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 $[0,\infty)$

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 $\Omega \subset \mathbb{R}^n$ 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
$\mathrm{C}^\infty(\Omega)$	the intersection $\bigcap_{m\in\mathbb{N}_0} \mathrm{C}^m(\Omega)$ for a non-empty open subset $\Omega \subset \mathbb{R}^n$
$\mathrm{C}_c^\infty(\Omega)$	the space of test functions on a non-empty open subset $\Omega \subset \mathbb{R}^n$ – i.e. of all functions in $\mathrm{C}^\infty(\Omega)$ 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 $\overline{\Omega}$
$L^p(\Omega, \nu)$	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 $\Omega \subset \mathbb{R}^n$ 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 $k \in \mathbb{N}_0$
$W^{k,p}_0(\Omega)$	the closure of the space of test functions $C^\infty_c(\Omega)$ 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 $k \in \mathbb{N}_0$
$H^k_0(\Omega)$	short for $W^{k,2}_0(\Omega)$, 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

Spectral theory

$\sigma(A)$	spectrum 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

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

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

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 $j \in \{1,\dots,n\}$ 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 $\lambda\,\mathrm{id} - A$; where $\mathrm{id} \in \mathbb{C}^{n\times n}$, 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 $|\lambda| > \mathrm{r}(A)$, 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 $\mathrm{r}(A/\lambda) < 1$, 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

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

Here we used that $\|A^{k+1}/\lambda^{k+1}\| \to 0$ as $k \to \infty$ according to Proposition 1.2.2 since $\mathrm{r}(A/\lambda) < 1$. 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 $x : [0,\infty) \to \mathbb{C}^{n\times n}$, $t \mapsto e^{tA}x_0$ 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

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

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 $u^\mathrm{T} A v \geq 0$.

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

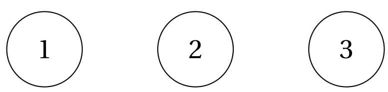

Figure 1.3.1: Eucalyptus trees in an Australian forest (some imagination is required).

- 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 $\mathrm{r}(P) = 1$ 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 $\mathrm{r}(A) > 0$. Use the Jordan normal form of A to show that there exists a number $c > 0$ such that $\|A^k\| \le c\,\mathrm{r}(A)^k(1 + k^{n-1})$ 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 $\|e^{tA}\| \le ce^{t\,\mathrm{s}(A)}(1 + t^{n-1})$ 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 $\alpha \geq 0$ 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 $\alpha \geq 0$.

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 its 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 $x \geq 0$ 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 $\leq$ defines only a partial order on $\mathbb{R}^n$ when $n \geq 2$: there exist vectors $x \in \mathbb{R}^n$ that satisfy neither $x \geq 0$ nor $x \leq 0$. 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 6, 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 6, 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*

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

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

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

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 diagonal 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, 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

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

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 $n_1 + \dots + n_\ell$ 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 $A|_{\operatorname{rg} P}$

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

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

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

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 is called **semisimple** if its geometric and algebraic multiplicities 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 $\lambda \in \sigma(A)$.*[1]

(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).*

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

(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.*

(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* $P = u v^{\mathrm{T}}$.

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) ⇔ (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) ⇔ (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) ⇔ (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) ⇒ (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) ⇒ (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, $P = \alpha u v^{\mathrm{T}}$ 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 $P = u v^{\mathrm{T}}$ 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 $\big(e^{t(A+B)}\big)_{t\geq 0}$ 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 $\big(e^{t(A+B)}\big)_{t\geq 0}$ 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 $\big(e^{t(A+B)}\big)_{t\geq 0}$ 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 $\big(e^{t(A+B)}\big)_{t\geq 0}$ is eventually positive.*

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

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

Proof. "(iii) ⇒ (i)": In this case, $\big(e^{t(A+B)}\big)_{t\geq 0}$ 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 $\mathrm{s}(A) = 0$ 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 $v \succeq \mathbb{1}$ and all $\mu > 0$, since strongly positive vectors are dense in $\mathbb{R}^n_+$. Thus let us fix such a vector $v \succeq \mathbb{1}$ 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 $\mathrm{s}(A) = 0$ 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

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

(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 [DGK16, Theorem 6.1]. The fact that eventual positivity implies positivity for 2×2 matrix semigroups (Exercise 2.5) was observed in [DGK16, 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 be a continuous map from $\operatorname{dom}(A)$ to X 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 both (x_k) and (Ax_k) are Cauchy in X, 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' converges uniformly to some $g \in \mathrm{C}([-1,1])$. For every $x \in [-1,1]$ it follows that

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

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, in particular, 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 $f_k(x) = \left(x^2 + \frac{1}{k}\right)^{1/2}$ converges uniformly to f, and thus, in particular, with respect to $\|\cdot\|_p$. Moreover,

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

for all $x \in [-1,1]$. Since (Af_k) converges pointwise almost everywhere to the signum function, the dominated convergence theorem yields 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** $A'\colon \operatorname{dom}(A') \subseteq Y' \to X'$ is defined by

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

where x' in the second line is the vector that occurs in the definition of $\operatorname{dom}(A')$.[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 $f : \Omega \to \mathbb{R}$ 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 integration 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 $\partial^\alpha : L^p(\Omega) \supseteq \operatorname{dom}(\partial^\alpha) \to L^p(\Omega)$ is closed when its domain is chosen as large as possible:

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

(in the first line of this formula we used the shortcut introduced in Definition 3.2.4(b)). In particular, $\operatorname{dom}(\partial^\alpha)$ 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

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

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 norm $\|\cdot\|_{V_1}, \dots, \|\cdot\|_{V_n}$, 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, $\mathrm{dist}(\mu, \sigma(A)) \geq \|\mathcal{R}(\mu, A)\|^{-1}$.

(b) Let $\lambda \in \rho(A)$ and $x \in \mathrm{dom}(A)$. Clearly, $(A - \lambda)\mathcal{R}(\lambda, A)x = x = \mathcal{R}(\lambda, A)(A - \lambda)x$ 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 $A : X \supseteq \mathrm{dom}(A) \to X$ be a closed operator. The **spectral bound** A is defined as

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

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 $A\colon \mathrm{C}([0,1]) \supseteq \mathrm{dom}(A) \to \mathrm{C}([0,1])$, $Af = f'$ 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 $A\colon \mathrm{C}_0((0,1]) \supseteq \mathrm{dom}(A) \to \mathrm{C}_0((0,1])$, $Af := f'$, where

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

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 $u \in L^1_{\text{loc}}(\Omega)$. We say that Δu **exists weakly** or, briefly and by a slight abuse of notation, that $\Delta u \in L^1_{\text{loc}}(\Omega)$, if there exists a (necessarily unique) function $g \in L^1_{\text{loc}}(\Omega)$ 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

$$\begin{aligned} \operatorname{dom}(\Delta_{\max}) &:= \{u \in L^2(\Omega) : \Delta u \in L^2(\Omega)\}, & \Delta_{\max} u &:= \Delta u, \\ \operatorname{dom}(\Delta_{\text{Dir}}) &:= \operatorname{dom}(\Delta_{\max}) \cap H_0^1(\Omega), & \Delta_{\text{Dir}} u &:= \Delta u. \end{aligned}$$

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 $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$ and $\Delta_{\mathrm{Dir}} u = g$ 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 $\rho(\Delta_{\mathrm{Dir}})$.

Note that the intersection with $H_0^1(\Omega)$ in the definition of $\operatorname{dom}(\Delta_{\mathrm{Dir}})$ 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, $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$ and $\Delta_{\mathrm{Dir}} u = g$ if and only if $u \in \operatorname{dom}(\Delta_{\max}) \cap H_0^1(\Omega)$ 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 $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$ and $\Delta_{\mathrm{Dir}} u = g$ 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, $\operatorname{dom}(\Delta_{\max})$ 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 $\operatorname{dom}(\Delta_{\mathrm{Dir}})$. 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 $u \in \operatorname{dom}(\Delta_{\mathrm{Dir}})$. 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 $\left(v \mid \lambda u - g\right)_{L^2} = -(\nabla v \mid \nabla u)_{L^2}$ for all $v \in H_0^1(\Omega)$, which is in turn equivalent to $u \in H_0^1(\Omega)$ and

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

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

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

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 $\left(\cdot \mid g\right)_{L^2}$ 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 $f \in L^1_{\text{loc}}(\mathbb{R}^n) \cap C^1(\mathbb{R}^n \setminus \{0\})$[10] and assume that there exist $g_1, \dots g_n \in L^1_{\text{loc}}(\mathbb{R}^n)$ 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 $\varphi \in C^\infty_c(\mathbb{R}^n)$, 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 $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|_2$. 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 $A : X \supseteq \operatorname{dom}(A) \to Y$ be a densely defined linear operator.

(a) Prove that the dual operator $A' : Y' \supseteq \operatorname{dom}(A') \to X'$ (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 $\operatorname{dom}\big((\lambda + A)'\big) = \operatorname{dom}\big(A'\big)$ and $\operatorname{dom}\big((\alpha A)'\big) = \operatorname{dom}\big(A'\big)$ and that

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

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 $\|x\|_X \le \|(\lambda - A)x\|_X \, \big\|\mathcal{R}(\lambda, A')\big\|$ 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 $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and f has a representative whose restriction to the set $\mathbb{R}^n \setminus \{0\}$ is C^1.

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.

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

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

with respect to the $W^{k,p}$ norm. It is straightforward to check that $H^{k,p}(\Omega) \subseteq W^{k,p}(\Omega)$, but this leaves the obvious question of whether the converse inclusion holds. It took some time before this problem was finally resolved in the affirmative in 1964 by Norman Meyers and James Serrin in their iconic paper [MS64]. We say a few more words about this 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 topic can be found, for example, in [Lue82, Chapter 2]. From the perspective of the general 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* $\operatorname{supp}(f * g) \subseteq \overline{\operatorname{supp} f + \operatorname{supp} g}$*, 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* $1 \leq p < \infty$.

(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* $1 \leq p < \infty$.

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* $f \in L^1_{\text{loc}}(\Omega)$ *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

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

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

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

for all sufficiently small $t > 0$ and all $x \in \Omega_k$, where we have used 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 $f_k = 0$ 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 $(u_m) \subset \mathrm{C}_c^\infty(\mathbb{R}^n)$.

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 $\zeta \in \mathrm{C}_c^\infty(\mathbb{R}^n)$ 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 $u_m \in \mathrm{C}_c^\infty(\mathbb{R}^n)$ 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 $C := \max_{|\beta| \le k} \|\partial^\beta \zeta\|_{L^\infty(\mathbb{R}^n)} < \infty$. 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

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

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

Likwise, 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 $\mathrm{C}^k(\mathbb{R}^{n-1};\mathbb{R})$.

An example of a region in $\mathbb{R}^2$ with Lipschitz boundary is illustrated in Figure 3.B.1. In practice, note that the open balls $B(\xi_0, r)$ may be replaced by other kinds of open sets, e.g. open cubes, according to what is most convenient.

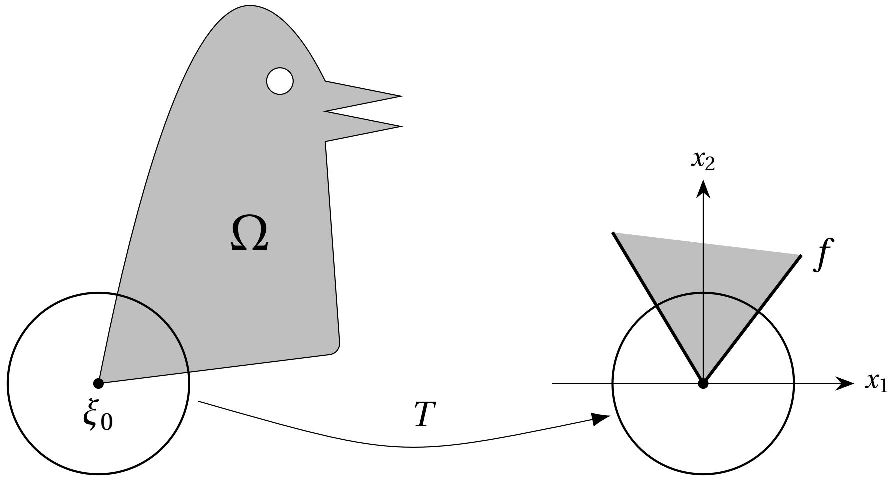

Figure 3.B.1: A region with Lipschitz boundary and local coordinates around ξ_0.

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*

$$\|\operatorname{Tr} u\|_{L^p(\partial\Omega)} \leq C\,\|u\|_{W^{1,p}(\Omega)}$$

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.

Bibliography

[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.

[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.

[CD13] Alexander P. Campbell and Daniel Daners. Linear algebra via complex analysis. *Amer. Math. Monthly*, 120(10):877–892, 2013.

[DG18] Daniel Daners and Jochen Glück. Towards a perturbation theory for eventually positive semigroups. *J. Operator Theory*, 79(2):345–372, 2018.

[DGK16] Daniel Daners, Jochen Glück, and James B. Kennedy. Eventually positive semigroups of linear operators. *J. Math. Anal. Appl.*, 433(2):1561–1593, 2016.

[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.

[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.

[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.

[Gra14] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Grad. Texts Math.* New York, NY: Springer, 3rd ed. edition, 2014.

[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.

[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.

[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.

[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.

[KT17] Michael Kasigwa and Michael J. Tsatsomeros. Eventual cone invariance. *Electron. J. Linear Algebra*, 32:204–216, 2017.

[Leo09] Giovanni Leoni. *A first course in Sobolev spaces*, volume 105 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2009.

[Lue82] Jesper Luetzen. *The prehistory of the theory of distributions*, volume 7 of *Stud. Hist. Math. Phys. Sci.* Springer-Verlag, New York, NY, 1982.

[Mac00] Charles R. MacCluer. The many proofs and applications of Perron's theorem. *SIAM Review*, 42(3):487–498, 2000.

[MS64] Norman G. Meyers and James Serrin. $H = W$. *Proc. Natl. Acad. Sci. USA*, 51:1055–1056, 1964.

[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.

[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.

[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.

[Sen06] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Ser. Stat. New York, NY: Springer, revised reprint of the 2nd ed. edition, 2006.

[Soo19] Aivar Sootla. Properties of eventually positive linear input-output systems. *IET Control Theory Appl.*, 13(7):891–897, 2019.

[Str03] Robert S. Strichartz. *A guide to distribution theory and Fourier transforms.* River Edge, NJ: World Scientific, 2003.

[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.

[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).