29th Internet Seminar

Eventual Positivity

Sahiba Arora, Jochen Glück, Jonathan Mui

Lecture Notes
Winter term 2025/26

Contents

Introduction

Prerequisites

This course was designed for postgraduate students (Masters and PhD) and advanced Bachelor students with the following prerequisite knowledge:

- Calculus/analysis in one and several variables;
- Linear algebra (in particular eigenvalues and the Jordan normal form of matrices);
- An introduction to real analysis, measure and integration theory (in particular, familiarity with L^p spaces);
- An introduction to functional analysis (Banach spaces, Hilbert spaces, bounded linear operators); and
- An introduction to complex analysis (holomorphic functions, complex path integrals and Cauchy's integral formula, Laurent series).

Acknowledgements

We thank Alexander Wierzba for sending us the font used in these lecture notes.

Nomenclature

The following table gives an overview of important symbols used in the lectures. This table will be updated every week as we introduce new notation.

Elementary notation

$\mathbb{N}$	set of strictly positive integers, i.e. $\mathbb{N} = \{1,2,3,\dots\}$
$\mathbb{N}_0$	set of integers that are ≥ 0, i.e. $\mathbb{N}_0 = \{0,1,2,\dots\}$
$\mathbb{R}_+$	alternative notation for the interval $[0,\infty)$
$B_{<r}(x)$	open ball with centre x and radius r, where the metric space is clear from context
$B_{\leq r}(x)$	closed ball with centre x and radius r, where the metric space is clear from context

Function spaces

$\mathbb{1}$	the vector in $\mathbb{R}^n$ whose entries are all 1, or the constant function with value 1 on a set that is clear from the context
$\mathrm{F}(S;\mathbb{R})$	real-valued analogue of any function space $\mathrm{F}(S)$ that occurs in the following list and consists of complex-valued functions on a set S
$\mathrm{C}(K)$	space of complex-valued continuous functions on a compact metric space (or compact topological Hausdorff space) K; endowed with the sup norm
$\mathrm{C}_0(\Omega)$	space of complex-valued continuous functions on a non-empty open set $\Omega \subseteq \mathbb{R}^n$ which approach 0 at $\partial\Omega$ and at ∞; endowed with the sup norm
$\mathrm{C}^m(\Omega)$	space of complex-valued m-times continuously differentiable functions on a non-empty open subset $\Omega \subseteq \mathbb{R}^n$ for $m \in \mathbb{N}_0$
$\mathrm{C}(\Omega)$	short for $\mathrm{C}^0(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ is non-empty and open; in contrast to $\mathrm{C}(K)$ for compact K we do not endow $\mathrm{C}(\Omega)$ with a specific norm or topology, unless otherwise stated

$C^\infty(\Omega)$	the intersection $\bigcap_{m\in\mathbb{N}_0} C^m(\Omega)$ for a non-empty open subset $\Omega \subseteq \mathbb{R}^n$
$C_c^\infty(\Omega)$	the space of test functions on a non-empty open subset $\Omega \subseteq \mathbb{R}^n$ – i.e. of all functions in $C^\infty(\Omega)$ whose support is a compact subset of Ω
$C^m(\overline{\Omega})$	the space of m-times continuously differentiable functions such that every derivative up to order m can be extended to a continuous function on $\overline{\Omega}$, for a non-empty, bounded, open set $\Omega \subset \mathbb{R}^n$; note that the space depends on Ω, not only on its closure $\overline{\Omega}$
$C_b^m(\Omega)$	for $m = 0$, the space of bounded continuous functions on a non-empty set $\Omega \subseteq \mathbb{R}^n$ (not necessarily bounded); for $m \in \mathbb{N}$, the space of m-times continuously differentiable functions on Ω such that every derivative up to order m is in $C_b(\Omega)$
$L^p(\Omega,\nu)$	complex-valued L^p-space over the measure space (Ω,ν) for $p \in [1,\infty]$
$L^p(\Omega)$	complex-valued L^p-space over a measurable subset $\Omega \subseteq \mathbb{R}^n$ that is endowed with the Lebesgue measure for $p \in [1,\infty]$
$L^p_{\mathrm{loc}}(\Omega)$	space of (equivalence classes of) complex-valued measurable functions on a non-empty open set $\Omega \subseteq \mathbb{R}^n$ that are locally in L^p, for $p \in [1,\infty]$
$W^{k,p}(\Omega)$	the Sobolev space of complex-valued functions on a non-empty open set $\Omega \subseteq \mathbb{R}^n$ whose weak derivatives up to order k all exist and are in $L^p(\Omega)$; for $p \in [1,\infty]$ and $k \in \mathbb{N}_0$
$W_0^{k,p}(\Omega)$	the closure of the space of test functions $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$, for $p \in [1,\infty)$ and $k \in \mathbb{N}_0$
$H^k(\Omega)$	short for $W^{k,2}(\Omega)$, for $k \in \mathbb{N}_0$
$H_0^k(\Omega)$	short for $W_0^{k,2}(\Omega)$, for $k \in \mathbb{N}_0$

Banach spaces and linear operators

id	the identity matrix in $\mathbb{C}^{n\times n}$ or the identity operator on a normed space that is clear from the context
X'	norm dual space of a normed space X; depending on the context, elements of X' are denoted by $x', y', \dots$ or by $\varphi, \psi, \dots$
$\langle \varphi, x \rangle$	stands for $\varphi(x)$, where $x \in X$ and $\varphi \in X'$ for a normed space X
A'	the dual operator $A'\colon Y' \supseteq \operatorname{dom}(A') \to X'$ of a densely defined linear operator $A\colon X \supseteq \operatorname{dom}(A) \to Y$ between to Banach spaces X and Y over the same scalar field
$(\cdot \mid \cdot)$	inner product on a Hilbert space; antilinear in the *first* component

$\mathcal{L}(X,Y)$ — space of bounded linear operators between two normed spaces X and Y over the same scalar field

$\mathcal{L}(X)$ — short for $\mathcal{L}(X,X)$, where X is a normed space

$Y \hookrightarrow X$ — continuous embedding of a Banach space Y into a Banach space X, i.e. $Y \subseteq X$ and there exists $C > 0$ such that $\|y\|_X \le C\|y\|_Y$ for all $y \in Y$

Spectral theory

$\sigma(A)$ — spectrum of a closed linear operator A

$\sigma_{\mathrm{pnt}}(A)$ — point spectrum, i.e. the set of eigenvalues of a closed linear operator A

$\rho(A)$ — resolvent set of a closed linear operator A, i.e. $\rho(A) := \mathbb{C} \setminus \sigma(A)$

$\mathcal{R}(\lambda, A)$ — resolvent of a closed linear operator A at a point $\lambda \in \rho(A)$, i.e. $\mathcal{R}(\lambda, A) := (\lambda - A)^{-1}$

$\mathrm{r}(A)$ — spectral radius of a bounded linear operator A, defined by the formula $\mathrm{r}(A) := \max\{|\lambda| : \lambda \in \sigma(A)\} \in [0,\infty)$

$\mathrm{s}(A)$ — spectral bound of a closed linear operator A, defined by the formula $\mathrm{s}(A) := \sup\{\mathrm{Re}\,\lambda : \lambda \in \sigma(A)\} \in [-\infty,\infty]$

Ordered structures on vector spaces

V_+ — positive cone of a vector lattice V

$x \preceq y$ — $cx \le y$ for some number $c > 0$ (equivalently, $x \le cy$ for some number $c > 0$)

$y \succeq x$ — alternative notation for $x \preceq y$

$\sup S$ — supremum of a subset S of a partially ordered space X

$\inf S$ — infimum of a subset S of a partially ordered space X

$x \vee y$ — $\sup\{x, y\}$

$x \wedge y$ — $\inf\{x, y\}$

x^+ — $x \vee 0$

x^- — $x \wedge 0$

$|x|$ — $x \vee (-x)$

$V_{\mathbb{C}}$ — complexification of a real vector space V

$E_{\mathbb{R}}$ — real part of a complex Banach lattice E

$\operatorname{dom}(A)_{\mathbb{R}}$	the real part of the domain of a linear operator $A\colon E \supseteq \operatorname{dom}(A) \to F$ between Banach lattices E and F, i.e. $\operatorname{dom}(A)_{\mathbb{R}} := \operatorname{dom}(A) \cap E_{\mathbb{R}}$
$f \gneq 0$	$f \ge 0$ but $f \ne 0$
$0 \lneq f$	alternative notation for $f \gneq 0$
E_u	principal ideal generated by a positive element u of a Banach lattice E
$\lVert\cdot\rVert_{E_u}$	gauge norm on the principal ideal E_u
E^{φ}	AL-space generated by a strictly positive functional φ on a Banach lattice E
$\lVert\cdot\rVert_{E^{\varphi}}$	the norm $\langle \varphi, \lvert\cdot\rvert \rangle$, where $\varphi \in E'$ is strictly positive

Chapter 1

Positive matrices and matrix semigroups

The topic of the ISEM 29 is the interplay between dynamical systems (more specifically: differential equations), sign preservation, and operator theory. The material in the first two chapters develops the essence of the theory in finite dimensions. In Chapter 1, we study positive matrices and matrix exponential functions, and show how the positivity affects their eigenvalues and eigenvectors. The titular subject, *eventual positivity*, makes an appearance in Chapter 2.

1.1 Positive matrices and the standard order on $\mathbb{R}^n$

As a foundation for everything that follows, we endow the spaces $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ with the following partial order.

Definition 1.1.1 (The order and the cone on $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$)**.**

(a) For $x, y \in \mathbb{R}^n$ we write $x \le y$ if this inequality holds componentwise, i.e. if $x_k \le y_k$ for every index k. As usual we use the notation $y \ge x$ synonymously with $x \le y$.

Vectors $x \in \mathbb{R}^n$ that satisfy $x \ge 0$ [1] are called the **positive** elements of $\mathbb{R}^n$, and the set $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \ge 0\}$ of all positive vectors is called the **positive cone** of $\mathbb{R}^n$.

(b) We use the same conventions for matrices: for $A, B \in \mathbb{R}^{m\times n}$ we write $A \le B$ (or $B \ge A$) if $A_{jk} \le B_{jk}$ for all indices j, k.

A matrix $A \in \mathbb{R}^{m\times n}$ is called **positive** if $A \ge 0$, and the set $\mathbb{R}^{m\times n}_+ := \{A \in \mathbb{R}^{m\times n} : A \ge 0\}$ of positive matrices is called the **positive cone** in $\mathbb{R}^{m\times n}$.

Note that Definition 1.1.1(a) can be considered a special case of part (b) if we identify $\mathbb{R}^n$ with $\mathbb{R}^{n\times 1}$. The relation $\le$ is a partial order on $\mathbb{R}^n$ and is compatible with its vector

[1] We always write 0 for the zero vector when the corresponding space is clear from context.

space structure in the following sense: if $x \le y$ for $x, y \in \mathbb{R}^n$, then

$$\alpha x \le \alpha y \qquad \text{and} \qquad x + z \le y + z$$

for all numbers $\alpha \in [0, \infty)$ and all vectors $z \in \mathbb{R}^n$. Analogous statements hold for the partial order $\le$ on $\mathbb{R}^{m \times n}$.

Remark 1.1.2 (Terminology: positive vectors)**.** At first glance, it might be surprising that our definition of 'positivity' is inconsistent with its common meaning for real numbers: in English, a number $\alpha \in \mathbb{R}$ is usually called positive if $\alpha > 0$. Yet, the vector $0 \in \mathbb{R}^n$ is positive in the sense of Definition 1.1.1. For $n = 1$ this means, in particular, that the real number 0 is positive in the sense of Definition 1.1.1.

Nevertheless, our usage of 'positive' is standard in the theory of Banach lattices, which we use frequently from Chapter 4 on. For readers who take pleasure in terminological digressions, a more thorough discussion is provided in the notes at the end of this chapter.

From an operator-theoretic perspective, it is desirable to describe positivity of matrices in terms of how they act as linear maps. We do this in the next proposition.

Proposition 1.1.3. *For a matrix $A \in \mathbb{R}^{m \times n}$, the following are equivalent:*

(i) *A is positive, i.e. $A \in \mathbb{R}_+^{m \times n}$.*

(ii) *$A(\mathbb{R}_+^n) \subseteq \mathbb{R}_+^m$.*

(iii) *A acts monotonically, i.e. if $x, y \in \mathbb{R}^n$ satisfy $x \le y$, then $Ax \le Ay$.*

Proof. "(i) ⇒ (ii)": This is clear from the definition of the matrix-vector product.

"(ii) ⇒ (iii)": Assume that (ii) holds and let $x, y \in \mathbb{R}^n$ satisfy $x \le y$. Then $y - x \in \mathbb{R}_+^n$ and hence $Ay - Ax = A(y - x) \in \mathbb{R}_+^m$, which implies that $Ax \le Ay$.

"(iii) ⇒ (i)": Assume that (iii) holds. For $j \in \{1, \dots, n\}$ and the canonical unit vector $e_j \in \mathbb{R}^n$ one has $0 \le e_j$ and thus $0 = A0 \le Ae_j$. Since Ae_j is the j-th column of A and j was arbitrary, we conclude that all entries of A are ≥ 0. ☐

Since we defined the order relation $\le$ by comparing vectors (and matrices) entrywise, it is natural to generalise the modulus from scalars to vectors in the same way:

Definition 1.1.4 (The modulus of vectors and matrices)**.** For every vector $x \in \mathbb{C}^n$ and every matrix $A \in \mathbb{C}^{m \times n}$ we define the matrix $|A| \in \mathbb{R}_+^{m \times n}$ and the vector $|x| \in \mathbb{R}_+^n$ by taking the entrywise modulus of x and A, i.e.

$$|x|_j := |x_j| \qquad \text{and} \qquad |A|_{jk} := |A_{jk}|$$

for all indices j and k.

The modulus has a submultiplicative property, which is very useful to prove estimates for positive matrices.

Proposition 1.1.5 (Submultiplicativity of the modulus)**.** *Let* $A \in \mathbb{C}^{m\times n}$ *and* $x \in \mathbb{C}^n$*.*

(a) *One has* $|Ax| \le |A|\,|x|$*.*

(b) *In particular, if* $A \in \mathbb{R}_+^{m\times n}$*, then* $|Ax| \le A|x|$*.*

Proof. (a) One can check the inequality entrywise: for every $j \in \{1,\dots,m\}$ one has

$$|Ax|_j = \left|(Ax)_j\right| = \left|\sum_{k=1}^n A_{jk}x_k\right| \le \sum_{k=1}^n \left|A_{jk}\right| |x_k| = \left(|A|\,|x|\right)_j .$$

(b) For positive A, one has $|A| = A$, so the claim follows from part (a). □

Remark 1.1.6 (Norms on $\mathbb{C}^n$)**.** In the following we often work with norms on $\mathbb{C}^n$. While they are all equivalent, we assume throughout that $\mathbb{C}^n$ is endowed with a norm that satisfies $\||x|\| = \|x\|$ for all $x \in \mathbb{C}^n$ as well as $\|x\| \le \|y\|$ for all $x, y \in \mathbb{R}^n$ with $0 \le x \le y$ – this is sometimes more convenient in estimates. For instance, the p-norm has this property for every $p \in [1,\infty]$.

1.2 The spectrum of positive matrices

An intriguing feature of positive matrices is that their eigenvalues and eigenvectors enjoy a variety of remarkable properties. This is the content of the classical **Perron-Frobenius theorem**, which we study in this section. This theorem is only a first instance of one of the most important themes of the course: the interaction between positivity and the spectrum of linear operators. We take this opportunity to introduce some fundamental concepts and tools in spectral theory.

Definition 1.2.1 (Spectrum and spectral radius)**.** Let $A \in \mathbb{C}^{n\times n}$. The set $\sigma(A) \subseteq \mathbb{C}$ that consists of all eigenvalues of A is called the **spectrum** of A, and the number

$$\mathrm{r}(A) := \max\left\{|\lambda| : \lambda \in \sigma(A)\right\} \in [0,\infty)$$

is called the **spectral radius** of A.

The spectral radius determines whether the powers of a matrix converge to 0 as the exponent tends to ∞. More precisely, one has the following equivalence.

Proposition 1.2.2 (Convergence to 0 of matrix powers)**.** *For every matrix* $A \in \mathbb{C}^{n\times n}$*, the following are equivalent:*

(i) $\mathrm{r}(A) < 1$.

(ii) $A^k \to 0$ *as* $k \to \infty$*.*

(iii) *There exist numbers* $\eta \in [0,1)$ *and* $c \ge 0$ *such that* $\left\|A^k\right\| \le c\eta^k$ *for each* $k \in \mathbb{N}_0$*.*

Proof. "(i) ⇒ (iii)": The implication is clear if $\mathrm{r}(A) = 0$, hence we assume $\mathrm{r}(A) > 0$. One can then show, using the Jordan normal form of A, that $\|A^k\| \le \tilde{c}\,\mathrm{r}(A)^k(1 + k^{n-1})$ for a number $\tilde{c} \ge 0$ and all $k \in \mathbb{N}_0$; see Exercise 1.4(c). So the claim follows by taking any $\eta \in (\mathrm{r}(A), 1)$ and using that $\frac{\mathrm{r}(A)^k}{\eta^k}$ decays exponentially.

"(iii) ⇒ (ii)": This implication is obvious.

"(ii) ⇒ (i)": Let λ be an eigenvalue of A with $|\lambda| = \mathrm{r}(A)$ associated to an eigenvector z of norm one. One has $|\lambda|^k = |\lambda|^k \|z\| = \|A^k z\| \to 0$ as $k \to \infty$. Thus, $\mathrm{r}(A) = |\lambda| < 1$. □

To formulate the statement of some parts of the Perron-Frobenius theorem we need the following stronger notion of positivity.

Definition 1.2.3 (Strong positivity in finite dimensions)**.** A vector $x \in \mathbb{R}^n$ is called **strongly positive** if $x_k > 0$ for all $k \in \{1, \dots, n\}$. Similarly, a matrix $A \in \mathbb{R}^{m \times n}$ is called **strongly positive** if $A_{jk} > 0$ for all indices j, k.

Observe that the strongly positive vectors in $\mathbb{R}^n$ are precisely the points in the interior of the positive cone $\mathbb{R}^n_+$. Similarly as in Proposition 1.1.3, strong positivity of matrices can also be interpreted in terms of their actions as linear mappings: a matrix $A \in \mathbb{R}^{m \times n}$ is strongly positive if and only if it maps every $0 \ne x \in \mathbb{R}^n_+$ to a strongly positive vector.

It is convenient to have a notation for strong positivity. The following has the advantage that it can easily be generalised to the infinite-dimensional setting in later chapters.

Notation 1.2.4 (Inequality up to a factor)**.**

(a) For two vectors $x, y \in \mathbb{R}^n$ we write $x \preceq y$ or equivalently $y \succeq x$ if there exists a number $c > 0$ such that $cx \le y$ (equivalently, if there exists a number $c > 0$ such that $x \le cy$).

(b) We let $\mathbb{1} \in \mathbb{R}^n$ denote the vector with every entry equal to 1. Hence, a vector $x \in \mathbb{R}^n$ is strongly positive if and only if $x \succeq \mathbb{1}$.

The main result of this section is the following classical theorem about the eigenvalues and eigenvectors of positive matrices.

Theorem 1.2.5 (Perron–Frobenius)**.** *Let* $0 \le A \in \mathbb{R}^{n \times n}$.

(a) *The spectral radius* $\mathrm{r}(A)$ *is an eigenvalue of* A *with an eigenvector* $x \ge 0$.

(b) *If all diagonal entries of* A *are non-zero, then* $\mathrm{r}(A) > 0$, *and* $\mathrm{r}(A)$ *is a* ***radially strictly dominant*** *eigenvalue in the sense that* $|\lambda| < \mathrm{r}(A)$ *for all other eigenvalues* λ *of* A.

(c) *If* A *is even strongly positive, then* $\mathrm{r}(A) > 0$, *the eigenvalue* $\mathrm{r}(A)$ *of* A *is algebraically simple,*[2] *and its eigenspace is spanned by a strongly positive vector.*

[2] Recall that the **algebraic multiplicity** of an eigenvalue λ of A is the dimension of the generalised eigenspace $\bigcup_{k=1}^{n} \ker(\lambda - A)^k$. The eigenvalue λ is called **algebraically simple** if its algebraic multiplicity is one.

The Perron–Frobenius theorem is useful to study the behaviour of A^k of a positive matrix A as $k \to \infty$. A concrete application to Markov chains is explored in Exercise 1.3.

Various proofs of the theorem and variations thereof are known; see e.g. the survey article [Mac00] for some nice bedtime reading. The proof we present has a strong functional analytic flavour and already anticipates several ideas and arguments that occur again in the infinite-dimensional case – strongly relying on properties of the resolvent. We define and study this object now and finally use it to prove Theorem 1.2.5.

Definition 1.2.6 (The resolvent of a matrix)**.** Let $A \in \mathbb{C}^{n\times n}$. The complement of its spectrum, i.e. $\rho(A) := \mathbb{C} \setminus \sigma(A)$, is called the **resolvent set** of A. The mapping

$$\mathcal{R}(\cdot, A)\colon \rho(A) \to \mathbb{C}^{n\times n}, \qquad \lambda \mapsto \mathcal{R}(\lambda, A) := (\lambda - A)^{-1}$$

is called the **resolvent** of A.

In the preceding definition, we used the notation $\lambda - A$, which is shorthand for $\lambda \operatorname{id} - A$; where $\operatorname{id} \in \mathbb{C}^{n\times n}$, denotes the identity matrix of the same dimension as A.

To state the following proposition we need the concept of a vector-valued analytic functions. In finite dimensions this is easy: a mapping from an open subset of $\mathbb{C}$ to $\mathbb{C}^n$ or to $\mathbb{C}^{n\times n}$ is called **analytic** or **holomorphic** if every component of the mapping is analytic.

Proposition 1.2.7 (Properties of resolvents)**.** *Let $A \in \mathbb{C}^{n\times n}$.*

(a) *The resolvent $\mathcal{R}(\cdot, A)\colon \rho(A) \to \mathbb{C}^{n\times n}$ is analytic.*

(b) *For $\lambda \in \mathbb{C}$ with $|\lambda| > \operatorname{r}(A)$, the resolvent can be represented as the* ***Neumann series***

$$\mathcal{R}(\lambda, A) = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}},$$

which converges absolutely in $\mathbb{C}^{n\times n}$ (with respect to any norm).

Proof. (a) It follows from Cramer's rule for the inverse of a matrix that, for all indices j, k, the matrix entry $\mathcal{R}(\cdot, A)_{jk}\colon \rho(A) \to \mathbb{C}$ is a rational function and thus analytic.

(b) As $\operatorname{r}(A/\lambda) < 1$, so by Proposition 1.2.2, there exist numbers $\eta \in [0,1)$ and $c \geq 0$ such that $\|A^k/\lambda^k\| \leq c\eta^k$ for every $k \in \mathbb{N}_0$. Thus, $\sum_{k=0}^{\infty} \left\| \frac{A^k}{\lambda^{k+1}} \right\| < \infty$, and hence the series converges absolutely in $\mathbb{C}^{n\times n}$. To show the resolvent formula, we compute

$$(\lambda - A) \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}} = \lim_{K\to\infty} \sum_{k=0}^{K} \left(\frac{A^k}{\lambda^k} - \frac{A^{k+1}}{\lambda^{k+1}} \right) = \lim_{K\to\infty} \left(\operatorname{id} - \frac{A^{K+1}}{\lambda^{K+1}} \right) = \operatorname{id}.$$

Here we used that $\|A^{k+1}/\lambda^{k+1}\| \to 0$ as $k \to \infty$ according to Proposition 1.2.2 since $\operatorname{r}(A/\lambda) < 1$. After multiplying by $\mathcal{R}(\lambda, A)$, we obtain the claimed formula. □

Finally, we need the following lemma about simplicity of eigenvalues. It is illuminating to check explicitly how the assumption $v^{\mathrm{T}} u \neq 0$ below fails for a 2×2 Jordan block.

Lemma 1.2.8 (Algebraic simplicity from geometric simplicity)**.** *Let $\lambda \in \mathbb{C}$ be a geometrically simple eigenvalue*[3] *of some $A \in \mathbb{C}^{n\times n}$. If there exist eigenvectors u and v of A and A^{T} respectively for the eigenvalue λ satisfying $v^{\mathrm{T}} u \neq 0$, then λ is even an algebraically simple eigenvalue of A.*

Proof. Let $x \in \mathbb{C}^n$. It suffices to show that if $(\lambda - A)^2 x = 0$, then $(\lambda - A)x = 0$, so assume that $(\lambda - A)^2 x = 0$. Since $(\lambda - A)x$ is in the eigenspace $\ker(\lambda - A)$ which is spanned by u, there exists a scalar $\alpha \in \mathbb{C}$ such that $(\lambda - A)x = \alpha u$. Hence,

$$\alpha v^{\mathrm{T}} u = v^{\mathrm{T}}(\lambda - A)x = \left(\left(\lambda - A^{\mathrm{T}}\right) v\right)^{\mathrm{T}} x = 0.$$

As $v^{\mathrm{T}} u \neq 0$, this implies that $\alpha = 0$, so $(\lambda - A)x = 0$, as claimed. □

Now we have all the tools that we need to prove the Perron–Frobenius theorem.

Proof of Theorem 1.2.5. (a) We first consider the case $\sigma(A) = \{0\}$. In this case, one has $\mathrm{r}(A) = 0 \in \sigma(A)$. Moreover, there exists an integer $k \geq 1$ such that $A^k = 0$. Choose any non-zero vector $y \in \mathbb{R}^n_+$ and let $j \in \{0, 1, \dots, k-1\}$ be the maximal number for which $x := A^j y \neq 0$. Then x is positive since A^j is positive, and $x \in \ker A$.

Now we consider the more interesting case where $\sigma(A) \neq \{0\}$ and hence $\mathrm{r}(A) > 0$.

Choose an eigenvalue λ of A with modulus $|\lambda| = \mathrm{r}(A)$ and let $z \in \mathbb{C}^n$ be an eigenvector of norm 1 corresponding to λ. For every $s > 1$ one has $\mathcal{R}(s\lambda, A)z = \frac{1}{s\lambda - \lambda} z$, and thus

$$\begin{aligned}\frac{1}{(s-1)\,\mathrm{r}(A)}\,|z| &= \left|\frac{1}{s\lambda - \lambda} z\right| = |\mathcal{R}(s\lambda, A)z| = \left|\sum_{k=0}^{\infty} \frac{A^k}{(s\lambda)^{k+1}} z\right| \\ &\leq \sum_{k=0}^{\infty} \frac{\left|A^k\right|}{|s\lambda|^{k+1}}\,|z| = \sum_{k=0}^{\infty} \frac{A^k}{(s\,\mathrm{r}(A))^{k+1}}\,|z| = \mathcal{R}(s\,\mathrm{r}(A), A)\,|z|;\end{aligned}$$

where the penultimate equality uses the positivity of A^k (Proposition 1.1.5). Here we have twice used the Neumann series representation of the resolvent (Proposition 1.2.7(b)), which is applicable because $|s\lambda|, |s\,\mathrm{r}(A)| > \mathrm{r}(A)$.

If we take norms in the inequality $\frac{1}{(s-1)\,\mathrm{r}(A)}\,|z| \leq \mathcal{R}(s\,\mathrm{r}(A), A)\,|z|$ that we just proved, we get $\frac{1}{(s-1)\,\mathrm{r}(A)} \leq \|\mathcal{R}(s\,\mathrm{r}(A), A)\|$ (see the properties of the norm in Remark 1.1.6), so $\|\mathcal{R}(s\,\mathrm{r}(A), A)\| \to \infty$ as $s \downarrow 1$. By continuity of the resolvent (Proposition 1.2.7(a)), it follows that $\mathrm{r}(A)$ is not in the resolvent set and is thus an eigenvalue of A.

It remains to show the existence of an eigenvector $x \in \mathbb{R}^n_+$ for the eigenvalue $\mathrm{r}(A)$. Consider any sequence (s_k) in $(1, \infty)$ that converges to 1; for each index k we define

$$\alpha_k := \|\mathcal{R}(s_k\,\mathrm{r}(A), A)\,|z|\| \qquad \text{and} \qquad x_k := \frac{\mathcal{R}(s_k\,\mathrm{r}(A), A)\,|z|}{\alpha_k}.$$

[3] Recall that the **geometric multiplicity** of an eigenvalue λ of A is the dimension of the eigenspace $\ker(\lambda - A)$. The eigenvalue λ is called **geometrically simple** if its geometric multiplicity is one.

We have already seen that $\alpha_k \to \infty$ as $k \to \infty$ and that $x_k \geq 0$ for each k. Moreover,

$$\begin{aligned}(A - \mathrm{r}(A))x_k &= (A - s_k \mathrm{r}(A))x_k + (s_k \mathrm{r}(A) - \mathrm{r}(A))x_k \\ &= -\frac{|z|}{\alpha_k} + (s_k - 1)\,\mathrm{r}(A)x_k \to 0.\end{aligned}$$

Since $\|x_k\| = 1$ for all k and as the unit sphere in $\mathbb{C}^n$ is compact, there exists a subsequence (x_{k_j}) of (x_k) that converges to a non-zero vector $x \in \mathbb{R}^n_+$. Thus, $(A - \mathrm{r}(A))x = \lim_{j\to\infty}(A - \mathrm{r}(A))x_{k_j} = 0$, and so x is an eigenvector of A for the eigenvalue $\mathrm{r}(A)$.

(b) Assume now that all diagonal entries of A are non-zero. Then we can find a number $\delta > 0$ such that $A - \delta \geq 0$.[4] Consider the spectral radius $r := \mathrm{r}(A-\delta)$ of $A - \delta$. Since $A - \delta$ is positive, we can apply (a) to this matrix and thus see that r is an eigenvalue of $A - \delta$ and so $r + \delta$ is an eigenvalue of A. In particular, $0 < r + \delta \leq \mathrm{r}(A)$.

On the other hand, as all eigenvalues of $A - \delta$ are contained in the closed disk $B_{\leq r}(0)$ with radius r and centre 0, so all eigenvalues of A are contained in the disk $B_{\leq r}(\delta)$ with radius r and centre δ. Therefore, $\mathrm{r}(A) \leq r + \delta$. It follows that $\mathrm{r}(A) = r + \delta$. But the circle with radius $r + \delta$ and centre 0 intersects the disk $B_{\leq r}(\delta)$ only in the point $r + \delta$, so A has no further eigenvalue of modulus $r + \delta = \mathrm{r}(A)$.

(c) Finally, assume that A is strongly positive. For every eigenvector $x \in \mathbb{R}^n_+$ of A corresponding to the eigenvalue $\mathrm{r}(A)$ – which exists according to (a) – one has $\mathrm{r}(A)x = Ax \succeq \mathbb{1}$. Hence $\mathrm{r}(A) > 0$ and $x \succeq \mathbb{1}$.

Next we show that the eigenvalue $\mathrm{r}(A)$ is geometrically simple. To this end, let $x \succeq \mathbb{1}$ be an eigenvector of A for the eigenvalue $\mathrm{r}(A)$ and let $y \in \mathbb{R}^n$ be any other eigenvector for the same eigenvalue. Then there exists a number $\gamma \in \mathbb{R} \setminus \{0\}$ such that $x - \gamma y$ is positive, but has at least one component that is 0. If $x - \gamma y$ were non-zero, it would be an eigenvector of A for the eigenvalue $\mathrm{r}(A)$, which would imply $x - \gamma y \succeq \mathbb{1}$, as we have just seen. Thus, $x - \gamma y = 0$, so y is a multiple of x. This proves the geometric simplicity of the eigenvalue $\mathrm{r}(A)$.

To see that $\mathrm{r}(A)$ is algebraically simple, we now use Lemma 1.2.8. By applying (a) to the transposed matrix A^{T}, one gets an eigenvector $y \geq 0$ of A^{T} for the eigenvalue $\mathrm{r}(A^{\mathrm{T}}) = \mathrm{r}(A)$. As $y \neq 0$ and $x \succeq \mathbb{1}$, one has $y^{\mathrm{T}}x > 0$, so Lemma 1.2.8 is applicable and shows that the geometric simplicity of $\mathrm{r}(A)$ implies the algebraic simplicity. □

1.3 Positive matrix semigroups

The powers A^k of a square matrix give the solutions $x\colon \mathbb{N}_0 \to \mathbb{C}^n$ to the difference equation $x(k) = Ax(k-1)$ for $k \in \mathbb{N}$. As in the scalar case, it is natural to study the continuous time analogue of this dynamical system, i.e. the differential equation $\dot{x}(t) = Ax(t)$ with $x\colon [0,\infty) \to \mathbb{C}^n$. For this, one uses the matrix exponential function.

[4] Let us recall here the convention $A - \delta := A - \delta\,\mathrm{id}$ that we first used Definition 1.2.6.

Definition 1.3.1 (Matrix exponential function)**.** For every $A \in \mathbb{C}^{n\times n}$ one defines

$$e^A := \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathbb{C}^{n\times n},$$

where the series converges absolutely in $\mathbb{C}^{n\times n}$.

We first discuss a number of essential properties of the matrix exponential function, in particular its relation to linear differential equations. Positivity takes the stage back in Theorems 1.3.8 and 1.3.9.

Proposition 1.3.2 (Properties of the matrix exponential function)**.** *The matrix exponential function has the following properties:*

(a) $e^0 = \mathrm{id}$.

(b) *The matrix exponential function* $\exp\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, $A \mapsto e^A$ *is continuous.*

(c) *For fixed* $A \in \mathbb{C}^{n\times n}$, *the mapping* $\mathbb{C} \to \mathbb{C}^{n\times n}$, $z \mapsto e^{zA}$ *is differentiable, and hence analytic, with derivative* $\frac{\mathrm{d}}{\mathrm{d}z} e^{zA} = Ae^{zA} = e^{zA}A$ *at each* $z \in \mathbb{C}$.

(d) *If two matrices* $A, B \in \mathbb{C}^{n\times n}$ *satisfy* $AB = BA$, *then* $e^{A+B} = e^A e^B$.

Proof. (a) This follows readily from the definition of the matrix exponential function.

(b) Let $A, B \in \mathbb{C}^{n\times n}$. An induction argument yields the geometric sum formula

$$A^k - B^k = \sum_{j=0}^{k-1} A^j (A-B) B^{k-1-j}$$

for all integers $k \geq 1$. On the right hand side, it is important to have $A - B$ in the middle since A and B are not assumed to commute. Thus we can estimate $\left\|A^k - B^k\right\| \leq k\alpha^{k-1} \|A - B\|$ with $\alpha := \max\{\|A\|, \|B\|\}$. The continuity now follows from

$$\left\|e^A - e^B\right\| \leq \sum_{k=1}^{\infty} \frac{1}{k!} \left\|A^k - B^k\right\| \leq \|A - B\| \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{(k-1)!} = e^{\alpha} \|A - B\|.$$

(c) This can be shown in the same way as for the scalar-valued exponential function.

(d) One can prove this by using the Cauchy product formula for infinite series, as in the scalar-valued case. Readers familiar with the uniqueness theorem for ordinary differential equations might also find the following alternative proof insightful.

Consider the functions $X_1, X_2 \colon \mathbb{R} \to \mathbb{C}^{n\times n}$ that are given by

$$X_1(t) = e^{t(A+B)} \qquad \text{and} \qquad X_2(t) = e^{tA}e^{tB}$$

for all $t \in \mathbb{R}$. According to (a) and (c) the function X_1 solves the initial value problem

$$\begin{cases} \dot{X}(t) = (A+B)X(t) & \text{for all } t \in \mathbb{R}, \\ X(0) = \mathrm{id}. \end{cases}$$

On the other hand, as A and B commute, the definition of the matrix exponential function implies that B also commutes with e^{tA} for all $t \in \mathbb{R}$. This together with (c) and the product rule for differentiation implies that X_2 solves the same initial value problem. Hence, by the uniqueness theorem for linear initial value problems it follows that $X_1(t) = X_2(t)$ for all $t \in \mathbb{R}$. For $t = 1$ this gives the claim. □

Proposition 1.3.2(c) has the following consequence, which is the main reason why one is interested in matrix exponential functions.

Corollary 1.3.3 (Solutions to linear differential equations)**.** *Let $A \in \mathbb{C}^{n\times n}$ and $x_0 \in \mathbb{C}^n$. Then the function $x : [0,\infty) \to \mathbb{C}^n$, $t \mapsto e^{tA}x_0$ satisfies the initial value problem*

$$\begin{cases} \dot{x}(t) = Ax(t) & \text{for all } t \in [0,\infty), \\ x(0) = x_0. \end{cases}$$

From the uniqueness theorem for ordinary differential equations, the function x is in fact the only solution to the initial value problem in Corollary 1.3.3.

For a matrix $A \in \mathbb{C}^{n\times n}$, Corollary 1.3.3 shows that the matrix family $(e^{tA})_{t\geq 0}$ is a quite fundamental object. Hence, it gets its own name, which is inspired by the property $e^{(s+t)A} = e^{sA}e^{tA}$ for all $s,t \geq 0$ that follows from Proposition 1.3.2(d).

Definition 1.3.4 (Matrix semigroups and positivity)**.**

(a) Let $A \in \mathbb{C}^{n\times n}$. The family $(e^{tA})_{t\geq 0}$ is called the **matrix semigroup** generated by A.

(b) Let $A \in \mathbb{R}^{n\times n}$. Then $(e^{tA})_{t\geq 0}$ is called **positive** if $e^{tA} \geq 0$ for all $t \in [0,\infty)$.

We have seen (in Proposition 1.2.2) that the spectral radius of a matrix A determines the long-term behaviour of the powers A^k. For the matrix semigroup $(e^{tA})_{t\geq 0}$, a similar role is played by the so-called **spectral bound**.

Definition 1.3.5 (The spectral bound of a matrix)**.** Let $A \in \mathbb{C}^{n\times n}$. The number

$$\mathrm{s}(A) := \max\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\}$$

is called the **spectral bound** of A.

Proposition 1.3.6 (Convergence to 0 of matrix semigroups)**.** *For each matrix $A \in \mathbb{C}^{n\times n}$, the following are equivalent:*

(i) $\mathrm{s}(A) < 0$.

(ii) $e^{tA} \to 0$ *as* $t \to \infty$.

(iii) *There exist numbers* $\mu < 0$ *and* $c \geq 0$ *such that* $\|e^{tA}\| \leq ce^{t\mu}$ *for each* $t \geq 0$.

Proof. "(i) ⇒ (iii)": As in the proof of Proposition 1.2.2, this can be deduced using the Jordan normal form of A. We refer to Exercise 1.4(d) for a detailed discussion; cf. proof of Proposition 1.2.2.

"(iii) ⇒ (ii)": This implication is obvious.

"(ii) ⇒ (i)": Let λ be an eigenvalue of A with real part $\operatorname{Re}\lambda = \operatorname{s}(A)$ and an associated eigenvector $z \in \mathbb{C}^n$ of norm one. For every $k \in \mathbb{N}_0$ one has $A^k z = \lambda^k z$ and thus, $e^{tA}z = e^{t\lambda}z$ for every $t \geq 0$ by the definition of the matrix exponential function. So $e^{t\operatorname{s}(A)} = e^{t\operatorname{Re}\lambda}\|z\| = \|e^{tA}z\| \to 0$ for each $t \to \infty$, which shows that $\operatorname{s}(A) < 0$. □

The Neumann series representation of the resolvent of a matrix A (given in Proposition 1.2.7(b)) has the following analogue in continuous time.

Lemma 1.3.7 (Laplace transform representation of the resolvent)**.** *Let* $A \in \mathbb{C}^{n\times n}$. *For every* $\lambda \in \mathbb{C}$ *that satisfies* $\operatorname{Re}\lambda > \operatorname{s}(A)$ *one has*

$$\mathcal{R}(\lambda, A) = \int_0^\infty e^{-t\lambda}e^{tA}\,\mathrm{d}t,$$

where the integral converges absolutely.

Proof. Let $\operatorname{Re}\lambda > \operatorname{s}(A)$. Then $\operatorname{s}(A-\lambda) < 0$ and so by Proposition 1.3.6, there are numbers $\mu < 0$ and $c \geq 0$ such that $\|e^{-t\lambda}e^{tA}\| \leq ce^{t\mu}$ for all $t \geq 0$. Hence, the integral indeed converges absolutely.

To prove that the integral equals $\mathcal{R}(\lambda, A)$, observe that

$$(\lambda - A)\int_0^\infty e^{-t\lambda}e^{tA}\,\mathrm{d}t = \lim_{T\to\infty} -\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}e^{t(A-\lambda)}\,\mathrm{d}t = \lim_{T\to\infty}\left(-e^{T(A-\lambda)} + \mathrm{id}\right) = \mathrm{id};$$

where the last equality uses again that $\operatorname{s}(A-\lambda) < 0$, which indeed gives $e^{T(A-\lambda)} \to 0$ as $T \to \infty$ according to Proposition 1.3.6. □

Except in special cases in small dimensions, it is typically not possible to explicitly compute e^{tA} for a given matrix A. Fortunately, one can check positivity of the semigroup $(e^{tA})_{t\geq 0}$ purely in terms of A, as condition (iv) in the following theorem shows.

Theorem 1.3.8 (Characterisation of positive matrix semigroups)**.** *Let* $A \in \mathbb{R}^{n\times n}$. *The following are equivalent:*

(i) $e^{tA} \geq 0$ *for all real numbers* $t \geq 0$.

(ii) *For all real numbers* $\lambda > \operatorname{s}(A)$ *one has* $\mathcal{R}(\lambda, A) \geq 0$.

(iii) *For all sufficiently large real numbers* $\lambda > \operatorname{s}(A)$ *one has* $\mathcal{R}(\lambda, A) \geq 0$.

(iv) *All off-diagonal entries of* A *are in* $[0,\infty)$, *i.e.* $A_{jk} \geq 0$ *for all indices* $j \neq k$.

Proof. "(i) ⇒ (ii)": This follows from the representation of the resolvent $\mathcal{R}(\lambda, A)$ as the Laplace transform of the semigroup $(e^{tA})_{t\geq 0}$ given in Lemma 1.3.7.

"(ii) ⇒ (iii)": This implication is obvious.

"(iii) ⇒ (iv)": Consider numbers $\lambda \in \mathbb{R}$ that satisfy $\lambda > \mathrm{r}(A)$. The Neumann series representation of the resolvent (Proposition 1.2.7(b)) shows that

$$\lambda^2 \mathcal{R}(\lambda, A) - \lambda \,\mathrm{id} = \sum_{k=1}^{\infty} \frac{A^k}{\lambda^{k-1}} \to A$$

as $\lambda \to \infty$. For indices $j \neq k$ one thus gets

$$A_{jk} = \lim_{\lambda\to\infty} \big(\lambda^2 \mathcal{R}(\lambda, A) - \lambda \,\mathrm{id}\big)_{jk} = \lim_{\lambda\to\infty} \lambda^2 \mathcal{R}(\lambda, A)_{jk} \geq 0,$$

where the last inequality follows from (iii).

"(iv) ⇒ (i)": As (iv) holds there exists a number $c \in \mathbb{R}$ such that $A + c\,\mathrm{id} \geq 0$. So

$$0 \leq e^{t(A+c\,\mathrm{id})} = e^{tc\,\mathrm{id}} e^{tA} = e^{tc} e^{tA}$$

for all $t \in [0,\infty)$, where the inequality at the beginning follows from $A + c\,\mathrm{id} \geq 0$ and the definition of matrix exponential function, and the first equality follows from Proposition 1.3.2(d). Division by the numbers $e^{tc} \in (0,\infty)$ yields (i). □

Several other equivalent conditions for positivity of matrix semigroups can be found in Exercise 1.2. We conclude this lecture with a Perron–Frobenius type theorem for positive matrix semigroups. It is remarkable that in part (b) of the following theorem, no additional assumption on A is needed. This is in sharp contrast to situation for single operators, where we needed an additional assumption in Theorem 1.2.5(b).

Theorem 1.3.9 (Perron–Frobenius for positive matrix semigroups)**.** *Let $A \in \mathbb{R}^{n\times n}$ and assume the matrix semigroup $(e^{tA})_{t\geq 0}$ is positive.*

(a) $\mathrm{s}(A)$ *is an eigenvalue of A and there exists a corresponding eigenvector $x \geq 0$.*

(b) $\mathrm{s}(A)$ *is a* ***strictly dominant*** *eigenvalue of A in the sense that* $\operatorname{Re}\lambda < \mathrm{s}(A)$ *for all* $\lambda \in \sigma(A) \setminus \{\mathrm{s}(A)\}$.

Proof. Since $(e^{tA})_{t\geq 0}$ is positive, there exists $c \in \mathbb{R}$ such that $A + c \geq 0$ by Theorem 1.3.8. Therefore by the Perron–Frobenius theorem for positive matrices (Theorem 1.2.5), the spectral radius $\mathrm{r}(A+c)$ is an eigenvalue of $A+c$ with a positive eigenvector. Consequently, it equals $\mathrm{s}(A+c)$ and is a strictly dominant eigenvalue of $A+c$. The assertions thus follow from $\sigma(A+c) = \sigma(A) + c$. □

We end this chapter by pointing out that a similar result as in Theorem 1.2.5(c) can also be proved for matrix semigroups if e^{tA} is strongly positive for every $t > 0$. We do not discuss this further at this point, but a result in the next chapter, Theorem 2.3.1, will contain this as a special case.

Exercises for Chapter 1

Exercise 1.1. Let $A, B \in \mathbb{C}^{n\times n}$.

(a) Give an example to show that $e^{A+B} = e^A e^B$ does not imply $AB = BA$.

(b) If there exists $\varepsilon > 0$ such that $e^{t(A+B)} = e^{tA}e^{tB}$ for all $t \in [0, \varepsilon)$, then show $AB = BA$.

Exercise 1.2 (Continuation of Theorem 1.3.8)**.** Let $A \in \mathbb{R}^{n\times n}$ be given. Prove that the following are equivalent:

(iv) All off-diagonal entries of A are in $[0, \infty)$, i.e. $A_{jk} \geq 0$ for all indices $j \neq k$.

(v) The matrix A satisfies the *positive minimum principle*, i.e. for all $u \in \mathbb{R}^n_+$ and all $k \in \{1, \dots, n\}$ with $u_k = 0$ one has $(Au)_k \geq 0$.

(vi) The matrix A is *cross positive*, i.e. for all $u, v \in \mathbb{R}^n_+$ with $u^{\mathrm{T}} v = 0$ one has $u^{\mathrm{T}} Av \geq 0$.

(vii) The matrix A satisfies the *Beurling–Deny criterion*, i.e. for every $u \in \mathbb{R}^n$ one has $(u^-)^{\mathrm{T}} Au^+ \geq 0$, where

$$(u^+)_k := \begin{cases} u_k & \text{if } u_k \geq 0, \\ 0 & \text{if } u_k < 0 \end{cases}$$

for all $k \in \{1, \dots, n\}$, and where $u^- := (-u)^+$.

(viii) The matrix A satisfies the *Arendt–Kato inequality*, i.e. for all $u \in \mathbb{R}^n$ and all indices k with $u_k \geq 0$ one has $(Au^+)_k \geq (Au)_k$.

Exercise 1.3. The koala (*Phascolarctos cinereus*) is a notoriously lazy animal, sleeping up to 20 hours a day. It is also a very picky eater. Suppose that a particular koala has 3 favourite eucalyptus trees, arranged as in Figure 1.3.1.

For $i, j \in \{1, 2, 3\}$, let P_{ij} denote the probability that the koala will eat at tree i the following day given that it has eaten at tree j today. Consider the following model:

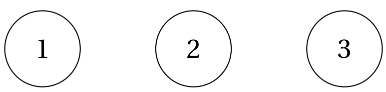

Figure 1.3.1: Eucalyptus trees in an Australian forest (some imagination is required).

- With probability $q \in (0,1)$, the koala will stay at the same tree the following day.
- Since it is lazy, the koala will only move to adjacent trees. Hence, if it has eaten at tree 2 on one day, it will move to either tree 1 or 3 the next day with equal probability (or otherwise stay in place). On the other hand, if it has eaten at tree 1 or 3, it will only move to tree 2 (or otherwise stay in place).

(a) For the probabilities P_{ij} described above, write down the matrix $P = (P_{ij})_{1\le i,j\le 3}$, which is called the **transition matrix** of the model, in terms of q. Explain what the (i,j)-th entry of the matrix powers P^k represents.

(b) Show that $\operatorname{r}(P) = 1$ and that 1 is a strictly dominant eigenvalue.

(c) In the long run, what can you say about the proportion of days the koala spends at each tree?

Exercise 1.4. Let $\lambda_0 \in \mathbb{C}$ and let $J_0 \in \mathbb{C}^{n_0\times n_0}$ denote the Jordan block

$$J_0 = \begin{pmatrix} \lambda_0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}.$$

(a) Find and prove a formula for the matrix J_0^k for every $k \in \mathbb{N}_0$.

(b) Find and prove a formula for the matrix e^{tJ_0} for every $t \in [0,\infty)$.

Hint: First consider the case $\lambda_0 = 0$ and then use Proposition 1.3.2(d).

(c) Let $A \in \mathbb{C}^{n\times n}$ such that $\operatorname{r}(A) > 0$. Use the Jordan normal form of A to show that there exists a number $c > 0$ such that $\|A^k\| \le c\operatorname{r}(A)^k(1+k^{n-1})$ for each $k \in \mathbb{N}_0$.

(d) Let $A \in \mathbb{C}^{n\times n}$. Use the Jordan normal form of A to show that there exists a number $c > 0$ such that $\|e^{tA}\| \le ce^{t\operatorname{s}(A)}(1+t^{n-1})$ for each $t \in [0,\infty)$.

Notes for Chapter 1

Positivity versus non-negativity

As promised in Remark 1.1.2, we now discuss the terminology *positive* and the related question of whether 0 is considered positive in a bit more detail.

Real numbers: It is remarkable that even for real numbers, the meaning of the term *positive* depends on the language. The convention that positivity of a real number α means $\alpha > 0$ – while the property $\alpha \ge 0$ is often referred to as α being *non-negative* – is common in English and, for instance, also in German. On the other hand, in French the adjective *positif* typically refers to a number $\alpha \ge 0$.

As the real numbers are defined as an ordered field with a number of additional properties, it is worthwhile to also take a brief look at conventions in the theory of ordered groups and fields. Unsurprisingly, the French meaning of 'positive' is employed by Bourbaki in their definition of ordered groups [Bou07, p. A VI.4]. The same convention is then used in the English translation [Bou03, p. A VI.4].

Linear algebra: A substantial amount of literature studies order properties of $\mathbb{R}^n$ and matrices, in particular in relation to the Perron–Frobenius theorem and its applications. In this field, it seems to be most common to call a vector x *positive* if $x_k > 0$ for all indices k (note that we call this property *strongly positive* in Definition 1.2.3). Vectors $x \ge 0$ are typically referred to as *non-negative* vectors in this part of the literature. This terminology has the advantage that it is consistent with the standard conventions for real numbers in English.

Care must be taken, though, since the coordinate-wise relation $\le$ defines only a partial order on $\mathbb{R}^n$ when $n \ge 2$: there exist vectors $x \in \mathbb{R}^n$ that satisfy neither $x \ge 0$ nor $x \le 0$. If one gives in to the temptation to call a vector $x \in \mathbb{R}^n$ *negative* if $-x$ is positive, then the common terminology in linear algebra leads to the following situation: x is negative if and only if $x_k < 0$ for all indices k and thus we have a linguistically unpleasant situation where the assertion "x is non-negative" is inequivalent to "x is not negative".

Real-valued functions: Some parts of the literature adopt a convention similar to the one in linear algebra: a function $f : \Omega \to \mathbb{R}$ defined on a set Ω is called *non-negative* if $f(\omega) \geq 0$ for all $\omega \in \Omega$ or, for instance in the setting of L^p-spaces, for almost all $\omega \in \Omega$. Accordingly, f is then called *positive* if $f(\omega) > 0$ for (almost) all $\omega \in \Omega$.

This adaptation of the finite-dimensional perspective comes with an additional caveat in infinite dimensions that only becomes apparent when one develops a systematic theory of ordered spaces in infinite dimensions. The property $f(\omega) > 0$ for (almost) all $\omega \in \Omega$ has very different consequences depending on the surrounding space. For example, in $C(K)$ it implies that f dominates a strictly positive constant on K, whereas in $L^p(\Omega)$ it does not. We elaborate on this later in Chapter 7, when we have enough Banach lattice theory available.

Elements of ordered vector spaces and Banach lattices: In the theory of ordered vector spaces and Banach lattices, which we introduce in Chapter 4, it is common to call a vector x *positive* if $x \geq 0$; in particular, the zero vector is positive. We follow this convention since we frequently use Banach lattice theory later on. To maintain consistency throughout these notes, we have adopted the same convention in the finite-dimensional setting, as can be seen in Definition 1.1.1.

Perron–Frobenius and friends

The story of the Perron–Frobenius theorem, and the theory of non-negative matrices[5] in general, has a surprising beginning. At the turn of the 20th century, at the University of Munich (LMU), Oskar Perron was studying the problem of convergence of continued fraction algorithms (German: *Kettenbruchalgorithmen*), following the work of his colleague Alfred Pringsheim. His breakthrough was to reduce the problem to a study of the eigenvalue equation for specific matrices with positive entries (although he did not use this terminology). Consequently Perron was able to simplify the convergence criteria from earlier works (such as those of Pringsheim), and moreover, his methods could be extended to treat the more general case of Jacobi algorithms. This became the subject of his *Habilitation* paper [Per07a], published in *Mathematische Annalen* in 1907.

Clearly Perron recognised the utility of his methods beyond their original purpose and the potential for a systematic theory, for he then followed up with the article *Zur Theorie der Matrices*, which was also published in the *Mathematische Annalen* [Per07b]. In this work, Perron's main theorem corresponds more or less to Theorem 1.2.5(c) in this chapter. Moreover, he could derive the same conclusions under the weaker assumption that $A \geq 0$ and A^k is positive (in our terminology, strongly positive) for some $k \in \mathbb{N}$. However, he expressed dissatisfaction with his rather convoluted argument to achieve this generalisation, and in addition he left open the possibility that a larger class of non-negative matrices could satisfy the conclusions of his theorem.

At this point, Frobenius enters the story. In a series of three papers [Fro08, Fro09, Fro12], he manages to resolve the issues raised by Perron. In a 1908 paper and its sequel in 1909, he proves strengthened versions of Perron's results for positive matrices

[5] In this historical account, we use the classical terminology from linear algebra as explained at the beginning of these notes.

using thoroughly linear-algebraic techniques, especially the properties of determinants. Then, in the 1912 paper, Frobenius extends his investigations to encompass irreducible non-negative matrices. Here, *irreducible* refers to matrices that cannot be put into block upper-triangular form via simultaneous row or column permutations. We encourage readers who are interested in further details (both mathematical and biographical) of the history of the Perron-Frobenius theorem to consult the article [Haw08].

The study of non-negative matrices, stemming from the ideas of Perron and Frobenius, has proved to be very fruitful, and has found diverse applications in the natural and social sciences, from population models (e.g. Leslie matrices) to queuing theory, to input-output models in economics (e.g. the Leontief model), to Google's PageRank algorithm. Thus, the literature on 'Perron–Frobenius theory' is vast. Two texts that are now considered classical include [Sen06], which deals with applications to probability theory and in particular Markov chains, and [BP94], which is notable for its systematic study of positivity with respect to general cones in $\mathbb{R}^n$.

Finally, we cannot omit a mention of the fairly recent monograph [BKFR17], which was based on the material of the 17th Internet Seminar (2013–2014). Part I of that book contains an accessible exposition of the Perron–Frobenius theorem (in particular, Frobenius' contributions), properties of (positive) matrix exponential functions, and numerous applications.

Chapter 2

Eventual positivity in finite dimensions

In this chapter, we encounter our main protagonist, eventual positivity (Section 2.2). We shall see that many spectral properties of positive matrix semigroups remain true for eventually positive ones, but to prove this, we first need more advanced tools from spectral theory (Section 2.1). Remarkably though, already in the finite dimensional setting, there are significant differences between the positive and the eventually positive case. The characterisation of eventual positivity of a matrix semigroup $(e^{tA})_{t\geq 0}$ in terms of A (Section 2.3) has a different flavour than for positive semigroups, and the differences become even clearer when it comes to perturbation theory (Section 2.4).

2.1 Prelude: Spectral decomposition of matrices

For the analysis of matrix powers and exponentials in Exercise 1.4, you have already encountered a very useful tool: the Jordan normal form. In this section, the same tool is used to derive a variety of spectral properties, so let us fix the notation for it.

Let $A \in \mathbb{C}^{n\times n}$. By a coordinate transform, one can achieve that A is in **Jordan normal form**. This means that A can be written in block diagonal form as

$$A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix} \tag{2.1.1}$$

for matrices $J_k \in \mathbb{C}^{n_k \times n_k}$ that are called **Jordan blocks**, i.e. each of them has the form

$$J_k = \begin{pmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{pmatrix} \tag{2.1.2}$$

for a number $\lambda_k \in \mathbb{C}$. The numbers $\lambda_1, \dots, \lambda_m$ are the eigenvalues of A, where the number of occurrences of each eigenvalue in this list coincides with its geometric multiplicity. In other words, the geometric multiplicity of λ_k is the number of Jordan blocks associated to λ_k. On the other hand, the algebraic multiplicity of λ_k is the sum of the sizes of all Jordan blocks corresponding to λ_k. We set

$$N_k := J_k - \lambda_k = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},$$

which is nilpotent of index n_k.

For an open set $\emptyset \neq \Omega \subseteq \mathbb{C}$, a point $\lambda \in \Omega$, and an analytic function $f \colon \Omega \setminus \{\lambda\} \to \mathbb{C}^{m \times n}$, the point λ is called a **pole** of f if for some $p \in \mathbb{N}$, the $\lim_{\mu \to \lambda} (\mu - \lambda)^p f(\mu)$ exists and is non-zero. In this case, p is unique and is called the **pole order** of λ. By considering the entries of f, one can see that a similar Laurent series expansion as in the scalar-valued case continues to hold.

Proposition 2.1.1 (Eigenvalues are poles of the resolvent)**.** *Let $\lambda \in \mathbb{C}$ be an eigenvalue of a matrix $A \in \mathbb{C}^{n \times n}$. If the largest Jordan block of A corresponding to λ has size $p \geq 1$, then*

(a) *λ is a pole of the resolvent $\mathcal{R}(\cdot, A)$ with pole order p, i.e. there exist matrices $Q_j \in \mathbb{C}^{n \times n}$ for $j \geq -p+1$ such that $Q_{-p+1} \neq 0$ and*[1]

$$(\mu - A)^{-1} = \sum_{j=-p}^{\infty} Q_{j+1} (\mu - \lambda)^j$$

for all $\mu \neq \lambda$ that are sufficiently close to λ.

(b) *The range of Q_{-p+1} is contained in the eigenspace* $\ker(\lambda - A)$.

Proof. (a) We may assume that A is in Jordan normal form (2.1.1). For every $\mu \in \rho(A)$ one then has

$$(\mu - A)^{-1} = \begin{pmatrix} (\mu - J_1)^{-1} & & \\ & \ddots & \\ & & (\mu - J_m)^{-1} \end{pmatrix}.$$

For each $k \in \{1, \dots, m\}$ the matrix $N_k = J_k - \lambda \in \mathbb{C}^{n_k \times n_k}$ is nilpotent, so by the Neumann series representation of the resolvent (Proposition 1.2.7(b)) one gets from $(\mu - J_k)^{-1} = (\mu - \lambda_k)^{-1} \left(\mathrm{id} - (\mu - \lambda_k)^{-1} N_k\right)^{-1}$ that

$$(\mu - J_k)^{-1} = \sum_{j=0}^{n_k - 1} \frac{N_k^j}{(\mu - \lambda_k)^{j+1}} = \sum_{j=-n_k}^{-1} N_k^{-(j+1)} (\mu - \lambda_k)^j. \tag{2.1.3}$$

[1] Note that our enumeration of the coefficients Q_j is shifted by 1 compared to the usual enumeration of coefficients of Laurent series. A first indication that this index shift is a good idea can be found in formula (2.1.3), where the same shifted index occurs. When we discuss the Laurent series of resolvents in infinite dimension in Chapter 6 the shifted index will make many formulae much easier.

One can see in (2.1.3) that $\mu \mapsto (\mu - J_k)^{-1}$ does not have a pole at λ if $\lambda \neq \lambda_k$ and that it has a pole of order n_k at λ if $\lambda = \lambda_k$. Thus, λ is a pole of $\mathcal{R}(\cdot, A)$ of order $\max\{n_k : k \in \{1,\dots,m\} \text{ with } \lambda_k = \lambda\} = p$ and the coefficients Q_{j+1} are block diagonal matrices whose entries can be seen in (2.1.3). In particular, $Q_{-p+1} \neq 0$.

(b) For each index k, $\operatorname{rg}(N_k^{p-1}) \subseteq \ker(\lambda_k - J_k)$ because

$$(\lambda_k - J_k)N_k^{p-1} = -N_k^p = 0.$$

From the proof of (a), we know that the non-zero blocks of Q_{-p+1} are the matrices N_k^{p-1} for exactly those k for which $\lambda_k = \lambda$ and $n_k = p$. So $\operatorname{rg} Q_{-p+1} \subseteq \ker(\lambda - A)$. □

The formula (2.1.3) has a number of useful consequences that we discuss now. We need the following crucial observation about contour integration in complex analysis.

Proposition 2.1.2. *Let $\lambda \in \mathbb{C}$ and let γ be a closed C^1-path in $\mathbb{C} \setminus \{\lambda\}$.*

(a) *For each integer $j \neq 1$ one has $\oint_\gamma \frac{1}{(\mu-\lambda)^j}\,\mathrm{d}\mu = 0$.*

(b) *If γ encircles λ precisely once anticlockwise, then $\frac{1}{2\pi \mathrm{i}}\oint_\gamma \frac{1}{\mu-\lambda}\,\mathrm{d}\mu = 1$.*

(c) *If γ does not encircle λ, then $\oint_\gamma \frac{1}{\mu-\lambda}\,\mathrm{d}\mu = 0$.*

Proposition 2.1.2 can be applied to every entry of $(\mu - J_k)^{-1}$ in formula (2.1.3). Doing this for all the Jordan blocks of A, one immediately obtains the following.

Corollary 2.1.3. *Let $A \in \mathbb{C}^{n\times n}$ and let γ be a closed C^1-path in $\mathbb{C} \setminus \sigma(A)$.*

(a) *If γ encircles each eigenvalue of A precisely once, then $\frac{1}{2\pi \mathrm{i}}\oint_\gamma (\mu - A)^{-1}\,\mathrm{d}\mu = \mathrm{id}$.*

(b) *If γ does not encircle any eigenvalue of A, then $\oint_\gamma (\mu - A)^{-1}\,\mathrm{d}\mu = 0$.*

Corollary 2.1.3 now comes in handy to prove the representation formula, and hence the uniqueness, in the following proposition.

Proposition 2.1.4 (Spectral decomposition)**.** *Let $A \in \mathbb{C}^{n\times n}$ and let $\sigma_0 \subseteq \sigma(A)$. There exists precisely one projection $P \in \mathbb{C}^{n\times n}$ that has the following properties:*

(a) *P commutes with A.*

(b) *The restrictions of A to the range and the kernel of P have the spectra*

$$\sigma\left(A|_{\operatorname{rg} P}\right) = \sigma_0 \qquad \text{and} \qquad \sigma\left(A|_{\ker P}\right) = \sigma(A) \setminus \sigma_0.$$

Moreover, P is given by the formula

$$P = \frac{1}{2\pi \mathrm{i}} \oint_\gamma (\mu - A)^{-1}\,\mathrm{d}\mu \tag{2.1.4}$$

for any closed C^1-path γ in $\mathbb{C}$ that encircles each element of σ_0 precisely once, but no element of $\sigma(A) \setminus \sigma_0$.

Proof. Existence: After a coordinate transform, we may assume that A is in Jordan normal form, i.e. that it is given by the formula (2.1.1); we use the notation specified next to this formula. After ordering the eigenvalues $\lambda_1, \dots, \lambda_m$ appropriately, we can find an $\ell \in \{0, 1, \dots, m\}$ such that $\sigma_0 = \{\lambda_k : 1 \le k \le \ell\}$ and $\sigma(A) \setminus \sigma_0 = \{\lambda_k : \ell < k \le m\}$; we allow $\ell = 0$ to account for the case $\sigma_0 = \emptyset$. Now, let $P \in \mathbb{C}^{n \times n}$ be the projection onto the first $n_1 + \cdots + n_\ell$ components of $\mathbb{C}^n$. Then P has the required properties.

Uniqueness and integral formula: Let $P \in \mathbb{C}^{n \times n}$ be a projection that satisfies (a) and (b) and let the complex path γ have the properties specified at the end of the proposition. It suffices to show that P is given by the claimed path integral.

Since P commutes with A, it also commutes with $(\mu - A)^{-1}$ for every $\mu \in \rho(A)$ and one has $(\mu - A)^{-1}|_{\operatorname{rg} P} = (\mu - A|_{\operatorname{rg} P})^{-1}$ for all such μ; the same also holds for the restriction to $\ker P$. So Corollary 2.1.3(a) applied to $A|_{\operatorname{rg} P}$ gives[2]

$$\frac{1}{2\pi \mathrm{i}} \oint_\gamma (\mu - A)^{-1} \,\mathrm{d}\mu|_{\operatorname{rg} P} = \frac{1}{2\pi \mathrm{i}} \oint_\gamma (\mu - A|_{\operatorname{rg} P})^{-1} \,\mathrm{d}\mu = \mathrm{id}_{\operatorname{rg} P} = P|_{\operatorname{rg} P}$$

and similarly, Corollary 2.1.3(b) applied to $A|_{\ker P}$ gives

$$\frac{1}{2\pi \mathrm{i}} \oint_\gamma (\mu - A)^{-1} \,\mathrm{d}\mu|_{\ker P} = \frac{1}{2\pi \mathrm{i}} \oint_\gamma (\mu - A|_{\ker P})^{-1} \,\mathrm{d}\mu = 0 = P|_{\ker P}.$$

This shows the claimed formula for P. ☐

Definition 2.1.5 (Spectral projections)**.** In the situation of Proposition 2.1.4, the projection P is called the **spectral projection** of A associated to σ_0.

In the situation of Proposition 2.1.4, note that the complementary projection $1 - P$ also commutes with A and satisfies $\operatorname{rg}(1-P) = \ker P$ and $\ker(1-P) = \operatorname{rg} P$. Hence, it follows that $1 - P$ is the spectral projection of A associated to $\sigma(A) \setminus \sigma_0$.

Recall that an eigenvalue λ of a square matrix $A \in \mathbb{C}^{n \times n}$ is called **semisimple** if the generalised eigenspace $\bigcup_{k=1}^n \ker(\lambda - A)^k$ coincides with the eigenspace $\ker(\lambda - A)$, i.e. if the geometric and algebraic multiplicities of λ coincide.

Proposition 2.1.6. *Let $A \in \mathbb{C}^{n \times n}$ and let P be the spectral projection of A associated to an eigenvalue $\lambda \in \sigma(A)$.*[3]

(a) *P is equal to the coefficient Q_0 of the term $(\mu - \lambda)^{-1}$ in the Laurent series expansion of $(\mu - A)^{-1}$ in Proposition 2.1.1(a).*

(b) *$\operatorname{rg} P = \bigcup_{k=1}^n \ker(\lambda - A)^k$, i.e. the range of P coincides with the generalised eigenspace of λ. In particular, $\dim \operatorname{rg} P$ is the algebraic multiplicity of the eigenvalue λ.*

[2]Note that we are slightly imprecise here. We use several notions and properties for linear maps between finite dimensional vector spaces now, although we introduced them only for matrices. This does not cause problems since all those concepts are consistent with linear changes of the coordinates and can thus be transferred to linear maps by choosing a basis of the underlying vector space.

[3]This is an informal way of saying that P is the spectral projection associated to the set $\{\lambda\}$.

(c) *The following are equivalent:*

(i) *The eigenvalue λ is a first order pole of the resolvent $\mathcal{R}(\cdot, A)$.*

(ii) *The eigenvalue λ is a first order pole of the dual resolvent $\mathcal{R}(\cdot, A^{\mathrm{T}})$.*

(iii) *The limit $\lim_{\mu\to\lambda}(\mu-\lambda)\mathcal{R}(\mu, A)$ exists.*

(iv) *The eigenvalue λ is semisimple.*

(v) *The range of P consists of eigenvectors, i.e. $\operatorname{rg} P = \ker(\lambda - A)$.*

If the equivalent conditions (i)–(v) are satisfied, then $\lim_{\mu\to\lambda}(\mu-\lambda)\mathcal{R}(\mu, A) = P$.

(d) *If $\ker(\lambda - A) = \operatorname{span}\{u\}$ and $v \in \ker(\lambda - A^{\mathrm{T}})$ satisfy $v^{\mathrm{T}} u = 1$, then $P = uv^{\mathrm{T}}$.*

Proof. (a) This follows from the integral representation of P in Proposition 2.1.4 and from the integration formula in Proposition 2.1.2(a).

(b) This follows from how P was constructed in the existence argument in the proof of Proposition 2.1.4.

(c) If λ is a first order pole, then $\lim_{\mu\to\lambda}(\mu-\lambda)\mathcal{R}(\mu, A) = Q_0$ by the Laurent series expansion of $\mathcal{R}(\cdot, A)$ about λ (Proposition 2.1.1(a)). Moreover, $Q_0 = P$ according to (a). Let us now prove the claimed equivalence.

"(i) $\Leftrightarrow$ (ii)": One has $\mathcal{R}(\mu, A^{\mathrm{T}}) = \mathcal{R}(\mu, A)^{\mathrm{T}}$ for all $\mu \in \rho(A^{\mathrm{T}}) = \rho(A)$, so one can see the claimed equivalence by taking the transposition operation out of the Laurent series expansion in Proposition 2.1.1(a).

"(i) $\Leftrightarrow$ (iii)": For a scalar-valued holomorphic function f that has an isolated singularity at λ it is a standard result from complex analysis that λ is a first order pole of f if and only if $\lim_{\mu\to\lambda}(\mu-\lambda)f(\mu)$ exists. The claim now follows from applying this fact to all components of the resolvent.

"(i) $\Leftrightarrow$ (iv)": Since λ is semisimple if and only if every Jordan block corresponding to λ has size 1, which is equivalent to λ having pole order one by Proposition 2.1.1.

"(i) $\Rightarrow$ (v)": Note that $\ker(\lambda - A) \subseteq \operatorname{rg} P = \operatorname{rg} Q_0$ by (b) and (a). If the pole order at λ is 1, the converse inclusion $\operatorname{rg} Q_0 \subseteq \ker(\lambda - A)$ follows from Proposition 2.1.1(b).

"(v) $\Rightarrow$ (iv)": The geometric multiplicity of λ is $\dim\ker(\lambda - A)$ and the algebraic multiplicity of λ is $\dim\operatorname{rg} P$, according to (b). It follows from (v) that the two are equal.

(d) The assumptions ensure that λ is even algebraically simple (hence, semisimple) due to Lemma 1.2.8. Now (c) implies $\operatorname{rg} P$ is spanned by u. By applying the same argument to A^{T} – whose spectral projection for the eigenvalue λ is P^{T} due to formula (2.1.4) – we also see that $\operatorname{rg} P^{\mathrm{T}}$ is spanned by v. Thus, $P = \alpha uv^{\mathrm{T}}$ for a scalar α. As P is a projection and $v^{\mathrm{T}} u = 1$, it follows that $\alpha = 1$. □

Theorem 2.1.7 (Spectral mapping theorem for the matrix exponential function)**.** *Let $A \in \mathbb{C}^{n\times n}$ and let $t \in \mathbb{R}$. Then one has $\sigma(e^{tA}) = e^{t\sigma(A)}$. More precisely:*

(a) *If $\lambda \in \mathbb{C}$ is an eigenvalue of A with eigenvector $z \in \mathbb{C}^n$, then $e^{t\lambda}$ is an eigenvalue of e^{tA} with eigenvector z.*

(b) *If $\mu \in \sigma(e^{tA})$, then there exists $\lambda \in \sigma(A)$ such that $\mu = e^{t\lambda}$.*

Proof. (a) It follows from $Az = \lambda z$ that $(tA)^j z = (t\lambda)^j z$ for all integer $j \geq 0$, hence

$$e^{tA} z = \sum_{j=0}^{\infty} \frac{(tA)^j}{j!} z = \sum_{j=0}^{\infty} \frac{(t\lambda)^j}{j!} z = e^{t\lambda} z.$$

(b) After a coordinate transformation, we may assume A is in Jordan normal form (2.1.1). Then by Exercise 1.4(b), e^{tA} is an upper triangular matrix whose eigenvalues are $e^{t\lambda_1}, \dots, e^{t\lambda_m}$, where $\lambda_1, \dots, \lambda_m$ are the eigenvalues of A (counted with their geometric multiplicity). Hence $\mu = e^{t\lambda}$ for some $\lambda \in \sigma(A)$. □

2.2 Eventually positive matrix semigroups

Our main objects of study for the rest of Chapter 2 are matrix semigroups $(e^{tA})_{t\geq 0}$ which are positive for all sufficiently large times.

Definition 2.2.1 (Eventually (strongly) positive matrix semigroups)**.** Let $A \in \mathbb{R}^{n\times n}$.

(a) The matrix semigroup $(e^{tA})_{t\geq 0}$ is called **eventually positive** if there exists $t_0 \geq 0$ such that $e^{tA} \geq 0$ for all $t \in [t_0, \infty)$.

(b) The matrix semigroup $(e^{tA})_{t\geq 0}$ is called **eventually strongly positive** if there exists $t_0 \geq 0$ such that $e^{tA} x \succeq \mathbb{1}$ for all $0 \neq x \in \mathbb{R}^n_+$ and all $t \in [t_0, \infty)$.

This definition uses Notation 1.2.4 again, i.e. for a given t, the inequality $e^{tA} x \succeq \mathbb{1}$ means that there exists a number $c > 0$ such that $e^{tA} x \geq c\,\mathbb{1}$. Observe that c can a priori depend on t.

Examples 2.2.2.

(a) The matrix

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

is nilpotent and satisfies $A^k = 0$ for all $k \geq 3$. Moreover,

$$A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \text{so} \qquad e^{tA} = \begin{pmatrix} 1 & t & \frac{t^2}{2} - t \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$$

for all $t \geq 0$. Thus, $(e^{tA})_{t\geq 0}$ is eventually positive but not eventually strongly positive. Since $\frac{t^2}{2} - t < 0$ for $t \in (0,2)$, the semigroup is not positive. Alternatively, this follows from the characterisation of positive semigroups in Theorem 1.3.8 since the off-diagonal entry $A_{13} = -1$ is strictly negative.

(b) Consider the matrix

$$A := U\begin{pmatrix}0 & & \\ & -1 & \\ & & -9\end{pmatrix}U^{\mathrm{T}} = \begin{pmatrix}-2 & -1 & 3\\ -1 & -2 & 3\\ 3 & 3 & -6\end{pmatrix},$$

where $U := (u_1, u_2, u_3) \in \mathbb{R}^{3\times 3}$ is the orthogonal matrix with the columns

$$u_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1 & 1 & 1\end{pmatrix}^{\mathrm{T}}, \quad u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1 & 0\end{pmatrix}^{\mathrm{T}}, \quad u_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}1 & 1 & -2\end{pmatrix}^{\mathrm{T}}.$$

Let $P \in \mathbb{R}^{3\times 3}$ be the projection onto the first component of $\mathbb{R}^3$. Then $e^{tA} \to UPU^{\mathrm{T}} = u_1 u_1^{\mathrm{T}}$ as $t \to \infty$. Since every entry of $u_1 u_1^{\mathrm{T}}$ is $\frac{1}{3}$, it follows that $(e^{tA})_{t\geq 0}$ is eventually strongly positive. However, A has some negative off-diagonal entries, so by the characterisation of positive semigroups in Theorem 1.3.8, $(e^{tA})_{t\geq 0}$ is not positive.

Parts (a) and (b) of the following result show that Theorem 1.3.9 about positive semigroups continues to hold in the eventually positive case. We also add a third property (c) which we will use to study perturbation theory in Theorem 2.4.2.

Theorem 2.2.3 (Perron–Frobenius for eventually positive matrix semigroups)**.** *Let $A \in \mathbb{R}^{n\times n}$ be such that $(e^{tA})_{t\geq 0}$ is eventually positive. The following assertions hold:*

(a) *The spectral bound* $\mathrm{s}(A)$ *is an eigenvalue of A with an eigenvector $x \geq 0$.*

(b) $\mathrm{s}(A)$ *is a strictly dominant eigenvalue of A, i.e.* $\operatorname{Re}\lambda < \mathrm{s}(A)$ *for all* $\lambda \in \sigma(A)\setminus\{\mathrm{s}(A)\}$.

(c) *If* $\mathrm{s}(A)$ *is semisimple, then its associated spectral projection is positive.*

Proof. We may replace A with $A - \mathrm{s}(A)$. This does not affect the eventual positivity of the semigroup since $e^{t(A-\mathrm{s}(A))} = e^{-t\mathrm{s}(A)}e^{tA}$ for all $t \in [0,\infty)$ by Proposition 1.3.2(d), and it allows us to assume $\mathrm{s}(A) = 0$. Let $t_0 \in [0,\infty)$ be such that $e^{tA} \geq 0$ for all $t \geq t_0$.

(a) and (b) We first prove that 0 is an eigenvalue of A and that there are no non-zero eigenvalues of A on the imaginary line. Since $\mathrm{s}(A) = 0$, it follows from the spectral mapping theorem for the matrix exponential function (Theorem 2.1.7) that e^{tA} has spectral radius 1 for each $t \geq 0$. Let $i\beta \in i\mathbb{R}$ be an eigenvalue of A. Again by Theorem 2.1.7, $e^{i\beta t}$ is an eigenvalue of e^{tA} for each $t \in [0,\infty)$. We need to show that $\beta = 0$.

For each index $j \in \{1,\dots,n\}$ one has $(e^{0A})_{jj} = 1$, so by the uniqueness theorem for analytic functions, the set $\bigcup_{j=1}^{n}\left\{t \in [0,\infty) : (e^{tA})_{jj} = 0\right\}$ does not accumulate in $[0,\infty)$. Thus, there exist times $t_2 > t_1 \geq t_0$ such that for each $t \in [t_1, t_2]$, the diagonal entries of e^{tA} are strictly positive. According to Theorem 1.2.5(b) this implies, for each such t, that the spectral radius 1 is a radially strictly dominant eigenvalue of e^{tA}, i.e. the matrix e^{tA} does not have eigenvalues on the unit circle except for the number 1. Thus, $e^{i\beta t} = 1$ for all $t \in [t_1, t_2]$, which implies that $\beta = 0$, as claimed.

Now we show the existence of a positive eigenvector of A for the eigenvalue 0. With the notation of the Laurent series expansion of $(\mu - A)^{-1}$ about the eigenvalue 0 in

Proposition 2.1.1(a), one gets $Q_{-p+1} = \lim_{\mu\to 0}\mu^p(\mu - A)^{-1}$. Using the Laplace transform representation of the resolvent (Lemma 1.3.7) yields

$$Q_{-p+1} = \lim_{\mu\downarrow 0}\Big(\underbrace{\mu^p\int_0^{t_0} e^{-t\mu}e^{tA}\,\mathrm{d}t}_{\to 0} + \underbrace{\mu^p\int_{t_0}^{\infty} e^{-t\mu}e^{tA}\,\mathrm{d}t}_{\geq 0}\Big) \geq 0.$$

Since Q_{-p+1} is non-zero and $\mathbb{R}^n_+$ spans $\mathbb{R}^n$, we can find a vector $x \in \mathbb{R}^n_+$ such that $0 \leq Q_{-p+1}x \neq 0$. According to Proposition 2.1.1(b) that range of Q_{-p+1} is contained in $\ker A$, so $Q_{-p+1}x$ is indeed a positive eigenvector of A for the eigenvalue $\mathrm{s}(A) = 0$.

(c) Due to semisimplicity, Proposition 2.1.6(a) and (c) give $P = Q_0 = Q_{-p+1} \geq 0$. □

Example 2.2.4. In Theorem 2.2.3(c), the spectral projection can fail to be positive if $\mathrm{s}(A)$ is not a semisimple eigenvalue. Indeed, let

$$A := T\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & -2\end{pmatrix}T^{-1} = \begin{pmatrix}-1 & 1 & 0\\ 1 & -1 & 0\\ 1 & 0 & 0\end{pmatrix} \qquad \text{for} \qquad T := \begin{pmatrix}0 & 1 & 2\\ 0 & 1 & -2\\ 1 & 0 & -1\end{pmatrix}$$

Since all off-diagonal entries of A are ≥ 0, the semigroup $(e^{tA})_{t\geq 0}$ is positive (Theorem 1.3.8). The given Jordan normal form of A shows that $\mathrm{s}(A) = 0$ is not semisimple. The spectral projection of A associated to the eigenvalue 0 is

$$P = T\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}T^{-1} = \frac{1}{4}\begin{pmatrix}2 & 2 & 0\\ 2 & 2 & 0\\ 1 & -1 & 4\end{pmatrix} \not\geq 0.$$

2.3 Characterisation

By Theorem 1.3.8, positivity of $(e^{tA})_{t\geq 0}$ is equivalent to the positivity of $\mathcal{R}(\lambda, A)$ for all $\lambda \in (\mathrm{s}(A), \infty)$. Parts (i) and (ii) of the next theorem provide a related characterisation for eventual strong positivity. Parts (iii) and (iv) show that eventual strong positivity can be characterised in terms of Perron–Frobenius like spectral properties. In this sense, Perron–Frobenius theory is closer to eventual positivity than to positivity.

Theorem 2.3.1. *Let $A \in \mathbb{R}^{n\times n}$. The following are equivalent:*

(i) $(e^{tA})_{t\geq 0}$ *is eventually strongly positive.*

(ii) $\mathrm{s}(A)$ *is a strictly dominant eigenvalue of A and there exists $\lambda_0 > \mathrm{s}(A)$ such that $\mathcal{R}(\lambda, A)$ is strongly positive for all $\lambda \in (\mathrm{s}(A), \lambda_0)$.*

(iii) $\mathrm{s}(A)$ *is a strictly dominant eigenvalue of A and the associated spectral projection is strongly positive.*

(iv) $\mathrm{s}(A)$ *is a strictly dominant eigenvalue of A and the eigenspaces* $\ker(\mathrm{s}(A) - A)$ *and* $\ker(\mathrm{s}(A) - A^{\mathrm{T}})$ *are spanned by strongly positive vectors.*

If these equivalent assertions hold, then the eigenvalue $\mathrm{s}(A)$ *is even algebraically simple.*

Proof. As usual, we may assume $\mathrm{s}(A) = 0$. Note that if the spectral projection P associated to 0 is strongly positive, then the eigenvalue 0 of A is algebraically simple (and in turn, semisimple): indeed, if P is strongly positive, then Perron–Frobenius (Theorem 1.2.5(c)) guarantees that $\mathrm{r}(P) > 0$ and $\ker(\mathrm{r}(P) - P)$ is spanned by a strongly positive vector. On the other hand, $\mathrm{r}(P) = 1$, as P is a non-zero projection. Consequently, $\operatorname{rg} P = \ker(1 - P)$ is one-dimensional. As $\dim \operatorname{rg} P$ is the algebraic multiplicity of the eigenvalue $\mathrm{s}(A) = 0$ (Proposition 2.1.6(b)), it follows that 0 is algebraically simple.

"(i) ⇒ (iv)": By Theorem 2.2.3 (Perron–Frobenius for eventually positive matrix semigroups) the spectral bound 0 is a strictly dominant eigenvalue of A. Choose $t_0 > 0$ such that the matrix $e^{t_0 A}$ is strongly positive. Due to the spectral mapping theorem for the matrix exponential function (Theorem 2.1.7(a)), one has $\{0\} \neq \ker A \subseteq \ker(1 - e^{t_0 A})$, and the latter space is spanned by a strongly positive vector according to the Perron–Frobenius theorem for strongly positive matrices (Theorem 1.2.5(c)). Thus, $\ker A = \ker(1 - e^{t_0 A})$, which proves the claim for $\ker A$. The same argument applies to $\ker(A^{\mathrm{T}})$, since $e^{t_0 A^{\mathrm{T}}} = (e^{t_0 A})^{\mathrm{T}}$ is also strongly positive.

"(iv) ⇒ (iii)": By assumption, $\ker A$ and $\ker A^T$ are spanned by strongly positive vectors u, v respectively. Replacing u by a scalar multiple, we may assume that $v^{\mathrm{T}} u = 1$. Proposition 2.1.6(d) now yields $P = uv^{\mathrm{T}}$ is strongly positive.

"(iii) ⇒ (i)": We have seen that the strong positivity of P implies that 0 is semisimple. This ensures $\operatorname{rg} P = \ker A$ due to Proposition 2.1.6(c). The spectral mapping theorem (Theorem 2.1.7) thus implies that e^{tA} acts as the identity matrix on $\operatorname{rg} P$.

Also, since 0 is a strictly dominant eigenvalue of A, all eigenvalues of $A|_{\ker P}$ have strictly negative real part by Proposition 2.1.4(b). Therefore, $e^{tA}|_{\ker P} \to 0$ as $t \to \infty$ according to Proposition 1.3.6. Consequently, $e^{tA} = e^{tA}P + e^{tA}(1 - P) \to P$ as $t \to \infty$. The strong positivity of P hence implies the eventual strong positivity of $(e^{tA})_{t \geq 0}$.

"(ii) ⇒ (iv)": Let $\lambda \in (0, \lambda_0) = (\mathrm{s}(A), \lambda_0)$. One has

$$\sigma\big(\mathcal{R}(\lambda, A)\big) = \left\{ \frac{1}{\lambda - \mu} : \mu \in \sigma(A) \right\}.$$

As $\mathrm{s}(A) = 0 \in \sigma(A)$, it follows that $\mathrm{r}(\mathcal{R}(\lambda, A)) = \frac{1}{\lambda}$ is an eigenvalue of $\mathcal{R}(\lambda, A)$.

Observe that $\ker A = \ker\big(\frac{1}{\lambda} - \mathcal{R}(\lambda, A)\big)$ and the latter space is spanned by a strongly positive vector according to the Perron–Frobenius theorem for strongly positive matrices (Theorem 1.2.5(c)). The same argument can be applied to A^{T}.

"(iii) ⇒ (ii)": Since P is strongly positive, 0 is semisimple as already shown. Proposition 2.1.6(c) thus gives $\lim_{\mu \to 0} \mu \mathcal{R}(\mu, A) = P$. As P is strongly positive, this implies that $\mathcal{R}(\mu, A)$ is also strongly positive for all $\mu > 0$ that are sufficiently close to 0. □

2.4 Perturbations

We conclude this chapter with a sneak peek of the perturbation theory for eventually positive semigroups. By **perturbations** – more precisely, additive perturbations – we mean the following: given two matrices $A, B \in \mathbb{C}^{n\times n}$, we study which properties of the semigroup $(e^{tA})_{t\geq 0}$ are inherited by the **perturbed** semigroup $\left(e^{t(A+B)}\right)_{t\geq 0}$ if B has sufficiently nice properties. In other words, B is viewed as a perturbation of A , and our goal is to determine which semigroup properties are preserved under such perturbations.

A simple instance of such a perturbation result is the fact that positive perturbations do not destroy the positivity of a semigroup. This is a particular case of the following.

Proposition 2.4.1. *Let $A, B \in \mathbb{R}^{n\times n}$ and assume that $(e^{tA})_{t\geq 0}$ is positive. If all off-diagonal entries of B are ≥ 0, then the perturbed semigroup $\left(e^{t(A+B)}\right)_{t\geq 0}$ is also positive.*

Proof. By Theorem 1.3.8, the semigroup generated by a matrix $C \in \mathbb{R}^{n\times n}$ is positive if and only if all off-diagonal entries of C are ≥ 0. The assertion is now immediate. ☐

It is natural to ask whether a similar perturbation result holds for eventually positive semigroups: if $(e^{tA})_{t\geq 0}$ is eventually positive and $B \in \mathbb{R}^{n\times n}$ is a positive matrix, does it follow that the perturbed semigroup $\left(e^{t(A+B)}\right)_{t\geq 0}$ is also eventually positive? The answer to this question is quite surprising (and perhaps disappointing): unless the unperturbed semigroup is already positive, there always exists a positive perturbation that destroys the eventual positivity. We prove this in the following theorem.

Theorem 2.4.2. *Let $A \in \mathbb{R}^{n\times n}$. The following are equivalent.*

(i) *For every $B \in \mathbb{R}_+^{n\times n}$, the semigroup $\left(e^{t(A+B)}\right)_{t\geq 0}$ is eventually positive.*

(ii) *For every $B \in \mathbb{R}_+^{n\times n}$ of rank ≤ 1, the semigroup $\left(e^{t(A+B)}\right)_{t\geq 0}$ is eventually positive.*

(iii) *The semigroup $(e^{tA})_{t\geq 0}$ is positive.*

Proof. "(iii) ⇒ (i)": In this case, $\left(e^{t(A+B)}\right)_{t\geq 0}$ is even positive (Proposition 2.4.1).

"(i) ⇒ (ii)": This is trivial.

"(ii) ⇒ (iii)": This is the surprising part. The key ingredient is the Sherman–Morrison–Woodbury formula for rank-one perturbations of resolvents, presented in Exercise 2.3.

As before, we assume $\operatorname{s}(A) = 0$ without loss of generality. According to Theorem 1.3.8, it suffices to prove that $\mathcal{R}(\mu, A) \geq 0$ for all $\mu > 0$. To achieve this, in fact it suffices to prove that $v^{\mathrm{T}}\mathcal{R}(\mu, A) \geq 0$ for all $v \geq \mathbb{1}$ and all $\mu > 0$, since strongly positive vectors are dense in $\mathbb{R}_+^n$. Thus let us fix such a vector $v \geq \mathbb{1}$ and a number $\mu > 0$.

Firstly, observe that assumption (ii) with $B = 0$ implies that $(e^{tA})_{t\geq 0}$ is eventually positive. By Theorem 2.2.3, we deduce that $\operatorname{s}(A) = 0$ is an eigenvalue of A with an eigenvector $u \geq 0$. Thus $v^{\mathrm{T}}u > 0$ and there exists $\alpha > 0$ such that $\alpha v^{\mathrm{T}}u = \mu$.

Consider the rank-one matrix $B := \alpha u v^{\mathrm{T}}$. By Exercise 2.3(b), we have $\mathrm{s}(A+\alpha u v^{\mathrm{T}}) = \mu$, it is a semisimple eigenvalue of $A+B$, and formula (2.4.2) (with $\lambda_0 = 0$) yields

$$(\lambda-\mu)\mathcal{R}(\lambda, A+\alpha u v^{\mathrm{T}}) = (\lambda-\mu)\mathcal{R}(\lambda, A) + \alpha u v^{\mathrm{T}}\mathcal{R}(\lambda, A)$$

for all $\lambda > \mu$. Due to semisimplicity, Proposition 2.1.6(c) ensures that the spectral projection corresponding to the eigenvalue μ of $A+B$ is given by

$$\lim_{\lambda\downarrow\mu}(\lambda-\mu)\mathcal{R}(\lambda, A+\alpha u v^{\mathrm{T}}) = \alpha u v^{\mathrm{T}}\mathcal{R}(\mu, A).$$

By hypothesis, $A+B$ generates an eventually positive semigroup, so this projection is positive by Theorem 2.2.3(c). As $u \geq 0$ and non-zero, it follows that $v^{\mathrm{T}}\mathcal{R}(\mu, A) \geq 0$. □

Theorem 2.4.2 is not quite the end of the story. The notion "perturbation" already suggests that one is often interested in perturbations that are small in some sense. Furthermore, the above theorem leaves open the possibility that perhaps a more positive result (pun intended) holds for eventual *strong* positivity. As it turns out, one can show that eventual strong positivity of a semigroup $(e^{tA})_{t\geq 0}$ is preserved by all perturbations $B \geq 0$ that are sufficiently small in operator norm. In fact, such a result holds even in infinite dimensions, as we will see in Chapter 13, where perturbation theory for eventually positive semigroups is developed in greater depth.

Exercises for Chapter 2

Exercise 2.1. Consider the matrix

$$A := T \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix} T^{-1} \in \mathbb{R}^{4\times 4}$$

for an invertible matrix $T \in \mathbb{R}^{4\times 4}$.

(a) Compute the Laurent series expansion of $\mathcal{R}(\cdot, A)$ about the spectral value 1 and the associated spectral projection. What is the pole order of $\mathcal{R}(\cdot, A)$ at 1?

(b) Compute the Laurent series expansion of $\mathcal{R}(\cdot, A)$ about the spectral value 0 and the associated spectral projection. What is the pole order of $\mathcal{R}(\cdot, A)$ at 0?

Does one have equality of the subspaces in Proposition 2.1.1(b)?

(c) Find a T such that $(e^{tA})_{t\geq 0}$ is eventually strongly positive.

Exercise 2.2. Prove the following assertions are equivalent for $A \in \mathbb{R}^{n\times n}$:

(i) $(e^{tA})_{t\geq 0}$ is eventually positive.

(ii) For every $0 \leq x \in \mathbb{R}^n$, there exists $t_0 = t_0(x) \geq 0$ such that $e^{tA}x \geq 0$ for all $t \geq t_0$.

(iii) For every $0 \leq x \in \mathbb{R}^n$ and $0 \leq y \in \mathbb{R}^n$, there exists $t_0 = t_0(x, y) \geq 0$ such that $y^{\mathrm{T}} e^{tA} x \geq 0$ for all $t \geq t_0$.

Exercise 2.3 (Sherman–Morrison–Woodbury formula)**.** Let $A \in \mathbb{C}^{n\times n}$ and let $u, v \in \mathbb{C}^n$.

(a) If A is invertible, prove that $A - uv^{\mathrm{T}}$ is invertible if and only if $v^{\mathrm{T}} A^{-1} u \neq 1$, and in this case it holds that

$$(A - uv^{\mathrm{T}})^{-1} = A^{-1} + \frac{1}{1 - v^{\mathrm{T}} A^{-1} u} A^{-1} u v^{\mathrm{T}} A^{-1}. \tag{2.4.1}$$

(b) Let $\lambda \in \rho(A)$ and let u be an eigenvector corresponding to an eigenvalue $\lambda_0 \in \mathbb{C}$ of A. Deduce that $\lambda \in \rho(A + uv^{\mathrm{T}})$ if and only if $\lambda \neq \lambda_0 + v^{\mathrm{T}} u$, and in this case

$$\mathcal{R}(\lambda, A + uv^{\mathrm{T}}) = \mathcal{R}(\lambda, A) + \frac{1}{\lambda - (\lambda_0 + v^{\mathrm{T}} u)} uv^{\mathrm{T}} \mathcal{R}(\lambda, A). \tag{2.4.2}$$

Also show that in the other case, i.e. $\lambda = \lambda_0 + v^{\mathrm{T}} u$, the number λ is a semisimple eigenvalue of $A + uv^{\mathrm{T}}$.

Exercise 2.4 (Another characterisation of eventual strong positivity)**.** Let $A \in \mathbb{R}^{n\times n}$.

(a) Assume that there exists $c \in \mathbb{R}$ and $k_0 \in \mathbb{N}$ such that $(A + cI)^k$ is strongly positive for all integers $k \geq k_0$. Show that $(e^{tA})_{t\geq 0}$ is eventually strongly positive.

(b) Suppose $B \in \mathbb{C}^{n\times n}$ is a matrix such that $\mathrm{r}(B) > 0$ is a semisimple and radially strictly dominant eigenvalue (see Theorem 1.2.5(b)). Prove that $\left(\frac{B}{\mathrm{r}(B)}\right)^k$ converges to the spectral projection associated with $\mathrm{r}(B)$ as $k \to \infty$. [*Hint:* Jordan normal form.]

(c) Assume that $(e^{tA})_{t\geq 0}$ is eventually strongly positive. Use part (b) to deduce that there exists $k_0 \in \mathbb{N}$ and $c \in \mathbb{R}$ such that $(A + cI)^k$ is strongly positive for all $k \geq k_0$.

Exercise 2.5 (Eventual (?) positivity in two dimensions)**.** Let $A \in \mathbb{R}^{2\times 2}$. Show that if $(e^{tA})_{t\geq 0}$ is eventually (strongly) positive, then $(e^{tA})_{t\geq 0}$ is (strongly) positive.

Hint: as a first step, think about what $\sigma(A)$ could look like.

Notes for Chapter 2

Spectral projections

For the historical development of the functional calculus for linear operators – which contains spectral projections as a special case – we refer to Section 5.2.1 in Pietsch's monograph [Pie07] about the history of Banach spaces and linear operators. A very accessible presentation of spectral projections of matrices and, more generally, of eigenvalue theory via complex analysis techniques can be found, for instance, in [CD13] (an updated version with minor corrections is available on Daniel Daners' webpage).

Eventual positivity in finite dimensions

Matrices with eventually positive powers

The predecessors of the eventual positivity theory in finite dimensions can be found in various results about matrices A whose powers A^k are positive for some, or all sufficiently large, $k \in \mathbb{N}$. Matrices with a positive power were, for instance, studied in [Bra61], and more recently in [TCDF15]. Matrices for which a polynomial $p(A)$ is positive are studied in [Sen06]. Some early papers on matrices with eventually positive powers, such as [Fri78, ZT99], were motivated by inverse spectral problems, i.e. the question of which sets in $\mathbb{C}$ can be realised as the spectrum of matrices with certain prescribed properties.

In the 21st century, the literature on eventually positive matrices has grown quickly. Of particular interest were spectral properties of such matrices in the spirit of the Perron–Frobenius theorem, e.g. in the papers [TRH01, JT04, Nou06, ES08, ES09]. In particular, in [Nou06, Theorem 2.2] one can find a discrete-time analogue of the equivalence of (i) and (iv) in Theorem 2.3.1. Matrices with eventually positive powers did not occur in the lecture notes, but they appear in Exercise 2.4. For further references about matrices with eventually positive powers we refer to [Glü16, Section 6.4]; most of the preceding two paragraphs is also taken from this reference.

Eventually positive matrix semigroups

As for the continuous-time case, eventual positivity of matrix semigroups was studied by Noutsos and Tsatsomeros in [NT08]. The equivalence of (i) and (iv) in Theorem 2.3.1 as well as the characterisation of eventually strongly positive matrix semigroups in Exercise 2.4 appeared in [NT08, Theorem 3.3]; however, the approach outlined in the exercise follows [DGK16b, Theorem 6.1]. The fact that eventual positivity implies positivity for 2×2 matrix semigroups (Exercise 2.5) was observed in [DGK16b, Proposition 6.2].

Perturbation theory

Perturbation theory for matrices with eventually positive powers was studied in [SA17]. In particular, [SA17, Proposition 3.6 and Theorem 3.7] contain discrete-time analogues of the equivalence of (i) and (iii) in Theorem 2.4.2. For the continuous-time case, perturbation theory was first studied in [DG18], though with a focus on the infinite-dimensional case that we consider later. A specific finite-dimensional result in this article is [DG18, Proposition 4.6], which shows that the set of all $A \in \mathbb{R}^{n\times n}$ for which $(e^{tA})_{t\geq 0}$ is eventually strongly positive is open in $\mathbb{R}^{n\times n}$. Theorem 2.4.2 is a finite-dimensional version of [DG18, Theorem 2.3]; we were able to slightly weaken the assumptions in this theorem. Example 2.2.2(b) is also taken from [DG18, Example 2.1]. There, an explicit rank-one operator B is given with the property that $A + sB$ generates an eventually positive semigroup for $s \in [0,4)$, but the eventual positivity is lost for $s > 4$.

Eventual positivity with respect to other cones

Positivity of matrices and matrix semigroups has often been studied with respect to general cones. Naturally, this can also be done for eventual positivity; see, for instance, [Kas17, KT17, Soo19]. The following phenomenon in this context is remarkable: Exercise 2.2 suggests that eventual positivity for a dynamical system can be defined 'individually' (i.e. for individual orbits $x \mapsto e^{tA}x$) or 'uniformly' on the level of operators (as in Definition 2.2.1). However, the result of that exercise states that these two notions are equivalent for matrix semigroups. This is a product of two separate features: finite dimensionality, and geometric properties of the positive cone $\mathbb{R}^n_+$. As we see later, the equivalence of individual and uniform eventual positivity fails in infinite dimensions. More surprisingly, it also fails in finite dimensions if we consider positivity with respect to other cones, as was shown in [GH23, Example 3.1].

Chapter 3

Unbounded operators and their spectra

In Chapters 1 and 2 we studied (eventual) positivity properties for matrix semigroups $(e^{tA})_{t\geq 0}$ and resolvents $\mathcal{R}(\lambda, A) = (\lambda - A)^{-1}$ of matrices. These objects can be used to solve different types of equations in $\mathbb{R}^n$. For $x_0 \in \mathbb{C}^n$, the mapping $t \mapsto e^{tA}x_0$ solves a linear differential equation (Corollary 1.3.3) and the vector $u := \mathcal{R}(\lambda, A)x_0$ solves the linear equation $(\lambda - A)u = x_0$ in $\mathbb{C}^n$.

From now on and throughout the rest of the course, we study the analogous equations in infinite dimensions. The equations of interest are typically partial differential equations, so the matrix A will be replaced by a differential operator on a Banach space. As we will see, these are typically unbounded operators, and hence it is the purpose of the current chapter to give an introduction to the theory of unbounded operators.

3.1 Unbounded operators

Differential operators are operators that map every function f from a suitable function space to a new function that involves the (partial) derivatives of f. Such operators cannot be defined everywhere on classical function spaces such as $\mathrm{C}([0,1])$, because not every continuous function has a derivative. This motivates the definition of linear operators that are only defined on a vector subspace of a given Banach space.

Definition 3.1.1 (Linear operators)**.** Let X, Y be Banach spaces over the same scalar field.

(a) A **linear operator**, or briefly, an **operator**, between X and Y is a linear mapping $A\colon \operatorname{dom}(A) \to Y$, where $\operatorname{dom}(A)$ is a vector subspace of X. We briefly write $A\colon X \supseteq \operatorname{dom}(A) \to Y$ for such an operator. If $X = Y$, we say that A is an operator *on* X. The space $\operatorname{dom}(A)$ is called the **domain of** A.

Now, let $A\colon X \supseteq \operatorname{dom}(A) \to Y$ be a linear operator.

(b) The operator A is said to be **everywhere defined** if $\operatorname{dom}(A) = X$. It is said to be **densely defined** or to have a **dense domain** if $\operatorname{dom}(A)$ is dense in X.

(c) The operator A is called **closed** if its graph $\{(x, Ax) : x \in \operatorname{dom}(A)\}$ is closed in $X \times Y$.

(d) A norm on $\operatorname{dom}(A)$ is called a **graph norm** of A if it is equivalent to the norm

$$\|\cdot\|_A : \operatorname{dom}(A) \to [0,\infty), \qquad x \mapsto \|x\|_X + \|Ax\|_Y .$$

When the domain $\operatorname{dom}(A)$ of an operator is endowed with a graph norm, then the inclusion map $\operatorname{dom}(A) \hookrightarrow X$ is obviously continuous. Note that a linear operator $A\colon X \supseteq \operatorname{dom}(A) \to Y$ is, in general, not a continuous map from $\operatorname{dom}(A)$ to Y if $\operatorname{dom}(A)$ is endowed with the norm induced by X; hence, one often refers to such operators as **unbounded operators**. However, A is clearly continuous when $\operatorname{dom}(A)$ is endowed with a graph norm of A. If one wants to apply the theory of bounded linear operators between Banach spaces to A, ideally $\operatorname{dom}(A)$ would be a Banach space with respect to some (hence, every) graph norm of A. We now prove that this is the case if and only if A is closed.

Proposition 3.1.2 (Characterisation of closedness)**.** *Let X, Y be Banach spaces and let $A\colon X \supseteq \operatorname{dom}(A) \to Y$ be a linear operator. The following are equivalent:*

(i) *The operator A is closed.*

(ii) *If a sequence (x_k) in $\operatorname{dom}(A)$ converges (with respect to the X-norm) to a point $x \in X$ and (Ax_k) converges to a point $y \in Y$, then $x \in \operatorname{dom}(A)$ and $Ax = y$.*

(iii) *The domain $\operatorname{dom}(A)$ is complete (hence, a Banach space) with respect to some (equivalently, every) graph norm.*

Proof. "(i) ⇔ (ii)": This follows directly from the definition of closed operators.

"(ii) ⇒ (iii)": Let (x_k) be a Cauchy sequence in $\operatorname{dom}(A)$ with respect to $\|\cdot\|_A$. Then (x_k) and (Ax_k) are Cauchy in X and Y respectively, so there exist $x \in X$, $y \in Y$ such that $(x_k) \to x$ in X and $(Ax_k) \to y$ in Y. By (ii), $x \in \operatorname{dom}(A)$ and $Ax = y$, and thus

$$\|x_k - x\|_A = \|x_k - x\|_X + \|A(x_k - x)\|_Y = \|x_k - x\|_X + \left\|Ax_k - y\right\|_Y \to 0.$$

"(iii) ⇒ (ii)": Assume (iii) and let (x_k), x, y be as in (ii). Then (x_k) and (Ax_k) are Cauchy in X and Y respectively. Hence, (x_k) is Cauchy with respect to $\|\cdot\|_A$ and thus converges to a point $w \in \operatorname{dom}(A)$ with respect to $\|\cdot\|_A$. In particular, one also has $x_k \to w$ with respect to $\|\cdot\|_X$, so $w = x$. Hence, $x \in \operatorname{dom}(A)$.

On the other hand, the convergence of $x_k \to x$ with respect to $\|\cdot\|_A$ also implies that $Ax_k \to Ax$ and therefore $Ax = y$. □

We briefly recall one of the fundamental results in functional analysis, the closed graph theorem. In the terminology from Definition 3.1.1, it can be phrased as follows.

Theorem 3.1.3 (Closed graph theorem)**.** *Let X, Y be Banach spaces and consider a linear operator $A : X \supseteq \operatorname{dom}(A) \to Y$. If A is closed and everywhere defined, then A is continuous.*

A simple but illustrative class of closed operators that are not everywhere defined are operators that act by multiplication with an unbounded function. These are easy to work with and thus useful to get a first intuition for many concepts in operator theory. You will investigate this further in Exercise 3.2. Our actual objects of interest, though, are differential operators. Let us start with two simple one-dimensional examples.

Examples 3.1.4 (Differential operators on an interval)**.**

(a) Let $A\colon \mathrm{C}([-1,1]) \supseteq \operatorname{dom}(A) := \mathrm{C}^1([-1,1]) \to \mathrm{C}([-1,1])$ be given by $Af := f'$ for all $f \in \operatorname{dom}(A)$. Then A is densely defined and closed.

(b) Let $p \in [1,\infty)$ and let $A\colon L^p(-1,1) \supseteq \operatorname{dom}(A) := \mathrm{C}^1([-1,1]) \to L^p(-1,1)$ be given by $Af := f'$ for all $f \in \operatorname{dom}(A)$. Then A is densely defined, but not closed.

Proof. (a) It follows from the Weierstraß approximation theorem that $\mathrm{C}^1([-1,1])$ is dense in $\mathrm{C}([-1,1])$, so A is densely defined. To show closedness, let (f_k) be a sequence in $\mathrm{C}^1([-1,1])$ that converges uniformly to $f \in \mathrm{C}([-1,1])$ and assume that the derivatives f_k' converge uniformly to some $g \in \mathrm{C}([-1,1])$. For every $x \in [-1,1]$ it follows that

$$f(x) - f(0) = \lim_{k\to\infty} \big(f_k(x) - f_k(0)\big) = \lim_{k\to\infty} \int_0^x f_k'(y)\,\mathrm{d}y = \int_0^x g(y)\,\mathrm{d}y.$$

Thus, $f \in \mathrm{C}^1([-1,1]) = \operatorname{dom}(A)$ and $Af = f' = g$, so A is closed.

(b) As in (a) the Weierstraß approximation theorem shows that $\mathrm{C}^1([-1,1])$ is dense in $\mathrm{C}([-1,1])$ with respect to $\|\cdot\|_\infty$ and thus also with respect to $\|\cdot\|_p$. As $\mathrm{C}([-1,1])$ is dense in $L^p(-1,1)$ it follows that the same is true for $\mathrm{C}^1([-1,1]) = \operatorname{dom}(A)$.

To see that A is not closed, let $f \in L^p(-1,1)$ denote the modulus function, i.e. $f(x) = |x|$ for all $x \in [-1,1]$. The sequence (f_k) in $\operatorname{dom}(A)$ given by $f_k(x) = \big(x^2 + \frac{1}{k}\big)^{1/2}$ converges uniformly to f, and thus, in particular, with respect to $\|\cdot\|_p$. Moreover,

$$(Af_k)(x) = f_k'(x) = x\left(x^2 + \frac{1}{k}\right)^{-1/2} \quad \text{with} \quad \big|Af_k(x)\big| \le 1$$

for all $x \in [-1,1]$. Since (Af_k) converges pointwise almost everywhere to the signum function, the dominated convergence theorem gives that this convergence also holds in L^p. As $f \notin \operatorname{dom}(A)$, thus A is not closed. □

We wrap up this introductory section with another crucial tool in operator theory.

Definition 3.1.5 (Dual operators)**.** Let X, Y be Banach spaces and let $A\colon \operatorname{dom}(A) \subseteq X \to Y$ be densely defined. The **dual operator** $A'\colon Y' \supseteq \operatorname{dom}\big(A'\big) \to X'$ is defined by

$$\begin{aligned} \operatorname{dom}\big(A'\big) &:= \big\{y' \in Y' \mid \exists\, x' \in X'\colon \langle y', Ax\rangle = \langle x', x\rangle \ \forall\, x \in \operatorname{dom}(A)\big\} \\ A'y' &:= x'; \end{aligned}$$

where x' in the second line is the vector that occurs in the definition of $\operatorname{dom}\big(A'\big)$.[1]

We will see more of dual operators – and also their relation to adjoint operators on Hilbert spaces – in Exercise 3.5 and from Chapter 6 onwards.

[1] Note that x' is unique by density of $\operatorname{dom}(A)$ in X.

3.2 Weak derivatives and Sobolev spaces

Example 3.1.4(b) illustrates that differential operators on L^p are not closed when their domain is a space of continuously differentiable functions. The proof showed that, for instance, the modulus function causes such problems: it is not differentiable, but this non-differentiability cannot be "seen" from within an L^p space. To obtain closed differential operators on L^p, one needs a weaker concept of differentiability. This is the topic of the present section. Let us first recall the standard multi-index notation to denote higher order partial derivatives.

Notation 3.2.1. Let $n \in \mathbb{N}$ and let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open. A vector $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ is called a **multi-index**, and its **order** is defined as $|\alpha| := \sum_{j=1}^n \alpha_j$. We write

$$\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$$

for every function $f : \Omega \to \mathbb{C}$ that is continuously differentiable of order $|\alpha|$. With this notation, we have $\partial^{e_j} f = \partial_j f$, where $e_j \in \mathbb{N}_0^n$ denotes the j-th canonical unit vector.

To generalise classical derivatives to a larger class of functions, the following two function spaces are useful.

Definition 3.2.2 (Test functions and local L^p-spaces)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$.

(a) By $\mathrm{C}_\mathrm{c}^\infty(\Omega)$, we denote the space of all **test functions** on Ω, i.e. functions $f : \Omega \to \mathbb{C}$ that are differentiable of every order and vanish outside a compact subset of Ω.

(b) Let $p \in [1, \infty]$. By $L^p_{\mathrm{loc}}(\Omega)$ we denote the space of all Lebesgue measurable $f : \Omega \to \mathbb{C}$ that satisfy $f|_K \in L^p(K)$ for every compact set $\emptyset \neq K \subseteq \Omega$; here we identify functions that are equal almost everywhere on Ω.

Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open. Note that for all $p \in [1, \infty]$ one has

$$L^1_{\mathrm{loc}}(\Omega) \supseteq L^p_{\mathrm{loc}}(\Omega) \supseteq L^p(\Omega) + \mathrm{C}(\Omega).$$

In particular, $L^1_{\mathrm{loc}}(\Omega)$ is the largest of all function spaces that we consider on Ω. By integrating – i.e. "testing" – against test functions, one can determine the derivatives of a smooth function.

Proposition 3.2.3. *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open.*

(a) *For $f, g \in L^1_{\mathrm{loc}}(\Omega)$ one has $f = g$ if and only if $\int_\Omega f\varphi \,\mathrm{d}x = \int_\Omega g\varphi \,\mathrm{d}x$ for all $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$.*

(b) *Let $\alpha \in \mathbb{N}_0^n$ and $f \in \mathrm{C}^{|\alpha|}(\Omega)$. Then $\partial^\alpha f$ is the unique element of $L^1_{\mathrm{loc}}(\Omega)$ satisfying*

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

for all $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$.

Proof. (a) This result is sometimes called the fundamental lemma of the calculus of variations. Its proof relies on the fact that there exist sufficiently many test functions on Ω and on techniques from measure theory. For readers interested in the details we provide a proof in supplementary Section 3.A.

(b) Let $f \in C^{|\alpha|}(\Omega)$ and $\varphi \in C_c^\infty(\Omega)$. By extending the function $f\varphi$ by the value 0 outside of Ω, we obtain a function in $C^{|\alpha|}(\mathbb{R}^n)$ that vanishes outside a compact set. For fixed $k \in \{1, \dots, n\}$, (one-dimensional) integration by parts then shows that $\int_\Omega (\partial_k f)\, \varphi \,\mathrm{d}x = -\int_\Omega f\, \partial_k \varphi \,\mathrm{d}x$. By applying this equality α_1 times for the index $k = 1$, then α_2 times for the index $k = 2$, and so on, we obtain the required formula. The fact that $\partial^\alpha f$ is the only function in $L^1_{\mathrm{loc}}(\Omega)$ that satisfies this equality follows from (a). □

Proposition 3.2.3(b) characterises the partial derivatives of a function f by integrating f against derivatives of test functions. This shows us a path to defining generalised derivatives for a large class of functions. Since properties that rely on testing against functionals are often called **weak properties** in functional analysis, these generalised derivatives are called weak derivatives.

Definition 3.2.4 (Weak derivative)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open, $\alpha \in \mathbb{N}_0^n$ and $f \in L^1_{\mathrm{loc}}(\Omega)$.

(a) We say that f has a **weak αth derivative** if there exists a $g \in L^1_{\mathrm{loc}}(\Omega)$ that satisfies

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

for all $\varphi \in C_c^\infty(\Omega)$. In this case g – which is unique due to Proposition 3.2.3(a)) – is called the **weak αth derivative of** f and is denoted by $g =: \partial^\alpha f$.

As in the classical case, we also use the notation $\partial_j f := \partial^{e_j} f$ for weak derivatives; cf. Notation 3.2.1.

(b) We often just write $\partial^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$ as a shortcut for "f has an αth weak derivative". If $V \subseteq L^1_{\mathrm{loc}}(\Omega)$ is any subset we write $\partial^\alpha f \in V$ as a shortcut for "f has an αth weak derivative and $\partial^\alpha f \in V$."

It follows from Proposition 3.2.3(b) that every function $f \in C^{|\alpha|}(\Omega)$ has a weak αth derivative which coincides with the classical derivative $\partial^\alpha f$. Hence, the notation for weak derivatives is consistent with the notation for classical derivatives. By using weak derivatives we can now fix the issue observed in Example 3.1.4(b) that classical derivatives are not closed operators on L^p-spaces.

Example 3.2.5. Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open, $\alpha \in \mathbb{N}_0^n$ and $p \in [1, \infty]$. The differential operator $\partial^\alpha : L^p(\Omega) \supseteq \mathrm{dom}(\partial^\alpha) \to L^p(\Omega)$ is closed when its domain is chosen as large as possible:

$$\begin{aligned} \mathrm{dom}\left(\partial^\alpha\right) &:= \left\{ f \in L^p(\Omega) \mid \partial^\alpha f \in L^p(\Omega) \right\} \\ &= \left\{ f \in L^p(\Omega) \mid f \text{ has a weak } \alpha\text{th derivative and } \partial^\alpha f \in L^p(\Omega) \right\} \end{aligned}$$

(in the first line of this formula we used the shortcut introduced in Definition 3.2.4(b)). In particular, $\mathrm{dom}(\partial^\alpha)$ is a Banach space when endowed with the graph norm $\|\cdot\|_{\partial^\alpha}$ (or any other graph norm of ∂^α).

Proof. Let (f_k) be a sequence in $\operatorname{dom}(\partial^\alpha)$ that converges in p-norm to a function $f \in L^p(\Omega)$, and assume also that $\partial^\alpha f_k \to g \in L^p(\Omega)$. For every $\varphi \in C_c^\infty(\Omega)$ one has

$$\int_\Omega g\varphi \,\mathrm{d}x = \lim_{k\to\infty} \int_\Omega (\partial^\alpha f_k)\,\varphi \,\mathrm{d}x = \lim_{k\to\infty} (-1)^{|\alpha|} \int_\Omega f_k \partial^\alpha \varphi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega f \partial^\alpha \varphi \,\mathrm{d}x,$$

so f has an αth weak derivative and this derivative is $g \in L^p(\Omega)$. In other words, $f \in \operatorname{dom}(\partial^\alpha)$ and $\partial^\alpha f = g$. The completeness of $\operatorname{dom}(\partial^\alpha)$ with respect to every graph norm of ∂^α is, according to Proposition 3.1.2, a consequence of the closedness of ∂^α. □

If one requires $\partial^\alpha f \in L^p(\Omega)$ not only for one multi-index α, but for all α up to a given order, one arrives at the following class of spaces.

Definition 3.2.6 (Sobolev spaces)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open, $p \in [1,\infty]$ and $k \in \mathbb{N}_0$.

(a) The **Sobolev space** of order k with integrability index p is defined as

$$W^{k,p}(\Omega) := \left\{ f \in L^p(\Omega) \mid \partial^\alpha f \in L^p(\Omega) \text{ for all } \alpha \in \mathbb{N}_0^n \text{ of order } |\alpha| \le k \right\}$$

and is endowed with the norm

$$\|f\|_{W^{k,p}} := \begin{cases} \left(\sum_{|\alpha|\le k} \|\partial^\alpha f\|_p^p \right)^{1/p} & \text{if } 1 \le p < \infty, \\ \max_{|\alpha|\le k} \|\partial^\alpha f\|_\infty & \text{if } p = \infty. \end{cases}$$

(b) When $p = 2$, we write $H^k(\Omega) := W^{k,2}(\Omega)$ and $\|\cdot\|_{H^k} := \|\cdot\|_{W^{k,2}}$. We endow $H^k(\Omega) = W^{k,2}(\Omega)$ with the inner product[2,3]

$$(f \mid g)_{H^k} = \sum_{|\alpha|\le k} (\partial^\alpha f \mid \partial^\alpha g)_{L^2}.$$

(c) For $p \neq \infty$, we define[4] $W_0^{k,p}(\Omega)$ as the closure of the space $C_c^\infty(\Omega)$ of test functions in $W^{k,p}(\Omega)$. Again, when $p = 2$, we write $H_0^k(\Omega) := W_0^{k,2}(\Omega)$.

Note that one has $W^{0,p}(\Omega) = L^p(\Omega)$. After Example 3.2.5 it should not come as a surprise that the Sobolev spaces are Banach spaces.

Proposition 3.2.7. *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open, $p \in [1,\infty]$ and $k \in \mathbb{N}_0$. The Sobolev space $W^{k,p}(\Omega)$ is a Banach space. In particular, $H^k(\Omega)$ is a Hilbert space.*

Proof. For every $\alpha \in \mathbb{N}_0^n$, let $\operatorname{dom}(\partial^\alpha)$ be defined as in Example 3.2.5. Then $W^{k,p}(\Omega) = \bigcap_{|\alpha|\le k} \operatorname{dom}(\partial^\alpha)$ and $\|\cdot\|_{W^{k,p}}$ is equivalent to the norm $\|f\| := \sum_{|\alpha|\le k} \|f\|_{\partial^\alpha}$ on $W^{k,p}(\Omega)$. Since $\operatorname{dom}(\partial^\alpha)$ is complete with respect to the graph norm $\|\cdot\|_{\partial^\alpha}$ (Example 3.2.5), the claim is an immediate consequence of the Lemma 3.2.8 below. □

[2] Which is indeed an inner product, as one can easily check, and which induces the norm $\|\cdot\|_{H^k}$.

[3] Throughout the course we will follow the convention from physics that inner products on complex spaces are linear in the second argument (rather than in the first).

[4] We do not need to define the space $W_0^{k,\infty}(\Omega)$ in this course. The curious reader should note that there are (at least) two non-equivalent definitions; see e.g. [Leo09, Remark 11.15].

Lemma 3.2.8. *Let X be a Banach space and let $V_1, \dots, V_n \subseteq X$ be vector subspaces that are Banach spaces with respect to norms $\|\cdot\|_{V_1}, \dots, \|\cdot\|_{V_n}$, respectively. Assume that the inclusion map $(V_k, \|\cdot\|_{V_k}) \to (X, \|\cdot\|_X)$ is continuous for each k. Then $V := V_1 \cap \dots \cap V_n$ is a Banach space with respect to the norm $\|v\|_V := \|v\|_{V_1} + \dots + \|v\|_{V_n}$.*

The proof of the lemma is a small exercise in functional analysis, which we omit. In this section we have seen how to define some closed differential operators on L^p. Closedness of operators is important in spectral theory, as we explain in the next section.

3.3 Spectrum and resolvent

Similarly as for matrices and bounded linear operators on Banach spaces, one can also define spectral values for unbounded operators. However, one must now be careful to always take the domain of the operator into account.

Definition 3.3.1 (Spectrum and resolvent)**.** Let X be a complex Banach space and let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed linear operator.

(a) The **spectrum** of A is the set

$$\sigma(A) := \{\lambda \in \mathbb{C} : \lambda - A\colon \operatorname{dom}(A) \to X \text{ is not bijective}\},$$

where $\lambda - A := \lambda \operatorname{id} - A$ and with $\operatorname{id}\colon \operatorname{dom}(A) \to X$ denoting the inclusion map. The elements of $\sigma(A)$ are called the **spectral values** of A.

(b) The complement $\rho(A) := \mathbb{C} \setminus \sigma(A)$ of the spectrum is called the **resolvent set** of A. For every $\lambda \in \rho(A)$, the **resolvent** of A at λ is the bounded operator $\mathcal{R}(\lambda, A) : X \to \operatorname{dom}(A)$ defined as $\mathcal{R}(\lambda, A) := (\lambda - A)^{-1}$.

Closedness of A was assumed in Definition 3.3.1 for the following reason: Endow $\operatorname{dom}(A)$ with a graph norm; then $\operatorname{dom}(A)$ is a Banach space since A is closed. For $\lambda \in \rho(A)$, the bijection $\lambda - A\colon \operatorname{dom}(A) \to X$ is continuous. By the bounded inverse theorem, the resolvent operator $\mathcal{R}(\lambda, A)$ is continuous from X to $\operatorname{dom}(A)$ and thus, in particular, from X to X since the inclusion map $\operatorname{dom}(A) \to X$ is continuous.

Proposition 3.3.2 (Basic properties of the spectrum and the resolvent)**.** *Let X be a complex Banach space and let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed linear operator.*

(a) *Let $\mu \in \rho(A)$. Each $\lambda \in \mathbb{C}$ with $|\lambda - \mu| < \|\mathcal{R}(\mu, A)\|^{-1}$ satisfies $\lambda \in \rho(A)$ with*

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

In particular, $\rho(A)$ is open in $\mathbb{C}$ and one has $\|\mathcal{R}(\mu, A)\| \geq \frac{1}{\operatorname{dist}(\mu, \sigma(A))}$.

(b) *The resolvent commutes with A, i.e. $A\mathcal{R}(\mu, A)x = \mathcal{R}(\mu, A)Ax$ for each $\mu \in \rho(A)$ and $x \in \operatorname{dom}(A)$.*

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

$$\mathcal{R}(\lambda, A) - \mathcal{R}(\mu, A) = (\mu - \lambda)\mathcal{R}(\lambda, A)\mathcal{R}(\mu, A).$$

In particular, the resolvent operators commute.

Proof. (a) Let $\lambda \in \mathbb{C}$ with $|\lambda - \mu| < \|\mathcal{R}(\mu, A)\|^{-1}$. Then $\|(\mu - \lambda)\mathcal{R}(\mu, A)\| < 1$ and so the operator $\mathrm{id} - (\mu - \lambda)\mathcal{R}(\mu, A) \colon X \to X$ is invertible. Thus the identity

$$\lambda - A = \mu - A + \lambda - \mu = \left[\mathrm{id} - (\mu - \lambda)\mathcal{R}(\mu, A)\right](\mu - A)$$

implies that $\lambda \in \rho(A)$ and the claimed series expansion holds. This immediately gives that $\rho(A)$ is open. Moreover, it shows that every $\lambda \in \sigma(A)$ satisfies $|\lambda - \mu| \geq \|\mathcal{R}(\mu, A)\|^{-1}$ and hence, $\operatorname{dist}(\mu, \sigma(A)) \geq \|\mathcal{R}(\mu, A)\|^{-1}$.

(b) Let $\lambda \in \rho(A)$ and $x \in \operatorname{dom}(A)$. Clearly, $(\lambda - A)\mathcal{R}(\lambda, A)x = x = \mathcal{R}(\lambda, A)(\lambda - A)x$ and $\lambda\mathcal{R}(\lambda, A)x = \mathcal{R}(\lambda, A)\lambda x$. Adding those equalities gives the claim.

(c) The resolvent identity can be obtained immediately from the identities

$$\begin{aligned} \mathcal{R}(\lambda, A) &= \mathcal{R}(\lambda, A)\left[\mu\mathcal{R}(\mu, A) - A\mathcal{R}(\mu, A)\right] \\ \text{and} \quad \mathcal{R}(\mu, A) &= \left[\lambda\mathcal{R}(\lambda, A) - A\mathcal{R}(\lambda, A)\right]\mathcal{R}(\mu, A) \end{aligned}$$

which hold for all $\lambda, \mu \in \rho(A)$. □

Definition 3.3.3 (Spectral bound)**.** Let X be a complex Banach space and let $A : X \supseteq \operatorname{dom}(A) \to X$ be a closed operator. The **spectral bound** of A is defined as

$$\mathrm{s}(A) := \sup\left\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\right\} \in [-\infty, \infty];$$

with the convention $\sup \emptyset = -\infty$.

In general, studying the spectrum of multiplication operators is quite insightful; a simple example is presented in Exercise 3.2. Meanwhile, in the rest of the chapter, we consider the spectrum of a number of concrete differential operators.

Examples 3.3.4 (The spectrum of differential operators on an interval)**.**

(a) Consider the closed operator $A \colon \mathrm{C}([0,1]) \supseteq \operatorname{dom}(A) \to \mathrm{C}([0,1])$, $Af = f'$ that we already studied in Example 3.1.4(a).[5] One has $\sigma(A) = \mathbb{C}$ and thus, $\mathrm{s}(A) = \infty$.

(b) We now consider the operator from (a) on a smaller space: Let $\mathrm{C}_0((0,1])$ denote the space of continuous complex-valued functions on $[0,1]$ that vanish at 0 and let $A \colon \mathrm{C}_0((0,1]) \supseteq \operatorname{dom}(A) \to \mathrm{C}_0((0,1])$, $Af := f'$, where

$$\operatorname{dom}(A) := \left\{f \in \mathrm{C}^1([0,1]) \cap \mathrm{C}_0((0,1]) : f' \in \mathrm{C}_0((0,1])\right\}.$$

Then A is closed, $\sigma(A) = \emptyset$, and thus $\mathrm{s}(A) = -\infty$.

[5]We work on a different interval now, but clearly this does not affect the proof of the closedness of A.

Proof. (a) Let $\lambda \in \mathbb{C}$ and consider the function $f \in C^1([0,1])$, $f(x) = e^{\lambda x}$. Then $(\lambda - A)f = 0$, so $\lambda - A$ is not injective. Hence, $\lambda \in \sigma(A)$.

(b) The closedness of A follows from the closedness of the operator in (a). To show $\sigma(A) = \emptyset$, let $\lambda \in \mathbb{C}$ and $g \in C_0((0,1])$. A function $f : [0,1] \to \mathbb{C}$ is in $\operatorname{dom}(A)$ and solves the equation $(\lambda - A)f = g$ if and only if $f \in C^1([0,1])$ and solves the initial value problem

$$f' = \lambda f - g \qquad \text{and} \qquad f(0) = 0.$$

From the theory of linear ordinary differential equations, $f(x) = -\int_0^x e^{\lambda(x-y)} g(y)\,\mathrm{d}y$ is the unique function with those properties. So $\lambda - A$ is bijective, i.e. $\lambda \in \rho(A)$, and $\mathcal{R}(\lambda, A)g(x) = -\int_0^x e^{\lambda(x-y)} g(y)\,\mathrm{d}y$ for all $g \in C_0((0,1])$. □

As our final example in this section we consider the Laplace operator. For an open set $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ and a function $u \in C^2(\Omega)$, the Laplace operator applied to u is defined as the trace of the Hessian matrix of u:

$$\Delta u := \sum_{j=1}^n \partial_j^2 u.$$

We want to define Δ as an operator on $L^p(\Omega)$. To this end, we proceed analogously to Definition 3.2.4 and define Δ in a weak sense.

Definition 3.3.5 (The weak Laplace operator)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ and $u \in L^1_{\mathrm{loc}}(\Omega)$. We say that Δu **exists weakly** or, briefly and by a slight abuse of notation, that $\Delta u \in L^1_{\mathrm{loc}}(\Omega)$, if there exists a (necessarily unique) function $g \in L^1_{\mathrm{loc}}(\Omega)$ that such

$$\int_\Omega g\varphi \,\mathrm{d}x = \int_\Omega u \Delta\varphi \,\mathrm{d}x$$

for every test function $\varphi \in C_c^\infty(\Omega)$. In this case we set $\Delta u := g$.

Similarly as in Example 3.2.5, we could consider the Laplace operator on all functions $u \in L^p(\Omega)$ that satisfy $\Delta u \in L^p(\Omega)$. There are two caveats though. First, the L^p theory turns out to be substantially more involved. We thus stick to the simpler case $p = 2$ for now and return to the general case later. Second, without specifying additional conditions, a similar phenomenon as in Example 3.3.4(a) occurs – the spectrum of the Laplace operator is $\mathbb{C}$ in most cases. A common way to resolve this problem is to impose boundary conditions on the functions in the domain of the operators. Many choices of boundary conditions occur in PDE theory. For now, we focus on one of the simplest cases, which is **Dirichlet boundary conditions**.

Example 3.3.6 (The Dirichlet Laplacian on $L^2(\Omega)$)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open. We define the **maximal Laplace operator** and the **Dirichlet Laplace operator** on $L^2(\Omega)$ by

$$\operatorname{dom}(\Delta_{\max}) := \{u \in L^2(\Omega) : \Delta u \in L^2(\Omega)\}, \qquad \Delta_{\max} u := \Delta u,$$
$$\operatorname{dom}(\Delta_{\mathrm{Dir}}) := \operatorname{dom}(\Delta_{\max}) \cap H_0^1(\Omega), \qquad \Delta_{\mathrm{Dir}} u := \Delta u.$$

They have the following properties:

(a) Both operators $\Delta_{\max}$ and Δ_{Dir} are closed and densely defined.

(b) Let $u, g \in L^2(\Omega)$. Then $u \in \text{dom}(\Delta_{\text{Dir}})$ and $\Delta_{\text{Dir}} u = g$ if and only if $u \in H_0^1(\Omega)$ and[6]

$$\big(v \mid g\big)_{L^2} = -(\nabla v \mid \nabla u)_{L^2} \qquad \text{for all } v \in H_0^1(\Omega).$$

(c) Every $\lambda \in (0,\infty)$ is in the resolvent set $\rho(\Delta_{\text{Dir}})$.

Note that the intersection with $H_0^1(\Omega)$ in the definition of $\text{dom}(\Delta_{\text{Dir}})$ means that we consider only functions that "vanish" on the boundary $\partial\Omega$. A precise formulation of this reasoning requires the theory of boundary traces of Sobolev functions, which we do not discuss in the main text. However, the interested readers can find a brief overview in supplementary Section 3.B, in particular in Theorem 3.B.3.

Proof of Example 3.3.6(a)–(c).

(b) By definition, $u \in \text{dom}(\Delta_{\text{Dir}})$ and $\Delta_{\text{Dir}} u = g$ if and only if $u \in \text{dom}(\Delta_{\max}) \cap H_0^1(\Omega)$ and $\Delta u = g$. The latter is equivalent to $u \in H_0^1(\Omega)$ and $\big(v \mid g\big)_{L^2} = (\Delta v \mid u)_{L^2}$ for all $v \in C_c^\infty(\Omega)$, since a function $v \in C_c^\infty(\Omega)$ if and only if its complex conjugate $\overline{v} \in C_c^\infty(\Omega)$.

Note that for each $u \in H_0^1(\Omega)$, we have $(\Delta v \mid u)_{L^2} = -(\nabla v \mid \nabla u)_{L^2}$ for all $v \in C_c^\infty(\Omega)$ by Definition 3.2.4(a). Whence $u \in \text{dom}(\Delta_{\text{Dir}})$ and $\Delta_{\text{Dir}} u = g$ if and only if $u \in H_0^1(\Omega)$ and

$$\big(v \mid g\big)_{L^2} = -(\nabla v \mid \nabla u)_{L^2} \qquad \text{for all } v \in C_c^\infty(\Omega).$$

However, the validity of the above for all $v \in C_c^\infty(\Omega)$ is equivalent to its validity for all $v \in H_0^1(\Omega)$, since both sides are continuous in v with respect to the H^1-norm and since $C_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$ with respect to this norm.

(a) Both operators $\Delta_{\max}$ and Δ_{Dir} are densely defined since their domains contain the dense subspace $C_c^\infty(\Omega)$. The fact that $\Delta_{\max}$ is closed follows by precisely the same argument as the closedness of the differential operator ∂^α in Example 3.2.5 (but using Definition 3.3.5 in place of Definition 3.2.4).

Substituting $v := u$ in the equality in (b), we observe that[7,8]

$$\|\nabla u\|_{L^2}^2 \le \|u\|_{L^2}\,\|\Delta u\|_{L^2} \tag{3.3.1}$$

from the Cauchy–Schwarz inequality in $L^2(\Omega)$.

Since $\Delta_{\max}$ is closed, $\text{dom}(\Delta_{\max})$ is complete with respect to the graph norm $\|u\|_{\Delta_{\max}} = \|\Delta u\|_2 + \|u\|_2$. As the Sobolev space $H_0^1(\Omega)$ is also complete (by Proposition 3.2.7), it follows from Lemma 3.2.8 that the norm $u \mapsto \|\Delta u\|_2 + \|\nabla u\|_2 + \|u\|_2$ is complete on $\text{dom}(\Delta_{\text{Dir}})$. Moreover, this norm is equivalent to every graph norm of Δ_{Dir} since the inequality (3.3.1) implies that $\|\Delta u\|_2^2 + \|\nabla u\|_2^2 + \|u\|_2^2 \le (\|\Delta u\|_2 + \|u\|_2)^2$ for each $u \in \text{dom}(\Delta_{\text{Dir}})$. Hence, Δ_{Dir} is also closed.

[6] For $v \in H^1(\Omega)$ we use the same notation $\nabla v := (\partial_1 v, \dots, \partial_n v)^{\mathrm{T}} \in L^2(\Omega; \mathbb{C}^n)$ as for classical derivatives.
[7] Here $\|\nabla u\|_{L^2}^2 := \int_\Omega \|\nabla v(x)\|_2^2 \, \mathrm{d}x$, where $\|\nabla v(x)\|_2$ is the Euclidean norm of the vector $\nabla v(x) \in \mathbb{C}^n$.
[8] This is a very simple example of an **interpolation inequality**.

(c) Let $\lambda \in (0,\infty)$ and let $g \in L^2(\Omega)$. For every $u \in L^2(\Omega)$ the conditions $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$ and $(\lambda - \Delta_{\mathrm{Dir}})u = g$ are – according to (b) – equivalent to $u \in H_0^1(\Omega)$ and $\big(v \mid \lambda u - g\big)_{L^2} = -(\nabla v \mid \nabla u)_{L^2}$ for all $v \in H_0^1(\Omega)$, which is in turn equivalent to $u \in H_0^1(\Omega)$ and

$$\lambda\,(v \mid u)_{L^2} + (\nabla v \mid \nabla u)_{L^2} = \big(v \mid g\big)_{L^2} \qquad \text{for all } v \in H_0^1(\Omega). \tag{3.3.2}$$

Observe that for every $u \in H_0^1(\Omega)$ one has

$$\min\{1,\lambda\}\big(\|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2\big) \le \lambda\,\|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2 \le \max\{1,\lambda\}\big(\|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2\big).$$

Hence the norm induced by the inner product

$$\mathfrak{a}(v,u) := \lambda\,(v \mid u)_{L^2} + (\nabla v \mid \nabla u)_{L^2}$$

is equivalent to the standard H^1-norm on $H_0^1(\Omega)$. As $\big(\cdot \mid g\big)_{L^2}$ is a continuous antilinear functional[9] on $H_0^1(\Omega)$, the Riesz representation theorem on Hilbert spaces gives a unique $u \in H_0^1(\Omega)$ satisfying (3.3.2). Consequently, $\lambda \in \rho(\Delta_{\mathrm{Dir}})$. □

The spectral information given in Example 3.3.6(c) is far from optimal. We shall see more about the spectrum of Δ_{Dir} as we proceed.

[9] Recall that a map $\varphi\colon X \to \mathbb{C}$ from a complex Banach space X to $\mathbb{C}$ is called **antilinear** if $\varphi(x + \alpha y) = \varphi(x) + \overline{\alpha}\varphi(y)$ for all $x, y \in X$ and all $\alpha \in \mathbb{C}$.

Exercises for Chapter 3

Exercise 3.1 (The derivative at a point is not closed)**.** Define the operator

$$A\colon \mathrm{C}([0,1]) \supseteq \operatorname{dom}(A) := \mathrm{C}^1([0,1]) \to \mathbb{C}, \qquad Af := f'(1).$$

Prove that A is densely defined but not closed.

Exercise 3.2 (Multiplication operators)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open and let $\mathrm{C}_0(\Omega)$ denote the space of continuous functions $f : \Omega \to \mathbb{C}$ with the following property: for every $\varepsilon > 0$ there exists a compact set $K \subseteq \Omega$ such that $\left|f(x)\right| \leq \varepsilon$ for all $x \in \Omega \setminus K$. This is a Banach space with respect to the sup norm $\|\cdot\|_\infty$.

Let $h : \Omega \to \mathbb{C}$ be continuous and define the operator M_h on $\mathrm{C}_0(\Omega)$ by

$$\operatorname{dom}(M_h) := \left\{f \in \mathrm{C}_0(\Omega) : hf \in \mathrm{C}_0(\Omega)\right\}, \qquad M_h f := hf.$$

(a) Show that M_h is closed and densely defined.

(b) Prove that M_h is everywhere defined if and only if h is bounded.

(c) Show that $\sigma(M_h) = \overline{h(\Omega)}$.

Exercise 3.3. This exercise applies the closed graph theorem (Theorem 3.1.3).

(a) Let X, Y, Z be Banach spaces and consider linear operators $X \xrightarrow{T} Y \xrightarrow{J} Z$, where J is injective. Show that if J and JT are continuous, then so is T.

(b) Let (Ω, μ) be a finite measure space and $1 \leq p \leq q \leq \infty$. Let $T\colon L^p(\Omega,\mu) \to L^p(\Omega,\mu)$ be a bounded linear operator whose range is contained in $L^q(\Omega,\mu)$. Show that T is continuous from $L^p(\Omega,\mu)$ to $L^q(\Omega,\mu)$.

(c) Let X be a Banach space, let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed linear operator and $T\colon X \to X$ a bounded linear operator such that $T(X) \subseteq \operatorname{dom}(A)$. Show that T is continuous from X to $\operatorname{dom}(A)$ if $\operatorname{dom}(A)$ is endowed with a graph norm.

Exercise 3.4.

(a) Consider the function $f \in L^1_{\mathrm{loc}}(\mathbb{R})$ given by $f(x) = |x|$. Show that f is weakly differentiable and compute its weak derivative.

(b) Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) \cap \mathrm{C}^1(\mathbb{R}^n \setminus \{0\})$[10] and assume that there exist $g_1, \dots g_n \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ coinciding respectively with the classical partial derivatives $\partial_1 f, \dots, \partial_n f$ on $\mathbb{R}^n \setminus \{0\}$. Assume in addition that $\int_{\|x\|_2 = r} |f(x)| \, \mathrm{d}\sigma(x) \to 0$ as $r \downarrow 0$, where σ denote the surface measure of the ball with radius r in $\mathbb{R}^n$.

Show that f is weakly differentiable with weak derivatives $g_1, \dots, g_n$.

Hint: Given a test function $\varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^n)$, let $R > 0$ be such that $\operatorname{supp} \varphi \subseteq B_{<R}(0)$. Apply the divergence theorem on a 'shell' $\{x \in \mathbb{R}^n : r \le \|x\|_2 \le R\}$, and let $r \downarrow 0$.

(c) Let $f \colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|_2$. Show that f is weakly differentiable and compute its weak derivatives $\partial_1 f, \dots, \partial_n f$.

Exercise 3.5. Let X, Y be Banach spaces over the same field and let $A \colon X \supseteq \operatorname{dom}(A) \to Y$ be a densely defined linear operator.

(a) Prove that the dual operator $A' \colon Y' \supseteq \operatorname{dom}(A') \to X'$ (Definition 3.1.5) is closed.

(b) Give an example where A' is not densely defined.

Suggestion: Take $X = Y = \ell^1$ and take A to be a multiplication operator, analogous to Exercise 3.2.

(c) If $X = Y$, show that $\operatorname{dom}\big((\lambda + A)'\big) = \operatorname{dom}(A')$ and $\operatorname{dom}\big((\alpha A)'\big) = \operatorname{dom}(A')$ and that[11]

$$(\lambda + A)' = \lambda + A' \qquad \text{and} \qquad (\alpha A)' = \alpha A'$$

for all scalars λ, α with $\alpha \neq 0$. What goes wrong if $\alpha = 0$?

Assume now that $X = Y$, that the scalar field is $\mathbb{C}$, and that A is closed.

(d) Show that if $\lambda \in \rho(A)$, then also $\lambda \in \rho(A')$ and $\mathcal{R}(\lambda, A') = \mathcal{R}(\lambda, A)'$, where the latter operator denotes the dual of $\mathcal{R}(\lambda, A) \in \mathcal{L}(X)$.

(e) Conversely, show that if $\lambda \in \rho(A')$, then also $\lambda \in \rho(A)$.

Hints: First show that $\|x\|_X \le \|(\lambda - A)x\|_X \, \|\mathcal{R}(\lambda, A')\|$ for all $x \in \operatorname{dom}(A)$. Then derive that there exists a (λ-dependent) number $c > 0$ such that $\|x\|_A \le c \, \|(\lambda - A)x\|_X$ for all $x \in \operatorname{dom}(A)$. Hence deduce that $\lambda - A$ is injective and has closed range.

[10] Strictly speaking, this means that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and f has a representative whose restriction to the set $\mathbb{R}^n \setminus \{0\}$ is C^1.

[11] For a linear operator $A \colon X \supseteq \operatorname{dom}(A) \to Y$ and a scalar α one defines the operator $\alpha A \colon X \supseteq \operatorname{dom}(\alpha A) \to Y$ by $\operatorname{dom}(\alpha A) := \operatorname{dom}(A)$ and $(\alpha A)x := \alpha(Ax)$ for all $x \in \operatorname{dom}(\alpha A) = \operatorname{dom}(A)$.

Notes for Chapter 3

Unbounded operators and their spectra

> It is a fact of life that many of the most important operators which occur in mathematical physics are not bounded. — [RS80, p. 249]

As the above quote from the classic text of Reed and Simon suggests, unbounded operators are among the most fundamental objects in functional analysis. Their importance was already recognised in the early days of quantum mechanics: the so-called **position operator** that describes – no terminological surprise here – the position of a particle, acts as a multiplication with an unbounded function on the space $L^2(\mathbb{R})$. And the so-called **momentum operator** acts as a first order differential operator on the same space; see [Hal13, Chapter 3] for an accessible introduction to the relevant physical concepts, tailored to mathematicians. Similarly, motivated by the study of differential operators arising in (classical and quantum) physics, the spectral theory of unbounded operators has long been a fruitful subject. For applications in quantum physics, we again refer the interested reader to [Hal13] and [RS80, Chapter VIII].

While we are in the realm of physics, we point out that the expression $(\nabla v \mid \nabla u)_{L^2}$ plays an important role in the study of the Laplacian in the weak formulation, as shown in Example 3.3.6(b). The map $(v,u) \mapsto (\nabla v \mid \nabla u)_{L^2}$ is a **sesquilinear form**, and the associated **quadratic form** $u \mapsto \|\nabla u\|_{L^2}^2$ often has the physical interpretation of an 'energy'. Sesquilinear form methods provide another approach to the study of differential operators, and unsurprisingly are widely used in mathematical physics and the calculus of variations. We will discuss some basic aspects of these methods in Chapter 5.

A word on the closed graph theorem

The usual proofs of the closed graph theorem for Banach spaces (Theorem 3.1.3), found in many standard texts on functional analysis, rely on the **Baire category theorem**. It turns out that it is not necessary to use this theorem, as shown, for instance, in [Kes21]; cf. [Kes17]. In addition, the latter reference illustrates the equivalence between closed graph theorem, open mapping theorem, bounded inverse theorem, and the uniform boundedness principle in Banach spaces. This shows that completeness is the underlying principle in these foundational results. We also refer the reader to [Rot94], where a version of the Hahn-Banach theorem is given and employed to prove the uniform boundedness principle and the open mapping theorem.

Weak derivatives and Sobolev spaces

An alternative but quite natural way to define a space of weakly differentiable functions is via limits of classically differentiable functions. To be precise, we define the norm $\|\cdot\|_{W^{k,p}}$ exactly as in Definition 3.2.6, and let $H^{k,p}(\Omega)$ denote the completion of the space

$$\{u \in C^{\infty}(\Omega) : \|u\|_{W^{k,p}} < \infty\}$$

with respect to the $W^{k,p}$ norm. While the inclusion that $H^{k,p}(\Omega) \subseteq W^{k,p}(\Omega)$ is straightforward to check, it took some time before the converse inclusion was finally settled in 1964 by Norman Meyers and James Serrin in their iconic paper [MS64]. We comment on this further in the Supplement 3.A.

The weak derivative in Definition 3.2.4 is often also called the **distributional derivative**. This is a fundamental notion in the theory of distributions, also known as generalised functions. The idea of extending differentiation beyond the classical setting originates well before the 20th century; a detailed historical account of this can be found in [Lue82, Chapter 2]. From the perspective of the theory of distributions, elements of a Sobolev space are simply 'well-behaved' distributions, where the element itself and all its distributional derivatives up to a given order can be represented by L^p functions.

One can also define Sobolev spaces via the Fourier transform, as is typically done in harmonic analysis. For readers interested in this approach, there is a wide selection of good literature, including [Str03, Chapter 8], [Gru09, Part II], and [Gra14, Chapter 1].

Encore: if you want to know more...

3.A Regularisation of functions

In this supplementary section, we use the following notation: given a measurable function $h : \mathbb{R}^n \to \mathbb{R}$, for each $t > 0$ we set

$$h_t(x) := \frac{1}{t^n} h\left(\frac{x}{t}\right). \tag{3.A.1}$$

Definition 3.A.1 (Mollifiers). A **mollifier** is a family $\{\rho^t : t > 0\}$ of functions $\rho^t : \mathbb{R}^n \to \mathbb{R}$ satisfying the following properties for each $t > 0$:

(1) $\rho^t \in C_c^\infty(\mathbb{R}^n)$ and ρ^t is supported in $B_{\le t}(0)$; and

(2) $\rho^t \ge 0$ on $\mathbb{R}^n$ and $\int_{\mathbb{R}^n} \rho^t(x)\, \mathrm{d}x = 1$.

We define a special test function $\theta \in C_c^\infty(\mathbb{R}^n)$ by

$$\theta(x) := \begin{cases} c \exp\left(-\frac{1}{1-|x|^2}\right) & |x| < 1 \\ 0 & |x| \ge 1, \end{cases} \tag{3.A.2}$$

where the constant $c > 0$ is chosen so that $\int_{\mathbb{R}^n} \theta(x)\, \mathrm{d}x = 1$. Then the family $\{\theta_t : t > 0\}$ (using the notation (3.A.1)) is called the **standard mollifier**.

The key features of mollifiers are that they consist of very smooth functions, and most crucially, by taking $t \downarrow 0$, the support of ρ^t can be made as small as desired. This allows us to **regularise** non-smooth functions via convolutions while maintaining precise control of the support. We summarise some standard facts about convolutions and regularisations below, and refer to the literature (e.g. [Bre11, Section 4.4]) for the proofs.

Proposition 3.A.2 (An analysis toolkit).

(a) (Young's inequality) *Let $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$ for $1 \le p \le \infty$. Then $f * g \in L^p(\mathbb{R}^n)$ with* $\operatorname{supp}(f * g) \subseteq \operatorname{supp} f + \operatorname{supp} g$, *and*

$$\|f * g\|_p \le \|f\|_1 \|g\|_p.$$

(b) (Regularisation) *For all $g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$, one has $\varphi * g \in C^\infty(\mathbb{R}^n)$ with*

$$\partial^\alpha(\varphi * g) = (\partial^\alpha \varphi) * g \quad \textit{for all } \alpha \in \mathbb{N}_0^n.$$

(c) (Approximate identity) *Let* $\{\rho^t : t > 0\}$ *be a mollifier. Then* $\lim_{t\to 0} \|\rho^t * f - f\|_p = 0$ *for all* $f \in L^p(\mathbb{R}^n)$ *with* $1 \le p < \infty$.

(d) (Density of test functions) *For every open set* $\emptyset \neq \Omega \subseteq \mathbb{R}^n$*, the space of test functions* $C_c^\infty(\Omega)$ *is dense in* $L^p(\Omega)$ *for* $1 \le p < \infty$.

We can now complete the proof of Proposition 3.2.3 from the main text.

Theorem 3.A.3 (Fundamental lemma of the calculus of variations)**.** *Let* $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ *be open. Suppose* $f \in L^1_{\mathrm{loc}}(\Omega)$ *satisfies*

$$\int_\Omega f\varphi \,\mathrm{d}x = 0 \quad \textit{for all } \varphi \in C_c^\infty(\Omega). \tag{3.A.3}$$

Then $f = 0$ *in* Ω.

Proof. Choose increasing compact subsets Ω_k so that $\Omega = \bigcup_{k\in\mathbb{N}} \Omega_k$; for instance, take

$$\Omega_k := \left\{ x \in \Omega : |x| \le k \text{ and } \operatorname{dist}(x, \partial\Omega) \ge \frac{1}{k} \right\}.$$

It suffices to prove that $f(x) = 0$ for almost every $x \in \Omega_k$ for each $k \in \mathbb{N}$. We do this via an approximation argument.

Since $f \in L^1_{\mathrm{loc}}(\Omega)$, we have $f_k := f\,\mathbb{1}_{\Omega_{k+1}} \in L^1(\Omega)$ for each $k \in \mathbb{N}$. Denote by $\tilde{f}_k$ the extension of f_k by 0 outside Ω_{k+1}, and note that $\tilde{f}_k \in L^1(\mathbb{R}^n)$. Since $\Omega_k \subset \Omega_{k+1} \subset \Omega$ with strict inclusions, for all sufficiently small $t > 0$ (depending on k) the support of $\theta_t(x - \cdot)$ lies in Ω_{k+1} for all $x \in \Omega_k$, and therefore $\theta_t(x - \cdot) \in C_c^\infty(\Omega)$ for all $x \in \Omega_k$. Consequently

$$(\theta_t * \tilde{f}_k)(x) = \int_{\mathbb{R}^n} \tilde{f}_k(y)\theta_t(x-y)\,\mathrm{d}y = \int_\Omega f(y)\theta_t(x-y)\,\mathrm{d}y = 0$$

for all sufficiently small $t > 0$ and all $x \in \Omega_k$, where we have used the property that $\tilde{f}_k = f$ on Ω_{k+1}, followed by assumption (3.A.3) in the last equality. By Proposition 3.A.2(c), we conclude

$$0 = \lim_{t\downarrow 0}(\theta_t * \tilde{f}_k) = \tilde{f}_k$$

in $L^1(\mathbb{R}^n)$, which implies that $f|_{\Omega_k} = 0$ as desired. □

The following theorem relies on a clever use of partition of unity and regularisation.

Theorem 3.A.4 (Meyers, Serrin)**.** *Let* $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ *be open, and let* $k \in \mathbb{N}$ *and* $p \in [1, \infty)$*. Then* $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ *is dense in* $W^{k,p}(\Omega)$.

The original 1964 paper of Meyers and Serrin [MS64] arguably has one of the most iconic titles in the field of analysis: $H = W$. It is merely one-and-a-half pages long with an extremely brief proof. The reader interested in a detailed proof is thus advised to consult more recent literature, e.g. [GT01, Theorem 7.9] or [Leo09, Theorem 11.24]. Having said that, the following simple yet important corollary is worth presenting in detail.

Corollary 3.A.5. *Let $k \in \mathbb{N}$ and $p \in [1,\infty)$. Then*

$$W_0^{k,p}(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n).$$

Proof. The non-trivial inclusion to prove is $W^{k,p}(\mathbb{R}^n) \subseteq W_0^{k,p}(\mathbb{R}^n)$, i.e. to show that every $u \in W^{k,p}(\mathbb{R}^n)$ can be approximated in the $W^{k,p}$ norm by a sequence $(u_m) \subset \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$.

By the Meyers-Serrin Theorem (Theorem 3.A.4), it suffices to prove the claim for $u \in \mathrm{C}^\infty(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$. For this purpose, fix a function $\zeta \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ such that $0 \le \zeta \le 1$ on $\mathbb{R}^n$ and $\zeta \equiv 1$ for $|x| \le 1$, and set

$$u_m(x) := \zeta\left(\frac{x}{m}\right) u(x), \qquad x \in \mathbb{R}^n$$

for $m \in \mathbb{N}$. Clearly $u_m \in \mathrm{C}_\mathrm{c}^\infty(\mathbb{R}^n)$ for each $m \in \mathbb{N}$, and the dominated convergence theorem yields $u_m \to u$ in $L^p(\mathbb{R}^n)$. The generalised Leibniz rule yields

$$\partial^\alpha u_m = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta \big(\zeta(\cdot/m)\big) \partial^{\alpha-\beta} u = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \frac{1}{m^{|\beta|}} (\partial^\beta \zeta)(\cdot/m) \partial^{\alpha-\beta} u$$

for every multi-index $|\alpha| \le k$. If $\beta = (0,\dots,0)$, then again by dominated convergence we obtain $\zeta(\frac{\cdot}{m})\partial^\alpha u \to \partial^\alpha u$ in $L^p(\mathbb{R}^n)$. For $\beta \ne 0$, observe that

$$\int_{\mathbb{R}^n} \left| \partial^\beta \big(\zeta(x/m)\big) \partial^{\alpha-\beta} u(x) \right|^p \,\mathrm{d}x \le \frac{C^p}{m^{|\beta| p}} \int_{\mathbb{R}^n} \left| \partial^{\alpha-\beta} u(x) \right|^p \,\mathrm{d}x \to 0$$

as $m \to \infty$, where $C := \max_{|\beta| \le k} \left\| \partial^\beta \zeta \right\|_{L^\infty(\mathbb{R}^n)} < \infty$. Altogether, we have shown that $\partial^\alpha u_m \to \partial^\alpha u$ in L^p for every multi-index $|\alpha| \le k$, and thus $u_m \to u$ in $W^{k,p}(\mathbb{R}^n)$. □

3.B Traces of $W^{1,p}$ functions

In Example 3.3.6, we introduced the Dirichlet Laplace operator Δ_{Dir} with domain contained in $H_0^1(\Omega)$, a space which intuitively encodes the boundary condition '$u = 0$ on $\partial\Omega$'. In Theorem 3.B.3 below, we make precise the sense in which a function in the Sobolev space $W_0^{1,p}(\Omega)$ 'vanishes' on the boundary $\partial\Omega$. This is achieved via the theory of **traces**. Before we proceed, we need to understand 'regularity' of the boundary of subsets of $\mathbb{R}^n$.

A **rigid motion** of $\mathbb{R}^n$ is a map $T : \mathbb{R}^n \to \mathbb{R}^n$ of the form $T(x) = Rx + c$, for a rotation R and fixed $c \in \mathbb{R}^n$. Using this notion, we can say what it means for the boundary of an open set $\Omega \subseteq \mathbb{R}^n$ to be Lipschitz continuous. Intuitively, this means that the boundary looks locally like the graph of a scalar-valued Lipschitz function in $n-1$ variables.

Definition 3.B.1. Let $n \ge 2$, and let $\emptyset \ne \Omega \subset \mathbb{R}^n$ be open. We say that the boundary $\partial\Omega$ is **Lipschitz continuous** (or simply **Lipschitz**) if for every $\xi_0 \in \partial\Omega$, there exists a rigid motion T with $T(\xi_0) = 0$, a Lipschitz continuous function $f : \mathbb{R}^{n-1} \to \mathbb{R}$, and $r > 0$ such that

$$T\big(\Omega \cap B_{<r}(\xi_0)\big) = \big\{x \in B_{<r}(0) : x_n > f(x_1,\dots,x_{n-1})\big\}$$

where the **local coordinates** are given by $x := T(\xi)$ for all $\xi \in \Omega \cap B_{<r}(\xi_0)$.

Likewise, given $k \in \mathbb{N}_0 \cup \{\infty\}$, we say that $\partial\Omega$ is **of class** C^k if the functions f occurring above belong to $\mathrm{C}^k(\mathbb{R}^{n-1};\mathbb{R})$.

When we assume boundary regularity in the main text we stick to C^k since it has an equivalent definition that is less technical to state (Definition 5.3.1). An example of a region in $\mathbb{R}^2$ with Lipschitz boundary is illustrated in Figure 3.B.1.

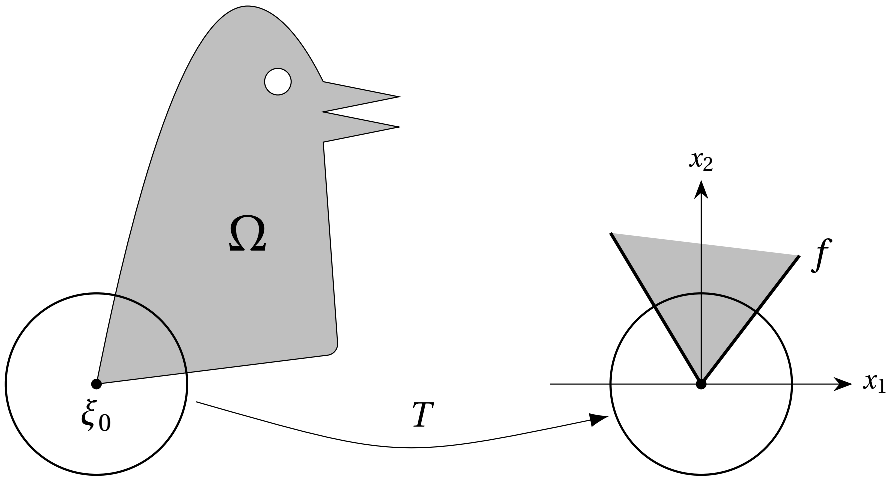

Figure 3.B.1: A region with Lipschitz boundary and local coordinates around ξ_0.

In the theorem below, the boundary $\partial\Omega$ is equipped with the $(d-1)$-dimensional Hausdorff measure.

Theorem 3.B.2 (Trace operator)**.** *Let $n \geq 2$, $1 \leq p < \infty$, and let $\emptyset \neq \Omega \subset \mathbb{R}^n$ be an open set with bounded and Lipschitz continuous boundary $\partial\Omega$. There exists a unique linear operator* $\operatorname{Tr}\colon W^{1,p}(\Omega) \to L^p(\partial\Omega)$, *called the* ***trace operator****, with the following properties:*

(a) $\operatorname{Tr} u = u|_{\partial\Omega}$ *for all* $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$.

(b) (Trace inequality) *There exists* $C > 0$ *(depending only on* Ω*) such that*

$$\|\operatorname{Tr} u\|_{L^p(\partial\Omega)} \leq C\,\|u\|_{W^{1,p}(\Omega)} \qquad \text{for all } u \in W^{1,p}(\Omega).$$

For a proof, we refer to [Leo09, Theorem 18.1]. Note that in this reference, the result is stated more generally for open sets with unbounded boundary, but it is easy to see that it reduces to our statement when $\partial\Omega$ is bounded.

The space $W_0^{1,p}(\Omega)$ can now be characterised as the kernel of the trace operator.

Theorem 3.B.3 (Characterisation of $W_0^{1,p}(\Omega)$)**.** *Let $n \geq 2$, $1 \leq p < \infty$, and let $\emptyset \neq \Omega \subset \mathbb{R}^n$ be an open set with bounded and Lipschitz continuous boundary $\partial\Omega$. For all $u \in W^{1,p}(\Omega)$, the following assertions are equivalent:*

(i) $\operatorname{Tr} u = 0$.

(ii) $u \in W_0^{1,p}(\Omega)$.

We refer to [Leo09, Theorem 18.7] for the proof.

Chapter 4

Ordered function spaces and Banach lattices

After studying (eventual) positivity for matrices in Chapters 1 and 2, we encountered a selection of unbounded operators in infinite-dimensional spaces in Chapter 3. It is our goal to analyse positivity properties related to such operators. In particular, the question of when the resolvent of an unbounded operator is positive will be important.

This requires an order structure on the underlying Banach spaces. All examples in Chapter 3 were defined on function spaces, where it is natural to consider the pointwise (almost everywhere) order. Yet, it turns out that this order behaves quite differently on, for instance, L^p-spaces compared to $C_0(\Omega)$. In this chapter, we introduce the general framework of **Banach lattices** that includes many classical function spaces and allows us to develop the subsequent theory on all such spaces simultaneously.[1]

4.1 Real Banach lattices and function spaces

We recall a few elementary notions about partially ordered sets. Consider a set X together with a **partial order** $\le$ on X, which means that $\le$ is a reflexive, anti-symmetric and transitive relation on X. Let $S \subseteq X$. An element $b \in X$ is called an **upper bound** of S if $x \le b$ for all $x \in S$, and b is called a **lower bound** of S if $b \le x$ for all $x \in S$. The element $b \in X$ is called the **least upper bound** or the **supremum** of S if b is an upper bound of S and, in addition, $b \le c$ for all upper bounds c of S.[2] Similarly, b is called the **greatest lower bound** or the **infimum** of S if b is a lower bound of S and, in addition, $b \ge c$ for all lower bounds c of S. It is not difficult to see that the supremum or infimum of a set is unique if it exists. We denote the supremum and the infimum of S respectively by $\sup S$ and $\inf S$ whenever they exist. Finally, for two elements $x, y \in X$ we use the notation $x \vee y := \sup\{x, y\}$ and $x \wedge y := \inf\{x, y\}$ whenever they exist.

[1] Even beyond a unified treatment of different types of function spaces, there are good reasons to develop the theory in the general setting of Banach lattices. We explain this later on.

[2] In other words, b is an upper bound of S and a lower bound of the set of all upper bounds.

Definition 4.1.1 (Real vector lattices)**.** A **real vector lattice**[3] is a real vector space V together with a partial order $\le$ on V that satisfies the following properties:

(I) The order $\le$ is compatible with the linear structure of V, i.e. for all $x, y, z \in V$ that satisfy $x \le y$ and all real numbers $\alpha \in [0, \infty)$ one has $x + z \le y + z$ and $\alpha x \le \alpha y$.

(II) Any two elements $x, y \in V$ have a supremum $x \vee y$ in V.

A vector subspace W of V is called a **vector sublattice** of V if $x \vee y \in W$ for all $x, y \in W$.

Note that property (I) implies that $x \le y$ if and only if $-y \le -x$. One can easily deduce from the definition of a vector lattice V that for all $x, y \in V$ the infimum $x \wedge y$ also exists and is equal to $-\big((-x) \vee (-y)\big)$. We frequently use the following concepts in vector lattices.

Definition 4.1.2. Let V be a real vector lattice.

(a) The set $V_+ := \{x \in V : x \ge 0\}$ is called the **positive cone** or, briefly, the **cone** of V.

The elements of V_+ are called the **positive** elements of V.

(b) For all $x \in V$, we call the elements

$$x^+ := x \vee 0, \qquad x^- := (-x)^+, \qquad |x| := x \vee (-x)$$

of V_+ the **positive part**, the **negative part**, and the **modulus** of x respectively.[4]

Observe that for elements x, y of a vector lattice V, one has $x \le y$ if and only if $y - x \in V_+$. Moreover, for each $x \in V$ the negative part is given by $x^- = -(x \wedge 0)$.

Proposition 4.1.3 (Algebraic properties in vector lattices)**.** *Let x, y, z be elements of a real vector lattice V. The following identities hold:*

(a) $x \vee y = -\big((-x) \wedge (-y)\big)$.

(b) *Translation property:* $(x \vee y) + z = (x + z) \vee (y + z)$ *and* $(x \wedge y) + z = (x + z) \wedge (y + z)$.

(c) *Scaling property:* $\alpha(x \vee y) = (\alpha x) \vee (\alpha y)$ *and* $\alpha(x \wedge y) = (\alpha x) \wedge (\alpha y)$ *for all scalars* $\alpha > 0$, *and* $|\alpha x| = |\alpha|\,|x|$ *for all scalars* $\alpha \in \mathbb{R}$.

(d) *Distributive law:* $(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z)$ *and* $(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z)$.

(e) *Triangle inequalities:* $\left|x + y\right| \le |x| + \left|y\right|$ *and* $\left|x - y\right| \ge \big||x| - |y|\big|$.

(f) $x \vee y = \frac{1}{2}\big(x + y + |x - y|\big)$ *and* $x \wedge y = \frac{1}{2}\big(x + y - |x - y|\big)$.

(g) $x = x^+ - x^-$ *and* $|x| = x^+ + x^-$.

[3] Also known as a **Riesz space**, especially in the Dutch tradition.
[4] One can also call $|x|$ the **absolute value** of x.

Proof. (a)–(e) Assertions (a)–(c) and (e) are straightforward to verify directly from the definitions. The non-trivial inequality in the distributive law (d) can be obtained[5] by first showing that $x \vee y + x \wedge y = x + y$ for all $x, y \in V$; see (f) below.

(f) Using the definition of the modulus and properties (b) and (c), observe that

$$x + y + |x - y| = x + y + (x - y) \vee (y - x) = (2x) \vee (2y) = 2(x \vee y).$$

Similarly, using (a) in addition, we find

$$\begin{aligned} x + y - |x - y| &= x + y - [(x - y) \vee (y - x)] \\ &= x + y + (y - x) \wedge (x - y) = (2y) \wedge (2x) = 2(x \wedge y). \end{aligned}$$

(g) By setting $y = 0$ in part (f), we obtain

$$x^+ = x \vee 0 = \tfrac{1}{2}(x + |x|) \quad \text{and} \quad x^- = (-x) \vee 0 = -(x \wedge 0) = -\tfrac{1}{2}(x - |x|).$$

Adding and subtracting these yields the claim. □

Examples 4.1.4. We discuss some standard examples of vector lattices.

(a) We consider $\mathbb{R}^n$ with the componentwise ordering introduced in Chapter 1. It is easy to check that the supremum of any two vectors exists and is given by

$$(x \vee y)_i = x_i \vee y_i \quad \forall\, x, y \in \mathbb{R}^n, i = 1, \ldots, n;$$

where $\vee$ on the right-hand side is simply the usual supremum in $\mathbb{R}$, and likewise for the infimum. With these operations, it is clear that $\mathbb{R}^n$ is a real vector lattice. Equivalently, one can use the definition of the modulus introduced in Definition 1.1.4 and recover the lattice operations via the identities in Proposition 4.1.3(f), which of course hold for real numbers.

(b) The space $\mathrm{C}(\Omega;\mathbb{R})$ of continuous real-valued functions defined on $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ is a real vector lattice with the natural pointwise ordering and lattice operations, i.e. $f \le g$ if and only if $f(x) \le g(x)$ for all $x \in \Omega$, $|f|(x) := |f(x)|$, and so on.

(c) Let (Ω, μ) be a measure space. The Lebesgue spaces $L^p(\Omega, \mu; \mathbb{R})$, $p \in [1, \infty]$ admit a partial order $f \le g$ if and only if $f(\omega) \le g(\omega)$ for μ-a.e. ω and a modulus $|f|(\omega) := |f(\omega)|$. The lattice operations are again expressed via the identities in Proposition 4.1.3(f), and thus we obtain a vector lattice structure on $L^p(\Omega, \mu; \mathbb{R})$.

(d) Now for a perhaps slightly surprising example: For each open set $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ and each $p \in [1, \infty]$, the Sobolev space $W^{1,p}(\Omega; \mathbb{R})$ is a vector sublattice of $L^p(\Omega; \mathbb{R})$ and

$$\partial_k(u^+) = \mathbb{1}_{\{u>0\}}\, \partial_k u, \qquad \partial_k(u^-) = -\,\mathbb{1}_{\{u<0\}}\, \partial_k u, \qquad \partial_k(|u|) = \operatorname{sgn}(u)\, \partial_k u$$

[5] Note that the vector lattice structure is important here. There exist lattices (i.e. partially ordered sets where each pair of elements has a supremum and an infimum) in which the distributive law does not hold.

for every $u \in W^{1,p}(\Omega;\mathbb{R})$ and every $k \in \{1,\dots,n\}$. This is a famous result due to the Italian mathematician Guido Stampacchia, often referred to as the **Stampacchia lemma**. To avoid a detour, we do not prove this result here. Readers interested in the details can find a proof in Supplement 4.B.

(e) If $p \neq \infty$, then the Sobolev space $W_0^{1,p}(\Omega;\mathbb{R})$ (recall that we did not define $W_0^{1,\infty}$) is a vector sublattice of $W^{1,p}(\Omega;\mathbb{R})$ (and thus of $L^p(\Omega;\mathbb{R})$ by (d)). The intuition here is quite clear: we think of $W_0^{1,p}(\Omega;\mathbb{R})$ as the subspace of those $u \in W^{1,p}(\Omega;\mathbb{R})$ that vanish at the boundary $\partial\Omega$. If u vanishes at the boundary, it seems plausible that $|u|$ does as well. A rigorous proof requires more theory of Sobolev spaces, so we do not show it in the lectures. Readers interested in the details can find them in [AU23, Theorem 6.37 and 6.39].

We now specialise to vector lattices which are also Banach spaces, so that all the standard tools of functional analysis are available at our disposal.

Definition 4.1.5 (Real Banach lattices)**.** A **real Banach lattice** is a Banach space E over $\mathbb{R}$ with a partial order $\le$ that turns E into a real vector lattice that is compatible with the norm in the following sense: whenever $x, y \in E$ satisfy $|x| \le |y|$, then $\|x\| \le \|y\|$.

Norms on real vector lattices satisfying the above compatibility condition are called **lattice norms**. Note that instead of the assumption that $|x| \le |y|$ implies $\|x\| \le \|y\|$, one can equivalently require that the following two properties are satisfied:

(a) For all $x \in E$ one has $\|x\| = \||x|\|$.

(b) For all $x, y \in E$ the inequalities $0 \le x \le y$ imply $\|x\| \le \|y\|$.

Let us now discuss which of the vector lattices in Examples 4.1.4 are Banach lattices.

Examples 4.1.6.

(a) For a given $p \in [1,\infty]$, it is easy to verify that the ℓ^p norm on $\mathbb{R}^n$, defined by

$$\|x\|_p := \begin{cases} \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} & \text{if } 1 \le p < \infty \\ \max_{k=1,\dots,n} |x_k| & \text{if } p = \infty, \end{cases}$$

is a lattice norm on $\mathbb{R}^n$. Since $\mathbb{R}^n$ is complete with respect to any of the above norms, it follows that $(\mathbb{R}^n, \|\cdot\|_p)$ is a real Banach lattice.

However, one can also construct norms on $\mathbb{R}^n$ which are not lattice norms. For instance, if $n = 2$, consider the norm

$$\|x\|_W := |x_1| + |x_2 - x_1|.$$

Observe that the vectors $a = (1,-1)^{\mathrm{T}}$ and $b = (1,1)^{\mathrm{T}}$ have the same modulus, but $\|a\|_W = 3 > \|b\|_W = 1$.

(b) If $\Omega = K$ is compact, then clearly the supremum norm $\|f\|_\infty := \sup_{x\in K} |f(x)|$ is a lattice norm on $\mathrm{C}(K;\mathbb{R})$. As $\mathrm{C}(K;\mathbb{R})$ is complete with respect to the supremum norm, it is a real Banach lattice.

(c) It is also clear that the L^p norms are lattice norms, and thus the $L^p(\Omega,\mu;\mathbb{R})$ spaces are real Banach lattices.

(d) The $W^{1,p}$ norms are *not* lattice norms. Since the $W^{1,p}$ norm involves the (weak) partial derivatives of functions, it is easy to see that it does not satisfy the lattice norm property (e.g. one can consider smooth, bounded functions with oscillations). Thus the Sobolev spaces $W^{1,p}(\Omega;\mathbb{R})$ are examples of real vector lattices which are Banach spaces but not real Banach lattices.

As the definition suggests, the setting of a Banach lattice allows order, lattice, and topological structures to fit together in a useful way. Let us makes this precise.

Proposition 4.1.7. *Let E be a real Banach lattice.*

(a) *The mappings $x \mapsto |x|$, $x \mapsto x^+$, and $x \mapsto x^-$ are continuous from E to E.*

(b) *The mappings $(x,y) \mapsto x \vee y$ and $(x,y) \mapsto x \wedge y$ are continuous from $E \times E$ to E.*

(c) *The positive cone E_+ is closed.*

(d) *If two sequences (x_n) and (y_n) in E converge to points $x, y \in E$ respectively, such that $x_n \le y_n$ for all indices n, then $x \le y$.*

Proof. (a) The continuity of $|\cdot|$ is an immediate consequence of the reverse triangle inequality (Proposition 4.1.3(e)) and the fact that the norm is a lattice norm (Definition 4.1.5). The continuity of the positive and negative parts is then immediate, since $x^+ = \frac{1}{2}(x + |x|)$ and $x^- = \frac{1}{2}(|x| - x)$ for all $x \in E$ (Proposition 4.1.3(g)).

(b) This follows from (a) and the representation formula of $\vee$ and $\wedge$ in Proposition 4.1.3(f).

(c) If $(x_n) \subset E_+$ is a sequence converging to $x \in E$, then $x_n^- = 0$ for all $n \in \mathbb{N}$, and thus $x^- = \lim_{n\to\infty} x_n^- = 0$ by (a). Hence, $x \ge 0$.

(d) This is an immediate consequence of (c). □

4.2 Complex Banach lattices

In all the previous chapters, spectral theory has been a recurring theme, and for this reason, it is important to consider vector spaces over $\mathbb{C}$. The theory of vector lattices, introduced in Section 4.1, is however a theory over the real field. To solve this issue we now study how to *complexify* a real vector space and, in a particular, a Banach lattice.

Definition 4.2.1 (Complexification of a real vector space)**.** Let V be a real vector space. We give the Cartesian product $V \times V$ the structure of a complex vector space by defining:

(i) $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ for all $(x_1, y_1), (x_2, y_2) \in V \times V$;

(ii) $(\alpha + i\beta)(x, y) = (\alpha x - \beta y, \beta x + \alpha y)$ for all $(x, y) \in V \times V$ and $\alpha, \beta \in \mathbb{R}$.

The resulting vector space is called the **complexification** of V, and is denoted as $V_{\mathbb{C}}$ or $V + iV$. Furthermore, given $(x, y) \in V_{\mathbb{C}}$, we define **complex conjugation** by $\overline{(x, y)} := (x, -y)$.

Via the injective $\mathbb{R}$-linear map $V \to V_{\mathbb{C}}$, $x \mapsto (x, 0)$, we identify V with an $\mathbb{R}$-linear subspace of $V_{\mathbb{C}}$. Thus, $V_{\mathbb{C}}$ becomes the direct sum (over $\mathbb{R}$) of its subspaces V and iV. As a consequence, every $z \in V_{\mathbb{C}}$ can be written uniquely in the form $z = x + iy$ with $x, y \in V$. We make this explicit in the following.

Notation 4.2.2. In the setting of Definition 4.2.1, we denote each $z = (x, y) \in V_{\mathbb{C}}$ by

$$z = x + iy.$$

The **real** and **imaginary parts** of z are defined by $\operatorname{Re} z = x$ and $\operatorname{Im} z = y$, respectively. In particular, $\overline{z} = x - iy$.

For a real Banach lattice E it is natural to extend the modulus function $|\cdot| : E \to E_+$ to the complexification $E_{\mathbb{C}}$, just as the modulus on $\mathbb{R}$ can be extended to $\mathbb{C}$. We deviate from the standard construction in the literature and follow a more axiomatic approach.

Definition 4.2.3 (Complex modulus function)**.** Let E be a real Banach lattice and let $E_{\mathbb{C}}$ be its vector space complexification. A **complex modulus function** is a function $|\cdot|_{\mathbb{C}} : E_{\mathbb{C}} \to E_+$ that satisfies the following axioms:

(I) **Compatibility with the real modulus**: For all $x \in E$ one has $|x|_{\mathbb{C}} = |x|$.

(II) **Triangle inequality:** For all $z, w \in E_{\mathbb{C}}$ one has $|z + w|_{\mathbb{C}} \le |z|_{\mathbb{C}} + |w|_{\mathbb{C}}$.

(III) **Absolute homogeneity:** For all $z \in E_{\mathbb{C}}$ and all $\alpha \in \mathbb{C}$ one has $|\alpha z|_{\mathbb{C}} = |\alpha|\,|z|_{\mathbb{C}}$.

To obtain the existence of a modulus function on the complexification of a Banach lattice, one takes inspiration from the following formulae in $\mathbb{C}$. Given $z \in \mathbb{C}$, one can write z in the polar form $z = e^{i\varphi}|z|$ for a number $\varphi \in [0, 2\pi)$. On one hand, this gives

$$|z| = \sup_{\theta \in [0, 2\pi]} \operatorname{Re}(e^{i\theta} z). \tag{4.2.1}$$

and on the other,

$$\int_0^{2\pi} \left|\operatorname{Re}(e^{i\theta} z)\right| \, d\theta = |z| \int_0^{2\pi} \left|\operatorname{Re} e^{i(\varphi+\theta)}\right| \, d\theta = |z| \int_0^{2\pi} |\cos(\theta)| \, d\theta = 4|z|, \tag{4.2.2}$$

where have we used the 2π-periodicity of cos for the second equality. As only the modulus of real numbers occurs in (4.2.2), the formula shows us a potential way to construct a complex modulus function on a Banach lattice.

Theorem 4.2.4 (Existence and uniqueness of the complex modulus)**.** *Let E be a real Banach lattice and $E_{\mathbb{C}}$ its vector space complexification. There exists precisely one complex modulus function $|\cdot|_{\mathbb{C}} : E_{\mathbb{C}} \to E_+$. Moreover, it can be represented by the formulae*

$$|z|_{\mathbb{C}} = \frac{1}{4}\int_0^{2\pi} \left|\mathrm{Re}(e^{\mathrm{i}\theta} z)\right| \,\mathrm{d}\theta = \sup_{\theta\in[0,2\pi]} \mathrm{Re}(e^{\mathrm{i}\theta} z)$$

for all $z \in E_{\mathbb{C}}$.

The supremum in the theorem is understood within the vector lattice E (note that it is not at all obvious that this supremum exists since it is a supremum over an infinite set) and the integral can either be interpreted as a vector-valued Riemann integral or as a Bochner integral. Since we will often need integrals of functions with values in Banach spaces, we provide a brief overview of them in Appendix 4.A for readers who are not yet familiar with this concept.

We prove the existence result in Theorem 4.2.4 by showing that the integral formula in the theorem defines a complex modulus function. The proof of the uniqueness and of the supremum formula require more advanced tools from Banach lattice theory. Thus we do not prove it here and refer to [MW74, Theorem 2.2] instead.[6]

Proof of the existence and the integral formula in Theorem 4.2.4. For every $z \in E_{\mathbb{C}}$ we define $|z|_{\mathbb{C}} := \frac{1}{4}\int_0^{2\pi} \left|\mathrm{Re}(e^{\mathrm{i}\theta} z)\right| \,\mathrm{d}\theta$ for all $z \in \mathbb{C}$. Observe that the integral exists since $\theta \mapsto \left|\mathrm{Re}(e^{\mathrm{i}\theta} z)\right| = |\cos\theta \,\mathrm{Re}\, z - \sin\theta \,\mathrm{Im}\, z|$ is continuous from $[0, 2\pi]$ to E. We now prove that $|\cdot|_{\mathbb{C}}$ satisfies the axioms from Definition 4.2.3.

(I) For every $x \in E$ one has

$$|x|_{\mathbb{C}} = \frac{1}{4}\int_0^{2\pi} \left|\mathrm{Re}(e^{\mathrm{i}\theta} x)\right| \,\mathrm{d}\theta = \frac{1}{4}\int_0^{2\pi} |\cos\theta|\,|x| \,\mathrm{d}\theta = |x|.$$

(II) This follows readily from the triangle inequality of the modulus on the real Banach lattice E (Proposition 4.1.3(e)) since integration preserves the order structure on E.

(III) This follows from the same substitution that we used in formula (4.2.2). □

Given the existence and uniqueness of the complex modulus we can now define complex Banach lattices. Instead of using the notation E and $E_{\mathbb{C}}$ from above, it is often more convenient to write $E_{\mathbb{R}}$ and E. Moreover, since the complex modulus is consistent with the real one by definition, we will simply use the notation $|\cdot|$ from now on.

Definition 4.2.5 (Complex Banach lattices)**.**

(a) A **complex Banach lattice** E is a vector space complexification of a real Banach lattice $E_{\mathbb{R}}$ together with a complex modulus function $|\cdot| : E \to (E_{\mathbb{R}})_+$ and the norm $\|\cdot\|$ on E given by $\|z\| := \||z|\|_{E_{\mathbb{R}}}$.

The real Banach lattice $E_{\mathbb{R}}$ is called the **real part** of E.

[6] Unfortunately this paper is only available in German and it seems that its results have not appeared in a book. . . yet.

(b) By a **Banach lattice** we mean a real or complex Banach lattice.

It is straightforward to check that the norm on a complex Banach lattice is indeed a norm and that it is complete, i.e. a complex Banach lattice is indeed a Banach space. By the definition of a real Banach lattice one has $\|x\|_{E_\mathbb{R}} = \||x|\|_{E_\mathbb{R}} = \|x\|$ for every $x \in E_\mathbb{R}$, so we can actually denote the norms on E and $E_\mathbb{R}$ by the same symbol $\|\cdot\|$.

Remark 4.2.6 (Inequalities and the cone in a complex Banach lattice)**.** Let E be a complex Banach lattice. Its **positive cone** is defined as $E_+ := (E_\mathbb{R})_+$. If we write $x \le y$ or $y \ge x$ for two vectors $x, y \in E$, we mean tacitly that $x, y \in E_\mathbb{R}$.

Before we discuss some examples, let us note the following properties of the modulus in a complex Banach lattice.

Proposition 4.2.7. *Let E be a complex Banach lattice. Then one has $|\overline{z}| = |z|$ as well as $|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z|$ for all $z \in E$.*

Proof. Let $z \in E$. For all $\theta \in \mathbb{R}$ one has $\operatorname{Re}(e^{\mathrm{i}\theta}\overline{z}) = \operatorname{Re}\left(\overline{e^{-\mathrm{i}\theta} z}\right) = \operatorname{Re}(e^{-\mathrm{i}\theta} z)$, so it follows from the integral formula for $|z|$ in Theorem 4.2.4 that

$$|\overline{z}| = \frac{1}{4}\int_0^{2\pi} \left|\operatorname{Re}(e^{-\mathrm{i}\theta} z)\right| \,\mathrm{d}\theta = \frac{1}{4}\int_0^{2\pi} \left|\operatorname{Re}(e^{\mathrm{i}\varphi} z)\right| \,\mathrm{d}\varphi = |z|,$$

where the second equality uses the substitution $\varphi := 2\pi - \theta$. By using the axioms of the complex modulus (Definition 4.2.3) we thus get

$$2\,|\operatorname{Re} z| = |z + \overline{z}| \overset{\text{(II)}}{\le} |z| + |\overline{z}| = 2\,|z|,$$

and therefore, by axiom (III), $|\operatorname{Im} z| = |\operatorname{Re}(-\mathrm{i}z)| \le |-\mathrm{i}z| = |z|$. □

Examples 4.2.8.

(a) Let K be a compact metric (or topological) space. Endowed with the pointwise modulus function and the sup norm $\|\cdot\|_\infty$, $\mathrm{C}(K) = \mathrm{C}(K;\mathbb{C})$ is a complex Banach lattice with real part $\mathrm{C}(K;\mathbb{R})$.

Indeed, $\mathrm{C}(K)$ is the vector space complexification (as in Definition 4.2.1) of $\mathrm{C}(K;\mathbb{R})$ and the pointwise modulus function obviously satisfies the axioms from Definition 4.2.3. Finally, one clearly has $\|f\|_\infty = \||f|\|_\infty$.

(b) Let (Ω, μ) be a measure space and $1 \le p \le \infty$. Similarly as in (a) one can readily see that $L^p(\Omega,\mu) = L^p(\Omega,\mu;\mathbb{C})$ is a complex Banach lattice with real part $L^p(\Omega,\mu;\mathbb{R})$.

We close this section by introducing the following property of linear operators.

Definition 4.2.9 (Real operators)**.** Let E, F be complex Banach lattices. Given a linear operator $A\colon E \supseteq \operatorname{dom}(A) \to F$, set $\operatorname{dom}(A)_\mathbb{R} := \operatorname{dom}(A) \cap E_\mathbb{R}$. Then A is called **real** if

$$\operatorname{dom}(A) = \operatorname{dom}(A)_\mathbb{R} + \mathrm{i}\operatorname{dom}(A)_\mathbb{R} \quad \text{and} \quad A(\operatorname{dom}(A)_\mathbb{R}) \subseteq F_\mathbb{R}.$$

Observe that if A is everywhere defined, then A is real if and only if $A(E_\mathbb{R}) \subseteq F_\mathbb{R}$. Typical examples of unbounded real operators are differential operators with real coefficients on L^p-spaces with a Sobolev space as domain.

4.3 Positive operators

It follows from Proposition 1.1.3 that the following definition generalises positivity of matrices to linear operators.

Definition 4.3.1 (Positive operators)**.**

(a) A linear map $T\colon V \to W$ between two real vector lattices V and W is called **positive** if $TV_+ \subseteq W_+$.

(b) Similarly, a linear map $T\colon E \to F$ between two complex Banach lattices is called **positive** if $TE_+ \subseteq F_+$.

Note that part (a) of the definition includes operators between real Banach lattices, but it sometimes convenient to have the notion “positive operator” available in the more general case of vector lattices. We distinguished the cases (a) and (b) in the definition since we did not define the general concept of **complex vector lattices**.

Observe that a positive linear operator T preserves inequalities, i.e. if $x \le y$, then also $Tx \le Ty$. If the scalar field is complex, every positive operator is real, since the positive cone in a complex Banach lattice E spans the real part $E_{\mathbb{R}}$.

The following generalises Proposition 1.1.5 from matrices to positive operators.

Proposition 4.3.2. *Let E, F be Banach lattices over the same scalar field. For each positive linear operator $T \in \mathcal{L}(E,F)$ and each $x \in E$ one has $|Tx| \le T|x|$.*

Proof. First assume that the scalar field is $\mathbb{R}$. For each $x \in E$ it then follows from $\pm x \le |x|$ and the positivity of T that $\pm Tx \le T|x|$, so $|Tx| = Tx \vee (-Tx) \le T|x|$.

Now let the scalar field be $\mathbb{C}$. By applying the real case to the restriction $T|_{E_{\mathbb{R}}}$ one gets $|Tx| \le T|x|$ for every $x \in E_{\mathbb{R}}$. For the general case $x \in E$ we thus have

$$|Tx| = \frac{1}{4}\int_0^{2\pi} \left|\operatorname{Re}(e^{\mathrm{i}\theta}Tx)\right| \,\mathrm{d}\theta = \frac{1}{4}\int_0^{2\pi} \left|T\operatorname{Re}(e^{\mathrm{i}\theta}x)\right| \,\mathrm{d}\theta \le \frac{1}{4}\int_0^{2\pi} T\left|\operatorname{Re}(e^{\mathrm{i}\theta}x)\right| \,\mathrm{d}\theta = T|x|,$$

where the second equality holds since T is real, and the inequality in the middle uses the established inequality for vectors in $E_{\mathbb{R}}$. □

In the study of positive operators on Banach lattices, it is mandatory to present the following delightful result.

Theorem 4.3.3 (Automatic continuity of positive operators)**.** *Let E, F be Banach lattices over the same scalar field and let $T\colon E \to F$ be a positive linear map. Then T is continuous.*

Proof. We consider the real case; the complex case is then an easy consequence. Assume that T is not continuous. As every $x \in E$ is of the form $x = x^+ - x^-$, we can find a sequence (x_n) in E_+ such that $\|x_n\| \le 1$ and $\|Tx_n\| \ge n^3$ for all $n \in \mathbb{N}$. Observe that $x := \sum_{n=1}^{\infty} \frac{x_n}{n^2}$ is a well-defined element of E – as the series converges absolutely and E is complete – that satisfies $0 \le \frac{x_n}{n^2} \le x$ for all $n \in \mathbb{N}$. Thus, using the positivity of T and the lattice norm property, we find $\|Tx\| \ge \left\|T\left(\frac{x_n}{n^2}\right)\right\| \ge \frac{1}{n^2}\cdot n^3 = n$ for all $n \in \mathbb{N}$, a contradiction. □

4.4 Dual spaces of Banach lattices

Given a real Banach lattice E, we introduce a partial order $\le$ on E' by defining $\varphi \le \psi$ for two functionals if $\langle \varphi, x\rangle \le \langle \psi, x\rangle$ for all $x \in E_+$. Functionals satisfying $\varphi \ge 0$ are called **positive**; these are precisely the functionals positive in the operator sense (Definition 4.3.1). The set of all positive functionals in E' is denoted by E'_+ and is called the **dual cone of** E'_+.

Example 4.4.1 (The dual cone of $\mathbb{R}^n_+$)**.** Endow $E = \mathbb{R}^n$ with the standard order and the Euclidean norm and identify E' with the space $\mathbb{R}^{1\times n}$ of row vectors.

(a) A functional $\varphi \in E'$ is positive (in the sense defined above this example) if and only if all its entries are ≥ 0 when φ is viewed as an element of $\mathbb{R}^{1\times n}$.

Hence, the positive cone E'_+ agrees with the standard cone $\mathbb{R}^{1\times n}_+$. In particular, E' is a Banach lattice.

(b) For every $\varphi \in E'$ and every $x \in E_+$ one has $\langle |\varphi|, x\rangle = \max\{|\langle \varphi, y\rangle| : y \in \mathbb{R}^n, |y| \le x\}$.

Proof. (a) This follows immediately from the characterisation of positive matrices in terms of the action in Proposition 1.1.3.

(b) Let $\varphi \in E' = \mathbb{R}^{1\times n}$ and $x \in E_+ = \mathbb{R}^n_+$. Since the order on E' is the standard order, the lattice operations on it are given componentwise (see Example 4.1.4(a)).

To obtain the claimed equality, observe that for every $y \in \mathbb{R}^n$ with $|y| \le x$ one has $\langle |\varphi|, x\rangle = |\varphi| x \ge |\varphi||y| \ge |\varphi y| = |\langle \varphi, y\rangle|$, where the second inequality uses Proposition 1.1.5(a). On the other hand, if one chooses $y \in \mathbb{R}^n$ with $y_k = x_k$ if $\varphi_k \ge 0$ and $y_k = -x_k$ if $\varphi_k < 0$, then $|y| = x$ and $\langle |\varphi|, x\rangle = \langle \varphi, y\rangle$. ☐

The preceding example motivates the following general result about duals of Banach lattices. For a complex Banach lattice E with real part $E_{\mathbb{R}}$, observe that a functional $\varphi \in E'$ is real in the sense of Definition 4.2.9 if and only if $\varphi(E_{\mathbb{R}}) \subseteq \mathbb{R}$.

Theorem 4.4.2 (Duals of Banach lattices)**.**

(a) *If E is a real Banach lattice, then so is E'.*

(b) *Let E be a complex Banach lattice with real part $E_{\mathbb{R}}$. Then E' is a complex Banach lattice whose real part consists of the real functionals in E' and can be identified with $(E_{\mathbb{R}})'$ by restricting each such functional to $E_{\mathbb{R}}$.*

In both cases one has the ***Riesz–Kantorovich formula***

$$\langle |\varphi|, x\rangle = \sup\{|\langle \varphi, y\rangle| \,:\, y \in E,\ |y| \le x\} = \sup\{\operatorname{Re}\langle \varphi, y\rangle \,:\, y \in E,\ |y| \le x\}$$

for all $\varphi \in E'$ and $x \in E_+$, and the inequality $|\langle \varphi, z\rangle| \le \langle |\varphi|, |z|\rangle$ for all $\varphi \in E'$ and $z \in E$.

For the proof of Theorem 4.4.2 in the real case, we refer to the classical books on Banach lattices, e.g. [MN91, Theorem 1.3.2 and Proposition 1.3.7]. For the complex case see e.g. [MN91, Proposition 2.2.6] and the sentence thereafter, or [Sch74, Corollary 3 to Theorem IV.1.8].

Remark 4.4.3 (The dual cone of complex Banach lattice)**.** If E is a complex Banach lattice, then by Remark 4.2.6 and Theorem 4.4.2(b), the cone E'_+ consists of the real functionals φ whose restriction to $E_\mathbb{R}$ is positive. One readily checks that these are precisely all the $\varphi \in E'$ that are positive operators from E to $\mathbb{C}$ in the sense of Definition 4.3.1(b), so E'_+ consists precisely of the positive functionals in the complex case as well.

Finally, we discuss how positivity of operators behaves with respect to duality. To this end, we need the following consequence of the Hahn–Banach separation theorem.

Proposition 4.4.4. *Let E be a Banach lattice. An element $x \in E$ is positive if and only if $\langle \varphi, x \rangle \geq 0$ for all $\varphi \in E'_+$.*

Proof. *"⇒"*: This implication is obvious.

"⇐": Assume that $x \notin E_+$. Since the convex set E_+ is closed (Proposition 4.1.7(c)), the Hahn-Banach separation theorem (see e.g. [Bre11, Theorem 1.7]) shows that there exists a functional $\varphi \in E'$ that strictly separates $\{x\}$ from E_+, i.e. there exists $\alpha \in \mathbb{R}$ such that $\langle \varphi, x \rangle < \alpha < \langle \varphi, y \rangle$ for all $y \in E_+$. In particular, taking $y = 0$, we get $\alpha < 0$.

We show that φ is positive. To this end, fix an arbitrary $0 \neq y \in E_+$. Then $\lambda^{-1} y \in E_+$ for all $\lambda > 0$, and thus $\lambda\alpha < \langle \varphi, y \rangle$ for all $\lambda > 0$. As $\lambda \downarrow 0$, we obtain $0 \leq \langle \varphi, y \rangle$. Thus $\varphi \in E'_+$ with $\langle \varphi, x \rangle < 0$. □

Corollary 4.4.5. *Let E, F be Banach lattices over the same field and $T \in \mathcal{L}(E, F)$. Then T is positive if and only if T' is positive.*

Proof. The implication "⇒" follows immediately from the definitions, and the implication "⇐" is a consequence of Proposition 4.4.4. □

Exercises for Chapter 4

Exercise 4.1 (Getting acquainted with lattice operations)**.**

(a) Let V be a real vector lattice and let $x, y \in V$. Derive directly from the axioms (I) and (II) in Definition 4.1.1 and from the definition of the modulus in Definition 4.1.2 that the triangle inequality $\left|x+y\right| \le |x| + \left|y\right|$ holds for all $x, y \in V$ (i.e. prove the first part of Proposition 4.1.3(e)).

(b) Consider the real Banach lattice $C([0,1];\mathbb{R})$ and the functions $f_n \in C([0,1];\mathbb{R})$ given by $f_n(x) = x^{1/n}$ for each $n \in \mathbb{N}$ and all $x \in [0,1]$.

Show that the set $\{f_n : n \in \mathbb{N}\}$ has a supremum f in $C([0,1];\mathbb{R})$. Does the equality $f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ hold for all $x \in [0,1]$?

(c) Find a sequence (g_n) in $C([0,1];\mathbb{R})$ that satisfies $0 \le g_n \le \mathbb{1}$ for all n and such that the set $\{g_n : n \in \mathbb{N}\}$ does not have a supremum in the space $C([0,1];\mathbb{R})$.

Exercise 4.2 (C^1 is not a vector lattice)**.** Endow the space $C^1([-1,1];\mathbb{R})$ with the partial order $\le$ inherited from $C([-1,1];\mathbb{R})$, i.e. functions are compared pointwise. Note that this space satisfies property (I) in Definition 4.1.1.

(a) Starter: Show that $C^1([-1,1];\mathbb{R})$ is not a vector sublattice of $C([-1,1];\mathbb{R})$.

(b) Show that $C^1([-1,1];\mathbb{R})$ is not a vector lattice.

Exercise 4.3 (Positivity of the resolvent for Δ_{Dir})**.** Let $\emptyset \ne \Omega \subseteq \mathbb{R}^n$ be open and consider the Dirichlet Laplacian $\Delta_{\mathrm{Dir}} \colon L^2(\Omega) \supseteq \operatorname{dom}(\Delta_{\mathrm{Dir}}) \to L^2(\Omega)$ from Example 3.3.6 given by

$$\operatorname{dom}(\Delta_{\mathrm{Dir}}) := \left\{u \in H^1_0(\Omega) : \Delta u \in L^2(\Omega)\right\}, \qquad \Delta_{\mathrm{Dir}} u := \Delta u.$$

Let $\lambda \in (0,\infty)$. According to Example 3.3.6(c) one has $\lambda \in \rho(\Delta_{\mathrm{Dir}})$. In this exercise we show that the resolvent $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) \colon L^2(\Omega) \to L^2(\Omega)$ is a positive operator.

(a) Let $0 \le g \in L^2(\Omega)$ and set $u := \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) g \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$. Derive from the description of Δ_{Dir} in Example 3.3.6(b) that

$$\lambda\,(v \mid u)_{L^2} + (\nabla v \mid \nabla u)_{L^2} = \left(v \mid g\right)_{L^2} \tag{4.4.1}$$

for all $v \in H^1_0(\Omega)$. Are we allowed to substitute $v := u^-$ in this equation?

(b) Show that $u^- = 0$ and conclude that $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) \geq 0$.

Exercise 4.4 (Range of a positive projection)**.**

(a) Let V be a real vector lattice and $P \in \mathcal{L}(V)$ a positive projection.[7] Show that $\operatorname{rg} P$ is a vector lattice with respect to the order inherited from V and the corresponding cone $(\operatorname{rg} P)_+ := \{v \in \operatorname{rg} P : v \geq 0\}$ satisfies $(\operatorname{rg} P)_+ = \operatorname{rg} P \cap V_+ = P(V_+)$.

Give an explicit formula for the modulus in $\operatorname{rg} P$. Is $\operatorname{rg} P$ always a vector sublattice?

(b) Let E be a real Banach lattice and $P \in \mathcal{L}(E)$ a positive projection. Show that $\|x\|_{\operatorname{rg} P} := \|P|x|\|$ for all $x \in E$ defines an equivalent norm on $\operatorname{rg} P$ with respect to which $\operatorname{rg} P$ is a Banach lattice.

(c) Let E be a complex Banach lattice and $P \in \mathcal{L}(E)$ a positive projection. Observe that $\operatorname{rg} P$ is the vector space complexification of the Banach lattice $P(E_{\mathbb{R}})$ and thus becomes a complex Banach lattice by Definition 4.2.5.

Give an explicit formula for the complex modulus function $\operatorname{rg} P \to P(E_{\mathbb{R}})$ and for the norm that turns $\operatorname{rg} P$ into a complex Banach lattice.

(d) Let $A \subseteq \mathrm{C}([0,1];\mathbb{R})$ denote the two-dimensional subspace consisting of all affine functions endowed with the norm and the order inherited from $\mathrm{C}([0,1];\mathbb{R})$.

Show that A is not a vector sublattice of $\mathrm{C}([0,1];\mathbb{R})$ but is itself a Banach lattice.

Exercise 4.5. In this exercise, we will use some general notions from order theory. A **directed set** is a non-empty set J together with a reflexive and transitive relation $\leq$ such that for all $a, b \in J$, there exists $c \in J$ such that $a \leq c$ and $b \leq c$. A **net** in a set X is a function $x\colon J \to X$, where J is a directed set; it is customary to write $(x_j)_{j\in J}$ instead of x.[8]

Let E be a complex Banach lattice and let $(T_j)_{j\in J}$ be a net in the space $\mathcal{L}(E)$ of bounded linear operators on E. Assume that $(T_j)_{j\in J}$ is an **individually eventually positive net** in the sense that for each $f \in E_+$, there exists $j_0 \in J$ such that $T_j f \geq 0$ for all $j \geq j_0$.

(a) Show that for each $f \in E_{\mathbb{R}}$, there exists $j_1 \in J$ such that $\left|T_j f\right| \leq T_j \left|f\right|$ for all $j \geq j_1$.

(b) Assume now that there exists a countable set $I \subseteq J$ that is **majorising**, meaning that for each $j \in J$ there exists a $j_1 \in I$ that satisfies $j_1 \geq j$. Show that $(T_j)_{j\in J}$ is **uniformly eventually real** in the sense that there exists $j_0 \in J$ such that T_j is a real operator for all $j \in J$ with $j \geq j_0$.

Hint: Use Baire's theorem.

[7] Recall that a linear map is called a **projection** if it satisfies $P^2 = P$.

[8] Note that every sequence is a net with $J = \mathbb{N}$, endowed with its usual order.

Notes for Chapter 4

Vector lattices, Banach lattices, and ordered vector spaces

Standard literature

Vector lattices are also called **Riesz spaces** in the literature. The theory dates back to the first half of the twentieth century, but we refrain from trying to give a historical account of its development. Standard books on the topic include: the two volumes by Luxemburg and Zaanen [LZ71] and Zaanen [Zaa83] – the first of which focuses on vector lattices without norms while the latter has more material about Banach lattices; Schaefer's classical monograph [Sch74] on Banach lattices; the reprint [AB06] of a classical book by Aliprantis and Burkinshaw from the '80s; the book of Meyer–Nieberg [MN91], which might sometimes be a bit less heavy on notation and technicalities; and more recently, the introductory textbook [Zaa97] by Zaanen. We should also mention the book [Wnu99] by Wnuk which gives a comprehensive treatment of the important class of Banach lattices with **order continuous norm**.

Terminological remarks

As Section 4.1 shows, the theory of vector lattices is most naturally developed over the real field. It is thus customary in the literature to simply call "vector lattice" and "Banach lattice" what we call a "real vector lattice" and "real Banach lattice", respectively. Since we use the complex scalar field frequently and also discuss a number of results that are true over both scalar fields, we find it more natural to use the term "Banach lattice" for both cases simultaneously and specify the field whenever necessary.

Ordered vector spaces and ordered Banach spaces

Many classes of vector spaces and Banach spaces carry a canonical structure that does not turn them into vector lattices. Typical examples are spaces of differentiable functions like $C^k(\Omega)$ (see for instance Exercise 4.2), higher order Sobolev spaces [AN09, Example 2.3(d)], and the self-adjoint parts of non-commutative C^*-algebras; the latter are never lattice ordered by a classical result of Sherman [She51, Theorems 1 and 2]. Still, these spaces are often Banach spaces. Important results on such ordered Banach spaces can, for instance, be found in the first part of the classical paper [BR84] by Batty and

Robinson, in the book [KLS89] by Krasnosel'skii, Lifshits, and Sobolev, and [AT07, Section 2.5] of an excellent book by Aliprantis and Tourky. More recently, the German translation [Wul17] by Weber of two classical Russian books by Wulich appeared and contains some notes on more recent development of the field.

A very recent topic is the study of ordered vector spaces by embedding them into vector lattices – this is the theory of **pre-Riesz spaces** which is presented in the book [KvG19] by Kalauch and van Gaans.

As pointed out in Example 4.1.4(d), the Stampacchia lemma implies that the first order Sobolev spaces $W^{1,p}(\Omega;\mathbb{R})$ are vector lattices. However, they are not Banach lattices according to Example 4.1.6(d), and the same argument – considering quickly oscillating functions – shows that they do not become Banach lattices with respect to any equivalent norm. However, they are vector lattices and Banach spaces at the same time and their positive cone is closed. It seems that, from an abstract point of view, such "lattice-ordered Banach spaces" have not been studied systematically in the literature, but some interesting results about them were proved by Borwein and Yost in [BY84].

Complexification

Existence of a complex modulus function

Complexifications of Banach lattices go back to Lotz who developed them to study the spectrum of positive operators [Lot68]. His approach was based on the formula (4.2.1) for complex numbers: For an element z of the vector space complexification $E_{\mathbb{C}}$ of a real Banach lattice E he defined the complex modulus as

$$|z|_{\mathbb{C}} := \sup_{\theta\in[0,2\pi]} \operatorname{Re}(e^{\mathrm{i}\theta} z) \tag{4.4.2}$$

and then showed that it has the properties (I)–(III) stipulated in Definition 4.2.3. The approach still seems to be the most common one in the literature. Its main obstacle is that it is not at all clear why this supremum exists – note that the definition of a vector lattice only requires the existence of the supremum of sets of two elements (and hence of finite sets). The most common way to solve this problem is to use a representation result due to Kakutani about a class of Banach lattices, called **AM-spaces**. This result enables reducing the situation to spaces of continuous functions where one can explicitly check that the supremum exists. We refer for instance to [Sch74, Section II.11] for the details of this procedure. A different approach is used in [Zaa97, Chapter 6] where the existence of the supremum is proved intrinsically – i.e. without employing a representation result – by using a certain type of convergence that is often called **relative uniform convergence**.

By contrast, our existence proof for a complex modulus function relies on the integral formula (4.2.2). To the best of our knowledge, this formula has not been used in the construction of complex vector lattices so far – which is remarkable since it actually gives existence under more general assumptions than the more common procedures which require the vector lattice to be **relatively uniformly complete**.[9]

While formula (4.2.2) does not seem to have found its way into the standard literature on Banach lattices, representing the modulus in this way is known to be a useful technique in related parts of functional analysis. For instance, consider a bounded linear operator T on a real-valued L^p-space $L^p(\Omega,\mu;\mathbb{R})$ and let $T_\mathbb{C}$ denote its canonical extension to the complex-valued space $L^p(\Omega,\mu)$. Then we can prove that $\|T\| = \|T_\mathbb{C}\|$. This is not obvious at all, but it can be proved by using an L^p-version of the integral representation of the complex modulus, see e.g. [Fen98, Proposition 2.1.1 and Remark 2.1.1]

Uniqueness of the complex modulus function

Uniqueness of the complex modulus function was proved by Mittelmeyer and Wolff in [MW74, Theorem 2.2]. It is valid under much weaker assumptions than existence results – one only needs the underlying real vector lattice to be **Archimedean**, a property that is typically satisfied in non-pathological examples. Remarkably, the theory of complex Banach lattices can essentially be developed without discussing uniqueness. Indeed, the standard books on Banach lattices define the complex modulus by the formula in (4.4.2) (or a version thereof) and then derive the property that one would expect from a modulus function without ever discussing uniqueness. From this perspective, one could analogously define the theory of complex Banach lattices by defining the complex modulus in terms of the integral formula in Theorem 4.2.4 without bothering about uniqueness.

Apart from purely theoretical interest, the one point where knowledge of the uniqueness of the complex modulus function is quite handy, is to check that concrete examples of complex functions spaces are indeed complex Banach lattices. The efficiency of this approach is demonstrated in Examples 4.2.8. Without the uniqueness result, showing that the common function spaces are indeed complex Banach lattices requires a bit more work. For L^p-spaces one can for instance do this by showing that it suffices to take the supremum in $\sup_{\theta\in[0,2\pi]} \operatorname{Re}(e^{i\theta}z)$ over a countable subset of $[0,2\pi]$ and then by working pointwise almost everywhere. Alternatively, if one takes our integral formula approach, it can be shown that L^p-spaces are complex Banach lattices by using results about the almost-everywhere evaluation of L^p-valued Bochner integrals (see e.g. [HvNVW16, Proposition 1.2.25 and Remark 1.2.26]).

More on complexifications

The norm on a complex Banach lattice is closely related to tensor products of Banach spaces as shown in [vN97].

Complexifications can be defined not only for Banach lattices, but also make sense for real Banach spaces. In this case one does, of course, not define a complex modulus function, but focuses on ways to extend the norm on a real Banach space to its vector space complexification. We refer to the paper [MST99] for an overview.

[9] Note that we have not presented the results in the main text under optimal assumptions by any means, to keep the presentation concise and accessible and in order to focus on the key ideas.

Finally, we point out that our definition of the complexification of real vector spaces (Definition 4.2.1) is rather explicit and not axiomatic. This makes the definition itself rather straightforward to digest, but it has the disadvantage that it requires certain identifications when one considers a vector space that is not explicitly given as $V \times V$ as the complexification of the real vector space V. Such identifications already occur for instance in Examples 4.2.8 and they become more explicit in Exercise 4.4(c) and, in particular, in Theorem 4.4.2(b). For a presentation of complexifications that is more axiomatic and thus makes all identifications explicit, we refer to [Glü16, Appendix C].

The Riesz–Kantorovich formula(s)

The Riesz–Kantorovich formula from Theorem 4.4.2 also holds more generally for so-called **regular** operators between a Banach lattice E and a **Dedekind complete** Banach lattice F. For the case of real scalars, see e.g. [AB06, Theorem 1.18]; there, one can also find similar formulas for the positive and the negative part of a regular operator. For the complex case, we refer to [Sch74, Definition IV.1.7 and Theorem IV.1.8].

Appendices

4.A Vector-valued integrals

In this appendix, we present some essential aspects of integration theory for functions with values in a Banach space. A good overview of elementary concepts can be found in [AE01, Kapitel X.1 & X.2], and [HvNVW16, Chapter 1] offers a comprehensive introduction. The stochastically-minded reader can also consult [DPZ14, Section 1.1].

To keep things simple, we only consider σ-finite measure spaces throughout this section, although much of the theory can also be developed in more generality. Recall that a measure space $(\Omega, \mathscr{A}, \mu)$ is called **σ-finite** if there exist countably many measurable sets $A_n \in \mathscr{A}$ with $\mu(A_n) < \infty$ and $\bigcup_{n\in\mathbb{N}} A_n = \Omega$.

Measurability

As a first reaction, one might think: "what's the big deal?" If $(\Omega, \mathscr{A})$ is a measurable space and X is a Banach space, then surely it makes sense to say that a function $f\colon \Omega \to X$ is measurable if and only if $f^{-1}(B) \in \mathscr{A}$ for every Borel subset B in X. It turns out that this natural definition is not as useful as it looks. Often in functional analysis, the norm dual X' is used to study problems on X by reducing to the case of scalar-valued functions. This is supported by the following simple result.

Proposition 4.A.1. *Let X be a separable Banach space, and let $\mathscr{A}$ be the σ-algebra generated by all sets of the form*

$$\{x \in X : \langle x', x\rangle \le \alpha\}, \quad x' \in X', \alpha \in \mathbb{R}. \tag{4.A.1}$$

Then $\mathscr{A}$ coincides with the Borel σ-algebra $\mathscr{B}(X)$.

Proof. The inclusion $\mathscr{A} \subseteq \mathscr{B}(X)$ is immediate, since each set of the form (4.A.1) is closed and hence in $\mathscr{B}(X)$. To prove the converse, it suffices to show that $\mathscr{A}$ contains all open balls in X. Since X is separable, there is a countable norming sequence $(x'_n) \subset X'$, i.e.

$$\|x\| = \sup_{n\in\mathbb{N}} \left|\langle x'_n, x\rangle\right| \quad \forall\, x \in X.$$

Consequently, for arbitrary $x \in X$ and $r > 0$, it holds that

$$B_{<r}(x) = \bigcup_{m=1}^{\infty} B_{\le r(1-1/m)}(x) = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \left\{y \in X : \left|\langle x'_n, y - x\rangle\right| \le r\left(1 - \tfrac{1}{m}\right)\right\} \in \mathscr{A}. \qquad \square$$

However, if the Banach space X is not separable, then it may happen that the σ-algebra generated by X' is strictly smaller than the Borel σ-algebra $\mathscr{B}(X)$ (so in some sense $\mathscr{B}(X)$ is 'too large'). In this case, the measurability properties on X cannot be effectively determined via linear functionals.

On the other hand, we often study measurability and integrability of scalar-valued functions via approximation by simple functions. These considerations motivate the following.

Definition 4.A.2. Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space and X a Banach space.

(a) A function $\varphi\colon \Omega \to X$ is called **μ-simple** if it satisfies: (I) $\varphi(\Omega)$ is a finite subset of X; (II) $\varphi^{-1}(\{x\}) \in \mathscr{A}$ for all $x \in X$; and (III) $\mu\big(\varphi^{-1}(X \setminus \{0\})\big) < \infty$.

(b) A function $f\colon \Omega \to X$ is called **strongly μ-measurable** if there is a sequence $(f_n)_{n\in\mathbb{N}}$ of μ-simple functions such that $\lim_{n\in\mathbb{N}} f_n(\omega) = f(\omega)$ for μ-almost every $\omega \in \Omega$.

   In the case $X = \mathbb{C}$, it is common to omit the word 'strongly' and simply speak of μ-measurable functions $f\colon \Omega \to \mathbb{C}$.

(c) A function $f\colon \Omega \to X$ is called **weakly μ-measurable** if the $\mathbb{C}$-valued map $x' \circ f = \langle x', f(\cdot)\rangle$ is μ-measurable for all $x' \in X'$.

Remark 4.A.3. Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space and X a Banach space. Every μ-simple function $\varphi\colon \Omega \to X$ can be written uniquely in the form

$$\varphi = \sum_{k=1}^{N} \mathbb{1}_{A_k} x_k \tag{4.A.2}$$

with distinct elements $\{x_1, \dots, x_N\} \subset X$, and where the sets $A_k \in \mathscr{A}$ satisfy $\mu(A_k) < \infty$ for each $k = 1, \dots, N$ and are pairwise disjoint, i.e. $A_j \cap A_k = \emptyset$ whenever $j \neq k$. Indeed, let $\{x_1, \dots, x_N\} \subset X$ be the disinct non-zero values of φ, and set $A_k := \varphi^{-1}(\{x_k\})$. Clearly the A_k's are pairwise disjoint and $\mu(A_k) < \infty$ for each $k = 1, \dots, N$. The uniqueness of the representation (4.A.2) can then be easily verified.

Definition 4.A.2 does not require the Banach space X to be separable. This is important, since non-separable Banach spaces (such as $L^\infty(\Omega)$ for an open subset $\emptyset \neq \Omega \subseteq \mathbb{R}^n$) are also useful and appear in many applications. Nevertheless, Proposition 4.A.1 gives a hint that if we want a good integration theory, separability should not be far away.

Definition 4.A.4. Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space and X a Banach space. A function $f\colon \Omega \to X$ is called **μ-almost separably valued** if there exists a separable subset[10] X_0 of X such that $f(\omega) \in X_0$ for μ-almost every $\omega \in \Omega$.

We can now formulate the fundamental theorem of Pettis.

[10] Naturally, we equip X_0 with the relative topology inherited from X. Note that some texts require X_0 to be a closed, separable vector subspace of X. However, it can be shown that the two definitions are equivalent.

Theorem 4.A.5 (Pettis)**.** *Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space and X a Banach space. For a function $f\colon \Omega \to X$, the following assertions are equivalent:*

(i) *f is strongly μ-measurable;*

(ii) *f is μ-almost separably valued and weakly μ-measurable.*

For the proof, we refer to [HvNVW16, Theorem 1.1.20].

The Bochner integral

For strongly μ-measurable functions with values in a Banach space, one can define an integral as follows.

Definition 4.A.6 (The Bochner integral)**.** Let $(\Omega, \mathscr{A}, \mu)$ be a σ-finite measure space and X a Banach space.

(a) For a μ-simple function $\varphi\colon \Omega \to X$, $\varphi = \sum_{k=1}^{N} \mathbb{1}_{A_k} x_k$, we define

$$\int_\Omega \varphi \,\mathrm{d}\mu := \sum_{k=1}^{N} \mu(A_k) x_k. \tag{4.A.3}$$

(b) Let $f\colon \Omega \to X$ be a strongly μ-measurable function. We say that f is **Bochner integrable** (with respect to μ) if there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of μ-simple functions such that

$$\lim_{n\to\infty} \int_\Omega \left\| f_n - f \right\| \,\mathrm{d}\mu = 0.$$

In this case, we define

$$\int_\Omega f \,\mathrm{d}\mu := \lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu. \tag{4.A.4}$$

Remarks 4.A.7.

(a) By the uniqueness assertion in Remark 4.A.3, the integral for simple functions (4.A.3) is well-defined. Moreover, it is clear that $\left\| \int_\Omega \varphi \,\mathrm{d}\mu \right\| \le \int_\Omega \left\| \varphi \right\| \,\mathrm{d}\mu$, and it is straightforward to verify that additivity $\int_\Omega \varphi \,\mathrm{d}\mu + \int_\Omega \psi \,\mathrm{d}\mu = \int_\Omega \varphi + \psi \,\mathrm{d}\mu$ holds for all simple functions $\varphi, \psi\colon \Omega \to X$ (cf. [AE01, Kapitel X.2, Bemerkungen 2.1] for details).

(b) To see that (4.A.4) is well-defined, first note that the map $\Omega \ni \omega \mapsto \left\| f_n(\omega) - f(\omega) \right\| \in [0,\infty)$ is μ-measurable, since it is a composition of the strongly μ-measurable function $f_n - f$ with the continuous function $\|\cdot\|\colon X \to [0,\infty)$. By the properties of the integral for simple functions mentioned above, one has

$$\left\| \int_\Omega f_n \,\mathrm{d}\mu - \int_\Omega f_m \,\mathrm{d}\mu \right\| \le \int_\Omega \left\| f_n - f_m \right\| \,\mathrm{d}\mu \le \int_\Omega \left\| f_n - f \right\| \,\mathrm{d}\mu + \int_\Omega \left\| f_m - f \right\| \,\mathrm{d}\mu$$

for all $m, n \in \mathbb{N}$. Hence if f is Bochner integrable, then the integrals $\int_\Omega f_n \,\mathrm{d}\mu$ form a Cauchy sequence in X, and thus converge to a unique element.

Theorem 4.A.8 (Bochner)**.** *Let* $(\Omega, \mathscr{A}, \mu)$ *be a* σ*-finite measure space and* X *a Banach space. A strongly* μ*-measurable function* $f\colon \Omega \to X$ *is Bochner integrable if and only if*

$$\int_\Omega \|f\| \, \mathrm{d}\mu < \infty.$$

In this case, it holds that

$$\left\| \int_\Omega f \, \mathrm{d}\mu \right\| \le \int_\Omega \|f\| \, \mathrm{d}\mu.$$

We refer to [HvNVW16, Proposition 1.2.2] for the proof. In the case that the measure space is an interval $I \subseteq \mathbb{R}$ with the Lebesgue measure, proofs of Theorems 4.A.5 and 4.A.8 can also be found in [ABHN11, Theorem 1.1.1] and [ABHN11, 1.1.4] respectively.

Example 4.A.9. Let λ^n be the Lebesgue measure on $\mathbb{R}^n$, $K \subset \mathbb{R}^n$ a non-empty compact subset, and X a Banach space. Then every continuous function $f\colon K \to X$ is strongly λ^n-measurable. Indeed, the image $f(K)$ is a compact subset of X, and hence separable, so in particular f is λ^n-almost separably valued. (More generally, if $\Omega \subseteq \mathbb{R}^n$ can be covered by countably many compact sets, then every continuous function $f\colon \Omega \to X$ is separably valued). For each $x' \in X'$, the $\mathbb{C}$-valued map $x' \circ f$ is continuous (as a composition of continuous maps) and hence λ^n-measurable. The conclusion now follows from Theorem 4.A.5.

Moreover, Theorem 4.A.8 shows that every continuous function $f\colon K \to X$ is also Bochner integrable, since one has

$$\int_K \|f(\omega)\| \, \mathrm{d}\lambda^n \le \sup_{\omega \in K} \|f(\omega)\| \cdot \lambda^n(K) < \infty.$$

Finally, we present a useful result that explains how closed linear operators interact with Bochner integrals.

Theorem 4.A.10 (Hille)**.** *Let* $(\Omega, \mathscr{A}, \mu)$ *be a* σ*-finite measure space and* X, Y *Banach spaces. Suppose that* $f\colon \Omega \to X$ *is Bochner integrable, and* $A\colon X \supseteq \operatorname{dom}(A) \to Y$ *is a closed linear operator. Assume that* $f(\omega) \in \operatorname{dom}(A)$ *for* μ*-almost every* $\omega \in \Omega$ *and that* $A \circ f\colon \Omega \to Y$ *(defined for* μ*-almost every* $\omega \in \Omega$*) is Bochner integrable. Then* f *is Bochner integrable as a* $\operatorname{dom}(A)$*-valued function,* $\int_\Omega f \, \mathrm{d}\mu \in \operatorname{dom}(A)$*, and*

$$A \int_\Omega f \, \mathrm{d}\mu = \int_\Omega A \circ f \, \mathrm{d}\mu. \tag{4.A.5}$$

We refer to [HvNVW16, Theorem 1.2.4] for the proof. Observe that if $A \in \mathcal{L}(X, Y)$, then (4.A.5) can be derived easily from Definition 4.A.6 alone.

Encore: if you want to know more…

4.B The lattice structure of $W^{1,p}$

In Example 4.1.4(d), it was mentioned that the Sobolev spaces $W^{1,p}(\Omega)$ are vector lattices. In this appendix, we give a proof of this fact, which is a consequence of Theorem 4.B.3 due to Stampacchia. This is not merely a curiosity for vector lattice theory, but it turns out to be a very useful property for PDE theory in general.

As preparation, we require a Sobolev space version of the chain rule (which is also a very useful tool in itself).

Proposition 4.B.1 (A chain rule for $W^{1,p}$)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open, and $u \in W^{1,p}(\Omega)$ for some $p \in [1,\infty]$. Let $G \in \mathrm{C}^1(\mathbb{R};\mathbb{R})$ satisfy $G' \in L^\infty(\mathbb{R})$ and $G(0) = 0$. Then the composition $G \circ u$ belongs to $W^{1,p}(\Omega)$ and*

$$\partial_k(G \circ u) = (G' \circ u)\partial_k u, \qquad k = 1, \dots, n. \tag{4.B.1}$$

Proof. The assumptions on G and the mean value theorem imply that $|G(x)| \leq M|x|$ for all $x \in \mathbb{R}$, with $M = \|G'\|_{L^\infty(\mathbb{R})}$. Hence $|G \circ u| \leq M|u|$, which shows that $G \circ u \in L^p(\Omega)$. Since $G' \in L^\infty(\mathbb{R})$, we also have $(G' \circ u)\partial_k u \in L^p(\Omega)$ for all $k \in \{1, \dots, n\}$.

We treat the case $p < \infty$ first; this allows us to leverage the Meyers-Serrin theorem (Theorem 3.A.4), and choose a sequence $(u_n) \in \mathrm{C}^\infty(\Omega) \cap W^{1,p}(\Omega)$ such that $u_n \to u$ in $W^{1,p}(\Omega)$. Hence, by the classical integration by parts and chain rule, we obtain

$$\int_\Omega (G \circ u_n)\partial_k \varphi \,\mathrm{d}x = -\int_\Omega (G' \circ u_n)(\partial_k u_n)\varphi \,\mathrm{d}x \tag{4.B.2}$$

for all $n \in \mathbb{N}$ and all $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$. The inequality $|G \circ u_n - G \circ u| \leq M|u_n - u|$ implies that $G \circ u_n \to G \circ u$ in $L^p(\Omega)$. Furthermore, since G' is continuous, $G' \circ u_n$ converges pointwise a.e. to $G' \circ u$, and hence $(G' \circ u_n)\partial_k u_n \to (G' \circ u)\partial_k u$ in $L^p(\Omega)$ by dominated convergence. By letting $n \to \infty$ in (4.B.2), we therefore obtain

$$\int_\Omega (G \circ u)\partial_k \varphi \,\mathrm{d}x = -\int_\Omega (G' \circ u)(\partial_k u)\varphi \,\mathrm{d}x$$

for all $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$. This proves (4.B.1) in the case $p \neq \infty$.

The case $p = \infty$ is handled via a simple trick. Given $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$, choose a bounded subset Ω' such that $\operatorname{supp}\varphi \subseteq \Omega'$ and $\overline{\Omega'} \subseteq \Omega$. Then it follows that $u \in W^{1,p}(\Omega')$ for any $1 \leq p < \infty$, and the previous arguments can be applied. □

Remark 4.B.2. The assumption that $G(0) = 0$ in Proposition 4.B.1 is only needed to obtain $G \circ u \in L^p(\Omega)$ (as the proof readily reveals). If one is not concerned about p-integrability, this assumption on G can be dropped; cf. [GT01, Lemma 7.5].

Theorem 4.B.3 (Stampacchia)**.** *Let $\emptyset \neq \Omega \subset \mathbb{R}^n$ be open and $p \in [1,\infty]$. If $u \in W^{1,p}(\Omega)$, then $u^+, u^-, |u| \in W^{1,p}(\Omega)$ with*

$$\partial_k(u^+) = \begin{cases} \partial_k u & \text{on } u > 0 \\ 0 & \text{on } u \leq 0, \end{cases} \qquad \partial_k(u^-) = \begin{cases} 0 & \text{on } u \geq 0 \\ -\partial_k u & \text{on } u < 0, \end{cases} \qquad \partial_k |u| = \begin{cases} \partial_k u & \text{on } u > 0 \\ 0 & \text{on } u = 0 \\ -\partial_k u & \text{on } u < 0, \end{cases}$$

for all $k \in \{1, \ldots, d\}$. In particular, $W^{1,p}(\Omega)$ is a vector sublattice of $L^p(\Omega)$.

Proof. For every $\varepsilon > 0$, we define a smooth approximation of the function $x \mapsto x^+$ by

$$G_\varepsilon(x) := \begin{cases} (x^2 + \varepsilon^2)^{1/2} - \varepsilon & x > 0 \\ 0 & x \leq 0. \end{cases}$$

We use Proposition 4.B.1 and then integrate by parts to find

$$\int_\Omega (G_\varepsilon \circ u) \partial_k \varphi \, \mathrm{d}x = -\int_\Omega \frac{u \partial_k u}{(u^2 + \varepsilon^2)^{1/2}} \varphi \, \mathrm{d}x = -\int_{\{u>0\}} \frac{u \partial_k u}{(u^2 + \varepsilon^2)^{1/2}} \varphi \, \mathrm{d}x$$

for all $\varphi \in \mathrm{C}_\mathrm{c}^\infty(\Omega)$. Since $G_\varepsilon \circ u \to u^+$ pointwise a.e. in Ω and $u(u^2 + \varepsilon^2)^{-1/2} \to 1$ pointwise a.e. in the subset $\{u > 0\}$ as $\varepsilon \downarrow 0$, the above equalities in the limit yield

$$\int_\Omega u^+ \partial_k \varphi \, \mathrm{d}x = -\int_{\{u>0\}} (\partial_k u) \varphi \, \mathrm{d}x.$$

This proves the required formula for $\partial_k(u^+)$.

The formulas for $\partial_k(u^-)$ and $\partial_k |u|$ now follow immediately from the identities $u^- = (-u)^+$ and $|u| = u^+ + u^-$. The lattice operations of supremum and infimum, naturally inherited from $L^p(\Omega)$, can then be expressed in terms of the modulus, as shown in Proposition 4.1.3(f). This proves that $W^{1,p}(\Omega)$ is a vector sublattice of $L^p(\Omega)$. □

Remark 4.B.4. One can show that, for $p \in [1,\infty)$, the lattice operations $u \mapsto u^+$, $u \mapsto u^-$ and $u \mapsto |u|$ are continuous on $W^{1,p}(\Omega)$ (this does not follow from Proposition 4.1.7, which only yields continuity on $L^p(\Omega)$). The curious reader may consult, for instance, [AU23, Theorem 6.37], for the proof when $p = 2$ that can be easily adapted for $p \in [1,\infty)$. Alternatively, see [BY84, Example 2 and the paragraph thereafter].

However, the situation is more complicated on $W^{1,\infty}$. This space coincides with the space of bounded Lipschitz continuous functions, see e.g. [Eva10, Theorem 4, Section 5.8]. The lattice operations on it are continuous at the point 0, but one can easily construct examples to see that they are not continuous on the whole space.

Chapter 5

Positive and eventually positive solutions to PDEs

5.1 Positivity via sesquilinear forms

In Example 3.3.6(b), we described the domain and the action of the Dirichlet Laplacian Δ_{Dir} in terms of the expression $\int_\Omega \overline{\nabla v} \cdot \nabla u \, \mathrm{d}x$, defined for elements $u, v \in H_0^1(\Omega)$. This is a special case of a *sesquilinear form*. It is the purpose of this section to develop some key aspects in the theory of sesquilinear forms, which yield useful tools to study linear operators on Hilbert spaces. In addition, these so-called *form methods* are well-suited to the study of positivity, as shown in Theorem 5.1.7 below.

Definition 5.1.1 (Sesquilinear forms)**.** Let V be a complex Hilbert space.

(a) A map $\mathfrak{a}\colon V \times V \to \mathbb{C}$ is called a **sesquilinear form** if it is antilinear in the first component[1] and linear in the second.

(b) A sesquilinear form $\mathfrak{a}\colon V \times V \to \mathbb{C}$ is called **bounded** if there exists a number $c \geq 0$ such that $|\mathfrak{a}(v, w)| \leq c \, \|v\|_V \, \|w\|_V$ for all $v, w \in V$.

A fundamental result is that bounded sesquilinear forms can be represented by bounded linear operators.

Lemma 5.1.2 (Lax-Milgram)**.** *Let V be a complex Hilbert space and let $\mathfrak{a}\colon V \times V \to \mathbb{C}$ be a bounded sesquilinear form.*

(a) *There exists a unique $A \in \mathcal{L}(V)$ that satisfies $\mathfrak{a}(w, v) = -(w \mid Av)_V$ for all $v, w \in V$.*[2]

(b) *Moreover, if $\operatorname{Re} \mathfrak{a}(v, v) \geq \delta \, \|v\|_V^2$ for a number $\delta > 0$ and all $v \in V$, then A is bijective.*

[1] This is consistent with our convention for inner products.

[2] We put a minus sign here to be consistent with Definition 5.1.3 below.

Proof. (a) The uniqueness is clear. For existence, note that for each $v \in V$, $-\mathfrak{a}(\cdot, v)$ is a bounded antilinear functional on V. By the Riesz–Fréchet theorem, there exists $Av \in V$ such that $-\mathfrak{a}(w, v) = (w \mid Av)_V$ for all $v, w \in V$. Clearly, $v \mapsto Av$ is linear and

$$|(w \mid Av)_V| = |\mathfrak{a}(w, v)| \le c \, \|w\|_V \, \|v\|_V \quad \text{for all } w \in V$$

for some constant $c \ge 0$. Hence, $\|Av\|_V \le c\,\|v\|_V$ for all $v \in V$, so A is bounded.

(b) We may assume that $V \ne \{0\}$. So, boundedness of A ensures that $\sigma(A)$ is non-empty and compact. Let $\lambda \in \sigma(A)$ with maximal real part. It suffices to show that $\operatorname{Re}\lambda < 0$, since this implies $0 \notin \sigma(A)$. As λ lies on the boundary of $\sigma(A)$, there exists an **approximate eigenvector** of A for λ, i.e. a normalised sequence (v_n) in V such that $\lambda v_n - Av_n \to 0$ (see Exercise 5.2). It follows that $\operatorname{Re}\lambda \le -\delta$, since

$$0 \leftarrow \operatorname{Re}(\lambda v_n - Av_n \mid v_n)_V = \operatorname{Re}\lambda + \operatorname{Re}\mathfrak{a}(v_n, v_n) \ge \operatorname{Re}\lambda + \delta. \qquad \square$$

For the Dirichlet Laplacian, given $u, f \in L^2(\Omega)$, we have $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$ and $\Delta_{\mathrm{Dir}} u = f$ if and only if $u \in H_0^1(\Omega)$ and $\mathfrak{a}(v, u) := (\nabla v \mid \nabla u)_{L^2} = -\big(v \mid f\big)_{L^2(\Omega)}$ for all $v \in H_0^1(\Omega)$ (Example 3.3.6(b)). This observation serves as a blueprint for a general way to obtain operators from forms.

Definition 5.1.3 (Operators induced by forms)**.** Let V, H be complex Hilbert spaces such that V embeds continuously and densely into H. Let $\mathfrak{a}\colon V \times V \to \mathbb{C}$ be a bounded sesquilinear form on V. The **operator associated to** $\mathfrak{a}$, $A\colon H \supseteq \operatorname{dom}(A) \to H$, is given by

$$\begin{aligned} \operatorname{dom}(A) &:= \big\{u \in V : \exists f \in H,\ \forall v \in V\colon \mathfrak{a}(v, u) = -\big(v \mid f\big)_H\big\}, \\ Au &:= f, \end{aligned}$$

where $f \in H$ in the definition of Au is the vector that occurs in the definition of $\operatorname{dom}(A)$.[3]

There are two differences between Lemma 5.1.2 and Definition 5.1.3. In the lemma the inner product used to defined A is $(\cdot \mid \cdot)_V$ while it is $(\cdot \mid \cdot)_H$ in the definition. Consequently, the operator A in the lemma maps from V to V, while it maps from the smaller space $\operatorname{dom}(A)$ to the larger space H in the definition. We now show that the operator A from Definition 5.1.3 is quite well-behaved if $\mathfrak{a}$ satisfies a so-called ellipticity estimate.

Theorem 5.1.4 (Properties of operators induced by forms)**.** *Let V, H be complex Hilbert spaces such that V embeds continuously and densely into H. Let $\mathfrak{a}\colon V \times V \to \mathbb{C}$ be a bounded sesquilinear form on V. Assume that $\mathfrak{a}$ satisfies the* ***ellipticity estimate***

$$\operatorname{Re}\mathfrak{a}(v, v) + \mu\,\|v\|_H^2 \ge \delta\,\|v\|_V^2 \tag{5.1.1}$$

for some numbers $\mu \in \mathbb{R}$ and $\delta > 0$ and for all $v \in V$. Then the operator $A\colon H \supseteq \operatorname{dom}(A) \to H$ associated to $\mathfrak{a}$ has the following properties:

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

[3] Observe that f is uniquely determined since V is dense in H. Thus, A is indeed well-defined.

(b) *One has* $\mathrm{s}(A) \le \mu$ *and every* $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > \mu$ *satisfies* $\|\mathcal{R}(\lambda, A)\|_{H \leftarrow H} \le \frac{1}{\operatorname{Re}\lambda - \mu}$.

(c) *If the form* $\mathfrak{a}$ *is* ***symmetric****, i.e.* $\mathfrak{a}(u,v) = \overline{\mathfrak{a}(v,u)}$ *for all* $u, v \in V$*, then* $\sigma(A) \subseteq \mathbb{R}$.[4]

Proof. (a) Let $(u_n) \subseteq \operatorname{dom}(A)$ converge in H to $u \in H$ and assume (Au_n) converges to $v \in H$. Since $\mathfrak{a}(u_n - u_m, u_n - u_m) = (u_n - u_m \mid Au_m - Au_n)_H$, it follows from the ellipticity estimate (5.1.1) that (u_n) is Cauchy in V, hence convergent in V. The embedding $V \hookrightarrow H$ yields $u \in V$. The boundedness of $\mathfrak{a}$ and the convergence $u_n \to u$ in V imply $\mathfrak{a}(w,u) = \lim_{n\to\infty} \mathfrak{a}(w, u_n) = -(w \mid v)_H$ for every $w \in V$. Hence, $u \in \operatorname{dom}(A)$ and $Au = v$, so A is indeed closed.

We now show (b). The density of $\operatorname{dom}(A)$ then follows from Exercise 5.1.

(b) Let $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge \mu$. We first show that $\lambda - A \colon \operatorname{dom}(A) \to H$ is surjective, so let $f \in H$. Since $\big(\cdot \mid f\big)_H$ defines a bounded antilinear form on V, the Riesz–Fréchet representation theorem gives a vector $v_0 \in V$ such that $(u \mid v_0)_V = \big(u \mid f\big)_H$ for all $u \in V$. Now consider the bounded sesquilinear form $\mathfrak{b}\colon V \times V \to \mathbb{C}$, defined by $\mathfrak{b}(u,v) := \lambda\,(u \mid v)_H + \mathfrak{a}(u,v)$, which satisfies $\operatorname{Re}\mathfrak{b}(v,v) \ge \delta\,\|v\|_V^2$ for all $v \in V$, and let $B \in \mathcal{L}(V)$ be the associated bijective operator (Lemma 5.1.2(b)). Then $w_0 := -B^{-1} v_0 \in V$ satisfies

$$\mathfrak{a}(u, w_0) + \lambda\,(u \mid w_0)_H = \mathfrak{b}(u, w_0) = (u \mid v_0)_V = \big(u \mid f\big)_H$$

for all $u \in V$. Hence, $w_0 \in \operatorname{dom}(A)$ and $(\lambda - A)w_0 = f$, so $\lambda - A$ is indeed surjective.

On the other hand, $\lambda - A$ is injective, hence bijective, because

$$\begin{aligned} \|v\|_H \, \|(\lambda - A)v\|_H &\ge \operatorname{Re}(v \mid \lambda v - Av)_H \\ &= \operatorname{Re}\lambda\,\|v\|_H^2 + \operatorname{Re}\mathfrak{a}(v,v) \ge (\operatorname{Re}\lambda - \mu)\,\|v\|_H^2 + \delta\,\|v\|_V^2 . \end{aligned}$$

In fact, this even gives that $\|(\lambda - A)v\|_H \ge (\operatorname{Re}\lambda - \mu)\,\|v\|_H$ for all $v \in \operatorname{dom}(A)$, which implies the desired resolvent estimate if $\operatorname{Re}\lambda > \mu$.

(c) Consider the form $\mathfrak{b}\colon V \times V \to \mathbb{C}$ given by $\mathfrak{b}(u,v) := \mathfrak{a}(u,v) + \mu\,(u \mid v)_H$. The operator B associated to $\mathfrak{b}$ can easily be checked to satisfy $B = A - \mu$ (with the same domain as A). So it suffices to show that $\sigma(B) \subseteq \mathbb{R}$.

The symmetry of $\mathfrak{a}$ implies that $\mathfrak{a}(v,v) \in \mathbb{R}$ and thus $\mathfrak{b}(v,v) \ge \delta\,\|v\|_V^2$ for all $v \in V$. Now take a complex number γ with $\operatorname{Re}\gamma > 0$. Then the form $\gamma\mathfrak{b}\colon V \times V \to \mathbb{C}$, which is associated to the operator γB, satisfies

$$\operatorname{Re}\big(\gamma\mathfrak{b}(v,v)\big) = (\operatorname{Re}\gamma)\mathfrak{b}(v,v) \ge \operatorname{Re}\gamma\,\delta\,\|v\|_V^2$$

for all $v \in V$; here we used again that $\mathfrak{b}(v,v) \in \mathbb{R}$. We thus conclude from (b) that $\mathrm{s}(\gamma B) \le 0$, so $\operatorname{Re}(\gamma\lambda) \le 0$ for every $\lambda \in \sigma(B)$. As this is true for each $\gamma \in \mathbb{C}$ with $\operatorname{Re}\gamma > 0$, it follows that $\sigma(B) \subseteq (-\infty, 0]$. □

[4] In fact, if $\mathfrak{a}$ is symmetric, then one can show that A is a so-called **self-adjoint** operator, but we shall not discuss this further at this point.

We frequently use Theorem 5.1.4 to study examples, starting with Exercises 5.4 and 5.6 and Example 5.4.3. In Definition 4.2.9 we introduced real operators on complex Banach lattices. Analogously we now define real sesquilinear forms.

Definition 5.1.5. Let V be a Hilbert space that embeds continuously and densely into $L^2(\Omega,\nu)$ for a σ-finite measure-space (Ω,ν). A sesquilinear form $\mathfrak{a}\colon V \times V \to \mathbb{C}$ is called **real** if $V_{\mathbb{R}} := V \cap L^2(\Omega,\nu;\mathbb{R})$ satisfies $V = V_{\mathbb{R}} + \mathrm{i}V_{\mathbb{R}}$ and $\mathfrak{a}(u,v) \in \mathbb{R}$ for all $u, v \in V_{\mathbb{R}}$.

It is easy to see that the condition $V = V_{\mathbb{R}} + \mathrm{i}V_{\mathbb{R}}$ in the previous definition is equivalent to the condition $\operatorname{Re} v \in V_{\mathbb{R}}$ for all $v \in V$. The following proposition is a straightforward consequence of Definitions 4.2.9, 5.1.3, and 5.1.5.

Proposition 5.1.6 (Real forms induce real operators)**.** *In the situation of Definition 5.1.5, if $\mathfrak{a}$ is bounded and real, then the associated operator $A\colon L^2(\Omega,\nu) \supseteq \operatorname{dom}(A) \to L^2(\Omega,\nu)$ is also real.*

For operators constructed via real forms, the following result makes it often quite easy to check whether the resolvent is positive for all sufficiently large real numbers. The result is a substantial generalisation of Exercise 4.3. The inequality $\mathfrak{a}(v^-, v^+) \le 0$ in assertion (ii) is an infinite-dimensional version of property (vii) in Exercise 1.2.

Theorem 5.1.7 (The Beurling–Deny criterion for positivity of resolvents)**.** *Under the assumptions of Theorem 5.1.4, let $H = L^2(\Omega,\nu)$ for a σ-finite measure space (Ω,ν). Assume in addition that $\mathfrak{a}$ is real. Then the following are equivalent for the associated operator $A\colon L^2(\Omega,\nu) \supseteq \operatorname{dom}(A) \to L^2(\Omega,\nu)$:*

(i) $\mathcal{R}(\lambda, A) \ge 0$ *for all* $\lambda \in [\mu,\infty)$.

(ii) $V_{\mathbb{R}} := V \cap L^2(\Omega,\nu;\mathbb{R})$ *is a sublattice of* $L^2(\Omega,\nu;\mathbb{R})$ *and* $\mathfrak{a}(v^-, v^+) \le 0$ *for all* $v \in V_{\mathbb{R}}$.

Proof. Observe that, as $\mathfrak{a}$ is real by assumption, so is A according to Proposition 5.1.6. This easily implies that $\mathcal{R}(\lambda, A)$ is also a real operator for every $\lambda \in \mathbb{R} \cap \rho(A)$.

"(ii) ⇒ (i)": Let $\lambda \in [\mu,\infty)$, $0 \le f \in L^2(\Omega)$, and set $v := \mathcal{R}(\lambda, A)f$. We know already that $v \in V_{\mathbb{R}}$, and need to show that $v \ge 0$. To this end, we now proceed similarly as in Exercise 4.3. Since $v^- \in V$, one can compute using $\left(v^- \mid v^+\right)_H = 0$ that

$$\begin{aligned} 0 \le \left(v^- \mid f\right)_H &= (v^- \mid (\lambda - A)v)_H \\ &= -\lambda \|v^-\|_H^2 + \mathfrak{a}(v^-, v) \le -\mu \|v^-\|_H^2 - \mathfrak{a}(v^-, v^-) \le -\delta \|v^-\|_V^2; \end{aligned}$$

where the inequalities respectively used $\mathfrak{a}(v^-, v^+) \le 0$ and ellipticity estimate (5.1.1). Hence $v^- = 0$, which yields $v \ge 0$.

"(i) ⇒ (ii)": Let $v \in V_{\mathbb{R}}$. To show that $V_{\mathbb{R}}$ is a sublattice, it suffices to prove that $v^+ \in V$. For this we use the following abstract regularisation technique: set $w_n := n\mathcal{R}(n, A)v^+$ for all $n > \mu^+$. Then $(w_n) \subseteq \operatorname{dom}(A) \subseteq V$ and, as shown in Exercise 5.1, $w_n \to v^+$ in H. In particular, (w_n) is bounded in H. Assume for a moment that (w_n) is also bounded in V. As V is reflexive, a subsequence (w_{n_k}) converges weakly to some $w \in V$ within V, hence within H. The uniqueness of weak limits in H then ensures $v^+ = w \in V$.

Let us thus prove that the norms $\|w_n\|_V$ are uniformly bounded. Observe that

$$\begin{aligned}\mathfrak{a}(v - w_n, w_n) &= -(v - w_n \mid Aw_n)_H = n\left(v - w_n \mid v^+ - w_n\right)_H \\ &= n\left(\left\|v^+ - w_n\right\|_H^2 + (v^- \mid w_n)_H\right) \geq 0\end{aligned}$$

since $w_n \geq 0$ due to (i). Thus, $\mathfrak{a}(w_n, w_n) \leq \mathfrak{a}(v, w_n)$ for all n. Using the ellipticity estimate (5.1.1), this implies

$$\begin{aligned}\delta \|w_n\|_V^2 &\leq \mathfrak{a}(w_n, w_n) + \mu \|w_n\|_H^2 \\ &\leq \mathfrak{a}(v, w_n) + \mu^+ \|w_n\|_H^2 \leq c\|v\|_V \|w_n\|_V + \mu^+ d \|w_n\|_H \|w_n\|_V\end{aligned}$$

for constants $c, d \geq 0$ and all indices $n > \mu^+$; for the last inequality, we used that $\mathfrak{a}$ is bounded and $V \hookrightarrow H$. Dividing by $\|w_n\|_V$ yields that (w_n) is bounded in V.

It remains to show the second claim in (ii), i.e. that $\mathfrak{a}(v^-, v^+) \leq 0$. Note that

$$\mathfrak{a}(v^-, w_{n_k}) = -\left(v^- \mid Aw_{n_k}\right)_H = -n\left(v^- \mid w_{n_k} - v^+\right)_H = -n\left(v^- \mid w_{n_k}\right)_H \leq 0$$

for each k. Since $\mathfrak{a}(v^-, \cdot)$ is a bounded linear functional on V, the weak convergence of (w_{n_k}) to v^+ in V implies that $0 \geq \mathfrak{a}(v^-, w_{n_k}) \to \mathfrak{a}(v^-, v^+)$. □

As a first application of Theorem 5.1.7, we will discuss Laplace operators with non-local boundary conditions in Exercise 5.6 and Example 5.4.3.

5.2 The maximum principle

In this and the next section, we present a different way to obtain positivity of solution operators to certain PDEs, based on the **maximum principle**. As a motivation, consider a function $v \in \mathrm{C}^2([0,1];\mathbb{R})$ that satisfies $v'' \geq 0$. This means that v is convex, and hence its maximal value is attained at least at one of the boundary points of $\{0,1\}$.

The maximum principle generalises this to a larger class of operators, also on higher dimensional domains. There, one cannot directly work with convexity of v, but the Laplace operator also has a related property that translates well into more general situations: if $0 \leq v \in \mathrm{C}^2([0,1];\mathbb{R})$ vanishes at a point $x \in (0,1)$, then $\Delta v(x) \geq 0$. This property is captured by assumption (1) in the following theorem. For a first intuition, one should think in Theorem 5.2.1 of the situation where $S := \Omega$ is an open set in $\mathbb{R}^n$ and $M = \overline{\Omega}$. Non-open S will become relevant in Chapter 7 (Theorem 7.2.1 and Example 7.2.2)

Theorem 5.2.1 (An abstract maximum principle)**.** *Let (M, d) be a metric space and let $\emptyset \neq S \subseteq M$ be relatively compact.*[5] *Let $D \subseteq \mathrm{C}(\overline{S};\mathbb{R})$ be a vector subspace such that $\mathbb{1} := \mathbb{1}_{\overline{S}} \in D$ and let $A\colon D \to \mathbb{R}^S$ be a linear map with the following properties:*

[5] Recall that a subset of a metric space is called **relatively compact** if its closure is compact.

(1) *The map A satisfies the* ***positive minimum principle on*** *S, i.e. for each $x \in S$ and each function $0 \le u \in D$ one has the implication*

$$u(x) = 0 \qquad \Longrightarrow \qquad (Au)(x) \ge 0.$$

(2) *One has $A\mathbb{1} \le 0$ and there exists a function $0 \le w \in D$ with $(Aw)(x) > 0$ for all $x \in S$.*

Let $v \in D$ attain at least one value in $[0,\infty)$ and satisfy $Av \ge 0$. Then $\partial S \ne \emptyset$ and v attains its maximum at ∂S.

Before the proof, we show Theorem 5.2.1 in action for a classic PDE example.

Example 5.2.2 (The weak maximum principle for the Laplace operator)**.** Let $\emptyset \ne \Omega \subseteq \mathbb{R}^n$ be open and bounded and let $c \in (-\infty, 0]$. Assume that $v \in C(\overline{\Omega};\mathbb{R}) \cap C^2(\Omega;\mathbb{R})$ satisfies $\Delta v(x) + cv(x) \ge 0$ for all $x \in \Omega$.

If v has at least one value in $[0,\infty)$ then v attains its maximum on $\partial\Omega$. In particular, if v vanishes on $\partial\Omega$, then $v \le 0$ in Ω.

Proof. We apply Theorem 5.2.1 to the set $S := \Omega$ in the metric space $M := \overline{\Omega}$. Let $D := C(\overline{\Omega};\mathbb{R}) \cap C^2(\Omega;\mathbb{R}) \subseteq C(\overline{\Omega};\mathbb{R})$ and define $A\colon D \to \mathbb{R}^\Omega$ by $Au := (\Delta + c)u|_\Omega$ for all $u \in D$. It suffices to show that A satisfies the assumptions (1) and (2) of Theorem 5.2.1.

(1) Let $x \in \Omega$ and let $0 \le u \in D$ satisfy $u(x) = 0$. Then u has a global minimum at x. As u is C^2 in a neighbourhood of x, it follows that the Hessian matrix $Hu(x)$ is positive semidefinite. So its trace satisfies $\operatorname{tr}\big(Hu(x)\big) \ge 0$ and thus,

$$(Au)(x) = \Delta u(x) + cu(x) = \operatorname{tr}\big(Hu(x)\big) \ge 0.$$

(2) Clearly, $A\mathbb{1} = c\mathbb{1} \le 0$. Let $w \in D$ be given by $w(x) = e^{\alpha x_1}$ for all $x \in \overline{\Omega}$ and a real number α that satisfies $\alpha^2 > -c$. Then $(Aw)(x) = (\alpha^2 + c)w(x) > 0$ for all $x \in \Omega$. ☐

We will discuss below in Example 5.3.6 how the weak maximum principle is related to the positivity of $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})$. Theorem 5.2.1 can be applied to further types of differential operators, see Exercise 5.3. For the proof of the theorem, we use the following lemma.

Lemma 5.2.3. *Let (M, d) be a compact metric space and let $\emptyset \ne K \subseteq M$ be a compact subset. Let (f_k) be a sequence in $C(M;\mathbb{R})$ that converges uniformly to a function $f \in C(M;\mathbb{R})$. If each f_k attains its maximum on K, then so does f.*

Proof. By assumption, for each $k \in \mathbb{N}$ there is a point $x_k \in K$ such that f_k attains its maximum at x_k. As K is compact, by passing to a subsequence we may assume that (x_k) converges to a point $x^* \in K$. For every $x \in M$, by continuity of f one then has

$$f(x) \le f_k(x) + \|f - f_k\|_\infty \le f_k(x_k) + \|f - f_k\|_\infty \le f(x_k) + 2\|f - f_k\|_\infty \xrightarrow{k\to\infty} f(x^*). \qquad \square$$

Proof of Theorem 5.2.1. We first make a preliminary observation: If a function $u \in D$ attains its maximum at a point $x_1 \in S$ and satisfies $u(x_1) \geq 0$, then $(Au)(x_1) \leq 0$.

Indeed, set $h := u(x_1)\mathbb{1} - u$. Then $0 \leq h \in D$ and $h(x_1) = 0$. According to assumption (1), A satisfies the positive minimum principle in S, so it follows that $(Ah)(x_1) \geq 0$ and thus $(Au)(x_1) \leq u(x_1) \cdot (A\mathbb{1})(x_1)$. As $u(x_1) \geq 0$ and since $(A\mathbb{1})(x_1) \leq 0$ by assumption (2), it follows that $(Au)(x_1) \leq 0$, as claimed.

Now, define $v_k := v + \frac{1}{k}w \in D$ for each $k \in \mathbb{N}$. For each $k \in \mathbb{N}$ choose a point $x_k \in \overline{S}$ where v_k attains its maximum. Fix an index $k \in \mathbb{N}$ for the moment. Since $w \geq 0$ and by assumption v is not everywhere < 0 in $\overline{S}$, one has $v_k(x_k) \geq 0$. Moreover, $Av \geq 0$ implies $(Av_k)(x) \geq \frac{1}{k}(Aw)(x) > 0$ for all $x \in S$, so it follows from the preliminary observation that $x_k \notin S$, so in particular $x_k \in \partial S$. So each v_k attains its maximum at ∂S and Lemma 5.2.3 shows that the same is true for v. □

5.3 Intermezzo: Regularity of solutions

In this section we briefly discuss – without proofs, but with a few intuitive explanations – that the solution u of the equation $(\lambda - \Delta_{\mathrm{Dir}})u = f$ has better regularity than f. We need this later in the course; for the moment we focus on the following motivation.

We know from Exercise 4.3 that the resolvent $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})\colon L^2(\Omega) \to L^2(\Omega)$ is positive if $\lambda > 0$, and the technique from this exercise was generalised to a very broad class of operators in the Beurling–Deny criterion in Theorem 5.1.7. It is natural to ask whether one can alternatively use the maximum principle from Section 5.2 to obtain positivity of $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})$. There is a difficulty with this approach: for $f \in L^2(\Omega)$, the function $u = \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})f$ will in general only be in $\operatorname{dom}(\Delta_{\mathrm{Dir}})$, but the maximum principle (Example 5.2.2) requires more from u, namely to be continuous on $\overline{\Omega}$ and C^2 on Ω.

The question thus arises: under what conditions is the solution u sufficiently smooth? This is a fundamental question of the **regularity theory** for PDEs, which is a rather subtle subject: not only does the answer depend on f, but also on the geometric properties of Ω. For this reason, it is essential to be able to quantify 'smoothness' of the boundary.

Definition 5.3.1. Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be bounded and open. For $k \in \mathbb{N}$, we say that Ω has C^k **boundary**, or equivalently, that $\partial\Omega$ is of **class** C^k, if there exists $\Phi \in \mathrm{C}^k(\mathbb{R}^n;\mathbb{R})$ such that[6]

$$\Omega = \{x \in \mathbb{R}^n : \Phi(x) > 0\}$$

and $\nabla\Phi(x) \neq 0$ for all $x \in \partial\Omega$. It follows that $\partial\Omega$ is the level set $[\Phi = 0]$.

We can now state the major regularity result for the Dirichlet Laplacian Δ_{Dir} on $L^2(\Omega)$.

Theorem 5.3.2 (Elliptic regularity for the Dirichlet Laplacian)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open with a bounded C^2 boundary. Let $\lambda > 0$. Then the resolvent $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})$ of the Dirichlet Laplacian on $L^2(\Omega)$ has the following properties:*

[6] Some readers may be familiar with the definition of C^k boundary via local charts. We comment on this further in the notes at the end of the chapter.

(a) $\operatorname{dom}(\Delta_{\mathrm{Dir}}) = \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) L^2(\Omega) \subseteq H^2(\Omega)$.

(b) *More generally, if* Ω *has* C^{k+2} *boundary, then one has* $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) H^k(\Omega) \subseteq H^{k+2}(\Omega)$ *for every* $k \in \mathbb{N}_0$.

This is a non-trivial result that is best left for a dedicated course in PDEs. Thus we do not prove it here and refer instead to the literature, e.g. [Bre11, Theorem 9.25]. However, it is instructive to discuss a few special cases of assertion (a):

Remarks 5.3.3. In the setting of Theorem 5.3.2, let $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$. Assertion (a) of the theorem says that $u \in H^2(\Omega)$. Let us explain this in the following simpler situations:

(a) If $n = 1$, the fact that $u \in H^2(\Omega)$ is not surprising at all: by the definition of $\operatorname{dom}(\Delta_{\mathrm{Dir}})$, one has $u \in H^1_0(\Omega)$ and $\Delta u \in L^2(\Omega)$, so the first and second weak derivatives of u exist and are in $L^2(\Omega)$. Thus $u \in H^2(\Omega)$. The point here is that, since Ω is one-dimensional, there is only one second derivative and it equals Δu.

(b) Now let $n \geq 2$ but consider $\Omega = \mathbb{R}^n$. Again, we know $u \in H^1_0(\mathbb{R}^n)$ and $\Delta u \in L^2(\mathbb{R}^n)$. But why does $\Delta u \in L^2(\mathbb{R}^n)$ imply that the weak derivatives $\partial_j \partial_k u$ exist[7] and are in $L^2(\mathbb{R}^n)$ for all j, k? The key insight is as follows:

Assume for a moment that also $u \in \mathrm{C}^\infty_c(\mathbb{R}^n)$. One readily checks that Δ_{Dir} acts as the classical Laplace operator on u. Moreover, integration by parts with respect to the jth and the kth variables shows that $\int_{\mathbb{R}^n} (\partial_j^2 \overline{u})(\partial_k^2 u)\, \mathrm{d}x = \int_{\mathbb{R}^n} \left|\partial_j \partial_k u\right|^2 \mathrm{d}x$ for all indices j, k. Hence, $\|\Delta_{\mathrm{Dir}} u\|_{L^2}^2 = \sum_{j,k=1}^n \left\|\partial_j \partial_k u\right\|_{L^2}^2$ and thus,

$$\|\Delta_{\mathrm{Dir}} u\|_{L^2}^2 + \|u\|_{H^1}^2 = \|u\|_{H^2}^2 \quad \text{for all } u \in \mathrm{C}^\infty_c(\mathbb{R}^n). \tag{5.3.1}$$

One can prove – although we shall not do this here – that $\mathrm{C}^\infty_c(\mathbb{R}^n)$ is dense in $\operatorname{dom}(\Delta_{\mathrm{Dir}})$ with respect to $\|\cdot\|_{\Delta_{\mathrm{Dir}}}$, so (5.3.1) implies that $\operatorname{dom}(\Delta_{\mathrm{Dir}}) \subseteq H^2(\mathbb{R}^n)$.

(c) On domains $\Omega \subseteq \mathbb{R}^n$, things are more involved. Clearly, (5.3.1) holds for $u \in C^\infty_c(\Omega)$, but this space will typically not be dense in $\operatorname{dom}(\Delta_{\mathrm{Dir}})$. Indeed, if u is in the closure of $\mathrm{C}^\infty_c(\Omega)$ within $\operatorname{dom}(\Delta_{\mathrm{Dir}})$, then (5.3.1) implies $u \in H^2_0(\Omega)$, which is a proper subspace of $\operatorname{dom}(\Delta_{\mathrm{Dir}})$ in general.

Theorem 5.3.2 is particularly useful when combined with the following result that connects weak to classical differentiability.

Theorem 5.3.4 (A Sobolev embedding theorem)**.** *Let* $n \geq 2$ *and let* $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ *be open and bounded with* C^1 *boundary. Let* $k \in \mathbb{N}$ *and* $p \in [1, \infty]$ *satisfy* $n < kp$ *and let* $m \in \mathbb{N}_0$ *satisfy* $m < k - \frac{n}{p}$. *Then* $W^{k,p}(\Omega) \hookrightarrow \mathrm{C}^m(\overline{\Omega})$.

[7] We use here the common notation $\partial_j \partial_k u$ for the weak partial derivative with respect to the kth and the jth variable. In terms of the multi-index notation used in Definition 3.2.4 this means $\partial_j \partial_k u := \partial^{e_j + e_k} u$.

For $n = 1$ a bit more is true, see Theorem 5.3.7(b) below. A proof of Theorem 5.3.4 can be found, for example, in [Eva10, Section 5.6.3]. However, the result is true under weaker regularity assumptions on Ω; we present some details for the interested reader in Theorem 5.B.8 of the supplementary Section 5.B.

Theorems 5.3.2 and 5.3.4 can be used to derive positivity of $\mathcal{R}(\lambda, \Delta_{\text{Dir}})$, in some cases, from the maximum principle (as a classical alternative to the form approach). To demonstrate this, we use the following properties.

Proposition 5.3.5. *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open and $1 \leq p < \infty$.*

(a) *The cone of positive test functions, $C_c^\infty(\Omega) \cap L^p(\Omega)_+$, is dense in $L^p(\Omega)_+$.*

(b) *Let $n \geq 2$ and assume that Ω is bounded and has C^1 boundary. Then every $u \in C(\overline{\Omega}) \cap W_0^{1,p}(\Omega)$ vanishes on $\partial\Omega$.*

Again, we refrain from showing the proofs. Assertion (a) can be obtained by means of the standard cut-off and regularisation technique which is, for instance, also used in Supplement 3.A. Assertion (b) reflects the intuition discussed after Example 3.3.6: functions in $H_0^1(\Omega)$ "vanish" on $\partial\Omega$. The proof is based on the so-called trace operator that is explained in a bit more detail in Supplement 3.B, where we also discuss Ω with less regular boundary.

The conclusion of the next example is already known from Exercise 4.3, even without regularity or boundedness assumptions on Ω. Yet, the example seems worthwhile as it demonstrates the connection between the maximum principle and positive resolvents.

Example 5.3.6 (Positivity of the resolvent of the Dirichlet Laplacian, again)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^{k+2} boundary, where $k \in \mathbb{N}$ satisfies $k > \frac{n}{2}$. Then for every $\lambda > 0$ the resolvent of the Dirichlet Laplacian on $L^2(\Omega)$ satisfies $\mathcal{R}(\lambda, \Delta_{\text{Dir}}) \geq 0$.

Proof. Fix $\lambda > 0$. Since $\frac{n}{2} < k$, we have $2 < (k+2) - \frac{n}{2}$, and thus Theorem 5.3.2(b) combined with the Sobolev embedding theorem 5.3.4 yields

$$\mathcal{R}(\lambda, \Delta_{\text{Dir}})\, C_c^\infty(\Omega) \subseteq H^{k+2}(\Omega) \subseteq C^2(\overline{\Omega}).$$

By Proposition 5.3.5(a), it suffices to show that $\mathcal{R}(\lambda, \Delta_{\text{Dir}})$ maps $C_c^\infty(\Omega) \cap L^2(\Omega)_+$ into $L^2(\Omega)_+$.

Let $0 \leq f \in C_c^\infty(\Omega)$ and set $u := \mathcal{R}(\lambda, \Delta_{\text{Dir}}) f$. Then $u \in C^2(\overline{\Omega})$ as shown above. In particular, u is continuous on $\overline{\Omega}$. Since $u \in \operatorname{dom}(\Delta_{\text{Dir}}) \subseteq H_0^1(\Omega)$, we conclude that u vanishes on $\partial\Omega$ (Proposition 5.3.5(b)). Since $(\Delta_{\text{Dir}} - \lambda)(-u) = f \geq 0$, the weak maximum principle for the Laplace operator (Example 5.2.2) implies that $-u \leq 0$, so $u \geq 0$. □

The observant reader might have noticed that the regularity improvement in Theorem 5.3.2(b) is more than we needed in Example 5.3.6: here, it suffices to know that the regularity of $\mathcal{R}(\lambda, \Delta_{\text{Dir}}) f$ is not worse than that of f, i.e. if $\mathcal{R}(\lambda, \Delta_{\text{Dir}}) H^k(\Omega) \subseteq H^k(\Omega)$ for each $k \geq 0$. We will, however, use the stronger result of Theorem 5.3.2 in later examples.

We end this subsection with the following theorem on the regularity of Sobolev functions in dimension 1. It shows that, to a large extent, one can compute with weak derivatives in one dimension similarly as with classical derivatives.

Theorem 5.3.7. *Let $\emptyset \neq I = (a,b) \subseteq \mathbb{R}$ be a bounded open interval and let $p \in [1,\infty]$.*

(a) Fundamental theorem of calculus: *If $u \in W^{1,p}(I)$, then u has a representative $\tilde{u}$ that is continuous on $\bar{I}$ such that*

$$\tilde{u}(x) = \tilde{u}(y) + \int_y^x u'(t)\,\mathrm{d}t \qquad \text{for } x,y \in \bar{I}. \tag{5.3.2}$$

(b) Sobolev embedding: *For all $k \in \mathbb{N}$, we have the continuous embedding $W^{k,p}(I) \hookrightarrow \mathrm{C}^{k-1}(\bar{I})$. More precisely, $W^{k,p}(I) \hookrightarrow L^\infty(I)$ and every $u \in W^{k,p}(I)$ has a representative in $\mathrm{C}^{k-1}(\bar{I})$.*[8]

(c) Integration by parts: *If $u, v \in W^{1,p}(I)$, then $uv \in W^{1,p}(I)$ and*

$$\int_y^x uv'\,\mathrm{d}t = \big(u(x)v(x) - u(y)v(y)\big) - \int_y^x u'v\,\mathrm{d}t, \qquad \text{for } x,y \in [a,b].$$

Throughout the course, Theorem 5.3.7 will be used in a variety of examples. Hence, instead of a mere reference to the literature, we present the interested reader with proofs of parts (a) and (b) and a brief sketch of the proof of (c) in Supplement 5.A.

5.4 Positivity close to the spectral bound

For the Dirichlet Laplacian $\Delta_{\mathrm{Dir}} \colon L^2(\Omega) \supseteq \operatorname{dom}(\Delta_{\mathrm{Dir}}) \to L^2(\Omega)$ on an open set $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ we know from Exercise 4.3 that $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) \geq 0$ for all $\lambda > 0$. If Ω is bounded, one can even show that $\mathrm{s}(\Delta_{\mathrm{Dir}}) < 0$[9], which raises the question of whether one also has positivity of the resolvent at all points $\lambda \in (\mathrm{s}(\Delta_{\mathrm{Dir}}), 0]$. This answer is affirmative, and one way to see this is the following general theorem (for $Q = 0$).

For elements x, y of a Banach lattice E we continue to use the convention introduced in Notation 1.2.4: we write $x \preceq y$ if there exists a number $c > 0$ such that $x \leq cy$. For two operators $T, S \in \mathcal{L}(E,F)$ between Banach lattices, we write $S \leq T$ if $T - S \geq 0$ and, in the case of complex scalars, both operators are real. Naturally, we also write $S \preceq T$ if there is a number $c > 0$ such that $S \leq cT$.

Theorem 5.4.1. *Let $A \colon E \supseteq \operatorname{dom}(A) \to E$ be a closed linear operator on a complex Banach lattice E and let $0 \leq Q \in \mathcal{L}(E)$ be a projection. Let $\lambda_0, \mu_0 \in \mathbb{R}$ satisfy $\lambda_0 < \mu_0$ and $(\lambda_0, \mu_0] \subseteq \rho(A)$. If $\mathcal{R}(\mu_0, A) \succeq Q$, then*

$$\mathcal{R}(\mu, A) \succeq Q \quad \textit{for all} \quad \mu \in (\lambda_0, \mu_0].$$

Proof. Consider the set $U := \{\mu \in (\lambda_0, \mu_0) : \mathcal{R}(\lambda, A) \succeq Q \text{ for all } \lambda \in (\mu, \mu_0)\}$. We want to show that $U = (\lambda_0, \mu_0)$. As the latter is connected, it suffices to show that U is non-empty,

[8] From now on we identify each $u \in W^{k,p}(I)$ with its representative in $\mathrm{C}^{k-1}(\bar{I})$.

[9] This is a consequence of the **Poincaré inequality**; however, we leave this subject aside for now.

open, and closed within (λ_0, μ_0). The closedness is straightforward to check. The non-emptiness follows from the Taylor expansion formula of $\mathcal{R}(\cdot, A)$ in Proposition 3.3.2(a): for all $\lambda < \mu_0$ sufficiently close to μ_0 one has

$$\mathcal{R}(\lambda, A) = \sum_{k=0}^{\infty} \underbrace{(\mu_0 - \lambda)^k}_{\geq 0} \underbrace{\mathcal{R}(\mu_0, A)^{k+1}}_{\geq Q^{k+1} = Q} \geq Q.$$

The same Taylor series argument can be used to show that if $\mu \in U$, then a left neighbourhood of μ is also in U. Since $[\mu, \mu_0) \subseteq U$ as well, U is indeed open. □

The fact that Q can be an arbitrary positive projection in Theorem 5.4.1 will become relevant later. For now, we note that even the case $Q = 0$ can be interesting:

Example 5.4.2 (Positivity of the resolvent for the Dirichlet Laplacian left of zero)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open. The Dirichlet Laplacian $\Delta_{\text{Dir}} \colon L^2(\Omega) \supseteq \operatorname{dom}(\Delta_{\text{Dir}}) \to L^2(\Omega)$ satisfies $\sigma(\Delta_{\text{Dir}}) \subseteq (-\infty, 0]$ and $\mathcal{R}(\lambda, \Delta_{\text{Dir}}) \geq 0$ for all $\lambda \in (\mathrm{s}(\Delta_{\text{Dir}}), \infty)$.

Proof. The property $\sigma(\Delta_{\text{Dir}}) \subseteq (-\infty, 0]$ is proved in Exercise 5.4. Moreover, you showed in Exercise 4.3 that $\mathcal{R}(\lambda, A) \geq 0$ for all $\lambda \in (0, \infty)$. By applying Theorem 5.4.1 for $Q = 0$ and any $\mu_0 > 0$, one sees that this inequality remains true for all $\lambda > \mathrm{s}(\Delta_{\text{Dir}})$. □

Theorem 5.4.1 for $Q = 0$ shows that if the resolvent of an operator is positive at one point μ_0, then it is also positive on the left of μ_0 up to the next spectral value on the real axis. On the other hand, the theorem gives no information on positivity on the right of μ_0. The next example shows that there are even second-order differential operators A whose resolvent is positive in a right neighbourhood of $\mathrm{s}(A)$, but not on all of $(\mathrm{s}(A), \infty)$.

Example 5.4.3 (A Laplacian with non-local boundary conditions)**.** Consider the sesquilinear form $\mathfrak{a} \colon H^1(0,1) \times H^1(0,1) \to \mathbb{C}$ defined by

$$\mathfrak{a}(v, u) := \int_0^1 \overline{v'} u' \, \mathrm{d}x + \frac{1}{2} \begin{pmatrix} \overline{v(0)} & \overline{v(1)} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \end{pmatrix}. \tag{5.4.1}$$

Denote the associated operator by $\Delta_B \colon L^2(0,1) \supseteq \operatorname{dom}(\Delta_B) \to L^2(0,1)$.

(a) The operator Δ_B is closed and acts as the weak second derivative on its domain

$$\operatorname{dom}(\Delta_B) = \left\{ u \in H^2(0,1) : \begin{pmatrix} -u'(0) \\ u'(1) \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \end{pmatrix} \right\}.$$

(b) One has $\sigma(\Delta_B) \subseteq (-\infty, 0)$.

(c) One has $\mathcal{R}(\lambda, \Delta_B) \geq 0$ for all $\lambda \in (\mathrm{s}(\Delta_B), 0]$, but not for all $\lambda > 0$.

Proof. The form $\mathfrak{a}$ is the one given in Exercise 5.6 for the choice $B := -\frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ (carefully note the minus sign in the form in that exercise).

(a) This is shown for general B in Exercise 5.6.

(b) As B is a self-adjoint matrix, the form $\mathfrak{a}$ is symmetric in the sense of Theorem 5.1.4(c), and hence $\sigma(\Delta_B) \subseteq \mathbb{R}$. Moreover, as B is negative semidefinite, one can use any number $\mu > 0$ in the ellipticity estimate in Exercise 5.6(b). Hence, $\mathrm{s}(\Delta_B) \le 0$. It remains to show that $0 \notin \sigma(\Delta_B)$. To this end, we prove that $-\Delta_B\colon \operatorname{dom}(\Delta_B) \to L^2(0,1)$ is bijective.

If $u \in \ker(-\Delta_B)$, then $u \in H^2(0,1)$ and $u'' = 0$. Thus, in fact $u \in H^k(0,1)$ for all $k \in \mathbb{N}$, in particular $u \in \mathrm{C}^2([0,1])$ by Theorem 5.3.7(b). By using classical derivatives one can now immediately check that $u'' = 0$ and the boundary conditions in $\operatorname{dom}(\Delta_B)$ imply $u = 0$. So $-\Delta_B$ is injective.

On the other hand, let $f \in L^2(0,1)$. Define a function $u \in L^2(0,1)$ by

$$u(x) = \frac{1}{4}\int_0^1 f(z)\,\mathrm{d}z + \frac{1}{2}\int_x^1\int_0^y f(z)\,\mathrm{d}z\,\mathrm{d}y + \frac{1}{2}\int_0^x\int_y^1 f(z)\,\mathrm{d}z\,\mathrm{d}y \tag{5.4.2}$$

for all $x \in [0,1]$. By using the fundamental theorem of calculus for H^2 (see Theorem 5.3.7(a)) one can check that $u \in \operatorname{dom}(\Delta_B)$ and $-\Delta_B u = f$. So $-\Delta_B$ is surjective.

(c) Clearly, if f is positive, then so is u in formula (5.4.2), and hence $\mathcal{R}(0,\Delta_B)$ is a positive operator. Theorem 5.4.1 for $Q = 0$ thus implies that $\mathcal{R}(\mu, A) \ge 0$ for all $\mu \in (\mathrm{s}(A), 0]$. On the other hand, the matrix B has a strictly negative off-diagonal entry, so Exercise 5.6(c) shows that $\mathcal{R}(\lambda,\Delta_B) \not\ge 0$ for some $\lambda > \mathrm{s}(\Delta_B)$. □

The boundary conditions in $\operatorname{dom}(\Delta_B)$ in Example 5.4.3 are a simple example of **non-local Robin boundary conditions**. One can interpret them in more physical terms: the outward flux through the boundary points, represented by the vector $(-u'(0)\;\; u'(1))^{\mathrm{T}}$, is equal to the average of the boundary values.

In Example 5.4.3 the resolvent $\mathcal{R}(\lambda,\Delta_B)$ is positive in a right neighbourhood of the spectral bound, but not for large $\lambda \in \mathbb{R}$. We have already encountered a similar behaviour in finite dimensions in Theorem 2.3.1(ii). In infinite dimensions we will study this phenomenon in more generality in the next chapter.

Exercises for Chapter 5

Exercise 5.1 (Approximation by resolvents)**.** Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed operator on a complex Banach space X. Assume that there exist $\lambda_0 \in \mathbb{R}$, $C \geq 0$ such that

$$[\lambda_0, \infty) \subseteq \rho(A) \qquad \text{and} \qquad \|\lambda \mathcal{R}(\lambda, A)\| \leq C \quad \text{for all } \lambda \in [\lambda_0, \infty).$$

Prove the following statements:

(a) $\lambda \mathcal{R}(\lambda, A)x \to x$ as $\lambda \to \infty$ for all $x \in \overline{\operatorname{dom}(A)}$.

(b) If X is reflexive[10], then A is densely defined.

Hints: For $x \in X$, use the equality $n\mathcal{R}(n, A)x - x = A\mathcal{R}(n, A)x$ for all integers $n \geq \lambda_0$. Also observe that the graph of A is a closed convex subset of $X \times X$ and is hence weakly closed by the Hahn–Banach separation theorem.

Exercise 5.2 (Approximate eigenvectors for the boundary of the spectrum)**.** Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed operator on a complex Banach space X. Fix $\lambda \in \partial\sigma(A)$ and let (λ_n) be a sequence in $\rho(A)$ that converges to λ. Show that there exists a sequence $(y_n) \subseteq X$ such that $\|y_n\| = 1$ for all $n \in \mathbb{N}$ and $(\lambda - A)y_n \to 0$ in X as $n \to \infty$. Such a sequence is called an **approximate eigenvector** for the spectral value λ.

Hint: Use Proposition 3.3.2 to first obtain a sequence $(x_n) \subseteq X$ with $\|x_n\| = 1$ and $\|\mathcal{R}(\lambda_n, A)x_n\| \to \infty$ as $n \to \infty$.

Exercise 5.3 (The maximum principle for first order differential operators)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be a bounded open set such that $0 \notin \Omega$, and let $b : \Omega \to \mathbb{R}^n$ be a vector field such that

$$x^{\mathrm{T}} b(x) > 0 \quad \text{for all } x \in \Omega. \tag{5.4.3}$$

For $D := \mathrm{C}(\overline{\Omega}; \mathbb{R}) \cap \mathrm{C}^1(\Omega; \mathbb{R})$, consider the linear operator $A : D \to \mathbb{R}^\Omega$ defined by

$$Af := b^{\mathrm{T}} \nabla f, \qquad f \in D.$$

Show that if $u \in D$ satisfies $Au \geq 0$ in Ω, then u attains its maximum on $\partial\Omega$.

Hint: use (5.4.3) to design a function $0 \leq w \in D$ such that $Aw(x) > 0$ for all $x \in \Omega$.

[10] Recall that a Banach space X is reflexive if and only if every bounded sequence in X has a weakly convergent subsequence (due to the Eberlein–Šmulian theorem).

Exercise 5.4. Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open. According to Example 3.3.6(c) no strictly positive number belongs to the spectrum of the Dirichlet Laplacian $\Delta_{\mathrm{Dir}} \colon L^2(\Omega) \supseteq \operatorname{dom}(\Delta_{\mathrm{Dir}}) \to L^2(\Omega)$. Use Theorem 5.1.4 and the characterisation of Δ_{Dir} in Example 3.3.6(b) to show that even $\sigma(\Delta_{\mathrm{Dir}}) \subseteq (-\infty, 0]$.

Exercise 5.5 (Estimates via maximum principle)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be bounded and open. For $\lambda > 0$ and $f \in \mathrm{C}(\overline{\Omega})_+$, assume $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$ solves the boundary value problem

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

Prove that $0 \leq u(x) \leq \frac{1}{\lambda} \|f\|_\infty$ for all $x \in \overline{\Omega}$.

Exercise 5.6 (Laplacian with non-local boundary conditions)**.** Let $B \in \mathbb{R}^{2\times 2}$, and consider the sesquilinear form $\mathfrak{a} \colon H^1(0,1) \times H^1(0,1) \to \mathbb{C}$ defined by

$$\mathfrak{a}(v,u) := \int_0^1 \overline{v'} u' \, \mathrm{d}x - \begin{pmatrix} \overline{v(0)} & \overline{v(1)} \end{pmatrix} B \begin{pmatrix} u(0) \\ u(1) \end{pmatrix}.$$

Let $\Delta_B \colon L^2(0,1) \supseteq \operatorname{dom}(\Delta_B) \to L^2(0,1)$ denote the operator associated to $\mathfrak{a}$.

(a) Prove that

$$\operatorname{dom}(\Delta_B) = \left\{ u \in H^2(0,1) : \begin{pmatrix} -u'(0) \\ u'(1) \end{pmatrix} = B \begin{pmatrix} u(0) \\ u(1) \end{pmatrix} \right\}$$

and that $\Delta_B u$ is the second weak derivative of u for every $u \in \operatorname{dom}(\Delta_B)$.

Hint: Use the integration by parts formula in Theorem 5.3.7(c).

(b) Show that there exist $\mu \in \mathbb{R}$ and $c > 0$ such that

$$\operatorname{Re} \mathfrak{a}(u,u) + \mu \|u\|^2_{L^2(0,1)} \geq c \|u\|^2_{H^1(0,1)} \quad \text{for all } u \in H^1(0,1).$$

Conclude that Δ_B is closed and that $\mathrm{s}(\Delta_B) \leq \mu$.

(c) Prove that the following are equivalent:

(i) There exists $\lambda_0 \geq \mathrm{s}(\Delta_B)$ such that $\mathcal{R}(\lambda, \Delta_B) \geq 0$ for all $\lambda \geq \lambda_0$.

(ii) $\mathcal{R}(\lambda, \Delta_B) \geq 0$ for all $\lambda > \mathrm{s}(\Delta_B)$

(iii) All off-diagonal entries of the matrix B are positive, i.e. $B_{jk} \geq 0$ for $j \neq k$.

Hint: use Stampacchia's lemma (Example 4.1.4(d)), the Beurling-Deny criterion (Theorem 5.1.7), and Theorem 5.4.1.

Notes for Chapter 5

Positivity via forms

The characterisation of positivity via forms (Theorem 5.1.7) can be extended to a much more general setting: for a closed convex set C in a Hilbert space H, a similar criterion can be given to characterise the invariance of C under $\lambda\mathcal{R}(\lambda, A)$ for all sufficiently large λ. This is due to Ouhabaz and is explained in his book [Ouh05].

Example 5.4.3 is, up to minor changes, taken from [DGK16a, Theorem 6.11]. The motivation for studying it in detail in this reference was an observation by Akhlil (private communication; see also [Akh18, Section 3]), who noted that the property $\mathcal{R}(\lambda, \Delta_B) \geq 0$ does not hold for all $\lambda > \mathrm{s}(\Delta_B)$ in this example.

The maximum principle

The maximum principle in its many guises is a fundamental tool for the analysis of so-called *elliptic* and *parabolic* PDEs, and techniques based on the maximum principle have been developed to a high degree of sophistication. Two much-loved references on this subject include the books of Protter and Weinberger [PW84], and Gilbarg and Trudinger [GT01]. More recent books, which focus on nonlinear equations and reflect modern research trends, include those of Fraenkel [Fra11] and Pucci and Serrin [PS07].

The abstract perspective on the maximum principle presented in this chapter seems to be quite uncommon in the PDE literature, but not entirely without precedent. For example, linear differential operators which satisfy the positive minimum principle (Condition (1) of Theorem 5.2.1) are characterised in quite a general setting in [DL00, Chapter V, §5]. Closely related, but from a completely different perspective, is the minimum (or maximum) principle for generators of Markov processes, often attributed to Dynkin; see for example [Dyn65, Chapter V, §3] or [Sch21, Lemma 7.20].

Regularity of the boundary

In Section 5.3, we have stated results for domains $\Omega \subseteq \mathbb{R}^n$ with C^k boundary ($k \geq 1$), where C^k regularity is defined using level sets (Definition 5.3.1). It is just as reasonable to define regularity of the boundary in terms of local charts: in short, $\partial\Omega$ is of class C^k if it is locally the graph of a function $f \in \mathrm{C}^k(\mathbb{R}^{n-1}; \mathbb{R})$. This is precisely how we defined Lipschitz boundaries in Definition 3.B.1. The former definition is efficient, but lacks the general scope of the latter. However, for C^k regularity ($k \geq 1$), the two definitions are equivalent. The reader should not be surprised to learn that this is due to the implicit function theorem, although it is still quite some effort to present the arguments rigorously; a proof may be found, for instance, in [Hen05, Theorem 1.3].

Many properties of functions and function spaces relevant for PDE analysis depend subtly on boundary regularity. For the Sobolev embedding theorems, Lipschitz regularity is sufficient for many applications, and yields all the 'standard' embeddings as discussed in Supplement 5.B. For a much more thorough investigation of optimal geometric conditions and finer embedding theorems, the classic text [AF03, Chapter 4] remains an indispensable reference.

In contrast, within the class of Lipschitz domains, the elliptic regularity for the Dirichlet Laplacian (Theorem 5.3.2) holds only in special cases, and is *not* true in general. The study of regularity of PDE solutions on 'rough' domains is challenging and very much an active research area. A well-known reference on this subject is the monograph of Grisvard [Gri11], which has a rather 'classical PDE' flavour. However, much progress in this subject has also relied on techniques from harmonic analysis; see, for instance, the treatise of Kenig [Ken94].

Encore: if you want to know more...

5.A Sobolev spaces over intervals

In this section, we present some essential facts about Sobolev functions in dimension 1, beginning with a version of the fundamental theorem of calculus (Theorem 5.3.7(a) in the main text).

Theorem 5.A.1. *Let $\emptyset \neq I \subseteq \mathbb{R}$ be an open interval (not necessarily bounded), and let $p \in [1, \infty]$. If $u \in W^{1,p}(I)$, then u has a continuous representative $\tilde{u}$ such that*

$$\tilde{u}(x) = \tilde{u}(y) + \int_y^x u'(t)\,\mathrm{d}t \qquad \forall\, x, y \in I. \tag{5.A.1}$$

The proof will be established via two lemmas of a technical nature.

Lemma 5.A.2. *Let $f \in L^1_{\mathrm{loc}}(I)$, fix $x_0 \in I$, and define*

$$g(x) := \int_{x_0}^x f(t)\,\mathrm{d}t, \qquad x \in I.$$

Then $g \in \mathrm{C}(I)$ and g is weakly differentiable with $g' = f$.

Proof. Write $I = (a, b)$ and let $\varphi \in \mathrm{C}_c^\infty(I)$. Then

$$\int_I g(x)\varphi'(x)\,\mathrm{d}x = -\int_a^{x_0} \left(\int_x^{x_0} f(t)\,\mathrm{d}t \right) \varphi'(x)\,\mathrm{d}x + \int_{x_0}^b \left(\int_{x_0}^x f(t)\,\mathrm{d}t \right) \varphi'(x)\,\mathrm{d}x.$$

We now compute the right hand side of the above equality using Fubini's theorem, which yields

$$-\int_a^{x_0} \int_a^t \varphi'(x)\,\mathrm{d}x f(t)\,\mathrm{d}t + \int_{x_0}^b \int_t^b \varphi'(x)\,\mathrm{d}x f(t)\,\mathrm{d}t = -\int_a^b \varphi(t) f(t)\,\mathrm{d}t.$$

Hence $\int_I g\varphi'\,\mathrm{d}x = -\int_I f\varphi\,\mathrm{d}x$ for all $\varphi \in \mathrm{C}_c^\infty(I)$, and the lemma is proved. □

Lemma 5.A.3. *Let $f \in L^1_{\mathrm{loc}}(I)$ satisfy*

$$\int_I f\varphi'\,\mathrm{d}x = 0 \qquad \forall\, \varphi \in \mathrm{C}_c^\infty(I). \tag{5.A.2}$$

Then f is constant a.e. in I.

Proof. We follow [Bre11, Lemma 8.1], which features a very clever trick. Fix a function $w \in C_c^\infty(I)$ such that $\int_I w \, dx = 1$. Given $\psi \in C_c^\infty(I)$, we define

$$\varphi(x) := \int_a^x \psi(t) - \left(\int_I \psi \, ds\right) w(t) \, dt.$$

Clearly $\varphi \in C^\infty(I)$, and moreover $\varphi(a) = \varphi(b) = 0$ by construction. Since ψ, w both have compact support, it follows that $\varphi \in C_c^\infty(I)$. Hence, the assumption (5.A.2) yields

$$0 = \int_I f\varphi' \, dx = \int_I f \left[\psi - \left(\int_I \psi \, dt\right) w\right] dx = \int_I \psi \left[f - \left(\int_I f w \, dt\right)\right] dx.$$

Since $\psi \in C_c^\infty(I)$ was arbitrary, we deduce that $f - \int_I f w \, dt = 0$ a.e. in I. □

Proof of Theorem 5.A.1. Let $u \in W^{1,p}(I)$ and fix $y \in I$. By Lemma 5.A.2, the function defined by

$$v(x) := \int_y^x u'(t) \, dt, \qquad x \in I$$

is continuous on I and weakly differentiable with $v' = u'$. By definition of weak derivatives, it therefore holds that $\int_I u\varphi' \, dx = \int_I v\varphi' \, dx$ for all $\varphi \in C_c^\infty(I)$. However, this implies that $u - v = c$ a.e. in I for some constant $c \in \mathbb{C}$ by Lemma 5.A.3. It follows that $\tilde{u} := v + c$ is a continuous representative of u. Since $v(y) = 0$, we deduce that $c = \tilde{u}(y)$, and hence $\tilde{u}$ satisfies (5.A.1). □

Theorem 5.A.4 (Sobolev embedding in dimension 1)**.** *Let $p \in [1,\infty]$ and let $\emptyset \neq I \subseteq \mathbb{R}$ be an open interval. Then $W^{k,p}(I) \hookrightarrow C_b^{k-1}(I)$; more precisely, every $u \in W^{k,p}(I)$ has a representative $\tilde{u} \in C^{k-1}(I)$ such that $\partial^\alpha \tilde{u} \in C_b(I)$ for all multi-indices $|\alpha| \leq k-1$.*

If I is a bounded interval, the theorem above is a direct consequence of Theorem 5.A.1 applied to u and all weak derivatives of order $\leq k-1$. In particular, this completes the proof of Theorem 5.3.7(b) in the main text. If the interval is unbounded, it is not obvious why the continuous representative is bounded, and hence additional tools are needed. A detailed proof in this general case can be found in [Bre11, Theorem 8.8].

We conclude this section by stating a Sobolev version for the product rule and integration by parts.

Proposition 5.A.5. *Let $\emptyset \neq I = (a,b) \subset \mathbb{R}$ be a bounded interval, let $p \in [1,\infty]$, and suppose $u, v \in W^{1,p}(I)$. Then $uv \in W^{1,p}(I)$ with $(uv)' = u'v + uv'$. Moreover, the classical integration by parts formula holds:*

$$\int_y^x u'v \, dt = \big(u(x)v(x) - u(y)v(y)\big) - \int_y^x uv' \, dt, \qquad \forall\, x, y \in [a,b].$$

Once it is known that functions in $W^{1,p}(I)$ can be approximated by functions in $C^1(\bar{I})$, then the product rule can be proved by first using the classical product rule, then passing to the limit. The integration by parts formula follows immediately from integrating the product rule. Note that the point evaluations are well-defined, thanks to Theorem 5.A.1. We refer to [Bre11, Corollary 8.10] for the full details.

5.B Sobolev embedding theorems

This section is a brief (!) overview of **Sobolev embedding theorems**, which play a fundamental role in the analysis of partial differential equations and the calculus of variations. A detailed treatment of this subject is usually part of a dedicated course in PDE theory. Thus, we do not show many proofs, but rather focus on the overall strategy and key ideas.

Before we begin, recall that a Banach space Y is said to **embed continuously** into a Banach space X, written as $Y \hookrightarrow X$, if $Y \subseteq X$ and there exists $C > 0$ such that $\|y\|_X \leq C\|y\|_Y$ for all $y \in Y$.

As a basic intuition, we may consider Sobolev spaces to be a compromise between L^p spaces and the classical spaces of continuously differentiable functions. From this perspective, given $k \in \mathbb{N}$, $p \in [1,\infty]$, and $u \in W^{k,p}(\Omega)$, there are two natural questions:

1. (*Improved integrability*) Since u is 'better than the average' L^p function, does it also belong to some $L^q(\Omega)$ with $q > p$?

2. (*Recovery of classical functions*) If u has sufficiently many weak derivatives, does it then have a continuous representative, or even an m-times continuously differentiable representative for some $m \in \mathbb{N}$?

Theorem 5.A.4 gives a satisfying answer in dimension 1. In higher dimensions, however, the answers to the same questions are surprisingly subtle, and depend on the precise relationship between the parameters k, p, the dimension n, and even on geometric properties of the set Ω. Fortunately, much can be learned from the simplest case $\Omega = \mathbb{R}^n$, which is already non-trivial. We begin with the Sobolev–Gagliardo–Nirenberg inequality in $\mathbb{R}^n$ (Theorem 5.B.1) for the Sobolev spaces $W^{1,p}(\mathbb{R}^n)$ with $1 \leq p < n$. In brief, the result asserts that a function in $W^{1,p}$ automatically has better integrability and belongs to L^q for a precisely determined $q > p$.

Theorem 5.B.1 (Sobolev, Gagliardo, Nirenberg). *Let $1 \leq p < n$, and define the* ***Sobolev exponent***

$$p^\star := \frac{np}{n-p}.$$

Then $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^{p^\star}(\mathbb{R}^n)$; in particular, there exists a constant $C > 0$ such that

$$\|u\|_{p^\star} \leq C\|\nabla u\|_p \qquad \forall\, u \in W^{1,p}(\mathbb{R}^n). \tag{5.B.1}$$

This is a standard result in PDE analysis and thus can be found in many textbooks. Some classic choices include [GT01, Theorem 7.10], [Eva10, Section 5.6, Theorem 1], and [Bre11, Theorem 9.9]. The following simple argument helps to get some intuition for the result. We *suppose* for the moment that an estimate of the form

$$\|u\|_q \leq C\|\nabla u\|_p$$

holds for some $q \in (p,\infty)$ and for all $u \in C_c^\infty(\mathbb{R}^n) \subseteq W^{1,p}(\mathbb{R}^n)$. Since $\mathbb{R}^n$ is invariant under the dilation maps $x \mapsto \lambda x$ for all $\lambda > 0$, the estimate should also hold for the dilated functions $x \mapsto u_\lambda(x) := u(\lambda x)$. We compute

$$\frac{1}{\lambda^{n/q}} \|u\|_q = \|u_\lambda\|_q \le C \|\nabla u_\lambda\|_p = C\lambda^{1-n/p} \|\nabla u\|_p$$

and hence

$$\|u\|_q \le C\lambda^{1-(\frac{n}{p}-\frac{n}{q})} \|\nabla u\|_p, \quad \forall \lambda > 0.$$

In order for the right hand side of the latter inequality to remain bounded for $\lambda \to 0$ and also $\lambda \to \infty$, the exponent in λ must be 0, and hence we require $\frac{1}{n} = \frac{1}{p} - \frac{1}{q}$. Upon solving for q, we obtain $q = \frac{np}{n-p} = p^\star$. The necessity of the condition $p < n$ is also evident.

Theorem 5.B.1 has a standard follow-up result.

Corollary 5.B.2. *Let $n \in \mathbb{N}$.*

(a) *If $1 \le p < n$, then $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for all $q \in [p, p^\star]$.*

(b) *If $n \ge 2$, then $W^{1,n}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for every $q \in [n, \infty)$.*

The first part of the corollary is a consequence of the following interpolation inequality, which is proved by a direct application of Hölder's inequality: if $v \in L^r(\mathbb{R}^n) \cap L^s(\mathbb{R}^n)$ with $1 \le r \le s \le \infty$, then $v \in L^p(\mathbb{R}^n)$ for all $p \in [r, s]$ and

$$\|v\|_p \le \|v\|_s^{1-\theta} \|v\|_r^\theta \tag{5.B.2}$$

where $\theta \in (0,1)$ satisfies $\frac{1}{p} = \frac{(1-\theta)}{s} + \frac{\theta}{r}$. Part (b) can be proved by recycling some steps from the proof of Theorem 5.B.1 followed by an induction argument; see e.g. [Bre11, Corollary 9.11] or [Leo09, Exercise 12.37] for details. Although it is a tempting conjecture (since $p^\star \uparrow \infty$ as $p \uparrow n$), one does *not* achieve the embedding $W^{1,n}(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n)$, as is explained in Remark 5.B.5(a) below.

The remaining case $p > n$ is handled by the theorem of Morrey.

Theorem 5.B.3 (Morrey)**.** *Let $n < p < \infty$. Then $u \in W^{1,p}(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n)$; moreover, there exists a constant $C > 0$ such that for every $u \in W^{1,p}(\mathbb{R}^n)$, one has*

$$|u(x) - u(y)| \le C|x-y|^{1-\frac{n}{p}} \|\nabla u\|_p \quad \text{a.e. } x, y \in \mathbb{R}^n \tag{5.B.3}$$

In particular, every $u \in W^{1,p}(\mathbb{R}^n)$ has a bounded, continuous representative.[11]

Again, we refer to standard literature for the proof, for instance [Bre11, Theorem 9.12], [Eva10, Section 5.6, Theorem 4] and [GT01, Theorem 7.10].

The above results can be applied inductively to obtain embeddings for higher-order Sobolev spaces. Part (c) of the following result is especially important for PDE applications, as it establishes a relationship between weak and classical derivatives.

[11] Readers familiar with the Hölder spaces will observe further that the continuous representative belongs to $C^{0,\alpha}(\mathbb{R}^n)$ with $\alpha = 1 - n/p$.

Corollary 5.B.4 (Summary of Sobolev embeddings in $\mathbb{R}^n$)**.** *Let $k, n \in \mathbb{N}$ with $n \geq 2$ and $1 \leq p < \infty$. One has the following continuous embeddings:*

(a) *If $kp < n$, then $W^{k,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for all $q \in [p, p^\star(k)]$ where $p^\star(k) := \frac{np}{n-kp}$.*

(b) *If $kp = n$, then $W^{k,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for all $q \in [p, \infty)$.*

(c) *If $kp > n$ and $m \in \mathbb{N}_0$ satisfies $m < k - \frac{n}{p}$, then $W^{k,p}(\mathbb{R}^n) \hookrightarrow \mathrm{C}_\mathrm{b}^m(\mathbb{R}^n)$; more precisely, every $u \in W^{k,p}(\mathbb{R}^n)$ has a representative $\tilde{u} \in \mathrm{C}^m(\mathbb{R}^n)$ such that $\partial^\alpha \tilde{u} \in \mathrm{C}_\mathrm{b}(\mathbb{R}^n)$ for all multi-indices $|\alpha| \leq m$.*

Proof. We provide the details, since they are often omitted in PDE books. For brevity, we write D to denote any first-order partial derivative and omit writing $\mathbb{R}^n$ from the spaces.

(a) Theorem 5.B.1 yields the result for $k = 1$. Now assume that the claim holds for $kp < n$ for some $k \geq 1$, and let $u \in W^{k+1,p}$ with $(k+1)p < n$. Then $u, Du \in W^{k,p} \hookrightarrow L^{p^\star(k)}$ by the induction hypothesis. Since $(k+1)p < n$ implies $p^\star(k) < n$, therefore $u \in W^{1,p^\star(k)} \hookrightarrow L^{(p^\star(k))^\star} = L^{p^\star(k+1)}$ by Theorem 5.B.1, and the inductive step is complete. We then obtain $W^{k,p} \hookrightarrow L^q$ for all $q \in [p, p^\star(k)]$ by the interpolation inequality (5.B.2).

(b) Corollary 5.B.2 covers the case of $k = 1$, and hence we may assume that $u \in W^{k,\frac{n}{k}}$ for some $k \geq 2$. We have $u, Du \in W^{k-1,\frac{n}{k}} \hookrightarrow L^{(n/k)^\star(k-1)} = L^n$ by part (a), since $(k-1)\frac{n}{k} < n$. Thus $u \in W^{1,n} \hookrightarrow L^q$ for all $q \in [p, \infty)$ by Corollary 5.B.2.

(c) Write $k = m + \ell$, where $\ell \in \mathbb{N}$ satisfies $\ell - 1 \leq \frac{n}{p} < \ell$. We treat the case $m = 0$ first, so $k = \ell$. Then $u, Du \in W^{k-1,p}$ where $(k-1)p \leq n$. If $(k-1)p < n$, it follows from part (a) that $W^{k-1,p} \hookrightarrow L^r$, where $r = p^\star(k-1) > n$. Consequently $u \in W^{1,r}$ with $r > n$, so Theorem 5.B.3 yields that $u \in \mathrm{C}_\mathrm{b}$. If $(k-1)p = n$, then part (b) yields that $W^{k-1,p} \hookrightarrow L^r$ for all $r \in [p, \infty)$. By choosing $r > p \vee n$, we obtain from Theorem 5.B.3 again that $u \in W^{1,r} \hookrightarrow \mathrm{C}_\mathrm{b}$.

In the general case $m \geq 1$, the above conclusion holds for $\partial^\alpha u$ for all multi-indices $|\alpha| \leq m$, and thus $u \in \mathrm{C}_\mathrm{b}^m$. □

Let us comment on some exceptional cases to highlight some of the many hidden subtleties in Sobolev embedding theorems.

Remarks 5.B.5.

(a) It is important to note that Corollary 5.B.2 does *not* assert that $W^{1,n}(\mathbb{R}^n) \subseteq L^\infty(\mathbb{R}^n)$. For $n \geq 2$, this is false! Here is a standard example: define a function $u \in \mathrm{C}^\infty(\mathbb{R}^n \setminus \{0\}; \mathbb{R})$ such that

$$u(x) = \log\left(\log\left(\frac{4}{\|x\|_2}\right)\right), \qquad \|x\|_2 < 1,$$

and $u(x) = 0$ for all $\|x\|_2 \geq 2$. Clearly $u \notin L^\infty(\mathbb{R}^n)$, due to the singularity at $x = 0$. It is then a slightly tedious exercise to check that $u \in W^{1,n}(\mathbb{R}^n)$ nevertheless.

In fact, by further and even more tedious calculations, the same example works to show that $W^{k,p}(\mathbb{R}^n) \not\subseteq L^\infty(\mathbb{R}^n)$ if $kp = n$ and $p > 1$; see [AF03, Example 4.43].

(b) The case $p = 1$ and $k = n$ is special: we have $W^{n,1}(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n)$ (and hence $W^{n,1}(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$ for all $1 \le p \le \infty$ by interpolation). Indeed, if $u \in C_c^\infty(\mathbb{R}^n)$ and $x = (x_1, \dots, x_n)$ is arbitrary, we have

$$u(x) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} (\partial_1 \partial_2 \cdots \partial_n u)(t_1, \dots, t_n) \, dt_1 \cdots dt_n,$$

and hence

$$\|u\|_{L^\infty(\mathbb{R}^n)} \le \|\partial_1 \partial_2 \cdots \partial_n u\|_{L^1(\mathbb{R}^n)} \le \|u\|_{W^{n,1}(\mathbb{R}^n)}.$$

The inequality for general $u \in W^{n,1}(\mathbb{R}^n)$ follows, as usual, by taking an approximating sequence of test functions (u_k) such that $u_k \to u$ in $W^{n,1}$ and $u_k(x) \to u(x)$ a.e. $x \in \mathbb{R}^n$, cf. Corollary 3.A.5.

(c) Theorem 5.B.3 does not extend to $p = \infty$, since the typical proofs require the approximation result Corollary 3.A.5. Nevertheless, it is true that a function $u \colon \mathbb{R}^n \to \mathbb{R}$ is (uniformly) Lipschitz continuous if and only if $u \in W^{1,\infty}(\mathbb{R}^n)$; a proof may be found, for instance, in [Eva10, Theorem 4, Section 5.8].

Corollary 5.B.4 gives a rather satisfactory picture of the Sobolev embeddings in $\mathbb{R}^n$. We now turn to the case of domains, i.e. open subsets $\Omega \subseteq \mathbb{R}^n$, where the situation is not so neat. It turns out that the results are highly influenced by geometric properties of the boundary $\partial\Omega$, and many different approaches are available for the analysis of such properties. Here, we choose to present the *extension technique*, which can be formulated easily in operator-theoretic terms (see Theorem 5.B.7 below).

The key idea is conceptually quite simple.[12] Given a non-empty open subset $\Omega \subseteq \mathbb{R}^n$, we ask if there is a method to extend each function $u \in W^{k,p}(\Omega)$ to some $\tilde{u} \in W^{k,p}(\mathbb{R}^n)$, for which the embedding theorems in $\mathbb{R}^n$ can then be applied. Any such method must necessarily take into account the local behaviour of u near the boundary $\partial\Omega$, in order to preserve the values of u within Ω while allowing weak differentiability on $\mathbb{R}^n$. The following example, which is a simplified version of [Maz11, Example 2 in Section 1.5.1], shows that Ω cannot be arbitrary if this is to work.

Example 5.B.6. Consider the set

$$\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_1 < 1, 0 < x_2 < x_1^3\}$$

and define $u(x) := x_1^{1-\beta}$ with $1 < \beta < 2$. By elementary calculations and using the result of Exercise 3.4(b), one checks that $u \in H^1(\Omega) = W^{1,2}(\Omega)$.

Suppose that u can be extended to a function $\tilde{u} \in H^1(\mathbb{R}^2)$. By Corollary 5.B.2(b), it holds that $\tilde{u} \in L^q(\mathbb{R}^2)$ for all $q \ge 2$. But then

$$\int_{\mathbb{R}^2} |\tilde{u}(x)|^q \, dx \ge \int_\Omega |u(x)|^q \, dx = \int_0^1 \int_0^{x_1^3} x_1^{(1-\beta)q} \, dx_2 \, dx_1 = \int_0^1 x_1^{3-(\beta-1)q} \, dx_1,$$

[12] But technically quite involved!

and we can choose $q > 2$ sufficiently large so that the final integral does not converge. This is a contradiction, and thus we deduce that no such extension $\tilde{u}$ exists.

The previous example illustrates what happens when the boundary $\partial\Omega$ is badly behaved: at the point $(0,0)$, the set Ω has a **cusp**, which allows functions to blow-up while still being in a Sobolev space. However, for the extended function, the boundary is 'removed' and membership in a Sobolev space now excludes very wild behaviour.

It turns out that Lipschitz boundary (see Definition 3.B.1) is the sweet spot for the extension technique.

Theorem 5.B.7 (Sobolev extension operator)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. For every $k \in \mathbb{N}$ and $1 \le p \le \infty$, there exists a bounded linear operator $E\colon W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$ such that $(Eu)(x) = u(x)$ for a.e. $x \in \Omega$. In particular, there exists a constant $C > 0$ such that*

$$\|Eu\|_{W^{k,p}(\mathbb{R}^n)} \le C\,\|u\|_{W^{k,p}(\Omega)} \tag{5.B.4}$$

for $u \in W^{k,p}(\Omega)$.

There is a large variety of techniques for creating extension operators; see e.g. [AF03, Section 5.4] and the references therein. In the statement of Theorem 5.B.7, the operator E a priori depends on the parameters k and p, as well as Ω of course. However, a remarkable theorem of Elias Stein shows that in fact there is a *universal* construction, in the sense that given a bounded domain with Lipschitz boundary, there is an extension operator E that works simultaneously for all $k \in \mathbb{N}$ and $1 \le p \le \infty$. We refer the reader to Stein's original monograph [Ste70, Chapter 6, §3] or [Leo09, Theorem 13.17] for the details of this beautiful result.

Using the heavy machinery of Theorem 5.B.7, it is now almost trivial to write down Sobolev embeddings for bounded Lipschitz domains. We highlight in particular the embedding into C^m spaces, since it is used in the main text.

Theorem 5.B.8. *Let $n \ge 2$ and let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Then all the embeddings of Corollary 5.B.4 remain true when $\mathbb{R}^n$ is replaced by Ω. In particular, if $kp > n$ and $m \in \mathbb{N}_0$ satisfies $m < k - \frac{n}{p}$, then $W^{k,p}(\Omega) \hookrightarrow \mathrm{C}^m(\overline{\Omega})$.*

Proof. If $u \in W^{k,p}(\Omega)$, then by Corollary 5.B.4 and Theorem 5.B.7 we obtain

$$\|u\|_{L^q(\Omega)} \le \|Eu\|_{L^q(\mathbb{R}^n)} \le C\,\|Eu\|_{W^{k,p}(\mathbb{R}^n)} \le \tilde{C}\,\|u\|_{W^{k,p}(\Omega)}$$

in the case $kp \le n$, where $\tilde{C}$ is the product of C in (5.B.4) with the implied embedding constants in Corollary 5.B.4. Likewise, in the case $kp > n$ we obtain $\|u\|_{\mathrm{C}^m(\overline{\Omega})} \le \tilde{C}\,\|u\|_{W^{k,p}(\Omega)}$. □

Chapter 6

Eventually positive resolvents and their spectral properties

We know by now that the resolvent of the Dirichlet Laplacian is positive everywhere on the right of the spectral bound (Example 5.4.2) and that this is closely related to estimates for sesquilinear forms (Exercise 4.3) and to the classical maximum principle (Example 5.3.6). The same techniques can be used to prove similar results for many second order elliptic differential operator with "nice" boundary conditions.

Therefore, the phenomenon observed at the end of the previous chapter (Example 5.4.3) might be all the more surprising: choosing slightly uncommon boundary conditions can result in the resolvent being positive in a right neighbourhood of the spectral bound, but not everywhere up to ∞. As we proceed, we shall see that this **eventual positivity** – where "eventual" means "as one moves towards the spectral bound from the right" – occurs in many more examples. Consequently, we lay the groundwork for a general theory of eventual positivity in this chapter.

6.1 Eventually positive resolvents

The following concept is at the heart of this and the next two chapters.

Definition 6.1.1 (Eventual positivity of resolvents)**.** Let $A\colon E \supseteq \operatorname{dom}(A) \to E$ be a closed operator on a complex Banach lattice E. Let $\lambda_0 \in \mathbb{R}$ be a spectral value of A such that a right neighbourhood of λ_0 is contained in $\rho(A)$. Let $u \in E_+$ and $0 \le Q \in \mathcal{L}(E)$.

(a) $\mathcal{R}(\cdot, A)$ is said to be **individually eventually positive with respect to** u **at** λ_0 if for each $0 \lneq f \in E$ one has $\mathcal{R}(\lambda, A)f \ge u$ for all λ in a (f-dependent) right neighbourhood of λ_0.

(b) $\mathcal{R}(\cdot, A)$ is said to be **uniformly eventually positive with respect to** Q **at** λ_0 if one has $\mathcal{R}(\lambda, A) \ge Q$ for all λ in a right neighbourhood of λ_0.

The case $u \gneq 0$ and $Q \gneq 0$ in Definition 6.1.1 will become relevant in Chapter 7. In the present chapter we focus on the weakest choices of u and Q, namely $u = 0$ and $Q = 0$.

The Laplace operator with non-local boundary conditions in Example 5.4.3 is uniformly eventually positive with respect to 0 at its spectral bound, as shown in that example. It is natural to wonder whether individual and uniform eventual positivity are in fact equivalent, but here is a counterexample.

Example 6.1.2 (Individual versus uniform eventual positivity)**.** There exists an operator $A \in \mathcal{L}\big(C([-1,1]\big)$ with spectrum $\sigma(A) = \{0,-1,-3\}$ and the following properties:

(a) $\mathcal{R}(\cdot, A)$ is individually eventually positive with respect to 0 at $\lambda_0 := 0$.

(b) $\mathcal{R}(\lambda, A)$ is not uniformly eventually positive with respect to 0 at $\lambda_0 = 0$.

Proof. Let $\varphi \in C([-1,1])'$ be given by $\langle \varphi, f\rangle = \frac{1}{2}\int_{-1}^{1} f(\omega)\,\mathrm{d}\omega$. Let $S : \ker\varphi \to \ker\varphi$ be the reflection operator given by $(Sf)(\omega) = f(-\omega)$ for all $f \in \ker\varphi$ and all $\omega \in [-1,1]$. We define A as the block diagonal operator

$$A := \begin{pmatrix} 0 & 0 \\ 0 & -S-2 \end{pmatrix}$$

with respect to the decomposition $C([-1,1]) = \mathbb{C}\,\mathbb{1} \oplus \ker\varphi$. In other words, $A\mathbb{1} = 0$ and $Af = (-S-2)f$ for all $f \in \ker\varphi$. Clearly, 1 and -1 are eigenvalues of S, so it follows from $S^2 = \mathrm{id}_{\ker\varphi}$ and the spectral mapping theorem for polynomials that $\sigma(S) = \{-1,1\}$. Hence, $\sigma(A) = \sigma(0) \cup \sigma(-S-2) = \{0,-1,-3\}$. We now prove (a) and (b).

(a) By using $S^2 = \mathrm{id}_{\ker\varphi}$ one can readily check that, for $\lambda \in \rho(A)$,

$$\mathcal{R}(\lambda, A) = \begin{pmatrix} \frac{1}{\lambda} & 0 \\ 0 & \frac{\lambda+2-S}{(\lambda+2)^2-1} \end{pmatrix}.$$

Now let $0 \lneq f \in C([-1,1])$ and write f as $f = \langle \varphi, f\rangle \mathbb{1} + g$ for a function $g \in \ker\varphi$. Of course, $\lambda\mathcal{R}(\lambda,A)f = \langle\varphi, f\rangle\mathbb{1} + \lambda\mathcal{R}(\lambda,A)g \to \langle\varphi,f\rangle\mathbb{1}$ with respect to $\|\cdot\|_\infty$ for $\lambda \downarrow 0$. Since $\lambda\mathcal{R}(\lambda,A)f$ is real for real λ and $\langle\varphi,f\rangle\mathbb{1}$ is a constant function with value $\langle\varphi,f\rangle > 0$, it follows that $\mathcal{R}(\lambda,A)f \geq 0$ for all $\lambda > 0$ that are sufficiently close to 0.

(b) Fix a number $\lambda > 0$; we show that $\mathcal{R}(\lambda, A) \not\geq 0$. For each $\varepsilon > 0$, one can choose $f_\varepsilon \in C([-1,1])_+$ that satisfies $f_\varepsilon(1) = 1$, $f_\varepsilon(-1) = 0$ and $\langle\varphi, f_\varepsilon\rangle = \varepsilon$. Again, we write f_ε as $f_\varepsilon = \varepsilon\,\mathbb{1} + g_\varepsilon$ for $g_\varepsilon \in \ker\varphi$. Since $g_\varepsilon(-1) = -\varepsilon$ and $g_\varepsilon(1) = 1-\varepsilon$, we get

$$\mathcal{R}(\lambda,A)f_\varepsilon(-1) = \frac{\varepsilon}{\lambda} + \frac{(\lambda+2)g_\varepsilon(-1) - g_\varepsilon(1)}{(\lambda+2)^2-1} \xrightarrow{\varepsilon\downarrow 0} \frac{-1}{(\lambda+2)^2-1} < 0.$$

So there exists $\varepsilon > 0$ such that $\mathcal{R}(\lambda,A)f_\varepsilon(-1) < 0$. Thus $\mathcal{R}(\lambda,A) \not\geq 0$ because $f_\varepsilon \geq 0$. □

6.2 Intermezzo: Eigenvalues and poles of the resolvent

We now study eventual positivity of resolvents with similar spectral theoretic techniques as in the finite-dimensional case. In this section, we present some spectral theoretic machinery in infinite dimensions that resembles, to a certain extent, the tools from Section 2.1. For this to be possible, compactness will have to play an essential role.

In addition, we will need standard results for holomorphic functions with values in a Banach space. Readers not familiar with this theory can find a brief summary thereof in Appendix 6.A.

Proposition 6.2.1 (The resolvent is analytic)**.** *Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed linear operator on a complex Banach space X. The resolvent mapping $\mu \mapsto \mathcal{R}(\mu, A)$ is analytic on $\rho(A)$ as a function with values in $\mathcal{L}(X)$.*

Proof. This follows from the series expansion of the resolvent in Proposition 3.3.2(a) and the characterisation of analyticity in terms of the Taylor series (Theorem 6.A.4). □

The following is an infinite-dimensional analogue of Proposition 2.1.4.

Proposition 6.2.2 (Spectral decomposition)**.** *Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed linear operator on a complex Banach space X and let $\sigma_0 \subseteq \sigma(A)$ such that σ_0 is compact and $\sigma(A) \setminus \sigma_0$ is closed. Then there exists a unique projection $P \in \mathcal{L}(X)$ with the following properties:*

(a) $\operatorname{rg} P \subseteq \operatorname{dom}(A)$ *and* $PAx = APx$ *for all* $x \in \operatorname{dom}(A)$.

(b) $\sigma\left(A|_{\operatorname{rg} P}\right) = \sigma_0$ *and* $\sigma\left(A|_{\ker P}\right) = \sigma(A) \setminus \sigma_0$.[1]

Moreover, for any closed C^1-cycle γ in $\mathbb{C}$ that encircles each element of σ_0 precisely once, but no element of $\sigma(A) \setminus \sigma_0$, we have

$$P = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \mathcal{R}(\mu, A) \, \mathrm{d}\mu.$$

The notion of **cycle** that occurs in Proposition 6.2.2 means a formal sum of finitely many paths. This is necessary in infinite dimensions, since the spectrum need not consist of isolated points and hence, no single path with the required properties might exist. We do not present the proof of Proposition 6.2.2 and instead refer to the literature, e.g. [EN00, Proposition IV.1.16]. However, in the case that σ_0 is an isolated singleton which is also a pole of the resolvent, the analysis appears much like in the finite dimensional case (see Theorem 6.2.6 below).

Definition 6.2.3 (Spectral projections)**.** In the situation of Proposition 6.2.2, the projection P is called the **spectral projection** of A associated to σ_0.

[1] Here, $A|_{\operatorname{rg} P} \in \mathcal{L}(\operatorname{rg} P)$ and $A|_{\ker P}$ acts as a closed operator on $\ker P$ with domain $\operatorname{dom}(A) \cap \ker P$.

After defining spectral projections in finite dimensions (Definition 2.1.5) we mentioned that the roles of σ_0 and $\sigma(A) \setminus \sigma_0$ are symmetric – i.e. swapping them gives the complementary projection. In infinite-dimensions this is, in general, only true if $A \in \mathcal{L}(X)$. For unbounded A, the set $\sigma(A) \setminus \sigma_0$ can be unbounded, so that no associated spectral projection is defined. Even if the set is bounded, the spectral projections of $\sigma(A) \setminus \sigma_0$ and σ_0 need not add up to id_X – indeed, this is the case whenever $\sigma_0 := \sigma(A) = \emptyset$.

Definition 6.2.4 (Powers of unbounded operators)**.** Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a linear operator on a Banach space X. The powers of A are defined inductively by

$$A^0 := \mathrm{id}, \quad A^{n+1}x := A(A^n x) \quad \text{for } x \in \operatorname{dom}\left(A^{n+1}\right) := \left\{x \in \operatorname{dom}\left(A^n\right) : A^n x \in \operatorname{dom}(A)\right\}.$$

Of course, $\operatorname{dom}(A^n)$ is a decreasing sequence of subspaces. Moreover, $A^m A^n = A^{m+n}$ whenever the composition is defined. Exercise 6.2 explores such properties.

Definition 6.2.5 ((Generalised) eigenspaces and semisimplicity)**.** Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a linear operator on a complex Banach space X and let $\lambda \in \mathbb{C}$.

(a) λ is called an **eigenvalue of** A if $\ker(\lambda - A) \neq \{0\}$. In this case, $\ker(\lambda - A)$ is called the associated **eigenspace** and its non-zero elements are called the associated **eigenvectors**. The dimension of $\ker(\lambda - A)$ is called the **geometric multiplicity** of λ.

The **point spectrum** $\sigma_{\mathrm{pnt}}(A)$ is the set of all eigenvalues of A.

(b) If λ is an eigenvalue of A, then the subspace $\bigcup_{k\in\mathbb{N}} \ker(\lambda - A)^k$ of $\operatorname{dom}(A)$ is called the associated **generalised eigenspace**. Its dimension is called the **algebraic multiplicity** of λ.

(c) If λ is an eigenvalue of A, it is called a **semisimple eigenvalue** if its generalised eigenspace coincides with its eigenspace.

Note that $\ker(\lambda - A) \subseteq \ker(\lambda - A)^2 \subseteq \dots$, so the generalised eigenspace is indeed a subspace of $\operatorname{dom}(A)$. In finite dimensions, semisimplicity of an eigenvalue λ is obviously equivalent to equality of its geometric and algebraic multiplicities. However, in infinite dimensions, it can happen that both multiplicities are ∞ without equality of the eigenspace and the generalised eigenspace.

For a spectral value that is a pole of the resolvent, many of the properties from the finite-dimensional setting (Proposition 2.1.6) remain true. We summarise the most important ones for our purposes in the following theorem.

Theorem 6.2.6 (Poles of the resolvent)**.** *Let A be a closed operator on a complex Banach space X and let $\lambda \in \sigma(A)$ be a pole of the resolvent $\mathcal{R}(\cdot, A) : \rho(A) \to \mathcal{L}(X)$ of order $p \in \mathbb{N}$. Let*

$$\mathcal{R}(\mu, A) = \sum_{k=-p}^{\infty} Q_{k+1}(\mu - \lambda)^k.$$

denote the Laurent series expansion of the resolvent about λ with coefficients $Q_{k+1} \in \mathcal{L}(X)$.[2]

[2]Note that $Q_{-p+1} \neq 0$ since p is the pole order of λ.

(a) *One has* $\{0\} \neq \operatorname{rg} Q_{-p+1} \subseteq \ker(\lambda - A)$. *In particular,* λ *is an eigenvalue of* A.

(b) *One has* $\operatorname{rg} Q_0 = \ker(\lambda - A)^k$ *for all* $k \geq p$, *so* $\operatorname{rg} Q_0$ *is the generalised eigenspace of* A *for the eigenvalue* λ. *Thus, the eigenvalue* λ *is semisimple if and only if* $p = 1$.

(c) *The coefficient* $Q_0 \in \mathcal{L}(X)$ *is the spectral projection of* A *associated with* λ.

(d) *If* A *is densely defined, then* λ *is also a pole of* $\mathcal{R}(\cdot, A')$ *of order* p *with the Laurent series expansion*

$$\mathcal{R}(\mu, A') = \sum_{n=-p}^{\infty} Q'_{k+1}(\mu - \lambda)^k.$$

In particular, λ *is also an eigenvalue of* A' *with the associated spectral projection* Q'_0.

Proof of (c) *and* (d). (c) As in Proposition 2.1.6(a), this follows from contour integral formula of the spectral projection in Proposition 6.2.2 and the scalar-valued integral formulas in Proposition 2.1.2.

(d) This follows from (a) by using the fact $\mathcal{R}(\cdot, A') = \mathcal{R}(\cdot, A)'$ from Exercise 3.5(d). ☐

The proofs of (a) and (b) rely on a detailed analysis of the coefficients Q_k of the Laurent series expansion. To stay on track, we again refrain from discussing those arguments here. Readers fond of complex analysis or looking for a deeper understanding of spectral theory, can find the full analysis in Supplement 6.B, in Theorems 6.B.1 and 6.B.3.

In practice, a very convenient way to check that a spectral value is a pole of the resolvent is to use the following concept and Theorem 6.2.9 below.

Definition 6.2.7 (Compact resolvent)**.** A closed operator A on a complex Banach space X is said to have **compact resolvent** if there exists $\lambda \in \rho(A)$ such that $\mathcal{R}(\lambda, A) : X \to X$ is a compact operator.

Proposition 6.2.8. *Let* $A\colon X \supseteq \operatorname{dom}(A) \to X$ *be a closed operator on a complex Banach space* X. *The following are equivalent:*

(i) *A has compact resolvent.*

(ii) $\rho(A) \neq \emptyset$ *and the operator* $\mathcal{R}(\lambda, A)\colon X \to X$ *is compact for each* $\lambda \in \rho(A)$.

(iii) $\rho(A) \neq \emptyset$ *and embedding* $\operatorname{dom}(A) \hookrightarrow X$ *is compact.*

Proof. "(i) ⇒ (iii)": Let $\lambda \in \rho(A)$ be such that $\mathcal{R}(\lambda, A)\colon X \to X$ is compact. Composing with $\lambda - A \in \mathcal{L}(\operatorname{dom}(A), X)$ yields that $\operatorname{id}_{\operatorname{dom}(A)}\colon \operatorname{dom}(A) \to X$ is compact.

"(iii) ⇒ (ii)": Let $\lambda \in \rho(A)$. If the embedding $\operatorname{dom}(A) \hookrightarrow X$ is compact, then its composition with $\mathcal{R}(\lambda, A) \in \mathcal{L}(X, \operatorname{dom}(A))$ ensures that $\mathcal{R}(\lambda, A)\colon X \to X$ is compact.

"(ii) ⇒ (i)": This implication is obvious. ☐

It can happen that $\operatorname{dom}(A)$ embeds compactly into X but $\rho(A) = \emptyset$. For instance, this is the case for the first order differential operator on $C([0,1])$ from Example 3.3.4(a), where the compact embedding follows from the Arzelà–Ascoli theorem.

In examples, compactness of the resolvent is often established by using a compact embedding result for Sobolev spaces to check condition (iii) in Proposition 6.2.8. The reason why we are interested in operators with compact resolvent is that they have particularly nice spectral properties that are reminiscent of the finite-dimensional case.

Theorem 6.2.9 (Compact resolvent and spectrum)**.** *Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed operator on a complex Banach space X. If A has compact resolvent, then*

(a) *$\sigma(A)$ consists only of eigenvalues.*

In fact, each spectral value of A is a pole of the resolvent and an eigenvalue of finite algebraic multiplicity.

(b) *$\sigma(A)$ has no accumulation points in $\mathbb{C}$.*

One can either derive Theorem 6.2.9 from the Riesz–Schauder spectral theory of compact operators by using that the resolvent satisfies appropriate spectral mapping results, or one can obtain it as a special case of the so-called **analytic Fredholm theory**, see e.g. [GGK90, Section XI.8]. Theorem 6.2.9 will only be useful for our purposes if there are any spectral values at all.[3] For operators associated to symmetric sesquilinear forms, this is guaranteed by part (c) of the following result.

Proposition 6.2.10 (Compact embedding of form domains)**.** *Let the Hilbert spaces V, H, the sesquilinear form $\mathfrak{a}\colon V \times V \to \mathbb{C}$, and its associated operator $A\colon H \supseteq \operatorname{dom}(A) \to H$ satisfy the assumptions of Theorem 5.1.4. Then:*

(a) *The inclusion map $\operatorname{dom}(A) \hookrightarrow V$ is continuous.*

(b) *If the embedding $V \hookrightarrow H$ is compact, then A has compact resolvent.*

(c) *If the form $\mathfrak{a}$ is symmetric, then $\sigma(A) \neq \emptyset$.*[4]

Proof of (a) and (b). (a) Recall that $\operatorname{dom}(A) \subseteq V \subseteq H$ and that all three spaces are Banach spaces. The embeddings $\operatorname{dom}(A) \hookrightarrow H$ and $V \hookrightarrow H$ are continuous, so the closed graph theorem implies that the inclusion map from $\operatorname{dom}(A)$ into V is also continuous (Exercise 3.3(a)).

(b) Now assume that $V \hookrightarrow H$ is compact. Then it follows from (a) that $\operatorname{dom}(A) \hookrightarrow H$ is compact. Moreover, one has $\rho(A) \neq \emptyset$ by Theorem 5.1.4(b). □

We will not prove part (c) in the main text, but the interested reader can find a proof in Theorem 6.C.5(a) in Supplement 6.C.

[3] Note that there are operators with compact resolvent and empty spectrum (Example 3.3.4(b)).

[4] Readers familiar with form methods might recognise that this is a special case of the fact that self-adjoint operators have non-empty spectrum.

6.3 Spectral consequences of eventual positivity

To apply the spectral theoretic results from the previous sections in concrete examples, one needs tools to check that an operator has compact resolvent. Compact embedding theorems for Sobolev spaces are very useful for this purpose. The following result in one dimension is easy to prove, given the properties from Theorem 5.3.7.

Theorem 6.3.1. *Let $\emptyset \neq I \subset \mathbb{R}$ be an open bounded interval. For all $1 < p \leq \infty$, the embeddings $W^{1,p}(I) \hookrightarrow \mathrm{C}(\bar{I})$ and $W^{1,p}(I) \hookrightarrow L^p(I)$ are compact.*

Proof. Let $B = \{u \in W^{1,p}(I) : \|u\|_{W^{1,p}} \leq 1\}$ be the closed unit ball in $W^{1,p}(I)$. By the embedding $W^{1,p}(I) \hookrightarrow \mathrm{C}(\bar{I})$ from Theorem 5.3.7(b), we deduce that B is bounded in $\mathrm{C}(\bar{I})$. By the fundamental theorem of calculus for Sobolev functions (Theorem 5.3.7(a)) and Hölder's inequality, it follows for all $u \in B$ that

$$\left|u(x) - u(y)\right| \leq \int_y^x \left|u'(t)\right| \,\mathrm{d}t \leq \left\|u'\right\|_{L^p(I)} \left|x-y\right|^{1-1/p} \qquad \forall\, x, y \in \bar{I}. \tag{6.3.1}$$

Since $\left\|u'\right\|_{L^p(I)} \leq 1$ for all $u \in B$ and $p > 1$, the above inequality shows that B is equicontinuous. Hence, the Arzelà-Ascoli theorem implies that B is compact in $\mathrm{C}(\bar{I})$. Finally, the embedding $\mathrm{C}(\bar{I}) \hookrightarrow L^p(I)$ is continuous, and the remaining assertion follows. □

The case $p = 1$ is different and will be treated in Exercise 6.5. Here is a first example of how compactness of the resolvent can be used to determine the spectrum of an operator.

Example 6.3.2. Consider the Dirichlet Laplacian $\Delta_{\mathrm{Dir}} \colon L^2(0,\pi) \supseteq \operatorname{dom}(\Delta_{\mathrm{Dir}}) \to L^2(0,\pi)$.

(a) One has $\sigma(\Delta_{\mathrm{Dir}}) = \sigma_{\mathrm{pnt}}(\Delta_{\mathrm{Dir}}) = \left\{-k^2 : k \in \mathbb{N}\right\}$.

(b) The eigenspace $\ker(-k^2 - \Delta_{\mathrm{Dir}})$ is spanned by $\sin(k\,\cdot)$ for all $k \in \mathbb{N}$.

In particular, $\mathrm{s}(\Delta_{\mathrm{Dir}}) = -1$ is an eigenvalue with a positive eigenvector.

Proof. One readily checks for every $k \in \mathbb{N}$ that $\sin(k\,\cdot)$ is indeed in $\operatorname{dom}(\Delta_{\mathrm{Dir}})$ and is an eigenvector of Δ_{Dir} for the eigenvalue $-k^2$. Let us show that, conversely, every spectral value is an eigenvalue of the form $-k^2$ and that its eigenspace is spanned by $\sin(k\,\cdot)$.

First, we deduce from Example 3.3.6(b) that Δ_{Dir} is the operator associated to a sesquilinear form $\mathfrak{a}$ on $H_0^1(0,\pi)$ and that $\mathfrak{a}$ satisfies the assumptions of Theorem 5.1.4. The embedding $H_0^1(0,\pi) \hookrightarrow L^2(0,\pi)$ is compact (Theorem 6.3.1) and thus, Δ_{Dir} has compact resolvent (Proposition 6.2.10).

We already know that $\sigma(\Delta_{\mathrm{Dir}}) \subseteq (-\infty, 0]$ (Example 5.4.2) and the compactness of the resolvent implies that every spectral value is actually an eigenvalue (Theorem 6.2.9(a)). Now, let $\lambda \in (-\infty, 0]$ be an eigenvalue of Δ_{Dir} with a corresponding eigenvector $0 \neq u \in \operatorname{dom}(\Delta_{\mathrm{Dir}}) \subseteq H^2(0,\pi)$. Then $u'' = \Delta_{\mathrm{Dir}} u = \lambda u \in H^2(0,\pi)$. A simple inductive argument[5] yields $u \in H^k(0,\pi) \subseteq \mathrm{C}^{k-1}([0,\pi])$ for all $k \in \mathbb{N}$, where the latter inclusion is due to the one-dimensional Sobolev embedding theorem 5.3.7(b).

[5] This is often called **bootstrapping** in the PDE literature.

Classical ODE theory can thus be applied to $u'' = \lambda u$ and yields that $u = \alpha\cos(\sqrt{-\lambda}\,\cdot) + \beta\sin(\sqrt{-\lambda}\,\cdot)$ for suitable scalars $\alpha, \beta \in \mathbb{C}$. As $u \in H_0^1(\Omega)$ one has $u(0) = u(\pi) = 0$ (Exercise 6.3). The equality $u(0) = 0$ implies $\alpha = 0$, so $0 \neq u = \beta\sin(\sqrt{-\lambda}\,\cdot)$. The equality $u(\pi) = 0$ now gives $\sqrt{-\lambda} \in \mathbb{N}$, i.e. there exists $k \in \mathbb{N}$ such that $\lambda = -k^2$ and $u = \beta\sin(k\,\cdot)$. □

We already know from Example 5.4.2 that the resolvent $\mathcal{R}(\cdot, \Delta_{\mathrm{Dir}})$ is positive on the right of the spectral bound $\mathrm{s}(\Delta_{\mathrm{Dir}})$. In Example 6.3.2 we have now seen by a direct computation that the eigenvalue $\mathrm{s}(\Delta_{\mathrm{Dir}})$ has a positive eigenvector. Thinking back to the Perron–Frobenius type results in finite dimensions (Theorems 1.3.8 and 1.3.9 in the positive case, Theorems 2.2.3 and 2.3.1 in the eventually positive case), one might wonder whether the existence of a positive eigenvector is a consequence of (eventual) positivity of the resolvent in general infinite-dimensional situations, as well. The answer is 'yes':

Theorem 6.3.3 (Eigenvectors for eventually positive resolvents)**.** *Let A be a closed operator on a complex Banach lattice E and let $\lambda \in \mathbb{R}$ be a pole of the resolvent $\mathcal{R}(\cdot, A)$.*[6] *If $\mathcal{R}(\cdot, A)$ is individually eventually positive with respect to* 0 *at λ, then:*

(a) *There exists a positive eigenvector $v \in E$ of A for the eigenvalue λ.*

(b) *If the eigenvalue λ is semisimple, then its associated spectral projection is positive.*

(c) *If A is densely defined, then A' has a positive eigenvector $\varphi \in E'$ for the eigenvalue λ.*

Proof. Let $p \in \mathbb{N}$ denote the pole order of λ and consider the the Laurent series expansion

$$\mathcal{R}(\mu, A) = \sum_{k=-p}^{\infty} Q_{k+1}(\mu - \lambda)^k$$

of the resolvent about λ, where $Q_{-p+1} \neq 0$ and where the range of Q_{-p+1} is contained in the eigenspace $\ker(\lambda - A)$ according to Theorem 6.2.6(a). One has $Q_{-p+1} = \lim_{\mu\to\lambda}(\mu - \lambda)^p \mathcal{R}(\mu, A)$. For every $x \in E_+$ it thus follows that $Q_{-p+1}x = \lim_{\mu\downarrow\lambda}(\mu-\lambda)^p\mathcal{R}(\mu, A)x \geq 0$ due to the eventual positivity of $\mathcal{R}(\cdot, A)$. So Q_{-p+1} is a positive operator.

(a) Since E_+ spans E and $Q_{-p+1} \neq 0$ there exists a vector $x \in E_+$ such that $v := Q_{-p+1}x \neq 0$. Hence, v is a positive eigenvector of A for the eigenvalue λ.

(b) If the eigenvalue λ is semisimple, then $p = 1$ and $0 \leq Q_{-p+1} = Q_0$ is the spectral projection (Theorem 6.2.6(b) and (c)).

(c) By Theorem 6.2.6(a) and (d), λ is an eigenvalue of A' and $\mathrm{rg}(Q'_{-p+1}) \subseteq \ker(\lambda - A')$. As the dual cone E'_+ spans E' and $Q'_{-p+1} \neq 0$, there exists $\psi \in E'_+$ such that $0 \neq \varphi := Q'_{-p+1}\psi \in \ker(\lambda - A')$. Also $\varphi \geq 0$ because $Q'_{-p+1} \geq 0$ (Corollary 4.4.5). □

Let us demonstrate Theorem 6.3.3 in two examples.

[6] Hence, λ is an eigenvalue of A and, if A is densely defined, also an eigenvalue of A' (Theorem 6.2.6(a) and (d)).

Example 6.3.4. Let $\Delta_B \colon L^2(0,1) \supseteq \operatorname{dom}(\Delta_B) \to L^2(0,1)$ denote the Laplace operator with non-local boundary conditions from Example 5.4.3. Then $-\infty < \mathrm{s}(\Delta_B) < 0$ and $\mathrm{s}(\Delta_B)$ is an eigenvalue of Δ_B with a positive eigenvector.

Proof. We know from Example 5.4.3 that $\mathrm{s}(\Delta_B) < 0$ and $\mathcal{R}(\lambda, \Delta_B) \geq 0$ for all $\lambda \in (\mathrm{s}(\Delta_B), 0]$. Moreover, Proposition 6.2.10(b) shows that Δ_B has compact resolvent since the embedding of the associated form domain $H^1(0,1)$ into $L^2(0,1)$ is compact (Theorem 6.3.1). Thus, all spectral values of Δ_B are poles of the resolvent (Theorem 6.2.9(a)).

As the form that we used to define Δ_B is symmetric, one has $\sigma(\Delta_B) \neq \emptyset$ (Proposition 6.2.10(c)), equivalently $\mathrm{s}(\Delta_B) > -\infty$. As $\sigma(\Delta_B) \subseteq (-\infty, 0)$ (Example 5.4.3) is non-empty and closed, $\mathrm{s}(\Delta_B) \in \sigma(\Delta_B)$ and hence is a pole of $\mathcal{R}(\cdot, \Delta_B)$. So we can apply Theorem 6.3.3(a) to see that there exists a positive eigenvector for $\mathrm{s}(\Delta_B)$. □

In contrast to the simple case on the interval that was treated in Example 6.3.2, for the Dirichlet Laplacian on general domains in $\mathbb{R}^n$, the eigenvalues or eigenvectors cannot be explicitly computed. Yet a positive eigenvector corresponding to $\mathrm{s}(\Delta_{\mathrm{Dir}})$ exists:

Example 6.3.5 (Dirichlet Laplacian on domains)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open and bounded. The Dirichlet Laplacian Δ_{Dir} on $L^2(\Omega)$ has compact resolvent, and its spectral bound $\mathrm{s}(\Delta_{\mathrm{Dir}})$ is not $-\infty$ and is an eigenvalue with a positive eigenvector.

Proof. The argument is the same as in Example 6.3.4, except that one now has to use compactness of Sobolev embeddings in several dimensions (Theorem 6.3.6 below). □

The following result is a version of Theorem 6.3.1 in several dimensions. It was used in the previous example and will be frequently useful in the rest of the course.

Theorem 6.3.6 (Compact embeddings of Sobolev spaces)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open and bounded and let $p \in [1, \infty)$.*

(a) *The embedding $W_0^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is compact.*

(b) *If Ω has C^1-boundary, then the embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is compact.*

The proof of Theorem 6.3.6 usually relies on the Fréchet-Kolmogorov compactness theorem (see Exercise 6.5). Details may be found in [Bre11, Theorem 9.16].

6.4 The left neighbourhood of spectral values

For an operator A, eventual positivity of $\mathcal{R}(\cdot, A)$ at a point $\lambda_0 \in \mathbb{R}$ means, loosely speaking, that the solution u to $(\lambda - A)u = f$ is positive if $f \geq 0$ and λ is in a right neighbourhood of λ_0. It is natural to ask what happens on the left of λ_0. A glance at the case $E = \mathbb{C}$ suggests that one might expect negativity of solutions there. It turns out that this is the correct idea in principle, but there are also quite a few subtleties and surprises.

Definition 6.4.1 (Eventual negativity of resolvents)**.** Let $A \colon E \supseteq \operatorname{dom}(A) \to E$ be a closed operator on a complex Banach lattice E. Let $\lambda_0 \in \mathbb{R}$ be a spectral value of A such that a left neighbourhood of λ_0 is contained in $\rho(A)$. Let $u \in E_+$ and $0 \leq Q \in \mathcal{L}(E)$.

(a) $\mathcal{R}(\cdot, A)$ is said to be **individually eventually negative with respect to u at λ_0** if for each $0 \lneq f \in E$, one has $\mathcal{R}(\lambda, A)f \preceq -u$ for all λ in an (f-dependent) left neighbourhood of λ_0.

(b) $\mathcal{R}(\cdot, A)$ is said to be **uniformly eventually negative with respect to Q at λ_0** if one has $\mathcal{R}(\lambda, A) \preceq -Q$ for all λ in a left neighbourhood of λ_0.

We will see in the subsequent chapters that eventual positivity indeed occurs in various examples. In the present section we focus on the opposite phenomenon: a rather surprising result that says that, in many cases, eventual positivity and negativity are not possible at the same time. We need the following notion from Banach lattice theory.

Definition 6.4.2 (Ideals in vector lattices)**.** Let E be a real vector lattice or a (real or complex) Banach lattice.[7] An **ideal** in E is a vector subspace $I \subseteq E$ such that, for all $x, y \in E$, the properties $y \in I$ and $|x| \le |y|$ imply that $x \in I$.

Note that a vector subspace $I \subseteq E$ is an ideal if and only if I is a vector sublattice and $0 \le x \le y \in I$ implies $x \in I$. Let us give examples of ideals in our favourite Banach lattices.

Examples 6.4.3.

(a) Let (Ω, μ) be a σ-finite measure space and let $\Omega' \subseteq \Omega$ be measurable. For $p \in [1, \infty]$, it is easy to check that

$$I_{\Omega'} := \left\{ f \in L^p(\Omega, \mu) : f|_{\Omega'} = 0 \right\}$$

is a closed ideal of $L^p(\Omega, \mu)$. If $p \ne \infty$ one can prove that, in fact, all closed ideals in $L^p(\Omega, \mu)$ are of this form [BKFR17, Proposition 10.15].

(b) Let (Ω, μ) be a σ-finite measure space. Clearly, $L^\infty(\Omega, \mu)$ is an ideal in $L^p(\Omega, \mu)$ for each $p \in [1, \infty]$. It is not closed unless $p = \infty$ or $L^p(\Omega, \mu)$ is finite-dimensional.

(c) Let K be a compact metric space [8] and let $J \subseteq K$ be a closed set. Then

$$I_J := \left\{ f \in \mathrm{C}(K) : f|_J = 0 \right\}$$

is a closed ideal of $\mathrm{C}(K)$. Again, one can prove that all closed ideals in $\mathrm{C}(K)$ are of this form [BKFR17, Proposition 10.13].

(d) Let $\emptyset \ne \Omega \subseteq \mathbb{R}^n$ be an open set, and $p \in [1, \infty)$. Then $W_0^{1,p}(\Omega; \mathbb{R})$ is an ideal in $W^{1,p}(\Omega; \mathbb{R})$. A proof for the case $p = 2$, which can easily be adapted for general $p \in [1, \infty)$, can be found in [AU23, Theorem 6.39]. An important ingredient is the fact that the lattice operations are continuous on $W^{1,p}(\Omega)$, cf. Remark 4.B.4.

In many examples, one can use the following theorem to exclude eventually negative behaviour. The assumption $\operatorname{dom}(A^m) \subseteq I$ and the conclusion $\operatorname{dom}(A) \subseteq I$ is closely related to Sobolev embedding theorems, as we shall see in Example 6.4.6.

[7] This slightly strange assumption is merely due to the fact that we did not define complex vector lattices.
[8] Or more generally, a compact Hausdorff topological space.

Theorem 6.4.4. *Let $A\colon E \supseteq \operatorname{dom}(A) \to E$ be a closed operator on a complex Banach lattice E and let $\lambda_0 \in \mathbb{R}$ be an isolated spectral value of A. Let $I \subseteq E$ be an ideal (not necessarily closed) and assume:*

(1) $\operatorname{dom}(A^m) \subseteq I$ *for some* $m \in \mathbb{N}$.

(2) $\mathcal{R}(\cdot, A)$ *is uniformly eventually positive and negative with respect to* 0 *at* λ_0.

Then $\operatorname{dom}(A) \subseteq I$.

In Chapter 7 we will show that the theorem remains true, up to minor changes, if one only assumes individual eventual positivity and negativity in (2); but this requires a bit more preparation. In the rest of Chapter 6 we give the rather simple proof of Theorem 6.4.4 and discuss an example. We use the following finite expansion of the resolvent.

Lemma 6.4.5. *Let A be a closed operator on a complex Banach space X and $n \in \mathbb{N}_0$. Then*

$$\mathcal{R}(\lambda, A) = \sum_{k=1}^{n} (\mu-\lambda)^{k-1}\mathcal{R}(\mu, A)^k + \mathcal{R}(\lambda, A)(\mu-\lambda)^n\mathcal{R}(\mu, A)^n \qquad \text{for all } \lambda, \mu \in \rho(A).$$

Proof. This follows by iterating the resolvent identity (Proposition 3.3.2(c)). □

Proof of Theorem 6.4.4. By assumption (2) there are $\lambda, \mu \in \rho(A) \in \mathbb{R}$ that satisfy $\lambda < \lambda_0 < \mu$ such that $\mathcal{R}(\lambda, A) \le 0$ and $\mathcal{R}(\mu, A) \ge 0$. Abbreviating $S := \mathcal{R}(\lambda, A)(\mu-\lambda)^{m-1}\mathcal{R}(\mu, A)^{m-1}$, the finite resolvent expansion in Lemma 6.4.5 gives

$$0 \le -\mathcal{R}(\lambda, A) = -\sum_{k=1}^{m-1} (\mu-\lambda)^{k-1}\mathcal{R}(\mu, A)^k \; - \; S \le -S.$$

For every $f \in E_+$ we conclude that $0 \le -\mathcal{R}(\lambda, A)f \le -Sf \in I$, since S maps into $\operatorname{dom}(A^m)$ and thus, by assumption (1), into I. As I is an ideal in E, it follows that $\mathcal{R}(\lambda, A)f \in I$. Since E_+ spans E, this shows that $\operatorname{dom}(A) = \operatorname{rg}\mathcal{R}(\lambda, A) \subseteq I$, as claimed. □

Example 6.4.6. Let $\emptyset \ne \Omega \subseteq \mathbb{R}^n$ be open and bounded with C^{2m} boundary, where $m \in \mathbb{N}$ satisfies $m > \frac{n}{4}$, and consider the Dirichlet Laplacian Δ_{Dir} on $L^2(\Omega)$. If $n \ge 4$, the resolvent $\mathcal{R}(\cdot, \Delta_{\mathrm{Dir}})$ is not uniformly eventually negative with respect to 0 at $\mathrm{s}(\Delta_{\mathrm{Dir}})$.

Proof. We apply the contrapositive of Theorem 6.4.4, for the ideal $I := L^\infty(\Omega)$ in $L^2(\Omega)$. Assumption (1) of the theorem is satisfied: by iterating Theorem 5.3.2(b) one obtains $\operatorname{dom}\left(\Delta_{\mathrm{Dir}}^m\right) \subseteq H^{2m}(\Omega)$ for each $m \in \mathbb{N}$, so the Sobolev embedding theorem 5.3.4 ensures that $\operatorname{dom}\left(\Delta_{\mathrm{Dir}}^m\right) \subseteq L^\infty(\Omega)$ whenever $n < 4m$. Moreover, the first part of assumption (2) holds since $\mathcal{R}(\cdot, \Delta_{\mathrm{Dir}})$ is uniformly eventually positive with respect to 0 at $\mathrm{s}(\Delta_{\mathrm{Dir}})$ by Example 5.4.2.

To see that $\mathcal{R}(\cdot, \Delta_{\mathrm{Dir}})$ is not uniformly eventually negative with respect to 0 at $\mathrm{s}(\Delta_{\mathrm{Dir}})$, we observe that the conclusion of theorem is not satisfied. Indeed, it is a classical fact that $H^2(\Omega) \cap H^1_0(\Omega)$ does not embed into $L^\infty(\Omega)$ for $n \ge 4$, i.e. $\operatorname{dom}(\Delta_{\mathrm{Dir}}) \not\subseteq L^\infty(\Omega)$; for instance, the famous example mentioned in Remark 5.B.5(a) can be modified to ensure that the Dirichlet boundary conditions are satisfied. □

Exercises for Chapter 6

Exercise 6.1. Give an example of a closed operator A on a complex Banach space X and a set $\emptyset \neq \sigma_0 \subsetneq \sigma(A)$ such that σ_0 and $\sigma(A) \setminus \sigma_0$ are both compact yet the spectral projections corresponding to σ_0 and $\sigma(A) \setminus \sigma_0$ do not add up to id_X.

Exercise 6.2. Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be an operator on a Banach space X.

(a) Let $m, n \in \mathbb{N}_0$ and $x \in \operatorname{dom}\left(A^{m+n}\right)$. Show that $x \in \operatorname{dom}(A^n)$, $A^n x \in \operatorname{dom}(A^m)$, and

$$A^m(A^n x) = A^{m+n} x.$$

(b) Let $\lambda \in \mathbb{C}$. Show that $\operatorname{dom}((\lambda - A)^n) = \operatorname{dom}(A^n)$ for all $n \in \mathbb{N}$.

Give an example of

(c) an operator A on a Banach space X such that $\operatorname{dom}\left(A^2\right) \neq \operatorname{dom}(A)$;

(d) a closed operator A on a Banach space X such that A^2 is not closed.

Exercise 6.3 (A characterisation of $W_0^{1,p}(I)$)**.** Let $p \in [1, \infty)$ and let $a, b \in \mathbb{R}$, $a < b$. In this exercise we show that

$$W_0^{1,p}(a,b) = \left\{u \in W^{1,p}(a,b) : u(a) = u(b) = 0\right\}$$

This is the one-dimensional version of a result that we have mentioned multiple times.

(a) Show the inclusion "$\subseteq$" using the continuity of the embedding $W^{1,p}(a,b) \hookrightarrow \mathrm{C}([a,b])$.

(b) Show the inclusion "$\supseteq$". You may use that there is a function $w \in \mathrm{C}^\infty([a,b])$ that is constantly 0 in some right neighbourhood of a and constantly 1 in some left neighbourhood of b.

Hint: For $u \in W^{1,p}(a,b)$ with $u(a) = u(b) = 0$, take test functions φ_n that converge in L^p to u'. Then consider the functions $(1-w)u_n + w v_n$, where

$$u_n(x) := \int_a^x \varphi_n(s)\,\mathrm{d}s \qquad \text{and} \qquad v_n(x) := \int_b^x \varphi_n(s)\,\mathrm{d}s$$

for all $x \in (a,b)$.

Exercise 6.4. Consider the operator $A\colon \mathrm{C}([0,1]) \supseteq \operatorname{dom}(A) \to \mathrm{C}([0,1])$ given by

$$\begin{aligned} \operatorname{dom}(A) &:= \{u \in \mathrm{C}^1([0,1]) : u'(0) = u'(1)\}, \\ Au &= u'. \end{aligned}$$

(a) Prove that A is closed, densely defined, and has compact resolvent.[9]

(b) Compute all eigenvalues of A.

(c) Compute the eigenspace $\ker A$ and the generalised eigenspace $\bigcup_{n\in\mathbb{N}} \ker A^n$. Show that both spaces are spanned by positive vectors.

(d) Compute the eigenspace $\ker A'$ and the generalised eigenspace $\bigcup_{n\in\mathbb{N}} \ker(A')^n$ of the dual operator A' on $\mathrm{C}([0,1])'$ (Definition 3.1.5).

Hint: First use Theorem 6.2.6(b) and (d) to determine $\dim \bigcup_{n\in\mathbb{N}} \ker(A')^n$.

(e) Is $\mathcal{R}(\cdot, A)$ individually eventually positive with respect to 0 at the spectral value 0? Is it individually eventually negative with respect to 0 at 0?

Exercise 6.5.

(a) Show that the embedding $W^{1,1}(0,1) \hookrightarrow \mathrm{C}([0,1])$ is not compact.

(b) Show that the embedding $W^{1,1}(0,1) \hookrightarrow L^1(0,1)$ is compact.

Hint: By the Fréchet–Kolmogorov compactness theorem, a subset $F \subseteq L^1(0,1)$ is relatively compact if and only if

$$\sup_{f\in F} \int_{(0,1)} \left| f(s+h) - f(s) \right| \,\mathrm{d}s \to 0$$

as $h \to 0$. Here, one extends each f by 0 outside of $(0,1)$ to always make sense of the integral.

[9] Beware that it does not suffice for the compactness of the resolvent to note that the embedding $\mathrm{C}^1([0,1]) \hookrightarrow \mathrm{C}([0,1])$ is compact.

Notes for Chapter 6

Eventual positivity and negativity and the (anti-)maximum principle

As demonstrated in Example 5.3.6, positivity of $\mathcal{R}(\cdot, \Delta_{\text{Dir}})$ is closely related to the classical maximum principle. Inspired by this link – which holds also for more general second order operators with a variety of boundary conditions – some authors simply use the term **maximum principle** to refer to what we call eventual positivity of the resolvent (or versions thereof). Consequently, eventual negativity of the resolvent is then referred to as an **anti-maximum principle**, for instance by Clément & Sweers [CS00, CS01] and Grunau & Sweers [SG01]. An abstract operator theoretic approach to anti-maximum principles is due to Takáč [Tak96].

The notion **eventually positive resolvent** was used in [DGK16b, DGK16a], where it mainly served as a tool to better understand eventually positive operator semigroups – a topic that we will discuss in later chapters. Example 6.1.2 is taken from [DGK16b, Example 5.7]. A similar example where the operator has compact resolvent can be found in [DGK16b, Example 5.8]. Theorem 6.3.3 is, up to a few modifications, taken from [DG17, Theorem 3.1], which in turn generalised results from [DGK16a]. The idea for Theorem 6.4.4 stems from [AG22, AG23], although we now presented it in a somewhat different perspective. It is remarkable that the positivity and negativity of solutions to the equation $(\lambda - A)u = f$ for λ in a neighbourhood of the spectral value λ_0, together with the a priori regularity assumption $\operatorname{dom}(A^m) \subseteq I$, leads to improved regularity of such solutions, which is captured by the property $\operatorname{dom}(A) \subseteq I$.

Appendices

6.A Vector-valued analytic functions

This appendix collects some basic facts on complex-analytic mappings taking values in Banach spaces. We will mainly apply the concepts and results from this appendix to the resolvents of linear operators, which are analytic maps according to Proposition 6.2.1.

Definition 6.A.1. Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open and let X be a complex Banach space. A function $f : \Omega \to X$ is called **analytic** (or **holomorphic**) if

$$f'(z_0) := \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists in X for each $z_0 \in \Omega$.

Note that analytic implies continuous and in fact, as in the scalar-valued case, if f is analytic, then so is f'. Hence, one can iteratively define the the k-th derivatives $f^{(k)}$ for $k \in \mathbb{N}_0$ (with the usual convention $f^{(0)} := f$). Moreover, if $f : \Omega \to X$ is analytic, then it is also **weakly analytic**, i.e. $x' \circ f$ is analytic for each $x' \in X'$ with

$$(x' \circ f)'(z_0) = \langle x', f'(z_0) \rangle .$$

It turns out that analyticity and weak analyticity are equivalent notions. In fact, even more is true (Theorem 6.A.6).

As in the scalar-valued case, contour integrals are an important concept for vector-valued analytic functions. They are defined in precisely the same way, where the occurring integral is a Bochner integral. To describe spectral projections in full generality, one needs the contour integral not only over closed C^1-curves, but also over formal sums thereof – these are called C^1**-cycles**. For a C^1-cycle γ and a point $z_0 \in \mathbb{C}$ we write, by slight abuse of notation, $z_0 \not\in \gamma$ to say that z_0 does not lie in the image of any of the curves of γ.

A contour integral over a cycle γ is naturally defined as the sum over the contour integrals of all the curves it contains. The **winding number** of γ around a point $z_0 \not\in \gamma$ is thus also defined. We say that γ **encircles** z_0 **once** if this winding number is 1 and we say that γ **does not encircle** z_0 if the winding number if 0.

Proposition 6.A.2 (Cauchy's integral theorem for vector-valued functions)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, let X be a complex Banach space, and let $f : \Omega \to X$ be analytic. Let γ be a C^1-cycle in Ω that does not encircle any point in $\mathbb{C} \setminus \Omega$.*

(a) *One has* $\oint_\gamma f(z)\,\mathrm{d}z = 0$.

(b) *If* $z_0 \in \Omega$ *satisfies* $z_0 \notin \gamma$ *and is encircled once by* γ, *then*

$$f(z_0) = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(z)}{z - z_0}\,\mathrm{d}z.$$

The identity theorem for vector-valued analytic functions can be worded as follows.

Proposition 6.A.3 (Identity theorem for analytic functions)**.** *Let* $\emptyset \neq \Omega \subseteq \mathbb{C}$ *be open and connected,* X *be a Banach space,* $Y \subseteq X$ *a closed vector subspace, and let* $f : \Omega \to X$ *be analytic. Let* $(z_n) \subseteq \Omega$ *be a convergent sequence such that* $(f(z_n)) \subseteq Y$. *If* $\lim_{n\to\infty} z_n \in \Omega$, *then* $f(z) \in Y$ *for all* $z \in \Omega$.

This follows from the scalar valued identity theorem by testing against elements of X'. Note that the 'usual' statement of the identity theorem is recovered by taking $Y = \{0\}$.

Analogous to the scalar-valued case, we obtain Taylor and Laurent series expansions for vector-valued analytic functions.

Theorem 6.A.4 (Taylor expansion)**.** *Let* $\emptyset \neq \Omega \subseteq \mathbb{C}$ *be open let* X *be a complex Banach space, and let* $f : \Omega \to X$. *The following are equivalent:*

(i) *The function* f *is analytic on* Ω.

(ii) *For each* $z_0 \in \mathbb{C}$, *there exist* $(a_k) \subseteq X$ *and radius* $r \in (0, \operatorname{dist}(z_0, \partial\Omega))$ *such that* $f(z) = \sum_{k=0}^\infty a_{k+1}(z - z_0)^k$ *for all* $z \in B_{\leq r}(z_0)$, *with absolute uniform convergence on this disk.*

(iii) *For each* $z_0 \in \mathbb{C}$ *there exists* $(a_k) \subseteq X$ *such that* $f(z) = \sum_{k=0}^\infty a_{k+1}(z - z_0)^k$ *for all* $z \in B_{<\operatorname{dist}(z_0,\partial\Omega)}(z_0)$, *with absolute uniform convergence on compact subsets of this disk.*

If these equivalent conditions are satisfied, the coefficients $a_1, a_2, \ldots$ *in* (ii) *and* (iii) *do not depend on the choice of* r *and are given by*

$$a_{k+1} = \frac{1}{2\pi \mathrm{i}} \oint_{|z-z_0|=r} \frac{f(z)}{(z-z_0)^{k+1}}\,\mathrm{d}z = \frac{f^{(k)}(z_0)}{k!}$$

for all $k \in \mathbb{N}_0$ *and any* $0 < r < \operatorname{dist}(z_0, \partial\Omega)$.

Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, let X be a complex Banach space and let $f : \Omega \setminus \{z_0\} \to X$ be analytic. As in the scalar-valued case, a point $z_0 \in \mathbb{C}$ is called a **pole** of f if z_0 is an isolated point of $\mathbb{C} \setminus \Omega$ and there exists a $p \in \mathbb{N}$ such that $\lim_{z \to z_0} (z - z_o)^p f(z)$ exists (in norm) and is non-zero. In this case, p is uniquely determined and is called the **order** of the pole z_0.

Theorem 6.A.5 (Laurent expansion)**.** *Let* $\emptyset \neq \Omega \subseteq \mathbb{C}$ *be open, let* X *be a complex Banach space, let* $z_0 \in \Omega$ *and let* $f : \Omega \setminus \{z_0\} \to X$ *be analytic.*

(a) *Then there exists* $r > 0$ *such that* $B_{\le r}(z_0) \setminus \{z_0\} \subseteq \Omega$

$$f(z) = \sum_{k=-\infty}^{\infty} a_{k+1}(z - z_0)^k \qquad \text{for all } z \in B_{\le r}(z_0) \setminus \{z_0\}$$

with absolute uniform convergence on compact subsets of the disk, where

$$a_{k+1} = \frac{1}{2\pi \mathrm{i}} \oint_{|z-z_0|=r} \frac{f(z)}{(z - z_0)^{k+1}} \,\mathrm{d}z \qquad \text{for all } k \in \mathbb{Z}.$$

(b) z_0 *is a pole of* f *if and only if there exists* $p \in \mathbb{N}$ *such that the coefficients* a_{k+1} *from* (a) *satisfy* $a_{-p+1} \neq 0$ *and* $a_{-k+1} = 0$ *for all* $k > p$. *In this case,* p *is the order of the pole.*

Finally, we expand a bit more on the connection between analyticity and weak analyticity. We call a function $f\colon M \to X$ from a metric space M into a Banach space X **locally bounded** if every point $z_0 \in M$ has a neighbourhood U such that $\sup_{z\in U} \|f(z)\| < \infty$. Recall that a vector subspace Y of a dual Banach space X' is weak*-dense in X' if and only if it is **separating**, i.e. $\langle x', x\rangle = 0$ for a vector $x \in X$ and all $x' \in Y$ implies $x = 0$.

Theorem 6.A.6. *Let* $\emptyset \neq \Omega \subseteq \mathbb{C}$ *be open, let* X *be a Banach space, and let* $f\colon \Omega \to X$ *be locally bounded. The following are equivalent.*

(i) f *is analytic.*

(ii) $x' \circ f$ *is analytic for all* x' *in a weak*-dense subspace of* X'.

Note that if the condition (ii) in the theorem is satisfied for all $x' \in X'$, then the local boundedness of f follows automatically from the uniform boundedness theorem. For the proof we refer for instance to [ABHN11, Theorem A.7].

Corollary 6.A.7. *Let* $\emptyset \neq \Omega \subseteq \mathbb{C}$ *be open, let* X, Y *be Banach spaces, and let* $f : \Omega \to \mathcal{L}(X, Y)$ *be locally bounded. The following are equivalent.*

(i) f *is analytic.*

(ii) $f(\cdot)x$ *is analytic for all* x *in a dense subspace of* X.

(iii) $\langle y', f(\cdot)x\rangle$ *is analytic for all* x *in a dense subspace of* X *and all* y' *in a weak*-dense subspace of* Y'.

As before, we note that if condition (ii) in the corollary is satisfied for all $x \in X$, or if (iii) is satisfied for all $x \in X$ and for all $y' \in Y'$, then the local boundedness assumption on f is automatically satisfied due to the uniform boundedness principle.

Encore: if you want to know more...

6.B Isolated singularities of the resolvent

In this supplemental section, we expand a bit on the background of Theorem 6.2.6. The essence of the ideas is in the following Theorem 6.B.1, which analyses the coefficients of the Laurent series expansion of the resolvent around an isolated spectral value. Thus, the missing parts of the proof of Theorem 6.2.6 are then deduced in Theorem 6.B.3 below.

For a closed linear operator $A : X \supseteq \operatorname{dom}(A) \to X$ on a complex Banach space X, we use the notation $\operatorname{dom}(A^\infty) := \bigcap_{j\in\mathbb{N}_0} \operatorname{dom}\left(A^j\right)$.

Theorem 6.B.1 (Isolated singularities of the resolvent)**.** *Let $A : X \supseteq \operatorname{dom}(A) \to X$ be a closed linear operator on a complex Banach space X and let $\lambda \in \sigma(A)$ be an isolated point in $\sigma(A)$. Let*

$$\mathcal{R}(\mu, A) = \sum_{k=-\infty}^{\infty} Q_{k+1}(\mu-\lambda)^k.$$

denote the Laurent series expansion of the resolvent about λ. Then the operators $Q_k \in \mathcal{L}(X)$ commute, and the following assertions hold:

(a) *For each $k \in \mathbb{Z}$ one has $\operatorname{rg} Q_k \subseteq \operatorname{dom}(A)$ and $Q_k Ax = AQ_k x$ for all $x \in \operatorname{dom}(A)$.*

(b) *For all $k \in \mathbb{N}$ one has*

$$Q_k = (-1)^{k+1} Q_1^k \qquad \text{and} \qquad Q_{-k} = (Q_{-1})^k.$$

(c) *$Q_k Q_{-\ell} = 0$ for all $k, \ell \in \mathbb{N}$.*

(d) *Q_0 is a projection and satisfies*

$$Q_0 Q_k = 0 \qquad \text{and} \qquad Q_0 Q_{-k} = Q_{-k} \qquad \text{for all } k \in \mathbb{N}.$$

(e) *For every $k \in \mathbb{Z}$ one has*

$$(\lambda - A)Q_k = \begin{cases} -Q_{k-1}, & \text{if } k \neq 1, \\ \operatorname{id} - Q_0, & \text{if } k = 1. \end{cases}$$

(f) *The spectral radius $\operatorname{r}(Q_{-1})$ is 0, and hence $\operatorname{r}(Q_{-k}) = 0$ for all $k \in \mathbb{N}$.*

(g) *If* $q \geq 1$ *and* $Q_{-q} = 0$, *then* λ *is a pole of the resolvent, and its pole order is at most* q.

(h) *If* $d := \dim(\operatorname{rg} Q_0) < \infty$, *then* λ *is a pole of the resolvent of order at most* d.

(i) $\operatorname{rg} Q_{-k} \subseteq \operatorname{dom}(A^\infty)$ *for all* $k \in \mathbb{N}_0$.

Proof. First, we observe, according to Theorem 6.A.5(a), that

$$Q_k = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{\mathcal{R}(\mu, A)}{(\mu - \lambda)^k} \, \mathrm{d}\mu \tag{6.B.1}$$

for all $k \in \mathbb{Z}$, where γ denotes any sufficiently small circle about λ which is oriented anticlockwise. Since the resolvent operators of A all commute, this readily implies that the Q_k mutually commute.

(a) Endow $\operatorname{dom}(A)$ with a graph norm $\|\cdot\|_A$. This renders $\operatorname{dom}(A)$ a Banach space, since A is closed.

For every $\lambda \in \rho(A)$, the operator $\mathcal{R}(\lambda, A) : X \to \operatorname{dom}(A)$ is continuous by the closed graph theorem. Moreover, the mapping $\mathcal{R}(\cdot, A) : \rho(A) \to \mathcal{L}(X, \operatorname{dom}(A))$ is continuous; this follows from the preceding sentence together with the fact that the mapping is continuous with values in $\mathcal{L}(X)$ and the resolvent identity. Consequently, by Example 4.A.9, the integrand in (6.B.1) is Bochner integrable[10] with values in $\mathcal{L}(X, \operatorname{dom}(A))$. As the latter space embeds continuously into $\mathcal{L}(X)$, it follows that the integrals in both spaces coincide. Hence, $Q_k \in \mathcal{L}(X, \operatorname{dom}(A))$ and $Q_k X \subseteq \operatorname{dom}(A)$ for each $k \in \mathbb{Z}$.

For every $x \in \operatorname{dom}(A)$ and every $\mu \in \rho(A)$, one has $A\mathcal{R}(\mu, A)x = \mathcal{R}(\mu, A)Ax$. By applying the equality (6.B.1), and together with the facts that the integral in this equality can be interpreted as a Riemann integral in $\mathcal{L}(X, \operatorname{dom}(A))$ and that A is continuous from $\operatorname{dom}(A)$ to X, we thus obtain $Q_k A x = A Q_k x$ for every $k \in \mathbb{Z}$.

(b), (c), *and* (d) Let $k_1, k_2 \in \mathbb{Z}$ and consider formula (6.B.1) for two sufficiently small concentric circles γ_1, γ_2 with centre λ, where γ_2 has larger radius than γ_1. Then

$$\begin{aligned}
Q_{k_2} Q_{k_1} &= \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_2} \oint_{\gamma_1} \frac{\mathcal{R}(\mu_2, A)\mathcal{R}(\mu_1, A)}{(\mu_2 - \lambda)^{k_2} (\mu_1 - \lambda)^{k_1}} \, \mathrm{d}\mu_1 \, \mathrm{d}\mu_2 \\
&= \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_2} \oint_{\gamma_1} \frac{\mathcal{R}(\mu_2, A) - \mathcal{R}(\mu_1, A)}{(\mu_1 - \mu_2)(\mu_2 - \lambda)^{k_2} (\mu_1 - \lambda)^{k_1}} \, \mathrm{d}\mu_1 \, \mathrm{d}\mu_2 \\
&= \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_2} \frac{\mathcal{R}(\mu_2, A)}{(\mu_2 - \lambda)^{k_2}} \oint_{\gamma_1} \frac{1}{(\mu_1 - \mu_2)(\mu_1 - \lambda)^{k_1}} \, \mathrm{d}\mu_1 \, \mathrm{d}\mu_2 \\
&\quad - \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_1} \frac{\mathcal{R}(\mu_1, A)}{(\mu_1 - \lambda)^{k_1}} \oint_{\gamma_2} \frac{1}{(\mu_1 - \mu_2)(\mu_2 - \lambda)^{k_2}} \, \mathrm{d}\mu_2 \, \mathrm{d}\mu_1
\end{aligned}$$

By employing the residue theorem, one can compute the above integrals (where one has to distinguish several cases based on the signs of k_1 and k_2) and thus obtain the formulas claimed in (b), (c), and (d). We omit the computations.

[10] In fact, the integral also exists as a Riemann integral.

(e) Fix $k \in \mathbb{Z}$. For every $\mu \in \rho(A)$ one has $(\lambda - A)\mathcal{R}(\mu, A) = (\lambda - \mu)\mathcal{R}(\mu, A) + \mathrm{id}$. Since the integral in the formula (6.B.1) can be interpreted as a Riemann integral in $\mathcal{L}(X, \operatorname{dom}(A))$ and since $\lambda - A : \operatorname{dom}(A) \to X$ is continuous, it follows that

$$
\begin{aligned}
(\lambda - A)Q_k &= \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{(\lambda - \mu)\mathcal{R}(\mu, A) + \mathrm{id}}{(\mu - \lambda)^k} \,\mathrm{d}\lambda \\
&= \frac{1}{2\pi \mathrm{i}} \oint_\gamma -\frac{\mathcal{R}(\mu, A)}{(\mu - \lambda)^{k-1}} \,\mathrm{d}\lambda + \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{\mathrm{id}}{(\mu - \lambda)^k} \,\mathrm{d}\mu.
\end{aligned}
$$

The first summand above is $-Q_{k-1}$, and the second summand is equal to 0 if $k \neq 1$, and equal to id if $k = 1$.

(f) Let $\varepsilon > 0$ be sufficiently small, such that (6.B.1) holds for the circle γ with radius ε about λ. For every integer $k \geq 1$ it follows from (b) that

$$
\left\| Q_{-1}^k \right\| = \|Q_{-k}\| \leq \frac{1}{2\pi} \oint_\gamma \frac{\left\|\mathcal{R}(\mu, A)\right\|}{\varepsilon^{-k}} \,\mathrm{d}\left|\mu\right| \leq \sup_{\lambda \in \gamma} \left\|\mathcal{R}(\mu, A)\right\| \varepsilon^{k+1}.
$$

We now use the spectral radius formula: by taking the k-th root and letting $k \to \infty$, we thus see that $\mathrm{r}(Q_{-1}) \leq \varepsilon$. This shows that $\mathrm{r}(Q_{-1}) = 0$, as claimed.

As $Q_{-k} = Q_{-1}^k$ for all $k \geq 1$, it follows from the spectral mapping theorem for polynomials that $\mathrm{r}(Q_{-k}) = 0$ for all $k \geq 1$ as well.

(g) If $Q_{-q} = 0$, then it follows from (b) that $Q_{-(q+j)} = Q_{-q}Q_{-j} = 0$ for all $j \geq 0$ as well; this shows the claim.

(h) The operator Q_{-1} commutes with Q_0, so it leaves the range of Q_0 invariant. Moreover, Q_{-1} has spectral radius 0, so its restriction to $\operatorname{rg} Q_0$ is nilpotent; more precisely, the d-th power of this restriction is 0. Hence, $Q_{-d} = Q_{-1}^d = Q_{-1}^d Q_0 = 0$, so according to (g), λ is indeed a pole of order at most d.

(i) Fix an integer $k \leq 0$. We show by induction over n that $Q_{-k}X \subseteq \operatorname{dom}(A^n)$ for each $n \in \mathbb{N}$. In (a) we proved the claim for $n = 1$, so assume now that the claim holds for some $n \in \mathbb{N}$. Let $x \in X$. It follows from (d) that $Q_{-k} = Q_{-k}Q_0$, and from (a) that $\operatorname{rg} Q_0 \subseteq \operatorname{dom}(A)$. Hence,

$$
A^n Q_{-k} x \in A^n Q_{-k}(\operatorname{rg} Q_0) = A^{n-1} Q_{-k} A(\operatorname{rg} Q_0),
$$

where we used the formula from (a) for the equality (which is possible since $\operatorname{rg} Q_0 \subseteq \operatorname{dom}(A)$). Since $Q_{-k}X \subseteq \operatorname{dom}(A^n)$ by the induction hypothesis, it follows that one has $A^{n-1}Q_{-k}A(\operatorname{rg} Q_0) \subseteq \operatorname{dom}(A)$. Thus we have shown that $A^n Q_{-k} x \in \operatorname{dom}(A)$, which implies $Q_{-k}x \in \operatorname{dom}\left(A^{n+1}\right)$ as claimed. □

Note that property (b) in Theorem 6.B.1 implies that λ is a pole of $\mathcal{R}(\cdot, A)$ if and only if Q_{-1} is nilpotent. In this case, the pole order is the smallest integer $q \geq 1$ such that $(Q_{-1})^q = 0$.

Remark 6.B.2 (The action of Q_{-1})**.** In the situation of Theorem 6.B.1 it follows from (d) and (e) that

$$Q_{-1}Q_0 = Q_{-1} = (A-\lambda)Q_0.$$

In other words, on $\operatorname{rg} Q_0$ the operator Q_{-1} acts as the operator $A-\lambda$.

Let us now demonstrate how Theorem 6.2.6 in the main text follows from Theorem 6.B.1. Assertions (c) and (d) of Theorem 6.2.6 have already been shown in the main text, so we focus on (a) and (b) here. For easier reference from within the proof, we state those two parts of the theorem here again.

Theorem 6.B.3 (Poles of the resolvent)**.** *Let $A\colon X \supseteq \operatorname{dom}(A) \to X$ be a closed operator on a complex Banach space X and let $\lambda \in \sigma(A)$ be a pole of the resolvent $\mathcal{R}(\cdot, A) : \rho(A) \to \mathcal{L}(X)$ of order $p \in \mathbb{N}$. Let*

$$\mathcal{R}(\mu, A) = \sum_{k=-p}^{\infty} Q_{k+1}(\mu-\lambda)^k$$

denote the Laurent series expansion of the resolvent about λ with coefficients $Q_{k+1} \in \mathcal{L}(X)$ and with $Q_{-p+1} \neq 0$.

(a) *One has $\{0\} \neq \operatorname{rg} Q_{-p+1} \subseteq \ker(\lambda - A)$. In particular, λ is an eigenvalue of A.*

(b) *One has $\operatorname{rg} Q_0 = \ker(\lambda - A)^k$ for all $k \geq p$, so $\operatorname{rg} Q_0$ is the generalised eigenspace of A for the eigenvalue λ. Thus, the eigenvalue λ is semisimple if and only if $p = 1$.*

Proof. (a) Since Q_{-p+1} is non-zero, so is its range. Moreover, Q_{-p+1} maps into $\operatorname{dom}(A)$ according to Theorem 6.B.1(e), and part (e) of the same theorem shows that $(\lambda - A)Q_{-p+1} = -Q_{-p} = 0$.

(b) Fix $k \geq p$. It follows from Theorem 6.B.1(e) that

$$(\lambda - A)^k Q_0 = (-1)^k Q_{-k} = 0,$$

so $\operatorname{rg} Q_0 \subseteq \ker\big((\lambda - A)^k\big)$. Conversely, let $x \in \ker\big((\lambda - A)^k\big)$. Since Q_0 is a projection, so is $\operatorname{id} - Q_0$, and hence it follows from Theorem 6.B.1(e) that

$$(\operatorname{id} - Q_0)x = (\operatorname{id} - Q_0)^k x = Q_1^k(\lambda - A)^k x = 0.$$

This proves that $x \in \operatorname{rg} Q_0$.[11] ☐

A classical example of an operator with an isolated spectral value that is not a pole of the resolvent, is the Volterra operator:

[11] Note that this argument could just as well be used to show directly that $\ker\big((\lambda - A)^k\big) \subseteq \operatorname{rg} Q_0$ for every $k \geq 1$ rather than just for every $k \geq p$. However, this is not important here, since the spaces $\ker\big((\lambda - A)^k\big)$ are increasing with respect to k anyway.

Example 6.B.4 (An essential singularity of the resolvent)**.** Consider the **Volterra operator**, $A\colon \mathrm{C}([0,1]) \to \mathrm{C}([0,1])$ defined by

$$Af(x) := \int_0^x f(t)\,\mathrm{d}t, \qquad x \in [0,1].$$

It is straightforward to show by induction and Fubini's theorem that

$$A^n f(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,\mathrm{d}t, \qquad x \in [0,1].$$

This formula easily yields the operator norm estimate

$$\|A^n\|_{\mathcal{L}(\mathrm{C}([0,1]))} \le \frac{1}{n!} \qquad \forall\, n \in \mathbb{N}.$$

Since $(n!)^{1/n} \to \infty$, the spectral radius formula yields

$$\mathrm{r}(A) = \lim_{n\to\infty} \|A^n\|^{1/n} \le \lim_{n\to\infty} \frac{1}{(n!)^{1/n}} = 0,$$

and thus we have shown that $\sigma(A) = \{0\}$.

The spectral value 0 is not an eigenvalue. Indeed, the operator A is injective: if $Af = 0$ for some $f \in \mathrm{C}([0,1])$, then we may differentiate the equation to find that $f = 0$. However, 0 is an essential singularity of the resolvent of A. To see this, we observe that $\lambda - A$ is invertible for every $\lambda \in \mathbb{C} \setminus \{0\}$ and $\mathcal{R}(\lambda, A)$ is represented by the Neumann series:

$$\mathcal{R}(\lambda, A) = \lambda^{-1}(\mathrm{id} - \lambda^{-1}A)^{-1} = \sum_{k=0}^{\infty} A^k \lambda^{-(k+1)}.$$

Therefore, in the notation of Theorem 6.B.1, we have $Q_{-k} = A^k$ for all $k \in \mathbb{N}_0$, and the Laurent expansion of $\mathcal{R}(\lambda, A)$ around 0 has infinitely many non-zero terms in its singular part.

6.C Positivity of leading eigenvectors via sesquilinear forms

We have seen in Section 5.1 that sesquilinear forms are quite useful for constructing linear operators and establishing properties of their spectrum. Moreover, the Beurling–Deny criterion (Theorem 5.1.7) shows that they can be used to characterise positivity of the resolvent of an operator everywhere on the right of its spectral bound. In this supplement, we show how this line of thought can be developed further, in particular for symmetric forms. We first establish how the **numerical range** of a form is related to the spectral bound of the associated operator A. Then we show that for symmetric forms which satisfy the Beurling–Deny criterion, positivity of a leading eigenvector can also be established directly by form methods as an alternative to Theorem 6.3.3.

Definition 6.C.1 (Numerical range of a sesquilinear form)**.** Let V be a complex Hilbert space and $\mathfrak{a}\colon V \times V \to \mathbb{C}$ be a sesquilinear form. The set

$$\mathrm{W}(\mathfrak{a}) := \left\{\mathfrak{a}(u,u) : u \in V,\ \|u\|_H = 1\right\} \subseteq \mathbb{C}$$

is called the **numerical range of** $\mathfrak{a}$.

Recall again from Theorem 5.1.4(c) that the form $\mathfrak{a}$ is called **symmetric** if $\mathfrak{a}(u,v) = \overline{\mathfrak{a}(v,u)}$ for all $u, v \in V$. Let us note that $\mathfrak{a}$ is symmetric if and only if $\mathrm{W}(\mathfrak{a}) \subseteq \mathbb{R}$. Indeed, the symmetry implies that $\mathfrak{a}(u,u) = \overline{\mathfrak{a}(u,u)}$ and thus $\mathfrak{a}(u,u) \in \mathbb{R}$ for all $u \in V$. The converse implication follows from the polarisation identity for sesquilinear forms.

We will now prove a variety of results under the following general assumptions.

Setting 6.C.2. Let V, H be complex Hilbert spaces with the dense embedding $V \hookrightarrow H$. Let $\mathfrak{a}\colon V \times V \to \mathbb{C}$ be a bounded sesquilinear form which satisfies the ellipticity estimate

$$\operatorname{Re}\mathfrak{a}(u,u) + \mu \|u\|_H^2 \geq \delta \|u\|_V^2 \qquad \forall\, u \in V \tag{6.C.1}$$

for some $\mu \in \mathbb{R}$ and $\delta > 0$. Denote by $A\colon H \supseteq \operatorname{dom}(A) \to H$ the associated operator.

Lemma 6.C.3. *In Setting 6.C.2, assume that* $\operatorname{Re}\mathfrak{a}(u,u) \geq 0$ *for all* $u \in V$ *– i.e. the numerical range* $\mathrm{W}(\mathfrak{a})$ *is contained in the closed right half plane. Then for every* $\varepsilon > 0$*, there exists* $\delta_\varepsilon > 0$ *such that*

$$\operatorname{Re}\mathfrak{a}(u,u) + \varepsilon \|u\|_H^2 \geq \delta_\varepsilon \|u\|_V^2 \qquad \forall\, u \in V.$$

Proof. We proceed by contradiction. Assume that there exists $\varepsilon_0 > 0$ such that for every $n \in \mathbb{N}$, there exists $u_n \in V$ with $\|u_n\|_V = 1$ such that

$$\operatorname{Re}\mathfrak{a}(u_n,u_n) + \varepsilon_0 \|u_n\|_H^2 < \frac{1}{n} \qquad \forall\, n \in \mathbb{N}.$$

Since $\operatorname{Re}\mathfrak{a}(u_n,u_n) \geq 0$, it follows that $\|u_n\|_H \to 0$ and $\operatorname{Re}\mathfrak{a}(u_n,u_n) \to 0$ as $n \to \infty$. This contradicts the ellipticity estimate (6.C.1) because $\|u_n\|_V = 1$ for all n. ☐

Proposition 6.C.4. *In Setting 6.C.2, if* $\operatorname{Re}\mathfrak{a}(u,u) \geq 0$ *for all* $u \in V$*, then* $\mathrm{s}(A) \leq 0$.

Proof. For every $\varepsilon > 0$, it follows from Lemma 6.C.3 and Theorem 5.1.4(b) that $\mathrm{s}(A) \leq \varepsilon$, and thus $\mathrm{s}(A) \leq 0$. ☐

The first part of the following theorem yields a proof for Proposition 6.2.10(c).

Theorem 6.C.5. *In Setting 6.C.2, assume that* $\mathfrak{a}$ *is symmetric (and hence* $\sigma(A) \subseteq \mathbb{R}$ *by Theorem 5.1.4(c)). Then the following assertions hold:*

(a) $\mathrm{s}(A) = -\inf \mathrm{W}(\mathfrak{a}) \in \sigma(A)$*. In particular,* $\sigma(A) \neq \emptyset$.

(b) *Let* $v \in V$ *such that* $\|u\|_H = 1$*. Then* $\mathfrak{a}(v,v) = -\inf \mathrm{W}(\mathfrak{a})$ *if and only if* $-\inf \mathrm{W}(\mathfrak{a})$ *is an eigenvalue of* A *and* v *is a corresponding eigenvector.*[12]

[12] And hence, in particular, that $v \in \operatorname{dom}(A)$.

(c) *Every eigenvalue of A is semisimple.*

In the proof we use that if a sesquilinear form $\mathfrak{a}\colon V \times V \to \mathbb{C}$ is symmetric and satisfies $\mathfrak{a}(u,u) \geq 0$ for all $u \in V$ – equivalently, $\mathrm{W}(\mathfrak{a}) \subseteq [0,\infty)$ – then it satisfies the Cauchy–Schwarz inequality

$$|\mathfrak{a}(u,v)| \leq \mathfrak{a}(u,u)^{1/2}\mathfrak{a}(v,v)^{1/2}$$

for all $u, v \in V$. The proof is the same as for inner products.

Proof of Theorem 6.C.5. (a) We assume without loss of generality that $\inf \mathrm{W}(\mathfrak{a}) = 0$. As shown in Proposition 6.C.4, one then has $\mathrm{s}(A) \leq 0$, so it remains to show that 0 is a spectral value.

Assume to the contrary that $0 \in \rho(A)$. Then $\mathcal{R}(0,A)$ is a well-defined bounded linear operator from H to $\operatorname{dom}(A)$. Since $\operatorname{dom}(A)$ embeds continuously into V, it follows from the closed graph theorem that $\mathcal{R}(0,A) \in \mathcal{L}(V)$.

As $\inf \mathrm{W}(\mathfrak{a}) = 0$, we can find a sequence (v_n) in V such that $\|v_n\|_H = 1$ for all n and $\mathfrak{a}(v_n, v_n) \to 0$. Since $\mathfrak{a}$ is bounded, there exists a $c \geq 0$ such that, for all $n \in \mathbb{N}$, one has

$$\mathfrak{a}\big(\mathcal{R}(0,A)v_n, \mathcal{R}(0,A)v_n\big)^{1/2} \leq c\,\|\mathcal{R}(0,A)v_n\|_V \leq c\,\|\mathcal{R}(0,A)\|_{\mathcal{L}(V)} < \infty.$$

Thus,

$$\begin{aligned}\|v_n\|_H^2 &= -(v_n \mid A\mathcal{R}(0,A)v_n)_H = \mathfrak{a}(v_n, \mathcal{R}(0,A)v_n)\\ &\leq \mathfrak{a}(v_n,v_n)^{1/2}\mathfrak{a}\big(\mathcal{R}(0,A)v_n, \mathcal{R}(0,A)v_n\big)^{1/2} \leq \mathfrak{a}(v_n,v_n)^{1/2} c\,\|\mathcal{R}(0,A)\|_{\mathcal{L}(V)}\end{aligned}$$

where the first inequality uses the Cauchy–Schwarz inequality mentioned before the proof. Hence, $\mathfrak{a}(v_n, v_n) \to 0$ implies that $\|v_n\|_H^2 \to 0$, which is absurd.

(b) As in (a), we assume without loss of generality that $\inf \mathrm{W}(\mathfrak{a}) = 0$.

If $v \in \operatorname{dom}(A)$ and $Av = 0$, then $\mathfrak{a}(v,v) = -(v \mid Av)_H = 0$. Now assume conversely that $\mathfrak{a}(v,v) = 0$, and let $w \in V$ be arbitrary. Again by the Cauchy–Schwarz inequality, we have

$$|\mathfrak{a}(w,v)| \leq \mathfrak{a}(w,w)^{1/2}\mathfrak{a}(v,v)^{1/2} = 0.$$

Hence $\mathfrak{a}(w,v) = 0$ for all $w \in V$, which implies $v \in \operatorname{dom}(A)$ and $Av = 0$.

(c) Let $\lambda \in \mathbb{R}$ be an eigenvalue of A. Now we shift the form so that, without loss of generality, $\lambda = 0$. Let $v \in \ker A^2$. It suffices to show that $v \in \ker A$. One has

$$\|Av\|_H^2 = (Av \mid Av)_H = -\mathfrak{a}(Av, v) = -\overline{\mathfrak{a}(v, Av)} = \overline{(v \mid A^2 v)}_H = 0,$$

where the penultimate equality uses that $Av \in \operatorname{dom}(A)$ and hence $Av \in V$. Thus $Av = 0$, as claimed. □

The assertion of Theorem 6.C.5(b) can be rephrased by saying that a non-zero vector $v \in V$ is in $\ker(\mathrm{s}(A) - A)$ if and only if it minimises the **Rayleigh quotient**, a nonlinear functional on V defined by

$$Q(v) := \frac{\mathfrak{a}(v,v)}{\|v\|_H^2}.$$

Now we give a form based proof for the existence of a positive eigenvector for $\mathrm{s}(A)$ that was promised at the beginning of this supplemental section.

Theorem 6.C.6. *In Setting 6.C.2, assume that $H = L^2(\Omega,\nu)$ for a σ-finite measure space (Ω,ν). Let $\mathfrak{a}$ be real and symmetric and assume that $\mathcal{R}(\lambda, A) \geq 0$ for all sufficiently large $\lambda > \mathrm{s}(A)$.*[13] *If $\mathrm{s}(A)$ is an eigenvalue of A, then it has a positive eigenvector.*

Proof. Without loss of generality, we assume $\mathrm{s}(A) = 0$ and hence, $\inf \mathrm{W}(\mathfrak{a}) = 0$ by Theorem 6.C.5(b). It follows from the Beurling–Deny criterion (Theorem 5.1.7) that the real part $V_{\mathbb{R}} := V \cap L^2(\Omega,\nu;\mathbb{R})$ of the form domain V is a vector sublattice of $L^2(\Omega,\nu;\mathbb{R})$ and that $\mathfrak{a}(v^-, v^+) \leq 0$ for all $v \in V_{\mathbb{R}}$.

Now let $0 \neq v \in \ker A$, and write $v = v_1 + \mathrm{i} v_2$ for real-valued functions v_1, v_2. Since the form $\mathfrak{a}$ is real, so is the operator A (Proposition 5.1.6) and hence, $v_1, v_2 \in \operatorname{dom}(A)$. Due to Theorem 6.C.5(b) we thus have

$$0 = \mathfrak{a}(v,v) = \mathfrak{a}(v_1,v_1) + \mathfrak{a}(v_2,v_2) + \mathrm{i}\mathfrak{a}(v_1,v_2) - \mathrm{i}\mathfrak{a}(v_2,v_1).$$

By taking real parts and using that $\inf \mathrm{W}(\mathfrak{a}) = 0$, one thus obtains $\mathfrak{a}(v_1,v_1) = \mathfrak{a}(v_2,v_2) = 0$. Since v is non-zero, so is at least one of the vectors v_1, v_2. Thus by Theorem 6.C.5(b), we have found a real eigenvector of A for the eigenvalue 0. Let us call this eigenvector v from now on. Again Theorem 6.C.5(b) gives

$$0 = \mathfrak{a}(v,v) = \mathfrak{a}(v^+,v^+) + \mathfrak{a}(v^-,v^-) - 2\mathfrak{a}(v^-,v^+) \geq \mathfrak{a}(v^+,v^+) + \mathfrak{a}(v^-,v^-),$$

where we use that $\mathfrak{a}$ is real to get $\operatorname{Re}\mathfrak{a}(v^-,v^+) = \mathfrak{a}(v^-,v^+)$. Hence $\mathfrak{a}(v^+,v^+) = \mathfrak{a}(v^-,v^-) = 0$ since $\inf \mathrm{W}(\mathfrak{a}) = 0$. At least one of the vectors v^+, v^- is non-zero and is thus in $\ker A$ by Theorem 6.C.5(b). ☐

In contrast to Theorem 6.3.3, observe that $\mathrm{s}(A)$ in Theorem 6.C.6 need not be an isolated spectral value.

[13] And hence for all $\lambda > \mathrm{s}(A)$ by Theorem 5.4.1.

Chapter 7

Criteria for eventual positivity of resolvents: the individual case

In Chapter 6, we introduced some flexibility in the definition of eventually positive resolvents (Definition 6.1.1), enabling us to consider both the weak inequality ≥ 0 as well as a lower bound $\succeq u$ with respect to some non-zero positive vector u. We refer to the latter case informally as **strong positivity**, in analogy with the property $\succeq \mathbb{1}$ in finite dimensions. The present and the subsequent chapter are devoted to a study of eventual strong positivity of resolvents.

7.1 Banach lattice overture: Principal ideals and quasi-interior points

It turns out that a useful generalisation of the constant vector $\mathbb{1} \in \mathbb{R}^n$ is the notion of a **quasi-interior point** in a Banach lattice.

Definition 7.1.1 (Principal ideals and quasi-interior points)**.** Let E be a Banach lattice and let $u \in E_+$.

(a) The set $E_u := \{x \in E : |x| \leq u\}$ is called the **principal ideal** in E generated by u.[1] The mapping $\|\cdot\|_{E_u} : E_u \to [0,\infty)$ given by

$$\|x\|_{E_u} := \inf\{c \in [0,\infty) : |x| \leq cu\}$$

is called the **gauge norm** with respect to u.

(b) The vector u is called a **quasi-interior point** of E_+ if E_u is dense in E.

If the surrounding Banach lattice and its positive cone are clear from context, we will sometimes say "u is a quasi-interior point" as a shorthand for "u is a quasi-interior point of E_+."

[1] Of course, principal ideals are ideals.

The following example is simple but very instructive.

Example 7.1.2. Let (Ω, ν) be a finite measure space, so that $\mathbb{1}_\Omega \in L^p(\Omega, \nu)$ for all $p \in [1, \infty]$. Then it is easy to see that $(L^p(\Omega, \nu))_{\mathbb{1}_\Omega} = L^\infty(\Omega, \nu)$ for all $p \in [1, \infty]$, and the corresponding gauge norm is precisely the essential supremum norm.

Note that if $p \neq \infty$, then $L^\infty(\Omega, \nu)$ is not a closed ideal in $L^p(\Omega, \nu)$.

Below, we use the not particularly surprising observation that an element x of a Banach lattice is 0 if $|x| = 0$. In the real case this follows from $0 \le x^+, x^- \le |x|$ and $x = x^+ - x^-$. In the complex case it then follows from Proposition 4.2.7.

Proposition 7.1.3. *Let E, F be Banach lattices over the same scalar field, let $u \in E_+$ be a quasi-interior point, and let $0 \le T \in \mathcal{L}(E, F)$. If $Tu = 0$, then $T = 0$.*

Proof. The positivity of T implies that $|Tx| \le T|x| \le Tu = 0$ (Proposition 4.3.2) and thus $Tx = 0$ for all $x \in E$ that satisfy $|x| \le u$. But the span of such x is dense in E owing to the fact that u is a quasi-interior point. □

Proposition 7.1.4. *Let E be a Banach lattice. A vector $u \in E_+$ is a quasi-interior point if and only if $\langle \psi, u \rangle > 0$ for each $0 \lneq \psi \in E'$.*

Proof. "⇒": This follows from Proposition 7.1.3.

"⇐": If u is not a quasi-interior point of E_+, then by the Hahn-Banach theorem, there exists $\psi \in E' \setminus \{0\}$ such that $\langle \psi, x \rangle = 0$ for each $x \in \overline{E_u}$. Thus, $0 \lneq |\psi| \in E'$ satisfies

$$\langle |\psi|, u \rangle = \sup_{|x| \le u} |\langle \psi, x \rangle| = 0$$

due to the Riesz-Kantorovich formula (Theorem 4.4.2). □

Examples 7.1.5. The following examples are discussed in detail in Exercise 7.1.

(a) Let (Ω, μ) be a σ-finite measure space. The quasi-interior points of $L^p(\Omega, \mu)_+$ for $p \in [1, \infty)$ are exactly those functions in $L^p(\Omega, \mu)$ that are strictly positive almost everywhere.

(b) Let (Ω, μ) be a σ-finite measure space. The quasi-interior points of $L^\infty(\Omega, \mu)_+$ are exactly those $f \in L^\infty(\Omega, \mu)$ that satisfy $f \succeq \mathbb{1}_\Omega$.

(c) Let K be a compact metric space[2]. A function $u \in C(K)_+$ is a quasi-interior point if and only if $u(x) > 0$ for all $x \in K$ if and only if $u \succeq \mathbb{1}_K$.

(d) Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open[3]. The quasi-interior points of $C_0(\Omega)_+$ are precisely those functions $u \in C_0(\Omega)$ that satisfy $u(x) > 0$ for all $x \in \Omega$.

[2] Or more generally, a compact Hausdorff topological space.

[3] Or more generally, let $\emptyset \neq \Omega$ be a locally compact Hausdorff topological space.

(e) In particular, in finite dimensions the quasi-interior points are exactly the strongly positive vectors; see Definition 1.2.3.

Proposition 7.1.6. *Let E be a Banach lattice and $u \in E_+$. The principal ideal E_u equipped with the gauge norm is a Banach lattice that embeds continuously into E.*

For the proof, one can first show the proposition in the real case and then derive the complex case from it. Writing down all the details is a bit tedious, but does not require any surprising ideas, so we refrain from doing so.

7.2 Strong positivity properties of the Dirichlet Laplacian

In this section, we revisit the maximum principle of Chapter 5, and demonstrate how it may be used to prove strong positivity properties for certain PDEs. For simplicity, we once again focus on the Dirichlet Laplacian as the leading example. While this material is part of classical PDE theory, it falls neatly into our abstract framework nonetheless.

Theorem 7.2.1 (Location of strict maxima)**.** *Let (M, d) be a metric space and let $\emptyset \neq S \subseteq M$ be relatively compact. Let $D \subseteq \mathrm{C}(\overline{S}; \mathbb{R})$ be a vector subspace such that $\mathbb{1} := \mathbb{1}_{\overline{S}} \in D$ and let $A\colon D \to \mathbb{R}^S$ be a linear map with the same properties as in Theorem 5.2.1, i.e.*

(1) *The map A satisfies the positive minimum principle on S, i.e. for each $x \in S$ and each function $0 \le u \in D$ one has the implication*

$$u(x) = 0 \qquad \Longrightarrow \qquad (Au)(x) \ge 0.$$

(2) *One has $A\mathbb{1} \le 0$ and there exists a function $0 \le w \in D$ with $(Aw)(x) > 0$ for all $x \in S$.*

(3) *Let $x_0 \in \partial S$ and assume that the function w from assumption (2) vanishes at all points in ∂S that are sufficiently close to x_0*

Let $v \in D$ attain at least one value in $[0, \infty)$ and satisfy $Av \ge 0$ in S. If v has a strict global maximum at x_0, then x_0 is not in S.

Proof. Assume to the contrary that $x_0 \in S$. Since, by assumption, v has a strict global maximum at x_0 and w vanishes at all points of ∂S that are close to x_0 (assumption (3)), we can find a number $\varepsilon > 0$ such that $v(x_0) \ge v(x) + \varepsilon w(x)$ for all $x \in \partial S$. Set $h := v + \varepsilon w - v(x_0)\,\mathbb{1}$. Then $h(x) \le 0$ for all $x \in \partial S$ and $h(x_0) = 0$. Moreover,

$$(Ah)(x) = (Av)(x) + \varepsilon (Aw)(x) - v(x_0)(A\mathbb{1})(x) \ge (Av)(x) + \varepsilon (Aw)(x) > 0$$

for all $x \in S$, where we used for the first inequality that $v(x_0) \ge 0$ and $(A\mathbb{1})(x) \le 0$. It follows from Theorem 5.2.1, applied to the function h, that h attains its maximum on ∂S. Since $h \le 0$ on ∂S, we conclude that $h \le 0$ in S.

Hence, the function $0 \le -h \in D$ satisfies $-h(x_0) = 0$ but $\big(A(-h)\big)(x_0) = -(Ah)(x_0) < 0$, contradicting the positive minimum principle (assumption (1)) at the point $x_0 \in S$. □

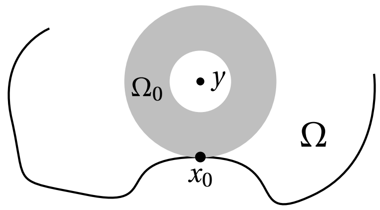

Figure 7.2.1: Geometric conditions in the Hopf lemma.

As an application of the previous theorem, we prove a simple version of the classical Hopf boundary lemma. It is perhaps surprising that we obtain this result in the same framework as for the classical maximum principle, and thus we state it as an example.

Example 7.2.2 (Hopf boundary lemma)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open[4] and let $c \leq 0$ be a real number. Let $v \in C^1(\overline{\Omega};\mathbb{R}) \cap C^2(\Omega;\mathbb{R})$ be such that $\Delta v(x) + cv(x) \geq 0$ for all $x \in \Omega$ and assume that $x_0 \in \partial\Omega$ has the following properties:

- **Strict maximum at x_0:** One has $v(x_0) \geq 0$ and $v(x_0) > v(x)$ for all $x \in \Omega$.
- **Interior ball condition:** There exists an open ball $B \subseteq \Omega$ with $x_0 \in \partial B$.

If ν denotes the outer unit normal of the ball B, then $\partial_\nu v(x_0) > 0$.

Proof. Geometric setup: By the interior ball condition, there exist $y \in \Omega$ and $R > 0$ such that $B_{<R}(y) \subset \Omega$ and $x_0 \in \partial B_{<R}(y)$. By decreasing R and moving y a bit towards x_0 if necessary, we may assume that $\partial\Omega \cap \partial B_{<R}(y) = \{x_0\}$. Choose an arbitrary $r \in (0,R)$ and consider the annular region $\Omega_0 = B_{<R}(y) \setminus B_{\leq r}(y)$; see Figure 7.2.1 for a visual aid. Set $S := \Omega_0 \cup \{x_0\}$.

Let D be the space of restrictions of all functions in $C^1(\overline{\Omega};\mathbb{R}) \cap C^2(\Omega;\mathbb{R})$ to $\overline{S}$. Clearly $\mathbb{1}_{\overline{S}} \in D$. Define a linear operator $A\colon D \to \mathbb{R}^S$ by

$$(Af)(x) := \begin{cases} \Delta f(x) + cf(x) & x \in \Omega_0 \\ -\partial_\nu f(x) & x = x_0 \end{cases} \qquad \forall\, f \in D.$$

We verify that the assumptions of Theorem 7.2.1 are satisfied:

(1) Suppose $0 \leq f \in D$ such that $f(x) = 0$ for some $x \in S$. We want to show $(Af)(x) \geq 0$ If $x = x_0$, then $(Af)(x) = -(\partial_\nu f)(x) \geq 0$ (observe that $-\partial_\nu$ is the inward normal derivative). On the other hand, if $x \in \Omega_0$, the conclusion $(Af)(x) \geq 0$ follows from the fact that the operator $\Delta + c$ satisfies the positive minimum principle in Ω_0 (verified in the proof of Example 5.2.2).

[4] Not necessarily bounded.

(2) Of course, $A\mathbb{1}_{\overline{S}} \le 0$. For all $x \in \overline{S}$, define $w\colon \overline{S} \to \mathbb{R}$ by

$$w(x) := e^{-\alpha\|x-y\|_2^2} - e^{-\alpha R^2}$$

for some $\alpha > 0$ to be determined later. We have $0 \le w \in D$, and since $c \le 0$, we deduce

$$Aw(x) = \left(4\alpha^2\,\|x-y\|_2^2 - 2\alpha n + c\right)e^{-\alpha\|x-y\|_2^2} - ce^{-\alpha R^2} \ge \left(4\alpha^2 r^2 - 2\alpha n + c\right)e^{-\alpha\|x-y\|_2^2}$$

for all $x \in \Omega_0$. Hence we can choose $\alpha > 0$ sufficiently large so that $(Aw)(x) > 0$ for all $x \in \Omega_0$. We further note that $\nu(x_0) = \frac{x_0 - y}{R}$, which yields

$$(Aw)(x_0) = -\partial_\nu w(x_0) = -\nabla w(x_0)\cdot\nu(x_0) = 2\alpha R e^{-\alpha R^2} > 0.$$

Therefore $(Aw)(x) > 0$ for all $x \in S$.

(3) Note that w vanishes on $\partial B_{<R}(y)$, in particular, it vanishes at all points in ∂S sufficiently close to x_0.

Finally, the restriction of v to $\overline{S}$ lies in D, attains a positive value and a strict global maximum at $x_0 \in S$. Thus, the location of the strict maximum in Theorem 7.2.1 implies that $Av \not\ge 0$ on S. But, $Av \ge 0$ on Ω_0 by assumption. Consequently, $\partial_\nu v(x_0) = -(Av)(x_0) > 0$. □

Corollary 7.2.3 (Strong maximum principle for the Laplace operator)**.** *Let $\emptyset \ne \Omega \subseteq \mathbb{R}^n$ be connected and open*[5] *and let $c \in (-\infty, 0]$. Let $v \in \mathrm{C}^2(\Omega;\mathbb{R})$ satisfy $\Delta v(x) + cv(x) \ge 0$ for all $x \in \Omega$. If v attains a positive maximum at an interior point of Ω, then v is constant in Ω.*

Proof. Assume for contradiction that v is non-constant and attains a positive maximum $M \ge 0$ at an interior point of Ω. The set $\Omega^- := \{x \in \Omega : v(x) < M\}$ is non-empty and open, and $\partial\Omega^- \cap \Omega \ne \emptyset$ because Ω is connected. Choose a point $y \in \Omega^-$ such that $\operatorname{dist}(y, \partial\Omega^-) < \operatorname{dist}(y, \partial\Omega)$, and let B be the largest ball centred at y contained entirely in Ω^-. By choice of B, the boundary ∂B touches $\partial\Omega^-$, so $v(x_0) = M$ for some point $x_0 \in \partial B$, while $v(x) < M$ for all $x \in B$. In other words, v has a strict maximum at x_0.

As B trivially satisfies the internal ball condition at x_0, the Hopf boundary lemma (Example 7.2.2) ensures that $\partial_\nu v(x_0) > 0$. But x_0 is an interior point of Ω and v attains a maximum there, so that $\nabla v(x_0) = 0$, a contradiction. □

Example 7.2.4 (Strong positivity for the Dirichlet Laplacian)**.** Let $\emptyset \ne \Omega \subseteq \mathbb{R}^n$ be open, bounded, and connected. We assume that Ω has C^{2k} boundary for some integer $k > \frac{n}{4} + 1$. Consider the quasi-interior point u of $L^2(\Omega)_+$ given by $u(x) := \operatorname{dist}(x, \partial\Omega)$ for all $x \in \Omega$.

The resolvent of the Dirichlet Laplace operator $\Delta_{\mathrm{Dir}}\colon L^2(\Omega) \supseteq \operatorname{dom}(\Delta_{\mathrm{Dir}}) \to L^2(\Omega)$ has the following properties at every $\lambda > \mathrm{s}(\Delta_{\mathrm{Dir}})$:

(a) For every $f \in L^2(\Omega)$, the function $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k f$ is in $\mathrm{C}^2(\overline{\Omega})$ and thus continuous on $\overline{\Omega}$ and satisfies $\left|\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k f\right| \le u$. In particular $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k$ maps into the principal ideal $L^2(\Omega)_u$.

[5] Not necessarily bounded.

(b) In addition, if $0 \lneq f \in L^2(\Omega)$, then $u \preceq \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})f$.

Proof. (a) Let $f \in L^2(\Omega)$. We first observe that $g := \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k f \in C^2(\overline{\Omega})$. Indeed, one has $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k L^2(\Omega) \subseteq H^{2k}(\Omega) \subseteq C^2(\overline{\Omega})$, where the first inclusion follows from elliptic regularity for the Dirichlet Laplacian (Theorem 5.3.2), and the second from the Sobolev embedding theorem 5.3.4 since Ω has C^{2k} boundary and $2k > \frac{n}{2} + 2$.

In particular, g is continuous on $\overline{\Omega}$, so it remains to show that $|g| \preceq u$. To this end, fix a point $x \in \Omega$.

The compactness of $\partial\Omega$ yields the existence of some $x_0 \in \partial\Omega$ such that $\operatorname{dist}(x, \partial\Omega) = \|x - x_0\|_2$. Observe that $g(x_0) = 0$; indeed, g vanishes on all of $\partial\Omega$ since it is a continuous element of $\operatorname{dom}(A)$ and thus of $H_0^1(\Omega)$ (cf. Proposition 5.3.5(b)). Since $g \in C^1(\overline{\Omega})$, the fundamental theorem of calculus yields

$$g(x) = \int_0^1 (\nabla g)\big((1-t)x_0 + tx\big)\,\mathrm{d}t \cdot (x - x_0),$$

which implies the estimate

$$\big|g(x)\big| \le \big\|\nabla g\big\|_{C(\overline{\Omega})} \|x - x_0\|_2 = \big\|\nabla g\big\|_{C(\overline{\Omega})} \operatorname{dist}(x, \partial\Omega).$$

Hence $|g| \preceq u$ as claimed.

(b) Now assume $0 \lneq f \in L^2(\Omega)$. We show the lower bound $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})f \succeq u$ in two steps.

Step 1: We assume in addition that $\lambda \ge 0$ and show that $g := \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k f \succeq u$.

We already know that $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}}) \ge 0$ (either from Example 5.3.6 or by the Beurling-Deny criterion of Theorem 5.1.7), so $g \ge 0$. Next, we show that $g(x) > 0$ for every $x \in \Omega$, so fix such an x and assume that $g(x) = 0$.

One has $(\Delta_{\mathrm{Dir}} - \lambda)(-g) = \mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^{k-1} f \ge 0$. Since $\lambda \ge 0$ and since $g \in C^2(\overline{\Omega})$ according to the proof of (a), the strong maximum principle (Corollary 7.2.3) is applicable to $-g$. As $-g \le 0$ and $-g(x) = 0$, it follows that $-g$ is constant and thus $g = 0$. But this is absurd since $\mathcal{R}(\lambda, \Delta_{\mathrm{Dir}})^k$ is injective and $f \neq 0$.

Now we consider the behaviour of g close to $\partial\Omega$ separately from the behaviour away from the boundary. To this end, let $\delta > 0$ be a number that we determine later and consider the compact set

$$\Omega_\delta := \{x \in \Omega : u(x) \ge \delta\}.$$

One has $g|_{\Omega_\delta} \succeq \mathbb{1}_{\Omega_\delta} \succeq u|_{\Omega_\delta}$ since $g(x) > 0$ for all $x \in \Omega$, so it remains to show that $g|_{\Omega \setminus \Omega_\delta} \succeq u|_{\Omega \setminus \Omega_\delta}$.

To this end we use the following geometric fact: since Ω has C^m boundary for some $m \ge 2$, the interior ball condition is satisfied at every boundary point $x_0 \in \partial\Omega$[6]. By the Hopf boundary lemma (Example 7.2.2), we have $-\partial_\nu g(x_0) > 0$ for all

[6] An intuitive 'proof' can be deduced from the second-order Taylor expansion of a C^m function; see [GT01, pp. 354–355] for a proof using the notion of boundary curvatures.

$x_0 \in \partial\Omega$. Hence, by the compactness of $\partial\Omega$ and the continuity of ∇g on $\overline{\Omega}$, there exists a constant $c > 0$ such that $-\partial_\nu g \geq c$ on $\partial\Omega$.

Now let $x \in \Omega \setminus \Omega_\delta$. Choose $x_0 \in \partial\Omega$ such that $u(x) = \|x - x_0\|_2 < \delta$. Since $g(x_0) = 0$, Taylor's formula with remainder gives the estimate

$$\begin{aligned} g(x) &= \underbrace{\nabla g(x_0)^{\mathrm{T}}(x - x_0)}_{= -\partial_\nu g(x_0)\|x - x_0\|_2} + \int_0^1 (1-t)(x-x_0)^{\mathrm{T}}(Hg)\big((1-t)x_0 + tx\big)(x - x_0)\,\mathrm{d}t \\ &\geq c\,\|x - x_0\|_2 - \frac{1}{2}\,\|Hg\|_{C(\overline{\Omega})}\,\|x - x_0\|_2^2 \geq \Big(c - \frac{\delta}{2}\,\|Hg\|_{C(\overline{\Omega})}\Big)u(x), \end{aligned}$$

where Hg is the Hessian matrix of g. So if δ is chosen sufficiently small, then $g(x) \geq \frac{c}{2}u(x)$ holds for all $x \in \Omega \setminus \Omega_\delta$.

Step 2: Consider a number $\mu \in (\mathrm{s}(\Delta_{\mathrm{Dir}}), \lambda)$. If μ is sufficiently close to λ, the Taylor series expansion of the resolvent (Proposition 3.3.2(a)) gives

$$\mathcal{R}(\mu, \Delta_{\mathrm{Dir}})f = \sum_{j=0}^{\infty} (\lambda - \mu)^j \mathcal{R}(\mu, \Delta_{\mathrm{Dir}})^{j+1} f \geq (\lambda - \mu)^{k-1} g \geq u,$$

where we used that $\mathcal{R}(\mu, \Delta_{\mathrm{Dir}}) \geq 0$. Finally, one can repeat the argument from the proof of Theorem 5.4.1 to see the estimate $\mathcal{R}(\mu, \Delta_{\mathrm{Dir}})f \succeq u$. ☐

We will use Example 7.2.4 at the end of the next section (Example 7.3.8) to obtain further knowledge about the eigenspace associated to the spectral bound of Δ_{Dir}.

7.3 Characterisation of individual eventual strong positivity

After the deep dive into the positivity properties of the Dirichlet Laplacian in the previous section, we continue with developing the theory of eventually positive resolvents – and we shall meet Δ_{Dir} again before the end of the section. From Theorem 6.3.3 we know that, for semisimple eigenvalues, eventual positivity guarantees that the spectral projection is positive. For eventual positivity with respect to a quasi-interior point, it turns out even a characterisation can be given in terms of the spectral projection (Theorem 7.3.6).

Notation 7.3.1. Let X be a Banach space. For $u \in X$ and $\varphi \in X'$, we define the operator

$$u \otimes \varphi := \langle \varphi, \cdot \rangle\, u \in \mathcal{L}(X).$$

Clearly, the operator $u \otimes \varphi$ has rank 1 unless $u = 0$ or $\varphi = 0$. It is not difficult to check that its operator norm is $\|u \otimes \varphi\| = \|u\|\,\|\varphi\|$ and that the operator is a non-zero projection if and only if $\langle \varphi, u \rangle = 1$.

A functional $\varphi \in E'$ on a Banach lattice E is called **strictly positive** if $\langle \varphi, f \rangle > 0$ for all $0 \lneq f \in E$.

Proposition 7.3.2. *Let E be a Banach lattice, let $P \in \mathcal{L}(E)$ be a projection, let $0 \lneq u \in E$ and assume that $Pf \succeq u$ for all $0 \lneq f \in E$. Then there is a strictly positive functional $\varphi \in E'$ such that $P = (Pu) \otimes \varphi$.*

Proof. First observe that the assumption implies $P \gneq 0$ and that $Pu \succeq u$. Moreover, for each $0 \lneq f \in E$ one has $Pf \succeq u$ by assumption, and hence, by applying the projection P again, $Pf \succeq Pu$.

Now consider a vector $0 \le w \in \operatorname{rg} P$. We show that w is a multiple of Pu. Note that $w - tPu \not\geq 0$ for all sufficiently large $t > 0$. Indeed, if we could find numbers $t_n \to \infty$ such that $w - t_n Pu \ge 0$ for all n, then dividing by t_n and letting $n \to \infty$ would yield $0 \ge Pu$, which is absurd as Pu is positive and non-zero. On the other hand, $w - tPu \ge 0$ for $t \le 0$. Hence, there exists a maximal $t_0 \in \mathbb{R}$ for which $w - t_0 Pu \ge 0$. If $w - t_0 Pu \gneq 0$, it follows that $w - t_0 Pu \succeq Pu$, which contradicts the maximality of t_0. So $w - t_0 Pu = 0$.

Since E_+ spans E, the space $\operatorname{rg} P$ is spanned by its positive elements, so it is actually spanned by Pu. As P is a projection, it follows that there is a non-zero functional $\varphi \in E'$ such that $P = (Pu) \otimes \varphi$. Since P and Pu are positive, so is φ. Finally, as $Px \gneq 0$ for every $x \gneq 0$, it follows that φ is strictly positive. □

We now look at a couple of auxiliary results that aid in the proof of Theorem 7.3.6.

Lemma 7.3.3. *Let A be a closed operator on a complex Banach space X and let $\lambda \in \sigma(A)$.*

(a) *If λ is a geometrically simple eigenvalue and there exist $v \in \ker(\lambda - A)$ and $\psi \in \ker(\lambda - A')$ such that $\langle \psi, v \rangle \neq 0$, then λ is algebraically simple.*

(b) *Suppose that $X = E$ is a Banach lattice and $\lambda \in \mathbb{R}$ is a pole of the resolvent $\mathcal{R}(\cdot, A)$ of order $p \in \mathbb{N}$ such that the coefficient Q_{-p+1} of $(\mu - \lambda)^{-p}$ in the Laurent series expansion of $\mathcal{R}(\cdot, A)$ is positive.*

   If $\ker(\lambda - A)$ contains a quasi-interior point of E_+, then $p = 1$.

Proof. (a) The proof of Lemma 1.2.8 carries over mutatis mutandis.

(b) Let $v \in \ker(\lambda - A)$ be a quasi-interior point of E_+. Since Q_{-p+1} is positive (by assumption) and non-zero (by Theorem 6.2.6(a)), Proposition 7.1.3 implies that $Q_{-p+1} v \neq 0$. Employing $v \in \ker(\lambda - A)$, we obtain that

$$\lim_{\mu \to \lambda} (\mu - \lambda)^{p-1} v = \lim_{\mu \to \lambda} (\mu - \lambda)^p \mathcal{R}(\mu, A) v = Q_{-p+1} v \neq 0.$$

As a consequence, $p = 1$. □

Lemma 7.3.4. *Let $\lambda \in \mathbb{R}$ be a spectral value of a closed operator $A \colon X \supseteq \operatorname{dom}(A) \to X$ on a complex Banach space X. If λ is a first order pole of the resolvent $\mathcal{R}(\cdot, A)$, then*

$$\lim_{\mu \to \lambda} \left\| (\mu - \lambda)^m \mathcal{R}(\mu, A)^m - P \right\|_{\mathcal{L}\left(X, \operatorname{dom}(A^m)\right)} = 0 \quad \text{for all } m \in \mathbb{N};$$

where P denotes the spectral projection of A associated to λ.

Proof. As λ is a first order pole, $\lim_{\mu \to \lambda} (\mu - \lambda)^m \mathcal{R}(\mu, A)^m = P$ in $\mathcal{L}(X)$ and $\operatorname{rg} P = \ker(\lambda - A)$ by Theorem 6.2.6(b) and (c). Now using $(\lambda - A)\mathcal{R}(\mu, A) = (\lambda - \mu)\mathcal{R}(\mu, A) + \mathrm{id}$, we obtain

$$(\lambda - A)^m (\mu - \lambda)^m \mathcal{R}(\mu, A)^m = (\mu - \lambda)^m \big((\lambda - \mu)\mathcal{R}(\mu, A) + \mathrm{id}\big)^m \to 0 = (\lambda - A)^m P$$

as $\mu \to \lambda$, from which the assertion follows. □

Lemma 7.3.5. *Let $(x_j)_{j\in J}$ be a net*[7] *of elements in the real part $E_\mathbb{R}$ of a Banach lattice E and let $u \in E$ be a quasi-interior point of E_+. Let $x \in E$ be such that x_j converges to x in the Banach lattice E_u.*

If there exists $c > 0$ such that $x \geq cu$, then for each $\varepsilon \in (0, c)$, there exists $j_0 \in J$ such that $x_j \geq \varepsilon u$ for all $j \geq j_0$.

Proof. For each $\varepsilon \in (0, c)$, we can find $j_0 \in J$ such that for each $j \geq j_0$, we have $\|x_j - x\|_u \leq c - \varepsilon$ and in turn, $|x_j - x| \leq (c-\varepsilon)u$ by the definition of gauge norm. Since each x_j is real, we obtain that $x_j \geq x - (c-\varepsilon)u \geq \varepsilon u$ for all $j \geq j_0$. □

Theorem 7.3.6. *Let $A\colon E \supseteq \operatorname{dom}(A) \to E$ be a closed, densely defined, and real operator on a complex Banach lattice E. Let $\lambda \in \sigma(A) \cap \mathbb{R}$ be a pole of the resolvent $\mathcal{R}(\cdot, A)$ and let $u \in E_+$ be a quasi-interior point. Consider the following assertions:*

(i) *The resolvent $\mathcal{R}(\cdot, A)$ is individually eventually positive with respect to u at λ.*

(ii) *The spectral projection P associated to λ satisfies $Pf \succeq u$ whenever $0 \lneq f \in E$.*

(iii) *The eigenspace $\ker(\lambda - A)$ is spanned by a vector $v \succeq u$ and $\ker(\lambda - A')$ contains a strictly positive functional ψ.*

Each of them implies that λ is algebraically simple and hence a first order pole of $\mathcal{R}(\cdot, A)$.[8] *One has (i) ⇒ (ii) ⇔ (iii), and if $\operatorname{dom}(A) \subseteq E_u$, then all three assertions are equivalent.*

Proof. "(ii) ⇒ (iii)": By Proposition 7.3.2, $P = (Pu) \otimes \varphi$ for a strictly positive functional $\varphi \in E'$. In particular, $\operatorname{rg} P$ is one-dimensional.

According to Theorem 6.2.6(b) and (c), $\operatorname{rg} P$ coincides with the generalised eigenspace of λ, so it follows that λ is algebraically, and hence geometrically simple, and hence a first order pole. In particular, there exists $0 \lneq v \in \ker(\lambda - A) = \operatorname{rg} P$ with $v = Pv \succeq u$.

Moreover, as P is the spectral projection of A corresponding to λ, P' is the spectral projection of A' corresponding to λ (Theorem 6.2.6(d)). Thus, $\ker(\lambda - A')$ contains the strictly positive functional φ.

"(iii) ⇒ (ii)": Firstly, Lemma 7.3.3(a) ensures that λ is even algebraically simple. Therefore, we obtain from Theorem 6.2.6(b) and (c) that $\operatorname{rg} P = \ker(\lambda - A)$. From the same theorem, we also have $\operatorname{rg} P' = \ker(\lambda - A')$.

Now, if $0 \lneq f \in E$, then there exists $\alpha \in \mathbb{C}$ such that $Pf = \alpha v$. Actually, as A is real, and hence so is P, we get $\alpha \in \mathbb{R}$. We claim that $\alpha > 0$. If not, then

$$0 \lneq f \leq f - \alpha v = f - Pf \in \ker P.$$

Since $\psi \in \ker(\lambda - A') = \operatorname{rg} P'$, we have $\psi = P'\varphi$ for some $\varphi \in E'$. This implies that $\langle \psi, f - \alpha v\rangle = \langle \varphi, P(f - \alpha v)\rangle = 0$, contradicting the strict positivity of ψ.

[7] Recall the definition of a net from Exercise 4.5.
[8] By Theorem 6.2.6(b).

"(i) $\Rightarrow$ (ii)": From Theorem 6.3.3, there exist $0 \lneq v \in \ker(\lambda - A)$ and $0 \lneq \psi \in \ker(\lambda - A')$. By assumption, there exists $\mu > \lambda$ such that $v = (\mu - \lambda)\mathcal{R}(\mu, A)v \succeq u$. In particular, v is also a quasi-interior point of E_+. Furthermore, ψ is strictly positive. Indeed, if $0 \lneq f \in E$, choose $\mu > \lambda$ such that $\mathcal{R}(\mu, A)f \succeq u$. Then

$$\langle \psi, f \rangle = \langle (\mu - \lambda)\mathcal{R}(\mu, A')\psi, f \rangle = (\mu - \lambda)\langle \psi, \mathcal{R}(\mu, A)f \rangle \succeq \langle \psi, u \rangle > 0$$

because u is a quasi-interior point (Proposition 7.1.4).

Next, let $p \in \mathbb{N}$ be the pole order of λ. Then $(\mu - \lambda)^p \mathcal{R}(\mu, A)$ converges to Q_{-p+1} as $\mu \downarrow \lambda$ and the resolvent is individually eventually positive at λ, hence $Q_{-p+1} \geq 0$. Lemma 7.3.3(b) thus ensures that $p = 1$.

Owing to Theorem 6.2.6, $Q_0 = Q_{-p+1} \gneq 0$ is the spectral projection of A associated to λ, $\operatorname{rg} Q_0 = \ker(\lambda - A)$, and $\operatorname{rg} Q_0' = \ker(\lambda - A')$. The last equality ensures that $Q_0'\psi = \psi$. As ψ is strictly positive, this implies $\ker Q_0$ does not contain any positive non-zero elements. So if $0 \lneq f \in E$, then $Q_0 f \gneq 0$ and in turn, there exists $\mu > \lambda$ such that

$$Q_0 f = (\mu - \lambda)\mathcal{R}(\mu, A)Q_0 f \succeq u.$$

Lastly, assume that $\operatorname{dom}(A) \subseteq E_u$.

"(ii) $\Rightarrow$ (i)": As already observed, λ is a first order pole of the resolvent $\mathcal{R}(\cdot, A)$ and hence by Lemma 7.3.4, $(\mu - \lambda)\mathcal{R}(\mu, A) \to P$ in $\mathcal{L}(E, \operatorname{dom}(A))$ as $\mu \downarrow \lambda$. Since $\operatorname{dom}(A) \subseteq E_u$, this convergence even holds in $\mathcal{L}(E, E_u)$ thanks to the closed graph theorem.

Thus for $0 \lneq f \in E$, the net $(\mu - \lambda)\mathcal{R}(\mu, A)f$ converges to $Pf \succeq u$ in E_u. The assertion thus follows by an application of Lemma 7.3.5. ☐

As a natural follow-up to Theorem 7.3.6, one may ask whether the eventual positivity of the resolvent can be obtained from the spectral assertions without assuming $\operatorname{dom}(A) \subseteq E_u$. Changing the state space to $L^p(-1, 1)$ in Example 6.1.2 for $1 \leq p < \infty$ (un)fortunately, refutes this; readers interested in details of the computation can find it in [DGK16a, Example 5.4]. Let us observe next that eventual negativity also fits into the framework of Theorem 7.3.6.

Corollary 7.3.7. *In the situation of Theorem 7.3.6, assume that* $\operatorname{dom}(A) \subseteq E_u$. *Then the assertions (i)–(iii) are also equivalent to the following property.*

(iv) *The resolvent $\mathcal{R}(\cdot, A)$ is individually eventually negative with respect to u at λ.*

Proof. Without loss of generality let $\lambda = 0$. Replacing A with $-A$ in the theorem, assertion (iii) remains unchanged, but (i) becomes individual eventual positivity of $\mathcal{R}(\cdot, -A)$ with respect to u at 0, which is equivalent to (iv). ☐

An application of Theorem 7.3.6 and Corollary 7.3.7 to a fourth-order differential operator is discussed in Exercise 7.2. In our final example in this chapter we revisit the most prominent (and most classical) example so far, the Dirichlet Laplacian.

On bounded domains in $\mathbb{R}^n$ we already know that the Dirichlet Laplacian on L^2 has a positive eigenvector for the eigenvalue $s(\Delta_{\text{Dir}})$. Using the abstract results established in the present section and the concrete results from Section 7.2 we can now show much more. In one dimension we already knew assertions (a) and (b) in the following example from a concrete computation (Example 6.3.2); such an explicit computation is, of course, not possible on general domains in dimension ≥ 2.

Example 7.3.8 (The leading eigenfunction of the Dirichlet Laplacian)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{R}^n$ be open, bounded, and connected. We assume that Ω has C^{2k} boundary for some integer $k > \frac{n}{4} + 1$. Consider the quasi-interior point u of $L^2(\Omega)_+$ given by $u(x) := \operatorname{dist}(x, \partial\Omega)$ for all $x \in \Omega$. The spectral bound $s(\Delta_{\text{Dir}})$ of the Dirichlet Laplace operator $\Delta_{\text{Dir}} \colon L^2(\Omega) \supseteq \operatorname{dom}(\Delta_{\text{Dir}}) \to L^2(\Omega)$ has the following properties:

(a) $\ker(s(\Delta_{\text{Dir}}) - \Delta_{\text{Dir}})$ is spanned by a positive function v that satisfies $u \preceq v \preceq u$.

(b) One has $s(\Delta_{\text{Dir}}) < 0$.[9]

(c) If $n = 1$, then $\mathcal{R}(\cdot, \Delta_{\text{Dir}})$ is individually eventually negative with respect to u at $s(\Delta_{\text{Dir}})$.

Proof. (a) Let us abbreviate $\lambda_0 := s(\Delta_{\text{Dir}})$. Since Δ_{Dir} has compact resolvent and $-\infty < \lambda_0 \leq 0$ (Example 6.3.5), λ_0 is a pole of the resolvent and an eigenvalue (Theorem 6.2.9).

As shown in Example 7.2.4 we have $\mathcal{R}(\lambda, \Delta_{\text{Dir}}) f \succeq u$ for all $\lambda > \lambda_0$ and all $0 \lneq f \in L^2(\Omega)$. Hence, Theorem 7.3.6 implies that $\ker(\lambda_0 - \Delta_{\text{Dir}})$ is spanned by a function $v \succeq u$. On the other hand, for $\lambda > \lambda_0$ one has $(\lambda - \lambda_0)^{-k} v = \mathcal{R}(\lambda, \Delta_{\text{Dir}})^k v$, and the latter vector is in the principal ideal E_u according to Example 7.2.4. Thus, $v \preceq u$.

(b) We use the notation from the proof of (a). Assume for a contradiction that $\lambda_0 = 0$. Then $\Delta_{\text{Dir}} v = 0$. One has $v \in \operatorname{rg} \mathcal{R}(\lambda, A)^k \subseteq C^2(\overline{\Omega})$, where the inclusion was shown in Example 7.2.4. Hence, we can apply the maximum principle (Example 5.2.2) to conclude that v obtains its maximum at $\partial\Omega$. But v vanishes on $\partial\Omega$ and v is positive, so $v = 0$, a contradiction.

(c) We apply Theorem 7.3.6. As pointed out in the proof of (a), condition (i) of the theorem is satisfied. Moreover, since $n = 1$ one has

$$\operatorname{dom}(\Delta_{\text{Dir}}) = H^2(\Omega) \cap H_0^1(\Omega) \subseteq C^1(\overline{\Omega}) \cap C_0(\Omega) \subseteq L^2(\Omega)_u,$$

so Corollary 7.3.7 can be applied and gives the claimed eventual negativity. □

We now know that $\mathcal{R}(\cdot, \Delta_{\text{Dir}})$ is individually eventually negative at $s(\Delta_{\text{Dir}})$ if $n = 1$ and that it is not uniformly eventually negative there when $n \geq 4$ (Example 6.4.6). This obviously leaves a gap, which we will close later on.

[9] As pointed out before, this also follows from the **Poincaré inequality** under more general assumptions on Ω, but here we give a proof based on the maximum principle.

Exercises for Chapter 7

Exercise 7.1 (Quasi-interior points)**.**

(a) Let E be a Banach lattice. Show that $u \in E_+$ is a quasi-interior point if and only if $(nu) \wedge f \to f$ as $n \to \infty$ for every $f \in E_+$.

(b) Let $p \in [1, \infty)$ and let (Ω, μ) be a σ-finite measure space.

Show that $u \in L^p(\Omega, \mu)_+$ is a quasi-interior point if and only if $u(\omega) > 0$ for almost all $\omega \in \Omega$. Also show that $v \in L^\infty(\Omega, \mu)_+$ is a quasi-interior point if and only if $v \geq \mathbb{1}$.

(c) Let K be a compact metric space. Show that $u \in C(K)_+$ is a quasi-interior point if and only if $u(x) > 0$ for all $x \in K$, if and only if $u \geq \mathbb{1}$.

(d) Let $\emptyset \neq \Omega \subseteq \mathbb{R}^d$ be an open set. Prove that a function $u \in C_0(\Omega)_+$ is a quasi-interior point if and only if $u(x) > 0$ for all $x \in \Omega$.

(e) Give an example of a Banach lattice E with no quasi-interior points.

(f) Let E be a Banach lattice and let $\varphi \in E'_+$. Which of the following conditions implies the other?

(i) φ is strictly positive.

(ii) φ is a quasi-interior point of E'_+.

Exercise 7.2 (A fourth order operator on an interval)**.**

(a) Consider the space $L^2(0,1)$ and endow its vector subspace

$$V := \left\{ v \in H^2(0,1) : v'(0) = v'(1) = 0 \right\}$$

with the H^2-norm. Define the sesquilinear form

$$\mathfrak{a} \colon V \times V \to \mathbb{C}, \qquad \mathfrak{a}(u,v) := \int_0^1 \overline{u''}\, v'' \, \mathrm{d}x.$$

Show that $\mathfrak{a}$ satisfies all assumptions of Theorem 5.1.4, where μ can be chosen to be any number > 0.

(b) Let $A \colon L^2(0,1) \supseteq \operatorname{dom}(A) \to L^2(0,1)$ denote the operator associated to $\mathfrak{a}$. Show that there exists $\lambda_0 > 0$ such that $\mathcal{R}(\lambda, A) \not\geq 0$ for all $\lambda \in [\lambda_0, \infty)$.

(c) Compute $\operatorname{dom}(A)$ and Au for all $u \in \operatorname{dom}(A)$.

(d) Show that $\ker A$ and $\ker(A')$ are spanned by $\mathbb{1}$, that $\operatorname{s}(A) = 0$, and that A has compact resolvent.

(e) Prove that $\mathcal{R}(\cdot, A)$ is individually eventually positive with respect to $\mathbb{1}$ at 0.

Exercise 7.3 (The positive minimum principle and higher order operators)**.** Let K be a compact metric space. Let $A\colon \mathrm{C}(K) \supseteq \operatorname{dom}(A) \to \mathrm{C}(K)$ be closed and densely defined, and assume that $\operatorname{s}(A) < \infty$ and $\mathcal{R}(\lambda, A) \geq 0$ for all $\lambda > \operatorname{s}(A)$.

(a) Show that A is real.

(b) Fix a point $\mu \in (\operatorname{s}(A), \infty)$ and set $u := \mathcal{R}(\mu, A)\,\mathbb{1}$. Prove that that u is a quasi-interior point of $\mathrm{C}(K)_+$, i.e. that $u(x) > 0$ for each $x \in K$.

Show furthermore that $\mathcal{R}(\lambda, A)u \leq \frac{1}{\lambda - \mu} u$ for all $\lambda > \mu$.

(c) Show for every $v \in \operatorname{dom}(A)$ that $\lambda\big(\lambda\mathcal{R}(\lambda, A) - \mathrm{id}\big)v \to Av$ in $\mathrm{C}(K)$ as $\lambda \to \infty$.

Hint: Use Exercise 5.1(a).

(d) Prove that, for each $x_0 \in K$, the operator A satisfies the **positive minimum principle at** $x_0 \in K$, i.e. for each $0 \leq v \in \operatorname{dom}(A)$ with $v(x_0) = 0$ one has $(Av)(x_0) \geq 0$.[10]

(e) Let $m \geq 3$ be an integer. Consider a densely defined closed operator $B\colon \mathrm{C}([-1,1]) \supseteq \operatorname{dom}(B) \to \mathrm{C}([-1,1])$. Assume that $\operatorname{dom}(B)$ contains all test functions on $(-1,1)$ and that $Bv = v^{(m)}$ for each such test function v.

Show that B does not satisfy the positive minimum principle for any $x_0 \in (-1,1)$.

[10]Note that this is the same property that was assumed in Theorems 5.2.1 and 7.2.1.

Notes for Chapter 7

Quasi-interior points

Quasi-interior points were introduced by Schaefer [Sch60] as a generalisation of points in the topological interior of the cone. The main motivation was that cones in infinite-dimensional spaces often have empty interior, but quasi-interior points still have many useful properties of interior points. Quasi-interior points are, for instance, useful in the study of so-called **irreducible** operators, see e.g. [MN91, Section 4.2]. In the literature, the symbol $u \gg 0$ is sometimes used to denote that u is a quasi-interior point.

We point out that the characterisation of quasi-interior points via positive linear functionals (Proposition 7.1.4) fails in the general setting of so-called **ordered Banach spaces**. See [GW20, Section 2.2] for a detailed discussion of this topic.

The Hopf boundary point lemma

The arguments that we gave to prove the Hopf boundary lemma (Example 7.2.2) are, in principle, almost the same as one may find in standard PDE books. What is unusual about our approach is that we phrased it in the abstract setting of Theorem 7.2.1, which extends the setting of the abstract maximum principle from Theorem 5.2.1 by an additional assumption on w.

To encode the inner normal derivative $-\partial_\nu$ at x_0 into the action of the operator A in the proof of Hopf's boundary lemma does not seem to be a common approach. This somewhat unconventional structure of A, without a clear theoretical explanation for its occurrence, is one indication – among others – that the abstract versions of the maximum principle in Theorems 5.2.1 and 7.2.1 are not yet in a really satisfactory state.

Readers with an inclination towards PDE theory may also object to the strong assumptions on the regularity of the boundary of the domain Ω in Example 7.2.4. This is due to our Sobolev-space approach, which starts with very little regularity (merely L^2 functions) and heavily depends on the Sobolev embedding theorems. An equally well-established approach is the so-called **Schauder theory**, which works with spaces of Hölder continuous functions and classical derivatives. In short, if Ω is a bounded domain with $C^{2,\alpha}$ boundary (for some $\alpha \in (0,1)$), $f \in C^{0,\alpha}(\overline{\Omega})$, and if u solves $\lambda u - \Delta u = f$ in Ω and $u = 0$ on $\partial\Omega$ in the classical sense, then $u \in C^{2,\alpha}(\overline{\Omega})$; see [GT01, Theorem 6.19]. Thus, the order of differentiability of the PDE solution and the boundary agree.

Individually eventually positive resolvents

The characterisation of individual eventual positivity with respect to u in Theorem 7.3.6 stems from [DGK16a, Sections 3 and 4]. The fact that the implication from (i) to the other assertions holds even without the domination assumption $\operatorname{dom}(A) \subseteq E_u$ was proved in [DG17, Section 4]. The proof that we presented for this implication is a bit different in that it avoids using properties of quasi-interior points that are more strongly tied to the lattice structure of the surrounding space (in particular, the properties discussed in Supplement 7.A). This might turn out beneficial in potential generalisations of eventual positivity theory to ordered Banach spaces.

The fact that the domination assumption $\operatorname{dom}(A) \subseteq E_u$ is in fact necessary in many cases in order to have individual eventual positivity and negativity at the same time, was shown in [AG23]. In these notes, we simplified the proof and improved the result by removing the assumption that the vector u satisfies $u \preceq v \preceq u$ for an eigenvector of A.

Encore: if you want to know more...

7.A More on quasi-interior points

In the Banach lattice $\mathbb{R}$, every non-zero positive element is a quasi-interior point. Here, we show that this cannot happen in any dimension larger than 1 (Proposition 7.A.4). The underlying idea is related to Proposition 7.3.2, but is technically a bit more involved.

Definition 7.A.1 (Disjointness of vectors)**.** Two vectors x, y of a Banach lattice E are called **disjoint** if $|x| \wedge |y| = 0$.

Example 7.A.2. If x is any vector in a Banach lattice E, then x^+ and x^- are disjoint elements, and thus $x = x^+ - x^-$ is a *disjoint decomposition*. Indeed, one has

$$x^+ \wedge x^- = [(x^+ - x^-) \wedge (x^- - x^-)] + x^- = (x \wedge 0) + x^- = -x^- + x^- = 0$$

by elementary properties of vector lattice operations (Proposition 4.1.3).

Lemma 7.A.3. *If E is a Banach lattice with dimension $\dim E \geq 2$, then there exist two disjoint non-zero elements $x, y \in E_+$.*

Proof. As $\dim E \geq 2$ and as the linear span of E_+ equals E, we can find two linearly independent vectors $u, v \in E_+$; in particular, $u, v \neq 0$. Observe that there exists a real number $t_0 \in (0, \infty)$ such that neither $u \leq t_0 v$ nor $u \geq t_0 v$ holds. Indeed, both the sets

$$\{t \in (0, \infty) : u \leq tv\} \quad \text{and} \quad \{t \in (0, \infty) : u \geq tv\}$$

are not equal to $(0, \infty)$ since $u, v \neq 0$, are relatively closed in $(0, \infty)$ since E_+ is closed, and are disjoint since u, v are linearly independent. Hence, the union of those two sets cannot be $(0, \infty)$ since $(0, \infty)$ is connected.

Let us now replace v with $t_0 v$. Then u, v are non-zero vectors in E_+ that satisfy $u \not\leq v$ and $u \not\geq v$. Hence, the positive vectors $x := u - u \wedge v$ and $y := v - u \wedge v$ are non-zero and they are disjoint since

$$0 \leq x \wedge y = (u - u \wedge v) \wedge (v - u \wedge v) = u \wedge v - u \wedge v = 0. \qquad \square$$

Proposition 7.A.4. *If every positive non-zero element of a Banach lattice E is a quasi-interior point, then $\dim E \leq 1$.*

Proof. Suppose $\dim E \geq 2$. By Lemma 7.A.3, there exist disjoint non-zero vectors $x, y \in E_+$. Every element of E_y is then disjoint to x: indeed, for each $z \in E_y$ there exists a number $c \geq 1$ such that $|z| \leq cy$. Hence, $x \wedge |z| = 0$ follows from

$$0 \leq x \wedge |z| \leq x \wedge (cy) \leq (cx) \wedge (cy) = c(x \wedge y) = 0,$$

From continuity of the lattice operations (Proposition 4.1.7), it follows that each element of the closure $\overline{E_y}$ is disjoint to x. Since x is non-zero, it is not disjoint to itself and thus, $x \notin \overline{E_y}$. Therefore, y is not a quasi-interior point of E_+. □

Chapter 8

Criteria for eventual positivity of resolvents: the uniform case

After proving sufficient conditions for individual eventual positivity (and negativity) of resolvents in the previous chapter, we now turn to the uniform case. A key assumption to get individual eventual positivity in Theorem 7.3.6 was the property $\operatorname{dom}(A) \subseteq E_u$, since it gives that the resolvent maps into the principal ideal E_u. For the uniform case, we need a stronger property of the resolvent. In Sections 8.1 and 8.2 we set the stage for this, before we turn to the main theorem in Section 8.3.

8.1 Banach lattice overture: Norms induced by functionals

To obtain criteria for uniformly eventually positive resolvents, we need a few more tools from Banach lattice theory. This section introduces a construction dual to the principal ideals discussed in Section 7.1. The following example serves as motivation.

Example 8.1.1. Let (Ω, μ) be a finite measure space and let $p, p' \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{p'} = 1$. As usual we identify $L^{p'}(\Omega, \mu)$ with the dual space $\big(L^p(\Omega, \mu)\big)'$ – note that this identification respects the order structure.

The function $\mathbb{1} \in L^{p'}(\Omega, \mu)$ acts as a strictly positive functional on the Banach lattice $L^p(\Omega, \mu)$, and for each $f \in L^p(\Omega, \mu)$ one has

$$\langle \mathbb{1}, |f| \rangle = \int_\Omega |f| \, \mathrm{d}\mu = \|f\|_{L^1}.$$

Hence $\langle \mathbb{1}, |\cdot| \rangle$ is a norm on $L^p(\Omega, \mu)$. Of course, the norm completion of $L^p(\Omega, \mu)$ with respect to this norm is the Banach lattice $L^1(\Omega, \mu)$.

Now we take this example and turn it into a general construction on Banach lattices. If E, F are Banach lattices over the same field, a linear map $J \colon E \to F$ is called a **lattice homomorphism** if $|Jx| = J|x|$ for all $x \in E$. Observe that every lattice homomorphism is positive (and hence continuous by Theorem 4.3.3). A bijective lattice homomorphism is called a **lattice isomorphism**. This is justified, since one easily checks that the inverse is also a lattice homomorphism.

Proposition 8.1.2 (AL-spaces generated by functionals)**.** *Let E be a Banach lattice, let $\varphi \in E'_+$ be a strictly positive functional, and consider the norm $\|\cdot\|_{E^\varphi} := \langle \varphi, |\cdot| \rangle$ on E. There exists, up to an isometric lattice isomorphism, precisely one Banach lattice E^φ over the same field as E with the following properties:*

(a) *As a Banach space, E^φ is the norm completion of the normed space $(E, \|\cdot\|_{E^\varphi})$.*[1]

(b) *The inclusion map $E \hookrightarrow E^\varphi$ is a lattice homomorphism.*

*The Banach lattice E^φ is called the **AL-space generated by** φ.*

The terminology for E^φ is due to the fact that the norm on this space can be readily seen to be additive on the positive cone, and Banach lattices with this property are often called **AL-spaces**. In the real case, the proposition can be checked by showing that E^φ is a Banach lattice with the order induced by $\overline{E_+}^{\|\cdot\|_{E^\varphi}}$ and the embedding $E \hookrightarrow E^\varphi$ is a lattice homomorphism; see, for instance, the beginning of Section IV.3 in [Sch74]. The complex case can then be derived from the real one.

The reason why we call the construction of E^φ dual to the construction of principal ideals is explained in Exercise 8.1.

8.2 Smoothing properties of operators

Let (Ω_1, μ_1) and (Ω_2, μ_2) be finite measure spaces and let $k \in L^\infty(\Omega_1 \times \Omega_2, \mu_1 \otimes \mu_2)$. Then

$$\begin{aligned} T_k \colon L^1(\Omega_2, \mu_2) &\to L^\infty(\Omega_1, \mu_1) \\ f &\mapsto \int_\Omega k(\cdot, y) f(y) \, \mathrm{d}\mu_2(y) \end{aligned}$$

defines a bounded linear operator. In fact, the **Dunford-Pettis theorem** says that every $T \in \mathcal{L}(L^1(\Omega_2, \mu_2), L^\infty(\Omega_1, \mu_1))$ is of this form; see for instance [Are06, Theorem 4.1.1] from the lecture notes of the 9th Internet Seminar. If k is real-valued, then T_k is real and it follows from $|k| \le \|k\|_\infty \mathbb{1}_{\Omega_1 \times \Omega_2}$ that

$$\pm T_k \le \|k\|_\infty \mathbb{1}_{\Omega_1} \otimes \mathbb{1}_{\Omega_2};$$

where we use the notation for rank-1 operators introduced in Notation 7.3.1. Now let $p_1, p_2 \in [1, \infty]$. If a real operator $T \colon L^{p_2}(\Omega_2, \mu_2) \to L^{p_1}(\Omega_1, \mu_1)$ extends to an operator $L^1(\Omega_2, \mu_2) \to L^\infty(\Omega_1, \mu_1)$, then the extended operator has the form described above. Thus, the extension and hence T itself are dominated above and below by multiples of $\mathbb{1}_{\Omega_1} \otimes \mathbb{1}_{\Omega_2}$.

We now generalise these observations to abstract Banach lattices. To recover the situation above in the following theorem, observe that the principal ideal generated by $\mathbb{1}_{\Omega_1}$ in $L^{p_1}(\Omega_1, \mu_1)$ is $L^\infty(\Omega_1, \mu_1)$, and that AL-space generated by the functional $\mathbb{1}_{\Omega_2} \in \left(L^{p_2}(\Omega_2, \mu_2)\right)'$ equals $L^1(\Omega_2, \mu_2)$.

[1] Recall that the norm completion of a normed space is unique up to isometric Banach space isomorphism.

Theorem 8.2.1. *Let E, F be Banach lattices, $u \in F_+$, and $\varphi \in E'$ a strictly positive functional. For every real operator $T \in \mathcal{L}(E, F)$, the following are equivalent:*

(i) *T extends to a bounded linear operator $\widetilde{T} \in \mathcal{L}(E^\varphi, F_u)$.*

(ii) *There exists a constant $c \geq 0$ such that $\pm T \leq c\, u \otimes \varphi$; in short, $\pm T \preceq u \otimes \varphi$.*

(iii) *There exists a constant $c' \geq 0$ such that $|Tx| \leq c' \langle \varphi, |x| \rangle\, u$ for all $x \in E$.*

If any of the above assertions hold, then $c = c' = \|\widetilde{T}\|_{F_u \leftarrow E^\varphi}$.

Proof. Note that the inclusions $j : E \hookrightarrow E^\varphi$ and $k : F_u \hookrightarrow F$ are lattice homomorphisms.

"(i) ⇔ (iii)": If (i) holds, then Diagram (8.2.1) commutes and for every $x \in E$,[2]

$$\left\|\widetilde{T} j x\right\|_{F_u} \leq c' \left\|jx\right\|_{E^\varphi} = c' \langle \varphi, \left|jx\right| \rangle = c' \langle \varphi, \left|f\right| \rangle;$$

where $c' = \|\widetilde{T}\|_{F_u \leftarrow E^\varphi}$. In turn, $\left|\widetilde{T} j x\right| \leq c' \langle \varphi, |x| \rangle\, u$. Consequently

$$|Tx| = \left|k \widetilde{T} j x\right| = \left|\widetilde{T} j x\right| \leq c' \langle \varphi, |x| \rangle\, u$$

which is (iii).

$$\begin{array}{ccc} E^\varphi & \xrightarrow{\ \widetilde{T}\ } & F_u \\ {\scriptstyle j}\uparrow & & \downarrow{\scriptstyle k} \\ E & \xrightarrow[\ T\ ]{} & F \end{array} \tag{8.2.1}$$

Conversely, if (iii) holds, then from the definitions of the norms on F_u and E^φ, $\left\|\widetilde{T} x\right\|_{F_u} \leq c' \|x\|_{E^\varphi}$ for all $x \in E$. By density of E in $(E^\varphi, \|\cdot\|_{E^\varphi})$, assertion (i) follows.

"(ii) ⇔ (iii)": The inequality in (iii) implies $-c \langle \varphi, x \rangle\, u \leq -|Tx| \leq Tx \leq |Tx| \leq c \langle \varphi, x \rangle\, u$ for all $x \in E_+$, which immediately yields (ii).

Conversely, assume that (ii) holds. Then for every $x \in E_\mathbb{R}$, we have

$$\left|Tx^+\right| \leq c \langle \varphi, x^+ \rangle\, u \quad \text{and} \quad \left|Tx^-\right| \leq c \langle \varphi, x^- \rangle\, u.$$

Thus $|Tx| \leq \left|Tx^+\right| + \left|Tx^-\right| \leq c \langle \varphi, |x| \rangle\, u$. For each $z \in E$ and $\theta \in [0, 2\pi]$, this gives

$$\left|\operatorname{Re}(e^{i\theta} Tz)\right| = \left|T \operatorname{Re}(e^{i\theta} z)\right| \leq c \left\langle \varphi, \left|\operatorname{Re}(e^{i\theta} z)\right| \right\rangle u;$$

the first equality uses that T is a real operator. Recalling our construction of the complex modulus function (Theorem 4.2.4), we deduce

$$\begin{aligned} |Tz| = \frac{1}{4} \int_0^{2\pi} \left|\operatorname{Re}(e^{i\theta} Tz)\right| \,\mathrm{d}\theta &\leq \frac{1}{4} \int_0^{2\pi} c \left\langle \varphi, \left|\operatorname{Re}(e^{i\theta} z)\right| \right\rangle u \,\mathrm{d}\theta \\ &= c \left\langle \varphi, \frac{1}{4} \int_0^{2\pi} \left|\operatorname{Re}(e^{i\theta} z)\right| \,\mathrm{d}\theta \right\rangle u = c \langle \varphi, |z| \rangle\, u, \end{aligned}$$

and thus the proof is complete. □

[2] In order to keep the notation reasonable, we do not show explicitly in which spaces the various moduli are taken.

Condition (i) in Theorem 8.2.1 easily gives the following consequence.

Corollary 8.2.2. *Let E be a Banach lattice, let $u \in E_+$, and let $\varphi \in E'$ be a strictly positive functional. If two real operators $T_1, T_2 \in \mathcal{L}(E)$ satisfy the equivalent assertions of Theorem 8.2.1 with $E = F$, then so do $\alpha_1 T_1 + \alpha_2 T_2$ and $T_1 S T_2$ for all $\alpha_1, \alpha_2 \in \mathbb{R}$ and for all real operators $S \in \mathcal{L}(E)$.*

The equivalent assertions in Theorem 8.2.1 are actually closely related to assumptions of the type $\operatorname{dom}(A) \subseteq E_u$ that occurred in Chapter 7.1. The following result makes the connection more explicit.

Proposition 8.2.3. *Let E, F be Banach lattices, let $\varphi \in E'$ be a strictly positive functional, and let $T \in \mathcal{L}(E,F)$. The following assertions are equivalent:*

(i) *T extends to a bounded linear operator $\widetilde{T}\colon E^\varphi \to F$.*

(ii) *The range of $T'\colon F' \to E'$ is contained in the principal ideal $(E')_\varphi$.*

In this case,

$$\|\widetilde{T}\|_{F\leftarrow E^\varphi} = \|T'\|_{(E')_\varphi \leftarrow F'}.$$

Proof. "(i) $\Rightarrow$ (ii)": For all $\psi \in F'$ with $\|\psi\|_{F'} \le 1$ and all $y \in E$, we have

$$|\langle \psi, Ty\rangle| \le \|Ty\|_F \le \|\widetilde{T}\|_{F\leftarrow E^\varphi} \langle \varphi, |y|\rangle.$$

Now let $x \in E_+$ be arbitrary. The Riesz-Kantorovich formula (Theorem 4.4.2) implies

$$\begin{aligned}\langle |T'\psi|, x\rangle = \sup_{|y|\le x} |\langle T'\psi, y\rangle| &= \sup_{|y|\le x} |\langle \psi, Ty\rangle| \\ &\le \|\widetilde{T}\|_{F\leftarrow E^\varphi} \sup_{|y|\le x} \langle \varphi, |y|\rangle = \|\widetilde{T}\|_{F\leftarrow E^\varphi} \langle \varphi, x\rangle.\end{aligned}$$

This proves that $|T'\psi| \le \|\widetilde{T}\|_{F\leftarrow E^\varphi} \|\psi\|_{F'} \varphi$ for all $\psi \in F'$, and therefore $\operatorname{rg} T' \subseteq (E')_\varphi$ with $\|T'\|_{(E')_\varphi \leftarrow F'} \le \|\widetilde{T}\|_{F\leftarrow E^\varphi}$.

"(ii) $\Rightarrow$ (i)": The closed graph theorem implies that $T'\colon F' \to (E')_\varphi$ is bounded. It follows by the definition of gauge norm that $|T'\psi| \le \|T'\psi\|_{(E')_\varphi} \varphi$ for all $\psi \in F'$. Hence

$$|\langle \psi, Tx\rangle| = |\langle T'\psi, x\rangle| \le \langle |T'\psi|, |x|\rangle \le \|T'\psi\|_{(E')_\varphi} \langle \varphi, |x|\rangle = \|T'\psi\|_{(E')_\varphi} \|x\|_{E^\varphi}$$

for all $x \in E$ and $\psi' \in F'$, where the first inequality is a direct consequence of the Riesz-Kantorovich formula. After taking the supremum over all $\psi \in F'$ with $\|\psi\|_{F'} \le 1$, we find

$$\|Tx\|_F \le \|T'\|_{(E')_\varphi \leftarrow F'} \|x\|_{E^\varphi}$$

for all $x \in E$. Since by definition E is dense in E^φ with respect to $\|\cdot\|_{E^\varphi}$, it follows that T extends to a bounded linear operator $\widetilde{T}\colon E^\varphi \to F$ with $\|\widetilde{T}\|_{F\leftarrow E^\varphi} \le \|T'\|_{(E')_\varphi \leftarrow F'}$. □

Corollary 8.2.4. *Let E be a Banach lattice, let $u \in E_+$ and let $\varphi \in E'_+$ be a strictly positive functional. If $T_1, T_2, S \in \mathcal{L}(E)$ are real operators and satisfy*

$$\operatorname{rg} T_1 \subseteq E_u \qquad \text{and} \qquad \operatorname{rg} T_2' \subseteq (E')_\varphi,$$

then $T_1 S T_2$ satisfies the equivalent assertions of Theorem 8.2.1.

Proof. The closed graph theorem implies that $T_1 \in \mathcal{L}(E, E_u)$, while Proposition 8.2.3 shows that T_2 extends to an operator $\widetilde{T}_2 \in \mathcal{L}(E^\varphi, E)$. Hence $T_1 S T_2 : E \to E$ extends to the bounded linear operator $T_1 S \widetilde{T}_2 : E^\varphi \to E_u$, i.e. assertion (i) of Theorem 8.2.1 is fulfilled. □

Example 8.2.5. Consider the Banach lattice $E = L^2(0,1)$, and identify E' with E. Define a continuous function $G: [0,1]^2 \to [0,\infty)$ by $G(x,y) := x \wedge y - xy$. Let $T \in \mathcal{L}(E)$ be given by

$$Tf = \int_0^1 G(\cdot, y) f(y) \,\mathrm{d}y$$

for all $f \in E$. Consider the quasi-interior point $u \in L^2(0,1)_+$ given by $u(x) = x(1-x)$. Then T has the following properties:

(a) $\operatorname{rg} T \subseteq E_u$ and $\operatorname{rg} T' \subseteq E_u$. Thus, $T^2 \preceq u \otimes u$ according to Corollary 8.2.4.

(b) However, $T \not\preceq u \otimes u$.

Proof. (a) Since the kernel is symmetric (i.e. $G(x,y) = G(y,x)$), it suffices to show $\operatorname{rg} T \subseteq E_u$. For $x, y \in (0,1)$,

$$0 \le u(x)^{-1} G(x,y) = \begin{cases} x^{-1} y, & \text{if } y \le x, \\ (1-x)^{-1}(1-y), & \text{if } y \ge x, \end{cases}$$

hence $u(x)^{-1} G(x, \cdot) \le 1$ for all $x \in (0,1)$. Therefore, for each $f \in E$ we have

$$\left|(Tf)(x)\right| \le u(x) \int_0^1 \left|f(y)\right| \,\mathrm{d}y \le u(x) \left\|f\right\|_2,$$

which proves $Tf \in E_u$.

(b) For $\delta > 0$ we have $\frac{G(\delta,\delta)}{u(\delta)u(\delta)} = \frac{1}{\delta(1-\delta)} \to \infty$ as $\delta \downarrow 0$, in turn $T \not\preceq u \otimes u$. □

8.3 A sufficient condition for uniform eventual positivity

Our key assumption to get a sufficient condition for uniform eventual positivity (or negativity) of resolvents is that the resolvent satisfies a kernel estimate as described for general operators in Theorem 8.2.1 above. Let us first note that the validity of such an estimate does not depend on the point that one considers within the resolvent set.

Proposition 8.3.1. *Let $A: E \supseteq \operatorname{dom}(A) \to E$ be a closed, densely defined, and real operator on a complex Banach lattice E. Let $u \in E_+$ and let $\varphi \in E'_+$. If there exists a number $\lambda \in \rho(A) \cap \mathbb{R}$ such that $\pm \mathcal{R}(\lambda, A) \preceq u \otimes \varphi$, then the same is true for all $\lambda \in \rho(A) \cap \mathbb{R}$.*

Proof. If the given estimate is true for at least one number $\lambda \in \rho(A) \cap \mathbb{R}$, then $\operatorname{dom}(A) \subseteq E_u$ and $\operatorname{dom}(A') \subseteq (E')_\varphi$. For all other $\tilde{\lambda} \in \rho(A) \cap \mathbb{R}$, one has the resolvent identity

$$\mathcal{R}(\tilde{\lambda}, A) = \mathcal{R}(\lambda, A) + (\lambda - \tilde{\lambda})\mathcal{R}(\tilde{\lambda}, A)\mathcal{R}(\lambda, A),$$

which gives the claim, since $\pm\mathcal{R}(\tilde{\lambda}, A)\mathcal{R}(\lambda, A) \preceq u \otimes \varphi$ by Corollary 8.2.4. □

The main result of this chapter is the following theorem which contains the sufficient conditions for uniform eventual positivity (and negativity) promised in the introduction.

Theorem 8.3.2. *Let $A : E \supseteq \operatorname{dom}(A) \to E$ be a closed, densely defined, and real operator on a complex Banach lattice E. Let $\lambda \in \sigma(A) \cap \mathbb{R}$ be a pole of the resolvent $\mathcal{R}(\cdot, A)$ and assume the following properties:*

(1) *The eigenspace* $\ker(\lambda - A)$ *is spanned by a quasi-interior point* $u \in E_+$.

(2) *The dual eigenspace* $\ker(\lambda - A')$ *contains a strictly positive functional* φ.

(3) *There exists a number* $\lambda_1 \in \rho(A) \cap \mathbb{R}$ *such that* $\pm\mathcal{R}(\lambda_1, A) \preceq u \otimes \varphi$.

Then $\mathcal{R}(\cdot, A)$ is uniformly eventually positive with respect to $u \otimes \varphi$ at λ and uniformly eventually negative with respect to $u \otimes \varphi$ at λ.

Proof. Without loss of generality, we assume that $\lambda = 0$ and $\langle \varphi, u \rangle = 1$.

Step 1: Since φ is a strictly positive eigenvector of A' and $u \in \ker A$ is a quasi-interior point of E_+, we know from Theorem 7.3.6 that 0 is a first order pole and the associated spectral projection P satisfies $Pf \succeq u$ for all $0 \lneq u \in E$. Thus P is a rank-one projection with $Pu = u$ and $P'\varphi = \varphi$. Consequently, $P = u \otimes \varphi$.

Step 2: Let $\mu \in \rho(A) \cap \mathbb{R}$. By the finite expansion of the resolvent (Lemma 6.4.5 for $n = 2$),

$$\mathcal{R}(\mu, A) = \mathcal{R}(\lambda_1, A) + (\lambda_1 - \mu)\mathcal{R}(\lambda_1, A)^2 + R_\mu,$$

where $R_\mu := (\lambda_1 - \mu)^2 \mathcal{R}(\lambda_1, A)\mathcal{R}(\mu, A)\mathcal{R}(\lambda_1, A)$. By assumption, $\mathcal{R}(\lambda_1, A)$ satisfies the equivalent assertions of Theorem 8.2.1, hence so does $\mathcal{R}(\mu, A)$ by Corollary 8.2.2.

Step 3: Using $\lambda_1 \mathcal{R}(\lambda_1, A)P = P$, we can write

$$\mu R_\mu - P = \mathcal{R}(\lambda_1, A)\big((\lambda_1 - \mu)^2 \mu\mathcal{R}(\mu, A) - \lambda_1^2 P\big)\mathcal{R}(\lambda_1, A).$$

Therefore, $\mu R_\mu - P$ also satisfies the equivalent assertions of Theorem 8.2.1 by Corollary 8.2.2. Combining this with the fact $\mu\mathcal{R}(\mu, A) \to P$ in $\mathcal{L}(E)$ as $\mu \to 0$, it follows that $\mu R_\mu \to P$ in $\mathcal{L}(E^\varphi, E_u)$ as $\mu \to 0$.

Step 4 : Steps 2 and 3 together give that even $\mu\mathcal{R}(\mu, A) \to P$ in $\mathcal{L}(E^\varphi, E_u)$ as $\mu \to 0$. Since $P = u \otimes \varphi$ by Step 1, it follows that $\mu\mathcal{R}(\mu, A) \succeq u \otimes \varphi$ for all μ in a neighbourhood of 0. Thus, $\mathcal{R}(\cdot, A)$ is uniformly eventually positive with respect to $u \otimes \varphi$ at 0. □

We now discuss three examples for uniform eventual positivity and negativity, each of them on a bounded interval. For the first one, everything can be computed explicitly, for the other two we use Theorem 8.3.2.

Example 8.3.3 (A first order differential operator)**.** Let $p \in [1,\infty)$. Consider the Banach lattice $E = L^p(0,1)$ and its dual space $E' = L^{p'}(0,1)$, where $p' \in (1,\infty]$ satisfies $\frac{1}{p} + \frac{1}{p'} = 1$. The closed operator A_0 on $L^p(0,1)$ given by

$$\begin{aligned} \operatorname{dom}(A_0) &:= \{f \in W^{1,p}(0,1) : f(0) = f(1)\} \\ A_0 f &:= f' \end{aligned}$$

has the following properties.

(a) A_0 has compact resolvent and its spectrum is $\sigma(A_0) = 2\pi \mathrm{i}\mathbb{Z}$.

(b) The resolvent of A_0 satisfies

$$\mathcal{R}(\mu, A_0) \preceq -\mathbb{1} \otimes \mathbb{1} \;\text{ if } \mu \in (-\infty, 0) \qquad \text{and} \qquad \mathcal{R}(\mu, A_0) \succeq \mathbb{1} \otimes \mathbb{1} \;\text{ if } \mu \in (0, \infty).$$

Proof. (a) One can verify that for each $\mu \in \mathbb{C} \setminus 2\pi\mathrm{i}\mathbb{Z}$, the integral operator with kernel

$$K_\mu(x,y) = e^{\mu(x-y)}\left(-\mathbb{1}_{[y \le x]} + \frac{1}{1-e^{-\mu}}\right). \tag{8.3.1}$$

is inverse to $\mu - A_0$. Hence, $\mu \in \rho(A_0)$ and $\mathcal{R}(\mu, A_0)$ is the integral operator with kernel K_μ. On the other hand, for $\mu \in 2\pi\mathrm{i}\mathbb{Z}$ the function $v \in L^p(0,1)$ given by $v(x) = e^{\mu x}$ is clearly in $\operatorname{dom}(A_0)$ and satisfies $A_0 v = \mu v$, so indeed $\sigma(A_0) = 2\pi\mathrm{i}\mathbb{Z}$. The compactness of the resolvent follows, for instance, from the compactness of the embedding $W^{1,p}(0,1) \hookrightarrow L^p(0,1)$ (Theorem 6.3.1 and Exercise 6.5).

(b) Let $\mu \in \mathbb{R} \setminus \{0\}$. We use that the resolvent $\mathcal{R}(\mu, A_0)$ has the integral kernel given by formula (8.3.1). If $\mu < 0$ the summand $1/(1-e^{-\mu})$, and hence $K_\mu(x,y)$, is strictly negative. By continuity, there exists $c := -\max_{x,y \in [0,1]} K_\mu(x,y) > 0$ such that

$$(\mathcal{R}(\mu, A_0)f)(x) = \int_0^1 K_\mu(x,y) f(y)\,\mathrm{d}y \le -c \int_0^1 f(y)\,\mathrm{d}y = -c\big((\mathbb{1} \otimes \mathbb{1})f\big)(x)$$

for all $0 \le f \in E$. The proof in the case $\mu > 0$ is similar. □

Example 8.3.4 (A third order differential operator)**.** On the Banach lattice $E = L^2(0,1)$, consider the closed operator

$$\begin{aligned} \operatorname{dom}(A) &:= \{f \in H^3(0,1) : f^{(k)}(0) = f^{(k)}(1)\ \forall\, k = 0,1,2\} \\ Af &:= f''', \end{aligned} \tag{8.3.2}$$

and identify E' with E. Then $0 \in \sigma(A)$ is an isolated spectral value of A and $\mathcal{R}(\cdot, A)$ is uniformly eventually positive with respect to $\mathbb{1} \otimes \mathbb{1}$ at 0.

Proof. If A_0 denotes the operator from Example 8.3.3 for $p = 2$, one has $A = A_0^3$. Since $\sigma(A_0) \subseteq \mathrm{i}\mathbb{R}$, this implies that $\rho(A) \neq \emptyset$ and that A has compact resolvent. We verify the assumptions of Theorem 8.3.2.

(1) *and* (2) Observe that $\mathbb{1} \in \ker A$ and all eigenfunctions of A corresponding to the eigenvalue 0 are polynomials of degree at most 2. The periodic boundary conditions imposed in (8.3.2) ensure that they are constant. In other words, $\ker A$ is spanned by $\mathbb{1}$. Since $A' = -A$, the dual eigenspace $\ker A'$ is also spanned by $\mathbb{1}$.

(3) On E, consider the closed operator

$$\begin{aligned} \operatorname{dom}(B) &:= \{u \in H^2(0,1) : u^{(k)}(0) = u^{(k)}(1)\ \forall\, k = 0,1\} \\ Bf &:= f + f' + f'' \end{aligned}$$

Then $1 - A = (1 - A_0)B$ and $1 \in \rho(A) \cap \rho(A_0)$; where A_0 denotes the operator in Example 8.3.3. Therefore $0 \in \rho(B)$ and

$$\mathcal{R}(0,-B)E \subseteq H^2(0,1) \subseteq L^\infty(0,1) = E_{\mathbb{1}},$$

by the Sobolev embedding in Theorem 5.3.7(b). Observing $A_0' = -A_0$, we similarly obtain $\mathcal{R}(1,A_0)'E' \subseteq (E')_{\mathbb{1}}$. Since $\mathcal{R}(1,A) = \mathcal{R}(0,-B)\mathcal{R}(1,A_0)$, it follows from Corollary 8.2.4 that $\pm\mathcal{R}(1,A) \preceq \mathbb{1} \otimes \mathbb{1}$. □

Example 8.3.5 (The Laplacian with non-local boundary conditions, revisited)**.** Consider the Laplace operator $\Delta_B : L^2(0,1) \supseteq \operatorname{dom}(\Delta_B) \to L^2(0,1)$ with non-local boundary conditions from Examples 5.4.3 and 6.3.4, whose domain is

$$\operatorname{dom}(\Delta_B) = \left\{ u \in H^2(0,1) : \begin{pmatrix} -u'(0) \\ u'(1) \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} u(0) \\ u(1) \end{pmatrix} \right\}.$$

Then $\mathcal{R}(\cdot,\Delta_B)$ is uniformly eventually positive and negative with respect to $\mathbb{1} \otimes \mathbb{1}$ at $\mathrm{s}(\Delta_B)$.

Proof. We verify the assumptions of Theorem 8.3.2. Recall from Example 5.4.3 that $\mathrm{s}(\Delta_B) < 0$ and the explicit formula for the resolvent of Δ_B at 0 is given by

$$\mathcal{R}(0,\Delta_B)f(x) = \frac{1}{4}\int_0^1 f(z)\,\mathrm{d}z + \frac{1}{2}\int_x^1\int_0^y f(z)\,\mathrm{d}z\,\mathrm{d}y + \frac{1}{2}\int_0^x\int_y^1 f(z)\,\mathrm{d}z\,\mathrm{d}y \tag{8.3.3}$$

for all $f \in L^2(0,1)$.

Step 1: If $0 \lneq f \in L^2(0,1)$, then formula (8.3.3) shows that

$$\mathcal{R}(0,\Delta_B)f(x) \geq \frac{1}{4}\int_0^1 f(z)\,\mathrm{d}z = \frac{1}{4}\big((\mathbb{1} \otimes \mathbb{1})f\big)(x)$$

for all $x \in [0,1]$, which implies $\mathcal{R}(0,\Delta_B) \succeq \mathbb{1} \otimes \mathbb{1}$. Applying Theorem 5.4.1 with $Q = \mathbb{1} \otimes \mathbb{1}$, we obtain that $\mathcal{R}(\mu,\Delta_B) \succeq \mathbb{1} \otimes \mathbb{1}$ for all $\mu \in (\mathrm{s}(\Delta_B), 0]$. This shows that $\mathcal{R}(\cdot,\Delta_B)$ is uniformly eventually positive with respect to $\mathbb{1}$ at $\mathrm{s}(\Delta_B)$.

Theorem 7.3.6 now implies that $\ker(\mathrm{s}(\Delta_B) - \Delta_B)$ is spanned by a vector $v \succeq \mathbb{1}$, and the dual eigenspace $\ker(\mathrm{s}(\Delta_B) - \Delta_B')$ contains a strictly positive functional φ. The property $v \succeq \mathbb{1}$ clearly shows that v is a quasi-interior point.

Step 2: Formula (8.3.3) also directly yields the estimate

$$\left|\mathcal{R}(0,\Delta_B)f\right| \le \left(\frac{1}{4}+\frac{1}{2}+\frac{1}{2}\right)\|f\|_{L^1(0,1)} = \frac{5}{4}\langle \mathbb{1}, |f|\rangle \mathbb{1}$$

for all $f \in L^2(0,1)$. Thus $\pm\mathcal{R}(0,\Delta_B) \preceq \mathbb{1}\otimes\mathbb{1}$.

We have thus verified all the conditions of Theorem 8.3.2 and conclude that $\mathcal{R}(\cdot,\Delta_B)$ is also uniformly eventually negative with respect to $\mathbb{1}\otimes\mathbb{1}$ at $\mathrm{s}(\Delta_B)$. □

8.4 Intermezzo: Hilbert space adjoints vs. Banach space duals

Our extensive use of form methods to construct linear operators on Hilbert spaces makes it worthwhile to spend a short intermezzo on clarifying the relation between dual operators on Banach space, adjoint operators on Hilbert spaces, and adjoints of sesquilinear forms. We use this in Example 8.5.2 in the next section. The definition of adjoint operators on a Hilbert space H is very similar to that of dual operators (Definition 3.1.5).

Definition 8.4.1 (Adjoint operators and forms)**.** Let V, H be complex Hilbert spaces.

(a) Let $A\colon H \supseteq \operatorname{dom}(A) \to H$ be densely defined linear operator. The **adjoint operator** $A^*\colon H \supseteq \operatorname{dom}(A^*) \to H$ is defined by

$$\begin{aligned} \operatorname{dom}\left(A^*\right) &:= \left\{x \in H \mid \exists\, y \in H\colon\ (x \mid Av) = \left(y \mid v\right)\ \forall\, v \in \operatorname{dom}(A)\right\} \\ A^* x &:= y, \end{aligned}$$

where y in the second line is the vector that occurs in the definition of $\operatorname{dom}(A^*)$.[3]

The operator A is called **self-adjoint** if $A^* = A$.

(b) Let $\mathfrak{a}\colon V\times V \to \mathbb{C}$ be a sesquilinear form. The form $\mathfrak{a}^*\colon V\times V\to\mathbb{C}$ given by $\mathfrak{a}^*(u,v) := \overline{\mathfrak{a}(v,u)}$ for all $u, v \in V$ is called the **adjoint form** of $\mathfrak{a}$.

Recall from Theorem 5.1.4(c) that a sesquilinear form $\mathfrak{a}$ is called **symmetric** if $\mathfrak{a}(u,v) = \overline{\mathfrak{a}(v,u)}$ for all $u, v \in V$. In other words, $\mathfrak{a}$ is symmetric if and only if $\mathfrak{a}^* = \mathfrak{a}$.

The adjoint and the dual of a Hilbert space operator are related as follows.

Proposition 8.4.2. *Let $A\colon H \supseteq \operatorname{dom}(A) \to H$ be densely defined linear operator on a complex Hilbert space H. Consider the mapping $J\colon H \to H'$, $x \mapsto (x \mid \cdot)$, which is an anti-linear isometric bijection by the Riesz–Fréchet theorem. Then $J\big(\operatorname{dom}(A^*)\big) = \operatorname{dom}\left(A'\right)$ and the following diagram commutes:*

$$\begin{array}{ccccc} H & \supseteq & \operatorname{dom}(A^*) & \xrightarrow{A^*} & H \\ \downarrow J & & \downarrow J & & \downarrow J \\ H' & \supseteq & \operatorname{dom}\left(A'\right) & \xrightarrow[A']{} & H' \end{array}$$

[3]Observe that y is unique by density of $\operatorname{dom}(A)$ in H.

Proof. This follows by chasing the definitions. □

Let (Ω, ν) be a σ-finite measure space. Observe that there are two common ways to identify $L^2(\Omega, \nu)$ with its dual space. When thinking mainly about Hilbert space, one typically identifies each $f \in L^2(\Omega, \nu)$ with the functional $\left(f \mid \cdot\right) = \int_\Omega \overline{f} \cdot d\nu$. This identification is actually the anti-linear isomorphism from the Riesz-Fréchet representation theorem that we called J in Proposition 8.4.2. On the other hand, for $p \in [1, \infty)$ and $p' \in (1, \infty]$ with $1/p + 1/p' = 1$, it is more common to identify $L^{p'}(\Omega, \nu)$ with the dual space $\left(L^p(\Omega, \nu)\right)'$ by identifying each function $f \in L^{p'}(\Omega, \nu)$ with the functional $\int_\Omega f \cdot d\nu$. Note that the isomorphism that maps f to this functional is linear rather than antilinear. However, for real-valued functions f both isomorphisms associate the same functional to f.

In particular, if A is a real operator and $\lambda \in \mathbb{R}$ then the real-valued elements of $\ker(\lambda - A^*)$ and $\ker(\lambda - A')$ coincide under those identifications. This is often useful to find out information on the dual operator A' since A^* can be described, again, by form method:

Proposition 8.4.3 (The adjoint operator via the adjoint form)**.** *Under the assumptions of Theorem 5.1.4, let $A\colon H \supseteq \operatorname{dom}(A) \to H$ be the operator associated to the form $\mathfrak{a}\colon V \times V \to \mathbb{C}$. Then A^* is the operator associated to the adjoint form $\mathfrak{a}^*$. In particular, if $\mathfrak{a}$ is symmetric, then A is self-adjoint.*

Giving in to the belief that this proposition will not appear too surprising, we refrain from discussing the proof and instead return to kernel estimates and eventual positivity.

8.5 Kernel estimates for resolvents via forms

We close this chapter with a tool to check the key assumption $\pm\mathcal{R}(\lambda_1, A) \preceq u \otimes \varphi$ in Theorem 8.3.2 if the operator A is associated to a sesquilinear form on L^2.

Proposition 8.5.1. *Let $H = L^2(\Omega, \nu)$ for a σ-finite measure space (Ω, ν) and let V be a complex Hilbert space such that V embeds continuously and densely into H. Let $\mathfrak{a}\colon V \times V \to \mathbb{C}$ be a bounded, real sesquilinear form on V that satisfies the ellipticity estimate*

$$\operatorname{Re}\mathfrak{a}(v, v) + \mu \|v\|_H^2 \geq \delta \|v\|_V^2$$

for some numbers $\mu \in \mathbb{R}$ and $\delta > 0$ and for all $v \in V$.

If $u \in E_+$ satisfies[4] *$V \subseteq H_u$, then the operator A associated with $\mathfrak{a}$ satisfies $\pm\mathcal{R}(\lambda, A) \preceq u \otimes u$ for one, hence all, $\lambda \in \rho(A) \cap \mathbb{R}$.*

Proof. By shifting the form we may assume that $\mu < 0$. Hence, $0 \in \rho(A)$ according to Theorem 5.1.4(b). Note that A is a real operator since $\mathfrak{a}$ is assumed to be real; hence, so is $\mathcal{R}(0, A)$. Let $x \in H$ and set $w := \mathcal{R}(0, A)x \in \operatorname{dom}(A) \subseteq V$. The inclusion map from V into H_u is continuous by the closed graph theorem, say with norm c. Hence,

$$\|w\|_{H_u}^2 \leq c^2 \|w\|_V^2 \leq \frac{c^2}{\delta} \operatorname{Re}\mathfrak{a}(w, w) \leq \frac{c^2}{\delta} |\mathfrak{a}(w, w)|$$

[4] This implies that u is a quasi-interior point since V is dense in H.

$$= \frac{c^2}{\delta} |(w \mid x)_H| \le \frac{c^2}{\delta} \|w\|_{H_u} \big(u \mid |x|\big)_H = \frac{c^2}{\delta} \|w\|_{H_u} \langle u, |x|\rangle,$$

so $\|w\|_{H_u} \le \frac{c^2}{\delta} \langle u, |x|\rangle$. In other words, $|\mathcal{R}(0,A)x| = |w| \le \frac{c^2}{\delta} \langle u, |x|\rangle\, u$. According to Theorem 8.2.1 this means that $\pm\mathcal{R}(0,A) \preceq u \otimes u$. □

Example 8.5.2 (The Neumann Laplacian on an interval)**.** Consider the sesquilinear form $\mathfrak{a}\colon H^1(0,\pi) \times H^1(0,\pi) \to \mathbb{C}$, $\mathfrak{a}(v,w) := \big(v' \mid w'\big)_{L^2}$ on $L^2(0,\pi)$. Its associated operator Δ_{Neu} acts as the weak second derivative on the domain

$$\operatorname{dom}(\Delta_{\text{Neu}}) = \big\{u \in H^2(0,\pi) : u'(0) = u'(\pi) = 0\big\}.$$

The operator Δ_{Neu} is called the **Neumann Laplace operator** and has the following properties.

(a) Δ_{Neu} has compact resolvent.

(b) $\operatorname{s}(\Delta_{\text{Neu}}) = 0$ is an eigenvalue and $\ker \Delta_{\text{Neu}}$ and $\ker \Delta'_{\text{Neu}}$ are both spanned by $\mathbb{1}$.

(c) For every $\lambda > 0$ one has $\mathcal{R}(\lambda, \Delta_{\text{Neu}}) \ge 0$.

(d) The resolvent $\mathcal{R}(\cdot, \Delta_{\text{Neu}})$ is uniformly eventually positive and negative with respect to $\mathbb{1} \otimes \mathbb{1}$ at 0.

Proof. The fact that Δ_{Neu} has the claimed domain and acts as the weak second derivative, is a special case of Exercise 5.6(a). Let us show that Δ_{Neu} has the claimed properties.

(b) For each $\mu > 0$, we have

$$\operatorname{Re}\mathfrak{a}(v,v) + \mu \|v\|_{L^2}^2 = \big\|v'\big\|_{L^2}^2 + \mu \|v\|_{L^2}^2 \ge \min\{1,\mu\} \|v\|_{H^1}^2$$

for all $v \in H^1(0,\pi)$. Therefore by Theorem 5.1.4, the associated operator is closed and densely defined, and $\operatorname{s}(\Delta_{\text{Neu}}) \le 0$. In fact, $\operatorname{s}(\Delta_{\text{Neu}}) = 0$, as one can easily check that 0 is an eigenvalue and the corresponding eigenspace is spanned by $\mathbb{1}$. Since the form $\mathfrak{a}$ is symmetric, the operator Δ_{Neu} is self-adjoint (Proposition 8.4.3) and hence, the dual eigenspace is also spanned by $\mathbb{1}$.

(a) The embedding $H^1(0,\pi) \hookrightarrow L^2(0,\pi)$ is compact by Theorem 6.3.1 and we have shown in the proof of (b) that the assumptions of Theorem 5.1.4 are fulfilled. It follows that Δ_{Neu} has compact resolvent by Proposition 6.2.10(b).

(c) Recall from Example 4.1.4(d) that $H^1(0,\pi;\mathbb{R}) = H^1(0,\pi) \cap L^2((0,\pi);\mathbb{R})$ is a sublattice of $L^2((0,\pi);\mathbb{R})$ and

$$\mathfrak{a}(v^-, v^+) = \big(-\mathbb{1}_{v<0}\, v' \mid \mathbb{1}_{v>0}\, v'\big)_{L^2} = 0$$

for all $v \in H^1(0,\pi;\mathbb{R})$. The Beurling–Deny criterion (Theorem 5.1.7) therefore implies the assertion.

(d) By the one-dimensional Sobolev embeddings in Theorem 5.3.7(b), we know that $H^1(0,\pi) \subseteq L^\infty(0,\pi) = L^2(0,\pi)_{\mathbb{1}}$. As a result $\pm\mathcal{R}(\lambda, \Delta_{\text{Neu}}) \preceq \mathbb{1} \otimes \mathbb{1}$ for all $\lambda > 0$ by Proposition 8.5.1. Thus, all assumptions of Theorem 8.3.2 are fulfilled with $u = \varphi = \mathbb{1}$, whence the assertion follows. □

Exercises for Chapter 8

Exercise 8.1 (AL-spaces generated by functionals vs. principal ideals)**.** Let E be a Banach lattice and $\varphi \in E'$ a strictly positive functional. Show that the Banach lattices $(E^\varphi)'$ and $(E')_\varphi$ are isomorphic.

More precisely, show that $J\colon (E^\varphi)' \to E'$, $\psi \mapsto \psi|_E$ maps the dual space $(E^\varphi)'$ bijectively to the principal ideal $(E')_\varphi$ and that $J\psi \geq 0$ if and only if $\psi \geq 0$.

Exercise 8.2 (A fourth order operator on an interval, continued)**.** Consider the fourth order differential operator $A\colon L^2(0,1) \supseteq \operatorname{dom}(A) \to L^2(0,1)$ from Exercise 7.2. We know from that exercise that $\mathcal{R}(\cdot, A)$ is individually eventually positive with respect to $\mathbb{1}$ at the spectral bound 0. Now we improve this result.

(a) Show that $\mathcal{R}(\cdot, A)$ is uniformly eventually positive with respect to $\mathbb{1} \otimes \mathbb{1}$ at 0.

(b) Is $\mathcal{R}(\cdot, A)$ also uniformly eventually negative with respect to $\mathbb{1} \otimes \mathbb{1}$ at 0?

Exercise 8.3 (Yet another fourth order operator on an interval)**.** Consider the Dirichlet Laplace operator $\Delta_{\mathrm{Dir}}\colon L^2(0,\pi) \supseteq \operatorname{dom}(\Delta_{\mathrm{Dir}}) \to L^2(0,\pi)$. Define $A := -\Delta_{\mathrm{Dir}}^2$.

(a) Show that $\sigma(A) \subseteq (-\infty, 0)$ and that A has compact resolvent.

(b) Prove that $\mathcal{R}(\cdot, A)$ is uniformly eventually positive and negative with respect to $\sin \otimes \sin$ at $\mathrm{s}(A)$.

Exercise 8.4 (A first order operator on a space of continuous functions)**.** In this exercise we revisit the differential operator A_0 from Example 8.3.3, but now on the space $\mathrm{C}([0,1])$. Let $A_0\colon \mathrm{C}([0,1]) \supseteq \operatorname{dom}(A_0) \to \mathrm{C}([0,1])$ be given by

$$\begin{aligned} \operatorname{dom}(A_0) &:= \left\{ f \in \mathrm{C}^1([0,1]) : f(0) = f(1) \right\} \\ A_0 f &:= f'. \end{aligned}$$

(a) Show that A_0 has compact resolvent, that $\sigma(A_0) = 2\pi \mathrm{i}\mathbb{Z}$, and that the resolvent at $\mu \in \rho(A_0)$ is given by the integration against the kernel from formula (8.3.1).

(b) Find a strictly positive functional $\varphi \in \mathrm{C}([0,1])$ that spans the kernel $\ker A'$ of the dual operator. Is φ is a quasi-interior point of $\mathrm{C}([0,1])'_+$?

(c) Prove that $\mathcal{R}(\mu, A_0) \preceq -\mathbb{1} \otimes \varphi$ for all $\mu \in (-\infty, 0)$ and $\mathcal{R}(\mu, A_0) \succeq \mathbb{1} \otimes \varphi$ for all $\mu \in (0, \infty)$ where φ is the functional from (b).

Notes for Chapter 8

Sufficient conditions for uniform eventual positivity

Theorem 8.3.2 is a special case of the main result of [AG22], which was in turn inspired by earlier results of Takáč [Tak96]. In fact, one can show the same conclusion as in Theorem 8.3.2 if the assumption $\pm\mathcal{R}(\lambda_1, A) \preceq u \otimes \varphi$ is replaced with the one-sided estimate $\mathcal{R}(\lambda_1, A) \succeq -u \otimes \varphi$ along with the weaker assumptions $\operatorname{dom}(A^m) \subseteq E_u$ and $\operatorname{dom}\big((A')^m\big) \subseteq (E')_\varphi$ for some $m \in \mathbb{N}$; see [AG22, Theorem 1.2].

For concrete differential operators, the domain embeddings usually follow from elliptic regularity results (up to the boundary) combined with Sobolev embedding theorems – a theme that we have already used multiple times for the Dirichlet Laplace operators. The lower estimate $\mathcal{R}(\lambda_1, A) \succeq -u \otimes \varphi$ can sometimes be shown for higher order differential operators as a consequence of the observation that, while the integral kernel of the resolvent – i.e. the **Green's function** – of such operators is not positive, its singularity typically is. Such concrete kernel estimates are, for instance, shown in [DMS05, Theorem 1.5], [GR10, Theorem 1], and [Pul15, Theorem 4.1]. It would be very desirable to have an general operator theoretic explanations for this kind of behaviour, but we are currently not aware of any such abstract explanation.

Automatic compactness

The assumption in Theorem 8.3.2 that λ be a pole of the resolvent, is actually redundant. In fact, the assumption $\pm\mathcal{R}(\lambda_1, A) \preceq u \otimes \varphi$ implies that $\mathcal{R}(\lambda_1, A)^3$ is compact. Indeed, it is a classical result in Banach lattice theory that if $0 \le S_k \le T_k$ for linear operators and for indices $k \in \{1, 2, 3\}$, then compactness of all T_k imply that $S_3 S_2 S_1$ is compact [MN91, Corollary 3.7.14]. From this one can readily derives that if $\pm S \preceq T$ and T is compact, then S^3 is compact; see [AG22, Corollary 2.7] for details. Since $\mathcal{R}(\lambda_1, A)^3$ is compact, one can derive from analytic Fredholm theory that every spectral value of A is a pole of $\mathcal{R}(\cdot\,, A)$.

Bibliography

[AB06] Charalambos D. Aliprantis and Owen Burkinshaw. *Positive operators*. Berlin: Springer, reprint of the 1985 original edition, 2006.

[ABHN11] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser/Springer Basel AG, Basel, second edition, 2011.

[AE01] Herbert Amann and Joachim Escher. *Analysis III*. Grundstud. Math. Basel: Birkhäuser, 2001.

[AF03] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[AG22] Sahiba Arora and Jochen Glück. An operator theoretic approach to uniform (anti-)maximum principles. *J. Differential Equations*, 310:164–197, 2022.

[AG23] Sahiba Arora and Jochen Glück. A characterization of the individual maximum and anti-maximum principle. *Math. Z.*, 305(2):Paper No. 24, 17, 2023.

[Akh18] Khalid Akhlil. Locality and domination of semigroups. *Result. Math.*, 73(2):11, 2018. Id/No 59.

[AN09] Wolfgang Arendt and Robin Nittka. Equivalent complete norms and positivity. *Arch. Math.*, 92(5):414–427, 2009.

[Are06] Wolfgang Arendt. Heat kernels: ISEM 2005/6, 2006. Available at https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/arendt/downloads/internetseminar.pdf.

[AT07] Charalambos D. Aliprantis and Rabee Tourky. *Cones and duality*, volume 84 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2007.

[AU23] Wolfgang Arendt and Karsten Urban. *Partial differential equations. An introduction to analytical and numerical methods. Translated from the German by James B. Kennedy*, volume 294 of *Grad. Texts Math.* Cham: Springer, 2023.

[BKFR17] András Bátkai, Marjeta Kramar Fijavž, and Abdelaziz Rhandi. *Positive operator semigroups*, volume 257 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer, Cham, 2017. From finite to infinite dimensions. With a foreword by Rainer Nagel and Ulf Schlotterbeck.

[Bou03] Nicolas Bourbaki. *Elements of mathematics. Algebra II. Chapters 4–7. Transl. from the French by P. M. Cohn and J. Howie.* Berlin: Springer, reprint of the 1990 English translation edition, 2003.

[Bou07] Nicolas Bourbaki. *Éléments de mathématique. Algèbre. Chapitres 4 à 7.* Berlin: Springer, reprint of the 1981 original edition, 2007.

[BP94] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Class. Appl. Math.* Philadelphia, PA: SIAM, 1994.

[BR84] Charles J. K. Batty and Derek W. Robinson. Positive one-parameter semigroups on ordered Banach spaces. *Acta Appl. Math.*, 2:221–296, 1984.

[Bra61] Alfred Brauer. On the characteristic roots of power-positive matrices. *Duke Math. J.*, 28:439–445, 1961.

[Bre11] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011.

[BY84] Jonathan M. Borwein and David T. Yost. Absolute norms on vector lattices. *Proc. Edinb. Math. Soc., II. Ser.*, 27:215–222, 1984.

[CD13] Alexander P. Campbell and Daniel Daners. Linear algebra via complex analysis. *Amer. Math. Monthly*, 120(10):877–892, 2013.

[CS00] Philippe Clément and Guido Sweers. Uniform anti-maximum principles. *J. Differ. Equations*, 164(1):118–154, 2000.

[CS01] Philippe Clément and Guido Sweers. Uniform anti-maximum principle for polyharmonic boundary value problems. *Proc. Am. Math. Soc.*, 129(2):467–474, 2001.

[DG17] Daniel Daners and Jochen Glück. The role of domination and smoothing conditions in the theory of eventually positive semigroups. *Bull. Aust. Math. Soc.*, 96(2):286–298, 2017.

[DG18] Daniel Daners and Jochen Glück. Towards a perturbation theory for eventually positive semigroups. *J. Operator Theory*, 79(2):345–372, 2018.

[DGK16a] Daniel Daners, Jochen Glück, and James B. Kennedy. Eventually and asymptotically positive semigroups on Banach lattices. *J. Differential Equations*, 261(5):2607–2649, 2016.

[DGK16b] Daniel Daners, Jochen Glück, and James B. Kennedy. Eventually positive semigroups of linear operators. *J. Math. Anal. Appl.*, 433(2):1561–1593, 2016.

[DL00] Robert Dautray and Jacques-Louis Lions. *Mathematical analysis and numerical methods for science and technology. Volume 2: Functional and variational methods. With the collaboration of Michel Artola, Marc Authier, Philippe Bénilan, Michel Cessenat, Jean-Michel Combes, Hélène Lanchon, Bertrand Mercier, Claue Wild, Claude Zuily. Transl. from the French by Ian N. Sneddon.* Berlin: Springer, 2nd printing edition, 2000.

[DMS05] Anna Dall'Acqua, Christian Meister, and Guido Sweers. Separating positivity and regularity for fourth order Dirichlet problems in 2d-domains. 25(3):205–261, 2005.

[DPZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encycl. Math. Appl.* Cambridge: Cambridge University Press, 2nd ed. edition, 2014.

[Dyn65] Evgeniĭ B. Dynkin. *Markov processes. Vols. I, II. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra and G. Majone.*, volume 121/122 of *Grundlehren Math. Wiss.* Springer, Cham, 1965.

[EN00] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[ES08] Abed Elhashash and Daniel B. Szyld. On general matrices having the Perron-Frobenius property. *Electron. J. Linear Algebra*, 17:389–413, 2008.

[ES09] Abed Elhashash and Daniel B. Szyld. Two characterizations of matrices with the Perron-Frobenius property. *Numer. Linear Algebra Appl.*, 16(11-12):863–869, 2009.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2nd ed. edition, 2010.

[Fen98] Gero Fendler. On dilations and transference for continuous one-parameter semigroups of positive contractions on $\mathcal{L}^p$-spaces. *Ann. Univ. Sarav., Ser. Math.*, 9(1):1–97, 1998.

[Fra11] L. Edward Fraenkel. *An introduction to maximum principles and symmetry in elliptic problems*, volume 128 of *Camb. Tracts Math.* Cambridge: Cambridge University Press, reprint of the 2000 hardback edition edition, 2011.

[Fri78] Shmuel Friedland. On an inverse problem for nonnegative and eventually nonnegative matrices. *Israel J. Math.*, 29(1):43–60, 1978.

[Fro08] Georg Frobenius. Über Matrizen aus positiven Elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 471–476, 1908.

[Fro09] Georg Frobenius. Über Matrizen aus positiven Elementen. II. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 514–518, 1909.

[Fro12] Georg Frobenius. Über Matrizen aus nicht negativen Elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 456–477, 1912.

[GGK90] Israel Gohberg, Seymour Goldberg, and Marinus A. Kaashoek. *Classes of linear operators. Vol. I*, volume 49 of *Oper. Theory: Adv. Appl.* Basel etc.: Birkhäuser Verlag, 1990.

[GH23] Jochen Glück and Julian Hölz. Eventual cone invariance revisited. *Linear Algebra Appl.*, 675:274–293, 2023.

[Glü16] Jochen Glück. *Invariant Sets and Long Time Behaviour of Operator Semigrouos.* PhD thesis, Ulm University, 2016.

[GR10] Hans-Christoph Grunau and Frédéric Robert. Positivity and almost positivity of biharmonic Green's functions under Dirichlet boundary conditions. *Archive for Rational Mechanics and Analysis*, 195(3):865–898, Mar 2010.

[Gra14] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Grad. Texts Math.* New York, NY: Springer, 3rd ed. edition, 2014.

[Gri11] Pierre Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Class. Appl. Math.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), reprint of the 1985 hardback ed. edition, 2011.

[Gru09] Gerd Grubb. *Distributions and operators*, volume 252 of *Grad. Texts Math.* New York, NY: Springer, 2009.

[GT01] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[GW20] Jochen Glück and Martin R. Weber. Almost interior points in ordered Banach spaces and the long-term behaviour of strongly positive operator semigroups. *Stud. Math.*, 254(3):237–263, 2020.

[Hal13] Brian C. Hall. *Quantum theory for mathematicians*, volume 267 of *Grad. Texts Math.* New York, NY: Springer, 2013.

[Haw08] Thomas Hawkins. Continued fractions and the origins of the Perron-Frobenius theorem. *Arch. Hist. Exact Sci.*, 62(6):655–717, 2008.

[Hen05] Dan Henry. *Perturbation of the boundary in boundary-value problems of partial differential equations. With editorial assistance from Jack Hale and Antônio Luiz Pereira*, volume 318 of *Lond. Math. Soc. Lect. Note Ser.* Cambridge: Cambridge University Press, 2005.

[HvNVW16] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces. Volume I. Martingales and Littlewood-Paley theory*, volume 63 of *Ergeb. Math. Grenzgeb., 3. Folge.* Cham: Springer, 2016.

[JT04] Charles R. Johnson and Pablo Tarazaga. On matrices with Perron-Frobenius properties and some negative entries. *Positivity*, 8(4):327–338, 2004.

[Kas17] Michael Kasigwa. *Eventual Cone Invariance.* ProQuest LLC, Ann Arbor, MI, 2017. Thesis (Ph.D.)–Washington State University.

[Ken94] Carlos E. Kenig. *Harmonic analysis techniques for second order elliptic boundary value problems: dedicated to the memory of Professor Antoni Zygmund*, volume 83 of *Reg. Conf. Ser. Math.* Providence, RI: American Mathematical Society, 1994.

[Kes17] Srinivasan Kesavan. A note on the grand theorems of functional analysis. *Math. Newsl., Ramanujan Math. Soc.*, 27(3):188–191, 2017.

[Kes21] Srinivasan Kesavan. The grand theorems of functional analysis revisited: a Baire-free approach. *Math. Newsl., Ramanujan Math. Soc.*, 31(3):89–93, 2021.

[KLS89] Mark A. Krasnosel'skii, Evgeniĭ A. Lifshits, and Mark V. Sobolev. *Positive linear systems. - The method of positive operators - Transl. from the Russian by Jürgen Appell*, volume 5 of *Sigma Ser. Appl. Math.* Berlin: Heldermann-Verlag, 1989.

[KT17] Michael Kasigwa and Michael J. Tsatsomeros. Eventual cone invariance. *Electron. J. Linear Algebra*, 32:204–216, 2017.

[KvG19] Anke Kalauch and Onno van Gaans. *Pre-Riesz spaces*, volume 66 of *De Gruyter Expo. Math.* Berlin: De Gruyter, 2019.

[Leo09] Giovanni Leoni. *A first course in Sobolev spaces*, volume 105 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2009.

[Lot68] Heinrich P. Lotz. Über das Spektrum positiver Operatoren. *Math. Z.*, 108:15–32, 1968.

[Lue82] Jesper Luetzen. *The prehistory of the theory of distributions*, volume 7 of *Stud. Hist. Math. Phys. Sci.* Springer-Verlag, New York, NY, 1982.

[LZ71] Wilhelmus A. J. Luxemburg and Adriaan C. Zaanen. *Riesz spaces. Vol. I*, volume 1 of *North-Holland Math. Libr.* Elsevier (North-Holland), Amsterdam, 1971.

[Mac00] Charles R. MacCluer. The many proofs and applications of Perron's theorem. *SIAM Review*, 42(3):487–498, 2000.

[Maz11] Vladimir G. Maz'ya. *Sobolev spaces. With applications to elliptic partial differential equations. Transl. from the Russian by T. O. Shaposhnikova*, volume 342 of *Grundlehren Math. Wiss.* Berlin: Springer, 2nd revised and augmented ed. edition, 2011.

[MN91] Peter Meyer-Nieberg. *Banach lattices.* Universitext. Springer-Verlag, Berlin, 1991.

[MS64] Norman G. Meyers and James Serrin. $H = W$. *Proc. Natl. Acad. Sci. USA*, 51:1055–1056, 1964.

[MST99] Gustavo A. Muñoz, Yannis Sarantopoulos, and Andrew Tonge. Complexifications of real Banach spaces, polynomials and multilinear maps. *Stud. Math.*, 134(1):1–33, 1999.

[MW74] Günter Mittelmeyer and Manfred Wolff. Über den Absolutbetrag auf komplexen Vektorverbänden. *Math. Z.*, 137:87–92, 1974.

[Nou06] Dimitrios Noutsos. On Perron-Frobenius property of matrices having some negative entries. *Linear Algebra Appl.*, 412(2-3):132–153, 2006.

[NT08] Dimitrios Noutsos and Michael J. Tsatsomeros. Reachability and holdability of nonnegative states. *SIAM J. Matrix Anal. Appl.*, 30(2):700–712, 2008.

[Ouh05] El Maati Ouhabaz. *Analysis of heat equations on domains*, volume 31 of *London Mathematical Society Monographs Series.* Princeton University Press, Princeton, NJ, 2005.

[Per07a] Oskar Perron. Grundlagen für eine theorie des jacobischen kettenbruchalgorithmus. *Math. Ann.*, 64(1):1–76, 1907.

[Per07b] Oskar Perron. Zur Theorie der Matrices. *Math. Ann.*, 64(2):248–263, 1907.

[Pie07] Albrecht Pietsch. *History of Banach spaces and linear operators.* Birkhäuser Boston, Inc., Boston, MA, 2007.

[PS07] Patrizia Pucci and James Serrin. *The maximum principle*, volume 73 of *Prog. Nonlinear Differ. Equ. Appl.* Basel: Birkhäuser, 2007.

[Pul15] Ludwig Pulst. *Dominance of positivity of the Green's function associated to a perturbed polyharmonic Dirichlet boundary value problem by pointwise estimates.* PhD thesis, Otto-von-Guericke-Universität Magdeburg, 2015. DOI: 10.25673/4208.

[PW84] Murray H. Protter and Hans F. Weinberger. Maximum principles in differential equations. Corr. reprint. New York etc.: Springer-Verlag. X, 261 p. DM 79.00 (1984)., 1984.

[Rot94] Walter Roth. A combined approach to the fundamental theorems for normed spaces. *Bull. Inst. Math., Acad. Sin.*, 22(1):83–89, 1994.

[RS80] Michael Reed and Barry Simon. Methods of modern mathematical physics. I: Functional analysis. Rev. and enl. ed. New York etc.: Academic Press, A Subsidiary of Harcourt Brace Jovanovich, Publishers, XV, 400 p. $ 24.00 (1980)., 1980.

[SA17] Fatemeh Shakeri and Rahim Alizadeh. Nonnegative and eventually positive matrices. *Linear Algebra Appl.*, 519:19–26, 2017.

[Sch60] Helmut H. Schaefer. Some spectral properties of positive linear operators. *Pac. J. Math.*, 10:1009–1019, 1960.

[Sch74] Helmut H. Schaefer. *Banach lattices and positive operators.* Die Grundlehren der mathematischen Wissenschaften, Band 215. Springer-Verlag, New York-Heidelberg, 1974.

[Sch21] René L. Schilling. *Brownian motion. A guide to random processes and stochastic calculus. With a chapter on simulation by Björn Böttcher.* De Gruyter Grad. Berlin: De Gruyter, 3rd revised and extended edition edition, 2021.

[Sen06] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Ser. Stat. New York, NY: Springer, revised reprint of the 2nd ed. edition, 2006.

[SG01] Guido Sweers and Hans-Christoph Grunau. Optimal conditions for anti-maximum principles. *Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser.*, 30(3-4):499–513, 2001.

[She51] Seymour Sherman. Order in operator algebras. *Am. J. Math.*, 73:227–232, 1951.

[Soo19] Aivar Sootla. Properties of eventually positive linear input-output systems. *IET Control Theory Appl.*, 13(7):891–897, 2019.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*, volume 30 of *Princeton Math. Ser.* Princeton University Press, Princeton, NJ, 1970.

[Str03] Robert S. Strichartz. *A guide to distribution theory and Fourier transforms.* River Edge, NJ: World Scientific, 2003.

[Tak96] Peter Takáč. An abstract form of maximum and anti-maximum principles of Hopf's type. *J. Math. Anal. Appl.*, 201(2):339–364, 1996.

[TCDF15] Francesco Tudisco, Valerio Cardinali, and Carmine Di Fiore. On complex power nonnegative matrices. *Linear Algebra Appl.*, 471:449–468, 2015.

[TRH01] Pablo Tarazaga, Marcos Raydan, and Ana Hurman. Perron-Frobenius theorem for matrices with some negative entries. *Linear Algebra Appl.*, 328(1-3):57–68, 2001.

[vN97] Jan M. A. M. van Neerven. The norm of a complex Banach lattice. *Positivity*, 1(4):381–390, 1997.

[Wnu99] Witold Wnuk. *Banach lattices with order continuous norms.* Warsaw: Polish Scientific Publishers PWN, 1999.

[Wul17] Boris Zacharowitsch Wulich. *Geometrie der Kegel: in normierten Räumen.* De Gruyter Stud. Berlin: De Gruyter, 2017.

[Zaa83] Adriaan C. Zaanen. *Riesz spaces II*, volume 30 of *North-Holland Math. Libr.* Elsevier (North-Holland), Amsterdam, 1983.

[Zaa97] Adriaan C. Zaanen. *Introduction to operator theory in Riesz spaces.* Springer-Verlag, Berlin, 1997.

[ZT99] Boris G. Zaslavsky and Bit-Shun Tam. On the Jordan form of an irreducible matrix with eventually non-negative powers. *Linear Algebra Appl.*, 302/303:303–330, 1999. Special issue dedicated to Hans Schneider (Madison, WI, 1998).