17th Internet Seminar on Evolution Equations 2013/14: Positive Operator Semigroups and Applications

András Bátkai, Marjeta Kramar Fijavž, Abdelaziz Rhandi

February 8, 2014

Lecture 1

An Invitation to Positive Matrices

1.1 Motivating examples

Let us start with three different examples where positive matrices make an appearance. We will see that various properties can be described by studying the behavior of such matrices.

The Fibonacci Sequence

Consider the following, well-known sequence: $f_0 = 0$, $f_1 = 1$,

$$f_{n+1} = f_n + f_{n-1}.$$

Introducing the new sequence $g_n := f_{n-1}$, we obtain the following system

$$\begin{aligned} f_{n+1} &= f_n + g_n, \\ g_{n+1} &= f_n, \end{aligned}$$

with $f_1 = 1$ and $g_1 = 0$. Hence, using vectorial notation and the well established connection between systems of linear equations and matrices, we see that the above system is equivalent to

$$\begin{pmatrix} f_{n+1} \\ g_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_n \\ g_n \end{pmatrix}.$$

Repeating the above argument, we see that

$$\begin{pmatrix} f_{n+1} \\ g_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_n \\ g_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 \begin{pmatrix} f_{n-1} \\ g_{n-1} \end{pmatrix} = \ldots = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} f_1 \\ g_1 \end{pmatrix}. \tag{1.1}$$

Hence, the properties of the numerical sequence (f_n) are strongly related to the properties of a matrix sequence (A^n) with $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Graphs and Markov chains

A *graph* $G = (V, E)$ is a mathematical structure that models relations between objects. Objects are represented by the set of *vertices* V while every *edge* $e \in E$ corresponds to a pairwise relation between two vertices: $e = \{u, v\}$ (if the relation is symmetric) or $e = (u, v)$, $u, v \in V$. In the first case we call the graph *undirected* and in the latter *directed*. The simplest way to represent a graph is by drawing a picture, see Figure 1.1.

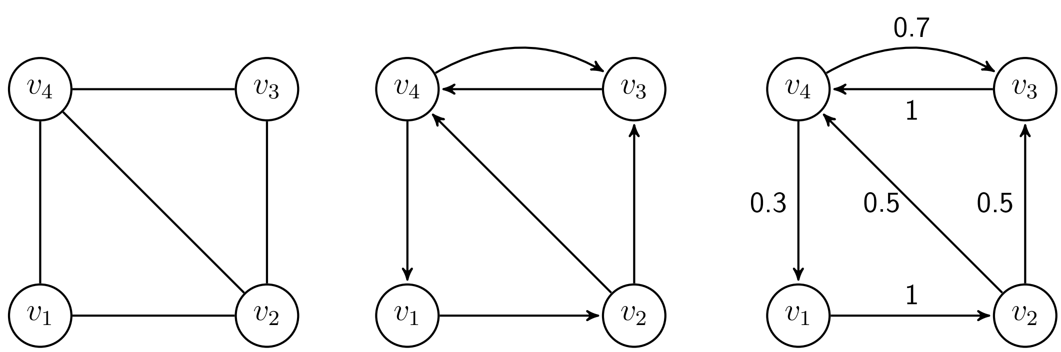
To study properties of the graph or even some dynamical process on it, it is however more useful to describe it in terms of matrices. Let us assume that the graph G is finite with the set of vertices $V = \{v_1, \dots, v_n\}$. The $n \times n$ *adjacency matrix* $A = (a_{ij})$ of G is defined as

$$a_{ij} > 0 \text{ if } (v_i, v_j) \in E \quad \text{and} \quad a_{ij} = 0 \text{ otherwise.}$$

If the nonzero elements of A all equal 1, we say that G is an *unweighted* graph, otherwise G is *weighted* with *weights* a_{ij} corresponding to the edges (v_i, v_j).

Example 1.1.1. The adjacency matrices of the three graphs, depicted in Figure 1.1 are:

$$\text{(a)} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad \text{(b)} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad \text{(c)} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 1 \\ 0.3 & 0 & 0.7 & 0 \end{pmatrix}$$



(a) Nondirected graph (b) Directed graph (c) Weighted directed graph

Figure 1.1: Examples of graphs

Note that the adjacency matrix of an undirected graph is always symmetric. An interesting property of the adjacency matrix of an unweighted graph is the following: computing its k-th power, $(A^k)_{ij}$ gives the number of (directed or undirected) walks of length k from vertex v_i to vertex v_j (see Exercise 2).

Consider now a discrete finite homogeneous *Markov chain* with state space V and *transition matrix* $P = (p_{ij})$. Its entries $p_{ij} \in [0, 1]$ represent *transition probabilities* to move from state v_i to v_j in one step. P is a positive and *row stochastic matrix*, i.e. $\sum_{j=1}^{n} p_{ij} = 1$ for all $i = 1, \dots, n$.

P is actually an adjacency matrix of a directed weighted graph G with the set of vertices V and edges given by the entries of P:

$$(v_i, v_j) \in E \iff p_{ij} > 0.$$

Note that G might have loops but does not have multiple directed edges. A Markov process is in fact a random walk on graph G.

Now $(P^k)_{ij}$ yields the probability of reaching state v_j from state v_i in k steps. Moreover, various long-term properties of such Markov processes, such as periodicity, ergodicity, the existance of a stationary distribution, etc., can be seen by observing the entries of P^k as $k \to \infty$.

The Competitive Markets Model

Suppose n similar commodities are competing for the consumer's money. Excess demand (i.e. demand minus supply) f_i for commodity i is approximately a linear function of prices p less equilibrium prices p^0, i.e.,

$$f_i \approx \sum_{j=1}^{n} a_{ij}(p_j - p_j^0).$$

Since higher prices for one commodity will increase excess demand for the others, $a_{ij} \geq 0$ for $i \neq j$ and $a_{ii} < 0$.

The rate of price readjustment, once disturbed from equilibrium, must be proportional to excess demand, i.e.,

$$\dot{p} = KA(p - p^0), \tag{1.2}$$

where $K = \operatorname{diag}(k_1, \dots, k_n)$, a diagonal matrix of positive *adjustment speeds*, $A = (a_{ij})$ and $p = (p_j)$. Thus future prices are given by

$$p(t) = p^0 + \mathrm{e}^{tKA}c, \tag{1.3}$$

where $c = p(0) - p^0$.

If we start with a (non-equilibrium) price $p(0)$, will prices eventually return to the equilibrium p^0 or will (some) prices stay away from p^0, oscillate or become unbounded? We will see that the answers to these questions depend on the spectral properties of the matrix KA.

One of the main topics of this course will be to analyze the behavior of matrix sequences of the form (A^n) as $n \to +\infty$ or the behavior of matrix exponentials of the form "e^{tA}" as $t \to +\infty$. Special attention will be paid to matrices with nonnegative entries, the so called *positive matrices*. We will apply the obtained results to various biological, economical or mathematical problems.

1.2 Convergence

Since the asymptotics of A^k and of e^{tA} is our ultimate goal, it is indispensable that we make precise what we understand by convergence in $X = \mathbb{C}^n$ or in $\mathcal{L}(X) = M_n(\mathbb{C})$.

The most natural concept of convergence in $\mathbb{C}^n$ is coordinatewise convergence, or convergence in every coordinate, a notion which explains itself. However, the question immediately arises if a change of coordinates touches this sort of convergence. Moreover, with respect to sequences of $n \times n$ matrices, a formally different kind of convergence seems natural, namely, pointwise convergence on $\mathbb{C}^n$. This means for a sequence $(A_k)_{k\in\mathbb{N}}$ in $M_n(\mathbb{C})$ the convergence of $(A_k x)_{k\in\mathbb{N}}$ for every $x \in \mathbb{C}^n$.

The following proposition tells us that these concepts all amount to the same thing, and that norms are an excellent instrument to deal with problems related to convergence. For basic information about norms, see Appendix A, Section A.1.

Proposition 1.2.1. *(a)*

$$M_n(\mathbb{C}) \ni A = (\alpha_{ij}) \mapsto \begin{cases} \max_{i,j} |\alpha_{ij}| =: \|A\|_\infty \\ \max_j \sum_{i=1}^n |\alpha_{ij}| =: \|A\|_1 \end{cases}$$

are norms on $M_n(\mathbb{C})$ *with*

$$\|A\|_\infty \le \|A\|_1 \le n\|A\|_\infty$$

for all $A \in M_n(\mathbb{C})$.

(b) Let $A_k := \left(\alpha_{ij}^{(k)}\right)$, $k \in \mathbb{N}$ *be a sequence in* $M_n(\mathbb{C})$, $A_0 = \left(\alpha_{ij}^{(0)}\right) \in M_n(\mathbb{C})$. *Then the following statements are equivalent:*

(i) $A_k x \to A_0 x$ *coordinate-wise, as* $k \to \infty$, *for every* $x \in X = \mathbb{C}^n$.

(ii) $\alpha_{ij}^{(k)} \to \alpha_{ij}^{(0)}$ *as* $k \to \infty$ *for all* $1 \le i, j \le n$.

(iii) $\|A_k - A_0\|_\infty \to 0$ *as* $k \to \infty$.

(iv) $\|A_k - A_0\|_1 \to 0$ *as* $k \to \infty$.

Proof. See Exercise 5 □

There are many possibilities other than those considered above to define norms, and correspondingly, there exists a large variety of potentially different notions of convergence. Proposition 1.2.1, however, indicates that the latter might always coincide. This is actually true (at least in finite dimensional vector spaces) by the following result, due to A. Tikhonov, which will allow us to define convergence without specific reference to coordinates (or entries). For a proof of this result and further discussions, see Theorem A.1.5 in Appendix A.

Theorem 1.2.2 (Tikhonov). *Let X be a finite dimensional (complex) vector space, and let $\|\cdot\|$ and $|||\cdot|||$ be norms on X. Then there exist constants $m, M > 0$ such that*

$$m\,|||x||| \leq \|x\| \leq M\,|||x||| \tag{1.4}$$

for all $x \in X$.

We call two norms $\|\cdot\|$ and $|||\cdot|||$ *equivalent* if (1.4) holds for suitable constants $m, M > 0$. Equivalent norms possess the same convergent sequences and the same Cauchy sequences. Therefore, in finite dimensions, Tikhonov's Theorem 1.2.2 allows us now to define convergence as follows.

Definition 1.2.3. Let X be a vector space of finite dimension and $(x_k)_{k\in\mathbb{N}}$ a sequence in X.

(a) We say that (x_k) is a *Cauchy sequence*, if (x_k) is a Cauchy sequence for one, and hence for every norm on X. Recall that a sequence is called a Cauchy sequence for a norm $\|\cdot\|$, if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for every $k, m > N$, $\|x_k - x_m\| < \varepsilon$.

(b) We say that (x_k) *converges*, as $k \to \infty$, to an element $x_0 \in X$, if $\|x_k - x_0\| \to 0$ as $k \to \infty$ for one, and hence for every norm on X.

We will now give a more specific description of convergence which will be the clue to our subsequent discussions. To this end, let (x_k) be a sequence in X. If $\{z_1, z_2, \ldots, z_n\}$ is a basis for X, there exist uniquely determined scalars $\xi_i^{(k)}$, $1 \leq i \leq n$, such that

$$x_k = \sum_{i=1}^{n} \xi_i^{(k)} z_i.$$

We call the sequences $\left(\xi_1^{(k)}\right)_{k\in\mathbb{N}}, \ldots, \left(\xi_n^{(k)}\right)_{k\in\mathbb{N}}$, the *coordinate sequences* of $(x_k)_{k\in\mathbb{N}}$ with respect to the basis $\{z_1, \ldots, z_n\}$.

Theorem 1.2.4. *For a sequence (x_k) in X, the following are equivalent.*

(a) (x_k) is a Cauchy sequence.

(b) (x_k) converges, as $k \to \infty$, to an element $x_0 \in X$.

(c) For every basis $\{z_1, z_2, \ldots, z_n\}$ of X, the coordinate sequences of (x_k) with respect to $\{z_1, z_2, \ldots, z_n\}$ converge, as $k \to \infty$.

In this case, the element x_0 in (b) is uniquely determined and the coordinates of x_0 with respect to $\{z_1, z_2, \ldots, z_n\}$ are the limits of the respective coordinate sequences.

Proof. The statement is an immediate consequence of the fact that for any basis $\{z_1, z_2, \ldots, z_n\}$ of X,

$$X \ni x \mapsto \max\left\{|\xi_i| \,:\, x = \sum \xi_i z_i\right\}$$

is a norm. □

We introduce yet another notion related to norms. Again, X is a vector space of finite dimension.

Definition 1.2.5. A subset $\Omega \subset X$ is called *bounded*, if for one, hence for every norm $\|\cdot\|$ on X,

$$\sup_{x\in\Omega} \|x\| < \infty.$$

It is clear that boundedness of a sequence (x_n) in X means boundedness of the set $\{x_n : n \in \mathbb{N}\}$, and that this is equivalent to boundedness of all coordinate sequences of (x_n) with respect to one, hence all, basis of X.

Although there appears to be no need to discriminate between different norms on a finite dimensional vector space, the equivalence of all norms on such a space gives, on the other hand, the possibility to discuss convergence of a specific sequence by using a particularly favorable norm. This was the key point in the proof of the previous theorem. Apart from that, we will occasionally have to distinguish, even on the finite dimensional vector space $\mathcal{L}(X)$, between norms in general and norms that come from the underlying vector space X.

Definition 1.2.6. Let $A \in \mathcal{L}(X)$ and let $\|\cdot\|$ be a norm on X. The corresponding *operator norm* on $\mathcal{L}(X)$, again denoted by $\|\cdot\|$, is defined as

$$\|A\| := \sup\{\|Ax\|, x \in X, \|x\| \leq 1\} = \max\{\|Ax\|, x \in X, \|x\| \leq 1\}.$$

Alternative descriptions of the operator norm coming from a norm $\|\cdot\|$ on X are the following:

$$\begin{aligned} \|A\| &= \sup\{\|Ax\| : \|x\| = 1\} \\ &= \sup\left\{\frac{\|Ax\|}{\|x\|} : x \neq 0\right\} \\ &= \min\{M : \|Ax\| \leq M\|x\| \text{ for all } x \in X\}. \end{aligned}$$

The relation

$$\|Ax\| \leq \|A\|\|x\|, \qquad x \in X,$$

is a reformulation of the last two expressions above. In particular, for $A, B \in \mathcal{L}(X)$,

$$\|ABx\| = \|A(Bx)\| \leq \|A\|\ \|Bx\| \leq \|A\|\ \|B\|\ \|x\|$$

for all $x \in X$. Hence, $\|AB\| \leq \|A\|\ \|B\|$ whenever $\|\cdot\|$ is an operator norm. We note in particular that for any operator norm $\|\cdot\|$ on $\mathcal{L}(X)$, the identity must have norm 1. Thus, if $\|\cdot\|$ is an operator norm, the norm

$$A \mapsto \mu\|A\|, \qquad \mu > 0,$$

cannot be an operator norm unless $\mu = 1$.

Another interesting feature of operator norms is that, for such a norm $\|\cdot\|$, necessarily $|\lambda| \leq \|A\|$ for every eigenvalue λ. This is obvious since for any eigenvector x belonging to λ we have $\|Ax\| = |\lambda|\ \|x\|$. For later reference, we note this as a corollary.

Corollary 1.2.7. *Let $\|\cdot\|$ be an operator norm on $\mathcal{L}(X)$, $A \in \mathcal{L}(X)$ and λ an eigenvalue of A. Then $\|A\| \geq |\lambda|$.*

Generally, the operator norm is difficult to calculate explicitly. We give two important examples where it can be calculated. Recall that for $1 \leq p < \infty$ and $x \in X = \mathbb{C}^n$, $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$, is called the *p-norm* on X, the ∞-*norm* is defined as $\|x\|_\infty := \max_{1\leq i\leq n} |x_i|$.

Example 1.2.8. (i) The operator norm of a matrix $A = (a_{ij}) \in M_n(\mathbb{C})$ corresponding to the ∞– norm on $\mathbb{C}^n$ is the *row norm*

$$\|A\|_\infty := \max\{\|Ax\|_\infty : x \in \mathbb{C}^n, \|x\|_\infty \leq 1\} = \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|.$$

(ii) The operator norm on $M_n(\mathbb{C})$ corresponding to the 1– norm on $\mathbb{C}^n$ is the *column norm*

$$\|A\|_1 := \max\{\|Ax\|_1 : x \in \mathbb{C}^n, \|x\|_1 \leq 1\} = \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|.$$

Proposition 1.2.9. *For a sequence $(A_k)_{k\in\mathbb{N}} \subseteq \mathcal{L}(X)$ the following assertions are equivalent.*

(a) The sequence $(A_k x)_{k\in\mathbb{N}}$ converges in X for all $x \in X$.

(b) The sequence $(A_k)_{k\in\mathbb{N}}$ converges in $\mathcal{L}(X)$.

(c) For any basis of X and the corresponding matrix representation $A_k = (a_{ij}^{(k)})$ the entries $a_{ij}^{(k)}$ converge in $\mathbb{C}$ for every pair (i, j) as $k \to \infty$.

The proof of this proposition is left as an exercise.

1.3 Notes and Remarks

For more notions and results form graph theory we refer to [1], [2] or [4]. The model presented in Subsection 1.1 is taken from [3, p.492]. We recall some necessary results from linear algebra in Appendix A.

1.4 Exercises

1. Determine the eigenvalues and eigenvectors of the matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, corresponding to the Fibonacci sequence (f_n) (see Section 1.1). Using this find a formula for A^n and f_n.

2. An undirected graph G is depicted in the Figure 1.2.

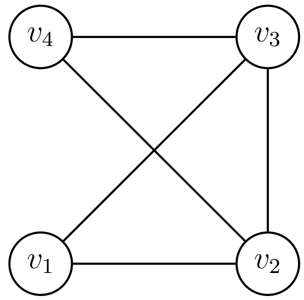


Figure 1.2: The graph G from Exercise 2.

(a) Write the adjacency matrix A of the graph G and compute A^2 and A^3.

(b) Show that $(A^k)_{ij}$ gives the number of walks of length k from vertex v_i to vertex v_j.

(c) Choose directions of the edges of G to obtain a directed graph $\widetilde{G}$. Repeat (a) and (b) for $\widetilde{G}$ instead of G.

(d) * Show that (b) holds for any unweighted graph.

3. Let A be the adjacency matrix of a simple undirected graph G (by simple we mean unweighted, without loops or multiple edges). Show the following spectral properties of A.

   (a) The sum of all eigenvalues of A equals 0.

   (b) The sum of all squares of eigenvalues of A equals $2m$ where m is the number of edges in G.

   (c) The sum of all cubes of eigenvalues of A equals $6t$ where t is the number of triangles in G.

   (Hint: use the trace and Exercise 2(d).)

4. (a) Show that the formulae in Example A.1.2 of Appendix A define norms on $\mathbb{C}^n$.

   (b) Verify the following relations among the norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$:

$$\begin{aligned}
&\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2,\\
&\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty,\\
&\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty.
\end{aligned}$$

   (c) Show that $\lim_{p\to\infty}\|x\|_p = \|x\|_\infty$.

5. Prove the statements of Proposition 1.2.1.

6. Verify the equivalence of the alternative descriptions of the operator norm in Definition 1.2.6 and prove the statements of Example 1.2.8.

7. Prove Proposition 1.2.9.

8. Solve the exercises in Appendix A.

Bibliography

[1] A. Bondy and U.S.R. Murty, *Graph Theory*, Graduate Texts in Math., vol. 244, Springer-Verlag, 2008.

[2] Ch. D. Godsil and G. Royle, *Algebraic Graph Theory*, Graduate Texts in Math., vol. 207, Springer-Verlag, 2008.

[3] C.R. MacCluer, *The Many Proofs and Applications of Perron's Theorem*, SIAM Rev. **42** (2000), 487–498.

[4] D.B. West, *Introduction to Graph Theory*, Prentice-Hall, 2001.

Lecture 2

Functional Calculus

2.1 Polynomials

In this section we propose a method of constructing functions of a linear operator $A \in \mathcal{L}(X)$, where X is a finite dimensional vector space. Of special interest are, in view of reducing subspaces (see Appendix A, Section A.1.2), projections commuting with A. We will look for such projections among the *polynomials* in A, i.e., the linear combinations of powers of A. Indeed, if such a projection is a polynomial, it automatically commutes with A and hence its range is a reducing subspace.
We denote the set of all *polynomials* in A by $\mathcal{P}_A$, i.e.

$$\mathcal{P}_A := \left\{ \sum_{i=0}^{m} \alpha_i A^i : \alpha_i \in \mathbb{C}, m \in \mathbb{N} \right\} \subset \mathcal{L}(X), \tag{2.1}$$

and if $p(x) = \sum_{i=0}^{m} \alpha_i x^i$ is a polynomial, we write

$$p(A) := \sum_{i=0}^{m} \alpha_i A^i$$

and say that the operator $p(A)$ is obtained by *plugging A into p*. From this definition it is clear that the mapping

$$\Phi_A : \mathbb{C}[x] \ni p \mapsto p(A) \in \mathcal{P}_A$$

is a homomorphism from the algebra $\mathbb{C}[x]$ of all polynomials onto $\mathcal{P}_A$. For those readers who are acquainted with the notion of the factor map, we remark that

$$\hat{\Phi}_A : {}^{\mathbb{C}[x]}\big/_{\ker \Phi_A} \to \mathcal{P}_A \tag{2.2}$$

becomes an algebraic isomorphism.

Exercise 2.1.1. Let $A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Calculate $\dim \mathcal{P}_A$.

In order to find polynomials such that $p(A) = (p(A))^2 = p^2(A)$, i.e., $p(A)$ is a projection, we have a closer look at the algebraic structure of $\ker \Phi_A$.

Proposition 2.1.2. *There exists a unique polynomial $m_A \in \mathbb{C}[x]$ of degree ≥ 1 and with leading coefficient 1 such that the following holds:*

A polynomial $p \in \mathbb{C}[x]$ belongs to $\ker \Phi_A$, i.e., $p(A) = 0$, if and only if $p = m_A \cdot q$ for some $q \in \mathbb{C}[x]$.

Proof. Take a nonzero polynomial $m \in \ker \Phi_A$ with minimal degree and leading coefficient 1. If $p = m \cdot q$ for some $q \in \mathbb{C}[x]$, then $p(A) = m(A) \cdot q(A) = 0$, hence $p \in \ker \Phi_A$.

On the other hand, if $p \in \mathbb{C}[x]$ is any polynomial with $p(A) = 0$, then the division of p by m gives

$$p = m \cdot q + r$$

for polynomials q and r satisfying $\deg r < \deg m$.

Thus

$$0 = p(A) = m(A) \cdot q(A) + r(A) = r(A)$$

and $r \in \ker \Phi_A$. Since m has minimal degree in $\ker \Phi_A$, we conclude that $r = 0$. This shows that $m_A := m$ is the desired polynomial.
Uniqueness follows since the leading coefficient was supposed to be 1. □

We call this polynomial m_A generating the kernel of Φ_A the *minimal polynomial* of A. For polynomials $p, q, r \in \mathbb{C}[X]$ we use the notation

$$p \equiv q \mod r \iff p - r = s \cdot r \text{ for some } s \in \mathbb{C}[x].$$

The following is an immediate consequence of Proposition 2.1.2.

Corollary 2.1.3. *For $p, q \in \mathbb{C}[x]$, we have $p(A) = q(A)$ if and only if*

$$p \equiv q \mod m_A.$$

In particular, $p(A)$ is a projection if and only if $p^2 \equiv p \mod m_A$.

Denoting by

$$\lambda_1, \dots, \lambda_m$$

the zeros of the minimal polynomial m_A, and by

$$\nu_1, \dots, \nu_m$$

their respective multiplicities (as zeros of m_A), we see that

$$m_A(z) = (z - \lambda_1)^{\nu_1}(z - \lambda_2)^{\nu_2} \cdots (z - \lambda_m)^{\nu_m}. \tag{2.3}$$

Lemma 2.1.4. *For $p, q \in \mathbb{C}[x]$, we have*

$$p \equiv q \mod m_A \iff \text{The derivatives of } p \text{ and } q \text{ satisfy } p^{(\nu)}(\lambda_i) = q^{(\nu)}(\lambda_i) \text{ for } i = 1, \dots, m \text{ and } \nu = 0, \dots, \nu_i - 1. \tag{2.4}$$

Proof. Follows directly from Lemma A.2.2 in the Appendix A. □

We use this characterization of the equality $p(A) = q(A)$ to extend the domain of Φ_A and hence to be able to define more general functions of A. Before proceeding, let us make the following observation.

Remark 2.1.5. For the exponential function

$$f(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

it seems to be natural to define the exponential function of a matrix A with the formula

$$e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

This is a completely legitimate way and we will also justify this formula later on (see Corollary 2.2.7). However, since $\mathcal{L}(X)$ is an n^2-dimensional vector space (where $n = \dim X$), in the above infinite series there are at most n^2 linearly independent terms. Hence, e^A is actually a polynomial in A, of course with rather complicated coefficients. This observation justifies our procedure to look for the operator $f(A)$ among the polynomials of A.

Before going on with the abstract argument let us illustrate our last point on two examples of simple 2×2 matrices where we are able to calculate the power series.

Examples 2.1.6. (i) Let $A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then, using the above formula and the fact that $A^2 = 0$, we see that

$$e^{tA} = I + tA = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

This is clearly a linear polynomial of A.

(ii) Let $A := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Then, using the above formula and the fact that

$$(tA)^n = \begin{pmatrix} t^n & 0 \\ 0 & (2t)^n \end{pmatrix},$$

we see that

$$e^{tA} = \begin{pmatrix} e^t & 0 \\ 0 & e^{2t} \end{pmatrix} = e^t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + e^{2t} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = e^t(2I - A) + e^{2t}(A - I).$$

This is again a linear polynomial of A.

Exercise 2.1.7. Calculate e^{tA_i} for

$$A_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

2.2 Smooth functions

Let $A \in \mathcal{L}(X)$ be an operator on the finite dimensional vector space X with minimal polynomial m_A. As before, $\lambda_1, \ldots, \lambda_m$ are the roots of the minimal polynomial with corresponding multiplicities $\nu_1, \ldots, \nu_m$. We denote the set of functions, that are defined on an open neighborhood containing all λ_i and are infinitely differentiable, with

(2.5)
$$C_A^\infty := \{f : D(f) \subset \mathbb{C} \to \mathbb{C} \,:\, \exists U \subset \mathbb{C} \text{ open}, U \subset D(f), \{\lambda_1, \ldots, \lambda_m\} \subset U, f|_U \in C^\infty\}.$$

Here $D(f)$ denotes the domain of f. C_A^∞ is the set of functions for which we would like to define the functional calculus.

Definition 2.2.1 (Functional Calculus)**.** Let $f \in C_A^\infty$. We then set

$$f(A) := \Phi_A(p_f) = p_f(A),$$

where p_f is an *interpolation polynomial*[1] for f in the sense that the derivatives of f satisfy

$$f^{(\nu)}(\lambda_i) = p_f^{(\nu)}(\lambda_i)$$

for $i = 1, \ldots, m$ and $\nu = 0, \ldots, \nu_i - 1$.

Again, we say that $f(A)$ is the result of *plugging A into the function f*. Hence, we extended our function Φ_A, defined on $\mathbb{C}[x]$, to the function $\tilde{\Phi}_A$ on C_A^∞ with the formula

$$\tilde{\Phi}_A(f) := \Phi_A(p_f) = p_f(A).$$

We now collect useful properties of $\tilde{\Phi}_A$ that follow directly from Definition 2.2.1, Corollary 2.1.3, and Lemma 2.1.4.

Lemma 2.2.2. *With the above notation the following holds.*

(i) The definition of $\tilde{\Phi}_A(f)$ does not depend on the particular choice of the interpolation polynomial p_f.

(ii) The map $\tilde{\Phi}_A$ is an extension of Φ_A.

(iii) The map $\tilde{\Phi}_A$ is an algebra homomorphism in the sense that

$$\tilde{\Phi}_A(\lambda f + \mu g) = \lambda \tilde{\Phi}_A(f) + \mu \tilde{\Phi}_A(g),$$
$$\tilde{\Phi}_A(f \cdot g) = \tilde{\Phi}_A(f) \cdot \tilde{\Phi}_A(g)$$

for $\lambda, \mu \in \mathbb{C}$ and functions $f, g \in C_A^\infty$.

[1] For more information on interpolation polynomials see Appendix A, Section A.2

At first glance not much seems to be gained since the range of $\tilde{\Phi}_A$ is still $\mathcal{P}_A$. However, the domain of $\tilde{\Phi}_A$ is now much larger and contains many more functions, in particular easily identifiable, non-trivial idempotents. In fact, any characteristic function $\mathbb{1}_U$ of an open set $U \subset \mathbb{C}$ such that none of the points $\lambda_1, \dots, \lambda_m$ lies on the boundary of U is an idempotent and belongs to the domain of $\tilde{\Phi}_A$. As a consequence of Corollary 2.1.3, $\mathbb{1}_U(A)$ must be a projection contained in $\mathcal{P}_A$, hence commuting with A.

We now pick a particular set of such projections. Let $U_1, \dots, U_m$ be open subsets of $\mathbb{C}$ satisfying

(i) $\lambda_i \in U_i$ for $i = 1, \dots, m$, and

(ii) $U_i \cap U_j = \emptyset$ for $i \neq j$.

Writing $\mathbb{1}_i$ for the characteristic function of U_i, we obtain projections

$$P_i := \mathbb{1}_i(A) \in \mathcal{P}_A \text{ for } i = 1, \dots, m, \tag{2.6}$$

whose range will be denoted by

$$X_i := \operatorname{im} P_i. \tag{2.7}$$

We remark that P_i is independent of the specific choice of U_i, and $P_i \neq 0$ for all i.

The following is now the fundamental structure theorem for linear operators on finite dimensional vector spaces.

Theorem 2.2.3. *Let X be a finite dimensional vector space and let $A \in \mathcal{L}(X)$ with minimal polynomial m_A having zeros $\lambda_1, \dots, \lambda_m$ with respective multiplicities $\nu_1, \dots, \nu_m$. If we take P_i and X_i as in* (2.6) *and* (2.7)*, respectively, then*

$$X = X_1 \oplus \cdots \oplus X_m$$

is a direct sum decomposition into A-invariant subspaces such that the restriction of $\lambda_i - A$ to X_i is nilpotent of order ν_i for $i = 1, \dots, m$.

Proof. Since $\mathbb{1}_i(\sum_{j\neq i} \mathbb{1}_j) = 0$ for any i, $P_i(\sum_{j\neq i} P_j) = 0$ and hence $X_i \cap [\bigoplus_{j\neq i} X_j] = \{0\}$. Moreover, since $\{\lambda_1, \dots, \lambda_m\} \subset \bigcup_{i=1}^m U_i =: U$, we conclude that

$$\sum_{i=1}^m P_i = \tilde{\Phi}_A\left(\sum_{i=1}^m \mathbb{1}_i\right) = \tilde{\Phi}_A(\mathbb{1}_U) = I,$$

hence $X_1 \oplus \cdots \oplus X_m = X$.

We recall that a matrix B is called *nilpotent of order k*, if $B^k = 0$ but $B^{k-1} \neq 0$. Observe now that, for any fixed i, $g_i(\lambda) = (\lambda_i - \lambda)^{\nu_i}\mathbb{1}_i(\lambda)$ is a function in the domain of $\tilde{\Phi}_A$ that coincides in each of the points $\lambda_1, ..., \lambda_m$ with the zero function $\mathbb{0}$, including all relevant derivatives. By the properties of $\tilde{\Phi}_A$ stated in Lemma 2.2.2 we must have

$$(\lambda_i - A)^{\nu_i} P_i = \big((\lambda_i - \lambda)^{\nu_i}\mathbb{1}_i\big)(A) = \mathbb{0}(A) = 0$$

for $i = 1, \dots, m$ and $\lambda \in \mathbb{C}$. On the other hand, the function

$$f_i(\lambda) := (\lambda_i - \lambda)^{\nu_i - 1} \mathbb{1}_i(\lambda) \ , \quad \lambda \in \mathbb{C},$$

does *not* satisfy

$$f_i^{(\nu_i - 1)}(\lambda_i) = 0,$$

hence

$$(\lambda_i - A)^{\nu_i - 1} P_i = f_i(A) \neq 0,$$

proving that $(\lambda_i - A)$ is nilpotent of order ν_i on X_i. □

The zeros $\lambda_1, \dots, \lambda_m$ of the minimal polynomial m_A can now be identified with the *eigenvalues* of A, i.e., with those $\lambda \in \mathbb{C}$ for which $(\lambda - A)x = 0$ for some $0 \neq x \in X$.

Corollary 2.2.4. *Under the above assumptions the points $\lambda_1, \dots, \lambda_m$ are exactly those $\lambda \in \mathbb{C}$ for which $\lambda - A$ is not invertible.*

Proof. Every λ_i is an eigenvalue of A with an eigenvector contained in X_i. Otherwise, $\lambda_i - A$ would be bijective (see Proposition A.1.14), hence could not be nilpotent on X_i (which is $\neq \{0\}$ since $P_i \neq 0$).

There are no other eigenvalues of A since for any $\mu \in \mathbb{C}$ distinct from all $\lambda_1, \dots, \lambda_m$, the function $f(\lambda) := \frac{1}{\mu - \lambda}$ is the inverse of $f^{-1}(\lambda) := \mu - \lambda$, thus the operator $f(A)$ is the inverse of $f^{-1}(A) = \mu - A$, and hence $\mu - A$ is invertible. □

After these preparations we are now able to express the operator $f(A)$, f in the domain of $\tilde{\Phi}_A$, by the action of $(A - \lambda_i)$ on the subspaces X_i and by the values of (the derivatives of) f in λ_i.

Theorem 2.2.5. *Let $A \in \mathcal{L}(X)$ with eigenvalues $\lambda_1, \dots, \lambda_m$ and respective multiplicities $\nu_1, \dots, \nu_m$, and define the projections P_i as in* (2.6). *For every function $f \in C_A^\infty$ one has*

$$f(A) = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \frac{f^{(\nu)}(\lambda_i)}{\nu!} (A - \lambda_i)^\nu P_i. \tag{2.8}$$

In particular, the following holds.

(i) $$R(\mu, A) : = (\mu - A)^{-1} = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \frac{(A - \lambda_i)^\nu}{(\mu - \lambda_i)^{\nu+1}} P_i \quad \textit{for } \mu \notin \{\lambda_1, \dots, \lambda_m\}.$$

(ii) $$e^{tA} : = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \frac{e^{t\lambda_i} t^\nu}{\nu!} (A - \lambda_i)^\nu P_i \qquad \textit{for } t \in \mathbb{R}.$$

(iii) $$A^k : = \sum_{i=1}^{m} \sum_{\nu=0}^{\min\{\nu_i - 1, k\}} \binom{k}{\nu} \lambda_i^{k-\nu} (A - \lambda_i)^\nu P_i \quad \textit{for } k \in \mathbb{N}.$$

Proof. The function

$$g(\lambda) := \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i-1} \frac{f^{(\nu)}(\lambda_i)}{\nu!}(\lambda-\lambda_i)^\nu \mathbb{1}_i(\lambda), \quad \lambda \in \mathbb{C},$$

coincides with f, including all relevant derivatives, on all of the points $\lambda_1, \ldots, \lambda_m$. By Lemma 2.2.2 we therefore obtain $f(A) = g(A)$.
The special cases (i),(ii) and (iii) follow by taking $f(\lambda)$ as $(\mu-\lambda)^{-1}$, $\mathrm{e}^{t\lambda}$, and λ^k, respectively. □

The following result tells us that our functional calculus is in accordance with questions of convergence. It will lead us to alternatives to formulas (ii) and (iii) in Theorem 2.2.5 involving infinite series.

Proposition 2.2.6. *Let $A \in \mathcal{L}(X)$ and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions in C_A^∞ that converges pointwise, including the relevant derivatives, at each of the eigenvalues $\lambda_1, \ldots, \lambda_m$, to a function $f \in C_A^\infty$. Then the operators $f_n(A)$ converge to $f(A)$ in $\mathcal{L}(X)$.*

Proof. By hypothesis, we have

$$f_n^{(\nu)}(\lambda_i) \to f^{(\nu)}(\lambda_i)$$

for all $1 \le i \le m$ and $0 \le \nu \le \nu_i - 1$ as $n \to \infty$.
Let $\|\cdot\|$ be a norm on $\mathcal{L}(X)$. Then, by (2.8),

$$\begin{aligned}\|f_n(A) - f(A)\| &= \left\| \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i-1} \left(f_n^{(\nu)}(\lambda_i) - f^{(\nu)}(\lambda_i)\right) \frac{(A-\lambda_i)^\nu}{\nu!} P_i \right\| \\ &\le \left(\sup_{\substack{0\le\nu\le\nu_i-1 \\ 1\le i\le m}} \left| f_n^{(\nu)}(\lambda_i) - f^{(\nu)}(\lambda_i)\right| \right) \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i-1} \left\| \frac{(A-\lambda_i)^\nu}{\nu!} P_i \right\|.\end{aligned}$$

Since $(f_n)_{n\in\mathbb{N}}$ converges, including the relevant derivatives, at each of the points λ_i to a function f, the right hand side tends to 0 as $n \to \infty$. □

With the help of this proposition, we can see that our construction of the functional calculus yields the same as the power series for the exponential function.

Corollary 2.2.7. *Let $A \in \mathcal{L}(X)$ and $t \in \mathbb{R}$. Then*

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{t^k A^k}{k!}.$$

Proof. Put $f_n(\lambda) := \sum_{k=0}^{n} \frac{t^k \lambda^k}{k!}$, $\lambda \in \mathbb{C}$. Since f_n converges, as $n \to \infty$, in all derivatives to $e^{\lambda t} =: f(\lambda)$, the assertion follows from Proposition 2.2.6. □

For later reference, let us introduce the notation $r(A) := \max_{1\le i\le m} |\lambda_i|$ and call this the *spectral radius* of the operator A.

Corollary 2.2.8. *Let $A \in \mathcal{L}(X)$ and $|\lambda| > \max_{1\le i\le m} |\lambda_i| = r(A)$. Then*

$$R(\lambda, A) = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}. \tag{2.9}$$

Proof. Under the hypothesis on λ, we have

$$\frac{1}{\lambda - z} = \frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{z}{\lambda}\right)^k = \sum_{k=0}^{\infty}\frac{z^k}{\lambda^{k+1}}$$

for $|z| < |\lambda|$. Let $f_n(\lambda) := \sum_{k=0}^{n} \frac{z^k}{\lambda^{k+1}}$ and note that $f_n \to 1/(\lambda - z)$ pointwise as $n \to \infty$ in all derivatives, hence the result follows again by Proposition 2.2.6. □

The above expression (2.9) is called the *Neumann series* for $R(\lambda, A)$.

Later on, we will frequently use the following observations which help to read the formulas of Theorem 2.2.5.

Lemma 2.2.9. *(i) Take $i \in \{1, ..., m\}$ and $0 \neq z \in X_i$. Then the set*

$$\left\{(A - \lambda_i)^\nu z : \nu = 0, \dots, \nu_i - 1\right\} \setminus \{0\}$$

is linearly independent in X_i.

(ii) The set

$$B_A := \left\{(A - \lambda_i)^\nu P_i : i = 1, \dots, m;\ \nu = 0, \dots, \nu_i - 1\right\}$$

is linearly independent in $\mathcal{L}(X)$.

Proof. (i) Since $A - \lambda_i$ is nilpotent of order ν_i in X_i, there exists, for every $0 \neq z \in X_i$, an exponent $\nu_0 \le \nu_i$ with $(A - \lambda_i)^{\nu_0 - 1} z \neq 0$ and $(A - \lambda_i)^\nu z = 0$ for $\nu \ge \nu_0$. An equation of the form

$$0 = \sum_{\nu=0}^{\nu_0 - 1} \alpha_\nu (A - \lambda_i)^\nu z$$

leads, after multiplication by $(A - \lambda_i)^{\nu_0 - 1}$, to $\alpha_0 = 0$. Subsequent multiplication with decreasing powers of $(A - \lambda_i)$ shows that all the other coefficients α_ν must also be zero.

(ii) As in the proof of (i), one shows that the set

$$B_A^{(i)} := \left\{(A - \lambda_i)^\nu P_i : \nu = 0, \dots, \nu_i - 1\right\}$$

is linearly independent in $\mathcal{L}(X)$ for each $1 \le i \le m$. If

$$\sum_{i=1}^{m}\sum_{\nu=0}^{\nu_i - 1} \alpha_\nu^{(i)} (A - \lambda_i)^\nu P_i = 0,$$

then for fixed $1 \le i \le m$ multiplication by P_i shows that the partial sum

$$\sum_{\nu=0}^{\nu_i - 1} \alpha_\nu^{(i)} (A - \lambda_i)^\nu P_i = 0$$

must vanish, thus $\alpha_\nu^{(i)} = 0$ for $\nu = 0, \dots, \nu_i - 1$ by the linear independence of $B_A^{(i)}$. □

2.3 Spectral theory

We now introduce the usual terminology in connection with the symbols appearing in Theorem 2.2.5. Let λ_i be an eigenvalue of the matrix A. We call $\ker(\lambda_i - A)$ the *eigenspace*, each $0 \neq x \in \ker(\lambda_i - A)$ an *eigenvector*, the projection P_i defined in (2.6) the *spectral projection*, and $X_i := P_i X$ the *spectral subspace* corresponding to λ_i.

The set $\{\lambda_1, \dots, \lambda_m\}$ of all eigenvalues of A is denoted by $\sigma(A)$ and called the *spectrum* of A. Its complement is the *resolvent set* $\rho(A) := \mathbb{C}\backslash\sigma(A)$ and the *resolvent* of A is the map

$$\rho(A) \ni \lambda \mapsto R(\lambda, A) := (\lambda - A)^{-1}.$$

Remark 2.3.1. For those readers who are familiar with elementary complex function theory, we mention that from the formula in Theorem 2.2.5(i) we see that the resolvent is a meromorphic function with poles of order ν_i in λ_i for $i = 1, \dots, m$. Since the principal part of the Laurent expansion of $R(\lambda, A)$ around λ_i is

$$\sum_{\nu=0}^{\nu_i - 1} \frac{(A - \lambda_i)^\nu}{(\lambda - \lambda_i)^{\nu+1}} P_i,$$

we see that P_i is just the residue of the resolvent $R(\cdot, A)$ at λ_i. This interpretation of Theorem 2.2.5(i) is not essential in our studies now. However, it will become important in the infinite dimensional situation. We will therefore adopt the corresponding terminology right away and henceforth speak of ν_i as the *order of the pole* λ_i of $R(\cdot, A)$.

The following is a useful characterization of X_i by means of A.

Lemma 2.3.2. *The spectral subspace of A corresponding to the eigenvalue λ_i is*

$$X_i = \ker(\lambda_i - A)^{\nu_i},$$

where ν_i is the order of the pole λ_i of $R(\cdot, A)$.

Proof. The inclusion $X_i \subset \ker(\lambda_i - A)^{\nu_i}$ follows since $(\lambda_i - A)$ is nilpotent of order ν_i on X_i by Theorem 2.2.3.

Suppose now that the inclusion is strict, i.e. there exists a nonzero $x \in \ker(\lambda_i - A)^{\nu_i} \backslash X_i$. Note that for some $j \neq i$ there is a nonzero $y := P_j x \in X_j \cap \ker(\lambda_i - A)^{\nu_i}$. Take the largest $p \in \mathbb{N}$ such that $z := (\lambda_i - A)^p y \neq 0$. Then $z \in X_j$ and $Az = \lambda_i z$, thus $(A - \lambda_j)^{\nu_j} z = (\lambda_i - \lambda_j)^{\nu_j} z \neq 0$ which is a contradiction. ☐

We formulate two special cases explicitly.

Corollary 2.3.3. *(i) If $\nu_i = 1$, then $A_{|X_i} = \lambda_i I_{X_i}$.*

(ii) If all $\nu_i = 1$, then A is diagonalizable.

We are now going to uncover more information on A and $f(A)$ that is hidden in Theorems 2.2.3 and 2.2.5. To that purpose we will frequently use the following result.

Theorem 2.3.4 (Spectral Mapping Theorem). *Let $A \in M_n(\mathbb{C})$ and $f \in C_A^\infty$. Then*

$$\sigma\big(f(A)\big) = f\big(\sigma(A)\big) =: \big\{f(\lambda) : \lambda \in \sigma(A)\big\}.$$

Proof. If $\mu \notin f(\sigma(A))$, then $\frac{1}{\mu - f(\lambda)} =: u(\lambda) \in C_A^\infty$ with $u \cdot (\mu - f) = \mathbb{1}$ on a neighborhood of $\sigma(A)$. Hence plugging A into $u \cdot (\mu - f)$ gives

$$u(A)\big(\mu - f(A)\big) = \big(\mu - f(A)\big)u(A) = I,$$

hence $\mu - f(A)$ is invertible and $\mu \notin \sigma(f(A))$.

On the other hand, if x_i is an eigenvector of A belonging to λ_i for some $1 \le i \le m$, then $P_j x_i = 0$ for all $j \neq i$ by Theorem 2.2.3, and Theorem 2.2.5 gives

$$f(A)x_i = f(\lambda_i)x_i,$$

hence $f(\lambda_i) \in \sigma(f(A))$. □

The following is a consequence of Lemma 2.3.2 and Theorem 2.3.4.

Corollary 2.3.5. *The matrix A is nilpotent (i.e., $A^k = 0$ for suitable $k \in \mathbb{N}$) if and only if $\sigma(A) = \{0\}$.*

Proof. If $A^k = 0$, then $\{0\} = \sigma(A^k) = [\sigma(A)]^k$, hence $\sigma(A) = \{0\}$. If, on the other hand, $\sigma(A) = \{0\}$, then the spectral projection corresponding to the only eigenvalue 0 must be equal to the identity and $A = A - 0$ nilpotent on the corresponding spectral subspace which is equal to X. □

Corollary 2.3.6. *(i) $\lambda_i - A$ is bijective on $\bigoplus_{j \neq i} X_j$.*

(ii) $\sigma(A_{|X_i}) = \{\lambda_i\}$, $\sigma(A_{|\bigoplus_{j\neq i} X_j}) = \sigma(A)\backslash\{\lambda_i\}$ for $1 \le i \le m$.

Proof. (i) By Corollary 2.3.5 the spectrum of $A - \lambda_i$ restricted to X_i is zero. Hence, $\sigma(A_{|X_i}) = \{\lambda_i\}$ by the Spectral Mapping Theorem 2.3.4. Since all eigenvectors belonging to λ_i are contained in X_i we have $\lambda_i \notin \sigma(A_{|\bigoplus_{j\neq i} X_j})$. Again, by Theorem 2.3.4, we obtain $0 \notin \sigma((A - \lambda_i)_{|\bigoplus_{j\neq i} X_j})$. This means that $A - \lambda_i$ is bijective on $\bigoplus_{j\neq i} X_j$.
(ii) By above we already have $\sigma(A_{|X_i}) = \{\lambda_i\}$ and $\sigma(A_{|\bigoplus_{j\neq i} X_j}) \subseteq \sigma(A)\backslash\{\lambda_i\}$. Since $\lambda_k \in \sigma(A_{|\bigoplus_{j\neq i} X_j})$ for $k \neq i$, the proof is complete. □

The argument used in the second part of the proof of the Spectral Mapping Theorem 2.3.4 gives the following.

Corollary 2.3.7. *The eigenspace of $f(A)$ belonging to $f(\lambda_i)$ contains all eigenvectors of A that belong to the eigenvalues λ_j with $f(\lambda_j) = f(\lambda_i)$.*

It should be noted, however, that the eigenspace of $f(A)$ to the eigenvalue $f(\lambda_i)$ need not be generated by eigenvectors of A, as the following simple example shows.

Example 2.3.8. The matrix $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ has spectrum $\sigma(A) = \{0\}$, the corresponding eigenspace has dimension 1. By contrast, since $A^2 = 0$, the eigenspace of A^2 to 0 is twodimensional.

The situation is much simpler for the spectral projections. First we need an auxiliary result.

Lemma 2.3.9. *Let $g \in C^\infty_{f(A)}$. Then $g \circ f \in C^\infty_A$ and*

$$(g \circ f)(A) = g(f(A)).$$

Proof. The hypothesis means that $g \in C^\infty$ on a neighborhood of $\sigma(f(A)) = f(\sigma(A))$, hence $g \circ f \in C^\infty$ on a neighborhood of $\sigma(A)$. The rest is straightforward since we can now assume that f and g are polynomials. □

Theorem 2.3.10. *Fix $1 \le i \le m$. The spectral projection of $f(A)$ belonging to $f(\lambda_i)$ is the sum of the spectral projections P_j of A with $f(\lambda_j) = f(\lambda_i)$.*

Proof. Let us denote by $P_i^f = \mathbb{1}_{V_i}(f(A))$ the spectral projection of $f(A)$ belonging to $f(\lambda_i)$, where $V_i \subset \mathbb{C}$ is an open set such that $f(\lambda_i) \in V_i$. Using this notation, we need to prove is that

$$P_i^f = \bigoplus_{f(\lambda_j)=f(\lambda_i)} P_j.$$

By Lemma 2.3.9 we have $P_i^f = (\mathbb{1}_{V_i} \circ f)(A)$, so the statement follows, since $\mathbb{1}_{V_i} \circ f$ is a characteristic function of a neighborhood of exactly those eigenvalues λ_j with $f(\lambda_j) = f(\lambda_i)$. □

The question arises whether there are simple and general relations between the pole orders of $R(\cdot, A)$ at λ_i and $R(\cdot, f(A))$ at $f(\lambda_i)$. However, the following examples show that both an increase or a decrease can occur.

Examples 2.3.11. (1) Let $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, $f(\lambda) = \lambda^2$. Then 0 is a pole of order 2 of $R(\cdot, A)$, but $0 = f(0)$ is a simple pole of $R(\cdot, f(A)) = R(\cdot, 0)$.

(2) Consider

$$A = \begin{pmatrix} 2\pi i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad f(\lambda) = \mathrm{e}^\lambda.$$

Then $2\pi i$ is a pole of order 1 of $R(\cdot, A)$, but $1 = f(2\pi i)$ is a pole of order 2 of $R(\cdot, f(A))$ since $f(A) = \mathrm{e}^A = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right)$.

We close this section with a result that clarifies, at least for the functions $f_t(\lambda) = \mathrm{e}^{t\lambda}$, $t \in \mathbb{R}$, the relation between eigenvectors or pole orders with respect to A and $f(A)$.

Theorem 2.3.12. *Let $t \neq 0$ and define e^{tA} as in Theorem 2.2.5 (ii).*

(i) *The eigenspace of* e^{tA} *at* $e^{t\lambda_i}$ *is the sum of those eigenspaces of* A *that belong to eigenvalues* λ_j *with* $e^{t\lambda_j} = e^{t\lambda_i}$.

(ii) *The pole order of* $e^{t\lambda_i}$ *with respect to* $R(\cdot, e^{tA})$ *dominates the pole order of* λ_i *with respect to* $R(\cdot, A)$.

(iii) *If all* λ_j *with* $e^{t\lambda_j} = e^{t\lambda_i}$ *are first order poles of* $R(\cdot, A)$, *then* $e^{t\lambda_i}$ *is a pole of order* 1 *of* $R(\cdot, e^{tA})$.

Proof. (i) Let x be an eigenvector of e^{tA} for $\mathrm{e}^{t\lambda_i}$ and some fixed i. Then $x \in \bigoplus_{\mathrm{e}^{t\lambda_j}=\mathrm{e}^{t\lambda_i}} X_j$ by Theorem 2.3.10. Let us denote by $i_1, \dots, i_r$ those j with $\mathrm{e}^{t\lambda_j} = \mathrm{e}^{t\lambda_i}$, so

$$x = \sum_{s=1}^{r} y_s \quad \text{with } y_s \in X_{i_s}.$$

We will show that actually $y_s \in \ker(A - \lambda_{i_s})$ for all s. We first observe that each X_{i_s} is invariant under e^{tA}, hence

$$\sum_{s=1}^{r} \mathrm{e}^{t\lambda_i} y_s = \mathrm{e}^{t\lambda_i} x = \mathrm{e}^{tA} x = \sum_{s=1}^{r} \mathrm{e}^{tA} y_s$$

implies $\mathrm{e}^{t\lambda_i} y_s = \mathrm{e}^{tA} y_s$ for $s = 1, \dots, r$. Now, for each fixed s, either $\nu_{i_s} = 1$ and we are done, or $\nu_{i_s} \geq 2$ and we can write

$$0 = \mathrm{e}^{tA} y_s - \mathrm{e}^{t\lambda_i} y_s = \sum_{\nu=1}^{\nu_{i_s}-1} \frac{\mathrm{e}^{t\lambda_{i_s}} t^\nu}{\nu!} (A - \lambda_{i_s})^\nu y_s.$$

If $(A-\lambda_{i_s}) y_s \neq 0$, the right hand side must also be nonzero by Lemma 2.2.9, a contradiction.

(ii) Generally speaking, for any operator B and any $\lambda_0 \in \sigma(B)$ the pole order of λ_0 with respect to $R(\cdot, B)$ is the order of nilpotency of $(B - \lambda_0)$ on X_0, the spectral subspace of B belonging to λ_0. Hence, for $z \in X_0$, this pole order is the maximal length of any chain

$$z, (B - \lambda_0) z, \dots, (B - \lambda_0)^\nu z \neq 0.$$

Going now back to A, suppose $\nu_i \geq 2$ and take $z \in X_i$ with $(A - \lambda_i)^{\nu_i - 1} z \neq 0$. Then

$$\left(\mathrm{e}^{tA} - \mathrm{e}^{t\lambda_i}\right) z = \sum_{\nu=1}^{\nu_i - 1} \frac{\mathrm{e}^{t\lambda_i} t^\nu}{\nu!} (A - \lambda_i)^\nu z =: z_1 \neq 0,$$

hence the pole order of $\mathrm{e}^{t\lambda_i}$ with respect to $R(\cdot, \mathrm{e}^{tA})$ is at least 2. If $\nu_i \geq 3$, we can apply $(A - \lambda_i)^{\nu_i - 2}$ to z_1, and obtain a vector $\neq 0$. Moreover,

$$\left(\mathrm{e}^{tA} - \mathrm{e}^{t\lambda_i}\right) z_1 = \sum_{\nu=1}^{\nu_i - 2} \frac{\mathrm{e}^{t\lambda_i} t^\nu}{\nu!} (A - \lambda_i)^\nu z_1 \neq 0,$$

hence the pole order of $\mathrm{e}^{t\lambda_i}$ is ≥ 3. This can be repeated $(\nu_i - 1)$-times and gives the conclusion.

(iii) This follows easily from Theorem 2.2.5(ii) and Theorem 2.3.10. □

2.4 Notes and Remarks

There are of course many equivalent ways to define functional calculus for matrices. The first one is an application of Jordan canonical form that is usually taught in the first linear algebra course. We have also already mentioned the approach with power series representation (see Remark 2.1.5 and Corollary 2.2.7), which has a drawback that it only works directly for entire functions f making the analysis of spectral projections extremely difficult. Using Cauchy formula we can also show that our definition of $f(A)$ agrees with the Dunford's Integral representation. I.e., for every $A \in \mathcal{L}(X)$ and $f \in C_A^\infty$ there is an open set $W \supset \sigma(A)$ such that

(i) $\overline{W}$ is contained in the domain of f.

(ii) ∂W consists of a finite number of smooth closed curves.

(iii) $f(A) = \dfrac{1}{2\pi i} \displaystyle\int_{\partial W} f(\lambda) R(\lambda, A)\, d\lambda.$

2.5 Exercises

1. Let A be a $n \times n$ diagonalizable matrix with m distinct eigenvalues $\lambda_1, \dots, \lambda_m$. Prove that in this case its spectral projections are of the form

$$P_i = \prod_{j \neq i} \frac{A - \lambda_j}{\lambda_i - \lambda_j}, \quad i = 1, \dots, m.$$

2. (a) Determine the eigenvalues λ_i and the corresponding multiplicities ν_i for the following matrices.

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

   (b) Discuss further matrices you find interesting.

3. Calculate e^{tA}, A^n and $\sin(tA)$ where $A = \begin{pmatrix} 6 & -1 \\ 3 & 2 \end{pmatrix}$. Discuss further matrices you find interesting.

4. Show that for $B = S^{-1}AS$, where S is an invertible matrix, and $f \in C_A^\infty$ it follows $f \in C_B^\infty$ and $f(B) = S^{-1} f(A) S$.

5. (a) Show that Theorem 2.2.3 does not hold in the case of real scalars (i.e., in the situation of $X = \mathbb{R}^n$, $\mathcal{L}(X) = M_n(\mathbb{R})$ and $\mathbb{R}[x]$).

(b) Which one of the arguments leading to Theorem 2.2.3 does not hold in the real case?

6. Let $C \in \mathrm{M}_n(\mathbb{C})$ with $\|C\| < 1$. Show that

$$(I - C)^{-1} = \sum_{n=0}^{\infty} C^n. \tag{2.10}$$

7. Using the result of the previous exercise, show the Neumann series representation (2.9) for $|\lambda| > \|A\|$.

8. Show, using the definition, the so called resolvent equations

$$AR(\lambda, A) = \lambda R(\lambda, A) - I \tag{2.11}$$

for $\lambda \in \rho(A)$, and

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A) \tag{2.12}$$

for $\lambda, \mu \in \rho(A)$.

9. Using the results of the previous exercises, show that for $|\lambda - \mu| < \frac{1}{\|R(\mu,A)\|}$, we have

$$R(\lambda, A) = \sum_{n=0}^{\infty} (\mu - \lambda)^n R(\mu, A)^{n+1}. \tag{2.13}$$

10. Give an alternative proof of Theorem 2.3.12 (ii), using the formula

$$R(\mu, \mathrm{e}^{tA}) = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \frac{g_\mu^{(\nu)}(\lambda_i)}{\nu!} (A - \lambda_i)^\nu P_i$$

with $g_\mu(\lambda) = (\mu - \mathrm{e}^{t\lambda})^{-1}$ ($\mu \notin \mathrm{e}^{t\sigma(A)}$). Can you determine the pole order in question?

11. * Find conditions on f such that the assertions of Theorem 2.3.12 hold for $f(A)$ instead of e^{tA}.

Lecture 3

Powers of Matrices

3.1 The coordinate sequences

We further on assume X to be a vector space with $\dim X = n < \infty$. We will employ the formulae obtained in Lecture 2 to study the asymptotic behavior of the powers of a given operator acting on X. In order to be consistent with the exponential function later on, we denote here the given operator by $T \in \mathcal{L}(X)$. We retain, however, the notation of Theorem 2.2.5 with respect to the eigenvalues $\lambda_1, \ldots, \lambda_m$, the multiplicities $\nu_1, \ldots, \nu_m$ (as zeroes of the minimal polynomial of T), and the spectral projections $P_1, \ldots, P_m$.
Thus, by Theorem 2.2.5.(iii), we have

$$T^k = \sum_{i=1}^{m} \sum_{\nu=0}^{\min\{\nu_i-1,k\}} \binom{k}{\nu} \lambda_i^{k-\nu} (T-\lambda_i)^{\nu} P_i \quad \text{for } k \in \mathbb{N}. \tag{3.1}$$

Example 3.1.1. Let (f_n) be the Fibonacci sequence considered as the motivating example in Section 1.1. We have seen that

$$\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{for} \qquad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad n = 0, 1, 2, \ldots.$$

It is not difficult to compute the eigenvalues of A,

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \qquad \text{and} \qquad \lambda_2 = \frac{1-\sqrt{5}}{2},$$

and the corresponding spectral projections

$$P_1 = \frac{\sqrt{5}}{5} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{-1+\sqrt{5}}{2} \end{pmatrix} \qquad \text{and} \qquad P_2 = \frac{\sqrt{5}}{5} \begin{pmatrix} \frac{-1+\sqrt{5}}{2} & -1 \\ -1 & \frac{1+\sqrt{5}}{2} \end{pmatrix}.$$

By formula (3.1) we obtain

$$A^n = \lambda_1^n P_1 + \lambda_2^n P_2 = \frac{\sqrt{5}}{5} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} & \left(\frac{1-\sqrt{5}}{2}\right)^{n}\left(\frac{1+\sqrt{5}}{2}\right) - \left(\frac{1+\sqrt{5}}{2}\right)^{n}\left(\frac{1-\sqrt{5}}{2}\right) \end{pmatrix}$$

and thus the explicit formula for the entries of the Fibonacci sequence is

$$f_n = \frac{\sqrt{5}}{5}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right), \quad n = 0, 1, 2, \ldots.$$

Our main interest, however, lies in the asymptotic behavior. In order to understand what happens with T^k as $k \to \infty$, we use the linear independence of the set

$$B_T := \{(T - \lambda_i)^\nu P_i : i = 1, \ldots, m;\ \nu = 0, \ldots, \nu_i - 1\}$$

in the vector space $\mathcal{L}(X)$ (cf. Lemma 2.2.9). If we extend this set to a basis $\mathbb{B}_T$ of $\mathcal{L}(X)$, then (3.1) means that the (non-zero) coordinates of T^k with respect to $\mathbb{B}_T$ are

$$\left\{\binom{k}{\nu}\lambda_i^{k-\nu} : i = 1, \ldots, m;\ \nu = 0, \ldots, \nu_i - 1\right\}$$

(since from now on we consider $k \to \infty$, we allow ourselves to simplify the upper bound of ν from $\min\{\nu_i - 1, k\}$ to $\nu_i - 1$). Likewise, if we are interested in the "orbit" $\{T^k x : k \in \mathbb{N}\}$ of a single element $x \in X$ under the powers T^k of T, we may use a basis of X containing the set

$$\{(T - \lambda_i)^\nu P_i x : i = 1, \ldots, n; \nu = 0, \ldots, \nu_i - 1\} \backslash \{0\}.$$

Again, the corresponding coordinate sequences are among the ones obtained above. Thus, since convergence in a finite dimensional vector space is convergence in every coordinate, no matter what basis is employed, the behavior of T^k (or of $T^k x$ for $x \in X$) as $k \to \infty$ is reflected by the behavior of the sequences

$$z_{\lambda,\nu}(k) := \binom{k}{\nu}\lambda^{k-\nu} \tag{3.2}$$

for $\lambda \in \sigma(T)$, $\nu = 0, \ldots, n-1$ (observe that $\nu_i \le n$ for each i). In case of convergence of all these sequences, $(T^k)_{k\in\mathbb{N}}$ (or $T^k x$ for a given $x \in X$) converges and the coordinates of the limit are obtained as the limits of the corresponding coordinate sequences.

The advantage of this approach is that the behavior of functions in (3.2) is easily understood and essentially depends on the modulus of λ.

- If $|\lambda| < 1$, then $z_{\lambda,\nu}(k) \to 0$ as $k \to \infty$ for all ν since $\lim_{k\to\infty} k^\nu \lambda^k = 0$.
- If $|\lambda| > 1$, then $|z_{\lambda,\nu}(k)| \to \infty$ as $k \to \infty$ for all ν.
- If $|\lambda| = 1$ and $\nu = 0$, then $z_{\lambda,0}(k) = \lambda^k$.
- If $|\lambda| = 1$ and $\nu \ge 1$, then $|z_{\lambda,\nu}(k)| \to \infty$ as $k \to \infty$.

Using these facts, we will be able to describe the asymptotics of $T^k x$ for $x \in X_i = \operatorname{rg} P_i$, depending on λ_i and ν_i.

3.1.1 The spectral radius

We begin our investigations with estimates for $\|T^k\|$ related to the *spectral radius*

$$r(T) := \max\{|\lambda| \,:\, \lambda \in \sigma(T)\}\,. \tag{3.3}$$

Lemma 3.1.2. *Let $\|\cdot\|$ be a norm on $\mathcal{L}(X)$, $T \in \mathcal{L}(X)$, $\mu > r(T)$. Then there exist constants $N > 0$, $M \geq 1$ such that for all $k \in \mathbb{N}$,*

$$N \cdot r(T)^k \leq \|T^k\| \leq M \cdot \mu^k.$$

If $\|\cdot\|$ is an operator norm, we can choose $N = 1$.

Proof. By the Spectral Mapping Theorem 2.3.4,

$$r(T^k) = r(T)^k,$$

hence the last part of the assertion is clear for any operator norm on $\mathcal{L}(X)$ by Corollary 1.2.7. Then the lower estimate

$$N \cdot r(T)^k \leq \|T^k\|$$

(with suitable $N > 0$) follows by equivalence of norms for finite dimensional spaces.

Turning to the upper estimate in the assertion, we note that the coordinates of T^k with respect to a basis of $\mathcal{L}(X)$ containing B_T are

$$z_{\lambda,\nu}(k) := \binom{k}{\nu}\lambda^{k-\nu}$$

where $1 \leq i \leq m$, $0 \leq \nu \leq \nu_i$. Since

$$\binom{k}{\nu} \leq k^\nu,$$

and since for $|\lambda| < 1$

$$k^\nu \lambda^{k-\nu} \to 0$$

as $k \to \infty$, for all $\nu \in \mathbb{N}$, the coordinate sequences of $\frac{T^k}{\mu^k}$ remain bounded as $k \to \infty$. Thus,

$$\|T^k\|_\infty \leq C\mu^k \qquad (k \in \mathbb{N})$$

for a suitable constant C, where $\|\cdot\|_\infty$ denotes the usual maximum norm with respect to coordinates belonging to B_T. The desired estimate for $\|T^k\|$ again follows by the equivalence of norms. □

It is immensely important to know that the spectral radius of T can be determined through the sequence $(\|T^k\|)_{k\in\mathbb{N}}$, regardless of the norm $\|\cdot\|$.

Proposition 3.1.3 (*Gelfand's formula*)**.** *For any $T \in \mathcal{L}(X)$ the following holds.*

(i) $r(T) = \lim_{k\to\infty} \|T^k\|^{\frac{1}{k}}$ *for any norm* $\|\cdot\|$ *on* $\mathcal{L}(X)$.

(ii) If $\|\cdot\|$ *is an operator norm on* $\mathcal{L}(X)$*, then* $r(T) = \inf_{k>0} \|T^k\|^{\frac{1}{k}}$.

Proof. The proof is an immediate consequence of the previous Lemma 3.1.2, which yields the estimate

$$N^{\frac{1}{k}} r(T) \le \|T^k\|^{\frac{1}{k}} \le M^{\frac{1}{k}} \mu$$

whenever $\mu > r(T)$ for suitable constants N, M. If $\|\cdot\|$ is an operator norm, then the (admissible) choice $N = 1$ in these estimates forces $r(T) = \inf_{k>0} \|T^k\|^{\frac{1}{k}}$. □

Let us now turn our attention back to estimates for $\|T^k x\|$ and suppose again that the same $\|\cdot\|$ denotes the corresponding operator norm. The special case when $x \in X_i$ for some i can be treated analogously to our previous considerations.

Proposition 3.1.4. *Let* $T \in \mathcal{L}(X)$ *and* $\|\cdot\|$ *any norm on* X*. Then for every* $\mu > |\lambda_i|$ $(1 \le i \le m)$ *there exists a number* $M \ge 1$ *such that*

$$\|T^k x\| \le M\mu^k \|x\|$$

for all $k \in \mathbb{N}$ *and* $x \in X_i$*. Further, if* $\lambda_i \neq 0$*, then for every* $0 < \rho < |\lambda_i|$ *there exists* $N > 0$ *such that*

$$N\rho^k \|x\| \le \|T^k x\| \le M\mu^k \|x\|$$

for all $k \in \mathbb{N}$ *and* $x \in X_i$*.*

Proof. Since $\|T^k x\| \le \|T^k\| \cdot \|x\|$, and since $\sigma(T|_{X_i}) = \{\lambda_i\}$, we obtain from Lemma 3.1.2 the existence of a constant M for every $\mu > |\lambda_i|$ such that

$$\|T^k x\| \le M\mu^k \|x\|$$

for all $k \in \mathbb{N}$ and $x \in X_i$.

On the other hand, if $\lambda_i \neq 0$, then $T|_{X_i}$ has an inverse S with $\sigma(S) = \{\lambda_i^{-1}\}$. So for every $0 < \rho < |\lambda_i|$, there exists $C \ge 1$ such that for all $z \in X_i$, we obtain

$$\|S^k z\| \le C\rho^{-k} \|z\|.$$

Choosing $z := T^k x$ for $k \in \mathbb{N}$ we have $S^k z = x$ and hence

$$\|x\| \le C\rho^{-k} \|z\|,$$

that is

$$\|T^k x\| = \|z\| \ge \frac{1}{C} \rho^k \|x\|.$$

□

3.2 Asymptotics

After this intermezzo on the spectral radius $r(T)$ we now describe the long-time behavior of the sequence $(T^k)_{k\in\mathbb{N}}$. We will consider various types of asymptotical behavior.

Definition 3.2.1. For $T \in \mathcal{L}(X)$ and any norm $\|\cdot\|$ on $\mathcal{L}(X)$ we call the sequence $(T^k)_{k\in\mathbb{N}}$

- *bounded*[1], if $\sup_{k\in\mathbb{N}} \|T^k\| < \infty$;
- *stable*[1], if $\lim_{k\to\infty} \|T^k\| = 0$;
- *convergent*, if $\lim_{k\to\infty} T^k = P$ for some $P \in \mathcal{L}(X)$;
- *periodic* with *period* p, if T is periodic with period p, i.e., $T^p = I$;
- *Cesàro summable* (or T is *mean-ergodic*), if the limit $\lim_{k\to\infty} \frac{1}{k}\sum_{l=0}^{k-1} T^l$ exists.

Remarks 3.2.2. 1. Note that boundedness of the sequence $(T^k)_{k\in\mathbb{N}}$ is equivalent to boundedness of all coordinate sequences $z_{\lambda,\nu}(k)$ with respect to $\mathbb{B}_T$.

2. If the sequence $(T^k)_{k\in\mathbb{N}}$ converges to P then

$$TP = T\lim_{k\to\infty} T^k = \lim_{k\to\infty} T^{k+1} = P = PT = \lim_{k\to\infty} T^{2k} = P^2,$$

i.e. P is a projection commuting with T.

3. If $(\alpha_k)_{k\in\mathbb{N}}$ is a sequence in a vector space, then $\alpha^{(k)} := \frac{1}{k}\sum_{l=0}^{k-1} \alpha_l$ is called the corresponding sequence of *Cesàro means*. Analogously, we call

$$T^{(k)} := \frac{1}{k}\sum_{l=0}^{k-1} T^l, \quad k \in \mathbb{N},$$

appearing in the definition above, the *Cesàro means* of T.

The asymptotical behavior of $(T^k)_{k\in\mathbb{N}}$ will depend essentially on the size of $r(T)$ compared to 1. By Γ we denote the unit circle in $\mathbb{C}$ and by $\Gamma_q = \{e^{\frac{2k\pi i}{q}} : k = 0, \ldots, q-1\}$ the set of all q-th rootsof unity in $\mathbb{C}$. We call $\lambda_0 > 0$ a *radially dominant eigenvalue*, if $\lambda_0 \in \sigma(T)$ and $|\lambda| < \lambda_0$ for all $\lambda \in \sigma(T)\backslash\{\lambda_0\}$.

Theorem 3.2.3. *For $T \in \mathcal{L}(X)$ the following assertions hold.*

(i) $(T^k)_{k\in\mathbb{N}}$ is stable if and only if $r(T) < 1$.

(ii) $(T^k)_{k\in\mathbb{N}}$ is bounded if and only if $r(T) \le 1$ and all eigenvalues of modulus 1 are simple poles of $R(\cdot, T)$.

[1]Note that working with ODEs or dynamical systems, a different terminology is also widely accepted: what we call "bounded" is often also called "stable", and what we call "stable" is also often called "asymptotically stable".

(iii) $(T^k)_{k\in\mathbb{N}}$ *is periodic with period* p *if and only if it is bounded and* $\sigma(T)\setminus\{0\}\subseteq\Gamma_p$.

(iv) $1\in\sigma(T)$ *and* $\lim_{k\to\infty} T^k = P_1$ *(*P_1 *denotes the spectral projection of* T *belonging to* 1*) if and only if* 1 *is a radially dominant eigenvalue of* T *which is a simple pole of the resolvent* $R(\cdot, T)$.

(v) $\sigma(T)\cap\Gamma=\emptyset$ *if and only if there exist* T*-invariant subspaces* X_s *and* X_u *such that* $X = X_s\oplus X_u$ *and*

$$\lim_{k\to\infty}\|T^k x\| = 0 \text{ for } x\in X_s \quad \text{and} \quad \lim_{k\to\infty}\|T^k x\| = \infty \text{ for } x\in X_u.$$

Proof. (i) This is a direct consequence of Lemma 3.1.2.

(ii) Assume that $(T^k)_{k\in\mathbb{N}}$ is bounded. Lemma 3.1.2 again yields $r(T)\le 1$. If an eigenvalue $\lambda\in\sigma(T)$ with $|\lambda|=1$ would have multiplicity $\nu>1$, then the corresponding coordinate sequence $z_{\lambda,\nu-1}(k)$ of T^k with respect to $\mathbb{B}_T$ would not be bounded, a contradiction.

Conversely, if $r(T)\le 1$ and all eigenvalues of modulus 1 are simple poles of the resolvent, then all the coordinate sequences of T^k with respect to $\mathbb{B}_T$ are bounded.

(iii) Let $(T^k)_{k\in\mathbb{N}}$ be bounded and $\sigma(T)\setminus\{0\}\subseteq\Gamma_p$. By (ii) all nonzero eigenvalues have multiplicity 1, thus the formula (3.1) is simplified as

$$T^k = \sum_{i=1}^{m}\lambda_i^k P_i \quad \text{for } k\in\mathbb{N}.$$

Since $\lambda_i^p = 1$ for all $i=1,\dots,m$, we have $T^p = I$.

If $(T^k)_{k\in\mathbb{N}}$ is periodic, i.e. $T^p = I$ for some $p\in\mathbb{N}$, it is bounded and for every $\lambda\in\sigma(T)$ we have $\lambda^p = 1$.

(iv) Let $\lim_{k\to\infty} T^k = P_1$. From (i) and (ii) follows that then $r(T)=1$. If there exists an eigenvalue $\lambda\ne 1$ with $|\lambda|=1$, then the corresponding coordinate sequence of T^k with respect to $\mathbb{B}_T$ contains $z_{\lambda,0}(k)=\lambda^k$, which does not converge as $k\to\infty$. Also, if 1 is not a simple pole, then T^k has coordinate $z_{\lambda,1}(k)=k$, which again does not converge as $k\to\infty$.

Conversely, if 1 is a radially dominant eigenvalue and a simple pole of the resolvent, then all coordinate sequences $z_{\lambda,0}(k)$ with respect to $\mathbb{B}_T$ converge. Since by (i) the coordinate sequences belonging to eigenvalues with $|\lambda|<1$ converge to 0, $T^k\to P_1$ as $k\to\infty$.

(v) Let

$$X_s := \bigoplus_{|\lambda_i|<1} X_i \qquad \text{and} \qquad X_u := \bigoplus_{|\lambda_i|>1} X_i,$$

and define the operators $T_s := T|_{X_s}$ and $T_u := T|_{X_u}$. Since $\sigma(T) = \sigma(T_s) \cup \sigma(T_u)$, the assertion follows using (i) and (ii) for the operators T_s and T_u.

□

Summing up we see that the limit $\lim_{k\to\infty} T^k$ exists only in either one of the following two situations:

- if $r(T) < 1$, then $T^k \to 0$ as $k \to \infty$, and
- if $r(T) = 1 = \lambda_1$ is radially dominant eigenvalue with multiplicity $\nu_1 = 1$, then $T^k \to P_1$ as $k \to \infty$.

3.2.1 Cesàro summability

In the remainder of this lecture we will discuss situations where a suitable subsequence of $(T^k)_{k\in\mathbb{N}}$ converges, or where the Cesàro means of T converge.

Lemma 3.2.4 (Kronecker)**.** *Let $\lambda_1, \dots, \lambda_k \in \Gamma$. There exists a sequence $(s_n)_{n\in\mathbb{N}}$, $s_n \in \mathbb{N}$, such that*

$$\lim_{n\to\infty} \lambda_i^{s_n} = 1$$

for all $1 \le i \le k$.

Proof. We give the proof for the situation $k = 2$. The general case can be obtained in the same manner.
On the torus $\Gamma \times \Gamma = \{(e^{i\varphi_1}, e^{i\varphi_2})\colon 0 \le \varphi_{1,2} < 2\pi\}$, we define a metric d by

$$d\left((e^{i\varphi_1}, e^{i\varphi_2}), (e^{i\zeta_1}, e^{i\zeta_2})\right) := \max\{|\varphi_1 - \zeta_1| (\text{mod } 2\pi), |\varphi_2 - \zeta_2| (\text{mod } 2\pi)\}.$$

Let $(\lambda_1, \lambda_2) \in \Gamma \times \Gamma$. If there exists a $k \in \mathbb{N}$ such that $\lambda_1^k = 1 = \lambda_2^k$, then taking $s_n := k \cdot n$ for all $n \in \mathbb{N}$ we obtain

$$(\lambda_1, \lambda_2)^{s_n} := (\lambda_1^{s_n}, \lambda_2^{s_n}) = (1, 1),$$

and the assertion of the lemma holds.

Now assume $\lambda_1 = e^{i\alpha}$ with $\frac{\alpha}{2\pi} \in \mathbb{R} \setminus \mathbb{Q}$. In this case all powers $(\lambda_1, \lambda_2)^k$, $k \in \mathbb{N}$, are pairwise different. Let $\xi_n := \frac{2\pi}{n}$ and consider all "boxes" of the form $\hat{\Gamma}_j \times \hat{\Gamma}_k$ where the sector $\hat{\Gamma}_j$ is defined as

$$\hat{\Gamma}_j := \left\{e^{i\xi}\colon (j-1)\xi_n \le \xi \le j\xi_n\right\}.$$

For $1 \le j, k \le n$ there are n^2 boxes of the form $\hat{\Gamma}_j \times \hat{\Gamma}_k$. Thus by Dirichlet's pigeonhole principle two of the n^2+1 powers $(\lambda_1, \lambda_2)^l$ with $0 \le l \le n^2$ are in the same box. If l and m are the exponents with $0 \le l < m \le n^2$ such that $(\lambda_1, \lambda_2)^l$ and $(\lambda_1, \lambda_2)^m$ are in the same box, then

$$d((\lambda_1, \lambda_2)^l, (\lambda_1, \lambda_2)^m) < \xi_n.$$

Observe that multiplication by a nonzero element of $\Gamma \times \Gamma$ is well defined and is an isometry for the metric d. Therefore,

$$d((1, 1), (\lambda_1, \lambda_2)^{m-l}) < \xi_n$$

The statement of the lemma now follows as $n \to \infty$. □

Using the same arguments as in the proof of Theorem 3.2.3 together with Lemma 3.2.4 we obtain the following result on the convergence of a subsequence of $(T^k)_{k\in\mathbb{N}}$ (see Exercise 2).

Theorem 3.2.5. *Let $T \in \mathcal{L}(X)$. Then the following assertions are equivalent.*

(i) There exists a subsequence of $(T^k)_{k\in\mathbb{N}}$ which converges, as $k \to \infty$, to some limit $P \neq 0$.

(ii) $r(T) = 1$ and all eigenvalues of T with modulus 1 are simple poles of the resolvent.

Definition 3.2.6. We call $T \in M_n(\mathbb{C})$ a *spectral contraction* if the conditions in Theorem 3.2.5 are satisfied.

The following is an alternative characterization of the spectral contraction that explains its name.

Theorem 3.2.7. *Let $T \in M_n(\mathbb{C})$. Then the following assertions are equivalent.*

(i) T is a spectral contraction.

(ii) There exists a norm $\|\cdot\|$ on $\mathbb{C}^n$, such that for the corresponding operator norm on $M_n(\mathbb{C})$ it holds $\|T^k\| = 1$ for all $k \in \mathbb{N}$.

(iii) $r(T) = 1$ and the sequence $\left(T^k\right)_{k\in\mathbb{N}}$ is bounded.

Proof. Let $|||\cdot|||$ be a norm on $\mathbb{C}^n$. If (i) holds, then $\left|\left|\left|T^k\right|\right|\right| \leq M, k \in \mathbb{N}$, for some $M < \infty$. Put

$$\|x\| := \sup_{k\in\mathbb{N}\cup\{0\}} \left|\left|\left|T^k x\right|\right|\right| .$$

Then $\|\cdot\|$ is a norm on $\mathbb{C}^n$ such that for the corresponding operator norm $\|T^k\| = 1$ for all $k \in \mathbb{N}$.

Further, if (ii) holds, then $r(T) = 1$ and, by the equivalence of norms $\left(\|T^k\|\right)_{k\in\mathbb{N}}$ is bounded for any norm $\|\cdot\|$ on $M_n(\mathbb{C})$.

Finally, assuming (iii) Theorem 3.2.3(ii) yields (ii) of Theorem 3.2.5 and the implication loop is closed. □

At the end let us briefly consider the convergence of the Cesàro means $T^{(k)} := \frac{1}{k}\sum_{\nu=0}^{k-1} T^\nu$. We will see that the sequence $(T^k)_{k\in\mathbb{N}}$ is Cesàro summable iff it is bounded, that is either if $r(T) < 1$ or T is a spectral contraction.

Theorem 3.2.8. *For $T \in \mathcal{L}(X)$ the following assertions are equivalent.*

(i) $(T^k)_{k\in\mathbb{N}}$ is Cesàro summable.

(ii) $\lim_{k\to\infty}\left(T^k/k\right) = 0$.

(iii) $(T^k)_{k\in\mathbb{N}}$ is bounded.

(iv) $r(T) \leq 1$ and each eigenvalue of modulus 1 is a simple pole of the resolvent.

In any of these equivalent cases the sequence $(T^{(k)})_{k\in\mathbb{N}}$ of Cesàro means of T converges to 0, if $1 \notin \sigma(T)$, and to the spectral projection P_1 belonging to 1, if $1 \in \sigma(T)$.

Proof. (i) $\Rightarrow$ (ii) follows since we can write $T^k = kT^{(k)} - (k-1)T^{(k-1)}$. (ii) $\Rightarrow$ (iiii) is clear and (iii) $\Rightarrow$ (iv) holds by Theorem 3.2.3(ii). It remains to show (iv) $\Rightarrow$ (i) and the assertion on the limit of the Cesàro means.

The sequence $(T^{(k)})_{k\in\mathbb{N}}$ has as coordinate sequences with respect to $\mathbb{B}_T$ the Cesàro means of the coordinate sequences of $(T^k)_{k\in\mathbb{N}}$:

$$z_{\lambda,\nu}^{(k)} := \frac{1}{k}\sum_{l=0}^{k-1} z_{\lambda,\nu}(l) = \frac{1}{k}\sum_{l=0}^{k-1}\binom{l}{\nu}\lambda^{l-\nu}.$$

These sequences clearly converge to 0, if $|\lambda| < 1$ and diverge, if $|\lambda| > 1$ (see the computation below). Furthermore, for $|\lambda| = 1$ the sequence $z_{\lambda,\nu}^{(k)}$ is bounded only in the case $\nu = 0$, i.e. λ is a simple pole of the resolvent.

If $\lambda \neq 1$, then compute

$$z_{\lambda,0}^{(k)} = \frac{1}{k}\sum_{l=0}^{k-1}\binom{l}{0}\lambda^{l} = \frac{1}{k}\frac{\lambda^k - 1}{\lambda - 1}$$

which tends to zero as $k \to \infty$, if $|\lambda| \leq 1, \lambda \neq 1$. On the other hand, for $\lambda = 1$ we have $z_{1,0}^{(k)} = 1$ for every $k \in \mathbb{N}$ and thus in this case $T^{(k)} \to P_1$ as $k \to \infty$. □

This means that if the sequence $(T^k)_{k\in\mathbb{N}}$ converges, its limit is the same as the Cesàro limit. On the other hand, there are Cesàro summable sequences which are not convergent, as we will see in the Exercises.

3.3 Notes and Remarks

The question of convergence of $(T^k)_{k\in\mathbb{N}}$ is immensely important for numerical methods. Let we just briefly mention a very simple method for computing a dominant eigenpair (λ_1, v_1) of a diagonalizable matrix $T \in M_n(\mathbb{R})$ with eigenvalues

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

We know by Theorem 3.2.3 that

$$\lim_{k\to\infty}\left(\frac{T}{\lambda_1}\right)^k = P_1,$$

where P_1 is the spectral projection of $\frac{T}{\lambda_1}$ corresponding to 1, and consequently

$$\lim_{k\to\infty}\left(\frac{T}{\lambda_1}\right)^k x_0 = P_1 x_0$$

for any $x_0 \in \mathbb{R}^n$. For $P_1 x_0 \neq 0$ therefore $\frac{T^k x_0}{\|T^k x_0\|}$ converges to an eigenvector associated with λ_1. The iterative algorithm basing on these considerations is called *power method*:

$$x_{n+1} = \frac{Tx_n}{\|Tx_n\|}. \tag{3.4}$$

3.4 Exercises

1. For matrices $A = (a_{ij})$ and $B = (b_{ij})$ we write $A \leq B$, if $a_{ij} \leq b_{ij}$ for all i, j, and we denote $|A| := (|a_{ij}|)$. Prove that if $|A| \leq B$ then the following inequalities concerning spectral radii holds
$$r(A) \leq r(|A|) \leq r(B).$$
2. Prove Theorem 3.2.5.
3. Show that, if there exists an operator norm $\|\cdot\|$ on $M_n(\mathbb{C})$ such that $\|T\| < 1$, then the sequence $(T^k)_{k\in\mathbb{N}}$ is stable.
4. For the following matrices compute powers T^k and evaluate $\lim_{k\to\infty} T^k$, if it exists.
$$\text{(a) } T = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad \text{(b) } T = \begin{pmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \qquad \text{(c) } T = \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$
5. Describe the asymptotic behavior of the sequence $(T^k)_{k\in\mathbb{N}}$ for the following special classes of matrices $T \in M_n(\mathbb{R})$.
   (i) T is idempotent (or involutary), i.e. $T^2 = I$.
   (ii) T is nilpotent, i.e. $T^q = 0$ for some $q \in \mathbb{N}$.
   (iii) T is unipotent, i.e. $T - I$ is nilpotent.
   (iv) T is orthogonal, i.e. $T^\top T = TT^\top = I$.
6. For a sequence $(a_k)_{k\in\mathbb{N}}$ in X the associated Cesàro sequence $\left(a^{(k)}\right)_{k\in\mathbb{N}}$ is defined by $a^{(k)} := \frac{1}{k}\sum_{i=1}^k a_i$.
   (i) Prove that if $(a_k)_{k\in\mathbb{N}}$ converges in X, also $\left(a^{(k)}\right)_{k\in\mathbb{N}}$ converges in X and the limits coincide.
   (ii) Find an example of a non-convergent sequence whose associated Cesàro sequence converges.
7. For each of the following matrices determine whether its powers are convergent or Cesàro summable. Evaluate the limit of each convergent matrix and the Cesàro limit of each summable matrix.
$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} -1/2 & 1/2 & -1/2 \\ 1 & 0 & -1/2 \\ 1 & -1 & 1/2 \end{pmatrix}.$$

Lecture 4

Matrix Exponential Function

4.1 Main properties

Let X be a n-dimensional vector space. By the *exponential function* of a complex matrix $A \in \mathcal{L}(X)$ (i.e. $A \in \mathrm{M}_n(\mathbb{C})$) we understand the mapping

$$\mathbb{R} \ni t \mapsto \exp(tA) = \mathrm{e}^{tA} \in \mathcal{L}(X).$$

Here, as explained in Section 2.2, $\exp(tA) = \mathrm{e}^{tA}$ stands for the matrix $f_t(A)$ with $f_t(\lambda) := \mathrm{e}^{t\lambda}$. Therefore, Theorem 2.2.5(ii) says that e^{tA} can be written as

$$\mathrm{e}^{tA} = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \frac{\mathrm{e}^{t\lambda_i} t^{\nu}}{\nu!} (A - \lambda_i)^{\nu} P_i, \tag{4.1}$$

where $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of A with corresponding multiplicities $\nu_1, \ldots, \nu_m$ (as roots of the minimal polynomial) and spectral projections $P_1, \ldots, P_m$. Alternatively, according to Corollary 2.2.7, the matrix e^{tA} is represented by the exponential series

$$\mathrm{e}^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}. \tag{4.2}$$

Formula (4.1) gives an easy access to the following properties of e^{tA}. Firstly, since $f_0(\lambda) = 1$, we have

$$f_0(A) = e^{0A} = I.$$

By the multiplicativity of the functional calculus and the fact that

$$f_{s+t}(\lambda) = f_s(\lambda) \cdot f_t(\lambda), \qquad \lambda \in \mathbb{C},$$

it follows that

$$\mathrm{e}^{(s+t)A} = \mathrm{e}^{sA} \cdot \mathrm{e}^{tA} \tag{4.3}$$

for $s, t \in \mathbb{R}$. Hence, $(\mathrm{e}^{tA})_{t \in \mathbb{R}}$ is a subgroup of the multiplicative semigroup $M_n(\mathbb{C})$, and the mapping $t \mapsto \mathrm{e}^{tA}$ is a homomorphism of the additive group $(\mathbb{R}, +)$ into $M_n(\mathbb{C})$.

Remark 4.1.1. It is usual to refer to these properties of $t \mapsto \mathrm{e}^{tA}$ by saying that $(\mathrm{e}^{tA})_{t\in\mathbb{R}}$ is the *matrix group generated by* A. If we consider only $t \geq 0$, we call $(\mathrm{e}^{tA})_{t\geq 0}$ a *matrix semigroup* generated by A

Furthermore, the function $t \mapsto \mathrm{e}^{tA}$ is differentiable.

Theorem 4.1.2. *The matrix exponential function $t \mapsto e^{tA}$ is differentiable on $\mathbb{R}$ with derivative*

$$\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A, \quad t \in \mathbb{R}. \tag{4.4}$$

Proof. Let $f(\lambda) := \lambda \mathrm{e}^{t\lambda}$ and observe that

$$\left(\lambda \mathrm{e}^{t\lambda}\right)^{(\nu)}(\lambda) = \left(\lambda t^{\nu} + \nu t^{\nu-1}\right)\mathrm{e}^{t\lambda}.$$

Hence applying Theorem 2.2.5 for $f(\lambda) = \lambda \mathrm{e}^{t\lambda}$ we obtain

$$\begin{aligned} A\mathrm{e}^{tA} &= \sum_{i=1}^{m}\sum_{\nu=0}^{\nu_i-1}\left(\lambda_i t^{\nu}\mathrm{e}^{t\lambda_i} + \nu t^{\nu-1}\mathrm{e}^{t\lambda_i}\right)\frac{(A-\lambda_i)^{\nu}}{\nu!}P_i \\ &= \sum_{i=1}^{m}\sum_{\nu=0}^{\nu_i-1}\left[\frac{d}{dt}(t^{\nu}\mathrm{e}^{t\lambda_i})\right]\frac{(A-\lambda_i)^{\nu}}{\nu!}P_i \\ &= \frac{d}{dt}\left(\sum_{i=1}^{m}\sum_{\nu=0}^{\nu_i-1}t^{\nu}\mathrm{e}^{t\lambda_i}\frac{(A-\lambda_i)^{\nu}}{\nu!}P_i\right) \\ &= \frac{d}{dt}\mathrm{e}^{tA}. \end{aligned}$$

Notice that since $\lambda \mathrm{e}^{t\lambda} = \mathrm{e}^{t\lambda}\lambda$, by the properties of functional calculus it follows $A\mathrm{e}^{tA} = \mathrm{e}^{tA}A$. □

The following consequence of formula (4.4) motivates our interest in the behavior of the function $t \mapsto \mathrm{e}^{tA}$ as $t \to \infty$.

Corollary 4.1.3. *Let $A = (a_{ij}) \in M_n(\mathbb{C})$. Then for each $x = (x_1, ..., x_n) \in \mathbb{C}^n$ the function*

$$t \mapsto e^{tA}x =: (x_1(t), ..., x_n(t))$$

is the unique solution of the system of differential equations

$$\begin{aligned} \frac{d}{dt}x_1(t) &= a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\ \frac{d}{dt}x_1(t) &= a_{21}x_1(t) + \cdots + a_{2n}x_n(t) \\ &\vdots \\ \frac{d}{dt}x_n(t) &= a_{n1}x_1(t) + \cdots + a_{nn}x_n(t), \end{aligned}$$

with initial condition

$$(x_1(0), ..., x_n(0)) = (x_1, ..., x_n).$$

Proof. A look at the differential quotient defining the derivative $\frac{d}{dt}(\mathrm{e}^{tA}x)$ and Theorem 4.1.2 teaches us that in fact

$$\frac{d}{dt}(\mathrm{e}^{tA}x) = \left(\frac{d}{dt}\mathrm{e}^{tA}\right)x = (A\mathrm{e}^{tA})x = A(\mathrm{e}^{tA}x)$$

for all $t \in \mathbb{R}$. Since $e^{0A} = I$, it follows that $\mathrm{e}^{tA}x$ is a solution of the above initial value problem. Now, let $x(t)$ be any solution and define $y(t) := e^{-tA}x(t)$. Then

$$\begin{aligned}\frac{d}{dt}y(t) &= \left(\frac{d}{dt}\mathrm{e}^{-tA}\right)x(t) + \mathrm{e}^{-tA}\frac{d}{dt}x(t) \\ &= -A\mathrm{e}^{-tA}x(t) + \mathrm{e}^{-tA}Ax(t) \\ &= 0.\end{aligned}$$

Therefore, $t \mapsto y(t) = \mathrm{e}^{-tA}x(t)$ is constant. Since for $t = 0$ we have $y(0) = x$, we conclude that $x(t) = \mathrm{e}^{tA}x$ for all $t \in \mathbb{R}$. □

Remark 4.1.4. In short, Corollary 4.1.3 tells us that the matrix semigroup generated by $A = (a_{ij})$ solves the initial value problem

$$\begin{cases}\dot{x}(t) = Ax(t), & t \geq 0, \\ x(0) = x_0\end{cases}$$

in the sense that the orbit $\{\mathrm{e}^{tA}x_0 : t \in \mathbb{R}\}$ of the initial value $x_0 \in \mathbb{C}^n$ is the unique solution of the problem. At present, we note that Theorem 4.1.2 remains true if t is allowed to run through $\mathbb{C}$, hence

$$z \mapsto \mathrm{e}^{zA}$$

is a holomorphic function on $\mathbb{C}$.

4.2 The coordinate functions

We intend to study the behavior of the function $t \mapsto \mathrm{e}^{tA}$ (or of $t \mapsto \mathrm{e}^{tA}x$ for a given $x \in X$), as $t \to \infty$, following the same pattern as we did in the previous lecture for the matrix powers. Nevertheless, a few comments seem to be appropriate.

While in Lecture 3 we studied the sequence $(T^k)_{k\in\mathbb{N}}$ and based our considerations on the characterization of convergence of the coordinate sequences given in Section 3.1, we will now have to deal with a function $t \mapsto \mathrm{e}^{tA}$ of the real variable t. We will formulate, without going into a detailed discussion, the following variation of the convergence properties discussed in Section 1.2.

If $t \mapsto y(t)$ is a real function with values in a finite dimensional vector space X with a basis $\{y_1, ..., y_n\}$, then $y(t) = \sum_{i=1}^{n} \eta_i(t) y_i$ with uniquely determined values of $\eta_i(t)$ for each $t \in \mathbb{R}$. We call the functions $t \mapsto \eta_i(t)$ the *coordinate functions* of $y(t)$ with respect to $\{y_1, ..., y_n\}$. Convergence of $y(t)$ as $t \to \infty$ in X is equivalent to the convergence of all coordinate functions $\eta_i(t)$, no matter what basis is employed, the coordinates of the limit being the limits of the respective coordinate functions.

In order to discuss the function $t \mapsto \mathrm{e}^{tA}$, in analogy to Lemma 2.2.9, we use a basis $\mathbb{B}_A$ of $M_n(\mathbb{C})$ containing the set

$$B_A := \left\{ \frac{(A - \lambda_i)^\nu}{\nu!} P_i : i = 1, \ldots, m;\ \nu = 0, \ldots, \nu_i - 1 \right\}.$$

By formula (4.1), the non–zero coordinate functions with respect to such a basis are

$$g_{\nu,\lambda_i}(t) := t^\nu \mathrm{e}^{t\lambda_i} \tag{4.5}$$

for $i = 1, ..., m$ and $\nu = 0, \ldots, \nu_i - 1$. Likewise, if we wish to study $\mathrm{e}^{tA}x$ for a given $x \in X$, we use a basis $\mathbb{B}_{A,x}$ of X containing the non–zero elements of

$$B_{A,x} := \left\{ \frac{(A - \lambda_i)^\nu x}{\nu!} : i = 1, \ldots, m;\ \nu = 0, \ldots, \nu_i - 1 \right\}.$$

Again, the coordinate functions of $\mathrm{e}^{tA}x$ with respect to such a basis are among the functions $g_{\nu,\lambda_i}(t)$ defined in (4.5).

The behavior of a function $g_{\nu,\lambda}(t) := t^\nu \mathrm{e}^{t\lambda}$ is easy to understand and essentially depends on the real part of λ. The following cases are possible.

- $\underline{\operatorname{Re}\lambda < 0}$. Then, for each fixed value of ν, $\mathrm{e}^{t\lambda} t^\nu \to 0$ as $t \to \infty$, the decay being exponential in the following sense:
  for any $0 < \delta < -\operatorname{Re}\lambda$ there is $M_\delta \geq 1$ such that
  $$\left|\mathrm{e}^{\lambda t} t^\nu\right| = t^\nu \mathrm{e}^{tRe\lambda} \leq M_\delta \mathrm{e}^{-\delta t} \quad \text{for all } t \geq 0.$$
- $\underline{\operatorname{Re}\lambda > 0}$. Then, for each fixed value of ν, $\left|\mathrm{e}^{t\lambda} t^\nu\right| \to \infty$ as $t \to \infty$, the increase being exponential (or exponentially bounded) in the following sense:
  for each $w > \operatorname{Re}\lambda > \delta > 0$ there is $M_w \geq 1$ such that
  $$\left|\mathrm{e}^{\lambda t} t^\nu\right| \leq M_w \mathrm{e}^{wt} \quad \text{for all } t \geq 0.$$
- $\underline{\operatorname{Re}\lambda = 0 \text{ and } \nu = 0}$. Then $\mathrm{e}^{t\lambda}$ is constant (for $\lambda = 0$) or periodic of period $\frac{2\pi i}{\lambda}$ (for $\lambda \neq 0$).
- $\underline{\operatorname{Re}\lambda = 0 \text{ and } \nu \geq 1}$. Then $\left|\mathrm{e}^{\lambda t} t^\nu\right| = t^\nu \to \infty$ as $t \to \infty$.

After these preparations, we now look at the behavior of $\mathrm{e}^{tA}x$ on the spectral subspaces $X_i = \operatorname{im} P_i$ of X.

Theorem 4.2.1. *Let $A \in \mathcal{L}(X)$, let $\|\cdot\|$ be a norm on X and fix an $i \in \{1, \dots, m\}$. Then the following assertions holds.*

(i) For every $\rho < \operatorname{Re}\lambda_i < w$ there exist $M \geq 1$ and $N > 0$ such that

$$N e^{\rho t}\|x\| \leq \left\|e^{tA}x\right\| \leq M e^{wt}\|x\|$$

for all $t \geq 0$ and all $x \in X_i$.

(ii) If $\operatorname{Re}\lambda_i = 0$, then

$$\left\{e^{tA}x : t \geq 0\right\}$$

is bounded for every $x \in X_i$ if and only if λ_i is a simple pole of $R(\cdot, A)$, i.e., if $\nu_i = 1$. In this case, $\mathrm{e}^{tA}x = \mathrm{e}^{t\lambda_i}x$ for every $x \in X_i$ and $t \geq 0$.

Proof. (i) By (4.1) we have for $x \in X_i$ that

$$\begin{aligned}\left\|\mathrm{e}^{tA}x\right\| &= \left\|\sum_{\nu=0}^{\nu_i-1} \mathrm{e}^{t\lambda_i}t^{\nu}\frac{(A-\lambda_i)^{\nu}}{\nu!}x\right\| \\ &\leq \sum_{\nu=0}^{\nu_i-1}\left\|\frac{(A-\lambda_i)^{\nu}}{\nu!}\right\|\left|\mathrm{e}^{t\lambda_i}t^{\nu}\right| \cdot \|x\| \\ &\leq M e^{\omega t}\|x\|\end{aligned}$$

for all $\omega > \operatorname{Re}\lambda_i$ and some $M \geq 1$.
Now, we apply (i) to $-A$ which has $-\lambda_i$ as eigenvalue with the same spectral projection P_i and spectral subspace X_i as before. Hence,

$$\left\|\mathrm{e}^{-tA}y\right\| \leq M\mathrm{e}^{-\rho t}\|y\|$$

for all $y \in X_i$, $t \geq 0$, and some $M \geq 1$. Since $\mathrm{e}^{-tA}\mathrm{e}^{tA} = I$, we find for every $x \in X_i$ an element $y \in X_i$ such that $x = \mathrm{e}^{-tA}y$. This implies

$$\left\|\mathrm{e}^{tA}x\right\| = \|y\| \geq \frac{1}{M}\mathrm{e}^{\rho t}\|x\|$$

for all $t \geq 0$.

(ii) If $\nu_i = 1$, then $\mathrm{e}^{tA}x = \mathrm{e}^{t\lambda_i}x$ for all $x \in X_i$. On the other hand, if $\nu_i \geq 2$, then $\ker(A - \lambda_i) \subsetneqq X_i$, hence there is an $x \in X_i$ with $(A - \lambda_i)x \neq 0$. The coordinate function of $(A - \lambda_i)x$ with respect to the basis element $x \in \mathbb{B}_{A,x}$ equals $te^{t\lambda_i}$ which is unbounded as $t \to \infty$. □

We now put all these information together to describe the action of e^{tA} on all of X. To that purpose it is convenient to introduce the following constant which plays the same role for the exponential function $t \mapsto \mathrm{e}^{tA}$ as the spectral radius $r(T)$ does for the powers $k \to T^k$ (see Section 3.1.1).

4.3 The spectral bound

Definition 4.3.1. For $A \in \mathcal{L}(X)$ the number

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

is called the *spectral bound* of A.

We note that the spectral bound of A can be determined from $\|\mathrm{e}^{tA}\|$ in the following way (compare with Proposition 3.1.3).

Proposition 4.3.2. *If $\|\cdot\|$ is any norm on $\mathcal{L}(X)$, then*

$$s(A) = \lim_{t\to\infty} \frac{1}{t} \log \|e^{tA}\|. \tag{4.6}$$

If $\|\cdot\|$ is an operator norm, then

$$s(A) = \inf_{t>0} \frac{1}{t} \log \|e^{tA}\|. \tag{4.7}$$

Proof. By equivalence of norms on $M_n(\mathbb{C})$ the limit

$$\lim_{t\to\infty} \frac{1}{t} \log \|\mathrm{e}^{tA}\|,$$

if it exists, does not depend on the specific norm. Hence, we can use the supremum norm $|\!|\!|\cdot|\!|\!|$ with respect to the basis $\mathbb{B}_A$ above. Then, we have

$$|\!|\!|\mathrm{e}^{tA}|\!|\!| = \left|t^\nu \mathrm{e}^{t\lambda_i}\right| = t^\nu \mathrm{e}^{t\operatorname{Re}\lambda_i}$$

for some $i \in \{1, ..., m\}$ and some $0 \le \nu \le n-1$. Hence,

$$\frac{1}{t} \log |\!|\!|\mathrm{e}^{tA}|\!|\!| = \frac{\nu}{t} \log t + \operatorname{Re}\lambda_i$$

for all $t > 0$ and i and ν as before. Since the function $t \mapsto \frac{\nu}{t}\log t$ tends to zero as $t \to \infty$, we obtain

$$\frac{\nu}{t} \log t + \operatorname{Re}\lambda_i \to \operatorname{Re}\lambda_i$$

as $t \to \infty$, and formula (4.6) follows.
The proof of formula (4.7) is left as an exercise. □

A repeated application of Theorem 4.2.1 yields the following Corollary.

Corollary 4.3.3. *Let $A \in \mathcal{L}(X)$ and let $\|\cdot\|$ be any norm on X. Then for every $w > \mathrm{s}(A)$ there is a constant $M \ge 1$ such that*

$$\left\|e^{tA}x\right\| \le M e^{wt}\|x\|$$

for all $t \geq 0$ and $x \in X$. Furthermore,

$$\mathrm{s}(A) = \omega_0$$

where

$$\omega_0 := \inf\big\{w \in \mathbb{R} : \exists\, M \geq 1 \text{ such that } \|e^{tA}\| \leq M e^{wt} \text{ for } t \geq 0\big\}. \tag{4.8}$$

Remark 4.3.4. The number ω_0 defined in (4.8) is known as the *growth bound* of the matrix semigroup $(\mathrm{e}^{tA})_{t\geq 0}$. Note the if X is an infinite dimensional vector space, the equality $s(A) = \omega_0$ remains valid only under certain additional assumptions on A.

4.4 Asymptotics

As in Section 3.2 we first define different kinds of long-time behavior of e^{tA}.

Definition 4.4.1. For $A \in \mathcal{L}(X)$ and any norm $\|\cdot\|$ on X we call $(\mathrm{e}^{tA})_{t\geq 0}$

- *bounded*[1], if $\sup_{t\geq 0} \|\mathrm{e}^{tA}\| < \infty$.
- *stable*, if $\lim_{t\to\infty} \|\mathrm{e}^{tA}\| = 0$.
- *exponentially stable*, if there exist $M \geq 1$ and $\varepsilon > 0$ such that

  $$\left\|\mathrm{e}^{tA}\right\| \leq M\mathrm{e}^{-\varepsilon t}$$

  for all $t \geq 0$.
- *convergent*, if $\lim_{t\to\infty} \mathrm{e}^{tA} = P$ for some $P \in \mathcal{L}(X)$.
- *periodic*, if $\mathrm{e}^{t_0 A} = I$ for some $t_0 > 0$.
- *hyperbolic*, if there exist A-invariant subspaces X_s and X_u, such that $X = X_s \oplus X_u$ and

  $$\begin{aligned} \left\|\mathrm{e}^{tA}x\right\| &\leq M\mathrm{e}^{-\varepsilon t}\|x\| \qquad &&\text{for } x \in X_s, \\ \left\|\mathrm{e}^{tA}x\right\| &\geq \frac{1}{M}\mathrm{e}^{\varepsilon t}\|x\| \qquad &&\text{for } x \in X_u, \end{aligned}$$

  for all $t \geq 0$ and some constants $M \geq 1$, $\varepsilon > 0$. The subspaces X_s and X_u are called *stable* and *unstable* subspace, respectively.

Remarks 4.4.2. Since pointwise and norm convergence on $\mathrm{M}_n(\mathbb{C})$ coincide, a statement about the long-time behavior of $\left\|\mathrm{e}^{tA}x\right\|$ for all $x \in X$ is equivalent to the same statement regarding $\left\|\mathrm{e}^{tA}\right\|$ for the appropriate operator norm.

[1]Note that working with ODEs or dynamical systems, a different terminology is also widely accepted: what we call "bounded" is often also called "stable", and what we call "stable" is also often called "asymptotically stable", and what we call "hyperbolic" is often called "exponential dichotomy".

Note that stability of a matrix semigroup is equivalent to exponential stability, see Exercise 5.

We now classify the asymptotic behavior of e^{tA} in terms of spectral properties of the matrix A. The following theorem can be proved analogously as Theorem 3.2.3 and is left as an exercise.

Theorem 4.4.3. *Let $A \in \mathcal{L}(X)$ and take any norm $\|\cdot\|$ on X.*

(i) $(e^{tA})_{t\geq 0}$ *is (exponentially) stable if and only if* $\mathrm{s}(A) < 0$.

(ii) $(e^{tA})_{t\geq 0}$ *is bounded if and only if* $\mathrm{s}(A) \leq 0$ *and all eigenvalues of A with real part equal to 0 are simple poles of the resolvent $R(\cdot, A)$.*

(iii) $(e^{tA})_{t\geq 0}$ *is periodic if and only if it is bounded and $\sigma(A) \subset 2\pi i\alpha\mathbb{Z}$ for some $\alpha \in \mathbb{R}$.*

(iv) $\lim_{t\to\infty} e^{tA} = P_1$ *(P_1 denotes the spectral projection of A belonging to the eigenvalue 0) if and only if* $\mathrm{s}(A) = 0$ *is a simple pole of the resolvent $R(\cdot, A)$ and $\sigma(A) \cap i\mathbb{R} = \{0\}$.*

(v) $(e^{tA})_{t\geq 0}$ *is hyperbolic if and only if $\sigma(A) \cap i\mathbb{R} = \emptyset$.*

Remark 4.4.4. Theorem 4.4.3 (i) is Liapunov's Stability Theorem proved in 1892.

Thus, in complete analogy to the situation in Section 3, convergence of e^{tA} as $t \to \infty$ is restricted to either one of the following situations.

- $\lim_{t\to\infty} \mathrm{e}^{tA} = 0$: this is the case if and only if $s(A) < 0$;
- $\lim_{t\to\infty} \mathrm{e}^{tA} = P_1$, where P_1 is the spectral projection belonging to $\lambda_1 = 0$: this is the case if and only if $s(A) = 0$, $\sigma(A) \cap i\mathbb{R} = \{0\}$, and 0 is a simple pole of the resolvent $R(\cdot, A)$.

Decomposing the space we can study stability concepts more in detail. One example of this approach is already the definition of hyperbolicity of a matrix semigroup. Let us use this approach to obtain another asymptotic property.

Definition 4.4.5. For $A \in \mathcal{L}(X)$ we call $(\mathrm{e}^{tA})_{t\geq 0}$ *asymptotically periodic* if there is a direct sum decomposition

$$X = X_0 \oplus X_1$$

into A-invariant subspaces X_0 and X_1 such that

(i) $\mathrm{e}^{tA}{}_{|X_0}$ is stable, i.e., $\lim_{t\to\infty} \mathrm{e}^{tA}x = 0$ for all $x \in X_0$, and

(ii) $\mathrm{e}^{tA}{}_{|X_1}$ is periodic, i.e., there exists $t_0 > 0$ such that $\mathrm{e}^{t_0 A}y = y$ for all $y \in X_1$.

Again, this property can be described by spectral properties of A.

Theorem 4.4.6. *For $A \in \mathcal{L}(X)$ the following assertions are equivalent.*

(a) $(e^{tA})_{t\geq 0}$ *is asymptotically periodic.*

(b) $(e^{tA})_{t\geq 0}$ *is bounded and* $\sigma(A) \cap i\mathbb{R} \subset 2\pi i\alpha\mathbb{Z}$ *for some* $\alpha \in \mathbb{R}$.

(c) $\mathrm{s}(A) \leq 0$, *the set* $\sigma(A) \cap i\mathbb{R}$ *consists of simple poles of the resolvent* $R(\cdot, A)$ *and is contained in* $2\pi i\alpha\mathbb{Z}$ *for some* $\alpha \in \mathbb{R}$.

Proof. (a)⇒(b). The boundedness of $(\mathrm{e}^{tA})_{t\geq 0}$ follows directly from (a). Let $X = X_0 \oplus X_1$ be the corresponding decomposition. Then

$$\sigma(A) = \sigma(A_{|X_0}) \cup \sigma(A_{|X_1}),$$

where $\sigma(A_{|X_0}) \subset \{\lambda \in \mathbb{C} : \mathrm{Re}\,\lambda < 0\}$ by Theorem 4.4.3 (i) and $\sigma(A_{|X_1}) \subset 2\pi i\alpha\mathbb{Z}$ for some $\alpha \in \mathbb{R}$ by Theorem 4.4.3 (iii).

(b)⇒(c) follows by Theorem 4.4.3 (ii).

Finally, to show (c)⇒(a) we define

$$X_0 := \bigoplus_{\mathrm{Re}\,\lambda_i < 0} X_i \qquad \text{and} \qquad X_1 := \bigoplus_{\mathrm{Re}\,\lambda_i = 0} X_i$$

and apply Theorem 4.4.3 (i) and (iii). □

We finally discuss the question under which conditions a subsequence or the Cesàro means of $(\mathrm{e}^{tA})_{t\geq 0}$ converge. First recall the definition of the spectral contraction in Definition 3.2.6.

Theorem 4.4.7. *Let* $A \in \mathcal{L}(X)$. *The following assertions are equivalent.*

(a) The operators e^{tA} *are spectral contractions for one/all* $t > 0$.

(b) $s(A) = 0$ *and all eigenvalues of* A *with real part equal to 0 are simple poles of the resolvent* $R(\lambda, A)$.

(c) There exists a sequence $(t_n)_{n\in\mathbb{N}}$ *of the form* $t_n := tk_n$, *where* $(k_n)_{n\in\mathbb{N}}$ *is a subsequence of* $(k)_{k\in\mathbb{N}}$, *such that* $(\mathrm{e}^{t_nA})_{n\in\mathbb{N}}$ *converges to some limit* $P \neq 0$ *for one/all* $t > 0$.

(d) There is an operator norm $|||\cdot|||$ *on* $\mathcal{L}(X)$ *such that* $\left|\left|\left|\mathrm{e}^{ktA}\right|\right|\right| = 1$ *for all* $k \in \mathbb{N}$ *and one/all* $t > 0$.

Proof. (a) $\iff$ (b) follows by Theorem 3.2.5 combined with the Spectral Mapping Theorem 2.3.4 and Theorem 2.3.12, (a) $\iff$ (c) again by Theorem 3.2.5, while (a) $\iff$ (d) holds by Theorem 3.2.7. □

Definition 4.4.8. We say that $(\mathrm{e}^{tA})_{t\geq 0}$ is a *spectral contraction semigroup*, if any of the equivalent assertions of Theorem 4.4.7 is true.

The following is the continuous analogue to the Cesàro means in Lecture 3.

Definition 4.4.9. Let $A \in \mathcal{L}(X)$. The matrices

$$C(r) := \frac{1}{r}\int_0^r \mathrm{e}^{sA} ds, \qquad r > 0,$$

are called the *Cesàro means* of the semigroup $(\mathrm{e}^{tA})_{t\geq 0}$. The semigroup $(\mathrm{e}^{tA})_{t\geq 0}$ is *mean ergodic* (or *Cesàro summable*), if $\lim_{r\to\infty} C(r)$ exists.

Theorem 4.4.10. *Let $A \in \mathcal{L}(X)$. The semigroup $(e^{tA})_{t\geq 0}$ is mean ergodic if and only if either $s(A) < 0$ or $(e^{tA})_{t\geq 0}$ is a spectral contraction semigroup. In the first case the Cesàro means $C(r)$ converge to 0 and in the second case to the spectral projection of A belonging to* 0.

Proof. First note that the coordinate functions of $C(r)$ with respect to $\mathbb{B}_A$ are of the form

$$g^{(r)}_{\nu,\lambda_i} := \frac{1}{r}\int_0^r g_{\nu,\lambda_i}(s)ds = \frac{1}{r}\int_0^r \mathrm{e}^{s\lambda_i} s^\nu ds. \tag{4.9}$$

Following the discussion on the page 42 we see that $g^{(r)}_{\nu,\lambda}$ converges only in two cases: either for $\operatorname{Re}\lambda < 0$ or for $\operatorname{Re}\lambda = 0 = \nu$. This proves the first assertion of the theorem.

Since for $\operatorname{Re}\lambda < 0$ we have $g^{(r)}_{\nu,\lambda} \to 0$ as $r \to \infty$, in the case when $s(A) < 0$ we obtain $C(r) \to 0$. On the other hand, in the case of spectral contraction semigroup, the only nonzero limit of the coordinate functions as $r \to \infty$ is $g^{(r)}_{0,0} = 1$. By (4.1), $C(r)$ then converges towards the spectral projection of A belonging to $\lambda = 0$. ☐

4.5 Exercises

1. Show that if $A, B \in \mathcal{L}(X)$ commute, then $\mathrm{e}^{t(A+B)} = \mathrm{e}^{tA}\mathrm{e}^{tB}$. Find an example to show that the commutativity assumption is necessary.

2. Let $B \in \mathcal{L}(X)$. Under which conditions is there an $A \in \mathcal{L}(X)$ such that $\mathrm{e}^{kA} = B^k$ for all $k \in \mathbb{N}$?

3. Prove formula (4.7).

4. Prove Theorem 4.4.3 (Hint: follow the proof of Theorem 3.2.3 and use Theorem 4.2.1).

5. Prove that for $A \in \mathcal{L}(X)$ the matrix semigroup e^{tA} is stable if and only if it is exponentially stable.

6. Show that $\{\mathrm{e}^{tA}x : t \in \mathbb{R}\}$ is bounded for every $x \in X$ if and only if $\sigma(A) \subset i\mathbb{R}$ and all eigenvalues are simple poles of the resolvent $R(\cdot, A)$.

7. Show that the semigroup $(\mathrm{e}^{tA})_{t\geq 0}$ is hyperbolic if and only if $\sigma\left(\mathrm{e}^{tA}\right) \cap \Gamma = \emptyset$ for some/all $t > 0$, where Γ denotes the unit circle in $\mathbb{C}$.

8. Compute the matrix exponential e^{tA} for

$$A = \begin{pmatrix} -a & b \\ a & -b \end{pmatrix} \text{ with } a + b \neq 0.$$

9. For every one of the following matrices A compute e^{tA} and discuss its asymptotic behavior.

$$(a) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad (b) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad (c) \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \qquad (d) \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \qquad (e) \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}.$$

Lecture 5

Positive Matrices

5.1 Positivity

We start by introducing the relevant notation and terminology.

The inequality $x \leq y$ for vectors $x = (\xi_1, \ldots, \xi_n), y = (\eta_1, \ldots, \eta_n) \in \mathbb{R}^n$ means that $\xi_i \leq \eta_i$ for all i. Similarly, for real matrices $T = (\tau_{ij})$ and $S = (\sigma_{ij})$, the notation $T \leq S$ means $\tau_{ij} \leq \sigma_{ij}$ for all i, j. As usual, the symbol "$<$" means "less but not equal", i.e., $x < y$ if $\xi_i \leq \eta_i$ for all i and there is an index j such that $\xi_j < \eta_j$.

Definition 5.1.1. A vector $x = (\xi_1, \ldots, \xi_n)$ (a matrix $T = (\tau_{ij})$) is called *positive*, if for all its coordinates it holds: $\xi_i \geq 0$ (for all its entries it holds: $\tau_{ij} \geq 0$). In this case, we write $x \geq 0$ ($T \geq 0$).

We point out that, in our terminology, a positive vector does not need to have *all* coordinates > 0 and a positive matrix does not need not to have *all* entries > 0. Likewise, a vector (matrix) which is > 0 may have many coordinates (entries) equal to 0, but at least one coordinate (entry) must be > 0.

Definition 5.1.2. A vector $x = (\xi_1, \ldots, \xi_n)$ (a matrix $T = (\tau_{ij})$) is called *strictly positive*,[1] if for all its coordinates it holds: $\xi_i > 0$ (for all its entries it holds: $\tau_{ij} > 0$). In this case, we write $x \gg 0$ ($T \gg 0$).

By the *absolute value* of a vector $x = (\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$ we mean the vector

$$|x| := \big(|\xi_1|, \ldots, |\xi_n|\big).$$

Similarly, for a matrix $T = (\tau_{ij}) \in \mathrm{M}_n(\mathbb{C})$ we call

$$|T| := \big(|\tau_{ij}|\big)$$

the *absolute value* of T.

The verification of the following lemma is quite straightforward, see Exercise 1.

[1] There is a different terminology that has its followers, calling vectors or matrices that are positive in our sense *non-negative* and reserving the term positive for what we call strictly positive. This may of course lead to misunderstandings and confusion but the coexistence of both terminologies is a fact.

Lemma 5.1.3. *Let $T \in M_n(\mathbb{C})$. Then the following properties hold.*

(i) $T \geq 0$ *if and only if* $Tx \geq 0$ *for all* $x \geq 0$.

(ii) $|Tx| \leq |T|\,|x|$, *hence* $|Tx| \leq T\,|x|$ *if* $T \geq 0$.

In the following, we will always use the maximum norm on $\mathbb{C}^n$,

$$\|x\| := \|x\|_\infty = \max_{1\leq i\leq n} |\xi_i|,$$

while on $\mathrm{M}_n(\mathbb{C})$ the corresponding operator norm (see Example 1.2.8) is used which is given as

$$\|T\| := \|T\|_\infty = \max_{\|x\|\leq 1} \|Tx\| = \max_{1\leq i\leq n} \sum_{j=1}^{n} |\tau_{ij}|.$$

The reason for this special choice becomes clear from the following lemma which is also easily verified, see Exercise 1.

Lemma 5.1.4. *(i) For* $x, y \in \mathbb{C}^n$ *the inequality* $|x| \leq |y|$ *implies* $\|x\| \leq \|y\|$, *in particular* $\|\,|x|\,\| = \|x\|$ *for all* x.

(ii) For $S, T \in M_n(\mathbb{C})$ *the inequality* $|S| \leq |T|$ *implies* $\|S\| \leq \|T\|$. *In particular,*

$$\|\,|T|\,\| = \|T\|$$

for all T, *and* $|S| \leq T$ *implies* $\|S\| \leq \|T\|$.

(iii) If $T \geq 0$, *then*

$$\|T\| = \|T\mathbb{1}\|,$$

where $\mathbb{1} := (1, 1, \ldots, 1)$.

The following observation on the resolvent has striking consequences.

Proposition 5.1.5. *Let* T *be a positive matrix with spectral radius* $\mathrm{r}(T)$.

(i) The resolvent $R(\mu, T)$ *is positive whenever* $\mu > \mathrm{r}(T)$.

(ii) If $|\mu| > \mathrm{r}(T)$, *then*

$$\big|R(\mu, T)\big| \leq R\big(|\mu|, T\big).$$

Proof. We use the Neumann series representation

$$R(\mu, T) = \sum_{k=0}^{\infty} \frac{T^k}{\mu^{k+1}} \tag{5.1}$$

for the resolvent, which is valid for $|\mu| > \mathrm{r}(T)$ by Corollary 2.2.8.

(i) If $T \geq 0$, then $T^k \geq 0$ for all k, hence for $\mu > \mathrm{r}(T)$, we have

$$R(\mu, T) = \lim_{N\to\infty} \sum_{k=0}^{N} \frac{T^k}{\mu^{k+1}} \geq 0$$

since the finite sums are positive and convergence holds in every entry.

(ii) We have for $|\mu| > r(T)$ that

$$\begin{aligned}\big|R(\mu, T)\big| &= \left|\lim_{N\to\infty} \sum_{k=0}^{N} \frac{T^k}{\mu^{k+1}}\right| \leq \lim_{N\to\infty} \sum_{k=0}^{N} \left|\frac{T^k}{\mu^{k+1}}\right| \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} \frac{T^k}{|\mu|^{k+1}} = R\big(|\mu|, T\big).\end{aligned}$$

□

The following result, known as the fundamental property of positive matrices, was discovered by O. Perron in 1907 and can be considered as the first major result in the theory of positive matrices.

Theorem 5.1.6. *If T is a positive matrix, then $\mathrm{r}(T)$ is an eigenvalue of T with positive eigenvector.*

Proof. Assertion (ii) of Proposition 5.1.5 and Lemma 5.1.4 imply

$$\big\|R(\mu, T)\big\| \leq \big\|R(|\mu|, T)\big\| \qquad \text{for } |\mu| > \mathrm{r}(T).$$

Recall the important formula for the resolvent proved in Theorem 2.2.5:

$$R(\mu, T) = (\mu - T)^{-1} = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \frac{(T-\lambda_i)^\nu}{(\mu - \lambda_i)^{\nu+1}} P_i \text{ for } \mu \notin \{\lambda_1, \dots, \lambda_m\}. \tag{5.2}$$

Here λ_i are the eigenvalues of T with respective multiplicities of the minimal polynomial ν_i and spectral projections P_i, $i = 1, \dots, m$.

Let now $\lambda_j \in \sigma(T)$ such that $|\lambda_j| = \mathrm{r}(T)$. Then $\|R(\mu, T)\| \to \infty$ whenever μ approaches λ_j. This is obtained immediately from (5.2) by looking at the sup-norm on $\mathrm{M}_n(\mathbb{C})$ with respect to a basis containing the set B_T in the sense of Lemma 2.2.9 (ii). Putting $\mu = s\lambda_j$ with $s > 1$ the above estimate yields

$$\big\|R(s\,\mathrm{r}(T), T)\big\| \geq \big\|R(s\lambda_j, T)\big\| \to \infty \quad \text{as } s \searrow 1,$$

hence $\mathrm{r}(T)$ must be an eigenvalue of T.

Finally, again by (5.2), we have

$$\lim_{\mu \downarrow \mathrm{r}(T)} R(\mu, T)\big(\mu - \mathrm{r}(T)\big)^{\nu_1} = \big(T - \mathrm{r}(T)\big)^{\nu_1 - 1} P_1,$$

where P_1 denotes the spectral projection corresponding to $\lambda_1 = \mathrm{r}(T)$ and ν_1 is the pole order of $R(\cdot, T)$ at $\mathrm{r}(T)$. Hence $(T - \mathrm{r}(T))^{\nu_1 - 1} P_1 \geq 0$ by Proposition 5.1.5 (i). Since $(T - \mathrm{r}(T))^{\nu_1 - 1} P_1 \neq 0$, there must be a *positive* vector y_1 with

$$x_1 := (T - \lambda_1)^{\nu_1 - 1} P_1 \, y_1 \neq 0.$$

Any such x_1 is a positive eigenvector of T belonging to $\mathrm{r}(T)$. □

Originally, Perron studied strictly positive matrices and thus obtained stronger results. Under the assumption on strict positivity of T he proved for $r = \mathrm{r}(T)$ that $r > 0$, it is a first order pole of the resolvent $R(\cdot, T)$, and the only eigenvalue of T of modulus r. The corresponding eigenspace is one-dimensional and spanned by a strictly positive vector.

However, as already Frobenius noticed, it is not the nonexistence of zero entries in a given matrix but the positions of them, that implies all these nice spectral properties. Frobenius defined the term irreducibility of a matrix, which we discuss in the next section.

5.2 Irreducibility

We now turn to the question under which conditions on a positive matrix T (other than strict positivity) the spectral radius $\mathrm{r}(T)$ is a first order pole of the resolvent, a property with important consequences for the behavior of the powers T^k as $k \to \infty$. The main consequences we have in mind will, however, be discussed in the next lecture and concern the situation where $T = \mathrm{e}^{tA}$.

The following property of T, again relatively easy to recognize from the matrix entries, will turn out to be sufficient. From now on we assume that $n > 1$.

Definition 5.2.1. A matrix $T \in \mathrm{M}_n(\mathbb{C})$ is called *reducible* if there exists a subspace

$$J_M := \big\{(\xi_1, \ldots, \xi_n) : \xi_i = 0 \text{ for } i \in M\big\} \subset \mathbb{C}^n$$

for some $\emptyset \neq M \subsetneqq \{1, \ldots, n\}$ which is invariant under T. If T is not reducible, it is called *irreducible*.

It is important to note that arbitrary coordinate transformations may destroy or produce irreducibility of a given matrix. However, a permutation of the canonical basis vectors of $\mathbb{C}^n$ does not touch it. So, T is reducible if and only if, after a reordering of the canonical basis vectors of $\mathbb{C}^n$, there is $1 \leq k < n$ such that

$$J_{M_k} := \big\{(\xi_1, \ldots, \xi_n) : \xi_1 = \cdots = \xi_k = 0\big\} \tag{5.3}$$

is invariant under T.

This leads to the following characterization, which can be applied easily to concrete matrices.

Lemma 5.2.2. *A matrix $T \in M_n(\mathbb{C})$ is reducible if and only if there exists a permutation matrix P such that*

$$S := PTP^{-1}$$

has block-triangular form

$$S = \begin{pmatrix} A & 0 \\ B & C \end{pmatrix}$$

for quadratic matrices A and C.

We will apply the notion of irreducibility mainly to positive matrices and list now the most important examples.

Examples 5.2.3. (i) A matrix $T = (\tau_{ij})_{n\times n}$ with all off-diagonal entries $\tau_{ij} > 0$ $(i \neq j)$ is irreducible (see Exercise 5).

(ii) If $T \geq 0$ is irreducible and $T \leq S$, then S is irreducible.

(iii) The permutation matrix

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}$$

is irreducible (see Exercise 5).

The following result shows that every matrix can, in a certain way, be decomposed into irreducible blocks.

Proposition 5.2.4. *For every matrix $T \in M_n(\mathbb{C})$ there exists a permutation matrix P such that $S := PTP^{-1}$ has block-triangular form*

$$\begin{pmatrix} T_{11} & * & \dots & * \\ & \ddots & \ddots & \vdots \\ 0 & & \ddots & * \\ & & & T_{mm} \end{pmatrix},$$

where the (square)diagonal blocks T_{ii} are all irreducible.

Proof. We prove the result by induction on n. The case $n = 1$ is clear. If $n > 1$, let us suppose that the result holds for matrices of size $\leq n-1$. If T is irreducible, there is nothing to prove. If T is reducible, a reordering of the canonical basis produces the form

$$\begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix},$$

where T_{11} is a $k \times k$ block for suitable $1 \leq k < n$. Since $k < n$, we can, by assumption, rearrange the first k basis vectors in such a way that T_{11} has block triangular form with irreducible diagonal blocks. Since T_{22} can be treated in the same way, we obtain the assertion. □

The following is our main result on *irreducible positive* matrices and was proved by F.G. Frobenius in 1912.

Theorem 5.2.5. *Let $T \in M_n(\mathbb{C})$ be an irreducible and positive matrix. Then the spectral radius $r := \mathrm{r}(T)$ satisfies $r > 0$, and r is a first order pole of the resolvent $R(\cdot, T)$. The corresponding eigenspace is one-dimensional and spanned by a strictly positive vector $z = (\zeta_1, \dots, \zeta_n)$, i.e., with $\zeta_i > 0$ for all i.*

Proof. By Theorem 5.1.6 there exists $0 < z = (\zeta_1, \dots, \zeta_n)$ such that $Tz = rz$. Suppose now that z is not strictly positive. After a reordering of the coordinates we may assume $\zeta_i = 0$ for $i = 1, \dots, k$ and $\zeta_i > 0$ for $i = k+1, \dots, n$. This implies that for every $y \in J_{M_k}$ (see (5.3)) there is a suitable $c > 0$ such that $|y| \leq c \cdot z$ holds. Thus

$$|Ty| \leq T|y| \leq c\,Tz = cr \cdot z,$$

which shows that $Ty \in J_{M_k}$, i.e., J_{M_k} is T-invariant. Since T was supposed to be irreducible, we must have $J_{M_k} = \{0\}$ which is a contradiction. Therefore, z must be strictly positive.

For the next step assume $r = 0$, hence $Tz = 0$ for a strictly positive vector z. As before, we conclude that

$$|Ty| = 0$$

for all $y \in \mathbb{C}^n$. Thus $T = 0$ which is not irreducible.

We now show that the eigenspace belonging to r is one-dimensional. Let

$$Ty = ry$$

for some $0 \neq y \in \mathbb{C}^n$. Since T is a positive, hence a real matrix, it follows that the real and imaginary parts of y are eigenvectors belonging to r. Therefore, we can assume $0 \neq y \in \mathbb{R}^n$. Since the eigenvector z found above is strictly positive, there exists a $c \in \mathbb{R}$ such that

$$x := z - cy$$

is positive, but not strictly positive. Since $Tx = rx$, it follows that the subspace J_M corresponding to the zero-coordinates of x is invariant under T, hence must be $\{0\}$. This implies $z = cy$.

Finally, we determine the pole order of r. Since $r > 0$, we may, after rescaling, take $r = 1$, hence

$$Tz = z$$

for some strictly positive $z = (\zeta_1, \dots, \zeta_n)$. Define $D_z := \mathrm{diag}(\zeta_1, \dots, \zeta_n)$ and

$$S := D_z^{-1} T D_z.$$

Then $S \geq 0$ and $S\mathbb{1} = \mathbb{1} = (1, ..., 1)$, hence $\|S\| = 1$ by Lemma 5.1.4(iii). This implies $\|S^k\| \leq 1$ and

$$\left\|T^k\right\| = \left\|D_z S^k D_z^{-1}\right\| \leq \|D_z\| \cdot \left\|D_z^{-1}\right\|$$

for all $k \in \mathbb{N}$, hence $(T^k)_{k\in\mathbb{N}}$ is bounded. By Theorem 3.2.3(ii) is then 1 a simple pole of the resolvent. □

The strictly positive vector z appearing in Theorem 5.2.5 is called the *Perron vector* for T and is unique up to multiplication by positive scalars (assuming $\|z\| = 1$, for example, we have uniqueness). For any positive irreducible matrix T also its transpose $T^\top$ has these properties and thus a strictly positive Perron eigenvector, that spans the one-dimensional eigenspace corresponding to $r(T) = r(T^\top)$ (see also Exercise 7).

We can now immediately use this result to describe the long-term behavior of the powers of a positive irreducible matrix.

Corollary 5.2.6. *Let T be a positive irreducible matrix with $r(T) = 1$ radially dominant. Then*

$$\lim_{k\to\infty} T^k = P_1 \gg 0,$$

where P_1 denotes the spectral projection belonging to the eigenvalue 1. Moreover P_1 is of the form

$$P_1 x = \langle x, y \rangle\, z, \quad x \in \mathbb{C}^n,$$

where $z \gg 0$ and $y \gg 0$ are the Perron vectors for T and $T^\top$, respectively, such that $\langle z, y \rangle = 1$.

Proof. The convergence of the powers T^k to the spectral projection P_1 is an immediate consequence of Theorems 3.2.3 and 5.2.5.

Let now $z \gg 0$ and $y \gg 0$ be the respective Perron vectors for T and $T^\top$ such that $\langle z, y \rangle = 1$ and define $Px := \langle x, y \rangle\, z$ for all $x \in \mathbb{C}^n$. It is clear that $P^2 = P \gg 0$ and $\operatorname{rg} P = \ker(T - I)$ is one-dimensional. Since

$$\langle Tx - x, y \rangle\, z = \left\langle x, T^\top y \right\rangle z - \langle x, y \rangle\, z = 0, \quad x \in \mathbb{C}^n,$$

we have $\operatorname{rg}(T - I) \subseteq \ker P$. By above, the dimensions of both subspaces equal $n - 1$, hence they are equal. This means that P is a projection to $\ker(T - I)$ along $\operatorname{rg}(T - I)$, therefore $P = P_1$. □

We showed almost the same properties for irreducible positive matrices as Perron has obtained for strictly positive matrices. However, we were not able to show that $r = r(T)$ is the only eigenvalue of modulus r.

5.3 Imprimitivity

The number of eigenvalues on the spectral circle has interesting impact on the asymptotic behavior of T^k for a positive irreducible matrix T.

Definition 5.3.1. *The boundary spectrum* of a matrix T with spectral radius $r = r(T)$ is the set

$$\sigma_b(T) := \{\lambda \in \mathbb{C}\colon |\lambda| = r\} \cap \sigma(T).$$

A positive irreducible matrix T with $\sigma_b(T) = \{r\}$ is called a *primitive matrix*. If a positive irreducible matrix has exactly $h > 1$ eigenvalues in the set $\sigma_b(T)$, it is called *imprimitive matrix* and h is referred to as *index of imprimitivity*.

A primitive matrix can be characterized by the following asymptotic behavior.

Proposition 5.3.2. *For a positive irreducible matrix T with spectral radius $r = r(T)$ the following assertions are equivalent.*

(i) T is primitive.

(ii) $\lim_{k\to\infty}(T/r)^k = P_1$ where P_1 denotes the spectral projection belonging to the eigenvalue r.

Moreover, in case any of the above assertions holds,

$$P_1 x = \langle x, y\rangle\, z, \quad x \in \mathbb{C}^n,$$

where $z \gg 0$ and $y \gg 0$ are the Perron vectors for T and $T^\top$, respectively, such that $\langle z, y\rangle = 1$.

Proof. Observe that T is primitive if and only if T/r is primitive, which is true if and only if $1 = r(T/r)$ is radially dominant. Now the implication from (i) to (ii) follows by Corollary 5.2.6 and the converse by Theorems 3.2.3 and 5.2.5. □

Next result was proved by Wielandt in 1950.

Lemma 5.3.3. *Let S be any complex matrix and T a positive irreducible matrix such that $|S| \le T$. Let $r = \mathrm{r}(T)$. Then for any $\lambda \in \sigma(S)$ we have $|\lambda| \le r$. Moreover,*

$$|\lambda| = r \quad \text{if and only if} \quad S = e^{i\varphi} D T D^{-1},$$

where $e^{i\varphi} = \lambda/r$ and D is a diagonal matrix with $|D| = I$. If we set $d_{11} = 1$, the matrix $D = \mathrm{diag}(d_{ii})$ is uniquely determined.

Proof. From Exercise 3.41 we know that $r(S) \le r(T)$.

Assume that $|\lambda| = r$ and let x be an eigenvector of S corresponding to the eigenvalue λ, i.e. $Sx = \lambda x$. Then

$$r|x| = |\lambda||x| = |\lambda x| = |Sx| \le |S||x| \le T|x|.$$

By Exercise 9 we have $T|x| = r|x|$, $|x| \gg 0$, and by above also $|S||x| = r|x|$, therefore $(T - |S|)|x| = 0$. Since $T - |S| \geq 0$ and $|x| \gg 0$, it follows $T = |S|$.

Let now $\lambda = |\lambda|\mathrm{e}^{i\varphi} = r\mathrm{e}^{i\varphi}$ and $x_k = |x_k|e^{i\theta_k}$ for some $\varphi, \theta_1, \ldots, \theta_n \in \mathbb{R}$ and define $D := \operatorname{diag}(\mathrm{e}^{i\theta_1} \ldots, \mathrm{e}^{i\theta_n})$. Then $x = D|x|$ and taking $V := \mathrm{e}^{-i\varphi}D^{-1}SD$ we have

$$V|x| = r|x| = T|x|.$$

Note that $|V| = |S| = T$, thus we obtain $(V - |V|)|x| = 0$. Taking only the real part of this equation and noting that $|V| \geq \operatorname{Re}(V)$ and $|x| \gg 0$ it follows that $\operatorname{Re}(V) = |V|$, which further implies $V = \operatorname{Re}(V) = |V| = T$, i.e. $S = \mathrm{e}^{i\varphi}DTD^{-1}$.

The converse implication obviously holds. □

Using Wielandt's Lemma we see that the eigenvalues on the spectral boundary of an imprimitive matrix are exactly the hth roots of the spectral radius. The following can be regarded as a continuation of Perron-Frobenius theorem 5.2.5.

Theorem 5.3.4. *Let T be an imprimitive matrix with index of imprimitivity h and with spectral radius $r = r(T)$. Then the following holds.*

(i) All eigenvalues of T of modulus r are simple poles of the resolvent and the corresponding eigenspaces are one-dimensional.

(ii) $\sigma_b(T) = \{r, r\omega, r\omega^2, \ldots, r\omega^{h-1}\}$, where $\omega = e^{2\pi i/h}$.

(iii) The whole spectrum $\sigma(T)$ is invariant under rotation about the origin through an angle $2\pi/h$, but not through any other positive smaller angle.

(iv) There exists a permutation matrix P such that

$$PTP^\top = \begin{pmatrix} 0 & T_{12} & 0 & \ldots & 0 \\ 0 & 0 & T_{23} & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & T_{h-1,h} \\ T_{h1} & 0 & \ldots & 0 & 0 \end{pmatrix}, \tag{5.4}$$

where the blocks on the main diagonal are square.

Proof. (i): Let $\sigma_b(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_h\}$ where $\lambda_k = r\mathrm{e}^{i\varphi_k}$, $k = 1, \ldots, h$. Wielandt's Lemma 5.3.3 with $S = T$ and $\lambda = \lambda_k$ yields

$$T = \mathrm{e}^{i\varphi_k} D_k T D_k^{-1}, \quad k = 1, \ldots, h, \tag{5.5}$$

showing that T and $\mathrm{e}^{i\varphi_k}T$ are similar. Let $Tz = rz$ where $z \gg 0$ is the Perron vector for T. Then for $z_k := D_k z$ we have $Tz_k = \lambda_k z_k$ and z_k is an eigenvector corresponding to the simple eigenvalue λ_k which is unique up to multiplication by scalars. This proves (i).

(ii): By (5.5) we also have

$$\begin{aligned} T &= \mathrm{e}^{i\varphi_{k_1}} D_{k_1} T D_{k_1}^{-1} = \mathrm{e}^{i\varphi_{k_1}} D_{k_1} \left(\mathrm{e}^{i\varphi_{k_2}} D_{k_2} T D_{k_2}^{-1}\right) D_{k_1}^{-1} \\ &= \mathrm{e}^{i(\varphi_{k_1}+\varphi_{k_2})} \left(D_{k_1} D_{k_2}\right) T \left(D_{k_1} D_{k_2}\right)^{-1}. \end{aligned} \tag{5.6}$$

Consequently, we obtain $\mathrm{e}^{i(\varphi_{k_1}+\varphi_{k_2})} \in \sigma_b(T)$ for any pair $k_1, k_2 \in \{1, \ldots, h\}$. Thus $\sigma_b(T)$ is a multiplicative abelian group of order h, yielding (ii).

(iii): Now let $\sigma(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and note that multilplying by $\omega = \mathrm{e}^{2\pi i/h}$ we have

$$\sigma(\omega T) = \{\omega\lambda_1, \omega\lambda_2, \ldots, \omega\lambda_n\}.$$

As above we obtain that ωT and T are similar, hence $\sigma(\omega T) = \sigma(T)$. On the other hand no rotation less then $2\pi/h$ keeps $\sigma_b(T)$ invariant, therefore the same holds for the whole spectrum $\sigma(T)$.

(iv): By (ii) the eigenvalues on the boundary are of the form $\lambda_k = r\omega^k$, $k = 0, \ldots, h-1$. For the matrices D_k from (5.5) and the Perron eigenvector z of T it follows from (5.6) that $D_{k_1} D_{k_2} z$ is an eigenvector of T corresponding to the eigenvalue $r\omega^{k_1+k_2}$. We may assume that the upper left entry of each diagonal matrix D_k is 1 and are therefore uniquely dertermined. Hence also the diagonal matrices D_k form a multiplicative abelian group of order h. In particular, $D_1^h = I_n$, so its main diagonal consists of hth roots of unity.

Let P be a permutation matrix such that

$$P D_1 P^\top = \operatorname{diag}(\omega^{m_1} I_{n_1}, \omega^{m_2} I_{n_2}, \ldots, \omega^{m_s} I_{n_s}),$$

where I_{n_j} are identity matrices of size $n_j \times n_j$, $\sum_{j=1}^s n_j = n$ and $0 = m_1 < m_2 < \cdots < m_s \le h-1$. Using the same permutation matrix P we obtain the block matrix

$$P T P^\top = \begin{pmatrix} T_{11} & T_{12} & \ldots & T_{1s} \\ T_{21} & T_{22} & \ldots & T_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ T_{s1} & T_{s2} & \ldots & T_{ss} \end{pmatrix},$$

where each block T_{pq} is of size $n_p \times n_q$. Now, equating the (p,q)-blocks on the both sides of the matrix equation

$$P T P^\top = \omega (P D_1 P^\top)(P T P^\top)(P D_1^{-1} P^\top)$$

we obtain a system of s^2 equations

$$T_{pq} = \omega^{1+m_p-m_q} T_{pq}, \quad p, q = 1, \ldots, s.$$

Therefore $T_{pq} \neq 0$ if and only if

$$m_q = m_p + 1 \mod h. \tag{5.7}$$

T is an irreducible matrix, hence for every p there is a q such that $m_q = m_p + 1 \mod h$. Since m_i, $i = 1, \ldots, s$ are strictly ordered numbers from the set $\{0, \ldots, h-1\}$, the only possibility is that $s = h$ and $m_k = k-1$, $k = 1, \ldots, h$. So the block matrix $PTP^\top$ has exactly h nonzero blocks: $T_{pq} \neq 0$ iff $q = p + 1 \mod h$. $\square$

We conclude with a description of asymptotic behavior of imprimitive matrices.

Definition 5.3.5. Let $T \in \mathcal{L}(X)$. We call $(T^k)_{k\in\mathbb{N}}$ *asymptotically periodic with period* p if there is a direct sum decomposition

$$X = X_0 \oplus X_1$$

into T-invariant subspaces X_0 and X_1 such that

(i) $T_{|X_0}$ is stable, i.e., $\lim_{k\to\infty} \|T^k x\| = 0$ for all $x \in X_0$, and

(ii) $T_{|X_1}$ is periodic with period p, i.e., $T^p y = y$ for all $y \in X_1$ and $p \in \mathbb{N}$ is the smallest natural number with this property[2].

Corollary 5.3.6. *Let T be an imprimitive matrix with index of imprimitivity h and with spectral radius $r = r(T)$. Then the sequence $\left((T/r)^k\right)_{k\in\mathbb{N}}$ is asymptotically periodic with period h.*

Proof. Let

$$X_0 := \bigoplus_{|\lambda_i|<r} X_i \qquad \text{and} \qquad X_1 := \bigoplus_{|\lambda_i|=r} X_i$$

and use Theorems 3.2.3 and 5.3.4 for the matrix T/r restricted to its invariant subspaces X_0 and X_1. □

5.4 Notes and Remarks

Theorem 5.1.6 can be found in the paper by O. Perron [2], which was the beginning of the general theory of positive matrices (or "matrices with non-negative entries", as Frobenius used to call them). Theorem 5.2.5 goes back to F.G. Frobenius [1]. Almost 40 years later, H. Wielandt [3] proved Theorem 5.3.3, bringing up a connection to Tübingen mathematical school.

The form of an imprimitive matrix presented in Theorem 5.3.4 was originally shown by Frobenius [1] and is known as the *Frobenius Form.*

5.5 Exercises

1. Prove Lemmas 5.1.3 and 5.1.4.

2. For the matrix

$$T = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix} \text{ where } a, b > 0 \text{ and } a + b = 1,$$

verify that $r(T)$ is an eigenvalue of T, and find the appropriate eigenvector.

[2]Note that we updated the definition of periodicity from Lecture 3: T is called *periodic*, if $T^p = I$ for some $p \in \mathbb{N}$.

3. Let $T \geq 0$. Prove the existence of a positive eigenvector x_0 belonging to the eigenvalue $r := \mathrm{r}(T)$ through the following steps:

   (i) There exists $y \geq 0$ such that
   $$\left\|R(r + \tfrac{1}{n}, T)y\right\| \to \infty \quad \text{as} \quad n \to \infty.$$

   (ii) The sequence
   $$y_n := \frac{R(r + \frac{1}{n}, T)y}{\left\|R(r + \frac{1}{n}, T)y\right\|}$$
   has a convergent subsequence $(y_{n_k})_{k \in \mathbb{N}}$.

   (iii) The sequence $(r - T)y_{n_k} \to 0$ as $k \to \infty$.

   (iv) The limit $x_0 := \lim_{k\to\infty} y_{n_k}$ is a positive eigenvector of T belonging to r.

4. Let $T \geq 0$. Then for $\mu \in \rho(T)$
   $$R(\mu, T) \geq 0 \quad \text{implies} \quad \mu > \mathrm{r}(T).$$
   (Hint: Let $x \geq 0$ be an eigenvector of T belonging to $\mathrm{r}(T)$. Then $R(\mu, T)x = (\mu - \mathrm{r}(T))^{-1}x$).

5. (i) A matrix $T = (\tau_{ij})_{n\times n}$ with all off-diagonal entries $\tau_{ij} \neq 0$ $(i \neq j)$ is irreducible.

   (ii) If $a_1, \dots, a_n \in \mathbb{C}$ are all non-zero, then
   $$\begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ 0 & 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & a_{n-1} \\ a_n & 0 & \dots & 0 & 0 \end{pmatrix}$$
   is irreducible.

6. Prove that for a positive matrix T the following conditions are equivalent.

   (i) T is irreducible.

   (ii) $R(\mu, T)x \gg 0$ for some $\mu > \mathrm{r}(T)$ and all $x > 0$.

   (iii) $R(\mu, T)x \gg 0$ for all $\mu > \mathrm{r}(T)$ and all $x > 0$.

7. Show that $T \geq 0$ is irreducible if and only if the eigenspaces of T and of $T^\top$ belonging to $\mathrm{r}(T) = \mathrm{r}(T^\top)$ are one dimensional and spanned by a strictly positive vector.

8. If $0 \leq T < S$ and T is irreducible, then $\mathrm{r}(T) < \mathrm{r}(S)$. In other words: The spectral radius is a strictly monotone function on the set of irreducible positive matrices.

9. Suppose that T is a positive irreducible matrix with $r = r(T)$ and $rx \leq Tx$ for some nonzero $x \geq 0$. Prove that then $Tx = rx$ and $x \gg 0$.

10. Let T be a positive irreducible matrix. Prove the following statements.

    (i) If the trace $\operatorname{tr} T > 0$, then T is primitive.

    (ii) T is primitive if and only if $T^m \gg 0$ for some $m \in \mathbb{N}$.

11. Verify irreducibility and imprimitivity of given positive matrix T and discuss the asymptotic behavior of the sequence $\left((T/r)^k\right)_{k\in\mathbb{N}}$.

$$(a) \quad T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad (b) \quad T = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Bibliography

[1] G. Frobenius, *Über Matrizen aus nicht negativen Elementen*, Sitzungsber. Preuß. Akad. Wiss. Phys.-Math. Kl. (1909), 456–477.

[2] O. Perron, *Zur Theorie der Matrizen*, Math. Ann. **64** (1907), 248–263.

[3] H. Wielandt, *Unzerlegbare, nicht negative matrizen*, Math. Z. **52** (1950), 642–648.

Lecture 6

Applications of Positive Matrices

6.1 Motivating examples revisited

We start by revisiting our motivating examples from Section 1.1 (apart from the Fibonacci sequence example, that was already revisited in Lecture 3).

6.1.1 Graphs

Let $G = (V, E)$ be a directed graph with n vertices $V = \{v_1, \ldots, v_n\}$ and directed edges $e = (v_i, v_j) \in E$. First let us explain some graph theoretical notions. A *directed path* in the graph G is a sequence of directed edges which connect a sequence of vertices in G. The *path length* equals the number of the edges in the path. A *directed cycle* is a directed path such that the start and the end vertex are the same. Graph G is called *strongly connected*, if for any pair of vertices $v_i \neq v_j$ there is a directed path in G from v_i to v_j.

Let $A = (a_{ij})$ be the adjacency matrix of G with entries

$$a_{ij} > 0 \text{ iff } (v_i, v_j) \in E \quad \text{and} \quad a_{ij} = 0 \text{ otherwise.}$$

Matrix A is clearly positive. There is a nice graph-theoretic description of its irreducibility.

Proposition 6.1.1. *The adjacency matrix A is irreducible if and only if G is strongly connected.*

Proof. Let G be a strongly connected graph and suppose A is reducible. By Lemma 5.2.2, we can partition the sets of vertices $V = V_1 \cup V_2$ into two disjoint subsets such that after relabeling the vertices we obtain block-triangular form for A

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \tag{6.1}$$

where the block A_{kl} for each $k, l \in \{1, 2\}$ corresponds to connections from the set of vertices V_k to the set V_l. Note that $A_{21} = 0$ implies that there are no direct edges from a vertex in V_2 to a vertex in V_1. Now choose arbitrary vertices $v_i \in V_2$ and $v_j \in V_1$. Since G is strongly connected, there exists a directed path in G from v_i to v_j, hence

$$a_{ii_1}a_{i_1i_2}\cdots a_{i_sj} \neq 0$$

for some $i_1, \dots, i_s \in \{1, \dots, n\}$. Observe that in this product there must be a nonzero entry with "mixed" indices, i.e. $a_{i_ki_l} \neq 0$ with $v_{i_k} \in V_2$ and $v_{i_l} \in V_1$, which contradicts (6.1). So, A must be irreducible.

The reverse implication can be proved in the same manner and is left as an exercise. □

As a corollary we obtain a nice combinatorial characterization of positive irreducible matrices. Note that every positive matrix can be seen as an adjacency matrix of a certain graph.

Corollary 6.1.2. *A positive matrix $A \in M_n(\mathbb{R})$, $n \geq 2$, is irreducible if and only if for every $i, j \in \{1, \dots, n\}$ there exists an $s \in \mathbb{N}$ such that $(A^s)_{ij} > 0$.*

We will illustrate another property of the adjacency matrix A in terms of the structure of the graph G. Recall from Theorem 5.3.4 that any imprimitive matrix with index of imprimitivity h can be written in the Frobenius form

$$PAP^\top = \begin{pmatrix} 0 & A_{12} & 0 & \dots & 0 \\ 0 & 0 & A_{23} & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & A_{h-1,h} \\ A_{h1} & 0 & \dots & 0 & 0 \end{pmatrix} \tag{6.2}$$

with square blocks on the main diagonal.

Lemma 6.1.3. *Let A be an imprimitive matrix with index of imprimitivity h and Frobenius form (6.2). Then $A_{12}A_{23}\cdots A_{h1}$ is a primitive matrix.*

Proof. The proof is left as an exercise (see Exercise 3). □

Proposition 6.1.4. *Let the adjacency matrix A be imprimitive with index of imprimitivity h. Then h equals the greatest common divisor of all cycle lengths in the graph G.*

Proof. Fix a vertex v_i of G. Let $\mathcal{L}_i$ be the set of all lengths of cycles through v_i and $d_i = \gcd\{l \mid l \in \mathcal{L}_i\}$. It is not difficult to see that $d_i = d$, the greatest common divisor of all cycle lengths in G (see Exercise 2).

The existence of a cycle in G through v_i is equivalent to condition $(A^l)_{ii} > 0$ for some $l \in \mathbb{N}$ and in this case $l \in \mathcal{L}_i$. Observe that $\mathcal{L}_i$ is closed under addition and hence contains all but finite numbers of positive multiples of d. Hence $(A^{kd})_{ii} > 0$ for all sufficiently large $k \in \mathbb{N}$ and $(A^s)_{ii} = 0$ if s is not a multiple of d.

Since A is an imprimitive matrix with index of imprimitivity h, we may assume it is of form (6.2). Clearly, only powers of A^h can have nonzero diagonal elements. Note that by Lemma 6.1.3 the square diagonal blocks of A^h consist of primitive matrices, hence $A^{mh} \gg 0$ for some $m \in \mathbb{N}$ (check Exercise 10 from Lecture 5). Therefore $(A^{mh})_{ii} > 0$ for all sufficiently large $m \in \mathbb{N}$ and $(A^s)_{ii} = 0$ if s is not a multiple of h.

All together we see that $h = d$. □

6.1.2 Markov chains

Now let a positive stochastic $n \times n$ matrix $P = (p_{ij})$ be the transition matrix of a discrete finite homogeneous Markov chain with the state space $V = \{v_1, \dots, v_n\}$. The *k-th step probability distribution vector* $p(k) = (p_i(k))$ is defined as a positive stochastic vector, i.e.

$$0 \le p_i(k) \le 1, \quad \sum_{i=1}^{n} p_i(k) = 1,$$

where $p_i(k)$ is the probability of Markov process being in the state v_i after k steps. By the Markov property, the k-th step distribution is determined from the initial distribution $p(0)$ and the transition matrix:

$$p(k)^\top = p(0)^\top P^k, \quad k \in \mathbb{N}.$$

Therefore the long-run (or limiting) probability distribution depends on the behavior of P^k for $k \to \infty$. Using our results from Lectures 3 and 5 we can describe it in terms of spectral properties of P.

Let us first state some nice spectral properties of P.

Lemma 6.1.5. *For the transition matrix P the following holds.*

(i) $r(P) = 1$ is an eigenvalue of P with the corresponding eigenvector $\mathbb{1}$.

(ii) All eigenvalues of P with modulus 1 are simple poles of the resolvent.

Proof. Follows easily from the fact that P is positive and row-stochastic matrix (cf. Exercise 5). □

As a consequence, P is always Cesàro summable with Cesàro means converging to the spectral projection of P belonging to 1 (cf. Theorem 3.2.8). The sequence P^k, however, does not necessarily converge as $k \to \infty$. This is true only, if 1 is a radially dominant eigenvalue (see Theorem 3.2.3) while the limit is in this case the same as the Cesàro limit.

The Cesàro means of $p(k)$ have a nice interpretation in the context of Markov chains. Pick a state v_j and define a sequence of random variables $(X_i)_{i=0}^{\infty}$ by

$$X_i = \begin{cases} 1, & \text{if the chain is in the state } v_j \text{ after } i \text{ steps,} \\ 0, & \text{otherwise.} \end{cases}$$

Then $\frac{1}{k}\sum_{i=0}^{k-1} X(i)$ represents the fraction of time that the state v_j is visited in $(k-1)$ steps. Since the expected value of each X_i is $E(X_i) = p_j(i)$, we have

$$E\left(\frac{1}{k}\sum_{i=0}^{k-1} X(i)\right) = \left(\frac{1}{k}\sum_{i=0}^{k-1} p^\top(i)\right)_j .$$

This means that the j-th component of the Cesàro limit vector represents the fraction of time that the chain spends in the state v_j in the long-run.

Assume now, that the matrix P is irreducible (i.e. all states v_i are reachable from each other in finite number of steps). In this case we have two possibilities.

- If P is a primitive matrix, then

  $$\lim_{k\to\infty} P^k = P_1 \text{ with } P_1 x = \langle x, y\rangle \mathbb{1} \quad \text{and} \quad \lim_{k\to\infty} p(k) = y, \tag{6.3}$$

  where y is the stochastic Perron vector for $P^\top$.

- If P is an imprimitive matrix, then the above limits do not exist. However, for the corresponding Cesàro means it holds

  $$\lim_{k\to\infty} P^{(k)} = P_1 \text{ with } P_1 x = \langle x, y\rangle \mathbb{1} \quad \text{and} \quad \lim_{k\to\infty} \frac{1}{k}\sum_{i=0}^{k-1} p(i) = y, \tag{6.4}$$

  where again y is the stochastic Perron vector for $P^\top$.

A Markov chain with an irreducible and imprimitive transition matrix is called *periodic*. In such a chain all states are visited periodically, with the period equaling the index of imprimitivity of P.

Note that the value of the (Cesàro) limit is independent of the initial distribution $p(0)$. Vector y in (6.3) and (6.4) is called the *stationary distribution vector* for the Markov chain. It is the unique stochastic vector satisfying $P^\top y = y$. Its components represent the long-run fraction of time that the chain spends in the corresponding state. If the limiting distribution exists, y_j equals the long-run probability of the chain being in the state v_j.

6.2 The Google Matrix

We shall demonstrate now that we encounter positive matrices and their Perron eigenvectors on everyday basis. We will look at mathematics behind Google[1], currently the world biggest web search engine.

Every web search engine must build its web-page repository and index the pages stored there in the best possible way. For this purpose they use crawler software that creates virtual robots, called spiders, that constantly travel the web. The spiders number each

[1]The name comes from the misspelled number googol $= 10^{100}$.

page, collect important data from it (such as title, key words, link names, anchors, etc.) and create an index of all visited pages. Now the pages have to be ranked according to their importance. When the user does an internet search it is desired that more relevant pages arrive at the beginning of the produced list. This is actually the most important and delicate step for a search engine. It is because of intelligent ranking that Google swept with its competitors as it appeared on the market. The core of Google is the ranking algorithm *PageRank*, that was developed in 1998 by Larry Page and Sergey Brin, then PhD students at Stanford University, California.

6.2.1 PageRank

Assume we have n web pages $W = \{W_k \mid k = 1, \ldots, n\}$. For a page W_k we denote by $I_k := \{i \mid W_i \to W_k\}$ the set of indices of all *inlinks* to W_k, by $O_k := \{j \mid W_k \to W_j\}$ the set of indices of all *outlinks* of W_k, and by $x_k \geq 0$ the *rank of the page* W_k. Now the question is, how to define x_k properly?

The answer of Page and Brin is: *A page is important if it is pointed to by other important pages.* Their formula for the rank is thus recursive:

$$x_k := \sum_{i \in I_k} \frac{x_i}{|O_i|}, \quad k = 1, \ldots, n. \tag{6.5}$$

Here it is assumed that a link from a page to itself does not count. The equation is recursive and it is not clear at this point whether it admits a solution.

Internet can be viewed as a huge directed graph with n vertices (= web pages) whose edges are hyperlinks. Let H be a weighted (transposed) adjacency matrix of this graph, called also the *hyperlink matrix*, with entries

$$H_{ij} = \frac{1}{|O_j|} \text{ iff } W_j \to W_i \quad \text{and} \quad H_{ij} = 0 \text{ otherwise.}$$

We can interpret the values H_{ij} as probabilities of accessing page W_i from page W_j. Collecting single ranks to a *ranking vector* $x := (x_k)_{1 \leq k \leq n}$, we can now write the recursive equation (6.5) as a matrix equation

$$x = Hx. \tag{6.6}$$

The solution vector, if it exists, is thus the fixed vector of the hyperlink matrix H. To assure uniqueness, we impose from now on that the ranking vector x is stochastic, i.e. $\|x\|_1 = 1$.

Note that H is a positive matrix, thus by Perron's Theorem 5.1.6, its spectral radius $r(H)$ is an eigenvalue of H with positive eigenvector. Note that H is a substochastic matrix, i.e. $\sum_{i=1}^{n} H_{ij} \leq 1$ for all j, hence $r(H) \leq 1$. Having (6.6) in mind, we wish $r(H) = 1$. Observe that the sum of non-zero rows actually equals 1, but H might have some zero rows which represent the so-called *dangling nodes*, that is, pages without outlinks. Brin and Page therefore suggested to adjust the matrix H: replace all zero rows with $(1/n, \ldots, 1/n)$. The adjusted matrix becomes stochastic and thus (6.6) with the modified matrix H has a solution. We can also interpret this adjustment. Imagine a random surfer traveling the web using hyperlinks, which he chooses randomly. At some point he might find himself at a dangling node. His way out is to randomly type an url and thus jump to any page with probability $1/n$.

In order to assure the uniqueness of the solution to (6.6), we would like H to be irreducible. By Proposition 6.1.1, H is irreducible if and only if the web is strongly connected, which is clearly a nonrealistic assumption. However, Brin and Page overcommed also this problem with a new adjustment: they replaced the matrix H by the *Google matrix*

$$G := \alpha H + (1-\alpha)S, \tag{6.7}$$

where $S = (1/n)_{n\times n}$ and $\alpha \in [0,1]$ is some fixed number. The interpretation of this adjustment is a continuation of the one above: a random surfer sometimes decides to jump to some other page directly by typing an url instead of following some hyperlink, even if he is not at the dangling node. The role of the parameter α is to balance between the original web structure given by H and fully connected web represented by S. We would of course like to weight the original hyperlink structure heavily and take α close to 1.

For any $\alpha \in [0,1)$, the Google matrix G is positive, irreducible and stochastic, hence Frobenius Theorem 5.2.5 guarantees that the equation $Gx = x$ has a unique strictly positive stochastic solution. Thus the desired ranking vector is nothing but the Perron vector for G!

6.2.2 Computation of the Perron vector

To compute the Perron vector for G we can use a very simple numerical method called the *power method* that was already mentioned at the end of Lecture 3. It is an iterative method defined by

$$x^{(k+1)} = Gx^{(k)}.$$

From this it follows $x^{(k+1)} = G^k x^{(0)}$, thus convergence of this process is assured by Corollary 5.2.6, independent on the choice of the initial vector $x^{(0)}$. Here it is important that 1 is a strictly dominant eigenvalue of the positive irreducible matrix G.

It is well known that the rate of the convergence of the power method is governed by the magnitude of the second eigenvalue $|\lambda_2|$ of the matrix. For the Google matrix we have $|\lambda_2| \leq \alpha$. This means the convergence is faster for smaller α. Since we argued above that α should be close to 1, one has to accept a compromise here. It is reported that Google uses $\alpha = 0.85$, the value set already by Brin and Page in 1998.

6.3 Age-structured population models

Plant, animal and human population models are typical examples for positive dynamical systems in which the state variables represent biomass, density, or the number of individuals of the population. Many of these models, in particular those describing predation, competition and symbiosis among species, are nonlinear and therefore deemed to investigations by other means. An important and still widely used exception is the well-known Leslie Model, which describes the time evolution of population in which fertility and survival rates of individuals strongly depend on their age. For this reason, such populations are called age-structured populations. In the Leslie model, the time is discrete and denotes the reproduction season (typically the year in case of mammals), while the variables $x_1(t), x_2(t), \ldots, x_n(t)$ represent the number of females (or individuals, or couples) of age $1, 2, \ldots, n$ at the beginning of year t.

In the simplest possible case one can describe the aging process by means of the equation

$$x_{i+1}(t+1) = s_i x_i(t), \quad i = 1, 2, \ldots, n-1,$$

where $s_i > 0$ is the survival coefficient at age i, that is, the fraction of females of age i that survive at least for 1 year. The first state equation takes into account the reproduction process, and is

$$x_1(t+1) = s_0(f_1 x_1(t) + f_2 x_2(t) + \ldots + f_n x_n(t)),$$

where $s_0 > 0$ is the survival coefficient during the first year of life and $f_i \geq 0$ is the fertility rate of females of age i, that is, the mean number of females born from each female of age i. These equations, originally proposed by Leslie, lead to a positive linear autonomous model

$$x(t+1) = Ax(t),$$

where the matrix A, called the *Leslie matrix*, is given as

$$A = \begin{pmatrix} s_0 f_1 & s_0 f_2 & \ldots & s_0 f_{n-1} & s_0 f_n \\ s_1 & 0 & \ldots & 0 & 0 \\ 0 & s_2 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & s_{n-1} & 0 \end{pmatrix}. \tag{6.8}$$

Though Leslie models appear to be crude on the first sight, they are extensively used for making demographic projections, i.e., forecasting

$$x(k) = A^k x(0)$$

given $x(0)$.

Let us comment on the usefulness of these models first. In Leslie models, survival and fertility rates depend exclusively on age. In reality, this is more or less true provided the individuals in each age class are not too many. In fact, as soon as the density of the individuals increases, some phenomena show up, which may reduce fertility and/or survival rates. For example, finding appropriate niches for reproduction becomes more difficult if the number of fertile individuals increases; the spreading of epidemics is favoured by high population densities; the search for food becomes more and more difficult as population increases, and so on. This means that Leslie models are well suited for describing the dynamics of populations doomed to extinction, that is, characterized by small densities $x_i(t)$ for which we can suppose that survival and fertility rates are constant as time evolves. Leslie models are also extremely effective yielding short term forecasts in growing populations.

Investigating the properties of the Leslie matrix, we see that it is positive and, if $f_n > 0$, it is also irreducible. Looking at the directed graph whose adjacency matrix is given by (6.8) and using Proposition 6.1.4 one easily obtains that the index of imprimitivity of the Leslie matrix equals

$$h = \gcd\left\{k \in \{1, \dots, n\} \colon f_k > 0\right\}.$$

Hence, if there are two consecutive ages with strictly positive fertility age, then the Leslie matrix is primitive.

The (normalized) Perron eigenvector of the Leslie matrix is called the *stable age structure*, which is roughly the asymptotic age distribution as time evolves. More precisely, we have the following result.

Proposition 6.3.1. *Consider the Leslie matrix A given in* (6.8) *with $f_n > 0$ and assume that A is a primitive matrix. Denote the Perron eigenvalue with $\lambda_1 = r(A)$ and the corresponding eigenvector with $x_1 \geq 0$. Then*

$$\lambda_1^{-k} A^k - P_1 \to 0$$

as $k \to \infty$, where P_1 is the projection to the one dimensional subspace spanned by x_1.

Let us note that in many applications it is better to structure the population not in age groups, but in so-called *stage groups*. As an example, we consider Virginias hunted black bear populations. Deep statistical analysis, where we omit the details, leads to the following table, which is here only reproduced to show the complexity of such problems.

Age Class	Average Reproduction / Year	Low Reproduction / Year	High Reproduction / Year	Average Annual Survival	Low Annual Survival	High Annual Survival
Cub	0.00	0.00	0.00	0.80	0.41	0.99
1-year-old	0.00	0.00	0.00	0.75	0.41	0.99
2-year-old	0.00	0.00	0.00	0.71	0.41	0.90
3-year-old	0.28	0.00	0.50	0.84	0.69	0.93
Adult	0.58	0.23	0.82	0.84	0.69	0.93

Figure 6.1: Input parameters for Leslie Matrix population model (based on females only) of Virginias hunted black bear populations as estimated between 1994-1999.

In this case, as we see, it is better to investigate the so-called stage-based Leslie model. Stage-based models are frequently used for long-lived species because data on specific ages are not available, demographic variables within age classes are not different, and individual age classes for a species that lives, for example, up to 30 years (like black bear), would result in matrices of sizes up to 30×30. Analysis of this table can lead to the following Leslie matrix, where various other effects have been taken into account, and which was used successfully in analysis done by biologists:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0.275 & 0.575 \\ 0.80 & 0 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0 & 0 \\ 0 & 0 & 0.71 & 0 & 0 \\ 0 & 0 & 0 & 0.84 & 0.84 \end{pmatrix}. \tag{6.9}$$

Here the last row stands for the whole adult stage, the element in the lower right corner of the matrix representing the rate of the adult population remaining alive after the year.

We now consider a second model, which is famous in the literature. The eastern wild turkey (Meleagris gallopavo silvestris) inhabits roughly the eastern half of the United States. Turkey hunting has a substantial economic effect in many rural communities. It is not only important because of the actual turkey hunting, but it also takes part in the development of the related industries of turkey-hunting clothes and equipment. Improvement of the knowledge of turkey population dynamics is important to formulate hunting regulations and other turkey management practices. A Leslie matrix model can be developed for the population dynamics of eastern wild turkeys in Iowa based on Iowa wild turkey studies. A three-stage model is chosen in order to simplify the modeling procedure. The first category is "poults", aged from 0 to 1, the second category is "yearlings", aged from 1 to 2, and the last category is "adults", aged 2 and older. Reproduction occurs from yearlings onwards. The time unit is one year. The Leslie-matrix obtained is

$$A = \begin{pmatrix} 0 & 0.880 & 1.860 \\ 0.445 & 0 & 0 \\ 0 & 0.616 & 0.610 \end{pmatrix}. \tag{6.10}$$

This grouping makes sense for example if there are regulations allowing only the adult population to be hunted, see Exercise 10.

6.4 Notes and Remarks

For further reading on the topic of the search engines and PageRank algorithm we recommend excellent monograph [3]. The modeling and investigation of age structured populations was initiated by Leslie in 1945 [5], and extended to stage structured populations by Lefkovitch [4]. Virginias hunted black bear populations is discussed in the PhD dissertation [2]. Much research about the rates of reproduction, mortality, and survival, and the movement of wild turkeys has been done by Dickson [1].

6.5 Exercises

1. Show the missing implication in the proof of Proposition 6.1.1.

2. Let G be a strongly connected graph and let v_i be a vertex of G, $\mathcal{L}_i$ the set of all lengths of cycles through v_i and
$$d_i = \gcd\{l \mid l \in \mathcal{L}_i\}.$$
Show that d_i equals the greatest common divisor of all cycle lengths in the graph G.

3. Prove Lemma 6.1.3.

4. Verify that the matrix
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$
is irreducible and imprimitive using graph-theoretical interpretations. Compute also its index of imprimitivity.

5. Show the statements of Lemma 6.1.5.

6. Argument the statements given in (6.3) and (6.4). Why is the limiting distribution independent of $p(0)$?

7. Find the limiting distribution for the Markov chain given by the transition matrix
$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 1/3 & 2/3 & 0 \end{pmatrix}.$$

8. Translate the PageRank algorithm into the language of Markov chains.

9. Compute the ranking vector for the web depicted in Figure 6.2. Choose several values for α and observe how does this choice affect the ranking and the computation time.

10. Consider the Leslie matrix (6.10) corresponding to the turkey population in Iowa. Use an appropriate computer software if necessary.

    (a) Calculate the Perron eigenvalue and the corresponding stable age structure. Is the population growing?

    (b) Assume we can change the survival rate of the adult population. How should we change the survival rate of the adult population to ensure that the Perron eigenvalue equals 1, meaning that the population remains balanced?

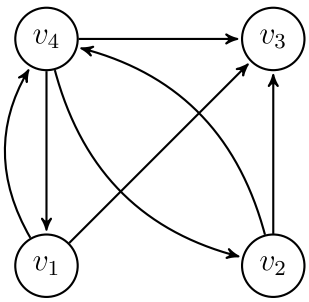


Figure 6.2: The web graph from Exercise 9.

(c) Using a 1977 survey, the age structure in a region in Iowa was estimated as $x_1(0) = 580$, $x_2(0) = 123$, $x_3(0) = 156$. How many adults should be hunted down at the end of the first year to ensure this decrease in the survival rate of adults?

11. What is the Perron eigenvalue and the corresponding stable age distribution of the the Leslie matrix (6.9) corresponding to the bear population? Is the population growing, balanced or dying out? Use an appropriate computer software if necessary.

12. To connect two topics of today, google further Leslie matrix models, for example for the annual bluegrass (poa annua) or the brown rat (rattus norvegicus) populations.

Bibliography

[1] J. G. Dickson, *The wild turkey: biology and management*, Stackpole Books, Harrisburg, PA, 1992.

[2] S. Klenzendorf, *Population dynamics of virginias hunted black bear population*, Ph.D. thesis, Virginia Polytechnic Institute, 2002.

[3] A. Langville and C.D. Meyer, *Google's pagerank and beyond: The science of search engine rankings*, Princeton University Press, 2006.

[4] L. P. Lefkovitch, *The study of population growth in organisms grouped by stages*, Biometrics **21** (1965), 1–18.

[5] P.H. Leslie, *On the use of matrices in certain population mathematics*, Biometrika **33** (1945), 183–212.

Lecture 7

Positive Matrix Semigroups and Applications

7.1 Positive Semigroups

In this section we combine the matrix exponential from Lecture 4 with the positivity from Lecture 5. More precisely, we consider *positive matrix semigroups* $(\mathrm{e}^{tA})_{t\geq 0}$, i.e., we assume that each e^{tA}, $t \geq 0$, is a positive matrix. As a first step, we characterize this property by the entries of A. In particular, we show that positivity of A is sufficient, but not necessary for this.

Let $A = (\alpha_{ij})_{n\times n}$ be given. We then obtain by Theorem 4.1.2 that

$$A = \lim_{t\downarrow 0} \frac{\mathrm{e}^{tA} - I}{t},$$

which means

$$\alpha_{ij} = \lim_{t\downarrow 0} \left\langle \frac{\mathrm{e}^{tA} e_j - e_j}{t}, e_i \right\rangle, \tag{7.1}$$

for $i, j = 1, \ldots, n$ and e_i the i-th unit vector in $\mathbb{C}^n$. If we denote the (i,j)-th entry of e^{tA} by $\tau_{ij}(t)$, then (7.1) means

$$\alpha_{ij} = \begin{cases} \lim_{t\downarrow 0} \frac{\tau_{ij}(t)}{t} & \text{for } i \neq j, \\ \lim_{t\downarrow 0} \frac{\tau_{ii}(t)-1}{t} & \text{for } i = j. \end{cases} \tag{7.2}$$

If $(\mathrm{e}^{tA})_{t\geq 0}$ is positive, i.e. $\tau_{ij}(t) \geq 0$ for all t, i and j, this implies

$$\begin{aligned} &\alpha_{ij} \geq 0 \quad \text{for } i \neq j, \text{ and} \\ &\alpha_{ii} \in \mathbb{R} \quad \text{for } i = j. \end{aligned}$$

We call such matrices *positive off-diagonal* and have shown the necessity part of the following characterization.

Theorem 7.1.1. *A matrix $A = (\alpha_{ij})_{n\times n} \in M_n(\mathbb{C})$ generates a positive semigroup if and only if it is real and positive off-diagonal.*

Proof. It remains to show the sufficiency of the condition.

Since A is real and positive off-diagonal, we can find $\rho \in \mathbb{R}$ such that

$$B_\rho := A + \rho I \geq 0 \tag{7.3}$$

(e.g., take $\rho := \max_{1\leq i\leq n} |\alpha_{ii}|$). Note that then also $\mathrm{e}^{tB_\rho} \geq 0$ for all $t \geq 0$. Applying the Functional Calculus introduced in Section 2.2 to the function $f(\lambda) := \mathrm{e}^{t\lambda - t\rho}$ we obtain

$$\begin{aligned}\mathrm{e}^{tA} &= \mathrm{e}^{[t(A+\rho I)-t\rho I]} \\ &= f(B_\rho) \\ &= \mathrm{e}^{-t\rho}\cdot \mathrm{e}^{tB_\rho} \geq 0\end{aligned}$$

for all $t \geq 0$. □

Let us mention another terminology here. If A is a real and positive off-diagonal matrix, $-A$ is also called a *Z-matrix*.

Since from $\mathrm{e}^{tA} \geq 0$ does not necessarily follow $A \geq 0$, Perron's Theorem 5.1.6 is not directly applicable and $r(A)$ may not be an eigenvalue of A. However, observing the positive matrix B_ρ defined in (7.3) we obtain an important property of the spectral bound $\mathrm{s}(A)$ of A.

Theorem 7.1.2. *If A generates a positive semigroup $(e^{tA})_{t\geq 0}$, then $\mathrm{s}(A)$ is a* strictly dominant *eigenvalue, i.e. $\mathrm{s}(A) \in \sigma(A)$ and*

$$\mathrm{Re}\,\lambda < \mathrm{s}(A)$$

for all other eigenvalues λ of A.

Proof. As already noticed in the proof of Theorem 7.1.1, $B_\rho := A + \rho I \geq 0$ for $\rho := \max_{1\leq i\leq n} |\alpha_{ii}|$. Perron's Theorem 5.1.6 yields $\mathrm{r}(B_\rho) \in \sigma(B_\rho)$. Evidently, $\mathrm{r}(B_\rho) = \mathrm{s}(B_\rho)$ which is strictly dominant in $\sigma(B_\rho)$. Since

$$\sigma(B_\rho) = \sigma(A + \rho I) = \sigma(A) + \rho,$$

and thus $\mathrm{s}(B_\rho) = \mathrm{s}(A + \rho I) = \mathrm{s}(A) + \rho$, we obtain that $\mathrm{s}(A) \in \sigma(A)$ and is strictly dominant. □

As we have seen in Theorem 4.4.3, $\mathrm{s}(A)$ determines the asymptotic behavior of e^{tA} as $t \to \infty$. The case $\mathrm{s}(A) < 0$, yielding stability of the semigroup, is of particular importance. In the case of positive semigroups, we thus obtain the following characterization of stability.

Corollary 7.1.3. *If A generates a positive semigroup $(e^{tA})_{t\geq 0}$, then the following assertions are equivalent.*

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

(b) The characteristic polynomial of A has no real root ≥ 0.

(c) The matrix $-A^{-1}$ exists and is positive.

(d) There exists $x \geq 0$ such that $Ax = -\mathbb{1}$.

(e) The semigroup $(e^{tA})_{t\geq 0}$ is exponentially stable.

Proof. The equivalence of (e) and (a) should be clear from Theorem 4.4.3. The equivalence (a) $\iff$ (b) follows from Theorem 7.1.2. Since $-A^{-1} = R(0, A)$, see Exercise 7.7.1 for (c) $\iff$ (a). (c)$\Rightarrow$(d) follows taking $x := -A^{-1}\mathbb{1}$.

We close the implications loop by showing (d)$\Rightarrow$(a). If $Ax = -\mathbb{1}$, then

$$x \geq x - \mathrm{e}^{tA}x = -A\int_0^t \mathrm{e}^{sA}x\, ds = \int_0^t \mathrm{e}^{sA}\mathbb{1}\, ds \text{ for all } t > 0.$$

Since e^{tA} is positive for all $t \geq 0$, it follows that the function $t \mapsto (\int_0^t \mathrm{e}^{sA}\mathbb{1}\, ds)$ is nondecreasing and satisfies

$$0 \leq \int_0^t \mathrm{e}^{sA}\mathbb{1}\, ds \leq x.$$

Hence, $\int_0^\infty \mathrm{e}^{sA}\mathbb{1}\, ds$ exists and by

$$A\int_0^t \mathrm{e}^{sA}\mathbb{1}\, ds = e^{tA}\mathbb{1} - \mathbb{1},$$

we have the existence of $\lim_{t\to+\infty} e^{tA}\mathbb{1}$. This implies $r(e^{tA}) \leq 1$, since $\|e^{tA}\|_\infty = \|e^{tA}\mathbb{1}\|_\infty$. Thus, $s(A) \leq 0$. Take now $\varepsilon > 0$. Then, (see Exercise 7.7.1),

$$\begin{aligned} 0 \leq R(\varepsilon, A)\mathbb{1} &= -R(\varepsilon, A)Ax \\ &= x - \varepsilon R(\varepsilon, A)x \leq x. \end{aligned}$$

So, we have the existence of $\lim_{\varepsilon\to 0} R(\varepsilon, A)\mathbb{1}$, since the function $\varepsilon \to R(\varepsilon, A)$ is nonincreasing. Therefore,

$$\lim_{\varepsilon\to 0} \|R(\varepsilon, A)\|_\infty$$

exists, and hence $s(A) < 0$. □

For a matrix A with above properties, $-A$ is in literature also known as a *nonsingular M-matrix*.

In the case when $\mathrm{s}(A) = 0$ the following is a consequence of Theorem 4.4.3.

Corollary 7.1.4. *Let A generate a positive semigroup $(e^{tA})_{t\geq 0}$ and assume $\mathrm{s}(A) = 0$. Then $\lim_{t\to\infty} e^{tA}$ exists if and only if 0 is a first order pole of the resolvent $R(\cdot, A)$. In this case, $\lim_{t\to\infty} e^{tA}$ is the spectral projection of A belonging to 0, and its range is the kernel of A.*

Again, it is important to assure that $\mathrm{s}(A)$ is a simple pole. As shown in Theorem 5.2.5, irreducibility can help for this purpose.

Theorem 7.1.5. *Let A generate a positive semigroup $(e^{tA})_{t\geq 0}$. Then any of the following conditions implies*

$$\lim_{t\to\infty} e^{tA} = P_1$$

for P_1 a projection of the form

$$P_1 x = \langle x, y\rangle\, z, \quad x \in \mathbb{C}^n,$$

with strictly positive vectors $y \gg 0$ and $z \gg 0$ such that $\langle z, y\rangle = 1$.

(i) A *is irreducible with* $\mathrm{s}(A) = 0$.

(i') A *is irreducible,* $(e^{tA})_{t\geq 0}$ *is bounded, and* $0 \in \sigma(A)$.

(ii) $e^{t_0 A}$ *is irreducible for some* $t_0 > 0$ *and* $\mathrm{s}(A) = 0$.

(ii') $e^{t_0 A}$ *is irreducible for some* $t_0 > 0$, $(e^{tA})_{t\geq 0}$ *is bounded, and* $0 \in \sigma(A)$.

(iii) $(e^{tA})_{t\geq 0}$ *is bounded,* $\mathrm{s}(A) = 0$, *and the kernels of* A *and* $A^\top$ *are* 1*-dimensional and spanned by strictly positive vectors.*

Proof. First observe that irreducibility of A is equivalent to irreducibility of e^{tA} for some/all $t > 0$ (see Exercise 4) and also to the condition on the eigenspaces of A and $A^\top$ belonging to $r(A) = r(A^\top)$ given in (iii) (see Exercise 5.5.7).

Next, if A is irreducible, then $B_\rho = A + \rho I$ from (7.3) is positive and irreducible, hence by Theorem 5.2.5 its spectral radius $r(B_\rho)$ is a first order pole of the resolvent $R(\cdot, B_\rho)$. Therefore also $s(A)$ is a first order pole of the resolvent $R(\cdot, A)$.

All assertions now follow by Corollary 7.1.4 and Theorem 4.4.3 (ii). The formula for $P_1 x$ can be verified as in the proof of Corollary 5.2.6. □

7.2 The Competitive Markets Model

As a first application let us revise the competitive markets model presented in Lecture 1. Recall that the dynamic of the prices $p(t)$ in this model is given by

$$p(t) = p^0 + \mathrm{e}^{tKA} c, \quad t \geq 0, \quad \text{where} \quad c = p(0) - p^0. \tag{7.4}$$

Here p^0 are equilibrium prices, $p(0)$ initial prices, $K = \mathrm{diag}(k_1, \ldots, k_n)$ a diagonal matrix of positive adjustment speeds while for the coefficients of the matrix $A = (a_{ij})$ we have

$$a_{ij} \geq 0 \text{ for } i \neq j \text{ and } a_{ii} < 0.$$

Using theory we developed so far we are able to study the behavior of the prices depending on spectral properties of the matrix KA.

Let us first examine in which case the prices eventually return to the equilibrium p^0. Assuming $p(0) \neq p^0$, by Theorem 4.4.3 this happens if and only if $\mathrm{s}(KA) < 0$. Moreover, since KA is real and positive off-diagonal, it generates a positive semigroup and $\mathrm{s}(KA)$ is the largest real eigenvalue of KA (cf. Theorems 7.1.1 and 7.1.2). Furthermore, by Corollary 7.1.3 we have

$$\begin{aligned} \mathrm{s}(KA) < 0 \quad &\Longleftrightarrow \quad (KA)^{-1} \leq 0 \\ &\Longleftrightarrow \quad A^{-1} \leq 0 \\ &\Longleftrightarrow \quad \mathrm{s}(A) < 0 \\ &\Longleftrightarrow \quad \text{there exists } x \geq 0 \text{ such that } Ax = -\mathbb{1}. \end{aligned}$$

Hence automatic return to the equilibrium p^0 can be checked by determining A^{-1} (if it exists; one positive entry in A^{-1} means: $\mathrm{s}(KA) \geq 0$), or by solving the equation $Ax = -\mathbb{1}$.

If $\mathrm{s}(A) > 0$, the prices will rise unboundedly. For $\mathrm{s}(A) = 0$ we have more possibilities: the prices may either converge to a new equilibrium or oscillate periodically.

Let us consider the case when A is strictly positive off-diagonal and $\mathrm{s}(A) = 0$. By Theorem 7.1.5 (see also Exercise 5.5.5) in this case e^{tKA} converges to a projection of the form $P = u \otimes v$ with $u, v \gg 0$. Hence $p(t)$ converges, as $t \to \infty$, to $p^0 + d$ where

$$d = \lim_{t \to \infty} \mathrm{e}^{tKA} c = \langle c, u \rangle \, v.$$

This produces the strange effect that for $\langle c, u \rangle > 0$ a new equilibrium $\tilde{p}^0 = p^0 + d$ develops with $d \gg 0$. On the other hand, $\langle c, u \rangle < 0$ produces a new equilibrium with $d \ll 0$.

7.3 Queueing models

Systems in which an operation is performed on individuals are frequently encountered in applications. Such systems are generally composed of two parts, the queue line and the service.

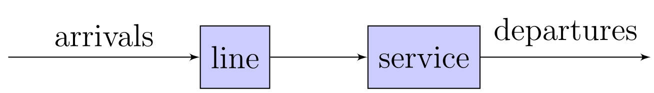


Figure 7.1: Structure of a queueing system

An airport with a queue of airplanes waiting for landing, an office where the papers for driving licence renewal are processed, or the waiting room of a service center (telefon, gas, electric, etc.) are typical examples characterized by a waiting time followed by a service. Such systems can be modeled when the statistics of the arrivals, the rule for selecting the next user, and the statistics of the service times are specified.

We will assume that the arrivals and the departures are random processes characterized by the property that the probability of one arrival or departure during a small time Δt is proportional to Δt itself, meaning that they are Poisson's processes. The proportionality coefficients, denoted by η and μ, respectively, depend on the total number of people in the system.

To build the equations governing the system, we make the following assumption. At time $t + \Delta t$ there are no users in the system, if one of the following happens: either there are no users in the system at time t and no user arrives during this time interval, or there is one user in the system at time t, this user leaves the system, and no other user arrives. We arrive at the equations

$$y_0(t + \Delta t) = y_0(t)(1 - \eta_0 \Delta t) + y_1(t)\mu_1 \Delta t(1 - \eta_1 \Delta t),$$

where $y_0(t)$ is the probability that no user is in the system at time t, and $y_1(t)$ is the probability that exactly one user is in the system at time t.

Taking the limit $\Delta t \to 0$, we arrive at the equation

$$\dot{y}_0 = -\eta_0 y_0(t) + \mu_1 y_1(t), \tag{7.5}$$

which is the first state equation of the system. With a similar reasoning, we get for $i > 0$ that

$$\dot{y}_i(t) = \eta_{i-1} y_{i-1}(t) - (\eta_i + \mu_i) y_i(t) + \mu_{i+1} y_{i+1}(t), \tag{7.6}$$

where $y_i(t)$ is the probability that i users are in the system at time t. This means that there are four possibilities of the system to change in the state i:

- There were $i - 1$ users and someone arrived;
- There were i users and someone arrived;
- There were i users and someone left;
- There were $i + 1$ users and someone left.

Such problems are naturally modeled by infinite dimensional systems. In many applications, however, we have a natural bound on the number of possible users entering the system on the whole, hence we get a finite dimensional system, and this is the case we are investigating now. Assuming that all the proportionality constants η_j and μ_j are non-zero, this all leads to a system of differential equations

$$\dot{y}(t) = Ay(t), \tag{7.7}$$

where $y = (y_0, y_1, \ldots, y_n)^T \in \mathbb{R}^{n+1}$ and

$$A = \begin{pmatrix} -\eta_0 & \mu_1 & 0 & 0 & \ldots & 0 & 0 \\ \eta_0 & -\eta_1 - \mu_1 & \mu_2 & 0 & \ldots & 0 & 0 \\ 0 & \eta_1 & -\eta_2 - \mu_2 & \mu_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \eta_{n-1} & -\mu_n \end{pmatrix}. \tag{7.8}$$

The matrix A is a so-called *band matrix*, has negative elements on the diagonal, positive elements below and above the diagonal, and zeros elsewhere. It follows that the matrix is irreducible and hence the positive and irreducible matrix $B_\rho = A + \rho I$ defined in (7.3) has unique strictly positive Perron eigenvector x_P. Note that x_P is also an eigenvector of A with the corresponding strictly dominant eigenvalue $\lambda_1 \in \mathbb{R}$. Moreover, since all the columns of A have a zero sum, $\lambda_1 = 0$. Since all the eigenvalues of A, except λ_1, have negative real part, Theorem 7.1.5 yields

$$\mathrm{e}^{tA} \to P_1 \qquad \text{as } t \to \infty.$$

Hence, the system will always converge to a stationary probability distribution given by the Perron eigenvector x_P, and the expected value of the clients in the queue (and the expected waiting time) can be estimated using this on the long run. This asymptotic distribution can be easily evaluated because A is almost diagonal. In fact, we easily see that

$$\bar{x}_{i+1} = \frac{\eta_i}{\mu_{i+1}} \bar{x}_i,$$

where $x_P = (\bar{x}_0, \dots, \bar{x}_n)^T$. Hence,

$$\bar{x}_i = \varphi_i \bar{x}_0,$$

where

$$\varphi_i = \frac{\eta_0 \eta_1 \cdots \eta_{i-1}}{\mu_1 \mu_2 \cdots \mu_i}.$$

Since x_P should be normalized to represent a probability distribution, its coordinates sum up to 1, hence

$$\bar{x}_0 = \frac{1}{1 + \sum_{i=1}^n \varphi_i}.$$

We may also assume that there are s servers and c people can get into the queue, i.e., $n = c + s$. Then the average number of people in the system is

$$\bar{n} = \sum_{i=0}^{n} i \bar{x}_i,$$

and we see

$$\bar{c} = \sum_{i=s+1}^{n} (i - s) \bar{x}_i$$

persons waiting in the queue.

7.4 Disease transition models

The spread and persistence of infectious diseases is a result of the complex interaction between the behaviour of individual epidemiological units (e.g. individual, city, county, etc.), disease characteristics and various control programmes that are aimed at halting disease transmission or bringing infection prevalence to as low a level as possible. The main aim of many models is to gain insight into how diseases transmit and to identify the most effective strategies for their prevention and control. The early work by Kermack and McKendrick from 1927 forms the basis of differential-equation-based models which lie at the heart of modern quantitative epidemiology. Traditionally, mathematical epidemiology is based on differential equation models and these operate on the basis of some strong simplifying assumptions about the behaviour of the individuals and the biology of the disease. A key component in any disease transmission model is the population contact structure, and in most cases, this is highly heterogeneous with strong correlations and non-trivial large scale structure.

We will consider here a really elementary, so-called "SIS" model, where individuals can have two possible states: susceptible (but healthy) or infected. Infected individuals can infect susceptible ones with a certain rate, and infected individuals recover with a certain rate, and can be infected afterward again. Usually there are natural time delays (like incubation) and non-linear dependencies in the model, or other stages (like immunized individuals), but we consider here a rather simple case.

Assume that there are n individuals in the model, and that there is a connection between them describing the connections where the infection can spread. Usually the graph is huge and some random graph models have to be used, but let us restrict ourselves to the case where we have a rather small group of individuals with a clear social network. So the assumptions on the model are the following:

Suppose the individuals are connected by a weighted undirected graph, with adjacency matrix $G = (w_{ij})$, where the weights $0 \leq w_{ij} \leq 1$ describe the strength of the connection. Each individual can recover with a rate μ and infect a connected person with a rate η. Denoting with $y_i(t)$ the probability of the i-th individual to be infected, a possible simple model for the change of this probability after a small time Δt is

$$y_i(t + \Delta t) = (1 - \mu\Delta t)y_i(t) + \eta\Delta t \sum_{j=1}^{n} w_{ij} y_j(t).$$

This leads to the differential equation

$$\dot{y}(t) = (\eta G - \mu I)y(t). \tag{7.9}$$

It is usual to assume that the graph is connected and has no loops. Hence, the matrix $A := \eta G - \mu I$ is irreducible. Theorem 7.1.5 is applicable if $\mu = \eta s(G)$.

Convergence to zero here is implied by the condition $s(G) < \frac{\mu}{\eta}$, and means that everyone recovers from the illness eventually, and convergence to a projection means that there is a stationary distribution of probabilities and the system tries to achieve this.

7.5 Discrete maximum principles

Maximum principles play an important part in many mathematical subdisciplines and it has been known long that there is a deep connection between maximum principles and positive semigroups. Discrete maximum principles play a prominent role in numerical analysis: when you discretize a differential operator, you not only want to get some good approximation result, but it is also important to preserve some qualitative properties of the underlying differential equation. Hence, if you discretize an elliptic problem, it is good to know whether the discretization also satisfies some kind of a maximum principle. The literature is vast even in the matrix case and we only mention here one illustrative example.

Let $A = (a_{ij})$ be a real $N \times N$ matrix and for a vector $y = (y_1, \ldots, y_N)^T$ let us use the notation $N^+(y) = \{i \in \{1, \ldots, N\} : y_i > 0\}$. We say that the matrix A satisfies the *discrete maximum principle (DMP)*, if

$$-Ax = y \tag{7.10}$$

with $y \geq 0$ implies $x \geq 0$ and

$$\max_{1 \leq i \leq N} x_i = \max_{i \in N^+(y)} x_i. \tag{7.11}$$

We present now a simple sufficient condition to ensure the discrete maximum principle.

Theorem 7.5.1. *Suppose that A generates an exponentially stable positive semigroup and that $A\mathbb{1} \ll 0$ where $\mathbb{1} = (1, \ldots, 1)^T$. Then the discrete maximum principle holds.*

Proof. Let $y > 0$ and $-Ax = y$. Then $x = -A^{-1}y > 0$, since $-A^{-1}$ exists and is positive by Corollary 7.1.3. Let $x_k = \max_{1 \leq i \leq N} x_i$. By assumption we have $\sum_{l=1}^N a_{kl} = (A\mathbb{1})_k < 0$ and using Theorem 7.1.1 we obtain

$$y_k = -\sum_{l=1}^{N} a_{kl} x_l = -a_{kk} x_k - \sum_{l \neq k} a_{kl} x_l \geq -x_k \left(a_{kk} + \sum_{l \neq k} a_{kl} \right) \geq 0, \tag{7.12}$$

hence $k \in N^+(y)$. □

7.6 Notes and Remarks

For nonsingular M-matrices we refer to the monograph by A. Berman and R.J. Plemmons [1].

For queuing systems see the excellent exposition in Feller [4, Section XVII.7].

For populations equations see the monograph by Diekmann and Heesterbeek [3], where all the problems we omitted here are discussed. Nowadays the investigation of disease transition on large networks is an important and active research field and we plan to come back to it in later lectures with more realistic models.

The investigation of discrete maximum principles was initiated in the late sixties and early seventies, see R. S. Varga [7], Ph. Ciarlet [2]. We follow here G. Stoyan [6], where a much more general statement is formulated. In the infinite dimensional setting, we shall discuss the connection of maximum principles and generation of positive semigroups. Here we recommend the paper by A. Kalauch [5] for a direct generalization to Banach lattices.

7.7 Exercises

1. If $\mathrm{e}^{tA} \geq 0$ for all $t \geq 0$ and $\mu \in \rho(A)$, then $R(\mu, A) \geq 0$ if and only if $\mu > \mathrm{s}(A)$.
2. Let $A \in \mathrm{M}_n(\mathbb{C})$. If $\operatorname{Re}\lambda > \mathrm{s}(A)$, then
$$R(\lambda, A) = \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{tA}\, dt.$$
3. If $\mathrm{e}^{tA} \geq 0$ for all $t \geq 0$, then $\mathrm{s}(A) \in \sigma(A)$ with a strictly positive eigenvector.
4. Let $A \in \mathrm{M}_n(\mathbb{C})$ generate a positive semigroup. Show that the following are equivalent.
   (a) A is irreducible.
   (b) $\mathrm{e}^{t_0 A}$ is irreducible for some $t_0 > 0$.
   (c) e^{tA} is irreducible for all $t > 0$.
   (d) The semigroup $(\mathrm{e}^{tA})_{t\geq 0}$ is irreducible in the following sense: For no subset $\emptyset \neq M \subsetneqq \{1, \dots, n\}$ the subspace J_M (cf. Definition 5.2.1) is invariant under each e^{tA} $(t \geq 0)$.
   (e) $\mathrm{e}^{tA} \gg 0$ for all $t > 0$.
5. Let A be positive off-diagonal and irreducible, $\mathrm{s}(A) = 0$. Show that then e^{tA} converges to a projection P of the form
$$P = u \otimes v$$
   with $u, v \gg 0$, $u \in \ker A^T$, $v \in \ker A$.
6. (a) Characterize those matrices $A \in \mathrm{M}_n(\mathbb{C})$ for which e^{tA} is positive for all $t \in \mathbb{R}$.
   (b) Find all positive periodic semigroups, and all positive, periodic, irreducible semigroups.
7. Consider the Competitive Markets Model given by (7.4).
   (i) List the conditions for the matrix A under which the prices will behave periodically.
   (ii) The price average in given time $T > 0$ can be expressed with the Cesàro mean
$$C(T) := \frac{1}{T}\int_0^T e^{tKA} dt.$$
   What can we say about the long-time behavior of the price averages depending on the properties of matrix A?

8. Consider the queuing system and give formula for the average frequency of arrivals, the fraction of time during which the system is not used, and the probability that a user cannot be served on arrival.

9. Consider a telephone system of a large company where calls are accepted if at least one line is free and rejected otherwise. Thus, the frequency of arrivals η is independent of the number of busy lines if at least one is free. Each accepted call engages the line for an average time $1/\mu$. Hence, $c = 0$, $s = n$, $\eta_i = \eta$ for $i = 0, 1, \ldots, n-1$, $\eta_n = 0$, and $\mu_i = i\mu$ for $i = 1, 2, \ldots, n$. Develop formulas for this system. Discuss some concrete examples.

10. Assume that we have a group of individuals who are arranged in the vertices of a graph and the disease can spread along the edges. Each individual can recover from he illness with a rate of $\mu = 1/4$. Discuss the role of the parameter η, if he graph is

    (a) a complete graph with 4 vertices,

    (b) a cycle of length 5 (regular pentagon),

    (c) a cube (8 vertices),

    (d) the graph from Exercise 1.4.2.

11. It is also possible to speak about maximum principles in connection with difference equations and boundary value problems. Let $a_i, b_i > 0$ and $c_i \geq a_i + b_i$ for $i = 0, 1, \ldots, N+1$. Define the difference operator

    $$Ly_i = a_i y_{i-1} - c_i y_i + b_i y_{i+1}$$

    for $i = 1, \ldots, N$.

    Let $\alpha, \beta \geq 0$ and consider the vector $0 \leq f \in \mathbb{R}^N$. Let $y \in \mathbb{R}^{N+2}$ be such that

    $$Ly_i = f_i \quad \text{for } i = 1, \ldots, N, \quad y_0 = \alpha,\ y_{N+1} = \beta$$

    holds, which we call a non-homogeneous boundary value problem. Show that

    $$\max_{i=0,\ldots,N+1} y_i = \max\{\alpha, \beta\}.$$

Bibliography

[1] A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, vol. 9, SIAM, 1994.

[2] Philippe G. Ciarlet, *Discrete maximum principle for finite-difference operators*, Aequationes Math. **4** (1970), 338–352.

[3] Odo Diekmann and J. A. P. Heesterbeek, *Mathematical epidemiology of infectious diseases*, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons Ltd., Chichester, 2000, Model building, analysis and interpretation.

[4] William Feller, *An introduction to probability theory and its applications. Vol. I*, Third edition, John Wiley & Sons Inc., New York, 1968.

[5] Anke Kalauch, *On maximum principles for M-operators*, Linear Algebra Appl. **371** (2003), 209–224.

[6] Gisbert Stoyan, *On maximum principles for monotone matrices*, Linear Algebra Appl. **78** (1986), 147–161.

[7] Richard S. Varga, *On a discrete maximum principle*, SIAM J. Numer. Anal. **3** (1966), 355–359.

Lecture 8

Positive Linear Systems

8.1 Introduction

The main goal of this lecture is to present an elementary introduction to positive linear systems. Our aim is twofold: on the one hand, we would like to present another application of positivity for those who are not experts in this field. On the other hand, we would like to make the connections also for students interested in control theory visible to see the usefulness of the subjects also later on. For simplicity, we present here continuous time systems, but the definitions and most results can be altered to the discrete time case in a straightforward way.

First we would like to set the stage and present the relevant notation, terminology and motivation. For the sake of simplicity, we will only refer to the case of time-invariant, finite dimensional input-output systems, which are described by the state equations of the form

$$(\Sigma(A,B,C)) \qquad \begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ x(0) &= x_0, \\ y(t) &= Cx(t), \end{aligned}$$

where the appearing objects are the following:

- $X = \mathbb{C}^n$ is the *state space*, $Y = \mathbb{C}^q$ is the observability space, and $U = \mathbb{C}^p$ is the control space,
- The function $x : \mathbb{R}^+ \to X$ is the state vector, the operator $A \in \mathcal{L}(X)$ is the state (or system) operator,
- The function $u : \mathbb{R}^+ \to U$ is the control, the operator $B \in \mathcal{L}(U, X)$ is the input (or control) operator,
- The function $y : \mathbb{R}^+ \to Y$ is the output (or observation), the operator $C \in \mathcal{L}(X, Y)$ is the output (or observation) operator,
- The vector $x_0 \in X$ is the initial value.

The interpretation of this set of equations is the following: There is a system described by a set of n equations and governed by the operator A. This is also referred to as the "free system", the system without intervention. The function u is the control we apply from the outside, and the operator B represents the action of u on the system. Finally, the function y is the set of parameters we are able to measure, and the measurement process is described by the observation operator C.

If we would like to stress the dependence of the solution x on the initial value x_0, then we shall write $x(t) = x(t; x_0)$ later on.

Before turning our attention to controllability concepts, let us make the following crucial observation and present a representation formula. Suppose that the control u is (at least...) locally integrable. Since x is the solution of an inhomogeneous linear differential equation, the variation of constants formula holds, hence,

$$x(t) = \mathrm{e}^{tA}x_0 + \int_0^t \mathrm{e}^{(t-s)A}Bu(s)\, ds \tag{8.1}$$

(see Exercise 1), yielding the formula

$$y(t) = C\mathrm{e}^{tA}x_0 + \int_0^t C\mathrm{e}^{(t-s)A}Bu(s)\, ds. \tag{8.2}$$

The function $h(t) = C\mathrm{e}^{tA}B$ is sometimes called the *impulse response*.

Often the control u is designed depending on the observation, and such systems are called *feedback systems*. If $u(t) = Ky(t)$, then the operator $K \in \mathcal{L}(Y, U)$ is called the feedback operator. Note that in this case we have the representation formula

$$x(t) = \mathrm{e}^{t(A+BKC)}x_0. \tag{8.3}$$

Now we list a few properties a time-invariant linear system can have, and which are important in view of applications.

Definition 8.1.1. A linear system $\Sigma(A, B, C)$ is said to be *externally positive*, if the output corresponding to the zero initial state is positive for every positive input function. In other words, $u(t) \geq 0$ implies $y(t) \geq 0$ if $x_0 = 0$.

Proposition 8.1.2. *A linear system is externally positive if and only if its impulse response is positive.*

Proof. By the representation formula (8.2) the sufficiency is clear. Suppose now that there is $t_0 > 0$ such that $h(t_0) = C\mathrm{e}^{t_0 A}B$ is not positive. Then, by continuity, at least one entry of $h(t)$ would be negative on a whole interval $[t_1, t_2]$. Thus, the appropriate entry of the output would be negative for every input function which is strictly positive in $[t - t_2, t - t_1]$ and zero elsewhere. Hence, the system cannot be externally positive. □

For a simple example of an externally positive system, see Exercise 2.

Definition 8.1.3. A linear system $\Sigma(A, B, C)$ is said to be *positive* (or internally positive), if the state and the output corresponding to a positive initial state is positive for every positive input function. In other words, $u(t) \geq 0$ and $x_0 \geq 0$ implies $x(t) \geq 0$ and $y(t) \geq 0$. The system is said to be *reducible*, if the matrix A is reducible, and *irreducible* otherwise.

Proposition 8.1.4. *A linear system is positive if and only if $B \geq 0$, $C \geq 0$, and A generates a positive matrix semigroup.*

Proof. Assume that the system is positive. Then, letting $x_0 = 0$ and $u(t) = u$, a nonnegative constant, we see that

$$0 \leq \frac{x(t)}{t} = \frac{1}{t} \int_0^t e^{(t-s)A} Bu \, ds = \left(\frac{1}{t} \int_0^t e^{sA} ds \right) Bu.$$

So, by letting $t \to 0$, we obtain the positivity of B.

Since $Cx_0 = Cx(0) = y(0) \geq 0$ for every $x_0 \geq 0$, the operator C has to be positive too. Finally, applying the zero control, we see that $x(t) = \mathrm{e}^{tA} x_0 \geq 0$ for every $x_0 \geq 0$.

To prove the converse implication, suppose that $B \geq 0$, $C \geq 0$, and that A generates a positive matrix semigroup. Taking $x_0 \geq 0$ and $u \geq 0$, we see that

$$\mathrm{e}^{tA} x_0 \geq 0$$

and that $Bu(s) \geq 0$ for each $s \in [0, t]$, hence $\mathrm{e}^{(t-s)A} Bu(s) \geq 0$, implying

$$\int_0^t \mathrm{e}^{(t-s)A} Bu(s) \, ds \geq 0,$$

which proves the statement by (8.1) and (8.2). □

Definition 8.1.5. A positive system is said to be *excitable*, if each state variable can be made strictly positive by applying an appropriate positive input to the system initially at rest. In other words, for each $i = 1, 2, \ldots, n$ there is a control u_i and a time t_i such that $x_i(t_i) > 0$ if $x_0 = 0$.

Excitable systems do have some remarkable properties. To be able to present some of them, we introduce some new concepts. To keep the presentation as simple as possible, we restrict ourselves for the rest of this section to the case $Y = U = \mathbb{C}$, i.e., we only consider one dimensional control and observation spaces.

The *influence graph* of the positive system $\Sigma(A, B, C)$ is a graph G with $n + 2$ nodes labeled with $0, 1, \ldots, n + 1$. The first (0) node is associated to the input u, and the last $(n + 1)$ is associated with the output y. The remaining nodes $(i = 1, \ldots, n)$ correspond to the state variables $x_1, \ldots, x_n$. The directed edges should represent somehow the influence among the variables and are constructed as follows:

- There is a directed edge from node 0 to node j if and only if $b_j > 0$, $j = 1, \ldots, n$.

- For $i, j = 1, \dots, n$, $i \neq j$, there is a directed edge from node i to node j if and only if $a_{ji} > 0$.
- There is a directed edge from node i to node $n+1$ if and only if $c_i > 0$, $i = 1, \dots, n$.

No other edges are present in the graph.

The corresponding graph matrices are constructed as follows: $\hat{A} = (\hat{a}_{ij})$, where $\hat{a}_{ij} = 1$ if and only if $i \neq j$ and $a_{ji} > 0$, otherwise $\hat{a}_{ij} = 0$. The row- and column- matrix $\hat{B}$ and $\hat{C}$, respectively, are constructed in a similar manner. The following $(n+2) \times (n+2)$ matrix

$$\mathbb{A} := \begin{pmatrix} 0 & \hat{B} & 0 \\ 0 & \hat{A} & \hat{C} \\ 0 & 0 & 0 \end{pmatrix}. \tag{8.4}$$

is thus the $0-1$ adjacency matrix of the unweighted influence graph.

Many properties of a positive linear system $\Sigma(A, B, C)$ can be described in terms of its influence graph. Observe, for example, that by Proposition 6.1.1 a system is irreducible if and only if the subgraph of its influence graph, consisting only from nodes $1, \dots, n$ and edges between them, is strongly connected.

We can also express excitability of the system in terms of the above graph matrices.

Proposition 8.1.6. *A positive linear system is excitable if and only if there exists at least one path from the input node* 0 *to each node* $i = 1, \dots, n$ *in the influence graph* G*, or, equivalently, if*

$$\hat{B} + \hat{B}\hat{A} + \cdots + \hat{B}\hat{A}^{n-1} \gg 0.$$

Proof. Excitability means that each state variable x_i can be influenced by the input u. This implies that there has to be at least one path from the node 0 to the node i in the influence graph G.

Note that the powers of the adjacency matrix $\mathbb{A}$ from (8.4) have the same block form:

$$\mathbb{A}^k = \begin{pmatrix} 0 & \hat{B}\hat{A}^{k-1} & \hat{B}\hat{A}^{k-2}\hat{C} \\ 0 & \hat{A}^k & \hat{A}^{k-1}\hat{C} \\ 0 & 0 & 0 \end{pmatrix}, \quad k \in \mathbb{N}.$$

Recall from Exercise 1.2 that the i^{th} component of the row-vector $\hat{B}\hat{A}^{k-1}$ represents the number of paths of length k from node 0 to node i, $i = 1, \dots, n$. Hence, there is a path to every node, if and only if

$$\hat{B} + \hat{B}\hat{A} + \cdots + \hat{B}\hat{A}^{n-1} \gg 0.$$

Assume that the positive system $\Sigma(A, B, C)$ is not excitable. Then, there exists $i \in \{1, \dots, n\}$ such for all t and all control $u \geq 0$,

$$x_i(t) = \left(\int_0^t e^{(t-s)A} Bu(s)\, ds \right)_i = 0.$$

By taking $u(t) = 1$, we obtain

$$b_i = \lim_{t\to 0} \left(\frac{1}{t} \int_0^t e^{sA} ds B \right)_i = 0.$$

On the other hand,

$$\dot{x}_i(t) = \left(Bu(t) - A \int_0^t e^{(t-s)A} Bu(s)\, ds \right)_i = 0.$$

This implies that

$$\left(A \int_0^t e^{(t-s)A} Bu(s)\, ds \right)_i = 0.$$

As above, by taking $u(t) = 1$, we obtain $(AB)_i = \sum_{j=1}^n a_{ij} b_j = 0$. Using the positivity, we deduce that $a_{ij} b_j = 0$ for all j (note that by above $b_i = 0$, so also $a_{ii} b_i = 0$). Hence for the graph matrices we have $(\hat{B}\hat{A})_i = 0$. By repeating the same arguments we obtain $(\hat{B}\hat{A}^k)_i = 0$ for all $k = 0, \dots, n-1$, and this ends the proof of the proposition. □

We consider now rather special constant inputs, $u(t) = \bar{u} > 0$.

Theorem 8.1.7. *An excitable and asymptotically stable positive linear system has a strictly positive equilibrium state.*

Proof. Since by asymptotic stability all the eigenvalues of A have negative real part, A is invertible and $\bar{x} := -A^{-1} B\bar{u}$ is the unique equilibrium of the system, which is asymptotically stable. We only have to show that it is strictly positive.

Suppose that there are indices i such that $\bar{x}_i = 0$, and collect these indices in the set $I := \{i \in \{1, \dots, n\} \,:\, \bar{x}_i = 0\}$. Then, since

$$A\bar{x} + B\bar{u} = 0,$$

we see that

$$\sum_{j \notin I} a_{ij} \bar{x}_j + b_i \bar{u} = 0 \qquad \text{for } i \in I.$$

This implies that $b_i = 0$ and that $a_{ij} = 0$ for $i \in I$ and $j \notin I$. Hence, there is no path from the input node 0 to vertices $i \in I$. Hence, in this case the system is not excitable.

□

8.2 Controllability

For simplicity, we consider here systems without observation, i.e., where $Y = X$ and $C = I$.

Definition 8.2.1. The system $\Sigma(A, B)$ is called *controllable in time* τ if for every initial value $x_0 \in X$ and every state $x_1 \in X$ there is a control u such that for the solution x we have $x(\tau; x_0) = x_1$.

We will shortly call a system *controllable*, if there exists a $\tau \geq 0$ such that it is controllable in time τ.

Lemma 8.2.2. *The system $\Sigma(A,B)$ is controllable in time τ if and only if every state $x_1 \in X$ can be reached from $x_0 = 0$ in time τ.*

Proof. We have only to prove the converse. Let us take $x_1 \in X$ and set $x_2 = x_1 - \mathrm{e}^{\tau A}x_0$. Then, by assumption, there is a control u such that $x_2 = x(\tau; 0)$. Then

$$x(\tau;0) = \int_0^\tau \mathrm{e}^{(\tau-s)A}Bu(s)\,ds = x_2,$$

hence

$$x(\tau;x_0) = \mathrm{e}^{\tau A}x_0 + \int_0^\tau \mathrm{e}^{(\tau-s)A}Bu(s)\,ds = \mathrm{e}^{\tau A}x_0 + x_2 = x_1.$$

□

To investigate the possible reachable states, we take a functional analytic point of view and introduce an operator which maps control functions into states which are reached from the origin by using this control.

Definition 8.2.3. Fix $\tau > 0$. The *controllability operator* $\mathcal{B}_\tau : L^1([0,\tau],U) \to X$ is defined by

$$\mathcal{B}_\tau(u) := \int_0^\tau \mathrm{e}^{(\tau-s)A}Bu(s)\,ds.$$

Hence, the system is controllable in time τ if and only if $\mathcal{B}_\tau$ is surjective.

Lemma 8.2.4. *The operator $\mathcal{B}_\tau$ has the following properties:*

1. *The operator $\mathcal{B}_\tau$ is linear.*
2. *The operator $\mathcal{B}_\tau : L^1([0,\tau],U) \to X$ is bounded, i.e.,*

$$\sup_{\|u\|\leq 1} \|B_\tau(u)\| < \infty.$$

The proof is left as Exercise 3.

Fortunately there is an important characterization of the range of $\mathcal{B}_\tau$.

Theorem 8.2.5. *For every $\tau > 0$ we have*

$$\mathrm{im}(\mathcal{B}_\tau) = \mathrm{span}\left\{x,\ Ax,\ A^2x, \ldots, A^{n-1}x \,:\, x \in \mathrm{im}(B)\right\}.$$

Proof. Let us introduce first some shorthand notation and denote for this proof

$$X_1 := \operatorname{span}\left\{x,\ Ax,\ A^2x, \dots, A^{n-1}x \ :\ x \in \operatorname{im}(B)\right\}.$$

Let us introduce two further spaces,

$$\begin{aligned} X_2^\tau &:= \operatorname{span}\left\{\mathrm{e}^{tA}y \ :\ 0 \le t \le \tau,\ y \in \operatorname{im}(B)\right\}, \\ X_3^\tau &:= \operatorname{span}\left\{\int_0^t \mathrm{e}^{sA}y\ ds \ :\ 0 \le t \le \tau,\ y \in \operatorname{im}(B)\right\}. \end{aligned}$$

Since step functions are dense in L^1, and using that in X every subspace is automatically closed, we can conclude by the continuity of $\mathcal{B}_\tau$ that

$$\operatorname{im}(\mathcal{B}_\tau) = X_3^\tau.$$

Observe also that, using Theorem 2.2.5, Formula (2.8) we see that e^{tA} is a polynomial in A of degree at most $n-1$, and thus, $X_2^\tau = X_1$.

Now let us take $y \in \operatorname{im}(B)$. Then $x(t) = \int_0^t \mathrm{e}^{sA}y\ ds \in X_3^\tau$ for $t \le \tau$. Clearly, all the derivatives of x lie in X_3^τ, hence

$$\begin{aligned} x(0) &= 0 \in X_3^\tau, \\ \dot{x}(0) &= y \in X_3^\tau, \\ \ddot{x}(0) &= Ay \in X_3^\tau, \\ &\text{etc.} \end{aligned}$$

Hence, $X_1 \subset X_3^\tau$. On the other hand, if $y \in \operatorname{im}(B)$, then $\mathrm{e}^{sA}y \in X_1$, implying that

$$\int_0^t \mathrm{e}^{sA}y\ ds \in X_1,$$

i.e., $X_3^\tau \subset X_1$. □

Corollary 8.2.6 (Kálmán criterion)**.** *For a control system $\Sigma(A, B)$ the following are equivalent.*

(i) The system is controllable in time τ for some $\tau > 0$.

(ii) The controllability operator $\mathcal{B}_\tau$ is surjective for every $\tau > 0$.

(iii) The rank condition $\operatorname{rank}(B, AB, A^2B, \dots, A^{n-1}B) = n$ *holds.*

In many applications it is natural to consider only positive initial values, positive controls, and expect the states of the system to be in the positive cone for all times. Hence, we restrict our investigations here to this case only.

The *reachability set* of a positive system X_τ^+ is defined as the set of points that can be reached from the origin by applying a positive control. In other words,

$$X_\tau^+ = \left\{ x \in X,\, x \geq 0 \,:\, \exists u \geq 0 \text{ s.t. } x = \int_0^\tau \mathrm{e}^{(\tau-s)A} Bu(s)\, ds \right\}.$$

By linearity and positivity of the operators, the set X_τ^+ is a convex cone. Actually, much more can be said.

Theorem 8.2.7. *The set X_τ^+ is a convex cone which is non-degenerate (i.e., it contains an open ball) if and only if the positive system $\Sigma(A,B)$ is controllable, i.e., the Kálmán rank condition holds.*

The proof is left as Exercise 7.

An important case is when X_τ^+ is actually the whole positive cone in X, which would correspond to the notion of "positive controllability", i.e., when each positive state can be reached by applying a positive control from the origin. This is a much more delicate question with more sophisticated results and we leave it to the project phase. See also Exercise 8 for a simple demonstration.

8.3 Stabilization

We restrict ourselves here again to systems without observation and where the control is given by a suitable feedback K.

Definition 8.3.1. A system $\Sigma(A,B)$ is called *stabilizable* if there is a feedback K such the the state converges to zero for every initial value, i.e.,

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} \mathrm{e}^{t(A+BK)} x_0 = 0 \tag{8.5}$$

for every $x_0 \in X$.

Note that by Theorem 4.4.3 a system is stabilizable if and only if there is a feedback K such that $s(A+BK) < 0$.

There is an important connection between stabilization and controllability. To this end, let us denote by $\lambda_1, \ldots \lambda_n$ the eigenvalues of A and X_i the corresponding spectral subspaces, and

$$X_+ := \oplus_{\operatorname{Re}\lambda_i \geq 0} X_i$$

be the subspace corresponding to the eigenvalues with positive real part, and denote by A_+ the part of A restricted to this subspace. Then B can also be decomposed into $B_+ : U \to X_+$ and $B_- : U \to X \ominus X_+$. The following is rather useful because stabilizability can be checked without actually constructing the feedback.

Theorem 8.3.2. *The system $\Sigma(A,B)$ is stabilizable if and only if the subsystem $\Sigma(A_+,B_+)$ is controllable.*

Proof. Assume first that there is a $K \in \mathcal{L}(X)$ such that the semigroup generated by $A+BK$ is exponentially stable. Suppose now that the system $\Sigma(A_+, B_+)$ is not controllable. Let $\tau > 0$.Then the range of $\mathcal{B}_\tau$ is not the whole space X_+, hence there is $z_0 \in X_+$ such that

$$z_0 \perp \int_0^\tau \mathrm{e}^{(\tau-s)A_+} B_+ u(s)\, ds \quad \text{for all } u \in L^1([0,\tau], U).$$

Note that by the variation of constants formula (8.1) we get

$$\mathrm{e}^{\tau(A+BK)} x = \mathrm{e}^{\tau A} x + \int_0^\tau \mathrm{e}^{(\tau-s)A} BK \mathrm{e}^{s(A+BK)} x\, ds,$$

hence

$$P_+ \mathrm{e}^{\tau(A+BK)} x = \mathrm{e}^{\tau A_+} x + \int_0^\tau \mathrm{e}^{(\tau-s)A_+} B_+ K \mathrm{e}^{s(A+BK)} x\, ds.$$

By the stability of the semigroup generated by $A + BK$ and by the choice of z_0 we see that

$$\langle \lim_{t\to\infty} P_+ \mathrm{e}^{\tau(A+BK)} x, z_0 \rangle = \langle \lim_{t\to\infty} \mathrm{e}^{\tau A_+} x, z_0 \rangle = 0.$$

But the eigenvalues of A_+ all have positive real parts, hence the last equality can only hold for all $x \in X_+$ if $z_0 = 0$.

The proof of the other direction is quite demanding and will not be presented here. We mention only that it is a consequence of the solvability of the so-called "pole placement problem". □

A positive system $\Sigma(A, B)$ is called *positively stabilizable*, if there is a positive feedback operator K such that (8.5) holds for every $x_0 \geq 0$.

Proposition 8.3.3. *A positive system is positively stabilizable if and only if it is stabilizable with a positive feedback.*

Proof. Note that for every element in X its real part and imaginary part can be represented as the difference of two positive elements in X^+. Since

$$\lim_{t\to\infty} \mathrm{e}^{t(A+BK)} x_0 = 0$$

for every x_0 is equivalent to

$$\lim_{t\to\infty} \mathrm{e}^{t(A+BK)} (x_1 - x_2) = 0$$

for every $x_1, x_2 \geq 0$, the statement follows. □

8.4 Notes and Remarks

There are many excellent introductions to systems and control theory. We have relied our preparations on the monograph by B. Jacob and H. Zwart [2] (which is based upon a former Internet Seminar), on the manuscript by V. Mehrmann [3], and on the monograph by J. Zabczyk [4].

Positivity aspects of control problems are discussed in the monograph by L. Farina and S. Rinaldi [1]. Many further interesting topics could be studied here, and in case we succeeded to make you curious, you can look them up in the above mentioned sources.

8.5 Exercises

1. Prove formula (8.1).

2. Suppose that $U = Y = \mathbb{C}$ and $X = \mathbb{C}^2$, and let

$$A = \begin{pmatrix} -a & -a \\ 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} a \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

   for a parameter $a > 0$. For which values of a will the system $\Sigma(A, B, C)$ be externally positive?

3. Prove the basic properties of the controllability operator $\mathcal{B}_\tau$ as stated in Lemma 8.2.4.

4. Show that a system $\Sigma(A, B)$ is controllable if and only if for every eigenvector v of A^T we have $vB \neq 0$.

5. Show that a system $\Sigma(A, B)$ is controllable if and only if $\operatorname{rank}(\lambda - A, B) = n$ for all $\lambda \in \mathbb{C}$.

6. Let $U = \mathbb{C}$ and $X = \mathbb{C}^2$, and consider

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -a \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

   with $a > 0$. What can you say about the reachability set X_τ^+ of this positive linear system? In other words, which states can be reached from the origin by applying a positive control u in time τ?

7. Show Theorem 8.2.7.

8. Let $U = \mathbb{C}$ and $X = \mathbb{C}^2$, and consider

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 \\ a \end{pmatrix}$$

   with $a > 0$. Is the system $\Sigma(A, B)$ stabilizable? Is it positively stabilizable?

9. Show that a system $\Sigma(A, B)$ is stabilizable if and only if $\operatorname{rank}(\lambda - A, B) = n$ for all $\operatorname{Re}\lambda \geq 0$.

Bibliography

[1] Lorenzo Farina and Sergio Rinaldi, *Positive linear systems*, Pure and Applied Mathematics (New York), Wiley-Interscience, New York, 2000, Theory and applications.

[2] Birgit Jacob and Hans J. Zwart, *Linear port-Hamiltonian systems on infinite-dimensional spaces*, Operator Theory: Advances and Applications, vol. 223, Birkhäuser/Springer Basel AG, Basel, 2012, Linear Operators and Linear Systems.

[3] V. Mehrmann, *Kontrolltheorie*, lecture notes, 2004.

[4] J. Zabczyk, *Mathematical Control Theory*, Systems & Control: Foundations & Applications, Birkhäuser Verlag, 1992.

Lecture 9

Banach Lattices

In this chapter we give a brief introduction to Banach lattices. To do so, we first try to distillate the essential ideas from finite dimensional theory.

9.1 Ordered Function Spaces

Let us first summarize the order structure of $\mathbb{R}^N$. The key point is that vectors in $\mathbb{R}^N$ can be usually identified with functions:

$$\mathbb{R}^N \equiv \{f : \{1, \dots, N\} \to \mathbb{R}\}.$$

Positivity of a vector is thus nothing but pointwise positivity of the representing function:

$$f \geq 0 \text{ if and only if } f(k) \geq 0 \text{ for all } k = 1, \dots, N.$$

Hence, if we have a vector space of real valued functions, it is natural to introduce an order relation on it by pointwise ordering. Let us illustrate this with the most important example.

If $K \subset \mathbb{R}^p$ is a compact set, then we may take the set of continuous functions

$$X = \mathrm{C}(K, \mathbb{R}) := \{f : K \to \mathbb{R} \, : \, f \text{ is continuous}\},$$

which is a Banach space with the norm

$$\|f\| = \|f\|_\infty = \max_{x \in K} |f(x)|.$$

The pointwise ordering in this case is

$$f \geq g \iff f(x) \geq g(x) \text{ for all } x \in K.$$

This clearly generalizes the finite dimensional case with $K = \{1, \dots, N\} \subset \mathbb{R}$ and the usual maximum norm.

It is straightforward from the definition that the ordering is compatible with the vector space operations in the sense that

$$f \leq g \text{ implies } f + h \leq g + h \text{ for all } h \in \mathrm{C}(K, \mathbb{R})$$

and

$$0 \leq f \text{ implies } 0 \leq tf \text{ for all } t \in \mathbb{R},\ t \geq 0.$$

We can also define supremum and infimum of two functions as

$$(f \vee g)(x) := \max\{f(x), g(x)\} \quad \text{and} \quad (f \wedge g)(x) := \min\{f(x), g(x)\}$$

for all $x \in K$. The positive and negative part and absolute value of a function can be then given as

$$f^+ := f \vee 0, \quad f^- := (-f) \vee 0, \quad |f| := f \vee (-f).$$

An important property of the positive and negative part of a function is that they live separate lives: if $f^+(x) \neq 0$, then $f^-(x) = 0$ and vice versa. This property is sometimes called orthogonality.

Note that the following properties also follow from the fact that we defined the order relation pointwise and that the order behaves nicely on the real numbers:

$$\begin{array}{rcl} f & = & f^+ - f^-, \\ |f| & = & f^+ + f^-, \\ f \leq g & \Longleftrightarrow & f^+ \leq g^+ \text{ and } g^- \leq f^-, \\ |f - g| & = & (f \vee g) - (f \wedge g), \\ |f| \leq |g| & \Longrightarrow & \|f\| \leq \|g\|. \end{array} \tag{9.1}$$

We suggest that you check these properties immediately.

Recall that for irreducibility in Lecture 5 (see Definition 5.2.1), we needed the invariance of a subspace of the form

$$J_M := \big\{(\xi_1, \dots, \xi_N) : \xi_i = 0 \text{ for } i \in M\big\} \subset \mathbb{R}^N$$

for some $\emptyset \neq M \subsetneqq \{1, \dots, N\}$. In analogy, we may define the following. Suppose that $A \subset K$ is a closed set and define

$$J_A := \{f \in \mathrm{C}(K, \mathbb{R}) \,:\, f(x) = 0 \text{ for all } x \in A\}\,. \tag{9.2}$$

Because of their special properties subspaces of the above form are also called *ideals*, see also Exercise 9.5.1. It is extremely important that ideals can be characterized by order theoretic concepts.

Proposition 9.1.1. *For a closed subspace $I \subset C(K, \mathbb{R})$ the following are equivalent:*

(a)

$$f \in I \textit{ implies } |f| \in I,$$

and

$$0 \le g \le f \in I \textit{ implies } g \in I.$$

(b) There is closed set $A \subset K$ such that $I = J_A$.

Proof. The case $I = \{0\}$ is trivial. Let us assume that $I \neq \{0\}$.

Clearly, if $I = J_A$ for a closed subset A, then the listed properties in (a) hold.

For the other direction, define

$$A := \{x \in K \,:\, f(x) = 0 \text{ for all } f \in I\}.$$

With this notation, clearly $I \subset J_A$. Now take a positive function

$$0 \neq f \in J_A, \quad f \geq 0.$$

Our aim is to show that $f \in I$. For a $\varepsilon > 0$ let $B^f_\varepsilon := \{x \in K \,:\, f(x) \geq \varepsilon\} =: [f \geq \varepsilon]$. The set B^f_ε is closed satisfying $B^f_\varepsilon \cap A = \emptyset$. Thus, for every $x \in B^f_\varepsilon$ there is $0 \le g_x \in I$ such that $g_x(x) > 0$. Since B^f_ε is compact, there are finitely many $x_1, \ldots, x_r \in B^f_\varepsilon$ such that

$$B^f_\varepsilon \subset [g_{x_1} > 0] \cup [g_{x_2} > 0] \cup \cdots \cup [g_{x_r} > 0],$$

where we used the notation $[g_{x_i} > 0] := \{x \in K \,:\, g_{x_i}(x) > 0\}$.

Now we construct an approximation of f in the set I. First observe that from (a) and properties (9.1) it follows that $f_1, f_2 \in I$ implies $f_1 \vee f_2 \in I$ and $f_1 \wedge f_2 \in I$. We define

$$g := g_{x_1} \vee g_{x_2} \vee \cdots \vee g_{x_r} \in I,$$

and take $\delta > 0$ such that $g(x) \geq \delta$ for all $x \in B^f_\varepsilon$. Then the function

$$h := f \wedge \left(\frac{\|f\|}{\delta} g\right) \in I$$

satisfies $0 \le h \le f$ and $h(x) = f(x)$ for all $x \in B^f_\varepsilon$. It follows by the definition of the set B^f_ε that $\|f - h\| \le \varepsilon$. Hence, for every $f \in J_A$ and every $\varepsilon > 0$ we found $h \in I$ such that h approximates f with an error less than ε. By the closedness of I we obtain the desired conclusion. □

A subspace I of $\mathrm{C}(K, \mathbb{R})$ is called *ideal* if the condition (a) in Proposition 9.1.1 is satisfied.

Another important observation concerning ideals is the following. Taking $f \geq 0$, we would like to build the smallest ideal containing f, and denote it by J_f. It is then really straightforward to check that

$$J_f = \bigcup_{n \in \mathbb{N}} [-nf, nf]$$

holds, where $[f_1, f_2] := \{g : f_1 \le g \le f_2\}$. We call J_f *the ideal generated by* f.

In some proofs (like in Corollary 7.1.3) strictly positive vectors (or the vector $\mathbb{1}$) played an important role. A natural observation is the fact that the ideal generated by a strictly positive function is the whole space $\mathrm{C}(K, \mathbb{R})$. A function with this property is sometimes also called an *order unit*. Unfortunately, as we shall see in the next section, not all important function spaces possess order units. We will be able to introduce a weaker notion that will be almost as satisfactory for our proofs, see Example 9.3.8 and the considerations before.

Finally, let us note that for statements in spectral theory we need complex vector spaces. Note however that

$$\mathrm{C}(K, \mathbb{C}) \cong \mathrm{C}(K, \mathbb{R}) \oplus i \cdot \mathrm{C}(K, \mathbb{R}),$$

meaning that for a complex valued continuous function its real part and imaginary parts are real valued continuous functions. Hence there will be no problems speaking about spectrum of an operator.

9.2 Vector lattices

Now we take an abstract point of view and try to axiomatise what we have seen in the previous section to have a terminology at hand which is applicable to many concrete situations. Our main examples, besides the finite dimensional vector spaces, are $\mathrm{C}(K)$ spaces, $\mathrm{L}^p(\Omega, \mu)$ spaces and $\mathrm{C}_0(\Omega)$ spaces (see Example 9.3.2 later on). If you are uncomfortable with abstract terminology, you should pick one of these spaces and have it in mind for the rest of this lecture.

We start by ordering. A non empty set M with a relation $\le$ is said to be an *ordered set* if the following conditions are satisfied.

i) $x \le x$ for every $x \in M$,

ii) $x \le y$ and $y \le x$ implies $x = y$, and

iii) $x \le y$ and $y \le z$ implies $x \le z$.

First examples of ordered sets are number sets: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

Having ordering at hand, we can consider boundedness. Let A be a subset of an ordered set M. The element $x \in M$ (resp. $z \in M$) is called *an upper bound* (resp. *lower bound*) of A if $y \le x$ for all $y \in A$ (resp. $z \le y$ for all $y \in A$). Moreover, if there is an upper bound (resp. lower bound) of A, then A is said *bounded from above* (resp. *bounded from below*). If A is bounded from above and from below, then it is called *order bounded set*.

We can introduce the concept of an interval analogous to the interval on the real line. Let $x, y \in M$ such that $x \le y$. We denote by

$$[x, y] := \{z \in M : x \le z \le y\}$$

the *order interval* between x and y. It is obvious that a subset A is order bounded if and only if it is contained in some order interval.

Definition 9.2.1. A real vector space E which is ordered by some order relation $\leq$ is called a *vector lattice* if any two elements $x, y \in E$ have a least upper bound denoted by $x \vee y = \sup(x, y)$ and a greatest lower bound denoted by $x \wedge y = \inf(x, y)$ and the following properties are satisfied.

(L1) $x \leq y$ implies $x + z \leq y + z$ for all $x, y, z \in E$,

(L2) $0 \leq x$ implies $0 \leq tx$ for all $x \in E$ and $t \in \mathbb{R}_+$.

Let E be a vector lattice. We denote by $E_+ := \{x \in E : 0 \leq x\}$ the *positive cone* of E. For $x \in E$, let us define

$$x^+ := x \vee 0, \quad x^- := (-x) \vee 0, \quad \text{and} \quad |x| := x \vee (-x)$$

the *positive part*, the *negative part*, and the *absolute value* of x, respectively.

Two elements $x, y \in E$ are called *orthogonal* (or *lattice disjoint*) (denoted by $x \perp y$) if $|x| \wedge |y| = 0$.

For a vector lattice E we have the following properties, which we will use frequently.

Proposition 9.2.2. *For all $x, y, z \in E$ the following assertions are satisfied.*

(i) $x + y = (x \vee y) + (x \wedge y)$.

(ii) $x \vee y = -(-x) \wedge (-y)$.

(iii) $(x \vee y) + z = (x + z) \vee (y + z)$ *and* $(x \wedge y) + z = (x + z) \wedge (y + z)$.

(iv) $(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z)$ *and* $(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z)$.

(v) For all $x, y, z \in E_+$ we have $(x + y) \wedge z \leq (x \wedge z) + (y \wedge z)$.

Proof. We shall only prove (i). The proof of the other properties is left as an exercise (cf. Exercise 2). We have $x \wedge y \leq y \Rightarrow x \leq x + y - x \wedge y$. In a similar way we have $y \leq x + y - x \wedge y$. Hence, $x \vee y \leq x + y - x \wedge y$, which gives

$$x \vee y + x \wedge y \leq x + y.$$

For the reverse inequality we note that $y \leq x \vee y \Rightarrow x + y - x \vee y \leq x$ and similarly $x + y - x \vee y \leq y$. Thus,

$$x + y - x \vee y \leq x \wedge y.$$

□

For positive, negative part, and absolute value of $x \in E$ we have the following useful properties (compare with properties (9.1) of functions in $\mathrm{C}(K, \mathbb{R})$).

Proposition 9.2.3. *If $x, y \in E$, then*

(i) $x = x^+ - x^-$;

(ii) $|x| = x^+ + x^-$;

(iii) $x^+ \perp x^-$ *and the decomposition of* x *into the difference of two orthogonal positive elements is unique;*

(iv) $x \leq y$ *is equivalent to* $x^+ \leq y^+$ *and* $y^- \leq x^-$;

(v) $|x - y| = (x \vee y) - (x \wedge y)$.

Proof. (i): Using (i) and (ii) of Proposition 9.2.2, we obtain

$$\begin{aligned} x &= x + 0 = x \vee 0 + x \wedge 0 \\ &= x \vee 0 - (-x) \vee 0 = x^+ - x^-. \end{aligned}$$

(ii): Applying Proposition 9.2.2-(iii) and (i) proved above we have

$$\begin{aligned} x \vee (-x) &= (2x \vee 0) - x = 2(x \vee 0) - x \\ &= 2x^+ - x^+ + x^- = x^+ + x^-. \end{aligned}$$

(iii): Let us prove first that $x^+ \wedge x^- = 0$. To this purpose we apply Proposition 9.2.2-(iii) and deduce

$$\begin{aligned} x^+ \wedge x^- &= (x^+ - x^-) \wedge 0 + x^- = (x \wedge 0) + x^- \\ &= -[(-x) \vee 0] + x^- = 0. \end{aligned}$$

Now, let $x = y - z$ with $y \wedge z = 0$. By (iii) and (i) of Proposition 9.2.2, we have $x^+ = (y - z) \vee 0 = y \vee z - z = (y + z - (y \wedge z)) - z = y$. In a similar way we get $x^- = z$.
(iv): It is easy to prove and is left as an exercise.
(v): This can be proved using the identities

$$\begin{aligned} x \vee y &= \frac{1}{2}(x + y + |x - y|) \\ x \wedge y &= \frac{1}{2}(x + y - |x - y|). \end{aligned}$$

□

9.3 Banach lattices

We have finally arrived to the main object of this lecture. In this section we will consider Banach spaces which are ordered and whose norm is compatible with this ordering. First, let us explain what we mean by compatible.

A norm on a vector lattice E is called a *lattice norm* if

$$|x| \leq |y| \text{ implies } \|x\| \leq \|y\| \quad \text{for } x, y \in E.$$

Definition 9.3.1. A *Banach lattice* is a real Banach space E endowed with an ordering $\leq$ such that $(E, \leq)$ is a vector lattice and the norm on E is a lattice norm.

We will see that this combination of properties of a complete normed vector space and a compatible ordering will lead us to many fruitful results.

As already announced, apart from finite dimensional vector spaces (such as $\mathbb{R}$ or $\mathbb{R}^N$), there are many interesting infinitely dimensional examples of Banach lattices.

Example 9.3.2. The following Banach spaces are Banach lattices for the pointwise (almost everywhere) ordering (see Exercise 9.5.4).

- $\mathrm{L}^p(\Omega, \mu)$, $1 \leq p \leq \infty$ endowed with the norm

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \, d\mu \right)^{\frac{1}{p}}, \quad \text{if } 1 \leq p < \infty,$$

  and

$$\|f\|_\infty = \inf\{M : |f(x)| \leq M \ \mu - \text{a.e.}\, x \in \Omega\} \quad \text{if } p = \infty,$$

  and with the order

$$f \geq 0 \iff f(x) \geq 0 \text{ for } \mu\text{-a.e. } x \in \Omega,$$

  where μ is σ-finite measure on a set Ω,

- $\mathrm{C}_0(\Omega)$, the space of all real-valued continuous functions vanishing at infinity, endowed with the supremum norm

$$\|f\|_\infty = \sup_{x \in \Omega} |f(x)|,$$

  and with the natural order

$$f \geq 0 \iff f(x) \geq 0 \text{ for all } x \in \Omega,$$

  where Ω is locally compact,

- $\mathrm{C}(K)$, the space of real-valued continuous functions on a compact set K, endowed with the supremum norm and with the order defined above were already investigated in Section 9.1.

Note that there are many naturally ordered function spaces which are not Banach lattices, see for example Exercise 9.5.7.

Let us now list some further nice properties of a Banach lattice.

Proposition 9.3.3. *For a Banach lattice E the following holds.*

(a) The lattice operations are continuous.

(b) The positive cone E_+ is closed, and

(c) order intervals are closed and bounded.

Proof. The proof is left as an exercise (see Exercise 5). □

The following nice property of Banach lattices is a consequence of the Hahn-Banach theorem (see Appendix B).

Proposition 9.3.4. *In a Banach lattice E every weakly convergent increasing sequence (x_n) is norm-convergent.*

Proof. Let $A := \{\sum_{i=1}^{n} a_i x_i : n \in \mathbb{N}, a_i \geq 0, a_1 + \cdots + a_n = 1\}$ be the convex hull of the set $\{x_n : n \in \mathbb{N}\}$. By the Hahn-Banach theorem, the norm-closure of A coincides with the weak closure. This implies that $x \in \overline{A}$, where $x :=$ weak $- \lim_{n\to\infty} x_n$. Thus, for $\varepsilon > 0$, there exist $y \in A$, i.e.

$$y = a_1 x_1 + \cdots + a_n x_n, \text{ with } a_1, \ldots, a_n \geq 0 \text{ and } a_1 + \cdots + a_n = 1,$$

such that $\|y - x\| < \varepsilon$. Since $y \leq x_k \leq x$, it follows that $\|x - x_k\| \leq \|x - y\| < \varepsilon$ for all $k \geq n$. □

Here we state a result we shall need later on.

Lemma 9.3.5. *Let E be a totally ordered (this means $x \in E \Rightarrow 0 \leq x$ or $x \leq 0$) real Banach lattice. Then* $\dim E \leq 1$.

Proof. Let $e \in E_+$ and $x \in E$. We consider the closed subsets $C_+ := \{\alpha \in \mathbb{R} : \alpha e \geq x\}$ and $C_- := \{\alpha \in \mathbb{R} : \alpha e \leq x\}$ of $\mathbb{R}$. It is obvious that C_+ and C_- are non empty and $C_+ \cup C_- = \mathbb{R}$. Since $\mathbb{R}$ is connected, it follows that $C_+ \cap C_- \neq \varnothing$. Hence there is $\alpha \in \mathbb{R}$ such that $x = \alpha e$. □

9.3.1 Sublattices and ideals

A vector subspace F of a vector lattice E is a *vector sublattice* if and only if the following are satisfied.

(1) $|x| \in F$ for all $x \in F$,

(2) $x^+ \in F$ or $x^- \in F$ for all $x \in F$.

A subspace I of a Banach lattice E is called *ideal* if

$$x \in I \text{ implies } |x| \in I \text{ and } 0 \leq y \leq x \in I \text{ implies } y \in I$$

(compare with Proposition 9.1.1!).

Consequently, a vector sublattice F is an ideal in E if $x \in F$ and $0 \leq y \leq x$ implies $y \in F$. Since the notions of sublattice and ideal are invariant under the formation of arbitrary intersections, there exists, for any subset M of E, a uniquely determined smallest sublattice (resp. ideal) of E containing M. This will be called *the sublattice (resp. the ideal) generated by M.*

We summarize all important properties of sublattices and ideals which we will need in the sequel.

Proposition 9.3.6. *If E is a Banach lattice, then the following properties hold.*

(i) The closure of every sublattice of E is a sublattice.

(ii) The closure of every ideal of E is an ideal.

(iii) For every $x \in E_+$, the ideal generated by $\{x\}$ is

$$E_x = \bigcup \{n[-x, x] : n \in \mathbb{N}\}.$$

Proof. The first two assertion follows from the continuity of the lattice operations, see Proposition 9.3.3. For the last assertion one can see easily that $I = \bigcup \{n[-x, x] : n \in \mathbb{N}\}$ is an ideal while any ideal included in I and containing x equals I. This means that $I = E_x$. □

For examples of closed ideals we again pay a visit to our function spaces, see also Exercise 9.5.10.

Example 9.3.7. (1) If $E = \mathrm{L}^p(\Omega, \mu)$, $1 \leq p < \infty$, where μ is a $\sigma-$finite measure, then the closed ideals in E are characterized as follows: A subspace I of E is a closed ideal if and only if there exists a measurable subset Y of Ω such that

$$I = \{\psi \in E : \psi(x) = 0 \text{ a.e. } x \in Y\}.$$

(2) If $E = \mathrm{C}_0(\Omega)$, where Ω is a locally compact topological space, then a subspace J of E is a closed ideal if and only if there is a closed subset A of Ω such that

$$J = \{\varphi \in E : \varphi(x) = 0 \text{ for all } x \in A\}.$$

Sometimes the Banach lattice E is generated by a single positive element. If $E_e = E$ holds for some $e \in E_+$ then e is called an *order unit*. If $\overline{E_e} = E$, then $e \in E_+$ is called a *quasi interior point* of E_+.
It follows that e is an order unit of E if and only if e is an interior point of E_+. Quasi interior points of the positive cone exist, for example, in every separable Banach lattice.

Example 9.3.8. (1) If $E = \mathrm{C}(K)$, where K is a compact set, then the function constant $\mathbb{1}$, $\mathbb{1}(x) \equiv 1$, is an order unit. In fact, for every $f \in E$, there is $n \in \mathbb{N}$ such that $\|f\|_\infty \leq n$. Hence, $|f(s)| \leq n\mathbb{1}(s)$ for all $s \in K$. This implies $f \in n[-\mathbb{1}, \mathbb{1}]$.

(2) If $E = L^p(\mu)$ with $\sigma-$finite measure μ and $1 \leq p < \infty$, then the quasi interior points of E_+ coincide with the $\mu-$a.e. strictly positive functions, while E_+ does not contain any interior point. Here measure spaces with atoms are excluded (see Exercise 9.5.11).

9.3.2 Complexification of real Banach lattices

It is often necessary to consider complex vector spaces (for instance in spectral theory). Therefore, we introduce the concept of a complex Banach lattice.

The complexification of a real Banach lattice E is the complex Banach space $E_{\mathbb{C}}$ whose elements are pairs $(x, y) \in E \times E$, with addition and scalar multiplication defined by $(x_0, y_0) + (x_1, y_1) := (x_0 + x_1, y_0 + y_1)$ and $(a + ib)(x, y) := (ax - by, ay + bx)$, and norm

$$\|(x, y)\| := \| \, |(x, y)| \, \|,$$

where

$$|(x, y)| := \sup_{0 \le \theta \le 2\pi} (x \sin\theta + y \cos\theta)$$

is the natural extension of the modulus $|\cdot|$ in E. One can show that the above supremum exists in E (cf. [4], p. 134). By identifying $(x, 0) \in E_{\mathbb{C}}$ with $x \in E$, E is isometrically isomorphic to a real linear subspace of $E_{\mathbb{C}}$. We write $0 \le x \in E_{\mathbb{C}}$ if and only if $x \in E_+$.

A complex Banach lattice is an ordered complex Banach space $(E_{\mathbb{C}}, \le)$ that arises as the complexification of a real Banach lattice E. The underlying real Banach lattice E is called the real part of $E_{\mathbb{C}}$ and is uniquely determined as the closed linear span of all $x \in (E_{\mathbb{C}})_+$.

Instead of the notation (x, y) for elements of $E_{\mathbb{C}}$, we usually write $x + iy$. The complex conjugate of an element $z = x + iy \in E_{\mathbb{C}}$ is the element $\overline{z} = x - iy$. We use also the notation $\mathrm{Re}(z) := x$ for $z = x + iy \in E_{\mathbb{C}}$. All concepts first introduced for real Banach lattices have a natural extension to complex Banach lattices.

9.4 Notes and Remarks

The investigation of ordered algebraic structures is a classical subject and of great interest in the literature, we mention here the classical monograph by L. Fuchs [2]. Most results of this Lecture can be found for example in the monographs by Schaefer [4], Meyer-Nieberg [3] or Aliprantis and Burkinshaw [1]. For Proposition 9.3.6 see [3, Propositions 1.1.5, 1.2.3, and 1.2.5]).

9.5 Exercises

1. Let $K \subset \mathbb{R}^p$ be a bounded, closed set. Show that a subspace $J \subset \mathrm{C}(K, \mathbb{R})$ is an algebraic ideal in the Banach algebra $\mathrm{C}(K)$ if and only if $J = J_A$ defined in (9.2) for a closed set $A \subset K$.

2. Prove the properties (ii)-(v) in Proposition 9.2.2 and (iv) in Proposition 9.2.3.

3. Let E be a vector lattice and x, y, $z \in E$. Prove

   (a) $x \vee y = \frac{1}{2}(x + y + |x - y|)$, and $x \wedge y = \frac{1}{2}(x + y - |x - y|)$.

(b) $|x| \vee |y| = \frac{1}{2}(|x+y| + |x-y|)$ and deduce that

$$|x| \wedge |y| = \frac{1}{2}\,||x+y| - |x-y||\,.$$

(c) Deduce that $x \perp y$ is equivalent to $|x-y| = |x+y|$.

(d) The triangle inequality: $||x| - |y|| \le |x+y| \le |x| + |y|$.

(e) Deduce that $x \perp y$ is equivalent to $|x| \vee |y| = |x| + |y|$ and in this case $||x| - |y|| = |x+y| = |x| + |y|$.

(f) Birkhoff's inequalities: $|x \vee z - y \vee z| \le |x-y|$ and $|x \wedge z - y \wedge z| \le |x-y|$.

4. Check that the examples in Example 9.3.2 are Banach lattices.

5. Prove Proposition 9.3.3 (Hint: use Birkhoff's inequalities from Exercise 9.5.3) .

6. Prove that a subspace I of a Banach lattice is an ideal if and only if

$$(\,x \in I,\ \ |y| \le |x|\,) \Longrightarrow y \in I.$$

7. Let us consider the Banach space $\mathrm{C}^1[0,1]$ of continuous differentiable functions on $[0,1]$ with the norm

$$\|f\| = \max_{s\in[0,1]} |f(s)| + \max_{s\in[0,1]} |f'(s)|$$

and the natural order $f \ge 0$ if $f(s) \ge 0$ for all $s \in [0,1]$. Prove that the above norm is not a lattice norm.

8. Consider $\mathrm{C}^1[0,1]$ equipped with the norm

$$\|f\| = \max_{s\in[0,1]} |f'(s)| + |f(0)|$$

and the order $f \ge 0$ whenever $f(0) \ge 0$ and $f' \ge 0$. Show that $E := (\mathrm{C}^1[0,1], \ge, \|\cdot\|)$ is a Banach lattice.

9. Let E be a Banach lattice. Use the Hahn-Banach theorem to prove that

(a) $0 \le x$ is equivalent to $\langle x, x^*\rangle \ge 0$ for all $x^* \in E^*_+$;

(b) for each $0 \lneqq x \in E$ there exists $x^* \in E^*_+$ such that $\|x^*\| = 1$ and $\langle x, x^*\rangle = \|x\|$.

10. Show the characterization of ideals in the spaces $\mathrm{L}^p(\Omega,\mu)$ and $\mathrm{C}_0(\Omega)$ as stated in Example 9.3.7.

11. Prove the characterization of quasi-order units in Example 9.3.8 (2).

Bibliography

[1] C.D. Aliprantis and O. Burkinshaw, *Positive Operators*, Academic Press, 1985.

[2] L. Fuchs, *Partially ordered algebraic systems*, Addison-Wesley, 1963.

[3] P. Meyer-Nieberg, *Banach Lattices*, Springer-Verlag, 1991.

[4] H.H. Schaefer, *Banach Lattices and Positive Operators*, Grundlehren Math. Wiss., vol. 215, Springer-Verlag, 1974.

Lecture 10

Positive Operators

10.1 Basic properties

This section is concerned with positive operators on Banach lattices. Let E, F be two complex Banach lattices. A linear operator T from E into F is called *positive* (notation: $T \geq 0$) if $TE_+ \subset F_+$, which is equivalent to

$$|Tx| \leq T|x| \quad \text{for all}\, x \in E, \tag{10.1}$$

see Exercise 10.5.1. All positive operators are bounded as the following theorem shows.

Theorem 10.1.1. *Every positive linear operator $T : E \to F$ is continuous.*

Proof. Assume that T is not bounded. Then there is $(x_n) \subset E$ such that $\|x_n\| = 1$ and $\|Tx_n\| \geq n^\gamma$ for each $n \in \mathbb{N}$ and some $\gamma > 2$. Since $|Tx_n| \leq T|x_n|$, one can assume that $x_n \geq 0$ for all $n \in \mathbb{N}$. From $\sum_{n=1}^\infty \frac{\|x_n\|}{n^{\gamma-1}} < \infty$, it follows that $\sum_{n=1}^\infty \frac{x_n}{n^{\gamma-1}}$ is norm convergent in E. Set $x = \sum_{n=1}^\infty \frac{x_n}{n^{\gamma-1}}$. Then,

$$0 \leq \frac{x_n}{n^{\gamma-1}} \leq x, \quad \forall n \in \mathbb{N}.$$

So,

$$n \leq \left\| T\left(\frac{x_n}{n^{\gamma-1}}\right)\right\| \leq \|Tx\| < \infty, \quad \forall n \in \mathbb{N},$$

which is a contradiction. Thus, $T \in \mathcal{L}(E, F)$. □

As a consequence we obtain the following corollary.

Corollary 10.1.2. *Let $\|\cdot\|_1$, $\|\cdot\|_2$ be two norms on a Banach space E. If $E_1 = (E, \|\cdot\|_1)$ and $E_2 = (E, \|\cdot\|_2)$ are both Banach lattices, then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

Proof. This follows from the positivity of the identity operators $I : E_1 \to E_2$ and $I : E_2 \to E_1$. □

We denote by $\mathcal{L}(E,F)_+$ the set of all positive linear operators from a Banach lattice E into a Banach lattice F. For positive operators one can prove the following properties.

Proposition 10.1.3. *Let $T \in \mathcal{L}(E,F)_+$. Then the following properties hold.*

(i) $(Tx)^+ \leq Tx^+$ *and* $(Tx)^- \leq Tx^-$ *for all* $x \in E_{\mathbb{R}}$.

(ii) $\|T\| = \sup\{\|Tx\| : x \in E_+, \|x\| \leq 1\}$.

(iii) If $S \in \mathcal{L}(E,F)$ *such that* $0 \leq S \leq T$ *(this means that* $0 \leq Sx \leq Tx$ *for all* $x \in E_+$*), then* $\|S\| \leq \|T\|$.

Proof. (i): Since for $x \in E$, $Tx = Tx^+ - Tx^- \leq Tx^+$ and $(Tx)^+ = Tx \vee 0$ we obtain $(Tx)^+ \leq Tx^+$. The second property follows from $Tx^+ - (Tx)^+ = Tx^- - (Tx)^-$.
(ii): Holds by (10.1).
(iii): Since $0 \leq S \leq T$ we have $|Sx| \leq S|x| \leq T|x|$ for all $x \in E$. The assertion now follows by (ii). □

Another nice property of positive operators is, that they have positive resolvent. The converse in not always true, see also Proposition 10.3.1.

Proposition 10.1.4. *Let $T \in \mathcal{L}(E)$ be a positive operator with spectral radius* $\mathrm{r}(T)$.

(i) The resolvent $R(\mu,T)$ *is positive whenever* $\mu > \mathrm{r}(T)$.

(ii) If $|\mu| > \mathrm{r}(T)$*, then*

$$\big|R(\mu,T)x\big| \leq R\big(|\mu|,T\big)|x|, \quad x \in E.$$

Proof. We use the Neumann series representation

$$R(\mu,T) = \sum_{n=0}^{\infty} \frac{T^n}{\mu^{n+1}}$$

for the resolvent, which is valid for $|\mu| > \mathrm{r}(T)$ by Proposition B.1.10.

(i) If $T \geq 0$, then $T^n \geq 0$ for all n, hence for $\mu > \mathrm{r}(T)$, we have

$$R(\mu,T) = \lim_{N\to\infty} \sum_{n=0}^{N} \frac{T^n}{\mu^{n+1}} \geq 0$$

since the finite sums are positive and convergence holds in every entry.

(ii) We have for $|\mu| > r(T)$ and $x \in E$ that

$$\begin{aligned}\big|R(\mu,T)x\big| &= \left|\lim_{N\to\infty} \sum_{n=0}^{N} \frac{T^n x}{\mu^{n+1}}\right| \leq \lim_{N\to\infty} \sum_{n=0}^{N} \left|\frac{T^n x}{\mu^{n+1}}\right| \\ &\leq \lim_{N\to\infty} \sum_{n=0}^{N} \frac{T^n}{|\mu|^{n+1}}|x| = R\big(|\mu|,T\big)|x|.\end{aligned}$$

□

The following is an easy version of Perron's Theorem 5.1.6 for the infinite dimensional case.

Theorem 10.1.5. *If $A \in \mathcal{L}(E)$ is positive, then* $\mathrm{r}(T) \in \sigma(T)$.

Proof. Assertion (ii) of Proposition 10.1.4 implies

$$\big\|R(\mu,T)\big\| \le \big\|R(|\mu|,T)\big\| \qquad \text{for } |\mu| > \mathrm{r}(T).$$

Let now $\lambda \in \sigma(T)$ such that $|\lambda| = \mathrm{r}(T)$. Then, it follows from Corollary B.1.12 that $\|R(\mu,T)\| \to \infty$ whenever μ approaches λ. Putting $\mu = s\lambda$ with $s > 1$ the above estimate yields

$$\big\|R(s\,\mathrm{r}(T),T)\big\| \ge \big\|R(s\lambda,T)\big\| \to \infty \quad \text{as } s \searrow 1,$$

hence, by Corollary B.1.12, $\mathrm{r}(T)$ must be in the spectrum of T. □

Note that in general we cannot say more because points of the spectrum may not be isolated or eigenvalues. It will be possible to generalize the statements concerning irreducibility under additional assumptions. We will do it in a later lecture when we generalize the concept of spectral projections.

10.1.1 Lattice homomorphism and signum operators

Let us first consider the Banach lattice of continuous functions on a compact set K. Positive operators on $C(K)$ with $T\mathbb{1} = \mathbb{1}$ are nothing but contraction operators.

Lemma 10.1.6. *Suppose that K is compact and $T : C(K) \to C(K)$ is a linear operator satisfying $T\mathbb{1} = \mathbb{1}$. Then $0 \le T$ if and only if $\|T\| \le 1$.*

Proof. If $0 \le T$, then

$$|Tf| \le T|f| \le T(\|f\|_\infty \mathbb{1}) = \|f\|_\infty \mathbb{1}.$$

Hence $\|T\| \le 1$.
To prove the converse, we first observe that

$$-\mathbb{1} \le f \le \mathbb{1} \Leftrightarrow \|f - ir\mathbb{1}\|_\infty \le \rho_r := \sqrt{1+r^2} \text{ for all } r \in \mathbb{R}. \tag{10.2}$$

Let $0 \le f \in C(K)$. Then there is $n \in \mathbb{N}$ such that with $0 \le f \le n\mathbb{1}$. Set $g = \frac{2}{n} f$. Then, $0 \le g \le 2\mathbb{1}$ and so, $-\mathbb{1} \le g - \mathbb{1} \le \mathbb{1}$. By (10.2) we have $\|g - \mathbb{1} - ir\mathbb{1}\|_\infty \le \rho_r$ for all $r \in \mathbb{R}$. Since $T\mathbb{1} = \mathbb{1}$ and $\|T\| \le 1$, $\|Tg - \mathbb{1} - ir\mathbb{1}\|_\infty \le \rho_r$ for all $r \in \mathbb{R}$. So by (10.2) we obtain $-\mathbb{1} \le Tg - \mathbb{1} \le \mathbb{1}$. This implies $0 \le Tg \le 2\mathbb{1}$ and hence $Tf \ge 0$. □

The following result, due to Kakutani, shows that for every $e \in E_+$ the generated ideal satisfies $E_e \cong C(K)$ for some compact K. Here, E_e is equipped with the norm $\|x\|_e := \inf\{\lambda > 0 : x \in \lambda[-e,e]\}$, $x \in E_e$. We recall that $T \in \mathcal{L}(E,F)$ is called *lattice homomorphism* if $|Tx| = T|x|$ for every $x \in E$, where E, F are Banach lattices.

Theorem 10.1.7. *Let $e \in E_+$ and take E_e the ideal generated by $\{e\}$. Let $B := \{x^* \in (E_e)^*_+ : \langle e, x^* \rangle = 1\}$ and K the set of all extreme points of B. Then K is $\sigma(E^*, E)$-compact and the mapping*

$$U_e : E_e \ni x \mapsto f_x \in C(K), \quad f_x(x^*) = \langle x, x^* \rangle, x^* \in K,$$

is an isometric lattice isomorphism.

If $|h|$ is a quasi interior point of E_+, then $E_{|h|}$ is a dense subspace of E isomorphic to a space $C(K)$. Let $U_{|h|}$ be the lattice isomorphism from Kakutani's theorem and let $\widetilde{h} := U_{|h|}h$. Then $|\widetilde{h}| = U_{|h|}|h| = \mathbb{1}$. Consider the operator

$$\widetilde{S}_0 : C(K) \to C(K);\ f \mapsto (\operatorname{sign} \widetilde{h}) f := \frac{\widetilde{h}}{|\widetilde{h}|} f = \widetilde{h} f,$$

and put $S_h := U_{|h|}^{-1} \widetilde{S}_0 U_{|h|}$. Then S_h is a linear mapping from $E_{|h|}$ into itself satisfying

(i) $S_h \overline{h} = |h|$,

(ii) $|S_h x| \leq |x|$ for every $x \in E_{|h|}$.

Since (ii) implies the continuity of S_h for the norm induced by E and $\overline{E_{|h|}} = E$, S_h can be uniquely extended to E. This extension will be also denoted by S_h and is called *signum operator* with respect to h.

We now give the following auxiliary result which we will need later on.

Lemma 10.1.8. *Let $T, R \in \mathcal{L}(E)$ and assume that $|h|$ is a quasi interior point of E_+. Suppose we have $Rh = h$, $T|h| = |h|$, and $|Rx| \leq T|x|$ for all $x \in E$. Then $T = S_h^{-1} R S_h$.*

Proof. It follows from $|Rx| \leq T|x|$, $x \in E$ that T is a positive operator. Since $T|h| = |h|$, $E_{|h|}$ is T-and R-invariant. Consider the operators $\widetilde{T} := U_{|h|} T U_{|h|}^{-1}$, $\widetilde{R} := U_{|h|} R U_{|h|}^{-1}$, and put $\widetilde{h} := U_{|h|} h$. We then have

$$\widetilde{Rh} = \widetilde{h},\ \widetilde{T}\mathbb{1} = \mathbb{1},\ |\widetilde{R} f| \leq \widetilde{T}|f| \text{ for all } f \in C(K). \tag{10.3}$$

Define $T_1 := \widetilde{S}_0^{-1} \widetilde{R} \widetilde{S}_0$, where $\widetilde{S}_0$ is the multiplication operator by $\widetilde{h}$ on $C(K)$ defined above. By (10.3) we have

$$\begin{aligned} T_1 \mathbb{1} &= \mathbb{1} \quad \text{and} \\ |T_1 f| &= |\widetilde{S}_0^{-1} \widetilde{R} \widetilde{S}_0 f| = |\widetilde{R} \widetilde{S}_0 f| \leq \widetilde{T} |\widetilde{S}_0 f| = \widetilde{T}|f| \end{aligned} \tag{10.4}$$

for all $f \in C(K)$. Hence $\|T_1\| \leq \|\widetilde{T}\| = \|\widetilde{T}\mathbb{1}\|_\infty = 1$. So by Lemma 10.1.6, T_1 is a positive operator and (10.4) implies that $0 \leq T_1 \leq \widetilde{T}$. Therefore, $\|\widetilde{T} - T_1\| = \|(\widetilde{T} - T_1)\mathbb{1}\|_\infty = 0$. Thus, $T_1 = \widetilde{T}$ and hence, $T = S_h^{-1} R S_h$. □

10.2 Exponential functions

We review in this section basic facts about exponential functions of bounded linear operators. Many results are quite analogous to the matrix case, and it is actually possible to prove them by using the same functional calculus argument. However, since building up the functional calculus would need different arguments, we leave that line aside. For the followings, let X be a Banach space and $A \in \mathcal{L}(X)$.

Definition 10.2.1. For $A \in \mathcal{L}(X)$ we define

$$\mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}$$

for each $t \geq 0$.

That this definition is well-formulated is the content of Exercise 10.5.2. Note that this also implies a bound

$$\|\mathrm{e}^{tA}\| \leq \mathrm{e}^{t\|A\|} \tag{10.5}$$

for $t \geq 0$.

Proposition 10.2.2. *For $A \in \mathcal{L}(X)$, the following properties hold for its exponential function:*

(i) The functional equation

$$e^{(t+s)A} = e^{tA} e^{sA}, \qquad e^{0A} = I \tag{10.6}$$

holds.

(ii) The function $\mathbb{R}_+ \ni t \mapsto e^{tA}$ is continuous.

(iii) The function $\mathbb{R}_+ \ni t \mapsto e^{tA} =: T(t)$ is differentiable and satisfies the differential equation

$$\begin{aligned} \frac{d}{dt} T(t) &= AT(t) \\ T(0) &= I. \end{aligned}$$

Proof. The proof is essentially to show that the standard manipulations with power series, as we are used to them in the scalar case, are justified. To show the functional equation, note that using the formula for the Cauchy product of infinite series we obtain

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \cdot \sum_{k=0}^{\infty} \frac{s^k A^k}{k!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{t^{n-k} A^{n-k}}{(n-k)!} \cdot \frac{s^k A^k}{k!} &= \sum_{n=0}^{\infty} \frac{A^n}{n!} \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} t^{n-k} s^k \\ &= \sum_{n=0}^{\infty} \frac{(t+s)^n A^n}{n!}, \end{aligned}$$

where we used the binomial theorem in the last step.

To show continuity, note that

$$\mathrm{e}^{(t+h)A} - \mathrm{e}^{tA} = \mathrm{e}^{tA}(\mathrm{e}^{hA} - I),$$

which shows that it is enough to prove that

$$\lim_{h\to 0} \mathrm{e}^{hA} = I.$$

This follows from

$$\|\mathrm{e}^{hA} - I\| = \left\| \sum_{k=1}^{\infty} \frac{h^k A^k}{k!} \right\| \le \sum_{k=1}^{\infty} \frac{|h|^k \cdot \|A\|^k}{k!} = \mathrm{e}^{|h|\cdot\|A\|} - I.$$

Finally, differentiability follows by a similar argument as

$$\begin{aligned} \frac{1}{h}\|\mathrm{e}^{hA} - I - hA\| = \left\| \sum_{k=2}^{\infty} \frac{h^{k-1} A^k}{k!} \right\| &\le \sum_{k=2}^{\infty} \frac{|h|^{k-1} \cdot \|A\|^k}{k!} \\ &= \frac{1}{|h|} \left\| \mathrm{e}^{|h|\cdot\|A\|} - 1 - |h| \cdot \|A\| \right\|, \end{aligned}$$

and the statement follows from the scalar case. □

The functional equation in (10.6) plays a crucial role and hence we give it a name.

Definition 10.2.3. A map $\mathbb{R}_+ \ni t \mapsto T(t) \in \mathcal{L}(X)$ is called a one-parameter operator semigroup (or*operator semigroup* or just *semigroup* for short), if

$$T(0) = I,$$

and

$$T(t+s) = T(t)T(s)$$

holds for all $t, s \ge 0$.

The most important property of continuous semigroups is that they are exponential functions.

Proposition 10.2.4. *Let $(T(t))_{t\ge 0}$ be a semigroup which is continuous. Then there is an operator $A \in \mathcal{L}(X)$ such that $T(t) = e^{tA}$.*

Proof. Since the function $t \mapsto T(t)$ is continuous and $T(0) = I$, the operators

$$V(t_0) := \int_0^{t_0} T(s)ds$$

are invertible for sufficiently small $t_0 > 0$ (see Exercise 10.5 5). For properties of vector valued Riemann integrals of continuous functions see the Appendix B, especially Proposition B.1.8. It follows that

$$T(t) = V(t_0)^{-1}V(t_0)T(t) = V(t_0)^{-1}\int_0^{t_0} T(t+s)ds = V(t_0)^{-1}\int_t^{t+t_0} T(s)ds$$
$$= V(t_0)^{-1}(V(t+t_0) - V(t))$$

holds for all $t \geq 0$. Since V is the integral function of a continuous function, it is differentiable and so is $T(\cdot)$. Let us denote $\frac{d}{dt}T(0) =: A$. Then $A \in \mathcal{L}(X)$ and it follows from the functional equation that

$$\frac{d}{dt}T(t) = \lim_{h\to 0}\frac{T(t+h)-T(t)}{h} = \lim_{h\to 0}\frac{T(h)-I}{h}T(t) = AT(t)$$

for all $t \geq 0$. Hence T satisfies a linear differential equation of the form

$$\frac{d}{dt}T(t) = AT(t)$$

with $T(0) = I$. But $S(t) = \mathrm{e}^{tA}$ also satisfies the same differential equation, and by linearity we obtain the equality $T(t) = S(t)$, see Exercise 10.5.3. □

We end this subsection by defining the *growth bound*

$$\omega_0(A) = \inf\left\{\omega \in \mathbb{R} : \exists M_\omega \text{ such that } \|e^{tA}\| \leq M_\omega e^{\omega t},\ \forall t \geq 0\right\}$$

for a given $A \in \mathcal{L}(X)$. Some useful properties of $\omega_0(A)$ are listed in Exercise 10.5 7.

10.3 Positive exponential functions

In the following, let E be a Banach lattice and $A \in \mathcal{L}(E)$. We would like to investigate the positivity and asymptotic properties of the exponential function of A. We start with a characterization through the resolvent of A, see Definition B.1.9 in Appendix.

Proposition 10.3.1. *The semigroup $T(t) = e^{tA}$ is positive if and only if*

$$R(\lambda, A) = (\lambda - A)^{-1} \geq 0$$

for all $\lambda > \omega_0(A)$.

Proof. Taking $\lambda > \omega_0(A)$, notice that since the function $t \mapsto \mathrm{e}^{t(A-\lambda)}$ is continuously differentiable, the fundamental theorem of calculus holds and

$$\mathrm{e}^{t(A-\lambda)} - I = (A-\lambda)\int_0^t \mathrm{e}^{-\lambda s}T(s)ds = \int_0^t \mathrm{e}^{-\lambda s}T(s)ds(A-\lambda) \tag{10.7}$$

holds. Hence, for $\lambda > \omega_0(A)$, the equality

$$-I = (A - \lambda)\int_0^\infty \mathrm{e}^{-\lambda s} T(s)ds = \int_0^\infty \mathrm{e}^{-\lambda s} T(s)ds(A - \lambda)$$

holds where we used the definition of $\omega_0(A)$ and the fact that $\lambda > \omega_0(A)$. This means that the resolvent of A is represented by

$$R(\lambda, A) = \int_0^\infty \mathrm{e}^{-\lambda s} T(s)ds \tag{10.8}$$

for $\lambda > \omega_0(A)$ proving the direction that when $T(\cdot)$ is a positive semigroup, then $R(\lambda, A)$ is positive for $\lambda > \omega_0(A)$.

For the other direction notice that by Exercise 10.5.4 the Euler formula

$$\lim_{n\to\infty}\left(I - \frac{t}{n}A\right)^{-n} = \mathrm{e}^{tA}$$

holds for $t \geq 0$. Since $(I - \frac{t}{n}A)^{-n} = \left(\frac{n}{t}R(\frac{n}{t}, A)\right)^n \geq 0$ for n sufficiently large by assumption, the positivity of the operators $T(t)$ follows. □

For $A \in \mathcal{L}(E)$ we define its *spectral bound* as

$$s(A) := \sup\{\mathrm{Re}\,\lambda \,:\, \lambda \in \sigma(A)\}. \tag{10.9}$$

It follows from the spectral mapping theorem for bounded linear operators, see Theorem B.1.13, and Exercise 10.5 7 that

$$\mathrm{s}(A) = \omega_0(A)$$

for $A \in \mathcal{L}(X)$. So, as a Corollary we obtain the following result which will be improved in a more general situation (without using the spectral mapping theorem) in a forthcoming lecture.

Proposition 10.3.2. *For a positive exponential function $T(t) = e^{tA}$ we have*

$$R(\lambda, A) = \int_0^\infty e^{-\lambda s} T(s)ds$$

for all $\lambda > s(A)$. Hence, for all $\lambda > s(A)$ we have $0 \leq R(\lambda, A)$.

Corollary 10.3.3. *For a positive exponential function $T(t) = e^{tA}$ we have*

$$s(A) \in \sigma(A).$$

Proof. The positivity of the operators $T(t)$ means that

$$|T(t)f| \leq T(t)|f|$$

holds for all $f \in E$ and $t \geq 0$. We obtain therefore that

$$|R(\lambda, A)f| \leq \int_0^\infty e^{-\operatorname{Re}\lambda s} T(s)|f| ds$$

for all $\operatorname{Re}\lambda > s(A)$ and $f \in E$. Hence,

$$\|R(\lambda, A)\| \leq \|R(\operatorname{Re}\lambda, A)\|.$$

Note that there is $\lambda_n \in \rho(A)$ such that $\operatorname{Re}\lambda_n \to s(A)$, $\operatorname{Re}\lambda > s(A)$ and $\|R(\lambda_n, A)\| \to \infty$. This implies $\|R(\operatorname{Re}\lambda_n, A)\| \to \infty$ and hence $s(A) \in \sigma(A)$ (see Corollary B.1.12). □

Compare the following with Corollary 7.1.3 from Lecture 7.

Corollary 10.3.4. *Let K be compact and $E = C(K)$. If $A \in \mathcal{L}(E)$ is such that it generates a positive semigroup $(e^{tA})_{t\geq 0}$, then the following assertions are equivalent.*

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

(b) The operator $-A^{-1}$ exists and is positive.

(c) There exists $0 \leq f \in E$ such that $Af = -\mathbb{1}$.

Proof. The equivalence (a) $\iff$ (b) follows from Proposition 10.3.2. Since $-A^{-1} = R(0, A)$, (b)$\Rightarrow$(c) follows taking $f := -A^{-1}\mathbb{1}$.

We close the proof loop by showing (c)$\Rightarrow$(a). Assume that $Af = -\mathbb{1}$ for some $0 \leq f \in E$. Then for $\lambda > \max\{\mathrm{s}(A), 0\}$ we have

$$\begin{aligned} 0 &\leq R(\lambda, A)\mathbb{1} = -AR(\lambda, A)f \\ &= f - \lambda R(\lambda, A)f \leq f. \end{aligned}$$

Hence,

$$\sup_{\lambda > \max\{\mathrm{s}(A), 0\}} \|R(\lambda, A)\| \leq \|f\|_\infty.$$

Since by Corollary 10.3.3, $\mathrm{s}(A) \in \sigma(A)$, it follows from Corollary B.1.12 that $\mathrm{s}(A) < 0$. □

10.4 Notes and Remarks

For more details on Kakutani's Theorem see e.g. [1]. Concerning Lemma 10.1.8 and signum operators we refer to [2, B-III].

10.5 Exercises

1. Show that an operator is positive, i.e., $TE_+ \subset F_+$, if and only if $|Tx| \leq T|x|$ holds for all $x \in X$.

2. (a) Show that in a Banach space every absolutely convergent series is convergent.
   (b) Use the previous result to show that e^{tA} is a well-defined bounded linear operator on X.

3. Let $A \in \mathcal{L}(X)$. Show that the Cauchy problem
$$\frac{d}{dt}F(t) = AF(t)$$
$$F(0) = B \in \mathcal{L}(X)$$
has a unique solution $F : \mathbb{R}_+ \to \mathcal{L}(X)$.

4. For $A \in \mathcal{L}(X)$ prove Euler's formulas:
$$\lim_{n\to\infty} \left(I + \frac{t}{n}A\right)^n = \mathrm{e}^{tA}$$
and
$$\lim_{n\to\infty} \left(I - \frac{t}{n}A\right)^{-n} = \mathrm{e}^{tA}$$
for $t \geq 0$.

5. Let $A \in \mathcal{L}(X)$. Show that $V(t_0) = \int_0^{t_0} e^{sA}ds$ is invertible for sufficiently small $t_0 > 0$.

6. For $A \in \mathcal{L}(X)$ prove that the spectrum $\sigma(A)$ of A is compact and nonempty. Deduce that the spectral radius $r(A) := \sup\{|\lambda| : \lambda \in \sigma(A)\}$ is finite and satisfies $r(A) \leq \|A\|$.

7. Let $A \in \mathcal{L}(X)$. Prove the following assertions:
   (a) $\omega_0(A) = \lim_{t\to+\infty} t^{-1} \log \|e^{tA}\| = \inf_{t>0} t^{-1} \log \|e^{tA}\|$.
   (b) $r(e^{tA}) = e^{t\omega_0(A)}$ for all $t \geq 0$.

8. Consider the Banach lattice $\mathrm{C}^1[0,1]$ equipped with the norm
$$\|f\| = \max_{s\in[0,1]} |f'(s)| + |f(0)|$$
and the order $f \geq 0$ whenever $f(0) \geq 0$ and $f' \geq 0$. Let us define the operator
$$(Tf)(t) := \int_0^t g(s)f(s)ds$$
with a given $g \in C[0,1]$. Calculate $\|T\|$. For which g is T positive?

Bibliography

[1] P. Meyer-Nieberg, *Banach Lattices*, Springer-Verlag, 1991.

[2] R. Nagel (ed.), *One-parameter Semigroups of Positive Operators*, Lect. Notes in Math., vol. 1184, Springer-Verlag, 1986.

Lecture 11

Operator Semigroups

11.1 Motivation

We first recall the following important facts from the previous lecture, see Propositions 10.2.2 and 10.2.4. Let $(T(t))_{t\geq 0}$ be a semigroup which is continuous. Then there is an operator $A \in \mathcal{L}(X)$ such that $T(t) = \mathrm{e}^{tA}$, and $u(t) = T(t)f$ solves the differential equation

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}u(t) = Au(t), & t \geq 0, \\ u(0) = f. \end{cases} \tag{11.1}$$

for all $f \in X$. It turns out, however, that there are important semigroups which do not satisfy this continuity property but a weaker one. We will also see that there are important differential equations of the form (11.1) where the operator A is no longer bounded. As a motivation, we work out an important example, the shift (semi)group.

Example 11.1.1. Take

$$X = \mathrm{C}_{\mathrm{ub}}(\mathbb{R}) := \big\{ f : \mathbb{R} \to \mathbb{R} \,:\, f \text{ is uniformly continuous and bounded} \big\},$$

which is a Banach space with the supremum norm

$$\|f\|_\infty := \sup_{s\in\mathbb{R}} |f(s)|.$$

The additive (semi)group structure of $\mathbb{R}$ naturally induces a semigroup on this Banach space by setting

$$(S(t)f)(s) = f(t+s), \quad \text{for } f \in X,\ s \in \mathbb{R},\ t \geq 0.$$

One readily sees that $S(t)$ is a bounded linear operator on X, in fact a linear isometry. The semigroup property follows immediately from the definition. From the uniform continuity of $f \in X$ we conclude that the mapping

$$t \mapsto S(t)f$$

is continuous on $X = \mathrm{C}_{\mathrm{ub}}(\mathbb{R})$, i.e., that $(S(t))_{t\geq 0}$ is a "strongly continuous semigroup" called the *left shift semigroup*. Note that $S(\cdot)$ is not continuous in the operator norm, because if

$$\|S(t) - I\| = \sup_{\|f\|\leq 1} \|S(t)f - f\| = \sup_{\|f\|\leq 1} \sup_{s\in\mathbb{R}} |f(s+t) - f(s)|$$

would converge to zero as $t \to 0$, then every function in the unit ball of $\mathrm{C}_{\mathrm{ub}}(\mathbb{R})$ would be uniformly equicontinuous, which is impossible (and, as a side remark, which would imply by the Arzelà-Ascoli Theorem that the unit ball of $\mathrm{C}_{\mathrm{ub}}(\mathbb{R})$ would be compact).

Let us investigate whether this semigroup $(S(t))_{t\geq 0}$ solves some initial value problem such as (11.1). The heuristics of exponential functions helps again: Given e^{tA} for a matrix $A \in \mathbb{R}^{d\times d}$, we can "calculate" the exponent by differentiating this exponential function at 0:

$$A = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA}\Big|_{t=0}.$$

What happens in the case of the shift semigroup $(S(t))_{t\geq 0}$? The semigroup $(S(t))_{t\geq 0}$ is not even continuous for the operator norm. So let us look at differentiability of the *orbit map* $t \mapsto S(t)f$ for some given $f \in X$, which is called *strong differentiability*. The limit

$$\lim_{h\to 0} \frac{1}{h}(S(h)f - f) = \lim_{h\to 0} \frac{f(h+\cdot) - f(\cdot)}{h}$$

must exist in the sup-norm of X. We immediately find a suitable candidate for the limit: Since the limit must exist pointwise on $\mathbb{R}$, it cannot be anything else but f'. Hence, the function f must be at least differentiable so that the limit can exist. For f differentiable with f' being uniformly continuous we have

$$\sup_{s\in\mathbb{R}} \left| \frac{f(h+s) - f(s)}{h} - f'(s) \right| = \sup_{s\in\mathbb{R}} \left| \frac{1}{h} \int_s^{s+h} \big(f'(r) - f'(s)\big) dr \right| \leq \varepsilon,$$

for all h with $|h| \leq \delta$, where $\delta > 0$ is sufficiently small, chosen for the arbitrarily given $\varepsilon > 0$ from the uniform continuity of f'. This shows that if $f, f' \in X$, then we have

$$\lim_{h\to 0} \left\| \frac{f(h+\cdot) - f(\cdot)}{h} - f'(\cdot) \right\|_\infty = \lim_{h\to 0} \sup_{s\in\mathbb{R}} \left| \frac{f(h+s) - f(s)}{h} - f'(s) \right| = 0.$$

Note that for the derivative of $S(t)f$ at arbitrary $t \in \mathbb{R}$ we obtain by the same argument

$$\frac{d}{dt}(S(t)f) = S(t)f'.$$

This means that for $f, f' \in X$ the orbit function $u(t) = S(t)f$ solves the differential equation

$$\begin{cases} \dot{u}(t) = Au(t), & t \in \mathbb{R}, \\ u(0) = f \in D(A), \end{cases}$$

where $(Af)(s) = f'(s)$ with domain

$$D(A) := \big\{ f : f, f' \in \mathrm{C}_{\mathrm{ub}}(\mathbb{R}) \big\}.$$

Clearly, A is not everywhere defined and cannot be extended to a bounded operator.

Definition 11.1.2. Let $T : [0, \infty) \to \mathcal{L}(X)$ be a mapping.

1. We say that T has the *semigroup property* if for all $t, s \in [0, \infty)$ the identities

$$T(t+s) = T(t)T(s) \quad \text{and}$$
$$T(0) = I \qquad \text{(i.e. the identity operator on } X)$$

   hold.

2. Suppose $Y \subseteq X$ is a linear subspace and for all $f \in Y$ the mapping

$$t \mapsto T(t)f \in X$$

   is continuous. Then T is called *strongly continuous on* Y. If $Y = X$ we just say *strongly continuous.*

3. If a strongly continuous mapping T possess the semigroup property, then $(T(t))_{t\geq 0}$ is called a *strongly continuous one-parameter semigroup* of bounded linear operators on the Banach space X. Often we shall abbreviate this terminology to C_0-*semigroup*[1].

11.2 Basic properties

Let us observe some elementary consequences of the semigroup property and the strong continuity, respectively. The first result we mention here reflects again the exponential function: Semigroups can grow at most exponentially.

Proposition 11.2.1. *(i) Let $T : [0, \infty) \to \mathcal{L}(X)$ be a strongly continuous function. Then for all $t \geq 0$ we have*

$$\sup_{s\in[0,t]} \|T(s)\| < \infty,$$

that is to say, T is locally bounded.

(ii) Let $(T(t))_{t\geq 0}$ be a C_0-semigroup. Then there are $M \geq 1$ and $\omega \in \mathbb{R}$ such that

$$\|T(t)\| \leq M e^{\omega t} \quad \textit{holds for all } t \geq 0.$$

We call the semigroup $(T(t))_{t\geq 0}$ of *type* (M, ω) if it satisfies the exponential estimate above with the particular constants M and ω. Note already here that the type of a semigroup may change if we pass to an equivalent norm on X.

[1] C_0 or $(C, 0)$ abbreviates Cesàro summable of order zero, which means the continuity property $\lim_{t\to 0} T(t)f = f$ for all $f \in X$.

Proof. (i): For a fixed $f \in X$, the mapping $T(\cdot)f$ is continuous on $[0, \infty)$, hence bounded on compact intervals $[0, t]$, i.e.,

$$\sup_{s \in [0,t]} \|T(s)f\| < \infty.$$

The uniform boundedness principle, see Appendix, Theorem B.1.1, implies the assertion.
(ii): By part (i) we have

$$M := \sup_{s \in [0,1]} \|T(s)\| < \infty.$$

Taking an arbitrary $t \geq 0$ we write $t = n + r$ with $n \in \mathbb{N}$ and $r \in [0, 1)$. From this representation and the semigroup property it follows that

$$\begin{aligned} \|T(t)\| &= \|T(r)T(1)^n\| \leq M\|T(1)^n\| \leq M\|T(1)\|^n \\ &\leq M(\|T(1)\| + 1)^n \leq M(\|T(1)\| + 1)^t = M\mathrm{e}^{\omega t} \end{aligned}$$

with $\omega = \log(\|T(1)\| + 1)$. □

Hence, orbits of C_0-semigroups are *exponentially bounded*. The *greatest lower bound* of these exponential bounds plays a special role in the theory.

Definition 11.2.2. For a C_0-semigroup $(T(t))_{t\geq 0}$ its *growth bound* is defined by

$$\omega_0(T) := \inf\bigl\{\omega \in \mathbb{R} : \text{ there is } M = M_\omega \geq 1 \text{ with } \|T(t)\| \leq M\mathrm{e}^{\omega t} \text{ for all } t \geq 0\bigr\}.$$

Remark 11.2.3.

1. A C_0-semigroup $(T(t))_{t\geq 0}$ is of type (M, ω) for all $\omega > \omega_0(T)$ and for some $M = M_\omega$. In general, however, it is *not* of type $(M, \omega_0(T))$ for any M. A simple example is the following. Let $X = \mathbb{C}^2$ and let the matrix semigroup given by

$$T(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

   Here $\omega_0 = 0$, but clearly T is not bounded, i.e., not of type $(M, 0)$ for any M.

2. For a matrix $A \in \mathbb{R}^{d\times d}$ we define $T(t) = \mathrm{e}^{tA}$. This semigroup $(T(t))_{t\geq 0}$ is of type $(1, \|A\|)$ as the trivial norm estimate

$$\|\mathrm{e}^{tA}\| \leq \mathrm{e}^{t\|A\|}$$

   shows. In contrast to this, in infinite dimensional situation it can happen that a semigroup is not of type $(1, \omega)$ for any ω, even though $\omega_0(T) = -\infty$. This is an extremely important fact, which causes major difficulties in many arguments. See Example 11.2.6 below for a simple demonstration.

The definition of a semigroup above comprises of the algebraic semigroup property, and the analytic property of strong continuity. We will see next that these two properties combine well, and we provide some means for verifying strong continuity.

Proposition 11.2.4. *(i) Let $T : [0,\infty) \to \mathcal{L}(X)$ be a locally bounded mapping with the semigroup property, and let $f \in X$. If the mapping $T(\cdot)f$ is* right continuous *at* 0*, i.e., $T(h)f \to f$ for $h \downarrow 0$, then it is continuous everywhere.*

(ii) A mapping T with the semigroup property is strongly continuous on X if and only if it is locally bounded and there is a dense subset $D \subseteq X$ on which T is strongly continuous.

Proof. (i): Fix $f \in X$ and $t > 0$, and set $M_t := \sup_{[0,t]} \|T(s)\|$. Then

$$T(t+h)f - T(t)f = T(t)(T(h)f - f), \quad \text{if } 0 < h,$$
$$T(t+h)f - T(t)f = T(t+h)(f - T(-h)f), \quad \text{if } -t < h < 0.$$

Summarizing, for $|h| \le t$ we obtain

$$\|T(t+h)f - T(t)f\| \le M_t \|f - T(|h|)f\|,$$

which converges to 0 for $|h| \to 0$ by the assumption.
(ii): In view of Proposition 11.2.1 one implication is straightforward. So we turn to the other one, and suppose that T is locally bounded and strongly continuous on a dense subspace D. Take an arbitrary $f \in X$ and $f_n \in D$ such that $f_n \to f$. Then, by local boundedness, for fixed $t_0 > 0$ we get that the functions $T(\cdot)f_n$ converge uniformly to the function $T(\cdot)f$ on $[0,t_0]$. Since uniform limits of continuous functions is continuous, the statement follows. □

Example 11.2.5. For $f \in \mathrm{L}^p(\mathbb{R})$ we define

$$(S(t)f)(s) := f(t+s) \quad \text{for } s \in \mathbb{R},\ t \ge 0.$$

Then $S(t)$ is a linear isometry on $\mathrm{L}^p(\mathbb{R})$. Moreover, S has the semigroup property. We call $(S(t))_{t\ge0}$ the *left shift semigroup* on $\mathrm{L}^p(\mathbb{R})$. Furthermore, for $p \in [1,\infty)$ the left shift semigroup $(S(t))_{t\ge0}$ is strongly continuous on $\mathrm{L}^p(\mathbb{R})$.

In fact, recall that the shift semigroup is strongly continuous on the space of bounded uniformly continuous functions and that the set of continuous functions with compact support is dense in $\mathrm{L}^p(\mathbb{R})$. Taking $f \in \mathrm{C}_c(\mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$ such that $\operatorname{supp} f \subset [\alpha, \beta]$, we see that

$$\|f(\cdot) - f(\cdot + t)\|_p^p = \int_{\mathbb{R}} |f(s) - f(s+t)|^p ds \le (\beta - \alpha) \sup_{s \in [\alpha,\beta]} |f(s) - f(s+t)|^p,$$

which goes to zero as $t \to 0$ by the uniform continuity of f, where $\|f\|_p := \int_{\mathbb{R}} |f(s)|^p ds$, $f \in \mathrm{L}^p(\mathbb{R})$.

Since $\|S(t)\| \le 1$, the statement follows by Proposition 11.2.4.

We have seen in Proposition 11.2.1 that a semigroup $(T(t))_{t\geq 0}$ is always exponentially bounded, meaning that an estimate of the type

$$\|T(t)\| \leq M\mathrm{e}^{\omega t}$$

holds. Let us give here an example to show that in the infinite dimensional case it is quite possible to have $M > 1$. The example will be a slight modification of the shift semigroup.

Example 11.2.6 (A bounded semigroup which is not a contraction)**.** Let us consider the Hilbert space $L^2((0,1),\mu)$, where μ denotes the measure defined by

$$\mu(A) := 2\lambda(A \cap (0, \tfrac{1}{2})) + \lambda(A \cap (\tfrac{1}{2}, 1))$$

for all Lebesgue measurable sets A where λ is the Lebesgue-measure. Furthermore, let $(S(t))_{t\geq 0}$ be the nilpotent left shift semigroup defined by

$$(S(t)f)(s) := \begin{cases} f(s+t), & \text{for } s+t \leq 1, \\ 0 & \text{for } s+t > 1. \end{cases}$$

Obviously, S satisfies the semigroup property and, since the norm $\|\cdot\|_\mu$ is equivalent to the norm $\|\cdot\|_\lambda$, the semigroup $(S(t))_{t\geq 0}$ is strongly continuous by similar arguments as in the previous example. Clearly, $\|S(t)\| \leq 2$. In addition we see that $S(t) = 0$ for all $t > 1$.

Finally, consider the function

$$f_t = \frac{1}{\sqrt{t}}\chi_{(\frac{1}{2},\frac{1}{2}+t)}$$

for $t \in (0, \frac{1}{2})$. It holds that $\|f_t\|_\mu = 1$ and

$$\|S(t)f_t\|_\mu = 2.$$

Hence, $\|S(t)\| = 2$ for $t \in (0,1)$.

11.3 The infinitesimal generator

One main message what we would like to transmit is that if we have a *semigroup*, then there is a *differential equation* so that the semigroup provides the solutions. Looking for the equation, we now consider the differentiability of orbit maps as in Example 11.1.1.

Lemma 11.3.1. *Take a semigroup $(T(t))_{t\geq 0}$ and an element $f \in X$. For the orbit map $u : t \mapsto T(t)f$, the following properties are equivalent:*

(i) u is differentiable on $[0,\infty)$,

(ii) u is right differentiable at 0.

If u is differentiable, then

$$\dot{u}(t) = T(t)\dot{u}(0).$$

Proof. We only have to show that (ii) implies (i). We proceed analogously to the proof of Proposition 11.2.4. First, we have

$$\begin{aligned}\lim_{h\downarrow 0} \tfrac{1}{h}\left(u(t+h) - u(t)\right) &= \lim_{h\downarrow 0} \tfrac{1}{h}\left(T(t+h)f - T(t)f\right) = T(t)\lim_{h\downarrow 0} \tfrac{1}{h}\left(T(h)f - f\right) \\ &= T(t)\lim_{h\downarrow 0} \tfrac{1}{h}\left(u(h) - u(0)\right) = T(t)\,\dot{u}(0),\end{aligned}$$

by the continuity of $T(t)$. Hence u is right differentiable on $[0,\infty)$.
On the other hand, for $-t \le h < 0$, we write

$$\begin{aligned}&\tfrac{1}{h}\left(T(t+h)f - T(t)f\right) - T(t)\dot{u}(0) \\ &\qquad = T(t+h)\left(\tfrac{1}{h}\left(f - T(-h)f\right) - \dot{u}(0)\right) + T(t+h)\dot{u}(0) - T(t)\dot{u}(0).\end{aligned}$$

As $h \uparrow 0$, the first term on the right-hand side converges to zero by the first part and by the boundedness of $\|T(t+h)\|$ for $h \in [-t,t]$. The other term converges to zero by the strong continuity of T. Hence u is also left differentiable, and its derivative is

$$\dot{u}(t) = T(t)\,\dot{u}(0)$$

for all $t \ge 0$. □

We thus see that the derivative $\dot{u}(0)$ of the orbit map $u(t) = T(t)f$ at $t = 0$ determines the derivative at each point $t \in [0,\infty)$. We give a name to the operator $f \mapsto \dot{u}(0)$.

Definition 11.3.2. The *infinitesimal generator*, or simply *generator* A of a C_0-semigroup $(T(t))_{t\ge 0}$ is defined as follows. Its *domain of definition* is given by

$$D(A) := \left\{f \in X : T(\cdot)f \text{ is differentiable in } [0,\infty)\right\},$$

and for $f \in D(A)$ we set

$$Af := \frac{\mathrm{d}}{\mathrm{d}t}T(t)f|_{t=0} = \lim_{t\downarrow 0} \tfrac{1}{t}\left(T(t)f - f\right).$$

As we hoped for, a semigroup yields solutions to some linear initial value problem in the Banach space X.

Proposition 11.3.3. *The generator A of a C_0-semigroup $(T(t))_{t\ge 0}$ has the following properties.*

(i) $A : D(A) \subseteq X \to X$ is a linear operator.

(ii) If $f \in D(A)$, then $T(t)f \in D(A)$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}T(t)f = T(t)Af = AT(t)f \quad \textit{for all } t \ge 0.$$

(iii) For a given $f \in D(A)$, the semigroup $(T(t))_{t\geq 0}$ provides the solutions to the initial value problem

$$\begin{cases} \dot{u}(t) = Au(t), \quad t \geq 0, \\ u(0) = f, \end{cases}$$

via $u(t) := T(t)f$.

Proof. (i): Linearity follows immediately from the definition because we take the limit of linear objects as $h \downarrow 0$.
(ii): Take $f \in D(A)$ and $t \geq 0$. We have to show that $T(\cdot)T(t)f$ is right differentiable at 0 with derivative $T(t)Af$. From the strong continuity of $T(t)$ we obtain

$$T(t)Af = T(t) \lim_{h\downarrow 0} \frac{T(h)f - f}{h} = \lim_{h\downarrow 0} \frac{T(h)T(t)f - T(t)f}{h}.$$

By the definition of A this further equals $AT(t)f$.
Part (iii) is just a reformulation of (ii). □

We now investigate infinitesimal generators further.

Proposition 11.3.4. *The generator A of a C_0-semigroup $(T(t))_{t\geq 0}$ has the following properties.*

(i) For all $t \geq 0$ and $f \in X$, one has

$$\int_0^t T(s)f\mathrm{d}s \in D(A),$$

where the integral has to be understood as the Riemann integral of the continuous function $s \mapsto T(s)f$, see Appendix B.

(ii) For all $t \geq 0$, one has

$$\begin{aligned} T(t)f - f &= A\int_0^t T(s)f\mathrm{d}s \quad \text{if } f \in X, \\ &= \int_0^t T(s)Af\mathrm{d}s \quad \text{if } f \in D(A). \end{aligned}$$

Proof. (i): For $g := \int_0^t T(s)f\mathrm{d}s$ we calculate the difference quotient

$$\begin{aligned} \frac{T(h)g - g}{h} &= \frac{1}{h}\Big(T(h)\int_0^t T(s)f\mathrm{d}s - \int_0^t T(s)f\mathrm{d}s\Big) = \frac{1}{h}\Big(\int_0^t T(h+s)f\mathrm{d}s - \int_0^t T(s)f\mathrm{d}s\Big) \\ &= \frac{1}{h}\Big(\int_h^{t+h} T(s)f\mathrm{d}s - \int_0^t T(s)f\mathrm{d}s\Big) = \frac{1}{h}\Big(\int_t^{t+h} T(s)f\mathrm{d}s - \int_0^h T(s)f\mathrm{d}s\Big). \end{aligned}$$

Since the integrands here are continuous, we can take limits as $h \downarrow 0$ and obtain

$$\lim_{h\downarrow 0} \frac{T(h)g - g}{h} = T(t)f - f.$$

This yields $g \in D(A)$ and $Ag = T(t)f - f$.
(ii): Taking $f \in D(A)$, by Proposition 11.3.3.(ii) we see that the identity $AT(t)f = T(t)Af$ holds, hence $v(t) := AT(t)f$ defines a continuous function. For $h > 0$ define the continuous functions $v_h(t) := \frac{1}{h}(T(t+h)f - T(t)f)$. Then we have the estimate

$$\|v_h(t) - v(t)\| \le \|T(t)\| \Big\| \frac{1}{h}(T(h)f - f) - Af \Big\|.$$

From this and the definition of A we conclude (by using the exponential boundedness of T) that v_h converges to v uniformly on every compact interval. This yields

$$\int_0^t v_h(s)\mathrm{d}s \to \int_0^t v(s)\mathrm{d}s \quad \text{as } h \downarrow 0.$$

Fortunately we have already calculated the limit of the left-hand side in part (i): It equals $T(t)f - f$. Hence,

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

which completes the proof. □

Before turning our attention to the most fundamental result of this section, let us introduce a new notation and define what a *closed operator* is. For a linear operator A defined on a linear subspace $D(A)$ of a Banach space X, we define the *graph norm* of A by

$$\|f\|_A := \|f\| + \|Af\| \quad \text{for } f \in D(A).$$

Then, indeed, $\|\cdot\|_A$ is a norm on $D(A)$[2]. The operator A is called *closed* if $D(A)$ is complete with respect to this graph norm, i.e., if $D(A)$ is a Banach space with this graph norm $\|\cdot\|_A$.

The following proposition yields simple yet useful reformulations of the closedness of a linear operator, we leave out the proof as Exercise 1.

Proposition 11.3.5. *Let A be a linear operator with domain $D(A)$ in X. The following assertions are equivalent.*

(i) A is a closed operator.

(ii) For every sequence $(f_n) \subseteq D(A)$ with $f_n \to f$ and $Af_n \to g$ in X for some $f, g \in X$ one has $f \in D(A)$ and $Af = g$.

[2] If X is a Hilbert space, it is customary to define the graph norm as $\|f\|_A^2 := \|f\|^2 + \|Af\|^2$, which makes $D(A)$ a pre-Hilbert space. Clearly, the two definitions yield equivalent norms.

If A is injective the properties above are further equivalent to the following:

(iii) The inverse A^{-1} (defined on the range of A) of A is a closed operator.

The main result of this section summarises the basic properties of the generator.

Theorem 11.3.6. *The generator of a C_0-semigroup is a closed and densely defined linear operator that determines the semigroup uniquely.*

Proof. To show the closedness of A, let $(f_n) \subseteq D(A)$ be a sequence and $f, g \in X$ such that

$$f_n \to f \text{ and } Af_n \to g.$$

We have to show that $f \in D(A)$ and $Af = g$.

For $t > 0$ we have

$$T(t)f_n - f_n = \int_0^t T(s)Af_n \mathrm{d}s$$

using Proposition 11.3.4. If we set $u_n(s) := T(s)Af_n$ and $u(s) := T(s)g$, then $u_n \to u$ uniformly on $[0, t]$ because T is locally bounded. So we can pass to the limit in the identity above, and obtain

$$T(t)f - f = \int_0^t T(s)g\mathrm{d}s.$$

From this we deduce that the function $t \mapsto T(t)f$ is differentiable at 0 with derivative $u(0) = g$. This means precisely that $f \in D(A)$ and $Af = g$ and implies therefore that A is a closed operator.

We now show that $D(A)$ is dense in X. Let $f \in X$ be arbitrary and define

$$v(t) := \frac{1}{t}\int_0^t T(s)f\mathrm{d}s \quad (t > 0).$$

By Proposition 11.3.4 we obtain that $v(t) \in D(A)$. Since the function $s \mapsto T(s)f$ is continuous, we have $v(t) \to T(0)f = f$ for $t \downarrow 0$.

Suppose $(S(t))_{t\geq 0}$ is a C_0-semigroup with the same generator A as $(T(t))_{t\geq 0}$. Let $f \in D(A)$ and $t > 0$ be fixed, and consider the function $u : [0, t] \to X$ given by $u(s) := T(t-s)S(s)f$. Then u is differentiable and its derivative is given by the product rule, see Appendix, Theorem B.1.4,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}u(s) &= (\tfrac{\mathrm{d}}{\mathrm{d}s}T(t-s))S(s)f + T(t-s)\tfrac{\mathrm{d}}{\mathrm{d}s}(S(s)f)\\ &= -AT(t-s)S(s)f + T(t-s)AS(s)f.\end{aligned}$$

Recall that the semigroup and its generator commute on $D(A)$, see Proposition 11.3.3.(2), we obtain that the right-hand term is 0, so u must be constant. This implies

$$S(t)f = u(t) = u(0) = T(t)f,$$

i.e., the bounded linear operators $S(t)$ and $T(t)$ coincide on the dense subspace $D(A)$, hence they must be equal everywhere. □

11.4 Notes and Remarks

Operator semigroups has been widely studied during the last decades and there are many monographs dealing with them. We mention here the excellent graduate texts by Engel and Nagel [1, 2], which motivated many parts of this manuscript. The first milestone in the theory was the opus of Hille and Phillips [4]. An important later reference is the dense book of Pazy [5], which was written from a PDE perspective and Goldstein [3], which contains lots of other applications as well.

11.5 Exercises

1. Show the properties of closed operators as listed in Proposition 11.3.5
2. Let $X = \ell^p$ for $p \in [1, \infty)$, and the operator A be given as

$$A(x_n) := (a_n x_n)$$

where $(a_n) \subset \mathbb{C}$ is a given sequence and with $D(A) := \{(x_n) \in \ell^p \,:\, (a_n x_n) \in \ell^p\}$. Show that:

   (a) $A \in \mathcal{L}(X)$ if and only if $(a_n) \in \ell^\infty$.

   (b) $D(A)$ is always dense.

   (c) A is always a closed operator.

   (d) A generates a C_0-semigroup $(T(t))_{t\geq 0}$ if and only if there is $\omega \in \mathbb{R}$ such that $\operatorname{Re} a_n \leq \omega$ for all $n \in \mathbb{N}$. In this case,

$$T(t)(x_n) = \left(\mathrm{e}^{ta_n} x_n\right).$$

   (e) Let $a_n = -n^2$. Then $(T(t))_{t\geq 0}$ is continuous in the operator norm on $(0, \infty)$, but not right continuous at $t = 0$.

3. Let $X = \mathrm{C}_0(\mathbb{R})$ and $q \in C(\mathbb{R})$. Consider the operator $(Af)(s) := q(s)f(s)$ with $D(A) := \{f \in X \,:\, qf \in X\}$ and make analogous statements as in the previous exercise. Prove these statements.
4. For a C_0-semigroup $(T(t))_{t\geq 0}$ and a boundedly invertible transformation $R \in \mathcal{L}(X)$ define $S(t) := R^{-1}T(t)R$. Prove that $(S(t))_{t\geq 0}$ is a C_0-semigroup as well. Determine its growth bound and its generator.
5. For a C_0-semigroup $(T(t))_{t\geq 0}$ and $z \in \mathbb{C}$ define $S(t) := \mathrm{e}^{tz}T(t)$. Prove that $(S(t))_{t\geq 0}$ is a C_0-semigroup, determine its growth bound and its generator.
6. For a C_0-semigroup $(T(t))_{t\geq 0}$ and $\alpha \geq 0$ define $S(t) := T(\alpha t)$. Prove that $(S(t))_{t\geq 0}$ is a C_0-semigroup, determine its growth bound and its generator.

7. Consider the closed subspace

$$\mathrm{C}_0([0,1)) := \big\{f \in \mathrm{C}([0,1]) \,:\, f(1)=0\big\}$$

of the Banach space $\mathrm{C}([0,1])$ of continuous functions on $[0,1]$. Define the nilpotent left shift semigroup on it and determine its generator.

8. Let $\mathrm{F_b}(\mathbb{R})$ denote the linear space of all bounded $\mathbb{R} \to \mathbb{R}$ functions. Define

$$(S(t)f)(s) := f(t+s) \quad \text{for } f \in \mathrm{F_b}(\mathbb{R}),\ s \in \mathbb{R},\ t \geq 0.$$

Prove that each of the following spaces is a Banach space with the supremum norm $\|\cdot\|_\infty$ and invariant under $S(t)$ for all $t \geq 0$. Is $(S(t))_{t\geq 0}$ a C_0-semigroup on these spaces?

   (a) $\mathrm{F_b}(\mathbb{R})$.
   (b) $\mathrm{C_b}(\mathbb{R})$ = the space bounded and continuous functions.
   (c) $\mathrm{C_0}(\mathbb{R})$ = the space bounded and continuous functions vanishing at infinity.

9. Determine the set of those $f \in \mathrm{C_{ub}}(\mathbb{R})$ for which $t \mapsto S(t)f$ is differentiable ($(S(t))_{t\geq 0}$ denotes the left shift semigroup).

10. Determine whether the following operators are closed or not:

   (a) $X := C[0,1]$, $Af(s) := \frac{1}{s(1-s)} f(s)$, $D(A) := \{f \in X \,:\, Af \in X\}$
   (b) $X := C[0,1]$, $Bf(s) := f'(s)$, $D(B) := \{f \in C^1[0,1] : f'(1) = 0\}$
   (c) $X := C[0,1]$, $Cf(s) := f'(s)$, $D(C) := \{f \in C^1[0,1] : f(0) = f(1)\}$
   (d) $X := C[0,1]$, $Df(s) := f''(s)$, $D(D) := C^2[0,1]$
   (e) $X := C[0,1]$, $Ef(s) := f''(s)$, $D(E) := \{f \in C^2[0,1],\ f(0) = f(1) = 0\}$
   (f) $X := C[0,1]$, $Ff(s) := f''(s)$, $D(F) := \{f \in C^2[0,1],\ f''(0) = 0\}$

11. Let $X = \mathrm{C_0}(\mathbb{R})$, $q \in \mathrm{C_b}(\mathbb{R})$, and

$$T(t)f(s) := e^{\int_{s-t}^{t} q(\tau)d\tau} \cdot f(s-t).$$

Show that $(T(t))_{t\geq 0}$ is a C_0-semigroup and identify its generator.

12. Let $X := L^p[1,\infty)$, $1 \leq p < \infty$ and $(T(t)f)(s) := f(se^t)$. Show that $(T(t))_{t\geq 0}$ is a C_0-semigroup and that $\omega_0(T) = -\frac{1}{p}$. Can you identify its generator?

13. Consider some semigroup examples appearing in this lecture and write down the corresponding abstract initial value problems. Can you associate partial differential equations to these initial value problems?

Bibliography

[1] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Math., vol. 194, Springer-Verlag, 2000.

[2] ———, *A Short Course on Operator Semigroups*, Springer-Verlag, 2006.

[3] J.A. Goldstein, *Semigroups of Operators and Applications*, Oxford University Press, 1985.

[4] E. Hille and R.S. Phillips, *Functional Analysis and Semigroups*, Amer. Math. Soc. Coll. Publ., vol. 31, Amer. Math. Soc., 1957.

[5] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Appl. Math. Sci., vol. 44, Springer-Verlag, 1983.

Lecture 12

Generation Properties

This week we investigate spectral properties of generators and characterize semigroup generators. Before doing so, however, let us analyze one fundamental example.

12.1 The Gaussian semigroup

Consider the heat equation on the entire line $\mathbb{R}$:

$$\begin{aligned} \partial_t w(t,x) &= \partial_{xx} w(t,x), \quad t \geq 0,\ x \in \mathbb{R}, \\ w(0,x) &= w_0(x), \quad x \in \mathbb{R}. \end{aligned} \tag{12.1}$$

Here w_0 is a function on $\mathbb{R}$ providing the initial heat profile. We seek the solution to this problem as an orbit map of some *semigroup*. To find a candidate for this semigroup we first make some formal computations by using the Fourier transform, which is given for $f \in \mathrm{L}^1(\mathbb{R})$ by the Fourier integral

$$\tilde{f}(\xi) := \mathcal{F}(f)(\xi) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-i\xi x} f(x) \mathrm{d}x.$$

Recall that the operator $\mathcal{F}$ maps differentiation to multiplication by the Fourier variable $i\xi$, i.e., $\mathcal{F}(\partial_x f(x))(\xi) = i\xi \mathcal{F}(f)(\xi)$. If we take the Fourier transform of equation (12.1) with respect to x and interchange the actions of $\mathcal{F}$ and ∂_t, we obtain

$$\begin{aligned} \partial_t \tilde{w}(t,\xi) &= -\xi^2 \tilde{w}(t,\xi) \quad t \geq 0,\ \xi \in \mathbb{R} \\ \tilde{w}(0,\xi) &= \tilde{w}_0(\xi), \quad \xi \in \mathbb{R}. \end{aligned}$$

This is an ordinary differential equation for $\tilde{w}$, which is easy to solve:

$$\tilde{w}(t,\xi) = \mathrm{e}^{-t|\xi|^2} \tilde{w}_0(\xi).$$

To get w back we take the inverse Fourier transform of this solution:

$$w(t,\cdot) = \mathcal{F}^{-1}(\tilde{w}(t,\cdot)) = \frac{1}{\sqrt{2\pi}} \mathcal{F}^{-1}(\mathrm{e}^{-t|\cdot|^2}) * \mathcal{F}^{-1}(\tilde{w}_0),$$

where we used that $\mathcal{F}^{-1}$ maps products to convolutions. At this point we only have to remember that

$$\mathcal{F}^{-1}(\mathrm{e}^{-t|\cdot|^2})(x) = \frac{1}{\sqrt{2t}}\mathrm{e}^{-\frac{|x|^2}{4t}}.$$

So if we set

$$g_t(x) := \frac{1}{\sqrt{4\pi t}}\mathrm{e}^{-\frac{|x|^2}{4t}} \quad (t > 0),$$

then the candidate for the solution to (12.1) is of the form

$$w(t) = g_t * w_0 \quad \text{for } t > 0.$$

Let us collect some fundamental properties of the function g_t.

Remark 12.1.1.

1. Consider the *standard Gaussian function*

   $$g(x) := \frac{1}{\sqrt{4\pi}}\mathrm{e}^{-\frac{x^2}{4}}.$$

   Then $g \geq 0$, $\|g\|_1 = 1$ and the function g belongs to $\mathrm{L}^p(\mathbb{R})$ for all $p \in [1, \infty]$.

2. We have $g_t(x) = \frac{1}{\sqrt{t}} g\big(\frac{x}{\sqrt{t}}\big)$, hence $g_t \geq 0$, $\|g_t\|_1 = 1$ and

   $$\lim_{t\downarrow 0} \int_{|x|>r} g_t(s)\, ds = 0 \quad \text{for all } r > 0 \text{ fixed.}$$

The function

$$G(t,x,y) := g_t(x-y) \quad (t > 0,\ x \in \mathbb{R},\ y \in \mathbb{R})$$

is called the *heat* or *Gaussian kernel* on $\mathbb{R}$ and gives rise to a semigroup, called the *heat* or *Gaussian semigroup.*

Proposition 12.1.2. *Let $p \in [1, \infty)$. For $f \in L^p(\mathbb{R})$ and $t > 0$ define*

$$(T(t)f)(x) := (g_t * f)(x) = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} f(y) e^{-\frac{(x-y)^2}{4t}} \mathrm{d}y = \int_{\mathbb{R}} f(y) G(t,x,y) \mathrm{d}y, \text{ and set}$$
$$T(0)f := f.$$

Then $T(t)$ is a linear contraction on $L^p(\mathbb{R})$, and $(T(t))_{t\geq 0}$ is a strongly continuous semigroup.

Proof. Let $f \in \mathrm{L}^p(\mathbb{R})$. By Young's inequality and since $g_t \in \mathrm{L}^1(\mathbb{R})$, we obtain that the convolution $g_t * f$ exists and

$$\|g_t * f\|_p \leq \|g_t\|_1 \cdot \|f\|_p = \|f\|_p.$$

In particular, $g_t * f$ belongs to $\mathrm{L}^p(\mathbb{R})$. Since linearity of $f \mapsto g_t * f$ follows immediately from the definition, we obtain that $T(t)$ is a linear contraction.

To prove the semigroup property, we employ the Fourier transform. Fixing $f \in \mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^p(\mathbb{R})$, we can take the Fourier transform of $g_t * (g_s * f)$ to obtain

$$\mathcal{F}(g_t * (g_s * f)) = \sqrt{2\pi}\mathcal{F}(g_t) \cdot \mathcal{F}(g_s * f) = (2\pi)\mathcal{F}(g_t) \cdot \mathcal{F}(g_s) \cdot \mathcal{F}(f).$$

Recall from above that

$$\mathcal{F}(g_t)(\xi) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t\xi^2}, \quad \text{therefore,} \quad \mathcal{F}(g_t)(\xi) \cdot \mathcal{F}(g_s)(\xi) = \frac{1}{2\pi}\mathrm{e}^{-(t+s)\xi^2} = \frac{1}{\sqrt{2\pi}}\mathcal{F}(g_{t+s})(\xi).$$

This yields

$$\mathcal{F}(g_t * (g_s * f)) = \sqrt{2\pi}\mathcal{F}(g_{t+s}) \cdot \mathcal{F}(f) = \mathcal{F}(g_{t+s} * f),$$

hence $g_t * (g_s * f) = g_{t+s} * f$. Therefore, the equality $T(t)T(s)f = T(t+s)f$ holds for $f \in \mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^p(\mathbb{R})$. By the continuity of the semigroup operators and by the denseness of this subspace in $\mathrm{L}^p(\mathbb{R})$, we obtain the equality everywhere.
From the properties of the function g_t listed in Remark 12.1.1.2, it follows that $g_t * f \to f$ in $\mathrm{L}^p(\mathbb{R})$ if $t \downarrow 0$. Hence the semigroup $(T(t))_{t\geq 0}$ is strongly continuous. □

12.2 Resolvent of a generator

We have seen in previous lectures that spectral analysis of matrices, more precisely, the determination of eigenvalues and eigenvectors, led to a construction of the semigroup generated by them. We investigate now some basic spectral properties of semigroup generators. Let us begin with the following fundamental spectral theoretic notions.

Definition 12.2.1. Let A be a closed linear operator defined on a linear subspace $D(A)$ of a Banach space X.

1. The *spectrum* of A is the set

$$\sigma(A) := \big\{\lambda \in \mathbb{C} : \lambda - A : D(A) \to X \text{ is not bijective}\big\}.$$

2. The *resolvent set* of A is $\rho(A) := \mathbb{C} \setminus \sigma(A)$, i.e.,

$$\rho(A) := \big\{\lambda \in \mathbb{C} : \lambda - A : D(A) \to X \text{ is bijective}\big\}.$$

3. If $\lambda \in \rho(A)$ then $\lambda - A$ is bijective, hence has an algebraic inverse $(\lambda - A)^{-1}$. We call this operator the resolvent of A at point λ and denote it by

$$R(\lambda, A) := (\lambda - A)^{-1}.$$

It is important to note that if $\lambda \in \rho(A)$, then its algebraic inverse

$$(\lambda - A)^{-1} : X \to D(A)$$

is defined on the entire X. Since A is closed so are $\lambda - A$ and its inverse. As consequence of the closed graph theorem, see Appendix, Theorem B.1.5, we immediately obtain that the operator $(\lambda - A)^{-1}$ is bounded.

Proposition 12.2.2. *For a closed linear operator A and for $\lambda \in \rho(A)$ we have*

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

Let us summarise also some fundamental properties of spectrum and the resolvent.

Proposition 12.2.3. *Let X be a Banach space and let A be a closed linear operator with domain $D(A) \subseteq X$. Then the following assertions are true:*

(i) The resolvent set $\rho(A)$ is open, hence its complement, the spectrum $\sigma(A)$, is closed.

(ii) The mapping

$$\rho(A) \ni \lambda \mapsto R(\lambda, A) \in \mathcal{L}(X)$$

is complex differentiable. Moreover, for $n \in \mathbb{N}$ we have

$$\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} R(\lambda, A) = (-1)^n n! R(\lambda, A)^{n+1}.$$

Proof. Statement (i) follows from the following Neumann series representation of the resolvent: For $\mu \in \rho(A)$ we have

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

Hence, for $|\lambda - \mu| < \frac{1}{\|R(\mu,A)\|}$, we obtain

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

Assertion (ii) follows from the power series representation in (i) and from the fact that a power series is always a Taylor series. □

The following result known as the spectral mapping theorem for the resolvent will be useful later on.

Proposition 12.2.4. *Let A be a closed linear operator with $\rho(A) \neq \emptyset$. Then for any $\lambda \in \rho(A)$,*

$$\sigma(R(\lambda, A)) \setminus \{0\} = \left\{ \frac{1}{\lambda - \mu} : \mu \in \sigma(A) \right\}.$$

Proof. For $0 \neq \alpha \in \mathbb{C}$ and $\lambda \in \rho(A)$, we have

$$\begin{aligned}(\alpha - R(\lambda, A))x &= \alpha \left[(\lambda - \frac{1}{\alpha}) - A \right] R(\lambda, A)x, \quad \forall x \in X, \\ &= \alpha R(\lambda, A) \left[(\lambda - \frac{1}{\alpha}) - A \right] x, \quad \forall x \in D(A).\end{aligned}$$

Thus, $\alpha \in \sigma(R(\lambda, A))$ if and only if $\lambda - \frac{1}{\alpha} \in \sigma(A)$. □

As a corollary one determines the spectral radius of $R(\lambda, A)$.

Corollary 12.2.5. *For $\lambda \in \rho(A)$ one has*

$$\operatorname{dist}(\lambda, \sigma(A)) = \frac{1}{r(R(\lambda, A))} \geq \frac{1}{\|R(\lambda, A)\|}.$$

Proof. Let $\lambda \in \rho(A)$. Then, By Proposition 12.2.4 we have

$$\begin{aligned} \operatorname{dist}(\lambda, \sigma(A)) &= \inf\{|\lambda - \mu| : \mu \in \sigma(A)\} \\ &= \left(\sup\left\{\left|\frac{1}{\lambda - \mu}\right| : \mu \in \sigma(A)\right\}\right)^{-1} \\ &= (\max\{|\alpha| : \alpha \in \sigma(R(\lambda, A))\})^{-1} \\ &= \frac{1}{r(R(\lambda, A))} \geq \frac{1}{\|R(\lambda, A)\|}. \end{aligned}$$

□

Our next aim is to prove that the resolvent set of a generator A is non-empty, and to relate the resolvent of A to the semigroup $(T(t))_{t\geq 0}$. The first step is provided by the next lemma.

Lemma 12.2.6. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup with generator A. Then for all $\lambda \in \mathbb{C}$ and $t > 0$ the following identities hold:*

$$\begin{aligned} e^{-\lambda t}T(t)f - f = (A - \lambda) \int_0^t e^{-\lambda s}T(s)f\, ds \qquad & \text{if } f \in X, \\ = \int_0^t e^{-\lambda s}T(s)(A - \lambda)f\, ds \qquad & \text{if } f \in D(A). \end{aligned}$$

Proof. Observe that $S(t) = \mathrm{e}^{-\lambda t}T(t)$ is also a strongly continuous semigroup with generator $B = A - \lambda$, see Exercise 11.5.5. Hence, we can apply Proposition 11.3.4.(ii). □

With the help of this lemma we obtain the following important relations between the resolvent of the generator and the semigroup.

Proposition 12.2.7. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup of type (M, ω) with generator A. Then the following assertions are true:*

(i) For all $f \in X$ and $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$ we have

$$R(\lambda, A)f = \int_0^\infty e^{-\lambda s}T(s)f\mathrm{d}s = \lim_{N\to\infty} \int_0^N e^{-\lambda s}T(s)f\mathrm{d}s. \tag{12.2}$$

(ii) For all $f \in X$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$ and $n \in \mathbb{N}$ we have

$$R(\lambda, A)^n f = \frac{1}{(n-1)!} \int_0^\infty s^{n-1} e^{-\lambda s}T(s)f\mathrm{d}s.$$

(iii) For all $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > \omega$ *we have*

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\operatorname{Re}\lambda - \omega)^n}. \tag{12.3}$$

Proof. From Lemma 12.2.6, closedness of A (see Theorem 11.3.6), and by taking limit as $t \to \infty$ we conclude that for $\operatorname{Re}\lambda > \omega$ we have

$$\begin{aligned} -f &= (A - \lambda) \int_0^\infty \mathrm{e}^{-\lambda s} T(s) f \, ds \qquad \text{if } f \in X, \\ &= \int_0^\infty \mathrm{e}^{-\lambda s} T(s)(A - \lambda) f \, ds \qquad \text{if } f \in D(A). \end{aligned}$$

Since this expression gives a bounded operator, (i) is proved. To show (ii) notice that

$$R(\lambda, A)^n f = \frac{(-1)^{n-1}}{(n-1)!} \frac{\mathrm{d}^{n-1}}{\mathrm{d}\lambda^{n-1}} R(\lambda, A) f = \frac{1}{(n-1)!} \int_0^\infty s^{n-1} \mathrm{e}^{-\lambda s} T(s) f \mathrm{d}s.$$

Finally, to see (iii) we make a norm estimate and obtain

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

which finishes the proof. □

An important property of positive semigroups now follows immediately.

Corollary 12.2.8. *Let E be a Banach lattice and $(T(t))_{t \geq 0}$ a positive semigroup. Then $R(\lambda, A) \geq 0$ for $\lambda > \omega_0(T)$.*

Proof. The statement is immediate from the Laplace transform representation of the resolvent (12.2). □

Let us summarise the above as follows:

If A is the generator of an operator semigroup $(T(t))_{t \geq 0}$, then it is closed, densely defined, and a suitable right half plane belongs to its resolvent set, where the estimate (12.3) holds. The resolvent operators are given by the *Laplace transform* (12.2) of the semigroup.

12.3 Generation theorems

In the previous section some quite nice properties of semigroups generators ant their resolvents were collected. It turns out that some of these properties actually characterise semigroup generators.

Theorem 12.3.1 (Hille-Yosida). *Let $(A, D(A))$ be a linear operator on a Banach space X. Then the following properties are equivalent.*

(i) $(A, D(A))$ *generates a C_0-semigroup* $(T(t))_{t\geq 0}$.

(ii) $(A, D(A))$ *is closed, densely defined, and there exist* $M \geq 1$ *and* $w \in \mathbb{R}$ *such that for every* $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > w$ *one has* $\lambda \in \rho(A)$ *and*

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\operatorname{Re}\lambda - w)^n} \quad \text{for all } n \in \mathbb{N}. \tag{12.4}$$

(iii) The Abstract Cauchy Problem

$$\text{(ACP)} \qquad \begin{cases} \dot{u}(t) = Au(t) & \text{for } t \geq 0, \\ u(0) = f \end{cases}$$

is well-posed *in the following sense: The domain $D(A)$ is dense and for each $f \in D(A)$ there exists a unique* classical solution $u = u(\cdot, f)$ *depending continuously on the initial value f. More precisely, $u(\cdot, f) : \mathbb{R}_+ \to X$ is continuously differentiable, $u(t, f) \in D(A)$ for all $t \geq 0$, (ACP) holds, and for every sequence $(f_n)_{n\in\mathbb{N}} \subset D(A)$ converging to 0 one has $\lim_{n\to\infty} u(t, f_n) = 0$ uniformly for t in compact intervals of $\mathbb{R}_+$.*

We will not prove this theorem here in its full generality but only follow through the main idea of its proof. As a first step we need the following approximation result.

Lemma 12.3.2. *Let A be a closed, densely defined operator. Assume that there are $M \geq 1$ and $\omega \in \mathbb{R}$ such that for all $\lambda > \omega$, we have $\lambda \in \rho(A)$ and $\|\lambda R(\lambda, A)\| \leq M$. Then*

(i) $\lambda R(\lambda, A)g \to g$ *for all* $g \in X$ *as* $\lambda \to \infty$, *and*

(ii) $\lambda AR(\lambda, A)f \to Af$ *for all* $f \in D(A)$ *as* $\lambda \to \infty$.

Proof. Taking $g \in D(A)$, we see that $\lambda R(\lambda, A)g = R(\lambda, A)Ag + g$. By assumption,

$$\|R(\lambda, A)Ag\| \leq \frac{M}{\lambda}\|Ag\|,$$

and hence $\lambda R(\lambda, A)g \to g$ as $\lambda \to \infty$. By the denseness of $D(A)$ and the boundedness, the convergence follows for all $g \in X$. The second statement is an immediate consequence of the first one, inserting $g = Af$, and the fact that A and $R(\lambda, A)$ commute. □

The operators $\lambda AR(\lambda, A)$, $\lambda > \omega$, are called *Yosida approximants.*

Proof of Hille-Yosida Theorem, special case. The core of the proof is the case $M = 1$, $\omega = 0$. The general situation can be reduced to this case by some scaling and renorming techniques, which we omit here (see Exercises 12.6.3 and 4).

Let

$$A_n := nAR(n, A) = n^2 R(n, A) - nI,$$

which are bounded operators for each $n \in \mathbb{N}$ and commute with each other. Consider the uniformly continuous semigroups given by

$$T_n(t) := \mathrm{e}^{tA_n}, \quad t \geq 0.$$

Since A_n converges to A pointwise on $D(A)$, we anticipate that the following properties hold.

- $T(t)f := \lim_{n\to\infty} T_n(t)f$ exists for each $f \in X$.
- $(T(t))_{t\geq 0}$ is a C_0-semigroup of contractions on X.
- This generator of this semigroup is $(A, D(A))$.

By establishing these statements we will complete the proof.

First of all, each $T_n(t)$ is a contraction semigroup, since

$$\|T_n(t)\| \leq \mathrm{e}^{-nt}\mathrm{e}^{\|n^2R(n,A)\|t} \leq \mathrm{e}^{-nt}\mathrm{e}^{nt} = 1 \qquad \text{for } t \geq 0.$$

So, again, it suffices to prove convergence just on $D(A)$. By the vector-valued version of the fundamental theorem of calculus applied to the functions

$$s \mapsto T_m(t-s)T_n(s)f$$

for $0 \leq s \leq t$, $f \in D(A)$, and $m, n \in \mathbb{N}$, and using the mutual commutativity of the semigroups $T_n(t)$ for all $n \in \mathbb{N}$, one has

$$\begin{aligned} T_n(t)x - T_m(t)f &= \int_0^t \left(T_m(t-s)T_n(s)f\right)\, ds \\ &= \int_0^t T_m(t-s)T_n(s)(A_nf - A_mf)\, ds. \end{aligned}$$

Accordingly,

$$\|T_n(t)f - T_m(t)f\| \leq t\,\|A_nf - A_mf\|.$$

Recall that $(A_nf)_{n\in\mathbb{N}}$ is a Cauchy sequence for each $f \in D(A)$. Therefore, $(T_n(t)f)_{n\in\mathbb{N}}$ converges uniformly on each interval $[0, t_0]$.

As a next step, the pointwise convergence of $(T_n(t)f)_{n\in\mathbb{N}}$ implies that the limit family $(T(t))_{t\geq 0}$ satisfies the semigroup property and consists of contractions. Moreover, for each $f \in D(A)$, the corresponding orbit map

$$\xi : t \mapsto T(t)f, \quad 0 \leq t \leq t_0,$$

is the uniform limit of continuous functions, and so is continuous itself. This suffices to obtain strong continuity via Proposition 11.2.4.

To finish the proof, denote by B the generator of $(T(t))_{t\geq 0}$ and fix $f \in D(A)$. On each compact interval $[0, t_0]$, the functions

$$\xi_n : t \mapsto T_n(t)f$$

converge uniformly to $\xi(\cdot)$, while the differentiated functions

$$\dot{\xi}_n : t \mapsto T_n(t)A_n f$$

converge uniformly to

$$\eta : t \mapsto T(t)Af.$$

This implies differentiability of ξ with $\dot{\xi}(0) = \eta(0)$, i.e., $D(A) \subset D(B)$ and $Af = Bf$ for $f \in D(A)$.
Now choose $\lambda > 0$. Then $\lambda - A$ is a bijection from $D(A)$ onto X, since $\lambda \in \rho(A)$ by assumption. On the other hand, B generates a contraction semigroup, and so $\lambda \in \rho(B)$. Hence, $\lambda - B$ is also a bijection from $D(B)$ onto X. But we have seen that $\lambda - B$ coincides with $\lambda - A$ on $D(A)$. This is possible only if $D(A) = D(B)$ and $A = B$. □

From this construction we immediately obtain a characterization of positive semigroups.

Corollary 12.3.3. *Let E be a Banach lattice, $(T(t))_{t\geq 0}$ a C_0-semigroup of contractions with generator A and suppose that $R(\lambda, A) \geq 0$ for λ sufficiently large. Then $T(t)$ is positive for all $t \geq 0$.*

Proof. Observe that for n sufficiently large, $R(n, A) \geq 0$, and hence

$$T_n(t) = \mathrm{e}^{-n}\mathrm{e}^{n^2 R(n,A)} \geq 0.$$

The statement then follows from the approximation argument in the proof of the Hille-Yosida theorem. □

12.4 Bounded perturbations

As an application of the generation theorem, let us mention one basic perturbation result. The idea behind perturbation theorems is always the same: We start with a generator A and assume that the operator B is "nice enough". Then $A + B$ generates a semigroup.

Theorem 12.4.1. *If A generates a semigroup $(T(t))_{t\geq 0}$ of type (M,ω) and $B \in \mathcal{L}(X)$, then $A+B$ with $D(A+B) = D(A)$ generates a semigroup $(S(t))_{t\geq 0}$ of type $(M, \omega + \|B\|)$.*

Proof. First we change to the operator to $A-\omega$ and then use the renorming procedure mentioned in the proof of the Hille-Yosida theorem and only consider without the loss of generality the case when A generates a semigroup of type $(1,0)$, i.e., a contraction semigroup.

As a next step, we show that the operator $A+B$ has non-empty resolvent set. More precisely, if $\lambda > 0$, we can use the identity

$$\lambda - A - B = (I - BR(\lambda, A))\,(\lambda - A), \tag{12.5}$$

showing that if $\|BR(\lambda, A)\| < 1$, then $\lambda \in \rho(A+B)$ and

$$R(\lambda, A+B) = R(\lambda, A) \sum_{n=0}^{\infty} (BR(\lambda, A))^n\,. \tag{12.6}$$

By assumption, A is a generator of a contraction semigroup, and hence $\lambda\|R(\lambda, A)\| \leq 1$. Hence, if $\lambda > \|B\|$, then $\lambda \in \rho(A+B)$ and (12.6) holds.

For the norm of this resolvent we see by crudely estimating the series term-by-term that

$$\|R(\lambda, A+B)\| = \frac{1}{\lambda} \sum_{n=0}^{\infty} \left(\frac{\|B\|}{\lambda}\right)^n = \frac{1}{\lambda} \cdot \frac{1}{1 - \frac{\|B\|}{\lambda}} = \frac{1}{\lambda - \|B\|}. \tag{12.7}$$

By the Hille-Yosida theorem, this means that the operator $A + B - \|B\|$ generates a contraction semigroup, and hence A generates a semigroup $(S(t))_{t\geq 0}$ satisfying

$$\|S(t)\| \leq \mathrm{e}^{t\|B\|},$$

which proves the statement. □

12.5 Notes and Remarks

For most parts of this lecture we refer again to the texts by Engel and Nagel [1, 2], the notes by Pazy [4], and Goldstein [3].

12.6 Exercises

1. Consider the operators from Exercise 11.5.10. Which of them are generators?

2. Prove that the generator of the Gaussian semigroup defined in Section 12.1 is given by

$$Af = f'', \quad f \in D(A) = \{f \in L^p(\mathbb{R}) : f'' \text{ (in the sense of distribution) is in } L^p(\mathbb{R})\}.$$

3. Let A be an operator satisfying the estimates (12.4). Without loss of generality (by considering $A-\omega$ instead of A) one can take $\omega = 0$ in (12.4). For every $\mu > 0$, define a new norm on X by
$$\|x\|_\mu := \sup_{n\geq 0} \|\mu^n R(\mu, A)^n x\|.$$
Show that:
   (a) $\|x\| \leq \|x\|_\mu \leq M\|x\|$, i.e., they are all equivalent to $\|\cdot\|$.
   (b) $\|\mu R(\mu, A)\|_\mu \leq 1$.
   (c) $\|\lambda R(\lambda, A)\|_\mu \leq 1$ for all $0 < \lambda \leq \mu$.
   (d) $\|\lambda^n R(\lambda, A)^n x\| \leq \|\lambda^n R(\lambda, A)^n x\|_\mu \leq \|x\|_\mu$ for all $0 < \lambda \leq \mu$ and $n \in \mathbb{N}$.
   (e) $\|x\|_\lambda \leq \|x\|_\mu$ for $0 < \lambda \leq \mu$.
4. Continuing and using the notation of the previous exercise, show that for the norm
$$|||x||| := \sup_{\mu>0} \|x\|_\mu$$
we have
   (a) $\|x\| \leq |||x||| \leq M\|x\|$
   (b) $|||\lambda R(\lambda, A)||| \leq 1$ for all $\lambda > 0$.
5. Let $F : \mathbb{R} \to \mathbb{R}$ be continuously differentiable with $\sup_{x\in\mathbb{R}} |F'(x)| < \infty$. Define the flow $\Phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ as the solution of the nonlinear ODE
$$\begin{cases} \dfrac{d}{dt}y(t) = F(y(t)) \\ \quad y(0) = s, \end{cases}$$
i.e., $\Phi(t, s) := y(t)$. Take $X := C_0(\mathbb{R})$ and define
$$\big(T(t)f\big)(s) := f\big(\Phi(t, s)\big)$$
for $t \geq 0$, $s \in \mathbb{R}$.
   (a) Show that $(T(t))_{t\geq 0}$ is a contraction semigroup (i.e., of type $(1, 0)$) and identify its generator.
   (b) What is the corresponding abstract Cauchy problem? Which partial differential equation can we associate with it? Relate the semigroup $(T(t))_{t\geq 0}$ to the method of characteristics.
6. Let $(T(t))_{t\geq 0}$ be a semigroup on the Banach space X with generator A. Prove that for all $f \in D(A^2)$ we have the Taylor formula
$$T(t)f = f + tAf + \int_0^t (t-s)T(s)A^2 f \mathrm{d}s.$$
Find a general Taylor formula for $f \in D(A^n)$.

7. Let $(T(t))_{t\geq 0}$ be a contraction semigroup on the Banach space X with generator A. Prove that
$$\|Af\|^2 \leq 4\|A^2 f\| \cdot \|f\|$$
holds for all $f \in D(A^2)$.

Bibliography

[1] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Math., vol. 194, Springer-Verlag, 2000.

[2] ______, *A Short Course on Operator Semigroups*, Springer-Verlag, 2006.

[3] J.A. Goldstein, *Semigroups of Operators and Applications*, Oxford University Press, 1985.

[4] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Appl. Math. Sci., vol. 44, Springer-Verlag, 1983.

Lecture 13

Spectral Theory for Positive Semigroups I

This week we are concerned with the remarkable spectral properties of positive semigroups on Banach lattices. Some beautiful facts we have observed in final dimensional case will reappear in infinite-diemnsional setting.
Throughout this chapter we suppose that E is a complex Banach lattice and denote by X a Banach space.

13.1 Stability of C_0-semigroups

We are interested in the asymptotic behaviour of the solution of the abstract Cauchy problem

$$(ACP) \qquad \begin{cases} \dot{u}(t) = Au(t), & t \geq 0, \\ u(0) = f \in X, \end{cases}$$

where A is the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ on X. Recall from Proposition 11.3.3 that the solution to this Cauchy problem is given by $u(t) := T(t)f$. In Lecture 11 we have also already defined the *growth bound* of the semigroup $(T(t))_{t\geq 0}$ as

$$\omega_0(T) := \inf\{\omega \in \mathbb{R} : \text{ there is } M = M_\omega \geq 1 \text{ with } \|T(t)\| \leq Me^{\omega t} \text{ for all } t \geq 0\}.$$

There is an important connection between the growth bound and the spectral radius of the semigroup operators.

Proposition 13.1.1. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup on X. Then*

(i) $\omega_0(T) = \lim_{t\to\infty} t^{-1} \log \|T(t)\| = \inf_{t>0} t^{-1} \log \|T(t)\|$.

(ii) For every $t \geq 0$, the spectral radius $r(T(t))$ of the operator $T(t)$ equals $e^{t\omega_0(T)}$.

Proof. (i): By Exercise 13.4.1 we have $\lim_{t\to\infty} t^{-1} \log \|T(t)\| = \inf_{t>0} t^{-1} \log \|T(t)\|$. Set

$$\eta := \inf_{t>0} t^{-1} \log \|T(t)\| = \lim_{t\to\infty} t^{-1} \log \|T(t)\|.$$

Then, $e^{\eta t} \leq \|T(t)\|$ for all $t \geq 0$. So, by the definition of $\omega_0(T)$ we deduce that $\eta \leq \omega_0(T)$. Take now $\omega > \eta$. Then there is a $\tau > 0$ such that

$$\frac{\log \|T(t)\|}{t} \leq \omega, \quad \text{for all } t \geq \tau.$$

Hence, $\|T(t)\| \leq e^{\omega t}$ for all $t \geq \tau$. Since $t \mapsto T(t)$ is bounded on $[0, \tau]$, it follows that $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$ and some constant $M \geq 1$. This implies that $\omega_0(T) \geq \eta$ and therefore $\omega_0(T) = \eta$. □

In the same way as in finite dimensions (cf. Definition 4.3.1) we define the *spectral bound* of A by

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

Motivated by the finite dimensional case, see Corollary 4.3.3, one may ask whether for a generator A of a C_0-semigroup $(T(t))_{t\geq 0}$ on X we have: $\omega_0(T) = s(A)$. We will, however, see later on that this equality is in general not true even for positive C_0-semigroups on a Banach lattice E.

We introduce now different stability concepts.

Definition 13.1.2. A C_0-semigroup $(T(t))_{t\geq 0}$ with generator A on X is called

(i) uniformly exponentially stable if $\omega_0(A) < 0$;

(ii) strongly stable if $\lim_{t\to\infty} \|T(t)f\| = 0$ for every $f \in X$.

It is clear that (i) implies (ii). The converse in general does not hold, as the following example shows.

Example 13.1.3. Let us consider the shift semigroup on $L^p(\mathbb{R})$, $1 \leq p < \infty$ defined by

$$(T(t)f)(s) := f(t+s), \quad t \geq 0, \ \text{a.e. } s \in \mathbb{R}.$$

Then $(T(t))_{t\geq 0}$ is a C_0-semigroup on $L^p(\mathbb{R})$ satisfying $\lim_{t\to\infty} T(t)f = 0$ for any $f \in L^p(\mathbb{R})$, since

$$\|T(t)f\|_p^p = \int_t^\infty |f(s)|^p \, ds, \quad t \geq 0, \ f \in L^p(\mathbb{R}).$$

On the other hand, by considering the function

$$f_t(s) = \mathbb{1}_{(t,t+1)}(s) = \begin{cases} 1, & \text{if } s \in (t, t+1), \\ 0, & \text{othewise,} \end{cases}$$

we have $\|T(t)f_t\|_p = 1$ and hence $\|T(t)\| = 1$.

The definition of the growth bound and Proposition 13.1.1 yield the following characterization of uniform exponential stability. Compare with Theorem 4.4.3 in the finite dimensional case.

Proposition 13.1.4. *For the generator A of a C_0-semigroup $(T(t))_{t\geq 0}$ on X, the following assertions are equivalent.*

(i) $\omega_0(T) < 0$, *i.e.,* $(T(t))_{t\geq 0}$ *is uniformly exponentially stable.*

(ii) $\lim_{t\to\infty} \|T(t)\| = 0$.

(iii) $\|T(t_0)\| < 1$ *for some* $t_0 > 0$.

(iv) $r(T(t_1)) < 1$ *for some* $t_1 > 0$.

Proof. The implications $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv)$ are straightforward.
$(iv) \Rightarrow (iii)$: Since $r(T(t_1)) = \lim_{k\to\infty} \|T(t_1 k)\|^{\frac{1}{k}} < 1$, it follows that there is $k_0 \in \mathbb{N}$ with $\|T(k_0 t_1)\| < 1$.
$(iii) \Rightarrow (i)$: For $\alpha := \|T(t_0)\| < 1$, $M := \sup_{0\leq s\leq t_0} \|T(s)\| \geq 1$ and $t = kt_0 + s$ with $s \in [0, t_0)$, we have

$$\|T(t)\| \leq \|T(s)\|\|T(t_0 k)\| \leq M\alpha^k = Me^{k\log\alpha}.$$

If we set $\varepsilon := \frac{-\log\alpha}{t_0} > 0$ (because $\alpha < 1$), then

$$\|T(t)\| \leq Me^{k\log\alpha} \leq Me^{-\varepsilon t}.$$

□

It is clear that if $\omega_0(T) < 0$, then there are constants $\varepsilon > 0$ and $M \geq 1$ such that

$$\| T(t) \| \leq Me^{-\varepsilon t}, \quad t \geq 0.$$

Hence, for every $p \in [1, \infty)$, $\int_0^\infty \|T(t)f\|^p dt < \infty$ for all $f \in X$. The following result due to Datko shows that the converse is also true.

Theorem 13.1.5. *A C_0-semigroup $(T(t))_{t\geq 0}$ on X is uniformly exponentially stable if and only if for some (and hence for every) $p \in [1, \infty)$,*

$$\int_0^\infty \|T(t)f\|^p dt < \infty$$

for all $f \in X$.

Proof. We only have to prove the converse. By Proposition 13.1.4 it suffices to prove that $\lim_{t\to\infty} \|T(t)\| = 0$. We note first that the set

$$\{\mathcal{T}_n f : n \in \mathbb{N}\} \subset L^p(\mathbb{R}_+, X)$$

is bounded for each $f \in X$, where $\mathcal{T}_n f := \mathbb{1}_{[0,n]}(\cdot)T(\cdot)f$. Thus, by the uniform boundedness principle (see Theorem B.1.1),

$$\int_0^t \|T(s)f\|^p \, ds \leq C^p \|f\|^p, \qquad \text{for all } f \in X,\ t \geq 0.$$

Since there are $M, \omega \in \mathbb{R}_+$ with $\|T(t)\| \leq Me^{\omega t}$, $t \geq 0$, we obtain

$$\begin{aligned}\frac{1-e^{-p\omega t}}{p\omega}\|T(t)f\|^p &= \int_0^t e^{-p\omega s}\|T(s)T(t-s)f\|^p\,ds \\ &\leq M^p \int_0^t \|T(t-s)f\|^p\,ds \\ &\leq M^p C^p \|f\|^p\end{aligned}$$

for all $f \in X$ and $t \geq 0$. Hence,

$$\|T(t)f\|^p \leq \frac{p\omega}{1-e^{-p\omega}} M^p C^p \|f\|^p \text{ for } f \in X \text{ and } t \geq 1.$$

Thus, there exists a constant $L > 0$ with $\|T(t)\| \leq L$ for all $t \geq 0$. Therefore,

$$\begin{aligned}t\|T(t)f\|^p &= \int_0^t \|T(t-s)T(s)f\|^p\,ds \\ &\leq L^p \int_0^t \|T(s)f\|^p\,ds \\ &\leq L^p C^p \|f\|^p\end{aligned}$$

for all $f \in X$ and $t \geq 0$. Thus,

$$\|T(t)\| \leq LCt^{-\frac{1}{p}}, \quad t > 0,$$

which implies $\lim_{t\to\infty} \|T(t)\| = 0$. □

In Hilbert spaces uniform exponential stability of a semigroup can be characterized nicely in terms of its generator as the following Gearhart-Prüß theorem shows.

Theorem 13.1.6. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup on a Hilbert space H with generator A. Then $(T(t))_{t\geq 0}$ is uniformly exponentially stable if and only if*

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 0\} \subseteq \rho(A) \quad \textit{and} \quad M := \sup_{\operatorname{Re}\lambda > 0} \|R(\lambda, A)\| < \infty.$$

Proof. Assume that $\omega_0(T) < 0$. Then $\int_0^\infty e^{-\lambda t}T(t)\,dt$ exists for all $\operatorname{Re}\lambda > 0$. So by Proposition 12.2.7, $\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 0\} \subseteq \rho(A)$ and $R(\lambda, A) = \int_0^\infty e^{-\lambda t}T(t)dt$. Therefore

$$\sup_{\operatorname{Re}\lambda > 0} \|R(\lambda, A)\| < \infty.$$

We now prove the converse. From Corollary 12.2.5 we know that

$$\operatorname{dist}(\lambda, \sigma(A)) \geq \frac{1}{\|R(\lambda, A)\|} \geq M^{-1}, \quad \text{for all } \operatorname{Re}\lambda > 0.$$

Thus, $i\mathbb{R} \subseteq \rho(A)$ and $\sup_{\mathrm{Re}\,\lambda\geq 0}\|R(\lambda,A)\| < \infty$. Let $\omega > |\omega_0(T)|+1$ and consider the C_0-semigroup $(T_{-\omega}(t))_{t\geq 0}$ defined by $T_{-\omega}(t) := e^{-\omega t}T(t)$, $t \geq 0$. By Proposition 12.2.7 we have

$$\begin{aligned} R(\omega+is,A)f &= R(is,A-\omega)f \\ &= \int_0^\infty e^{-ist}T_{-\omega}(t)f\ dt \\ &= \mathcal{F}(T_{-\omega}(\cdot)f)(s), \end{aligned}$$

where

$$\mathcal{F}f(s) := \int_{-\infty}^\infty e^{-ist}f(t)\ dt$$

denotes the Fourier transform from $L^2(\mathbb{R},H)$ into $L^2(\mathbb{R},H)$. Here we extend $T_{-\omega}(\cdot)$ to $\mathbb{R}$ by taking $T_{-\omega}(t)=0$ for $t<0$. Since $T_{-\omega}(\cdot)$ is uniformly exponentially stable, we obtain $T_{-\omega}(\cdot)x \in L^2(\mathbb{R},H)$. Then by applying Plancherel's Theorem B.1.14, we obtain

$$\int_{-\infty}^\infty \|R(\omega+is,A)f\|^2\ ds = 2\pi\int_0^\infty \|T_{-\omega}(t)f\|^2\ dt \leq L\|f\|^2$$

for some constant $L>0$ and all $f \in H$. The resolvent identity gives

$$R(is,A) = R(\omega+is,A) + \omega R(is,A)R(\omega+is,A), \quad \text{for all } s\in\mathbb{R}.$$

Hence, $\|R(is,A)f\| \leq (1+M\omega)\|R(\omega+is,A)f\|$ for $s\in\mathbb{R}$ and $f\in H$. This implies

$$\begin{aligned} \int_{-\infty}^\infty \|R(is,A)f\|^2\,ds &\leq (1+\omega M)^2\int_{-\infty}^\infty \|R(\omega+is,A)f\|^2\,ds \\ &\leq (1+\omega M)^2 L\|f\|^2. \end{aligned}$$

On the other hand, by the inverse Laplace transform formula, Theorem B.1.15, we know that

$$T(t)f = \frac{1}{2i\pi t}\lim_{n\to\infty}\int_{\omega-in}^{\omega+in} e^{\lambda t}R(\lambda,A)^2 f\,d\lambda, \quad t\geq 0, f\in D(A^2),$$

where $D(A^2) := \{f\in D(A) : Ax \in D(A)\}$. Then, by Cauchy's integral theorem,

$$\begin{aligned} (tT(t)f|g) &= \frac{1}{2i\pi}\int_{-\infty}^\infty e^{(\omega+is)t}\big(R(\omega+is,A)^2f|g\big)ds \\ &= \frac{1}{2i\pi}\int_{-\infty}^\infty e^{ist}\big(R(is,A)^2f|g\big)ds \\ &= \frac{1}{2i\pi}\int_{-\infty}^\infty e^{ist}\big(R(is,A)f|R(-is,A^*)g\big)ds \end{aligned}$$

for all $f\in D(A^2)$ and $g\in H$. As above one can see that

$$\int_{-\infty}^\infty \|R(is,A^*)g\|^2\,ds \leq (1+M\omega)^2L\|g\|^2, \quad g\in H.$$

Here we denote by $(\cdot|\cdot)$ the scalar product[1] in H. By applying the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}|(tT(t)f|g)| &\leq \frac{1}{2\pi}\left(\int_{-\infty}^{\infty}\|R(is,A)f\|^2\,ds\right)^{\frac{1}{2}}\left(\int_{-\infty}^{\infty}\|R(is,A^*)g\|^2\,ds\right)^{\frac{1}{2}}\\ &\leq \frac{(1+M\omega)^2L}{2\pi}\|f\|\|g\|\end{aligned}$$

for all $f \in D(A^2)$ and $g \in H$. Since $\overline{D(A^2)} = H$ (see Exercise 13.4.2), it follows that

$$\begin{aligned}\|tT(t)\| &= \sup\left\{|(tT(t)f|g)|; x,y \in D(A^2), \|f\| = \|g\| = 1\right\}\\ &\leq \frac{(1+M\omega)^2}{2\pi}L.\end{aligned}$$

Hence, $\lim_{t\to\infty}\|T(t)\| = 0$ and therefore, $\omega_0(T) < 0$. □

13.2 Spectral bounds for positive semigroups

In this section we characterize the spectral bound $s(A)$ of the generator A of a positive C_0−semigroup $(T(t))_{t\geq 0}$ on a complex Banach lattice E. We will see that $s(A)$ is always contained in $\sigma(A)$ provided that $\sigma(A) \neq \emptyset$. Compare this result with Perron's Theorem 5.1.6 for positive matrices!

To this end we first improve the integral representation formula for the resolvent for the case of positive semigroups (see Proposition 12.2.7).

Theorem 13.2.1. *Let A be the generator of a positive C_0−semigroup $(T(t))_{t\geq 0}$ on E. For $\operatorname{Re}\lambda > s(A)$ we have*

$$R(\lambda, A)f = \lim_{t\to\infty}\int_0^t e^{-\lambda s}T(s)f\,ds, \quad f \in E.$$

Moreover, $\int_0^t e^{-\lambda s}T(s)\,ds$ converges to $R(\lambda, A)$ with respect to the operator norm as $t \to \infty$.

Proof. Let $\lambda_0 > \omega_0(T)$ be fixed. We recall from Proposition 12.2.7 that $R(\lambda_0, A)f = \int_0^\infty e^{-\lambda_0 t}T(t)f\,dt$, and

$$R(\lambda_0, A)^{n+1}x = \frac{1}{n!}\int_0^\infty t^n e^{-\lambda_0 t}T(t)f\,dt$$

[1] We implemented here the suggestion of U. Groh from the discussion board to follow Bourbaki and denote scalar products by $(\cdot|\cdot)$.

for $n \in \mathbb{N}$ and $f \in E$. Let $\mu \in (s(A), \lambda_0)$, $f \in E_+$ and $f^* \in E^*_+$. By Corollary 12.2.5 one has $\frac{1}{\lambda_0 - \mu} > r(R(\lambda_0, A))$ and hence,

$$\begin{aligned}
\langle R(\mu, A)f, f^*\rangle &= \sum_{n=0}^{\infty} (\lambda_0 - \mu)^n \langle R(\lambda_0, A)^{n+1} f, f^*\rangle \\
&= \sum_{n=0}^{\infty} \int_0^{\infty} \frac{1}{n!} [(\lambda_0 - \mu)s]^n e^{-\lambda_0 s} \langle T(s)f, f^*\rangle \, ds \\
&= \int_0^{\infty} \left(\sum_{n=0}^{\infty} \frac{1}{n!} [(\lambda_0 - \mu)s]^n \right) e^{-\lambda_0 s} \langle T(s)f, f^*\rangle \, ds \\
&= \int_0^{\infty} e^{(\lambda_0 - \mu)s} e^{-\lambda_0 s} \langle T(s)f, f^*\rangle \, ds \\
&= \int_0^{\infty} e^{-\mu s} \langle T(s)f, f^*\rangle \, ds \\
&= \lim_{t \to \infty} \langle \int_0^t e^{-\mu s} T(s)f \, ds, f^*\rangle,
\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing[2] of E and E^*. Hence, for $f \in E_+$, $\left(\int_0^t e^{-\mu s} T(s)f \, ds \right)$ converges weakly to $R(\mu, A)f$ as $t \to \infty$. Since $f \in E_+$, and by the positivity of the semigroup $(T(t))_{t \geq 0}$, it follows that $\left(\int_0^t e^{-\mu s} T(s)f \, ds \right)_{t \geq 0}$ is monotone increasing and so, by Proposition 9.3.4, we have strong convergence. Thus,

$$\lim_{t \to \infty} \int_0^t e^{-\mu s} T(s)f \, ds = R(\mu, A)f$$

for all $f \in E_+$ and hence for all $f \in E$. If $\lambda = \mu + i\nu$ with $\mu, \nu \in \mathbb{R}$ and $\mu > s(A)$, then for any $f \in E$ and $f^* \in E^*$, we have

$$\left| \langle \int_r^t e^{-\lambda s} T(s)f \, ds, f^* \rangle \right| \leq \int_r^t e^{-\mu s} \langle T(s)|f|, |f^*|\rangle \, ds.$$

Hence,

$$\left\| \int_r^t e^{-\lambda s} T(s)f \, ds \right\| \leq \left\| \int_r^t e^{-\mu s} T(s)|f| \, ds \right\|,$$

which implies that

$$\lim_{t \to \infty} \int_0^t e^{-\lambda s} T(s)f \, ds \text{ exists for all } f \in E.$$

Then, using Lemma 12.2.6 one sees that

$$\lambda \in \rho(A) \text{ and } R(\lambda, A)f = \int_0^{\infty} e^{-\lambda t} T(t)f \, dt \quad \text{for all } f \in E.$$

[2] In other words, for a vector $f \in E$ and a linear functional $g^* \in E^*$, we have $\langle f, g^*\rangle = g^*(f)$.

It remains to prove that $\left(\int_0^\infty e^{-\lambda s}T(s)\,ds\right)$ converges in the operator norm as $t \to \infty$. We fix $\mu \in (s(A), \operatorname{Re}\lambda)$. As we have seen above, the function

$$\psi_{f,f^*} : s \mapsto e^{-\mu s}\langle T(s)f, f^*\rangle \text{ belongs to } L^1(\mathbb{R}_+) \quad \text{for all } f \in E, f^* \in E^*.$$

It follows from the closed graph theorem that the bilinear form

$$b : E \times E^* \to L^1(\mathbb{R}_+);\ (f, f^*) \mapsto \psi_{f,f^*}$$

is separately continuous and hence continuous. Thus, there exists a constant $M \geq 0$ such that

$$\int_0^\infty e^{-\mu s}|\langle T(s)f, f^*\rangle|\,ds \leq M\|f\|\|f^*\|, \quad f \in E, f^* \in E^*.$$

For $0 \leq t < r$ and $\varepsilon := \operatorname{Re}\lambda - \mu$ we have

$$\begin{aligned}\left|\int_t^r e^{-\lambda s}\langle T(s)f, f^*\rangle\,ds\right| &\leq \int_t^r e^{-(\operatorname{Re}\lambda-\mu)s}e^{-\mu s}|\langle T(s)f, f^*\rangle|\,ds \\ &\leq e^{-\varepsilon t}\int_t^r e^{-\mu s}|\langle T(s)f, f^*\rangle|\,ds \\ &\leq e^{-\varepsilon t}M\|f\|\|f^*\|.\end{aligned}$$

Hence, $\left\|\int_t^r e^{-\lambda s}T(s)\,ds\right\| \leq Me^{-\varepsilon t}$ and this implies that the function $t \mapsto \left(\int_0^t e^{-\lambda s}T(s)\,ds\right)$ satisfies the Cauchy condition in $\mathcal{L}(E)$ as $t \to \infty$. Thus, $\left(\int_0^\infty e^{-\lambda s}T(s)\,ds\right)$ converges in the operator norm. □

As an immediate consequence we obtain the following corollary.

Corollary 13.2.2. *Let A be the generator of a positive C_0-semigroup $(T(t))_{t\geq 0}$ on E. If $\operatorname{Re}\lambda > s(A)$, then*

$$|R(\lambda, A)f| \leq R(\operatorname{Re}\lambda, A)|f| \quad \textit{for all } f \in E.$$

Another important corollary is the following previously announced result.

Corollary 13.2.3. *If A is the generator of a positive C_0-semigroup $(T(t))_{t\geq 0}$ on E, and $s(A) > -\infty$, then $s(A) \in \sigma(A)$.*

Proof. Assume that $s(A) \in \rho(A)$. It follows from Corollary 13.2.2 that

$$|R(\lambda, A)f| \leq R(\operatorname{Re}\lambda, A)|f| \leq R(s(A), A)|f| \quad \text{for all } \operatorname{Re}\lambda > s(A),\ f \in E.$$

Hence the set $\{R(\lambda, A) : \operatorname{Re}\lambda > s(A)\}$ is uniformly bounded in $\mathcal{L}(E)$. Let

$$M := \sup_{\operatorname{Re}\lambda > s(A)} \|R(\lambda, A)\|.$$

Since

$$\|R(\lambda, A)\| \geq \frac{1}{\operatorname{dist}(\lambda, \sigma(A))}$$

for $\lambda \in \rho(A)$ (see Corollary 12.2.5), it follows that

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda = s(A)\} \subseteq \rho(A) \quad \text{and} \quad \|R(\lambda, A)\| \leq M \text{ for all } \lambda \text{ with } \operatorname{Re}\lambda = s(A).$$

Thus, since $\rho(A)$ is open,

$$\{\lambda \in \mathbb{C} : |\operatorname{Re}\lambda - s(A)| < M^{-1}\} \subseteq \rho(A).$$

This contradicts the definition of $s(A)$. □

We also obtain a relation between $s(A)$ and the positivity of the resolvent.

Corollary 13.2.4. *Suppose that A generates a positive C_0-semigroup $(T(t))_{t\geq 0}$ on E and $\lambda_0 \in \rho(A)$. Then the following assertions hold.*

(i) $R(\lambda_0, A)$ is positive if and only if $\lambda_0 > s(A)$.

(ii) If $\lambda > s(A)$, then $r(R(\lambda, A)) = \frac{1}{\lambda - s(A)}$.

Proof. (ii): This is a simple consequence of Corollary 13.2.3 and Proposition 12.2.4.
(i): Assume first that $R(\lambda_0, A) \geq 0$. So, one has

$$R(\lambda_0, A)E_{\mathbb{R}} \subset E_{\mathbb{R}},$$

where

$$E_{\mathbb{R}} := \{\operatorname{Re} f : f \in E\}.$$

Let $f \in E_{\mathbb{R}} \setminus \{0\}$. Then $g = R(\lambda_0, A)f \in E_{\mathbb{R}}$ and $\lambda_0 g - Ag = f$. So, $\lambda_0 g \in E_{\mathbb{R}}$ and hence $\lambda_0 \in \mathbb{R}$. On the other hand, Theorem 13.2.1 implies that $R(\lambda, A) \geq 0$ for all $\lambda > \max(\lambda_0, s(A))$ and hence

$$\begin{aligned} R(\lambda_0, A) &= R(\lambda, A) + (\lambda - \lambda_0)R(\lambda, A)R(\lambda_0, A) \\ &\geq R(\lambda, A) \geq 0 \end{aligned}$$

for all $\lambda > \max(\lambda_0, s(A))$. Therefore,

$$(\lambda - s(A))^{-1} = r(R(\lambda, A)) \leq \|R(\lambda, A)\| \leq \|R(\lambda_0, A)\|$$

for all $\lambda > \max(\lambda_0, s(A))$. But this is only true if $\lambda_0 > s(A)$.
The converse follows from Theorem 13.2.1. □

Remark 13.2.5. (a) It can be proved that the statement in Corollary 13.2.3 remains valid if one assumes only that A is a *resolvent positive* operator. So, one obtains as a corollary that if $R(\lambda_0, A) \geq 0$ for some $\lambda_0 \in \rho(A)$ then $\lambda_0 \in \mathbb{R}$ and $\lambda_0 > s(A)$, whenever A is a resolvent positive operator. Here we say that A is a resolvent positive operator if there is $\mu \in \mathbb{R}$ such that $(\mu, \infty) \subset \rho(A)$ and $R(\lambda, A) \geq 0$ for all $\lambda > \mu$, see Exercise 13.4.7. Recall that if A generates a positive C_0-semigroup, then A is a resolvent positive operator, see Corollary 12.2.8.

(b) As an immediate consequence of Corollary 13.2.4 we obtain

$$s(A) = \inf\{\lambda \in \rho(A) : R(\lambda, A) \geq 0\}$$

for the generator A of a positive C_0-semigroup on a Banach lattice E.

(c) If $E = \mathrm{C}(K)$, where K is a compact set, then $s(A) > -\infty$.
In fact, we know from Lemma 12.3.2 that $\lim_{\lambda\to\infty} \lambda R(\lambda, A)f = f$ for all $f \in E$. In particular we can find $\lambda_0 \in \mathbb{R}$ sufficiently large such that

$$\lambda_0 R(\lambda_0, A)\mathbb{1} \geq \frac{1}{2}\mathbb{1}.$$

Since $R(\lambda_0, A) \geq 0$, it follows that

$$R(\lambda_0, A)^n \mathbb{1} \geq \frac{1}{2\lambda_0}\mathbb{1} \quad \text{for all } n \in \mathbb{N}.$$

Thus,

$$r(R(\lambda_0, A)) = \lim_{n\to\infty} \|R(\lambda_0, A)^n\|^{\frac{1}{n}} \geq \frac{1}{2\lambda_0} > 0$$

and hence $\sigma(A) \neq \emptyset$.

The spectrum of a generator of a positive C_0-semigroup can be empty as the following examples show.

Example 13.2.6. (a) On $E := C_0[0, 1) := \{f \in C[0, 1] : f(1) = 0\}$ we consider the nilpotent C_0-semigroup $(T(t))_{t\geq 0}$ given by

$$(T(t)f)(x) = \begin{cases} f(x+t), & \text{if } x+t < 1, \\ 0, & \text{if } x+t \geq 1, \end{cases}$$

for $t \geq 0$, $x \in [0, 1]$ and $f \in E$. Then, $T(t) = 0$ for $t \geq 1$ and hence $\sigma(T(t)) = \{0\}$. So by the spectral inclusion theorem (see Exercise 13.4.3), $\sigma(A) = \emptyset$.

(b) Let $E := C_0[0, \infty) := \{f \in C(\mathbb{R}_+) : \lim_{t\to\infty} f(t) = 0\}$. On E we define the C_0-semigroup $(T(t))_{t\geq 0}$ by

$$(T(t)f)(x) := e^{-\frac{t^2}{2} - xt} f(x+t), \quad x, t \geq 0 \text{ and } f \in E.$$

Then, one can see that the generator A of $(T(t))_{t\geq 0}$ on E is given by

$$\begin{aligned} (Af)(x) &= f'(x) - xf(x),\ x \geq 0, \text{ and} \\ f \in D(A) &= \{f \in E : f \in C^1(\mathbb{R}_+) \text{ and } Af \in E\}. \end{aligned}$$

One proves that $\sigma(A) = \emptyset$, see Exercise 13.4.4.

However, for generators of positive C_0-groups the spectrum is always nonempty. This is given by the following corollary.

Corollary 13.2.7. *If A generates a positive C_0-group on a Banach lattice E, then*

$$\sigma(A) \neq \emptyset.$$

Proof. Assume that $\sigma(A) = \emptyset$. By Theorem 13.2.1 we have $R(\lambda, A) \geq 0$ for all $\lambda \in \mathbb{R}$. Again, one can apply the same theorem to $-A$ and obtains $R(\lambda, -A) \geq 0$ for all $\lambda \in \mathbb{R}$. But $R(\lambda, -A) = -R(-\lambda, A) \leq 0$ for all $\lambda \in \mathbb{R}$, and hence, $R(\lambda, -A) = 0$ for all $\lambda \in \mathbb{R}$. This contradicts the fact that $E \neq \{0\}$. □

We end this chapter by proving the existence of positive eigenfunction. We will see in Chapter 14 that such eigenfunction is unique (up to a scalar factor) if A generates an irreducible positive C_0-semigroup. Compare with Theorem 5.1.6 in the finite dimensional case.

Theorem 13.2.8. *Let A be a resolvent positive operator on E with compact resolvent such that $s(A) > -\infty$. Then there exists $0 \lneqq u \in D(A)$ such that $Au = s(A)u$.*

Proof. Replacing A with $A - s(A)$ one can assume without loss of generality that $s(A) = 0$. Since, by Exercise 13.4.7,

$$s(A) \in \sigma(A),$$

it follows from Corollary 12.2.5 that $\|R(\lambda_n, A)\| \to \infty$ as $n \to \infty$ for some $\lambda_n > 0$ with $\lim_{n\to\infty} \lambda_n = s(A) = 0$. By the uniform boundedness principle (see Theorem B.1.1) there is $f \in E$ such that

$$\lim_{n\to\infty} \|R(\lambda_n, A)f\| = \infty.$$

Since $|R(\lambda_n, A)f| \leq R(\lambda_n, A)|f|$ one can assume that $f \gneqq 0$. Take

$$u_n = \frac{R(\lambda_n, A)f}{\|R(\lambda_n, A)f\|}.$$

Then $0 \leq u_n \in D(A)$ with $\|u_n\| = 1$ and

$$\lim_{n\to\infty} \lambda_n u_n - Au_n = \lim_{n\to\infty} \|R(\lambda_n, A)f\|^{-1} f = 0.$$

Thus, (u_n) is bounded in the graph norm. Using the compactness of the resolvent of A we have that the embedding $D(A) \hookrightarrow E$ is compact and so we can assume that $\lim_{n\to\infty} u_n = u$ exists in E, considering a subsequence otherwise. Then $\|u\| = 1$, $u \geq 0$, and by the closedness of A we obtain $u \in D(A)$ and $Au = 0$. □

13.3 Notes and Remarks

The spectral theory and stability theory of semigroups is a broad subject and well documented in the literature. We refer to the monographs by Engel and Nagel [3], [4], by van Neerven [7], or Arendt, Batty, Hieber and Neubrander [1].

Datko's Theorem 13.1.5 appeared in [2]. The Gearhart-PrüßTheorem 13.1.6 was first proved by L. Gearhart [5] for contraction semigroups in Hilbert spaces using the functional calculus developed by Szőkefalvi-Nagy and Foiaş. The general case was then proved by J. Prüß [8]. There were many other proofs in the literature shortly thereafter available.

Theorem 13.2.8 is a variation of the famous Krein-Rutman theorem, see [6].

13.4 Exercises

1. Let $\zeta : [0,\infty) \to \mathbb{R}$ be bounded on compact intervals and subadditive, i.e., $\zeta(t+s) \leq \zeta(t)\zeta(s)$ for all t, $s \geq 0$. Prove that
$$\inf_{t>0} \frac{\zeta(t)}{t} = \lim_{t\to\infty} \frac{\zeta(t)}{t}.$$

2. Let A be the generator of a C_0-semigroup on a Banach space X. Show that $D(A^n)$ is dense in X for any $n \in \mathbb{N}$.

3. Let A be the generator of a C_0-semigroup in a Banach space X. Prove the spectral inclusion
$$e^{t\sigma(A)} \subset \sigma(T(t)), \qquad \text{for all } \ t \geq 0.$$

4. On $E := C_0[0,\infty)$ let us consider the family of operators
$$(T(t)f)(x) := e^{-\frac{t^2}{2}-xt} f(x+t), \quad x,t \geq 0 \text{ and } f \in E.$$
   (a) Show that $(T(t))_{t\geq 0}$ define a positive C_0-semigroup on E.
   (b) Prove that its generator is given by
$$\begin{aligned} (Af)(x) &= f'(x) - xf(x),\, x \geq 0, \text{ and} \\ f \in D(A) &= \{f \in E : f \in C^1(\mathbb{R}_+) \,\text{and}\, Af \in E\}. \end{aligned}$$
   (b) Prove that $\sigma(A) = \emptyset$.

5. Let $1 < p < q < \infty$ and $E := L^p[1,\infty) \cap L^q[1,\infty)$ the Banach lattice endowed with the norm
$$\|f\| := \|f\|_p + \|f\|_q, \quad f \in E.$$
Define the family of operators
$$T(t)f(s) = f(se^t) \quad s \geq 1,\, t \geq 0.$$

(a) Show that $(T(t))_{t\geq 0}$ define a positive C_0-semigroup on E with generator

$$\begin{aligned} (Af)(s) &= sf'(s),\ s \geq 1, \text{ and} \\ f \in D(A) &= \{f \in E : f \text{ absolutely continuous and } Af \in E\}. \end{aligned}$$

(b) Prove that $s(A) = -\frac{1}{p} < -\frac{1}{q} = \omega_0(T)$.

6. Let A be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach space X. Let us denote by $(T(t)^*)_{t\geq 0}$ the adjoint semigroup on E^*.

(a) Prove that the restriction $S(t) = T(t)^*_{|\overline{D(A^*)}}$ defines a C_0-semigroup on $\overline{D(A^*)}$.

(b) Prove that the generator B of $(S(t))_{t\geq 0}$ is given by

$$D(B) = \{f^* \in D(A^*) : A^*f^* \in \overline{D(A^*)}\} \text{ and } Bf^* = A^*f^* \text{ for } f^* \in D(B).$$

(c) Prove that $\sigma(B) = \sigma(A^*)$.

7. Let A be a resolvent positive operator on a Banach lattice E with $s(A) > -\infty$. Show that $s(A) \in \sigma(A)$.

Bibliography

[1] W. Arendt, C.J.K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser Verlag, 2011.

[2] R. Datko, *Extending a theorem of A.M. Liapunov to Hilbert space*, J. Math. Anal. Appl. **32** (1970), 610–616.

[3] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Math., vol. 194, Springer-Verlag, 2000.

[4] ______, *A Short Course on Operator Semigroups*, Springer-Verlag, 2006.

[5] L. Gearhart, *Spectral theory for contraction semigroups on Hilbert spaces*, Trans. Amer. Math. Soc. **236** (1978), 385–394.

[6] M. G. Kreĭn and M. A. Rutman, *Linear operators leaving invariant a cone in a Banach space*, Uspehi Matem. Nauk (N. S.) **3** (1948), no. 1(23), 3–95.

[7] J.M.A.A. van Neerven, *The Asymptotic Behaviour of Semigroups of Linear Operators*, Oper. Theory Adv. Appl., vol. 88, Birkhäuser Verlag, 1996.

[8] J. Prüss, *On the spectrum of C_0-semigroups*, Trans. Amer. Math. Soc. **284** (1984), 847–857.

Lecture 14

Spectral Theory for Positive Semigroups II

In this lecture we continue our investigations of spectral properties of positive C_0-semigroups on Banach lattices and show how the Perron-Frobenius theory can be generalized to the infinite dimensional setting.

14.1 The problem $\omega_0(T) = s(A)$ for positive semigroups

We proved in the finite dimensional setting that $s(A) = \omega_0(T)$, see Corollary 4.3.3. On the other hand, it follows from Proposition 12.2.7 that $s(A) \leq \omega_0(T)$ holds, whenever A generates a C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach space X. But, in the infinite dimensional case, $s(A) = \omega_0(T)$ is in general not true, even for positive C_0-semigroups.

Example 14.1.1. Consider the Banach lattice $E := C_0(\mathbb{R}_+) \cap L^1(\mathbb{R}_+, e^s ds)$ endowed with the norm

$$\|f\| := \sup_{s\geq 0} |f(s)| + \int_0^\infty |f(s)e^s\, ds =: \|f\|_\infty + \|f\|_1, \quad f \in E.$$

On E we define the translation semigroup

$$(T(t)f)(s) = f(s+t), \quad t, s \geq 0.$$

Then $(T(t))_{t\geq 0}$ is a positive C_0-semigroup. Its generator is given by

$$Af = f' \,\text{for}\, f \in D(A) = \{f \in E : f \in C^1(\mathbb{R}_+) \,\text{and}\, f' \in E\}.$$

Note that $\|T(t)\| = 1$ for all $t \geq 0$. Then $\omega_0(T) = 0$. On the other hand, the function $\varepsilon_\lambda(s) := e^{\lambda s}$ is an eigenfunction for A associated with λ provided that $\operatorname{Re}\lambda < -1$. Hence,

$$\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda \leq -1\} \subseteq \sigma(A).$$

Moreover, for $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > -1$ one sees that

$$\|\cdot\|_1 - \lim_{N\to\infty} \int_0^N e^{-\lambda s} T(s) f\, ds + \|\cdot\|_\infty - \lim_{N\to\infty} \int_0^N e^{-\lambda s} T(s) f\, ds$$

exists, because $\|T(s)f\|_1 \leq e^{-s}\|f\|_1$ for all $s \geq 0$, and $\int_0^\infty e^s |f(s)|\, ds < \infty$. Therefore $\int_0^\infty e^{-\lambda s} T(s) f\, ds$ exists in E for all $f \in E$. Thus, by Proposition 12.2.7, $\lambda \in \rho(A)$. Therefore,

$$\sigma(A) = \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda \leq -1\}, \text{ whence } s(A) = -1.$$

An other example of such situation is given by Exercise 13.4.5.

In this section we look for sufficient conditions implying the equality $\omega_0(T) = s(A)$ when A generates a positive C_0-semigroup $(T(t))_{t\geq 0}$ on E.

By applying Theorem 13.1.6 and Theorem 13.1.5 we obtain $\omega_0(T) = s(A)$ in the following cases.

Theorem 14.1.2. *Let A be the generator of a positive C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach lattice E. Then $\omega_0(A) = s(A)$ holds in the following cases.*

(i) E is a Hilbert space.

(ii) E is an AL-space, i.e. the norm satisfies: $\|f+g\| = \|f\| + \|g\|$ for all $f, g \in E_+$.

(iii) $E = C_0(\Omega)$ or $E = C(K)$, where Ω is locally compact and K is compact set.

Proof. (i): Fix $\mu > s(A)$. It follows by Corollary 13.2.2 that $\Lambda := \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 0\} \subseteq \rho(A-\mu)$ and

$$\|R(\lambda, A-\mu)\| \leq \|R(\operatorname{Re}\lambda, A-\mu)\| \leq \|R(\mu, A)\| \quad \text{for all } \lambda \in \Lambda.$$

So, by Theorem 13.1.6, we have $\omega_0(A) - \mu < 0$ and hence,

$$\omega_0(A) \leq s(A).$$

(ii): For $\lambda > s(A)$ and $f \in E_+$ we obtain from Theorem 13.2.1 that

$$\|R(\lambda, A)f\| = \left\| \int_0^\infty e^{-\lambda s} T(s) f\, ds \right\| = \int_0^\infty e^{-\lambda s} \|T(s) f\|\, ds,$$

where the second equality follows from the fact that the norm is additive on the positive cone. Hence,

$$\int_0^\infty \|(e^{-\lambda s} T(s)) f\|\, ds < \infty \quad \text{for all } f \in E.$$

So, by Theorem 13.1.5, we have $\omega_0(A) - \lambda < 0$ and thus

$$\omega_0(A) \leq s(A).$$

(iii): Is is easy to see that $\|f \vee g\| = \|f\| \vee \|g\|$ for all $f, g \in E_+$. Then, for $\gamma, \nu \in E_+^*$, we have

$$\begin{aligned} \langle f, \gamma\rangle + \langle g, \nu\rangle &\leq \langle f \vee g, \gamma \vee \nu\rangle| \\ &\leq \|\gamma + \nu\|\|f \vee g\| \\ &= \|\gamma + \nu\|(\|f\| \vee \|g\|), \quad f, g \in E_+. \end{aligned}$$

Hence, $\langle f, \gamma\rangle + \langle g, \nu\rangle \leq \|\gamma + \nu\|$ for all $f, g \in E_+$ with $\|f\| = \|g\| = 1$. It follows from the Hahn-Banach theorem that $\|\gamma\| + \|\nu\| \leq \|\gamma + \nu\|$ and hence,

$$\|\gamma\| + \|\nu\| = \|\gamma + \nu\|, \quad \gamma, \nu \in E_+.$$

This implies that E^* is an AL-space. If we set $F := \overline{D(A^*)}$, then it follows from Exercise 14.4.1 that F is a closed ideal and hence also an AL-space. On F we consider the positive C_0-semigroup $(S(t))_{t\geq 0}$ given by

$$S(t) := T(t)^*_{|F} \quad \text{for}\, t \geq 0,$$

and we denote by B its generator. Then B is the part of A^* in F, i.e.,

$$D(B) = \{\nu \in D(A^*) : A^*\nu \in F\} \text{ and } B\nu = A^*\nu \,\text{for}\, \nu \in D(B).$$

Moreover, one can show that

$$\sigma(B) = \sigma(A^*) = \sigma(A),$$

see Exercise 13.4.6.

Consequently, $s(B) = s(A)$ holds. Since B is the generator of the positive C_0-semigroup $S(\cdot)$ on the AL-space F, it follows from (ii) that $s(B) = \omega_0(B)$. Now, it suffices to prove that $\omega_0(B) = \omega_0(A)$. The inequality $\omega_0(B) \leq \omega_0(A)$ is trivial. Let $\omega > \omega_0(B)$, $f \in E$ and $\nu \in F$. Then we have

$$|\langle T(t)f, \nu\rangle| = |\langle f, S(t)\nu\rangle| \leq M\|f\|e^{\omega t}\|\nu\|$$

for $t \geq 0$ and some constant $M \geq 1$. On the other hand, since $f = \lim_{\lambda\to\infty} \lambda R(\lambda, A)f$ for all $f \in E$, we have $c := \limsup_{\lambda\to\infty} \lambda\|R(\lambda, A)\| < \infty$. Therefore,

$$\begin{aligned} |\langle T(t)f, \gamma\rangle| &= \lim_{\lambda\to\infty} |\langle \lambda R(\lambda, A)T(t)f, \gamma\rangle| \\ &= \lim_{\lambda\to\infty} |\langle T(t)f, \lambda R(\lambda, A^*)\gamma\rangle| \\ &\leq M\|f\|e^{\omega t} \limsup_{\lambda\to\infty} \lambda\|R(\lambda, A)^*\gamma\| \\ &\leq Mce^{\omega t}\|f\|\|\gamma\|, \qquad \gamma \in E^*. \end{aligned}$$

Consequently, $\|T(t)\| \leq Mce^{\omega t}$ for all $t \geq 0$ and hence $\omega_0(A) \leq \omega$ for all $\omega > \omega_0(B)$. Thus, we have shown that

$$\omega_0(B) = \omega_0(A).$$

□

14.2 Asymptotic behaviour of irreducible semigroups

In many concrete examples it can be shown that semigroups $(T(t))_{t\geq 0}$ which is not exponentially stable can possesses an *asynchronous exponential growth*. This means that there is a rank one projection P and constants $\varepsilon > 0$, $M \geq 1$ such that

$$\|e^{-s(A)t}T(t) - P\| \leq Me^{-\varepsilon t} \quad \text{for all}\, t \geq 0,$$

where A denotes the generator of $(T(t))_{t\geq 0}$.

In order to study such kind of behaviour we introduce the concept of irreducibility for positive C_0-semigroups.

Definition 14.2.1. A positive C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach lattice E with generator A is called *irreducible* if there is no $T(t)-$invariant closed ideal other than $\{0\}$ and E for all $t > 0$.

The following result gives a more concrete definition of irreducibility.

Proposition 14.2.2. *Let $(T(t))_{t\geq 0}$ be positive C_0-semigroup on a Banach lattice E with generator A. The following assertion are equivalent:*

(i) $(T(t))_{t\geq 0}$ is irreducible.

(ii) For some (and then for every) $\lambda > s(A)$, there is no $R(\lambda, A)-$invariant closed ideal except $\{0\}$ and E.

(iii) For some (and then for every) $\lambda > s(A)$, $R(\lambda, A)f$ is a quasi-interior point of E_+ for every $f \gneqq 0$.

Proof. Let us prove first that if for some $\lambda_0 > s(A)$, there is no $R(\lambda_0, A)-$invariant closed ideal except $\{0\}$ and E, then for every $\lambda > s(A)$, there is no $R(\lambda, A)-$invariant closed ideal except $\{0\}$ and E.
Let I be a closed ideal of E such that $R(\mu, A)I \subset I$ for $\mu > s(A)$. It follows from

$$0 \leq R(\lambda, A) \leq R(\mu, A), \quad \forall \lambda \geq \mu,$$

and the definition of ideals, that $R(\lambda, A)I \subset I$ for every $\lambda \geq \mu$.
On the other hand, for $\mu - \frac{1}{r(R(\mu,A))} < \lambda < \mu$, $R(\lambda, A)I \subset I$, since

$$R(\lambda, A) = R(\mu, A) \sum_{n=0}^{\infty} [(\mu - \lambda)R(\mu, A)]^n$$

and $R(\mu, A)I \subset I$. Iteration of the argument establishes that $R(\lambda, A)I \subset I$ for every $s(A) < \lambda < \mu$. This proves the above claim.
$(i) \Rightarrow (ii)$: Let $I \neq \{0\}$ be a closed ideal of E such that $R(\lambda, A)I \subset I$ for some (and then

for every) $\lambda > s(A)$. By the approximation formula (see the proof of Theorem 12.3.1 and Exercise 12.6.4)

$$T(t)f = \lim_{n\to\infty} e^{tA_n} f, \quad f \in E,$$

where $A_n = nAR(n, A) \in \mathcal{L}(E)$ the Yosida approximant, we deduce that $T(t)I \subset I$ for all $t > 0$. So, $I = E$.
$(ii) \Rightarrow (i)$: This follows from Theorem 13.2.1.
$(ii) \Rightarrow (iii)$: Let $\lambda > s(A)$, $0 \neq f \in E_+$ and consider $E_{R(\lambda,A)f} = \bigcup_{n\in\mathbb{N}}[-nR(\lambda, A)f, nR(\lambda, A)f]$ the ideal generated by $R(\lambda, A)f$. For $g \in E_{R(\lambda,A)f}$, it follows from the resolvent equation that

$$|R(\mu, A)g| \leq R(\mu, A)|g| \leq nR(\mu, A)R(\lambda, A)f \leq \frac{n}{\mu - \lambda} R(\lambda, A)f$$

for $\mu > \lambda$. Hence, $R(\mu, A)g \in E_{R(\lambda,A)f}$ for any $g \in E_{R(\lambda,A)f}$ and $\mu > \lambda$. Thus, $\{0\} \neq \overline{E_{R(\lambda,A)f}}$ is an $R(\mu, A)$-invariant closed ideal and hence $\overline{E_{R(\lambda,A)f}} = E$ which means that $R(\lambda, A)f$ is a quasi-interior point of E_+.
$(iii) \Rightarrow (ii)$: Let $\{0\} \neq I$ be an $R(\mu, A)$-invariant closed ideal for some $\mu > s(A)$. Let $0 \neq f \in E_+ \cap I$. So, for any $g \in E_{R(\mu,A)f}$ we have $|g| \leq nR(\mu, A)f$ for some $n \in \mathbb{N}$ and hence $g \in I$. This implies that $E_{R(\mu,A)f} \subset I$. Therefore, $E = \overline{E_{R(\mu,A)f}} = I$. □

Applying the characterization of closed ideals given in Example 9.3.7 (see Exercise 9.5.11) we obtain

Example 14.2.3. (a) Let $E := L^p(\Omega, \mu)$, $1 \leq p < \infty$ and $(T(t))_{t\geq 0}$ be a positive C_0-semigroup on E with generator A. Then, $(T(t))_{t\geq 0}$ is irreducible if and only if

$$0 \lneqq f \in E \Longrightarrow (R(\lambda, A)f)(s) > 0 \,\text{for a.e.}\, s \in \Omega \,\text{and some}\, \lambda > s(A).$$

(b) If $E := C_0(\Omega)$, where Ω is locally compact Hausdorff, and $(T(t))_{t\geq 0}$ a positive C_0-semigroup on E with generator A, then $(T(t))_{t\geq 0}$ is irreducible if and only if

$$0 \lneqq f \in E \Longrightarrow (R(\lambda, A)f)(s) > 0 \text{ for all}\, s \in \Omega \,\text{and some}\, \lambda > s(A).$$

We now state some consequences of irreducibility.

Proposition 14.2.4. *Assume that A is the generator of an irreducible C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach lattice E. Then the following assertions hold.*

(a) Every positive eigenvector of A is a quasi-interior point.

(b) Every positive eigenvector of A^* is strictly positive.

(c) If $\ker(s(A) - A^*)$ contains a positive element, then $\dim\ker(s(A) - A) \leq 1$.

(d) If $s(A)$ is a pole of the resolvent, then it has algebraic (and geometric) multiplicity equal to 1. The corresponding residue has the form $P_{s(A)} = u^* \otimes f$, where $f \in E$ is a positive eigenvector of A, $u^* \in E^*$ is a positive eigenvector of A^* and $\langle f, u^*\rangle = 1$.

Proof. (a) Let f be a positive eigenvector of A and $E_f := \cup_{n\in\mathbb{N}} n[-f,f]$ the ideal generated by f. If λ is such that $Af = \lambda f$, then $\lambda \in \mathbb{R}$. This follows from

$$f \geq 0 \text{ and } Af = \lim_{t\to 0^+} \frac{1}{t}(T(t)f - f).$$

Hence, $T(t)f = e^{\lambda t} f$ for $t \geq 0$. Thus, for $g \in E_f$,

$$|T(t)g| \leq T(t)|g| \leq nT(t)f = ne^{\lambda t} f, \quad t \geq 0.$$

Consequently, $T(t)E_f \subseteq E_f$ holds for all $t \geq 0$. Since $0 \neq f \in E_f$ and $(T(t))_{t\geq 0}$ is irreducible, it follows that $\overline{E_f} = E$.

(b) Let f^* be a positive eigenvector of A^* and λ its corresponding eigenvalue. By the same argument we have $\lambda \in \mathbb{R}$ and $T(t)^* f^* = e^{\lambda t} f^*$ for $t \geq 0$. Hence,

$$\langle |T(t)u|, f^*\rangle \leq \langle T(t)|u|, f^*\rangle = \langle |u|, e^{\lambda t} f^*\rangle, \quad u \in E, t \geq 0.$$

Thus, $I := \{u \in E : \langle |u|, f^*\rangle = 0\}$ is a $T(t)-$invariant closed ideal for all $t \geq 0$. Since $f^* \neq 0$ we have $I \subsetneq E$ and so by the irreducibility we obtain $I = \{0\}$. Therefore, $f^* > 0$.

(c) Let $0 \lneqq f^* \in \ker(s(A) - A^*)$. It follows from (b) that f^* is strictly positive. Assume that $\ker(s(A)-A) \neq \{0\}$. Then for $f \in \ker(s(A)-A)$ we have $e^{-s(A)t}T(t)f =: T_{-s(A)}(t)f = f$ and hence,

$$|f| = |T_{-s(A)}(t)f| \leq T_{-s(A)}(t)|f|, \quad t \geq 0.$$

Thus, for $t \geq 0$,

$$\begin{aligned} \langle |f|, f^*\rangle &\leq \langle T_{-s(A)}(t)|f|, f^*\rangle \\ &= \langle |f|, f^*\rangle. \end{aligned}$$

This implies that $\langle T_{-s(A)}(t)|f| - |f|, f^*\rangle = 0$, and since $f^* > 0$, we obtain $T_{-s(A)}(t)|f| = |f|$ for $t \geq 0$. Therefore,

$$|f| \in \ker(s(A) - A).$$

Since $(T_{-s(A)}(t)f)^+ \leq T_{-s(A)}(t)f^+$ (see Proposition 10.1.3), one can see by the same arguments as above that $f^+ \in \ker(s(A) - A)$ and $f^- \in \ker(s(A) - A)$. This implies that $F := E_{\mathbb{R}} \cap \ker(s(A) - A)$ is a real sublattice of E. For $f \in F$ we consider the ideal E_{f^+} (resp. E_{f^-}) generated by f^+ (resp. f^-). Then, E_{f^+} and E_{f^-} are $T_{-s(A)}(t)-$invariant for all $t \geq 0$. Since E_{f^+} and E_{f^-} are orthogonal, see Proposition 9.2.3, it follows from the irreducibility of $(T_{-s(A)}(t))_{t\geq 0}$ that $f^+ = 0$ or $f^- = 0$. Consequently, F is totally ordered. So by Lemma 9.3.5 we have

$$\dim F = \dim \ker(s(A) - A) = 1.$$

(d) We claim that if $s(A)$ is a pole of the resolvent, then there is an eigenvector $0 \lneqq f \in E$ of A corresponding to $s(A)$. Indeed, let k be the order of the pole $s(A)$ and $R_{-k} = \lim_{\lambda\to s(A)^+}(\lambda - s(A))^k R(\lambda, A)$ the corresponding residue. Then, $R_{-k} \neq 0$ and

$R_{-(k+1)} = 0$. Moreover, by Corollary 13.2.4 and (B.2), we have $R_{-k} \geq 0$. Hence, there is $0 \leq g \in E$ with $f := R_{-k}g \gneqq 0$. By the relation $R_{-(k+1)} = (A - s(A))R_{-k} = 0$ we obtain $(A - s(A))f = 0$. This proves the claim.

We can now use (a) to obtain $\overline{E_f} = E$. By taking the adjoint $R^*_{-(k+1)}$ of $R_{-(k+1)}$ and by the same computation as before one has, if $s(A)$ is a pole of the resolvent, then there is $0 \lneqq f^* \in \ker(s(A) - A^*)$. So by (c) we have $\dim\ker(s(A) - A) = 1$.

Now, assume that $k \geq 2$. Then we have

$$\begin{aligned} \langle f, f^*\rangle &= \langle R_{-k}g, f^*\rangle \\ &= \langle g, R^*_{-k}f^*\rangle \\ &= \langle g, R^*_{-(k+1)}(A^* - s(A))f^*\rangle \\ &= 0. \end{aligned}$$

Since $\overline{E_f} = E$, it follows that $\langle g, f^*\rangle = 0$ for all $g \in E_+$. This contradicts the assertion (b). Hence $k = 1$. From the inequality $m_g + k - 1 \leq m_a \leq m_g k$ (see Proposition B.1.16) we obtain

$$m_a = m_g = \dim P_{s(A)}E = \dim\ker(s(A) - A) = 1,$$

where we recall that $P_{s(A)} = R_{-1}$. Since $P_{s(A)}E \subseteq \ker(s(A) - A)$, it follows that

$$P_{s(A)}E = \ker(s(A) - A).$$

We now show the last part of Assertion (d). To this purpose, let $0 \lneqq f \in \ker(s(A) - A)$. Without loss of generality, we suppose that $\|f\| = 1$. Then $P_{s(A)}E = \mathrm{Span}\{f\}$, i.e. $P_{s(A)}g = \lambda f$ for some $\lambda \in \mathbb{C}$ and every $g \in E$. By the Hahn-Banach theorem (see Exercise 9.5.9) there exists $0 \leq g^* \in (\ker(s(A) - A))^*$ with $\|g^*\| = 1$ and $\langle f, g^*\rangle = \|f\| = 1$. Hence $\langle P_{s(A)}g, g^*\rangle = \lambda = \langle g, P^*_{s(A)}g^*\rangle$. If we put $u^* := P^*_{s(A)}g^* \geq 0$, then $P_{s(A)} = u^* \otimes f$ and $\langle f, u^*\rangle = \langle P_{s(A)}f, g^*\rangle = \langle f, g^*\rangle = 1$. This implies that $0 \lneqq u^* \in P^*_{s(A)}E^* \subseteq \ker(s(A) - A^*)$. So $u^* > 0$ by (b). This ends the proof of the proposition. □

The following result describes the eigenvalues of an irreducible semigroup which are contained in the boundary spectrum $\sigma_b(A) = \{\lambda \in \sigma(A) : \mathrm{Re}\,\lambda = s(A)\}$, where A is the corresponding generator.

Theorem 14.2.5. *Let $(T(t))_{t\geq 0}$ be an irreducible C_0-semigroup with generator A on a Banach lattice E. Assume that $s(A) = 0$ and there is $0 \lneqq f^* \in D(A^*)$ with $A^*f^* = 0$. If $\sigma_p(A) \cap i\mathbb{R} \neq \emptyset$, then the following assertions hold.*

(a) For $0 \neq h \in D(A)$ and $\alpha \in \mathbb{R}$ with $Ah = i\alpha h$, $|h|$ is a quasi-interior point and

$$S_h(D(A)) = D(A) \text{ and } S_h^{-1}AS_h = A + i\alpha$$

hold, where S_h is the signum operator defined in Section 10.1.1.

(b) $\dim\ker(\lambda - A) = 1$ for every $\lambda \in \sigma_p(A) \cap i\mathbb{R}$.

(c) $\sigma_p(A) \cap i\mathbb{R}$ *is an additive subgroup of* $i\mathbb{R}$.

(d) 0 is the only eigenvalue of A *admitting a positive eigenvector.*

Proof. We first remark that by Proposition 14.2.4.(b) we have $f^* > 0$ and $f^* = T(t)^* f^*$ for all $t \geq 0$.

(a) Assume that $Ah = i\alpha h$ for $0 \neq h \in D(A)$ and $\alpha \in \mathbb{R}$. Then $T(t)h = e^{i\alpha t}h$ and hence $|h| = |T(t)h| \leq T(t)|h|$. This implies that

$$T(t)|h| - |h| \geq 0 \quad \text{for all}\, t \geq 0.$$

On the other hand,

$$\begin{aligned} \langle T(t)|h| - |h|, f^* \rangle &= \langle |h|, T(t)^* f^* \rangle - \langle |h|, f^* \rangle \\ &= 0 \quad \text{for all}\, t \geq 0. \end{aligned}$$

Since $f^* > 0$, we obtain $T(t)|h| = |h|$ for all $t \geq 0$, which implies that $A|h| = 0$. So, by Proposition 14.2.4.(a), $|h|$ is a quasi-interior point. If we set $T_\alpha(t) := e^{-i\alpha t}T(t)$, $t \geq 0$, then $T(t)$ and $T_\alpha(t)$ satisfy the assumptions of Lemma 10.1.8 and hence

$$T(t) = S_h^{-1} T_\alpha(t) S_h, \quad t \geq 0.$$

Therefore, $S_h(D(A)) = D(A)$ and $A = S_h^{-1}(A - i\alpha)S_h$ and (a) is proved.

(b) It follows from (a) that $S_h : \ker(i\alpha + A) \to \ker A$ for $i\alpha \in \sigma_p(A) \cap i\mathbb{R}$. On the other hand, the proof of (a) implies that $\ker A \neq \{0\}$. So, by Proposition 14.2.4.(c), $\dim \ker A = 1$ and hence $\dim \ker(i\alpha - A) = 1$.

(c) Let $0 \neq h, g \in D(A)$, $\alpha, \beta \in \mathbb{R}$ such that $Ah = i\alpha h$ and $Ag = i\beta g$. By (a) we have

$$S_g^{-1} A S_g = A + i\beta \text{ and } S_h A S_h^{-1} = A - i\alpha.$$

Thus $A + i(\beta - \alpha) = S_h(A + i\beta)S_h^{-1} = S_h S_g^{-1} A S_g S_h^{-1}$ which implies that $\ker(A + i(\beta - \alpha)) = S_h S_g^{-1} \ker A \neq \{0\}$. Therefore

$$i(\beta - \alpha) \in \sigma_p(A).$$

(d) If $Af = \lambda f$, where $0 \lneqq f \in D(A)$, then

$$\lambda \langle f, f^* \rangle = \langle Af, f^* \rangle = \langle f, A^* f^* \rangle = 0.$$

Since $f^* > 0$, it follows that $\langle f, f^* \rangle > 0$. Hence, $\lambda = 0$. □

For irreducible semigroups we obtain the following description of the boundary spectrum.

Theorem 14.2.6. *Let* $(T(t))_{t \geq 0}$ *be an irreducible* C_0-*semigroup with generator* A *on a Banach lattice* E *and assume that* $s(A)$ *is a pole of the resolvent. Then there is* $\alpha \geq 0$ *such that*

$$\sigma_b(A) = s(A) + i\alpha\mathbb{Z}.$$

Moreover, $\sigma_b(A)$ *contains only algebraically simple poles.*

Proof. Without loss of generality we suppose that $s(A) = 0$. It can be shown that $\sigma_b(A) \subseteq \sigma_p(A)$, see Lemma B.1.15. Hence, $\sigma_b(A) = \sigma_p(A) \cap i\mathbb{R}$. By Proposition 14.2.4.(d) we obtain the existence of a positive eigenvector $f^* \in D(A^*)$ corresponding to the eigenvalue $s(A) = 0$. It follows from Theorem 14.2.5.(c) that $\sigma_b(A)$ is a subgroup of $(i\mathbb{R}, +)$. Since $\sigma_b(A)$ is closed and $s(A) = 0$ is an isolated point, we have

$$\sigma_b(A) = i\alpha\mathbb{Z} \qquad \text{for some}\, \alpha \geq 0.$$

Proposition 14.2.4.(d) implies that 0 is a simple pole and by Theorem 14.2.5.(a) we have, for $\lambda \in \rho(A)$,

$$R(\lambda + ik\alpha, A) = S_h^k R(\lambda, A) S_h^{-k} \qquad \text{for all}\, k \in \mathbb{Z}.$$

Therefore, $ik\alpha$ is a simple pole for each $k \in \mathbb{Z}$. This ends the proof of the theorem. □

We now give sufficient conditions for a C_0-semigroup to possess an asynchronous exponential growth. This result will be very useful for many applications.

Theorem 14.2.7. *Let $(T(t))_{t\geq 0}$ be an irreducible C_0-semigroup with generator A on a Banach lattice E. If $\omega_{ess}(T) < \omega_0(T)$, then there exists a quasi-interior point $0 \leq f \in E$, $0 < f^* \in E^*$ with $\langle f, f^* \rangle = 1$ such that*

$$\|e^{-s(A)t}T(t) - f^* \otimes f\| \leq Me^{-\varepsilon t} \qquad \textit{for all}\ t \geq 0,$$

and appropriate constants $M \geq 1$ and $\varepsilon > 0$.

Proof. Since $\omega_{ess}(T) < \omega_0(T)$, it follows from Proposition B.1.18 that

$$s(A) = \omega_0(T).$$

On the other hand, $\omega_{ess}(T) < \omega_0(T)$ implies that $r_{ess}(T(1)) < r(T(1))$. Hence, by Proposition B.1.17, $r(T(1))$ is a pole of the resolvent of $T(1)$. Hence $\omega_0(T) = s(A)$ is a pole of $R(\cdot, A)$. Thus, by Theorem 14.2.6, it follows that there exists $\alpha \geq 0$ such that $\sigma_b(A) = s(A) + i\alpha\mathbb{Z}$ and therefore $\sigma_b(A_{-\omega_0}) = i\alpha\mathbb{Z}$, where $A_{-\omega_0}$ denotes the generator of the semigroup $T_{-\omega_0}(t) := e^{-\omega_0(A)t}T(t)$, $t \geq 0$. Since $\omega_{ess}(T_{-\omega_0}) < 0$ and $\omega_0(T_{-\omega_0}) = 0$, we have, by Theorem B.1.19, that

$$\{\lambda \in \sigma(A_{-\omega_0}) : \operatorname{Re}\lambda \geq 0\} = \{\lambda \in \sigma(A_{-\omega_0}) : \operatorname{Re}\lambda = 0\} = \sigma_b(A_{-\omega_0})$$

is finite. Therefore $\sigma_b(A_{-\omega_0}) = \{0\}$. The theorem is now proved by applying Theorem B.1.19 and Proposition 14.2.4 to the rescaled semigroup $(T_{-\omega_0}(t))_{t\geq 0}$. □

14.3 Notes and Remarks

The content of Theorem 14.1.2 is due to Greiner and Nagel (Hilbert space case) [3] and Derdinger (the other cases) [2]. It is true also in L^p spaces due to a result by Weis [5]. For an elegant proof of Weis' Theorem see Arendt et al. [1, Theorem 5.3.6].

For the spectral theory of irreducible semigroups we refer to the classic monograph by W. Arendt, A. Grabosch, G. Greiner, U. Groh, H. Lotz, U. Moustakas, R. Nagel, F. Neubrander, and U. Schlotterbeck [4, Section B-III.3].

14.4 Exercises

1. We say that a Banach lattice E has order continuous norm if every monotone order bounded sequence of E is convergent.
   (a) Prove that every AL-space has order continuous norm.
   (b) Let A be a resolvent positive operator on a Banach lattice with order continuous norm E. Prove that $\overline{D(A)}$ is an ideal in E.

2. On $E := C_0(\mathbb{R}^d)$ we consider the d-dimensional heat semigroup defined by

$$(T(t)f)(x) = (g_t * f)(x) := \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4t}} f(y)dy \quad \text{for } t > 0 \text{ and}$$
$$T(0)f := f \in E,$$

   (a) Prove that its generator in E is given by

$$\begin{aligned} D(A) &= \{f \in C_0(\mathbb{R}^d) : \Delta f \in C_0(\mathbb{R}^d)\} \\ Af &= \Delta f, \quad f \in D(A), \end{aligned}$$

   where the equality $Af = \Delta f$ for $f \in D(A)$ is in the sense of distributions.
   (b) Prove that $\lim_{t\to\infty} T(t)f = 0$ for all $f \in E$.
   (c) Prove that A is injective but not surjective.

3. Let $(T(t))_{t\geq 0}$ be a C_0-semigroup on a Banach space X. Prove that the following assertion are equivalent:
   (a) $\omega_{ess}(T) < 0$.
   (b) $\|T(t_0) - K\| < 1$ for some $t_0 > 0$ and $K \in \mathcal{L}(X)$ compact.

4. On the Banach space $C(K)$ with $K = [-\infty, 0]$ we consider the operator

$$\begin{aligned} Af &= f' + mf \\ D(A) &= \{f \in C(K) : f \text{ is differentiable, } f' \in C(K) \text{ and } f'(0) = Lf\}, \end{aligned}$$

   where $m \in C(K)$ is real-valued and $L : C(K) \to \mathbb{R}$ a continuous linear form.
   (i) Show that A generates a C_0-semigroup $(T(t))_{t\geq 0}$ on $C(K)$.
   (ii) Prove that $(T(t))_{t\geq 0}$ is given by

$$T(t)f(s) = \begin{cases} e^{\int_s^0 m(\tau)\,d\tau} \Big(e^{(s+t)m(0)} f(0) \\ \quad + \int_0^{t+s} e^{\tau m(0)} LT(s+t-\tau)f\,d\tau\Big) & \text{for } s+t > 0, \\ e^{\int_s^{t+s} m(\tau)\,d\tau} f(t+s) & \text{for } s+t \leq 0. \end{cases}$$

(iii) By using Exercise 14.4.3, prove that $\omega_{ess}(T) < 0$ provided that $m(-\infty) < 0$.

5. Let us consider the transport operator

$$\begin{aligned} D(A) &= \left\{ f \in L^1(I \times V) : v\frac{\partial f}{\partial x} \in L^1(I \times V) \text{ and } \left\{ \begin{array}{l} f(0,v) = 0 \text{ if } v > 0, \\ f(1,v) = 0 \text{ if } v < 0 \end{array} \right\} \right. \\ (Af)(x,v) &= -v\frac{\partial f}{\partial x}(x,v), \end{aligned}$$

where $I = [0,1]$ and $V = \{v \in \mathbb{R} : 1 \le |v| \le 2\}$. Prove that A generates an irreducible C_0-semigroup on $L^1(I \times V)$.

6. On the Banach lattice $C[0,1]$ we consider the Laplace operator with Neumann boundary conditions

$$\begin{aligned} (Af)(x) &= f''(x), \quad x \in [0,1], \\ f \in D(A) &= \{f \in C^2[0,1] : f'(0) = f'(1) = 0\}. \end{aligned}$$

Prove that A generates an irreducible C_0-semigroup on $C[0,1]$.

7. Let $E = L^1[-1,0]$ and for $0 \le g \in L^\infty[-1,0]$ define the operator

$$Af := f',\ D(A) = \{f \in E : f' \in E \text{ and } f(0) = \int_{-1}^{0} f(s)g(s)\,ds\}.$$

(a) Show that A generates a positive C_0-semigroup $(T(t))_{t\ge 0}$ on E.

(b) Prove that $(T(t))_{t\ge 0}$ is not irreducible if and only if there exists $\varepsilon > 0$ such g vanishes a.e. on $[-1,-1+\varepsilon]$.

Bibliography

[1] W. Arendt, C.J.K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser Verlag, 2011.

[2] R. Derndinger, *Über das Spektrum positiver Generatoren*, Math. Z. **172** (1980), 281–293.

[3] G. Greiner and R. Nagel, *On the stability of strongly continuous semigroups of positive operators on $L^2(\mu)$*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **10** (1983), 257–262.

[4] R. Nagel (ed.), *One-parameter Semigroups of Positive Operators*, Lect. Notes in Math., vol. 1184, Springer-Verlag, 1986.

[5] L. Weis, *The stability of positive semigroups on L_p-spaces*, Proc. Amer. Math. Soc. **123** (1995), 3089–3094.

Lecture 15

An application to linear transport equations

In our final lecture we apply the Perron-Frobenius theory developed in Chapter 14 to a concrete example. We study the asymptotic behaviour of a one-dimensional linear transport equation. Before introducing the model we however need to state a formula for the perturbed semigroup.

15.1 Dyson-Phillips expansion

We know from Theorem 12.4.1 that $A + B$ with domain $D(A)$ generates a C_0-semigroup $(S(t))_{t\geq 0}$ on a Banach space X whenever A generates a C_0-semigroup $(T(t))_{t\geq 0}$ and $B \in \mathcal{L}(X)$. The perturbed semigroup $(S(t))_{t\geq 0}$ can be given by a series called *Dyson-Phillips expansion* as the following result shows.

Proposition 15.1.1. *Let A with domain $D(A)$ be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach space X. Let $B \in \mathcal{L}(X)$. Then the semigroup $(S(t))_{t\geq 0}$ generated by $A + B$ with domain $D(A)$ is given by the Dyson-Phillips expansion*

$$S(t) = \sum_{n=0}^{\infty} U_n(t), \quad t \geq 0, \tag{15.1}$$

where

$$U_0(t) = T(t) \text{ and } U_{n+1}(t) = \int_0^t U_n(t-s)BT(s)\,ds, \quad t \geq 0, n \in \mathbb{N}.$$

Moreover, the following variantion of constant formulas are satisfied :

$$\begin{aligned} S(t)f &= T(t)f + \int_0^t S(t-s)BT(s)f\,ds \\ &= T(t)f + \int_0^t T(t-s)BS(s)f\,ds, \quad f \in X,\, t \geq 0. \end{aligned}$$

Proof. Since $U_0(t) = T(t)$, we have $\|U_0(t)\| \le Me^{\omega t}$ for $t \ge 0$ and some constants $M \ge 1$ and $\omega \in \mathbb{R}$. Thus,

$$\begin{aligned} \|U_1(t)\| &\le Me^{\omega t} \int_0^t e^{-\omega s} \|BT(s)\| \, ds \\ &\le M^2 \|B\| t e^{\omega t} \quad \text{for all } t \ge 0. \end{aligned}$$

By induction one can see that

$$\|U_n(t)\| \le M \frac{(M\|B\|t)^n}{n!} e^{\omega t} \quad \text{for all } t \ge 0.$$

Therefore the series $\sum_{n=0}^{\infty} U_n(t)$ converges in $\mathcal{L}(X)$ uniformly on compact intervals of $\mathbb{R}_+$. Moreover,

$$\|\sum_{n=0}^{\infty} U_n(t)\| \le Me^{(\omega + M\|B\|)t} \quad \text{for all } t \ge 0.$$

Set $U(t) := \sum_{n=0}^{\infty} U_n(t)$ for $t \ge 0$. Since $t \mapsto U_n(t)f$ is continuous and the convergence is uniform on compact subsets of $\mathbb{R}_+$, we deduce that the function $t \mapsto U(t)f$ is continuous for any $f \in X$.

Let us prove now that $(U(t))_{t\ge 0}$ is a C_0-semigroup which coincides with the perturbed semigroup $(S(t))_{t \ge 0}$. For this purpose we let $f \in X$, $t, s \ge 0$, and first verify that

$$\sum_{k=0}^{n} U_k(s) U_{n-k}(t) f = U_n(t+s) f.$$

This is obviously true for $n = 0$. Assuming it is satisfied for some $n \in \mathbb{N}$ we compute

$$\begin{aligned} &\sum_{k=0}^{n+1} U_k(s) U_{n+1-k}(t) f \\ =\ &\sum_{k=0}^{n} U_k(s) \int_0^t U_{n-k}(t-r) BT(r) f \, dr + \int_0^s U_n(s-r) BT(r+t) f \, dr \\ =\ &\int_0^t U_n(s+t-r) BT(r) f \, dr + \int_t^{t+s} U_n(s+t-r) BT(r) f \, dr \\ =\ &\int_0^{t+s} U_n(s+t-r) BT(r) f \, dr = U_{n+1}(t+s) f, \quad t, s \ge 0,\ f \in X. \end{aligned}$$

Hence,

$$\begin{aligned} U(t)U(s)f &= \sum_{n=0}^{\infty} U_n(t) \sum_{m=0}^{\infty} U_m(s) f \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{j} U_k(t) U_{j-k}(s) f \\ &= \sum_{j=0}^{\infty} U_j(t+s) f = U(t+s) f, \quad t, s \ge 0,\ f \in X. \end{aligned}$$

So, $(U(t))_{t\geq 0}$ is a C_0-semigroup on X. Let us denote its generator by C with domain $D(C)$. Using the definition of $U(t)$ we obtain

$$\begin{aligned} U(t) &= T(t) + \sum_{n=0}^{\infty} U_{n+1}(t) \\ &= T(t) + \sum_{n=0}^{\infty} \int_0^t U_n(t-s)BT(s)\,ds \\ &= T(t) + \int_0^t \left(\sum_{n=0}^{\infty} U_n(t-s)BT(s) \right) ds \\ &= T(t) + \int_0^t U(t-s)BT(s)\,ds, \qquad \text{for all } t \geq 0. \end{aligned}$$

Hence, for $f \in D(A)$, we have

$$\begin{aligned} & \lim_{t\to 0} \left(\frac{1}{t} \int_0^t U(t-s)BT(s)f\,ds - Bf \right) \\ = & \lim_{t\to 0} \left(\frac{1}{t} \int_0^t U(t-s)B(T(s)f - f)\,ds + \frac{1}{t} \int_0^t (U(t-s)Bf - Bf)\,ds \right) \\ = & \lim_{t\to 0} \left(\frac{1}{t} \int_0^t U(t-s)B \int_0^s T(r)Af\,drds + \frac{1}{t} \int_0^t (U(s)Bf - Bf)\,ds \right) = 0. \end{aligned}$$

Thus, $f \in D(C)$ and $Cf = Af + Bf$. So, the assertion follows since C and $A+B$ are both generators. □

The above Dyson-Phillips expansion gives us the following positivity result.

Corollary 15.1.2. *If A generates a positive C_0-semigroup on a Banach lattice E and $B \in \mathcal{L}(E)$ is a positive operator, then the semigroup generated by $A+B$ is positive.*

15.2 Positive semigroups for linear transport equations

We would like to model the time evolution of the motion of neutrons in an absorbing and scattering homogeneous medium. This problem is known as the *reactor problem* and is given by the following integrodifferential equation.

$$\frac{\partial}{\partial t} u(t,x,v) = -\sum_{j=1}^{3} v_j \frac{\partial}{\partial x_j} u(t,x,v) - \sigma(x,v)u(t,x,v) + \int_V \zeta(x,v,v')u(t,x,v')\,dv', \tag{15.2}$$

where $u(t,x,v)$ represents the density distribution of the neutrons in terms of the space variable $x \in D \subseteq \mathbb{R}^3$, velocity $v \in V := \{v \in \mathbb{R}^3 : v_{min} \leq |v| \leq v_{max}\}$ at time t for

some minimal speed $v_{min} > 0$ and maximal speed $v_{max} < \infty$. Here D denotes the set describing the interior of the vessel in which neutron transport takes place. The medium D is surrounded by a total absorber (or by a vacuum if D is convex), and neutrons migrate in this volume, are scattered and absorbed by this material. We suppose that neutrons do not interact with each other.
The *free streaming term* $-\sum_{j=1}^{3} v_j \frac{\partial}{\partial x_j} u$ in (15.2) is responsible for the motion of the particles between collisions with the background material. The second term of the right-hand side of (15.2) corresponds to collisions including absorption σ, and the third term to scattering of neutrons: particles at the position x with the incoming speed v' generate particles at x with the outgoing speed v and the transition is governed by a scattering kernel $\zeta(x, v, v')$. The fact that $u(t, \cdot, \cdot)$ should describe a density suggests to require that $u(t, \cdot, \cdot)$ is an element of $L^1(D \times V)$ for all $t \geq 0$. Following this line and introducing the vector-valued function $u(t) := u(t, \cdot, \cdot)$, (15.2) is equivalent to the following abstract Cauchy problem

$$\begin{cases} u'(t) = (A + K_\zeta)u(t) := (A_0 - M_\sigma)u(t) + K_\zeta u(t), & t \geq 0, \\ u(0) \in D(A + K_\zeta). \end{cases}$$

Here $u(t) \in L^1(D \times V)$ and A_0 denotes the *free streaming operator* $-\sum_{j=1}^{3} v_j \frac{\partial}{\partial x_j}$ on a suitable domain. Moreover M_σ is the multiplication operator by σ and is called the *absorption operator*. The *scattering operator* K_ζ is defined by

$$(K_\zeta f)(x, v) := \int_V \zeta(x, v, v') f(x, v')\, dv', \qquad (x, v) \in D \times V, f \in L^1(D \times V).$$

The one-dimensional reactor problem

In order to simplify our exposition we study here only the one-dimensional case. The one-dimensional reactor problem is given by the following transport equation

$$\text{(TE)} \qquad \begin{cases} \frac{\partial u}{\partial t}(t, x, v) = -v \frac{\partial u}{\partial x}(t, x, v) - \sigma(x, v)u(t, x, v) + \int_V \zeta(x, v, v')u(t, x, v')\, dv', \\ \qquad\qquad\qquad\qquad\qquad\qquad t \geq 0, (x, v) \in D \times V, \\ u(t, 0, v) = 0 \text{ if } v > 0 \text{ and } u(t, 1, v) = 0 \text{ if } v < 0, \quad t \geq 0, \\ u(0, x, v) = f(x, v), \qquad (x, v) \in D \times V, \end{cases}$$

where $0 \leq \sigma \in L^\infty(D \times V)$, $0 \leq \zeta \in L^\infty(D \times V \times V)$, $D := [0, 1]$, and $V := \{v \in \mathbb{R} : v_{min} \leq |v| \leq v_{max}\}$ for given constants $0 < v_{min} < v_{max} < \infty$.
Supposing that the scattering kernel ζ satisfies

$$\text{(15.3)} \qquad \zeta(x, v, v') > 0 \qquad \text{for all } (x, v, v') \in D \times V \times V,$$

one can apply Theorem 14.2.7 and deduce the asymptotic behaviour of the semigroup solution to (TE).

Let us first rewrite (TE) as an abstract Cauchy problem on the Banach lattice $L^1(D \times V)$. So, we define the *free streaming* operator A_0 by

$$(A_0 f)(x,v) := -v\frac{\partial f}{\partial x}(x,v) \text{ with}$$
$$D(A_0) := \left\{ f \in L^1(D\times V) : v\frac{\partial f}{\partial x} \in L^1(D \times V),\ \begin{array}{ll} f(0,v)=0 & \text{if } v>0 \\ f(1,v)=0 & \text{if } v<0 \end{array} \right\},$$

the *absorption* operator

$$(M_\sigma f)(x,v) := \sigma(x,v) f(x,v), \quad (x,v) \in D \times V, f \in L^1(D\times V),$$

and the *scattering* operator

$$(K_\zeta f)(x,v) := \int_V \zeta(x,v,v') f(x,v')\, dv', \quad (x,v) \in D\times V, f \in L^1(D \times V).$$

We note that M_σ, $K_\zeta \in \mathcal{L}(L^1(D\times V))$ since $\sigma \in L^\infty(D\times V)$ and $\zeta \in L^\infty(D\times V\times V)$.

Let us study first the free streaming operator. By an easy computation one can see that $(0,\infty) \subseteq \rho(A_0)$ and

$$(R(\lambda, A_0) f)(x,v) = \begin{cases} \frac{1}{v}\int_0^x e^{-\frac{\lambda}{v}(x-x')} f(x',v)\, dx' & \text{if } v>0, \\ -\frac{1}{v}\int_x^1 e^{-\frac{\lambda}{v}(x-x')} f(x',v)\, dx' & \text{if } v<0, \end{cases} \tag{15.4}$$

for $(x,v) \in D\times V$ and $f \in L^1(D\times V)$. Hence,

$$(0,\infty) \subseteq \rho(A_0) \text{ and } \|R(\lambda, A_0)\| \le \frac{1}{\lambda} \qquad \text{for all } \lambda > 0.$$

Therefore, by Theorem 12.3.1, A_0 with domain $D(A_0)$ generates a C_0-semigroup $(T_0(t))_{t\ge 0}$ of contractions on $L^1(D\times V)$. Moreover, $(T_0(t))_{t\ge0}$ is positive since $R(\lambda, A_0) \ge 0$ for all $\lambda > 0$. On the other hand, it follows from (15.4) that

$$(R(\lambda, A_0) f)(x,v) = \int_0^\infty e^{-\lambda t} \mathbb{1}_D(x-vt) f(x-vt)\, dt$$

for $(x,v) \in D\times V$, $f \in L^1(D\times V)$, where $\mathbb{1}_D(x) = \begin{cases} 1 & \text{if } x \in D, \\ 0 & \text{if } x \notin D. \end{cases}$
So, by the uniqueness of the Laplace transform, we obtain

$$(T_0(t) f)(x,v) = \mathbb{1}_D(x-tv) f(x-tv, v), \quad (x,v) \in D\times V, f \in L^1(D\times V). \tag{15.5}$$

Moreover, since the absorption operator M_σ is bounded, it follows that

$$A := A_0 - M_\sigma \text{ with } D(A) = D(A_0)$$

generates the positive C_0-semigroup $(T(t))_{t\geq 0}$ given by

$$(T(t)f)(x,v) = e^{-\int_0^t \sigma(x-\tau v,v)\,d\tau}(T_0(t)f)(x,v), \tag{15.6}$$

for $(x,v)\in D\times V$, $f\in L^1(D\times V)$. The boundedness and the positivity of the scattering operator K_ζ implies that the transport operator $A+K_\zeta$ with domain $D(A_0)$ generates the positive C_0-semigroup $(S(t))_{t\geq 0}$ given by the Dyson-Phillips expansion (15.1), see Proposition 15.1.1 and Corollary 15.1.2. This semigroup will be called the *transport semigroup* and it satisfies the following properties:

Proposition 15.2.1. *For the streaming semigroup $(T(t))_{t\geq 0}$ and the transport semigroup $(S(t))_{t\geq 0}$ the following hold*

$$\begin{aligned} 0\leq T(t)\leq S(t) \qquad &\textit{for all } t\geq 0 \textit{ and}\\ \omega_0(A+K_\zeta) = s(A+&K_\zeta). \end{aligned} \tag{15.7}$$

Proof. The first assertion follows from the positivity of K_ζ and the Dyson-Phillips expansion (15.1). The second is a consequence of Theorem 14.1.2.(ii). ☐

We will see that the the transport semigroup $(S(t))_{t\geq 0}$ is irreducibile. First we show the following.

Lemma 15.2.2. *Let I be a closed ideal of $L^1(D\times V)$. If I is $S(\cdot)$-invariant, then I is invariant both under $T_0(\cdot)$ and K_ζ.*

Proof. Assume that I is $S(\cdot)$-invariant. Since $0\leq T(t)\leq S(t)$, $t\geq 0$, we deduce that I is $T(\cdot)$-invariant. Thus, by (15.6), it follows that I is $T_0(\cdot)$-invariant. By Proposition 15.1.1, we have

$$\lim_{t\downarrow 0}\frac{1}{t}(S(t)f - T(t)f) = \lim_{t\downarrow 0}\frac{1}{t}\int_0^t T(t-s)K_\zeta S(s)f\,ds = K_\zeta f$$

for $f\in L^1(D\times V)$. Since I is closed and invariant both under $S(\cdot)$ and $T(\cdot)$, we obtain that I is $K_\zeta-$invariant. ☐

Lemma 15.2.3. *For the transport semigroup $(S(t))_{t\geq 0}$ defined above the following properties hold.*

(i) The remainder $R_2(t) := \sum_{n=2}^\infty S_n(t)$, $t\geq 0$, of the Dyson-Phillips expansion (15.1) is a weakly compact operator on $L^1(D\times V)$.

(ii) If the scattering kernel ζ satisfies (15.3), then the transport semigroup $(S(t))_{t\geq 0}$ is irreducible.

Proof. For $0\leq f\in L^1(D\times V)$ and $t>0$ we have

$$\begin{aligned} (K_\zeta T(t)K_\zeta f)(x,v) &\leq (K_\zeta T_0(t)K_\zeta f)(x,v)\\ &\leq \|\zeta\|_\infty^2\int_V\int_V \mathbb{1}_D(x-tv'')f(x-tv'',v')\,dv''\,dv'\\ &\leq t^{-1}\|\zeta\|_\infty^2\int_V\int_D f(x',v')\,dx'\,dv'. \end{aligned}$$

Hence,

$$K_\zeta T(t) K_\zeta \leq \frac{\|\zeta\|_\infty^2}{t}(\mathbb{1}\otimes\mathbb{1}), \tag{15.8}$$

where $\mathbb{1}\otimes\mathbb{1}$ is the bounded linear operator defined by

$$(\mathbb{1}\otimes\mathbb{1})f := \left(\int_D\int_V f(x,v)\,dv\,dx\right)\mathbb{1}, \qquad f\in L^1(D\times V).$$

By using the definition of the terms $U_n(t)$ in the Dyson-Phillips series (15.1) one can see that

$$R_{n+1}(t) := \sum_{k=n+1}^{\infty} U_k(t) = \int_0^t T(t-s)K_\zeta R_n(s)\,ds, \quad t\geq 0, n\in\mathbb{N}.$$

In particular, $R_2(t) = \int_0^t\int_0^{t-s_2} T(s_1)K_\zeta T(s_2)K_\zeta S(t-s_1-s_2)\,ds_1\,ds_2$ for $t\geq 0$. Take $t>\varepsilon>0$ and consider

$$R_{2,\varepsilon}(t) := \int_\varepsilon^t\int_0^{t-s_2} T(s_1)(K_\zeta T(s_2)K_\zeta)S(t-s_1-s_2)\,ds_1\,ds_2.$$

Then it is easy to verify that

$$\lim_{\varepsilon\to 0}\|R_{2,\varepsilon}(t)-R_2(t)\| = 0 \quad \text{for all}\, t>0.$$

On the other hand, it follows from (15.8) that

$$R_{2,\varepsilon}(t) \leq \|\zeta\|_\infty^2 \int_\varepsilon^t\int_0^{t-s_2} \frac{1}{s_2}T(s_1)(\mathbb{1}\otimes\mathbb{1})S(t-s_1-s_2)\,ds_1\,ds_2.$$

From the definition of $(T_0(t))_{t\geq 0}$ and since $0\leq T(t)\leq T_0(t)$, one can see that $T(t)\mathbb{1}\otimes\mathbb{1}\leq \mathbb{1}\otimes\mathbb{1}$ for the order in $\mathcal{L}(L^1(D\times V))$. Now, for $0\leq f\in L^1(D\times V)$, and $s_1+s_2\leq t$, we obtain

$$\begin{aligned}(\mathbb{1}\otimes\mathbb{1})S(t-s_1-s_2)f &= \left(\int_D\int_V (S(t-s_1-s_2)f)(x,v)\,dv\,dx\right)\mathbb{1}\\ &\leq Me^{\omega(t-s_1-s_2)}\left(\int_D\int_V f(x,v)\,dv\,dx\right)\mathbb{1}\\ &= Me^{\omega(t-s_1-s_2)}(\mathbb{1}\otimes\mathbb{1})f,\end{aligned}$$

where $M\geq 1$ and $\omega\in\mathbb{R}$ are such that $\|S(t)\|\leq Me^{\omega t}$ for all $t\geq 0$. Consequently,

$$\begin{aligned}R_{2,\varepsilon}(t) &\leq M\|\zeta\|_\infty^2\left(\int_\varepsilon^t\frac{1}{s_2}\int_0^{t-s_2} e^{\omega(t-s_1-s_2)}\,ds_1\,ds_2\right)(\mathbb{1}\otimes\mathbb{1})\\ &= \frac{M\|\zeta\|_\infty^2}{\omega}\left(\int_\varepsilon^t \frac{e^{\omega(t-s_2)}-1}{s_2}\,ds_2\right)(\mathbb{1}\otimes\mathbb{1}),\end{aligned}$$

where $M \geq 1$ and $\omega \in \mathbb{R}$ are such that $\|S(t)\| \leq Me^{\omega t}$ for all $t \geq 0$. This implies that $R_{2,\varepsilon}(t)$ is dominated by a one-dimensional operator. So, by Proposition B.1.21, we obtain that $R_{2,\varepsilon}(t)$ is weakly compact and therefore $R_2(t)$ is weakly compact for all $t \geq 0$. This proves (i).
We recall that every closed ideal in $L^1(D \times V)$ has the form

$$I = \{f \in L^1(D \times V) : f \text{ vanish a.e. on } \Omega\}$$

for some measurable subset $\Omega \subseteq D \times V$. We suppose that I is $S(t)-$invariant for all $t \geq 0$. Then, by Lemma 15.2.2, I is $K_\zeta-$invariant. Assume that $\Omega \neq \emptyset$. Since $\mathbb{1}_{D\times V\setminus\Omega} \in I$, we obtain

$$\begin{aligned}(K_\zeta \mathbb{1}_{D\times V\setminus\Omega})(x,v) &= \int_V \zeta(x,v,v')\mathbb{1}_{D\times V\setminus\Omega}(x,v')\,dv' \\ &= \int_{V\setminus\Omega_x} \zeta(x,v,v')\,dv' = 0\end{aligned}$$

for $(x,v) \in \Omega$ and $\Omega_x := \{v \in V : (x,v) \in \Omega\}$. Since ζ is strictly positive, it follows that $\Omega_x = V$. Hence, $\Omega = Y \times V$ for some measurable subset Y of D.
On the other hand, again by Lemma 15.2.2, I is $T_0(t)-$invariant for all $t \geq 0$. Thus, I is $R(\lambda, A_0)-$invariant for all $\lambda > 0$. Hence, $(R(\lambda, A_0)\mathbb{1}_{D\times V\setminus\Omega})(x,v) = 0$ for a.e. $(x,v) \in \Omega$. So, by using (15.4), one can see that

$$\int_0^x \mathbb{1}_{D\setminus Y}(s)\,ds = 0 \text{ and } \int_x^1 \mathbb{1}_{D\setminus Y}(s)\,ds = 0.$$

Therefore, $\int_0^1 \mathbb{1}_{D\setminus Y}(s)\,ds = 0$ and this implies that $Y = D$. Consequently, $I = \{0\}$ or $I = L^1(D \times V)$ and (ii) is proved. □

We can now apply Theorem 14.2.7 to describe the asymptotic behaviour of the transport semigroup.

Theorem 15.2.4. *Assume that ζ satisfies (15.3). Then the transport semigroup $(S(t))_{t\geq 0}$ has balanced exponential growth. More precisely, there are two strictly positive functions $\varphi \in L^1(D \times V)$ and $\psi \in L^\infty(D \times V)$ satisfying $\int_{D\times V} \varphi(x,v)\psi(x,v)\,dv\,dx = 1$ such that*

$$\|e^{-s(A+K_\zeta)t}S(t) - \psi \otimes \varphi\| \leq Me^{-\varepsilon t}$$

for all $t \geq 0$ and some constants $M \geq 0$ and $\varepsilon > 0$.

Proof. Since $v_{min} > 0$, it follows that $(T(t))_{t\geq 0}$ is a nilpotent semigroup, i.e., there is $t_0 > 0$ such that

$$T(t) = 0 \text{ for all } t \geq t_0. \tag{15.9}$$

Here one can take $t_0 = \frac{1}{v_{min}}$. Hence, $r(T(t)) = r_{ess}(T(t)) = 0$ for all $t > 0$. So, by Lemma 15.2.3.(i) and Theorem B.1.23, we have

$$\omega_{ess}(S) = -\infty.$$

On the other hand, it follows from (15.9) that

$$U_1(t) = \int_0^t T(s)K_\zeta T(t-s)\,ds = 0 \text{ for all } t \geq 2t_0$$

and therefore

$$R_2(t) = S(t) \text{ for all } t \geq 2t_0.$$

So, by Lemma 15.2.3.(ii), we obtain that $R_2(t)$ is irreducible for all $t \geq 2t_0$. Now, one can apply Lemma 15.2.3.(i), Proposition B.1.20 and Theorem B.1.22, to obtain that $r(S(t)) = r(R_2(t)) > 0$ for all $t \geq 2t_0$. Therefore,

$$-\infty = \omega_{ess}(S) < \omega_0(S).$$

Applying Theorem 14.2.7 to the transport semigroup $(S(t))_{t\geq 0}$ we obtain the assertions. □

15.3 Notes and Remarks

The transport equation has been studied by many authors. Our presentation is based on the approach used in the papers [4, 7–9]. For more information see one of the monographs [1, 2, 5, 6].

As another nice application of our theory we mention the study of the asymptotic behavior of the solutions to the age-dependent population equation whose abstract form is given in Exercise 15.4.3. For a detailed study of this model we refer to [3, Sec. VI.4] and references therein.

15.4 Exercises

1. Let $(T(t))_{t\geq 0}$ be a C_0-semigroup with generator A. Moreover, assume that the mapping $t \mapsto T(t)$ is norm continuous for $t \geq t_0$ and some point $t_0 \geq 0$ and that $R(\lambda, A)T(t_0)$ is compact for some (and hence all) $\lambda \in \rho(A)$. Prove that the operators $T(t)$ are compact for all $t \geq t_0$.

2. Let us consider the operator defined on $L^1(\mathbb{R}_+)$ by

$$Af = -f' - \mu f,\ f \in D(A) := \{f \in L^1(\mathbb{R}_+) : f' \in L^1(\mathbb{R}_+) \text{ and } f(0) = \int_0^\infty \beta(a)f(a)\,da\},$$

where $\mu,\ \beta \in L^\infty(\mathbb{R}_+)$ are two nonnegative functions.

(a) Prove that A generates a positive C_0-semigroup $(T(t))_{t\geq 0}$.

(b) Prove that $(T(t))_{t\geq 0}$ is irreducible if and only if there is no $\tau \geq 0$ such that $\beta|_{(\tau,\infty)} = 0$ almost everywhere.

3. On $E := L^1[\alpha/2, 1]$ we consider the operators defined by

$$\begin{aligned} A_0 f &= -f' - (\mu + b) f \text{ with} \\ D(A_0) &= \{f \in E : f' \in E,\, f(\alpha/2) = 0\}, \\ Bf(s) &= \begin{cases} 4b(2s)f(2s) & \text{if } \alpha/2 \leq s \leq 1/2, \\ 0 & \text{if } 1/2 \leq s \leq 1 \end{cases} \end{aligned}$$

for $f \in E$, where $0 \leq \mu \in C[\alpha/2, 1]$ and b continuous with $b(s) > 0$ for $s \in (\alpha, 1)$ and $b(s) = 0$ otherwise.

(a) Prove that A_0 generates a positive semigroup $(T(t))_{t\geq 0}$ given by

$$T(t)f(s) = \begin{cases} e^{-\int_{s-t}^{s}(\mu(r)+b(r))\,dr} f(s-t) & \text{if } s - t > \alpha/2, \\ 0 & \text{otherwise.} \end{cases}$$

(b) Deduce that $A := A_0 + B$ with domain $D(A) = D(A_0)$ generates a positive C_0-semigroup $(S(t))_{t\geq 0}$ on E.

(c) Prove that $S(t)$ is compact for all $t > 1 - \alpha/2$ (use Exercise 15.4.1).

(d) Prove that $(S(t))_{t\geq 0}$ is irreducible on E.

(e) Deduce that $(S(t))_{t\geq 0}$ possess an asynchronous exponential growth. (Use Theorem B.1.22 and Theorem 14.2.7.)

Bibliography

[1] A. Belleni-Morante, *Applied Semigroups and Evolution Equations*, Oxford University Press, 1979.

[2] A. Belleni-Morante and A.C. McBride, *Applied Nonlinear Semigroups*, John Wiley & Sons, 1998.

[3] K.-J. Engel and R. Nagel, *A Short Course on Operator Semigroups*, Springer-Verlag, 2006.

[4] G. Greiner, *Spectral properties and asymptotic behavior of the linear transport equation*, Math. Z. **185** (1984), 167–177.

[5] J.Banasiak and L.Arlotti, *Perturbations of positive semigroups with applications*, Springer-Verlag, 2006.

[6] M. Mokhtar-Kharroubi, *Mathematical Topics in Neutron Transport Theory: New Aspects*, World Scientific, 1997.

[7] I. Vidav, *Spectra of perturbed semigroups with applications to transport theory*, J. Math. Anal. Appl. **30** (1970), 264–279.

[8] J. Voigt, *Positivity in time dependent linear transport theory*, Acta Appl. Math. **2** (1984), 311–331.

[9] ———, *Spectral properties of the neutron transport equation*, J. Math. Anal. Appl. **106** (1985), 140–153.

Appendix A

A.1 Linear algebra

We collect here the necessary notations and results from linear algebra. The set of $n \times n$ complex matrices will be denoted by $\mathrm{M}_n(\mathbb{C})$. We denote with $X := \mathbb{C}^n$ the usual euclidian vector space and use the identification $\mathrm{M}_n(\mathbb{C}) \cong \mathcal{L}(X)$. Here $\mathcal{L}(X)$ denotes the set of (continuous) linear transformations defined on X and mapping into X. Hence, we will view at matrices always as linear transformations or operators. This means especially that we will speak about the action of a matrix, or about the *kernel* or *image* of a matrix. These latter objects will be denoted by $\ker A$ and $\operatorname{im} A$, respectively, for a matrix $A \in M_n(\mathbb{C})$.

A.1.1 Norms

To be able to define convergence of vectors, there is a need for the notion of the length of a vector. A natural way to define the length of a vector in $\mathbb{R}^n$ is to use geometric intuition and for a vector $x = (x_1, x_2, \cdots, x_n)$ set its length as $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, the Euclidian length of a vector. However, it turns out that for different mathematical purposes different notions would serve more useful. This is crystallized in the abstract notion of the *norm* of a vector.

Definition A.1.1. A *norm* on X is a function $\|\cdot\|\colon X \to \mathbb{R}$ with the following properties.

(i) $\|x\| \geq 0$ for every $x \in X$ and $\|x\| = 0 \iff x = 0$;

(ii) $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in X$ and $\lambda \in \mathbb{C}$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

There are many norms on the space $X = \mathbb{C}^n$. Here we list some of them.

Example A.1.2. For $x = (x_1, x_2, \cdots, x_n) \in \mathbb{C}^n$ the following formulae define norms on $\mathbb{C}^n$.

(a) $\|x\|_2 := \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$ (the Euclidian or the 2-norm);

(b) $\|x\|_1 := |x_1| + |x_2| + \cdots + |x_n|$ (the taxicab or the 1-norm);

(c) $\|x\|_\infty := \max\{|x_1|, |x_2|, \dots, |x_n|\}$ (the maximum or the ∞-norm);

(d) $\|x\|_p := (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$ for $p \in \mathbb{R}$, $p \geq 1$ (the p-norm).

The p-norm is obviously a generalization of the Euclidian and taxicab norms. It is also not difficult to show that $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$. The Euclidian norm is induced by the scalar product:

$$\|x\|_2 = \sqrt{\langle x, x\rangle}.$$

Here $\langle \cdot, \cdot \rangle$ denotes the usual scalar product on $\mathbb{C}^n$, i.e. for $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n)$ we have

$$\langle x, y\rangle := x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}.$$

The following *Hölder's inequality* regarding scalar product and p-norm is very useful. For the case $p = q = 2$ it is actually the well known *Cauchy-Schwarz inequality*.

Lemma A.1.3 (Hölder's inequality)**.** *Let* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $x, y \in \mathbb{C}^n$*. Then*

$$|\langle x, y\rangle| \leq \|x\|_p \|y\|_q.$$

Any norm defines a metric $d_{\|\cdot\|}(x, y) := \|x - y\|$ on a vector space and hence the notions of convergence, continuity, etc. may depend on the specific norm we use. Therefore it is important to be able to compare different norms.

Definition A.1.4. The norms $\|\cdot\|$ and $|\!|\!|\cdot|\!|\!|$ on X are *equivalent*, if there exist constants $m, M > 0$ such that

$$m\, |\!|\!|x|\!|\!| \leq \|x\| \leq M\, |\!|\!|x|\!|\!| \tag{A.1}$$

for all $x \in X$.

One can easily see that the norms from Example A.1.2 are all equivalent. In fact, this is true for all the norms on a finite dimensional vector space, as Tikhonov proved in [3].

Theorem A.1.5 (Tikhonov)**.** *Any two norms on a finite dimensional (real or complex) vector space are equivalent.*

Proof. Clearly, it is enough to prove the statement of the theorem in the case where $X = \mathbb{C}^n$ and $|\!|\!|x|\!|\!| = \|x\|_\infty = \max\{|x_1|, |x_2|, \dots |x_n|\}$. We will make use of the fact that the $\|\cdot\|_\infty$-norm is complete and that the Bolzano-Weierstrass theorem holds, meaning that closed bounded sets are compact.

Let us introduce the constant

$$C = \max\{\|e_1\|, \|e_2\|, \dots, \|e_n\|\}$$

where the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \ e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

are the canonical basis vectors of X. We see one side of the required inequality immediately from the estimate

$$\|x\| = \|x_1 e_1 + x_2 e_2 + \dots x_n e_n\| \leq C\left(|x_1| + |x_2| + \dots + |x_n|\right) \leq Cn\|x\|_\infty.$$

Considering the set

$$K := \{x \in X \,:\, \|x\|_\infty = 1\},$$

and the function

$$f : K \to \mathbb{R}, \qquad f(x) = \|x\|,$$

we see that f is Lipschitz continuous since using the estimate before, we obtain that

$$|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq Cn\|x - y\|_\infty.$$

Hence, using that the set K is compact, by Extreme Value Theorem we obtain that there are vectors $z, w \in K$ such that for all $x \in K$, we obtain

$$\|z\| = f(z) \leq f(x) = \|x\| \leq f(w) = \|w\|.$$

Finally to conclude our proof, take $x \in X$ such that $x \neq 0$. Then $x = \|x\|_\infty x'$ with $x' = \frac{x}{\|x\|_\infty} \in K$. By the previous inequality, and using the notation $m := \|z\|$, $M := \|w\|$, we obtain

$$m \leq \|x'\| \leq M,$$

and hence, multiplying this inequality with $\|x\|_\infty$,

$$m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty.$$

□

It is important to note however that Tikhonov's Theorem does not hold in spaces of infinite dimension:

Exercise A.1.6. Let $Z := C[0,1]$ be the vector space of continuous $\mathbb{C}$-valued functions defined on the interval $[0,1]$ and put, for $f \in Z$,

$$\|f\| := \max_{t \in [0,1]} |f(t)|,$$

$$\|f\|_1 := \int_0^1 |f(t)| dt.$$

Then $\|f\|_1 \leq \|f\|$ for every f, but $\|\cdot\|_1$ and $\|\cdot\|$ are not equivalent.

A.1.2 Reducing subspaces and projections

We recall now some important notions which might be missing from some of the introductory linear algebra course. Let $A \in \mathcal{L}(X)$. A subspace $Y \subset X$ is called *invariant* under A, if

$$AY := \{Ay : y \in Y\} \subset Y.$$

A subspace $Y \subset X$ is called *reducing*, if there is a subspace $Z \subset X$ such that $X = Y \oplus Z$ and both Y and Z are invariant under A. Here the symbol "$\oplus$" denotes the direct sum of two subspaces, meaning that for every $x \in X$ there exists exactly one $y \in Y$ and one $z \in Z$ such that $x = y + z$. In this case we also say that Z *complements* Y. It is important to note that, if the subspace Y reduces A, there will, in general, be infinitely many subspaces complementing Y which are not invariant under A. So some care is needed when choosing the subspace Z. It follows readily from the definition that if a subspace Y reduces A, then the corresponding complementing subspace reduces A as well.

Let us illustrate the notion of a reducing subspace on a pair of examples.

Examples A.1.7. (i) Consider $X := \mathbb{C}^2$ and the matrix $A := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. It is easily seen that the subspaces

$$Y := \left\{ \begin{pmatrix} t \\ 0 \end{pmatrix} : t \in \mathbb{C} \right\}$$
$$Z := \left\{ \begin{pmatrix} 0 \\ t \end{pmatrix} : t \in \mathbb{C} \right\}$$

are invariant and that $X = Y \oplus Z$. Hence, these subspaces reduce the matrix A.

(ii) Consider $X := \mathbb{C}^2$ and the matrix $A := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. It is easily seen that the subspace

$$Y := \left\{ \begin{pmatrix} t \\ 0 \end{pmatrix} : t \in \mathbb{C} \right\}$$

is invariant. However, there is no other one-dimensional invariant subspace for A. Hence, Y is not a reducing subspace.

The existence of reducing subspaces is really convenient because it allows us to investigate the matrix on smaller subspaces. This is due to the fact that the matrix acts independently on each of its reducing subspaces. Reducing subspaces are of course closely related to projections.

Definition A.1.8. An operator $P \in \mathcal{L}(X)$ is called a *projection*, if $P^2 = P$.

Proposition A.1.9. *An operator $P \in \mathcal{L}(X)$ is a projection if and only if $X = \ker P \oplus \operatorname{im} P$ and $P|_{\operatorname{im} P} = I$.*

Proof. First note that $P|_{\operatorname{im} P} = I$ implies $P(Px) = Px$, i.e. $P^2 = P$.

Conversely, let $x \in X$ be chosen. Take $y \in \operatorname{im} P$ such that $y := Px$ and let $z := x - y$. Then obviously, $x = y + z$, $y \in \operatorname{im} P$ and $Pz = Px - Py = Px - P^2x = 0$. Hence, $z \in \ker P$. We can also deduce immediately from this line that if $y \in \operatorname{im} P$, then $Py = y$.

Finally, we only have to show that the obtained decomposition is unique. But if $x = y' + z'$ with $y' \in \operatorname{im} P$, $z' \in \ker P$, it follows that $Px = Py' + Pz' = y' = y$. Hence, $y' = y$ and $z' = z$. □

Exercise A.1.10. Consider the following matrices: $P_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $P_2 := \frac{1}{3}\begin{pmatrix} -1 & 2 \\ -2 & 4 \end{pmatrix}$. Show that they are projections. Can you find $\operatorname{im} P_i$ and $\ker P_i$ $(i = 1, 2)$? Can you give a geometric interpretation of the action of the matrices P_i?

Proposition A.1.11. *If $P \in M_n(\mathbb{C})$ is a projection, then $Q := I - P$ is also a projection. Further,* $\ker P = \operatorname{im} Q$ *and* $\operatorname{im} P = \ker Q$.

Proof. Clearly, using that P is a projection, we see that $Q^2 = (I - P)(I - P) = I - 2P + P^2 = I - P = Q$. Hence, Q is also a projection. We also see that

$$PQ = P(I - P) = P - P^2 = 0,$$
$$QP = (I - P)P = P - P^2 = 0.$$

Hence, $\operatorname{im} Q \subset \ker P$ and $\operatorname{im} P \subset \ker Q$. To show that here actually equality takes place, take $z \in \ker P$. Then $Qz = z - Pz = z$, which implies $z \in \operatorname{im} Q$. Similarly, if $y \in \ker Q$, then $Py = y - Qy = y$ implying $y \in \operatorname{im} P$. □

It is important to know that for every direct sum decomposition there is a corresponding projection.

Proposition A.1.12. *Assume that $Y, Z \subset X$ are subspaces such that $X = Y \oplus Z$. Then there is a unique projection $P \in M_n(\mathbb{C})$ such that* $\operatorname{im} P = Y$ *and* $\ker P = Z$.

Proof. Let us start from the decomposition $X = Y \oplus Z$. We define, in the spirit of Proposition A.1.9, $Px := y$ where $x = y + z$ is the unique decomposition with the property $y \in Y$ and $z \in Z$. Then for $y \in Y$, we obtain $Py = y$, for $z \in Z$ we see that $Pz = 0$. Hence, $P^2x = Py = y = Px$ meaning that P is really a projection. □

Finally, we close this summary with a connection between reducing subspaces and operators.

Proposition A.1.13. *Let $A \in \mathcal{L}(X)$ be given. The following are equivalent.*

(i) A subspace $Y \subset X$ reduces A with complementing reducing subspace $Z \subset X$.

(ii) The projection $P \in \mathcal{L}(X)$ with $Y = \operatorname{im} P$, $Z = \ker P$, commutes with A, i.e., $PA = AP$.

Proof. $(i) \Rightarrow (ii)$: If $X = Y \oplus Z$ with A-invariant subspaces Y and Z, then the projection P with range Y and kernel Z commutes with A. In fact, if $x \in X$ and $x = y + z$ is the unique decomposition of x into components $y \in Y$ and $z \in Z$, then

$$APx = Ay = PAy = PAy + PAz = PAx.$$

Here we used that for $y \in Y$, $Ay \in Y$ and hence $PAy = Ay$. Further, we also used that for $z \in Z$, $Az \in Z$, and hence $PAz = 0$.

$(ii) \Rightarrow (i)$: If $PA = AP$ and $\operatorname{im} P = PX =: Y$, then

$$AY = APX = PAX \subset PX = Y$$

and

$$A(I - P)X = (I - P)AX \subset (I - P)X = \ker P.$$

□

Let us recall here also a well-known property of linear mappings on finite dimensional spaces.

Proposition A.1.14. *For an operator $A \in \mathcal{L}(X)$ the following is equivalent*

$$A \text{ is injective} \iff A \text{ is surjective} \iff A \text{ is bijective.}$$

A.2 Complex functions and interpolation polynomials

We recall here some basic facts on the derivatives of polynomials, or, more generally, complex functions. We remind the reader that $\mathbb{C}$ is usually topologically identified with $\mathbb{R}^2$, hence, $\mathbb{C} \ni z_n \to z \in \mathbb{C}$ if and only if $\operatorname{Re} z_n \to \operatorname{Re} z$ and $\operatorname{Im} z_n \to \operatorname{Im} z$. This means that we can define neighborhoods and open sets just as we are used to it from multi-valued calculus. Let $U \subset \mathbb{C}$ be an open set, $f : U \to \mathbb{C}$, $z_0 \in U$. We call f *differentiable* at z_0, if the limit

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} =: f'(z_0)$$

exists. With the same proofs from elementary calculus we see that:

- If f, g are differentiable at z_0, then for $\alpha, \beta \in \mathbb{C}$ we have $(\alpha f + \beta g)'(z_0) = \alpha f'(z_0) + \beta g'(z_0)$.
- If $f(z) = z^m$, then f is differentiable and $f'(z) = m z^{m-1}$.
- If f, g are differentiable, then $(fg)' = f'g + fg'$.

Exercise A.2.1. Assume that the functions are at least m-times differentiable at a point z_0. Show that fg is also m-times differentiable and that

$$(fg)^{(m)}(z_0) = \sum_{i=0}^{m} \binom{m}{i} f^{(m-i)}(z_0) g^{(i)}(z_0). \tag{A.2}$$

Lemma A.2.2. *Let $z_1 \in \mathbb{C}$ be fixed. A polynomial p has the form $p(z) = (z - z_1)^m q(z)$ for some $m \in \mathbb{N}$ and $q \in \mathbb{C}[x]$, if and only if*

$$p^{(i)}(z_1) = 0 \qquad \text{for all } i = 0, 1, \ldots, m-1.$$

Proof. We can proceed by induction. For $m = 1$, the statement is well-known from elementary algebra. Assume now that we know the desired identity up to some number $m = k$ and want to know what happens for $m = k + 1$. From the induction assumption we see that $p(z) = (z - z_1)^k q_1(z)$ for some $q_1 \in \mathbb{C}[x]$. From Equation (A.2) we see that

$$p^{(k)}(z) = \sum_{i=0}^{k} \binom{k}{i} \frac{k!}{i!} (z - z_1)^i q_1^{(i)}(z) = (z - z_1) p_1(z) + k! q_1(z), \tag{A.3}$$

where p_1 is a suitable chosen polynomial, more precisely,

$$p_1(z) = \sum_{i=1}^{k} \binom{k}{i} \frac{k!}{i!} (z - z_1)^{i-1} q_1^{(i)}(z).$$

From Equation (A.3) we see that $p^{(k)}(z_1) = 0$ if and only if $q_1(z_1) = 0$. Hence, using again the well-known fact from elementary algebra, we see that this last is equivalent to the fact that

$$q_1(z) = (z - z_1) q(z)$$

for some polynomial $q \in \mathbb{C}[x]$ proving the assertion. □

Now we collect some information on interpolation polynomials. Given m distinct complex numbers $z_1, \ldots, z_m$ and another m complex numbers $w_1, \ldots, w_m$, one easily finds a polynomial p of degree $m - 1$ with $p(z_i) = w_i$, $i = 1, \ldots, m$:

$$p(z) := \sum_{k=1}^{m} w_k \prod_{\substack{1 \le j \le m \\ j \ne k}} \frac{z - z_j}{z_k - z_j},$$

known as the *Langrange polynomial.* The problem becomes more difficult if we would also like to prescribe the values of the derivatives of p at points $z_1, \ldots, z_m$. We obviously need higher order polynomials to fulfill this task. If for any z_i the value of p and its first $\nu_i - 1$ derivatives are given, $i = 1, \ldots, m$, then the minimal degree of the appropriate interpolation polynomial is $\nu_1 + \cdots + \nu_m - 1$ and can be achieved using *Hermite interpolation.* For more information we refer to [1, Section 5.2] and [2, Example 7.9.4].

Exercise A.2.3. Find some polynomials p with the following properties.

(i) $p(1) = 1$, $p(2) = 1$.

(ii) $p(1) = 1$, $p'(1) = 2$.

(iii) $p(1) = 1$, $p'(1) = 3$, $p(-1) = -1$.

(iv) $p(1) = 1$, $p'(1) = 3$, $p(-1) = -1$, $p'(-1) = 3$.

(v) $p(1) = 1$, $p'(1) = 3$, $p(-1) = -1$, $p'(-1) = 2$.

(vi) $p(1) = 1$, $p'(1) = 2$, $p''(1) = 3$.

Bibliography

[1] P. Lancaster, *Theory of Matrices*, Academic Press New York, 1969.

[2] C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[3] A. Tikhonov, *Ein Fixpunktsatz*, Math. Ann. **111** (1935), 767–776.

Appendix B

B.1 Functional analysis

This manuscript is written in a functional-analytic spirit. Its main objects are operators on Banach spaces, and from now on we use many results and techniques from functional analysis and operator theory. There are many excellent sources and we refer to textbooks like [5], [6], [11], [12], [13], [14], or [16]. However, for convenience we add this appendix, where we introduce our notation and list some results.

To start with, we introduce the following classical sequence and function spaces. Here, J is a real interval; $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$; and Ω, depending on the context, is a domain in $\mathbb{R}^n$, a locally compact metric space, or a measure space. The symbol X always stands for a Banach space. The canonical sequence spaces are the following:

$$\begin{aligned}
\ell^\infty(X) = \ell^\infty(\mathbb{N}, X) &:= \Big\{(x_n)_{n\in\mathbb{N}} \subset X : \sup_{n\in\mathbb{N}} \|x_n\| < \infty\Big\}, \|(x_n)_n\| := \sup_{n\in\mathbb{N}} \|x_n\|, \\
c(X) = c(\mathbb{N}, X) &:= \Big\{(x_n)_{n\in\mathbb{N}} \subset X : \lim_{n\to\infty} x_n \text{ exists}\Big\} \subset \ell^\infty(X), \\
c_0(X) = c_0(\mathbb{N}, X) &:= \Big\{(x_n)_{n\in\mathbb{N}} \subset X : \lim_{n\to\infty} x_n = 0\Big\} \subset c(X), \\
\ell^\infty = \ell^\infty(\mathbb{N}) &:= \ell^\infty(\mathbb{N}, \mathbb{C}), \\
c = c(\mathbb{N}) &:= c(\mathbb{N}, \mathbb{C}), \\
c_0 = c_0(\mathbb{N}) &:= c_0(\mathbb{N}, \mathbb{C}), \\
\ell^p = \ell^p(\mathbb{N}) &:= \ell^p(\mathbb{N}, \mathbb{C}) := \Big\{(x_n)_{n\in\mathbb{N}} \subset \mathbb{C} : \sum_{n\in\mathbb{N}} |x_n|^p < \infty\Big\}, \ \|(x_n)_n\| := \Big(\sum_{n\in\mathbb{N}} |x_n|^p\Big)^{\frac{1}{p}} \\
&\quad (p \in [1, \infty)).
\end{aligned}$$

Spaces of continuous functions are:

$$
\begin{aligned}
\mathrm{C}(K) &:= \{f : \Omega \to \mathbb{K} \mid f \text{ is continuous}\},\ \|f\|_\infty := \sup_{s\in K} |f(s)| \quad \text{(if } K \text{ is compact)},\\
\mathrm{C}_0(\Omega) &:= \{f \in \mathrm{C}(\Omega) : f \text{ vanishes at infinity}\},\\
\mathrm{C}_\mathrm{b}(\Omega) &:= \{f \in \mathrm{C}(\Omega) : f \text{ is bounded}\},\\
\mathrm{C}_\mathrm{c}(\Omega) &:= \{f \in \mathrm{C}(\Omega) : f \text{ has compact support}\},\\
\mathrm{C}_\mathrm{ub}(\Omega) &:= \{f \in \mathrm{C}(\Omega) : f \text{ is bounded and uniformly continuous}\},\\
\mathrm{AC}(J) &:= \{f : J \to \mathbb{K} \mid f \text{ is absolutely continuous}\},\\
\mathrm{C}^k(J) &:= \{f \in \mathrm{C}(J) : f \text{ is } k\text{-times continuously differentiable}\},\\
\mathrm{C}^\alpha(J) &:= \{f \in \mathrm{C}(J) : f \text{ is Hölder continuous of order } \alpha\},\\
\mathrm{C}^\infty(J) &:= \{f \in \mathrm{C}(J) : f \text{ is infinitely many times differentiable}\},\\
\mathrm{Lip}_\mathrm{u}(J) &:= \{f \in \mathrm{C}_\mathrm{ub}(\Omega) : f \text{ is Lipschitz continuous}\},\\
&\qquad \|f\|_\mathrm{Lip} := |f(0)| + \sup_{r\neq s} \left|\frac{f(r)-f(s)}{r-s}\right|.
\end{aligned}
$$

Spaces of integrable functions are:

$$
\begin{aligned}
\mathrm{L}^p(\Omega,\mu) &:= \{f : \Omega \to \mathbb{K} \mid f \text{ is } p\text{-integrable on } \Omega\},\ \|f\|_p := \Big(\int_\Omega |f|^p(s)\, d\mu(s)\Big)^{\frac{1}{p}},\\
\mathrm{L}^\infty(\Omega,\mu) &:= \{f : \Omega \to \mathbb{K} \mid f \text{ is measurable and } \mu\text{-essentially bounded}\},\ \|f\|_\infty := \operatorname{essup}|f|;\\
\mathrm{W}^{k,p}(\Omega) &:= \big\{f \in \mathrm{L}^p(\Omega) \,:\, f \text{ is } k\text{-times distributionally differentiable}\\
&\qquad\qquad \text{with } D^\alpha f \in \mathrm{L}^p(\Omega) \text{ for all } |\alpha| \le k\big\}.
\end{aligned}
$$

The strong operator topology

At this point, we do not want to give the definition of the strong operator topology, but just point out what *convergence* and *boundedness* mean in this setting.

Let X, Y be Banach spaces and let $(T_n) \subseteq \mathcal{L}(X,Y)$ be a sequence of bounded linear operators between X and Y. We say that the sequence (T_n) *converges strongly* to $T \in \mathcal{L}(X,Y)$, if

$$T_n x \to Tx \quad \text{holds in } Y \text{ as } n \to \infty \text{ for all } x \in X.$$

For the purposes of this course, this is the correct notion of convergence, being, as a matter of fact, nothing else than pointwise convergence.

A subset $\mathcal{K} \subseteq \mathcal{L}(X,Y)$ is called *strongly bounded* (or bounded poinwise) if for all $x \in X$ we have

$$\sup\big\{\|Tx\| \,:\, T \in \mathcal{K}\big\} < \infty.$$

Next, we list some classical functional analysis results concerning these two notions.

Theorem B.1.1 (Uniform Boundedness Principle)**.** *Let X, Y be Banach spaces and suppose $\mathcal{K} \subseteq \mathcal{L}(X, Y)$ is strongly bounded, i.e., for all $x \in X$ we have*

$$\sup\big\{\|Tx\| \,:\, T \in \mathcal{K}\big\} < \infty.$$

Then $\mathcal{K}$ is uniformly bounded*, that is*

$$\sup\big\{\|T\| \,:\, T \in \mathcal{K}\big\} < \infty.$$

This theorem has the following important consequence:

Theorem B.1.2. *Let X, Y be Banach spaces, and let $(T_n) \subseteq \mathcal{L}(X, Y)$ be a sequence such that $(T_n x) \subseteq Y$ converges for all $x \in X$. Then*

$$Tx := \lim_{n\to\infty} T_n x$$

defines a bounded linear operator on X.

Theorem B.1.3. *Let X, Y be Banach spaces, let $T \in \mathcal{L}(X, Y)$ and let $(T_n) \subseteq \mathcal{L}(X, Y)$ be a norm bounded sequence. Then the following assertions are equivalent:*

1. *For every $x \in X$ we have $T_n x \to Tx$ in X.*
2. *There is a dense subspace $D \subseteq X$ such that for all $x \in D$ we have $T_n x \to Tx$ in X.*
3. *For every compact set $K \subseteq X$ we have $T_n x \to Tx$ in X uniformly for $x \in K$.*

By adapting the classical proof of the product rule of differentiation and by making use of the theorem above one can easily prove the next result.

Theorem B.1.4 (Product rule)**.** *Let $u : [a, b] \to X$ be differentiable, and let $F : [a, b] \to \mathcal{L}(X, Y)$ be strongly continuous such that for every $t \in [a, b]$ the mapping*

$$Fu : s \mapsto F(s)u(t) \in Y$$

is differentiable. Then $s \mapsto F(s)u(s) \in Y$ is differentiable, and we have

$$(Fu)'(t) = F'(t) \cdot u(t) + F(t) \cdot u'(t).$$

An other important result we wish to recall from functional analysis is the closed graph theorem.

Theorem B.1.5 (Closed Graph Theorem)**.** *Let X be a Banach space, and let $A : X \to Y$ be a linear operator such that its graph is a closed subspace of $X \times X$ with dense domain $D(A)$ in X. Then A is bounded if and only if $D(A) = X$.*

The Hahn–Banach Theorem

Let X be a Banach space. A linear functional $\varphi : X \to \mathbb{C}$ is called *bounded* if there is a constant $M \geq 0$ such that

$$\|\varphi(f)\| \leq M\|f\| \quad \text{for all } f \in X.$$

The set

$$X' := \{\varphi \,:\, \varphi \text{ is a bounded linear functional on } X\}$$

of all bounded linear functionals is a linear space, and becomes a Banach space with the *functional norm*

$$\|\varphi\| := \sup_{\substack{f\in X\\ \|f\|\leq 1}} |\varphi(f)| = \sup_{\substack{f\in X\\ \|f\|\leq 1}} |\langle f, \varphi\rangle|.$$

Here we used the convenient notation $\varphi(f) = \langle f, \varphi\rangle$. If $\varphi \in X'$ then

$$|\langle f, \varphi\rangle| \leq \|\varphi\| \cdot \|f\|$$

holds for all $f \in X$. The space X' is called the *dual space* of X. That X' is large enough for every Banach space is highly non-trivial, and is actually the statement of the Hahn–Banach theorem (see [9] and [4]). Note however that in specific examples the dual space can be determined.

Theorem B.1.6 (Hahn–Banach)**.** *Let X be a Banach space, and let X' be its dual space. Then the following assertions are true:*

(i) For $f \in X$, $f \neq 0$, there is $\varphi \in X'$ with $\varphi(f) = \|f\|$ and $\|\varphi\| = 1$. Or, which is the same, for every $0 \neq f \in X$ there is $\varphi \in X'$ with $\varphi(f) = \|f\|^2 = \|\varphi\|^2$.

(ii) For $f, g \in X$ one has $f = g$ if and only if $\langle f, \varphi\rangle = \langle g, \varphi\rangle$ for all $\varphi \in X'$.

(iii) A subspace Y is dense in X if and only the zero functional is the only bounded linear functional that vanishes on Y.

The Banach–Alaoglu Theorem

Let $\varphi_n, \varphi \in X'$. We call φ_n *weak*-convergent* to φ if for all $f \in X$

$$\langle f, \varphi_n - \varphi\rangle \to 0 \quad \text{holds as } n \to \infty.$$

The functional φ is called the *weak*-limit* of the sequence, and if exists, then it is obviously unique. We call φ a *weak*-accumulation* point of the sequence (φ_n) if for all $f \in X$ and $\varepsilon > 0$ there is a subsequence (φ_{n_k}) with

$$|\langle f, \varphi_{n_k} - \varphi\rangle| \leq \varepsilon \quad \text{for all } k \in \mathbb{N}.$$

Obviously, if (φ_n) has a weak*-convergent subsequence φ, then φ is an accumulation point of the sequence. The converse implication is in general not true. The next rather weak formulation of a central result from functional analysis suffices for our purposes.

Theorem B.1.7 (Banach–Alaoglu, [1]). *Let X be a Banach space and consider its dual space. Let*

$$\mathcal{B}' := \{\varphi \in X' : \|\varphi\| \le 1\} \subseteq X'$$

be the unit ball in X'. Then every sequence $(\varphi_n) \subseteq \mathcal{B}'$ has a weak-accumulation point in $\mathcal{B}'$. If X is reflexive or separable, then every sequence $(\varphi_n) \subseteq \mathcal{B}'$ has a weak*-convergent subsequence with limit in $\mathcal{B}'$.*

The Riemann integral

Denote by $\mathrm{C}([a,b];X)$ the space of continuous X-valued functions on $[a,b]$, which becomes a Banach space with the supremum norm. For a continuous function $u \in \mathrm{C}([a,b];X)$ we define its *Riemann integral* by approximation through Riemann sums. Let us briefly sketch the idea how to do this. For $P = \{a = t_1 < t_2 < \cdots < t_n = b\} \subseteq [a,b]$ we set

$$\delta(P) = \max\{t_{j+1} - t_j : j = 0, \dots, n-1\},$$

and call P a *partition* of $[a,b]$ and $\delta(P)$ the *mesh* of P. We define the *Riemann sum* of u corresponding to the partition P by

$$S(P,u) := \sum_{j=0}^{n-1} u(t_j)(t_{j+1} - t_j),$$

where n is the number of elements in P. From the uniform continuity of u on the compact interval $[a,b]$ it follows that there exists $x_0 \in X$ such that $S(P,u)$ converges to x_0 if $\delta(P) \to 0$. More precisely, for all $\varepsilon > 0$ there is $\delta > 0$ such that

$$\|S(P,u) - x_0\| < \varepsilon$$

whenever $\delta(P) < \delta$. We call this $x_0 \in X$ the Riemann integral of f an denote it by

$$\int_a^b u(s)\, ds.$$

The Riemann integral enjoys all the usual properties known for scalar valued functions. Some of them are collected in the next proposition.

Proposition B.1.8.
1. *There is a sequence of Riemann sums $S(P_n, u)$ with $\delta(P_n) \to 0$ converging to the Riemann integral of u.*
2. *The Riemann integral is a bounded linear operator on the space $C([a,b];X)$ with values in X.*
3. *If $T \in \mathcal{L}(X,Y)$, then*

$$T \int_a^b u(s)\, ds = \int_a^b Tu(s)\, ds.$$

4. If $u : [a, b] \to X$ *is continuous, then*

$$v(t) := \int_0^t u(s)\, ds$$

is differentiable with derivative u*.*

5. If $u : [a, b] \to X$ *is continuously differentiable, then*

$$u(b) - u(a) = \int_a^b u'(s)\, ds$$

holds.

For the proof of these assertions one can take the standard route valid for scalar-valued functions.

Spectrum and resolvent

Definition B.1.9. Let A be (not necessarily bounded) linear operator defined on a linear subspace $D(A)$ of a Banach space X.

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

$$\sigma(A) := \{\lambda \in \mathbb{C} : \lambda - A : D(A) \to X \text{ is not bijective or the inverse is not continuous}\}.$$

(ii) The *resolvent set* of A is $\rho(A) := \mathbb{C} \setminus \sigma(A)$, i.e.,

$$\rho(A) := \{\lambda \in \mathbb{C} : \lambda - A : D(A) \to X \text{ is bijective with continuous inverse}\}.$$

(iii) If $\lambda \in \rho(A)$ then the operator $\lambda - A$ is injective, hence has an algebraic inverse $(\lambda - A)^{-1}$. We call this operator the resolvent of A at point λ and denote it by

$$R(\lambda, A) := (\lambda - A)^{-1}.$$

Note that if $\lambda \in \rho(A)$, the operator $\lambda - A$ is both injective and surjective, i.e., its algebraic inverse

$$(\lambda - A)^{-1} : X \to D(A)$$

is defined on the entire X.

Let us recall also the next fundamental properties of spectrum and the resolvent.

Proposition B.1.10. *Let* X *be a Banach space and let* A *be a linear operator with domain* $D(A) \subseteq X$*. Then the following assertions are true:*

(a) The resolvent set $\rho(A)$ *is open, hence its complement, the spectrum* $\sigma(A)$ *is closed.*

(b) *The mapping*

$$\rho(A) \ni \lambda \mapsto R(\lambda, A) \in \mathcal{L}(X)$$

is complex differentiable. Moreover, for $n \in \mathbb{N}$ we have

$$\frac{d^n}{d\lambda^n} R(\lambda, A) = (-1)^n n! R(\lambda, A)^{n+1}.$$

(c) *If $A \in \mathcal{L}(X)$, then for every $\lambda \in \mathbb{C}$ with $|\lambda| > r(A)$ we have $\lambda \in \rho(A)$.*

(d) *Let $\lambda_n \in \rho(A)$ with $\lim_{n\to\infty} \lambda_n = \lambda_0$. Then $\lambda_0 \in \sigma(A)$ if and only if*

$$\lim_{n\to\infty} \|R(\lambda_n, A)\| = \infty.$$

Proof. Statement (a) follows from the following Neumann series representation of the resolvent: For $\mu \in \rho(A)$ and $\lambda \in \mathbb{C}$ with $|\lambda - \mu| < \frac{1}{\|R(\mu,A)\|}$, we have $\lambda \in \rho(A)$ and

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

Assertion (b) follows from the above power series representation and from the fact that a power series is always a Taylor series.

Assertion (c) follows by similar Neumann series arguments as (a) since

$$(\lambda - A) = \lambda \left(I - \frac{A}{\lambda} \right),$$

formally

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

and the series converges if $|\lambda| > \limsup \|A^n\|^{1/n} =: r(A)$.

To show (d) first assume that $\lambda_0 \in \rho(A)$. The resolvent map is continuous and remains bounded on the compact set $\{\lambda_n : n \geq 0\}$, which contradicts the assertion on the limit of $\|R(\lambda_n, A)\|$, hence $\lambda_0 \in \sigma(A)$. For the converse implication observe that our considerations at the beginning of the proof yield $\|R(\mu, A)\| \geq \frac{1}{\operatorname{dist}(\mu, \sigma(A))}$ for all $\mu \in \rho(A)$ (see also Corollary B.1.12). □

The following result is known as the spectral mapping theorem for the resolvent.

Proposition B.1.11. *Let X be a Banach space and let A be a linear operator with domain $D(A) \subseteq X$ such that $\rho(A) \neq \emptyset$. Then, for any $\lambda \in \rho(A)$,*

$$\sigma(R(\lambda, A)) \setminus \{0\} = \left\{ \frac{1}{\lambda - \mu} : \mu \in \sigma(A) \right\}.$$

Proof. For $0 \neq \alpha \in \mathbb{C}$ and $\lambda \in \rho(A)$, we have

$$\begin{aligned}(\alpha - R(\lambda, A))x &= \alpha \left[\left(\lambda - \frac{1}{\alpha} \right) - A \right] R(\lambda, A)x, \quad \forall x \in X, \\ &= \alpha R(\lambda, A) \left[\left(\lambda - \frac{1}{\alpha} \right) - A \right] x, \quad \forall x \in D(A).\end{aligned}$$

This implies that $\alpha \in \sigma(R(\lambda, A))$ if and only if $\lambda - \frac{1}{\alpha} \in \sigma(A)$. □

As a corollary we obtain

Corollary B.1.12. *For any $\lambda \in \rho(A)$ the following holds:*

$$\operatorname{dist}(\lambda, \sigma(A)) = \frac{1}{r(R(\lambda, A))} \geq \frac{1}{\|R(\lambda, A)\|}.$$

Proof. For $\lambda \in \rho(A)$, it follows from the above proposition that

$$\begin{aligned}\operatorname{dist}(\lambda, \sigma(A)) &= \inf\{|\lambda - \mu| : \mu \in \sigma(A)\} \\ &= \left(\sup \left\{ \left| \frac{1}{\lambda - \mu} \right| : \mu \in \sigma(A) \right\} \right)^{-1} \\ &= (\max\{|\alpha| : \alpha \in \sigma(R(\lambda, A))\})^{-1} \\ &= \frac{1}{r(R(\lambda, A))} \geq \frac{1}{\|R(\lambda, A)\|}.\end{aligned}$$

□

Let us now recall the following spectral mapping theorems (see [6, Theorem VII.3.11] and [7, Theorem IV.3.7]).

Theorem B.1.13. *For $A \in \mathcal{L}(X)$ one has*

$$\sigma(e^{tA}) = e^{t\sigma(A)} = \{e^{t\lambda} : \lambda \in \sigma(A)\}.$$

Theorem B.1.14. *Let $(T(t))_{t \geq 0}$ be a C_0-semigroup with generator A on Banach space X. Then*

$$\sigma_p(T(t)) \setminus \{0\} = e^{t\sigma_p(A)}, \quad t \geq 0.$$

Here $\sigma_p(A) := \{\lambda \in \mathbb{C} : (\lambda - A) \text{ is not injective } \}$ denotes the *point spectrum* of A.

We end this subsection by characterizing the boundary spectrum

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

of the generator A of an irreducible C_0-semigroup $(T(t))_{t \geq 0}$ on a Banach lattice E. For the proof see [7, p. 315].

Lemma B.1.15. *Let A be the generator of an irreducible C_0-semigroup $(T(t))_{t \geq 0}$ on a Banach lattice E. If $s(A)$ is a pole of $R(\cdot, A)$ then $\sigma_b(A) \subset \sigma_p(A)$.*

Isolated singularities

Let A be a closed linear operator on a Banach space X. Assume that μ is isolated in $\sigma(A)$. Then the complex differentiable function $\lambda \mapsto R(\lambda, A)$ can be expanded as a Laurent series

$$R(\lambda, A) = \sum_{n=-\infty}^{\infty} (\lambda - \mu)^n R_n \quad \text{for } 0 < |\lambda - \mu| < \eta$$

with a sufficiently small $\eta > 0$, where

$$R_n := \frac{1}{2\pi i} \int_\gamma \frac{R(\lambda, A)}{(\lambda - \mu)^{n+1}}\, d\lambda, \quad n \in \mathbb{Z}. \tag{B.1}$$

Here γ is the positively oriented boundary of the disc centered at μ and radius $\eta/2$. The coefficient R_{-1} is the spectral projection P_μ corresponding to the decomposition $\sigma(A) = \{\mu\} \cup (\sigma(A) \setminus \{\mu\})$ and is called the *residue* of $R(\cdot, A)$ at η. It follows from (B.1) that

$$R_{-(n+1)} = (A - \mu)^n P_\mu, \quad n \in \mathbb{N}.$$

If there is $k \in \mathbb{N}$ such that $R_{-k} \neq 0$ and $R_{-(k+1)} = 0$, then μ is called a *pole* of $R(\cdot, A)$ of order k. In this case one obtains

$$R_{-k} = \lim_{\lambda \to \mu^+} (\lambda - \mu)^k R(\lambda, A). \tag{B.2}$$

We call $m_a := \dim rg\, P_\mu$ the *algebraic multiplicity* and $m_g := \dim ker(\mu - A)$ the *geometric multiplicity* of μ. If $m_a = 1$ we call μ an algebraically simple (or first order) pole of $R(\cdot, A)$. It can be proved that if μ is a pole of order $k \in \mathbb{N}$ then

Proposition B.1.16. *(i)* $m_g + k - 1 \leq m_a \leq k m_g$, *and*

(ii) $\mu \in \sigma_p(A)$ *and* $rg\, P_\mu = ker\,(\mu - A)^k$.

For proofs of these properties we refer to [8, Chap. II], [10, III.5], or [14, V.10].

The essential spectrum

We start with some definitions. A bounded operator $S \in \mathcal{L}(E)$ is called a *Fredholm operator* if there is $T \in \mathcal{L}(E)$ such that $I - TS$ and $I - ST$ are compact. We denote by

$$\sigma_{ess}(S) = \mathbb{C} \setminus \rho_F(S)$$

the *essential spectrum* of S, where

$$\rho_F(S) := \{\lambda \in \mathbb{C} : (\lambda - S) \text{ is a Fredholm operator}\}.$$

The *Calkin algebra* $\mathcal{C}(E) := \mathcal{L}(E)/\mathcal{K}(E)$ equipped with the quotient norm

$$\|S\|_{ess} := \|S + \mathcal{K}(E)\| = \operatorname{dist}(S, \mathcal{K}(E)) = \inf\{\|S - K\| : K \in \mathcal{K}(E)\}$$

is a Banach algebra with unit. The essential spectrum of $S \in \mathcal{L}(E)$ can also be defined as the spectrum of $S + \mathcal{K}(E)$ in the Banach algebra $\mathcal{C}(E)$. This implies that, for $S \in \mathcal{L}(E)$, $\sigma_{ess}(S)$ is non-empty and compact.

For $S \in \mathcal{L}(E)$ we define the *essential spectral radius* by

$$r_{ess}(S) := r(S + \mathcal{K}(E)) = \max\{|\lambda| : \lambda \in \sigma_{ess}(S)\}.$$

Since $(S + \mathcal{K}(E))^n = S^n + \mathcal{K}(E)$ for $n \in \mathbb{N}$, we have $r_{ess}(S) = \lim_{n\to\infty} \|S^n\|_{ess}^{\frac{1}{n}}$ and consequently,

$$r_{ess}(s + K) = r_{ess}(S), \quad \text{for every}\, K \in \mathcal{K}(E).$$

If we denote by

$$Pol(S) := \{\lambda \in \mathbb{C} : \lambda \quad \text{is a pole of finite algebraic multiplicity of}\, R(\cdot, S)\},$$

then one can prove that $Pol(S) \subseteq \rho_F(S)$ and an element of the unbounded connected component of $\rho_F(S)$ either is in $\rho(S)$ or a pole of finite algebraic multiplicity. For details concerning the essential spectrum we refer to [10, IV.5.6], [8, Chap. XVII]. Thus we obtain the following characterization.

Proposition B.1.17. *For $S \in \mathcal{L}(E)$ the essential spectral radius is given by*

$$r_{ess}(S) = \inf\{r > 0 : \lambda \in \sigma(S), |\lambda| > r \;\; \textit{and} \;\; \lambda \in \; Pol(S)\}.$$

Proof. If we set

$$a := \inf\{r > 0 : \lambda \in \sigma(S), |\lambda| > r \,\text{and}\, \lambda \in \; Pol(S)\},$$

then for all $\varepsilon > 0$ there is $r_\varepsilon > 0$ such that

$$\{\lambda \in \sigma(S) : |\lambda| > r_\varepsilon\} \subseteq Pol(S)$$

and $r_\varepsilon - \varepsilon \leq a$. On the other hand, we know that there is $\lambda_0 \in \sigma_{ess}(S)$ with $r_{ess}(S) = |\lambda_0|$. If we suppose that $r_{ess}(S) > r_\varepsilon$, then $\lambda_0 \in Pol(S)$. This implies that $\lambda_0 \in \rho_F(S)$, which is a contradiction. Hence, $r_{ess}(S) \leq r_\varepsilon \leq a + \varepsilon$. Thus, $r_{ess}(S) \leq a$.

To show the other inequality first note that

$$\{\lambda \in \sigma(S) : |\lambda| > r_{ess}(S)\} \subseteq \rho_F(S).$$

Therefore,

$$\{\lambda \in \sigma(S) : |\lambda| > r_{ess}(S)\} \subseteq Pol(S).$$

Consequently, $a \leq r_{ess}(S)$ and the proposition in proved. □

We define the *essential growth bound* $\omega_{ess}(T)$ of a C_0-semigroup $(T(t))_{t\geq 0}$ with generator A ad the growth bound of the quotient semigroup $T(\cdot) + \mathcal{K}(E)$ on $\mathcal{C}(E)$, i.e.,

$$\omega_{ess}(T) := \inf\{\omega \in \mathbb{R} : \exists M > 0 \,\text{such that}\, \|T(t)\|_{ess} \leq Me^{\omega t}, \forall\, t \geq 0\}.$$

Then, for all $t_0 > 0$, one can see that

$$\omega_{ess}(T) = \frac{\log r_{ess}(T(t_0))}{t_0} = \lim_{t\to\infty} \frac{\log \|T(t)\|_{ess}}{t}. \tag{B.3}$$

The following result gives the relationship between $\omega_{ess}(T)$ and $\omega_0(T)$.

Proposition B.1.18. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup with generator A on a Banach space E. Then one has*

$$\omega_0(T) = \max\{s(A), \omega_{ess}(T)\}.$$

Proof. If $\omega_{ess}(T) < \omega_0(T)$, then $r_{ess}(T(1)) < r(T(1))$. Let $\lambda \in \sigma(T(1))$ such that $|\lambda| = r(T(1))$. So by Proposition B.1.17, λ is an eigenvalue of $T(1)$ and by the spectral mapping theorem for the point spectrum, Theorem B.1.14, there is $\lambda_1 \in \sigma_p(A)$ with $e^{\lambda_1} = \lambda$. Therefore, $\operatorname{Re}\lambda_1 = \omega_0(T)$ and thus $\omega_0(T) = s(A)$. □

By using the essential growth bound one can deduce important consequences for the asymptotic behaviour, cf. [7, Theorem V.3.7].

Theorem B.1.19. *Let A be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach space X such that $\omega_{ess}(T) < 0$. Then the following assertions hold.*

(a) The set $\{\lambda \in \sigma(A) : \operatorname{Re}\lambda \geq 0\}$ is finite (or empty) and consists of poles of $R(\cdot, A)$ of finite algebraic multiplicity.
Denoting these poles by $\lambda_1, \dots, \lambda_m$, the corresponding spectral projections $P_1, \dots, P_m$ and the order of the poles $k_1, \dots, k_m$, we have

(b) $T(t) = T_1(t) + \ldots + T_m(t) + R(t)$, where

$$T_n(t) := e^{\lambda_n t} \sum_{j=0}^{k_n-1} \frac{t^j}{j!}(A-\lambda_n)^j P_n, \quad n = 1, \dots, m,$$

and

$$\|R(t)\| \leq M e^{-\varepsilon t} \quad \text{for some } \varepsilon > 0, M \geq 1 \text{ and all } t \geq 0.$$

We end this subsection by recalling a perturbation theorem for the essential spectral radius of C_0-semigroups due to J. Voigt [15]. To this purpose let us recall some definitions. An operator $B \in \mathcal{L}(E)$ is called *strictly power compact* if there is $n \in \mathbb{N}$ such that $(BT)^n$ is compact for all $T \in \mathcal{L}(E)$. An operator $T \in \mathcal{L}(E)$ is said to be *weakly compact* if for every norm bounded sequence (f_n) of E the sequence (Tf_n) has a weakly convergent subsequence in E.

Proposition B.1.20. *If E is an L^1-space then the product of two weakly compact operators is compact. In particular, every weakly compact operator is strictly power compact.*

See [6, Corollary VI.8.13]. For the following result we refer for example to [2, Theorem 5.25 and Theorem 5.31].

Proposition B.1.21. *Let (Ω, Σ, μ) be a $\sigma-$finite, positive measure space and S, T be two bounded linear operators on $L^1(\Omega, \mu)$. Then the following assertions hold.*

(a) The set of all weakly compact operators is a norm-closed subset of $\mathcal{L}(L^1(\Omega, \mu))$.

(b) If T is weakly compact and $0 \leq S \leq T$, then S is also weakly compact.

The following theorem is due to B. de Pagter.

Theorem B.1.22. *Let E be a Banach lattice. If $0 \leq T \in \mathcal{L}(E)$ is an irreducible compact operator, then $r(T) > 0$.*

The following theorem gives the relationship between the essential spectrum of the perturbed and the unperturbed semigroups (see [15]).

Theorem B.1.23. *Let A be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach space E and $B \in \mathcal{L}(E)$. Let $(S(t))_{t\geq 0}$ the C_0-semigroup generated by $A + B$. Assume that there exists $n \in \mathbb{N}$ and a sequence $(t_k) \subset \mathbb{R}_+$, $t_k \to \infty$, such that the remainder $R_n(t_k) := \sum_{p=n} S_p(t_k)$ of the Dyson-Phillips expansion(15.1) at t_k is strictly power compact for all $k \in \mathbb{N}$. Then*

$$r_{ess}(S(t)) \leq r_{ess}(T(t)), \quad t \geq 0.$$

Vector valued Laplace and Fourier transforms

The general reference for the following facts is the monograph by Arendt, Batty, Hieber, Neubrander [3].

Let X be a (complex) Banach space and $I \subset \mathbb{R}$ an interval. A function $F : I \to X$ is said to be a *step function*, if there is $n \in \mathbb{N}$ and $I_i \subset I$ Lebesgue measurable sets and $x_i \in X$ for all $i = 1, \dots, n$ such that

$$F(t) = \sum_{i=1}^{n} f_i \mathbb{1}_{I_i}.$$

A function $F : I \to X$ is called *measurable*, if there is a sequence $F_n : I \to X$ of step functions such that

$$F(t) = \lim_{n\to\infty} F_n(t)$$

for almost every $t \in I$. Here "almost every" is to be understood in the sense on Lebesgue measure in I.

We say that $F \in \mathrm{L}^p(I, X)$, if $F : I \to X$ is measurable and $\|F(\cdot)\| \in \mathrm{L}^p(I, \mathbb{R})$ for $p \in [1, \infty)$, and define

$$\|F\|_p := \left(\int_I \|F(t)\|^p \, dt \right)^{\frac{1}{p}}.$$

Then, using more or less straightforward arguments, one can show that $\mathrm{L}^p(I, X)$ enjoys similar properties to the scalar valued Lebesgue spaces.

Especially, if H is a Hilbert space, then $\mathrm{L}^2(I, H)$ becomes a Hilbert space with the scalar product

$$(F|G) := \int_I (F(t)|G(t))\, dt.$$

The integral of a step function F is defined as

$$\int_I F := \sum_{i=0}^{n} F_i \lambda(I_i).$$

A function $F : I \to X$ is said to be *Bochner integrable*, if there is a sequence of step functions $F_n : I \to X$ such that $\int_I \|F(t) - F_n(t)\|\, dt \to 0$ as $n \to \infty$. The integral of a Bochner integrable function can be defined then as

$$\int_I F := \lim_{n\to\infty} \int_I F_n.$$

It can be shown that the integral is well-defined and that if F is Bochner integrable, then $F \in \mathrm{L}^1(I, X)$ and

$$\left\| \int_I F \right\| \le \int \|F(t)\|\, dt.$$

The *Fourier transform* of a function $F \in \mathrm{L}^1(\mathbb{R}, X)$ can be defined as

$$\mathcal{F}F(s) := \int_{-\infty}^{\infty} e^{-ist} F(t)\ dt$$

and it can be shown (for example, by taking scalar products and reducing to the scalar case) that if H is a Hilbert space, then for $F \in \mathrm{L}^1(\mathbb{R}, H) \cap \mathrm{L}^1(\mathbb{R}, H)$, we have

$$\int_{\mathbb{R}} \|\mathcal{F}(F)(t)\|^2\, dt = 2\pi \int_{\mathbb{R}} \|F(t)|^2\, dt,$$

yielding to the vector valued version of *Plancherel's Theorem*.

Theorem B.1.24. *The Fourier thansform extends uniquely to a bounded linear operator on $L^2(\mathbb{R}, H)$, and the operator $\frac{1}{\sqrt{2\pi}}\mathcal{F}$ is unitary.*

If $F : \mathbb{R}_+ \to X$ is measurable and exponentially bounded (meaning that there is $M \ge 0$ and $\omega \in \mathbb{R}$ such that $|F(t)| \le M\mathrm{e}^{\omega t}$), then we can define its Laplace transform analogously to the scalar case as

$$\mathcal{L}(F)(\lambda) := \int_0^{\infty} \mathrm{e}^{-\lambda t} F(t)\, dt$$

for $\operatorname{Re}\lambda > \omega$.

We have seen in Proposition 12.2.7 that if $(T(t))_{t\ge0}$ is a C_0-semigroup of type (M, ω) with generator A, then

$$R(\lambda, A)f = \mathcal{L}(T(\cdot)f)(\lambda)$$

for $\operatorname{Re}\lambda > \omega$.

It can be shown by applying the theory of Laplace transforms that the following inversion formula hold, see Engel and Nagel [7, Corollary III.5.16].

Theorem B.1.25. *Let A generate a strongly continuous semigroup $(T(t))_{t\geq 0}$. Then for all $\omega > \omega_0(T)$ the representation formula*

$$T(t)f = \frac{1}{2i\pi t} \lim_{n\to\infty} \int_{\omega-in}^{\omega+in} e^{\lambda t} R(\lambda, A)^2 f \, d\lambda, \quad t \geq 0, f \in D(A^2)$$

holds.

Bibliography

[1] L. Alaoglu, *Weak topologies of normed linear spaces*, Ann. of Math. **41** (1940), 252–267.

[2] C.D. Aliprantis and O. Burkinshaw, *Positive Operators*, Academic Press, 1985.

[3] W. Arendt, C.J.K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser Verlag, 2011.

[4] S. Banach, *Sur les fonctionelles linéaires*, Studia Math. **1** (1929), 211–216.

[5] J.B. Conway, *A Course in Functional Analysis*, Graduate Texts in Math., vol. 96, Springer-Verlag, 1985.

[6] N. Dunford and J.T. Schwartz, *Linear Operators I. General Theory*, Interscience Publisher, 1958.

[7] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Math., vol. 194, Springer-Verlag, 2000.

[8] I. Gohberg, S. Goldberg, and M.A. Kaashoek, *Classes of Linear Operators I*, Oper. Theory Adv. Appl., vol. 49, Birkhäuser Verlag, 1990.

[9] H. Hahn, *Über lineare gleichungssysteme in linearen räumen*, J. Reine Angew. Math. **157** (1927), 214–229.

[10] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Grundlehren Math. Wiss., vol. 132, Springer-Verlag, 1980.

[11] S. Lang, *Real and Functional Analysis*, Graduate Texts in Math., vol. 142, Springer-Verlag, 1993.

[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I. Functional Analysis*, Academic Press, New York, 1972.

[13] W. Rudin, *Functional Analysis*, McGraw-Hill, 1973.

[14] A.E. Taylor and D.C. Lay, *Introduction to Functional Analysis*, 2nd ed., John Wiley & Sons, 1980.

[15] J. Voigt, *A perturbation theorem for the essential spectral radius of strongly continuous semigroups*, Monatsh. Math. **90** (1980), 153–161.

[16] K. Yosida, *Functional Analysis*, Grundlehren Math. Wiss., vol. 123, Springer-Verlag, 1965.