $\sigma(T_1)$ and $\sigma(T_2)$ are disjoint.

Let $\lambda \in \rho(T_1) \cap \rho(T_2)$. For $x \in X$, we have unique $x_j \in X_j$ with $x = x_1 + x_2$. If $\lambda x - Tx = 0$, then $0 = \lambda x_1 - T_1 x_1 + \lambda x_2 - T_2 x_2 \in X_1 \oplus X_2$ so that x_j is contained in $\mathrm{N}(\lambda I - T_j) = \{0\}$ for $j \in \{1,2\}$; i.e., $x = 0$. Given $y \in X$, we define $x_j = R(\lambda, T_j)y_j \in X_j$ for $j \in \{1,2\}$. Setting $x = x_1 + x_2$, we derive

$$\lambda x - Tx = \lambda x_1 - T_1 x_1 + \lambda x_2 - T_2 x_2 = y_1 + y_2 = y.$$

We have proved that $\lambda \in \rho(T)$, $R(\lambda, T)\restriction_{X_j} = R(\lambda, T_j)$, and

$$\sigma_1 \dot\cup \sigma_2 = \sigma(T) \subseteq \sigma(T_1) \dot\cup \sigma(T_2) \subseteq \sigma_1 \dot\cup \sigma_2.$$

Together with $\sigma(T_j) \subseteq \sigma_j$, it follows $\sigma(T_j) = \sigma_j$ for $j \in \{1,2\}$. □

We use the above concept to refine the results in Example 5.4 about the long-term behavior of e^{tA}, by studying its *exponential dichotomy*.

Example 5.6. In the setting of Example 5.4, assume that $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. We thus obtain closed sets $\sigma_1 \subseteq \mathbb{C}_-$ and $\sigma_2 \subseteq \mathbb{C}_+$ with $\sigma(A) = \sigma_1 \dot\cup \sigma_2$. Let P be the spectral projection of A for σ_1. We define A_1 and A_2 as the restrictions of A to $X_1 = \mathrm{R}(P)$ and $X_2 = \mathrm{N}(P)$, respectively, as in Theorem 5.5. Then $\mathrm{e}^{tA}\restriction_{X_j} = \mathrm{e}^{tA_j} : X_j \to X_j$ and there are constants $\delta, N > 0$ with

$$\|\mathrm{e}^{tA_1}\| \leqslant N\mathrm{e}^{-\delta t} \quad \text{and} \quad \|\mathrm{e}^{-tA_2}\| \leqslant N\mathrm{e}^{-\delta t}, \qquad t \geqslant 0.$$

In other words, X can be decomposed into e^{tA}-invariant subspaces on which e^{tA} decays exponentially in forward and in backward time, respectively.

Proof. Let γ, γ_1, and h be given as in Theorem 5.5 and its proof. For $x \in X_1$ we compute

$$\begin{aligned} \mathrm{e}^{tA}x = \mathrm{e}^{tA}Px = (f_t h)(A)x &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma_1} \mathrm{e}^{t\lambda} R(\lambda, A)x \,\mathrm{d}\lambda \\ &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma_1} \mathrm{e}^{t\lambda} R(\lambda, A_1)x \,\mathrm{d}\lambda = \mathrm{e}^{tA_1}x, \end{aligned}$$

where $f_t(\lambda) = \mathrm{e}^{t\lambda}$ for $t \in \mathbb{R}$. In the same way one derives $\mathrm{e}^{tA}x = \mathrm{e}^{tA_2}x$ for all $x \in X_2$ and $t \in \mathbb{R}$. Since $\sigma(A_1) = \sigma_1$, we obtain $\mathrm{s}(A_1) < 0$ and hence Example 5.4 shows that $\|\mathrm{e}^{tA}x_1\| \leqslant M\mathrm{e}^{-\omega t}\|x_1\|$ for all $t \geqslant 0$ and $x_1 \in X_1$ and some constants $M, \omega > 0$.

We have $\sigma(A_2) = \sigma_2$ and so $\mathrm{s}(-A_2) < 0$. Note that the curve $\tilde\gamma = -\gamma$ is admissible for $\sigma(-A) = -\sigma(A)$. Substituting $\mu = -\lambda$, we conclude that

$$\mathrm{e}^{-tA} = \frac{1}{2\pi \mathrm{i}} \int_{\gamma} \mathrm{e}^{-t\lambda} (\lambda I - A)^{-1} \,\mathrm{d}\lambda = \frac{1}{2\pi \mathrm{i}} \int_{\tilde\gamma} \mathrm{e}^{t\mu} (\mu I - (-A))^{-1} \,\mathrm{d}\mu = \mathrm{e}^{t(-A)}$$

for all $t \in \mathbb{R}$. For $x_2 \in X_2$ we thus obtain

$$\mathrm{e}^{-tA_2}x_2 = \mathrm{e}^{-tA}x_2 = \mathrm{e}^{t(-A)}x_2 = \mathrm{e}^{t(-A_2)}x_2$$

so that $\|\mathrm{e}^{-tA_2}x_2\| \leqslant M'\mathrm{e}^{-\omega' t}\|x_2\|$ for all $t \geqslant 0$ and some $M', \omega' > 0$. □

5.2. Sectorial operators

We extend the above results to certain unbounded operators A, restricting ourselves to the exponential e^{tA}.[1] For $\phi \in (0,\pi)$ we define the open *sector*

$$\Sigma_\phi = \{\lambda \in \mathbb{C}\backslash\{0\} \,|\, |\arg \lambda| < \phi\}.$$

Note that $\Sigma_\pi = \mathbb{C}\backslash\mathbb{R}_{\leqslant 0}$ and $\Sigma_{\pi/2} = \mathbb{C}_+$.

Definition 5.7. *A closed operator A is called* sectorial of angle $\phi \in (0,\pi)$ *if there is a constant $K > 0$ such that $\Sigma_\phi \subseteq \rho(A)$ and*

$$\|R(\lambda, A)\| \leqslant \frac{K}{|\lambda|}, \qquad \lambda \in \Sigma_\phi.$$

In the literature several small variations of the above definition are used. Note that a sectorial operator of angle ϕ is also sectorial of angle $\phi' \in (0,\phi)$. In applications often arise operators A such that $A-\omega I$ is sectorial for some $\omega \in \mathbb{R}$, cf. Remark 5.14. We discuss several core examples.

Example 5.8. Let A be self-adjoint on the Hilbert space X. We further suppose that $\sigma(A) \subseteq \mathbb{R}_{\leqslant 0}$. Then A is sectorial of every angle $\phi < \pi$.

Proof. Let $\phi \in (\frac{\pi}{2}, \pi)$ and $\lambda \in \Sigma_\phi$. Since $R(\lambda, A)' = R(\overline{\lambda}, A)$ by (4.3), the operator $R(\lambda, A)$ is normal. Propositions 4.3 and 1.20 then yield

$$\|R(\lambda, A)\| = \mathrm{r}(R(\lambda, A)) = \mathrm{d}(\lambda, \sigma(A))^{-1} \leqslant \mathrm{d}(\lambda, \mathbb{R}_{\leqslant 0})^{-1} = \begin{cases} \frac{1}{|\lambda|}, & \operatorname{Re}\lambda \geqslant 0, \\ \frac{1}{|\operatorname{Im}\lambda|}, & \operatorname{Re}\lambda < 0. \end{cases}$$

If $\operatorname{Re}\lambda < 0$, we can write $\lambda = |\lambda|\mathrm{e}^{\pm \mathrm{i}\theta}$ for some $\theta \in (\frac{\pi}{2}, \phi)$. We then have $\frac{|\operatorname{Im}\lambda|}{|\lambda|} = |\sin\theta| \geqslant \sin\phi > 0$, and thus

$$\|R(\lambda, A)\| \leqslant \frac{1/\sin\phi}{|\lambda|} =: \frac{K_\phi}{|\lambda|}, \qquad \lambda \in \Sigma_\phi. \qquad \square$$

Note that $K_\phi \to \infty$ as $\phi \to \pi$ in the above example.

Example 5.9. Let $X = C([0,1])$ and $Au = u''$ with $D(A) = \{u \in C^2([0,1]) \,|\, u(0) = u(1) = 0\}$. Then A is sectorial of every angle $\phi < \pi$.

Proof. We recall from Example 2.16 that $\Sigma_\pi \subseteq \rho(A)$ and that for $\lambda = \mu^2$ with $\mu \in \mathbb{C}_+$ the resolvent is given by

$$R(\mu^2, A)f(t) = a(f,\mu)\mathrm{e}^{\mu t} + b(f,\mu)\mathrm{e}^{-\mu t} + \frac{1}{2\mu}\int_0^1 \mathrm{e}^{-\mu|t-s|} f(s)\,\mathrm{d}s$$

for $t \in [0,1]$, $f \in X$, and the numbers

$$\begin{pmatrix} a(f,\mu) \\ b(f,\mu) \end{pmatrix} = \frac{1}{2\mu(\mathrm{e}^{-\mu} - \mathrm{e}^{\mu})} \begin{pmatrix} \mathrm{e}^{-\mu}\int_0^1 (\mathrm{e}^{\mu s} - \mathrm{e}^{-\mu s}) f(s)\,\mathrm{d}s \\ \int_0^1 (\mathrm{e}^{\mu}\mathrm{e}^{-\mu s} - \mathrm{e}^{-\mu}\mathrm{e}^{\mu s}) f(s)\,\mathrm{d}s \end{pmatrix}.$$

Fix $\phi \in (\frac{\pi}{2}, \pi)$. Take $\lambda \in \Sigma_\phi$. We obtain $\mu \in \Sigma_{\phi/2}$ and thus $\mu = |\mu|\,\mathrm{e}^{\mathrm{i}\theta}$ with $0 \leqslant |\theta| < \phi/2$ and $\operatorname{Re}\mu = |\mu|\cos\theta \geqslant |\mu|\cos\phi/2$. So we can estimate

$$\|R(\lambda, A)f\|_\infty \leqslant |a(f,\mu)|\,\mathrm{e}^{\operatorname{Re}\mu} + |b(f,\mu)| + \frac{\|f\|_\infty}{2|\mu|} \sup_{t\in[0,1]} \int_{t-1}^{t} \mathrm{e}^{-\operatorname{Re}\mu|r|}\,\mathrm{d}r$$

[1] See [**KW**] for a detailed study of a corresponding functional calculus.

$$
\begin{aligned}
&\leqslant \frac{\|f\|_\infty}{2|\mu|(\mathrm{e}^{\operatorname{Re}\mu}-\mathrm{e}^{-\operatorname{Re}\mu})}\Big(\int_0^1 \big(\mathrm{e}^{\operatorname{Re}\mu s}+\mathrm{e}^{-\operatorname{Re}\mu s}\big)\,\mathrm{d}s \\
&\qquad + \int_0^1 (\mathrm{e}^{\operatorname{Re}\mu}\mathrm{e}^{-\operatorname{Re}\mu s}+\mathrm{e}^{-\operatorname{Re}\mu}\mathrm{e}^{\operatorname{Re}\mu s})\,\mathrm{d}s\Big) + \frac{\|f\|_\infty}{|\mu|\operatorname{Re}\mu} \\
&= \frac{\|f\|_\infty}{2\operatorname{Re}\mu|\mu|(\mathrm{e}^{\operatorname{Re}\mu}-\mathrm{e}^{-\operatorname{Re}\mu})}\big((\mathrm{e}^{\operatorname{Re}\mu}-1+1-\mathrm{e}^{-\operatorname{Re}\mu}) \\
&\qquad + \mathrm{e}^{\operatorname{Re}\mu}(1-\mathrm{e}^{-\operatorname{Re}\mu})+\mathrm{e}^{-\operatorname{Re}\mu}(\mathrm{e}^{\operatorname{Re}\mu}-1)\big) + \frac{\|f\|_\infty}{|\mu|\operatorname{Re}\mu} \\
&\leqslant \frac{\frac{1}{\cos(\phi/2)}}{|\mu|^2}\|f\|_\infty\Big(\frac{(\mathrm{e}^{\operatorname{Re}\mu}-\mathrm{e}^{-\operatorname{Re}\mu})+(\mathrm{e}^{\operatorname{Re}\mu}-\mathrm{e}^{-\operatorname{Re}\mu})}{2(\mathrm{e}^{\operatorname{Re}\mu}-\mathrm{e}^{-\operatorname{Re}\mu})}+1\Big) \\
&= \frac{\frac{2}{\cos(\phi/2)}}{|\lambda|}\|f\|_\infty. \qquad \square
\end{aligned}
$$

Note that $\overline{D(A)} = \{u \in X \,|\, u(0) = u(1) = 0\}$ is not equal to X for the above 'Dirichlet–Laplacian', cf. Example 1.19 in [**FA**].

Similarly one can show that the 'Neumann–Laplacian' $A_1 u = u''$ with

$$D(A_1) = \big\{u \in C^2([0,1]) \,\big|\, u'(0) = u'(1) = 0\big\}$$

is sectorial for every angle $\phi < \pi$ with on $X = C([0,1])$. Moreover, its spectrum is given by $\sigma(A_1) = \sigma_{\mathrm{p}}(A_1) = \{-\pi^2 k^2 \,|\, k \in \mathbb{N}_0\}$ with eigenfunctions $u_k(t) = \cos(k\pi t)$. (See exercises.) Here $D(A_1)$ is dense in X.

Example 5.10. Let $X = L^p(\mathbb{R})$, $1 \leqslant p < \infty$, and $Au = \partial u$ for $\mathrm{D}(A) = W^{1,p}(\mathbb{R})$. Then A is sectorial of every angle $\phi < \frac{\pi}{2}$.

Proof. Example 3.43 says that $\sigma(A) = \mathrm{i}\mathbb{R}$ and $\|R(\lambda, A)\| \leqslant \frac{1}{\operatorname{Re}\lambda}$ for $\operatorname{Re}\lambda > 0$. If $\phi \in (0, \frac{\pi}{2})$ and $\lambda \in \Sigma_\phi$, we have $|\operatorname{Re}\lambda| \geqslant |\lambda|\cos\phi$ and hence

$$\|R(\lambda, A)\| \leqslant \frac{1/\cos\phi}{|\lambda|}.$$

Because of its spectrum, A is not sectorial of angle $\phi \geqslant \frac{\pi}{2}$. $\square$

For a typical elliptic partial differential operator we now sketch how one can construct sectorial operators on $L^2(U)$ via 'form methods', cf. Chapter 6 of [**Ka**] or Chapter 1 of [**Ou**].

Example 5.11. Let $U \subseteq \mathbb{R}^m$ be open and bounded, and the coefficients $a_{jk} \in L^\infty(U, \mathbb{C})$ for $j, k \in \{1, \ldots, m\}$ be *strictly accretive*; i.e.,

$$\operatorname{Re}\sum\nolimits_{j,k=1}^m a_{jk}(x) z_j \overline{z_k} \geqslant \eta |z|_2^2 \tag{5.5}$$

for some $\eta > 0$, all $z \in \mathbb{C}^m$, and a.e. $x \in U$. We write $a = (a_{jk})_{j,k}$. Let $E = L^2(U)$ and $V = W_0^{1,2}(U)$, where $(\cdot|\cdot) = (\cdot|\cdot)_{L^2}$ and $\|\cdot\| = \|\cdot\|_2$. Using Poincaré's equality in Theorem 3.36, we equip V with the equivalent norm $\||\nabla v|_2\|_2 =: \|v\|_V$. We now define the sesquilinear form

$$\underline{a} : V \times V \to \mathbb{C}; \quad \underline{a}(v, w) = \sum\nolimits_{j,k=1}^m \int_U \partial_j v\, a_{jk} \partial_k \overline{w}\,\mathrm{d}x.$$

Note that $\underline{a}$ is bounded; i.e., $|\underline{a}(v, w)| \leqslant \|a\|_\infty \|v\|_V \|w\|_V$. Each $f \in E$ yields an element ψ of $V^\star$ given by $\psi(v) = -\int_U v\overline{f}\,\mathrm{d}x = -(v|f)$. The *Lax–Milgram lemma* Theorem 1.51 in [**EE**] thus provides a unique function $u_f \in V$ such that $\underline{a}(v, u_f) = \psi(v)$ for all $v \in V$. Moreover, the map $\psi \mapsto u_f$ is antilinear, and hence $f \mapsto u_f$ is linear. We now define

$$\mathrm{D}(A) = \big\{u \in V \,\big|\, \exists\, w \in E \;\forall v \in V : -\underline{a}(v, u) = (v|w)\big\}, \quad Au = w.$$

Therefore $u = u_f$ belongs to $\mathrm{D}(A)$ and satisfies $Au = f$ and $-\underline{a}(v, u) = (v|Au) = (v|f)$ for all $v \in V$. Moreover, $A : \mathrm{D}(A) \to E$ is bijective. Using Theorem 3.36 with constant $c =: c_P$ and (5.5), we estimate $u = A^{-1}f$ by

$$\|u\|^2 \leqslant c_P^2 \||\nabla u|_2\|^2 \leqslant c_P^2 \eta^{-1} \operatorname{Re} \underline{a}(u, u) \leqslant c_P^2 \eta^{-1} |(u|f)| \leqslant c_P^2 \eta^{-1} \|u\| \|f\|$$

for $f \in E$; i.e., $A^{-1} : E \to E$ is bounded. In particular, A is closed.

We show that $\mathrm{D}(A)$ is dense in E. Indeed, take $\varphi \in E$ with $(v|\varphi) = 0$ for all $v \in \mathrm{D}(A)$. We have to check that $\varphi = 0$ in view of the projection Theorem 3.8 in [**FA**]. Inserting $\psi := A^{-1}\varphi \in \mathrm{D}(A)$, we derive

$$0 = (\psi|A\psi) = -\underline{a}(\psi, \psi) = -\operatorname{Re} \underline{a}(\psi, \psi) \leqslant -\eta \|\psi\|_V^2$$

from (5.5), so that ψ and thus φ are zero.

Note that the adjoint matrix $a' := \overline{a}^T$ also satisfies (5.5) and that its associated form a' on V is given by $\underline{a}'(v, w) = \overline{\underline{a}(w, v)}$. It thus induces an invertible operator $\tilde{A}$, too. For $w \in \mathrm{D}(\tilde{A})$ and $u \in \mathrm{D}(A)$, we infer

$$(Au|w) = \overline{(w|Au)} = -\overline{\underline{a}(w, u)} = -\underline{a}'(u, w) = (u|\tilde{A}w);$$

i.e., $\tilde{A} \subseteq A'$. Their invertibility implies $\tilde{A} = A'$ by Remark 4.5. In particular, A is self-adjoint if and only if $a' = a$, e.g., if a is real and symmetric.

To obtain sectoriality, we take $v \in V$ and $u \in \mathrm{D}(A)$ with $\|u\| = 1$. The properties of $\underline{a}$ imply

$$|\operatorname{Im} \underline{a}(v, v)| \leqslant |\underline{a}(v, v)| \leqslant \|a\|_\infty \|v\|_V^2 \leqslant \eta^{-1} \|a\|_\infty \operatorname{Re} \underline{a}(v, v),$$

and so $\underline{a}(v, v)$ belongs to $\overline{\Sigma_\theta}$ with $\theta := \arctan(\eta^{-1}\|a\|_\infty) \in (0, \pi/2)$. Fix $\varphi \in (\theta, \pi/2)$ and set $\phi = \pi - \varphi > \pi/2$. Take $\lambda \in \Sigma_\phi$. Then $-\lambda$ belongs to $\mathbb{C}\backslash\Sigma_\varphi$. Observe that $\mathrm{d}(-\lambda, \overline{\Sigma_\theta})$ is larger than $|\lambda| \sin(\varphi - \theta)$. These facts lead to the crucial lower bound

$$\|\lambda u - Au\| \geqslant |(u|\lambda u - Au)| = |-\lambda - \underline{a}(u, u)| \geqslant |\lambda| \sin(\varphi - \theta). \tag{5.6}$$

Hence, λ is not an element of $\sigma_{\mathrm{ap}}(A)$. Similarly we obtain $\lambda \notin \sigma_{\mathrm{ap}}(A')$, so that λ is contained in $\rho(A)$ by Theorem 1.24. Inequality (5.6) finally implies that A is sectorial of angle $\phi > \pi/2$. $\Diamond$

The above example can be extended to operators with lower-order terms and with other boundary conditions. With considerably more effort one can also establish similar sectoriality results on L^p-spaces. See [**Ou**] for a comprehensive account of this (mainly functional analytic) theory.

In Example 5.11 we obtain the inlusion $\mathrm{D}(A) \subseteq W_0^{1,2}(U)$. For 'very bad' coefficients or domains one cannot expect more. If $a \in W^{1,\infty}$ and $\partial U \in C^{1-}$, integration by parts (use Theorem 3.41) yields that A extends the operator $A_0 u = \operatorname{div}(a\nabla u)$ with $\mathrm{D}(A_0) = W^{2,2}(U) \cap W_0^{1,2}(U)$. If $\partial U \in C^2$, one can show equality here, but the proof needs methods from partial differential equations and is quite technical. Details can be found in the references in Example 3.49 and in the text following it, which also deal with $p \in [1, \infty]$.

Let A be sectorial of angle $\phi \in (\frac{\pi}{2}, \pi)$ with constant K. Take any $r > 0$ and $\theta \in (\frac{\pi}{2}, \phi)$. We define

$$\begin{aligned}
\Gamma_1 &= \{\lambda = \gamma_1(s) = (-s)\mathrm{e}^{-\mathrm{i}\theta} \,|\, -\infty < s \leqslant -r\},\\
\Gamma_2 &= \{\lambda = \gamma_2(\alpha) = r\mathrm{e}^{\mathrm{i}\alpha} \,|\, -\theta \leqslant \alpha \leqslant \theta\},\\
\Gamma_3 &= \{\lambda = \gamma_3(s) = s\mathrm{e}^{\mathrm{i}\theta} \,|\, r \leqslant s < \infty\},\\
\Gamma &= \Gamma(r,\theta) = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3.
\end{aligned}$$

For $t > 0$, we introduce the operator

$$\mathrm{e}^{tA} = \frac{1}{2\pi\mathrm{i}} \int_\Gamma \mathrm{e}^{t\lambda} R(\lambda, A)\,\mathrm{d}\lambda = \lim_{R\to\infty} \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_R} \mathrm{e}^{t\lambda} R(\lambda, A)\,\mathrm{d}\lambda, \tag{5.7}$$

where $\Gamma_R = \Gamma \cap \overline{B}(0, R)$ for $R > r$. We first have to show that the limit in (5.7) exists in $\mathcal{B}(X)$.

LEMMA 5.12. *Under the above assumptions, the integral in* (5.7) *converges absolutely in $\mathcal{B}(X)$ and gives an operator $\mathrm{e}^{tA} \in \mathcal{B}(X)$ which does not depend on the choice of $r > 0$ and $\theta \in (\frac{\pi}{2}, \phi)$. We also have $\|\mathrm{e}^{tA}\| \leqslant M$ for all $t > 0$ and a constant $M = M(K, \theta) > 0$.*

PROOF. Since $\|R(\lambda, A)\| \leqslant \frac{K}{|\lambda|}$ on Γ and $\cos\theta < 0$, we can estimate

$$\begin{aligned}
\left|\int_{\Gamma_R} \|\mathrm{e}^{t\lambda} R(\lambda, A)\|\,\mathrm{d}\lambda\right| &\leqslant K \int_r^R \frac{\exp(ts\,\mathrm{Re}\,\mathrm{e}^{-\mathrm{i}\theta})}{|s\mathrm{e}^{-\mathrm{i}\theta}|} |\mathrm{e}^{-\mathrm{i}\theta}|\,\mathrm{d}s\\
&\quad + K\int_{-\theta}^{\theta} \frac{\exp(tr\,\mathrm{Re}\,\mathrm{e}^{\mathrm{i}\alpha})}{|r\mathrm{e}^{\mathrm{i}\alpha}|} |\mathrm{i}r\mathrm{e}^{\mathrm{i}\alpha}|\,\mathrm{d}\alpha\\
&\quad + K\int_r^R \frac{\exp(ts\,\mathrm{Re}\,\mathrm{e}^{\mathrm{i}\theta})}{|s\mathrm{e}^{\mathrm{i}\theta}|} |\mathrm{e}^{\mathrm{i}\theta}|\,\mathrm{d}s\\
&\leqslant K\left(2\int_r^\infty \frac{\mathrm{e}^{ts\cos\theta}}{s}\,\mathrm{d}s + \int_{-\theta}^{\theta} \mathrm{e}^{tr\cos\alpha}\,\mathrm{d}\alpha\right)\\
&\leqslant K\left(2\int_{rt|\cos\theta|}^\infty \frac{\mathrm{e}^{-\sigma}}{\sigma}(-t\cos\theta)\frac{\mathrm{d}\sigma}{-t\cos\theta} + 2\theta\mathrm{e}^{tr}\right)\\
&=: Kc(r, t, \theta) < \infty
\end{aligned}$$

for $R > r$ and $t > 0$, substituting $\sigma = -st\cos\theta > 0$. Thus the limit in (5.7) exists absolutely in $\mathcal{B}(X)$ by the majorant criterium, and $\|\mathrm{e}^{tA}\| \leqslant Kc(r, t, \theta)$. If we take $r = 1/t$, then $c(1/t, t, \theta) =: c(\theta)$ does not depend on $t > 0$.

So it remains to check that the integral in (5.7) is independent of $r > 0$ and $\theta \in (\frac{\pi}{2}, \phi)$. To this aim, we define $\Gamma' = \Gamma(r', \theta')$ for some $r' > 0$ and $\theta' \in (\frac{\pi}{2}, \phi)$, where we may assume that $\theta' \geqslant \theta$. We further set $\Gamma'_R = \Gamma' \cap \overline{B}(0, R)$ and choose $R > r, r'$. Let C_R^+ and C_R^- be the circle arcs from the endpoint of Γ_R to that of Γ'_R in $\{\mathrm{Im}\,\lambda > 0\}$ and $\{\mathrm{Im}\,\lambda < 0\}$, respectively. (If $\theta = \theta'$, then $C_R^\pm$ contain just one point.) Then $S_R = \Gamma_R \cup C_R^+ \cup (-\Gamma'_R) \cup (-C_R^-)$ is a closed curve in the starshaped domain Σ_ϕ. So (5.1) shows that

$$\int_{S_R} \mathrm{e}^{t\lambda} R(\lambda, A)\,\mathrm{d}\lambda = 0.$$

We further estimate

$$\Big\| \int_{C_R^+} \mathrm{e}^{t\lambda} R(\lambda, A)\, \mathrm{d}\lambda \Big\| \leqslant \int_\theta^{\theta'} \mathrm{e}^{tR\,\mathrm{Re}\,\mathrm{e}^{\mathrm{i}\alpha}} \frac{K}{|R\mathrm{e}^{\mathrm{i}\alpha}|} |\mathrm{i}R\mathrm{e}^{\mathrm{i}\alpha}|\, \mathrm{d}\alpha \leqslant K(\theta'-\theta)\mathrm{e}^{tR\cos\theta} \ \to 0,$$

as $R \to \infty$ since $\cos\theta < 0$, and analogously for C_R^-. So we conclude that

$$\begin{aligned}\int_\Gamma \mathrm{e}^{t\lambda} R(\lambda, A)\, \mathrm{d}\lambda &= \lim_{R\to\infty} \int_{\Gamma_R} \mathrm{e}^{t\lambda} R(\lambda, A)\, \mathrm{d}\lambda = \lim_{R\to\infty} \int_{\Gamma'_R} \mathrm{e}^{t\lambda} R(\lambda, A)\, \mathrm{d}\lambda \\ &= \int_{\Gamma'} \mathrm{e}^{t\lambda} R(\lambda, A)\, \mathrm{d}\lambda. \qquad \square\end{aligned}$$

We next establish some of the fundamental properties of the operators e^{tA}. In view of these results one calls $(\mathrm{e}^{tA})_{t\geqslant 0}$ the 'holomorphic semigroup generated by A'. Actually, the theorem admits a converse. We refer to Section 2.1 of [**Lu**] for this and related facts. (See also Section 2.3 of [**EE**]).

THEOREM 5.13. *Let A be sectorial of angle $\phi > \frac{\pi}{2}$. Define e^{tA} as in (5.7) for $t > 0$, and set $\mathrm{e}^{0A} = I$. Then the following assertions hold.*

a) $\mathrm{e}^{tA}\mathrm{e}^{sA} = \mathrm{e}^{sA}\mathrm{e}^{tA} = \mathrm{e}^{(t+s)A}$ for all $t, s \geqslant 0$.

b) The map $t \mapsto \mathrm{e}^{tA}$ belongs to $C^1(\mathbb{R}_+, \mathcal{B}(X))$. Moreover, $\mathrm{e}^{tA}X \subseteq \mathrm{D}(A)$, $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = A\mathrm{e}^{tA}$ and $\|A\mathrm{e}^{tA}\| \leqslant \frac{C}{t}$ for a constant $C > 0$ and all $t > 0$. We also have $A\mathrm{e}^{tA}x = \mathrm{e}^{tA}Ax$ for all $x \in \mathrm{D}(A)$ and $t \geqslant 0$.

c) Let $x \in X$. Then $\mathrm{e}^{tA}x$ converges as $t \to 0$ in X if and only if $x \in \overline{\mathrm{D}(A)}$. In this case, $\mathrm{e}^{tA}x$ tends to x as $t \to 0$.

PROOF.[2] a) Let $t, s > 0$. Take $0 < r < r'$ and $\frac{\pi}{2} < \theta' < \theta < \phi$. Set $\Gamma = \Gamma(r, \theta)$ and $\Gamma' = \Gamma(r', \theta')$. Using the resolvent equation and Fubini's theorem, we compute

$$\begin{aligned}\mathrm{e}^{tA}\mathrm{e}^{sA} &= \frac{1}{(2\pi\mathrm{i})^2} \int_\Gamma \mathrm{e}^{t\lambda} \int_{\Gamma'} \mathrm{e}^{s\mu} R(\lambda, A)R(\mu, A)\, \mathrm{d}\mu\, \mathrm{d}\lambda \\ &= \frac{1}{2\pi\mathrm{i}} \int_\Gamma \mathrm{e}^{t\lambda} R(\lambda, A) \frac{1}{2\pi\mathrm{i}} \int_{\Gamma'} \frac{\mathrm{e}^{s\mu}}{\mu-\lambda}\, \mathrm{d}\mu\, \mathrm{d}\lambda \\ &\quad + \frac{1}{2\pi\mathrm{i}} \int_{\Gamma'} \mathrm{e}^{s\mu} R(\mu, A) \frac{1}{2\pi\mathrm{i}} \int_\Gamma \frac{\mathrm{e}^{t\lambda}}{\lambda-\mu}\, \mathrm{d}\lambda\, \mathrm{d}\mu.\end{aligned}$$

Fix $\lambda \in \Gamma$ and take $R > \max\{r, r', |\lambda|\}$. We set $C'_R = \{z = R\mathrm{e}^{\mathrm{i}\alpha} \,|\, \theta' \leqslant \alpha \leqslant 2\pi - \theta'\}$ and $S'_R = \Gamma'_R \cup C'_R$. Since $\mathrm{n}(S'_R, \lambda) = 1$, Cauchy's formula (5.2) yields

$$\frac{1}{2\pi\mathrm{i}} \int_{S'_R} \frac{\mathrm{e}^{s\mu}}{\mu-\lambda}\, \mathrm{d}\mu = \mathrm{e}^{s\lambda}.$$

As in Lemma 5.12, we further compute

$$\int_{\Gamma'_R} \frac{\mathrm{e}^{s\mu}}{\mu-\lambda}\, \mathrm{d}\mu \ \longrightarrow \ \int_{\Gamma'} \frac{\mathrm{e}^{s\mu}}{\mu-\lambda}\, \mathrm{d}\mu \qquad \text{and}$$

$$\left| \int_{C'_R} \frac{\mathrm{e}^{s\mu}}{\mu-\lambda}\, \mathrm{d}\mu \right| \leqslant 2\pi R \sup_{\mu\in C'_R} \frac{\mathrm{e}^{s\,\mathrm{Re}\,\mu}}{|\mu-\lambda|} \leqslant \mathrm{e}^{sR\cos\theta'} \frac{2\pi R}{R-|\lambda|} \ \longrightarrow \ 0$$

[2] Not shown in the lectures.

as $R \to \infty$. Consequently,

$$\mathrm{e}^{s\lambda} = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma'} \frac{\mathrm{e}^{s\mu}}{\mu - \lambda}\,\mathrm{d}\mu.$$

Closing Γ_R with the circle arc $C_R = \{z = R\mathrm{e}^{\mathrm{i}\alpha} \,|\, \theta \leqslant \alpha \leqslant 2\pi - \theta\}$ for sufficiently large $R > r$, one verifies in the same way that

$$0 = \int_\Gamma \frac{\mathrm{e}^{\lambda t}}{\lambda - \mu}\,\mathrm{d}\lambda$$

since $\mathrm{n}(\Gamma_R \cup C_R, \mu) = 0$. We thus conclude that

$$\mathrm{e}^{tA}\mathrm{e}^{sA} = \frac{1}{2\pi\mathrm{i}} \int_\Gamma \mathrm{e}^{\lambda t}\mathrm{e}^{s\lambda} R(\lambda, A)\,\mathrm{d}\lambda = \mathrm{e}^{(t+s)A} = \mathrm{e}^{sA}\mathrm{e}^{tA}.$$

b) Let $x \in X$, $t > 0$, $\varepsilon > 0$, and $R > r$. Observe that the Riemann sums for $\int_{\Gamma_R} \mathrm{e}^{t\lambda} R(\lambda, A)\,\mathrm{d}\lambda$ converge in $[\mathrm{D}(A)]$ since $\lambda \mapsto R(\lambda, A)$ is continuous in $\mathcal{B}(X, [\mathrm{D}(A)])$. We thus obtain

$$\begin{aligned} A \int_{\Gamma_R} \mathrm{e}^{\lambda t} R(\lambda, A)\,\mathrm{d}\lambda &= \int_{\Gamma_R} \mathrm{e}^{t\lambda} AR(\lambda, A)\,\mathrm{d}\lambda \\ &= \int_{\Gamma_R} \mathrm{e}^{\lambda t} \lambda R(\lambda, A)\,\mathrm{d}\lambda - \int_{\Gamma_R} \mathrm{e}^{t\lambda}\,\mathrm{d}\lambda\, I. \end{aligned} \tag{5.8}$$

Take again $C_R = \{\mu = R\mathrm{e}^{\mathrm{i}\alpha} \,|\, \theta \leqslant \alpha \leqslant 2\pi - \theta\}$. Using (5.1), one shows as in part a) the limit

$$\left|\int_{\Gamma_R} \mathrm{e}^{t\lambda}\,\mathrm{d}\lambda\right| = \left|-\int_{C_R} \mathrm{e}^{t\lambda}\,\mathrm{d}\lambda\right| \leqslant 2\pi R \sup_{\theta \leqslant \alpha \leqslant 2\pi - \theta} \mathrm{e}^{tR\cos\alpha} \leqslant 2\pi R\,\mathrm{e}^{\varepsilon R\cos\theta} \longrightarrow 0$$

as $R \to \infty$, uniformly for $t \geqslant \varepsilon$. Moreover, as in the proof of Lemma 5.12 (with $r = 1/t$) we estimate

$$\begin{aligned} \left|\int_{\Gamma_R} \|\lambda\mathrm{e}^{t\lambda} R(\lambda, A)\|\,\mathrm{d}\lambda\right| &\leqslant K\Big(2\int_{\frac{1}{t}}^{\infty} \frac{s}{s}\,\mathrm{e}^{ts\cos\theta}\,\mathrm{d}s + \int_{-\theta}^{\theta} r\mathrm{e}^{\cos\alpha} d\alpha\Big) \\ &\leqslant \frac{2K}{t|\cos\theta|} + \frac{2\mathrm{e}K\theta}{t} =: \frac{C'}{t}. \end{aligned}$$

Therefore, the right-hand side of (5.8) tends to

$$\int_\Gamma \lambda\mathrm{e}^{\lambda t} R(\lambda, A)\,\mathrm{d}\lambda$$

as $R \to \infty$. Since A is closed, it follows that $\mathrm{e}^{tA}X \subseteq \mathrm{D}(A)$ and

$$A\mathrm{e}^{tA} = \frac{1}{2\pi\mathrm{i}} \int_\Gamma \lambda\mathrm{e}^{\lambda t} R(\lambda, A)\,\mathrm{d}\lambda, \qquad \|A\mathrm{e}^{tA}\| \leqslant \frac{C'}{2\pi t}$$

for all $t > 0$. In a similar way one sees that

$$\left|\int_{\Gamma\setminus\Gamma_R} \lambda\mathrm{e}^{t\lambda} R(\lambda, A)\,\mathrm{d}\lambda\right| \leqslant 2K \int_R^\infty \mathrm{e}^{ts\cos\theta}\,\mathrm{d}s \leqslant \frac{2K}{\varepsilon|\cos\theta|}\mathrm{e}^{R\varepsilon\cos\theta} \longrightarrow 0$$

as $R \to \infty$, uniformly for $t \geqslant \varepsilon$. As a result,

$$\int_{\Gamma_R} \lambda\mathrm{e}^{\lambda t} R(\lambda, A)\,\mathrm{d}\lambda = \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Gamma_R} \mathrm{e}^{\lambda t} R(\lambda, A)\,\mathrm{d}\lambda$$

converges in $\mathcal{B}(X)$ uniformly for $t \geqslant \varepsilon$, and so $t \mapsto \mathrm{e}^{tA} \in \mathcal{B}(X)$ is continuously differentiable for $t > 0$ with $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = A\mathrm{e}^{tA}$. For $x \in \mathrm{D}(A)$, we further obtain

$$A\mathrm{e}^{tA}x = \lim_{R\to\infty} \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_R} \mathrm{e}^{\lambda t} R(\lambda, A)Ax \,\mathrm{d}\lambda = \mathrm{e}^{tA}Ax.$$

c) Let $x \in \mathrm{D}(A)$, $R > r$, and $t > 0$. As in step a), from Cauchy's formula (5.2) we derive

$$\frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \frac{\mathrm{e}^{\lambda t}}{\lambda} \,\mathrm{d}\lambda = \lim_{R\to\infty} \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_R} \frac{\mathrm{e}^{\lambda t}}{\lambda - 0} \,\mathrm{d}\lambda = 1$$

Observing that $\lambda R(\lambda, A)x - x = R(\lambda, A)Ax$, we conclude that

$$\mathrm{e}^{tA}x - x = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \mathrm{e}^{\lambda t}\Big(R(\lambda, A) - \frac{1}{\lambda}\Big)x \,\mathrm{d}\lambda = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \frac{\mathrm{e}^{\lambda t}}{\lambda} R(\lambda, A)Ax \,\mathrm{d}\lambda.$$

Because the integrand is bounded by $\frac{c}{|\lambda|^2}$ on Γ for all $t \in (0, 1]$, Lebesgue's convergence theorem implies the existence of the limit

$$\lim_{t\to 0} \mathrm{e}^{tA}x - x = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \frac{1}{\lambda} R(\lambda, A)Ax \,\mathrm{d}\lambda =: z.$$

Let $K_R = \{R\mathrm{e}^{\mathrm{i}\alpha} \,\big|\, -\theta \leqslant \alpha \leqslant \theta\}$. Cauchy's theorem (5.1) shows that

$$\int_{\Gamma_R \cup (-K_R)} \frac{1}{\lambda} R(\lambda, A)Ax \,\mathrm{d}\lambda = 0.$$

Since also

$$\left\| \int_{-K_R} \frac{1}{\lambda} R(\lambda, A)Ax \,\mathrm{d}\lambda \right\| \leqslant \frac{2\pi RK}{R^2} \,\|Ax\| \;\longrightarrow\; 0$$

as $R \to \infty$, we arrive at $z = 0$. Because of the uniform boundedness of e^{tA}, it follows that $\mathrm{e}^{tA}x \to x$ as $t \to 0$ for all $x \in \overline{\mathrm{D}(A)}$.

Conversely, if $\mathrm{e}^{tA}x \to y$ as $t \to 0$, then y belongs to $\overline{\mathrm{D}(A)}$ by assertion b). Moreover, $R(1, A)\mathrm{e}^{tA}x = \mathrm{e}^{tA}R(1, A)x$ tends to $R(1, A)x$ as $t \to 0$ because of $R(1, A)x \in \mathrm{D}(A)$. It follows $R(1, A)y = R(1, A)x$, and so $x = y \in \overline{\mathrm{D}(A)}$. □

Remark 5.14. Let $A - \omega I = A_\omega$ be sectorial of angle greater than $\frac{\pi}{2}$ for some $\omega \in \mathbb{R}$. We then compute

$$\mathrm{e}^{\omega t}\mathrm{e}^{tA_\omega} = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \mathrm{e}^{t(\lambda+\omega)} R(\lambda + \omega, A) \,\mathrm{d}\lambda = \frac{1}{2\pi\mathrm{i}} \int_{\omega+\Gamma} \mathrm{e}^{\mu t} R(\mu, A) \,\mathrm{d}\mu =: \mathrm{e}^{tA}$$

for $t > 0$. For $\mathrm{e}^{tA} = \mathrm{e}^{\omega t}\mathrm{e}^{tA_\omega}$ one obtains similar properties as for $\omega = 0$. ◊

We now solve the evolution equation (5.9) governed by a sectorial operator A with angle $\phi > \pi/2$. Such problems are called of 'parabolic type' since diffusion problems are typical applications, see Example 5.16. In contrast to the Schrödinger equation (4.8) we can allow for initial values in $\overline{\mathrm{D}(A)}$.

Corollary 5.15. *Let A be sectorial of angle $\phi > \pi/2$ and let $u_0 \in \overline{\mathrm{D}(A)}$. Then $u(t) = \mathrm{e}^{tA}u_0$, $t \geqslant 0$, is the unique solution in $C^1(\mathbb{R}_+, X) \cap C(\mathbb{R}_+, [\mathrm{D}(A)]) \cap C(\mathbb{R}_{\geqslant 0}, X)$ of the initial value problem*

$$u'(t) = Au(t), \quad t > 0, \qquad u(0) = u_0. \tag{5.9}$$

Proof. Existence follows from Theorem 5.13. Let v be another solution of (5.9). Let $0 < \varepsilon \leqslant s \leqslant t-\varepsilon < t$. Theorem 5.13 then implies that

$$\tfrac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{(t-s)A}v(s) = -\mathrm{e}^{(t-s)A}Av(s) + \mathrm{e}^{(t-s)A}v'(s) = 0.$$

As in Example 4.20, this fact yields $\mathrm{e}^{(t-\varepsilon)A}v(\varepsilon) = \mathrm{e}^{\varepsilon A}v(t-\varepsilon)$. Letting $\varepsilon \to 0$, one infers $\mathrm{e}^{tA}u_0 = v(t)$ as $\tau \mapsto \mathrm{e}^{\tau A}x$ is continuous for $\tau \geqslant 0$ and $x \in \overline{\mathrm{D}(A)}$. $\square$

We only give one of the possible examples.

Example 5.16. Let $X = C([0,1])$ and $A\varphi = \varphi''$ with $\mathrm{D}(A) = \{\varphi \in C^2([0,1]) \,|\, \varphi(0) = \varphi(1) = 0\}$. Let $u_0 \in C_0(0,1) = \overline{\mathrm{D}(A)}$. Then the function $u(t) = \mathrm{e}^{tA}u_0$, $t \geqslant 0$, belongs to

$$C(\mathbb{R}_{\geqslant 0}, C([0,1])) \cap C(\mathbb{R}_+, C^2([0,1])) \cap C^1(\mathbb{R}_+, C([0,1]))$$

and uniquely solves the partial differential equation

$$\begin{aligned} \partial_t u(t,x) &= \partial_{xx}u(t,x), && t>0,\ x\in[0,1],\\ u(t,0) &= u(t,1) = 0, && t \geqslant 0, \\ u(0,x) &= u_0(x), && x\in[0,1]. \end{aligned} \tag{5.10}$$

$\Diamond$

Formula (5.7) and Theorem 5.13 lead to a theory for parabolic problems which is similar to ordinary differential equations, see [**Lu**] or [**EE**] and [**nEE**]. We only add a basic theorem on the long-term behavior extending Example 5.4. It can be applied to (5.10) by Examples 5.9 and 3.46.

Theorem 5.17. *Let A be sectorial of angle $\phi > \pi/2$ and satisfy* $\mathrm{s}(A) := \sup\{\operatorname{Re}\lambda \,|\, \lambda \in \sigma(A)\} < -\delta < 0$. *Then there is a constant $N \geqslant 1$ such that* $\|\mathrm{e}^{tA}\| \leqslant N\mathrm{e}^{-\delta t}$ *for all $t \geqslant 0$.*

Proof.[3] The assumptions imply that $A_{-\delta} = A + \delta I$ is sectorial of some angle $\psi \in (\pi/2, \phi)$. Take $\Gamma = \Gamma(r,\theta)$ with $r > 0$ and $\theta \in (\pi/2, \psi)$. Remark 5.14 then yields that $\mathrm{e}^{tA} = e^{-\delta t}\mathrm{e}^{tA_{-\delta}}$, where however e^{tA} is defined by the curve integral (5.7) on the shifted path $\Gamma' := -\delta + \Gamma$. Lemma 5.12 shows that $\mathrm{e}^{tA_{-\delta}}$ is uniformly bounded for $t \geqslant 0$. It thus remains to verify

$$\int_{-\delta+\Gamma} \mathrm{e}^{t\mu}R(\mu,A)\,\mathrm{d}\mu = \int_\Gamma \mathrm{e}^{t\lambda}R(\lambda,A)\,\mathrm{d}\lambda, \qquad t>0. \tag{5.11}$$

To this end, let $R > r$ and $S_R^\pm$ be the horizontal line segments connecting the end points on Γ_R and Γ_R' in $\{\operatorname{Im}\lambda > 0\}$ and $\{\operatorname{Im}\lambda < 0\}$, respectively. Let $C_R = \Gamma_R \cup S_R^+ \cup (-\Gamma') \cup (-S_R^-)$. This path is contained in $\rho(A)$ and $\mathrm{n}(C_R, z) = 0$ for all $z \in \sigma(A)$. Cauchy's theorem (5.1) now implies

$$\int_{C_R} \mathrm{e}^{t\lambda}R(\lambda,A)\,\mathrm{d}\lambda = 0.$$

Observe that the segments $S_R^\pm$ have fixed length δ and that $\operatorname{Re}\lambda \leqslant R\cos\theta < 0$ and $|\lambda| \geqslant R$ for all $\lambda \in S_R^\pm$. Because of $\theta < \phi$, the sets $S_R^\pm$ belong to Σ_ϕ for all sufficiently large R. We can thus estimate

$$\Big\| \int_{S_R^\pm} \mathrm{e}^{t\lambda}R(\lambda,A)\,\mathrm{d}\lambda \Big\| \leqslant \frac{\delta K}{R}\,\mathrm{e}^{Rt\cos\theta}.$$

Since the right-hand side vanishes as $R \to \infty$, we have shown (5.11). $\square$

[3] Not shown in the lectures.

Bibliography

[AF] R.A. Adams and J.F. Fournier, *Sobolev Spaces.* 2nd edition. Academic Press, 2007.

[AE] H. Amann and J. Escher, *Analysis III.* Birkhäuser, Basel, 2008.

[Co1] J.B. Conway, *Functions of One Complex Variable.* 2nd edition. Springer, 1978.

[Co2] J.B. Conway, *A Course in Functional Analysis.* 2nd edition. Springer, 1990.

[DS] N. Dunford and J.T. Schwartz, *Linear Operators, Part 2, Spectral Theory, Self Adjoint Operators in Hilbert Space.* Wiley-Interscience, 1963.

[Ev] L.C. Evans, *Partial Differential Equations.* 2nd edition. American Math. Soc., 2010.

[GT] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order.* 2nd edition. Springer, 1998.

[Gr] L. Grafakos, *Modern Fourier Analysis.* 2nd edition. Springer, 2009.

[Ka] T.Kato, *Perturbation Theory for Linear Operators.* Reprint of the 1980 edition. Springer, 1995.

[Ke] O.D. Kellogg, *Foundations of Potential Theory.* Springer, 1967.

[KW] P.C. Kunstmann and L. Weis, *Maximal L_p–Regularity for Parabolic Problems, Fourier Multiplier Theorems and H^∞–Calculus*, in: M. Iannelli, R. Nagel and S. Piazzera (Eds.): 'Functional Analytic Methods for Evolution Equations', Springer, 2004, pp. 65–311.

[La] S. Lang, *Algebra.* Revised 3rd edition. Springer, 2002.

[Lu] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Birkhäuser, 1995.

[Ma] V. Maz'ya, *Sobolev Spaces with Applications to Elliptic Partial Differential Equations.* 2nd, revised and augmented edition. Springer, 2011.

[Ou] E.M. Ouhabaz, *Analysis of Heat Equations on Domains*, Princeton University Press, 2005.

[Ne] J. Nečas, *Direct Methods in the Theory of Elliptic Equations.* Corrected 2nd printing. Springer, 2012.

[PW] M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations.* Corrected reprint of 2nd printing, Springer, 1984.

[RS] M. Reed and B. Simon, *Functional Analysis.* Revised and enlarged edition. Academic Press, 1975.

[Ru1] W. Rudin, *Real and Complex Analysis.* McGraw-Hill, 1987.

[Ru2] W. Rudin, *Functional Analysis.* McGraw-Hill, 1991.

[Sa] B. Sandstede, *Stability of travelling waves.* In B. Fiedler (Ed.): 'Handbook of Dynamical Systems, Vol. 2,' North–Holland, 2002, pp. 983–1055.

[FA] R. Schnaubelt, *Functional Analysis.* Class Notes, Karlsruhe, 2018.

[EE] R. Schnaubelt, *Evolution Equations.* Class Notes, Karlsruhe, 2020.

[nEE] R. Schnaubelt, *Nonlinear Evolution Equations.* Class Notes, Karlsruhe, 2019.

[Si] B. Simon, *Operator Theory. A Comprehensive Course in Analysis, Part 4.* American Mathematical Society, 2015.

[St] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, 1970.

[Ta] H. Tanabe, *Functional Analytic Methods for Partial Differential Equations.* Dekker, 1997.

[Tr] H. Triebel, *Higher Analysis.* Barth Verlag, 1997.

[We] D. Werner, *Funktionalanalysis.* 5te erweiterte Auflage. Springer, 2005.