Eigenvalues of Matrices and Graphs

Der Fakultät für Mathematik und Informatik
der Universität Leipzig
eingereichte

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr.rer.nat.)

im Fachgebiet

M A T H E M A T I K

vorgelegt von

Diplomphysiker Mario Thüne, M.A.
geboren am 21.06.1979 in Leinefelde

Die Annahme der Dissertation wurde empfohlen von:

1. Prof. Dr. Jürgen Jost (MPI-MiS, Leipzig, Deutschland)
2. Prof. Dr. Leonid Bunimovich (Georgia Tech, Atlanta, USA)

Die Verleihung des akademischen Grades erfolgt auf Beschluss des Rates der Fakultät
für Mathematik und Informatik vom 27.02.2013 mit dem Gesamtprädikat cum laude.

Acknowledgement

I would like to thank my supervisor Prof. Dr. Jürgen Jost for his support, his confidence and his patience and first of all for some profound and inspiring lectures. I would also like to thank my colleagues and the staff at the MPI-MIS in Leipzig for their help and kindness.

There are many other people who have supported me in various ways. To all of them I would like to express my sincere gratitude.

My final thanks are reserved for my family, especially Linda and Hanna.

Ich widme diese Arbeit meinen Großeltern

(dedicated to my grandparents)

Zusammenfassung

Das Zusammenspiel von Spektrum und Struktur von Graphen ist das Thema der drei Kapitel dieser Arbeit.

Im ersten Kapitel werden die Eigenwerte von zwei komplexwertigen Matrizen, von denen eine die prinzipale Submatrix der anderen ist, in Beziehung gesetzt mit Hilfe eines annihilierenden Polynoms. Dies wird erweitert zu Matrizen, deren Einträge Polynome oder rationale Funktionen in Lambda sind. Die Erweiterung kann zudem als Generalisierung anderer bereits bekannter Techniken, die die Verkleinerung einer Matrix ohne Verlust der spektralen Information zum Ziel haben, aufgefasst werden. Einige Aspekte, die für die mögliche Anwendung zur Verringerung des numerischen Aufwands gewöhnlicher Eigenwertprobleme relevant sind, werden diskutiert.

Das zweite Kapitel betrachtet die Erweiterung des bekannten Konzepts der Equitable Partition auf komplexe Matrizen. Es enthält eine Methode, um ein Eigenproblem in die zwei (im nicht trivialen Fall) kleineren unabhängigen Eigenprobleme des Frontdivisors und des dazu komplementären Faktors zerlegt. Das Verfahren ist einfach, stabil und hat eine Komplexität von $O\left(N^2\right)$. Für die Behandlung mehrerer solcher Zerlegungen, die geordnet sind im Sinne des "feiner als", wird eine Möglichkeit vorgeschlagen vorhandene Hermitizität zu erhalten. Mögliche Generalisierungen der Equitable Partition werden betrachtet und ein einfaches Verfahren zum Auffinden einer Equitable Partition einer komplexen Matrix wird angegeben.

Das dritte Kapitel behandelt isospektrale und unitär äquivalente Graphen. Vorgestellt wird eine einfache Methode zur Erzeugung von Paaren von unitär äquivalenten Graphen, die sich als eine Verallgemeinerung des bekannten GM-Switchings auffassen lässt. Weiterhin wird eine Algebra betrachtet, die von der Adjazenzmatrix erzeugt wird und in einer ähnlichen Beziehung zur 1-dimensionalen Weisfeiler-Lehman-Methode steht wie die von dieser Matrix erzeugte Algebra kohärenter Konfigurationen zur 2-dimensionalen. Diese Algebra enthält unter anderem die Gradmatrix, verschiedene Laplace-Matrizen und die Seidel-Matrix des Graphen. In einfacher Weise können Graphpaare erzeugt werden, die bezüglich dieser Algebra simultan unitär äquivalent sind.

Abstract

The interplay between spectrum and structure of graphs is the recurring theme of the three more or less independent chapters of this thesis.

The first chapter provides a method to relate the eigensolutions of two matrices, one being the principal submatrix of the other, via an arbitrary annihilating polynomial. This is extended to lambda-matrices and to matrices the entries of which are rational functions in one variable. The extension may be interpreted as a possible generalization of other known techniques which aim at reducing the size of a matrix while preserving the spectral information. Several aspects of an application in order to reduce the computational costs of ordinary eigenvalue problems are discussed.

The second chapter considers the straightforward extension of the well known concept of equitable partitions to weighted graphs, i.e. complex matrices. It provides a method to divide the eigenproblem into smaller parts corresponding to the front divisor and its complementary factor in an easy and stable way with complexity $O\left(N^2\right)$. The exploitation of several equitable partitions ordered by refinement is discussed and a suggestion is made that preserves hermiticity if present. Some generalizations of equitable partitions are considered and a basic procedure for finding an equitable partition of complex matrices is given.

The third chapter deals with isospectral and unitary equivalent graphs. It introduces a construction for unitary equivalent graphs which contains the well known GM-switching as a special case. It also considers an algebra of graph matrices generated by the adjacency matrix that corresponds to the 1-dimensional Weisfeiler-Lehman stabilizer in a way that mimics the correspondence of the coherent closure and the 2-dimensional Weisfeiler-Lehman stabilizer. The algebra contains the degree matrix, the (combinatorial, signless and normalized) Laplacian and the Seidel matrix. An easy construction produces graph pairs that are simultaneously unitary equivalent w.r.t. that algebra.

Contents

Chapter 0

Introduction

The motivation for this thesis is the utilization of the entire spectrum of matrices as a tool for graph and network analysis. Two aspects will be considered. The first one is the wish for exploiting structure in order to save computational costs in the case of large networks. The second aspect is the determination of a network by the spectrum of its representing matrices.

Our main interest here are matrices arising from empirical networks. The spectral investigation of those usually employs only selected parts of the spectrum, prominently the few largest (in modulus) eigenvalues as in the various forms of spectral clustering since those are computable with moderate effort. Using for instance variants of the power iteration, one may additionally exploit sparsity to find those few eigenvalues. The theoretical analysis of large graphs usually assumes some global structure, e.g. a factorization as a graph product, or considers specific classes of graphs e.g. trees. A complementary approach considers ensembles of random graphs. With those assumptions the whole spectrum can often be described. The spectra of empirical networks seem to lie somewhere in between. The deviation even from slightly advanced theoretical structures is often to far to exploit those for spectral computation. On the other hand, there is typically structure on a smaller scale [4], [5], which indicates that spectral analysis may reveal important structure. However those structures typically cause large eigenspaces for certain eigenvalues, which is a problem for iterative methods for the computation of spectra like the power iteration

It is well known that many structural features must be shared by two graphs having the same spectrum, for instance the number of vertices, edges and triangles or more generally the number of walks of any given length. Although a network is in general not determined by its spectrum, the fractions of graphs sharing the same spectrum is believed to be rather small for large graphs [79], [36].

The first two chapters consider two ways to derive the spectrum of a matrix using structural information. One aims at controlling the change of the spectrum of a given matrix after adding or removing some vertices and the other exploits the spectral relation of a matrix and a quotient obtained of it by a careful coarsening of its vertex set. The third chapter considers the equivalence of graphs w.r.t. matrix representations and their spectra.

Notational Remark: Eigenvalues, Roots and Multiplicities Let $f(\lambda)$ be a complex valued function and let $\mathbf{X}$ be a complex square matrix. By $\rho(f(\lambda))$ we denote the multiset of roots of $f(\lambda)$ and by $\sigma(\mathbf{X})$ we denote the multiset of eigenvalues of $\mathbf{X}$. We frequently use the identity $\sigma(\mathbf{X}) = \rho(\det(\mathbf{X} - \lambda\mathbf{I}))$. Let $\alpha = \{x, x, x, y, y, z\}$ and $\beta = \{x, y\}$ be multisets. We indicate the multiplicity by superscripts and use the value 0 to distinguish an object which is not an element, i.e. $\beta = \{x^1, y^1, z^0\}$. We join two multisets by adding up the multiplicities, i.e. $\alpha + \beta = \{x^4, y^3, z^1\}$, and we remove a subset by subtraction, i.e. $\alpha - \beta = \{x^2, y^1, z^1\}$. When we add up n copies of the same multiset γ we may abbreviate with γ^n, i.e. $\alpha + \alpha = \alpha^2$. To denote the multiplicity of an element c in a multiset γ we use $m_c(\gamma)$. When we refer to an element λ in the spectrum of a square matrix $\mathbf{X}$, we may abbreviate $m_\lambda(\sigma(\mathbf{X}))$ with $m_\lambda(\mathbf{X})$.

Chapter 1

Perturbation of Matrices by Adding or Removing Nodes

In this chapter we aim at a relation between the spectrum of a matrix and the spectrum of one of its principal submatrices using an annihilating polynomial for one of the two and the spectrum of a third matrix which is derived thereof. The size of the third matrix is the product of the difference in size of the first and second one and the incremented degree of the polynomial. It is intrinsically expressible as a lambda-matrix, which allows for some flexibility in computing its spectrum. The technique works for complex matrices and is easily generalized to complex lambda-matrices. It utilizes a generalization of the adjoint matrix in order to apply the method of Schur complements to lambda-matrices.

1.1 Introduction

Let $\mathbf{A}$ and $\mathbf{M}$ be matrices of size N and $N + 1$, respectively, s.t. $\mathbf{A}$ is a *principal submatrix* of $\mathbf{M}$ obtained by removing a row and a column with the same index. In the interpretation of weighted adjacency matrices of graphs $\mathbf{A}$ is obtained by deleting a node of $\mathbf{M}$. One might equivalently say that $\mathbf{M}$ is obtained by adding a node to $\mathbf{A}$. Let $\lambda_1 \leq \ldots \leq \lambda_N$ and $\mu_1 \leq \ldots \leq \mu_{N+1}$ be the eigenvalues of $\mathbf{A}$ and $\mathbf{M}$, respectively. For the hermitian case the Cauchy interlacing theorem states that $\mu_i \leq \lambda_i \leq \mu_{i+1}$, $i = 1, \ldots, N$, and by induction it is easy to see that

$$\mu_i \leq \lambda_i \leq \mu_{i+s}, \; i = 1, \ldots, N \tag{1.1}$$

holds for hermitian matrix $\mathbf{M}$ of size $N + s$ with eigenvalues $\mu_1 \leq \ldots \leq \mu_{N+s}$ and $\mathbf{A}$ obtained by deleting s accordingly indexed rows and columns of $\mathbf{M}$ [45], [35]. Unfortunately, sound generalizations of (1.1) are not straightforward. They can be done, for instance, by reformulation as a sufficient and necessary condition for an imbedding problem on normal matrices considering suitable orderings on $\mathbb{C}$, e.g. [27], [75], [63]. However, orderings necessary to generalize (1.1) as a relation between spectra of arbitrary matrices and their submatrices might be impracticable. In [55] (1.1) is generalized by restriction to the class of totally nonnegative matrices.
By (1.1) we have the following

Corollary 1. *Let $\mathbf{A}$ and $\mathbf{M}$ be hermitian matrices s.t. $\mathbf{A}$ is a principal submatrix of $\mathbf{M}$. Then $|m_\lambda(\mathbf{A}) - m_\lambda(\mathbf{M})| \leq s \; \forall \lambda \in \mathbb{R}$.*

Note the symmetry in $\mathbf{A}$ and $\mathbf{M}$. (1.1) and corollary 1 provide bounds on eigenvalues and multiplicities. For hermitian matrices some part of the spectrum of $\mathbf{M}$ (or $\mathbf{A}$) may already be determined if some eigenvalues of $\mathbf{A}$ (or $\mathbf{M}$) have high multiplicity and the number of added (removed) nodes is small. Further information may determine the spectrum completely.

Motivated by those results we proved a generalization of corollary 1 for general complex matrices and developed a method that allows to determine the spectrum of a complex matrix $\mathbf{M}$ by the spectrum of a principal submatrix $\mathbf{A}$, the roots of an annihilating polynomial for $\mathbf{A}$ and the spectrum of a lambda-matrix of size less than $\mathbf{M}$. In a similar way the spectrum of $\mathbf{M}$ together with a scalar polynomial and a matrix polynomial determines the spectrum of $\mathbf{A}$. The determination is not only for a fraction or the distinct values but the complete multiset of eigenvalues. Since a version of the adjoint of a matrix is utilized we refer to it as an *adjoint technique*. We give a short outline of the main idea.

Assume that the spectrum of $\mathbf{A}$ is known and that we are given an *annihilating polynomial* $\mathrm{a}(x)$ for it, i.e. $\mathrm{a}(\mathbf{A}) = 0$, for instance the characteristic or the minimal polynomial. From $\mathrm{a}(x)$ we derive a polynomial $\mathrm{p}(x, y)$ s.t.

$$\mathrm{p}(\lambda, \mathbf{A})(\lambda \mathbf{I}_N - \mathbf{A}) = \mathrm{a}(\lambda)$$

as described below. The node set of $\mathbf{M}$ of size $(N + s)$ is partitioned into two parts whereas the first one of size N induces $\mathbf{A}$. This yields a partition of $\mathbf{M} = \left(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix} \right)$ into a 2×2 block matrix. Using $\mathrm{a}(x)$, $\mathrm{p}(x, y)$ and the partition of $\mathbf{M}$ we obtain the matrix

$$\mathcal{L}(\lambda) = \mathbf{C}\,\mathrm{p}(\mathbf{A}, \lambda)\,\mathbf{B} + \mathrm{a}(\lambda)(\mathbf{D} - \lambda \mathbf{I}_s)$$

of size $s \times s$ the entries of which are polynomials in λ. According to theorem 2 below the spectrum of $\mathbf{M}$ is completely determined by

$$\sigma(\mathbf{M}) = \sigma(\mathbf{A}) - \rho^s(\mathrm{a}) + \rho(\det(\mathcal{L})). \tag{1.2}$$

Thus, given $\mathrm{a}(x)$ the spectrum of $\mathbf{M}$ is reduced to some known terms and a polynomial eigenvalue problem (PEP) posed by $\mathcal{L}(\lambda)$ of degree $(m + 1)$ where m is the degree of the polynomial a. By a well known technique this PEP can be transformed into a standard eigenvalue problem of size $s(m + 1)$. Our method is particularly useful when $s(m + 1) \ll (N + s)$. For the converse case of determining the spectrum of the submatrix $\mathbf{A}$ given the spectrum and an annihilating polynomial of $\mathbf{M}$ a similar formula holds according to theorem 3. After having found a proof for (1.2) in the case of hermitian $\mathbf{M}$, proving the generalization to complex $\mathbf{M}$ seemed rather tedious due to technical subtleties for instance arising from the fact that our initial ansatz explicitly uses eigenvectors although complex matrices might be defective. Eventually, a short proof was found using the properties of $\mathrm{p}(x, y; \mathrm{a})$ and well known manipulations of determinants. However, in order to motivate the use of $\mathrm{p}(x, y; \mathrm{a})$ and

since it allows for the recovery of eigenvectors, the exposition starts with the eigenvector ansatz.

By chance, the author noticed a recent method by Bunimovich and Webb for sharpening spectral bounds and reducing the size of graphs weighted by rational functions called *isospectral graph reduction* [13], [14]. We give a reformulation of one of their results, compare [12, Theorem 3.5. on page 5], using our terminology to be developed below, which covers one essential idea in their work.

Theorem 1 (Bunimovich, Webb). *Let $\mathbb{W}$ be the field of complex rational functions in λ. Let $\mathcal{A} \in \mathbb{W}^{N \times N}$, $\mathcal{B} \in \mathbb{W}^{N \times s}$, $\mathcal{C} \in \mathbb{W}^{s \times N}$, $\mathcal{D} \in \mathbb{W}^{s \times s}$ and*

$$\mathcal{M} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \in \mathbb{W}^{(N+s) \times (N+s)}.$$

Assume

$$\det(\mathcal{M}) \not\equiv 0 \ and \ \det(\mathcal{A}) \not\equiv 0. \tag{1.3}$$

Let the entries $m_{vw}(\lambda)$ of

$$\mathcal{M}_{BW} = \mathcal{M} + \lambda \mathbf{I}_{(N+s)}. \tag{1.4}$$

be arc weights of a digraph Γ s.t. the arc pointing from vertex v to vertex w has weight $m_{vw}(\lambda)$. Let an arc pointing from v to w be present in Γ if and only if $m_{vw}(\lambda) \not\equiv 0$. Let

$$\mathcal{A}_{BW} = (\mathcal{A} + \lambda \mathbf{I}_N) \ and \ \mathcal{D}_{BW} = (\mathcal{D} + \lambda \mathbf{I}_s).$$

Denote the entries of $\mathcal{A}_{BW}$ by a_{vw}. Suppose that the directed subgraph of Γ underlying the nonzero structure of the submatrix $\mathcal{A}_{BW}$ of $\mathcal{M}_{BW}$ is acyclic, with loops allowed and let

$$\tilde{\mathcal{A}} = (\lambda \mathbf{I}_N - \mathrm{diag}(\mathcal{A}_{BW}))^{-1} (\mathcal{A}_{BW} - \mathrm{diag}(\mathcal{A}_{BW})).$$

Let n be the smallest natural power s.t. the nilpotent matrix $\tilde{\mathcal{A}}$ vanishes. Let

$$\mathcal{L}_{BW}(\lambda) = \mathbf{C} \sum_{k=0}^{n-1} \tilde{\mathcal{A}}^k (\lambda \mathbf{I}_N - \mathrm{diag}(\mathcal{A}_{BW}))^{-1} \mathcal{B} + \mathcal{D}_{BW}$$

and

$$\mathcal{L}(\lambda) = \mathcal{L}_{BW}(\lambda) - \lambda \mathbf{I}_s.$$

Then the spectrum of $\mathcal{L}_{BW}$, that is the roots of $\det(\mathcal{L})$, coincides with the spectrum of $\mathcal{M}_{BW}$, that is the roots of $\det(\mathcal{M})$, with the possible exception of elements of the set

$$\mathcal{N} = \bigcup_{i=1}^{N} \{\lambda_0 \in \mathbf{C} | a_{vv}(\lambda) - \lambda \ has \ a \ root \ at \ \lambda_0 \ or \ a_{vw} \ has \ a \ pole \ at \ \lambda_0\}.$$

The matrix $\mathcal{L}_{BW}$ can be interpreted as the weighted adjacency matrix of a graph which has less nodes than Γ. If a subset of the vertex set of Γ induces an acyclic graph, it is called a *structural set* [12]. Note that for undirected graphs structural set are always independent sets (with arbitrary loops, though). Note also that a single vertex is a structural set, thus a graph may be reduced to any positive size. For brevity we omitted that a further inspection on the known set $\mathcal{N}$ will eventually reveal the complete spectrum of $\mathcal{M}_{BW}$. In order to incorporate theorem 1 we developed a framework which generalizes their and our approach. It will, for example, allow for the reduction via a subgraph of Γ which has the same non zero weight on each arc except for arbitrary loop weights. The strength of this more general framework comes from the fact that it relies on algebraic properties of the weighted adjacency matrix instead on combinatorial ones alone.

The structure of the chapter   The rest of this chapter is organized as follows. Since our initial motivation was the possible reduction of a large standard eigenproblem to a small polynomial eigenproblem in order to safe computational costs, we give a short and incomplete introduction to lambda-matrices and the solution of polynomial eigenproblems. The two main points are that any standard eigenvalue problem may be interpreted as a root problem for the determinant of a (monic) linear lambda-matrix and that any root problem of a lambda-matrix with invertible leading coefficient can be transformed into a standard eigenproblem without effort. After that starting with two examples we develop the reduction technique using the eigenvector ansatz and prove our main result for spectra of complex matrices and then for latent values of lambda-matrices. In both cases we also consider a converse variant of the method which uses an annihilating polynomial of a matrix to reduce the eigenproblem of a principal submatrix. We then generalize our approach to a wider set of exploitable polynomials and to matrices the entries of which are complex rationale functions in λ. We conclude by considering aspects regarding applications.

1.2 Matrix Polynomials and Lambda-Matrices

In this section we only recall well known facts about lambda-matrices, which can be found for example in [47], [33] or [53].

1.2.1 Eigenvalues of Polynomial Matrices

The expression

$$\sum_{k=0}^{d} \mathbf{A}_k \mathbf{X}^k \tag{1.5}$$

with $\mathbf{A}_k \in \mathbb{C}^{M \times N}$, $k = 0, \ldots, d$ and $\mathbf{X} \in \mathbb{C}^{N \times N}$ is called a *matrix polynomial* (or a *polynomial matrix*). It is a generalization of the ordinary polynomial, which arises for $M = N = 1$. We are only interested in the square case $M = N$ and the special indeterminate $\mathbf{X} = \lambda\mathbf{I}$. The expression

$$\mathcal{L}(\lambda) = \sum_{k=0}^{d} \mathbf{A}_k \lambda^k \ , \ \mathbf{A}_k \in \mathbb{C}^{N \times N} \ , \ k = 0, \ldots, d \tag{1.6}$$

is called a *lambda-matrix*. The problem of finding numbers λ and vectors $\mathbf{v} \neq \mathbf{0}$ s.t. $\det(\mathcal{L}(\lambda)) = 0$ and $\mathcal{L}(\lambda)\mathbf{v} = \mathbf{0}$ is known as the *polynomial eigenvalue problem* (PEP). The so-called *generalized eigenvalue problem* corresponds to a lambda-matrix of degree $d = 1$. A generalized eigenvalue problem with $\mathbf{A}_1 = \mathbf{I}$ is called an emphordinary eigenvalue problem. The roots of $\det(\mathcal{L}(\lambda))$ are called the *latent values* of $\mathcal{L}(\lambda)$. Note that there is a difference between an eigenvalue and a latent root. The constant lambda-matrix $\mathbf{A}_0$ for instance has no latent values although it certainly has eigenvalues. However, in practical applications the latent values of $\mathbf{A}_1\lambda + \mathbf{A}_0$ are usually called eigenvalues to the generalized eigenproblem given by the matrix pencil $(-\mathbf{A}_0, \mathbf{A}_1)$. Similarly, in the context of polynomial eigenproblems one usually calls the multiset $\rho(\det(\mathcal{L}(\lambda)))$ the spectrum and its elements the eigenvalues of the PEP posed by the lambda-matrix $\mathcal{L}(\lambda)$. We will adopt this convention when the context is clear. However, in order to prevent confusion we will always stick to the notation

$$\sigma(\mathbf{X}) = \rho(\det(\lambda\mathbf{I} - \mathbf{X}))$$

for a square matrix $\mathbf{X}$ the entries of which are functions in λ where $\rho(\det(\mathbf{X}))$ refers to the multiset of latent roots of $\mathbf{X}$.

A lambda-matrix is called *regular* if $\exists\, \lambda \in \mathbb{C}$ s.t. $\det(\mathcal{L}(0)) \neq 0$ and *singular* otherwise, which is a generalization of the same notions for constant matrices. If and only if $\det(\mathcal{L}(0))$ is independent of λ and not zero we call the lambda-matrix *unimodular*. $\left(\begin{smallmatrix} 1 & -\lambda \\ 0 & 2 \end{smallmatrix}\right)$ is an example of a non constant unimodular matrix. There are several notions of unimodularity. Note that we do not require that the determinant has to be in the set $\{-1, 1\}$, but we allow for any value in $\mathbb{C} \setminus 0$. Unimodular matrices are of great practical importance since they allow for elementary operations on lambda-matrices without changing the spectrum. Since $\det(\mathcal{L}(0)) = \det(\mathbf{A}_0)$ a PEP has eigenvalue zero if and only if $\mathbf{A}_0$ is singular. Using the *reversed lambda-matrix*

$$\mathcal{L}_{rev}\left(\frac{1}{\lambda}\right) = \left(\frac{1}{\lambda}\right)^d \mathcal{L}(\lambda) = \sum_{k=0}^{d} \mathbf{A}_{d-k}\left(\frac{1}{\lambda}\right)^k \tag{1.7}$$

one sees that a PEP may have *infinite eigenvalues* only if $\mathbf{A}_d$ is singular. A lambda-matrix is called *monic* if $\mathbf{A}_d = \mathbf{I}_N$. Obviously, a lambda-matrix with regular leading coefficient shares all properties of a monic lambda-matrix up to a multiplication by an invertible constant matrix. Due to the possibility of

infinite values solutions of PEPs are sometimes given as pairs of numbers with a certain ratio, $\lambda = \frac{\alpha}{\beta}$, where $\beta = 0$ corresponds to an infinite eigenvalue. Using such a pair the generalized eigenproblem $\mathbf{A}\mathbf{v} = \lambda \mathbf{B}\mathbf{v}$ becomes $\beta \mathbf{A}\mathbf{v} = \alpha \mathbf{B}\mathbf{v}$. Since our intended application does not involve infinite eigenvalues we will not use this extended notation. In fact, applying the reduction method of this chapter to eigenproblems of complex matrices results in a lambda-matrix which is monic (up to a scalar factor).

1.2.2 Linearization of Polynomial Eigenvalue Problems

We are interested in the solution of PEPs, especially monic PEPs, i.e. we want to find $\lambda \in \mathbb{C}$ and $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^N$ s.t.

$$\mathcal{L}(\lambda)\mathbf{v} = \mathbf{0} \text{ with } \mathcal{L}(\lambda) = \sum_{k=0}^{d} \mathbf{A}_k \lambda^k \ , \ \mathbf{A}_k \in \mathbb{C}^{N \times N} \tag{1.8}$$

in particular with the further restriction

$$\mathbf{A}_d = \mathbf{I}_N.$$

There are algorithms that work directly on this problem, for instance [66] or [83]. Another and perhaps the best known way to solve PEPs is via so-called *linearizations*, which transform the PEP into a generalized eigenproblem $\mathbf{Z}_0 \mathbf{w} = -\lambda \mathbf{Z}_1 \mathbf{w}$ of size dN that is equivalent to it. Any admissible linearization can be expressed as

$$\mathbf{Z}_1 \lambda + \mathbf{Z}_0 = \mathcal{X}(\lambda) \begin{pmatrix} \mathcal{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N(d-1)} \end{pmatrix} \mathcal{Y}(\lambda) \tag{1.9}$$

with unimodular matrices $\mathcal{X}(\lambda)$ and $\mathcal{Y}(\lambda)$. Considering the determinant on both sides one sees that the eigenvalues of a PEP are exactly the eigenvalues of its linearization. A thorough treatment of linearizations can be found e.g. in [3], [57], [40], [39] or [52]. The classical approach utilizes block companion matrices, which are well known for scalar polynomials. The widely used so-called *first companion form* is given by

$$\mathbf{Z}_0 = \begin{pmatrix} -\mathbf{A}_{d-1} & \cdots & -\mathbf{A}_1 & -\mathbf{A}_0 \\ \mathbf{I} & & \mathbf{0} & \mathbf{0} \\ & \ddots & & \vdots \\ \mathbf{0} & & \mathbf{I} & \mathbf{0} \end{pmatrix},$$

$$\mathbf{Z}_1 = -\begin{pmatrix} \mathbf{A}_d & & & \mathbf{0} \\ & \mathbf{I} & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{I} \end{pmatrix} \text{ and } \mathbf{w} = \begin{pmatrix} \lambda^{d-1}\mathbf{v} \\ \vdots \\ \lambda\mathbf{v} \\ \mathbf{v} \end{pmatrix}. \tag{1.10}$$

In case of a monic PEP we essentially end up with an ordinary eigenvalue problem. It is easy to see that any solution of (1.8) provides a solution for (1.10) and that any solution of (1.10) obeys the block structure of the companion form and therefore immediately provides a solution for (1.8).

It is a disadvantage of the companion form that it in general does not reflect special properties of the PEP like all coefficients being hermitian. More general linearizations allow for instance transformations into block symmetric generalized eigenproblems so that in the case of hermitian lambda-matrices the linearization is also hermitian although moniticity is not preserved in general. Algorithms for the linear (generalized) eigenproblem may be found for example in [61], [34] or [72].

1.2.3 Miscellaneous Facts: Annihilating Polynomials, Triangularization and Extension to Rational Functions

By the Cayley-Hamilton theorem any lambda matrix is annihilated by its characteristic polynomial. To be more precise, the formal expression

$$\det\left(\mathcal{L}\left(\lambda\right) - t\mathbf{I}\right) = \sum_i a_i\left(\lambda\right) t^i \in R\left[t\right]$$

vanishes for all λ if we replace the indeterminate t by $\mathcal{L}\left(\lambda\right)$ and the coefficients $a_i\left(\lambda\right)$ by $a_i\left(\lambda\right)\mathbf{I}$. As a consequence for any lambda-matrix there exist annihilating polynomials in $R\left[X\right]$ where R is the ring of complex scalar polynomials in λ.

It is known, [33], [77] and possibly [76], that for any regular lambda-matrix $\mathcal{L}\left(\lambda\right)$ there are unimodular matrices $\mathcal{U}\left(\lambda\right)$ and $\mathcal{V}\left(\lambda\right)$ and an upper triangular lambda-matrix $\mathcal{T}\left(\lambda\right)$ of the same degree as $\mathcal{L}$, i.e. $\mathcal{T}\left(\lambda\right) = \sum_{i=0}^d \mathbf{T}_i\lambda^i$ with $\mathbf{T}_i$ upper triangular, s.t.

$$\mathcal{U}\left(\lambda\right)\mathcal{L}\left(\lambda\right)\mathcal{V}\left(\lambda\right) = \mathcal{T}\left(\lambda\right).$$

It is easy to see that for any triangular lambda-matrix $\mathcal{T}$ the polynomial

$$\prod_{i=1}^N \left(\mathcal{T} - \mathcal{T}_{ii}\mathbf{I}\right)$$

where $\mathcal{T}_{ii}$ denotes the i-th diagonal entry of $\mathcal{T}$ vanishes.

Let $\mathbb{W}$ denote the field of all complex rational functions in one variable. Every element in $\mathbb{W}$ can be expressed as a fraction of two elements of $\mathbb{C}\left[X\right]$, the commutative ring of complex polynomials, with the convention that factors shared by the numerator and denominator are removed. By the following standard technique several results for lambda-matrices may be extended to matrices with entries in $\mathbb{W}$. Let $\mathcal{M}\left(\lambda\right) \in \mathbb{W}^{N \times N}$ with entries $\mathcal{M}_{ij}\left(\lambda\right) = \frac{\mathrm{p}_{ij}\left(\lambda\right)}{\mathrm{q}_{ij}\left(\lambda\right)}$. Let $\mathcal{Q}_i\left(\lambda\right)$ be the least common multiple of all denominators in row i. Let $\mathcal{Q}\left(\lambda\right) = \mathrm{diag}\left(\mathcal{Q}_i\left(\lambda\right)\right) = \mathcal{Q}_i\left(\lambda\right)\delta_{ij}$, then $\mathcal{Q}\mathcal{M}$ is a lambda-matrix. Since

$\det(\mathcal{Q}\mathcal{M}) = \det(\mathcal{Q})\det(\mathcal{M})$ the latent roots of $\mathcal{M}$ and $\mathcal{Q}\mathcal{M}$ coincide with the exception of the latent roots of $\mathcal{Q}$. Several variants of this technique may be considered, for instance $\mathcal{Q}_i(\lambda)$ being just the product $\prod_j q_{ij}(\lambda)$, which corresponds to the *polynomial extension* in [13], or employing a column wise definition and right multiplication of $\mathcal{Q}$ or choosing simply $\mathcal{Q}(\lambda) = \prod_{ij} q_{ij}(\lambda) \mathbf{I}_N$.

1.3 Examples Considering Complex Matrices

1.3.1 First Example: Independent Set

To get familiar with the notation we consider the most basic example for our method and assume an empty principal submatrix. However, the results already generalize to submatrices which are constant diagonal. A more representative example will be given in the next subsection.

Consider the constant zero matrix $\mathbf{0}_N$ of size N. Its minimal polynomial is given by

$$\mathrm{minpol}(x; \mathbf{0}_N) = x.$$

We add a node by appending a row and a column, which results in the block structured matrix

$$\mathbf{M} = \left(\begin{array}{cc} \mathbf{0} & \mathbf{b} \\ \mathbf{c}' & d \end{array} \right) \in \mathbb{C}^{(N+1)\times(N+1)}$$

with $\mathbf{b}, \mathbf{c} \in \mathbb{C}^{N\times 1}$ and $d \in \mathbb{C}$. If $\mathbf{c} = \mathbf{b}$ and $d \in \mathbb{R}$ then $\mathbf{M}$ is hermitian. If $\mathbf{A}$ is the adjacency matrix of a graph, the nodes inducing $\mathbf{A}$, i.e. all initial nodes, form an independent set. Let $\left(\lambda, \left(\begin{smallmatrix} \mathbf{v} \\ w \end{smallmatrix}\right)\right)$ be an eigenpair of $\mathbf{M}$ s.t. the partition of the eigenvector conforms with the block structure of $\mathbf{M}$. Due to the partition the eigenproblem $\mathbf{M}\left(\begin{smallmatrix} \mathbf{v} \\ w \end{smallmatrix}\right) = \lambda\left(\begin{smallmatrix} \mathbf{v} \\ w \end{smallmatrix}\right)$ splits into a system of two equations

$$\mathbf{b}w = \lambda\mathbf{v}$$
$$\mathbf{c}'\mathbf{v} + dw = \lambda w.$$

Multiplying the last equation by λ, inserting the first one and performing some simple manipulations yields

$$\mathcal{L}(\lambda)\mathbf{w} = \left[\lambda^2 - d\lambda - \mathbf{c}'\mathbf{b}\right]w = 0. \tag{1.11}$$

Note, that $w = 0$ implies $\lambda = 0$ for any eigenvector of $\mathbf{M}$. If $\lambda_i \neq 0$, $i = 1, 2$, which is equivalent to $\mathbf{c}'\mathbf{b} \neq 0$, the corresponding eigenvectors of $\mathbf{M}$ are given by $\left(\begin{smallmatrix} \mathbf{b} \\ \lambda_i \end{smallmatrix}\right)$, $i = 1, 2$. The remaining N-1 eigenvalues of $\mathbf{M}$ can be determined to be zero using $\dim(\ker(\mathbf{c}')) \geq N$-1. Since $m_0(\mathbf{0}_N) = N$, there is a $(N$-1)-dimensional subspace $\mathcal{V}$ of the eigenspace to the eigenvalue 0 which is also a subset of $\ker(\mathbf{c}')$. Any vector of the form $\left(\begin{smallmatrix} \mathbf{v} \\ 0 \end{smallmatrix}\right)$ with $\mathbf{v} \in \mathcal{V}$ is an eigenvector of $\mathbf{M}$ to the eigenvalue 0. Therefore $m_0(\mathbf{M})$-1 $\geq N$-1. In the hermitian case we can refer to corollary 1 and $m_0(\mathbf{0}_N) = N$ for the same result. Excluding $\mathbf{c}'\mathbf{b} = 0$ we end up with the simple formula

$$\sigma(\mathbf{M}) = \left\{0^{N-s}\right\} + \rho(\det(\mathcal{L})) \tag{1.12}$$

with $s = 1$. In the case $\mathbf{c'b} = 0$ the matrix $\mathbf{M}$ becomes defective and our ansatz relying on eigenvectors does not apply properly. However, since defectiveness is an extremely sensitive matrix property in the sense that it might get lost even for tiny perturbations, it is not surprising that (1.2) is also correct for the case $\mathbf{c'b} = 0$ as will be proven below. Recall that we cold have solved (1.11) by solving the eigenproblem for

$$\begin{pmatrix} d & \mathbf{c'b} \\ 1 & 0 \end{pmatrix}.$$

Note that for large N the costs for constructing this matrix are linear in N while those of finding its spectrum are fixed and small.

As indicated by (1.12) the above scheme easily carries over to the more general case of adding $s \geq 1$ nodes. In order to notate this generalization properly, w becomes the vector $\mathbf{w} \in \mathbb{C}^s$, $\mathbf{b}$, $\mathbf{c'}$ and d become matrices $\mathbf{B} \in \mathbb{C}^{N \times s}$, $\mathbf{C} \in \mathbb{C}^{s \times N}$ and $\mathbf{D} \in \mathbb{C}^{s \times s}$ and the coefficients in the polynomial of (1.11) become square matrices. Thus, we have

$$\mathcal{L}(\lambda) \mathbf{w} = \left[\lambda^2 - \mathbf{D}\lambda - \mathbf{CB} \right] \mathbf{w} = 0. \tag{1.13}$$

instead of (1.11).

As shown in the previous section the spectrum of the lambda-matrix $\mathcal{L}$ is equivalent to the spectrum of its first companion form. For our example we have

$$\det \left[\begin{pmatrix} \mathbf{D} & \mathbf{CB} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} - \lambda \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \right] = 0, \tag{1.14}$$

which is a standard eigenproblem of size $2s \times 2s$. In the more general case of complex matrices with minimal polynomial of degree $m \geq 1$ we essentially will end up with a polynomial in λ of degree $m + 1$ with matrix coefficients. The linearization, for example by the first companion form, will be of size $(m + 1) s$. We now prove (1.12) for the general case $s \geq 1$ which employs (1.13).

Proof of (1.12). Considering

$$0 = \det (\mathbf{M} - \lambda \mathbf{I_{N+s}}) = \det \begin{pmatrix} -\lambda \mathbf{I}_N & \mathbf{B} \\ \mathbf{C} & \mathbf{D} - \lambda \mathbf{I}_s \end{pmatrix},$$

we set

$$\begin{aligned} 0 &= \det \begin{pmatrix} -\lambda \mathbf{I}_N & \mathbf{B} \\ \mathbf{C} & \mathbf{D} - \lambda \mathbf{I}_s \end{pmatrix} \lambda^s \\ &= \det \begin{pmatrix} -\lambda \mathbf{I}_N & \lambda \mathbf{B} \\ \mathbf{C} & \lambda \mathbf{D} - \lambda^2 \mathbf{I}_s \end{pmatrix} \\ &= \det \begin{pmatrix} -\lambda \mathbf{I}_N & \lambda \mathbf{B} \\ \mathbf{C} & \lambda \mathbf{D} - \lambda^2 \mathbf{I}_s \end{pmatrix} \det \begin{pmatrix} \mathbf{I}_N & \mathbf{B} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} \\ &= \det \begin{pmatrix} -\lambda \mathbf{I}_N & \mathbf{0} \\ \mathbf{C} & \mathbf{CB} + \lambda \mathbf{D} - \lambda^2 \mathbf{I}_s \end{pmatrix}. \end{aligned}$$

Therefore,

$$\sigma\left(\mathbf{M}\right) + \left\{0^s\right\} = \left\{0^N\right\} + \rho\left(\det\left(\mathcal{L}\right)\right).$$

$$\square$$

remarks If $\mathbf{M}$ shows a scaled identity, $r\mathbf{I}_N$, instead of $\mathbf{0}$ in its upper diagonal block, we may consider the sift $\mathbf{M} - r\mathbf{I}_{(N+s)}$, which implies the shift $\mathbf{A} - r\mathbf{I}_N$. The original eigenvalues of $\mathbf{M}$ and $\mathbf{A}$ are recovered by a simple back shift. Equivalently, one can apply the general method using a polynomial of the form $\mathrm{a}\left(x; r\mathbf{I}\right) = r_1 x - r_0$, $\frac{r_0}{r_1} = r$, yielding the lambda-matrix

$$\mathcal{L}\left(\lambda; \mathrm{a}\right) = \text{-}r_1\mathbf{I}_s\lambda^2 + \left(r_1\mathbf{D} - r_0\mathbf{I}_s\right)\lambda + r_1\mathbf{CB} + r_0\mathbf{D} \tag{1.15}$$

and the relation

$$\sigma\left(\mathbf{M}\right) = \left\{\left(\frac{\text{-}r_0}{r_1}\right)^{N\text{-}s}\right\} + \rho\left(\det\left(\mathcal{L}\right)\right). \tag{1.16}$$

1.3.2 Second Example: Quadratic Annihilating Polynomial

In this section we show how to utilize an annihilating polynomial of degree two for a principal submatrix to simplify the eigenproblem of a matrix. This case already captures the most important aspects of the general procedure.
Let

$$\mathrm{a}\left(x; \mathbf{A}\right) = r_2 x^2 + r_1 x + r_0 = r_2\left(x - \rho_1\right)\left(x - \rho_2\right) \quad , \quad r_2 \neq 0 \tag{1.17}$$

be an annihilating polynomial for $\mathbf{A}$, with roots ρ_1 and ρ_2. Adding s nodes to $\mathbf{A}$ by appending rows and columns results in the block structured matrix

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in \mathbb{C}^{N\times N}.$$

Let $\left(\lambda, \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}\right)$ be a right-eigenpair of $\mathbf{M}$ s.t. the partition of the eigenvector conforms with the block structure of $\mathbf{M}$. Due to the partition the eigenproblem $\mathbf{M}\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \lambda\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}$ splits into a system of two equations

$$\mathbf{Av} + \mathbf{Bw} = \lambda\mathbf{v} \tag{1.18}$$

$$\mathbf{Cv} + \mathbf{Dw} = \lambda\mathbf{w}. \tag{1.19}$$

Note from (1.18) that $\mathbf{w} = \mathbf{0}$ implies that $\mathbf{v}$ is an eigenvector of $\mathbf{A}$. Multiplying the latter equation by λ and inserting the former one yields

$$\mathbf{CAv} + \mathbf{CBw} + \lambda\mathbf{Dw} = \lambda^2\mathbf{w}. \tag{1.20}$$

Multiplying again by λ and making the same substitution gives

$$\mathbf{CA}^2\mathbf{v} + \mathbf{CABw} + \lambda\mathbf{CBw} + \lambda^2\mathbf{Dw} = \lambda^3\mathbf{w}. \tag{1.21}$$

A linear combination, $r_2(1.21)+r_1(1.20)+r_0(1.19)$, according to (1.17) yields.

$$\mathbf{C}\, \mathrm{a}\left(\mathbf{A};\mathbf{A}\right)\mathbf{v} + \mathbf{C}\left[r_2\mathbf{A} + \left(r_2\lambda + r_1\right)\mathbf{I}_s\right]\mathbf{B} + \mathrm{a}\left(\lambda;\mathbf{A}\right)\left(\mathbf{D} - \mathbf{I}_s\right) = 0 \qquad (1.22)$$

The first term on the left-hand side cancels and after rearranging we obtain the cubic PEP

$$\mathcal{L}\left(\lambda\right)\mathbf{w} = \sum_{k=0}^{3}\mathbf{A}_k\lambda^k\mathbf{w} \qquad (1.23)$$

with coefficients

$$
\begin{aligned}
\mathbf{A}_0 &= & r_2\mathbf{CAB} + r_1\mathbf{CB} + r_0\mathbf{D} \\
\mathbf{A}_1 &= & r_2\mathbf{CB} & + r_1\mathbf{D} & - r_0\mathbf{I}_s \\
\mathbf{A}_2 &= & r_2\mathbf{D} & - r_1\mathbf{I}_s \\
\mathbf{A}_3 &= -r_2\mathbf{I}_s,
\end{aligned}
\qquad (1.24)
$$

which gets monic after divison by $-r_2$ and is therefore equivalent—by linearization—to a standard eigenvalue problem of size $3s \times 3s$.
We now proof the formula

$$\sigma\left(\mathbf{M}\right) + \{\rho_1^s, \rho_2^s\} = \sigma\left(\mathbf{A}\right) + \rho\left(\det\left(\mathcal{L}\right)\right). \qquad (1.25)$$

Proof of (1.25). Consider

$$\det\begin{pmatrix} \mathbf{A} - \lambda\mathbf{I}_N & \mathbf{B} \\ \mathbf{C} & \mathbf{D} - \lambda\mathbf{I}_s \end{pmatrix}\left(r_2\lambda^2 + r_1\lambda + r_0\right)^s \qquad (1.26)$$

$$= \det\begin{pmatrix} \mathbf{A} - \lambda\mathbf{I}_N & \left(r_2\lambda^2 + r_1\lambda + r_0\right)\mathbf{B} \\ \mathbf{C} & \mathrm{a}\left(\lambda;\mathbf{A}\right)\left(\mathbf{D} - \lambda\mathbf{I}_s\right) \end{pmatrix}$$

$$= \det\begin{pmatrix} \mathbf{A} - \lambda\mathbf{I}_N & \left(r_2\lambda^2 + r_1\lambda + r_0\right)\mathbf{B} \\ \mathbf{C} & \mathrm{a}\left(\lambda;\mathbf{A}\right)\left(\mathbf{D} - \lambda\mathbf{I}_s\right) \end{pmatrix}$$

$$\times \det\begin{pmatrix} \mathbf{I}_N & \left(\left(r_2\lambda + r_1\right)\mathbf{I}_N + r_2\mathbf{A}\right)\mathbf{B} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix}$$

$$= \det\begin{pmatrix} \mathbf{A} - \lambda\mathbf{I}_N & \mathbf{0} \\ \mathbf{C} & \mathcal{L}\left(\lambda\right) \end{pmatrix}. \qquad (1.27)$$

The roots of the left-hand side, (1.26), are given by $\sigma\left(\mathbf{M}\right) + \{\rho_1^s, \rho_2^s\}$ and the roots of the right-hand side, (1.27), are given by $\sigma\left(\mathbf{A}\right) + \rho\left(\det\left(\mathcal{L}\right)\right)$. $\square$

The term $\left(\left(r_2\lambda + r_1\right)\mathbf{I}_N + r_2\mathbf{A}\right)$ in the second postmultiplication may be motivated by the middle term of the left-hand side of (1.22).

Remarks Since $\mathrm{a}\left(x;\mathbf{A}\right)$ annihilates $\mathbf{A}$, the minimal polynomial of $\mathbf{A}$ must divide it. However, it is not required that all roots of $\mathrm{a}\left(x;\mathbf{A}\right)$ are eigenvalues of $\mathbf{A}$. The condition that the leading coefficient should not vanish is not essential. For $r_2 = 0$ (1.23) automatically turns into (1.15) as can be observed

from (1.24). However, in this case the factorization in (1.17) does not apply and one has to replace the pair ρ_1 and ρ_2 by the actual root $-\frac{r_0}{r_1}$. When that is taken into account, then (1.25) turns into (1.16).

The example of this subsection covers the case of exploiting a clique, a set of mutually adjacent vertices, of size N in a simple graph Γ of size $(N + s)$ in order to compute the spectrum of its adjacency matrix $\mathbf{M}$. This corresponds to $\mathbf{A} = \mathbf{J}_N - \mathbf{I}_N$ and, for instance, a $(x) = (x - N + 1)(x + 1)$. Using (1.25), (1.23) and (1.24) one reduces the eigenproblem of $\mathbf{M}$ to a cubic lambda-matrix of size s the linearization of which has size $3s$. However, since for $\mathbf{A}$ the eigenvector $\mathbf{j}_N$ to the eigenvalue $(N\text{-}1)$ is known, we can deflate it by an unitary transformation $\mathbf{Q}$ s.t.

$$\mathbf{Q}'\mathbf{A}\mathbf{Q} = \begin{pmatrix} \tilde{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & N\text{-}1 \end{pmatrix} \tag{1.28}$$

Since the multiplicity of $(N\text{-}1)$ is 1 all eigenvalues of $\tilde{\mathbf{A}}$ are given by the second eigenvalue of $\mathbf{A}$, hence $\tilde{\mathbf{A}} = -\mathbf{I}_{N\text{-}1}$. We consider

$$\tilde{\mathbf{M}} = \begin{pmatrix} \mathbf{Q}' & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} \mathbf{M} \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} = \begin{pmatrix} \mathbf{Q}'\mathbf{A}\mathbf{Q} & \mathbf{Q}'\mathbf{B} \\ \mathbf{C}\mathbf{Q} & \mathbf{D} \end{pmatrix}, \tag{1.29}$$

which has the following form

$$\tilde{\mathbf{M}} = \left(\begin{array}{ccc|c|ccc} & & & 0 & \mathtt{x} & \cdots & \mathtt{x} \\ & \text{-}\mathbf{I}_{N\text{-}1} & & \vdots & \vdots & \ddots & \vdots \\ & & & 0 & \mathtt{x} & \cdots & \mathtt{x} \\ \hline 0 & \cdots & 0 & N\text{-}1 & \mathtt{x} & \cdots & \mathtt{x} \\ \hline \mathtt{x} & \cdots & \mathtt{x} & \mathtt{x} & & & \\ \vdots & \ddots & \vdots & \vdots & & \mathbf{D} & \\ \mathtt{x} & \cdots & \mathtt{x} & \mathtt{x} & & & \end{array} \right) \tag{1.30}$$

If we utilize a 2×2 block partition of $\tilde{\mathbf{M}}$ such that $\tilde{\mathbf{A}}$ becomes the upper diagonal block, we may apply (1.16) to reduce the eigenproblem of $\tilde{\mathbf{M}}$ to that of a quadratic lambda-matrix of size $(s + 1)$ the linearization of which has size $2(s + 1)$. Particularly in the case of $s > 2$ this might be an advantage compared to the straightforward reduction. However, in general we might not deflate each known eigenpair of $\mathbf{A}$. For example, the deflation of an eigenvalue -1 of the submatrix $\mathbf{J}_N - \mathbf{I}_N$ does not allow for the use of a linear polynomial a (x) if $N \geq 3$. Hence, the lambda-matrix of the transformed and repartitioned problem would still be cubic.

If additionally the graph is regular, we may consider the adjacency matrix of the complement graph. In this case we obtain a quadratic lambda-matrix without incrementing its size. We only have to take into account some simple relations between the spectrum of a regular graph and its complement. Note that it is in general harmful to deflate an eigenpair of $\mathbf{M}$ since this typically changes the spectrum of submatrices completely.

1.4 The Main Theorem

1.4.1 Useful Identities

Polynomial Identities and the Adjoint   Let $q(x) = \sum_{k=0}^{m} r_k x^k$ be an univariate complex polynomial. We define

Definition 1. $p_k(y, z) = \sum_{i=1}^{k} y^{i-1} z^{k-i}$, $k \in \mathbb{N}$.

The identity $y\, p_k(y, z) - p_k(y, z)\, z = y^k - z^k$ is easily proven. In the commutative case, $yz = zy$, we have

$$(y - z)\, p_k(y, z) = y^k - z^k. \tag{1.31}$$

We also define

Definition 2. $p(y, z; q) = \sum_{k=1}^{m} r_k\, p_k(y, z)$.

Thus for commuting y and z

$$(y - z)\, p(y, z; q) = q(y) - q(z). \tag{1.32}$$

p can be obtained by the formal substitution

$$x^k \rightarrow \frac{1}{k+1} \sum_{i=0}^{k} y^i z^{k-i}$$

in the first derivative of $q(x)$. An important relation holds for commuting y and z s.t. $q(z) \neq q(y) = 0$. In this case

$$(y - z)^{-1} = p(y, z; q)\, \frac{1}{q(z)}\, .$$

Our variables will usually be of the form $y = \mathbf{X}$ and $z = \lambda \mathbf{I}$ with square matrix $\mathbf{X}$ and complex scalar λ. Let $\operatorname{char}(x; \mathbf{X})$ and $\operatorname{minpol}(x; \mathbf{X})$ be the characteristic and the minimal polynomial of $\mathbf{X}$, respectively. The functions

$$B(\lambda) = p(\lambda \mathbf{I}, \mathbf{X}; \operatorname{char}) \tag{1.33}$$

and

$$C(\lambda) = p(\lambda \mathbf{I}, \mathbf{X}; \operatorname{minpol}) \tag{1.34}$$

are called the *adjoint matrix* and the *reduced adjoint matrix*, respectively, by Gantmacher [28, equations (31) and (55) on pp. 84/91]. $B(0)$ is also known as the *adjoint* or the *adjugate* of $\mathbf{X}$ up to different sign conventions. It turned out that the method in this chapter is most easily understood by just considering a generalization of (1.33) and (1.34) to all annihilating polynomials of $\mathbf{X}$, which also works for an extension of $\mathbf{X}$ to matrices the entries of which are functions

in lambda. We conclude with the following observation [28]. If $a\left(x; \mathbf{X}\right)$ is any annihilating polynomial of $\mathbf{X}$, then we have according to (1.32)

$$\left(\mathbf{X} - \lambda\mathbf{I}\right) p\left(\mathbf{X}, \lambda\mathbf{I}; a\right) = -a\left(\lambda; \mathbf{X}\right)\mathbf{I}. \tag{1.35}$$

If λ is an eigenvalue of $\mathbf{X}$, then the right-hand side of (1.35) vanishes, which implies that the columns of $p\left(\mathbf{X}, \lambda\mathbf{I}; a\right)$ are eigenvectors of $\mathbf{X}$ to the eigenvalue λ.

Block Determinants and the Schur Complement   Let R be a commutative ring with identity. We recall four properties of determinants of block matrices with elements in R. $\mathbf{A}$ and $\mathbf{D}$ are square matrix blocks not necessarily of the same size. Each $\mathbf{I}$ stands for an identity matrix. $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{X}$ are rectangular of suitable size. All partitions are conformable. Let $n = \dim\left(\mathbf{D}\right)$ and $\mu \in R$.

- $\det\left(\mathbf{A}\mathbf{X}\right) = \det\left(\mathbf{A}\right)\det\left(\mathbf{X}\right)$.

- $\det\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} = \det\left(\mathbf{A}\right)\det\left(\mathbf{D}\right)$,

- $\det\begin{pmatrix} \mathbf{A} & \mu\mathbf{B} \\ \mathbf{C} & \mu\mathbf{D} \end{pmatrix} = \det\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mu\mathbf{C} & \mu\mathbf{D} \end{pmatrix} = \mu^{n}\det\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$,

- $\det\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}\det\begin{pmatrix} \mathbf{I} & \mathbf{X} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \det\begin{pmatrix} \mathbf{A} & \mathbf{A}\mathbf{X} + \mathbf{B} \\ \mathbf{C} & \mathbf{C}\mathbf{X} + \mathbf{D} \end{pmatrix}$.

A common application of the last relation assumes that $\mathbf{A}$ is invertible and uses $\mathbf{X} = -\mathbf{A}^{-1}\mathbf{B}$ in order to be able to apply the first relation. In this case the lower right matrix block on the right-hand side becomes $\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$, which is also called a Schur complement [18], [86]. Our method may be seen as an application of the technique of Schur complements to lambda-matrices with the help of a generalization of the adjoint for obtaining an inverse.

1.4.2 A Recursion

Let $\mathbf{M}$ be a complex square matrix partitioned into a 2×2 block form. We consider the block partitioned eigenproblem

$$\mathbf{M}\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \lambda\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}, \tag{1.36}$$

which leads to

$$\mathbf{A}\mathbf{v} + \mathbf{B}\mathbf{w} = \lambda\mathbf{v} \tag{1.37}$$

$$\mathbf{C}\mathbf{v} + \mathbf{D}\mathbf{w} = \lambda\mathbf{w} \tag{1.38}$$

Multiplying the last equation by λ and inserting the first one yields

$$\mathbf{C}\mathbf{A}\mathbf{v} + \mathbf{C}\mathbf{B}\mathbf{w} + \lambda\mathbf{D}\mathbf{w} = \lambda^{2}\mathbf{w}. \tag{1.39}$$

Repeating the last step we get a sequence of equations. The k-th element reads as

$$\mathbf{C}\mathbf{A}^k\mathbf{v} + \mathbf{C}\sum_{i=1}^{k}\lambda^{i-1}\mathbf{A}^{k-i}\mathbf{B}\mathbf{w} + \lambda^k\mathbf{D}\mathbf{w} = \lambda^{k+1}\mathbf{w}, \tag{1.40}$$

which is equivalent to

$$\mathbf{C}\mathbf{A}^k\mathbf{v} + \mathbf{C}\,\mathrm{p}_k\left(\mathbf{A},\lambda\right)\mathbf{B}\mathbf{w} + \lambda^k\mathbf{D}\mathbf{w} = \lambda^{k+1}\mathbf{w}. \tag{1.41}$$

When we utilize a linear combination of these equations s.t. the term involving $\mathbf{v}$ vanishes, we will obtain a PEP.

1.4.3 The Associated Lambda-Matrix

Let

$$\mathrm{a}\left(x;\mathbf{A}\right) = \sum_{k=0}^{m} r_k x^k \ , \ r_m \neq 0 \tag{1.42}$$

be an annihilating polynomial of $\mathbf{A}$, which implies

$$\lambda \in \sigma\left(\mathbf{A}\right) \Rightarrow \lambda \in \rho\left(\mathrm{a}\left(x;\mathbf{A}\right)\right).$$

We consider a linear combination of (1.40) or equivalently (1.41) with $k = 0,\ldots,d$ using the coefficients r_k of (1.42). Since $\mathrm{a}\left(\mathbf{A};\mathbf{A}\right) = \mathbf{0}$, the term

$$\mathbf{C}\sum_{k=0}^{m} r_k\mathbf{A}^k\mathbf{v} \qquad \text{in} \qquad \sum_{k=0}^{m} r_k\lambda^{k+1}\mathbf{w} = \lambda\,\mathrm{a}\left(\lambda;\mathbf{A}\right)\mathbf{w}$$

vanishes, and we are left with

$$\mathbf{0} = \sum_{k=0}^{m} r_k\left[\mathbf{C}\sum_{i=1}^{k}\lambda^{i-1}\mathbf{A}^{k-i}\mathbf{B} + \lambda^k\mathbf{D} - \lambda^{k+1}\mathbf{I}_s\right]\mathbf{w} = \left[\sum_{k=0}^{m+1}\mathbf{A}_k\lambda^k\right]\mathbf{w} \tag{1.43}$$

where

$$\mathbf{A}_{m+1} = -r_m\mathbf{I}_s,$$
$$\mathbf{A}_k = \sum_{i=0}^{m-1-k} r_{k+1+i}\mathbf{C}\mathbf{A}^i\mathbf{B} + r_k\mathbf{D} - r_{k-1}\mathbf{I}_s \ , \ k = 1,\ldots,m,$$
$$\mathbf{A}_0 = \sum_{i=0}^{m-1} r_{i+1}\mathbf{C}\mathbf{A}^i\mathbf{B} + r_0\mathbf{D}, \tag{1.44}$$

which is a monic (up to a constant factor) PEP of degree $m + 1$ that can be solved by linearization. We call

$$\mathcal{L}\left(\lambda;\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathrm{a}\right) = \sum_{k=0}^{m+1}\mathbf{A}_k\lambda^k = \mathbf{C}\,\mathrm{p}\left(\mathbf{A},\lambda\mathbf{I};\mathrm{a}\right)\mathbf{B} + \mathrm{a}\left(\lambda;\mathbf{A}\right)\left(\mathbf{D} - \lambda\mathbf{I}\right) \tag{1.45}$$

the associated lambda-matrix and (1.43) the associated PEP for the block structured matrix $\mathbf{M} = \left(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix} \right)$ with annihilating polynomial $\mathrm{a}\,(x; \mathbf{A})$. Note that for hermitian (real) $\mathbf{M}$ the associated lambda-matrix is hermitian (real) if $\mathrm{a}\,(x; \mathbf{A})$ is real.

1.4.4 The Main Theorem for Complex Matrices

Theorem 2. *Let* $\mathbf{A} \in \mathbf{C}^{N \times N}$ *be annihilated by the polynomial*

$$\mathrm{a}\,(x; \mathbf{A}) = \sum_{k=0}^{m} r_k x^k \ , \ r_m \neq 0$$

of degree m. *Let* $\mathbf{B} \in \mathbb{C}^{N \times s}$, $\mathbf{C} \in \mathbb{C}^{s \times N}$, $\mathbf{D} \in \mathbb{C}^{s \times s}$ *and*

$$\mathbf{M} = \left(\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{array} \right).$$

Let

$$\mathcal{L}\,(\lambda; \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathrm{a}) = \mathbf{C}\,\mathrm{p}\,(\mathbf{A}, \lambda \mathbf{I}; \mathrm{a})\,\mathbf{B} + \mathrm{a}\,(\lambda)\,(\mathbf{D} - \lambda \mathbf{I})$$

be the associated lambda-matrix. Then, $\mathcal{L}$ *has a linearization as a standard eigenproblem and the spectra of* $\mathbf{A}$ *and* $\mathbf{M}$ *are related by*

$$\sigma\,(\mathbf{M}) = \sigma\,(\mathbf{A}) - \rho^s\,(\mathrm{a}\,(\lambda)) + \rho\,(\det\,(\mathcal{L}\,(\lambda)))\,.$$

Proof. According to (1.44), the first coefficient of $\mathcal{L}$ is $-r_m \mathbf{I}_s$. Thus, multiplication by the unimodular matrix $-\frac{1}{r_m}\mathbf{I}_s$ yields a monic lambda-matrix with monic linearization using the first companion form.
The spectrum of $\mathbf{M}$ are the roots of

$$\det\,\left(\mathbf{M} - \lambda \mathbf{I}_{(N+s)}\right) = \det \left(\begin{array}{cc} \mathbf{A} - \lambda \mathbf{I}_N & \mathbf{B} \\ \mathbf{C} & \mathbf{D} - \lambda \mathbf{I}_s \end{array} \right).$$

Thus,

$$\begin{aligned}
0 &= \det \left(\begin{array}{cc} \mathbf{A} - \lambda \mathbf{I}_N & \mathbf{B} \\ \mathbf{C} & \mathbf{D} - \lambda \mathbf{I}_s \end{array} \right) \det\,(\mathrm{a}\,(\lambda)\,\mathbf{I}_s) \\
&= \det \left(\begin{array}{cc} \mathbf{A} - \lambda \mathbf{I}_N & \mathrm{a}\,(\lambda)\,\mathbf{B} \\ \mathbf{C} & \mathrm{a}\,(\lambda)\,(\mathbf{D} - \lambda \mathbf{I}_s) \end{array} \right) \\
&= \det \left(\begin{array}{cc} \mathbf{A} - \lambda \mathbf{I}_N & \mathrm{a}\,(\lambda)\,\mathbf{B} \\ \mathbf{C} & \mathrm{a}\,(\lambda)\,(\mathbf{D} - \lambda \mathbf{I}_s) \end{array} \right) \det \left(\begin{array}{cc} \mathbf{I}_N & \mathrm{p}\,(\mathbf{A}, \lambda; \mathrm{a})\,\mathbf{B} \\ \mathbf{0} & \mathbf{I}_s \end{array} \right) \\
&= \det \left(\begin{array}{cc} \mathbf{A} - \lambda \mathbf{I}_N & \mathbf{0} \\ \mathbf{C} & \mathbf{C}\,\mathrm{p}\,(\mathbf{A}, \lambda; \mathrm{a})\,\mathbf{B} + \mathrm{a}\,(\lambda)\,(\mathbf{D} - \lambda \mathbf{I}_s) \end{array} \right) \\
&= \det\,(\mathbf{A} - \lambda \mathbf{I}_N) \det\,(\mathcal{L}\,(\lambda))
\end{aligned}$$

and therefore,

$$\sigma\,(\mathbf{M}) + \rho^s\,(\mathrm{a}\,(\lambda)) = \sigma\,(\mathbf{A}) + \rho\,(\det\,(\mathcal{L}\,(\lambda)))\,.$$

$\square$

Since $\mathcal{L}$ has a linearization as a standard eigenproblem, infinite eigenvalues are excluded. Obviously, if we impose the rather harmless restriction $r_m = -1$, then we may already build up a monic linearized form of $\mathcal{L}$. In this case the reduction procedure would involve only complex matrices and scalar polynomials. However, utilizing lambda matrices allows for a larger set of available algorithms and may be used to exploit hermiticity or other structural properties.

We may reformulate the eigenvalue problem of $\mathbf{M}$ as the equivalent latent root problem of $\mathbf{M} - \lambda\mathbf{I}$, which leads to a generalization of theorem 2 in section 1.6.

1.4.5 Recovery of Eigenvectors

We give three relations between right-eigenvectors of $\mathbf{A}$, those of the associated lambda-matrix $\mathcal{L}$ and those of the complex matrix $\mathbf{M}$. The case of left-eigenvectors is similar. Let $a(x; \mathbf{A}) \in \mathbb{C}[x]$ be the annihilating polynomial for $\mathbf{A}$ used to construct the associated lambda-matrix. As in the recursion above any eigenvector of $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ to an eigenvalue λ can be partitioned as $\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}$ where $\mathbf{w}$ is an eigenvector of $\mathcal{L}$ to the same eigenvalue λ or vanishes. Using (1.36),(1.37) and (1.38) one sees that

$$\left(\lambda, \begin{pmatrix} \mathbf{v} \\ \mathbf{0} \end{pmatrix} \right) \text{ is an eigenpair of } \mathbf{M} \text{ if and only if}$$
$$(\lambda, \mathbf{v}) \text{ is an eigenpair of } \mathbf{A} \text{ and } \mathbf{v} \in \ker(\mathbf{C}). \tag{1.46}$$

According to (1.37) we have

$$(\mathbf{A} - \lambda\mathbf{I}_N)\,\mathbf{v} = \text{-}\mathbf{B}\mathbf{w}. \tag{1.47}$$

Assuming $\mathbf{w} = 0$ turns (1.47) into an eigenproblem of $\mathbf{A}$ which may be solved to check for eigenvectors of $\mathbf{M}$ with vanishing lower vector block according to (1.46). Additionally, one can insert any eigenpair $(\lambda, \mathbf{w})$ of the associated lambda-matrix and solve the system, which becomes underdetermined for $\lambda \in \sigma(\mathbf{A})$, to find $\mathbf{v}$.

However, any $\lambda \notin \rho(a(x; \mathbf{A}))$ implies $\lambda \notin \sigma(\mathbf{A})$ hence non vanishing $\mathbf{w}$ and of course $a(\lambda) \neq 0$. Premultiplication of (1.47) by $p(\mathbf{A}, \lambda\mathbf{I}_N; a)$ and division by $a(\lambda)$ leads to

$$\mathbf{v} = \frac{1}{a(\lambda)}\,p(\mathbf{A}, \lambda; a)\,\mathbf{B}\mathbf{w}, \tag{1.48}$$

which directly relates $\mathbf{v}$ to $\mathbf{w}$. Note that $\mathbf{v}$ is not an eigenvector but a first eigenvector block, therefore the scaling by $\frac{1}{a(\lambda)}$ must not be omitted.

1.5 The Spectrum of a Principal Submatrix

Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ be a principal submatrix of $\mathbf{M} \in \mathbb{C}^{(N+s) \times (N+s)}$. Theorem 2 allows for the exploitation of an annihilating polynomial for $\mathbf{A}$ to determine

the spectrum of $\mathbf{M}$. Now we consider the converse, i.e. we will exploit an annihilating polynomial for $\mathbf{M}$ to obtain information about the spectrum of its principal submatrix $\mathbf{A}$. We largely maintain the notation of the previous sections and start with the 2×2 block matrix

$$\mathbf{M} = \left(\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{array} \right).$$

Let

$$\Pi_1 = \left(\begin{array}{cc} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array} \right) \qquad \text{and} \qquad \Pi_2 = \left(\begin{array}{cc} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \end{array} \right) \tag{1.49}$$

be block matrices which are partitioned conformably to the partition of $\mathbf{M}$. Let

$$\left(\mathbf{M}^k \right)_{ij} = \Pi_i \mathbf{M}^k \Pi_j \quad , \quad i, j = 1, 2 \tag{1.50}$$

denote the blocks of the conformally partitioned k-th power of $\mathbf{M}$. Using

$$\left(\mathbf{M}^{k+1} \right)_{ij} = \left(\mathbf{M}^k \right)_{i1} \left(\mathbf{M}^1 \right)_{1j} + \left(\mathbf{M}^k \right)_{i2} \left(\mathbf{M}^1 \right)_{2j} \, , \, i, j = 1, 2 \, , \, k \geq 0 \tag{1.51}$$

it is easy to prove that for $k \geq 1$

$$\left(\mathbf{M}^k \right)_{11} = \left(\mathbf{M}^{k-1} \right)_{11} \mathbf{A} + \sum_{i=0}^{k-2} \left(\mathbf{M}^i \right)_{22} \mathbf{B} \mathbf{D}^{k-2-i} \mathbf{C} \tag{1.52}$$

$$\left(\mathbf{M}^k \right)_{21} = \sum_{i=0}^{k-1} \left(\mathbf{M}^i \right)_{22} \mathbf{C} \mathbf{A}^{k-1-i}. \tag{1.53}$$

Note that each term in the expansion of the recursion on the right-hand side of (1.52) either ends on $\mathbf{C}$ followed by a power of $\mathbf{A}$, including $\mathbf{A}^0 = \mathbf{I}_N$, or is $\mathbf{A}^k$, and note that each summand on the right-hand side of (1.53) ends on $\mathbf{C}$ followed by a power of $\mathbf{A}$, again including $\mathbf{I}_N$.

Now, let $a(x; \mathbf{M}) = \sum_{k=0}^{m} r_k x^k$ be an annihilating polynomial for $\mathbf{M}$ and let $(\lambda, \mathbf{v})$ be an eigenpair of $\mathbf{A}$. Since $a(\mathbf{M}; \mathbf{M}) = \mathbf{0}$, we have

$$\left(\begin{array}{c} \mathbf{0} \\ \mathbf{0} \end{array} \right) = a(\mathbf{M}; \mathbf{M}) \left(\begin{array}{c} \mathbf{v} \\ \mathbf{0} \end{array} \right) = \left(\begin{array}{c} \sum_{k=0}^{m} r_k \left(\mathbf{M}^k \right)_{11} \mathbf{v} \\ \sum_{k=0}^{m} r_k \left(\mathbf{M}^k \right)_{21} \mathbf{v} \end{array} \right),$$

which gives the following two equations for $\mathbf{v}$.

$$\mathbf{0} = \sum_{k=0}^{m} r_k \left(\mathbf{M}^k \right)_{11} \mathbf{v} \tag{1.54}$$

$$\mathbf{0} = \sum_{k=0}^{m} r_k \left(\mathbf{M}^k \right)_{21} \mathbf{v} \tag{1.55}$$

If $\mathbf{v} \in \ker(\mathbf{C})$ then we obtain by (1.52) and (1.54)

$$\mathbf{0} = \mathrm{a}(\lambda; \mathbf{M})\,\mathbf{v},$$

Since we may have chosen $\mathrm{a}(x; \mathbf{M})$ to be the characteristic polynomial of $\mathbf{M}$ or one of its divisors, this implies that λ is an eigenvalue of $\mathbf{M}$. Thus,

$$\mathbf{v} \in \ker(\mathbf{C}) \Rightarrow \lambda \in \sigma(\mathbf{M}), \qquad (1.56)$$

which can also be deduced from (1.46). If $\mathbf{v} \notin \ker(\mathbf{C})$ then $\mathbf{C}\mathbf{v} \neq \mathbf{0}$ and we obtain by (1.53) and (1.55)

$$\mathbf{0} = \Pi_2\,\mathrm{a}(\mathbf{M}; \mathbf{M})\,\Pi_1 \mathbf{v} = \sum_{k=0}^{m} r_k \left(\mathbf{M}^k\right)_{21} \mathbf{v}$$

$$= \sum_{k=0}^{m} r_k \sum_{i=0}^{k-1} \left(\mathbf{M}^i\right)_{22} \mathbf{C}\mathbf{A}^{k-1-i} \mathbf{v} \qquad (1.57)$$

$$= \sum_{k=0}^{m} r_k \sum_{i=0}^{k-1} \left(\mathbf{M}^i\right)_{22} \lambda^{k-1-i} \mathbf{C}\mathbf{v} \qquad (1.58)$$

$$= \Pi_2\,\mathrm{p}(\mathbf{M}, \lambda; \mathrm{a})\,\Pi_2 \mathbf{C}\mathbf{v} \qquad (1.59)$$

$$= \sum_{k=0}^{m-1} \left(\sum_{i=0}^{m-1-k} r_{k+i+1} \left(\mathbf{M}^i\right)_{22} \right) \lambda^k \mathbf{C}\mathbf{v} \qquad (1.60)$$

and therefore,

$$\lambda \in \rho \left(\det \left[\sum_{k=0}^{m-1} \left(\sum_{i=0}^{m-1-k} r_{k+i+1} \left(\mathbf{M}^i\right)_{22} \right) \lambda^k \right] \right). \qquad (1.61)$$

For convenience we will use the abbreviations

$$\mathbf{K}_k = \mathbf{K}_k(\mathrm{a}, \mathbf{M}) = \sum_{i=0}^{m-1-k} r_{k+i+1} \left(\mathbf{M}^i\right)_{22}. \qquad (1.62)$$

and

$$\mathcal{K}(\lambda) = \sum_{k=0}^{m-1} \mathbf{K}_k \lambda^k = \Pi_2\,\mathrm{p}(\mathbf{M}, \lambda; \mathrm{a})\,\Pi_2. \qquad (1.63)$$

Thus,

$$\mathbf{v} \notin \ker(\mathbf{C}) \Rightarrow \lambda \in \rho\left(\det\left(\mathcal{K}(\lambda)\right)\right). \qquad (1.64)$$

We summarize that

$$\lambda \in \sigma(\mathbf{M}) \cup \rho\left(\det\left(\mathcal{K}(\lambda)\right)\right)$$

for any eigenpair $\left(\lambda, \mathbf{v}\right)$, hence for any eigenvalue of $\mathbf{A}$. Thus, the eigenvalue problem of $\mathbf{A}$ is 'reduced' to some extend to that of $\mathbf{M}$ and a latent root problem of the lambda-matrix $\mathcal{K}(\lambda)$ of size s. Theorem 3 below gives a stronger

result.

Note that $\mathcal{K}(\lambda)$ is monic (up to the factor r_m), which excludes infinite eigenvalues. More important for the next theorem is the fact that moniticity ensures regularity. Since $\mathcal{K}(\lambda)$ is of finite size and finite degree and monic, $\det(\mathcal{K}(\lambda))$ is a polynomial of finite degree at least 1, hence $\rho(\det(\mathcal{K}(\lambda)))$ is finite.

Theorem 3. *Let*

$$\mathbf{M} = \left(\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{array} \right)$$

be a 2×2 *block matrix of size* $(N + s)$ *with* $\mathbf{A} \in \mathbb{C}^{N \times N}$, $\mathbf{B} \in \mathbb{C}^{N \times s}$, $\mathbf{C} \in \mathbb{C}^{s \times N}$ *and* $\mathbf{D} \in \mathbb{C}^{s \times s}$. *Let*

$$\mathrm{a}(x; \mathbf{M}) = \sum_{k=0}^{m} r_k x^k$$

be an annihilating polynomial for $\mathbf{M}$. *Let* $\left(\mathbf{M}^k\right)_{22}$ *be the lower diagonal block of the conformally partitioned matrix* $\mathbf{M}^k$, *which is the k-th power of* $\mathbf{M}$, *and let*

$$\mathbf{K}_k = \sum_{i=0}^{m-1-k} r_{k+i+1} \left(\mathbf{M}^i\right)_{22}$$

$$\mathcal{K}(\lambda) = \sum_{k=0}^{m-1} \mathbf{K}_k \lambda^k.$$

Then

$$\sigma(\mathbf{A}) = \sigma(\mathbf{M}) - \rho^s(\mathrm{a}(\lambda; \mathbf{M})) + \rho(\mathcal{K}(\lambda)).$$

Proof. Let $M = (N + s)$ and let

$$\mathrm{p}(\lambda, \mathbf{M}; \mathrm{a}) = \left(\begin{array}{cc} \mathcal{P}_{11}(\lambda) & \mathcal{P}_{12}(\lambda) \\ \mathcal{P}_{21}(\lambda) & \mathcal{P}_{22}(\lambda) \end{array} \right)$$

where we employ a partition which is conformable to that of $\mathbf{M}$. Recall that $\mathcal{P}_{22}(\lambda) = \Pi_2 \, \mathrm{p}(\mathbf{M}, \lambda; \mathrm{a}) \, \Pi_2 = \mathcal{K}(\lambda)$. By the theorem of Cayley-Hamilton $\mathcal{K}(\lambda)$ with any fixed λ is annihilated by its characteristic polynomial

$$\mathrm{char}(x; \mathcal{K}(\lambda)) = \det(x\mathbf{I}_s - \mathcal{K}(\lambda))$$

where the coefficients of char depend on λ, and we have

$$\mathcal{K} \, \mathrm{p}(0, \mathcal{K}; \mathrm{char}) = -\mathrm{char}(0; \mathcal{K}) \, \mathbf{I}_s = (-1)^{s+1} \det(\mathcal{K}) \, \mathbf{I}_s. \qquad (1.65)$$

According to (1.32) and since $\mathrm{a}(\mathbf{M}; \mathbf{M}) = \mathbf{0}$, we have

$$(\lambda \mathbf{I}_M - \mathbf{M}) \left(\begin{array}{cc} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{K} \end{array} \right) = \mathrm{a}(\lambda; \mathbf{M}) \, \mathbf{I}_M. \qquad (1.66)$$

We postmultiply both sides with

$$\begin{pmatrix} (-1)^s \det(\mathcal{K}) \mathbf{I}_N & \mathbf{0} \\ \mathrm{p}(0, \mathcal{K}; \mathrm{char}) \mathcal{P}_{21} & \mathbf{I}_s \end{pmatrix} \tag{1.67}$$

and set

$$\mathcal{P}_0 = (-1)^s \det(\mathcal{K}) \mathcal{P}_{11} + \mathcal{P}_{12} \, \mathrm{p}(0, \mathcal{K}; \mathrm{char}) \mathcal{P}_{21}$$

reaching

$$(\lambda \mathbf{I}_M - \mathbf{M}) \begin{pmatrix} \mathcal{P}_0 & \mathcal{P}_{12} \\ \mathbf{0} & \mathcal{K} \end{pmatrix} = \mathrm{a}(\lambda; \mathbf{M}) \begin{pmatrix} (-1)^s \det(\mathcal{K}) \mathbf{I}_N & \mathbf{0} \\ \mathrm{p}(0, \mathcal{K}; \mathrm{char}) \mathcal{P}_{21} & \mathbf{I}_s \end{pmatrix}. \tag{1.68}$$

Since $\mathbf{K}_{m-1} = r_m \mathbf{I}_m$ for the leading coefficient of $\mathcal{K}$, we have $\det(\mathcal{K}) \not\equiv 0$, i.e. the matrix (1.67) is regular. The lower left matrix block in the second matrix on the left-hand side of (1.68) vanishes because of (1.65). From (1.68) we deduce

$$\det(\lambda \mathbf{I}_M - \mathbf{M}) \det(\mathcal{P}_0) \det(\mathcal{K}(\lambda)) = (\mathrm{a}(\lambda))^{(N+s)} (-1)^{sN} \det(\mathcal{K}(\lambda))^N \tag{1.69}$$

and by considering the upper left diagonal block

$$(\mathbf{I}_N \lambda - \mathbf{A}) \mathcal{P}_0 = \mathrm{a}(\lambda) (-1)^s \det(\mathcal{K}(\lambda)) \mathbf{I}_N,$$

hence,

$$\det(\mathbf{I}_N \lambda - \mathbf{A}) \det(\mathcal{P}_0) = (\mathrm{a}(\lambda))^N (-1)^{sN} \det(\mathcal{K}(\lambda))^N. \tag{1.70}$$

Multiplying (1.69) on either side by $\det(\mathbf{I}_N \lambda - \mathbf{A})$ and canceling terms which are equal according to (1.70) yields

$$\det(\lambda \mathbf{I}_M - \mathbf{M}) \det(\mathcal{K}(\lambda)) = (\mathrm{a}(\lambda))^s \det(\lambda \mathbf{I}_N - \mathbf{A}). \tag{1.71}$$

Considering the roots of the polynomials in (1.71) we find

$$\sigma(\mathbf{M}) + \rho(\det(\mathcal{K}(\lambda))) = \rho^s(\mathrm{a}(\lambda)) + \sigma(\mathbf{A}) \ , \ \forall \lambda \in \mathbb{C}$$

$$\square$$

Eigenvectors We can recover eigenpairs $(\lambda, \mathbf{v})$ of $\mathbf{A}$ with $\mathbf{v} \in \ker(\mathbf{C})$ from eigenpairs of $\mathbf{M}$ using (1.46). In order to do so, one has to consider the whole eigenspace to a given eigenvalue of $\mathbf{M}$, take into account the subspace of all eigenvectors whose first block is in $\ker(\mathbf{C})$ and remove the subspace of those whose first block vanishes.

For those eigenvectors which obey $\mathbf{v} \notin \ker(\mathbf{C})$ we can utilize (1.59) in the form of

$$\mathcal{K}(\lambda) \mathbf{w} = \mathbf{0} \ , \quad \mathbf{w} = \mathbf{C} \mathbf{v}. \tag{1.72}$$

A latent vector $\mathbf{w}$ of the lambda-matrix $\mathcal{K}$ to a given latent root can be inserted in the second equation of (1.72) in order to solve it for eigenvectors $\mathbf{v}$ of $\mathbf{A}$.

1.6 Generalization to Lambda-Matrices, Isospectral Graph Reduction and Further Generalizations

Our previous results in particular theorem 2 which hold for complex matrices will be generalized to matrices the entries of which are univariate functions. We remind the reader of the formal difference between an eigenvalue and a latent root. One instance of our notation may need further description. Since the annihilating polynomial of a lambda-matrix has coefficients which are polynomials in lambda, it is an element of $(\mathbb{C}[\lambda])[x]$. However, when the indeterminate x is replaced by an element of $\mathbb{C}[\lambda]$, we interpret the resulting function as an element of $\mathbb{C}[\lambda]$.

1.6.1 The Main Theorem in the Case of Lambda-Matrices

Theorem 4. *Let $\mathcal{A}(\lambda)$ be a lambda-matrix of size N. Let $f(\lambda)$ be a polynomial in λ with complex coefficients. Set*

$$\mathcal{A}_1(\lambda) = \mathcal{A}(\lambda) + f(\lambda)\,\mathbf{I}_N.$$

Let

$$\mathrm{a}\left(x; \mathcal{A}_1(\lambda)\right) = \sum_{k=0}^{m} r_k(\lambda)\, x^k$$

be an annihilating polynomial of $\mathcal{A}_1$ with coefficients in $\mathbb{C}[\lambda]$. Let $\mathcal{A}$ be the left upper principal submatrix of

$$\mathcal{M}(\lambda) = \begin{pmatrix} \mathcal{A}(\lambda) & \mathcal{B}(\lambda) \\ \mathcal{C}(\lambda) & \mathcal{D}(\lambda) \end{pmatrix},$$

which is a lambda-matrix of size $N + s$. Let

$$\bar{a}(\lambda) = \mathrm{a}\left(f(\lambda), \mathcal{A}_1(\lambda)\right) \in \mathbb{C}[\lambda].$$

and let the associated lambda-matrix be given as

$$\mathcal{L}(\lambda) = \mathcal{C}(\lambda)\,\mathrm{p}\left(\mathcal{A}_1(\lambda), f(\lambda); \mathrm{a}\right)\mathcal{B}(\lambda) + \bar{a}(\lambda)\,\mathcal{D}(\lambda).$$

Then

$$\rho\left(\det(\mathcal{M})\right) = \rho\left(\det(\mathcal{A})\right) - \rho^s(\bar{a}) + \rho\left(\det(\mathcal{L})\right).$$

Proof.

$$
\begin{aligned}
0 &= \det\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \det\left(\mathrm{a}\left(f, \mathcal{A}_1\right)\mathbf{I}_s\right) \\
&= \det\begin{pmatrix} \mathcal{A}_1 - f\mathbf{I}_N & \mathrm{a}\left(f, \mathcal{A}_1\right)\mathcal{B} \\ \mathcal{C} & \mathrm{a}\left(f, \mathcal{A}_1\right)\mathcal{D} \end{pmatrix} \det\begin{pmatrix} \mathbf{I}_N & \mathrm{p}\left(\mathcal{A}_1, f; \mathrm{a}\right)\mathcal{B} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} \\
&= \det\begin{pmatrix} \mathcal{A} & \mathbf{0} \\ \mathcal{C} & \left[\mathcal{C}\,\mathrm{p}\left(\mathcal{A}_1, f; \mathrm{a}\right)\mathcal{B} + \mathrm{a}\left(f, \mathcal{A}_1\right)\mathcal{D}\right] \end{pmatrix} \\
&= \det(\mathcal{A})\det(\mathcal{L})
\end{aligned}
$$

and therefore,

$$\rho\left(\det\left(\mathcal{M}\right)\right) + \rho^s\left(\mathrm{a}\left(f, \mathcal{A}_1\right)\right) = \rho\left(\det\left(\mathcal{A}\right)\right) + \rho\left(\det\left(\mathcal{L}\right)\right).$$

$\square$

Since we may choose $\mathcal{D}$ to have any degree and any leading coefficient, the leading coefficient of $\mathcal{L}$ may be singular. Putting that aside, theorem 2 can be seen as a corollary of theorem 4. The freedom in decomposing $\mathcal{A}$ as a sum allows for some flexibility in finding suitable annihilating polynomials. After the next subsection we give two further possible generalizations of such a decomposition.

1.6.2 Principal Submatrices of Lambda-Matrices

Theorem 5. *Let $\mathcal{A}\left(\lambda\right)$, a lambda-matrix of size N, be a principal submatrix of the lambda-matrix $\mathcal{M}\left(\lambda\right)$, which is a 2×2 block matrix of size $M = \left(N+s\right)$ partitioned as*

$$\mathcal{M}\left(\lambda\right) = \left(\begin{array}{cc} \mathcal{A}\left(\lambda\right) & \mathcal{B}\left(\lambda\right) \\ \mathcal{C}\left(\lambda\right) & \mathcal{D}\left(\lambda\right) \end{array}\right).$$

Let $f\left(\lambda\right)$ be a polynomial in λ with complex coefficients. Set

$$\mathcal{M}_1\left(\lambda\right) = \mathcal{M}\left(\lambda\right) + f\left(\lambda\right)\mathbf{I}_M. \tag{1.73}$$

Let

$$\mathrm{a}\left(x; \mathcal{M}_1\left(\lambda\right)\right) = \sum_{k=0}^{m} r_k\left(\lambda\right) x^k$$

be an annihilating polynomial of $\mathcal{M}_1$ with coefficients in $\mathbb{C}\left[\lambda\right]$. Let

$$\mathrm{p}\left(x, \mathcal{M}_1\left(\lambda\right); \mathrm{a}\right) = \left(\begin{array}{cc} \mathcal{P}_{11}\left(x, \lambda\right) & \mathcal{P}_{12}\left(x, \lambda\right) \\ \mathcal{P}_{21}\left(x, \lambda\right) & \mathcal{P}_{22}\left(x, \lambda\right) \end{array}\right) \tag{1.74}$$

be partitioned conformable to $\mathcal{M}$, abbreviate

$$\bar{a}\left(\lambda\right) = \mathrm{a}\left(f\left(\lambda\right); \mathcal{M}_1\left(\lambda\right)\right) \text{ and } \mathcal{K}\left(\lambda\right) = \mathcal{P}_{22}\left(f\left(\lambda\right), \lambda\right)$$

and let $\mathcal{K}\left(\lambda\right)$ be regular. Then

$$\rho\left(\mathcal{A}\right) = \rho\left(\mathcal{M}\right) - \rho^s\left(\bar{a}\right) + \rho\left(\mathcal{K}\right).$$

Proof. Since

$$\left(\mathcal{M}_1 - f\mathbf{I}_M\right)\mathrm{p}\left(f, \mathcal{M}_1; \mathrm{a}\right) = \text{-}\,\mathrm{a}\left(f; \mathcal{M}_1\right)\mathbf{I}_M,$$

we have

$$\mathcal{M}\left(\begin{array}{cc} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{K} \end{array}\right) = \text{-}\bar{a}\mathbf{I}_M. \tag{1.75}$$

Since the lambda-matrix $\mathcal{K}$ is of finite size, of finite degree and regular, its root set is of finite size. For the following we fix a λ which obeys

$$\lambda \notin \rho(\mathcal{K}). \tag{1.76}$$

Thus, $\mathcal{K}(\lambda)$ is invertible. We postmultiply both sides of (1.75) with

$$\begin{pmatrix} \mathbf{I}_N & \mathbf{0} \\ -\mathcal{K}^{-1}\mathcal{P}_{21} & \mathbf{I}_s \end{pmatrix} \tag{1.77}$$

and set

$$\mathcal{P}_0 = \mathcal{P}_{11} - \mathcal{P}_{12}\mathcal{K}^{-1}\mathcal{P}_{21}.$$

This yields

$$\mathcal{M} \begin{pmatrix} \mathcal{P}_0 & \mathcal{P}_{12} \\ \mathbf{0} & \mathcal{K} \end{pmatrix} = -\bar{a} \begin{pmatrix} \mathbf{I}_N & \mathbf{0} \\ -\mathcal{K}^{-1}\mathcal{P}_{21} & \mathbf{I}_s \end{pmatrix}. \tag{1.78}$$

From (1.78) we deduce

$$\det(\mathcal{M})\det(\mathcal{P}_0)\det(\mathcal{K}) = (-1)^M \bar{a}^M \tag{1.79}$$

and by considering the upper left diagonal block

$$\det(\mathcal{A})\det(\mathcal{P}_0) = (-1)^N \bar{a}^N. \tag{1.80}$$

Multiplying (1.79) on either side by $\det(\mathcal{A})$ and canceling terms which are equal according to (1.80) yields

$$\det(\mathcal{M}(\lambda))\det(\mathcal{K}(\lambda)) = (-1)^s (\bar{a}(\lambda))^s \det(\mathcal{A}(\lambda)). \tag{1.81}$$

We have proven that equation (1.81) is valid for all λ obeying assumption (1.76). Since both sides of (1.81) are finite polynomials, we may drop this assumption. Hence,

$$\rho(\mathcal{M}) + \rho(\det(\mathcal{K})) = \rho^s(\bar{a}) + \rho(\mathcal{A}) \ , \ \forall \lambda \in \mathbb{C}$$

$\square$

Note the slight difference in the proof of theorem 5 compared to that of theorem 3, which is actually a corollary. A similar argument is used in the next subsection.

1.6.3 Isospectral Graph Reduction

In this subsection we consider a formal extension of theorem 4 which unifies our approach and a result by Bunimovich and Webb [14], which we reformulated as theorem 1. We call a matrix $\mathbf{X}(\lambda)$ the entries of which are functions in λ regular if $\exists \lambda$ s.t. $\mathbf{X}(\lambda)$ is defined and $\det(\mathbf{X}(\lambda)) \neq 0$.

Lemma 1. *Let $\mathbb{W}$ be the set of all complex rational functions in λ. Let $\mathcal{A} \in \mathbb{W}^{N \times N}$, $\mathcal{B} \in \mathbb{W}^{N \times s}$, $\mathcal{C} \in \mathbb{W}^{s \times N}$, $\mathcal{D} \in \mathbb{W}^{s \times s}$ and $\mathcal{M} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix}$. Let $\mathcal{A}$ and $\mathcal{M}$ be regular. Find a scalar function $f(x)$, matrices $\mathcal{A}_i \in \mathbb{W}^{N \times N}$, $i \in \{1, 2, 3\}$ and a polynomial $\mathrm{a}(x)$ with coefficients in $\mathbb{W}$ s.t.*

- *$\mathcal{A}_1 + \mathcal{A}_2 = \mathcal{A}$*

- *$\mathcal{A}_3$ is regular*

- *$\mathrm{a}\left(\mathcal{A}_3^{-1}\mathcal{A}_1\right) = \mathbf{0}_N$*

- *$\mathcal{A}_3^{-1}\mathcal{A}_2 = -f(\lambda)\mathbf{I}_N$.*

Then

$$\rho\left(\det\left(\mathcal{M}\right)\right) + \rho^s\left(\mathrm{a}\left(f\left(\lambda\right)\right)\right) - \rho\left(\det\left(\mathcal{A}\right)\right) =$$
$$\rho\left(\det\left[\mathcal{C}\,\mathrm{p}\left(\mathcal{A}_3^{-1}\mathcal{A}_1, f\left(\lambda\right); \mathrm{a}\right)\mathcal{A}_3^{-1}\mathcal{B} + \mathrm{a}\left(f\left(\lambda\right)\right)\mathcal{D}\right]\right)$$

with the possible exception of elements of

$$\mathcal{N} = \left\{\lambda \mid \mathcal{A}_3^{-1} \text{ or } \mathcal{A}_3 \text{ or } \mathcal{A}_1 \text{ or } \mathcal{M} \text{ or } \mathrm{a}\left(f\left(\lambda\right)\right) \text{ is not defined}\right\}.$$

Proof. We use the abbreviations $\tilde{\mathcal{X}} = \mathcal{A}_3^{-1}\mathcal{X}$ and $\tilde{a}(\lambda) = \mathrm{a}\left(f(\lambda)\right)$. Assume that $\mathcal{M}$ and $\mathcal{A}_1$ are defined. Assume that $\mathcal{A}_3$ is invertible, i.e. $\mathcal{A}_3$ is defined and $\det\left(\mathcal{A}_3\right) \neq 0$. Assume that $\bar{a}(\lambda)$ is defined. For all λ not violating those assumptions the following equations are valid

$$0 = \det\left(\mathcal{M}\right)\det\left(\tilde{a}\mathbf{I}_s\right)$$
$$= \det\left(\mathcal{A}_3\right)\det\begin{pmatrix} \tilde{\mathcal{A}}_1 - f\mathbf{I}_N & \tilde{a}\tilde{\mathcal{B}} \\ \mathcal{C} & \tilde{a}\mathcal{D} \end{pmatrix}$$
$$= \det\left(\mathcal{A}_3\right)\det\begin{pmatrix} \tilde{\mathcal{A}}_1 - f\mathbf{I}_N & \tilde{a}\tilde{\mathcal{B}} \\ \mathcal{C} & \tilde{a}\mathcal{D} \end{pmatrix}\det\begin{pmatrix} \mathbf{I}_N & \mathrm{p}\left(\tilde{\mathcal{A}}_1, f; \mathrm{a}\right)\tilde{\mathcal{B}} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix}$$
$$= \det\left(\mathcal{A}_3\right)\det\begin{pmatrix} \tilde{\mathcal{A}}_1 - f\mathbf{I}_N & \mathbf{0} \\ \mathcal{C} & \mathcal{C}\,\mathrm{p}\left(\tilde{\mathcal{A}}_1, f; \mathrm{a}\right)\tilde{\mathcal{B}} + \tilde{a}\mathcal{D} \end{pmatrix}$$
$$= \det\begin{pmatrix} \mathcal{A} & \mathbf{0} \\ \mathcal{C} & \mathcal{C}\,\mathrm{p}\left(\tilde{\mathcal{A}}_1, f; \mathrm{a}\right)\tilde{\mathcal{B}} + \tilde{a}\mathcal{D} \end{pmatrix}.$$

$\square$

The approach of Bunimovich and Webb arises with the assumption that

$$\mathcal{A} \text{ is a triangular matrix (up to a suitable permutation)}$$

and the choice

$$f \equiv 1, \ \mathcal{A}_2 = \text{diag}\left(\mathcal{A}\right), \ \mathcal{A}_3 = \text{-}\mathcal{A}_2 \text{ and } a\left(x\right) = x^n.$$

Theorem 4 arises with the restriction

$$\mathcal{M} \in \mathbb{C}\left[\lambda\right]^{(N+s)\times(N+s)}$$

and the choice

$$f\left(\lambda\right) \in \mathbb{C}\left[\lambda\right], \ \mathcal{A}_2 = \text{-}f\left(\lambda\right)\mathbf{I}_N, \ \mathcal{A}_3 = \mathbf{I}_N \text{ and } a\left(x\right) = a\left(x; \mathcal{A} + f\left(\lambda\right)\mathbf{I}_N\right)$$

That is, $a\left(x\right)$ may be any annihilating polynomial of $\mathcal{A} + f\left(\lambda\right)\mathbf{I}_N$.

Example Let $\mathcal{A} \in \mathbb{W}^{N \times N}$ be principal submatrix of $\mathcal{M} = \left(\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right) \in \mathbb{W}^{(N+s)\times(N+s)}$. Let $\mathcal{A}$ be regular and decomposable as

$$\mathcal{A}\left(\lambda\right) = w\left(\lambda\right)\mathbf{J} - \mathcal{T}\left(\lambda\right) \tag{1.82}$$

where $w\left(\lambda\right) \in \mathbb{W}$, $\mathbf{J} = \mathbf{j}_N \mathbf{j}'_N$ is the matrix with all entries 1 and $\mathcal{T} = t_{vv}\delta_{vw}$ is a diagonal matrix with entries $0 \neq t_{vv}\left(\lambda\right) \in \mathbb{W}$. Let

$$g\left(\lambda\right) = \sum_v \frac{1}{t_{vv}\left(\lambda\right)}. \tag{1.83}$$

One easily verifies that $a\left(w\mathcal{T}^{-1}\mathcal{J}; w\mathcal{T}^{-1}\mathcal{J}\right) = \mathbf{0}$ for

$$a\left(x; w\mathcal{T}^{-1}\mathbf{J}\right) = x^2 - g\left(\lambda\right)w\left(\lambda\right)x. \tag{1.84}$$

Finally, let

$$\mathcal{L} = \mathcal{C}\left(w\mathcal{T}^{-1}\mathbf{J}\mathcal{T}^{-1} + \left(1 - gw\right)\mathcal{T}^{-1}\right)\mathcal{B} + \left(1 - gw\right)\mathcal{D}. \tag{1.85}$$

Then, according to lemma 1, we have

$$\rho\left(\det\left(\mathcal{M}\right)\right) + \rho\left(1 - gw\right) = \rho\left(\det\left(\mathcal{A}\right)\right) + \rho\left(\det\left(\mathcal{L}\right)\right). \tag{1.86}$$

1.6.4 Further Generalizations

Natural generalizations may employ other properties of determinants in order to reduce the size of eigenproblems. An example is the well known rule for equally sized $N \times N$ matrices blocks $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$. Assuming

$$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}. \tag{1.87}$$

we have

$$\det\left(\mathcal{M}\right) = \det\left(\begin{array}{cc} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{array}\right) = \det\left(\mathcal{D}\mathcal{A} - \mathcal{C}\mathcal{B}\right). \tag{1.88}$$

If $\mathcal{M}$ is, for instance, a standard monic linear eigenproblem, then (1.88) transforms it into a monic quadratic eigenproblem.Although, this might be generalized to other pairs of commuting blocks and to $k \times k$ block matrices, the assumptions are quite restrictive. In particular, they depend on the off-diagonal block $\mathcal{B}$.

As a conclusion for this section we consider a generalization which may point out the main aspects of theorem 4 and lemma 1. We consider $\mathcal{M} = \left(\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix} \right)$ of size $(N + s)$. W.l.o.g. we aim at canceling the right upper off-diagonal block using properties of $\mathbf{A}$ but imposing as little assumptions on the other blocks as possible. As an ansatz we pre- and postmultiply with conformally partitioned matrices with non singular diagonal blocks s.t.

$$\begin{pmatrix} \mathcal{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_N & \mathcal{Y} \\ \mathbf{0} & \mathcal{Z} \end{pmatrix} = \begin{pmatrix} \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{A}\mathcal{Y} + \mathcal{X}\mathcal{B}\mathcal{Z} \\ \mathcal{C} & \mathcal{C}\mathcal{Y} + \mathcal{D}\mathcal{Z} \end{pmatrix}. \quad (1.89)$$

Thus, in case of

$$\mathcal{X}\mathcal{A}\mathcal{Y} + \mathcal{X}\mathcal{B}\mathcal{Z} = \mathbf{0} \quad (1.90)$$

one has

$$\det(\mathcal{M}) \det(\mathcal{Z}) = \det(\mathcal{A}) \det(\mathcal{C}\mathcal{Y} + \mathcal{D}\mathcal{Z}). \quad (1.91)$$

Although, there is some flexibility in choosing $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$, it is not entirely trivial which conditions on $\mathcal{A}$ might be useful and there is still a dependency on $\mathcal{B}$. However, $\mathcal{A} = q(\lambda)\mathbf{I}_N$ with arbitrary $q(\lambda)$, for instance, allows for $\mathcal{Y} = \mathcal{B}$ and $\mathcal{Z} = q(\lambda)\mathbf{I}_s$.

Canceling B   In order to get rid of any dependency on $\mathcal{B}$ in solving (1.90) one can set

$$\mathcal{Z} = -h(\lambda)\mathbf{I}_s \quad (1.92)$$

with arbitrary non zero function h, which allows to impose the factorization

$$\mathcal{Y} = \mathcal{R}\mathcal{X}\mathcal{B} \quad (1.93)$$

and the following equation for $h(\lambda)$, $\mathcal{R}$ and $\mathcal{X}$

$$\mathcal{X}\mathcal{A}\mathcal{R} = h(\lambda)\mathbf{I}_N. \quad (1.94)$$

Since $h(\lambda)$ and $\mathcal{X}$ have to be regular, $\mathcal{A}\mathcal{R}$ must not be singular.

Smith Form   (1.94) is for example solvable via the Smith form. One can employ a factorization of $\mathcal{X}$ and $\mathcal{R}$ as follows

$$\mathcal{X}\mathcal{A}\mathcal{R} = \mathcal{X}_1\mathcal{X}_2\mathcal{A}\mathcal{R}_1\mathcal{R}_2 = \mathcal{X}_1\mathcal{S}\mathcal{R}_2 \quad (1.95)$$

If $\mathcal{A}$ is regular, there are unimodular matrices $\mathcal{X}_2$ and $\mathcal{R}_1$ s.t. $\mathcal{S} = \mathcal{X}_2\mathcal{A}\mathcal{R}_1$ is a diagonal matrix s.t. the diagonal element with largest degree, say $g(\lambda)$, is

divided by all other diagonal elements, g_{ii}. Thus, one may choose h to be a multiple of g, say

$$h(\lambda) = q(\lambda) g(\lambda) \tag{1.96}$$

and $\mathcal{X}_1$ and $\mathcal{R}_2$ to be diagonal matrices with diagonal entries x_{ii} and r_{ii} s.t.

$$x_{ii}(\lambda) g_{ii}(\lambda) r_{ii}(\lambda) = h(\lambda). \tag{1.97}$$

The General Adjoint Technique In order to apply the adjoint technique to equation (1.94) one may factorize

$$\mathcal{R} = \mathcal{U}\mathcal{P}. \tag{1.98}$$

The crucial step is decomposing $\mathcal{X}\mathcal{A}\mathcal{U}$ as

$$\mathcal{X}\mathcal{A}\mathcal{U} = \mathcal{F} - f(\lambda) \mathbf{I}_N \tag{1.99}$$

with arbitrary $f(\lambda)$ and finally introducing an annihilating polynomial $\mathrm{a}(x; \mathcal{F})$ for $\mathcal{F}$ of degree $m > 0$ s.t. $\mathrm{a}(\mathcal{F}; \mathcal{F}) = \mathbf{0}$.
The quadruple $(\mathcal{X}, \mathcal{U}, f(\lambda), \mathrm{a}(x))$ provides a solution for (1.90) via

$$\mathcal{Y} = \mathcal{U}\,\mathrm{p}(x; \mathrm{a})\,\mathcal{X}\mathcal{B} \text{ and } \mathcal{Z} = \mathrm{a}(f(\lambda))\,\mathbf{I}_s.$$

However, it is easy to see that all quadruples $(\mathcal{X}_1, \mathcal{X}_2, f(\lambda), \mathrm{a}(x))$ with $\mathcal{X}_2\mathcal{X}_1 = \mathcal{X}$ are equivalent.
In conclusion, one may find a reduction of $\mathcal{M}$ via a regular principal submatrix $\mathcal{A}$ if one finds a regular matrix $\mathcal{X}$, a shift factor $f(\lambda)$ and a polynomial $\mathrm{a}(x)$ the coefficient of which are functions in λ s.t. $\mathrm{a}(\mathcal{X}\mathcal{A} + f(\lambda)\mathbf{I}_N) = \mathbf{0}$.
Thus, in a sense, the ansatz (1.89) corresponds to lemma 1. It would be interesting to find other ways to solve (1.90) with or without restrictions on $\mathcal{B}$. It may also be worth to exploit more complex pre- and postmultipliers in (1.89), possibly with mild conditions on $\mathcal{B}$, $\mathcal{C}$ or $\mathcal{D}$.

1.7 Application

The main advantage of our method is the reduction of eigenproblems w.r.t. their size. A feasible application is given in the work of Bunimovich and Webb, who use their reduction method to possibly sharpen spectral bounds for matrices with coefficients in $\mathbb{W}[\lambda]$. The initial motivation for us is the reduction of costs for computing the spectrum of a complex matrix $\mathbf{M} \in \mathbb{C}^{(N+s)\times(N+s)}$ by exploiting a large highly structured principal submatrix $\mathbf{A} \in \mathbb{C}^{N\times N}$. We call $\mathbf{A}$ large if s is small compared to N and highly structured if the degree m of its minimal polynomial (or any other known annihilating polynomial) is small, which implies eigenvalues of high (geometric) multiplicity. A rough quantitative criterion for the feasibility of our method is provided by

$$\kappa < (N + s) \tag{1.100}$$

where

Definition 3. $\kappa := (m+1)\, s$.

In this case, the size of the eigenvalue problem posed by the linearization of the associated lambda-matrix $\mathcal{L}$, i.e. κ, is smaller than that of the eigenproblem of $\mathbf{M}$. κ may be used as a general measure of reduction for given s and N and we will do so in our considerations below. However, κ clearly does not capture important computational aspects such as numerical stability, sparsity or hermiticity. For instance, as already pointed out, the linearization of an hermitian eigenproblems will in general not be hermitian. Which is a great disadvantage since available algorithms are more numerous, faster and more reliable for hermitian problems. However, this might be compensated if κ is small enough.

Costs for Constructing the Associated Lambda-Matrix We aim at an upper bound for the costs of computing the coefficients $\mathbf{A}_k$ of the associated lambda-matrix $\mathcal{L}$ by counting the number of elementary FLoating-point OPerations (FLOP) without distinction between summation and multiplication, which are the only operations needed. The matrix-manipulations are considered in the straight forward way without refering to sophisticated methods such as the Strassen algorithm or exploitation of sparsity.

The costs for computing $\mathbf{A}^{i+1}\mathbf{B}$ given $\mathbf{A}$ and $\mathbf{A}^i\mathbf{B}$ are less then $2sN^2$ FLOPs. The costs for multiplying $\mathbf{C}$ and $\mathbf{A}^i\mathbf{B}$ are less then $2s^2N$ FLOPs. Thus the costs for computing the m matrices $\mathbf{CA}^i\mathbf{B}$ are less then $2sN\,(m\,(N+s) - N)$. The costs for a linear combination of $\mathbf{D}$ and $\mathbf{I}$ are $s^2 + 2s$ FLOPs. Since the costs for multiplying a matrix of size $s \times s$ with a constant and adding the result to another matrix of the same size are $2s^2$ FLOPs, the additional costs for computing $\mathbf{A}_{k \leq m}$ are $2\,(m-k)\,s^2$ FLOPs.

Therefore, the costs for computing all coefficients of the associated lambda-matrix are less then

$$
\begin{aligned}
\text{cost} &= 2sN\,(m\,(N+s) - N) + \sum_{k=0}^{m} \left[\left(s^2 + 2s\right) + 2\,(m-k)\,s^2 \right] \\
&= 2s\,(m+1)\,N^2 - 4sN^2 + 2s^2 mN + (m+1)\,s\,\left[2 + (m+1)\,s\right] \\
&= 2\kappa N^2 - 4sN^2 + 2s\kappa N - 2s^2 N + \kappa^2 + 2s\kappa \\
&= 2\,(\kappa - 2s)\,N^2 + 2s\,(\kappa - s)\,N + O\left(\kappa^2\right).
\end{aligned}
\tag{1.101}
$$

When (1.100) applies, we get cost $\leq O\left(\kappa N^2\right)$. The costs for solving, for instance by linearization, the polynomial eigenproblem posed by $\mathcal{L}$ is $O(\kappa^3)$ FLOPs, whereas finding the spectrum of $\mathbf{M}$ amounts for $O\left((N+s)^3\right)$ FLOPs. Thus, for small κ the costs for constructing $\mathcal{L}$ dominate those for solving it and are less than the costs for solving the original eigenproblem. This curios effect may be illustrated by the example given above

$$
\mathbf{M} = \begin{pmatrix} \mathbf{0} & \mathbf{b} \\ \mathbf{c}' & d \end{pmatrix} \ , \ \mathcal{L} = \begin{pmatrix} d & \mathbf{c}'\mathbf{b} \\ 1 & 0 \end{pmatrix}
\tag{1.102}
$$

Here, the linearization of $\mathcal{L}$ can be solved exact by hand. However the construction is linear in N due to the scalar product $\mathbf{c}'\mathbf{b}$.

The storage costs for $\mathcal{L}$ are obviously of order ms^2. We implicitly assumed that we temporarily keep all matrices $\mathbf{A}^i\mathbf{B}$ while constructing $\mathcal{L}$. This amounts for additional storage of order msN. Thus, when (1.100) applies, the total storage costs are of order κN, whereas the storage for $\mathbf{M}$ is of order $(N+s)^2$. The analysis shows that our method is efficient if (1.100) applies.

Cliques and Independent Sets   Our method crucially depends on feasible principal submatrices which are hard to detect for an arbitrary matrix of size $M = (N + s)$. Even the modest assumption, $s = 1$, yields M possible candidates. However, assuming $m = 1$ might be very useful for sparse matrices. According to (1.100) our method efficiently applies if there is a set of more then $\frac{1}{2}M$ nodes which are independent and without loops i.e. which induce an empty principal submatrix. The initial eigenproblem then reduces to a monic quadratic lambda-matrix the linearization of which has size $2s$ as shown in our first example in section 1.3.1. We call such a set a *large independent set*. Note that proper subsets of a large independent sets may still be large independent sets. Although finding a *maximum independent set* is a hard task and still bad approximated, even a simple greedy algorithm might be practicle in the more modest task of finding a large independent set if such sets exist and are sufficiently large [2], [78], [19].

A similar reasoning applies to cliques in a graph, which are combinatorially equivalent to independent sets on its complement. However, exploiting them to find the spectrum of an adjacency matrix easier requires cliques of size more then $\frac{2}{3}M$ since here $m = 2$. Thus, cliques and independent sets are not equivalence w.r.t. their spectral properties. As already discussed in section 1.3.2, we can circumvented that asymmetry by similarity deflation of the eigenvector $\mathbf{j}_N$ of the clique and repartition or, in the case of a regular graph, by just considering its complement, which both amounts to reducing the degree of the minimal polynomial.

Bound on Multiplicities   According to theorem 2 and 4 and the notation therein we have the following two relations

$$\sigma\left(\mathbf{M}\right) + \rho^s\left(\mathrm{a}\left(\lambda;\mathbf{A}\right)\right) = \sigma\left(\mathbf{A}\right) + \rho\left(\det\left(\mathcal{L}\left(\lambda\right)\right)\right) \tag{1.103}$$

$$\sigma\left(\mathbf{A}\right) + \rho^s\left(\mathrm{a}\left(\lambda;\mathbf{M}\right)\right) = \sigma\left(\mathbf{M}\right) + \rho\left(\det\left(\mathcal{K}\left(\lambda\right)\right)\right). \tag{1.104}$$

We assume that $\mathrm{a}\left(\lambda;\mathbf{A}\right)$ and $\mathrm{a}\left(\lambda;\mathbf{M}\right)$ are the minimal polynomials of their corresponding matrices. Let λ denote an arbitrary eigenvalue of $\mathbf{A}$ or $\mathbf{M}$, let $c_\lambda\left(\mathbf{A}\right)$ $(c_\lambda\left(\mathbf{M}\right))$ denote its exponent in the minimal polynomial of $\mathbf{A}$ $(\mathbf{M})$, which is zero if it is not an eigenvalue, and let $d_\lambda\left(\mathcal{L}\right)$ and $d_\lambda\left(\mathcal{K}\right)$ denote its multiplicity in the set of latent roots of those two lambda-matrices. Then we

have

$$m_\lambda\left(\mathbf{M}\right) + sc_\lambda\left(\mathbf{A}\right) = m_\lambda\left(\mathbf{A}\right) + d_\lambda\left(\mathcal{L}\right) \geq 0 \tag{1.105}$$

$$m_\lambda\left(\mathbf{A}\right) + sc_\lambda\left(\mathbf{M}\right) = m_\lambda\left(\mathbf{M}\right) + d_\lambda\left(\mathcal{K}\right) \geq 0. \tag{1.106}$$

Adding both equations yields

$$sc_\lambda\left(\mathbf{A}\right) + sc_\lambda\left(\mathbf{M}\right) = d_\lambda\left(\mathcal{L}\right) + d_\lambda\left(\mathcal{K}\right) \tag{1.107}$$

Since $d_\lambda\left(\mathcal{L}\right)$ and $d_\lambda\left(\mathcal{K}\right)$ are nonnegative, we have

$$m_\lambda\left(\mathbf{M}\right) - m_\lambda\left(\mathbf{A}\right) = d_\lambda\left(\mathcal{L}\right) - sc_\lambda\left(\mathbf{A}\right) \leq sc_\lambda\left(\mathbf{M}\right) \tag{1.108}$$

$$m_\lambda\left(\mathbf{A}\right) - m_\lambda\left(\mathbf{M}\right) = d_\lambda\left(\mathcal{K}\right) - sc_\lambda\left(\mathbf{M}\right) \leq sc_\lambda\left(\mathbf{A}\right) \tag{1.109}$$

hence,

$$\left|m_\lambda\left(\mathbf{M}\right) - m_\lambda\left(\mathbf{A}\right)\right| \leq s\max\left(c_\lambda\left(\mathbf{M}\right), c_\lambda\left(\mathbf{A}\right)\right). \tag{1.110}$$

which is a generalization of corollary 1. As an application one can think of an eigenvalue λ s.t $c_\lambda\left(\mathbf{M}\right) \leq 1$. If $\mathbf{A}$ is a principal submatrix which happens to be hermitian, i.e. $c_\lambda\left(\mathbf{A}\right) \leq 1$, then we have $\left|m_\lambda\left(\mathbf{M}\right) - m_\lambda\left(\mathbf{A}\right)\right| \leq s$. The author is not aware of the same or better bounds elsewhere.

Shrinking the Degree of the Annihilating Polynomial and Deflation of Eigenpairs According to (1.100) it is an advantage if m is small. Therefore, it is desirable to decrease the degree m of a $(x; \mathbf{A})$, the given annihilating polynomial for $\mathbf{A}$. One way to do so is using $\sigma\left(\mathbf{A}\right)$ and $\rho\left(\mathrm{a}\right)$ to cancel linear factors of a s.t. no multiplicity of a given root exceeds the corresponding eigenvalue multiplicity. Of course, for diagonalizable matrices the spectrum already determines the minimal polynomial, which is, in a sense, the optimal choice for a. However, for further reduction we aim at exploiting given eigenpairs of $\mathbf{A}$ to reduce $\kappa = (m + 1)s$.

Let $\mathbf{v}$ be a real eigenvector of the real matrix $\mathbf{X}$ with eigenvalue λ. In order to exploit this eigenpair we can employ, for instance, the following Householder transformation

$$\mathbf{H} = \mathbf{I} - 2\frac{\left(\mathbf{v} + \beta\left|\mathbf{v}\right|\mathbf{e}\right)\left(\mathbf{v} + \beta\left|\mathbf{v}\right|\mathbf{e}\right)'}{\left(\mathbf{v} + \beta\left|\mathbf{v}\right|\mathbf{e}\right)'\left(\mathbf{v} + \beta\left|\mathbf{v}\right|\mathbf{e}\right)} \tag{1.111}$$

where $\beta \in \{-1, 1\}$ is arbitrary and $\mathbf{e}$ is the indicator vector for the position on the diagonal where the deflated eigenvalue will occur, usually the first or the last one. This generalizes to the complex case (compare section 2.2.1) and to the case of deflating several eigenpairs at once with minor adjustments. In fact, it is well known that for a complex square matrix $\mathbf{X}$ with eigenvector λ there is always an invertible transformation $\mathbf{Q}$ s.t.

$$\hat{\mathbf{X}} = \mathbf{Q}^{-1}\mathbf{X}\mathbf{Q} = \left(\begin{array}{c|c} \tilde{\mathbf{X}} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \mathtt{x} \ \cdots \ \mathtt{x} & \lambda \end{array}\right), \ \sigma\big(\hat{\mathbf{X}}\big) = \sigma\left(\mathbf{X}\right) = \sigma\big(\tilde{\mathbf{X}}\big) + \{\lambda\}.$$

The most important observation here is that any annihilating polynomial of $\mathbf{X}$ is an annihilating polynomial for $\tilde{\mathbf{X}}$ since we applied a similarity transformation. This implies

$$\text{degree of } \operatorname{minpol}\big(x; \tilde{\mathbf{X}}\big) \leq \text{ degree of } \operatorname{minpol}\big(x; \mathbf{X}\big) \tag{1.112}$$

In the case of $\mathbf{X}$ being $\mathbf{I}_N$, $N > 1$ the bound is trivially tight. However, it is easy to see that equality does not hold, for instance, if λ is an eigenvalue of $\mathbf{X}$ with multiplicity 1. Note that the usual application of a deflation exploits the smaller size of $\tilde{\mathbf{X}}$ compared to $\mathbf{X}$ and its partial independence of the lower diagonal block because of the upper right zero off-diagonal block. Thus, seeking for the spectrum of a matrix $\mathbf{X}$, one would usually deflate a given eigenpair by a similarity transformation to reach a reducible matrix. In contrast, in our method a similarity transformation on $\mathbf{M} = \big(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\big)$ is potentially harmful, for it changes the spectrum of its principal submatrices $\mathbf{A}$ in an unpredictable way.

Fortunately, there are cases where we can exploit known eigenpairs via deflation. Let $\big(\begin{smallmatrix} \mathbf{v} \\ \mathbf{w} \end{smallmatrix}\big)$ be a vector partitioned conformally with $\mathbf{M} = \big(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\big)$. We distinguish three not disjoint cases

(i) $\mathbf{v} = \mathbf{0}$ and $\big(\begin{smallmatrix} \mathbf{0} \\ \mathbf{w} \end{smallmatrix}\big)$ is an eigenvector for $\mathbf{M}$ to the eigenvalue λ

(ii) $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ to the eigenvalue λ

(iii) $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ and $\big(\begin{smallmatrix} \mathbf{v} \\ \mathbf{w} \end{smallmatrix}\big)$ is an eigenvector of $\mathbf{M}$.

In the first case there is a deflating similarity transformation which is the identity on $\mathbf{A}$. Here, s, the size of the associated lambda-matrix $\mathcal{L}$, can be reduced.

$$\widehat{\mathbf{M}}_{(i)} = \begin{pmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} = \left(\begin{array}{c|c|c} \mathbf{A} & \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \\ \vdots & \ddots & \vdots \\ \mathtt{x} & \cdots & \mathtt{x} \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \\ \vdots & \ddots & \vdots \\ \mathtt{x} & \cdots & \mathtt{x} \end{matrix} & \tilde{\mathbf{D}} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \end{matrix} & \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \end{matrix} & \lambda \end{array} \right) \tag{1.113}$$

Let $\widetilde{\mathbf{M}}$ be obtained by removing the last row and the last column of $\widehat{\mathbf{M}}_{(i)}$. We can apply our method to $\widetilde{\mathbf{M}}$ to find the spectrum of $\mathbf{M}$ except the one eigenvalue λ which is already known. Since we removed a column of $\mathbf{BQ}$ and of $\hat{\mathbf{D}} = \mathbf{Q}^{-1}\mathbf{DQ}$ and a row of $\mathbf{Q}^{-1}\mathbf{C}$ and of $\hat{\mathbf{D}}$ but still can utilize the same annihilating polynomial of $\mathbf{A}$, we decreased κ by $(m+1)$

In the second case, (ii), we can employ a block diagonal similarity transformation s.t. the block corresponding to $\mathbf{A}$ is a deflating similarity transformation

on $\mathbf{A}$.

$$\widehat{\mathbf{M}}_{(ii)} = \begin{pmatrix} \mathbf{Q}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{pmatrix} = \left(\begin{array}{c|c} \tilde{\mathbf{A}} & \begin{matrix} 0 & \mathtt{x} & \cdots & \mathtt{x} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \mathtt{x} & \cdots & \mathtt{x} \end{matrix} \\ \hline \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \\ \mathtt{x} & \cdots & \mathtt{x} \\ \vdots & \ddots & \vdots \\ \mathtt{x} & \cdots & \mathtt{x} \end{matrix} \begin{matrix} \lambda \\ \mathtt{x} \\ \vdots \\ \mathtt{x} \end{matrix} & \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \\ \\ \hat{\mathbf{D}} \\ \\ \end{matrix} \end{array} \right) \quad (1.114)$$

Here, the matrix $\mathbf{T}$ is an arbitrary invertible transformation. If we repartition $\widehat{\mathbf{M}}_{(ii)}$ s.t. $\tilde{\mathbf{A}}$ becomes the upper left diagonal block, then we have incremented s. However, under certain conditions a linear factor corresponding to the root λ can be canceled, which means that we can decrement the degree m of the annihilating polynomial. A sufficient condition for this to be allowed is given as an easy corollary.

Corollary 2. *Let $\mathbf{A}$ be a complex square matrix of size N. Let $\mathrm{a}\,(x; \mathbf{A})$ be an annihilating polynomial for $\mathbf{A}$. Let λ be an eigenvalue of $\mathbf{A}$. Let $c_\lambda\,(\rho\,(\mathrm{a}))$ denote the multiplicity of λ as a root of $\mathrm{a}\,(x)$ and let $m_\lambda\,(\mathbf{A})$ denote the multiplicity of λ as an eigenvalue of $\mathbf{A}$. Finally, let $\mathbf{Q}$ be a similarity transformation s.t.*

$$\hat{\mathbf{A}} = \mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} = \left(\begin{array}{c|c} \tilde{\mathbf{A}} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \end{matrix} & \lambda \end{array} \right)$$

where $\tilde{\mathbf{A}}$ is square of size $(N-1)$. If

$$c_\lambda\,(\rho\,(\mathrm{a})) \geq m_\lambda\,(\mathbf{A})$$

then there is a divisor $\tilde{\mathrm{a}}\,(x)$ of the polynomial $\mathrm{a}\,(x)$ such that

$$\mathrm{a}\,(x) = \tilde{\mathrm{a}}\,(x)\,(x - \lambda)$$

and $\tilde{\mathrm{a}}\,(x)$ is an annihilating polynomial for $\tilde{\mathbf{A}}$.

Incrementing s by repartition of $\widehat{\mathbf{M}}_{(ii)}$ and incrementing m by canceling a factor in a decreases κ if $s \geq m$. In the special case when $\begin{pmatrix} \mathtt{v} \\ \mathbf{0} \end{pmatrix}$ happens to be an eigenvector of $\mathbf{M}$, the transformation in 1.114 yields

$$\widehat{\mathbf{M}}_{(ii)*} = \begin{pmatrix} \mathbf{Q}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_s \end{pmatrix} = \left(\begin{array}{c|c} \tilde{\mathbf{A}} & \begin{matrix} 0 & \mathtt{x} & \cdots & \mathtt{x} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \mathtt{x} & \cdots & \mathtt{x} \end{matrix} \\ \hline \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \\ \mathtt{x} & \cdots & \mathtt{x} \\ \vdots & \ddots & \vdots \\ \mathtt{x} & \cdots & \mathtt{x} \end{matrix} \begin{matrix} \lambda \\ 0 \\ \vdots \\ 0 \end{matrix} & \begin{matrix} \mathtt{x} & \cdots & \mathtt{x} \\ \\ \mathbf{D} \\ \\ \end{matrix} \end{array} \right) \quad (1.115)$$

Here, one may just remove the row and column intersecting at the position of the deflated eigenvalue λ without increasing s.

We now consider (iii). An arbitrary similarity transformation on $\mathbf{M}$ in order to deflate the eigenvalue possibly destroys the spectrum of $\mathbf{A}$. An exception to this was given a few lines above. As a precaution we can, of course, treat (iii) as a special case of (ii). As an alternative we can think of a Wielandt deflation [67], which allows for manipulating the eigenvalues to the given eigenvectors. Note that this method may have unpredictable effects on the minimal polynomial if not handled with care. We therefore treat only a special case where $\mathbf{M}$ is hermitian.

Let $\mathbf{M}$ be hermitian. Let λ be a simple eigenvalue, i.e. with multiplicity 1, of $\mathbf{A}$ to the eigenvector $\mathbf{v}$. and let μ be the eigenvalue of $\mathbf{M}$ to the eigenvector $\begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix}$. Let

$$\mathbf{A}_W = \mathbf{A} + \gamma_0 \frac{1}{|\mathbf{v}|} \mathbf{v} \mathbf{v}' \text{ with } \gamma_0(\gamma) = \gamma \frac{|\mathbf{v}|}{|\mathbf{v}| + |\mathbf{w}|} \tag{1.116}$$

be the right upper block of the matrix

$$\mathbf{M}_W = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} + \gamma \frac{1}{|\mathbf{v}| + |\mathbf{w}|} \begin{pmatrix} \mathbf{v}\mathbf{v}' & \mathbf{v}\mathbf{w}' \\ \mathbf{w}\mathbf{v}' & \mathbf{w}\mathbf{w}' \end{pmatrix} \tag{1.117}$$

Then

$$\sigma(\mathbf{A}_W) + \{\lambda\} = \sigma(\mathbf{A}) + \{(\lambda + \gamma_0)\} \tag{1.118}$$
$$\sigma(\mathbf{M}_W) + \{\mu\} = \sigma(\mathbf{M}) + \{(\mu + \gamma)\}. \tag{1.119}$$

This is shown as follows. Let $\mathbf{V}$ be a unitary eigenvector matrix for $\mathbf{A}$ with first column being $\mathbf{v}$. By unitarity $\mathbf{v}$ is orthogonal to all columns of $\mathbf{V}$ except the first one. Therefore, all other columns of $\mathbf{V}$ are still eigenvectors of $\mathbf{A}_W$ to the same eigenvalue. One verifies using (1.116) that $\mathbf{v}$ is an eigenvector of $\mathbf{A}_W$ to the eigenvalue $\lambda + \gamma_0$. A similar reasoning applies for $\mathbf{M}$. Since $\mathbf{A}_W$ has an unitary eigenvector matrix, it is hermitian.

Thus, via the parameter γ we may shift λ and μ simultaneously without effecting the rest of the spectrum of either $\mathbf{A}$ or $\mathbf{M}$. Assume that $N \geq 2$ and let λ' be any eigenvalue of $\mathbf{A}_W$ s.t. $\lambda' \neq \lambda$, which exists since λ is simple. If we choose γ s.t. $\lambda + \gamma_0(\gamma) = \lambda'$ then the number of distinct eigenvalues of $\mathbf{A}_W$ is less than that of $\mathbf{A}$. Since for hermitian matrices the number of distinct eigenvalues is the degree of the minimal polynomial, m is decremented.

Since did not change matrix sizes s is preserved. Thus, by Wielandt deflation we can decrease κ by m.

Chapter 2

Equitable Partitions

This chapter deals with equitable partitions. We solve the practical problem of removing the eigenvalues of the front divisor from the spectrum of a complex matrix. Special cases like adjacency matrices of graphs can be derived without difficulties. The eigenproblem is divided into two smaller eigenproblems by a unitary block diagonal Householder transformation. The author is not aware of this method elsewhere. The costs are quadratic in matrix size with small constants. The back transformation of eigenvectors is easy. A procedure to properly utilize several equitable partitions which are ordered by refinement is discussed. Generalizations of equitable partitions and the necessary extensions of the method are considered.

2.1 Introduction

Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ be a complex square matrix. Let $\Pi = (c_1, \ldots, c_k)$ be a simultaneous (disjoint and exhaustive) partition of its rows and columns, respectively, into k cells. When $\mathbf{A}$ is seen as the weighted adjacency matrix of a graph, Π is a partition of its node set. Let n_i be the size of the i-th cell. Let $\mathbf{A}_{ij} \in \mathbf{C}^{n_i \times n_j}$ be the matrix block in $\mathbf{A}$ induced by the i-th cell of rows and the j-th cell of columns. We call $\mathbf{A}$ *row equitable* (or often just *equitable*) w.r.t. Π if

$$\forall\, 1 \leq i, j \leq k \quad \exists\, \theta_{ij} \in \mathbb{C} \quad \text{s.t.} \quad \mathbf{A}_{ij}\, \mathbf{j}_{n_i} = \theta_{ij}\, \mathbf{j}_{n_j} \tag{2.1}$$

where $\mathbf{j}_n = \{1\}^n$. In that case we will also speak of Π as an *(row)equitable partition* of $\mathbf{A}$. The discrete partition, which consists of N distinct, non empty cells of size 1, is always equitable. We call $\mathbf{A}$ *column equitable* (w.r.t. Π) if $\mathbf{A}'$ is row equitable (w.r.t. Π). It is easy to see that a row equitable hermitian matrix is always column equitable. The matrix $\Theta = (\theta_{ij})$ is known as the *quotient graph* or the *front divisor* of $\mathbf{A}$ w.r.t. Π.

Equitable partitions are well known in graph theory, and the concept is studied under several names e.g. front divisor [21], *exact coloration* [10], *graph fibration* [9], *quasi-block-stochastic matrices* [51], *1-dimensional Weisfeiler-Lehman stabilizer* [16]. See also [71], [84] and others. In applications $\mathbf{A}$ is often restricted to be the (binary) adjacency matrix of a graph. In this case the constant θ_{ij} has the interpretation of denoting the number of neighbours in cell j for any node in cell i. It is convenient to define an *indicator matrix*

$\mathbf{B} = (b_{vi}) \in \{0,1\}^{N \times k}$ of a partition by

$$b_{vi} = \begin{cases} 1 & \text{, if node } v \text{ is in cell } i \\ 0 & \text{, else} \end{cases} . \tag{2.2}$$

For instance the indiscrete (or singleton) partition, which has only one cell, yields $\mathbf{B} = \mathbf{j}_N$. The indicator matrix of the discrete partition (or partition of singletons) is a permutation matrix. Using the indicator matrix we can express (2.1) by

$$\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{\Theta}. \tag{2.3}$$

The most famous property of equitable partitions is that the characteristic polynomial of a matrix is divided by the characteristic polynomial of its front divisor i.e.

$$\sigma(\mathbf{\Theta}) \subset \sigma(\mathbf{A}).$$

Moreover, a simple premultiplication by $\mathbf{B}$ immediately lifts any ('column' or 'right') eigenvector of $\mathbf{\Theta}$ to an eigenvector of $\mathbf{A}$. Thus, given an equitable partition a subset of the eigenpairs of $\mathbf{A}$ can be obtained by solving the (smaller) eigenproblem of $\mathbf{\Theta}$. How one can use equitable partitions to actually reduce the eigenproblem for the adjacency matrix of a graph completely to the eigenproblem for two smaller matrices is shown for instance in [32]. The generalization to complex matrices is straightforward. Using the normalized indicator matrix $\tilde{\mathbf{B}} = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-\frac{1}{2}}$ there is always a matrix $\mathbf{C} \in \mathbb{C}^{N \times (N-k)}$ s.t. $\mathbf{Q} = (\tilde{\mathbf{B}}|\mathbf{C})$ is unitary. One shows that for $\mathbf{B}$ indicating an equitable partition of $\mathbf{A}$ the matrix

$$\tilde{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F} \end{pmatrix}$$

gets reducible with $\mathbf{E} = \tilde{\mathbf{B}}'\mathbf{\Theta}\tilde{\mathbf{B}}$ being similar to the front divisor and independent of $\mathbf{C}$. Since we applied a similarity transformation, the spectrum of $\mathbf{A}$ is preserved. We will refer to $\mathbf{F}$, which depends on the particular choice of $\mathbf{C}$, as a *factor*. One shows that all possible factors are similar and that

$$\sigma(\mathbf{A}) = \sigma(\mathbf{E}) + \sigma(\mathbf{F}).$$

Although it is sufficient to ensure that $\mathbf{Q}$ is invertible, the stronger requirement of unitarity has numerical advantages and preserves hermiticity if present.
Due to the involved matrix-matrix-multiplication this approach yields a worst case complexity which is cubic in N when today's methods are applied. At first glance this is acceptable for finding an equitable partition has a worst case complexity of $O(N^3)$. However, in the binary case, $\mathbf{A} \in \{0,1\}^{N \times N}$, the available software, for instance *nauty* [60], is pretty fast on average. There might also be some heuristics depending on the empirical background of a given matrix. Furthermore, the complexity to prove (or disprove) that a given partition is equitable requires only $O(m)$ floating point operations where m

is the number of non zero matrix entries. Therefore, it would be desirable to exploit an equitable partition to reduce a matrix with computational costs that are less then cubic in N.

There are at least two ways in speeding up the matrix-multiplication. One may apply intriguing methods for fast matrix-multiplication e.g. Strassens algorithm, or one may consider multiplications with special matrices e.g. Fourier matrices, $O\left(N^2 \log N\right)$, diagonal matrices, $O\left(N^2\right)$, or matrices of low rank r, $O\left(rN^2\right)$. Using block diagonal matrices with Householder matrices as diagonal blocks we follow the second approach reaching a complexity of $O\left(N^2\right)$. In fact, considering our slightly generalized notion of equitable partition applicable to complex valued rectangular matrices of size $N_1 \times N_2$ we reach a complexity of $O\left(N_1 N_2\right)$.

We sketch our main idea in the example case of the usual equitable partition (2.1). Given an equitable partition we find Householder matrices $\mathbf{H}_i$ s.t.

$$\mathbf{H}_i \mathbf{j}_{n_i} \propto \mathbf{e}_{n_i} = (1, 0, \ldots, 0)' \quad \text{and} \quad \mathbf{H}_i \mathbf{e}_{n_i} \propto \mathbf{j}_{n_i}.$$

Applying the $\mathbf{H}_i$ to the matrix blocks $\mathbf{A}_{ij}$ yields

$$\left[\mathbf{H}_i \mathbf{A}_{ij} \mathbf{H}_j'\right] \mathbf{e}_{n_j} \propto \mathbf{e}_{n_i}.$$

Thus, after transformation each block $\left[\mathbf{H}_i \mathbf{A}_{ij} \mathbf{H}_j'\right]$ has block triangular form with an upper left subblock of size 1. This implies that $\mathbf{A}$ is transformed into an implicitly block triangular matrix. In the hermitian case we obtain an implicitly block diagonal matrix with hermitian diagonal blocks. Implicitly here means up to a known permutation.

The chapter is organized as follows. First, we recall some properties of unitary elementary matrices, which are generalizations of the well known Householder transformation. Then, as an outline for the main ideas behind our method we consider three well established techniques for numerical computations all of them using Householder matrices: the deflation of an eigenpair, the deflation of a singular value and the **QR** decomposition of a rectangular matrix. Then we show how to efficiently block triangularize a matrix exploiting (2.1). We then give an example and some further remarks. After that we discuss a generalization of equitable partitions which is three-fold

- we replace the $\mathbf{j}_{n_i}$ in (2.1) by a vector $\mathbf{v}(i) \in \mathbb{C}^{n_i}$ associated to the i-th cell

- we may allow a violation of (2.1) for some pairs (i, j).

- we consider different partitions for the column and row set of $\mathbf{A}$ ultimately enabling us to treat rectangular matrices.

We end the chapter with partitions which have a perturbed equitability and some remarks regarding the search for equitable partitions.

2.2 Extended Householder Transformations

Applying a diagonal matrix or a rank-1 matrix to a vector both yields a complexity of $O(N)$. Hence the same complexity pertains to matrices which are finite linear combinations of diagonal and rank-1 matrices and particularly for the class of *elementary matrices* [44], which are rank-1 updates of the identity. Another class of matrices which are important for applications are *unitary matrices* because their inverse is available without floating point operations and all their eigenvalues have norm 1, which ensures numerical stability. The intersection of both classes, the *unitary elementary matrices*, combine those properties making them a powerful tool in numerical computation. Those matrices are sometimes called *Householder matrices* although this term may also denote a subclass of it.

This section aims at the transformation given by definitions (7) and (6) as one example of established complex generalizations of real Householder matrices [43], [54] and introduces *extended Householder matrices* as a block diagonal version of it. These two transformations will be the main tool of our method and used throughout the rest of the chapter.

2.2.1 Unitary Elementary Matrices

An complex elementary matrix has the form

Definition 4. $\mathbf{E}(\mathbf{u}, \mathbf{v}, \sigma) = \mathbf{I}_N - \sigma \mathbf{u} \mathbf{v}' \in \mathbb{C}^{N \times N}$.

Unitary elementary matrices are necessarily [73] of the form

Definition 5. $\mathbf{U}(\mathbf{u}, \gamma) = \mathbf{E}\left(\mathbf{u}, \mathbf{u}, 2 \frac{(\mathbf{u}'\mathbf{u})^+}{1 + \mathrm{i}\gamma}\right)$, $\gamma \in \mathbb{R}$,

where $|\mathbf{u}| = \sqrt{\mathbf{u}'\mathbf{u}}$ and $c^+ = \begin{cases} c^{-1} & , \ c \neq 0 \\ 0 & , \ c = 0 \end{cases}$ for $c \in \mathbb{C}$.

Using the definition, it is easy to verify that

$$\mathbf{U}'(\mathbf{u}, \gamma) = \mathbf{U}(\mathbf{u}, -\gamma) \ \text{ and } \ \mathbf{U}(c\mathbf{u}, \gamma) = \mathbf{U}(\mathbf{u}, \gamma) \ \text{ for } 0 \neq c \in \mathbb{C}. \tag{2.4}$$

We seek an unitary elementary matrix mapping $\mathbf{x} \neq \mathbf{0}$ in the direction of $\mathbf{y} \neq \mathbf{0}$ i.e.

$$\mathbf{U}(\mathbf{u}, \gamma)\mathbf{x} + \alpha \mathbf{y} = \mathbf{0} \ \text{ with } 0 \neq \alpha \in \mathbb{C}, \tag{2.5}$$

which implies that $\mathbf{u}$ is a linear combination of $\mathbf{x}$ and $\mathbf{y}$. Namely,

$$\mathbf{x} + \alpha \mathbf{y} = 2 \frac{(\mathbf{u}'\mathbf{u})^+}{1 + \mathrm{i}\gamma} (\mathbf{u}'\mathbf{x}) \mathbf{u}. \tag{2.6}$$

Since according to (2.4) scaling of $\mathbf{u}$ does not change $\mathbf{U}(\mathbf{u}, \gamma)$, we are free to choose $\mathbf{u} = \mathbf{x} + \alpha \mathbf{y}$, which yields

$$1 + \mathrm{i}\gamma = 2(\mathbf{u}'\mathbf{u})^+ \mathbf{u}'\mathbf{x}. \tag{2.7}$$

Using the property of being unitary in (2.5) we find

$$|\mathbf{x}| = |\alpha| \, |\mathbf{y}|, \tag{2.8}$$

thus, $|\mathbf{x}|$ and $|\mathbf{y}|$ determine the norm of α and we have

$$\gamma = \mathrm{Im}\left(\overline{\alpha}\mathbf{y}'\mathbf{x}\right)\left(\mathbf{x}'\mathbf{x} + \mathrm{Re}\left(\overline{\alpha}\mathbf{y}'\mathbf{x}\right)\right)^+. \tag{2.9}$$

Using the abbreviation $\beta = \frac{\alpha}{|\alpha|}$ we obtain

$$\mathbf{U}\left(\mathbf{x} + \beta\frac{|\mathbf{x}|}{|\mathbf{y}|}\mathbf{y}, -\mathrm{Im}\left(\beta\mathbf{x}'\mathbf{y}\right)\left(|\mathbf{x}|\,|\mathbf{y}| + \mathrm{Re}\left(\beta\mathbf{x}'\mathbf{y}\right)\right)^+\right)\mathbf{x} + \beta\frac{|\mathbf{x}|}{|\mathbf{y}|}\mathbf{y} = 0. \tag{2.10}$$

Fortunately, we only have to consider the special case $\mathbf{y} = \mathbf{e}_N = (1, 0, \ldots, 0)'$. Abbreviating

$$\mathbf{H}\left(\mathbf{x}, \beta\right) = \mathbf{U}\left(\mathbf{x} + \beta\,|\mathbf{x}|\,\mathbf{e}_N, -\mathrm{Im}\left(\beta\,\overline{x^1}\right)\left(|\mathbf{x}| + \mathrm{Re}\left(\beta\,\overline{x^1}\right)\right)^+\right), \tag{2.11}$$

we observe the properties

$$\mathbf{H}\left(\mathbf{x}, \beta\right)\mathbf{x} = -\beta\,|\mathbf{x}|\,\mathbf{e}_N \quad, \quad \mathbf{H}'\left(\mathbf{x}\right)\mathbf{e}_N = -\frac{\overline{\beta}}{|\mathbf{x}|}\mathbf{x}. \tag{2.12}$$

Although its norm is determined to be 1, the particular choice of β is arbitrary so far. We may choose

Definition 6. $\beta\left(\mathbf{x}\right) = \begin{cases} \exp\left(2\pi\mathrm{i}\arg\left(x^1\right)\right) &, x^1 \neq 0 \\ 1 &, x^1 = 0 \end{cases}, \quad \mathbf{x} \in \mathbb{C}^N,$

which ensures selfadjointness of $\mathbf{H}\left(\mathbf{x}, \beta\right)$ because then $\mathrm{Im}\left(\beta\,\overline{x^1}\right) = 0$ and thus $\gamma = 0$ in (2.11). It also supports numerical stability since it maximizes the norm of $\mathbf{x} + \beta\,|\mathbf{x}|\,\mathbf{e}_N$ in (2.11). We will henceforth use the matrix

Definition 7.

$$\mathbf{H}\left(\mathbf{x}\right) = \mathbf{H}\left(\mathbf{x}, \beta\left(\mathbf{x}\right)\right) = \mathbf{I} - 2\frac{\left(\mathbf{x} + \beta\,|\mathbf{x}|\,\mathbf{e}_N\right)\left(\mathbf{x} + \beta\,|\mathbf{x}|\,\mathbf{e}_N\right)'}{\left(\mathbf{x} + \beta\,|\mathbf{x}|\,\mathbf{e}_N\right)'\left(\mathbf{x} + \beta\,|\mathbf{x}|\,\mathbf{e}_N\right)}.$$

This is equivalent to

$$\mathbf{H}\left(\mathbf{x}\right) = \mathbf{I} - \frac{|\mathbf{x}|}{|\mathbf{x}| + |x^1|}\left(\frac{\mathbf{x}}{|\mathbf{x}|} + \beta\mathbf{e}_N\right)\left(\frac{\mathbf{x}}{|\mathbf{x}|} + \beta\mathbf{e}_N\right)', \quad \beta = \beta\left(\mathbf{x}\right). \tag{2.13}$$

Using the definition one easily shows that $\mathbf{H}\left(\mathbf{x}\right)$ is independent of a scaling of its argument $\mathbf{x}$,

$$\mathbf{H}\left(\mu\mathbf{x}\right) = \mathbf{H}\left(\mathbf{x}\right), \quad \mu \neq 0. \tag{2.14}$$

It is a crucial observation that $\mathbf{H}(\mathbf{x})$ does in general not 'commute' with a permutation matrix $\mathbf{P}$ acting on the entries of $\mathbf{x}$

$$\mathbf{H}(\mathbf{P}\mathbf{x}) \neq \mathbf{P}\mathbf{H}(\mathbf{x})\mathbf{P}' \text{ in general.} \tag{2.15}$$

However, equality does hold in (2.15) when the first entry is fixed. In particular, $\mathbf{H}(\mathbf{j}_n)$ is even independent of any permutation of its argument.

The matrix $\mathbf{H}(\mathbf{x})$ provides a numerical stable transformation since it is unitary, and it is easy to apply and most easily inverted since it is elementary and selfadjoint.

2.2.2 Extended Householder Matrices

We introduce *elementary block diagonal matrices* which are of the form

$$\mathbf{E}(\mathbf{u}, \mathbf{v}, \Pi, \sigma) = \operatorname{diag}\Big(\mathbf{E}\big(\mathbf{u}(1), \mathbf{v}(1), \sigma_1\big), \ldots, \mathbf{E}\big(\mathbf{u}(k), \mathbf{v}(k), \sigma_k\big)\Big) \tag{2.16}$$

where each diagonal block is an elementary matrix. The vectors $\mathbf{u}$ and $\mathbf{v}$ are of the same length N. The subvectors $\mathbf{u}(i)$ $(\mathbf{v}(i))$ are induced by the i-th cell of the labeled partition $\Pi = (c_1, \ldots, c_k)$ of the entries of $\mathbf{u}$ $(\mathbf{v})$. We will only use the special case

Definition 8. $\mathbf{H}(\mathbf{x}, \Pi) = \operatorname{diag}\Big(\mathbf{H}\big(\mathbf{x}(1)\big), \ldots, \mathbf{H}\big(\mathbf{x}(k)\big)\Big)$

and refer to it with *extended Householder matrix*. The labeling of Π already determines the entries of $\mathbf{x}(i)$. To avoid ambiguity and since the matrix $\mathbf{H}\big(\mathbf{x}(i)\big)$ crucially depends on the first entry of $\mathbf{x}(i)$ we explicitly specify that the sequence of the entries of $\mathbf{x}(i)$ already is such that the permutation Φ_Π of definition 9 below is the identity, which implies that the entries of $\mathbf{x}$ are ordered s.t. Π entails an explicit block structure on $\mathbf{x}$. Thus, $\mathbf{x}(i)$ is simply the i-th vector block of $\mathbf{\Phi}_\Pi(\mathbf{x})$.

Extended Householder matrices should not be confused with *block Householder transformations*, which are of the form $\mathbf{I}_N - \mathbf{X}\mathbf{Y}'$ with $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{N \times k}$ or variants of that and are a compact way to describe products of Householder matrices [8], [23]. They are used to replace matrix-vector by matrix-matrix multiplications to save communications on memory distributed parallel computers. Actually, extended Householder matrices can be described as a special case of that. When $\mathbf{x}$ is a vector canonically partitioned by Π with indicator matrix $\mathbf{B}$, form $\tilde{\mathbf{x}}$ by

$$\tilde{\mathbf{x}}(i) = \mathbf{x}(i) + \beta(\mathbf{x}(i)) |\mathbf{x}(i)| \, \mathbf{e}_{n_i}.$$

Set $\mathbf{Z} = \operatorname{diag}(\tilde{\mathbf{x}})\mathbf{B}$ and $\mathbf{X} = \sqrt{2}\,\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-\frac{1}{2}}$. Then

$$\mathbf{H}(\mathbf{x}, \Pi) = \mathbf{I} - \mathbf{X}\mathbf{X}'.$$

However, describing $\mathbf{H}(\mathbf{x}, \Pi)$ that way, i.e. ignoring its special structure, is not an advantage.

Since multiplying an elementary matrix of size N to a vector of length N has complexity $O(N)$, the complexity for multiplying an extended Householder matrix to a vector is also linear in its size, hence the complexity of multiplying a rectangular matrix of size $N_1 \times N_2$ from the left and from the right each by an extended Householder matrix of appropriate size is of order $O(N_1 N_2)$.

2.3 Three Well Known Techniques Using Householder Matrices

2.3.1 Deflating an Eigenpair

A scalar λ and a vector $\mathbf{v} \in \mathbb{C}^N$ with $|\mathbf{v}| \neq 0$ form an right-eigenpair of the square matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ if and only if

$$\mathbf{A}\mathbf{v} = \mathbf{v}\lambda. \tag{2.17}$$

Given an eigenvector the corresponding eigenvalue is uniquely determined. Eigenvectors are not unique, even if the eigenspace associated to λ is 1-dimensional and even if they are normalized due to an arbitrary phase factor which is still present as an arbitrary sign flip in the real case. The transformation

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{v}) \mathbf{A} \mathbf{H}'(\mathbf{v}) \tag{2.18}$$

yields a block triangularization of $\mathbf{A}$ because

$$\tilde{\mathbf{A}}\mathbf{e}_N = \frac{-\bar{\beta}(\mathbf{v})}{|\mathbf{v}|} \mathbf{H}(\mathbf{v}) \mathbf{A}\mathbf{v} = \frac{-\bar{\beta}(\mathbf{v})}{|\mathbf{v}|} \lambda \mathbf{H}(\mathbf{v}) \mathbf{v} = \lambda \mathbf{e}_N,$$

which implies

$$\tilde{\mathbf{A}} = \begin{pmatrix} \lambda & \mathbf{X} \\ \mathbf{0} & \end{pmatrix} \text{ with } \mathbf{X} \in \mathbb{C}^{N \times (N-1)}. \tag{2.19}$$

2.3.2 Deflating a Singular Value

A scalar λ is called a singular value of the matrix $\mathbf{A} \in \mathbb{C}^{N_1 \times N_2}$ if and only if there exist vectors $\mathbf{u}_1 \in \mathbb{C}^{N_1}$, $\mathbf{u}_2 \in \mathbb{C}^{N_2}$ s.t.

$$|\mathbf{u}_1| \mathbf{A}\mathbf{u}_2 = \mathbf{u}_1 \lambda |\mathbf{u}_2| \ , \ |\mathbf{u}_2| \mathbf{A}'\mathbf{u}_1 = \mathbf{u}_2 \lambda |\mathbf{u}_1| \tag{2.20}$$

with $0 \neq |\mathbf{u}_1|, |\mathbf{u}_2|$ and $0 \leq \lambda \in \mathbf{R}$. The normalization of $\mathbf{u}_1$ and $\mathbf{u}_2$ excludes the null vector. The norm of λ is already determined by (2.20). Choosing the phase of λ to be zero is a convenient restriction and ultimately yields the uniqueness of the multi set of singular values for a given matrix. If $\mathbf{u}_1$ and $\mathbf{u}_2$ are normed to unit length, they are called a left and right singular vector, respectively. Singular vectors are not unique, since there is at least an arbitrary phase vector. The transformation

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{u}_1) \mathbf{A} \mathbf{H}'(\mathbf{u}_2) \tag{2.21}$$

yields a block diagonalization of $\mathbf{A}$ because $\tilde{\mathbf{A}}\mathbf{e}_{N_2} = \mathbf{e}_{N_1}\lambda$ and $\tilde{\mathbf{A}}'\mathbf{e}_{N_1} = \mathbf{e}_{N_2}\lambda$ i.e.

$$\tilde{\mathbf{A}} = \left(\begin{array}{cc} \lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{X} \end{array} \right) \text{ with } \mathbf{X} \in \mathbb{C}^{(N_1-1)\times(N_2-1)}. \tag{2.22}$$

2.3.3 QR-Decomposition

Suppose we have a pair of vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$, at hand which may not be singular vectors i.e. may not obey (2.20), but fulfill the weaker condition

$$|\mathbf{u}_1|\,\mathbf{A}\mathbf{u}_2 = \mathbf{u}_1\lambda\,|\mathbf{u}_2| \text{ and } \mathbf{u}_1, \mathbf{u}_2 \neq \mathbf{0}. \tag{2.23}$$

for some $\lambda \in \mathbb{C}$. For later purposes it will be inconvenient to fix the phase of λ. Note that (2.17) arises out of (2.23) for $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{v}$. Using the same transformations as above we may not obtain block diagonal form, but since

$$\mathbf{H}(\mathbf{u}_1)\,\mathbf{A}\mathbf{H}'(\mathbf{u}_2)\,\mathbf{e}_{N_2} = \mathbf{e}_{N_1}\lambda \tag{2.24}$$

block triangular form of $\tilde{\mathbf{A}}$ is ensured. (2.24) being a consequence of (2.23) is essential to our approach. It is easy to find a pair of vectors obeying (2.23). One may for instance set $\mathbf{u}_1$ to be a multiple of $\mathbf{A}\mathbf{u}_2$ with arbitrarily chosen $\mathbf{u}_2$ that is not the null vector. Moreover, for the choice $\mathbf{u}_2 = \mathbf{e}_{N_2}$ the vector $\mathbf{u}_1$ is given by a multiple of the first column of $\mathbf{A}$. This is the basic idea of the common Householder approach to the **QR**-decomposition, which aims at a factorization of $\mathbf{A}$ into a product of an unitary and an upper triangular matrix.

2.4 Block Triangularizing by Equitable Partitions

2.4.1 Suitable Indexing

Let $\Pi_0 = (c_1, \ldots, c_k)$ be a simultaneous partition of the row and the column set of the complex $N \times N$ matrix $\mathbf{A}_0$ into $k < N$ cells. If Π_0 is an unlabeled equitable partition we may choose a labeling arbitrarily. Let $\mathbf{A}_{0,ij}$ be the matrix subblock of $\mathbf{A}_0$ induced by the i-th row and the j-th column cell. If we choose a permutation Φ of the node set of $\mathbf{A}_0$ s.t.

$$i < j \text{ and } v \in c_i \text{ and } w \in c_j \text{ implies } \Phi(v) < \Phi(w),$$

then the implicit block structure of $\mathbf{A}_0$ becomes explicit. We call $\mathbf{\Phi}'\mathbf{A}_0\mathbf{\Phi}$ where $\mathbf{\Phi}$ is the permutation matrix representing Φ *suitably indexed*. To avoid ambiguity we may define a standard permutation Φ_{Π_0}.

Definition 9. *Let $\Pi_0 = (c_1, \ldots, c_k)$ be a labeled partition of $[1, \ldots, N]$. The permutation $\Phi_{\Pi_0} : [1, \ldots, N] \to [1, \ldots, N]$ is defined by*
$\forall\, i, j \in [1, \ldots, k]\ \forall\, v, w \in [1, \ldots, N]$

(i) $i < j$ and $v \in c_i$ and $w \in c_j$ implies $\Phi_{\Pi_0}(v) < \Phi_{\Pi_0}(w)$

(ii) *$v < w$ and $v, w \in c_i$ implies $\Phi_{\Pi_0}(v) < \Phi_{\Pi_0}(w)$.*

Let $\boldsymbol{\Phi}_{\Pi_0}$ be the matrix form of Φ_{Π_0}. Definition 9 is chosen s.t. $\boldsymbol{\Phi}_{\Pi_0} = \mathbf{I}_N$ if and only if Π_0 entails an explicit block structure with accordingly labeled blocks on $\mathbf{A}_0$. We will denote by Π the partition of

$$\mathbf{A} = \boldsymbol{\Phi}'_{\Pi_0} \mathbf{A}_0 \boldsymbol{\Phi}_{\Pi_0} \tag{2.25}$$

which corresponds to the partition Π_0 of $\mathbf{A}_0$. Since permutations are unitary transformations,

$$\sigma(\mathbf{A}) = \sigma(\mathbf{A}_0). \tag{2.26}$$

Therefore, we can restrict our discussion to suitably indexed matrices w.l.o.g.

2.4.2 Block Triangularization

Let $\mathbf{A}_{ij}$ be the matrix subblock of the suitably indexed matrix $\mathbf{A}$ induced by the i-th row and the j-th column cell of the (labeled) equitable partition Π with indicator matrix $\mathbf{B}$. Let $n_i = |c_i|$ and

$$\mathbf{N} = \mathrm{diag}(n_1, \dots, n_k) = n_i \delta_{ij} = \mathbf{B}' \mathbf{B}. \tag{2.27}$$

We have with $\mathbf{j}_n = \{1\}^n$

$$\forall\, 1 \leq i, j \leq k \quad \mathbf{A}_{ij} \mathbf{j}_{n_j} = \mathbf{j}_{n_i} \theta_{ij}. \tag{2.28}$$

Note that (2.28) is a variant of (2.23). When we apply the same partition Π conformably on the rows and columns of

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{j}_N, \Pi)\, \mathbf{A} \mathbf{H}(\mathbf{j}_N, \Pi), \tag{2.29}$$

then each of its blocks has the form

$$\tilde{\mathbf{A}}_{ij} = \mathbf{H}(\mathbf{j}_{n_i})\, \mathbf{A}_{ij} \mathbf{H}(\mathbf{j}_{n_j}). \tag{2.30}$$

Using relations (2.12) and condition (2.28) we obtain

$$\tilde{\mathbf{A}}_{ij} \mathbf{e}_{n_j} = \mathbf{e}_{n_i} \theta_{ij} \frac{\sqrt{n_i}}{\sqrt{n_j}}, \tag{2.31}$$

which shows that each block is block triangular s.t. the upper left diagonal subblock is always a scalar. Again note the similarity to (2.24). We interpret $\mathbf{E} = \mathbf{N}^{\frac{1}{2}} \boldsymbol{\Theta} \mathbf{N}^{-\frac{1}{2}} \in \mathbf{C}^{k \times k}$, i.e. $e_{ij} = \theta_{ij} \frac{\sqrt{n_i}}{\sqrt{n_j}}$, as a normalized version of the front divisor $\boldsymbol{\Theta} = (\theta_{ij})$
The special structure of each block of $\tilde{\mathbf{A}}$ allows for an explicit block triangularization. There are readily available permutations $\boldsymbol{\Omega}$ s.t.

$$\boldsymbol{\Omega}' \tilde{\mathbf{A}} \boldsymbol{\Omega} = \begin{pmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F} \end{pmatrix} \tag{2.32}$$

where we call $\mathbf{F}$ a factor. Those permutations are any bijective extension to $[1, \dots, N] \to [1, \dots, N]$ of the function Ω_0 that we define via its inverse

Definition 10.

$$\Omega_0^{-1} : [1, \ldots, k] \to \operatorname{dom}(\Omega_0) \subset [1, \ldots, N] ,$$

$$\Omega_0^{-1}(i) = 1 + \sum_{j=1}^{i-1} n_j.$$

Ω_0 just maps the first nodes of the cells canonically into the first k positions of all nodes. To avoid ambiguity we explicitly define a standard extension of Ω_0.

Definition 11. *Let $\Pi = (c_1, \ldots, c_k)$ be a labeled partition of the ordered set $[1, \ldots, N]$ s.t. $i < j$ and $v \in c_i$ and $w \in c_j$ implies $v < w$. Let $n_i = |c_i|$. The permutation $\Omega_\Pi : [1, \ldots, N] \to [1, \ldots, N]$ is defined by*

$$\Omega_\Pi\Big(m_i + \sum_{j=1}^{i-1} n_j\Big) = \begin{cases} i & , m_i = 1 \\ k - i + m_i + \sum_{j=1}^{i-1} n_j & , m_i = 2, \ldots, n_i \end{cases} , \; i = 1, \ldots, k.$$

We summarize this section in the following theorem

Theorem 6. *Let $\Pi = (c_1, \ldots, c_k)$ be a labeled equitable partition of the square matrix $\mathbf{A}$ of size N. Let $\mathbf{A}$ be suitably indexed according to Π. Let $\mathbf{B}$ be the indicator matrix and let $\boldsymbol{\Theta}$ be the corresponding front divisor. Then*

$$\mathbf{A}_\Pi = \boldsymbol{\Omega}_\Pi' \mathbf{H}(\mathbf{j}_N, \Pi) \, \mathbf{A} \mathbf{H}(\mathbf{j}_N, \Pi) \, \boldsymbol{\Omega}_\Pi = \begin{pmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F} \end{pmatrix}$$

is block triangular with

$$\mathbf{E} = (\mathbf{B}'\mathbf{B})^{\frac{1}{2}} \, \boldsymbol{\Theta} \, (\mathbf{B}'\mathbf{B})^{-\frac{1}{2}}$$

and

$$\sigma(\mathbf{A}) = \sigma(\mathbf{E}) + \sigma(\mathbf{F}).$$

Proof. The theorem follows from the construction of $\mathbf{H}(\mathbf{j}_N, \Pi)$ and $\boldsymbol{\Omega}_\Pi$ and the observation that $\mathbf{H}(\mathbf{j}_N, \Pi) \boldsymbol{\Omega}_\Pi$ is unitary. $\square$

Since

$$\mathbf{A}_\Pi' = \boldsymbol{\Omega}_\Pi' \mathbf{H}(\mathbf{j}_N, \Pi) \, \mathbf{A}' \mathbf{H}(\mathbf{j}_N, \Pi) \, \boldsymbol{\Omega}_\Pi, \tag{2.33}$$

partitions which are simultaneously row and column equitable lead to a block diagonalization. Furthermore, since $\mathbf{A}' = \mathbf{A}$ implies $\mathbf{A}_\Pi' = \mathbf{A}_\Pi$, equitable hermitian matrices are diagonalized with diagonal blocks being again hermitian.

2.4.3 Recovery of Eigenvectors

Let Π_0 be an equitable partition of $\mathbf{A}_0$. Let $\mathbf{v}_\Pi$ be an eigenvector of the reduced problem $\mathbf{A}_\Pi$ to the eigenvalue λ i.e.

$$\mathbf{A}_\Pi \mathbf{v}_\Pi = \mathbf{\Omega}_\Pi' \mathbf{H}(\mathbf{j}_N, \Pi)' \mathbf{A} \mathbf{H}(\mathbf{j}_N, \Pi) \mathbf{\Omega}_\Pi = \lambda \mathbf{v}_\Pi. \tag{2.34}$$

Then

$$\mathbf{v} = \mathbf{H}(\mathbf{j}_N, \Pi) \mathbf{\Omega}_\Pi \mathbf{v}_\Pi \tag{2.35}$$

and

$$\mathbf{v}_0 = \mathbf{\Phi}_{\Pi_0} \mathbf{v} \tag{2.36}$$

are eigenvectors of $\mathbf{A}$ and $\mathbf{A}_0$, respectively, to the same eigenvalue λ, which can be seen by premultiplication of $\mathbf{H}(\mathbf{j}_N, \Pi) \mathbf{\Omega}_\Pi$ to (2.34) and using (2.25). We put the main idea in the following theorem, which is easily proved the same way.

Theorem 7. *Let* $\mathbf{A}$, Π *and* $\mathbf{A}_\Pi$ *be the same as in theorem 6. Let the columns of* $\mathbf{V}_\Pi$ *be right-eigenvectors of* $\mathbf{A}_\Pi$. *Then the columns of*

$$\mathbf{V} = \mathbf{H}(\mathbf{j}_N, \Pi) \mathbf{\Omega}_\Pi \mathbf{V}_\Pi \tag{2.37}$$

are eigenvectors of $\mathbf{A}$ *to the same eigenvalues, respectively.*

2.4.4 Example

One verifies that

$$\mathbf{A}_0 = \begin{pmatrix} 1 & 2 & 3 & 3 & 3 & 2 \\ 2 & 4 & 3 & 1 & 2 & 1 \\ 3 & 3 & 1 & 4 & 1 & 1 \\ 3 & 1 & 4 & 0 & 2 & 3 \\ 3 & 2 & 1 & 2 & 3 & 2 \\ 2 & 1 & 1 & 3 & 2 & 4 \end{pmatrix}$$

is row equitable w.r.t. $\Pi_0 = (1|2, 6|3, 4, 5)$. Using the permutation

$$\mathbf{P}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

we can transform it into the suitably labeled form

$$\mathbf{A} = \mathbf{P}_0' \mathbf{A}_0 \mathbf{P}_0 = \begin{pmatrix} 1 & 2 & 2 & 3 & 3 & 3 \\ 2 & 4 & 1 & 1 & 2 & 3 \\ 2 & 1 & 4 & 3 & 2 & 1 \\ 3 & 1 & 3 & 0 & 2 & 4 \\ 3 & 2 & 2 & 2 & 3 & 1 \\ 3 & 3 & 1 & 4 & 1 & 1 \end{pmatrix},$$

which is row equitable w.r.t. $\Pi = (1|2,3|4,5,6)$ with front divisor

$$\Theta = \begin{pmatrix} 1 & 4 & 9 \\ 2 & 5 & 6 \\ 3 & 4 & 6 \end{pmatrix}.$$

One can use

$$\mathbf{H}_1 = \mathbf{H}(\mathbf{j}_1) = -1,$$

$$\mathbf{H}_2 = \mathbf{H}(\mathbf{j}_2) = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

$$\mathbf{H}_3 = \mathbf{H}(\mathbf{j}_3) = -\frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{1+\sqrt{3}}{-2} & \frac{1-\sqrt{3}}{-2} \\ 1 & \frac{1-\sqrt{3}}{-2} & \frac{1+\sqrt{3}}{-2} \end{pmatrix}$$

and $\mathbf{H}(\mathbf{j}_6, \Pi) = \operatorname{diag}(\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3)$ to transform $\mathbf{A}$ s.t.

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{j}_6, \Pi)\,\mathbf{A}\mathbf{H}(\mathbf{j}_6, \Pi) = \begin{pmatrix} 1 & \frac{4}{\sqrt{2}} & 0 & \frac{9}{\sqrt{3}} & 0 & 0 \\ \frac{4}{\sqrt{2}} & 5 & 0 & 6\frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 3 & 0 & \text{-}3\text{+}\sqrt{3} & \text{-}3\text{-}\sqrt{3} \\ \frac{9}{\sqrt{3}} & 6\frac{\sqrt{2}}{\sqrt{3}} & 0 & 6 & 0 & 0 \\ 0 & 0 & \text{-}3\text{+}\sqrt{3} & 0 & \sqrt{3}\text{-}1 & \text{-}6 \\ 0 & 0 & \text{-}3\text{-}\sqrt{3} & 0 & \text{-}6 & \text{-}\sqrt{3}\text{-}1 \end{pmatrix}.$$

Using the permutation $\tilde{\mathbf{P}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ we obtain the matrix

$$\mathbf{A}_\Pi = \tilde{\mathbf{P}}' \tilde{\mathbf{A}} \tilde{\mathbf{P}} = \begin{pmatrix} 1 & \frac{4}{\sqrt{2}} & \frac{9}{\sqrt{3}} & 0 & 0 & 0 \\ \frac{4}{\sqrt{2}} & 5 & 6\frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 & 0 \\ \frac{9}{\sqrt{3}} & 6\frac{\sqrt{2}}{\sqrt{3}} & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & \text{-}3\text{+}\sqrt{3} & \text{-}3\text{-}\sqrt{3} \\ 0 & 0 & 0 & \text{-}3\text{+}\sqrt{3} & \sqrt{3}\text{-}1 & \text{-}6 \\ 0 & 0 & 0 & \text{-}3\text{-}\sqrt{3} & \text{-}6 & \text{-}\sqrt{3}\text{-}1 \end{pmatrix},$$

which is reducible. Since $\mathbf{A}$ is hermitian, the partition Π is row and column equitable and we actually obtain a block diagonal form. Note that both blocks are hermitian although the front divisor Θ is not hermitian. One verifies that

$$\mathbf{E} = \begin{pmatrix} 1 & \frac{4}{\sqrt{2}} & \frac{9}{\sqrt{3}} \\ \frac{4}{\sqrt{2}} & 5 & 6\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{9}{\sqrt{3}} & 6\frac{\sqrt{2}}{\sqrt{3}} & 6 \end{pmatrix} = \operatorname{diag}(1,2,3)^{\frac{1}{2}} \Theta \operatorname{diag}(1,2,3)^{-\frac{1}{2}}.$$

Let $\mathbf{F}$ denote the lower diagonal block of $\mathbf{A}_\Pi$. Let $\mathbf{V_E}$ and $\mathbf{V_F}$ be the eigenvector matrices of $\mathbf{E}$ and $\mathbf{F}$, respectively. Then one verifies that

$$\mathbf{V} = \mathbf{P}'_0 \mathbf{H}\left(\mathbf{j}_6, \Pi\right) \tilde{\mathbf{P}}' \begin{pmatrix} \mathbf{V_E} & \mathbf{0} \\ \mathbf{0} & \mathbf{V_F} \end{pmatrix}$$

is an eigenvector matrix of $\mathbf{A}$. Note that the transformation is not difficult and that $\mathbf{V}$ needs more storage than $\mathbf{V_E}$ and $\mathbf{V_F}$ together. Note that $\mathbf{P}_0$ (via definition 9), $\mathbf{H}$ and $\mathbf{P}$ (via definition 11) are determined by Π_0 which can be stored as a vector. Due to the small size the blocks of $\mathbf{H}\left(\mathbf{j}_6, \Pi\right)$ were given explicitly as dense matrices. For larger problems one would prefer the usual sparse form as a rank-1-update of the identity.

2.4.5 Exploiting an Ordered Set of Equitable Partitions

Let Π and Π' be two partitions of the same set S. We call Π' a *refinement* of Π if any two elements v and w which are not in the same cell of Π must not be in the same cell of Π'. Thus, the refined Π can be interpreted as a partition of the cells of Π'. The refinement relation induces a partial ordering on the set of all partitions of S. In this subsection we consider a subset of all equitable partitions of the node set of a matrix s.t. the refinement relation induces a total order. We do not consider the exploitation of equitable partitions which are not related by refinement.

Let $\{\Pi_\alpha\}_{\alpha=1,\ldots,A}$ be an ordered set of equitable partitions of $\mathbf{A}$ with front divisors $\mathbf{\Theta}_\alpha$ s.t. Π_α is a refinement of $\Pi_{\alpha-1}$ for $2 \leq \alpha \leq A$. Let $\mathbf{B}_\alpha$ be the indicator matrices for the partition Π_α of the elements of the set. Let Π_α^γ be the partition induced by Π_α on the cells of Π_γ with $\alpha > \gamma$. Let $\mathbf{B}_\alpha^\gamma$ be the corresponding indicator matrix. Let $\mathbf{B}_\alpha^0 = \mathbf{B}_\alpha$ and $\mathbf{B}_\alpha^\alpha = \mathbf{I}$. From the properties of indicator matrices and those of refinements it is easy to see that the following two rules apply where juxtaposition is considered as to stand for the usual matrix product

$$\mathbf{B}_\alpha^\gamma = \mathbf{B}_{\gamma+1}^\gamma \mathbf{B}_{\gamma+2}^{\gamma+1} \ldots \mathbf{B}_\alpha^{\alpha-1} \ , \ \alpha \geq \gamma \tag{2.38}$$

and simply

$$\mathbf{B}_\alpha^0 = \mathbf{B}_\gamma^0 \mathbf{B}_\alpha^\gamma \ , \ \alpha \geq \gamma. \tag{2.39}$$

Using $\mathbf{A}\mathbf{B}_\alpha = \mathbf{B}_\alpha \mathbf{\Theta}_\alpha$ we observe

$$\mathbf{A}\mathbf{B}_\alpha^0 = \mathbf{A}\mathbf{B}_\gamma^0 \mathbf{B}_\alpha^\gamma = \mathbf{B}_\gamma^0 \mathbf{B}_\alpha^\gamma \mathbf{\Theta}_\alpha \tag{2.40}$$

$$= \mathbf{B}_\gamma^0 \mathbf{\Theta}_\gamma \mathbf{B}_\alpha^\gamma \tag{2.41}$$

The equality of (2.40) and (2.41) gives an overdetermined system of equations. However, since $\mathbf{B}_\gamma^0$ is an indicator matrix there are just some repeated lines and we can deduce

$$\mathbf{\Theta}_\gamma \mathbf{B}_\alpha^\gamma = \mathbf{B}_\alpha^\gamma \mathbf{\Theta}_\alpha. \tag{2.42}$$

Thus, $\boldsymbol{\Theta}_\alpha$ is a front divisor of $\boldsymbol{\Theta}_\gamma$, which is a front divisor of $\mathbf{A}$. We therefore have a kind of a chain of front divisors which induces a chain of spectra inclusions

$$\sigma\left(\mathbf{A}\right) \subset \sigma\left(\boldsymbol{\Theta}_\gamma\right) \subset \sigma\left(\boldsymbol{\Theta}_\alpha\right) \ , \ \gamma < \alpha. \tag{2.43}$$

It is obviously desirable to exploit such chains for block triangularization. We consider two ways to do that. The first one operates on the chain of the unnormalized front divisors. Although this is maybe the more intuitive approach, we recommend the second one, which operates on the normalized front divisors and has numerical advantages.

Suppose we apply our block triangularization method using Π_1. As shown above this results in a matrix of the following form where $\mathbf{N}_1 = \mathbf{B}_1' \mathbf{B}_1$

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{E}_1 & \mathbf{G}_1 \\ \mathbf{0} & \mathbf{F}_1 \end{pmatrix} \ , \ \mathbf{E}_1 = \mathbf{N}_1^{\frac{1}{2}} \boldsymbol{\Theta}_1 \mathbf{N}_1^{-\frac{1}{2}}.$$

However, at this point we seem to stuck, for in general the matrix blocks of $\mathbf{E}_1$ induced by the indicator matrix $\mathbf{B}_2^1$ do not have constant row sum.

Proceeding On Theta   We may consider replacing $\mathbf{E}_1$ by $\boldsymbol{\Theta}_1$, which can be done by a similarity transformation utilizing $\mathbf{N}_1$

$$\begin{pmatrix} \mathbf{N}_1^{-\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{E}_1 & \mathbf{G}_1 \\ \mathbf{0} & \mathbf{F}_1 \end{pmatrix} \begin{pmatrix} \mathbf{N}_1^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Theta}_1 & \mathbf{N}_1^{-\frac{1}{2}} \mathbf{G}_1 \\ \mathbf{0} & \mathbf{F}_1 \end{pmatrix}.$$

Although this seems to be a clever solution, there are drawbacks. Suppose, for a moment, $\mathbf{A}$ is hermitian. Applying our method we will always obtain a normalized hermitian front divisor and a hermitian factor. But proceeding on the unnormalized front divisor the next transformation will in general not yield hermitian blocks. It might even introduce off-diagonal blocks. We give a short example for this undesired effect. The following hermitian matrix

$$\begin{pmatrix} 4 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \tag{2.44}$$

has equitable partitions $\Pi_1 = (1|2,3)$ and $\Pi_2 = (1,2,3)$ and we have $\Pi_2^1 = (1|2)$ Using $\mathbf{H}_1 = \mathbf{H}\left(\mathbf{j}_1\right)$ and $\mathbf{H}_2 = \mathbf{H}\left(\mathbf{j}_2\right)$, which are given explicitly in the example above, we obtain

$$\begin{pmatrix} \mathbf{H}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \begin{pmatrix} 4 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} \mathbf{H}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} = \begin{pmatrix} 4 & \sqrt{2} & 0 \\ \sqrt{2} & 5 & 0 \\ 0 & 0 & \text{-}1 \end{pmatrix} \tag{2.45}$$

We may utilize $\mathbf{N}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ to get

$$\begin{pmatrix} \mathbf{N}_1 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}^{\frac{1}{2}} \begin{pmatrix} 4 & \sqrt{2} & 0 \\ \sqrt{2} & 5 & 0 \\ 0 & 0 & \text{-}1 \end{pmatrix} \begin{pmatrix} \mathbf{N}_1 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}^{-\frac{1}{2}} = \begin{pmatrix} 4 & 2 & 0 \\ 1 & 5 & 0 \\ 0 & 0 & \text{-}1 \end{pmatrix} \tag{2.46}$$

When we proceed on the unnormalized front divisor $\left(\begin{smallmatrix} 4 & 2 \\ 1 & 5 \end{smallmatrix}\right)$ we obtain

$$\begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 0 \\ 1 & 5 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} 6 & \text{-1} & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \text{-1} \end{pmatrix} \qquad (2.47)$$

Of course, the initial matrix is actually diagonizable, but that is not mirrored in the reduced form, due to the transformation in (2.46) being not hermitian. This is clearly undesirable if one thinks of larger matrices and further refinements.

The Alternative Our proposed alternative approach allows for the application of a hermitian transformation at each step. It is based on a simple observation. Let Θ be a front divisor via the partition Π. Let $\mathbf{N}$ be the diagonal matrix with diagonal entries the size of the cells of Π and let $\mathbf{E} = \mathbf{N}^{\frac{1}{2}} \Theta \mathbf{N}^{-\frac{1}{2}}$ be the normalized front divisor. Let Π' be a an equitable partition of Θ, which is assumed to be suitably labeled. Introduce $\mathbf{d} = \mathbf{N}^{\frac{1}{2}} \mathbf{j}$ and let $\mathbf{d}(i)$ denote the i-th block of $\mathbf{d}$ according to partition Π'. Then

$$\Theta_{ij}\mathbf{j}_{n_j} = \mathbf{j}_{n_i}\lambda_{ij} \Leftrightarrow \mathbf{E}_{ij}\mathbf{d}(j) = \mathbf{d}(i)\lambda_{ij} \qquad (2.48)$$

where the blocks are induced by the cells of Π'. Recall definition 7. We have

$$\mathbf{H}(\mathbf{d}(j))\mathbf{E}_{ij}\mathbf{H}(\mathbf{d}(j))\mathbf{e}_{n_j} = \lambda_{ij}\frac{\sqrt{n_i}}{\sqrt{n_j}}\mathbf{e}_{n_i}. \qquad (2.49)$$

Thus, we can obtain the normalized front divisor for Θ by working only on its own normalized version $\mathbf{E}$. Moreover, we have

$$\sigma\left(\mathbf{H}(\mathbf{d}, \Pi)\mathbf{E}\,\mathbf{H}(\mathbf{d}, \Pi)\right) = \sigma\left(\mathbf{H}(\mathbf{j}_N, \Pi)\,\Theta\,\mathbf{H}(\mathbf{j}_N, \Pi)\right). \qquad (2.50)$$

This enables us to compute the normalized front divisors successively by unitary transformations.
In our example we should therefore proceed from (2.45) with the computation of $\mathbf{d}_1 = \mathbf{N}_1^{\frac{1}{2}}\mathbf{j} = \left(\begin{smallmatrix} 1 \\ \sqrt{2} \end{smallmatrix}\right)$ and then use $\mathbf{H}(\mathbf{d}_1) = \frac{-1}{\sqrt{3}}\left(\begin{smallmatrix} 1 & \sqrt{2} \\ \sqrt{2} & \text{-1} \end{smallmatrix}\right)$ to obtain

$$\begin{pmatrix} \mathbf{H}(\mathbf{d}_1) & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} 4 & \sqrt{2} & 0 \\ \sqrt{2} & 5 & 0 \\ 0 & 0 & \text{-1} \end{pmatrix} \begin{pmatrix} \mathbf{H}(\mathbf{d}_1) & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \text{-1} \end{pmatrix}. \quad (2.51)$$

In a sense, the transformation in (2.51) combines the transformations in (2.46) and (2.47) in a way that avoids loss of unitarity.
We briefly complete our notation. By $\mathbf{E}_\alpha^\gamma$ we denote the normalization of Θ_α seen as a front divisor of Θ_γ. However, we are mostly interested in $\mathbf{E}_\alpha = \mathbf{E}_\alpha^0$, which is the normalization of Θ_α seen as a front divisor of $\mathbf{A}$. As declared above we recommend the successive reduction that operates on the normalized front divisors, which allows for an entirely unitary transformation. Considering

that, we refer with $\mathbf{F}_\alpha^\gamma$ to the factor which arises from the one-step-reduction of $\mathbf{E}_\gamma$ by Π_α^γ.

In each step of the successive reduction we subdivide the current upper left diagonal block into a 2×2 block matrix with zero lower left block. Considering the development of the block diagonal we obtain the following sequence starting with $\mathbf{A}$.

$$(\mathbf{A}), (\mathbf{E_1}, \mathbf{F_1}), \left(\mathbf{E_2}, \mathbf{F_2^1}, \mathbf{F_1}\right), \left(\mathbf{E_3}, \mathbf{F_3^2}, \mathbf{F_2^1}, \mathbf{F_1}\right), \ldots, \left(\mathbf{E}_A, \mathbf{F}_A^{A\text{-}1}, \ldots, \mathbf{F_1}\right).$$

Since the used transformations are unitary we have

$$\sigma(\mathbf{A}) = \sigma(\mathbf{E}_\alpha) + \sum_{\gamma=1}^{\alpha} \sigma\left(\mathbf{F}_\gamma^{\gamma\text{-}1}\right) \ \forall \alpha, \ 1 \leq \alpha \leq A.$$

2.4.6 Discussion

Searching for a triangular form is a standard technique in matrix computations since such a form usually corresponds to a (partial) decoupling of subsystems which allows for parallelization and reduces error caused by interference. A possible approach uses two steps. First, one searches for suitable permutations to block triangularize a (sparse) matrix, [74] or [62], which is finding the weakly connected components of the directed graph underlying the structure of non zero elements. Second, one computes a complete triangularization for each diagonal block by applying suitable transformations, as for example in the QR-decomposition or the Schur decomposition. Our approach can be seen as an intermediate step which exploits some algebraic rather then combinatorial structure of the matrix. We utilize a cheap transformation instead of a permutation to block triangularize a row equitable matrix yielding the advantage of being applicable to non sparse matrices. The used transformations are numerically stable, easily reversed and computationally cheap as will be seen below. Note that although complex matrices can be processed, all computations are real in the case of real matrices.

Computational Costs of the Transformation The costs for forming the extended Householder matrix are the accumulated costs for forming its diagonal blocks, which are matrices of rank one added to the identity matrix. Therefore they are bounded by $\sum_i O\left(n_i^2\right) \leq O\left(\left(\sum_i n_i\right)^2\right) = O\left(N^2\right)$. Applying a suitably sized elementary matrix from either side to a $M \times N$ matrix accounts for costs of order $O(MN)$. Hence the transformation of each matrix block $\mathbf{A}_{ij}$ of $\mathbf{A}$ accounts for $O(n_i n_j)$. Since the transformations of the blocks are independent from each other, the total costs are of order $\sum_{ij} O(n_i n_j) = O\left(\sum_{ij} n_i n_j\right) = O\left(N^2\right)$, which is for large N negligible compared to the costs for solving the complete eigenproblem, $O\left(N^3\right)$. The recovery of the eigenvectors is done by (a permutation of rows and) only a

left multiplication of an extended Householder matrix. Therefore, the costs for that are also bounded by $O\left(N^2\right)$.

Since the transformations of the blocks $\mathbf{A}_{ij}$ are independent, the whole transformation is parallelizable to some extend. However, one has to take into account that the sizes of the matrix blocks are fixed and in general not equal.

Matrix Compression   Instead of simplifying eigencomputations and related problems, our block triangularization also has an application in matrix compression. Given the final matrix $\tilde{\mathbf{P}}\tilde{\mathbf{A}}\tilde{\mathbf{P}}'$, and Π_0 we are able to recover $\mathbf{\Phi}_\Pi$, $\mathbf{H}\left(\mathbf{j}_N, \Pi\right)$, $\mathbf{\Omega}_\Pi$ and therefore $\mathbf{A}$ easily. Recall that all three transformations are numerical stable and applicable with $O\left(N^2\right)$ costs. The storage of Π_0 requires at most storing an integer array of length N. It is easy to see that $k \leq (N\text{-}1)$ for the number of cells of Π_0 is sufficient to ensure that $\tilde{\mathbf{P}}\tilde{\mathbf{A}}\tilde{\mathbf{P}}'$ has at least $(N\text{-}1)$ zero entries. Of course, matrices in applications may be extremely sparse in which case the transformed matrix may actually have less zeros then the initial matrix. However, even for sparse matrices the eigenvector matrix tends to be rather dense, whereas the transformed eigenvector matrix bears the same block triangular form as the transformed matrix.

The applicability as a compression method (at least for the eigenvector matrix) has in some sense a negative implication. Because of its efficiency one expects that the structure it exploits should be rather rare. Indeed, non trivial equitable partitions are exceptions for random matrices. This is plausible from the reasoning that a small perturbation on a single entry of an equitable matrix may cause a cascade effect leaving a matrix with only trivial equitable partition. However, given an empirical background equitable partitions are frequent. For instance, any automorphism on a matrix, i.e. a permutation it commutes with, induces a partition, where the cells are given by the orbits of the permutation. The induced matrix blocks apparently must be of constant row (and column) sum. The existence of non trivial automorphisms is not necessary for equitable partitions as can be deduced from our example above.

Normalized and Unnormalized Front Divisor   The normalization of the front divisor is inherent in our method. That this should not be considered an artefact can be seen from the discussion above concerning the exploitation of several equitable partitions ordered by refinement and the preservation of hermiticity. In the generalization to weighted equitable partitions below the normalized front divisor will also have an invariance property. Thus, although Θ seems to be more intuitive, $\mathbf{E}$ is perhaps more suitable for a practical application. However, when dealing with rational (integer) matrices, the unnormalized front divisor may be preferable since it is also rational (integer), whereas the normalization of the front divisor in general introduces irrational numbers due to the involved square root.

2.5 Generalizations

This section introduces three ways of generalizing equitable partitions. This will at first be done separately considering each of them as a possible generalization of the established concept of (row) equitability to avoid notational confusion as long as possible.

2.5.1 Relaxing the Constant Vector

Let Π be a labeled simultaneous partition of the row and the column set of $\mathbf{A}$ into k cells. It entails an explicit block structure on $\mathbf{A}$ when rows and columns are suitably indexed. W.l.o.g. we assume such an indexing. Let n_i denote the size of cell i. Let $\{\mathbf{v}(i)\}_{i=1,\ldots,k}$ be a set of non zero vectors such that $\mathbf{v}(i) \in \mathbf{C}^{n_i}$. Let $\mathbf{v} = \begin{pmatrix} \mathbf{v}(1) \\ \vdots \\ \mathbf{v}(k) \end{pmatrix}$. We call the pair $[\mathbf{v}, \Pi]$ a *weighted (row) equitable partition* if

$$\forall \ 1 \leq i, j \leq k \quad \mathbf{A}_{ij}\mathbf{v}(j) = \mathbf{v}(i)\theta_{ij} \tag{2.52}$$

where $\mathbf{A}_{ij}$ is the matrix block of $\mathbf{A}$ induced by row cell i and column cell j and $\mathbf{\Theta} = (\theta_{ij})$ is called the front divisor of $\mathbf{A}$ w.r.t. $[\mathbf{v}, \Pi]$. The choice $[\mathbf{j}_N, \Pi]$ corresponds to the classical notion of equitability. When we apply the same partition Π conformallly on the rows and columns of

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{v}, \Pi)\mathbf{A}\mathbf{H}(\mathbf{v}, \Pi) \tag{2.53}$$

where $\mathbf{H}(\mathbf{v}, \Pi)$ is an extended Householder transformation, then each of its blocks has the form

$$\tilde{\mathbf{A}}_{ij} = \mathbf{H}(\mathbf{v}(i))\mathbf{A}_{ij}\mathbf{H}(\mathbf{v}(j)). \tag{2.54}$$

Hence

$$\tilde{\mathbf{A}}_{ij}\mathbf{e}_{n_j} = \mathbf{e}_{n_i}\theta_{ij}\frac{|\mathbf{v}(i)|}{|\mathbf{v}(j)|}\frac{\beta(\mathbf{v}(i))}{\beta(\mathbf{v}(j))}, \tag{2.55}$$

which shows that each block is block triangular s.t. the upper left diagonal subblock is always a scalar. The special structure of each block of $\tilde{\mathbf{A}}$ allows for an explicit block triangularization. In order to obtain that form, one may use Ω_Π from definition 11. Analogous to (2.32) one obtains

$$\mathbf{\Omega}'_\Pi \tilde{\mathbf{A}} \mathbf{\Omega}_\Pi = \begin{pmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F} \end{pmatrix} \tag{2.56}$$

where $\mathbf{F} \in \mathbf{C}^{(N\text{-}k)\times(N\text{-}k)}$ is called a factor and

$$\mathbf{E} = \mathbf{N}_\beta \mathbf{N}^{\frac{1}{2}} \mathbf{\Theta} \mathbf{N}^{-\frac{1}{2}} \mathbf{N}'_\beta \in \mathbf{C}^{k\times k} \text{ with } \mathbf{N}_\beta = \text{diag}\left(\beta(\mathbf{v}(1)),\ldots,\beta(\mathbf{v}(k))\right) \tag{2.57}$$

is a normalized version of the front divisor $\boldsymbol{\Theta}$. Because of the block triangular structure we have

$$\sigma\left(\mathbf{A}\right) = \sigma\left(\mathbf{E}\right) + \sigma\left(\mathbf{F}\right).$$

We think that weighted equitable partitions are in a way a straightforward generalization of equitable partitions and its exploitation in section 2.4. We therefore just mention that hermiticity is again preserved, that eigenvectors can be recovered by left multiplication of $\mathbf{H}\left(\mathbf{v}, \Pi\right)\mathbf{P}$ and that the complexity is still of order $O\left(N^2\right)$. Instead, we consider subtleties which are due to some freedom in the entries of the $\mathbf{v}\left(i\right)$.

Scaling a Weighted Equitable Partition   Let $\{\mathbf{v}\left(i\right)\}_{i=1,\ldots,k}$ and the labeled partition $\Pi = (c_i, \ldots, c_k)$ define a weighted equitable partition of $\mathbf{A}$ with front divisor Θ. Replacing $\mathbf{v}\left(i\right)$ by $\mu_i\mathbf{v}\left(i\right)$, $\mu_i \in \mathbb{C} \setminus \{0\}$, still yields a weighted equitable partition of $\mathbf{A}$, but with different front divisor. In contrast, using (2.57) one verifies that the normalized front divisor $\mathbf{E}$ is independent of such a scaling. However, only the absolute value of its entries e_{ij} are unique. Due to the definition of β its phase might depend on the indexing of the elements of $\mathbf{v}\left(i\right)$. We do not provide a general rule how to determine a unique phase in order to avoid expensive computations but this flaw might be rather harmless in most applications. In particular, this dependency is irrelevant if all entries have the same phase factor, as in the case of the unweighted equitable partition.

Quasi-Block-Stochastic Matrices and Pseudo-Regular Graphs   The concept of quasi-block-stochastic matrices of Kuich [51] as a generalization of quasi-stochastic matrices [37] bears a close resemblance to weighted equitable partitions. The minor difference is that for quasi-block-stochastic matrices it is required that the first entry of each $\mathbf{v}\left(i\right)$ has to be 1, which is a way to make the normalized front divisor unique. Kuich also describes how to exploit this structure to trangularize a (real) matrix by a (real) similarity transformation using a theorem of Haynsworth [38] and it seems that he implicitly thinks of transformations that are computationally cheap, but doing so he seems not to consider unitarity or other precautions for numerical stability.
Another similar concept is used by Fiol and Carriga and is called *pseudo-regular* partitions. It considers matrices with positive eigenvector $\mathbf{v}$ [58, pp. 278/9]. The partition Π of the matrix is pseudo-regular if $[\mathbf{v}, \Pi]$ is weighted equitable. Since $\mathbf{v}$ is fixed up to a positive scale factor, the pseudo-quotient (i.e. front divisor) is unique.

2.5.2 Allowing Non Equitable Blocks

Let $\mathcal{P}_{bt}\left(\mathbf{A}\right)$ be the set of all labeled partitions of $\mathbf{A} \in \mathbb{C}^{N \times N}$ that induce a block triangular form i.e. if t denotes the number of cells of such a partition

$\bar{\Pi} \in \mathcal{P}_{bt}$, we have after suitable permutation

$$\boldsymbol{\Phi}'_{\bar{\Pi}} \mathbf{A} \boldsymbol{\Phi}_{\bar{\Pi}} = \begin{pmatrix} \mathbf{A}_{11} & \dots & \mathbf{A}_{1t} \\ & \ddots & \vdots \\ \mathbf{0} & & \mathbf{A}_{tt} \end{pmatrix}.$$

We explicitly include $t = 1$ and we allow that $\mathbf{A}_{ii}$ is again block triangular. Let $\bar{\Pi} \in \mathcal{P}_{bt}$ with diagonal blocks $\mathbf{A}_{ii}$, $i = 1, \dots, t$, and let Π be a refinement of $\bar{\Pi}$. We consider the restriction of Π to the i-th cell of $\bar{\Pi}$, which induces a partition of $\mathbf{A}_{ii}$. If each diagonal block $\mathbf{A}_{ii}$ is equitable w.r.t. the corresponding restriction of Π we call Π *block-wise (row) equitable*. In other words we call a partition Π block-wise (row) equitable if it is a refinement of a partition in $\mathcal{P}_{bt}$ and (row) equitable on each diagonal block induced by that coarser partition. We give an example of a block-wise equitable but not equitable partition. Consider

$$\mathbf{A} = \begin{pmatrix} \text{-2} & 1 & 3 & 0 \\ \text{-7} & 5 & 4 & 0 \\ 0 & 0 & 2 & 3 \\ \hline 0 & 0 & 0 & 5 \end{pmatrix}$$

One verifies that $\Pi = (1, 2, 3 | 4)$ is not equitable, for the upper right off-diagonal block is not a constant column. However, Π is block-wise equitable.

Since we allow for non constant row sum on some blocks we have to generalize the notion of a front divisor. We do so by defining the (normalized) front divisor for arbitrary partitions.

Definition 12. *Let $\Pi = (c_1, \dots, c_k)$ be a partition of $\mathbf{A} \in \mathbb{C}^{N \times N}$ and let $n_i = |c_i|$. Let $\mathbf{A}_{ij}$ be the matrix block of $\mathbf{A}$ induced by row cell i and column cell j. Let $\mathbf{e}_n$ be the column vector of length n which is zero on each position except the first one where he shows 1. Let*

$$e_{ij} = \mathbf{e}'_{n_i} \mathbf{H}\left(\mathbf{j}_{n_i}\right) \mathbf{A} \mathbf{H}\left(\mathbf{j}_{n_j}\right) \mathbf{e}_{n_j} = \frac{\sum_{v,w} \mathbf{A}_{ij,vw}}{\sqrt{n_i n_j}}.$$

Let

$$\theta_{ij} = e_{ij} \frac{\sqrt{n_j}}{\sqrt{n_i}}.$$

Then

$$\mathbf{E} = (e_{ij}) \quad and \quad \boldsymbol{\Theta} = (\theta_{ij})$$

are called the normalized front divisor and the (unnormalized) front divisor, respectively.

Note that definition 12 implies

$$\theta_{ij} = \frac{1}{n_i} \sum_{v,w} \mathbf{A}_{ij,vw} = \frac{1}{n_i} \mathbf{j}_{n_i} \mathbf{A}_{ij} \mathbf{j}_{n_j},$$

which can be interpreted as the average row sum of the matrix block $\mathbf{A}_{ij}$. It is applicable to any partition of $\mathbf{A}$ and is equivalent to that used in [35, corollary 2.1.] for symmetric matrices.

Let Π be a block-wise equitable partition s.t. it is equitable on the diagonal blocks of the coarser block triangular partition $\bar{\Pi}$. In order to exploit Π one has to reduce each diagonal block induced by $\bar{\Pi}$ separately. We give an example. Let $\mathbf{A} = \left(\begin{smallmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{smallmatrix} \right)$ be suitably labeled. Let Π be a partition s.t. all matrix blocks induced by Π on $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ have constant row sum, but with no restriction on the induced subblocks of $\mathbf{A}_{12}$. Therefore, Π is block-wise equitable. For $i = 1, 2$ let $\mathbf{Q}_i$ be a unitary transformation that block triangularizes $\mathbf{A}_{ii}$ exploiting its equitable partition given by the corresponding restriction of Π, i.e.

$$\mathbf{Q}_i' \mathbf{A}_{ii} \mathbf{Q}_i = \left(\begin{array}{cc} \mathbf{E}_i & \mathbf{G}_i \\ \mathbf{0} & \mathbf{F}_i \end{array} \right).$$

Then

$$\tilde{\mathbf{A}} = \left(\begin{array}{cc} \mathbf{Q}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2' \end{array} \right) \left(\begin{array}{cc} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{array} \right) \left(\begin{array}{cc} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{array} \right) = \left(\begin{array}{cccc} \mathbf{E}_1 & \mathbf{G}_1 & \mathbf{X}_{11} & \mathbf{X}_{12} \\ & \mathbf{F}_1 & \mathbf{X}_{21} & \mathbf{X}_{22} \\ \mathbf{0} & & \mathbf{E}_2 & \mathbf{G}_2 \\ & & & \mathbf{F}_2 \end{array} \right)$$

is the corresponding reduced matrix. The normalized front divisor w.r.t. Π is given by $\left(\begin{smallmatrix} \mathbf{E}_1 & \mathbf{X}_{11} \\ \mathbf{0} & \mathbf{E}_2 \end{smallmatrix} \right)$. Because of the block triangular structure the spectrum of the normalized front divisor is contained in the spectrum of $\mathbf{A}$. However, due to the relaxation that the partition may induce some non equitable matrix blocks, not all eigenvectors of the front divisor can be lifted to those of $\mathbf{A}$ simply by multiplication with the indicator matrix.

2.5.3 Separating Row and Column Partition

Let Π_1 be a labeled partition of the row set and Π_2 a labeled partition of the column set of $\mathbf{A} \in \mathbb{C}^{N_1 \times N_2}$ with k_1 and k_2 cells, respectively. They entail an explicit block structure on $\mathbf{A}$ when rows and columns are suitably indexed. If each of those blocks $\mathbf{A}_{ij}$ has constant row sum i.e.

$$\forall\, 1 \leq i \leq k_1 \,\, \forall\, 1 \leq j \leq k_2 \,\, \mathbf{A}_{ij} \mathbf{j}_{n_j} = \mathbf{j}_{n_i} \theta_{ij} \tag{2.58}$$

we call the pair Π_1 and Π_2 a *splitted row equitable* or just a *splitted equitable* partition w.r.t. $\mathbf{A}$ with *front block* $\Theta = (\theta_{ij})$. Interpreting Π_1 (Π_2) as a partition of the column (row) set of $\mathbf{A}'$, we call this pair *splitted column equitable* w.r.t. $\mathbf{A}$ if it is row equitable w.r.t. $\mathbf{A}'$. Note that (2.58) is a variant of (2.23). Since reorderings can be obtained by cheap application of permutation matrices, which are unitary, we may assume w.l.o.g. that $\mathbf{A}$ is already suitably indexed.

When we apply the same partitions Π_1 and Π_2 on the rows and columns of

$$\tilde{\mathbf{A}} = \mathbf{H}\left(\mathbf{j}_{N_1}, \Pi_1 \right) \mathbf{A} \mathbf{H} \left(\mathbf{j}_{N_2}, \Pi_2 \right) \tag{2.59}$$

then each of its blocks has the form

$$\tilde{\mathbf{A}}_{ij} = \mathbf{H}\left(\mathbf{j}_{n_i}\right)\mathbf{A}_{ij}\mathbf{H}\left(\mathbf{j}_{n_j}\right) \tag{2.60}$$

hence

$$\tilde{\mathbf{A}}_{ij}\mathbf{e}_{n_i} = \mathbf{e}_{n_j}\theta_{ij}\frac{\sqrt{n_i}}{\sqrt{n_j}}, \tag{2.61}$$

which shows that each block is block triangular. Since the first diagonal block is always a scalar, there are readily available permutations Ω_{Π_1} and Ω_{Π_2} s.t.

$$\Omega'_{\Pi_1}\tilde{\mathbf{A}}\Omega_{\Pi_2} = \begin{pmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F} \end{pmatrix} \tag{2.62}$$

where $\mathbf{E}\in\mathbf{C}^{k_1\times k_2}$ with $e_{ij}=\theta_{ij}\frac{\sqrt{n_i}}{\sqrt{n_j}}$.

Thus, exploiting a row equitable partition we always obtain a block triangularization. We call a matrix of the form $\left(\begin{smallmatrix}\mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F}\end{smallmatrix}\right)$ *weak upper block triangular*. We call the block triangularization *strong* for the special case $k_1\leq k_2$. It is easy to see that partitions which are simultaneously splitted row equitable and splitted column equitable yield a block diagonalization, i.e. $\mathbf{G}=\mathbf{0}$ in (2.62). In this case, the singular values of the front block are contained in the set of singular values of $\mathbf{A}$, i.e.

$$\sigma\left(\Theta'\Theta\right) = \sigma\left(\mathbf{E}'\mathbf{E}\right) \subset \sigma\left(\mathbf{A}'\mathbf{A}\right) \text{ for splitted row and column equitable partition.}$$

However, if $\mathbf{A}$ and Θ are square matrices, the spectrum of the latter is in general not a subset of the former, since $\mathbf{H}\left(\mathbf{j}_{N_1},\Pi_1\right)$ is not the inverse of $\mathbf{H}\left(\mathbf{j}_{N_2},\Pi_2\right)$ if $\Pi_1\neq\Pi_2$. Thus, the characteristic polynomial of a matrix is in general not divided by the characteristic polynomial of one of its front blocks.

2.5.4 The General Case and Quasi-Equitable Partitions

Let $\mathbf{A}\in\mathbb{C}^{N_1\times N_2}$. Let $\mathcal{P}_{bt}$ be the set of all pairs $(\bar{\Pi}_1,\bar{\Pi}_2)$, where $\bar{\Pi}_1$ is a labeled partition of the row set and $\bar{\Pi}_2$ is a labeled partition of the column set of $\mathbf{A}$, s.t. the cells of these partitions induce (after suitable indexing) a block triangular form on $\mathbf{A}$. Let $(\bar{\Pi}_1,\bar{\Pi}_2)\in\mathcal{P}_{bt}$. Let Π_1 be a refinement of $\bar{\Pi}_1$ with k_1 cells and Π_2 a refinement of $\bar{\Pi}_2$ with k_2 cells. Let $\mathbf{A}$ be indexed s.t. Π_1 and Π_2 entail an explicit block structure. Let $\mathbf{v}_1\in\mathbb{C}^{N_1}$ and $\mathbf{v}_2\in\mathbb{C}^{N_2}$ s.t. the vector block $\mathbf{v}_1\left(i\right)$ induced by the i-th row cell and the vector block $\mathbf{v}_2\left(j\right)$ induced by the j-th column cell are non zero and

$$\mathbf{A}_{ij}\mathbf{v}_2\left(j\right) = \mathbf{v}_1\left(i\right)\theta_{ij} \tag{2.63}$$

for all matrix blocks $\mathbf{A}_{ij}$ which are subblocks of diagonal blocks of the block triangular form given by $(\bar{\Pi}_1,\bar{\Pi}_2)\in\mathcal{P}_{bt}$. We call $\Theta=(\theta_{ij})$ with

$$\theta_{ij} = \frac{\mathbf{v}_1\left(i\right)'\mathbf{A}_{ij}\mathbf{v}_2\left(j\right)}{\mathbf{v}\left(i\right)'\mathbf{v}\left(i\right)}$$

a front block. We call a quadruple $[\mathbf{v}_1, \pi_1, \mathbf{v}_2, \Pi_2]$ that obeys the conditions above a *general row equitable* partition with front block $\boldsymbol{\Theta} = (\theta_{ij})$. Considering the discussions in this section we have that

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{v}_1, \Pi_1) \, \mathbf{A} \mathbf{H}(\mathbf{v}_2, \Pi_2)$$

is a block triangular refinement of the block triangular form induced by $(\bar{\Pi}_1, \bar{\Pi}_2) \in \mathcal{P}_{bt}$ after an appropriate reindexing. When we impose the three further conditions

(i) $\mathbf{v}_1 = \mathbf{j}_{N_1}$ and $\mathbf{v}_2 = \mathbf{j}_{N_2}$,

(ii) $\bar{\Pi}_1$ and $\bar{\Pi}_2$ are singleton partitions,

(iii) $N_1 = N_2 = N$, $\bar{\Pi}_1 = \bar{\Pi}_2 = \bar{\Pi}$, $\Pi_1 = \Pi_2 = \Pi$ and $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{v}$,

the classical equitable partition arises. Withdrawing condition (i), (ii) or (iii), we obtain weighted, block-wise or splitted equitable partitions, respectively. Although general equitable partitions can be exploited for (refined) block triangularization of rectangular matrices, this notion is too general for the purpose of eigenvalue computation. We therefore introduce a less general concept which we call quasi-equitable partitions. Those are given by the general equitable partitions which obey condition (iii). For quasi-equitable partitions we have

$$\sigma(\Theta) \subset \sigma(\mathbf{A}). \tag{2.64}$$

and we may call the front block a front divisor.

2.5.5 Relaxations of Equitable Partitions

In our notion of quasi-equitable partitions the spectrum of the front divisor $\boldsymbol{\Theta}$ is always a subset of the spectrum of the given matrix $\mathbf{A}$. There are concepts in network analysis which can be described as generalizations of (2.1) which abandon that relation. For instance, Kate and Ravindran introduced *epsilon equitable* partitions for (an adjacency matrix $\mathbf{A}$ of) a simple graph [49]. Let $\Pi = (c_1, \dots, c_k)$ be a partition of the node set of $\mathbf{A}$. Let $\mathbf{A}_{ij}$ be induced by the i-th row cell and the j-th column cell. Let $\mathbf{r}_{ij} = \mathbf{A}_{ij} \mathbf{j}_{n_j}$ be a column vector of length $n_i = |c_i|$. If

$$\forall \, 1 \leq i,j \leq k \quad \max_{1 \leq v,w \leq n_i} |\mathbf{r}_{ij,v} - \mathbf{r}_{ij,w}| \leq \epsilon \geq 0$$

then Π is called ϵ-equitable. The classical equitable partition arises for $\epsilon = 0$. Another relaxation of (2.1) can be employed to describe the concept of regular equivalence [10]. It is defined by the restriction that for a partition Π any vector $\mathbf{r}_{ij} = \mathbf{A}_{ij} \mathbf{j}_{n_j}$ must have either no zero entry or all entries zero i.e.

$$\forall \, 1 \leq i,j \leq k \quad \prod_v \mathbf{r}_{ij,v} = 0 \Rightarrow \sum_v |\mathbf{r}_{ij,v}| = 0.$$

Both concepts are used to partition the node set of a graph (=assigning roles) according to structural properties and to derive a smaller graph (the quotient or image graph) which gives a condensed representation of essential relations between the cells (=roles) of that partition. However, in both generalizations there is no notion of a front divisor in the sense of spectra inclusion.

The concept considered in [11] can be described as a generalization of $\mathbf{B}$ in (2.3) which relaxes the requirement that there is only one entry per row for the indicator matrix. In order to obtain the generalization of (2.3) one assumes that a (partial) solution for the eigenproblem is available.

We will briefly consider another relaxation of equitable partitions. The main idea is as follows. By an equitable partition a matrix $\mathbf{A}$ can be unitarily transformed into the form $\left(\begin{smallmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{0} & \mathbf{F} \end{smallmatrix}\right)$ where $\mathbf{E}$ is the normalized front divisor. If the equitability is relaxed $\mathbf{A}$ is transformed into the form $\left(\begin{smallmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{D} & \mathbf{F} \end{smallmatrix}\right)$. Our aim is to bound the norm or the rank of $\mathbf{D}$.

Relaxation   Let $\mathbf{A} \in \mathbb{C}^{N \times N}$, let $\Pi = (c_1 \ldots, c_k)$ be a partition and let

$$\mathbf{r}_{ij} = \mathbf{A}_{ij}\mathbf{j}_{n_j} \text{ and } \mathbf{d}_{ij} = \mathbf{r}_{ij} - \theta_{ij}\mathbf{j}_{n_i} \text{ with } \theta_{ij} = \frac{\mathbf{j}'_{n_i}\mathbf{r}_{ij}}{n_i}. \tag{2.65}$$

We call Π *δ-equitable* iff $\sqrt{\sum_{ij} \mathbf{d}'_{ij}\mathbf{d}_{ij}} = \delta$. One shows easily that equitable partitions are always 0-equitable since then $\mathbf{r}_{ij}$ is a constant vector. The parameter δ is a measure for the deviation of a partition from being equitable. In order to see that we now derive two certain matrices $\boldsymbol{\Delta}$ and $\mathbf{D}$.

$$\boldsymbol{\Delta} = \begin{pmatrix} \mathbf{d}_{11} & \cdots & \mathbf{d}_{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{d}_{k1} & \cdots & \mathbf{d}_{kk} \end{pmatrix} \in \mathbb{C}^{N \times k}. \tag{2.66}$$

Let $\tilde{\mathbf{d}}^1_{ij}$ denote the first entry of

$$\tilde{\mathbf{d}}_{ij} = \mathbf{H}\left(\mathbf{j}_{n_i}\right)\mathbf{d}_{ij}.$$

Using the properties of $\mathbf{H}\left(\mathbf{j}_{n_i}\right)$ and $\mathbf{d}_{ij}$ it is easy to see that

$$\tilde{\mathbf{d}}^1_{ij} = \mathbf{e}'_{n_i}\tilde{\mathbf{d}}_{ij} = 0.$$

We denote the vector of length $\left(n_i\text{-}1\right)$ that is obtained by removing the first entry of $\tilde{\mathbf{d}}_{ij}$ with $\tilde{\mathbf{d}}^-_{ij}$ and define

$$\mathbf{D} = \begin{pmatrix} \tilde{\mathbf{d}}^-_{11} & \cdots & \tilde{\mathbf{d}}^-_{1k} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{d}}^-_{k1} & \cdots & \tilde{\mathbf{d}}^-_{kk} \end{pmatrix} \in \mathbb{C}^{(N\text{-}k) \times k}. \tag{2.67}$$

Thus we have

$$\mathbf{H}\left(\mathbf{j}_N, \Pi\right) \mathbf{\Delta} = \left(\begin{array}{c} \mathbf{0}_{kk} \\ \mathbf{D} \end{array} \right).$$

Since $\mathbf{H}$ is unitary, $\mathbf{\Delta}$ and $\mathbf{D}$ have the same singular values, which yields for δ-equitable partitions

$$|\mathbf{D}|_F = |\mathbf{\Delta}|_F = \sqrt{\text{trace}\left(\mathbf{\Delta}'\mathbf{\Delta}\right)} = \delta. \tag{2.68}$$

where $|\cdot|_F$ denotes the Frobenius norm, which is the square root of the sum of the squared singular values. After some lengthy calculation or by inspection one sees that

$$\mathbf{A}_\Pi = \Omega'_\Pi \mathbf{H}\left(\mathbf{j}_N, \Pi\right) \mathbf{A} \mathbf{H}\left(\mathbf{j}_N, \Pi\right) \Omega_\Pi = \left(\begin{array}{cc} \mathbf{E} & \mathbf{G} \\ \mathbf{D} & \mathbf{F} \end{array} \right)$$

where $\mathbf{E} \in \mathbb{C}^{k \times k}$ with $e_{ij} = \frac{1}{\sqrt{n_i n_j}} \theta_{ij} = \frac{1}{n_i} \mathbf{j}'_{n_i} \mathbf{r}_{ij}$. Let

$$\tilde{\sigma} = \sigma\left(\mathbf{E}\right) + \sigma\left(\mathbf{F}\right). \tag{2.69}$$

$\tilde{\sigma}$ and $\sigma\left(\mathbf{A}\right)$ are different in general for $\delta \neq 0$, but the deviation can be bounded in terms of δ. As an example we consider the bound of Weyl [25] for hermitian matrices. In this case we have $\mathbf{G} = \mathbf{D}'$ and

$$\mathbf{A}_\Pi = \left(\begin{array}{cc} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{array} \right) + \left(\begin{array}{cc} \mathbf{0} & \mathbf{D}' \\ \mathbf{D} & \mathbf{0} \end{array} \right). \tag{2.70}$$

We interpret the second matrix on the right-hand side as a perturbation. We can bound its largest absolute eigenvalue, i.e. its largest singular value, by δ. Let $\mu_1 \leq \ldots \leq \mu_N$ and $\lambda_1 \leq \ldots \leq \lambda_N$ be the elements of $\tilde{\sigma}$ and $\sigma\left(\mathbf{A}\right)$, respectively. Since $\sigma\left(\mathbf{A}_\Pi\right) = \sigma\left(\mathbf{A}\right)$

$$|\mu_i - \lambda_i| \leq \delta \;, \; 1 \leq i \leq N \tag{2.71}$$

by the Weyl inequalities. Many more general but less convenient eigenvalue bounds are known, e.g. [26]. One can think of other characterizations for approximate equitable partitions which have moderate computational costs, for instance the number of nonzero columns of $\mathbf{\Delta}$, which upper bounds the rank of $\mathbf{D}$, or the spectral norm of $\mathbf{\Delta}$ and $\mathbf{D}$, δ_{spec}, which is the largest singular value. This implies $\delta_{spec} \leq \delta$. However, the computation of δ_{spec} is typically more costly then that of δ.

2.6 Finding Equitable Partitions

We consider the search for an equitable partition according to (2.1). If the considered matrix is binary than one can utilize existing software, e.g. nauty [60], saucy [22] or bliss [48], which are developed for a more advanced task. They

label a graph canonical or compute its automorphism group and hence are able to solve a concrete graph isomorphism problem. The computation of the equitable partition is only implemented as a generally cheap and fast and most of the time very effective subroutine in order to prune the search tree. However, although they also consider vertex and edge labelings, they do not process weighted graphs, i.e. matrices with arbitrary real or complex entries. We will describe a basic approach for finding equitable partitions of a matrix $\mathbf{A}$ of size N which is rather algebraic. If $\mathbf{A}$ is a binary matrix, a combinatorial description in terms of counting and coloring is preferable. That can be found e.g. in [50].

Let Π be partition. We call a partition which is equivalent to Π or for which Π is a refinement *weakly coarser* than Π. Note that all partitions are weakly coarser than the discrete partition, which is always equitable. We call a partition *strongly coarser* than Π when it is weakly coarser than Π but not equivalent to Π. Now let Π with k cells be equitable with indicator matrix $\mathbf{B}$, and let the elements of the set $\mathcal{E}$ be the columns of $\mathbf{B}$. Let the nonnegative integer t be a time step. Let the elements of the set $\mathcal{E}(t)$ be the columns of $\mathbf{B}(t)$ which is an indicator matrix for the partition $\Pi(t)$ with k_t cells, and let $\Pi(t)$ be weakly coarser than Π.

The complex matrix $\mathbf{V}(t) = \mathbf{A}\mathbf{B}(t)$ is of size $N \times k_t$. Let $\{a_i\}(\alpha, t)$ be the distinct entries of column α of $\mathbf{V}(t)$. Let $\mathbf{e}_i(\alpha, t)$ be a binary vector which indicates each position of column α of $\mathbf{V}(t)$ where the entry $a_i(\alpha, t)$ occurs by a *one*. Let the union of those indicator vectors over all distinct entries of all columns of $\mathbf{V}(t)$ with $\mathcal{E}(t)$ be the set

$$\mathcal{E}(t+1) = \mathcal{E}(t) \cup \bigcup_{\alpha} \bigcup_{i} \mathbf{e}_i(\alpha, t).$$

It is easy to see that $\Pi(t)$ is weakly coarser than $\Pi(t+1)$. We show that $\Pi(t+1)$ is weakly coarser than Π. Since $\Pi(t)$ is weakly coarser than Π, every column of $\mathbf{B}(t)$ lies in the span of the columns of $\mathbf{B}$, which is an invariant subspace according to (2.3). Thus every column of $\mathbf{V}(t)$ is a linear combination of the columns of $\mathbf{B}$. Since those are binary vectors, the elements of $\mathcal{E}(t+1)$ lie also in the span of the columns of $\mathbf{B}$ by construction.

It is easy to see that $\mathcal{E}(t+1)$ is uniquely determined by $\mathcal{E}(t)$. We have

$$k_t \leq k_{t+1} \leq k \tag{2.72}$$

since $\mathcal{E}(t) \subset \mathcal{E}(t+1) \subset \mathcal{E}$. By (2.3) it is also easy to see that

$$\mathcal{E}(t) = \mathcal{E}(t+1) \iff \mathcal{E}(t) = \mathcal{E}(s) \ \forall \, s \geq t \iff k_t = k_{t+1} \iff \Pi(t) \text{ is equitable.}$$

Let

$$T = \min(t | k_t = k_{t+1}). \tag{2.73}$$

We call $\Pi(T)$ the coarsest equitable partition w.r.t. $\Pi(0)$. If $\mathbf{B}(0) = \mathbf{j}_N$ we call $\Pi(T)$ just the *coarsest equitable partition*. When the coarsest equitable

partition w.r.t. $\Pi(0)$ is found, we may impose a refinement on $\Pi(T)$ in order to find other equitable partitions. A more careful analysis reveals that a labeling for the elements of $\mathcal{E}(t)$ induces a labeling for the elements of $\mathcal{E}(t+1)$. This can be seen from the following reasoning. An ordering of the columns of $\mathbf{B}(t)$ induces an ordering for the columns of $\mathbf{V}(t)$, and since the possible values for the a_i are known and we therefore can impose an ordering on them a priori, this induces an ordering on the $\mathbf{e}_i$, hence for $\mathcal{E}(t+1)$.

One can process the columns of $\mathbf{B}(t)$ one at a time. As soon as the set $\mathcal{E}(t)$ increases a refined partition can be imposed. Columns which have already been considered and which have not changed after a refinement do not have to be considered twice. Thus, it is sufficient to consider single cells once. When a fast-matrix vector multiplication is available, the columns of $\mathbf{B}(t)$ may be processed in parallel considering a vector $\mathbf{w}(t)$ which is a linear combination of those. In order to be discriminating, the coefficients obviously have to be distinct (or *zero* else). That distinctness is not sufficient can be seen from the following example.

$$\mathbf{A}=\begin{pmatrix} 0 & 1 & 2 \\ 1 & 4 & 0 \\ 2 & 0 & 5 \end{pmatrix}, \ \mathcal{E}(t)=\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}, \ \mathbf{w}(t)=\begin{pmatrix} \text{-}1 \\ 1 \\ 1 \end{pmatrix} \rightarrow \mathbf{A}\mathbf{w}(t)=3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \quad (2.74)$$

Here, the particular choice of $\mathbf{w}(t)$ does not lead to a refinement. However, this is an exceptional case. Any other vector in the span of $\mathcal{E}(t)$ in (2.74) which is not a multiple of $(-1,1,1)'$ works fine. In application, at a given time step t one may choose a random linear combination of the columns of $\mathbf{B}(t)$ (see also [69]). If no refinement is obtained one may try another set of coefficients and eventually consider each column of the current indicator matrix $\mathbf{B}(t)$ separately to ensure equitability.

When dealing with real or complex entries of $\mathbf{A}$ in practical numerical computations, we have to take rounding errors into account. This means that we should refrain from imposing a refinement if the computed difference of two distinct entries in a column of $\mathbf{V}(t)$ is too small. In order to cope this problem, we employed a procedure in our numerical experiments which is basically as follows. Consider a column of $\mathbf{V}(t)$, divide the entries according to the current partition and sort them within each cell. If and only if the difference of consecutive elements of a given cell is greater than a certain reasonable threshold then impose a refinement. This procedure is designed to guarantee that any non refined partition found numerically is weakly coarser than the coarsest equitable partition w.r.t. $\mathbf{B}(0)$. However, it would be desirable to achieve that in a more efficient way.

Chapter 3

Isospectrality of Graphs

In this chapter we introduce a family of graph matrices which contains e.g. the adjacency matrix, the Laplacian or the Seidel matrix. We generalize the so called GM-switching technique and give general constructions for isosepctral graphs which may or may not utilize edge switching. They can be used to find graph pairs which are unitary equivalent (implying isospectral) w.r.t. the introduced matrix family.

3.1 Introduction

Let Γ_i, $i = 1, 2$, be two (di-)graphs with adjacency matrices $\mathbf{A}_i$, respectively. Let $\mathbf{X}(\Gamma_i)$ be any matrix representation e.g. adjacency matrix, Laplacian or Seidel matrix for Γ_i. We call Γ_1 and Γ_2 *isospectral* w.r.t. $\mathbf{X}$ if

$$\sigma(\mathbf{X}(\Gamma_1)) = \sigma(\mathbf{X}(\Gamma_2)). \tag{3.1}$$

In that case we also say that $\mathbf{X}(\Gamma_1)$ and $\mathbf{X}(\Gamma_2)$ are isospectral. We call Γ_1 and Γ_2 *unitary equivalent* w.r.t. $\mathbf{X}$ if there exists a unitary matrix $\mathbf{Q}$ s.t.

$$\mathbf{X}(\Gamma_2) = \mathbf{Q}'\mathbf{X}(\Gamma_1)\mathbf{Q}. \tag{3.2}$$

We call both, the matrix pair and the graph pair, *isomorphic* if there is even a permutation matrix $\mathbf{Q}$ s.t. (3.2). In fact, in that case it would be somewhat more appropriate to speak of $\mathbf{X}(\Gamma_i)$, $i = 1, 2$, as two isomorphic representations of the same graph $\Gamma_1 = \Gamma_2$. That no polynomial algorithm is known which can decide if two given arbitrary graphs have isomorphic matrix representations is at the heart of the famous graph isomorphism problem [64]. It is easy to see that

$$\text{isomorphic} \ \Rightarrow \ \text{unitary equivalent} \ \Rightarrow \ \text{isospectral}.$$

The second implication is a tautology for undirected graphs, which can be diagonalized by unitary matrices. However, the pair $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ shows that it can in general not be reversed for directed graphs. The first implication is well known to be in general not reversible. Counter examples are given by isospectral or cospectral graphs which are not isomorphic. The smallest pair of isospectral but not isomorphic graphs (w.r.t. the adjacency matrix) is given by the complete bipartite graph $K_{1,4}$ and the union of the cycle of length 4, C_4, and an isolated node K_1, with spectrum $\{-2, 0, 0, 0, 2\}$ [20]. Thus, although the spectrum of the adjacency matrix is a graph invariant, a property of a graph

that is independent of its representation, it does not distinguish all graphs. The same is true for other well known graph matrices, like the combinatorial Laplacian or the normalized Laplacian. An example of a set of graphs which are isospectral w.r.t. the latter graph matrix is the *night sky* $\mathcal{N}_{s,r}$. Each element of $\mathcal{N}_{s,r}$ has r edges and is the union of s star graphs with at least two vertices s.t. the number of all vertices is $r + s$. Since the spectrum w.r.t. the normalized Laplacian of a star graph with $n \geq 2$ vertices is $\left\{\text{-}1, 0^{n\text{-}2}, 1\right\}$, each element of $\mathcal{N}_{s,r}$ has spectrum $\left\{(\text{-}1)^s, 0^{r\text{-}s}, 1^s\right\}$ w.r.t. the normalized Laplacian. It is easy to see that e.g. $\mathcal{N}_{2,4}$ has two nonisomorphic elements. A much more general construction is given in [15].

An intriguing example of a graph invariant is the 2-dimensional Weisfeiler-Lehman-stabilizer or its equivalent, the coherent algebra generated by the adjacency matrix $\mathbf{A}$, the all-ones matrix $\mathbf{J}$ and the identity $\mathbf{I}$, which we will denote by $\mathrm{WL}_2(\mathbf{I}, \mathbf{A}, \mathbf{J})$. It is the closure of $\{\mathbf{I}, \mathbf{A}, \mathbf{J}\}$ w.r.t. complex conjugate transposition (i.e. adjoint), complex linear combination, the usual matrix product and the Hadamard-Schur (or point-wise) product [24], [41].

Structure of the Chapter   We introduce the matrix algebra $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$ generated by $\{\mathbf{I}, \mathbf{A}, \mathbf{J}\}$ and the closure operator WL_1 defined below, which corresponds to the 1-dimensional Weisfeiler-Lehman-stabilizer. Before that we recall a classical construction for isospectral graphs, Seidel switching, and its substantial extension by Godsil and McKay. We give a slight generalization for the latter and a general construction for unitary equivalent graphs. Later, we show that it can be used to produce graphs which are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$.

3.2 Isospectral graphs

3.2.1 Seidel-Switching

Using edge switching to produce isospectral graphs was proposed by Seidel [82]. The idea is as follows. Suppose $\mathbf{A}$ is the adjacency matrix of a graph Γ without loops. The matrix $\mathbf{S} = \mathbf{J} - 2\mathbf{A} - \mathbf{I}$ is called the Seidel matrix of Γ. We have

$$s_{vw} = \begin{cases} 0 & , & v = w \\ \text{-}1 & , & (v, w) \text{ is an edge in } \Gamma \\ 1 & , & (v, w) \text{ is not an edge in } \Gamma \end{cases} .$$

Suppose a disjoint partition of the vertex set, $V = V_1 \cup V_2$, which induces a block partition of

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \text{ and } \mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} .$$

Since $\mathbf{Q} = \left(\begin{smallmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{smallmatrix} \right)$ is orthogonal, the graph $\tilde{\Gamma}$ represented by the Seidel matrix

$$\tilde{\mathbf{S}} = \mathbf{QSQ} = \left(\begin{array}{cc} \mathbf{S}_{11} & \text{-}\mathbf{S}_{12} \\ \text{-}\mathbf{S}_{21} & \mathbf{S}_{22} \end{array} \right)$$

has the same Seidel spectrum as Γ. It is easy to verify that the adjacency matrix of $\tilde{\Gamma}$ is given by

$$\tilde{\mathbf{A}} = \left(\begin{array}{cc} \mathbf{A}_{11} & \mathbf{J} - \mathbf{A}_{12} \\ \mathbf{J} - \mathbf{A}_{21} & \mathbf{A}_{22} \end{array} \right). \tag{3.3}$$

We can describe that construction by a partition of the vertex set of Γ into two sets V_1 and V_2 indicated by $\mathbf{Q}$. The graph $\tilde{\Gamma}$ is obtained by removing all present edges between V_1 and V_2 and turning all nonpresent edges between that pair into present ones. This operation is often called edge switching.

3.2.2 GM-Switching and Q-Switching

The transformation of $\mathbf{A}$ into $\tilde{\mathbf{A}}$ in (3.3) of the last subsection does in general not correspond to a similarity transformation, hence does not preserve the adjacency spectrum. However, if $\mathbf{A}_{22}$ is of even size n and has constant row and column sum r and the sum of each row in $\mathbf{A}_{12}$ and the sum of each column in $\mathbf{A}_{21}$ is $\frac{n}{2}$, i.e. in the case of simple graphs V_2 is a $(r, \frac{n}{2})$-regular set [17], then one verifies that the transformation of $\mathbf{A}$ to $\tilde{\mathbf{A}}$ can be described by the unitary transformation $\left(\begin{smallmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}(n) \end{smallmatrix} \right)$, where

$$\mathbf{Q}\left(n\right) = \frac{2}{n}\mathbf{J}_{n,n} - \mathbf{I}_n. \tag{3.4}$$

A generalization of that idea leads to the well known construction of Godsil and McKay [31] [81] called GM-switching [36], which produces pairs of graphs that are isospectral with respect to the adjacency matrix and its complement. In this subsection we describe a slightly more general version, which we denote by Q-switching. Before that we point out the essential difference. A key observation is that the clever transformation used in GM-switching is block diagonal in correspondence with a partition of $\mathbf{A}$. Each diagonal block is of the form $\mathbf{Q}\left(n\right)$ with the exception of one block which shows the identity matrix $\mathbf{I}$. A trivial reformulation interprets the diagonal elements of $\mathbf{I}$ as $\mathbf{Q}\left(1\right)$, which corresponds to a refined partition of $\mathbf{A}$. However, we show that the occurrence of $\mathbf{Q}\left(1\right)$ as a diagonal block is not necessary. Another line of generalization can be found in [1].

Let $\Pi = (c_1, \ldots, c_k)$ be a partition of the first N positive integers $[1, \ldots, N]$ s.t. $v \in c_i$, $w \in c_j$ and $i < j$ implies $v < w$. Let $n_i = |c_i|$. Define the unitary block diagonal matrix

$$\mathbf{Q} = \left(\begin{array}{ccc} \mathbf{Q}\left(n_1\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{Q}\left(n_k\right) \end{array} \right) \text{ with } \mathbf{Q}\left(n\right) = \frac{2}{n}\mathbf{J}_{n,n} - \mathbf{I}_n \tag{3.5}$$

and let

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{k1} & \cdots & \mathbf{A}_{kk} \end{pmatrix} \text{ and } \widetilde{\mathbf{A}} = \mathbf{Q}\mathbf{A}\mathbf{Q} = \begin{pmatrix} \widetilde{\mathbf{A}}_{11} & \cdots & \widetilde{\mathbf{A}}_{1k} \\ \vdots & \ddots & \vdots \\ \widetilde{\mathbf{A}}_{k1} & \cdots & \widetilde{\mathbf{A}}_{kk} \end{pmatrix}$$

be matrices partitioned by Π s.t. the matrix blocks $\mathbf{A}_{ij}$ and $\widetilde{\mathbf{A}}_{ij}$ are induced by the i-th row cell and the j-th column cell. We have $\widetilde{\mathbf{A}}_{ij} = \mathbf{Q}(n_i)\mathbf{A}_{ij}\mathbf{Q}(n_j)$. Since $\mathbf{Q}$ is unitary, $\widetilde{\mathbf{A}}$ and $\mathbf{A}$ are unitary equivalent.

In order to derive a construction for isospectral adjacency matrices, we aim at finding conditions on a binary $\mathbf{A}$ s.t. a transformation by $\mathbf{Q}$ results in a binary matrix again. First, we consider some general properties of the transformation $\widetilde{\mathbf{M}} = \mathbf{Q}(n_i)\mathbf{M}\mathbf{Q}(n_j)$ on an arbitrary rectangular matrix $\mathbf{M} \in \mathbb{C}^{n_i \times n_j}$. Then, we consider the special case where $\mathbf{M}$ and $\widetilde{\mathbf{M}}$ are binary.

Complex case   We define

$$\mathbf{s}' = \frac{1}{n_i}\mathbf{j}'_{n_i}\mathbf{M}\left(\mathbf{Q}(n_j) - \mathbf{I}_{n_j}\right) \in \mathbb{C}^{1 \times n_j}$$

$$\mathbf{z} = \frac{1}{n_j}\left(\mathbf{Q}(n_i) - \mathbf{I}_{n_i}\right)\mathbf{M}\mathbf{j}_{n_j} \in \mathbb{C}^{n_i \times 1},$$

$$\mathbf{\Delta} = \widetilde{\mathbf{M}} - \mathbf{M} \text{ and } \operatorname{vol}(\mathbf{M}) = \mathbf{j}_{n_i}\mathbf{M}\mathbf{j}_{n_j}.$$

The function $\operatorname{vol}(\cdot)$ might be evaluated for any matrix in the obvious interpretation that it just sums up all entries. We collect some properties

(i) $\mathbf{\Delta} = \mathbf{j}_{n_i}\mathbf{s}' + \mathbf{z}\mathbf{j}'_{n_j}$,

(ii) $\mathbf{s}$ is a constant vector $\Leftrightarrow \mathbf{s} = \mathbf{0} \Leftrightarrow \mathbf{M}$ is column regular,
     $\mathbf{z}$ is a constant vector $\Leftrightarrow \mathbf{z} = \mathbf{0} \Leftrightarrow \mathbf{M}$ is row regular,

(iii) $\mathbf{s}'\mathbf{j}_{n_j} = 0 = \mathbf{j}'_{n_i}\mathbf{z}$ and $\operatorname{vol}(\mathbf{\Delta}) = 0$.

(iv) $\widetilde{\mathbf{M} + \widetilde{\mathbf{M}}} = \mathbf{Q}(n_i)\left(\mathbf{M} + \widetilde{\mathbf{M}}\right)\mathbf{Q}(n_j) = \mathbf{M} + \widetilde{\mathbf{M}}$

Proof. (i), (ii), (iii) and (iv) are easily verified using the definition of $\mathbf{Q}(n)$, $\mathbf{s}$, $\mathbf{z}$ and $\mathbf{\Delta}$. $\square$

From (i) we conclude that $\operatorname{rank}(\mathbf{\Delta}) \leq 2$. The next proposition distinguishes the three possible values for the rank of $\mathbf{\Delta}$.

Proposition 1.   • $\mathbf{\Delta} = \mathbf{0} \Leftrightarrow \mathbf{M}$ *is row and column regular.*

- rank $(\boldsymbol{\Delta}) = 1 \Leftrightarrow \mathbf{M}$ *is either row or column regular.*

- rank $(\boldsymbol{\Delta}) = 2 \Leftrightarrow \mathbf{M}$ *neither row nor column regular.*

Proof. All three statements follow from (i) and (ii). □

The Binary Case From now on we additionally assume that $\mathbf{M}$ *and* $\widetilde{\mathbf{M}}$ are (0,1)-matrices which implies $\boldsymbol{\Delta} \in \{-1, 0, 1\}^{n_i \times n_j}$. Considering a fixed row (column) of $\boldsymbol{\Delta}$ and using (i) we see that $\mathbf{s}$ ($\mathbf{z}$) has at most three different entries, say $a \leq b \leq c$ ($d \leq e \leq f$). According to proposition 1 we distinguish three cases.

Proposition 2. $\mathbf{M}$ *is row and column regular* $\Leftrightarrow \mathbf{M} = \widetilde{\mathbf{M}}$

Proof. This follows from the first statement in proposition 1. □

Proposition 3.
$\mathbf{M}$ *is row regular and not column regular* $\Leftrightarrow a = $ -1$, c = 1 \wedge d = e = f = 0$.
$\mathbf{M}$ *is column regular and not row regular* $\Leftrightarrow a = b = c = 0 \wedge d = $ -1$, f = 1$.

Proof. If $\mathbf{M}$ is row regular and not column regular, then $\mathbf{z} = \mathbf{0}$ and $\mathbf{s} \neq \mathbf{0}$ by (ii). Since $\mathbf{s}'\mathbf{j}_{n_j} = 0$, we have $a = $ -1 and $c = 1$. The second assertion is proved analogously. □

We consider the structure of a row but not column regular binary matrix $\mathbf{M}$ s.t. the transformed matrix is also binary. Since $a = $ -1$, c = 1$ and $\mathbf{z} = \mathbf{0}$, there is a permutation matrix $\mathbf{P}$ and possibly a row and column regular matrix $\mathbf{B}$ s.t.

$$\mathbf{s}' = \left(-\mathbf{j}'_{k_1}, \mathbf{0}'_{k_2}, \mathbf{j}'_{k_1}\right) \mathbf{P} \text{ with } 2k_1 + k_2 = n_j \, , \, k_1 \geq 1 \, , \, k_2 \geq 0 \qquad (3.6)$$

and

$$\mathbf{M} = (\mathbf{J}, \mathbf{B}, \mathbf{0}) \mathbf{P} \, , \, \widetilde{\mathbf{M}} = (\mathbf{0}, \mathbf{B}, \mathbf{J}) \mathbf{P} \qquad (3.7)$$

Considering $\mathbf{s}$ it is easy to see that $\mathrm{vol}\,(\mathbf{B}) = \frac{1}{2}k_2 n_i$. Thus, each column and each row of $\mathbf{B}$ contains as many zeros as ones. We summarize

$$\mathbf{B} \in \{0, 1\}^{n_i \times k_2} \text{ has constant row and column sum, } \mathrm{vol}\,(\mathbf{B}) = \frac{k_2 n_i}{2}. \quad (3.8)$$

One shows that conditions (3.6) and (3.8) suffice to ensure that a binary matrix block $\mathbf{M} = (\mathbf{J}, \mathbf{B}, \mathbf{0}) \mathbf{P}$ with first and third subblock of equal size transforms into the binary matrix block $\widetilde{\mathbf{M}} = (\mathbf{0}, \mathbf{B}, \mathbf{J}) \mathbf{P}$.
The case of a column but not row regular matrix $\mathbf{M}$ is analogous. One may exploit that $\mathbf{M}$ is column regular if and only if $\mathbf{M}'$ is row regular.

Proposition 4.
$\mathbf{M}$ *is neither row nor column regular* $\Leftrightarrow a < 0 < b = c \wedge d < 0 < e = f$

Proof. '⇒'. By proposition 1 rank $(\boldsymbol{\Delta}) = 2$. The assumption $a = b = c$ implies $\mathbf{s} = \mathbf{0}$, which implies rank $(\boldsymbol{\Delta}) \leq 1$. Analogous for the assumption $d = e = f$. By (i) and disregarding multiplicity of elements, we have

$$\{a + d, b + d, c + d, a + e, b + e, c + e, a + f, b + f, c + f\} \subset \{\text{-}1, 0, 1\}.$$

By that inclusion the assumption $a < b < c$ implies $d = e = f$, hence it implies rank $(\boldsymbol{\Delta}) \leq 1$. A similar argument applies for the assumption $d \neq e \neq f$. Therefore, both $\mathbf{s}$ and $\mathbf{z}$ have exactly two different entries. By (iii) both entries in each vector are of different sign.

'⇐'. By proposition 1. $a \neq b \wedge d \neq e$ implies rank $(\boldsymbol{\Delta}) = 2$ according to (i). □

Now we consider the structure of a neither row nor column regular binary $\mathbf{M}$. By proposition 4 and (i), $a + d = \text{-}1$, $0 = a + e = b + d$, $b + e = 1$. Hence, there are two permutations $\mathbf{P}_i$ and $\mathbf{P}_j$ s.t.

$$\boldsymbol{\Delta} = \mathbf{P}_i' \begin{pmatrix} -\mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} \end{pmatrix} \mathbf{P}_j$$

where the left upper diagonal block of $\boldsymbol{\Delta}$ is of size $m_i \times m_j$. This implies

$$\mathbf{M} = \mathbf{P}_i' \begin{pmatrix} \mathbf{J} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{pmatrix} \mathbf{P}_j \text{ and } \widetilde{\mathbf{M}} = \mathbf{P}_i' \begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{J} \end{pmatrix} \mathbf{P}_j. \tag{3.9}$$

with conformable partitions and

$$\mathbf{B} \in \{0, 1\}^{m_i \times (n_j - m_j)} \text{ and } \mathbf{C} \in \{0, 1\}^{(n_i - m_i) \times m_j} \text{ with } \begin{cases} 1 \leq m_i < n_i \\ 1 \leq m_j < n_j \end{cases}. \tag{3.10}$$

By (iv) and proposition 1, $\mathbf{M} + \widetilde{\mathbf{M}}$ is row and column regular. Therefore, $\mathbf{B}$ and $\mathbf{C}$ are row and column regular with

$$\delta^{\text{row}} (\mathbf{B}) + m_j = \delta^{\text{row}} (\mathbf{C}) + \frac{n_j}{2} \text{ and}$$
$$\delta^{\text{col}} (\mathbf{C}) + m_i = \delta^{\text{col}} (\mathbf{B}) + \frac{n_i}{2}. \tag{3.11}$$

where $\delta^{\text{row}} (\cdot)$ and $\delta^{\text{col}} (\cdot)$ give the constant row and column sum, respectively. By (iii) we also have,

$$m_i m_j = (n_i - m_i) (n_j - m_j). \tag{3.12}$$

One verifies that conditions (3.11) and (3.12) with (3.10) are already sufficient to ensure that the result of the transformation of a binary matrix block $\mathbf{M}$ formed as on the left side of (3.9) is again binary.

We remark that the partition Π of $\mathbf{A}$ is equitable if and only if $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{A}$, which follows from proposition (2). Since every partition of $\mathbf{J}$ is equitable,

Q always commutes with **J**. Therefore, every pair of graphs obtained by Q-switching is simultaneous unitary equivalent w.r.t. any polynomial in **I**, **A** and **J**. We also remark that we implicitly assumed that the presence (nonpresence) of a loop at vertex i is indicated by a *one* (a *zero*) at diagonal position ii to achieve strictly binary adjacency matrices. If the presence of a loop is indicated by an other number, we may impose the restriction that diagonal blocks of **A** always have to be row and column regular or simply not allow loops.

GM-Switching arises by the restriction that there exists at least one i s.t. $n_i = 1$ and that all matrix blocks having more than one row and more than one column must be row and column regular. To avoid some trivial cases one may add the assumption that there is at least one k s.t. n_k is even. Although nonisomorphism of the obtained graph pair is not guaranteed, it is a typical outcome [31].

In the next section we give three other constructions for unitary equivalent graphs the first two being generalized by the third one.

3.2.3 Further Constructions

Proposition 5. *Let* $\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix}$ *and* $\mathbf{Q}_+ = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_2 \end{pmatrix}$ *be*

unitary. Let

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \text{ and } \tilde{\mathbf{A}} = \mathbf{Q}'\mathbf{A}\mathbf{Q} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{pmatrix}$$

be adjacency matrices partitioned conformable with **Q**. *Let*

$$\mathbf{A}_+ = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{22} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{22} \end{pmatrix} \text{ and } \tilde{\mathbf{A}}_+ = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} & \tilde{\mathbf{A}}_{22} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} & \tilde{\mathbf{A}}_{22} \end{pmatrix}.$$

Then $\tilde{\mathbf{A}}_+ = \mathbf{Q}'_+\mathbf{A}_+\mathbf{Q}_+$

Proof. Exploit that $\tilde{\mathbf{A}}_{\mathbf{ij}} = \mathbf{Q}'_i\mathbf{A}_{ij}\mathbf{Q}_j$ for $i, j \in \{1, 2\}$. $\square$

Note that $\mathbf{Q}_{1/2}$ may itself be block diagonal and that the construction may be applied repeatedly. When **Q** yields a Q-switching on **A**, then $\mathbf{Q}_+$ yields a Q-switching on $\mathbf{A}_+$. The construction can be described conveniently by a Khatri-Rao product [56].

$$\mathbf{A}_+ = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \text{ and } \mathbf{Q}_+ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \circ \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix}$$

(3.13)

From this description generalizations can be inferred easily.

Proposition 6. *Let* $\mathbf{A}_i \in \{0,1\}^{n_i \times n_i}$ *be adjacency matrices and* $\mathbf{Q}_i \in \mathbb{C}^{n_i \times n_i}$ *be unitary matrices,* $i = 1, 2$ *and let* $\mathbf{B} \in \{0,1\}^{n_1 \times n_2}$ *s.t.*

$$\tilde{\mathbf{A}}_i = \mathbf{Q}_i' \mathbf{A}_i \mathbf{Q}_i \in \{0,1\}^{n_i \times n_i}, \ i = 1, 2, \ \tilde{\mathbf{B}} = \mathbf{Q}_1' \mathbf{B} \mathbf{Q}_2 \in \{0,1\}^{n_1 \times n_2}. \quad (3.14)$$

Let

$$\mathbf{A} = \left(\begin{array}{cc} \mathbf{A}_1 & \mathbf{B} \\ \mathbf{B}' & \mathbf{A}_2 \end{array} \right), \ \tilde{\mathbf{A}} = \left(\begin{array}{cc} \tilde{\mathbf{A}}_1 & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}' & \tilde{\mathbf{A}}_2 \end{array} \right) \ and \ \mathbf{Q} = \left(\begin{array}{cc} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{array} \right).$$

Then $\mathbf{A}$ *and* $\tilde{\mathbf{A}}$ *are adjacency matrices and* $\tilde{\mathbf{A}} = \mathbf{Q}' \mathbf{A} \mathbf{Q}$.

Proof. Exploit (3.14). $\qquad\qquad\square$

If both triples $\mathbf{A}_i$, $\mathbf{Q}_i$ and $\tilde{\mathbf{A}}_i$, $i = 1, 2$, correspond to a Q-switching, then the choice $\mathbf{B} = \mathbf{J}$ is always possible. Choosing $\mathbf{B}$ to be all ones works also, for instance, in the case of $\mathbf{Q}_1 = \mathbf{I}_{n_1}$ and regular $\mathbf{A}_2$ and regular $\tilde{\mathbf{A}}_2$ which implies $\mathbf{J}\mathbf{Q}_2 = \mathbf{J}$.

A construction generalizing the previous two is given in the following lemma.

Lemma 2. *Let* $\mathcal{U} = \{\mathbf{Q}_1, \ldots, \mathbf{Q}_k\}$ *be an ordered set of unitary matrices of sizes* $n_1, \ldots, n_k$, *respectively. Let* $\mathcal{A}$ *be a set of* k^2 *distinct subsets* $\mathcal{A}_{ij}$, $1 \le i, j \le k$. *Let the elements of each subset be ordered pairs of binary rectangular matrices* $(\mathbf{A}_{ij}, \tilde{\mathbf{A}}_{ij})$, *both of size* $n_i \times n_j$, *s.t.* $\tilde{\mathbf{A}}_{ij} = \mathbf{Q}_i' \mathbf{A}_{ij} \mathbf{Q}_j$. *Choose any sequence of numbers of the integer interval* $[1, k]$ *to form a word* ω *of length* N. *Denote the number at the i-th position of* ω *by* ω_i. *Let* $\mathbf{A}_+$ *($\tilde{\mathbf{A}}_+$) be a block matrix with block* ij *of size* $n_{\omega_i} \times n_{\omega_j}$, $1 \le i, j \le N$. *For all pairs* ij *choose a matrix pair in* $\mathcal{A}_{\omega_i \omega_j}$ *and let* $\mathbf{A}_{\omega_i \omega_j}$ *($\tilde{\mathbf{A}}_{\omega_i \omega_j}$) be the matrix block at* ij *of* $\mathbf{A}_+$ *($\tilde{\mathbf{A}}_+$). Replace the i-th 1 of the diagonal of the identity* $\mathbf{I}_N$ *by* $\mathbf{Q}_{w_i}$ *and each off-diagonal zero by a suitably sized all-zero matrix block to form the block diagonal matrix* $\mathbf{Q}_+$. *Then*

$$\mathbf{Q}_+' \mathbf{A}_+ \mathbf{Q}_+ = \tilde{\mathbf{A}}_+. \quad (3.15)$$

Proof. Since $\mathbf{Q}_+$ is block diagonal the ij matrix block of (3.15) is given as $\mathbf{Q}_{\omega_i}' \mathbf{A}_{\omega_i \omega_j} \mathbf{Q}_{\omega_j} = \tilde{\mathbf{A}}_{\omega_i \omega_j}$, which conforms to the properties of $\mathcal{A}$. $\qquad\square$

In order to build the sets $\mathcal{U}$ and $\mathcal{A}$ we may exploit the structure of some known graph pair as indicated by propositions 5 and 6. Note that $\mathcal{U}$ does not have to be known to construct the pair $\mathbf{A}_+$ and $\tilde{\mathbf{A}}_+$. It is convenient that a pair of suitably sized zero matrices is always an appropriate element of a set $\mathcal{A}_{ij}$. If all elements of $\mathcal{U}$ are block diagonal with diagonal blocks of the form $\frac{2}{n}\mathbf{J}_{nn} - \mathbf{I}_n$, then the set $\mathcal{A}$ can be constructed using the properties of Q-switching.

For a more concrete application of the construction consider isospectral pairs $(\mathbf{A}_i, \tilde{\mathbf{A}}_i)$ of adjacency matrices of regular simple graphs of sizes N_i, respectively. By hermiticity, regularity and isospectrality there are unitary matrices

$\mathbf{Q}_i$ s.t. $\tilde{\mathbf{A}} = \mathbf{Q}_i' \mathbf{A}_i \mathbf{Q}_i$ with $\mathbf{Q}_i \mathbf{j} = \mathbf{j}$. Therefore, we can construct $\mathbf{A}_+$ and $\tilde{\mathbf{A}}_+$ by

$$\mathbf{A}_{+,ii} = \mathbf{A}_i, \ \tilde{\mathbf{A}}_{+,ii} = \tilde{\mathbf{A}}_i \text{ and } \mathbf{A}_{+,ij} = \tilde{\mathbf{A}}_{ij} \in \{\mathbf{0}_{N_i N_j}, \mathbf{J}_{N_i N_j}\} \forall \ i \neq j.$$

Thus, by lemma 2 any according pair of *generalized compositions* [68] obtained from a family of isospectral pairs of regular simple graphs is unitary equivalent.

3.3 Algebras of Graph Matrices

This section introduces families of graph matrices generated by the adjacency matrix, which generalize some well studied examples of such families using ideas and notation from [24] and [41]. The next three subsections provide some basic definitions and introduce the closure operators WL_0, WL_1 and WL_2. The following section gives some properties of the first two algebras and particularly shows that the construction of lemma 2 can be used to build graphs which are simultaneous unitary equivalent w.r.t. those algebras.

3.3.1 $\mathbb{C}$-Valued Functions on a Set

Let X be a finite set and let $L(X)$ be the *set of $\mathbb{C}$-valued functions on X*. Since $\mathbb{C}$ is a $\mathbb{C}$-algebra, $L(X)$ is a $\mathbb{C}$-*algebra* under *pointwise operations*. We may identify $L(X)$ with the vector space $\mathbb{C}^{|X|}$ equipped with an additional pointwise multiplication and entries indexed by the elements of X. The *multiplicative identity*[1] is given by

$$\mathbf{j} \in L(X), \ \mathbf{j}(x) = 1 \ \forall x \in X.$$

A subalgebra is called *unital* if it contains $\mathbf{j}$. A vector $\mathbf{v} \in L(X)$ can be interpreted as a *coloring of a subset* of X, where $\mathbf{v}(x) = c \neq 0$ indicates that the element $x \in X$ belongs to a given subset and is colored with color c and $\mathbf{v}(x) = 0$ indicates the non presence of x in the given subset. Let $\mathcal{V}$ be a subalgebra of $L(X)$. Let $\mathbf{v} \in \mathcal{V}$ with entries $\mathbf{v}(x) \in \{c_i\}, c_i \in \mathbb{C}, c_i \neq c_{i' \neq i}$. Using

$$p_i(z) = \prod_{i' \neq i} (z - c_i') z$$

we find a decomposition of $\mathbf{v}$ as a linear combination of *idempotents*, namely $\mathbf{v} = \sum_{i, c_i \neq 0} c_i \mathbf{e}_i(\mathbf{v})$ with $\mathbf{e}_i(\mathbf{v}) = \frac{1}{p_i(c_i)} p_i(\mathbf{v}) \in \mathcal{V}$ indicating all elements of X colored with $c_i \neq 0$ by $\mathbf{v}$. The intersection of all those idempotents for all elements in $\mathcal{V}$ gives a unique set of idempotents $\mathcal{E} = \{\mathbf{e}_i\}$, $|\mathcal{E}| \leq |X|$, which are primitive w.r.t. $\mathcal{V}$ and are a basis of $\mathcal{V}$ seen as a vector space. We call $\mathcal{V}$ a *non excluding* subalgebra if $\forall x \in X \ \exists \mathbf{v} \in \mathcal{V}$ s.t. $\mathbf{v}(x) \neq 0$. By the following lemma non excluding is equivalent to unital for subalgebras of $L(X)$.

---

[1] When we consider the algebra $L(X \times X)$ we will denote the multiplicative identity by $\mathbf{J}$ with $\mathbf{J}(x, y) = 1 \ \forall (x, y) \in X \times X$ for notational convenience.

Proposition 7. *Let $\mathcal{V}$ be a subalgebra of $L(X)$. Let $\mathcal{E}$ denote the set of all primitive idempotents w.r.t. $\mathcal{V}$. Let $\mathbf{j} \in L(X)$ be the constant function assigning 1 to every element of X. The following properties are equivalent.*

- *(i) $\mathcal{V}$ is non excluding.*

- *(ii) $\mathcal{V}$ contains $\mathbf{j}$.*

Proof. '(ii)$\Rightarrow$(i)' is obvious. '(i)$\Leftarrow$(ii)'. Since $\mathcal{V}$ is non excluding, there exists $\mathbf{v} \in \mathcal{V}$ s.t. $\mathbf{v}(x) \neq 0$. Therefore, $\forall x$ there exists a primitive idempotent $\mathbf{e} \in \mathcal{E}$ s.t. $\mathbf{e}(x) \neq 0$, hence $\sum_{\mathbf{e} \in \mathcal{E}} \mathbf{e} = \mathbf{j}$. $\qquad\square$

Let f be the *set of all multivariate functions $\mathbb{C} \times \ldots \times \mathbb{C} \to \mathbb{C}$*. By (simultaneous) pointwise application we extend f to the set of multivariate functions $L(X) \times \ldots \times L(X) \to L(X)$ on $L(X)$ denoted by $f_{L(X)}$. It is easy to see that a subalgebra of $L(X)$ is closed w.r.t. $f_{L(X)}$ if and only if it contains $\mathbf{j}$.

Definition 13. *Let $\mathrm{U} \subset L(X)$. $\mathrm{Clos}(\mathrm{U})$ denotes the closure of U w.r.t. linear combination and pointwise multiplication.*

$\mathrm{Clos}(\mathrm{U}, \mathbf{j})$ is closed w.r.t. $f_{L(X)}$ and it is the smallest unital subalgebra of $L(X)$ containing U. By interpreting the idempotent basis of a unital subalgebra as a partition of the set X, the lattice of unital subalgebras of $L(X)$ is isomorphic to the lattice of partitions of X [41, lemma 1.1, p. 210].

3.3.2 Generalized Adjacency Matrix

We now consider the vertex set $V(\Gamma)$ and the edge set $E(\Gamma)$ of a graph Γ. Without any additional partition they constitute the trivial unital subalgebras of $L(V)$ and $L(E)$, respectively. A more advanced approach considers the unital subalgebra of $L(V \times V)$ given by $\mathrm{Clos}(\mathbf{A}, \mathbf{J})$, where $\mathbf{A}$ denotes the adjacency matrix of Γ and $\mathbf{J}$ the multiplicative identity in $L(X \times X)$. The idempotents are given by $\mathbf{A}$ and $\mathbf{J} - \mathbf{A}$ The spectrum of that algebra, i.e. the spectra of the matrix realizations of its elements, is basically equivalent to the spectrum of the *generalized adjacency matrix*. This matrix algebra is given by $\mathrm{Clos}(\mathbf{A}, \mathbf{I}, \mathbf{J})$ [46] [30] [80]. The idempotent basis of the generalized adjacency matrix is given by $\mathbf{I}$, $\mathbf{A}(\Gamma)$ and $\mathbf{A}(\bar{\Gamma})$, where $\mathbf{I} \in L(X \times X)$ assigns 1 to the elements of the set $\{(x,y) \,|\, x, y \in X, x = y\}$ and $\bar{\Gamma}$ denotes the complement graph of Γ. The generalized adjacency matrix is closely related to the main spectrum and the main angles [65], which is due to the following result in [80, contained in theorem 3, p. 4/5,].

Proposition 8. *Let Γ_1 and Γ_2 be a pair of simple graphs with adjacency matrices $\mathbf{A}(\Gamma_1)$ and $\mathbf{A}(\Gamma_2)$, respectively. The following statements are equivalent.*

- *Γ_1 and Γ_2 are isospectral w.r.t. $\mathbf{A}$, have and the same main angles.*

- *Γ_1 and Γ_2 are isospectral w.r.t. $\mathbf{A}$ and $\mathbf{J} - \mathbf{A}$*

- *There exists an orthogonal $\mathbf{Q}$ s.t. $\mathbf{Q}^T \mathbf{A}(\Gamma_1) \mathbf{Q} = \mathbf{A}(\Gamma_2)$ and $\mathbf{Q}\mathbf{j} = \mathbf{j}$.*

3.3.3 Matrices as Endomorphisms

We can interpret matrices in $\mathbb{C}^{|X| \times |X|}$ as linear maps acting on $L(X)$. Let M_X be the set of all $\mathbb{C}$-endomorphisms acting on $L(X)$ seen as a vector space. By interpreting the composition of endomorphisms as a product, M_X is a $\mathbb{C}$-algebra equivalent to the usual matrix algebra of square matrices indexed by $X \times X$, which is the vector space $\mathbb{C}^{|X| \times |X|}$ additionally equipped with the usual matrix multiplication. A subalgebra containing $\mathbf{I}$, the multiplicative identity on M_X, is called unital. We equip $L(X)$ with the usual inner product given by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{x \in X} \overline{\mathbf{v}(x)} \mathbf{w}(x) \text{ with } \mathbf{v}, \mathbf{w} \in L(X). \tag{3.16}$$

The complex conjugated transpose of $\mathbf{M} \in M_X$, denoted by $\mathbf{M}'$, satisfies

$$\langle \mathbf{v}, \mathbf{M}\mathbf{w} \rangle = \langle \mathbf{M}'\mathbf{v}, \mathbf{w} \rangle \; \forall \, \mathbf{v}, \mathbf{w} \in L(X). \tag{3.17}$$

The matrix $\mathbf{M}'$ defined by (3.17) is also called the adjoint of $\mathbf{M}$.

Definition 14. *Let $\mathcal{U} \subset M_X$. $\mathrm{WL}_0(\mathcal{U})$ is the smallest (by inclusion) subalgebra of M_X containing $\mathcal{U}$ that is closed w.r.t.*

- *the complex conjugated transpose.*

$L(X)$ can be embedded in M_X as the subalgebra $D_X = D(M_X)$ that is given by all diagonal matrices. By the bijective map

$$\mathrm{diag} : L(X) \to D_X \subset M_X, \quad \mathbf{v} \mapsto \mathbf{v}(x)\delta_{xy}, \quad x, y \in X \tag{3.18}$$

with inverse

$$\mathrm{diag}^{-1} : D_X \to L(X), \quad \mathbf{D} \mapsto \mathbf{Dj} \tag{3.19}$$

D_X is isomorphic to $L(X)$. For convenience we define the following operation:

Definition 15. $\deg : M_X \to D_X, \quad \mathbf{M} \mapsto \delta_{xy}\left(\sum_{z \in X} \mathbf{M}_{xz}\right), \quad x, y, z \in X.$

Since we can rewrite $\deg(\mathbf{M}) = \mathrm{diag}(\mathbf{Mj})$, $\deg$ is the identity on D_X. Note that $\deg(\mathbf{M}\,\mathrm{diag}(\mathbf{v})) = \mathrm{diag}(\mathbf{Mv})$. A subalgebra $\mathcal{M} \subset M_X$ is called closed w.r.t. $f_{L(X)}$ iff its subalgebra of diagonal matrices seen as a subalgebra of $L(X)$ is closed w.r.t. $f_{L(X)}$. Therefore, $\mathcal{M}$ is closed w.r.t. $f_{L(X)}$ if and only if it is unital, i.e. contains $\mathbf{I}$.

Concluding this section we introduce the closure operators WL_1 and WL_2.

Definition 16. *Let $\mathcal{U} \subset M_X$. $\mathrm{WL}_1(\mathcal{U})$ is the smallest subalgebra of M_X containing $\mathcal{U}$ that is closed w.r.t.*

- *the complex conjugated transpose*

- *the function* $\deg$.

Definition 17. *Let $\mathcal{U} \subset M_X$. $\mathrm{WL}_2(\mathcal{U})$ is the smallest subalgebra of M_X containing $\mathcal{U}$ that is closed w.r.t.*

- *the complex conjugated transpose*

- *the Hadamard (i.e. entry-wise) product.*

$\mathrm{WL}_2(\mathcal{U})$ is a subalgebra of M_X and a subalgebra of $L(X \times X)$. By the identity $\deg(\mathbf{M}) = \mathbf{I} \circ (\mathbf{M}\,\mathbf{J})$, $\mathrm{WL}_1(\mathbf{I}, \mathcal{U}.\mathbf{J})$ is a subalgebra of $\mathrm{WL}_2(\mathbf{I}, \mathcal{U}, \mathbf{J})$. Note that $\mathrm{WL}_2(\mathbf{I}, \mathcal{U}, \mathbf{J})$ is a coherent algebra.

3.4 Properties and Constructions

3.4.1 Simultaneous Unitary Equivalence

We aim at constructing pairs of (possibly directed) graphs that are simultaneous unitary equivalent w.r.t. families of matrices which are generated by a subset of $\{\mathbf{I}, \mathbf{A}, \mathbf{J}\}$, i.e. the identity w.r.t. the usual matrix multiplication, the adjacency matrix and the all-ones matrix, by use of the WL_0 or WL_1 closure. By the following lemma, which crucially depends on Spechts theorem [70], isospectrality and unitary equivalence are equivalent notions for those matrix families.

Lemma 3. *Let $\mathcal{U}$ be a subset of $\{\mathbf{I}, \mathbf{A}, \mathbf{J}\}$ where $\mathbf{A}$ is an adjacency matrix. Let $k \in \{0, 1, 2\}$. Two graphs are isospectral w.r.t. $\mathrm{WL}_k(\mathcal{U})$ if and only if they are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_k(\mathcal{U})$.*

Proof. '$\Rightarrow$'. Simultaneous isospectrality implies simultaneous equality of traces. Since all three closure operators respect the complex conjugated transpose and the usual matrix multiplication simultaneous unitary equivalence is implied by Spechts theorem. '$\Leftarrow$'. Trivial since unitary transformations preserve the spectrum. $\square$

We will also exploit the following four lemmas for simultaneous unitary equivalent matrix sets.

Lemma 4. *Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be matrix sets which are simultaneous unitary equivalent by the unitary transformation $\mathbf{Q}$ and let $\mathbf{J} \in \mathcal{U}_1$ and $\mathbf{J} \in \mathcal{U}_2$ be corresponding matrices, i.e.*

$$\mathbf{Q}'\mathcal{U}_1\mathbf{Q} = \mathcal{U}_2 \text{ and } \mathbf{Q}'\mathbf{J}\mathbf{Q} = \mathbf{J},$$

where $\mathbf{J}$ is the all-ones matrix. Then $\mathbf{Q}$ can be chosen s.t.

$$\mathbf{Q}\mathbf{j} = \mathbf{Q}'\mathbf{j} = \mathbf{j}.$$

Proof. Since $\mathbf{Q}'\mathbf{J}\mathbf{Q} = \mathbf{J}$ we have $\mathbf{J}\mathbf{Q} = \mathbf{Q}\mathbf{J}$. Multiplication from the left to $\mathbf{j}$ yields $(\mathbf{j}'\mathbf{Q}\mathbf{j})\,\mathbf{j} = (\mathbf{j}'\mathbf{j})\,\mathbf{Q}\mathbf{j}$. Therefore, $\mathbf{j}$ is an eigenvector of $\mathbf{Q}$ to the eigenvalue $\lambda = \frac{1}{N}\mathbf{j}'\mathbf{Q}\mathbf{j}$. In a similar way starting from $\mathbf{J}\mathbf{Q}' = \mathbf{Q}'\mathbf{J}$ one shows that $\mathbf{j}$ is an eigenvector of $\mathbf{Q}'$ to the eigenvalue $\bar{\lambda}$. By unitarity of $\mathbf{Q}$ we have $|\lambda| = 1$. Since $\mathbf{Q}'\mathcal{U}_1\mathbf{Q} = \mathcal{U}_2$ and $|\alpha| = 1$, $\alpha \in \mathbb{C}$, imply $(\alpha'\mathbf{Q}')\mathcal{U}_1\,(\alpha\mathbf{Q}) = \mathcal{U}_2$, we may choose $\mathbf{Q}$ s.t. $\lambda = 1$. $\qquad\square$

The next lemma is a kind of converse for the previous one.

Lemma 5. *Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be sets of matrices which are simultaneous unitary equivalent by the unitary transformation $\mathbf{Q}$ with $\mathbf{Q}\mathbf{j} = \mathbf{j}$. Then $\mathrm{WL}_0\,(\mathcal{U}_1, \mathbf{I}, \mathbf{J})$ and $\mathrm{WL}_0\,(\mathcal{U}_2, \mathbf{I}, \mathbf{J})$ are simultaneous unitary equivalent.*

Proof. It suffices to show that $\mathbf{Q}'\mathbf{I}\mathbf{Q} = \mathbf{I}$ and $\mathbf{Q}'\mathbf{J}\mathbf{Q} = \mathbf{J}$. The former is trivial and the latter follows from $\mathbf{j} = \mathbf{Q}\mathbf{j} = \mathbf{Q}'\mathbf{j}$. $\qquad\square$

Since the transformation guiding a Q-switching has always $\mathbf{j}$ as an eigenvector, we have by lemma 5 that graph pairs obtained by Q-switching are always simultaneous unitary equivalent w.r.t. $\mathrm{WL}_0\,(\mathbf{I}, \mathbf{A}, \mathbf{J})$. The next lemma shows that a certain block structure of simultaneous unitary equivalent matrix sets induces a block diagonal form for the unitary transformation.

Lemma 6. *Let $\mathbf{E}$ be an idempotent diagonal matrix, i.e. a matrix with all entries z e r o except possibly some o n e s on the diagonal. Let $\mathbf{P}$ be a permutation s.t. on the diagonal of $\mathbf{P}'\mathbf{E}\mathbf{P}$ from the upper left downwards there are all o n e s first and then all z e r o s. Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be simultaneous unitary equivalent by the unitary transformation $\mathbf{Q}$ and let $\mathbf{E} \in \mathcal{U}_1$ and $\mathbf{E} \in \mathcal{U}_2$ be corresponding matrices, i.e.*

$$\mathbf{Q}'\mathcal{U}_1\mathbf{Q} = \mathcal{U}_2 \ \text{and} \ \mathbf{Q}'\mathbf{E}\mathbf{Q} = \mathbf{E}.$$

Then $\mathbf{P}'\mathbf{Q}\mathbf{P}$ is an explicitly 2×2 block diagonal matrix with upper left block given by $\mathbf{P}'\mathbf{E}\mathbf{Q}\mathbf{E}\mathbf{P}$.

Proof. After applying a suitable simultaneous permutation at the rows and columns of the elements of $\mathcal{U}_1$ and $\mathcal{U}_2$ the matrix $\mathbf{E}$ has the desired form. Thus, w.l.o.g. we may assume $\mathbf{P} = \mathbf{I}$ for the rest of this proof. Let

$$\mathbf{E}_1 = \mathbf{E} \ , \ \mathbf{E}_2 = \mathbf{I} - \mathbf{E} \ \text{and} \ \mathbf{Q}_{ij} = \mathbf{E}_i\mathbf{Q}\mathbf{E}_j \ , \ 1 \leq i, j \leq 2.$$

We have

$$\mathbf{Q}'\mathbf{E}\mathbf{Q} = \mathbf{E} = \mathbf{Q}\mathbf{E}\mathbf{Q}'.$$

It follows from the left-hand side that $\mathbf{Q}_{12} = \mathbf{0}$ and $\mathbf{Q}'_{11}\mathbf{Q}_{11} = \mathbf{I}$ and from the right-hand side that $\mathbf{Q}_{21} = \mathbf{0}$ and $\mathbf{Q}'_{22}\mathbf{Q}_{22} = \mathbf{I}$. $\qquad\square$

There is a much less restrictive sufficient condition for the unitary transformation to be block diagonal (after suitable permutation) as can be seen from the following remark, which is given without proof.

Remark 1. *Let $\mathcal{U}_i$, $i = 1, 2$, and $\mathbf{Q}$ be as in lemma 6 and let $\mathbf{X}_i \in \mathcal{U}_i$, $i = 1, 2$, s.t. $\mathbf{X}_2 = \mathbf{Q}'\mathbf{X}_1\mathbf{Q}$. Let $\mathbf{X}_i$ be conformally partitioned 2×2 block matrices s.t. all blocks are zero except the upper right diagonal block which has full rank. Then $\mathbf{Q}$ is block diagonal conformable to the partition of $\mathbf{X}_{1/2}$.*

Lemma 7. *Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be sets of matrices which are simultaneous unitary equivalent by the unitary transformation $\mathbf{Q}$. Let $\mathbf{Q}$ be block diagonal. Let the block structure of $\mathbf{Q}$ be described by the partition Π with cells c_i. Let $\mathbf{E}_i$ be a binary diagonal matrix which has entry 1 at position vv if and only if $v \in c_i$. Then $\mathrm{WL}_0(\mathcal{U}_1, \mathbf{E}_i)$ and $\mathrm{WL}_0(\mathcal{U}_2, \mathbf{E}_i)$ are simultaneous unitary equivalent.*

Proof. It suffices to show that $\mathbf{Q}'\mathbf{E}_i\mathbf{Q} = \mathbf{E_i}$ which follows from the properties of $\mathbf{Q}$ and $\mathbf{E}_i$. $\qquad\square$

3.4.2 WL_0-**Closure**

We consider the subalgebra of $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$ which is generated by linear combinations of all nonnegative powers of $\mathbf{A}$. At least for undirected graphs it is known as the adjacency algebra [7, definition 2.4, p. 9] [6, section 6.2, p. 70]. Since $\mathbf{A}_{ij}^l$ gives the number of directed walks of length l from node i to node j [20, theorem 2.2.1, p. 24], we have [85, theorem 2.1, p. 3]

Proposition 9. *Let Γ be a digraph with at least two nodes and adjacency matrix $\mathbf{A}$. $\exists f(x) \in \mathbb{C}[x]$ s.t. $f(A)$ is positive if and only if Γ is strongly connected. $\mathrm{WL}_0(A)$ contains a positive matrix if and only if Γ is weakly connected.*

Proof. '$\Rightarrow$'. If Γ is not strongly connected, then $\mathbf{A}$ is reducible. Therefore, all nonnegative powers of $\mathbf{A}$ and also their linear combinations are also reducible, hence there is at least one position which is always zero. If Γ is not weakly connected, then $\mathbf{A}$ is proper block diagonal and hence all elements of $\mathrm{WL}_0(\mathbf{A})$ are proper block diagonal.'$\Leftarrow$'. If Γ is strongly connected, then by the Perron-Frobenius theorem $\mathbf{A}$ has a unique (up to scaling) pair of a positive left-eigenvector $\mathbf{v}$ and a positive right-eigenvector $\mathbf{w}$ corresponding to the same simple positive eigenvalue. Since the eigenvalue is simple, there is a polynomial f s.t. $f(\mathbf{A}) = \mathbf{w}\mathbf{v}'$. The Perron-Frobenius theorem can be employed similarly, when Γ is weakly connected since then $(\mathbf{A} + \mathbf{A}')$ is irreducible. $\qquad\square$

The next result is due to Hoffman [42].

Proposition 10. *Let Γ be a digraph of at least two nodes with adjacency matrix $\mathbf{A}$. $\exists f(x) \in \mathbb{C}[x]$ s.t. $f(A) = \mathbf{J}$ if and only if Γ is strongly connected and $\mathbf{A}$ is row and column regular.*

Proof. Let N be the number of nodes of Γ. '$\Rightarrow$'. By proposition 9, Γ is connected. Since $\mathbf{J}$ is a polynomial in $\mathbf{A}$, both matrices commute. Since $\mathbf{J}\mathbf{A}\mathbf{j} = (\mathbf{j}'\mathbf{A}\mathbf{j})\mathbf{j}$, $\mathbf{A}\mathbf{J}\mathbf{j} = N\mathbf{A}\mathbf{j}$ is a constant vector, hence $\mathbf{A}$ is row regular.

The column regularity is proved similarly. '⇐'. By regularity $\mathbf{j}'$ and $\mathbf{j}$ are left-eigenvector and right-eigenvector, respectively, to the same eigenvalue $\frac{1}{N}\mathbf{j}'\mathbf{A}\mathbf{j}$. By Perron-Frobenius this eigenvalue is simple since all entries of $\mathbf{j}$ are positive. Therefore, there is a polynomial $f(x) \in \mathbb{C}[x]$ s.t. $f(\mathbf{A}) = \mathbf{j}\mathbf{j}' = \mathbf{J}$. □

By proposition 9, $\mathbf{J} \in \mathrm{WL}_0(\mathbf{A}(\Gamma))$ implies that Γ is weakly connected, but it does not imply regularity. A counterexample is the directed graph with two nodes connected by a single arc with adjacency matrix $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, which generates the whole 2×2 complex matrix algebra. However, weakly connectedness does not imply $\mathbf{J} \in \mathrm{WL}_0(\mathbf{A}(\Gamma))$. Counterexamples are all non regular connected undirected graphs according to proposition 10.
Unitary equivalence w.r.t. $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$ implies unitary equivalence w.r.t. $\mathrm{WL}_0(\mathbf{I}, \mathbf{A})$, since the latter is a subalgebra of the former. The converse holds for regular connected undirected graphs by proposition 10 but is not true in general.

Construction   Using lemmata 2, 4 and 5 we can build a pair of graphs which are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$ when another graph pair with that property is provided. We give an example.
Let Γ_i, $i = 1, 2$, be graphs of size N which are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$ by the transformation $\mathbf{Q}$, which does not have to be known explicitly. According to lemma 2 consider the 9 sets $\mathcal{A}_{ij}$, $1 \le i, j \le 3$, with

$$\mathcal{A}_{ij} = \left\{ (\mathbf{F}(\Gamma_1), \mathbf{F}(\Gamma_2)) \,|\, \mathbf{F} \in \mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J}) \text{ and } \mathbf{F}(\Gamma_{1/2}) \text{ binary} \right\}, \forall 1 \le i, j \le 2,$$

$$\mathcal{A}_{13} = \mathcal{A}_{23} = \{(\mathbf{0}_{Nn}, \mathbf{0}_{Nn}), (\mathbf{J}_{Nn}, \mathbf{J}_{Nn})\}, \mathcal{A}_{31} = \mathcal{A}_{32} = \{(\mathbf{0}_{nN}, \mathbf{0}_{nN}), (\mathbf{J}_{nN}, \mathbf{J}_{nN})\}$$

and arbitrary set $\mathcal{A}_{33}$ of pairs of identical $n \times n$ binary matrices. The graph pair with adjacency matrices $(\mathbf{A}_+, \tilde{\mathbf{A}}_+)$, which are obtained by choosing an element of $\mathcal{A}_{ij}$ for all $1 \le i, j \le 3$ according to lemma 2, is simultaneous unitary equivalent w.r.t. $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$. By lemma 4 the unitary matrix

$$\mathbf{Q}_+ = \left(\begin{array}{ccc} \mathbf{Q} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_n \end{array} \right)$$

satisfies $\mathbf{Q}_+\mathbf{j} = \mathbf{j}$.
The sets $\mathcal{A}_{ij}$, $1 \le i, j \le 2$ contain e.g. $(\mathbf{J} - \mathbf{A}(\Gamma_1) - \mathbf{I}, \mathbf{J} - \mathbf{A}(\Gamma_2) - \mathbf{I})$ and $(\mathbf{A}(\Gamma_1), \mathbf{A}(\Gamma_2))$. The sets $\mathcal{A}_{13}$ and $\mathcal{A}_{23}$ ($\mathcal{A}_{31}$ and $\mathcal{A}_{32}$) can be extended by all pairs of binary matrices of size $N \times n$ ($n \times N$) which are related via left (right) multiplication of $\mathbf{Q}$.
An extension to the utilization of families of graph pairs that are isospectral w.r.t. $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$ is straightforward.

3.4.3 WL_1-**Closure**

We start with a basic proposition.

Proposition 11. *Let Γ be a digraph with at least two nodes and with adjacency matrix $\mathbf{A}$. $\mathrm{WL}_1(\mathbf{A})$ contains $\mathbf{I}$ if and only if Γ has no isolated nodes.*

Proof. '$\Rightarrow$'. If G contains isolated vertices, then Γ is disconnected and $\mathbf{A}$ is block diagonal. Since $\mathrm{WL}_1(\mathbf{A}(K_1)) \equiv \mathbf{0}$, $\mathbf{I} \notin \mathrm{WL}_1(\mathbf{A})$.'$\Leftarrow$'. If Γ contains no isolated vertices, then $\deg(\mathbf{A} + \mathbf{A}')$ has no zero diagonal element. Therefore its determinant is not zero, and by use of its characteristic polynomial and the Cayley-Hamilton theorem $\mathbf{I}$ is contained in the span of positive powers of $\deg(\mathbf{A} + \mathbf{A}')$. □

The following lemma shows the strong relation between the WL1 algebra and equitable partitions.

Lemma 8. *The idempotent orthogonal basis for the subalgebra of diagonal matrices of $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$ ($\mathrm{WL}_1(\mathbf{I}, \mathbf{A})$) indicates the coarsest row and column equitable partition of a digraph Γ with adjacency matrix $\mathbf{A}$.*

Proof. We consider $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$. The case $\mathrm{WL}_1(\mathbf{I}, \mathbf{A})$ is proven similar. Let the subalgebra of diagonal matrices be denoted with $\mathrm{D}(\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J}))$. Let the idempotent orthogonal basis be denoted with

$$\mathcal{E} = \{\mathbf{E}_i\} \subset \mathrm{D}(\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})) \subset \mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J}),$$

which has the property $\sum_i \mathbf{E}_i = \mathbf{I}$. If $\mathcal{E}$ does not indicate a row equitable partition of an arbitrary matrix $\mathbf{M} \in \mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$, then there is a pair ij s.t. $\deg(\mathbf{E}_i \mathbf{M} \mathbf{E}_j) \in \mathrm{D}(\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J}))$, which is orthogonal to $\mathbf{E}_{i' \neq i}$, is not a multiple of $\mathbf{E}_i$. That contradicts $\mathcal{E}$ being a basis of the diagonal subalgebra. Therefore, the idempotent basis $\mathcal{E}$ indicates a row equitable equitable partition. The same argument applies for column equitability.
Now, let $\mathcal{E}$ indicate a row and column equitable partition of $\mathbf{A}$, which is equivalent to $\deg(\mathbf{E}_i \mathbf{A} \mathbf{E}_j)$ being a multiple of $\mathbf{E}_i$ and $\deg(\mathbf{E}_j \mathbf{A}' \mathbf{E}_i)$ being a multiple of $\mathbf{E}_j \; \forall i, j$. It is easy to see that $\mathcal{E}$ indicates a row and column equitable partition for $\mathbf{J}$ and for any matrix in the span of $\mathcal{E}$. By induction the same partition is row and column equitable for all $\mathbf{M} \in \mathrm{WL}_1(\mathcal{E}, \mathbf{A}, \mathbf{J})$ since this is true for the generating set. Therefore, equitable partitions $\mathcal{E}$ are stable in the following sense. If and only if $\mathcal{E}$ indicates an equitable partition, then it is an idempotent basis for the diagonal subalgebra of $\mathrm{WL}_1(\mathcal{E}, \mathbf{A}, \mathbf{J})$.
Finally, let $\mathcal{E}$ be the coarsest row and column equitable refinement of $\mathbf{I}$. Then $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J}) = \mathrm{WL}_1(\mathcal{E}, \mathbf{A}, \mathbf{J})$ as an equivalence of sets. □

We now give a version of proposition 10 for the WL_1-closure.

Proposition 12. *Let Γ be a digraph with adjacency matrix $\mathbf{A}$. $\mathrm{WL}_1(\mathbf{A})$ contains $\mathbf{J}$ if and only if Γ is strongly connected.*

Proof. '⇒'. If Γ is not strongly connected, then $\mathbf{A}$ is proper block diagonal, and therefore all matrices in $\mathrm{WL}_1(\mathbf{A})$ are proper block diagonal, which contradicts $\mathbf{J} \in \mathrm{WL}_1(\mathbf{A})$. '⇐'. By proposition 11, $\mathbf{I} \in \mathrm{WL}_1(\mathbf{A})$. Let $\mathcal{E} = \{\mathbf{E}_i\}$ be a set if idempotent diagonal matrices indicating the coarsest equitable partition of $\mathbf{A}$. By lemma 8 we have $\mathcal{E} \subset \mathrm{WL}_1(\mathbf{A})$. Let $\mathbf{M} = \sum_{n=1}^{N} \mathbf{A}^n$. $\mathcal{E}$ also indicates an equitable partition of $\mathbf{M}$. Let

$$\mathbf{M}_{ij} = \mathbf{E}_i \mathbf{M} \mathbf{E}_j \text{ and } \mathbf{M}_{ij}\mathbf{j} = m_{ij}\mathbf{E}_i\mathbf{j}.$$

Since Γ is strongly connected, $\mathbf{M}$ is positive, hence $m_{ij} > 0 \ \forall i, j$. Therefore, m_{ii} is a simple eigenvalue of $\mathbf{M}_{ii}$ with unique eigenvector $\mathbf{E}_i\mathbf{j}$ by the theorem of Perron-Frobenius[2], hence $\mathrm{WL}_1(\mathbf{A})$ contains $\mathbf{E}_i\mathbf{J}\mathbf{E}_i$. And since

$$\frac{1}{m_{ij}}\mathbf{E}_i\mathbf{J}\mathbf{E}_i\mathbf{M}_{ij}\mathbf{E}_j\mathbf{J}\mathbf{E}_j = \mathbf{E}_i\mathbf{J}\mathbf{E}_j \neq \mathbf{0},$$

we have $\mathbf{J} = \sum_{ij} \mathbf{E}_i\mathbf{J}\mathbf{E}_j \in \mathrm{WL}_1(\mathbf{A})$. $\qquad\square$

Lemma 9. *Let Γ_1 and Γ_2 be a pair of graphs with coarsest column and row equitable partition Π_1 and Π_2, respectively. Then Γ_1 and Γ_2 are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$ if and only if there is a suitable indexing of Γ_1 and Γ_2 s.t. $\Pi_1 = \Pi_2 = \Pi$ with equal front divisors and there is a unitary matrix $\mathbf{Q}$ with $\mathbf{Q}\mathbf{j} = \mathbf{j}$ s.t.*

$$\mathbf{A}(\Gamma_2) = \mathbf{Q}'\mathbf{A}(\Gamma_1)\mathbf{Q}$$

which is block diagonal when partitioned conformable with Π.

Proof. '⇒'. Let $\mathcal{E} = \{\mathbf{E}_i\}$ be the idempotent basis for the diagonal subalgebra of $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$ and let $\mathbf{Q}$ be the simultaneous unitary transformation. According to lemma 8, $\mathcal{E}(\Gamma_1)$ and $\mathcal{E}(\Gamma_2)$ can be obtained by applying Π_1 and Π_2 to $\mathbf{I}$. Since $\mathcal{E}$ is an idempotent basis and since $\deg(\mathbf{E}) - \mathbf{E} = \mathbf{0}$ for each element $\mathbf{E} \in \mathcal{E}(\Gamma_1)$, we have $\mathbf{Q}'\mathcal{E}(\Gamma_1)\mathbf{Q} \subset \mathcal{E}(\Gamma_2)$ and by exchange of Γ_1 and Γ_2, $\mathbf{Q}'\mathcal{E}(\Gamma_1)\mathbf{Q} = \mathcal{E}(\Gamma_2)$. Therefore, there is an indexing s.t. $\mathbf{E}(\Gamma_1) = \mathbf{E}(\Gamma_2) \ \forall \mathbf{E} \in \mathcal{E}$. By lemma 6 $\mathbf{Q}$ is block diagonal w.r.t. the partition Π indicated by the set $\mathcal{E}$. Since Π is row and column equitable, corresponding matrix subblocks of $\mathbf{A}(\Gamma_1)$ and $\mathbf{A}(\Gamma_2)$ have constant row and column sum, which must be equal by isospectrality. Therefore, $\mathbf{Q}$ can be chosen s.t. $\mathbf{Q}\mathbf{j} = \mathbf{j}$. '⇐'. Since $\mathbf{Q}\mathbf{j} = \mathbf{j}$ we have $\mathbf{Q}'\mathbf{J}\mathbf{Q} = \mathbf{J}$. Let $\mathcal{E} = \{\mathbf{E}_i\}$ be a set of idempotent diagonal matrices indicating Π and let Γ_1 and Γ_2 be indexed w.r.t. Π. Since Π is equitable, $\deg(\mathbf{X})$ is a linear combination of $\mathcal{E}$ for each element $\mathbf{X} \in \mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$ by lemma 8. Therefore and since $\mathbf{Q}\mathbf{j} = \mathbf{j}$,

$$\deg(\mathbf{Q}'\mathbf{X}\mathbf{Q}) = \deg(\mathbf{X}) \ \forall \mathbf{X} \in \mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J}).$$

Thus, Γ_1 and Γ_2 are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_1(\mathbf{I}, \mathbf{A}, \mathbf{J})$ by the transformation $\mathbf{Q}$. $\qquad\square$

---

[2]Although $\mathbf{M}_{ii}$ is reducible when $\mathbf{E}_i \neq \mathbf{I}$, but it contains only one non vanishing irreducible diagonal block, which is all positive.

Construction   Analogous to the example at the end of the last subsection we can build a pair of graphs which are simultaneous unitary equivalent w.r.t. $\mathrm{WL}_1\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$ when another graph pair with that property is given. Using lemma 9 we can also exploit the coarsest equitable partition of such a pair to obtain smaller graphs by the removal of cells.

Further Properties   The algebra $\mathrm{WL}_1\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$ is sufficient to distinguish all trees by spectrum, i.e. there are no trees isospectral w.r.t. $\mathrm{WL}_1\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$ [59], and it contains several well known graph matrices. For instance,

- the adjacency matrix for the complement: $\mathbf{J} - \mathbf{A} - \mathbf{I}$

- the Seidelmatrix: $\mathbf{J} - 2\mathbf{A} - \mathbf{I}$

- the combinatorial(-) and signless(+) Laplacian: $\deg\left(\mathbf{A}\right) \pm \mathbf{A}$

- the normalized Laplacian: $\mathbf{I} - \left(\deg\left(\mathbf{A}\right)\right)^{-1}\mathbf{A}$

The latter one makes use of $f_{L(X)}$. Thus, lemma 2 and lemma 9 provide constructions for graph pairs which are unitary equivalent (isospectral) w.r.t. those graph matrices. However, the existence of at least one graph pair which is non isomorphic but simultaneous unitary equivalent w.r.t. $\mathrm{WL}_1\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$ is required to prevent triviality. This is done by the following proposition.

Proposition 13. *Any pair of isospectral regular simple graphs is* $\mathrm{WL}_1\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$-*isospectral.*

Proof. Since the adjacency matrices of simple graphs are self adjoint (symmetric), they can be diagonalized by unitary transformations. Therefore, they are unitary equivalent. By regularity both graphs have the same eigenvalue r to the shared eigenvector $\mathbf{j}$. From that it follows that the unitary transformation between both matrices must commute with $\mathbf{J}$. □

Examples of isospectral regular simple graph pairs are well known, e.g. [20].

3.4.4 WL_2-**Closure**

For the extension to the coherent algebra $\mathrm{WL}_2\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$ one has to consider that this algebra seen as a vector space has a basis which is orthogonal and idempotent w.r.t. the point-wise product and contains diagonal matrices which sum up to $\mathbf{I}$. This subset indicates a partition which is obeyed by all basis elements [41]. Therefore the simultaneous unitary transformation between two coherent algebras must be a bijective map between the two bases and also between the diagonal subsets of the bases. It is therefore block diagonal and has (as a matrix) a constant eigenvector. Examples of how this properties can be used to construct graphs simultaneous unitary equivalent w.r.t. $\mathrm{WL}_2\left(\mathbf{I}, \mathbf{A}, \mathbf{J}\right)$ can be derived from the *direct sum* and the *wreath product* in [24, G 2.1, p. 42 and G 4.1, p. 45].

3.4.5 Combined Graph Matrices

We conclude this chapter with the following theorem

Theorem 8. *Let $\mathbf{A}(\Gamma)$ denote the adjacency matrix of a digraph Γ. Let $\mathcal{E}(\Gamma)$ denote a set of binary diagonal matrices which sum up to $\mathbf{I}$ and indicate the coarsest equitable partition of Γ, and let $\mathcal{B}(\Gamma)$ denote the idempotent orthogonal (w.r.t. point-wise product) basis of its coherent closure $\mathrm{WL}_2(\mathbf{I}, \mathbf{A}, \mathbf{J})$. Let Γ_1 and Γ_2 be two digraphs. The simultaneous unitary equivalence of Γ_1 and Γ_2 w.r.t. $\mathrm{WL}_k(\mathbf{I}, \mathbf{A}, \mathbf{J})$ is equivalent to the simultaneous equivalence of the set*

- *$\{\mathbf{A}, \mathbf{J}\}$ for $k = 0$,*

- *$\{\mathcal{E}, \mathbf{A}, \mathbf{J}\}$ for $k = 1$,*

- *$\mathcal{B}$ for $k = 2$.*

Proof. '$k = 0$'. The algebra can be generated by $\{\mathbf{I}, \mathbf{A}, \mathbf{J}, \mathbf{A}'\}$ using only matrix product and linear combination. Unitary equivalence w.r.t. $\mathbf{A}$ is equivalent to unitary equivalence w.r.t. $\mathbf{A}'$ and that w.r.t. $\mathbf{I}$ is trivial. '$k = 1$'. According to (8), the algebra can be generated by multivariate polynomials in $\{\mathcal{E}, \mathbf{A}, \mathbf{A}', \mathbf{J}\}$ and unitary equivalence w.r.t. $\mathbf{A}$ implies that of $\mathbf{A}'$. '$k = 2$'. Obvious, since $\mathcal{B}$ is a basis. $\qquad\square$

Using a result from [29] those sets may be combined into a single matrix. The unitary equivalence w.r.t. this matrix is equivalent to the unitary equivalence of the corresponding algebra. For instance, in the case of $\mathrm{WL}_0(\mathbf{I}, \mathbf{A}, \mathbf{J})$ we may consider the unified graph matrix

$$\begin{pmatrix} \mathbf{0} & \mathbf{I} & \mathbf{J} & \mathbf{A} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Bibliography

[1] A. Abiad and W. H. Haemers. Cospectral graphs and regular orthogonal matrices of level 2. *Journal of Combinatorial Theory, Series A*, 19, 2012.

[2] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13:457–466, 1998.

[3] E. N. Antoniou and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electronic Journal of Linear Algebra*, 11:78–87, 2004.

[4] A. Banerjee and J. Jost. Spectral plot properties: towards a qualitative classification of networks. *Networks and Heterogeneous Media*, 3(2):395–411, 2008.

[5] A. Banerjee and J. Jost. Graph spectra as a systematic tool in computational biology. *Discrete Applied Mathematics*, 157(10):2425–2431, 2009.

[6] R. B. Bapat. *Graphs and Matrices*. Springer, 2011.

[7] N. Biggs. *Algebraic Graph Theory*. Cambridge University Press, 1993.

[8] C. Bischof and C. Van Loan. The wy representation for products of householder matrices. *SIAM Journal on Scientific and Statistical Computing*, 8(1):2–13, 1987.

[9] P. Boldi and S. Vigna. Fibrations of graphs. *Discrete Math*, 243, 2002.

[10] S. P. Borgatti and M. G. Everett. Regular equivalence: general theory. *Journal of Mathematical Sociology*, 19:29–52, 1994.

[11] U. Brandes and J. Lerner. Structural similarity: spectral methods for relaxed blockmodeling. *J. Classif.*, 27(3):279–306, 2010.

[12] L. A. Bunimovich and B. Z. Webb. Isospectral reductions. *preprint*, 2010. http://arxiv.org/abs/0909.0053.

[13] L. A. Bunimovich and B. Z. Webb. Isospectral graph reductions and improved estimates of matrices' spectra. *preprint*, 2011. http://arxiv.org/abs/0911.2453.

[14] L. A. Bunimovich and B. Z. Webb. Isospectral graph transformations, spectral equivalence, and global stability of dynamical networks. *Nonlinearity*, 25, 2012.

[15] S. Butler and J. Grout. A construction of cospectral graphs for the normalized Laplacian. *Electronic Journal of Combinatorics*, 18(1), 2011.

[16] J. Cai, M. Führer, and N. Immerman. An optimal lower bound on the number of variables for graphs. *Combinatorica*, 12(4):389–410, 1992.

[17] D. M. Cardoso and P. Rama. Spectral results on graphs with regularity constraints. *Linear Algebra and its Applications*, 423(1):90–98, 2006.

[18] D. Carlson. What are Schur complements, anyway? *Linear Algebra and its Applications*, 74:257–275, 1986.

[19] A. Coja-Oghlan. Finding large independent sets in polynomial expected time. *Combinatorics, Probability & Computing*, 15(5):731–751, 2006.

[20] D. Cvetković, P. Rowlinson, and S. Simić. *Eigenspaces of graphs*. Cambridge University Press, 1997.

[21] D. M. Cvetković, M. Doob, and H. Sachs. *Spectra of graphs: theory and application*. Pure and applied mathematics. Academic Press, 1980.

[22] P. T. Darga, K. A. Sakallah, and I. L. Markov. Faster symmetry discovery using sparsity of symmetries. In *Proceedings of the 45th annual Design Automation Conference*, DAC'08, pages 149–154. ACM, 2008.

[23] J. J. Dongarra, S. J. Hammarling, and D. C. Sorenson. Block reduction of matrices to condensed forms for eigenvalue computations. *Journal of Computational and Applied Mathematics*, 27:215–227, 1987.

[24] B. Weisfeiler (ed.). *On construction and identification of graphs*. Lecture notes in mathematics, 558. Springer, 1976.

[25] S. C. Eisenstat and I. C. F. Ipsen. Three absolute perturbation bounds for matrix eigenvalues imply relative bounds. *SIAM Journal on Matrix Analysis and Applications*, 20:149–158, 1998.

[26] L. Elsner. An optimal bound for the spectral variation of two matrices. *Linear Algebra and its Applications*, 71:77–80, 1985.

[27] K. Fan and G. Pall. Imbedding conditions for Hermitian and normal matrices. *Canadian Journal of Mathematics*, 9:298–304, 1957.

[28] F. R. Gantmacher. *Theory of Matrices*. Chelsea, 1960.

[29] T. G. Gerasimova, R. A. Horn, and V. V. Sergeichuk. Simultaneous unitary equivalences. *Linear Algebra and its Applications*, 2011. available online.

[30] C. D. Godsil and B. D. McKay. Spectral conditons for the reconstructibility of a graph. *Journal of Combinatorial Theory, Series B*, 30:285–289, 1981.

[31] C. D. Godsil and B. D. McKay. Constructing cospectral graphs. *Aequationes Mathematicae*, 28:257–268, 1982.

[32] C. D. Godsil and G. Royle. *Algebraic graph theory*. Graduate Texts in Mathematics. Springer, 2001.

[33] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, 1982.

[34] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins University Press, 3 edition, 1996.

[35] W. H. Haemers. Interlacing eigenvalues and graphs. *Linera Algebra and its Applications*, 227/228:593–616, 1995.

[36] W. H. Haemers and E. Spence. Enumeration of cospectral graphs. *European Journal of Combinatorics*, 25:199–211, 2004.

[37] E. V. Haynsworth. Quasi-stochastic matrices. *Duke Mathematical Journal*, 22(1), 1955.

[38] E. V. Haynsworth. Applications of a theorem on partitioned matrices. *J. Res. Bur. Stand.*, 63B(2), 1959.

[39] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur. Symmetric linearizations for matrix polynomials. *SIAM Journal of Matrix Analysis and Applications*, 29(1):143–159, 2006.

[40] N. J. Higham, D. S. Mackey, and F. Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM Journal on Matrix Analysis and Applications*, 28(4):1005–1028, 2006.

[41] D. G. Higman. Coherent algebras. *Linear Algebra and its Applications*, 93:209–239, 1987.

[42] A. J. Hoffman and M. H. McAndrew. The polynomial of a directed graph. *Proceedings of the American Mathematical Society*, 16(2):303–309, 1965.

[43] A. S. Householder. Unitary triangularization of a nonsymmetric matrix. *Journal of the ACM*, 5(4):339–342, 1958.

[44] A. S. Householder. *The theory of matrices in numerical analysis*. Dover books on advanced mathematics. Dover Publications, 1975.

[45] S. Hwang. Cauchy's interlace theorem for eigenvalues of Hermitian matrices. *American Mathematical Monthly*, 111(2):157–159, 2004.

[46] C. R. Johnson and M. Newman. A note on cospectral graphs. *Journal of Combinatorial Theory, Series B*, 28:96–103, 1980.

[47] J. E. Dennis Jr., J. F. Traub, and R. P. Weber. The algebraic theory of matrix polynomials. *SIAM Journal on Numerical Analysis*, 13(6):831–845, Dec. 1976.

[48] T. Juntilla and P. Kaski. Engineering an efficient canonical labeling tool for large sparse graphs. In *Proceedings of the Ninth Workshop on Algorithm Engineering and Experiments and the Fourth Workshop on Analytic Algorithms and Combinatorics*, pages 135–149. SIAM, 2007.

[49] K. Kate and B. Ravindran. Epsilon equitable partition; a positional analysis method. In *Proceedings of the 15th International Conference on Management of Data, 2009, Mysore, India*. Computer Society of India, 2009.

[50] D. L. Kreher and D. R. Stinson. *Combinatorial algorithms: generation, enumeration, and search*. CRC Press, 1999.

[51] W. Kuich. Quasi-block-stochastic matrices. *Pacific Journal of Mathematics*, 27(2):353–364, 1968.

[52] P. Lancaster. Linearization of regular matrix polynomials. *Electronic Journal of Linear Algebra*, 17:21–27, 2008.

[53] P. Lancaster and M. Tismenetsky. *The Theory of Matrices: With Applications*. Academic Press, 2 edition, 1985.

[54] R. B. Lehoucq. The computation of elementary unitary matrices. *ACM Transaction on Mathematical Software*, 22(4):393–400, 1996.

[55] C. Li and R. Mathias. Interlacing inequalities for totally nonnegative matrices. *Linear Algebra and its Applications*, 341:35–44, 2002.

[56] S. Liu and G. Trenkler. Hadamard, Khatri-Rao, Kronecker and other matrix products. *International Journal of Information and Systems Sciences*, 4(1):160–177, 2008.

[57] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM Journal of Matrix Analysis and Applications*, 28(4):971–1004, 2006.

[58] M.A.Fiol. Eigenvalue interlacing and weight parameters of graphs. *Linear Algebra and its Applications*, 290:275–301, 1999.

[59] B. McKay. On the spectral characterisation of trees. *Ars Combinatoria*, 3:219–232, 1977.

[60] B. D. McKay. nauty user's guide (version 1.5). Technical report, Australian National University, Department of Computer Science, http://cs.anu.edu.au/people/bdn/nauty/, 1990. Technical Report TR-CS-90-02.

[61] C. B. Moler and G. W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM Journal on Numerical Analysis*, 10(2):241–256, 1973.

[62] A. Pothen and C. Fan. Computing the block triangular form of a sparse matrix. *ACM Transaction on Mathematical Software*, 16(4):303–324, 1990.

[63] J. F. Queiró and A. L Duarte. Imbedding conditions for normal matrices. *Linear Algebra and its Applications*, 430:1806–1811, 2009.

[64] R. C. Read and D. G. Corneil. The graph isomorphism disease. *Journal of Graph Theory*, 1:339–363, 1977.

[65] P. Rowlinson. The main eigenvalues of a graph: a survey. *Applicable Analysis and Dsicrete Mathematics*, 1(2):455–471, 2007.

[66] A. Ruhe. Algorithms for the nonlinear eigenvalue problem. *SIAM Journal of Matrix Analysis and Applications*, 10(4):674–689, 1973.

[67] Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. Manchester University Press, 1992.

[68] A. Schwenk. Computing the characteristic polynomial of a graph. In R. Bari and F. Harary, editors, *Graphs and Combinatorics*, volume 406 of *Lecture Notes in Mathematics*, pages 153–172. Springer, 1974.

[69] M. K. Sparrow. A linear algorithm for computing automorphic equivalence classes: the numerical signature approach. *Social Networks*, 15:151–170, 1993.

[70] W. Specht. Zur theorie der matrizen. ii. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 50:19–23, 1940.

[71] P. F. Stadler and G. Tinhofer. Equitable partitions, coherent algebras and random walks: applications to the correlation structure of landscapes. *MATCH Communications in Mathematical and in Computer Chemistry*, 40:215–261, 1999.

[72] G. W. Stewart. *Matrix algorithms: Eigensystems, Band 2*. SIAM, 2001.

[73] X. Sun. On elementary unitary and ϕ-unitary transformations. Technical report, Duke University, 1995.

[74] R. Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2):146–160, 1972.

[75] R. C. Thompson. Principal submatrices of normal and Hermitian matrices. *Illinois Journal of Mathematics*, 10:296–308, 1966.

[76] F. Tisseur and I. Zaballa. Triangularization of matrix polynomials. Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012. In preparation.

[77] F. Tisseur and I. Zaballa. Triangularizing quadratic matrix polynomials. MIMS EPrint 2012.30, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012.

[78] Krauthgamer U. Feige R. Finding and certifying a large hidden clique in a semirandom graph. *Random Structures and Algorithms*, 16(2):195–208, 2000.

[79] E. R. van Dam and W. H. Haemers. Developments on spectral characterizations of graphs. *Discrete Mathematics*, 309:576–586, 2009.

[80] E. R. van Dam, W. H. Haemers, and J. H. Koolen. Cospectral graphs and the generalized adjacency matrix. *Linear Algebra and its Applications*, 423:33–41, 2007.

[81] E. R. van Dam, W. H. Haemers, J. H. Koolen, and E. Spence. Characterizing distance-regularity of graphs by the spectrum. *Journal of Combinatorial Theory, Series A*, 113:1805–1820, 2006.

[82] J. H. van Lint and J. J. Seidel. Equilateral point sets in elliptic geometry. *Proceedings. Koninklijke Nederlandse Akademie van Wetenschappen. Series A*, 69(3):335–348, 1966.

[83] H. Voss. An arnoldi method for nonlinear eigenvalue problems. Technical report, Dept.Math., TU Hamburg-Harburg, 2002.

[84] Y. Wu. Invariant subspace, determinant and characteristic polynomials. *Linear Algebra and its Applications*, 428:1753–1760, 2008.

[85] Y. Wu and A. Deng. Hoffman polynomials of nonnegative irreducible matrices and strongly connected digraphs. *Linear Algebra and its Applications*, 414:138–171, 2006.

[86] F. Zang. *The Schur Complement and Its Application*, volume 4 of *Numerical Methods and Algorithms*. Springer, 2005.

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den 31.08.2012

. . . .
(Mario Thüne)