Semi-Monic Operator Functions: Perron-Frobenius Theory, Factorization in Ordered Banach Algebras and Degree-Reductions

vorgelegt von
Master of Science in Mathematik
Paul Peter Kallus
Schweidnitz (Polen)

Von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:
Vorsitzender: Prof. Dr. Wilhelm Stannat
Gutachter: Prof. Dr. Etienne Emmrich
Gutachter: Prof. Dr. Karl-Heinz Förster
Gutachter: Prof. Dr. Béla Nagy (TU Budapest)
Gutachter: Prof. Dr. Jussi Behrndt (TU Graz)

Tag der wissenschaftlichen Aussprache: 15. April 2016

Berlin 2016

Acknowledgements

Firstly I want to thank Etienne Emmrich who unhesitatingly agreed to act as official advisor of this thesis and thus made it possible.
My sincerest thanks go to Karl-Heinz Förster. Not only did he spark my interest for this paricular field of research, he was the one who offered me the opportunity to write this thesis in the first place and helped shape the direction it took. During countless meetings he provided a seemingly unending supply of helpful suggestions and carefully looked over my work.
I also have to extend my thanks to Béla Nagy. Several fruitful discussions during the summer of 2015 helped shape the contents of Chapter 3.
I want to thank Jussi Behrndt, not only for agreeing to act as a referee and offering his support but also for his teaching. His lectures on functional analysis were a driving factor for me to specialize in this field of mathematics.

Lastly I want to thank my family and friends for supporting and encouraging me throughout the time I was working on this thesis.

Contents

Introduction

Focal point of this work are so called *semi-monic (operator) functions* of the form

$$q(\lambda) = \lambda^m e - a(\lambda) = \lambda^m e - \sum_{j \in \mathbb{Z}} \lambda^j a_j$$

for some $m \in \mathbb{N}$ defined on a suitable annulus and where the coefficients a_j are elements in a Banach algebra with unit e. Typically there will be some restrictions imposed on the coefficients a_j.

Semi-monic functions appear throughout the literature, mostly for the special case when $m = 1$ and the function $a(\cdot)$ is a polynomial. This is true in e.g. probability theory when studying Markov chains where the coefficients a_j are entrywise nonnegative matrices (see [BNS13, BLM02, GHT96, GHT98, Neu89]).

Comparison theorems for ordinary differential equations give rise to semi-monic matrix polynomials of degree 2. One is then interested in finding nonnegative solutions (see [BJW85]).

Semi-monic polynomials also appear in physics where the coefficients are then often self-adjoint or positive semi-definite operators in Hilbert spaces, for example in hydrodynamics when studying small motions of fluids in a container (see [AHKM03, KK01]).

In [Mar88] A. Markus implicitly treats semi-monic functions in Banach spaces when establishing conditions for when operator polynomials admit so called canonical factorizations.

Spectral properties of semi-monic functions were systematically investigated in [Har11].

Part of the thesis will be concerned with the *block numerical range* of semi-monic operator functions and operator polynomials. The block numerical range first appeared as the quadratic numerical range for operators in Hilbert spaces in [LMMT01]. For Hilbert spaces it was introduced in its full generality in [TW03, Tre08, Wag07], including its extension to operator functions. In the unpublished Bachelor's thesis [Kal11] the block numerical range was extended to bounded operators defined on a product space $X = X_1 \times \cdots \times X_n$ of n Banach spaces (combined via some p-norm) in the following way: A bounded operator A on X admits a block matrix representation

$$\begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{bmatrix}.$$

The block numerical range of A with respect to the particular product space, denoted $\Theta(A; X_1, \dots, X_n)$, is then defined as the set

$$\bigcup_{(f,u) \in S_{\mathrm{att}}(X_1,\dots,X_n)} \Sigma \begin{bmatrix} f_1(A_{11}u_1) & \dots & f_1(A_{1n}u_n) \\ \vdots & \ddots & \vdots \\ f_n(A_{n1}u_1) & \dots & f_n(A_{nn}u_n) \end{bmatrix}$$

where $\Sigma(\cdot)$ denotes the eigenvalues of the scalar matrix and

$$\begin{aligned} S_{\mathrm{att}}(X_1, \ldots, X_n) := \{(f, u) : & u = (u_i)_{i=1,\ldots,n} \in X_1 \times \cdots \times X_n, \\ & f = (f_i)_{i=1,\ldots,n} \in X_1' \times \cdots \times X_n', \\ & f_i(u_i) = \|f_i\| = \|u_i\| = 1\}. \end{aligned}$$

For $n = 1$ this is precisely the spatial numerical range of an operator on a Banach space (see [BD71]). A noteworthy property of the block numerical range is that its closure contains the spectrum of A and it itself is contained in the numerical range of A.
Recently the block numerical range has been applied to positive operators in Hilbert lattices (see [Rad13]) and Banach lattices (see [Rad15b, Rad15a]).
Applications of the block numerical range are diverse: It can for example be used to determine if an operator is block diagonalizable (see [KMM07]). Block operator matrices, which naturally appear with the block numerical range, play an important role in several fields, like the discretisation of PDEs (see [SAD$^+$00]) and evolution problems (see [EN00]) with applications in quantum mechanics (see [Tha92]), hydrodynamics (see [Cha61]) and magnetohydrodynamics (see [Lif89]). In these applications properties of the spectrum and numerical range of these block operator matrices are often of particular interest, which shows the usefulness of the block numerical range.

The main body of the thesis is structured as follows.
Chapter 1 is mainly concerned with proving a Perron-Frobenius type result for the block numerical range of irreducible semi-monic matrix polynomials with entrywise nonnegative coefficients, that is polynomials of the form

$$Q(\lambda) = \lambda^m I - A(\lambda) = \lambda^m I - \sum_{j=0}^{l} \lambda^j A_j$$

for $m, l \in \mathbb{N}$, where the A_j are entrywise nonnegative square matrices and their sum $A(1)$ is irreducible. We start by introducing the functions

$$\begin{aligned} \mathrm{spr}_A &: \rho \mapsto \sup_{|\lambda|=\rho} \text{spectral radius}(A(\lambda)), \\ \mathrm{nur}_A &: \rho \mapsto \sup_{|\lambda|=\rho} \text{numerical radius}(A(\lambda)), \\ \mathrm{bnr}_A &: \rho \mapsto \sup_{|\lambda|=\rho} \text{block numerical radius}(A(\lambda)) \end{aligned}$$

(the subscript refers to the function $A(\cdot)$). The functions spr_A and nur_A were used in [FH08, FN05a, FN05b, FN15, Har11] and [FK15] respectively to derive results on the spectrum and numerical range of semi-monic functions. The function bnr_A was introduced in [Kal11] and it is studied and used in Sections 1.1 and 1.2 to determine areas which are disjoint to the block numerical range of semi-monic operator functions.

In Section 1.3 we consider the infinite graph $G_m(A_l, \ldots, A_0)$ for $m \in \mathbb{N}$ and entrywise nonnegative matrices A_j and recall some of its properties. A similar concept first appeared in [GHT96] in the context of Markov chains, though a graph structure was not established. The graph was used in [Har11] and [FK15] to prove Perron-Frobenius type results for the spectrum and the numerical range of semi-monic matrix functions.
We generalize these results in our main theorem in Section 1.4 to the block numerical range: It says that on circles $\mathbb{T}_\beta$ centred at 0 with radii $\beta > 0$ satisfying

$$\mathrm{bnr}_A(\beta) = \beta^m$$

the points in the block numerical range of the semi-monic matrix polynomial $Q(\cdot)$ have a cyclic distribution, where the number of distinct values on $\mathbb{T}_\beta$ can be derived from the infinite graph $G_m(A_l, \ldots, A_0)$. Known Perron-Frobenius type results for the spectrum and the numerical range of single matrices (see e.g. [Min88] and [Iss66]) or monic matrix polynomials (see e.g. [PT04]) then appear as special cases of our theorem.
In Section 1.5 we look at monic matrix polynomials with entrywise nonnegative coefficients

$$Q(\lambda) = \lambda^m I - \sum_{j=0}^{m-1} \lambda^j A_j.$$

For such polynomials one can consider the companion matrix

$$C = \begin{bmatrix} A_{m-1} & \ldots & A_1 & A_0 \\ -I & & & 0 \\ & \ddots & & \vdots \\ & & -I & 0 \end{bmatrix}$$

which can be used to prove Perron-Frobenius type results in the monic case. We study how the graph G_C associated with C and the infinite graph $G_m(A_{m-1}, \ldots, A_0)$are related and conclude that the infinite graph can be seen as a generalization of G_C.
We close the chapter by considering semi-monic operator functions with positive semi-definite or normal operator coefficients. Extending results of [SW10, Wim11] we show that eigenvalues on certain circles around the origin are normal and semisimple.

Chapter 2 deals with factorizations of elements in (strongly) decomposing (Banach) algebras. An algebra $\mathcal{A}$ with unit e is said to be *decomposing* if it can be written as the direct sum of two subalgebras $\mathcal{A}_-$ and $\mathcal{A}_+$. It is said to be *strongly decomposing* if it can be written as the direct sum of three subalgebras

$$\mathcal{A} = \mathcal{A}_- \dot{+} \mathcal{A}_0 \dot{+} \mathcal{A}_+$$

where the subalgebras satisfy some additional conditions (e.g. elements in $\mathcal{A}_-$ and $\mathcal{A}_+$ are required to be quasi-nilpotent). As an example one might think of strictly lower triangular, diagonal and strictly upper triangular matrices. Also note that each strongly decomposing algebra is also a decomposing algebra.

Then it is known that an element $a \in \mathcal{A}$, which is in some sense close to the unit element, factorizes as

$$e - a = (e - a_-)(e - a_+)$$

in decomposing Banach algebras and as

$$e - a = (e - a_-)a_0(e - a_+)$$

in strongly decomposing Banach algebras, where the elements a_α, $\alpha = -, 0, +$, lie in the respective subalgebras $\mathcal{A}_\alpha$ and fulfil some additional invertibility conditions (see [GGK93, GKS03, Mar88]).
We introduce a third type of algebra, called *semi-strongly decomposing algebra*, which fits in between the two definitions above. In contrast to strongly decomposing algebras, it includes the important special case of the Wiener algebra.
In Section 2.1 we extend the concept to ordered algebras $\mathcal{A}$ with a cone $\mathcal{C}$ and define *((semi-)strongly) ordered decomposing algebras* from which we require that the natural projections onto their subalgebras leave the cone $\mathcal{C}$ invariant. A general way to construct strong decompositions of bounded operators on Banach spaces is given by the concept of *chains of projections* (see [GGK93]), and we extend it to the Banach algebra case in Section 2.1.1.
Section 2.2 contains our two main factorization results for positive elements with spectral radius smaller than 1 in either a decomposing or semi-strongly decomposing Banach algebra. The factorizations we obtain give elements $a_\pm$ and a_0^{-1} which are again nonnegative. Our theorems extend a result previously given [FN05b].
We continue by showing how the factorization results can be applied for M-matrices, Hilbert-Schmidt operators with nonnegative kernel function and positive operators in the Banach algebra of regular operators.
Finally we return to semi-monic functions in the form of semi-monic polynomials $q(\lambda) = \lambda^m - \sum_{j=0}^{l} \lambda^j a_j$ with coefficients in $\mathbb{C}$ to restate a result given in [FN05b] which follows more easily in our case: By first proving a factorization result in the Wiener algebra we are able to show that $q(\cdot)$ factorizes into

$$q(\lambda) = b(\lambda)q_0c(\lambda)$$

if and only if there exists a $\rho > 0$ such that

$$\mathrm{spr}_a(\rho) < \rho^m.$$

Here $b(\cdot)$ is monic of degree m and $c(\cdot)$ is semi-monic. Furthermore the spectrum of $q(\cdot)$ is divided between $b(\cdot)$ and $c(\cdot)$, with $b(\cdot)$ accounting for the spectrum inside the disc with radius ρ and $c(\cdot)$ for the spectrum outside of it.
As a corollary we obtain a version of Pellet's theorem for semi-monic polynomials with nonnegative coefficients.

The topic of Chapter 3 are *degree-reductions* of operator polynomials with the coefficients acting on Banach spaces. Linearisations of operator polynomials (for example via

the companion matrix displayed above) are a useful tool for studying properties of the original polynomial without having to deal with its higher degree (see [Rod89]). Degree-reductions generalize this concept by reducing the polynomial to an arbitrary degree. They were recently considered in [TDM14, TDD15, TDD] for matrix polynomials.
The aim of this chapter is mainly the study of a particular degree-reduction which we call *canonical degree-reduction*. For an operator polynomial

$$A(\lambda) = \lambda^l A_l + \cdots + A_0,$$

with coefficients acting between some Banach spaces X and Y, the canonical reduction to degree $\hat{l} < l$ is the operator polynomial (in block matrix form)

$$\hat{A}(\lambda) = \begin{bmatrix} A^{[\hat{l}]}(\lambda) & A_{l-\hat{l}-1} & A_{l-\hat{l}-2} & \cdots & A_1 & A_0 \\ -I_X & \lambda I_X & 0_X & \cdots & \cdots & 0_X \\ 0_X & -I_X & \lambda I_X & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0_X \\ 0_X & \cdots & \cdots & 0_X & -I_X & \lambda I_X \end{bmatrix},$$

where $A^{[\hat{l}]}(\cdot)$ is the $\hat{l}$-th Horner shift of $A(\cdot)$, i.e.

$$A^{[\hat{l}]}(\lambda) = \lambda^{\hat{l}} A_l + \cdots + \lambda A_{l-\hat{l}+1} + A_{l-\hat{l}}.$$

Note that for $\lambda \in \mathbb{C}$ the operator $\hat{A}(\lambda)$ goes from $X^{l-\hat{l}}$ to $Y \times X^{l-\hat{l}-1}$. These types of degree-reductions were introduced in [Nag07] and used in [Har11].
In Section 3.1 we look at divisions with remainder and show how the factors of a division with remainder of $\hat{A}(\cdot)$ can be recovered from a division with remainder of $A(\cdot)$ (and vice versa). In the case of a monic $A(\cdot)$ we show how a given division with remainder easily allows one to recover a spectral triple for $A(\cdot)$.
Consecutive degree-reductions of $A(\cdot)$ first to degree $\hat{l}_1$ and then to degree $\hat{l}_2 < \hat{l}_1$ never coincide with the degree-reduction that goes straight to $\hat{l}_2$. They are however closely related, and we examine their connection in Section 3.3 following a procedure outlined in [MX13] for matrix polynomials.
Lastly, we return once again to our semi-monic operator polynomials $Q(\cdot)$. Specifically, the canonical reduction $\hat{Q}(\cdot)$ to degree $l - m + 1$ is again semi-monic, but now the monic part has degree 1. Additionally the operator coefficients inherit properties such as nonnegativeness. We outline how a fixpoint iteration for the operator equation

$$\hat{Q}(C) = C - \tilde{A}(C) = 0$$

can be used to recover the factors of a division without remainder of $Q(\cdot)$.

The chapters are written to be mostly self-contained, with the exception of Sections 2.4 and 3.5.

Notation

- $\langle n \rangle$ the set of integers $1, \ldots, n$ and $\langle n \rangle_0 = \langle n \rangle \cup \{0\}$,
- $\mathbb{T}_\rho = \{\lambda \in \mathbb{C} : |\lambda| = \rho\}$ the circle in the complex plane $\mathbb{C}$ centred at the origin with radius $\rho > 0$,
- $\mathbb{A}_{\rho_1,\rho_2} = \{\lambda \in \mathbb{C} : \rho_1 < |\lambda| < \rho_2\}$ the open annulus in the complex plane $\mathbb{C}$ centred at the origin with radii $0 \leq \rho_1 < \rho_2$,
- $\mathbb{C}^{n,n}$ the set of complex $n \times n$ matrices and $\mathbb{R}_+^{n,n}$ the set of entrywise nonnegative $n \times n$ matrices,
- for $A = (a_{rs}) \in \mathbb{R}_+^{n,n}$ we also write $A \geq 0$ (iff $a_{rs} \geq 0$ for all $r, s \in \langle n \rangle$),
- for $x \in \mathbb{R}_+^n$ we write $x \gg 0$ iff all entries of x are strictly positive,
- for normed spaces X and Y denote by $\mathcal{L}(X)$ the set of bounded linear operators on X, and $\mathcal{L}(X, Y)$ the set of bounded operators from X to Y
- $\Sigma(B)$ the spectrum and $\mathrm{spr}(B)$ the spectral radius of B where either $B \in \mathcal{L}(X)$ or B an element in a Banach algebra,
- $\Sigma_p(B)$ the point spectrum or eigenvalues of $B \in \mathcal{L}(X)$,
- $\Theta(b; \mathcal{A})$ the numerical range and $\mathrm{nur}(b; \mathcal{A})$ the numerical radius of an element b in a Banach algebra $\mathcal{A}$,
- $\Theta(B; X)$ the spatial numerical range and $\mathrm{nur}(b; X)$ the spatial numerical radius of a bounded operator B on a Banach space X.

1. Semi-monic operator functions

The objects of main interest in this work will be so called *semi-monic operator functions* which we will now introduce. For that purpose let $\mathcal{A}$ be a Banach-algebra with unit e and $m \in \mathbb{N}$. Then we call

$$q(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \mapsto \mathcal{A} : \lambda \mapsto \lambda^m e - \sum_{j\in\mathbb{Z}} \lambda^j a_j =: \lambda^m e - a(\lambda) \tag{1.1}$$

a *semi-monic function* with monic part of degree m, where the $a_j \in \mathcal{A}$, $j \in \mathbb{Z}$ and the domain $\mathbb{A}_{\rho_1,\rho_2}$ is the maximal annulus on which the Laurent series function $a(\cdot)$ in (1.1) converges.

Remark 1.1. The term 'semi-monic' is derived from *monic polynomials*, i.e. polynomials of the form

$$q(\lambda) = \lambda^m e - \sum_{j=0}^{l} \lambda^j a_j$$

where $m, l \in \mathbb{N}$, $m > l$ (typically $l = m - 1$). Since we do not require for $m > l$ (and in fact allow $a(\cdot)$ to be a Laurent series function), we call operator functions of the form (1.1) semi-monic.

We will now introduce several real valued functions that are helpful in studying properties of semi-monic operator functions. First recall that the *spectrum* of an element $b \in \mathcal{A}$ is defined as

$$\Sigma(b) := \{\lambda \in \mathbb{C} : \lambda e - b \text{ is not invertible}\}$$

and the *spectral radius* as

$$\mathrm{spr}(b) := \max\{|\lambda| : \lambda \in \Sigma(b)\}.$$

The *numerical range* of $b \in \mathcal{A}$ as a Banach algebra element is defined as

$$\Theta(b, \mathcal{A}) := \{f(b) : f \in \mathcal{A}', f(e) = \|f\| = 1\},$$

with $\mathcal{A}'$ the dual of $\mathcal{A}$. The *numerical radius* is then defined as

$$\mathrm{nur}(b; \mathcal{A}) := \max\{|\lambda| : \lambda \in \Theta(b; \mathcal{A})\}.$$

Moreover, there holds $\mathrm{spr}(b) \leq \mathrm{nur}(b; \mathcal{A}) \leq \|b\|$ (see e.g. [BD71, Theorem 2.6]). We can define the aforementioned functions for $a(\cdot)$ as in (1.1):

$$\mathrm{spr}_a(\cdot) : (\rho_1, \rho_2) \to \mathbb{R}_+ : \rho \mapsto \sup_{|\lambda|=\rho} \mathrm{spr}(a(\lambda)), \tag{1.2}$$

$$\mathrm{nur}_a(\cdot) : (\rho_1, \rho_2) \to \mathbb{R}_+ : \rho \mapsto \sup_{|\lambda|=\rho} \mathrm{nur}(a(\lambda); \mathcal{A}), \tag{1.3}$$

$$\mathrm{norm}_a(\cdot) : (\rho_1, \rho_2) \to \mathbb{R}_+ : \rho \mapsto \sum_{j\in\mathbb{Z}} \rho^j \|a_j\|. \tag{1.4}$$

It follows immediately that $\operatorname{spr}_a(\rho) \le \operatorname{nur}_a(\rho) \le \operatorname{norm}_a(\rho)$ for all $\rho \in (\rho_1, \rho_2)$. The function $\operatorname{spr}_a(\cdot)$ was introduced and used in [FN05a] and [FN05b] and later [Har11] to derive spectral properties of semi-monic functions, $\operatorname{nur}_a(\cdot)$ was introduced in [FK15] and used to derive properties of the numerical range of semi-monic functions, while $\operatorname{norm}_a(\cdot)$ implicitly appeared in [Mar88, p.122] to formulate a sufficient condition for factorizations of semi-monic polynomials.

1.1. The block numerical range for semi-monic operator functions

We want to recall the block numerical range for operator functions with coefficients on a Banach space and introduce an analogue to the functions in (1.2) - (1.4) for the block numerical range. In order to do that we need to move our setting from Banach algebras to Banach spaces.
Knowledge of the block numerical range of operators on a Banach space will be required for the sequel. For the convenience of the reader we provide a self-contained introduction to the topic along with a collection of properties and their proofs in Appendix A. Nonetheless we will start this section with a very brief overview of the block numerical range of operators on Banach spaces which should enable the reader to follow the text.

Let $(X_1, \|\cdot\|_1), \ldots, (X_n, \|\cdot\|_n)$ be n Banach spaces and $X = X_1 \times \ldots \times X_n$ be the product space where the component norms are combined via some p-norm, $1 \le p < \infty$.
A bounded operator A on $X = X_1 \times \cdots \times X_n$ can then be identified with the block matrix operator

$$\begin{bmatrix} A_{11} & \ldots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \ldots & A_{nn} \end{bmatrix}$$

where $A_{rs} \in \mathcal{L}(X_s, X_r)$. We also need the set

$$S_{\mathrm{att}}(X_1, \ldots, X_n) = \{(f, u)_{i=1,\ldots,n} : (f_i, u_i) \in X_i' \times X_i, f_i(u_i) = \|f_i\| = \|u_i\| = 1\}.$$

The *block numerical range* of A with respect to the particular product space is then defined as the set

$$\Theta(A; X_1, \ldots, X_n) := \bigcup_{(f,u) \in S_{\mathrm{att}}(X_1, \ldots, X_d)} \Sigma \begin{bmatrix} f_1(A_{11}u_1) & \ldots & f_1(A_{1n}u_n) \\ \vdots & \ddots & \vdots \\ f_n(A_{n1}u_1) & \ldots & f_n(A_{nn}u_n) \end{bmatrix} \tag{1.5}$$

where $\Sigma(\cdot)$ denotes the eigenvalues of the scalar matrix. The *block numerical radius* of A is defined as

$$\operatorname{bnr}(A; X_1, \ldots, X_n) := \sup\{|\lambda| : \lambda \in \Theta(A; X_1, \ldots, X_n)\}.$$

A pair $(f, u) \in S_{\mathrm{att}}(X_1, \ldots, X_n)$ can be identified with a pair of operators F and U via

$$\begin{aligned} F &: X_1 \times \ldots \times X_n \to \mathbb{C}^n : (x_1, \ldots, x_n) \mapsto (f_1(x_1), \ldots, f_n(x_n)) \\ U &: \mathbb{C}^n \to X_1 \times \ldots \times X_n : (\lambda_1, \ldots, \lambda_n) \mapsto (\lambda_1 u_1, \ldots, \lambda_n u_n). \end{aligned}$$

The matrix in (1.5) can then be written as FAU which is sometimes more convenient. The set of of all of these operator pairs will be denoted $S_{\text{att}}^{op}(X_1,\dots,X_n)$. Note that there is a 1 to 1 correspondence between $S_{\text{att}}(X_1,\dots,X_n)$ and $S_{\text{att}}^{op}(X_1,\dots,X_n)$.
If we do not decompose X into a product space we end up with

$$\Theta(A;X) = \{f(Au) : (f,u) \in X' \times X, f(u) = \|f\| = \|u\| = 1\}$$

which is precisely the *spatial numerical range* of A as in [BD71]. The *spatial numerical radius* is denoted $\operatorname{nur}(A;X)$.

Remark 1.2. Note that we can also treat an operator $A \in \mathcal{L}(X)$ as an element of the Banach algebra $\mathcal{L}(X)$ and consider the numerical range in algebra sense

$$\Theta(A;\mathcal{L}(X)) = \{f(A) : f \in \mathcal{L}(X)', f(I_X) = \|f\| = 1\}$$

as in the previous section. These two definitions do not coincide. In general there only holds

$$\Theta(A;X) \subset \Theta(A;\mathcal{L}(X)).$$

However, it can be shown that $\Theta(A;\mathcal{L}(X))$ is the closed convex hull of $\Theta(A;X)$ (see [BD71, Theorem 9.4]). Therefore the (spatial) numerical radius of A is equal with respect to either definition and we will simply write $\operatorname{nur}(A)$.

We close this overview by collecting some important properties of the block numerical range of an operator A (proofs of which can be found in Appendix A):

- $\Sigma_p(A) \subseteq \Theta(A;X_1,\dots,X_n) \subset \Theta(A;X)$,
- $\Sigma(A) \subseteq \overline{\Theta(A;X_1,\dots,X_n)}$,
- $\operatorname{spr}(A) \leq \operatorname{bnr}(A;X_1,\dots,X_n) \leq \operatorname{nur}(A) \leq \|A\|$,
- $\Theta(A;X_1,\dots,X_n)$ consists of at most n connected components.

In the Banach space setting the semi-monic functions in (1.1) now read as

$$Q(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \mapsto \mathcal{L}(X) : \lambda \mapsto \lambda^m I - \sum_{j\in\mathbb{Z}} \lambda^j A_j =: \lambda^m I - A(\lambda) \tag{1.6}$$

with the $A_j \in \mathcal{L}(X)$.
Similarly to how the spectrum and spectral radius of an operator are extended to operator functions we define:

Definition 1.3. For an operator function $T(\cdot) : \Omega \to \mathcal{L}(X)$ define the block numerical range

$$\Theta(T(\cdot);X_1,\dots,X_n) := \{\lambda \in \Omega : 0 \in \Theta(T(\lambda);X_1,\dots,X_n)\},$$

and block numerical radius

$$\operatorname{bnr}(T(\cdot);X_1,\dots,X_n) := \sup\{|\lambda| : \lambda \in \Theta(T(\cdot);X_1,\dots,X_n)\}$$

where $\operatorname{bnr}(T(\cdot);X_1,\dots,X_n) := \infty$ should the supremum not exist.

Example 1.4. For an operator $A \in \mathcal{L}(X)$ and the operator polynomial $P_A(\cdot) : \mathbb{C} \to \mathcal{L}(X) : \lambda \mapsto \lambda - A$ we have $\Theta(P_A(\cdot); X_1, \dots, X_n) = \Theta(A; X_1, \dots, X_n)$.

Some numerically achieved illustrations of block numerical ranges of semi-monic matrix polynomials can be found in [Wag07, p.29, p.31].

Remark 1.5. For an operator function $T : \Omega \to \mathcal{L}(X)$ and $(F, U) \in S^{op}_{\text{att}}(X_1, \dots, X_n)$ we can construct a complex valued function $\det(FT(\cdot)U) : \Omega \to \mathbb{C}$. The block numerical range of T can then be characterized through

$$\lambda \in \Theta(T; X_1, \dots, X_n) \iff \exists (F, U) \in S^{op}_{\text{att}}(X_1, \dots, X_n) : \lambda \text{ is a zero of } \det(FT(\cdot)U).$$

Finally we can define for an analytic operator function as in (1.6)

$$\operatorname{bnr}_A(\cdot; X_1, \dots, X_n) : (\rho_1, \rho_2) \to \mathbb{R}_+ : \rho \mapsto \sup_{|\lambda| = \rho} \operatorname{bnr}(A(\lambda); X_1, \dots, X_n). \tag{1.7}$$

The functions in (1.2) - (1.4) are still well defined in the Banach space case of this section, we just need to use the Banach space definitions of the spectrum and the spatial numerical range.

Lemma 1.6. *For $A(\cdot)$ as in (1.6) there holds*

$$\operatorname{spr}_A(\rho) \leq \operatorname{bnr}_A(\rho; X_1, \dots, X_n) \leq \operatorname{nur}_A(\rho) \leq \operatorname{norm}_A(\rho)$$

for all $\rho \in (\rho_1, \rho_2)$.

Proof. The first inequality follows from Theorem A.13, the second inequality follows from Theorem A.16 and the third inequality is clear because $\operatorname{nur}(B) \leq \|B\|$ for a bounded operator B. □

Of particular interest will be inequalities (or equalities) of the form

$$\operatorname{bnr}_A(\rho; X_1, \dots, X_n) \leq (=) \rho^m \tag{1.8}$$

where the $m \in \mathbb{N}$ comes from the semi-monic operator function.

We will now extend results for semi-monic operator functions to the block numerical range that were previously shown for the spectrum in [Har11] and the numerical range in [FK15]. There is one more preliminary result that we need that appeared in [Har11].

Lemma 1.7. *For an analytic operator function $A(\cdot) : \mathbb{A}_{\rho_1, \rho_2} \to \mathcal{L}(X)$ the function*

$$\operatorname{spr}_A(\cdot) : (\rho_1, \rho_2) \to \mathbb{R} : \sigma \mapsto \sup_{|\lambda| = \sigma} \operatorname{spr}(A(\lambda)),$$

is geometrically convex, *that is for all $\sigma_1, \sigma_2 \in (\rho_1, \rho_2)$ and $\theta \in [0, 1]$ the functional inequality*

$$\operatorname{spr}_A(\sigma_1^\theta \sigma_2^{1-\theta}) \leq (\operatorname{spr}_A(\sigma_1))^\theta (\operatorname{spr}_A(\sigma_2))^{\theta - 1}$$

holds.

Proof. See [Har11, p.14]. □

This result naturally extends to the case of the block numerical range.

Lemma 1.8. *For an analytic operator function* $A : \mathbb{A}_{\rho_1,\rho_2} \to \mathcal{L}(X)$ *the function*

$$\operatorname{bnr}_A(\,\cdot\,; X_1, \dots, X_n) : (\rho_1, \rho_2) \to \mathbb{R} : \sigma \mapsto \sup_{|\lambda|=\sigma} \operatorname{bnr}(A(\lambda), X_1, \dots, X_n)$$

is geometrically convex, i.e. for $\sigma_1, \sigma_2 \in (\rho_1, \rho_2)$ *and* $\theta \in [0, 1]$ *there holds*

$$\operatorname{bnr}_A(\sigma_1^\theta \sigma_2^{1-\theta}; X_1, \dots, X_n) \le (\operatorname{bnr}_A(\sigma_1; X_1, \dots, X_n))^\theta (\operatorname{bnr}_A(\sigma_2; X_1, \dots, X_n))^{1-\theta}.$$

Proof. First note that $\Theta(B; X_1, \dots, X_n) = \bigcup_{(F,U)\in S_{\text{att}}^{op}} \Sigma(FBU)$ for an operator $B \in \mathcal{L}(X)$ implies

$$\operatorname{bnr}(B; X_1, \dots, X_n) = \sup_{(F,U)\in S_{\text{att}}^{op}} \operatorname{spr}(FBU).$$

Therefore it follows that for $\sigma_1, \sigma_2 \in (\rho_1, \rho_2)$ and $\theta \in [0, 1]$

$$\begin{aligned}
\operatorname{bnr}_A(\sigma_1^\theta \sigma_2^{1-\theta}; X_1, \dots, X_n) &= \sup_{|\lambda|=\sigma_1^\theta\sigma_2^{1-\theta}} \{\operatorname{nur}(A(\lambda); X_1, \dots, X_n)\} \\
&= \sup_{|\lambda|=\sigma_1^\theta\sigma_2^{1-\theta}} \left\{ \sup_{(F,U)\in S_{\text{att}}^{op}} \operatorname{spr}(FA(\lambda)U) \right\} \\
&= \sup_{(F,U)\in S_{\text{att}}^{op}} \left\{ \sup_{|\lambda|=\sigma_1^\theta\sigma_2^{1-\theta}} \operatorname{spr}(FA(\lambda)U) \right\} \\
&= \sup_{(F,U)\in S_{\text{att}}^{op}} \operatorname{spr}_{FA(\cdot)U}(\sigma_1^\theta \sigma_2^{1-\theta}) \\
&\le \sup_{(F,U)\in S_{\text{att}}^{op}} (\operatorname{spr}_{FA(\cdot)U}(\sigma_1))^\theta (\operatorname{spr}_{FA(\cdot)U}(\sigma_2))^{1-\theta} \\
&\le \left(\sup_{(F,U)\in S_{\text{att}}^{op}} \operatorname{spr}_{FA(\cdot)U}(\sigma_1) \right)^\theta \left(\sup_{(F,U)\in S_{\text{att}}^{op}} \operatorname{spr}_{FA(\cdot)U}(\sigma_2) \right)^{1-\theta} \\
&= (\operatorname{bnr}_A(\sigma_1; X_1, \dots, X_n))^\theta (\operatorname{bnr}_A(\sigma_2; X_1, \dots, X_n))^{1-\theta}
\end{aligned}$$

that is, $\operatorname{bnr}_A(\cdot)$ is geometrically convex. Here the second to last inequality follows from Lemma 1.7. □

The following proposition characterizes the behaviour of $\operatorname{bnr}_A(\cdot)$ on intervals.

Proposition 1.9. *Let* $A(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \to \mathcal{L}(X)$ *be an analytic operator function and* $\sigma_1, \sigma_2 \in (\rho_1, \rho_2)$ *with* $\sigma_1 \le \sigma_2$ *and* $\operatorname{bnr}_A(\sigma_j) = \sigma_j^m$ *for* $j = 1, 2$ *and some* $m \in \mathbb{N}$. *Then exactly one of the following assertions is true.*

(i) $\operatorname{bnr}_A(\sigma; X_1, \dots, X_n) = \sigma^m$ *for all* $\sigma \in (\sigma_1, \sigma_2)$.

(ii) $\operatorname{bnr}_A(\sigma; X_1, \dots, X_n) < \sigma^m$ *for all* $\sigma \in (\sigma_1, \sigma_2)$.

Proof. Note that the geometric convexity of $\mathrm{bnr}_A(\cdot)$ implies that $t \mapsto \log \mathrm{bnr}_A(e^t), t \in \mathbb{R}_+$ is convex. Since $t \mapsto \log(e^t)^m = mt$ is linear the result follows from well known facts about convex functions. $\square$

Remark 1.10. Proposition 1.9 also implies that the equality

$$\mathrm{bnr}_A(\beta; X_1, \ldots, X_n) = \beta^m$$

can be satisfied for none, one , two or for all $\beta \in (\sigma_1, \sigma_2)$. All of these cases can occur. For a thorough analysis of the function $\mathrm{spr}_A(\cdot)$ see [FN05a, Theorem 4.10].

The next two results provide conditions under which the block numerical range of semi-monic operator functions is disjoint to some annulus around the origin. See Figure 1 for an illustration.

Proposition 1.11. *Let $A(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \to \mathcal{L}(X)$ be an analytic operator function and $Q(\cdot)$ its corresponding semi-monic operator function for some $m \in \mathbb{N}$. Let $\sigma_1, \sigma_2 \in (\rho_1, \rho_2)$, $\sigma_1 < \sigma_2$, such that $\mathrm{bnr}_A(\sigma; X_1, \ldots, X_n) < \sigma^m$ for all $\sigma \in (\sigma_1, \sigma_2)$. Then*

$$\Theta(Q(\cdot); X_1, \ldots, X_n) \cap \mathbb{A}_{\sigma_1,\sigma_2} = \emptyset.$$

Proof. Assume there exists $\lambda_0 \in \Theta(Q(\cdot); X_1, \ldots, X_n) \cap \mathbb{A}_{\sigma_1,\sigma_2}$. This means that

$$\begin{aligned} & 0 \in \Theta(Q(\lambda_0); X_1, \ldots, X_n) \\ \Leftrightarrow \quad & 0 \in \Theta(\lambda_0^m - A(\lambda_0); X_1, \ldots, X_n) \\ \Leftrightarrow \quad & \lambda_0^m \in \Theta(A(\lambda_0); X_1, \ldots, X_n). \end{aligned}$$

Additionally we have $|\lambda_0| \in (\sigma_1, \sigma_2)$. It follows that

$$\begin{aligned} |\lambda_0|^m &\leq \mathrm{bnr}(A(\lambda_0); X_1, \ldots, X_n) \\ &\leq \sup_{|\lambda|=|\lambda_0|} \mathrm{bnr}(A(\lambda); X_1, \ldots, X_n) \\ &= \mathrm{bnr}_A(|\lambda_0|; X_1, \ldots, X_n) \\ &< |\lambda_0|^m \end{aligned}$$

which is a contradiction. $\square$

Theorem 1.12. *Let $A(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \to \mathcal{L}(X)$ be an analytic operator function and $Q(\cdot)$ its corresponding semi-monic operator function. Let $\sigma_1, \sigma_2 \in (\rho_1, \rho_2)$, $\sigma_1 < \sigma_2$, with $\mathrm{bnr}_A(\sigma_j; X_1, \ldots, X_n) = \sigma_j^m$ for $j = 1, 2$. If there exists a $\sigma \in (\sigma_1, \sigma_2)$ such that $\mathrm{bnr}_A(\sigma; X_1, \ldots, X_n) < \sigma^m$ then*

$$\Theta(Q(\cdot); X_1, \ldots, X_d) \cap \mathbb{A}_{\sigma_1,\sigma_2} = \emptyset.$$

Proof. Proposition 1.9 implies that $\mathrm{bnr}_A(\sigma; X_1, \ldots, X_n) < \sigma^m$ for all $\sigma \in (\sigma_1, \sigma_2)$. The assertion then follows from Proposition 1.11. $\square$

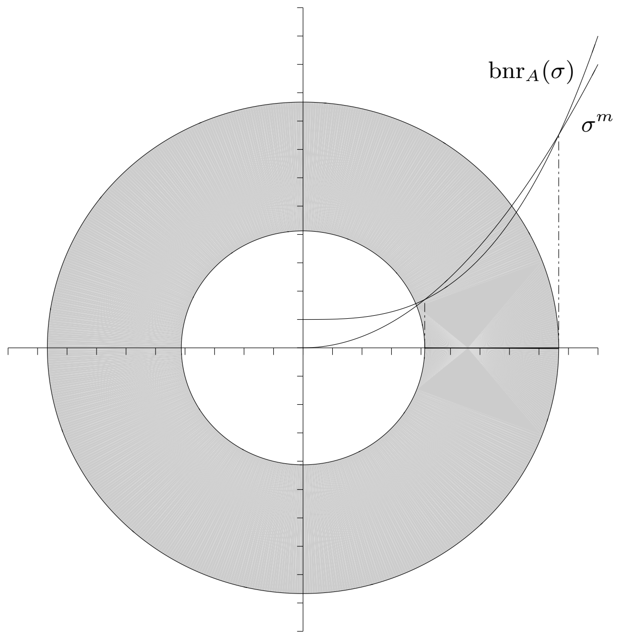


Figure 1: The open annulus characterized be the inequality $\mathrm{bnr}(A(\sigma); X_1, \ldots, X_n) < \sigma^m$ contains no block numerical range.

1.2. Semi-monic operator functions with nonnegative coefficients

We will now provide conditions under which the representation of the function $\mathrm{bnr}_A(\cdot)$ can be simplified. We first consider the case where the coefficients of the operator function $A(\cdot)$ are positive operators on a Banach lattice.
Let X_j be Banach lattices for $k = 1 \ldots, n$. Then the product space $X = X_1 \times \ldots \times X_n$ is also a Banach lattice where the order is defined component wise. Moreover by [Sch74, Proposition II.5.5] the dual spaces X_j' are again Banach lattices. Thus we can state the following lemma.

Lemma 1.13. *Let X_j be Banach lattices for $k = 1, \ldots, n$ and $(f, u) \in S_{att}(X_1, \ldots, X_n)$. Then also $(|f|, |u|) = [(|f_1|, |u_1|), \ldots, (|f_n|, |u_n|)] \in S_{att}(X_1, \ldots, X_n)$ and the corresponding operator pair $(|F|, |U|)$ is a pair of positive operators w.r.t. the standard order on $\mathbb{R}^n$.*
Moreover for a positive operator $A \in \mathcal{L}(X)$ the operator $|F|\, A\, |U|$ is again positive and

$$\|FAU\| \leq \|\, |F|\, A\, |U|\, \|.$$

Lemma 1.14. *Let $A(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \to \mathcal{L}(X)$ be an analytic operator function whose Laurent coefficients are positive operators in the Banach lattice $X = X_1 \times \ldots \times X_n$. Then for $(F, U) \in S^{op}_{att}(X_1, \ldots, X_n)$ there exists a $\gamma > 0$ such that*

$$\|\, |F|\, A(\lambda)\, |U|\, \| \leq \gamma \|\, |F|\, A(|\lambda|)\, |U|\, \|$$

for every $\lambda \in \mathbb{A}_{\rho_1,\rho_2}$.

Proof. This follows directly from the preceding lemma together with [Har11, Lemma 1.13]. $\square$

Proposition 1.15. *Let $A(\cdot) : \mathbb{A}_{\rho_1,\rho_2} \to \mathcal{L}(X)$ be an analytic operator function whose Laurent coefficients are positive operators in the Banach lattice $X = X_1 \times \ldots \times X_n$. Then for all $\rho \in (\rho_1, \rho_2)$ we have the equality*

$$\mathrm{bnr}_A(\rho, X_1, \ldots, X_n) = \mathrm{bnr}(A(\rho), X_1, \ldots, X_n).$$

Proof. Clearly $\mathrm{bnr}_A(\rho; X_1, \ldots, X_n) \geq \mathrm{bnr}(A(\rho); X_1, \ldots, X_n)$. The other direction follows from Lemma 1.13 and Lemma 1.14 via

$$\begin{aligned}
\mathrm{bnr}_A(\rho; X_1, \ldots, X_n) &= \sup_{|\lambda|=\rho} \mathrm{bnr}(A(\lambda); X_1, \ldots, X_n) \\
&= \sup_{|\lambda|=\rho} \sup_{(F,U)\in S^{op}_{\mathrm{att}}} \mathrm{spr}(FA(\lambda)U) \\
&= \sup_{|\lambda|=\rho} \sup_{(F,U)\in S^{op}_{\mathrm{att}}} \lim_{k\to\infty} \|(FA(\lambda)U)^k\|^{1/k} \\
&\leq \sup_{|\lambda|=\rho} \sup_{(F,U)\in S^{op}_{\mathrm{att}}} \lim_{k\to\infty} \|(|F|A(\lambda)|U|)^k\|^{1/k} \\
&\leq \sup_{|\lambda|=\rho} \sup_{(F,U)\in S^{op}_{\mathrm{att}}} \lim_{k\to\infty} \underbrace{\gamma^{1/k}}_{\to 1} \|(|F|A(|\lambda|)|U|)^k\|^{1/k} \\
&= \sup_{|\lambda|=\rho} \sup_{(F,U)\in S^{op}_{\mathrm{att}}} \mathrm{spr}(|F|A(\rho)|U|) \\
&= \sup_{(F,U)\in S^{op}_{\mathrm{att}},(F,U)=(|F|,|U|)} \mathrm{spr}(FA(\rho)U) \\
&\leq \sup_{(F,U)\in S^{op}_{\mathrm{att}}} \mathrm{spr}(FA(\rho)U) \\
&= \mathrm{bnr}(A(\rho); X_1, \ldots, X_n)
\end{aligned}$$

$\square$

The equation

$$\mathrm{bnr}_A(\beta; X_1, \ldots, X_n) = \beta^m$$

has a special significance for strictly monic matrix polynomials with entrywise nonnegative coefficients, i.e. polynomials

$$Q(\lambda) = \lambda^m I - A(\lambda) = \lambda^m I - \sum_{j=0}^{m-1} \lambda^j A_j, \tag{1.9}$$

where $m \in \mathbb{N}$ and the $A_j \in \mathbb{R}^{n,n}_+$. Since we are working with matrices the block numerical range will be formed with respect to the product space $\mathbb{C}^n = \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_p}$ (for an exposition on block numerical ranges of matrices see Section A.5). Since $Q(\cdot)$ is monic its numerical range is compact (see [PT04, p.12]) and thus its numerical radius $\mathrm{nur}(Q(\cdot)) = \sup\{|\lambda| : \lambda \in \Theta(Q(\cdot))\}$ is finite. Then by Theorem A.16 (which trivially extends to functions) its block numerical radius $\mathrm{bnr}(Q(\cdot); \mathbb{C}^{n_1} \times \ldots \times \mathbb{C}^{n_p})$ is also finite. We can then state the following result:

Lemma 1.16. *Let* $Q(\cdot)$ *be as in* (1.9) *and define* $\beta_{bnr} = \mathrm{bnr}(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$*. Then*

$$\mathrm{bnr}_A(\beta_{bnr}; \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) = \mathrm{bnr}(A(\beta_{bnr}); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) = \beta_{bnr}^m.$$

Proof. Since $\Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ is compact we can find a $\hat{\lambda} \in \Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ with $|\hat{\lambda}| = \beta_{bnr}$. Then by definition $\hat{\lambda}^m \in \Theta(A(\hat{\lambda}); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ and thus

$$\begin{aligned}\mathrm{bnr}_A(\beta_{bnr}; \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) &= \sup_{|\lambda| = \beta_{bnr}} \mathrm{bnr}(A(\lambda); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) \\ &\geq \mathrm{bnr}(A(\hat{\lambda}); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) \geq |\hat{\lambda}^m| = \beta_{bnr}^m.\end{aligned}$$

To see the other inequality assume that $\mathrm{bnr}_A(\beta_{bnr}; \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) > \beta_{bnr}^m$. Since $Q(\cdot)$ is monic it is easy to see that for large $\beta > 0$ there holds $\mathrm{bnr}_A(\beta; \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) < \beta^m$ and thus there needs to exist a $\tilde{\beta} > \beta_{bnr}$ satisfying

$$\mathrm{bnr}(A(\tilde{\beta}); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) = \mathrm{bnr}_A(\tilde{\beta}; \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) = \tilde{\beta}^m.$$

Since $A(\tilde{\beta})$ is entrywise nonnegative it follows from Proposition A.30.(i) that its block numerical radius is contained in its block numerical range, $\tilde{\beta}^m \in \Theta(A(\tilde{\beta}); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$. But then $\tilde{\beta} \in \Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ contradicting $\tilde{\beta} > \mathrm{bnr}(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$. ☐

1.3. The infinite graph $G_m(A_l, \ldots, A_0)$

A useful tool for studying semi-monic matrix polynomials is the *infinite graph*

$$G_m(A_l, \ldots, A_0) = (V; E)$$

for matrices $A_0, \ldots, A_l \in \mathbb{R}_+^{n,n}$ and $m \in \mathbb{N}$. We define the set of vertices of $G_m(A_l, \ldots, A_0)$ via

$$V = \{(r, p) : r \in \langle n \rangle, p \in \mathbb{Z}\}$$

and the set of its edges via

$$E = \{[(r, p), (s, q)] : A_{m-p+q}(r, s) > 0\},$$

where $A_{m+p-q}(r, s)$ denotes the entry with coordinates (r, s) in the matrix A_{m+p-q}. For $(r, p) \in V$ we call r the *phase* and p the *level* of the vertex. A sequence of edges

$$[(r_0, p_0), (r_1, p_1)], [(r_1, p_1), (r_2, p_2)], \ldots, [(r_{w-1}, p_{w-1}), (r_w, p_w)]$$

is called a *path* of length w connecting (r_0, p_0) with (r_w, p_w) and we might also write

$$(r_0, p_0) \to (r_1, p_1) \to \cdots \to (r_w, p_w).$$

For the above path the number $p_w - p_0$ is called its *level displacement*. Furthermore we call a path $(r_0, p_0) \to \cdots \to (r_w, p_w)$ a *phase cycle* if $r_0 = r_w$.

Then the *index of phase imprimitivity* of the graph $G_m(A_l, \ldots, A_0)$ is defined as the greatest common divisor (g.c.d.) of the level displacements of all of its phase cycles. In the case where every phase cycle has level displacement 0 (which can happen, see [FN05a, Example 4.3]) the index of phase imprimitivity is defined as 0. Moreover, the index of phase imprimitivity is defined to be nonnegative.

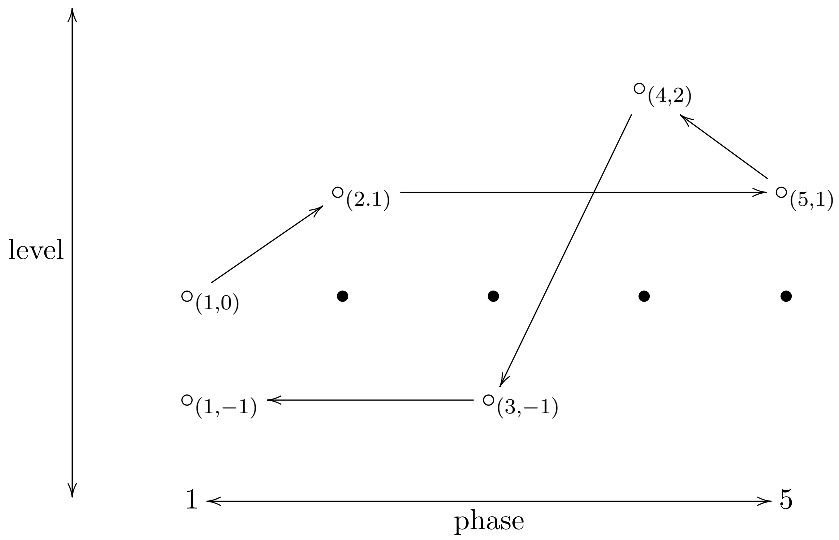
Example 1.17. Consider the matrices

$$A_4 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and let $m = 3$. Let us construct a path in the graph $G_3(A_4, A_3, A_0)$: Starting in the vertex $(r, p) = (1, 0)$ we see that only the matrix A_4 has a nonzero entry in its first row. Setting $q = 1$ we get $A_{m-p+q}(r, 2) = A_4(1, 2) > 0$ and therefore $[(1, 0), (2, 1)]$ is an edge in $G_3(A_4, A_3, A_0)$ (it is the only edge starting in $(1, 0)$). Continuing like that we get a path

$$(1, 0) \to (2, 1) \to (5, 1) \to (4, 2) \to (3, -1) \to (1, -1).$$

This path can be illustrated in a two-dimensional diagram as follows:



The path we constructed is in fact a phase cycle with level displacement -1. It is also essentially the only phase cycle in $G_3(A_4, A_3, A_0)$, as it can be easily seen that changing the starting level only 'shifts' the path up or down. As such the index of phase imprimitivity of $G_3(A_4, A_3, A_0)$ is 1.

A concept similar to the infinite graph above was considered in [GHT96, p.132]. Note that the graph associated with a single matrix $A_0 \in \mathbb{R}_+^{n,n}$ can be expressed by $G_1(A_0)$. Then the index of phase imprimitivity of $G_1(A_0)$ coincides with the usual index of imprimitivity of A_0 used in the Perron-Frobenius theory. Thus the graph $G_m(A_l, \ldots, A_0)$ can be seen as an extension of the usual graph associated with an entrywise nonnegative matrix.
In [Har11] the infinite graph was used to derive results for the spectrum of semi-monic Perron-Frobenius polynomials similar to our main result for the block numerical range

in the following section. We collect some properties of these infinite graphs which will be needed in the proof of our main result. The proofs are straight forward and can be found in [Har11, pp. 66-67].

Lemma 1.18. *For the graph $G_m(A_l, \dots, A_0)$ we have that $[(r,p),(s,q)] \in E$ implies that $[(r,p+u),(s,q+u)] \in E$ for all $u \in \mathbb{Z}$.*
Moreover for any $u \in \mathbb{Z}$ the paths $(r_0,p_0) \to (r_1,p_1) \to \cdots \to (r_w,p_w)$ and $(r_0,p_0+u) \to (r_1,p_1+u) \to \cdots \to (r_w,p_w+u)$ have the same level displacement.

Lemma 1.19. *Let $A_0, \dots, A_l \in \mathbb{R}_+^{n,n}$ and $m \in \mathbb{N}$. Then for $r, s \in \langle n \rangle$ the following assertions are equivalent:*

- *(i) There exists a path from r to s in the directed graph associated with the matrix $A_0 + \cdots + A_l \in \mathbb{R}_+^{n,n}$.*
- *(ii) For all $p \in \mathbb{Z}$ there exists a $q \in \mathbb{Z}$ such that there is a path from (r,p) to (s,q) in $G_m(A_l, \dots, A_0)$.*
- *(iii) For all $q \in \mathbb{Z}$ there exists a $p \in \mathbb{Z}$ such that there is a path from (r,p) to (s,q) in $G_m(A_l, \dots, A_0)$.*

For the next Lemma recall that a matrix $B \in \mathbb{R}_+^{n,n}$ is *irreducible* if and only if its associated directed graph is strongly connected, i.e. there exists a path between any two vertices.

Lemma 1.20. *Let $A_0, \dots, A_l \in \mathbb{R}_+^{n,n}$ and $m \in \mathbb{N}$. Then the following assertions are equivalent:*

- *(i) $A_0 + \cdots + A_l \in \mathbb{R}_+^{n,n}$ is irreducible.*
- *(ii) For all $r, s \in \langle n \rangle$ and $p \in \mathbb{Z}$ there exists a $q \in \mathbb{Z}$ such that there is a path from (r,p) to (s,q) in $G_m(A_l, \dots, A_0)$.*
- *(iii) For all $r, s \in \langle n \rangle$ and $q \in \mathbb{Z}$ there exists a $p \in \mathbb{Z}$ such that there is a path from (r,p) to (s,q) in $G_m(A_l, \dots, A_0)$.*

We will need one more lemma which is usually attributed to I. Schur. A proof can be found in [BR91, Lemma 3.4.2].

Lemma 1.21. *Let M be a nonempty set of integers which is closed under addition and let $d \in \mathbb{N}$ be the greatest common divisor of M. Then we have $kd \in M$ for all but finitely many $k \in \mathbb{N}$.*

1.4. A Perron-Frobenius type result for the block numerical range

In this section we will prove a Perron-Frobenius type result for the block numerical range of semi-monic matrix polynomials. Perron-Frobenius type results for a single matrix are known for the spectrum (see e.g. [HJ85] and [Min88, Chapter 1.4]) and the numerical range (see. e.g. [Iss66, MPT02, PT04]). More recently they were proven for the block

numerical range of a matrix (see [FH08]) and of an operator in a Hilbert lattice (see [Rad14]).
These results were extended to monic matrix polynomials in (see e.g. [PT04]) for the spectrum and in [MPT02] for the numerical range. They were further extended to the semi-monic case in [Har11] and [FK15] respectiviely. Recently an analogue was proved in [FN15] for the spectrum of semi-monic operator polynomials with coefficients in a Banach lattice.
Results for the block numerical range of monic polynomials can be found in [FH08] for matrix coefficients, and in [Rad14, RTW14] for operator coefficients in a Hilbert lattice.

The setting for this section is that of semi-monic matrix polynomials

$$Q(\lambda) = \lambda^m I - A(\lambda) = \lambda^m I - \sum_{j=0}^{l} \lambda^j A_j \tag{1.10}$$

where we assume the coefficients to be entrywise nonnegative, i.e. $A_j \in \mathbb{R}_+^{n,n}$. We will call $A(\cdot)$ *irreducible* if $\sum_{j=0}^{l} A_j$ is irreducible or if, equivalently, the real positive matrix $A(\beta)$ is irreducible for one (and then for all) $\beta > 0$.
The block numerical range will be formed with respect to the product space $\mathbb{C}^n = \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_p}$. For an exposition on block numerical ranges of matrices we point the reader to Section A.5. In particular recall the index sets $\mathcal{I}_k$ and the integer function $\iota(\cdot)$ from Definition A.28.
The main result, which is illustrated in Figure 2, is as follows.

Theorem 1.22. *Let $\mathbb{C}^n = \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_p}$ and $A(\lambda) = \sum_{j=0}^{l} \lambda^j A_j$ an irreducible matrix polynomial with entrywise nonnegative coefficients and $Q(\lambda) = \lambda^m I - A(\lambda)$ its corresponding semi-monic polynomial. Let further d be the index of phase imprimitivity of the graph $G_m(A_l, \ldots, A_0)$. Then for all $\beta > 0$ with*

$$\beta^m = \operatorname{bnr}(A(\beta); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) \tag{1.11}$$

the following statements hold:

(i) If $d = 0$ then $\mathbb{T}_\beta \subset \Theta(Q(\beta); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$.

(ii) If $d \geq 1$ then $\Theta(Q(\beta); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) \cap \mathbb{T}_\beta = \{\beta e^{2\pi i \frac{k}{d}} : k = 1, \ldots, d\}$.

Proof. We first prove (ii).
"$\subseteq$": Take an element of $\Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) \cap \mathbb{T}_\beta$, i.e. an $\beta\omega$ with $\omega \in \mathbb{T}_1$. Then there exists some $X \in S_{\text{att}}^{op}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ (with corresponding $x \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$) and some $u \in \mathbb{C}^p$ such that

$$(\beta\omega)^m u = X^* A(\beta\omega) X u. \tag{1.12}$$

We further have that

$$\operatorname{spr}(X^* A(\beta\omega) X) \leq \operatorname{bnr}(A(\beta\omega)) \leq \operatorname{bnr}(A(\beta)) = \beta^m.$$

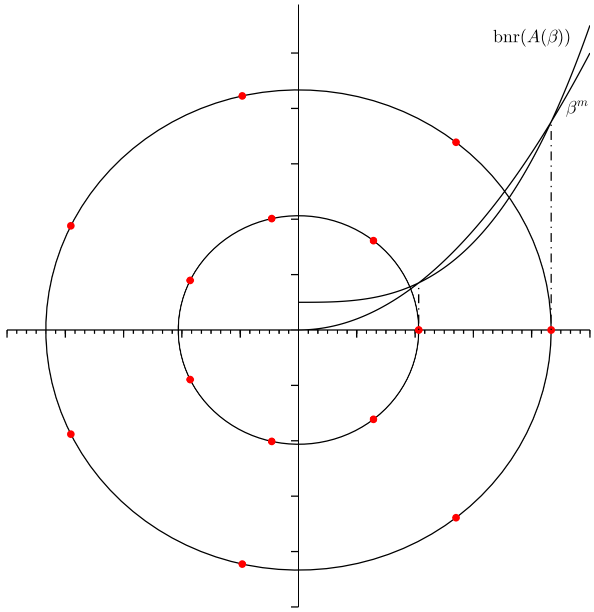


Figure 2: The dots represent values of $\Theta((Q(\cdot);\mathbb{C}^{n_1},\dots,\mathbb{C}^{n_p})$ on the circles with radii β satisfying $\beta^m = \mathrm{bnr}(A(\beta);\mathbb{C}^{n_1},\dots,\mathbb{C}^{n_p})$.

Noting that $|X| \in S^{op}_{\mathrm{att}}$ (with corresponding $|x| \in S_{\mathrm{att}}$) we then have

$$\beta^m = \mathrm{spr}(X^*A(\beta\omega)X) \le \mathrm{spr}(|X^*A(\beta\omega)X|) \le \mathrm{spr}(|X|^*A(\beta)|X|) \le \beta^m.$$

This implies $\mathrm{bnr}(A(\beta)) = \mathrm{spr}(|X|^*A(\beta)|X|)$ and since $A(\beta)$ is irreducible and $|X| \ge 0$ it follows from Proposition A.30.(iii) that $|x| \gg 0$. Then by Proposition A.30.(ii) the matrix $|X|^*A(\beta)|X|$ is irreducible. We also know that

$$\mathrm{spr}(|X|^*A(\beta)|X|)|u| = \beta^m|u| = |X^*A(\beta\omega)Xu| \le |X|^*A(\beta)|X|\,|u|$$

which implies together with the irreducibility of $|X|^*A(\beta)|X|$ that the vector $|u|$ is an eigenvector of $|X|^*A(\beta)|X|$ to the eigenvalue β^m and is thus strictly positive.
Setting $Y := |X|$ and $v := |u|$ this reads as

$$\beta^m v = Y^*A(\beta)Yv.$$

Taking a coordinate $h \in \langle p \rangle$ we can write

$$\begin{aligned}\beta^m v_h &= \sum_{k=1}^{p} (Y^* A(\beta) Y)_{hk} v_k \\ &= \sum_{j=1}^{l} \sum_{k=1}^{p} \beta^j (Y^* A_j Y)_{hk} v_k \\ &= \sum_{j=1}^{l} \sum_{k=1}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} \beta^j \bar{y}_r A_j(r,s) y_s v_k.\end{aligned}$$

Now dividing by the left-hand side we arrive at

$$1 = \sum_{j=1}^{l} \sum_{k=1}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} \beta^{j-m} \bar{y}_r A_j(r,s) y_s \frac{v_k}{v_h} \tag{1.13}$$

where the summands on the right-hand side are all nonnegative. Now doing the same for (1.12) we get

$$\begin{aligned}1 &= \sum_{j=1}^{l} \sum_{k=1}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} (\beta\omega)^{j-m} \bar{x}_r A_j(r,s) x_s \frac{u_k}{u_h} \\ &= \sum_{j=1}^{l} \sum_{k=1}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} \left[\beta^{j-m} \bar{y}_r A_j(r,s) y_s \frac{v_k}{v_h}\right] \left[\frac{\bar{x}_r}{\bar{y}_r} \frac{x_s}{y_s} \frac{u_k}{v_k} \frac{v_h}{u_h} \omega^{j-m}\right]. \end{aligned} \tag{1.14}$$

Moreover

$$\left|\frac{\bar{x}_r}{\bar{y}_r} \frac{x_s}{y_s} \frac{u_k}{v_k} \frac{v_h}{u_h} \omega^{j-m}\right| = \left|\frac{\bar{x}_r}{\bar{y}_r}\right| \left|\frac{x_s}{y_s}\right| \left|\frac{u_k}{v_k}\right| \left|\frac{v_h}{u_h}\right| \left|\omega^{j-m}\right| = 1. \tag{1.15}$$

Due to (1.13) we now see that (1.14) is a convex combination of numbers with absolute value 1 whose sum is equal to 1. This is only possible if for all r, s where $A_j(r,s) > 0$ we have that

$$1 = \frac{\bar{x}_r}{\bar{y}_r} \frac{x_s}{y_s} \frac{u_k}{v_k} \frac{v_h}{u_h} \omega^{j-m}.$$

Note that in the above equation $h = \iota(r)$ and $k = \iota(s)$. We can therefore write it as

$$\frac{y_s}{x_s} \frac{v_{\iota(s)}}{v_{\iota(r)}} = \frac{\bar{x}_r}{\bar{y}_r} \frac{u_{\iota(s)}}{u_{\iota(r)}} \omega^{j-m}. \tag{1.16}$$

Now take any phase cycle $(r_0, p_0) \to \ldots \to (r_w, p_w)$ in $G_m(A_l, \ldots, A_0)$ of some length $w \in \mathbb{N}$ (which is possible by Lemma 1.20) and denote its level displacement by $\tilde{d}$. Then $r_0 = r_w$ and $A_{m-p_{t-1}+p_t}(r_{t-1}, r_t) > 0$ for $t \in \langle w \rangle$. Setting $j_t = m - p_{t-1} + p_t$ we can then write $\tilde{d} = \sum_{t=1}^{w} j_t - m$ and with (1.16) it follows that

$$\begin{aligned}
\omega^{\tilde{d}} &= \frac{u_{\iota(r_0)}}{u_{\iota(r_w)}} \prod_{t=1}^{w} \omega^{j_t - m} \frac{u_{\iota(r_t)}}{u_{\iota(r_{t-1})}} \frac{\bar{x}_{r_{t-1}}}{\bar{y}_{r_{t-1}}} \prod_{t=1}^{w} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} \\
&= \frac{u_{\iota(r_0)}}{u_{\iota(r_w)}} \prod_{t=1}^{w} \frac{v_{\iota(r_t)}}{v_{\iota(r_{t-1})}} \frac{y_{r_t}}{x_{r_t}} \prod_{t=1}^{w} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} \\
&= \underbrace{\frac{u_{\iota(r_0)}}{u_{\iota(r_w)}} \frac{v_{\iota(r_w)}}{v_{\iota(r_0)}}}_{:=I} \underbrace{\prod_{t=1}^{w} \frac{y_{r_t}}{x_{r_t}} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}}}_{:=II}.
\end{aligned}$$

Since $r_0 = r_w$ we immediately see that $I = 1$. For part II calculate

$$II = \prod_{t=1}^{w} \frac{y_{r_t}}{x_{r_t}} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} = \prod_{t=1}^{w} \frac{y_{r_t}}{x_{r_t}} \frac{\bar{y}_{r_t}}{\bar{x}_{r_t}} = \prod_{t=1}^{w} \frac{|y_{r_t}|^2}{|x_{r_t}|^2} = 1,$$

where we again used that $r_0 = r_w$.
Thus $\omega^{\tilde{d}} = 1$. In order to see that then also $\omega^d = 1$ consider the set

$$M = \left\{ \begin{matrix} \text{the set of all level displacements of phase cycles} \\ \text{in } G_m(A_l, \ldots, A_0) \text{ and their sums} \end{matrix} \right\}$$

which is closed under addition. Obviously for an element $\hat{d} \in M$ there still holds $\omega^{\hat{d}} = 1$. Moreover the index of phase imprimitivity d is the greatest common divisor of M. Now by Lemma 1.21 there exists a $k \in \mathbb{N}$ such that kd and $(k+1)d$ are both in M, i.e. $\omega^{kd} = \omega^{(k+1)d} = 1$. Thus the increment d must also fulfil $\omega^d = 1$. It follows that

$$\beta\omega \in \{e^{2\pi i \frac{k}{d}} : k = 0, \ldots, d-1\}.$$

"$\supseteq$": For this inclusion choose an $\omega \in \mathbb{T}_1$ with $\omega^d = 1$. Since $A(\cdot)$ is irreducible by Proposition A.30 we can find a strictly positive $y \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ (with corresponding $Y \in S_{\text{att}}^{op}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$) and $v \in \mathbb{C}^p$ satisfying $\beta^m v = Y^* A(\beta) Y v$.
Our goal is to construct an $x \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ and $u \in \mathbb{C}^p$ satisfying $(\beta\omega)^m u = X^* A(\beta\omega) X u$. Set $u = v$ and $x_1 = y_1$. For $s \in \langle 2, n \rangle$ take a path $(r_0, p_0) \to \ldots \to (r_w, p_w)$ in $G_m(A_l, \ldots, A_0)$ such that $r_0 = 1$ and $r_w = s$ (which is possible by Lemma 1.20). Further by Lemma 1.18 we can assume w.l.o.g. that $p_w = 0$. Thus the path will have level displacement $-p_0$.
Now define x_s recursively via

$$x_{r_t} = y_{r_t} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} \omega^{p_{t-1} - p_t}, \qquad t \in \langle w \rangle.$$

Claim: The above construction is well defined, i.e. it is independent of the specific path.
To see the claim first note that $|x_{r_t}| = |y_{r_t}|$, $t \in \langle w \rangle_0$ and we can thus write

$$x_{r_t} = y_{r_t} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} \omega^{p_{t-1} - p_t} = y_{r_t} \frac{x_{r_{t-1}}}{y_{r_{t-1}}} \frac{|y_{r_{t-1}}|^2}{|x_{r_{t-1}}|^2} \omega^{p_{t-1} - p_t} = y_{r_t} \frac{x_{r_{t-1}}}{y_{r_{t-1}}} \omega^{p_{t-1} - p_t}.$$

Then

$$\begin{aligned}
x_s &= y_s \frac{x_{r_{w-1}}}{y_{r_{w-1}}} \omega^{p_{w-1}-p_w} \\
&= y_s \frac{y_{r_{w-1}}}{y_{r_{w-1}}} \frac{x_{r_{w-2}}}{y_{r_{w-2}}} \omega^{p_{w-2}-p_{w-1}} \omega^{p_{w-1}-p_w} \\
&\ \ \vdots \\
&= y_s \omega^{p_0 - p_w} = y_s \omega^{p_0}
\end{aligned}$$

Now take another path $(\tilde{r}_0, \tilde{p}_0) \to \ldots \to (\tilde{r}_{\tilde{w}}, \tilde{p}_{\tilde{w}})$ satisfying $\tilde{r}_0 = 1$ and $\tilde{r}_{\tilde{w}} = s$ and again w.l.o.g. $\tilde{p}_{\tilde{w}} = 0$ (so that this path has level displacement $-\tilde{p}_0$). This path gives rise to a $\tilde{x}_s$ and by the same calculation as above we get

$$\tilde{x}_s = y_s \omega^{\tilde{p}_0}.$$

In the last step consider a third path from $(s, 0)$ to $(1, \hat{d})$ with level displacement $\hat{d}$ (any such path will do). Attaching this path to either of the previous two paths gives phase cycles with level displacements $\hat{d} - p_0$ and $\hat{d} - \tilde{p}_0$ respectively. We can now divide x_s by $\tilde{x}_s$ to get

$$\frac{x_s}{\tilde{x}_s} = \omega^{p_0 - \tilde{p}_0} = \omega^{-(\hat{d} - p_0)} \omega^{\hat{d} - \tilde{p}_0} = 1,$$

where the last equality holds because both exponents are level displacements of phase cycles and thus divisible by d. The claim is proved.
The vector x was constructed in a way that whenever $A_j(r, s) > 0$ relation (1.16) is satisfied. We can thus write for $h \in \langle p \rangle$

$$\begin{aligned}
(X^* A(\beta\omega) X u)_h &= \sum_{j=0}^{l} (\beta\omega)^j (X^* A_j X u)_h \\
&= \sum_{j=0}^{l} \sum_{k=1}^{p} (\beta\omega)^j (X^* A_j X)_{hk} u_k \\
&= \sum_{j=0}^{l} \sum_{k=1}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} (\beta\omega)^m \bar{x}_r A_j(r, s) x_s u_k \\
&= \sum_{j=0}^{l} \sum_{k=1}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} \left[\beta^j \bar{y}_r A_j(r, s) y_s v_k\right] \left[\omega^{j-m} \frac{\bar{x}_r}{\bar{y}_r} \frac{x_s}{y_s} \frac{u_k}{v_k} \frac{v_h}{u_h}\right] \frac{u_h}{v_h} \omega^m \\
&= (Y^* A(\beta) Y v)_h \frac{u_h}{v_h} \omega^m \\
&= \beta^m v_h \frac{u_h}{v_h} \omega^m = (\beta\omega)^m u_h
\end{aligned}$$

which shows $\beta\omega \in \Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) \cap \mathbb{T}_\beta$.
(i) Note that $d = 0$ implies that in the proof of the reverse inclusion above we can choose any $\omega \in \mathbb{T}_1$. It then follows that $\mathbb{T}_\beta \subseteq \Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$. □

Remark 1.23. Theorem 1.22 is a generalisation of both [Har11, Theorem 4.23] for the spectral case and [FK15, Theorem 5.2] for the numerical range case. In fact if we form the block numerical range with respect to the trivial product space $\mathbb{C}^n$ or the maximal product space $\mathbb{C}^n = \mathbb{C} \times \cdots \times \mathbb{C}$, then those two results are exactly Theorem 1.22.
Lemma 1.16 implies that Theorem 1.22 is also a generalisation of analogous results for monic matrix polynomials. In particular it is a generalisation of [PT04, Theorem 3.3] for the spectral case and of [PT04, Corollary 5.6] for the numerical range case.

Remark 1.24. Equation (1.11) can be satisfied for several $\beta > 0$. However, the number of elements on the circles of radius β are the same for all $\beta > 0$ satisfying (1.11), since the number of elements depends only on the index of phase imprimitivity of $G_m(A_l, \dots, A_0)$. Equation (1.11) can further be satisfied with respect to different decompositions of the product space $\mathbb{C}^n$ (which might happen for different $\beta > 0$). In particular we can compare the cyclic distribution of block numerical ranges to the cyclic distribution of the spectrum and the numerical range. But as above the actual number of elements on circles $\mathbb{T}_\beta$ remains the same in all cases.
Finally, even changing the coefficient matrices A_j will not result in a different number of cyclic elements as long as the nonzero structure of the A_j's is preserved (though this will most likely result in changed radii $\beta > 0$).

Example 1.25. Let

$$Q_m(\lambda) = \lambda^m - A(\lambda) = \lambda^m - \lambda^4 A_4 - A_0$$

with

$$A_4 = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \end{bmatrix}$$

and $m \in \mathbb{N}$. Note that the matrix $A(1)$ is irreducible. Let us first consider the infinite graph $G_m(A_4, A_0)$. Since $A(1)$ contains exactly one nonzero element in every row and column there is effectively only on phase cycle in $G_m(A_4, A_0)$. Starting in the vertex $(1, 0)$. The cycle then is

$$(1, 0) \to (3, 4 - m) \to (2, 4 - 2m) \to (4, 8 - 3m) \to (1, 8 - 4m).$$

The level displacement of this cycle is $8 - 4m$. The index of phase imprimitivity is therefore $d = |8 - 4m|$.
For the block numerical range we choose the decomposition $\mathbb{C}^4 = \mathbb{C}^2 \times \mathbb{C}^2$. Let us first compute $\operatorname{bnr}(A(\beta); \mathbb{C}^2, \mathbb{C}^2)$ for $\beta > 0$: Take a nonnegative $x \in S_{\text{att}}(\mathbb{C}^2, \mathbb{C}^2)$ with

corresponding $X \in S^{op}_{\text{att}}(\mathbb{C}^2, \mathbb{C}^2)$. Then

$$\begin{aligned}
\operatorname{spr}(X^* A(\beta) X) &= \operatorname{spr}\left[\begin{array}{c|c} 0 & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^* \begin{bmatrix} 2\beta^4 & \\ & 2\beta^4 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \\ \hline \begin{bmatrix} x_3 \\ x_4 \end{bmatrix}^* \begin{bmatrix} & 0.5 \\ 0.5 & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & 0 \end{array}\right] \\
&= \operatorname{spr}\begin{bmatrix} 0 & 2\beta^4(x_1x_3 + x_2x_4) \\ 0.5(x_2x_3 + x_4x_1) & 0 \end{bmatrix} \\
&= \beta^2\sqrt{(x_1x_3 + x_2x_4)(x_3x_2 + x_4x_1)} \\
&= \beta^2\sqrt{x_1x_2 + x_3x_4}
\end{aligned}$$

The term under the root is maximal for $x_1 = x_2 = x_3 = x_4 = \sqrt{0.5}$ and then $\operatorname{bnr}(A(\beta); \mathbb{C}^2, \mathbb{C}^2) = \beta^2$. Therefore

$$\begin{aligned}
& \operatorname{bnr}(A(\beta); \mathbb{C}^2, \mathbb{C}^2) \overset{!}{=} \beta^m \\
\iff \quad & \beta^2 = \beta^m \\
\iff \quad & 1 = \beta
\end{aligned}$$

(where we left out the solution $\beta = 0$ since we require $\beta > 0$).
A similar calculation to the above shows that the peripheral block numerical range of $A(\beta)$ for any $\beta > 0$ is

$$\Theta(A(\beta); \mathbb{C}^2, \mathbb{C}^2) \cap \mathbb{T}_{\beta^2} = \{\beta^2 \exp(2\pi i \frac{k}{4}) : k = 0, \ldots, 3\}$$

(alternatively one could get this result by applying [FH08, Theorem 4.6]).
We claim that for any $\omega \in \mathbb{C}$ with $|\omega| = 1$ there holds

$$\Theta(A(\beta\omega); \mathbb{C}^2, \mathbb{C}^2) = \omega^2 \Theta(A(\beta); \mathbb{C}^2, \mathbb{C}^2).$$

To see this take any vector $x \in S_{\text{att}}(\mathbb{C}^2, \mathbb{C}^2)$ and write

$$\begin{aligned}
\Sigma(X^* A(\beta\omega) X) &= \Sigma\left[\begin{array}{c|c} 0 & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^* \begin{bmatrix} 2\beta^4\omega 4 & \\ & 2\beta^4\omega^4 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \\ \hline \begin{bmatrix} x_3 \\ x_4 \end{bmatrix}^* \begin{bmatrix} & 0.5 \\ 0.5 & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & 0 \end{array}\right] \\
&= \Sigma\begin{bmatrix} 0 & 2\beta^4\omega^4(\bar{x}_1x_3 + \bar{x}_2x_4) \\ 0.5(\bar{x}_3x_2 + \bar{x}_4x_1) & 0 \end{bmatrix}.
\end{aligned}$$

The claim now follows because

$$\begin{aligned}
& \Sigma\begin{bmatrix} 0 & 2\beta^4\omega^4(\bar{x}_1x_3 + \bar{x}_2x_4) \\ 0.5(\bar{x}_3x_2 + \bar{x}_4x_1) & 0 \end{bmatrix} \\
& \quad = \omega^2 \Sigma\begin{bmatrix} 0 & 2\beta^4(\bar{x}_1x_3 + \bar{x}_2x_4) \\ 0.5(\bar{x}_3x_2 + \bar{x}_4x_1) & 0 \end{bmatrix}.
\end{aligned}$$

It follows that

$$\Theta(A(\beta\omega);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_{\beta^2}=\{\beta^2\omega^2\exp(2\pi i\frac{k}{4}):k=0,\dots,3\}.$$

Now let $\beta=1$ such that $\operatorname{bnr}(A(\beta);\mathbb{C}^2,\mathbb{C}^2)=\beta^m$. We are interested in the distribution of

$$\Theta(Q_m(\cdot);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_1.$$

Our above calculations show that an element $\omega\in\mathbb{C}$, $|\omega|=1$ satisfies $\omega\in\Theta(Q_m(\cdot);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_1$ if and only if

$$\begin{aligned} & \omega^m\in\{\omega^2\exp(2\pi i\frac{k}{4}):k=0,\dots,3\} \\ \iff\quad & \omega^{m-2}\in\{\pm1,\pm i\}. \end{aligned}\tag{1.17}$$

This already implies a cyclic distribution. Let us now look at some specific values for m:

- $\mathbf{m=2}$: Condition (1.17) is satisfied for any ω with absolute value 1. Therefore

  $$\Theta(Q_2(\cdot);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_1=\mathbb{T}_1.$$

  This is consistent with Theorem 1.22 since the index of phase imprimitivity is $d=|8-2m|=0$.

- $\mathbf{m=3}$: Condition (1.17) implies that

  $$\Theta(Q_3(\cdot);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_1=\{\pm1,\pm i\}.$$

  This is again consistent with Theorem 1.22 since the index of phase imprimitivity is $d=4$.

- $\mathbf{m=4}$: Now condition (1.17) implies that

  $$\Theta(Q_4(\cdot);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_1=\{\exp(2\pi i\frac{k}{8}):k=0,\dots,7\}.$$

  As expected the index of phase imprimitivity is $d=8$.

- $\mathbf{m=5}$: We are now in the monic case. Here we have

  $$\Theta(Q_5(\cdot);\mathbb{C}^2,\mathbb{C}^2)\cap\mathbb{T}_1=\{\exp(2\pi i\frac{k}{12}):k=0,\dots,11\}$$

  and $d=12$.

By slightly altering the proof of Theorem 1.22 we can obtain a result about the rotation invariance of the whole block numerical range of semi-monic matrix polynomials. It is not necessary for equation (1.11) to be satisfied, but in turn we obtain a slightly weaker statement.

Theorem 1.26. *Let $A(\lambda) = \sum_{j=0}^{l} \lambda^j A_j$ be an irreducible matrix polynomial with entrywise nonnegative coefficients and $Q(\lambda) = \lambda^m - A(\lambda)$ its corresponding semi-monic polynomial for some $m \in \mathbb{N}$. Let further d be the index of phase imprimitivity of the associated graph $G_m(A_l, \ldots, A_0)$ and assume $d \geq 1$. Then $\Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ is invariant under rotation with the angle $\theta = \frac{2\pi}{d}$, i.e.*

$$\Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) = e^{i\frac{2\pi}{d}} \Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}).$$

Moreover if there exists a $\beta > 0$ such that $\operatorname{nur}(A(\beta)) = \beta^m$ then θ is the smallest such angle.

Proof. The proof is conceptually very similar to the second inclusion in part (ii) of the proof of Theorem 1.22. Let $\lambda \in \Theta(Q(\cdot); \mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ and $\omega \in \mathbb{T}_1$ with $\omega^d = 1$. Then there exist $y \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ (with corresponding $Y \in S_{\text{att}}^{op}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$) and $v \in \mathbb{C}^p$ satisfying $\beta^m v = Y^* A(\beta) Y v$. Our goal is to construct an $x \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ and $u \in \mathbb{C}^p$ satisfying $(\lambda\omega)^m u = X^* A(\beta\omega) X u$. Set $u = v$ and $x_1 = y_1$. For $s \in \langle 2, n\rangle$ take a path $(r_0, p_0) \to \ldots \to (r_w, p_w)$ in $G_m(A_l, \ldots, A_0)$ such that $r_0 = 1$ and $r_w = s$ (which is possible by Lemma 1.20). Further by Lemma 1.18 we can assume w.l.o.g. that $p_w = 0$. Thus the path will have level displacement $-p_0$.
Now define x_s recursively via

$$x_{r_t} = \begin{cases} y_{r_t} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} \omega^{p_{t-1} - p_t}, & \bar{x}_{r_{t-1}} \neq 0 \\ y_{r_t} \omega^{p_0 - p_{t-1}}, & \bar{x}_{r_{t-1}} = 0 \end{cases} \qquad t \in \langle w \rangle.$$

Claim: The above construction is well defined, i.e. it is independent of the specific path.
To see the claim first note that $|x_{r_t}| = |y_{r_t}|$, $t \in \langle w \rangle_0$ and we can thus write

$$x_{r_t} = y_{r_t} \frac{\bar{y}_{r_{t-1}}}{\bar{x}_{r_{t-1}}} \omega^{p_{t-1} - p_t} = y_{r_t} \frac{x_{r_{t-1}}}{y_{r_{t-1}}} \frac{|y_{r_{t-1}}|^2}{|x_{r_{t-1}}|^2} \omega^{p_{t-1} - p_t} = y_{r_t} \frac{x_{r_{t-1}}}{y_{r_{t-1}}} \omega^{p_{t-1} - p_t}$$

if $\bar{x}_{r_{t-1}} \neq 0$. Then

$$\begin{aligned} x_s &= y_s \frac{x_{r_{w-1}}}{y_{r_{w-1}}} \omega^{p_{w-1} - p_w} \\ &= y_s \frac{y_{r_{w-1}}}{y_{r_{w-1}}} \frac{x_{r_{w-2}}}{y_{r_{w-2}}} \omega^{p_{w-2} - p_{w-1}} \omega^{p_{w-1} - p_w} \\ &\vdots \\ &= y_s \omega^{p_0 - p_w} = y_s \omega^{p_0} \end{aligned}$$

Note that the above recursive expansion of x_s might stop early if the vector y has a zero entry, but the result will remain the same. Now take another path $(\tilde{r}_0, \tilde{p}_0) \to \ldots \to (\tilde{r}_{\tilde{w}}, \tilde{p}_{\tilde{w}})$ satisfying $\tilde{r}_0 = 1$ and $\tilde{r}_{\tilde{w}} = s$ and again w.l.o.g. $\tilde{p}_{\tilde{w}} = 0$ (so that this path has level displacement $-\tilde{p}_0$). This path gives rise to a $\tilde{x}_s$ and by the same calculation as above we get

$$\tilde{x}_s = y_s \omega^{\tilde{p}_0}.$$

In the last step consider a third path from $(s,0)$ to $(1,\hat{d})$ with level displacement $\hat{d}$ (any such path will do). Attaching this path to either of the previous two paths gives phase cycles with level displacements $\hat{d}-p_0$ and $\hat{d}-\tilde{p}_0$ respectively. We can now divide x_s by $\tilde{x}_s$ to get

$$\frac{x_s}{\tilde{x}_s} = \omega^{p_0-\tilde{p}_0} = \omega^{-(\hat{d}-p_0)}\omega^{\hat{d}-\tilde{p}_0} = 1,$$

where the last equality holds because both exponents are level displacements of phase cycles and thus divisible by d. The claim is proved.
The vector x was constructed in a way that whenever $A_j(r,s) > 0$ the following relation (which is the analogue to relation (1.16), taking into account that we chose $u = v$) is satisfied:

$$\omega^{j-m}\bar{x}_r x_s = \bar{y}_r y_s. \tag{1.18}$$

In the following equation we will use $\sum'$ to denote that we leave out all elements of the sum that are equal to zero. We can thus write for $h \in \langle p \rangle$

$$\begin{aligned}
(X^*A(\lambda\omega)Xu)_h &= \sum_{j=0}^{l}(\lambda\omega)^j(X^*A_jXu)_h \\
&= \sum_{j=0}^{l}{}' \sum_{k=1}^{p}{}'(\lambda\omega)^j(X^*A_jX)_{hk}u_k \\
&= \sum_{j=0}^{l}{}' \sum_{k=1}^{p}{}' \sum_{r\in\mathcal{I}_h}{}' \sum_{s\in\mathcal{I}_k}{}'(\lambda\omega)^m\bar{x}_r A_j(r,s)x_s u_k \\
&= \sum_{j=0}^{l}{}' \sum_{k=1}^{p}{}' \sum_{r\in\mathcal{I}_h}{}' \sum_{s\in\mathcal{I}_k}{}' \left[\lambda^j\bar{y}_r A_j(r,s)y_s v_k\right]\left[\omega^{j-m}\frac{\bar{x}_r}{\bar{y}_r}\frac{x_s}{y_s}\frac{u_k}{v_k}\frac{v_h}{u_h}\right]\frac{u_h}{v_h}\omega^m \\
&= (Y^*A(\lambda)Yv)_h\frac{u_h}{v_h}\omega^m \\
&= \lambda^m v_h\frac{u_h}{v_h}\omega^m = (\lambda\omega)^m u_h
\end{aligned}$$

which shows $\lambda\omega \in \Theta(Q(\cdot);\mathbb{C}^{n_1},\ldots,\mathbb{C}^{n_p})$.
The second assertion follows immediately from Theorem 1.22 since a smaller angle would imply that we would get additional elements of the block numerical range on the circle $\mathbb{T}_\beta$. $\square$

1.5. The special case of monic matrix polynomials

In this section we consider the case where $Q(\cdot)$ is a monic matrix polynomial, that is

$$Q(\lambda) = \lambda^m I - \lambda^{m-1}A_{m-1} - \cdots - A_0 = \lambda^m I - A(\lambda) \tag{1.19}$$

where $m \in \mathbb{N}$ and the $A_j \in \mathbb{R}_+^{n,n}$. For a matrix polynomial of this form we can define the (negative of the) companion matrix

$$C_Q := -\operatorname{Comp}[-A_{m-1}, \dots, -A_0] = \begin{bmatrix} A_{m-1} & \dots & A_1 & A_0 \\ I_n & & & \\ & \ddots & & \\ & & I_n & \end{bmatrix} \in \mathbb{R}^{mn,mn} \tag{1.20}$$

where I_n is the identity operator in $\mathbb{R}^{n,n}$. It is well known that C_Q is a linearisation of $Q(\cdot)$ (for more on degree-reductions and linearisations of operator polynomials such as the companion matrix see Section 3). Since C_Q is again entrywise nonnegative we can associate it with a directed graph

$$G_Q = \{V, E\}$$

with vertices $V = \{1, \dots, nm\}$ and edges given by the nonzero entries of C_Q.
In the Perron-Frobenius theory for matrix polynomials it is well known (see e.g. [PT04]) that, given C_Q is irreducible, the index of imprimitivity (or cyclic index = gcd of the lengths of all cycles in G_Q)) of the graph G_Q is equal to the number of distinct eigenvalues of $Q(\cdot)$ with maximal modulus.
Given that $A(\cdot)$ is irreducible (which follows from C_Q being irreducible) we can apply Theorem 1.22 with the one dimensional decomposition $\mathbb{C}^n = \mathbb{C} \times \dots \times \mathbb{C}$ to see that the number of eigenvalues with maximal modulus of $Q(\cdot)$ also coincides with the index of phase imprimitivity of the infinite graph $G_m(A_{m-1}, \dots, A_0)$.
Motivated by this we study how $G_m(A_{m-1}, \dots, A_0)$ and the graph associated with C_Q are related. We will see that the equality of the index of phase imprimitivity of $G_m(A_{m-1}, \dots, A_0)$ and the index of imprimitivity of G_Q holds even without any irreducibility assumptions.

Lemma 1.27. *Let $[(r_0, p_0), (r_1, p_1)]$ be an edge in $G_m(A_{m-1}, \dots, A_0)$. Then there exists a path $s_0 \to \dots \to s_{p_0 - p_1}$ of length $p_0 - p_1$ in G_Q such that $r_0 = s_0$ and $r_1 = s_{p_0 - p_q}$.*

Proof. Since $[(r_0, p_0), (r_1, p_1)]$ is an edge in $G_m(A_{m-1}, \dots, A_0)$ it follows that

$$A_{m-(p_0-p_1)}(r_0, r_1) > 0$$

(note that necessarily $p_0 - p_1 \in \langle m \rangle$). This implies that

$$C_Q(r_0, r_1 + (p_0 - p_1 - 1)n) > 0$$

and thus $[r_0, r_1 + (p_0 - p_1 - 1)n]$ is an edge in G_Q. If $p_0 - p_1 = 1$ this already completes the proof.
Otherwise, due to the identities on the lower minor diagonal of the matrix C_Q, we see that

$$[r_1 + (p_0 - p_1 - 1)n, r_1 + (p_0 - p_1 - 2)n]$$

is an edge in G_Q. Repeating this procedure we construct edges

$$\begin{gathered}[r_1 + (p_0 - p_1 - 2)n, r_1 + (p_0 - p_1 - 3)n] \\ \vdots \\ [r_1 + 2n, r_1 + n] \\ [r_1 + n, r_1].\end{gathered}$$

By setting $s_0 = r_0$ and $s_g = r_1 + (p_0 - p_1 - g)n$ for $g \in \langle 2, p_0 - p_1 \rangle$ we have constructed a path

$$s_0 \to s_1 \to \ldots \to s_{p_0 - p_1}$$

in G_Q of length $p_0 - p_1$. $\square$

Lemma 1.28. *Let $s_0 \to \ldots \to s_q$ be a cycle in G_Q. Then there exists a $g \in \langle q \rangle_0$ such that $s_g \in \langle n \rangle$.*

Proof. Assume that for all $i \in \langle q \rangle$ there holds $s_i \notin \langle n \rangle$, that is $s_i \in \langle n + 1, mn \rangle$. Then

$$C_Q(s_i, s) = 0 \quad \forall s \geq s_i$$

since the identities on the lower minor diagonal of C_Q are the only nonzero elements in these rows. That implies

$$s_0 < s_1 < \ldots < s_q$$

which contradicts $s_0 \to \ldots \to s_q$ being a cycle. $\square$

Theorem 1.29. *The following assertions hold:*

(i) Let $(r_0, p_0) \to \ldots \to (r_w, p_w)$ be a phase cycle in $G_m(A_{m-1}, \ldots, A_0)$ of length w and level displacement $r_w - r_0$. Then there exists a cycle $s_0 \to \ldots \to s_{p_0 - p_w}$ in G_Q of length $p_0 - p_w$ such that $r_0 = s_0 = s_{p_0 - p_w}$.

(ii) Let $s_0 \to \ldots \to s_q$ be a cycle in G_Q of length q. Then there exists a phase cycle $(r_0, p_0) \to \ldots \to (r_w, p_w)$ in $G_m(A_{m-1}, \ldots, A_0)$ with level displacement $p_w - p_0 = -q$.
Additionally if $s_0 \in \langle n \rangle$ then the phase cycle can be chose such that $s_0 = s_q = r_0 = r_w$.

Proof. (i) This follows immediately by consecutively applying Lemma 1.27 w times.
(ii) Due to Lemma 1.28 we can assume w.l.o.g. that $s_0 = s_q \in \langle n \rangle$. We can thus define $(r_0, p_0) = (s_0, 0)$. Now there exist $\tilde{s}_1 \in \langle n \rangle$ and $k_1 \in \langle m \rangle$ such that $s_1 = \tilde{s}_1 + (k_1 - 1)m$ and $A_{m-k}(s_0, \tilde{s}_1) > 0$.
Now similar to the proof of Lemma 1.29 it follows that $s_{k_1} = \tilde{s}_1$. Setting $p_1 = p_0 - k_1$ it follows that $m - k = m - (p_0 - p_1)$ and thus

$$A_{m-(p_0-p_1)}(r_0, s_{k_1}) > 0.$$

We can now set $(r_1, p_1) = (s_{k_1}, p_0 - k_1)$ and see that $[(r_0, p_0), (r_1, p_1)]$ is an edge in $G_m(A_{m-1}, \dots, A_0)$ with level displacement $p_1 - p_0 = -k_1$.
Continuing iteratively we find k_i's and set

$$(r_i, p_i) = \left(s_{\sum_{j=1}^{i} k_j}, p_{i-1} - k_i\right).$$

Denote by w the number of steps until we arrive at s_q. We have now defined a phase cycle

$$(r_0, p_0) \to (r_1, p_1) \to \ldots \to (r_w, p_w)$$

in $G_m(A_{m-1}, \dots, A_0)$ with level displacement

$$p_w - p_0 = \sum_{i=1}^{w} p_i - p_{i-1} = \sum_{i=1}^{w} -k_i = -q.$$

The last assertion follows directly from our construction. □

Corollary 1.30. *The index of phase imprimitivity of $G_m(A_{m-1}, \dots, A_0)$ coincides with the index of imprimitivity of G_Q.*

Proof. By Theorem 1.29 it follows that the sets

$$M_1 = \{|q| : \text{ there exists a phase cycle with level displacement } q \text{ in } G_m(A_{m-1}, \dots, A_0)\}$$

and

$$M_2 = \{q : \text{ there exists a path of length } q \text{ in } G_Q\}$$

coincide. Since the index of phase imprimitivity of $G_m(A_{m-1}, \dots, A_0)$ and the index of imprimitivity of G_Q are the greatest common divisors of the sets M_1 and M_2 respectively the assertion follows. □

1.6. Positive semi-definite and normal coefficients

In this section we will look at eigenvalues of semi-monic operator polynomials

$$Q(\lambda) = \lambda^m I - A(\lambda) = \lambda^m I - \sum_{j=0}^{l} \lambda^j A_j, \tag{1.21}$$

$A_j \in \mathcal{L}(H)$, H a Hilbert space, where the coefficients are either positive semi-definite or normal. In [Wim08] and [Wim11] the author showed that eigenvalues of maximal modulus of monic matrix polynomials are normal and semisimple for these coefficient classes (in [SW10] the result was extended to monic operator polynomials with positive semi-definite coefficients on a Hilbert space). We will do the same for semi-monic operator polynomials on circles with radii $\beta > 0$ satisfying

$$\mathrm{nur}_A(\beta) = \beta^m. \tag{1.22}$$

We first prove a result about scalar semi-monic polynomials.

Theorem 1.31. *Consider the scalar semi-monic polynomial*

$$q(\lambda) = \lambda^m - a(\lambda) = \lambda^m - \sum_{j=0}^{l} \lambda^j a_j,$$

with $a_j \geq 0$ and $m, l \in \mathbb{N}$ and assume that $q(\rho_0) > 0$ for some $\rho_0 > 0$. Let $\beta > 0$ be a positive root of $q(\cdot)$. Then all roots of $q(\cdot)$ on the circle $\mathbb{T}_\beta$ are simple.

Proof. Because the a_j are nonnegative it follows that the function

$$(0, \infty) \to (0, \infty) : \rho \mapsto a(\rho) = \sum_{j=0}^{l} \rho^j a_j$$

is geometrically convex (see e.g [Nic00, Proposition 2.4]). Since $q(\rho_0) > 0$ for some positive ρ we have $a(\rho) < \rho^m$. It then follows that $q(\rho)$ has at most two positive roots and that for such a root $\beta > 0$ there holds $q'(\beta) \neq 0$. To see this note that the real function

$$b(\cdot) : \mathbb{R} \to \mathbb{R} : \tau \mapsto \log(a(e^\tau))$$

is differentiable and convex and the function

$$l(\cdot) : \mathbb{R} \to \mathbb{R} : \tau \mapsto \log(e^\tau)^m = m\tau$$

is linear. The functions $b(\cdot)$ and $l(\cdot)$ are monotone transformations of $\rho \mapsto a(\rho)$ and $\rho \mapsto \rho^m$. Therefore $b(\log(\rho_0)) < l(\log(\rho_0))$. The convexity of $b(\cdot)$ now implies that it has at most two intersections with $l(\cdot)$ and that for such a root $\log(\beta)$ there holds

$$b'(\log(\beta)) \neq l'(\log(\beta)).$$

These conclusions relate 1 to 1 to the original setting and imply $a'(\beta) \neq m\beta^{m-1}$.
We can now apply [Har11, Proposition 4.27] for the scalar case to conclude that all roots of $q(\cdot)$ on the circle $\mathbb{T}_\beta$ are simple. □

We will also need some characterisation of eigenvalues which lie in the boundary of the block numerical range.

Theorem 1.32. *Let $\lambda \in \Sigma_p(Q(\cdot)) \cap \partial\Theta(Q(\cdot); H)$ be an eigenvalue. Then λ is a normal eigenvalue, i.e.*

$$\operatorname{Ker}(Q(\lambda)) = \operatorname{Ker}(Q(\lambda)^*). \tag{1.23}$$

Proof. This follows from [SW10, Theorem 3.5.(ii)]. While [SW10, Theorem 3.5.(ii)] formally assumes that $Q(\cdot)$ is self-adjoint, the proof still works if we assume that λ_0 is an eigenvalue. □

Recall that an eigenvalue $\lambda \in \Sigma_p(Q(\cdot))$ is semisimple if there exists no Jordan chain of length two, that is no two vectors $v, w \in H$ such that v is an eigenvector of λ and

$$Q'(\lambda)v + Q(\lambda)w = 0. \tag{1.24}$$

Remark 1.33. If λ is a normal eigenvector, i.e. $\mathrm{Ker}(Q(\lambda)) = \mathrm{Ker}(Q(\lambda)^*)$, we can multiply v^* from the left in the above equation and then semisimplicity of λ is equivalent to

$$v^* Q'(\lambda) v \neq 0 \tag{1.25}$$

for all eigenvectors v.

1.6.1. Positive semi-definite coefficients

Let us now deal with positive semi-definite coefficients.

Proposition 1.34. *Let $A(\cdot)$ be as in (1.21) with positive semi-definite coefficients A_j, $j = 0, \dots, l$. For $\beta > 0$ there holds*

$$\mathrm{spr}_A(\beta) = \mathrm{spr}(A(\beta)) = \mathrm{nur}(A(\beta)) = \mathrm{nur}_A(\beta). \tag{1.26}$$

Proof. The first equality follows from [Har11, Proposition 1.10]. The second equality holds because $A(\beta)$ is positive semi-definite (see e.g. [GR97, p.15]). For the last equality let $\lambda \in \mathbb{T}_\beta$ and $v \in H$, $v^*v = 1$ and write

$$|v^* A(\lambda) v| = \Big| \sum_{j=1}^{l} \lambda^j v^* A_j v \Big| \leq \sum_{j=1}^{l} |\lambda|^j |v^* A_j v| = \sum_{j=1}^{l} \beta^j v^* A_j v \leq \mathrm{nur}(A(\beta)).$$

The assumption now follows by taking the supremum over all λ and v. □

The next result is a generalization of [Wim08, Theorem 2.2] to semi-monic operator polynomials.

Theorem 1.35. *Let $Q(\cdot)$ be as in (1.21) with positive semi-definite coefficients A_j, $j = 0, \dots, l$. Let $\beta > 0$ such that $\mathrm{spr}(A(\beta)) = \mathrm{nur}(A(\beta)) = \beta^m$. Further assume that there exists an $\varepsilon > 0$ such that either*

$$\mathrm{nur}(A(\rho)) < \rho^m \qquad \textit{for all } \rho \in (\beta - \varepsilon, \beta) \tag{1.27}$$

or

$$\mathrm{nur}(A(\rho)) < \rho^m \qquad \textit{for all } \rho \in (\beta, \beta + \varepsilon). \tag{1.28}$$

Then every eigenvalue $\lambda_0 \in \Sigma_p(Q(\cdot)) \cap \mathbb{T}_\rho$ is normal and semisimple.

Proof. Without loss of generality assume that assumption (1.27) is satisfied. Let $\lambda_0 \in \Sigma_p(Q(\cdot)) \cap \mathbb{T}_\beta$ be an eigenvalue. From assumption (1.27) and Proposition 1.12 it follows that λ_0 lies in the boundary of $\Theta(Q(\cdot); H)$ and thus by Theorem 1.32 it follows that λ_0 is a normal eigenvalue.
For the semisimplicity let v be an eigenvector of λ_0 such that $v^*v = 1$. Define the function

$$q(\lambda) = \lambda^m - \sum_{j=0}^{l} \lambda^j v^* A_j v. \tag{1.29}$$

Since the A_j are positive semi-definite the coefficients of $q(\cdot)$ are real positives. By construction there holds $q(\lambda_0) = 0$. The next line shows that also $q(\beta) = 0$.

$$\beta^m = |\lambda_0^m| = \big|\sum_{j=0}^{l} \lambda_0^j v^* A_j v\big| \leq \sum_{j=0}^{l} \beta^j |v^* A_j v| = \sum_{j=0}^{l} \beta^j v^* A_j v \leq \beta^m \tag{1.30}$$

where the last inequality holds because $\mathrm{nur}(A(\beta)) = \beta^m$.
Now let $\rho \in (\beta - \varepsilon, \beta)$. By assumption (1.27) it follows that

$$\sum_{j=0}^{l} \rho^j v^* A v \leq \mathrm{nur}(A(\rho)) < \rho^m$$

and thus $q(\rho) > 0$. The function $q(\cdot)$ now fulfils the conditions of Theorem 1.31 and thus λ_0 is a simple root of $q(\cdot)$. This implies that $v^* Q'(\lambda_0) v \neq 0$ and by Remark 1.33 this is equivalent to λ_0 being semisimple. □

Remark 1.36. Roughly speaking assumptions (1.27) and (1.28) guarantee that the function $\mathrm{nur}(A(\cdot))$ 'cuts through' the function $\rho \mapsto \rho^m$ (as opposed to only touching it). This is crucial for Theorem 1.35 to hold. For a counter example where eigenvalues fail to be semisimple see [Har11, Example 2.15].

1.6.2. Normal coefficients

Let now

$$Q(\lambda) = \lambda^m - A(\lambda) = \lambda^m - \sum_{j=0}^{l} \lambda^j A_j \tag{1.31}$$

where $A_j \in \mathcal{L}(H)$ are normal. Further define

$$\bar{Q}(\lambda) = \lambda^m - \bar{A}(\lambda) = \lambda^m - \sum_{j=0}^{l} \lambda^j |A_j| \tag{1.32}$$

with $|A_j|$ the unique square root of $A_j^* A_j = A_j A_j^*$. Note that the coefficients of $\bar{A}(\cdot)$ are positive semi-definite.

Lemma 1.37. *Let $B \in \mathcal{L}(H)$ be normal and $v \in H$. Then*

$$|v^* B v| \leq v^* |B| v. \tag{1.33}$$

Proof. This is [Har11, Lemma 2.2]. □

Proposition 1.38. *There holds for $\beta > 0$*

$$\mathrm{spr}_A(\beta) \leq \mathrm{nur}_A(\beta) \leq \mathrm{nur}(\bar{A}(\beta)) = \mathrm{spr}(\bar{A}(\beta)). \tag{1.34}$$

Proof. The first inequality is clear. The last equality follows because $\bar{A}(\beta)$ is positive semi-definite. For the second inequality let $\lambda \in \mathbb{T}_1$ and $v \in H$, $v^*v = 1$. Then

$$|v^*A(\lambda)v| \le \sum_{j=0}^{l} |\lambda|^m |v^*A_jv| \le \sum_{j=0}^{l} \beta^m v^*|A_j|v \le \mathrm{nur}(\bar{A}(\beta))$$

where we used Lemma 1.37. The assertion now follows by taking the supremum over all λ and v. $\square$

Theorem 1.39. *Let $Q(\cdot)$ and $\bar{Q}(\cdot)$ as in* (1.31) *and* (1.32)*. Let $\beta > 0$ with $\beta^m = \mathrm{spr}_A(\beta) = \mathrm{spr}(\bar{A}(\beta))$ and assume there exists $\varepsilon > 0$ such that either*

$$\mathrm{nur}(\bar{A}(\rho)) < \rho^m \qquad \textit{for all } \rho \in (\beta - \varepsilon, \beta) \tag{1.35}$$

or

$$\mathrm{nur}(\bar{A}(\rho)) < \rho^m \qquad \textit{for all } \rho \in (\beta, \beta + \varepsilon). \tag{1.36}$$

Then

(i) $\mathrm{spr}_A(\beta) = \mathrm{nur}_A(\beta) = \beta^m$,

(ii) if $\lambda_0 \in \Sigma_p(Q) \cap \mathbb{T}_\beta$ then λ_0 is a normal eigenvalue.

Proof. (i) follows directly from Proposition 1.38. For (ii) assume without loss of generality that assumption (1.35) is satisfied. Then again by Proposition 1.38

$$\mathrm{nur}_A(\rho) \le \mathrm{nur}(\bar{A}(\rho)) < \rho^m \qquad \text{for all } \rho \in (\beta - \varepsilon, \beta).$$

It then follows from (i) and Proposition 1.12 that λ_0 lies in the boundary of $\Theta(Q(\cdot); H)$ and thus by Theorem 1.32 it follows that λ_0 is a normal eigenvalue. $\square$

2. Factorization in ordered Banach algebras

Multiplicative factorizations of elements in an algebra with respect to some additive decomposition of the algebra are well known and appear for example in [GGK93, GKS03, Mar88]. Additionally factorizations of polynomials have been studied in [Mar88].
We begin this chapter by stating some definitions and known results before we introduce an order structure in the next section.

Definition 2.1. (i) We call an algebra $\mathcal{A}$ a *decomposing algebra* if it contains two subalgebras $\mathcal{A}_+$ and $\mathcal{A}_-$ such that $\mathcal{A}$ is the direct sum of these subalgebras, i.e. $\mathcal{A} = \mathcal{A}_+ \dot{+} \mathcal{A}_-$.
We denote the natural projection onto $\mathcal{A}_+$ along $\mathcal{A}_-$ by P and $Q = Id_{\mathcal{A}} - P$.

(ii) We call $\mathcal{A}$ with unit e a *semi-strongly decomposing algebra* if it contains three subalgebras $\mathcal{A}_\pm$ and $\mathcal{A}_0$ such that $\mathcal{A} = \mathcal{A}_+ \dot{+} \mathcal{A}_0 \dot{+} \mathcal{A}_-$ (with natural projections P_+, P_0 and P_-) and

a) $e \in \mathcal{A}_0$, if $a_0 \in \mathcal{A}_0 \cap \mathcal{A}_{inv}$ then $a_0^{-1} \in \mathcal{A}_0$,

b) if $a_0 \in \mathcal{A}_0$ and $a_\pm \in \mathcal{A}_\pm$ then $a_0 a_\pm \in \mathcal{A}_\pm$ and $a_\pm a_0 \in \mathcal{A}_\pm$.

(iii) We call $\mathcal{A}$ with unit e a *strongly decomposing algebra* if it is a semi-strongly decomposing algebra and

c) if $a_\pm \in \mathcal{A}_\pm$ then $e - a_\pm \in \mathcal{A}_{inv}$ and $(e - a_\pm)^{-1} - e \in \mathcal{A}_\pm$.

If $\mathcal{A}$ is additionally a Banach algebra we call $\mathcal{A}$ a *((semi-)strongly) decomposing Banach algebra* if it is a ((semi-)strongly) decomposing algebra and the corresponding natural projections are continuous.

Part (i) appeared in [CG81, p. 34] and [GGK93, p. 806] and part (iii) in [GGK93, p. 545]. Note that we have the implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i), where in the second implication one needs to add $\mathcal{A}_0$ to one of the other two subalgebras. The reverse is in general not true as the following example illustrates.

Example 2.2. (i) Consider the Banach algebra $\mathbb{C}^n$ with component-wise multiplication as the algebra operation and unit $e = (1 \cdots 1)^T$. We can define the two closed subalgebras

$$\begin{aligned} \mathbb{C}^n_{odd} &:= \{(\lambda_j)_j \in \mathbb{C}^n : \lambda_j = 0 \text{ for } j \text{ even}\}, \\ \mathbb{C}^n_{even} &:= \{(\lambda_j)_j \in \mathbb{C}^n : \lambda_j = 0 \text{ for } j \text{ odd}\}. \end{aligned}$$

It is easy to see that $\mathbb{C}^n = \mathbb{C}^n_{odd} + \mathbb{C}^n_{even}$ is indeed a decomposing Banach algebra.
Note that the unit element is split up between the two subalgebras, therefore there is no straightforward way to further decompose $\mathbb{C}^n$ into a semi-strongly decomposing Banach algebra.

(ii) Let $\mathcal{A}$ be a Banach algebra with unit e. Then the Wiener algebra on the unit circle $\mathbb{T}$ with coefficients in $\mathcal{A}$ is defined as

$$W(\mathcal{A}) = \{a(\cdot) : a(\lambda) = \sum_{k=-\infty}^{\infty} \lambda^k a_k, \lambda \in \mathbb{T}, (a_k) \in \ell^1_{\mathbb{Z}}(\mathcal{A})\}$$

where $\ell^1_{\mathbb{Z}}(\mathcal{A}) = (a_k)_{k\in\mathbb{Z}} \subset \mathcal{A} : \sum_{k=-\infty}^{\infty} \|a_k\| < \infty\}$ and the algebra operations are the usual addition and pointwise multiplication of continuous functions. In other words $W(\mathcal{A})$ is the algebra of all continuous functions on $\mathbb{T}$ mapping into $\mathcal{A}$ whose series of Fourier coefficients is absolutely convergent. If we endow $W(\mathcal{A})$ with the norm

$$\|a(\cdot)\|_W = \sum_{k=-\infty}^{\infty} \|a_k\|$$

then $W(\mathcal{A})$ becomes a Banach algebra (since $\mathcal{A}$ is a Banach algebra) with unit $e_W(\cdot) \equiv e$ (see e.g. [Kat04, I.6.1 and VIII.2.9]).
The Wiener algebra is the direct sum of the subsets

$$\begin{aligned}
W_-(\mathcal{A}) &= \{a(\cdot) \in W(\mathcal{A}) : a_k = 0 \text{ for } k \geq 0\},\\
W_+(\mathcal{A}) &= \{a(\cdot) \in W(\mathcal{A}) : a_k = 0 \text{ for } k \leq 0\},\\
W_0(\mathcal{A}) &= \{a(\cdot) \in W(\mathcal{A}) : a(\cdot) \equiv a_0, a_0 \in \mathcal{A}\}.
\end{aligned}$$

Since the Fourier coefficients of the product of two functions are obtained via convolution, it follows that the three subsets $W_\alpha(\mathcal{A})$, $\alpha = -, 0, +$ are subalgebras of $W(\mathcal{A})$. They are also closed which follows from the $\ell^1_{\mathbb{Z}}$-norm. Definition 2.1.(ii).a) and Definition 2.1.(ii).b) are also clearly satisfied so that with the above decomposition the Wiener algebra becomes a semi-strongly decomposing Banach algebra.
However the Wiener algebra is not a strongly decomposing Banach algebra. The function $a(\lambda) = \lambda e$ belongs to $W_+(\mathcal{A})$ but clearly $e_W(\cdot) - a(\cdot)$ is not invertible in $W(\mathcal{A})$ since it has a root at $\lambda = 1$.

(iii) Consider the Banach algebra $\mathbb{C}^{n,n}$ of complex $n \times n$- matrices with the usual matrix multiplication and unit element the identity matrix I. Then $\mathbb{C}^{n,n}$ can be written as the direct sum of strictly lower triangular, diagonal and strictly upper triangular matrices. It is easy to check that $\mathbb{C}^{n,n}$ becomes a semi-strongly decomposing Banach algebra in this way. Indeed it even becomes a strongly decomposing Banach algebra since a strictly lower (respectively upper) triangular matrix A is nilpotent and thus $(I-A)^{-1}-I$ is again a strictly lower (respectively upper) triangular matrix by Neumann's series. Therefore Definition 2.1.(iii).c) is also satisfied.

The next proposition collects some known factorization results.

Proposition 2.3. *(i) Let $\mathcal{A} = \mathcal{A}_- \dot{+} \mathcal{A}_+$ be a decomposing Banach algebra with unit e and natural projections P and Q. If $a \in \mathcal{A}$ and $\|a\| < \min\{\|P\|^{-1}, \|Q\|^{-1}\}$, then $e - a$ admits a factorization*

$$e - a = (e - a_-)(e - a_+) \tag{2.1}$$

with $a_\pm \in \mathcal{A}_\pm$ and $e - a_\pm$ invertible with $(e - a_\pm)^{-1} - e \in \mathcal{A}_\pm$. Moreover such a factorization is unique.

(ii) Let $\mathcal{A} = \mathcal{A}_- \dot{+} \mathcal{A}_0 \dot{+} \mathcal{A}_+$ be a strongly decomposing algebra with natural projections P_-, P_0 and P_+. An element $a \in \mathcal{A}$ admits a factorization

$$e - a = (e - a_-)a_0(e - a_+) \tag{2.2}$$

with $a_\pm \in \mathcal{A}_\pm$ and $a_0 \in \mathcal{A}_0$ if and only if there exist elements $x_\pm \in \mathcal{A}_0 + \mathcal{A}_\pm$ satisfying

$$\begin{aligned} x_+ - Q_+(ax_+) &= Q_+(a), \\ x_- - Q_-(x_-a) &= Q_-(a), \end{aligned}$$

where $Q_\pm = P_0 + P_\pm$. Moreover such a factorization is unique.

Proof. (i) is the result given in [GGK93, Theorem XXIX.9.1]. A similar version is given in [Mar88, Theorem 23.3].
(ii) is [GGK93, Theorem XXII.8.2]. □

In [Mar88] A. S. Markus considers factorization results for operator functions and polynomials. In particular he implicitly states a factorization result for semi-monic polynomials which we restate here.

Proposition 2.4. *Let $1 \leq m < l$ and*

$$q(\lambda) = \lambda^m e - a(\lambda) = \lambda^m e - \sum_{j=0}^{l} \lambda^j a_j$$

with the a_j elements in a Banach algebra $\mathcal{A}$ and e the unit element. If

$$\operatorname{norm}_a(\rho) = \sum_{j=0}^{l} \rho^j \|a_j\| \leq \rho^m \tag{2.3}$$

for some $\rho > 0$, then $\lambda^{-m}q(\lambda)$ admits a canonical factorization with respect to the circle with radius $|\lambda| = \rho$, that is

$$\lambda^{-m}a(\lambda) = a_+(\lambda)(e + a_-(\lambda)), \qquad \lambda \in \mathbb{T}_\rho, \tag{2.4}$$

with continuous functions $a_\pm(\cdot)$ such that $a_+(\cdot)$ has a holomorphic extension to the disk $\mathbb{D}_{\leq\rho}$ and $a_-(\cdot)$ has a holomorphic extension to the outer disc $\mathbb{D}_{\geq\rho}$ which vanishes at infinity.

Proof. This is a slight reformulation of [Mar88, Corollary 23.5] where we emphasized the semi-monic structure of $q(\cdot)$. We further changed the setting from Banach spaces to Banach algebras since the proofs given also apply in this more general case. □

Markus then continues by specifying spectral properties of the divisors in (2.4). An order version of these results is given in Section 2.4.

Markus additionally considered other sufficient conditions for an operator polynomial $A(\cdot)$, with coefficients in $\mathcal{L}(H)$, H a Hilbert space, to admit a canonical factorization. It reads

$$\inf_{|\lambda|=\rho, \|x\|=1} |(A(\lambda)x, x)| > 0 \tag{2.5}$$

for some $\rho > 0$ and where $x \in H$. Then (2.5) and other conditions imply a factorization result; for an example see [Mar88, Theorem 26.12].
We want to show that conditions of the form (2.5) are satisfied for semi-monic functions if the function $\mathrm{nur}_a(\cdot)$ from (1.3) satisfies an inequality $\mathrm{nur}_a(\rho) < \rho^m$ for $\rho > 0$. We show this in three different settings.

(i) Let $\mathcal{A}$ be a Banach algebra with unity e and $q(\lambda) = \lambda^m e - a(\lambda)$ as in (1.1) with coefficients in $\mathcal{A}$. Recall that for $b \in \mathcal{A}$ the numerical range is defined as

$$\Theta(b; \mathcal{A}) = \{f(b) : f \in \mathcal{A}', f(e) = \|f\| = 1\}.$$

Then $\mathrm{nur}_a(\rho) = \sup_{|\lambda|=\rho} \mathrm{nur}(a(\lambda)) < \rho^m$ implies

$$\begin{aligned} &\inf\{|f(q(\lambda))| : |\lambda| = \rho, f \in \mathcal{A}', f(e) = \|f\| = 1\} \\ &\qquad = \inf\{|f(\lambda^m e - a(\lambda))| : |\lambda| = \rho, f \in \mathcal{A}', f(e) = \|f\| = 1\} \\ &\qquad = \inf\{|\lambda^m - f(a(\lambda))| : |\lambda| = \rho, f \in \mathcal{A}', f(e) = \|f\| = 1\} \\ &\qquad \geq \inf\{\rho^m - |f(a(\lambda))| : |\lambda| = \rho, f \in \mathcal{A}', f(e) = \|f\| = 1\} \\ &\qquad = \lambda^m - \sup_{|\lambda|=\rho} \mathrm{nur}(a(\lambda)) > 0. \end{aligned}$$

(ii) Let X be a Banach space and $Q(\lambda) = \lambda^m I - A(\lambda)$ as in (1.6) with coefficients in $\mathcal{L}(X)$. The spatial numerical range for $B \in \mathcal{L}(X)$ is

$$\begin{aligned} \Theta(B; X) &= \{f(Bx) : (f, x) \in S_{\mathrm{att}}(X)\} \\ &= \{f(Bx) : x \in X, f \in X', f(x) = \|f\| = \|x\| = 1\}. \end{aligned}$$

Then similarly to above the inequality $\mathrm{nur}_A(\rho) < \rho^m$ for a $\rho > 0$ implies

$$\begin{aligned} &\inf\{|f(Q(\lambda)x)| : |\lambda| = \rho, (f, x) \in S_{\mathrm{att}}(X)\} \\ &\qquad = \inf\{|f(\lambda^m x - A(\lambda)x)| : |\lambda| = \rho, (f, x) \in S_{\mathrm{att}}(X)\} > 0. \end{aligned}$$

(iii) Lastly we show the Hilbert space case. Let H be a Hilbert space and $Q(\lambda) = \lambda^m I - A(\lambda)$ be as above with coefficients in $\mathcal{L}(H)$. We have for $B \in \mathcal{L}(H)$

$$\Theta(B; H) = \{(Bx, x) : x \in H, \|x\| = 1\}.$$

Then $\mathrm{nur}_A(\rho) < \rho^m$ implies

$$\begin{aligned} &\inf\{|(Q(\lambda)x, x)| : |\lambda| = \rho, x \in H, \|x\| = 1\} \\ &\qquad = \inf\{|\lambda^m - (A(\lambda)x, x)| : |\lambda| = \rho, x \in H, \|x\| = 1\} > 0. \end{aligned}$$

2.1. Decomposing ordered Banach algebras

We will now introduce an order structure to the additive decompositions in the previous decomposition which will allow us to later state factorization results that take the order of the underlying algebra into account. This concept was previously examined in [FN05b] for decomposing Banach algebras, however we will be able to derive stronger factorization results.
Recall that an algebra $\mathcal{A}$ is called an *ordered algebra* if it contains a *cone*, that is, a subset $\mathcal{C} \subset \mathcal{A}$ satisfying

1. $\mathcal{C} + \mathcal{C} \subset \mathcal{C}$,

2. $\lambda\mathcal{C} \subset \mathcal{C}, \lambda > 0$.

If $\mathcal{C}$ additionally satisfies

3. $\mathcal{C} \cdot \mathcal{C} \subset \mathcal{C}$,

then it is called an *algebra cone.* If $\mathcal{A}$ contains a unit e then we also require

4. $e \in \mathcal{C}$.

For $a, b \in \mathcal{A}$ an algebra cone induces a partial ordering via

$$a \leq b \quad \text{if and only if} \quad b - a \in \mathcal{C}.$$

Moreover if $\mathcal{A}$ is a Banach algebra then the (algebra) cone $\mathcal{C}$ is called *normal* if there exists $\gamma > 0$ such that for every $a, b \in \mathcal{A}$ satisfying $0 \leq a \leq b$ there holds $\|a\|_{\mathcal{A}} \leq \gamma \|b\|_{\mathcal{A}}$.
The following proposition, which collects some known results, will prove useful throughout this chapter.

Proposition 2.5. *Let $\mathcal{A}$ be an ordered Banach algebra with closed normal algebra cone $\mathcal{C}$. Then*

(i) The spectral radius $\operatorname{spr}(\cdot)$ *is a monotone function on $\mathcal{C}$, i.e. if $0 \leq a \leq b$ then* $\operatorname{spr}(a) \leq \operatorname{spr}(b)$.

(ii) For all $a \in \mathcal{C}$ its spectral radius $\operatorname{spr}(a)$ *belongs to its spectrum $\Sigma(a)$.*

(iii) Let $a \in \mathcal{C}$ and $\lambda \in \mathbb{C}$. Then $\lambda \notin \Sigma(a)$ and $(\lambda e - a)^{-1} \in \mathcal{C}$ if and only if $\lambda \geq 0$ and $\operatorname{spr}(a) < \lambda$.

Proof. For the first two assertions see [RR89, Theorem 4.1.1 and Proposition 5.1]. The 'if' part of the last assertion is an immediate consequence of [KLS89, Theorem 25.1], and the 'only if' part follows by using Neumann's series. □

We can now introduce the order sensitive version of Definition 2.1

Definition 2.6. Let $\mathcal{A}$ be an ordered algebra with cone $\mathcal{C}$.

(i) We call $\mathcal{A}$ a *decomposing ordered algebra* if it is a decomposing algebra and $\mathcal{C}$ is invariant under the natural projections P and $Q = Id_{\mathcal{A}} - P$.

(ii) We call $\mathcal{A}$ a *semi-strongly decomposing ordered algebra* if it is a semi-strongly decomposing algebra and

a) $\mathcal{C}$ is invariant under the natural projections P_+, P_0 and P_-.

(iii) We call $\mathcal{A}$ a *strongly decomposing ordered algebra* if it is a strongly decomposing algebra and

a) $\mathcal{C}$ is invariant under the natural projections P_+, P_0 and P_-.

b) if $a_\pm \in \mathcal{A}_\pm \cap \mathcal{C}$ then $(e - a_\pm)^{-1} - e \in \mathcal{A}_\pm \cap \mathcal{C}$.

If $\mathcal{A}$ is additionally a Banach algebra we call $\mathcal{A}$ a *((semi-)strongly) decomposing ordered Banach algebra* if it is a ((semi-)strongly) decomposing ordered algebra and the corresponding natural projections are continuous.

The Banach algebras from Example 2.2 can all be endowed with an order structure such that they become ((semi-)strongly) decomposing ordered Banach algebras.

Example 2.7. (i) The Banach algebra $\mathbb{C}^n$ from Example 2.2.(i) can be endowed with the closed normal algebra cone $\mathbb{R}^n_+$. It is easy to see that $\mathbb{C}^n = \mathbb{C}^n_{odd} + \mathbb{C}^n_{even}$ becomes a decomposing ordered Banach algebra in this way.

(ii) For the Wiener algebra $W(\mathcal{A})$ from Example 2.2 assume that the coefficient algebra $\mathcal{A}$ contains an algebra cone $\mathcal{C}$. Then the order on $\mathcal{A}$ induces an order on $W(\mathcal{A})$ via the algebra cone

$$\mathcal{C}_W := \{a(\cdot) : a_k \in \mathcal{C} \text{ for all } k \in \mathbb{Z}\}.$$

The cone $\mathcal{C}_W$ will be normal and closed if and only if $\mathcal{C}$ is normal and closed. Since $\mathcal{C}_W$ is invariant under the natural projections onto the subalgebras, $W(\mathcal{A})$ becomes a semi-strongly decomposing ordered Banach algebra.
However, as before, $W(\mathcal{A})$ is not a strongly decomposing ordered Banach algebra.

(iii) For the Banach algebra $\mathbb{C}^{n,n}$ from Example 2.2.(iii) consider the closed normal algebra cone $\mathbb{R}^{n,n}_+$. Then it is easy to see that $\mathbb{C}^{n,n}$ becomes a semi-strongly decomposing ordered Banach algebra.
Furthermore Definition 2.6.(iii).b) is satisfied which makes $\mathbb{C}^{n,n}$ a strongly decomposing ordered Banach algebra. Indeed let L be strictly lower triangular matrix with nonnegative entries. Then via the Neumann series

$$(I - L)^{-1} - I = \sum_{j=1}^{n} L^j$$

which is again entrywise nonnegative.

The next section introduces a general method to define strongly decomposing ordered Banach algebras.

2.1.1. Chains of projections

For a vector space X one can introduce a partial order on the set of projections on X by setting

$$P_1 \leq_c P_2 \qquad \text{iff} \qquad \operatorname{Im}(P_1) \subset \operatorname{Im}(P_2) \text{ and } \operatorname{Ker}(P_2) \subset \operatorname{Ker}(P_1)$$

or equivalently

$$P_1 \leq_c P_2 \qquad \text{iff} \qquad P_1 P_2 = P_2 P_1 = P_1.$$

Then we call a set $\pi = \{P_0, P_1, \dots, P_n\}$ a *finite chain* on $\mathcal{A}$ if

$$0 = P_0 \leq_c P_1 \leq_c \dots \leq_c P_n = I.$$

We call a set $\mathbb{P} = \{P\}$ of projections of X a *chain* if $0, I \in \mathbb{P}$ and the elements in $\mathbb{P}$ are linearly ordered with respect to $\leq_c$. If X is additionally a Banach space then we require the projections to be continuous (for a thorough expositions on chains see [GGK93, Chapter XX]).

The concept of chains can also be introduced in an algebra $\mathcal{A}$ with unit e by setting

$$p_1 \leq_c p_2 \qquad \text{iff} \qquad p_1 p_2 = p_2 p_1 = p_1 \tag{2.6}$$

for two elements $p_1, p_2 \in \mathcal{A}$. Then (finite) chains are defined in exactly the same way as above where the identity operator I is replaced by the unit element $e \in \mathcal{A}$. For two chains $\mathbb{P}_1, \mathbb{P}_2$ we say that $\mathbb{P}_1$ *is finer than* $\mathbb{P}_2$ if $\mathbb{P}_2 \subseteq \mathbb{P}_1$.
If $\mathcal{A}$ is an ordered algebra with algebra cone $\mathcal{C}$ then following [Ale11] we call an element p *order idempotent* if

$$0 \leq p \leq e \qquad \text{and} \qquad p^2 = p$$

where $\leq$ refers to the order induced by $\mathcal{C}$. The next lemma shows that for order idempotents the order on chains as in (2.6) is equivalent to the algebra order induced by $\mathcal{C}$. Recall that the infimum of two elements $a, b \in \mathcal{A}$, denoted $a \wedge b$, is defined as the lowest upper bound of the set $\{a, b\}$ with respect to the order induced by $\mathcal{C}$ if it exists.

Lemma 2.8. *Let $\mathcal{A}$ be an ordered algebra and $p_1, p_2 \in \mathcal{A}$ be two order idempotents. Then*

$$p_1 \leq p_2 \qquad \textit{iff} \qquad p_1 \leq_c p_2.$$

Proof. By [Ale11, Lemma 2.1.(b)] there holds

$$p_1 p_2 = p_2 p_1 = p_1 \wedge p_2.$$

Since $p_1 \leq p_2$ if and only if $p_1 = p_1 \wedge p_2$ the assertion follows. □

Consequently we call a set $\mathbb{P}$ a *chain of order idempotents* if $\mathbb{P}$ is a chain whose elements are order idempotents and are linearly ordered with respect to $\leq$ (analogously for finite chains).

In the following we will recall constructions and results from [GGK93] which are stated for projections and operators on Banach spaces but can also be applied to our more general setting.
Let $\mathcal{A}$ be a Banach algebra, $\mathbb{P}$ a chain and $a \in \mathcal{A}$. Then for a finite subchain $\pi = \{0 = p_0, p_1, \dots, p_n = e\} \subset \mathbb{P}$ one can define

$$D(a, \pi) := \sum_{j=1}^{n} (\Delta p_j) a (\Delta p_j)$$

where $\Delta p_j = p_j - p_{j-1}$, $j = 1, \dots, n$. We can then define a *diagonal of a with respect to* $\mathbb{P}$ as

$$D(a, \mathbb{P}) := \lim_{\pi \subset \mathbb{P}} D(a, \pi), \tag{2.7}$$

assuming the limit exists in the norm of $\mathcal{A}$. Here the limit is to be understood as for each $\varepsilon > 0$ there exists a finite subchain $\pi_\varepsilon \subset \mathbb{P}$ such that

$$\|D(a, \mathbb{P}) - D(a, \pi)\| < \varepsilon$$

for all finite subchains $\pi \subset \mathbb{P}$ that are finer than π_ε. The set $\mathcal{A}(\mathbb{P})$ is then defined as the set of all $a \in \mathcal{A}$ that have a diagonal.
We need to introduce one more definition: A chain $\mathbb{P}$ is called *uniform* if there exists $\gamma > 0$ such that for every $a \in \mathcal{A}(\mathbb{P})$ and finite subchain $\pi \subset \mathbb{P}$ there holds

$$\|D(a, \pi)\| \leq \gamma \|a\|.$$

Note that in an ordered Banach algebra with normal algebra cone every chain of order idempotents is uniform.

Lemma 2.9. *If $\mathbb{P}$ is a uniform chain then the set $\mathcal{A}(\mathbb{P})$ is closed and it exists a constant $\gamma > 0$ such that for all $a \in \mathcal{A}(\mathbb{P})$ there holds*

$$\|D(a, \mathbb{P})\| \leq \gamma \|a\|.$$

Proof. This is [GGK93, Proposition XX.3.1]. □

We can now introduce the three sets.

$$\begin{aligned} \mathcal{A}_-(\mathbb{P}) &= \{a \in \mathcal{A}(\mathbb{P}) : pa = pap \text{ for all } p \in \mathbb{P}, D(a, \mathbb{P}) = 0\}, \\ \mathcal{A}_+(\mathbb{P}) &= \{a \in \mathcal{A}(\mathbb{P}) : ap = pap \text{ for all } p \in \mathbb{P}, D(a, \mathbb{P}) = 0\}, \\ \mathcal{A}_0(\mathbb{P}) &= \{a \in \mathcal{A}(\mathbb{P}) : ap = pa \text{ for all } p \in \mathbb{P}\}. \end{aligned} \tag{2.8}$$

Lemma 2.10. *If $\mathbb{P}$ is a uniform chain then the sets in* (2.8) *are closed subalgebras of* $\mathcal{A}$.

Proof. We show this exemplary for $\mathcal{A}_-(\mathbb{P})$. To see that it is a subalgebra let $a, b \in \mathcal{A}_-(\mathbb{P})$ and $p \in \mathbb{P}$. Then

$$pab = papb = papbp = pabp,$$

so the first property is satisfied. For any finite subchain $\pi \subset \mathbb{P}$ there holds

$$\begin{aligned}
D(a,\pi)D(b,\pi) &= \sum_{j=1}^{n} (\Delta p_j) a (\Delta p_j) \sum_{k=1}^{n} (\Delta p_k) b (\Delta p_k) \\
&= \sum_{j,k=1}^{n} (\Delta p_j) a (\Delta p_j)(\Delta p_k) b (\Delta p_k) \\
&= \sum_{j=1}^{n} (\Delta p_j) a (\Delta p_j)^2 b (\Delta p_j) \\
&= \sum_{j=1}^{n} (\Delta p_j) a (\Delta p_j) b (\Delta p_j) \\
&= \sum_{j=1}^{n} (\Delta p_j) ab (\Delta p_j) = D(ab, \pi)
\end{aligned}$$

because $(\Delta p_j)(\Delta p_k) = 0$ for $k \neq j$. Now let $\varepsilon > 0$, then there exists a finite subchain π_ε such that for all $\pi_\varepsilon \subset \pi \subset \mathbb{P}$ there holds

$$\|D(a,\mathbb{P}) - D(a,\pi)\| < \varepsilon, \qquad \|D(b,\mathbb{P}) - D(b,\pi)\| < \varepsilon.$$

Then also

$$\begin{aligned}
&\|D(a,\mathbb{P})D(b,\mathbb{P}) - D(ab,\pi)\| = \|D(a,\mathbb{P})D(b,\mathbb{P}) - D(a,\pi)D(b,\pi)\| \\
&\quad = \|D(a,\mathbb{P})D(b,\mathbb{P}) - D(a,\pi)D(b,\mathbb{P}) + D(a,\pi)D(b,\mathbb{P}) - D(a,\pi)D(b,\pi)\| \\
&\quad \leq \|D(a,\mathbb{P})D(b,\mathbb{P}) - D(a,\pi)D(b,\mathbb{P})\| + \|D(a,\pi)D(b,\mathbb{P}) - D(a,\pi)D(b,\pi)\| \\
&\quad \leq \underbrace{\|D(a,\mathbb{P}) - D(a,\pi)\|}_{\varepsilon} \underbrace{\|D(b,\mathbb{P})\|}_{\gamma\|b\|} + \underbrace{\|D(a,\pi)\|}_{\gamma\|a\|} \underbrace{\|D(b,\mathbb{P}) - D(b,\pi)\|}_{\varepsilon} \\
&\quad \leq \varepsilon\gamma(\|a\| + \|b\|)
\end{aligned}$$

where $\gamma > 0$ is the constant from Lemma 2.9. It follows that $D(ab, \mathbb{P})$ exists and

$$D(ab,\mathbb{P}) = D(a,\mathbb{P})D(b,\mathbb{P}) = 0$$

and thus $ab \in \mathcal{A}_-(\mathbb{P})$.
It remains to see that $\mathcal{A}_-(\mathbb{P})$ is closed. Multiplication is continuous in Banach algebras so that we only need concern ourselves with the condition $D(a,\mathbb{P}) = 0$. Let $(a_k)_{k\in\mathbb{N}} \subset \mathcal{A}_-(\mathbb{P})$, $a \in \mathcal{A}$ and $a_k \to a$. Since $\mathcal{A}(\mathbb{P})$ is closed we know that there exists a diagonal $D(a,\mathbb{P})$ of a. Now let $\varepsilon > 0$ and choose a $k_\varepsilon \in \mathbb{N}$ and a finite subchain $\pi_\varepsilon \subset \mathbb{P}$ such that

$$\|a - a_k\| < \varepsilon, \quad \|D(a,\mathbb{P}) - D(a,\pi)\| < \varepsilon, \quad \|D(a_k,\pi)\| < \varepsilon$$

for all $k \geq k_\varepsilon$ and finite subchains $\pi_\varepsilon \subset \pi \subset \mathbb{P}$. It then follows that

$$\|D(a,\mathbb{P})\| \leq \underbrace{\|D(a,\mathbb{P}) - D(a,\pi)\|}_{\leq \varepsilon} + \underbrace{\|D(a,\pi) - D(a_k,\pi)}_{=\|D(a-a_k,\pi)\| \leq \gamma\varepsilon}\| + \underbrace{\|D(a_k,\pi)\|}_{\leq \varepsilon} \leq (2+\gamma)\varepsilon$$

for all $k \geq k_\varepsilon$ and finite subchains $\pi_\varepsilon \subset \pi \subset \mathbb{P}$ (the γ comes once again from Lemma 2.9). We conclude that $\mathcal{A}_-(\mathbb{P})$ is closed. □

Similarly to the diagonal one can define for $a \in \mathcal{A}$ *lower and upper triangular parts* $L(a,\mathbb{P})$ and $U(a,\mathbb{P})$ with respect to $\mathbb{P}$ as the limits of

$$\lim_{\pi \in \mathbb{P}} \sum_{0 \leq l < k \leq n} (\Delta p_k) a (\Delta p_l) \quad \text{and} \quad \lim_{\pi \in \mathbb{P}} \sum_{0 \leq l < k \leq n} (\Delta p_l) a (\Delta p_k) \tag{2.9}$$

respectively if the limits exist.
Note that in general

$$\mathcal{A}_-(\mathbb{P}) + \mathcal{A}_0(\mathbb{P}) + \mathcal{A}_+(\mathbb{P}) \subsetneqq \mathcal{A}(\mathbb{P}) \subsetneqq \mathcal{A}$$

since the limits in (2.7) and (2.9) need not exist.
We can now state the main result about chains.

Proposition 2.11. *Let $\mathcal{A}$ be a Banach algebra with unit e and $\mathbb{P}$ a uniform chain and assume*

$$\mathcal{A} = \mathcal{A}_-(\mathbb{P}) + \mathcal{A}_0(\mathbb{P}) + \mathcal{A}_+(\mathbb{P}). \tag{2.10}$$

Then there holds:

(i) The sum in (2.10) *is direct and $\mathcal{A}$ is a strongly decomposing Banach algebra. For every element $a \in \mathcal{A}$ the limits $L(a,\mathbb{P}), D(a,\mathbb{P})$ and $U(a,\mathbb{P})$ exist, a can be written as*

$$a = L(a,\mathbb{P}) + D(a,\mathbb{P}) + U(a,\mathbb{P})$$

and $L(a,\mathbb{P}) \in \mathcal{A}_-(\mathbb{P})$, $D(a,\mathbb{P}) \in \mathcal{A}_0(\mathbb{P})$ and $U(a,\mathbb{P}) \in \mathcal{A}_+(\mathbb{P})$.

(ii) If additionally $\mathcal{A}$ is an ordered Banach algebra with normal closed algebra cone $\mathcal{C}$ and $\mathbb{P}$ is a chain of order idempotents then $\mathcal{A}$ is a strongly decomposing ordered Banach algebra.

Proof. Assertion (i) follows from [GGK93, Proposition XX.5.1] and [GGK93, Theorem XX.7.1]. For (ii) note that Definition 2.6.(ii).a) is satisfied since $\mathbb{P}$ is a chain of order idempotents and Definition 2.6.(iii).b) follows by the Neumann series since by [GGK93, Proposition XX.5.1] elements in $\mathcal{A}_\pm(\mathbb{P})$ are quasi-nilpotent. □

Example 2.12. (i) In $\mathbb{C}^n$ consider the canonical coordinate projections $\tilde{P}_0 = 0$ and $\tilde{P}_j$, $j = 1, \ldots, n$ given by

$$P_j : \mathbb{C}^n \to \mathbb{C}^n : (x_i)_{i=1,\ldots,n} \mapsto (\delta_{ij} x_i)_{i=1,\ldots,n}$$

where δ_{ij} is the usual Kronecker delta. Then, for $\mathbb{C}^{n,n}$ with cone $\mathbb{R}_+^{n,n}$, the set $\pi := \{P_k\}_{k=0,\dots,n}$ with $P_k = \sum_{j=0}^{k} \tilde{P}_j$ defines a finite chain of order idempotents. We will call π the *canonical chain on* $\mathbb{C}^{n,n}$. The associated decomposition decomposes $\mathbb{C}^{n,n}$ into strictly lower triangular, diagonal and strictly upper triangular matrices.
It is also possible to consider subchains of π which will result in a block triangular decomposition.

(ii) Consider the classical function spaces $L^p([a,b])$, $1 \le p \le \infty$. Then the projections

$$P_r : L^p([a,b]) \to L^p([a,b]) : f(\cdot) \mapsto \mathbf{1}_{[a,r]}(\cdot) f(\cdot), \qquad r \in [a,b]$$

give rise to the uniform chain $\mathbb{P} := \{P_r : a \le r \le b\}$ for the Banach algebra $\mathcal{L}(L^p([a,b]))$. If we restrict ourselves to an appropriate set of integral operators on $L^p([a,b])$ then the set of integral operators with positive kernel function forms an algebra cone. In this context $\mathbb{P}$ then becomes a chain of order idempotents. For precise conditions see Section 2.3.2.

Remark 2.13. We have seen in Example 2.7.(ii) that the Wiener algebra $W(\mathcal{A})$ is only a semi-strongly decomposing ordered Banach algebra. This implies that there does not exist a chain in $W(\mathcal{A})$ giving rise to the decomposition $W(\mathcal{A}) = W_-(\mathcal{A}) + W_0(\mathcal{A}) + W_+(\mathcal{A})$ as then by Proposition 2.11 it would have to be strongly decomposing.

2.2. Factorization in decomposing ordered Banach algebras

We present now the two main results of this section. The next theorem can be seen as an extension of [GGK93, Theorem XXIX.9.1] that takes the order structure of the underlying Banach algebra into account.

Theorem 2.14. *Let $\mathcal{A} = \mathcal{A}_+ \dot{+} \mathcal{A}_-$ be a decomposing ordered Banach algebra with unit e and closed normal algebra cone $\mathcal{C}$ where $e \in \mathcal{C}$. For a (positive) element $a \in \mathcal{C}$ are equivalent*

(i) $\operatorname{spr}(a) < 1$,

(ii) $e - a = (e - b_-)(e - b_+)$ *where* $b_\pm \in \mathcal{A}_\pm \cap \mathcal{C}$ *and* $\operatorname{spr}(b_\pm) < 1$.

The factorization in (ii) is unique.

Proof. Denote by P the projection in $\mathcal{A}$ onto $\mathcal{A}_+$ along $\mathcal{A}_-$ and $Q = Id_{\mathcal{A}} - P$.
(i) $\Rightarrow$ *(ii)*: For $n \in \mathbb{N}_0$

$$0 \le (PL_a)^n e \le a^n \quad \text{and} \quad 0 \le (R_a Q)^n e \le a^n$$

hold; indeed $0 \le (PL_a)^0 e = e = a^0$. Assume for some n the inequality $0 \le (PL_a)^n e \le a^n$ holds then $0 \le (PL_a)^{n+1} e \le PL_a(a^n) = P(a^{n+1}) = (Id_{\mathcal{A}} - Q)(a^{n+1}) \le a^{n+1}$ since P, Q and L_a all map the cone $\mathcal{C}$ into itself. The right inequalities follow similarly.
Since $\mathcal{C}$ is normal, there exists a $\gamma > 0$ such that

$$||(PL_a)^n e|| \leq \gamma ||a^n|| \quad \text{and} \quad ||(QR_a)^n e|| \leq \gamma ||a^n||, \quad n \in \mathbb{N}_0.$$

Now $\text{spr}(a) < 1$ implies that there exists a $\beta > 0$ and $\rho \in [0, 1[$ such that

$$||(PL_a)^n e|| \leq \beta \rho^n \quad \text{and} \quad ||(R_a Q)^n e|| \leq \beta \rho^n, \quad n \in \mathbb{N}_0.$$

Therefore

$$x_+ = \sum_{n=0}^{\infty} (PL_a)^{n+1} e = \sum_{n=0}^{\infty} (PL_a)^n (Pa)$$

and

$$x_- = \sum_{n=0}^{\infty} (QR_a)^{n+1} e = \sum_{n=0}^{\infty} (QR_a)^n (Qa)$$

exist, belong to $\mathcal{C}$ and satisfy

$$x_+ - (PL_a)x_+ = Pa \quad \text{and} \quad x_- - (QR_a)x_- = Qa.$$

Define

$$b_+ := P((e + x_+)a) = PR_a(e + x_+) \quad \text{and} \quad b_- := Q(a(e + x_-)) = QL_a(e + x_+).$$

Then $b_\pm \in \mathcal{A}_\pm \cap \mathcal{C}$. Now using $P + Q = Id_{\mathcal{A}}$ we see that the equations

$$(e - a)(e + x_+) = e - b_- \quad \text{and} \quad (e + x_-)(e - a) = e - b_+$$

hold, and the following equation holds:

$$(e - b_+)(e + x_+) = (e + x_-)(e - b_-).$$

Then

$$-b_+ + x_+ - b_+ x_+ = x_- - b_- - x_- b_- \in \mathcal{A}_- \cap \mathcal{A}_+ = \{0\}$$

and finally

$$(e - b_+)(e + x_+) = e = (e + x_-)(e - b_-).$$

By [GGK93, Lemma XXIX.9.2] (see the proof of part (d) of [GGK93, Theorem XXIX.9.1]) it follows that $e + x_+$ and $e - b_-$ are invertible and $(e + x_+)^{-1} = e - b_+$, therefore

$$(e - a) = (e - b_-)(e + x_+)^{-1} = (e - b_-)(e - b_+).$$

Further $(e-b_\pm)^{-1} = e+x_\pm \in \mathcal{C}$ which by Proposition 2.5.(iii) is equivalent to $\text{spr}(b_\pm) < 1$ since $b_\pm \in \mathcal{C}$ and the cone is normal.
$(ii) \Rightarrow (i)$ We obtain that $e - a \in \mathcal{A}_{inv}$ and $(e - a)^{-1} = (e - b_+)^{-1}(e - b_+)^{-1} \in \mathcal{C}$ which is equivalent to $\text{spr}(a) < 1$ again by Proposition 2.5.(iii) since $a \in \mathcal{C}$.
To see the uniqueness of the decomposition in (ii) we can apply the same proof as part (a) of the proof of [GGK93, Theorem XXIX.9.1]. $\square$

Remark 2.15. - In Theorem 2.14 the conditions (i) and (ii) are equivalent to:

(iii) The equations

$$x_+ - (PL_a)x_+ = Pa \quad \text{and} \quad x_- - (QR_a)x_- = Qa$$

have solutions in $\mathcal{C}$.

Further, for all solutions of these equations which belong to $\mathcal{C}$

$$x_+ \geq \sum_{n=0}^{\infty}(PL_a)^{n+1}e \quad \text{and} \quad x_- \geq \sum_{n=0}^{\infty}(QR_a)^{n+1}e$$

holds (see [FN05b]).

- The elements $b_\pm$ defined in the proof of Theorem 2.14 depend monotone increasing on the positive element a.

The next factorization result appeared implicitly in [FN05b, Theorem 4.1]. Our theorem can be seen as a stronger version.

Theorem 2.16. *Let $\mathcal{A} = \mathcal{A}_+ \dot{+} \mathcal{A}_0 \dot{+} \mathcal{A}_-$ be a semi-strongly decomposing ordered Banach algebra with unit e and closed normal algebra cone $\mathcal{C}$ where $e \in \mathcal{C}$. For a (positive) element $a \in \mathcal{C}$ are equivalent*

(i) $\operatorname{spr}(a) < 1$,

(ii) $e - a = (e - a_-)a_0(e - a_+)$ *where* $a_\pm \in \mathcal{A}_\pm \cap \mathcal{C}$, $\operatorname{spr}(a_\pm) < 1$, $a_0 \in \mathcal{A}_0$ *is invertible and* $a_0^{-1} \in \mathcal{C}$.

The factorization in (ii) is unique.

Proof. We define $\tilde{\mathcal{A}}_+ := \mathcal{A}_+ + \mathcal{A}_0$ and $\tilde{\mathcal{A}}_- := \mathcal{A}_-$, then $\mathcal{A} = \tilde{\mathcal{A}}_+ + \tilde{\mathcal{A}}_-$ and this sum is direct. Now we define $\tilde{P} = P_+ + P_0$ and $\tilde{Q} = P_-$, where P_+, P_0 and P_- are the projections corresponding to the semi-strong decomposition of $\mathcal{A}$. The cone $\mathcal{C}$ is invariant under these 5 projections.

(i) $\Rightarrow$ *(ii)* We can apply Theorem 2.14 and obtain $\tilde{b}_\pm \in \tilde{\mathcal{A}}_\pm \cap \mathcal{C}$ such that $\operatorname{spr}(\tilde{b}_\pm) < 1$ and $e - a = (e - \tilde{b}_-)(e - \tilde{b}_+)$. Then $\tilde{b}_+ \in \tilde{\mathcal{A}}_+ = \mathcal{A}_+ + \mathcal{A}_0$ implies $\tilde{b}_+ = b_+ + b_0$, where

$$b_+ = P_+(\tilde{b}_+) \in \mathcal{A}_+ \cap \mathcal{C} \quad \text{and} \quad b_0 = P_0(\tilde{b}_+) \in \mathcal{A}_0 \cap \mathcal{C}.$$

Now $0 \leq b_+ \leq \tilde{b}_+$ and therefore $\operatorname{spr}(b_+) \leq \operatorname{spr}(\tilde{b}_+) < 1$ and similarly $\operatorname{spr}(b_0) \leq \operatorname{spr}(\tilde{b}_+) < 1$. Therefore $e - b_+$ and $e - b_0$ are invertible and their inverses belong to $\mathcal{C}$.
Define

$$a_+ := (e - b_0)^{-1}b_+, \quad a_0 := e - b_0, \quad a_- := \tilde{b}_-.$$

Then

$$e - a = (e - a_-)a_0(e - a_+)$$

and

- $a_- \in \mathcal{A}_- \cap \mathcal{C}$ and $\operatorname{spr}(a_-) = \operatorname{spr}(\tilde{b}_-) < 1$,
- ${a_0}^{-1} = (e - b_0)^{-1} \in \mathcal{A}_0 \cap \mathcal{C}$ and therefore $a_0^{-1} \in \mathcal{C}$ and by Definition 2.1 $a_0 = (a_0^{-1})^{-1} \in \mathcal{A}_0$,
- $a_+ \in \mathcal{A}_+ \cap \mathcal{C}$ by Definition 2.1 and $(e - a_+)^{-1} = e + (e - \tilde{b}_+)^{-1} b_+ \in \mathcal{C}$ which is equivalent to $\operatorname{spr}(a_+) < 1$.

$(ii) \Rightarrow (i)$ follows as in the corresponding implication in the proof of Theorem 2.14.

For the uniqueness of the factorization in (ii) assume there exist two factorizations

$$(e - a) = (e - a_-)a_0(e - a_+) = (e - b_-)b_0(e - b_+).$$

Since $\operatorname{spr}(a_\pm), \operatorname{spr}(b_\pm) < 1$ it follows with Neumann's series that $(e - a_\pm), (e - b_\pm)$ are invertible and

$$(e - a-_\pm)^{-1} - e, (e - b_\pm)^{-1} - e \in \mathcal{A}_\pm.$$

Then also $a_0, b_0 \in \mathcal{A}_{inv}$ and we can write

$$(e - a_-)^{-1}(e - b_-) = a_0(e - a_+)(e - b_+)^{-1} b_0^{-1} \tag{2.11}$$

where all factors on the left-hand side belong to $\mathcal{A}_- + \mathcal{A}_0$ and all factors on the right-hand side belong to $\mathcal{A}_+ + \mathcal{A}_0$. It follows that the left-hand side belongs to $\mathcal{A}_0$. Moreover

$$(e - a_-)^{-1}(e - b_-) - e = \underbrace{(e - a_-)^{-1} - e}_{\in \mathcal{A}_-} - \underbrace{(e - a_-)^{-1} b}_{\in \mathcal{A}_-}$$

so that $(e - a_-)^{-1}(e - b_-) - e \in \mathcal{A}_- \cap \mathcal{A}_0 = \{0\}$. This implies $a_- = b_-$ and analogously one shows $a_+ = b_+$. Then also $a_0 = b_0$. $\square$

Remark 2.17. In the case when $\mathcal{A}$ is a strongly decomposing ordered Banach algebra one can replace assertion (ii) in Theorem 2.16 by the relaxed assertion

(ii)' $e - a = (e - a_-)a_0(e - a_+)$ where $a_\pm \in \mathcal{A}_\pm \cap \mathcal{C}$, $a_0 \in \mathcal{A}_0$ is invertible and $a_0^{-1} \in \mathcal{C}$.

The omitted assumption $\operatorname{spr}(a_\pm) < 1$ then follows from Proposition 2.5.(iii) since $a_\pm \in \mathcal{C}$ and by Definition 2.6.(iii).b) also $(e - a_\pm)^{-1} - e \in \mathcal{C}$.
This is not true if $\mathcal{A}$ is only a semi-strongly decomposing ordered Banach algebra. The implication (i) $\Rightarrow$ (ii)' obviously still holds but the reverse implication as well as the uniqueness of the decomposition do no longer hold.
To see this consider the Wiener algebra $W(\mathbb{C})$ from Example 2.7.(ii) with the complex numbers as coefficient algebra and closed normal algebra cone $\mathbb{R}_+$.
First look at

$$e_W(\lambda) - a(\lambda) = 1 - 5\lambda$$

which is in itself a factorization as in (ii)'. However, since $a(\lambda) = 5\lambda$, we have $\operatorname{spr}(a(\cdot)) = 5$ so that the reverse implication in Theorem 2.16 does not hold.
For the uniqueness consider

$$\begin{aligned} e_W(\lambda) - a(\lambda) &= 1 - (\alpha\beta\lambda^{-1} + 1 - \beta - \alpha\beta\gamma + \beta\gamma\lambda) \\ &= (1 - \alpha\lambda^{-1})\beta(1 - \gamma^{-1}) \\ &= (1 - \gamma^{-1}\lambda^{-1})(\alpha\beta\gamma)(1 - \alpha^{-1}\lambda^{-1}) \end{aligned}$$

with coefficients $\alpha, \beta, \gamma \in \mathbb{R}_+$. Choosing for example

$$\alpha, \gamma = 0.1, \qquad \beta = 0.5$$

it is easy to see that $\operatorname{spr}(a(\cdot)) \leq \|a(\cdot)\|_W < 1$ and that both factorization satisfy (ii)'. However only the first factorization also satisfies assumption (ii) and is thus the unique factorization in Theorem 2.16.

We present examples that show the necessity of some of the assumptions in Theorem 2.14 and Theorem 2.16.

Example 2.18. Set $\mathcal{A} = \mathbb{C}^{2\times 2}$ which can be written as $\mathcal{A} = \mathcal{A}_- \dot{+} \mathcal{A}_0 \dot{+} \mathcal{A}_+$ where $\mathcal{A}_-$ (resp. $\mathcal{A}_+$) denote the strictly lower (resp. strictly upper) triangular 2×2-matrices and $\mathcal{A}_0$ the diagonal 2×2-matrices, and consider the normal algebra cone $\mathcal{C}$ in $\mathcal{A}$ of entrywise nonnegative 2×2-matrices which makes $\mathcal{A}$ a strongly decomposing ordered Banach algebra. Then for a matrix $A \in \mathcal{A}$ a factorization

$$I - A = (I - A_-)A_0(I - A_+)$$

as in Theorem 2.16 coincides with an *LDU* factorization of the matrix $I - A$ where L (resp. U) is a lower (resp. upper) triangular matrix with all diagonal entries equal to 1 and D is a diagonal matrix. By [HJ85, Corollary 3.5.5] every invertible square matrix admits such a factorization if and only if its leading principal minors are not equal to zero. In that case the *LDU* factorization is unique.
(i) Take $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \in \mathcal{C}$ which can be factorized as

$$I - A = (I - A_-)A_0(I - A_+) := \left(I - \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix}\right) \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \left(I - \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix}\right).$$

However $A_-, A_0^{-1}, A_+ \notin \mathcal{C}$ in contrast to Theorem 2.16. Here $\operatorname{spr}(A) = 4 > 1$.

(ii) Set $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$ and note that $A^2 = 0$ and thus $\operatorname{spr}(A) = 0 < 1$. Then

$$I - A = \begin{bmatrix} 0 & -i \\ -i & 2 \end{bmatrix}$$

is invertible but the first leading principal minor is 0. As such the matrix $I - A$ does not admit an *LDU* factorization. Here $A \notin \mathcal{C}$.

Remark 2.19. We have written our factorizations such that $(e - a_-)$ is the left factor and $(e - a_+)$ is the right factor. We could simply swap the algebras $\mathcal{A}_-$ and $\mathcal{A}_+$ in the above proofs to obtain factorizations

$$e - a = (e - a_+)(e - a_-)$$

and

$$e - a = (e - a_+)a_0(e - a_-).$$

Note however, that such factorizations will in general result in different elements $a_\pm \in \mathcal{A}_\pm$ (i.e. the factors $(e - a_+)$ and $(e - a_-)$ do not commute). For example

$$\begin{aligned} I - \begin{bmatrix} 1/3 & 1/3 \\ 1/3 & 1/3 \end{bmatrix} &= \left(I - \begin{bmatrix} 0 & 0 \\ 1/2 & 0 \end{bmatrix} \right) \left(I - \begin{bmatrix} 1/3 & 1/3 \\ 0 & 1/2 \end{bmatrix} \right) \\ &\neq \left(I - \begin{bmatrix} 1/3 & 1/3 \\ 0 & 1/2 \end{bmatrix} \right) \left(I - \begin{bmatrix} 0 & 0 \\ 1/2 & 0 \end{bmatrix} \right) = I - \begin{bmatrix} 1/6 & 1/3 \\ 1/4 & 1/2 \end{bmatrix}. \end{aligned}$$

2.3. Applications

We now present various applications of the preceding two results.

2.3.1. M-matrices

Consider matrices

$$B = \beta I - A \tag{2.12}$$

with scalar $\beta > 0$ and $A \in \mathbb{R}_+^{n,n}$. A matrix of the form (2.12) is called an *M-matrix* if the spectral radius of A satisfies $\operatorname{spr}(A) \leq \beta$. An M-matrix is *nonsingular* if and only if $\operatorname{spr}(A) < \beta$.

Theorem 2.20. *Let $B \in \mathbb{C}^{n,n}$ be of the form* (2.12) *and $\mathbb{P}$ be a chain of order idempotents on $\mathbb{C}^n$. Then the following assertions are equivalent:*

(i) B is a nonsingular M-matrix,

(ii) there exist unique $(n \times n)$-matrices $B_+ = I - A_+, B_- = I - A_-$ and A_0 such that

$$B = B_- A_0 B_+,$$

$B_\pm$ are nonsingular M-matrices, $A_\pm \in \mathcal{A}_\pm(\mathbb{P}) \cap \mathbb{R}_+^{n,n}$, $A_0 \in \mathcal{A}_0(\mathbb{P})$ is invertible and $A_0^{-1} \in \mathbb{R}_+^{n,n}$.

Proof. Since B is of the form (2.12) it can be written as $B = \beta I - A$ with $A \in \mathbb{R}_+^{n,n}$. Note that B being a nonsingular M-matrix is equivalent to $\operatorname{spr}(\frac{1}{\beta}A) < 1$. Applying Theorem 2.16 to $\frac{1}{\beta}B$ gives a factorization

$$\frac{1}{\beta}B = B_- \tilde{A}_0 B_+.$$

Now define $A_0 := \beta \tilde{A}_0$ to get the assertion. $\square$

This equivalence is known for the case of the canonical chain on $\mathbb{C}^n$ from Example 2.12.(i), i.e. the chain $\mathbb{P} = \{P_k : k = 1, \ldots, n\}$ with $P_0 = 0$ and

$$P_k : \mathbb{C}^n \to \mathbb{C}^n : (x_j)_{j=1}^n \mapsto (y_j)_{j=1}^n,$$

where $y_j = x_j$ for $j = 1, \ldots, k$ and $y_j = 0$ for $j = k+1, \ldots, n$ (see e.g. [BP79, Theorem 2.3], [HJ91, p. 117]).

I. Kuo (see [Kuo77]) proved that for a singular irreducible M-matrix there exists an *LU*-factorization with respect to the canonical chain of projections on $\mathbb{C}^n$. This result also holds for arbitrary chains of order idempotents:

Theorem 2.21. *Let $B = \beta I - A \in \mathbb{C}^{n,n}$ be a singular irreducible M-matrix and $\mathbb{P}$ a chain of order idempotents on $\mathbb{C}^n$. Then B admits an LU-factorization*

$$B = B_- A_0 B_+$$

such that $B_\pm = I - A_\pm$ are M-matrices with $A_\pm \in \mathcal{A}_\pm(\mathbb{P})$ and $A_0 \in \mathcal{A}_0(\mathbb{P})$.

Proof. Without loss of generality assume that $\beta = 1$ (for details see the proof of Theorem 2.20). Note that an order idempotent projection in $\mathbb{C}^n$ is necessarily the sum of canonical coordinate projections in $\mathbb{C}^n$. Therefore there is a 1-to-1 correspondence between index sets $\mathcal{I} \subset \{1, \ldots, n\}$ and order idempotent projections

$$P_{\mathcal{I}} : \mathbb{C}^n \to \mathbb{C}^n : (z_k)_k \mapsto (z_k \delta_{\mathcal{I}}(k))_k,$$

with

$$\delta_{\mathcal{I}}(k) = \begin{cases} 1 & , k \in \mathcal{I} \\ 0 & , \text{otherwise} \end{cases}.$$

As such we can identify the matrix $P_{\mathcal{I}} A P_{\mathcal{I}}$ with the principal minor $A[\mathcal{I}, \mathcal{I}]$ of A. Since A is irreducible it follows by [BR97, Theorem 1.7.4] that for $\mathcal{I} \neq \{1, \ldots, n\}$

$$\operatorname{spr}(P_{\mathcal{I}} A P_{\mathcal{I}}) = \operatorname{spr}(A[\mathcal{I}, \mathcal{I}]) < \operatorname{spr}(A)$$

and $I - P_{\mathcal{I}} A P_{\mathcal{I}}$ is invertible.
Let $\mathbb{P} = \{0 = P_0 < P_1 < \ldots < P_l = I\}$ be the chain of order idempotents. We can now apply the same argument as in the reverse implication of the proof of [GGK93, Theorem XXII.1.1] to obtain the factorization with respect to $\mathbb{P}$

$$B = (I - A_-) A_0 (I - A_+)$$

with $A_\pm \in \mathcal{A}_\pm(\mathbb{P})$ and invertible $A_0 \in \mathcal{A}_0(\mathbb{P})$. Note that for the argument in [GGK93] we need $I - P_j A P_j$ to be invertible only for $j = 1, \ldots, l-1$ and not for $j = l$. $\square$

2.3.2. *LU*-factorizations of Hilbert-Schmidt perturbations of the identity

Let $H = L_2(\Omega)$ be the Hilbert space of square integrable functions on a domain Ω. We then consider the space of Hilbert-Schmidt integral operators $\mathcal{S}_2(H)$ on H, that is operators of the form

$$(Kf)(x) = \int_{y\in\Omega} k(x,y)f(y)dy, \qquad x \in \Omega$$

where the measurable kernel function k satisfies

$$\int_{x\in\Omega}\int_{y\in\Omega} |k(x,y)|^2 dydx < \infty. \tag{2.13}$$

Equation (2.13) gives rise to the Hilbert-Schmidt norm with which together $\mathcal{S}_2(H)$ becomes a Banach space and with the usual multiplication it becomes a Banach algebra without unit element. $\mathcal{S}_2(H)$ becomes an ordered Banach algebra with the algebra cone $\mathcal{C}_2 := \{K \in \mathcal{S}_2(H) : k \geq 0 \text{ a.s.}\}$. Note that $\mathcal{C}_2$ is normal since the Hilbert-Schmidt norm is monotone.
Let $\mathbb{P}$ be an orthogonal chain of order idempotents. Then by [GGK93, Corollary XX.8.2]

$$\mathcal{S}_2(H) = \mathbb{A}_{2,-}(H,\mathbb{P}) + \mathbb{A}_{2,0}(H,\mathbb{P}) + \mathbb{A}_{2,+}(H,\mathbb{P}) \tag{2.14}$$

gives rise to an additive decomposition as in Section 2.1.1. But since for our purposes we need $\mathbb{A}_{2,0}(H,\mathbb{P})$ to contain a unit element we need to modify $\mathcal{S}_2(H)$. Define

$$\mathbb{A}^{(2)}(H) := \{\lambda I_H + K : \lambda \in \mathbb{C}, K \in \mathcal{S}_2(H)\} \tag{2.15}$$

with norm

$$\|\lambda I_H + K\|_{(2)} := \max\{|\lambda|, \|K\|_2\}. \tag{2.16}$$

We can now introduce the decomposition

$$\mathbb{A}^{(2)}(H) = \mathbb{A}^{(2)}_-(H,\mathbb{P}) + \mathbb{A}^{(2)}_0(H,\mathbb{P}) + \mathbb{A}^{(2)}_+(H,\mathbb{P}) \tag{2.17}$$

where

$$\begin{aligned}
&\mathbb{A}^{(2)}_-(H,\mathbb{P}) = \mathbb{A}_{2,-}(H,\mathbb{P}),\\
&\mathbb{A}^{(2)}_+(H,\mathbb{P}) = \mathbb{A}_{2,+}(H,\mathbb{P}),\\
&\mathbb{A}^{(2)}_0(H,\mathbb{P}) = \{\lambda I_H + K : \lambda \in \mathbb{C}, K \in \mathbb{A}_{2,0}(H,\mathbb{P})\}.
\end{aligned}$$

In the same way we can modify the cone $\mathcal{C}_2$ to obtain an algebra cone with unit

$$\mathcal{C}^{(2)} := \{\lambda I_H + K : \lambda \geq 0, K \in \mathcal{C}_2\}.$$

Proposition 2.22. *Let $\mathbb{A}^{(2)}(H)$ with the normal algebra cone $\mathcal{C}^{(2)}$ be as above and let $\mathbb{P}$ be an orthogonal chain of order idempotents. Then*

$$\mathbb{A}^{(2)}(H) = \mathbb{A}_-^{(2)}(H,\mathbb{P}) + \mathbb{A}_0^{(2)}(H,\mathbb{P}) + \mathbb{A}_+^{(2)}(H,\mathbb{P}) \tag{2.18}$$

is a strongly decomposing ordered Banach algebra.

Proof. This follows from Proposition 2.11. □

We can now apply Theorem 2.16.

2.3.3. Factorization in the Banach algebra of regular operators

When decomposing operators with respect to a chain of projections as in Section 2.1.1 it turns out that not every bounded operator has a diagonal and thus cannot be decomposed (see Example 2.24). In this section we will define decompositions with respect to chains of order idempotents in a weaker sense for regular operators (which are a subset of bounded operators), so that we can still apply Theorem 2.16.

Let E be a real or the complexification of a real Banach lattice with order continuous norm (for reference see e.g. [MN91, Sch74, Wnu99]). The set $\mathcal{L}^r(E)$ of regular operators in E (= the linear span of all positive operators in E) is an order complete lattice algebra which is a Banach algebra with the r-norm $\|A\|_r = \| \, |A| \, \|_{\mathcal{L}(E)}$ for $A \in \mathcal{L}^r(E)$ (see [MN91, p.27], [Sch74, Proposition IV.1.3]) and $\mathcal{L}(E)_+$ is a closed normal algebra cone in $\mathcal{L}^r(E)$. Let $\mathbb{P}$ be a chain of order idempotents on E, i.e. all $P \in \mathbb{P}$ satisfy $0 \leq P \leq I$ in the order of $\mathcal{L}(E)$; which is equivalent to all $P \in \mathbb{P}$ being band projections on E (see [Sch74, p.61]).
For a finite subchain $\pi = \{0 = P_0, P_1, \ldots, P_n = I\} \subset \mathbb{P}$ (in the remainder of the section a small π will always denote a finite subchain of $\mathbb{P}$) and $A \in \mathcal{L}^r(E)$ we define the diagonal of A with respect to π as

$$D_r(A,\pi) := \sum_{j=1}^{n} (\Delta P_j) A (\Delta P_j) \quad \in \mathcal{L}^r(E) \tag{2.19}$$

where $(\Delta P_j) = P_j - P_{j-1}$ (see [GGK93, p.473]).
For $A \in \mathcal{L}(E)_+$ the system $\{D_r(A,\pi) : \pi \subset \mathbb{P}\}$ is downwards directed with respect to the refinements of subchains of $\mathbb{P}$; indeed if $\pi_1, \pi_2 \subset \mathbb{P}$ then

$$\pi_1 \subset \pi_2 \quad \Rightarrow \quad 0 \leq D_r(A,\pi_2) \leq D_r(A,\pi_1) \leq A.$$

In addition for a positive element $x \in E_+$ the set $\{D_r(A,\pi)x : \pi \in \mathbb{P}\} \subset E_+$ is also downwards directed. Since the norm on E is order continuous and E is Dedekind complete the vector $\inf_\pi D_r(A,\pi)x$ exists in E and $\lim_\pi D_r(A,\pi)x = \inf_\pi D_r(A,\pi)x$ where the limit exists with respect to the norm topology of E.
Each $x \in E$ can be written as a linear combination of at most four vectors in E_+ and each $A \in \mathcal{L}^r(E)$ is the linear combination of at most four operators in $\mathcal{L}(E)_+$, therefore:

For all $A \in \mathcal{L}^r(E)$ and all $x \in E$ the limit $\lim_\pi D_r(A,\pi)x$ exists. Now by the Banach-Steinhaus Theorem the operator

$$D_r(A,\mathbb{P}) : E \to E : x \mapsto D_r(A,\mathbb{P})x$$

exists in $\mathcal{L}^r(E)$ as a pointwise limit of the $D_r(A,\pi)$.
It follows that the map $\mathcal{L}^r(E) \to \mathcal{L}^r(E) : A \mapsto D_r(A,\mathbb{P})$ is linear, maps $\mathcal{L}(E)_+$ into itself and

$$|D_r(A,\mathbb{P})| \leq D_r(|A|,\mathbb{P}) \leq |A|$$

for all $A \in \mathcal{L}^r(E)$. We call $D_r(A,\mathbb{P})$ the diagonal of A with respect to $\mathbb{P}$.
In analogy to Section 2.1.1 we define the subspaces

$$\begin{aligned}
\mathbb{A}_{r,-}(E,\mathbb{P}) &= \{A \in \mathcal{L}^r(E) : PA = PAP \text{ for all } P \in \mathbb{P}, D_r(A,\mathbb{P}) = 0\}, \\
\mathbb{A}_{r,+}(E,\mathbb{P}) &= \{A \in \mathcal{L}^r(E) : AP = PAP \text{ for all } P \in \mathbb{P}, D_r(A,\mathbb{P}) = 0\}, \\
\mathbb{A}_{r,0}(E,\mathbb{P}) &= \{A \in \mathcal{L}^r(E) : AP = PA \text{ for all } P \in \mathbb{P}\}.
\end{aligned}$$

Proposition 2.23. *Let E ba a Banach lattice with order continuous norm and $\mathbb{P}$ be a chain of order idempotents on E. Then*

$$\mathcal{L}^r(E) = \mathbb{A}_{r,-}(E,\mathbb{P}) \dotplus \mathbb{A}_{r,0}(E,\mathbb{P}) \dotplus \mathbb{A}_{r,+}(E,\mathbb{P}) \tag{2.20}$$

is a semi-strongly decomposing ordered Banach algebra. The canonical projections are order idempotents.

Proof. For a finite subchain $\pi = \{0 = P_0, P_1, \ldots, P_n = I\} \subset \mathbb{P}$ and $A \in \mathcal{L}^r(E)$ we define

$$L_r(A,\pi) = \sum_{0 \leq j < k \leq n} (\Delta P_k)A(\Delta P_j) \quad \text{and} \quad U(A,\pi) = \sum_{0 \leq j < k \leq n} (\Delta P_j)A(\Delta P_k).$$

For $A \in \mathcal{L}(E)_+$ it follows that $0_E \leq L_r(A,\pi) \leq A$ and it is easy to see that the system $\{L_r(A,\pi) : \pi \in \mathcal{P}\}$ is upwards directed. Then for all $x \in E_+$ the set $\{L_r(A,\pi)x : \pi \in \mathcal{P}\}$ is upwards directed in E and $\lim_\pi L_r(A,\pi)x$ exists since $0 \leq L_r(A,\pi)x \leq Ax$ and the norm of E is order continuous. As in the case of the diagonal $D_r(A,\mathbb{P})$ it follows that $\lim_\pi L(A,\pi)x$ exists in the norm topology of E for all $A \in \mathcal{L}^r(E)$ and for all $x \in E$, therefore the operator

$$L_r(A,\mathbb{P}) : E \to E : x \mapsto \lim_\pi L_r(A,\pi)x$$

belongs to $\mathcal{L}^r(E)$ and $|L_r(A,\mathbb{P})| \leq L_r(|A|,\mathbb{P}) \leq |A|$. Similarly the operator

$$U_r(A,\mathbb{P}) : E \to E : x \mapsto \lim_\pi U_r(A,\pi)x$$

is well defined, belongs to $\mathcal{L}^r(E)$ and $|U_r(A,\mathbb{P})| \leq U_r(|A|,\mathbb{P}) \leq |A|$.
Now for all $A \in \mathcal{L}^r(E)$ there holds

$$A = L_r(A,\mathbb{P}) + D_r(A,\mathbb{P}) + U_r(A,\mathbb{P})$$

since for all $\pi \subset \mathbb{P}$

$$A = L_r(A,\pi) + D_r(A,\pi) + U_r(A,\pi).$$

We need to show that for all $A \in \mathcal{L}^r(E)$

$$L_r(A,\mathbb{P}) \in \mathbb{A}_{r,-}(E,\mathbb{P}), \quad D_r(A,\mathbb{P}) \in \mathbb{A}_{r,0}(E,\mathbb{P}) \quad \text{and} \quad U_r(A,\mathbb{P}) \in \mathbb{A}_{r,+}(E,\mathbb{P}). \qquad (2.21)$$

Consider the first assertion: Let $P \in \mathbb{P}$ and $\pi = \{0 = P_0, P_1, \ldots, P_n = I\} \subset \mathbb{P}$ a finite subchain, then

$$PL_r(A,\pi) = PL_r(A,\pi)P$$

if $P \in \pi$. Indeed let $P = P_l$ for some $0 \leq l \leq n$, then

$$P_l L_r(A,\pi) = \sum_{0 \leq j < k \leq l} (\Delta P_k) A (\Delta P_j) = P_l L_r(A,\pi) P_l.$$

Thus $PL_r(A,\mathbb{P}) = PL_r(A,\mathbb{P})P$ for all $P \in \mathbb{P}$. To see that $D_r(L_r(A,\mathbb{P}),\mathbb{P}) = 0$ holds for all $A \in \mathcal{L}^r(E)$ first note that $D_r(L_r(A,\pi),\pi) = 0$ for any finite subchain $\pi \subset \mathbb{P}$. Therefore for all $x \in E_+$ and all $A \in \mathcal{L}(E)_+$

$$\begin{aligned} 0 \leq D_r(L_r(A,\mathbb{P}),\pi)x &= \big(D_r(L_r(A,\mathbb{P}),\pi) - D_r(L_r(A,\pi),\pi)\big)x \\ &= D_r(L_r(A,\mathbb{P}) - L_r(A,\pi),\mathbb{P})x \\ &= \big(L_r(A,\mathbb{P}) - L_r(A,\pi)\big)x \end{aligned}$$

noting that $0 \leq L_r(A,\pi) \leq L_r(A,\mathbb{P}) \leq A$ since $A \in \mathcal{L}(E)_+$. This implies that

$$D_r(L_r(A,\mathbb{P}),\mathbb{P})x = 0$$

for all $x \in E_+$ and $A \in \mathcal{L}(E)_+$, but then this equally holds for all $x \in E$ and $A \in \mathcal{L}^r(E)$. Therefore $D_r(L_r(A,\mathbb{P}),\mathbb{P}) = 0$ for all $A \in \mathcal{L}^r(E)$ and $L_r(A,\mathbb{P}) \in \mathbb{A}_{r,-}(E,\mathbb{P})$.
The other two assertion in (2.21) follow similarly. In summary: The decomposition (2.20) holds.
Of course $I \in \mathbb{A}_{r,0}(E,\mathbb{P})$ and for $A \in \mathbb{A}_{r,0}(E,\mathbb{P}) \cap \mathcal{L}^r(E)_{inv}$ and $P \in \mathbb{P}$ the equality $PA = AP$ implies $A^{-1}P = PA^{-1}$, so that Definition 2.1.(ii).a) holds.
To see that Definition 2.1.(ii).b) holds it suffices to show that for $A \in \mathcal{L}^r(E)$ and $B \in \mathbb{A}_{r,0}(E,\mathbb{P})$ there holds

$$\begin{aligned} D_r(AB,\mathbb{P}) &= D_r(A,\mathbb{P})D_r(B,\mathbb{P}), \\ D_r(BA,\mathbb{P}) &= D_r(B,\mathbb{P})D_r(A,\mathbb{P}). \end{aligned}$$

Let us consider the first equation. For a finite subchain

$$\begin{aligned}
D_r(A,\pi)D_r(B,\pi) &= \sum_{j=0}^{n}(\Delta P_j)A(\Delta P_j)\sum_{k=0}^{n}(\Delta P_k)B(\Delta P_k) \\
&= \sum_{j,k=0}^{n}(\Delta P_j)A\underbrace{(\Delta P_j)(\Delta P_k)}_{=\delta_{jk}(\Delta P_j)}B(\Delta P_k) \\
&= \sum_{j=0}^{n}(\Delta P_j)A(\Delta P_j)B(\Delta P_j) \\
&= \sum_{j=0}^{n}(\Delta P_j)AB(\Delta P_j)(\Delta P_j) \\
&= \sum_{j=0}^{n}(\Delta P_j)AB(\Delta P_j) = D_r(AB,\pi).
\end{aligned}$$

Now for all $x \in E$

$$D_r(AB,\mathbb{P})x = \lim_{\pi} D_r(AB,\pi)x = \lim_{\pi} D_r(A,\pi)D_r(B,\pi)x = D_r(A,\mathbb{P})D_r(B,\mathbb{P})x$$

proving the assertion. $\square$

It is now possible to apply Theorem 2.16 to $\mathcal{L}^r(E) = \mathbb{A}_{r,-}(E,\mathbb{P}) \dot{+} \mathbb{A}_{r,0}(E,\mathbb{P}) \dot{+} \mathbb{A}_{r,+}(E,\mathbb{P})$.

One could ask whether the decomposition introduced in this section for regular operators is really a generalization of the theory in Section 2.1.1 for bounded operators. The next example shows that this is indeed the case by presenting an operator that has a diagonal as a regular operator but not as a bounded operator.
First we make an additional observation: By [TL80, Theorem IV.6.3] and [TL80, Problem IV.6.4] every operator A in the classical sequence spaces c_0 and ℓ_1 can be represented as an infinite matrix $A = (a_{rs})_{r,s\in\mathbb{N}}$. Additionally for $A \in \mathcal{L}(c_0)$ the operator norm of A can be represented by

$$\|A\| = \sup_r \sum_s |a_{rs}| = \|\,|A|\,\| = \|A\|_r$$

which shows that A is regular. A similar result holds for $A \in \mathcal{L}(\ell_1)$, where in the above equation one takes the supremum over the column sums instead. We conclude that every bounded operator in c_0 and in ℓ_1 is regular.
This is no longer true in ℓ_2. Even if a bounded operator admits a matrix representation $(a_{rs})_{r,s\in\mathbb{N}}$, the matrix $(|a_{rs}|)_{r,s\in\mathbb{N}}$ will in general not define a bounded operator. For an explicit example of such an operator see [HS78, Example 10.1].

Example 2.24. Let S be the left shift operator in c_0 or ℓ_1. For S we have the matrix representation

$$S \mathrel{\widehat{=}} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Consider the chain $\mathbb{P} = \{0, P_1, P_2, \dots, I\}$ where P_k is the projection on the first k coordinates.

We first show that S has a diagonal as a regular operator: Let $n \in \mathbb{N}$. It is easy to see that

$$\Delta P_n S \Delta P_n = 0$$

and

$$(1 - P_n)S(1 - P_n) \mathrel{\widehat{=}} \begin{bmatrix} 0 & \dots & 0 & & & & \\ \vdots & \ddots & \vdots & & & & \\ 0 & \dots & 0 & & & & \\ & & & 0 & 1 & & \\ & & & & 0 & 1 & \\ & & & & & \ddots & \ddots \end{bmatrix}.$$

For $x = (x_j)_{j \in \mathbb{N}} \in c_0$ it follows that

$$(1 - P_n)S(1 - P_n)x = \begin{bmatrix} 0 & \cdots & 0 & x_{n+1} & x_{n+2} & \cdots \end{bmatrix}^T \xrightarrow{n \to \infty} 0.$$

Then for any finite subchain $\pi_n = \{0, P_1, \dots, P_n, I\}$ we have

$$D(S, \pi_n)x = \sum_{j=1}^{n} \Delta P_j S \Delta P_j x = (1 - P_n)S(1 - P_n)x \xrightarrow{n \to \infty} 0.$$

It follows that $D_r(S, \mathbb{P}) = 0$.

However $\|(1 - P_n)S(1 - P_n)\| = 1$ and then also $\|D(S, \pi_n)\| = 1$. Therefore the limit $\lim_{n \to \infty} D(S; \pi_n)$ does not exist in the norm topology and S does not have a diagonal as in Section 2.1.1.

2.3.4. Wiener algebra and Laurent polynomials

Recall the Wiener algebra

$$W(\mathcal{A}) = W_-(\mathcal{A}) + W_0(\mathcal{A}) + W_+(\mathcal{A}) \tag{2.22}$$

with cone $\mathcal{C}_W$ from Example 2.7.(ii) where we showed that (2.22) is a semi-strongly decomposing ordered Banach algebra as long as the coefficient algebra $\mathcal{A}$ is an ordered Banach algebra with closed normal algebra cone $\mathcal{C}$.

The Wiener algebra is a subalgebra of $C(\mathbb{T}_1, \mathcal{A})$ (= the set of continuous functions from the unit circle into $\mathcal{A}$) endowed with the supremums norm $\|\cdot\|_\infty$. For $a(\cdot) \in W(\mathcal{A})$ there holds

$$\|a(\cdot)\|_\infty = \max_{|\lambda|=1} \|a(\lambda)\|_{\mathcal{A}} \leq \|a(\cdot)\|_W$$

where subscripts indicate the respective algebra. We may also by superscripts denote with regards to which algebra we take the spectral radius.

Now let us apply Theorem 2.16 to $W(\mathcal{A})$ to get:

Theorem 2.25. *Let A be an ordered Banach algebra with closed normal cone $\mathcal{C}$. Then for $a(\cdot) \in W(\mathcal{A}) \cap \mathcal{C}_W$ are equivalent:*

(i) $\operatorname{spr}^W(a(\cdot)) < 1$,

(ii) $e_W(\cdot) - a(\cdot) = (e_W(\cdot) - a_-(\cdot))a_0(\cdot)(e_W(\cdot) - a_+(\cdot))$, *where* $a_\pm(\cdot) \in W_\pm(\mathcal{A}) \cap \mathcal{C}_W$, $\operatorname{spr}^W(a_\pm(\cdot)) < 1$, $a_0(\cdot) \in W_0(\mathcal{A})$ *is invertible and* $a_0^{-1}(\cdot) \in \mathcal{C}_W$.

A factorization result for the Wiener algebra as a decomposing Banach algebra (without order) was previously established in [CG81, Theorem 6.1].
It is possible to obtain a stronger result if we restrict ourselves to Laurent polynomials, that is functions $a(\cdot) \in W(\mathcal{A})$ with only finitely many non-zero Fourier coefficients. We say that $a(\cdot)$ is of degree $\leq m \in \mathbb{N}$ if

$$a(\lambda) = \sum_{k=-m}^{m} \lambda^k a_k, \quad \lambda \in \mathbb{T}$$

where we do not require $a_{-m} \neq 0$ or $a_m \neq 0$. Note that $W(\mathcal{A})$ is a subalgebra of $C(\mathcal{A}, \mathbb{T})$ of continuous functions on the unit circle with supremums norm $\|\cdot\|_\infty$. Moreover $\|a(\cdot)\|_\infty \leq \|a(\cdot)\|_W$ for all $a(\cdot) \in W(\mathcal{A})$.
We will need two lemmas to state our result.

Lemma 2.26. *Let $\mathcal{A}$ be a Banach algebra and $a(\cdot) \in W(\mathcal{A})$ a Laurent polynomial of degree $m \in \mathbb{N}$. Then*

$$\|a(\cdot)\|_W \leq (2m+1)\|a(\cdot)\|_\infty \qquad \text{and} \qquad \operatorname{spr}^W(a(\cdot)) = \operatorname{spr}^\infty(a(\cdot)).$$

Proof. This is [FN05b, Lemma 4.3]. □

Lemma 2.27. *Let $\mathcal{A}$ be an ordered Banach algebra with closed normal algebra cone. Then there exists $\beta > 0$ such that*

$$\|a(\cdot)\|_\infty \leq \beta \|a(1)\|_A \qquad \text{and} \qquad \operatorname{spr}^\infty(a(\cdot)) = \operatorname{spr}^{\mathcal{A}}(a(1))$$

for all $a(\cdot) \in \mathcal{C}_W$.

Proof. This is [FN05b, Lemma 4.4]. □

Theorem 2.28. *Let $\mathcal{A}$ be an ordered Banach algebra with closed normal cone. Then for a Laurent polynomial $a(\cdot) \in W(\mathcal{A}) \cap \mathcal{C}_W$ of degree $\leq m$ are equivalent:*

(i) $\operatorname{spr}^{\mathcal{A}}(a(1)) < 1$,

(ii) $e_W(\cdot) - a(\cdot) = (e_W(\cdot) - a_-(\cdot))a_0(\cdot)(e_W(\cdot) - a_+(\cdot))$, *where* $a_\pm(\cdot) \in W_\pm(\mathcal{A}) \cap \mathcal{C}_W$, $\operatorname{spr}^W(a_\pm(\cdot)) = \operatorname{spr}^{\mathcal{A}}(a_\pm(1)) < 1$, $a_0(\cdot) \in W_0(\mathcal{A})$ *is invertible and* $a_0^{-1}(\cdot) \in \mathcal{C}_W$.

Proof. By Lemma 2.26 and 2.27 we have $\operatorname{spr}^W(a(\cdot)) = \operatorname{spr}^\infty(a(\cdot)) = \operatorname{spr}^{\mathcal{A}}(a(1)) < 1$ so that we can apply Theorem 2.16. □

2.4. Factorization of semi-monic polynomials

We now turn our attention back to semi-monic polynomials and apply the results of this section. Following results from [Mar88] and [FN05b] we achieve a factorization for semi-monic polynomials with coefficients in an algebra cone such that the factors have special spectral properties.

Theorem 2.29. *Let $\mathcal{A}$ be an ordered Banach algebra with identity e and $\mathcal{C}$ a closed normal algebra cone. Let*

$$q(\lambda) = \lambda^m e - a(\lambda) = \lambda^m e - \sum_{j=0}^{l} \lambda^j a_j,$$

with $a_j \in \mathcal{C}$, $j = 1, \dots, n$ and let $\rho > 0$. Then

$$\operatorname{spr}(a(\rho)) < \rho^m$$

if and only if there exists a unique factorization

$$q(\cdot) = b(\cdot) q_0 c(\cdot) \tag{2.23}$$

where

(i) $b(\cdot)$ is monic with degree m, i.e.

$$b(\lambda) = \lambda^m e - \sum_{j=0}^{m-1} \lambda^j b_j,$$

with $b_j \in \mathcal{C}$, $j = 0, \dots, m-1$ and $\Sigma(b(\cdot)) = \Sigma(q(\cdot)) \cap \{\lambda \in \mathbb{C} : |\lambda| < \rho\}$,

(ii) $q_0 \in \mathcal{A}$ is invertible and $q_0^{-1} \in \mathcal{C}$,

(iii) $c(\cdot)$ is a polynomial of the form

$$c(\lambda) = e - \sum_{j=1}^{l-m} \lambda^j c_j,$$

with $c_j \in \mathcal{C}$, $j = 1, \dots, m-l$ and $\Sigma(c(\cdot)) = \Sigma(q(\cdot)) \cap \{\lambda \in \mathbb{C} : |\lambda| > \rho\}$.

Proof. We first proof the 'only if' part: Define the function $\tilde{a}(\lambda) := \lambda^{-m}a(\lambda)$ which by (2.23) satisfies $\operatorname{spr}(\tilde{a}(\rho)) < 1$. Now setting $\tilde{a}_\rho(\lambda) := \tilde{a}(\rho\lambda)$ we obtain a function that satisfies the conditions of Theorem 2.28. Therefore $\tilde{a}_\rho(\cdot)$ factorizes as

$$e - \tilde{a}_\rho(\lambda) = (e - a_-(\lambda))a_0(e - a_+(\lambda)), \qquad \lambda \in \mathbb{T}$$

where $a_\pm(\cdot) \in W_\pm(\mathcal{A}) \cap \mathcal{C}_W$, $\operatorname{spr}(a_\pm(\cdot)) < 1$ and a_0 invertible with $a_0^{-1} \in \mathcal{C}$. Now

$$q(\rho\lambda) = \rho^m \lambda^m (e - \tilde{a}_\rho(\lambda)) = \rho^m(\lambda^m e - \lambda^m a_-(\lambda))a_0(e - a_+(\lambda)), \qquad \lambda \in \mathbb{T}.$$

Note that $a_-(\cdot)$ has an analytic extension outside of the unit circle $\mathbb{T}$ and vanishes at infinity whereas $(e - a_+(\cdot))$ has an analytic extension inside of $\mathbb{T}$. It then follows as in the proof of [Mar88, Theorem 22.11] that $b(\lambda) := \lambda^m - \lambda^m a_-(\rho^{-1}\lambda)$ is as in (i).
Now define $c(\lambda) := e - a_+(\rho^{-1}\lambda)$ and note that $(e - a_+(\cdot))^{-1} \in W(\mathcal{A}) \cap \mathcal{C}$ implies that $\Sigma(c(\cdot)) \subset \{\lambda \in \mathbb{C} : |\lambda| > \rho\}$. The assumptions in (iii) then follow immediately. The uniqueness of the decomposition follows from Theorem 2.16.
For the reverse implication, using the same notation as above, note that we can write

$$e - \tilde{a}_\rho(\lambda) = (e - a_-(\lambda))a_0(e - a_+(\lambda)), \qquad \lambda \in \mathbb{T}$$

where a_α, $\alpha = -, 0, +$ satisfy the conditions in Theorem 2.28.(ii). It then follows that $\operatorname{spr}(\tilde{a}_\rho(1)) < 1$ which in turn implies that $\operatorname{spr}(a(\rho)) < \rho^m$. ☐

The above theorem is the analogue of [FN05b, Theorem 5.1 and Corollary 5.2]. However, due to the more differentiated formulation in Theorem 2.16 compared to [FN05b, Theorem 4.1], it follows directly in our case.

As an immediate consequence we can formulate a version of Pellet's theorem for semi-monic entrywise nonnegative matrix polynomials. Our version is a modification of [Mel13, Theorem 3.3] where we require the matrix coefficients to be nonnegative but in turn achieve better bounds.

Theorem 2.30 (Pellet). *Let*

$$Q(\lambda) = \lambda^m I - \sum_{j=0}^{l} \lambda^j A_j$$

with the $A_j \in \mathbb{R}_+^{n,n}$, $j = 0, \dots, l$. *Assume there exist* $0 < \rho_1 < \rho_2$ *satisfying*

$$\operatorname{spr}(A(\rho_i)) = \rho_i^m, \qquad i = 1, 2 \tag{2.24}$$

and $\operatorname{spr}(A(\rho)) < \rho^m$ *for at least one* $\rho \in (\rho_1, \rho_2)$.
Then $Q(\cdot)$ *has exactly* nm *eigenvalues (counting multiplicities) in the closed disk* $\mathbb{D}_{\le \rho_1}$ *and no eigenvalues in the open annulus* $\mathbb{A}_{\rho_1,\rho_2}$.

Proof. By Theorem 1.12 (in the spectral case) $\Sigma(Q(\cdot)) \cap \mathbb{A}_{\rho_1,\rho_2} = \emptyset$.
By Theorem 2.29 we can factorize $Q(\cdot)$ as

$$Q(\cdot) = B(\cdot)Q_0C(\cdot)$$

where $B(\cdot)$ is a monic matrix polynomial of degree m and

$$\Sigma(B(\cdot)) = \Sigma(Q(\cdot)) \cap \{\lambda \in \mathbb{C} : |\lambda| < \rho\} = \Sigma(Q(\cdot)) \cap \mathbb{D}_{\leq \rho_1}.$$

Since $B(\cdot)$ is of degree m it has exactly nm eigenvalues (counting multiplicities), which implies the assertion. $\square$

Remark 2.31. Instead of the condition (2.24) A. Melman required in [Mel13, Theorem 3.3] that

$$\mathrm{norm}_A(\sigma_i) = \sigma_i^m, \qquad i = 1, 2$$

for some $0 < \sigma_1 < \sigma_2$. But since $\mathrm{spr}_A(\cdot) \leq \mathrm{norm}_A(\cdot)$ it follows that Theorem 2.30 offers stronger bounds if the $A_j \in \mathbb{R}_+^{n,n}$.

3. Degree-reduction of operator polynomials

Linearisations of matrix or operator polynomials are well studied in the literature (see [GLR82] for the matrix case and [Rod89] for the operator case), most famously via the companion form. They make it possible to acquire results on e.g. the eigenstructure of polynomials with arbitrary degree by only considering linear polynomials. It is then natural to generalize the concept of linearisations and reduce an operator polynomial to other degrees. Such degree reductions were previously considered in [Nag07] and [Har11] for operator polynomials.
More recently in [TDM14] the authors investigated degree reductions for matrix polynomials in the form of so called *ℓ-ififcations* and proved spectral relations between polynomials and their ℓ-ifications. In [TDD15] they proved existence results relating to degree reductions in a more abstract setting, and in [TDD] they derive an algorithm to construct certain ℓ-ifications using dual minimal bases.
In this section we will introduce a special type of degree reduction and prove a variety of related results.

Let X and Y be (real or complex) vector spaces, let $l \in \mathbb{N}$ and let $A_j : X \to Y$ be linear operators, $j = 1, \dots, l$. We consider the operator-valued polynomial

$$A(\lambda) = \sum_{j=0}^{l} \lambda^j A_j, \tag{3.1}$$

of degree l (where we do not require the leading coefficient to be nonzero). Inspired by the matrix case we call an operator polynomial $F(\cdot) : \mathbb{C} \to \mathcal{L}(X, Y)$ *unimodular* if $F(\lambda)$ is invertible for all $\lambda \in \mathbb{C}$.

Definition 3.1. Two operator polynomials $A(\cdot) : X_1 \to Y_1$ and $B(\cdot) : X_2 \to Y_2$ are said to be *extended unimodular equivalent* if for some Banach spaces Z_1, Z_2 there exist unimodular operator polynomials

$$E(\cdot) : \mathbb{C} \to \mathcal{L}(Y_2 \times Z_2, Y_1 \times Z_1),$$
$$F(\cdot) : \mathbb{C} \to \mathcal{L}(X_1 \times Z_1, X_2 \times Z_2)$$

such that

$$\operatorname{Diag}(A(\lambda), I_{Z_1}) = E(\lambda) \operatorname{Diag}(B(\lambda), I_{Z_2}) F(\lambda), \qquad \lambda \in \mathbb{C}.$$

We can now state the definition of a degree reduction.

Definition 3.2. Let $A(\lambda)$ be as in (3.1) with degree l and $B(\cdot) : \mathbb{C} \to \mathcal{L}(X_2, Y_2)$ an operator polynomial of degree $\hat{l} \in \mathbb{N}$. Then we call $B(\cdot)$ a *degree reduction of $A(\cdot)$ to degree $\hat{l}$* if $A(\cdot)$ and $B(\cdot)$ are extended unimodular equivalent.

We stated Definition 3.2 in a very general setting. In particular the degree $\hat{l}$ of the 'reduced' polynomial does not actually need to be smaller than the degree of $A(\cdot)$ (and

[TDD15] has results pertaining to this case).

We will now introduce a special type of degree reduction which we will use throughout the remainder of this chapter. For $\hat{l} \in \{1, \dots, l-1\}$ we define the *canonical* reduction of $A(\cdot)$ to degree $\hat{l}$ as the polynomial

$$\hat{A}(\lambda) = \lambda^{\hat{l}} \hat{A}_{\hat{l}} + \lambda^{\hat{l}-1} \hat{A}_{\hat{l}-1} + \dots + \lambda \hat{A}_1 + \hat{A}_0, \tag{3.2}$$

where the coefficients are defined as

$$\hat{A}_0 = \begin{bmatrix} A_{l-\hat{l}} & A_{l-\hat{l}-1} & \dots & A_1 & A_0 \\ -I_X & 0_X & \dots & \dots & 0_X \\ 0_X & -I_X & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_X & \dots & 0_X & -I_X & 0_X \end{bmatrix} \tag{3.3}$$

$$= \mathrm{Comp}[A_{l-\hat{l}}, A_{l-\hat{l}-1}, \dots, A_1, A_0], \tag{3.4}$$

$$\hat{A}_1 = \begin{bmatrix} A_{l-\hat{l}+1} & 0_{YX} & 0_{YX} & \dots & 0_{YX} \\ 0_X & I_X & 0_X & \dots & 0_X \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0_X \\ 0_X & \dots & \dots & 0_X & I_X \end{bmatrix} \tag{3.5}$$

$$= \mathrm{Diag}[A_{l-\hat{l}+1}, I_X, \dots, I_X] \tag{3.6}$$

and for $i = 2, \dots, \hat{l}$

$$\hat{A}_i = \begin{bmatrix} A_{l-\hat{l}+i} & 0_{YX} & \dots & 0_{YX} & 0_{YX} \\ 0_X & 0_X & \dots & 0_X & 0_X \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_X & 0_X & \dots & 0_X & 0_X \end{bmatrix} \tag{3.7}$$

$$= \mathrm{Diag}[A_{l-\hat{l}+i}, 0_X, \dots, 0_X], \tag{3.8}$$

where 0_X and I_X denote the zero and the identity operator in X, respectively, and 0_{YX} denotes the zero operator from X into Y.

The $(l-\hat{l}+1)\times(l-\hat{l}+1)$ block matrices $\hat{A}_i, i=0,1,\dots,\hat{l}$, are (the block matrix representiations of) linear operators from the product space $X^{l-\hat{l}+1}$ into the product space $Y\times X^{l-\hat{l}}$.
For $\hat{A}(\lambda)$ we have the following block matrix representation

$$\begin{bmatrix} A^{[\hat{l}]}(\lambda) & A_{l-\hat{l}-1} & A_{l-\hat{l}-2} & \dots & A_1 & A_0 \\ -I_X & \lambda I_X & 0_X & \dots & \dots & 0_X \\ 0_X & -I_X & \lambda I_X & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0_X \\ 0_X & \dots & \dots & 0_X & -I_X & \lambda I_X \end{bmatrix}, \tag{3.9}$$

where $A^{[\hat{l}]}(\cdot)$ is the $\hat{l}$-th Horner shift of $A(\cdot)$, i.e.

$$A^{[\hat{l}]}(\lambda)=\lambda^{\hat{l}}A_l+\dots+\lambda A_{l-\hat{l}+1}+A_{l-\hat{l}}.$$

Furthermore $\lambda\hat{A}_1+\hat{A}_0$ is the companion polynomial of the operator polynomial

$$\lambda^{l-\hat{l}+1}A_{l-\hat{l}+1}+\dots+A_0$$

and if $X=Y$ then $\hat{A}_0$ is the companion matrix of the monic operator polynomial

$$\lambda^{l-\hat{l}+1}I_X+\lambda^{l-\hat{l}}A_{l-\hat{l}}+\dots+A_0.$$

To simplify the notation we will from now on write 0 and I instead of 0_X and I_X, unless the domain and range spaces should not be clear from the context.
The following lemma shows that the term canonical degree reduction for $\hat{A}(\cdot)$ is justified.

Lemma 3.3. *The operator polynomial $\hat{A}(\cdot)$ is a degree reduction to degree $\hat{l}$ of $A(\cdot)$ as according to Definition 3.2. In particular $A(\cdot)$ and $\hat{A}(\cdot)$ are extended unimodular equivalent via*

$$E(\lambda)\hat{A}(\lambda)F(\lambda)=\operatorname{Diag}(A(\lambda),I_{l-\hat{l}})$$

where $I_{l-\hat{l}}$ is the identity operator on $X^{l-\hat{l}}$ and the unimodular operator polynomials

$$E(\lambda)=\begin{bmatrix} I & A^{[\hat{l}]}(\lambda) & A^{[\hat{l}+1]}(\lambda) & A^{[\hat{l}+2]}(\lambda) & \dots & A^{[l-2]}(\lambda) & A^{[l-1]}(\lambda) \\ & -I & -\lambda I & -\lambda^2 I & \dots & -\lambda^{l-\hat{l}-2}I & -\lambda^{l-\hat{l}-1}I \\ & & -I & -\lambda I & \dots & -\lambda^{l-\hat{l}-3}I & -\lambda^{l-\hat{l}-2}I \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & -I & -\lambda I & -\lambda^2 I \\ & & & & & -I & -\lambda I \\ & & & & & & -I \end{bmatrix}$$

and

$$F(\lambda) = \begin{bmatrix} \lambda^{l-\hat{l}} I & I & & & \\ \lambda^{l-\hat{l}-1} I & & I & & \\ \vdots & & & \ddots & \\ \lambda I & & & & I \\ I & 0 & \dots & \dots & 0 \end{bmatrix}$$

where $A^{[k]}(\cdot)$ is again the k-th Horner shift of $A(\cdot)$. Their inverses are

$$E^{-1}(\lambda) = \begin{bmatrix} I & A^{[\hat{l}]} & A_{l-\hat{l}-1} & \dots & A_1 \\ & -I & \lambda I & & \\ & & \ddots & \ddots & \\ & & & -I & \lambda I \\ & & & & -I \end{bmatrix}$$

and

$$F^{-1}(\lambda) = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 & I \\ I & & & & & -\lambda^{l-\hat{l}} I \\ & I & & & & -\lambda^{l-\hat{l}-1} I \\ & & \ddots & & & \vdots \\ & & & I & & -\lambda^2 I \\ & & & & I & -\lambda I \end{bmatrix}.$$

Proof. Straightforward calculation. □

For $\hat{l} = 1$, this is the *linearization* (or *companion form*) of $A(\cdot)$, see e.g. [Rod89]. Degree-reductions of this form were discussed in [Nag07] for matrix polynomials and later in [Har11] for polynomials with coefficients in a Banach algebra.

As an immediate consequence we can state a result about the finite spectrum of degree-reductions.

Corollary 3.4. *For $X = Y$ there holds*

$$\Sigma(A(\cdot)) = \Sigma(\hat{A}(\cdot)).$$

Proof. Since $E(\cdot)$ and $F(\cdot)$ are invertible for all $\lambda \in \mathbb{C}$ it follows that

$$\hat{A}(\lambda) \text{ invertible} \quad \Leftrightarrow \quad (A(\lambda) \oplus I_{X^{l-\hat{l}}}) \text{ invertible} \quad \Leftrightarrow \quad A(\lambda) \text{ invertible}. \qquad \square$$

Remark 3.5. In [TDM14] the authors distinguish between the *grade* and the *degree* of matrix polynomials. A matrix or operator polynomial

$$A(\lambda) = \sum_{j=0}^{l} \lambda^j A_j,$$

where the leading coefficients are allowed to be equal to zero, is then said to be of grade l; whereas the degree of $A(\lambda)$ corresponds to the highest nonzero coefficient as usual. This distinction has implications for the infinite eigenstructure of matrix polynomials.

3.0.1. Matrix polynomials and strong degree-reductions

For matrix polynomials $A(\lambda) = \sum_{j=0}^{l} \lambda^j A_j$ with coefficients $A_j \in \mathbb{C}^{n,n}$ the equality in Corollary 3.4 implies that the *finite elementary divisors* of $\hat{A}(\cdot)$ and $\mathrm{Diag}(A(\cdot), I_{(l-\hat{l})n})$ coincide (here $I_{(l-\hat{l})n}$ is the identity on $\mathbb{C}^{(l-\hat{l})n}$). Since $A(\cdot)$ need not be a monic polynomial it is of interest to also consider its *infinite elementary divisors*, i.e. the elementary divisors at zero of the reversed polynomial

$$\operatorname{rev} A(\lambda) := \sum_{j=0}^{l} \lambda^j A_{l-j} = \lambda^l A(\lambda^{-1}). \tag{3.10}$$

We state an important relation between degree reductions and reversed polynomials.

Lemma 3.6. *Let $A(\cdot)$ as above and $\hat{A}(\cdot)$ its canonical degree reduction to degree $\hat{l}$. Then the reversed polynomial* $\operatorname{rev} \hat{A}(\cdot)$ *is unimodular equivalent to* $\mathrm{Diag}(\operatorname{rev} A(\cdot), I_{(l-\hat{l})n})$. *In particular*

$$\tilde{E}(\lambda) \operatorname{rev} \hat{A}(\lambda) \tilde{F}(\lambda) = \mathrm{Diag}(\operatorname{rev} A(\cdot), I_{(l-\hat{l})n})$$

where

$$\tilde{E}(\lambda) = \begin{bmatrix} I & -\sum_{j=1}^{l-\hat{l}} \lambda^j A_{l-\hat{l}-j} & -\sum_{j=1}^{l-\hat{l}-1} \lambda^j A_{l-\hat{l}-1-j} & \dots & \dots & -(\lambda A_1 + \lambda^2 A_0) & -\lambda A_0 \\ 0 & \lambda^{l-\hat{l}-1} I & \lambda^{l-\hat{l}-2} I & \dots & \dots & \lambda I & I \\ 0 & \lambda^{l-\hat{l}-2} I & \lambda^{l-\hat{l}-3} I & & \iddots & \iddots & \\ \vdots & \vdots & \vdots & \iddots & \iddots & & \\ 0 & \lambda^2 I & \lambda I & I & & & \\ 0 & \lambda I & I & & & & \\ 0 & I & & & & & \end{bmatrix}$$

and

$$\tilde{F}(\lambda) = \begin{bmatrix} I & 0 & \dots & \dots & 0 \\ \lambda I & & & & I \\ \vdots & & & \iddots & \\ \lambda^{l-\hat{l}-1} I & & I & & \\ \lambda^{l-\hat{l}} I & I & & & \end{bmatrix}$$

are unimodular (block) matrix polynomials.

Proof. Straightforward calculation. $\square$

The lemma implies that the elementary divisors at 0 of $\operatorname{rev} \hat{A}(\cdot)$ are equal to the elementary divisors at 0 of $\operatorname{rev} A(\cdot)$ plus $(l-\hat{l})n$ copies of $\lambda^{\hat{l}-1}$. This motivates the next definition.

Definition 3.7. Let $A(\cdot)$ be a matrix polynomial of degree l. Then we call a matrix polynomial $\hat{A}(\cdot)$ of degree $\hat{l}$ a *strong* degree reduction (to degree $\hat{l}$) if $A(\cdot)$ is extended unimodular equivalent to $\hat{A}(\cdot)$ and the reversed polynomial $\operatorname{rev} A(\cdot)$ is extended unimodular equivalent to $\operatorname{rev} \hat{A}(\cdot)$.

This definition is consistent with the classical notion of strong linearisations and strong ℓ-ifications (see [TDM14]). It follows immediately that a canonical degree reduction to degree $\hat{l}$ is strong if and only if $\hat{l} = l$ or $\hat{l} = 1$. In [TDM14] the authors consider other forms of degree reductions which are strong for different values of $\hat{l}$. In [TDD15] they proved that a strong degree reduction exists for any $\hat{l}$ (even for $\hat{l} > l$), but no explicit formulas are given.

3.1. Division with remainder

For this section we assume that $A(\cdot)$ has full degree l. We will show how a division with remainder of the canonical degree reduction $\hat{A}(\cdot)$ can be recovered from a division with remainder of the original operator polynomial $A(\cdot)$ (and vice versa). This was previously investigated in [Har11] for the special case where division is possible without remainder.

Theorem 3.8. *Let $1 < \hat{l} < l$ and*

$$A(\lambda) = L(\lambda)D(\lambda) + R(\lambda), \tag{3.11}$$

where

$$L(\lambda) = \lambda^{\hat{l}-1}L_{\hat{l}-1} + \cdots + \lambda L_1 + L_0, \tag{3.12}$$

$$D(\lambda) = \lambda^{l-\hat{l}+1} + \lambda^{l-\hat{l}}D_{l-\hat{l}} + \cdots + \lambda D_1 + D_0, \tag{3.13}$$

$$R(\lambda) = \lambda^{l-\hat{l}}R_{l-\hat{l}} + \cdots + R_0 \tag{3.14}$$

with $L_j : X \to Y$, $D_j : X \to X$ and $R_j : X \to Y$.
Further let

$$\hat{A}(\lambda) = \tilde{L}(\lambda)(\lambda + \hat{D}_0) + \tilde{R}_0, \tag{3.15}$$

where $\hat{A}(\cdot)$ is the canonical degree reduction of $A(\cdot)$ to degree $\hat{l}$, $\hat{D}_0$ is the companion matrix of $D(\cdot)$, $\tilde{R}_0$ is independent of λ and

$$\tilde{L}(\lambda) = \lambda^{\hat{l}-1}\hat{L}_{\hat{l}-1} + \cdots + \lambda\hat{L}_1 + \hat{L}_0 \tag{3.16}$$

(note that, since the right divisors are monic, the coefficients of the operator polynomials $L(\cdot), \tilde{L}(\cdot), R(\cdot)$ and the matrix $\tilde{R}_0$ are uniquely determined by the principle of division with remainder; see [Rod89, Theorem 2.6.2]).
Then

$$\tilde{R}_0 = \begin{bmatrix} R_{l-\hat{l}} & R_{l-\hat{l}-1} & \dots & R_1 & R_0 \\ 0_X & 0_X & \dots & 0_X & 0_X \\ 0_X & 0_X & \dots & 0_X & 0_X \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_X & 0_X & \dots & 0_X & 0_X \end{bmatrix}, \tag{3.17}$$

for $j = 1, 2, \dots, \hat{l} - 1$

$$\tilde{L}_j = \begin{bmatrix} L_{j,l-\hat{l}} & L_{j,l-\hat{l}-1} & \dots & L_{j,1} & L_{j,0} \\ 0_X & 0_X & \dots & 0_X & 0_X \\ 0_X & 0_X & \dots & 0_X & 0_X \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_X & 0_X & \dots & 0_X & 0_X \end{bmatrix} \tag{3.18}$$

and

$$\tilde{L}_0 = \begin{bmatrix} L_{0,l-\hat{l}} & L_{0,l-\hat{l}-1} & L_{0,l-\hat{l}-2} & \dots & L_{0,0} \\ 0_X & I_X & 0_X & \dots & 0_X \\ 0_X & 0_X & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0_X \\ 0_X & 0_X & \dots & 0_X & I_X \end{bmatrix} \tag{3.19}$$

where

$$L_{j,l-\hat{l}} = L_j \tag{3.20}$$

and for $s = 0, 1, \dots, l - \hat{l} - 1$

$$L_{j,s} = -\sum_{i=-\infty}^{\infty} L_{j+i} D_{s-i}, \tag{3.21}$$

where we extend the index set for L_j *and* D_j *by setting*

$$L_j = \begin{cases} L_j & ,0 \le j \le \hat{l} - 1 \\ 0 & , otherwise \end{cases}, \quad D_j = \begin{cases} D_j & ,0 \le j \le l - \hat{l} \\ I & ,j = l - \hat{l} + 1 \\ 0 & , otherwise \end{cases}.$$

Proof. First note that

$$\begin{aligned} L_{0,0} D_0 + R_0 &= A_0, \\ L_{0,s} D_s - L_{0,s-1} + R_s &= A_s, \qquad\qquad 1 \le s \le l - \hat{l}, \end{aligned}$$

for $j = 1, \dots, \hat{l} - 1$

$$\begin{aligned}
L_{j,0} D_0 + L_{j-1,0} &= 0, \\
L_{j,l-\hat{l}} D_s - L_{j,s-1} + L_{j-1,s} &= 0, \qquad\qquad 1 \le s \le l - \hat{l} - 1, \\
L_{j,l-\hat{l}} D_{l-\hat{l}} - L_{j,l-\hat{l}-1} + L_{j-1,l-\hat{l}} &= A_{l-\hat{l}+j}
\end{aligned}$$

and

$$\begin{aligned}
L_{\hat{l}-1,l-\hat{l}} &= A_l, \\
L_{\hat{l}-1,s} &= 0, \qquad\qquad 0 \le s \le l - \hat{l} - 1.
\end{aligned}$$

It is then straight forward to check that

$$\begin{aligned}
\hat{A}_0 &= \tilde{L}_0 \hat{D}_0 + \tilde{R}_0, \\
\hat{A}_j &= \tilde{L}_j \hat{D}_0 + \tilde{L}_{j-1}, \qquad\qquad 1 \le j \le \hat{l} - 1, \\
\hat{A}_{\hat{l}} &= \tilde{L}_{\hat{l}-1}
\end{aligned}$$

which proves the assertion. □

A question that naturally arises in the context of divisions with remainder is how, given an operator polynomial $A(\cdot)$ of degree l and monic operator polynomial $D(\cdot)$ as in (3.13), one computes the unique operator polynomials $L(\cdot)$ and $R(\cdot)$ as in (3.12) and (3.14) such that

$$A(\lambda) = L(\lambda)D(\lambda) + R(\lambda).$$

In [Rod89, Theorem 2.6.2] L. Rodman answers this question using spectral triples. Theorem 3.8 gives rise to an alternative approach.
First consider the case when the divisor $D(\cdot)$ is linear, i.e.

$$L(\lambda) = \sum_{j=0}^{l-1} \lambda^j L_j, \qquad D(\lambda) = \lambda + D_0, \qquad R(\lambda) \equiv R_0.$$

Then the coefficients L_j are given by

$$L_j = \sum_{i=j+1}^{l} A_i (-D_0)^{i-j-1} \qquad , j = 0, \dots, l-1 \tag{3.22}$$

and R_0 is given by

$$R_0 = \sum_{i=0}^{l} A_i (-D_0)^i. \tag{3.23}$$

Theorem 3.8 then tells us that, given $A(\cdot)$ and a monic but not necessarily linear $D(\cdot)$, we can first compute the degree reduction (respectively linearisation) $\hat{A}(\cdot)$ and $\hat{D}_0$. Then use formulas (3.22) and (3.23) to compute $\tilde{L}(\cdot)$ and $\tilde{R}_0$. Since division with remainder is

unique we can now use the representations in (3.17) - (3.19) to obtain the coefficients of $L(\cdot)$ and $R(\cdot)$ (note that the coefficients of $L(\cdot)$ and $R(\cdot)$ explicitly appear as entries in the block matrices).
We illustrate the procedure in a diagram:

$$\begin{tikzcd}
A(\cdot), D(\cdot) \arrow[r, "\text{degree reduction}"] \arrow[d, "\text{[Rod89, Theorem 2.6.2]}"'] & \hat{A}(\cdot), \hat{D}_0 \arrow[d, "\text{(3.22) and (3.23)}"] \\
L(\cdot), R(\cdot) & \tilde{L}(\cdot), \tilde{R}_0 \arrow[l, "\text{Theorem 3.8}"']
\end{tikzcd}$$

The reverse direction in Theorem 3.8 is in general not true if we divide $\hat{A}(\cdot)$ by some arbitrary linear factor $(\lambda I + B)$ (i.e. B is not the companion matrix of a polynomial), because the coefficients of the quotient polynomial, divisor and rest will be of arbitrary structure.
This changes however if we consider only divisions without remainder.

Proposition 3.9. *Let $A(\cdot)$ and $\hat{A}(\cdot)$ be as in Theorem 3.8 and assume $\hat{A}(\cdot)$ has a factorization*

$$\hat{A}(\lambda) = \tilde{L}(\lambda)(\lambda I + \hat{D}_0)$$

for some operator polynomial

$$\tilde{L}(\lambda) = \sum_{j=0}^{\hat{l}-1} \lambda^j \tilde{L}_j$$

and block operator matrix $\hat{D}_0$. Then $\hat{D}_0$ is the companion (operator) matrix of a monic operator polynomial

$$D(\lambda) = \lambda^{l-\hat{l}+1} + \sum_{j=0}^{l-\hat{l}} \lambda^j D_j,$$

and the operator polynomial $A(\cdot)$ admits a factorization

$$A(\lambda) = \left(\sum_{j=0}^{\hat{l}-1} \lambda^j L_j \right) D(\lambda)$$

where L_j is the upper left entry in $\tilde{L}_j$.

Proof. This is [Har11, Theorem3.4.(ii)] together with [Har11, Remark 3.5.(ii)]. □

3.2. Spectral triples

In this section we consider only monic operator polynomials $A(\cdot)$, so that

$$A(\lambda) = \lambda^l I_X + \lambda^{l-1} A_{l-1} + \cdots + \lambda A_1 + A_0, \tag{3.24}$$

with $A_j : X \to X$ for $j = 0, \dots, l-1$.
Recall the definition of right spectral pairs and spectral triples: Let Z be a (real or complex) vector space. Then a pair (V, T), where $V : Z \to X$ and $T : Z \to Z$ is called a *right spectral pair* for $A(\cdot)$ if the operator

$$\begin{bmatrix} V \\ VT \\ \vdots \\ VT^{l-2} \\ VT^{l-1} \end{bmatrix} : Z \to X^l \tag{3.25}$$

is invertible and

$$A_0 V + A_1 VT + \cdots + A_{l-1} VT^{l-1} + VT^l = 0. \tag{3.26}$$

A triple (V, T, W) where $V : Z \to X$, $T : Z \to Z$ and $W : X \to Z$ is called a *spectral triple* for $A(\cdot)$ if (V, T) is a right spectral pair for $A(\cdot)$ and

$$\begin{bmatrix} V \\ VT \\ \vdots \\ VT^{l-1} \end{bmatrix} W = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \end{bmatrix}. \tag{3.27}$$

The next theorem, which gives a spectral triple with respect to a division of $A(\cdot)$, is motivated by an abstract representation result in [LŠ10, Equation (4)].

Theorem 3.10. *Let*

$$A(\lambda) = L(\lambda) D(\lambda) + R(\lambda), \tag{3.28}$$

where $D(\cdot)$ is monic with degree d (then necessarily $L(\cdot)$ is monic with degree $l-d$ and $\deg R(\cdot) = d-1$). Then the triplet of operators $[V_{l,d}, -T_{l,d}, W_{l,d}]$ is a spectral triple for $A(\cdot)$, where

$$V_{l,d} = \begin{bmatrix} 0 & \dots & 0 & I & 0 & \dots & 0 \end{bmatrix} \tag{3.29}$$

with I in the d-th entry,

$$T_{l,d} = \left[\begin{array}{cccc|cccc} D_{d-1} & \dots & D_1 & D_0 & 0 & \dots & 0 & -I \\ -I & & & 0 & 0 & \dots & 0 & 0 \\ & \ddots & & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & -I & 0 & 0 & \dots & 0 & 0 \\ \hline R_{d-1} & \dots & R_1 & R_0 & L_{l-d-1} & \dots & L_1 & L_0 \\ 0 & \dots & 0 & 0 & -I & & & 0 \\ \vdots & & \vdots & \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & 0 & & & -I & 0 \end{array}\right] \tag{3.30}$$

and

$$W_{l,d} = \begin{bmatrix} 0 & \dots & 0 & I & 0 & \dots & 0 \end{bmatrix}^T \tag{3.31}$$

with I in the $(d+1)$-th entry. In particular we have

$$V = V_{l,d} S_{l,d}, \qquad S_{l,d}\,\mathrm{Comp}[A_{l-1}, \dots, A_0] = T_{l,d} S_{l,d}, \qquad S_{l,d} W = W_{l,d} \tag{3.32}$$

where $V = \begin{bmatrix} 0 & \dots & 0 & I \end{bmatrix}$, $W = \begin{bmatrix} I & 0 & \dots & 0 \end{bmatrix}^T$ and $S_{l,d}$ is the invertible operator

$$S_{l,d} = \left[\begin{array}{cccccccccc} & & & & I & & & & & \\ & & & & & I & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & & I & \\ \hline I & D_{d-1} & D_{d-2} & \dots & D_0 & & & & & \\ & I & D_{d-1} & D_{d-2} & \dots & D_0 & & & & \\ & & \ddots & \ddots & \ddots & & \ddots & & & \\ & & & & I & D_{d-1} & D_{d-2} & \dots & D_0 & \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c} I \\ I \\ \ddots \\ I \end{array}}\right\} d \\ \left.\vphantom{\begin{array}{c} I \\ I \\ \ddots \\ I \end{array}}\right\} l-d \end{array} . \tag{3.33}$$

Proof. By [Rod89, Theorem 2.1.1] and its proof together with [Rod89, Proposition 2.5.1] it follows that $(V, -\mathrm{Comp}[A_{l-1}, \dots, A_0], W)$ is a spectral triple for $A(\cdot)$ and thus it is sufficient to show that the equations in (3.32) hold.
The first and last equations are obvious. For the second equation define

$$L_j = \begin{cases} L_j & ,0 \le j \le l-d-1 \\ I & ,j = l-d \\ 0 & ,otherwise \end{cases}, \quad D_j = \begin{cases} D_j & ,0 \le j \le d-1 \\ I & ,j = d \\ 0 & ,otherwise \end{cases}$$

$$R_j = \begin{cases} R_j & ,0 \le j \le d-1 \\ 0 & ,otherwise \end{cases}$$

for $j \in \mathbb{Z}$, so that we can write

$$A_j = \sum_{k=-\infty}^{\infty} L_{j-k} D_k + R_j, \quad 0 \le j \le l-1. \tag{3.34}$$

Keeping the above in mind one calculates that indeed

$$
\begin{aligned}
&S_{l,d}\,\mathrm{Comp}[A_{l-1},\dots,A_0] = \\
&\begin{bmatrix}
 & & & -I & & & & 0 \\
 & & & & -I & & & \vdots \\
 & & & & & \ddots & & \\
 & & & & & & -I & 0 \\
A_{l-1}-D_{d-1} & A_{l-2}-D_{d-2} & \dots & A_{l-d}-D_0 & A_{l-d-1} & \dots & A_1 & A_0 \\
-I & -D_{d-1} & \dots & D_1 & D_0 & & & 0 \\
 & \ddots & \ddots & & \ddots & \ddots & & \vdots \\
 & & -I & -D_{d-1} & \dots & D_1 & D_0 & 0
\end{bmatrix} \\
&= T_{l,d} S_{l,d}. \qquad \square
\end{aligned}
$$

3.3. Consecutive degree-reduction

One question that naturally arises in this context is the one of consecutive degree-reductions and their relation with the corresponding one-step degree-reduction.
To make this clear let $A(\cdot)$ be as in (3.1) and let $0 < l_2 < l_1 < l$. Then we can reduce the degree of $A(\cdot)$ to l_2 to obtain $A^{(l_2)}(\cdot)$. However we can also first reduce the degree of $A(\cdot)$ down to l_1 to obtain $A^{(l_1)}(\cdot)$ and then further decrease the degree to l_2 to obtain $(A^{(l_1)})^{(l_2)}(\cdot)$. We will denote this two-step degree-reduction by $A^{(l_1,l_2)}(\cdot)$.
It is clear that, despite having the same degree, there holds $A^{(l_2)}(\cdot) \neq A^{(l_1,l_2)}(\cdot)$ (since they do not even have the same block dimension). However we will see that $A^{(l_1,l_2)}(\cdot)$ is equivalent to the direct sum of $A^{(l_2)}(\cdot)$ and an operator polynomial with eigenvalues only at infinity.
The upcoming result is inspired by a procedure outlined in [MX13] for consecutive degree-reductions of degree 1 of matrix polynomials. We will first be looking at an example before moving to the general result.

Example 3.11. Consider the operator polynomial of degree $l = 7$, i.e.

$$A(\lambda) = \lambda^7 A_7 + \cdots \lambda A_1 + A_0 \tag{3.35}$$

and consecutive degree-reductions to $l_1 = 5$ and $l_2 = 2$. Then $A^{(5,2)}(\cdot)$ has the block matrix representation

$$A^{(5,2)} = \left[\begin{array}{ccc|ccc|ccc|ccc}
A^{[2]}(\lambda) & & & A_4 & & & A_3 & & & A_2 & A_1 & A_0 \\
 & & & & & & & I & & -I & & \\
 & & & & & & & & I & & -I & \\
\hline
-I & & & \lambda & & & & & & & & \\
 & -I & & & \lambda & & & & & & & \\
 & & -I & & & \lambda & & & & & & \\
\hline
 & & & -I & & & \lambda & & & & & \\
 & & & & -I & & & \lambda & & & & \\
 & & & & & -I & & & \lambda & & & \\
\hline
 & & & & & & -I & & & \lambda & & \\
 & & & & & & & -I & & & \lambda & \\
 & & & & & & & & -I & & & \lambda
\end{array}\right]. \tag{3.36}$$

By applying the permutation that moves the block rows and columns with indices $(2, 3, 5, 6, 8, 9)$ to the end we arrive at

$$A^{(5,2)} \cong \left[\begin{array}{cccccc|cccccc}
A^{[2]}(\lambda) & A_4 & A_3 & A_2 & A_1 & A_0 & & & & & & \\
-I & \lambda & & & & & & & & & & \\
 & -I & \lambda & & & & & & & & & \\
 & & -I & \lambda & & & & & & & & \\
 & & & & \lambda & & & & & & -I & \\
 & & & & & \lambda & & & & & & -I \\
\hline
 & & & -I & & & & & & & I & \\
 & & & & -I & & & & & & & I \\
 & & & & & & -I & & \lambda & & & \\
 & & & & & & & -I & & \lambda & & \\
 & & & & & & & & -I & & \lambda & \\
 & & & & & & & & & -I & & \lambda
\end{array}\right]. \tag{3.37}$$

Note how the lengths of the diagonal parts correspond to l_1 and l_2. We can now add block rows 7 and 8 to rows 5 and 6 respectively to obtain

$$A^{(5,2)} \cong \left[\begin{array}{cccccc|cccccc}
A^{[2]}(\lambda) & A_4 & A_3 & A_2 & A_1 & A_0 & & & & & & \\
-I & \lambda & & & & & & & & & & \\
 & -I & \lambda & & & & & & & & & \\
 & & -I & \lambda & & & & & & & & \\
 & & & -I & \lambda & & & & & & & \\
 & & & & -I & \lambda & & & & & & \\
\hline
 & & & -I & & & & & & & I & \\
 & & & & -I & & & & & & & I \\
 & & & & & & -I & & \lambda & & & \\
 & & & & & & & -I & & \lambda & & \\
 & & & & & & & & -I & & \lambda & \\
 & & & & & & & & & -I & & \lambda
\end{array}\right]. \tag{3.38}$$

The only thing left to do is removing the elements in the lower left quadrant which can be accomplished by a succession of elementary column and row additions: The first step is to add the columns 11 and 12 to the columns 5 and 6 respectively.

$$A^{(5,2)} \cong \left[\begin{array}{cccccc|cccccc}
A^{[2]}(\lambda) & A_4 & A_3 & A_2 & A_1 & A_0 & & & & & & \\
-I & \lambda & & & & & & & & & & \\
 & -I & \lambda & & & & & & & & & \\
 & & -I & \lambda & & & & & & & & \\
 & & & -I & \lambda & & & & & & & \\
 & & & & -I & \lambda & & & & & & \\
\hline
 & & & & & & & & & & I & \\
 & & & & & & & & & & & I \\
 & & & & & & -I & & \lambda & & & \\
 & & & & & & & -I & & \lambda & & \\
 & & & \lambda & & & & & -I & & \lambda & \\
 & & & & \lambda & & & & & -I & & \lambda
\end{array}\right]. \tag{3.39}$$

Now subtract rows 4 and 5 from rows 11 and 12 respectively.

$$A^{(5,2)} \cong \left[\begin{array}{cccccc|cccccc}
A^{[2]}(\lambda) & A_4 & A_3 & A_2 & A_1 & A_0 & & & & & & \\
-I & \lambda & & & & & & & & & & \\
 & -I & \lambda & & & & & & & & & \\
 & & -I & \lambda & & & & & & & & \\
 & & & -I & \lambda & & & & & & & \\
 & & & & -I & \lambda & & & & & & \\
\hline
 & & & & & & & & & & I & \\
 & & & & & & & & & & & I \\
 & & & & & & -I & & \lambda & & & \\
 & & & & & & & -I & & \lambda & & \\
 & & I & & & & & & -I & & \lambda & \\
 & & & I & & & & & & -I & & \lambda
\end{array}\right]. \tag{3.40}$$

At first it seems that not much was gained but we can repeat the procedure to move the block of identities up and left.

$$A^{(5,2)} \cong \left[\begin{array}{cccccc|cccccc}
A^{[2]}(\lambda) & A_4 & A_3 & A_2 & A_1 & A_0 & & & & & & \\
-I & \lambda & & & & & & & & & & \\
 & -I & \lambda & & & & & & & & & \\
 & & -I & \lambda & & & & & & & & \\
 & & & -I & \lambda & & & & & & & \\
 & & & & -I & \lambda & & & & & & \\
\hline
 & & & & & & & & & & I & \\
 & & & & & & & & & & & I \\
 & I & & & & & -I & & \lambda & & & \\
 & & I & & & & & -I & & \lambda & & \\
 & & & & & & & & -I & & \lambda & \\
 & & & & & & & & & -I & & \lambda
\end{array}\right]. \tag{3.41}$$

Now the block of identities can be eliminated by adding columns 7 and 8. Note that the size of the block of identities is related to the size of the first degree-reduction, while the number of times the block has to be moved is related to the size of the second degree reduction. We finally arrive at

$$A^{(5,2)} \cong \left[\begin{array}{cccccc|cccccc}
A^{[2]}(\lambda) & A_4 & A_3 & A_2 & A_1 & A_0 & & & & & & \\
-I & \lambda & & & & & & & & & & \\
 & -I & \lambda & & & & & & & & & \\
 & & -I & \lambda & & & & & & & & \\
 & & & -I & \lambda & & & & & & & \\
 & & & & -I & \lambda & & & & & & \\
\hline
 & & & & & & & & & & I & \\
 & & & & & & & & & & & I \\
 & & & & & & -I & & \lambda & & & \\
 & & & & & & & -I & & \lambda & & \\
 & & & & & & & & -I & & \lambda & \\
 & & & & & & & & & -I & & \lambda
\end{array}\right]. \tag{3.42}$$

The upper left block in (3.42) is the block matrix representation of $A^{(2)}(\lambda)$. The lower right part is an operator polynomial that has eigenvalues only at infinity.

From the procedure outlined above it is clear that these steps translate to general degree-reductions to degrees l_1 and l_2.

Theorem 3.12. *Let $A(\lambda) = \lambda^l A_l + \cdots + A_1 + A_0$ and $0 < l_2 < l_1 < l$. Then the block matrix representation of the consecutive degree-reduction $A^{(l_1,l_2)}(\cdot)$ is equivalent to*

$$\begin{bmatrix} C_{A^{(l_2)}}(\lambda) & 0 \\ 0 & K(\lambda) \end{bmatrix}, \tag{3.43}$$

where $C_{A^{(l_2)}}(\cdot)$ denotes the block matrix representation of the one-step degree-reduction $A^{(l_2)}(\cdot)$ and

$$K(\lambda) = \left[\begin{array}{ccccccc}
0 & & & & I & & \\
 & \ddots & & & & \ddots & \\
 & & 0 & & & & I \\
\hline
-I & & & \lambda & & & \\
 & \ddots & & & \ddots & & \\
 & & \ddots & & & \ddots & \\
 & & & -I & & & \lambda
\end{array}\right]
\begin{array}{l}
\left.\begin{array}{c} \\ \\ \\ \end{array}\right\} l - l_1 \\
\left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} (l_1 - l_2 - 1)(l - l_1)
\end{array}. \tag{3.44}$$

3.4. Degree-reduction and the block numerical range

In this section we will examine how the block numerical range of operator polynomials and their degree-reductions are related, extending results in [TW03] and [GLW09]. For

definitions and basic properties of the block numerical range see Appendix A and for the block numerical range of functions see Section 1.1.
A helpful tool in studying the invertability of operator matrices is the Schur complement:

Definition 3.13. Let Y_1, Y_2 be Banach spaces and $B_{ij} \in \mathcal{L}(Y_j, Y_i)$ for $i, j \in \{1, 2\}$. On $Y_1 \times Y_2$ define the operator B through

$$B := \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \in \mathcal{L}(Y_1 \times Y_2).$$

For B_{22} invertible define the *Schur complement S of B_{22} in B* as

$$S := B_{11} - B_{12}(B_{22})^{-1}B_{21}.$$

Proposition 3.14. *Let B as in Definition 3.13 with B_{22} invertible. Then*

(i) B is invertible $\quad\Leftrightarrow\quad S = B_{11} - B_{12}(B_{22})^{-1}B_{21}$ is invertible,

(ii) for B invertible the equality

$$B^{-1} = \begin{bmatrix} S^{-1} & -S^{-1}B_{12}B_{22}^{-1} \\ -B_{22}^{-1}B_{21}S^{-1} & B_{22}^{-1} + B_{22}^{-1}B_{21}S^{-1}B_{12}B_{22}^{-1} \end{bmatrix}$$

holds.

Proof. See [GGK93, p. 514]. □

Results that we are about to state for degree reduced polynomials are also true in a more general context:

Definition 3.15. For Banach spaces Y_1, Y_2 and $0 < \hat{l} < l$ define the operator polynomial

$$P_{(\hat{l})} : \mathbb{C} \to \mathcal{L}(Y_1 \times Y_2) : \lambda \mapsto \begin{bmatrix} C(\lambda) & C_{12} \\ C_{21} & \lambda - N \end{bmatrix}$$

where $C : \mathbb{C} \to \mathcal{L}(Y_1)$ is an operator polynomial of degree $\hat{l}$, $C_{ij} \in \mathcal{L}(Y_j, Y_i)$ for $i, j \in \{1, 2\}, i \neq j$ and $N \in \mathcal{L}(Y_2)$ nilpotent with $N^{l-\hat{l}} = 0$.

Lemma 3.16. *For $\lambda \neq 0$ the operator $\lambda - N$ is invertible. The Schur complement S_λ of $\lambda - N$ in $P_{(\hat{l})}(\lambda)$ is*

$$S_\lambda = \lambda^{-(l-\hat{l})} \left(\lambda^{l-\hat{l}} C(\lambda) - \sum_{j=0}^{l-\hat{l}-1} \lambda^j C_{12} N^{l-\hat{l}-j-1} C_{21} \right) =: \lambda^{-(l-\hat{l})} P(\lambda),$$

with an operator polynomial $P : \mathbb{C} \to \mathcal{L}(Y_1)$ of degree l.

Proof. First note that the series $\sum_{j=0}^{\infty} N^j = \sum_{j=0}^{l-\hat{l}-1} N^j$ converges. For $\lambda \neq 0$ it then follows with the Neumann series that

$$(\lambda - N)^{-1} = \sum_{j=0}^{l-\hat{l}-1} \lambda^{-(j+1)} N^j \in \mathcal{L}(Y_2).$$

For S_λ we then have

$$\begin{aligned}
S_\lambda &= C(\lambda) - C_{12}(\lambda - N)^{-1} C_{21} \\
&= C(\lambda) - \sum_{j=0}^{l-\hat{l}-1} \lambda^{-(j+1)} C_{12} N^j C_{21} \\
&= \lambda^{-(l-\hat{l})} \left(\lambda^{l-\hat{l}} C(\lambda) - \sum_{j=0}^{l-\hat{l}-1} \lambda^{l-\hat{l}-j-1} C_{12} N^j C_{21} \right) \\
&= \lambda^{-(l-\hat{l})} \left(\lambda^{l-\hat{l}} C(\lambda) - \sum_{j=0}^{l-\hat{l}-1} \lambda^{j} C_{12} N^{l-\hat{l}-j-1} C_{21} \right).
\end{aligned}$$

□

Remark 3.17. For an operator polynomial $A(\cdot)$ of degree l as in (3.1) we can identify its canonical degree $\hat{l}$ reduction $\hat{A}(\cdot)$ with $P_{(\hat{l})}(\cdot)$ via

$$\left[\begin{array}{c|cccc}
\sum_{j=0}^{\hat{l}} \lambda^j A_{l-\hat{l}+j} & A_{l-\hat{l}-1} & \dots & \dots & A_0 \\
\hline
-I & \lambda & & & \\
 & -I & \ddots & & \\
 & & \ddots & \ddots & \\
 & & & -I & \lambda
\end{array}\right] \mathrel{\widehat{=}} \left[\begin{array}{c|c} C(\lambda) & C_{12} \\ \hline C_{21} & \lambda - N \end{array}\right], \quad \lambda \in \mathbb{C}.$$

Through straightforward multiplication it then follows that $P(\cdot) = A(\cdot)$.

With the exception of the point $\lambda = 0$ a spectral equivalence as in Corollary 3.4 holds in this more general setting.

Corollary 3.18. *For $P_{(\hat{l})}(\cdot), P(\cdot)$ as in Lemma 3.16 and $\lambda \neq 0$ we get*

$$\lambda \in \Sigma(P_{(\hat{l})}(\cdot)) \quad \iff \quad \lambda \in \Sigma(P(\cdot)).$$

Proof. From Lemma 3.16 it immediately follows that for $\lambda \neq 0$ the operator $P_{(\hat{l})}(\lambda)$ is invertible if and only if its Schur complement $\lambda^{-(l-\hat{l})} P(\lambda)$ is invertible. □

To state a similar result for the block numerical range we need the Banach space Y_2 to be of finite dimension n so that we can decompose it into 1-dimensional spaces. We then consider a decomposition

$$Y = Y_1 \times Y_2 = (Y_{11} \times \ldots \times Y_{1d}) \times (Y_{21} \times \ldots \times Y_{2n})$$

where the Y_{2j} are 1-dimensional for $j \in \langle n \rangle$.

Theorem 3.19. *For $P_{(\hat{l})}(\cdot), P(\cdot)$ as in Lemma 3.16 and $\lambda \neq 0$ there holds*

$$\lambda \in \Theta(P(\cdot); Y_{11}, \ldots, Y_{1d}) \quad \Longleftrightarrow \quad \lambda \in \Theta(P_{(\hat{l})}; Y_{11}, \ldots, Y_{1d}, Y_{21}, \ldots, Y_{2n}).$$

Proof. For this whole proof let $\lambda \neq 0$. Now $\lambda \notin \Theta(P_{(\hat{l})}(\cdot); Y_{11}, \ldots, Y_{1d}, Y_{21}, \ldots, Y_{2n})$ if and only if for all $(F, U) \in S^{op}_{\text{att}}(Y_{11}, \ldots, Y_{1d})$ and $(G, V) \in S^{op}_{\text{att}}(Y_{21}, \ldots, Y_{2n})$ the operator

$$\begin{aligned}
&\begin{bmatrix} F & 0 \\ 0 & G \end{bmatrix} P_{(\hat{l})}(\lambda) \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \\
&= \begin{bmatrix} F & 0 \\ 0 & G \end{bmatrix} \begin{bmatrix} C(\lambda) & C_{12} \\ C_{21} & I\lambda - N \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \\
&= \begin{bmatrix} FC(\lambda)U & FC_{12}V \\ GC_{21}U & G(\lambda - N)V \end{bmatrix}
\end{aligned}$$

is invertible. Note that the diagonal entries of the operators

$$\begin{bmatrix} F & 0 \\ 0 & G \end{bmatrix} : Y_1 \times Y_2 \to \mathbb{C}^d \times \mathbb{C}^n, \qquad \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} : \mathbb{C}^d \times \mathbb{C}^n \to Y_1 \times Y_2$$

are mapping between different spaces. The zero operators are also to be interpreted between the corresponding spaces. However due to the 1-dimensionality imposed on the spaces Y_{2j} the operators G and V are invertible. Therefore $G(\lambda - N)V$ is also invertible and we can apply the Schur complement.

Now $\lambda \notin \Theta(P_{(\hat{l})}(\cdot); Y_{11}, \ldots, Y_{1d}, Y_{21}, \ldots, Y_{2n})$ if and only if for all $(F, U) \in S^{op}_{\text{att}}(Y_{11}, \ldots, Y_{1d})$ and $(G, V) \in S^{op}_{\text{att}}(Y_{21}, \ldots, Y_{2n})$ the operator

$$\begin{aligned}
&FC(\lambda)U - FC_{12}V(G(\lambda - N)V)^{-1}GC_{21}U \\
&= FC(\lambda)U - FC_{12}VV^{-1}(\lambda - N)^{-1}G^{-1}GC_{21}U \\
&= FC(\lambda)U - FC_{12}(\lambda - N)^{-1}C_{21}U \\
&= F(C(\lambda) - C_{12}(\lambda - N)^{-1}C_{21})U \\
&= F(\lambda^{-(l-\hat{l})}P(\lambda))U
\end{aligned}$$

is invertible. And this is the case if and only if $\lambda \notin \Theta(P(\cdot); Y_{11}, \ldots, Y_{1d})$. □

Back to our original operator polynomial $A(\cdot)$ as in (3.1) and its canonical degree reduction $\hat{A}(\cdot)$ to $\hat{l}$ this theorem translates as follows.

Theorem 3.20. *Let the operator polynomial $A(\cdot)$ be defined on a finite dimensional Banach space X with dimension n. Let $X_1, \ldots, X_d$ be an arbitrary (not necessarily 1-dimensional) decomposition of X and $\tilde{X}_1, \ldots, \tilde{X}_{(l-\hat{l})n}$ be a 1-dimensional decomposition of $X^{l-\hat{l}}$. Then for $\lambda \neq 0$ there holds*

$$\lambda \in \Theta(A(\cdot); X_1, \ldots, X_d) \quad \Longleftrightarrow \quad \lambda \in \Theta(\hat{A}(\cdot); X_1, \ldots, X_d, \tilde{X}_1, \ldots, \tilde{X}_{(l-\hat{l})n}).$$

Proof. This follows immediately from Remark 3.17 and Theorem 3.19. $\square$

Theorem 3.20 is a generalisation of [TW03, Theorem 5.2], which was posed in the Hilbert space setting and for linearisations (via the companion matrix) instead of degree reductions.
If we omit the finite dimensionality on X we can at least formulate an inclusion result extending [TW03, Theorem 5.1] and [Wag07, Proposition 3.15]. To make notation a little less cluttered we abbreviate the decomposition

$$X^{(l-\hat{l}+1)} := \underbrace{X_1, \dots, X_d, \dots, X_1, \dots, X_d}_{(l-\hat{l}+1)\text{-times}}.$$

Proposition 3.21. *Let $A(\cdot)$ and $\hat{A}(\cdot)$ be as above. Then*

$$\Theta(A(\cdot); X_1, \dots, X_d) \subset \Theta(\hat{A}(\cdot); X^{(l-\hat{l}+1)}).$$

Proof. Let $\lambda_0 \in \Theta(A(\cdot); X_1, \dots, X_d)$. Then there exists $(F, U) \in S_{\text{att}}^{op}(X_1, \dots, X_d)$ such that $0 \in \Sigma(FA(\lambda_0)U)$. We can regard $FA(\cdot)U$ as an operator polynomial

$$FA(\cdot)U : \mathbb{C} \to \mathcal{L}(\mathbb{C}^d) : \lambda \mapsto \sum_{j=0}^{l} \lambda^j F A_j U.$$

Then $\lambda \in \Sigma(FA(\cdot)U)$. Since $FA(\cdot)U$ is an operator polynomial of degree l it has a canonical degree $\hat{l}$ reduction

$$(\widehat{FA(\cdot)U}) : \mathbb{C} \to \mathcal{L}(\mathbb{C}^{d(l-\hat{l}+1)}).$$

Now with Proposition 3.4 we obtain $\Sigma(FA(\cdot)U) = \Sigma\big((\widehat{FA(\cdot)U})\big)$ and therefore $0 \in \Sigma\big((\widehat{FA(\lambda_0)}U)\big)$.
Now define the pair $(F^{(l-\hat{l}+1)}, U^{(l-\hat{l}+1)})$ through

$$F^{(l-\hat{l}+1)} := \underbrace{F \oplus \ldots \oplus F}_{(l-\hat{l}+1)\text{-times}}, \qquad U^{(l-\hat{l}+1)} := \underbrace{U \oplus \ldots \oplus U}_{(l-\hat{l}+1)\text{-times}}$$

resulting in $(F^{(l-\hat{l}+1)}, U^{(l-\hat{l}+1)}) \in S_{\text{att}}^{op}(X^{(l-\hat{l}+1)})$. In the same way as above we then get the operator polynomial $F^{(l-\hat{l}+1)}\hat{A}(\cdot)U^{(l-\hat{l}+1)} : \mathbb{C} \to \mathcal{L}(\mathbb{C}^{d(l-\hat{l}+1)})$. There now holds the equality

$$(\widehat{FA(\cdot)U}) = F^{(l-\hat{l}+1)}\hat{A}(\cdot)U^{(l-\hat{l}+1)}$$

and therefore

$$0 \in \Sigma\big((F\widehat{A(\lambda_0)}U)\big) = \Sigma\big(F^{(l-\hat{l}+1)}\hat{P}(\lambda_0)U^{(l-\hat{l}+1)}\big)$$

which yields $\lambda_0 \in \Theta(\hat{A}(\cdot); X^{(l-\hat{l}+1)})$. $\square$

Inclusions of this type in the case of linearisations were also considered [GLW09] where the authors used them to derive bounds of the block numerical range.

3.5. Degree-reduction of semi-monic operator polynomials

We return to our discussion of semi-monic operator polynomials

$$Q(\lambda) = \lambda^m I - A(\lambda) = \lambda^m I - \sum_{j=0}^{l} \lambda^j A_j \tag{3.45}$$

with $l > m$ and $A_j \in \mathcal{L}(X)$, $j = 1, \ldots, l$. Consider the degree-reduction of $Q(\lambda)$ to degree $l - m + 1$

$$\begin{aligned}
\hat{Q}(\lambda) &= \begin{bmatrix} \lambda I - A^{[l-m+1]}(\lambda) & -A_{m-2} & \ldots & -A_0 \\ -I & \lambda I & & \\ & \ddots & \ddots & \\ & & -I & \lambda I \end{bmatrix} \\
&= \lambda I_{X^m} - \begin{bmatrix} A^{[l-m+1]}(\lambda) & A_{m-2} & \ldots & A_0 \\ I & & & 0 \\ & \ddots & & \vdots \\ & & I & 0 \end{bmatrix} \\
&=: \lambda I_{X^m} - \tilde{A}(\lambda) = \lambda I_{X^m} - \sum_{j=0}^{l-m+1} \lambda^j \tilde{A}_j
\end{aligned} \tag{3.46}$$

where

$$\tilde{A}_0 = \begin{bmatrix} A_{m-1} & A_{m-2} & \ldots & A_0 \\ I & & & 0 \\ & \ddots & & \vdots \\ & & I & 0 \end{bmatrix},$$

$$\tilde{A}_j = \begin{bmatrix} A_{j+m-1} & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 \end{bmatrix}, \qquad j = 1, \ldots, l - m + 1.$$

We see that the reduced polynomial $\hat{Q}(\cdot)$ is again a semi-monic operator polynomial with linear monic part. In particular if the coefficients A_j are nonnegative then so are the coefficients $\tilde{A}_j$.

Remark 3.22. Note that in general the operator polynomial $\tilde{A}(\cdot)$ is not the degree-reduction of $A(\cdot)$ to degree $l - m + 1$.

3.5.1. Fix point iteration in semi-monic operator polynomials

For the degree reduced semi-monic operator polynomial $\hat{Q}(\cdot)$ as in (3.46) we can consider the operator equation

$$\hat{Q}(C) = C - \sum_{j=0}^{l-m+1} \tilde{A}_j C^j = 0. \tag{3.47}$$

By [Mar88, Lemma 22.9] the linear operator polynomial $(\lambda I - C)$ is then a right divisor of $\hat{Q}(\cdot)$. Consequently we will call an operator C satisfying (3.47) a *root* of $\hat{Q}(\cdot)$.

Remark 3.23. In (3.47) we multiply C from the right and end up with a right divisor. One could also multiply from the left and would then end up with a left divisor. While the results in the remainder of this section could also be stated for left divisors, we will only focus on right divisors.

In order to find a root of $\hat{Q}(\cdot)$ one might consider the fix point iteration given by

$$C_{k+1} := \sum_{j=0}^{l-m+1} \tilde{A}_j C_k^j \tag{3.48}$$

for some starting point C_0. Given that the iteration scheme converges (in the operator norm), then the limit point $\hat{C}$ will satisfy (3.47). W restate two results from [Har11] which give conditions under which the iteration scheme converges.

Proposition 3.24. *(i) Assume there exists a $\rho > 0$ such that*

$$\mathrm{norm}_{\tilde{A}}(\rho) = \sum_{j=0}^{l-m+1} \rho^j \|\tilde{A}_j\| < \rho.$$

Then the fix point iteration given by (3.48) *for $C_0 = 0$ converges.*

(ii) Let $Q(\cdot)$ be a semi-monic matrix polynomial with entrywise nonnegative matrix coefficients A_j. Then the canonical degree reduction $\hat{Q}(\cdot)$ as in (3.46) *has en entrywise nonnegative root if and only if the fix point iteration* (3.48) *with starting point $0 \leq C_0 \leq \tilde{A}_0$ converges.*

Proof. (i) is [Har11, Proposition 4.32]. (ii) follows from [Har11, Proposition 4.4] since the coefficients $\tilde{A}_j$ are entrywise nonnegative. □

Now assume that we have found a fix point $\hat{C}$ satisfying (3.47). Then $\hat{C}$ is a root of $\hat{Q}(\cdot)$ and we can write

$$\hat{Q}(\lambda) = \tilde{L}(\lambda)(\lambda I - \hat{C})$$

for some operator polynomial $\tilde{L}(\cdot)$. By Proposition 3.9 we conclude that $-\hat{C}$ is the companion matrix of a monic operator polynomial

$$D(\lambda) = \lambda^m I + \sum_{j=0}^{m-1} \lambda^j D_j$$

and $\hat{Q}(\cdot)$ factorizes as

$$\hat{Q}(\lambda) = \left(\sum_{j=0}^{l-m} \lambda^j L_j \right) D(\lambda),$$

where the coefficients L_j are determined by $\tilde{L}(\cdot)$ as in Proposition 3.9.
So in short we were able to find a monic right factor of $A(\cdot)$ by performing a fix point iteration for its canonical degree reduction.

3.5.2. Degree-reduction of irreducible semi-monic matrix polynomials

Now let $Q(\cdot), A(\cdot), \hat{Q}(\cdot)$ and $\tilde{A}(\cdot)$ as above but with the coefficients $A_j \in \mathbb{R}_+^{n,n}$. In Section 1.4 we have seen that semi-monic matrix polynomials with entrywise nonnegative coefficients are of particular interest if the sum of the coefficients $\sum_{j=0}^{m} A_j$ is irreducible. However, it is easy to see that the sum of the coefficients of the reduced polynomial, $\sum_{j=0}^{l-m+1} \tilde{A}_j = \tilde{A}(1)$, need no longer be irreducible (take any polynomial where $A_0 = 0$).

We will show that this problem can be remedied via a suitable permutation. The following algorithm describes how to obtain such a permutation.
For an $mn \times mn$-matrix B we will by $B[k, k]$ denote the leading principal minor of size $k \times k$, $1 \le k \le mn$.
The algorithm operates by 'removing' zero columns of $\tilde{A}(1)$ by moving them to the right via permutation (and leaving the order of the other columns unchanged). The same permutation is then applied to the rows.

Algorithm 3.25. Step 1: If the matrix $\tilde{A}(1)$ has no zero columns the algorithm stops. Otherwise note that $\tilde{A}(1)$ can only have zero columns in the last n coordinates. Let $s_1^1, \dots, s_{w_1}^1$ be the coordinates of these w_1 zero columns. Then apply the permutation that moves the columns to the right, leaving their order and the order of the remaining columns unchanged (i.e. the last w_1 columns of the permuted matrix are, from left to right, the previous columns $s_1^1, \dots, s_{w_1}^1$).
Now apply the same permutation to the rows of $\tilde{A}(1)$, resulting in a matrix B_1. If we denote by P_1 the corresponding permutation matrix, then $B_1 = P_1^T \tilde{A}(1) P$.
Step 2: If we restrict our view to the submatrix $B_1[mn - w_1, mn - w_1]$, it is possible that the permutation in step 1 has resulted in additional zero columns in $B_1[mn - w_1, mn - w_1]$ among the coordinates $(m-2)n + 1, \dots, (m-1)n$. Note however, that every such zero column must come from a zero column in step 1. To make this precise let $s_1^2, \dots, s_{w_2}^2$ be (the coordinates of) the zero columns in $B_1[mn - w_1, mn - w_1]$. Then for every $i \in \langle w_2 \rangle$ there must exist a $j \in \langle w_1 \rangle$ such that $s_i^2 = s_j^1 - n$. This additionally implies that $w_2 \le w_1$.
Now apply the permutation to B_1 that moves the zero columns with coordinates $s_1^2, \dots, s_{w_2}^2$ to the coordinates $mn - (w_1 + w_2) + 1, \dots, mn - w_1$ (leaving again the ordering intact). Apply the same permutation to the rows of B_2. If P_2 is the corresponding permutation matrix then define $B_2 = P_2^T B_1 P_2$.

$\vdots$

Step k: Repeat step 2 for the zero columns $s_1^k, \dots, s_{w_k}^k$ of the submatrix

$$B_{k-1}[mn - (w_1 + \cdots + w_{k-1}), mn - (w_1 + \cdots + w_{k-1})].$$

Stop: Terminate if no more zero columns appear. Output the permutation matrix $P = P_1 P_2 \cdots P_w$ where w is the number of steps the algorithm took.

The algorithm is guaranteed to terminate after at most $m-1$ steps (so the first n columns will not be permuted). This follows because the irreducibility of $A(1) = \sum_{j=0}^{l} A_j$ implies that for each $i \in \langle n \rangle$ at least one of the columns with coordinates $i, i+n, \dots, i+(m-1)n$ has nonzero entries among its first n entries.

Remark 3.26. We state some helpful observations:

(i) Let $s \in \langle m(n-1) \rangle$. Then the s-th column in $\tilde{A}(1)$ can only be removed by the algorithm if the column with coordinate $s+n$ was already removed by the algorithm.

(ii) Columns in $\tilde{A}(1)$ that have a nonzero entry among their first n entries will not be removed by the algorithm.

The next example illustrates the algorithm.

Example 3.27. Let $B_0 \in \mathbb{R}_+^{6,6}$ be of the form

$$B_0 = \begin{bmatrix} a & b & 0 & c & 0 & d \\ a & b & 0 & c & 0 & d \\ 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \end{bmatrix}$$

with some arbitrary positive numbers a, b, c, d. Then after the first step of Algorithm 3.25 we arrive at

$$B_1 = \left[\begin{array}{ccccc|c} a & b & 0 & c & d & 0 \\ a & b & 0 & c & d & 0 \\ 1 & & & & & \\ & 1 & & & & \\ & & & 1 & & \\ \hline & & 1 & & & \end{array}\right].$$

The next step leads to

$$B_2 = \left[\begin{array}{cccc|cc} a & b & c & d & 0 & 0 \\ a & b & c & d & 0 & 0 \\ & 1 & & & & \\ & & 1 & & & \\ \hline 1 & & & & & \\ & & & & 1 & \end{array}\right]$$

at which the algorithm concludes. Note that the upper left block is irreducible and the lower right block is nilpotent.

Our goal now is to show that the observation made at the end of Example 3.27 holds in general. Let P be the permutation matrix from Algorithm 3.25 where we assume that $\sum_{j=0}^{l} A_j$ is irreducible. Then we can identify

$$P^T \tilde{A}(1) P \mathrel{\widehat{=}} \left[\begin{array}{c|c} C & 0 \\ \hline R & N \end{array}\right] \tag{3.49}$$

where C and N are quadratic matrices and C has no zero columns. We claim that C is irreducible and that N is nilpotent. We denote by $p \in \mathbb{N}$ the size of C, i.e. $C \in \mathbb{R}_+^{p,p}$. Note that $n \leq p \leq mn$. Let further $\pi(\cdot) : \langle mn \rangle \to \langle mn \rangle$ the permutation function belonging to P.

Lemma 3.28. *Let $r_0 \to r_1 \to \cdots \to r_w$ be a path in the graph of $\tilde{A}(1)$. If $\pi(r_0) \in \langle p \rangle$ then $\pi(r_i) \in \langle p \rangle$ for $i = 1, \ldots, w$. In other words $\pi(r_0) \to \pi(r_1) \to \cdots \to \pi(r_w)$ is a path in the graph of C.*

Proof. We prove the assertion by induction: $\pi(r_0) \in \langle p \rangle$ by assumption. Now assume $\pi(r_i) \in \langle p \rangle$. If $r_i \in \langle n \rangle$ then the column of $\tilde{A}(1)$ with coordinate r_{i+1} has a nonzero entry among its first n entries. This implies (see Remark 3.26.(ii)) that $\pi(r_{i+1}) \in \langle p \rangle$. If $r_i \notin \langle n \rangle$ then the identities on the block minor diagonal of the matrix $\tilde{A}(1)$ imply that $r_{i+1} = r_i - n$. It then follows by Remark 3.26.(i) that $\pi(r_{i+1}) \in \langle p \rangle$. □

Lemma 3.29. *Let $r \in \langle mn \rangle$ and $s \in \langle n \rangle$ and assume that $\sum_{j=0}^{l} A_j$ is irreducible. Then there exists a path in the graph of $\tilde{A}(1)$ from r to s.*

Proof. Without loss of generality assume that $r \in \langle n \rangle$. Otherwise, by moving along the identities on the block minor diagonal of $\tilde{A}(1)$, we can construct a path

$$r \to r - n \to r - 2n \to \cdots \to \tilde{r}_0$$

such that $\tilde{r}_0 \in \langle n \rangle$.
Since the matrix $\sum_{j=0}^{l} A_j$ is irreducible, there exists a path in its corresponding graph from r to s. Denote this path by

$$r_0 \to r_1 \to \cdots \to r_w = \hat{r}.$$

Then we can, as in the beginning of the proof, construct a path from r_0 to r_w in the graph of $\tilde{A}(1)$. We demonstrate it for the first edge from r_0 to r_1: Since $A_j(r_0, r_1) > 0$ for some j, it follows that there exists a $k \in \langle m-1 \rangle_0$ such that $\tilde{A}(1)$ has a nonzero entry at coordinates $(r_0, r_1 + kn)$. Now, once again because of the identities on the block minor diagonal of $\tilde{A}(1)$, this implies the existence of the path

$$r_0 \to r_1 + kn \to r_1 + (k-1)n \to \cdots \to r_1 + n \to r_1. \qquad \square$$

Proposition 3.30. *Assume $\sum_{j=0}^{l} A_j$ is irreducible and let $C \in \mathbb{R}_+^{p,p}$ and $N \in \mathbb{R}_+^{mn-p,mn-p}$ be as in* (3.49). *Then C is irreducible and N is nilpotent.*

Proof. We first prove the irreducibility of C: Let $s_0, \bar{s} \in \langle p \rangle$. We need to show there exists a path from s_0 to $\bar{s}$ in the graph of C. Define $t_0 := \pi^{-1}(s_0)$ and $\bar{t} := \pi^{-1}(\bar{s})$. Our goal is to construct a path from t_0 to $\bar{t}$ in the graph of $\tilde{A}(1)$ and then apply Lemma 3.28. If $\bar{t} \in \langle n \rangle$ we can use Lemma 3.29 to get the existence of such a path.
Now assume $\bar{t} \notin \langle n \rangle$. Then there need to exist a $\hat{t} \in \langle n \rangle$ and a $k \in \langle m-1 \rangle_0$ such that

$$\hat{t} \to \bar{t} + kn \to \bar{t}(k-1)n \to \cdots \to \bar{t} + n \to \bar{t}$$

in the graph of $\tilde{A}(1)$. If such a path did not exist then, following Remark 3.26.(i), this would contradict $\pi(\bar{t}) = \bar{s} \in \langle p \rangle$. Together with Lemma 3.29 this gives us a path

$$t_0 \to \cdots \to \hat{t} \to \cdots \to \bar{t}$$

in the graph of $\tilde{A}(1)$. Now Lemma 3.28 tells us that

$$s_0 = \pi(t_0) \to \cdots \to \pi(\hat{t}) \to \cdots \to \pi(\bar{t}) = \bar{s}$$

is a path in the graph of C.

To see that N is nilpotent first note that the rightmost columns of N need to be zero columns. All other columns will contain a single 1 entry and be zero otherwise. Let r be the coordinate of a nonzero column in N. Then the column has coordinate $p+r$ in the bigger matrix $P^T \tilde{A}(1) P$. In the original matrix $\tilde{A}(1)$ it therefore had the coordinate $s := \pi^{-1}(p+r)$. In order to be removed by Algorithm 3.25 the column with coordinate $s+n$ must have been removed first (see Remark 3.26.(i)). The s-th column in $\tilde{A}(1)$ has its 1 entry at coordinate $s+n$. So when removing the column with coordinate $s+n$ by moving it to the right (and at the same time moving row $s+n$ down), the 1 entry in the column s also gets moved down to coordinate $\pi(s+n)$. After that it will not be affected by any other row operations.
From Algorithm 3.25 it follows that $\pi(s+n) > \pi(s) = p+r$ (since the algorithm adds to the zero columns from the left). It follows that the 1 entry of column $p+r$ lies in the strictly lower triangular part of $\tilde{A}(1)$. For the matrix N this implies that the 1 entry of the column r also lies in the strictly lower triangular part of N. Therefore N is a strictly lower triangular matrix and thus nilpotent. ☐

The proposition implies that the nonzero spectra of $\tilde{A}(1)$ and C coincide. The same holds for the semi-monic functions: Let P be the permutation matrix from Algorithm 3.25. Then we can identify

$$P^T \hat{Q}(\lambda) P = \lambda I_{mn} - P^T \tilde{A}(\lambda) P \mathrel{\widehat{=}} \left[\begin{array}{c|c} \lambda I_p & \\ \hline & \lambda I_{mn-p} \end{array}\right] - \left[\begin{array}{c|c} C(\lambda) & 0 \\ \hline R & N \end{array}\right]$$

where $C(\cdot)$ is a matrix polynomial of degree $l-m+1$ with entrywise nonnegative coefficients and $C(1)$ is irreducible. Now the semi-monic polynomial

$$Q_C(\lambda) = \lambda I_p - C(\lambda)$$

has the same nonzero spectrum as $\hat{Q}(\cdot)$ and thus also $Q(\lambda)$. But, in contrast to $\hat{Q}(\cdot)$, the irreducibility of $C(1)$ allows one to apply the results of Section 1.4 to $Q_C(\cdot)$.

A. The block numerical range

The content of this appendix is taken from [Kal11] with the exception of Section A.5 which was newly written for this thesis.

Let $(X_1, \|\cdot\|_{X_1}), \ldots, (X_d, \|\cdot\|_{X_d})$ be non trivial normed (Banach) spaces and consider the product space $X = X_1 \times \ldots \times X_d$. Equipped with one of the equivalent norms $\|\cdot\|_p$ defined through

$$\|\cdot\|_p : X \to \mathbb{R} : (x_1, \ldots, x_d)^T \mapsto \begin{cases} \sqrt[p]{\sum_{i=1}^d \|x_i\|_{X_i}^p}, & 1 \le p < \infty \\ \max_{i=1,\ldots,k} \|x_i\|_{X_i}, & p = \infty \end{cases}$$

X becomes a normed (Banach) space itself. Whenever we consider the space $\mathbb{C}^d$ it will be equipped with the corresponding p-norm.
An operator $A \in \mathcal{L}(X)$ can then be written as a so called *block operator* or *operator matrix*

$$\begin{bmatrix} A_{11} & \ldots & A_{1d} \\ \vdots & \ddots & \vdots \\ A_{d1} & \ldots & A_{dd} \end{bmatrix}$$

where $A_{rs} \in \mathcal{L}(X_s, X_r)$, for $r, s \in \langle d \rangle$. Introducing the natural 'projections'

$$P_k : X_1 \times \ldots \times X_d \to X_k : (x_1, \ldots, x_d)^T \mapsto x_k$$

and 'embeddings'

$$E_k : X_k \to X_1 \times \ldots \times X_d : x_k \mapsto (0, \ldots, 0, x_d, 0, \ldots, 0)^T$$

we can write A_{rs} explicitly as $A_{rs} = P_s A E_r$ for $k, r, s \in \langle d \rangle$.

A.1. Definitions and basic properties

Definition A.1. For a normed space Y we will denote the set of functionals which attain their norm on the unit sphere and corresponding points as

$$S_{\text{att}}(Y) := \{(f, u) : f \in Y', u \in Y, \|f\| = \|u\| = f(u) = 1\}.$$

Note that by the Hahn-Banach theorem we find for every normed $u \in X$ a normed functional $f \in X'$ with $f(u) = 1$ (the converse is in general only true in reflexive Banach spaces).
In order to work with block operators we extend this to the product space X via

$$S_{\text{att}}(X_1, \ldots, X_d) := S_{\text{att}}(X_1) \times \ldots \times S_{\text{att}}(X_d).$$

Given $\big((f_1, u_1), \ldots, (f_d, u_d)\big) \in S_{\text{att}}(X_1, \ldots, X_d)$ we define the two maps

$$\begin{aligned} F &: X_1 \times \ldots \times X_d \to \mathbb{C}^d, (x_1, \ldots, x_d) \mapsto (f_1(x_1), \ldots, f_d(x_d)) \\ U &: \mathbb{C}^d \to X_1 \times \ldots \times X_d, (\mu_1, \ldots, \mu_d) \mapsto (\mu_1 u_1, \ldots, \mu_d u_d) \end{aligned}$$

Let $S_{\text{att}}^{op}(X_1, \ldots, X_d)$ denote the set of all pairs (F, U) of these operators. Finally we define the *block numerical range* of an operator $A \in \mathcal{L}(X_1 \times \ldots \times X_d)$ as

$$\Theta(A; X_1, \ldots, X_d) := \bigcup_{(F,U) \in S_{\text{att}}^{op}(X_1,\ldots,X_d)} \Sigma(FAU).$$

If the underlying decomposition of the normed space is clear we will sometimes simply write S_{att}^{op} instead of $S_{\text{att}}^{op}(X_1, \ldots, X_d)$.

Remark A.2. - For $A \in \mathcal{L}(X)$ and $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ the operator FAU can be identified with the matrix

$$\begin{bmatrix} f_1(A_{11}u_1) & \ldots & f_1(A_{1d}u_d) \\ \vdots & \ddots & \vdots \\ f_d(A_{d1}u_1) & \ldots & f_d(A_{dd}u_d) \end{bmatrix}.$$

- The block numerical range can be characterized with the determinant on $\mathbb{C}^{d,d}$, i.e. for $\lambda \in \mathbb{C}$

$$\begin{aligned} \lambda \in \Theta(A; X_1, \ldots, X_d) &\Leftrightarrow \exists (F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d) : \det(FAU - \lambda) = 0 \\ &\Leftrightarrow \exists (F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d) : \det(F(A - \lambda)U) = 0. \end{aligned}$$

The second equivalence follows from $FU = I$.

- For every operator $A \in \mathcal{L}(X)$ the block numerical range is bounded by $\|A\|_X$ since

$$\|FAU\|_{\mathcal{L}(\mathbb{C}^d)} \leq \underbrace{\|F\|_{\mathcal{L}(X,\mathbb{C}^d)}}_{=1} \|A\|_X \underbrace{\|U\|_{\mathcal{L}(\mathbb{C}^d,X)}}_{=1} = \|A\|_X$$

for all $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$. The claim now follows because the spectrum of a matrix is bounded by its norm.

Let us now consider permutations of the order of the factors in $X_1 \times \ldots \times X_d$. For a permutation π of the set $\langle d \rangle$ define $X^\pi := X_{\pi(1)} \times \ldots \times X_{\pi(d)}$. The operator induced by π is

$$P : X \to X^\pi : \begin{bmatrix} u_1 \\ \vdots \\ u_d \end{bmatrix} \mapsto \begin{bmatrix} u_{\pi(1)} \\ \vdots \\ u_{\pi(d)} \end{bmatrix}$$

with the inverse operator P^{-1} induced by the inverse permutation. The operator $A^\pi := P^{-1}AP$ is then the equivalent of A in $\mathcal{L}(X^\pi)$.

Lemma A.3. *The block numerical range of an operator $A \in \mathcal{L}(X)$ is invariant under permutation of the factors X_i. That is for a permutation π of $\langle 1, d \rangle$ there holds*

$$\Theta(A^\pi; X_{\pi(1)}, \ldots, X_{\pi(d)}) = \Theta(A; X_1, \ldots, X_d).$$

Proof. Let $\lambda \in \Theta(A; X_1, \ldots, X_d)$ and $(F, U) \in S^{op}_{\text{att}}(X_1, \ldots, X_d)$ such that $\det(FAU - \lambda) = 0$. Then

$$0 = \det(FAU - \lambda) = \det(FPA^{\pi}P^{-1}U - \lambda).$$

Obviously $(FP, P^{-1}U) \in S^{op}_{\text{att}}(X_{\pi(1)}, \ldots, X_{\pi(d)})$ and thus $\lambda \in \Theta(A^{\pi}; X_{\pi(1)}, \ldots, X_{\pi(d)})$. The assertion now follows because the reverse inclusion is symmetric. □

Example A.4. (1) In case of a Hilbert space H with scalar product $(\cdot, \cdot)$ we can consider an orthogonal decomposition $H = H_1 \oplus \ldots \oplus H_d$. If we combine the factors H_i with the 2-norm we get the isometric isomorphism

$$H \cong H_1 \times \ldots \times H_d.$$

In Hilbert spaces by the Riesz isomorphism theorem there is exactly one normed $f_i \in H_i'$ for every normed $u_i \in H_i$ such that $f_i(u_i) = 1$ and it can be identified through the scalar product with $(\cdot, u_i)$. Therefore $S_{\text{att}}(H_1, \ldots, H_d)$ can be identified with the factor wise normed sphere $S(H_1, \ldots, H_d) := \{u \in H : \|u_i\|_{H_i} = 1\}$.
For an operator $A \in \mathcal{L}(H)$ the block numerical range then is

$$\begin{aligned} \Theta(A; H_1, \ldots, H_d) &= \bigcup_{(F,U) \in S^{op}_{\text{att}}(H_1,\ldots,H_d)} \Sigma(FAU) \\ &= \bigcup_{(f,u) \in S_{\text{att}}(H_1,\ldots,H_d)} \Sigma \begin{bmatrix} f_1(A_{11}u_1) & \ldots & f_1(A_{1d}u_d) \\ \vdots & \ddots & \vdots \\ f_d(A_{d1}u_1) & \ldots & f_d(A_{dd}u_d) \end{bmatrix} \\ &= \bigcup_{u \in S(H_1,\ldots,H_d)} \Sigma \begin{bmatrix} (A_{11}u_1, u_1) & \ldots & (A_{1d}u_d, u_1) \\ \vdots & \ddots & \vdots \\ (A_{d1}u_1, u_d) & \ldots & (A_{dd}u_d, u_d) \end{bmatrix}. \end{aligned}$$

This is just the block numerical range for operators on Hilbert spaces as introduced in [Wag07] or [Tre08]. The block numerical range for normed spaces as introduced here is therefore a generelazition of the known case on Hilbert spaces. Lemma A.3 guarantees that the block numerical range does not depend on the order of the orthogonal subspaces in the direct sum.

(2) Regarding the Hilbert space $\mathbb{C}^n$ a special type of orthogonal decomposition is the so called *component wise decomposition* as introduced in [FH08]. For it consider index sets $\mathcal{J}_i$ for $i = 1, \ldots, d$ such that

$$\mathcal{J}_1 \dot{\cup} \ldots \dot{\cup} \mathcal{J}_d = \langle 1, n \rangle.$$

Now define the spaces

$$X_i := \{(\nu_1, \ldots, \nu_n)^T \in \mathbb{C}^n : \nu_j = 0 \text{ for } j \notin \mathcal{J}_i\}.$$

Since the N_i were disjoint we get the decomposition $\mathbb{C}^n = X_1 \oplus \ldots \oplus X_d$.
The advantage of component wise decompositions is that the block operator representation of entrywise nonnegative matrices $A \in \mathbb{C}^{n\times n}$ with regard to these decompositions preserves the property of being entrywise nonnegative (this is in general not true if we consider an arbitrary orthogonal decomposition of $\mathbb{C}^n$). For more on the block numerical range in $\mathbb{C}^n$ see Section A.5.

(3) Let $a = t_0 < t_1 < \ldots < t_d = b \in \mathbb{R}$. The Banach space $X = L^p([a,b])$ for $p \in [0,\infty]$ can then be decomposed into

$$L^p([a,b]) \cong L^p([t_0,t_1]) \times L^p([t_1,t_2]) \times \ldots \times L^p([t_{d-1},t_d]),$$

where we have to combine the factors with the p-norm.

A.2. Spectral inclusion

We will now prove spectral inclusion results where the later results require the normed spaces X_i to be Banach spaces.

Lemma A.5. *There holds $\Sigma_p(A) \subseteq \Theta(A; X_1, \ldots, X_d)$.*

Proof. Let $\lambda \in \Sigma_p(A)$ then there exists $x = (x_1, \ldots, x_d)^T \in X_1 \times \ldots \times X_d$, such that $Ax = \lambda x$. Define normed $u_i \in X_i$ satisfying $\|x_i\|u_i := x_i$ for $i = 1, \ldots, d$. By the Hahn-Banach theorem there exist normed functionals $f_i \in X_i'$ such that $f_i(u_i) = \|u_i\| = 1$ for $i = 1, \ldots, d$, thus $\big((f_1,u_1), \ldots, (f_d,u_d)\big) \in S_{\text{att}}(X_1, \ldots, X_d)$. For the corresponding $(F,U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ there holds

$$FAU \begin{bmatrix} \|x_1\| \\ \vdots \\ \|x_d\| \end{bmatrix} = FA \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \lambda F \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \lambda \begin{bmatrix} \|x_1\| \\ \vdots \\ \|x_d\| \end{bmatrix},$$

i.e. $\lambda \in \Theta(A; X_1, \ldots, X_d)$. □

Lemma A.6. *Let $X_1, \ldots, X_d$ be Banach spaces and $X = X_1 \times \ldots \times X_d$ endowed with the p-norm and $X_1' \times \ldots \times X_d'$ with the conjugated q-norm (i.e. $p^{-1} + q^{-1} = 1$ with the usual convention $\frac{1}{\infty} = 0$). Then the map*

$$J : X_1' \times \ldots \times X_d' \to X', (f_1, \ldots, f_n) \mapsto f$$

with

$$f(x_1, \ldots, x_d)^T := \sum_{i=1}^{d} f_i(x_i)$$

is an isometric isomorphism.

Lemma A.7. *Let A' denote the adjoint operator of $A \in \mathcal{L}(X)$ and let $(\tilde{A}_{ij})_{ij}$ be the block operator representation of A' with regard to the decomposition $X_1' \times \ldots \times X_d'$. Then there holds $\tilde{A}_{ij} = A_{ji}'$, $i,j = 1, \ldots, d$. That is, the block operator of A' is the transpose of the component-wise adjoint of A.*

Proof. For $k = 1, \ldots, d$ let E_k denote the 'embedding' from X_k into $X_1 \times \ldots \times X_d$ and let P_k be the 'projection' from $X_1 \times \ldots \times X_d$ onto X_k. Furthermore let $\tilde{E}_k$ and $\tilde{P}_k$ denote their dual space counterparts. Using the isometric isomorphism from Lemma A.6 we obtain

$$\tilde{A}_{ij} = \tilde{P}_i \circ J^{-1} \circ A' \circ J \circ \tilde{E}_j.$$

Let $f_k \in X'_k$ and $(x_1, \ldots, x_d) \in X$ then

$$(f_k \circ P_k) \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = f_k(x_k) = [(J \circ \tilde{E}_k)(f_k)] \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}$$

i.e.

$$f_k \circ P_k = (J \circ \tilde{E}_k)(f_k). \tag{A.1}$$

Let $f = J^{-1}(f_1, \ldots, f_d) \in X'$ and $x_k \in X_k$ then

$$(f \circ E_k)(x_k) = f_k(x_k) = \left(\tilde{P}_k \begin{bmatrix} f_1 \\ \vdots \\ f_d \end{bmatrix} \right) (x_k) = [(\tilde{P}_k \circ J^{-1})(f)](x_k)$$

i.e.

$$f \circ E_k = (\tilde{P}_k \circ J^{-1})(f). \tag{A.2}$$

For $f_j \in X'_j$ we then have

$$\begin{aligned} \tilde{A}_{ij} f_j &= (\tilde{P}_i \circ J^{-1} \circ A' \circ J \circ \tilde{E}_j)(f_j) \\ &= (\tilde{P}_i \circ J^{-1})(A'[(J \circ \tilde{E}_j)(f_j)]) \\ &= (\tilde{P}_i \circ J^{-1})([(J \circ \tilde{E}_j)(f_j)] \circ A) \\ &\overset{(A.1)}{=} (\tilde{P} \circ J^{-1})(f_j \circ P_j \circ A) \\ &\overset{(A.2)}{=} f_j \circ P_j \circ A \circ E_j \\ &= f_j \circ A_{ji} = A'_{ji} f_j \end{aligned}$$

as desired. □

The following important theorem basically says that the norm attaining functionals lie dense in the dual space.

Theorem A.8 (Bishop-Phelps-Bollobás)**.** *Let X be a Banach space and $0 < \varepsilon < 1$. Let $f \in X'$, $\|f\| = 1$ and choose $v \in X$, $\|v\| \leq 1$ such that $|1 - f(v)| < \frac{\varepsilon^2}{4}$. Then there exists $u \in X$, $\|u\| = 1$ and $g \in X'$, $\|g\| = 1$ such that $g(u) = 1$ and $\|u - v\| < \varepsilon$ and $\|g - f\| < \varepsilon$.*

Proof. See [BD73, §16 Theorem 1]. □

Lemma A.9. *Let X be decomposed as before then*

$$\Theta(A; X_1, \dots, X_d) \subseteq \Theta(A'; X_1', \dots, X_d') \subseteq \overline{\Theta(A; X_1, \dots, X_d)}$$

holds.

Proof. For $i = 1, \dots, d$ let J_i denote the canonical embedding of X_i into its double dual. For the first inclusion let $\lambda \in \Theta(A; X_1, \dots, X_d)$ i.e. there exist $\big((f_1, u_1), \dots, (f_d, u_d)\big) \in S_{\text{att}}(X_1, \dots, X_d)$ and corresponding $(F, U) \in S_{\text{att}}^{op}(X_1, \dots, X_d)$ such that $\det(FAU - \lambda) = 0$. We define $\tilde{u}_i := f_i$ and $\tilde{f}_i := J_i(u_i)$. Then $\big((\tilde{f}_1, \tilde{u}_1), \dots, (\tilde{f}_d, \tilde{u}_d)\big) \in S_{\text{att}}(X_1', \dots, X_d')$ holds. Let $(\tilde{F}, \tilde{U}) \in S_{\text{att}}^{op}(X_1', \dots, X_d')$ denote the corresponding pair of operators. Using the last lemma we obtain

$$\begin{aligned}
\tilde{F} A' \tilde{U} &= \big(\tilde{f}_i(A_{ji}' \tilde{u}_j)\big)_{ij} \\
&= \big(\tilde{f}_i(\tilde{u}_j \circ A_{ji})\big)_{ij} \\
&= \big(\tilde{f}_i(f_j \circ A_{ji})\big)_{ij} \\
&= \big((f_j \circ A_{ji})(u_i)\big)_{ij} \\
&= (FAU)^T
\end{aligned}$$

Since $\Sigma(FAU)^T = \Sigma(FAU)$ we have $\lambda \in \Theta(A'; X_1', \dots, X_d')$.
For the latter inclusion let $\lambda \in \Theta(A'; X_1', \dots, X_d')$, i.e. there exist $(\tilde{f}, \tilde{u}) \in S_{\text{att}}(X_1', \dots X_d')$ and corresponding $(\tilde{F}, \tilde{U}) \in S_{\text{att}}^{op}(X_1', \dots, X_d')$ such that $\det(\tilde{F} A' \tilde{U} - \lambda) = 0$. Since the canonical embedding of the unit sphere of a Banach space into the unit sphere of its double dual is weak-* dense, there exist for all $i, j \in \langle 1, d \rangle$ and $0 < \varepsilon < 1$ elements $v_i \in X_i$ with $\|v_i\| \leq 1$ such that

$$\begin{aligned}
\big|\tilde{f}_i(\tilde{u}_i) - \big(J_i(v_i)\big)(\tilde{u}_i)\big| &< \frac{\varepsilon^2}{4}, \\
\big|\tilde{f}_i(A_{ji}' \tilde{u}_j) - \big(J_i(v_i)\big)(A_{ji}' \tilde{u}_j)\big| &< \varepsilon.
\end{aligned}$$

Therefore

$$|1 - \tilde{u}_i(v_i)| = \big|\tilde{f}_i(\tilde{u}_i) - \big(J_i(v_i)\big)(\tilde{u}_i)\big| < \frac{\varepsilon^2}{4}.$$

By the Bishops-Phelps-Bollobás theorem there exist $(f_i, u_i) \in S_{\text{att}}(X_i)$ with $\|f_i - \tilde{u}_i\| < \varepsilon$ and $\|u_i - v_i\| < \varepsilon$ for all $i \in \langle 1, d \rangle$, providing $(f, u) \in S_{\text{att}}(X_1, \dots, X_d)$ and corresponding $(F, U) \in S_{\text{att}}^{op}(X_1, \dots, X_d)$. Now define for $(v_1, \dots, v_d)^T$ as usual the operator

$$V : \mathbb{C}^d \to X : \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix} \mapsto \begin{bmatrix} \mu_1 v_1 \\ \vdots \\ \mu_d v_d \end{bmatrix}.$$

Note that since $\|v_i\| \leq 1$ for all $i \in \langle 1, d \rangle$ there holds $\|V\| < 1$. Now

$$\begin{aligned}
\|\tilde{F} A' \tilde{U} - FAU\| &\leq \|\tilde{F} A' \tilde{U} - \tilde{U} A' V\| + \|\tilde{U} AV - \tilde{U} AU\| + \|\tilde{U} AU - FAU\| \\
&\leq \underbrace{\Big\|\Big(\tilde{f}_i(A_{ji}' \tilde{u}_j) - \big(J_i(v_i)\big)(A_{ji}' \tilde{u}_j)\Big)_{ij}\Big\|}_{< c\varepsilon} + \underbrace{\|V - U\|}_{< \varepsilon} + \underbrace{\|\tilde{U} - F\|}_{< \varepsilon},
\end{aligned}$$

where c is a constant only depending on the norm on $\mathbb{C}^d$. Since ε was arbitrary we conclude that $\lambda \in \overline{\Theta(A; X_1, \ldots, X_d)}$. $\square$

Corollary A.10. *If $X_1, \ldots, X_d$ are reflexive, then*

$$\Theta(A; X_1, \ldots, X_d) = \Theta(A'; X_1', \ldots, X_d')$$

holds.

Lemma A.11. *Let X be a product of Banach spaces and $A \in \mathcal{L}(X)$. Then*

(i) $\Sigma_r(A) \subseteq \Sigma_p(A')$

(ii) $\Sigma_c(A) \subseteq \Sigma_{app}(A)$

hold. The approximate point spectrum Σ_{app} of A is defined by

$$\lambda \in \Sigma_{app}(A) \quad \Leftrightarrow \quad \exists (x_k)_k \subseteq X,\ \|x_k\| = 1 : Ax_k - \lambda x_k \xrightarrow{k\to\infty} 0.$$

Lemma A.12. *Let $M \in \mathbb{C}^{d\times d}$ be a matrix. If M is invertible then*

$$\|M^{-1}\| \leq \frac{\|M\|^{d-1}}{|\det M|}.$$

For $x \in \mathbb{C}^d$ with $\|x\| = 1$ we have

$$\operatorname{dist}(0, \Sigma(M)) \leq \sqrt[d]{\|M\|^{d-1}\,\|Mx\|}.$$

Proof. See [Tre08, Lemma 1.11.5]. $\square$

Theorem A.13. *Let X and A be as in Lemma A.11. Then*

(i) $\Sigma_p(A) \subseteq \Theta(A; X_1, \ldots, X_d)$

(ii) $\Sigma(A) \subseteq \overline{\Theta(A; X_1, \ldots, X_d)}$

hold.

Proof. Point (i) is Lemma A.5. For point (ii) let $\lambda \in \Sigma(A)$. If $\lambda \in \Sigma_p(A)$ the claim follows from point (i). If $\lambda \in \Sigma_r(A)$ then by Lemma A.11, point (i) and Lemma A.9

$$\lambda \in \Sigma_p(A') \subseteq \Theta(A'; X_1', \ldots, X_d') \subseteq \overline{\Theta(A; X_1, \ldots, X_d)}.$$

Lastly if $\lambda \in \Sigma_c(A)$ then again by Lemma A.11 $\lambda \in \Sigma_{\text{app}}(A)$. Thus there is a sequence

$$(x^{(k)})_k = \left(\left(x_1^{(k)}, \ldots, x_d^{(k)}\right)\right)_k \subset X,\ \|x^{(k)}\| = 1$$

such that $(A-\lambda)x^{(k)} \xrightarrow{k\to\infty} 0$. Choose $u_i^{(k)} \in X_i$, $\|u_i^{(k)}\| = 1$ such that $x_i^{(k)} = \|x_i^{(k)}\| u_i^{(k)}$ and $f_i^{(k)} \in X_i'$, $\|f_i^{(k)}\| = 1$ such that $f_i^{(k)}(u_i^{(k)}) = 1$ for $i = 1, \ldots, d$ and all $k \in \mathbb{N}$. Now we can construct the corresponding pair of operators $(F^{(k)}, U^{(k)}) \in S_{\text{att}}^{op}(X_1, \dots, X_d)$. Define $v^{(k)} := \big(\|x_1^{(k)}\|, \dots, \|x_d^{(k)}\|\big)$, then there holds

$$\begin{aligned}
\left(F^{(k)} A U^{(k)} - \lambda\right) v^{(k)} &= \left(F^{(k)} A U^{(k)} - F^{(k)} U^{(k)} \lambda\right) v^{(k)} \\
&= \left(F^{(k)} \left(A - \lambda\right) U^{(k)}\right) v^{(k)} \\
&= F^{(k)} \left(\left(A - \lambda\right) x^{(k)}\right) \xrightarrow{k \to \infty} 0.
\end{aligned}$$

Define $\varepsilon_k := \|(F^{(k)} A U^{(k)} - \lambda) v^{(k)}\|$. Using Lemma A.12 we obtain

$$\begin{aligned}
\operatorname{dist}\left(\lambda, \Sigma\big(F^{(k)} A U^{(k)}\big)\right) &= \operatorname{dist}\left(0, \Sigma\big(F^{(k)} A U^{(k)} - \lambda\big)\right) \\
&\leq \sqrt[d]{\left\|F^{(k)} A U^{(k)} - \lambda\right\|^{d-1} \varepsilon_k} \\
&\leq \sqrt[d]{\left(\left\|F^{(k)} A U^{(k)}\right\| + |\lambda|\right)^{d-1} \varepsilon_k} \\
&\leq \sqrt[d]{\left(\left\|F^{(k)}\right\| \cdot \|A\| \cdot \left\|U^{(k)}\right\| + |\lambda|\right)^{d-1} \varepsilon_k} \\
&\leq \sqrt[d]{\left(\|A\| + |\lambda|\right)^{d-1} \varepsilon_k} \quad \xrightarrow{k \to \infty} 0.
\end{aligned}$$

Therefore

$$\lambda \in \overline{\bigcup_{k \in \mathbb{N}} \Sigma\big(F^{(k)} A U^{(k)}\big)} \subseteq \overline{\bigcup_{(F,U) \in S_{\text{att}}^{op}} \Sigma\left(FAU\right)} = \overline{\Theta\left(A; X_1, \dots, X_d\right)}. \qquad \square$$

A.3. Refinement of the decomposition

Definition A.14. Let $d, \hat{d} \in \mathbb{N}$ and $X_1 \times \ldots \times X_d = \hat{X}_1 \times \ldots \times \hat{X}_{\hat{d}} = X$ be two decompositions of X. We call $\hat{X}_1, \dots, \hat{X}_{\hat{d}}$ a *refinement* of $X_1, \dots, X_d$ if $d \leq \hat{d}$ holds and if there exist indices $0 = i_0 < \ldots < i_d = \hat{d}$ such that

$$X_k = \hat{X}_{i_{k-1}+1} \times \ldots \times \hat{X}_{i_k}, \; k = 1, \dots, d.$$

Remark A.15. Let X be the the product space of d one dimensional spaces $X_1, \dots, X_d$. Then for every linear map $A \in \mathcal{L}(X)$

$$\Theta\left(A; X_1, \dots, X_d\right) = \Sigma\left(A\right) = \Sigma_p(A)$$

holds.

Theorem A.16. *Let $X_1, \dots, X_d$ be a refinement of $\hat{X}_1, \dots, \hat{X}_{\hat{d}}$, both product spaces endowed with the p-norm for $1 \leq p \leq \infty$ and $A \in \mathcal{L}(X)$. Then*

$$\Theta(A; \hat{X}_1, \dots, \hat{X}_{\hat{d}}) \subseteq \Theta(A; X_1, \dots, X_d)$$

holds.

Proof. Without loss of generality let $\hat{d} = d + 1$ and refine the first factor of X. The general case then follows from Lemma A.3 and by induction.
We only consider the case for $1 \leq p < \infty$. The proof for $p = \infty$ is similar. Define $X_1 = \hat{X}_0 \times \hat{X}_1$ and $X_k = \hat{X}_k$ for $k = 2, \ldots, d$. Let $\lambda \in \Theta(A; \hat{X}_0, \hat{X}_1, \hat{X}_2, \ldots, \hat{X}_d)$, then there exists $[(\hat{f}_0, \hat{u}_0), \ldots, (\hat{f}_d, \hat{u}_d)]^T \in S_{\text{att}}(\hat{X}_0, \hat{X}_1, \hat{X}_2, \ldots, \hat{X}_d)$ and the corresponding pair of operators $(\hat{F}, \hat{U}) \in S^{op}_{\text{att}}(\hat{X}_0, \hat{X}_1, \hat{X}_2, \ldots, \hat{X}_d)$ such that λ is an eigenvalue of $\hat{F}A\hat{U}$. Let $v = (\nu_0, \nu_1, \ldots, \nu_d)^T \in \mathbb{C}^{d+1}$ be an associated eigenvector with the property

$$|\nu_0|^p + |\nu_1|^p = 1 \tag{A.3}$$

which can be achieved by scaling as long as $|\nu_0| + |\nu_1| \neq 0$ (otherwise take an arbitrary eigenvector). We define two linear maps

$$\begin{aligned} P &: \mathbb{C}^{d+1} \to \mathbb{C}^d : (\omega_0, \ldots, \omega_d)^T \mapsto (\alpha\omega_0 + \beta\omega_1, \omega_2, \ldots, \omega_d)^T, \\ E &: \mathbb{C}^d \to \mathbb{C}^{d+1} : (\omega_1, \ldots, \omega_d)^T \mapsto (\nu_0\omega_1, \nu_1\omega_1, \omega_2, \ldots, \omega_d)^T \end{aligned}$$

where

$$\alpha := \frac{|\nu_0|^p}{\nu_0}, \qquad \beta := \frac{|\nu_1|^p}{\nu_1}.$$

If $\nu_0 = 0$ set $\alpha = 0$, if $\nu_1 = 0$ set $\beta = 0$. We now show that the operators $F := P\hat{F}$ and $U := \hat{U}E$ are in $S^{op}_{\text{att}}(X_1, \ldots, X_d)$:

regarding U: Let $w := (\omega_1, \ldots, \omega_d)^T \in \mathbb{C}^d$, then

$$Uw = \hat{U}Ew = \Big(\omega_1 \underbrace{\begin{pmatrix} \nu_0\hat{u}_0 \\ \nu_1\hat{u}_1 \end{pmatrix}}_{:=u_1}, \omega_2\hat{u}_2, \ldots, \omega_d\hat{u}_d\Big)^T.$$

So we calculate

$$\|u_1\|_p = \left\| \begin{pmatrix} \nu_0\hat{u}_0 \\ \nu_1\hat{u}_1 \end{pmatrix} \right\|_p = \sqrt[p]{|\nu_0|^p \|\hat{u}_0\|^p + |\nu_1|^p \|\hat{u}_1\|^p} \overset{(A.3)}{=} 1.$$

If $\nu_0 = \nu_1 = 0$ choose an arbitrary $u_1 \in X_1$, $\|u_1\| = 1$. Defining $u_k := \hat{u}_k$ we have $\|u_k\| = 1$ for $k = 2, \ldots, d$.

regarding F: Let $x := (x_0, \ldots, x_d)^T \in \hat{X}_0 \times \hat{X}_1 \times X_2 \times \ldots \times X_d$, then

$$Fx = P\hat{F}x = \Big(\underbrace{\left(\alpha\hat{f}_0, \beta\hat{f}_1\right)}_{:=f_1} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \hat{f}_2(x_2), \ldots, \hat{f}_d(x_d)\Big)^T.$$

Now

$$\begin{aligned}
\|f_1\| &= \left\|(\alpha\hat{f}_0, \beta\hat{f}_1)\right\| \\
&= \sup_{\substack{\|(x_0,x_1)\|_p=1 \\ (x_0,x_1)\in X_1}} \left|\left(\alpha\hat{f}_0, \beta\hat{f}_1\right)\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}\right| \\
&= \sup_{\|x_0\|^p+\|x_1\|^p=1} \left|\alpha\hat{f}_0(x_0) + \beta\hat{f}_1(x_1)\right| \\
&\leq \sup_{\|x_0\|^p+\|x_1\|^p=1} \frac{|\nu_0|^p}{|\nu_0|}\|\hat{f}_0\|\,\|x_0\| + \frac{|\nu_1|^p}{|\nu_1|}\|\hat{f}_1\|\,\|x_1\| \\
&= \sup_{\|x_0\|^p+\|x_1\|^p=1} |\nu_0|^{p-1}\|x_0\| + |\nu_1|^{p-1}\|x_1\| \\
&\leq \qquad |\nu_0|^p + |\nu_1|^p \overset{(A.3)}{=} 1.
\end{aligned}$$

If $\nu_0 = \nu_1 = 0$ construct $f_1 \in X_1'$ such that $f_1(u_1) = \|f_1\| = 1$. Otherwise

$$f_1(u_1) = \left(\alpha\hat{f}_0, \beta\hat{f}_1\right)\begin{pmatrix} \nu_0\hat{u}_0 \\ \nu_1\hat{u}_1 \end{pmatrix} = \alpha\nu_0\hat{f}_0(\hat{u}_0) + \beta\nu_1\hat{f}_1(\hat{u}_1) = \frac{|\nu_0|^p}{|\nu_0|}\nu_0 + \frac{|\nu_1|^p}{|\nu_1|}\nu_1 \overset{(A.3)}{=} 1.$$

Defining $f_k := \hat{f}_k$ we have $\|f_k\| = \|\hat{f}_k\| = 1$ and $f_k(u_k) = \hat{f}_k(\hat{u}_k) = 1$ for $k = 2, \ldots, d$.

Hence $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$. Finally

$$\begin{aligned}
FAU\begin{pmatrix} 1 \\ \nu_2 \\ \vdots \\ \nu_d \end{pmatrix} &= P\hat{F}A\hat{U}E\begin{pmatrix} 1 \\ \nu_2 \\ \vdots \\ \nu_d \end{pmatrix} = P\hat{F}A\hat{U}\begin{pmatrix} \nu_0 \\ \nu_1 \\ \vdots \\ \nu_d \end{pmatrix} = \lambda P\begin{pmatrix} \nu_0 \\ \nu_1 \\ \vdots \\ \nu_d \end{pmatrix} \\
&= \lambda\begin{pmatrix} |\nu_0|^p + |\nu_1|^p \\ \nu_2 \\ \vdots \\ \nu_d \end{pmatrix} \overset{(A.3)}{=} \lambda\begin{pmatrix} 1 \\ \nu_2 \\ \vdots \\ \nu_d \end{pmatrix}
\end{aligned}$$

holds (if $\nu_0 = \nu_1 = 0$ set the first entry of the vector to be zero instead of 1). That is $\lambda \in \Theta(A; X_1, \ldots, X_d)$. □

A.4. Continuity properties and connected components

In [BD71, §11 Theorem 4] it was shown that $S_{\text{att}}(Y)$ is connected for every Banach space Y in the product topology on $Y \times Y'$ given by the norm topology on Y and the weak-* topology on Y'. Using the corresponding product topology on $(X_1 \times X_1') \times \ldots \times (X_d \times X_d')$ this naturally extends to the following result.

Lemma A.17. *In the above mentioned topology $S_{att}(X_1, \dots, X_d)$ is connected.*

In order to be able to state some continuity results on S_{att} and the block numerical range we will use the following metrics.

- On the complex plane we will use the Hausdorff metric defined on compact subsets Ψ_1, Ψ_2 of $\mathbb{C}$ through

$$\mathrm{d}_H(\Psi_1, \Psi_2) := \max\left\{\sup_{\lambda\in\Psi_2} \operatorname{dist}(\lambda, \Psi_1), \sup_{\lambda\in\Psi_1} \operatorname{dist}(\lambda, \Psi_2)\right\}.$$

- Let $(X, \|\cdot\|)$ be a Banach space. Denote by $N(X)$ the set of norms on X equivalent to $\|\cdot\|$. In [BD73] F. F. Bonsall and J. Duncan introduced the following metric on $N(X)$: Let $p, q \in N(X)$. Since p and q are equivalent there exists $\delta > 0$ such that

$$\frac{1}{\delta} \leq \frac{p(x)}{q(x)} \leq \delta \qquad \forall x \in X \setminus \{0\}.$$

  We then define

$$\mathrm{d}_N(p, q) := \log\inf\left\{\delta \geq 1 : \frac{1}{\delta} \leq \frac{p(x)}{q(x)} \leq \delta \qquad \forall x \in X \setminus \{0\}\right\}.$$

  It is not hard to see that d_N forms a metric on $N(X)$.

- On $S_{\text{att}}(X_1, \dots, X_d)$ we use the norm

$$\|(f, u)\|_{S_{\text{att}}} := \sum_{i=1}^{d} \|f_i\|_{X_i'} + \|u_i\|_{X_i}, \qquad (f, u) \in S_{\text{att}}(X_1, \dots, x_d).$$

  Note that this norm is always defined in terms of the original norms on the factors, even if we change to equivalent norms.

The following statements about continuity are to be read in regard to these metrics.

Proposition A.18. *Let $X = X_1 \times \cdots \times X_d$ be Banach spaces and $A \in \mathcal{L}(X)$.*

(i) $S_{att}(X_1, \dots, X_d)$ depends continuously on equivalent norms on the factors X_i.

(ii) $\overline{\Theta(A; X_1, \dots, X_d)}$ depends continuously on equivalent norms on the factors X_i.

(iii) $\overline{\Theta(A; X_1, \dots, X_d)}$ depends continuously on the operator A.

Proof. (i) The proof for the case of a trivially decomposed space X was already done in [BD73, §18 Theorem 3]. This immediately extends to $S_{\text{att}}(X_1, \dots, X_d)$.

(ii) Let $A \in \mathcal{L}(X)$. First we show that the block numerical range depends continuously on S_{att}: For a decomposition $X_1, \ldots, X_d$ denote by $\hat{X}_1, \ldots, \hat{X}_d$ the same spaces endowed with equivalent norms such that given $\delta > 0$ we find for every $(f, u) \in S_{\mathrm{att}}(X_1, \ldots, X_d)$ a $(g, v) \in S_{\mathrm{att}}(\hat{X}_1, \ldots, \hat{X}_d)$ such that $\|(f - g, u - v)\|_{S_{\mathrm{att}}} < \delta$ (and vice versa). Now

$$\begin{aligned} \|FAU - GAV\| &\leq \|FAU - GAU\| + \|GAU - GAV\| \\ &\leq \|F - G\|\|A\|\|U\| + \|G\|\|A\|\|U - V\| \\ &= \big(\|F - G\| + \|U - V\|\big)\|A\| \\ &< \delta\|A\|. \end{aligned}$$

It follows that the eigenvalues of FAU and GAV differ by at most $\delta\|A\|$ and therefore

$$\mathrm{d}_H\left(\overline{\Theta(A; X_1, \ldots, X_d)}, \overline{\Theta(A; \hat{X}_1, \ldots, \hat{X}_d)}\right) \leq \delta\|A\|.$$

The assertion now follows from point (i).

(iii) For fixed $(F, U) \in S_{\mathrm{att}}^{op}(X_1, \ldots, X_d)$ the mapping $A \mapsto FAU$ is continuous for $A \in \mathcal{L}(X)$ because

$$\|FAU\|_{\mathcal{L}(\mathbb{C}^d)} \leq \|F\|_{\mathcal{L}(X, \mathbb{C}^d)}\|A\|_{\mathcal{L}(X)}\|U\|_{\mathcal{L}(\mathbb{C}^d, X)} = \|A\|_{\mathcal{L}(X)}.$$

Since the spectrum depends continuously on the operator on finite dimensional spaces it follows that the mapping $A \mapsto \Sigma(FAU)$ is continuous as well. This yields the assertion. $\square$

Remark A.19. The definition of $S_{\mathrm{att}}(X_1, \ldots, X_d)$ only depends on the norms $\|\cdot\|_{X_i}$ on the factors X_i. Therefore $S_{\mathrm{att}}(X_1, \ldots, X_d)$ as well as the block numerical range is invariant under choice of a different p-norm on the product space X.

It is a well known result that the numerical range of an operator on a Banach space is connected. This is not true in the case of block numerical ranges, however there holds a similar result. For that we first need to state a fact about matrix sets.

Lemma A.20. *If $\mathcal{M} \subset \mathbb{C}^{d\times d}$ is connected, then $\Sigma(\mathcal{M}) := \bigcup_{M\in\mathcal{M}} \Sigma(M)$ consists of at most d connected components. Moreover for every connected component $\Gamma \subset \Sigma(\mathcal{M})$ there exists $n_\Gamma \in \mathbb{N}$ such that every $M \in \mathcal{M}$ has exactly n_Γ eigenvalues in Γ (counting multiplicities).*

Proof. This is [Wag07, Proposition 1.10.(3)]. $\square$

Theorem A.21. *Let $X = X_1 \times \ldots \times X_d$ be Banach spaces and $A \in \mathcal{L}(X)$. Then the block numerical range $\Theta(A; X_1, \ldots, X_d)$ consists of at most d connected components. Furthermore if Θ_j denote the connected components of $\overline{\Theta(A; X_1, \ldots, X_d)}$ for $j = 1, \ldots, m$ with $m \leq d$, there exist integers n_j such that for each $(F, U) \in S_{att}^{op}(X_1, \ldots, X_d)$ exactly n_j eigenvalues (counting multiplicities) of the matrix FAU are contained in Θ_j.*

Proof. The natural map $S_{\mathrm{att}}(X_1,\dots,X_d) \to S_{\mathrm{att}}^{op}(X_1,\dots,X_d)$ is continuous. Additionally we have

$$\|F_0AU_0 - FAU\| = \|F_0AU_0 - F_0AU + F_0AU - FAU\| \leq \|A(U_0-U)\| + \|(F_0-F)A\|$$

for all $(F_0,U_0),(F,U) \in S_{\mathrm{att}}^{op}(X_1,\dots,X_d)$. Therefore the map $(F,U) \mapsto FAU \in \mathbb{C}^{d\times d}$ is continuous. Since $S_{\mathrm{att}}(X_1,\dots,X_d)$ is connected by Lemma A.17 the set $\{FAU \in \mathbb{C}^{d\times d} : (F,U) \in S_{\mathrm{att}}^{op}(X_1,\dots,X_d)\}$ is a connected subset of $\mathbb{C}^{d\times d}$. Both assertions now follow from Lemma A.20. $\square$

Although the above theorem is true for all equivalent norms on X, the actual number of connected components of the block numerical range may differ under different norms as the following example shows:

Example A.22. Let R be the right shift operator on the sequence space ℓ^∞ endowed with the supremums norm $\|(x_i)_i\|_\infty := \sup_{i\in\mathbb{N}} x_i$. The spectrum of R is the unit disc and $|f(Ru)| \leq \|R\|_\infty = 1$ for $(f,u) \in S_{\mathrm{att}}(\ell^\infty)$ thus

$$\{\lambda : |\lambda| < 1\} \subseteq \overline{\Theta(R;\ell^\infty)} \subseteq \{\lambda : |\lambda| \leq 1\}.$$

Therefore the closure of the block numerical range of the block operator

$$A := \begin{pmatrix} R-2.5 & 0 \\ 0 & R \end{pmatrix} \in \mathcal{L}(\ell^\infty \times \ell^\infty)$$

consists of two disjoint discs, i.e.

$$\Theta(A;\ell^\infty,\ell^\infty) = \{\lambda : |\lambda - 2.5| \leq 1\}\dot{\cup}\{\lambda : |\lambda| \leq 1\}.$$

The norm $\|(x_1,x_2,x_3,x_4,\dots)\|_{\mathrm{eq}} := \left\|(\frac{1}{2}x_1, x_2, \frac{1}{2}x_3, x_4,\dots)\right\|_\infty$ is equivalent to $\|\cdot\|_\infty$. Denote ℓ^∞ endowed with this norm by $\ell^\infty_{\mathrm{eq}}$. Define $u := (2,1,0,0,\dots)$ and $f : \ell^\infty_{\mathrm{eq}} \to \mathbb{C}$, $(x_i)_i \to x_2$. Then $\|u\|_{\mathrm{eq}} = \|f\|_{\mathrm{eq}} = f(u) = 1$ that is $(f,u) \in S_{\mathrm{att}}(\ell^\infty_{\mathrm{eq}})$. We have $|f(Ru)| = |f(0,2,1,0,0,\dots)| = 2$. Hence $2 \in \Theta(R;\ell^\infty_{\mathrm{eq}})$ thus $-0.5 \in \Theta(R-2.5;\ell^\infty_{\mathrm{eq}})$. Since $-0.5 \in \Sigma(R) \subseteq \overline{\Theta(R;\ell^\infty)}$ the intersection of the closed numerical ranges of R on ℓ^∞ and $R-2.5$ on $\ell^\infty_{\mathrm{eq}}$ are not empty. Therefore $\overline{\Theta(A;\ell^\infty_{\mathrm{eq}},\ell^\infty)}$ consists of only one component.
While we considered the infinite dimensional space ℓ^∞, note that the example works just as well in finite dimensions (where one then considers a cyclic right shift).

For an operator $A \in \mathcal{L}(X)$ such that its block operator representation $(A_{ij})_{ij}$ is upper or lower triangular it is easily seen that

$$\Theta(A;X_1,\dots,X_d) = \bigcup_{k=1}^{d} \Theta(A_{kk};X_k),$$

so in particular $\Theta(A_{kk};X_k) \subset \Theta(A;x_1,\dots,X_d)$ for $k \in \langle 1,d\rangle$. This is however not true in general (consider for example the matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathcal{L}(\mathbb{C}\times\mathbb{C})$). Under a certain dimensionality condition however the following holds.

Proposition A.23. *Let $X = X_1 \times \ldots \times X_d$ such that $\dim X_i \geq d$ for all $i \in \langle 1, d\rangle$. Then*

$$\Theta(A_{kk}; X_k) \subset \Theta(A; X_1, \ldots, X_d) \quad \forall k \in \langle 1, d\rangle.$$

In particular $\Theta(A_{kk}; X_k)$ is contained in a connected component of $\Theta(A; X_1, \ldots, X_d)$.

Proof. Due to Lemma A.3 we can without loss of generality assume that $k = 1$. Now let $\lambda \in \Theta(A_{11}; X_1)$, so there exists $(f_1, u_1) \in S_{\text{att}}(X_1)$ such that $f_1(A_{11}u_1) = \lambda$. We will now recursively find pairs $(f_i, u_i) \in S_{\text{att}}(X_i)$ for $i = 2, \ldots, d$ in the following way:
Define the space $V_i := \operatorname{span}\{A_{ij}u_j : j = 1, \ldots, i-1\} \subset X_i$. Because of the dimension condition we find a vector $x_i \in X_i$ such that $x_i \notin V_i$. Now consider the space $U_i := \operatorname{span}(V_i \cup \{x_i\})$. Since U_i is finite dimensional we now find a normed $u_i \in U_i$ with $u_i \notin V_i$ such that there exists a normed $\hat{f}_i \in U_i'$ such that $\hat{f}_i(V_i) = \{0\}$ and $\hat{f}_i(u_i) = 1$. By the Hahn-Banach theorem we can extend $\hat{f}_i$ to a norm preserving $f_i \in X_i'$ which yields the desired pair $(f_i, u_i) \in S_{\text{att}}(X_i)$.
Setting $f = (f_1, \ldots, f_d)^T$ and $u = (u_1, \ldots, u_d)^T$ we obtain $(f, u) \in S_{\text{att}}(X_1, \ldots, X_d)$. The corresponding $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ now has the property that FAU is an upper triangular matrix. Therefore

$$\Sigma(FAU) = \{f_i(A_{ii}u_i) : i \in \langle 1, d\rangle\}$$

and thus $\lambda = f_1(A_{11}u_1) \in \Theta(A; X_1, \ldots, X_d)$.
The last assertion follows since $\Theta(A_{kk}; X_k)$ is connected. □

The aquired knowledge on connected components allows us to estimate the resolvent of an operator $A \in \mathcal{L}(X)$ in analogy to [Tre08]. However we first need the following lemma which corresponds to [Tre08, Lemma 1.11.16].

Lemma A.24. *Let $X = X_1 \times \cdots \times X_d$ be Banach spaces and $A \in \mathcal{L}(X)$. If there exists a $\delta > 0$ such that for every $(F, U) \in S_{att}^{op}(X_1, \ldots, X_d)$*

$$\|FAUv\| \geq \delta\|v\| \qquad \forall v \in \mathbb{C}^n,$$

then

$$\|Ax\| \geq \delta\|x\| \qquad \forall x \in X.$$

Furthermore if A is invertible then $\|A^{-1}\| \leq \frac{1}{\delta}$.

Proof. Let $x = (x_1, \ldots, x_d)^T \in X$. Define $v = (\nu_1, \ldots, \nu_d)^T \in \mathbb{C}^n$ through $\nu_i := \|x_i\|$ and $u = (u_1, \ldots, u_d)^T \in X$ such that $\|u_i\| = 1$ and $\|x_i\|u_i := x_i$. Pick an $f \in X'$ such that $(f, u) \in S_{\text{att}}(X_1, \ldots, X_d)$ and let $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ be the corresponding operators. Now obviously $Ax = AUv$ and $\|x\|_X = \|v\|_{\mathbb{C}^n}$. We get

$$\|Ax\| = \underbrace{\|F\|}_{=1}\|Ax\| \geq \|FAx\| = \|FAUv\| \geq \delta\|v\| = \delta\|x\|.$$

Let now A be invertible. Then for all $x \in X$ by the above

$$\delta\|A^{-1}x\| \leq \|AA^{-1}x\| = \|x\|.$$

Supremum over all normed $x \in X$ yields the assertion. □

Theorem A.25. *Let $X = X_1 \times \cdots \times X_d$ be Banach spaces and $A \in \mathcal{L}(X)$. Let $\Theta_1, \ldots \Theta_m$ be the connected components of $\overline{\Theta(A; X_1, \ldots X_d)}$. There exist integers n_j for $j = 1, \ldots, m$ as in Theorem A.21 with $\sum_{j=1}^{m} n_j = d$ such that for every $\lambda \notin \overline{\Theta(A; X_1, \ldots X_d)}$ the inequality*

$$\|(A - \lambda)^{-1}\| \leq \frac{\|A - \lambda\|^{d-1}}{\prod_{j=1}^{m} \operatorname{dist}(\lambda, \Theta_j)^{n_j}}.$$

holds.

Proof. We first note that $A - \lambda$ as well as $FAU - \lambda$ for $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ are invertible for $\lambda \notin \overline{(A; X_1, \ldots X_d)}$. We choose the integers n_j as in Theorem A.21 so that for each $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ the matrix FAU has exactly n_j eigenvalues in Θ_j. For fixed $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$ let $\mu_1, \ldots, \mu_d$ be the eigenvalues of FAU. Then

$$|\det(FAU - \lambda)| = \prod_{i=1}^{d} |\mu_i - \lambda| \geq \prod_{j=1}^{m} \operatorname{dist}(\lambda, \Theta_j)^{n_j} > 0.$$

Together with Lemma A.12 this yields

$$\|(FAU - \lambda)^{-1}\| \leq \frac{\|FAU - \lambda\|^{d-1}}{|\det(FAU - \lambda)|} \leq \frac{\|A - \lambda\|^{d-1}}{\prod_{j=1}^{m} \operatorname{dist}(\lambda, \Theta_j)^{n_j}}$$

for all $(F, U) \in S_{\text{att}}^{op}(X_1, \ldots, X_d)$. The assertion now follows from Lemma A.24. $\square$

At the end of this section we give two consequences of Theorem A.25.

Corollary A.26. *Let $X = X_1 \times \cdots \times X_d$ be Banach spaces and $A \in \mathcal{L}(X)$. for every $\lambda \notin \overline{\Theta(A; X_1, \ldots X_d)}$ the inequality*

$$\|(A - \lambda)^{-1}\| \leq \frac{\|A - \lambda\|^{d-1}}{\operatorname{dist}(\lambda, \Theta(A; X_1, \ldots, X_d))^d}.$$

holds.

Corollary A.27. *Let X be a trivially decomposed Banach space and $A \in \mathcal{L}(X)$. For every $\lambda \notin \overline{\Theta(A; X)}$ the inequality*

$$\|(A - \lambda)^{-1}\| \leq \frac{1}{\operatorname{dist}(\lambda, \Theta(A; X))}$$

holds.

A.5. The block numerical range of positive matrices

When viewing the block numerical range of a matrix $A \in \mathbb{C}^{n,n}$ the question arises how to decompose the space $\mathbb{C}^n$. It seems natural to take some $n_1, \ldots, n_p \in \mathbb{N}$ such that $\sum n_i = n$. Then

$$\mathbb{C} = \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_p}.$$

When dealing with an element $x \in \mathbb{C} = \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_p}$ it is useful to be able to keep track of the product spaces a coordinate of the vector x lies in. The next definition makes this possible.

Definition A.28. Let $\mathbb{C} = \mathbb{C}^{n_1} \times \ldots \times \mathbb{C}^{n_p}$. Then define the integer sets

$$\mathcal{I}_k := \left\{ \sum_{j=0}^{k-1} n_j + 1, \sum_{j=0}^{k-1} n_j + 2, \ldots, \sum_{j=0}^{k} n_j - 1, \sum_{j=0}^{k} n_j \right\}, \quad k = 1, \ldots, p, \tag{A.4}$$

where we set $n_0 = 0$ and the functions

$$\iota : \langle n \rangle \to \langle p \rangle : i \mapsto k \quad \text{iff} \quad i \in \mathcal{I}_k. \tag{A.5}$$

A vector $x \in \mathbb{C}^n$ and A matrix $A \in \mathbb{C}^{n,n}$ can then be written in block form as

$$x = \begin{bmatrix} x_{\mathcal{I}_1} \\ \vdots \\ x_{\mathcal{I}_p} \end{bmatrix}, \qquad A = \begin{bmatrix} A_{\mathcal{I}_1,\mathcal{I}_1} & \ldots & A_{\mathcal{I}_1,\mathcal{I}_p} \\ \vdots & \ddots & \vdots \\ A_{\mathcal{I}_p,\mathcal{I}_1} & \ldots & A_{\mathcal{I}_p,\mathcal{I}_p} \end{bmatrix}$$

where $x_{\mathcal{I}_k} = (x_i)_{i \in \mathcal{I}_k}$ and $A_{\mathcal{I}_h,\mathcal{I}_k} \in \mathbb{C}^{n_h,h_k}$, $h, k \in \langle p \rangle$.
As in the Hilbert space case the set $S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ has a simplified form because by Riesz's lemma for any tuple $(f, x) \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ there holds $(f, x) = (x, x)$. In the matrix case we will therefore simply write $x \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ and for its associated operator $X \in S_{\text{att}}^{op}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$. It is instructive to write the definition down explicitly:

$$\begin{aligned} S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p}) &= \{(x_i) \in \mathbb{C}^n : x_{\mathcal{I}_k} \in S_{\text{att}}(\mathbb{C}^{n_k}), k = 1, \ldots, p\} \\ &= \{(x_i) \in \mathbb{C}^n : \|x_{\mathcal{I}_k}\|_{\mathbb{C}^{n_k}} = 1, k = 1, \ldots, p\}. \end{aligned}$$

For $x \in S_{\text{att}}(\mathbb{C}^{n_1}, \ldots, \mathbb{C}^{n_p})$ the associated operator $X : \mathbb{C}^n \to \mathbb{C}^p$ then has the matrix representation

$$X = \begin{bmatrix} x_{\mathcal{I}_1} & & & \\ & x_{\mathcal{I}_2} & & \\ & & \ddots & \\ & & & x_{\mathcal{I}_p} \end{bmatrix} \in \mathbb{C}^{n,p}.$$

For a matrix $A = (a_{rs})$ and a vector $u \in \mathbb{C}^p$ we then have

$$\begin{aligned} X^* A X u &= \begin{bmatrix} x_{\mathcal{I}_1}^* A_{\mathcal{I}_1,\mathcal{I}_1} x_{\mathcal{I}_1} & \ldots & x_{\mathcal{I}_1}^* A_{\mathcal{I}_1,\mathcal{I}_p} x_{\mathcal{I}_p} \\ \vdots & \ddots & \vdots \\ x_{\mathcal{I}_p}^* A_{\mathcal{I}_p,\mathcal{I}_1} x_{\mathcal{I}_1} & \ldots & x_{\mathcal{I}_p}^* A_{\mathcal{I}_p,\mathcal{I}_p} x_{\mathcal{I}_p} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_p \end{bmatrix} \\ &= \left[\sum_{k=0}^{p} \sum_{r \in \mathcal{I}_h} \sum_{s \in \mathcal{I}_k} \bar{x}_r A(r,s) x_s u_k \right]_{h=1,\ldots,p} \end{aligned}$$

Remark A.29. This type of decomposition of $\mathbb{C}^n$ is restrictive in the sense that the coordinates are grouped strictly sequentially. One might want to group them in any order as in Example A.4.(ii), i.e. for some disjoint integer sets $\mathcal{J}_k$, $k = 1, \dots, p$ such that $\bigcup \mathcal{J}_k = \langle n \rangle$ and calculate the block numerical range of a matrix A in the Hilbert space sense (this is the modus operandi in [FH08]).
However, one can simply take a permutation $\pi(\cdot)$ on $\mathbb{C}^n$ such that the coordinates of the permuted index sets

$$\mathcal{J}_k^{\pi} := \{\pi(i) : i \in \mathcal{J}_k\}$$

are then grouped sequentially. The block numerical range is then calculated for the permuted matrix $P^T AP$ (where P is the permutation matrix associated with $\pi(\cdot)$).

We will now turn to entrywise nonnegative matrices. Note that for $x \in S_{\mathrm{att}}(\mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p})$ the condition $x \gg 0$ (i.e. x is entrywise strictly nonnegative) does not imply that the associated matrix $X \in S^{op}_{\mathrm{att}}(\mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p})$ is also entrywise strictly nonnegative (though $X \geq 0$ still holds). It is therefore useful to use both representations.

Proposition A.30. *Let $A \in \mathbb{R}^{n,n}_+$ be an entrywise nonnegative matrix. Then*

(i) $\mathrm{bnr}(A; \mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p}) \in \Theta(A; \mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p})$.

If A is additionally irreducible and $x \in S_{att}(\mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p})$, $x \geq 0$ with corresponding $X \in S^{op}_{att}(\mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p})$, then

(ii) if $x \gg 0$ then $X^*AX \in \mathbb{R}^{p,p}_+$ is irreducible,

(iii) if $\mathrm{bnr}(A; \mathbb{C}^{n_1}, \dots, \mathbb{C}^{n_p}) = \mathrm{spr}(X^*AX)$ *then $x \gg 0$.*

Proof. (i) is [FH08, Proposition 3.1.2], (ii) and (iii) are [FH08, Proposition 4.1]. $\square$

References

[AHKM03] T. Y. Azizov, V. Hardt, N. D. Kopachevsky, and R. Mennicken. On the problem of small motions and normal oscillations of a viscous fluid in a partially filled container. *Mathematische Nachrichten*, 248/249:3–39, 2003.

[Ale11] E. A. Alekhno. The irreducibility in ordered Banach algebras. *Positivity*, 16:143 – 176, 2011.

[BD71] F. F. Bonsall and J. Duncan. *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, volume 2 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, London, 1971.

[BD73] F. F. Bonsall and J. Duncan. *Numerical Ranges II*, volume 10 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, London, 1973.

[BJW85] G. J. Butler, C. R. Johnson, and H. Wolkowicz. Nonnegative Solutions of a Quadratic Matrix Equation Arising from Comparison Theorems in Ordinary Differential Equations. *SIAM J. Alg. Disc. Meth.*, 6:47–53, 1985.

[BLM02] D. A. Bini, G. Latouche, and B. Meini. Solving Matrix Polynomial Equations Arising in Queueing Problems. *Linear Algebra Appl.*, 340:225–244, 2002.

[BNS13] D. Bini, V. Noferini, and M. Sharify. Locating the Eigenvalues of Matrix Polynomials. *SIAM. J. Matrix Anal. & Appl.*, 34(4):1708–1727, 2013.

[BP79] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Computer Sciences and Applied Mathematics. Academic Press, New York, 1979.

[BR91] R. A. Brualdi and H. J. Ryser. *Combinatorial Matrix Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1991.

[BR97] R. B. Bapat and T. E. S. Raghavan. *Nonnegative Matrices and Applications*, volume 64 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1997.

[CG81] K. F. Clancey and I. Gohberg. *Factoriation of Matrix Functions and Singular Integral Operators*, volume 3 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel and Boston, 1981.

[Cha61] S. Chandrasekhar. *Hydrodynamic and hydromagnetic stability*. The International Series of Monographs on Physics. Clarendon Press, Oxford, 1961.

[EN00] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Grad. Texts in Math.* Springer, New York, 2000.

[FH08] K.-H. Förster and N. Hartanto. On the Block Numerical Range of Nonnegative Matrices. *Operator Theory: Advances and Applications*, 188:113–133, 2008.

[FK15] K.-H. Förster and P. Kallus. Perron Frobenius Theorems for the Numerical Range of Semi-Monic Matrix Polynomials. *Linear Algebra an its Applications*, 2015.

[FN05a] K.-H. Förster and B. Nagy. On Nonmonic Quadratic Matrix Polynomials with Nonnegative Coefficients. *Operator Theory: Advances and Applications*, 162:145–163, 2005.

[FN05b] K.-H. Förster and B. Nagy. Spectral Properties of Operator Polynomials with Nonnegative Coefficients. *Operator Theory: Advances and Applications*, 163:147–162, 2005.

[FN15] K.-H. Förster and B. Nagy. On Operator Polynomials with Nonnegative Coefficients in a Banach Lattice. *Annales Univ. Sci. Budapest*, 58:37–54, 2015.

[GGK93] I. Gohberg, S. Goldberg, and A. Kaashoek. *Classes of Linear Operators Vol. II*, volume 63 of *Operator Theory: Advances and Applixations*. Birkhäuser, Basel, 1993.

[GHT96] H. R. Gail, S. L. Hantler, and B. A. Taylor. Spectral Analysis of $M/G/1$ and $G/M/1$ Type Markov Chains. *Adv. Appl. Prob.*, 28:114–165, 1996.

[GHT98] H. R. Gail, S. L. Hantler, and B. A. Taylor. Matrix-Geometric Invariant Measures for $G/M/1$ Type Markov Chains. *Commun. Statist. - Stochastic Models*, 14(3):537–569, 1998.

[GKS03] I. Gohberg, M. A. Kaashoek, and I. M. Spitkovsky. An Overview of Matrix Factorization Theory and Operator Applications. *Operator Theory: Advances and Applications*, 141:1 – 102, 2003.

[GLR82] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Computer Science and Applied Mathematics. Academic Press, New York, 1982.

[GLW09] H.-b. Guo, X.-g. Liu, and W.-g. Wang. The block numerical range of matrix polynomials. *Applied Mathemtics and Computation*, 211:265–283, 2009.

[GR97] K. E. Gustafson and D. K. M. Rao. *Numerical Range*. Springer, New York, 1997.

[Har11] N. V. Hartanto. Spectral Properties and Companion Forms of Operator- and Matrix Polynomials. PhD thesis, TU Berlin, 2011. unpublished.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[HJ91] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Sigma Series in Applied Mathematics. Cambridge University Press, Cambridge, 1991.

[HS78] P. R. Halmos and V. S. Sunder. *Bounded Integral Operators on L^2 Spaces.* Springer, Berlin, 1978.

[Iss66] J. N. Issos. The Field of Values of non-Negative Irreducible Marices. PhD thesis, Auburn University, 1966. unpublished.

[Kal11] P. Kallus. The Block Numerical Range of Operator Functions in Banach Spaces. Bachelors thesis, TU Berlin, 2011. unpublished.

[Kat04] Y. Katznelson. *An Introduction to Harmonic Analysis.* Cambridge University Press, Cambridge, third edition, 2004.

[KK01] N. D. Kopachevsky and S. G. Krein. *Operator Approach to Linear Problems of Hydrodynamics Vol. 1*, volume 128 of *Operator Theory: Advances and Applications.* Birkhäuser Verlag, Basel, 2001.

[KLS89] M. A. Krasnosel'skij, Je. A. Lifshits, and A. V. Sobolev. *Positive Linear Systems: The Method of Positive Operators*, volume 5 of *Sigma Series in Applied Mathematics.* Helderman, Berlin, 1989.

[KMM07] V. V. Kostrykin, K. A. Makarov, and A. K. Motolivov. Pertubation of spectra and spectral subspaces. *Transactions of the American Mathematical Society*, 359(1):77–89, 2007.

[Kuo77] I. Kuo. A Note on Factorizations of Singular M-Matrices. *Linear Algebra and its Applications*, 16:217 – 220, 1977.

[Lif89] A. E. Lifschitz. *Magnetohydrodynamics and Spectral Theory*, volume 4 of *Developments in Electromagnetic Theory and Applications.* Kluver Academic Publishers Group, Dordrecht, 1989.

[LMMT01] H. Langer, A. S. Markus, V. I. Matsaev, and C. Tretter. A New Concept for Block Operator Matrices: The Quadratic Numerical Range. *Linear Algebra Appl.*, 330(1-3):89–112, 2001.

[LŠ10] T. J. Laffey and H. Šmigoc. Nonnegatively realizable spectra with two positive eigenvalues. *Linear and Multilinear Algebra*, 58(8):1053–1069, 2010.

[Mar88] A. S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils*, volume 71 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, Rhode Island, 1988.

[Mel13] A. Melman. Generalization and variations of Pellet's theorem for matrix polynomials. *Linear algebra and its applications*, 439(5):1550–1567, 2013.

[Min88] H. Minc. *Nonnegative Matrices.* Wiley-Intersection Series in Discrete Mathematics and Optimization. Wiley, New York, 1988.

[MN91] P. Meyer-Nieberg. *Banach Lattices.* Springer, New York and Heidelberg, 1991.

[MPT02] J. Maroulas, P. J. Psarrakos, and M. J. Tsatsomeros. Perron–Frobenius type results on the numerical range. *Linear Algebra and its Applications*, 348:49–62, 2002.

[MX13] V. Mehrmann and H. Xu. Extraction of structured linear eigenvalue problems from structured polynomial eigenvalue problems. private communication, 2013.

[Nag07] B. Nagy. On the degree reduction of matrix polynomials. private communication, 2007.

[Neu89] M. F. Neuts. *Structured stochastic matrices of $M/G/1$ type and their applications*, volume 5 of *Probability: Pure and Applied.* Marcel Decker Inc., New York, 1989.

[Nic00] C. P. Niculescu. Convexity according to the geometric mean. *Mathematical Inequalities & Applications*, 3(2):155 – 167, 2000.

[PT04] P. J. Psarrakos and M. J. Tsatsomeros. A Primer of Perron-Frobenius theory for matrix polynomials. *Linear Algebra and its Applications*, 393:333–351, 2004.

[Rad13] A. Radl. The numerical range of positive operators on Hilbert lattices. *Integral Equations Operator Theory*, 75(4):459–472, 2013.

[Rad14] A. Radl. Perron-Frobenius type results for the block numerical range. *Proc. Appl. Math. Mech.*, 14:1001–1002, 2014.

[Rad15a] A. Radl. Contributions to the numerical range and some applications of semigroup theory. Habilitation, 2015. unpublished.

[Rad15b] A. Radl. The numerical range of positive operators on Banach lattices. *Positivity*, 19:603–623, 2015.

[Rod89] L. Rodman. *An Introduction to Operator Polynomials*, volume 38 of *Operator Theory: Advances and Applications.* Birkhäuser Verlag, Basel, 1989.

[RR89] H. Raubenheimer and S. Rode. Cones in Banach Algebras. *Indag. Math. (N. S.)*, 7(4):489 – 502, 1989.

[RTW14] A. Radl, C. Tretter, and M. Wagenhofer. The Block Numerical Range of Analytic Operator Functions. *Operators and Matrices*, 8(4):901–934, 2014.

[SAD$^+$00] Y. Saad, O. Axelsson, I.S. Duff, W.P. Tang, H.A. van der Vorst, and A. Wathen. Special Issue: Preconditioning Techniques for Large Sparse Matrix Problems in Industrial Applications, SPARSE '99. *Numerical Linear Algebra with Applications*, 7, 2000.

[Sch74] H. H. Schaefer. *Banach Lattices and Positive Operators.* Springer, Berlin and New York, 1974.

[SW10] J. Swoboda and H. K. Wimmer. Spectraloid operator polynomials, the approximate numerical range and an Eneström-Kakeya theorem in Hilbert space. *Studia Math.*, 198:279–300, 2010.

[TDD] F. De Teran, F. M. Dopico, and P. Van Dooren. Constructing strong l-ifications from dual minimal bases. submitted.

[TDD15] F. De Teran, F. M. Dopico, and P. Van Dooren. Matrix Polynomials with Completely Prescribed Eigenstructure. *SIAM Journal on Matrix Analysis and Applications*, 36:302–328, 2015.

[TDM14] F. De Teran, F. M. Dopico, and D. S. Mackey. Spectral equivalence of matrix polynomials and the Index Sum Theorem. *Linear Algebra and its Applications*, 459:264–333, 2014.

[Tha92] B. Thaller. *The Dyrac Equation.* Texts and Monographs in Physics. Springer, Berlin, 1992.

[TL80] A. E. Taylor and D. C. Lay. *Introduction to Funtional Analysis.* John Wiley & Sons, New York, second edition, 1980.

[Tre08] C. Tretter. *Spectral Theory of Block Operator Matrices and Applications.* Imperial College Press, London, 2008.

[TW03] C. Tretter and M. Wagenhofer. The Block Numerical Range of an $n \times n$ Block Operator Matrix. *SIAM J. Matrix Anal. Appl.*, 24(4):1003–1017, 2003.

[Wag07] M. Wagenhofer. Block Numerical Ranges. PhD thesis, Fachbereich 3 - Mathematik und Informatik der Universität Bremen, 2007. unpublished.

[Wim08] H. K. Wimmer. Polynomial matrices with hermitian coefficients and a generalization of the Eneström-Kakeya theorem. *Oper. Matrices*, 2:443–454, 2008.

[Wim11] H. K. Wimmer. Matrix polynomials with spectral radius equal to the numerical radius. *J. Math. Anal. Appl.*, 381:80–86, 2011.

[Wnu99] W. Wnuk. *Banach Lattices with Order Continuous Norms.* Polish Scientific Publishers PWN, Warsaw, 1999.