KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

# Structured Matrix Techniques for Orthogonal Rational Functions and Rational Krylov Methods

**Niel Van Buggenhout**

Supervisors:
Prof. dr. ir. M. Van Barel
Prof. dr. R. Vandebril

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Computer Science

October 2021

# Structured Matrix Techniques for Orthogonal Rational Functions and Rational Krylov Methods

**Niel VAN BUGGENHOUT**

Examination committee:
Prof. dr. ir. P. Wollants, chair
Prof. dr. ir. M. Van Barel, supervisor
Prof. dr. R. Vandebril, supervisor
Prof. dr. ir. K. Meerbergen
Prof. dr. ir. W. Michiels
Prof. dr. W. Van Assche
Prof. dr. P. Boito
(Università di Pisa)
Prof. dr. L. Reichel
(Kent State University)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Computer Science

October 2021

# Preface

This thesis is the result of 4 years of research conducted under the supervision of Marc Van Barel and Raf Vandebril. Here I want to take the time to express my gratitude to the people who played an important role in my life during the years of my PhD.

Nearing the end of my master studies I got more and more interested in doing a PhD. This is in part thanks to Marc, who supervised my master thesis at that time. When he offered to do a PhD under his supervision I did not have to think about it very long. I had only one request, to start in january so I could explore Europe for some months, to which Marc agreed. Afterwards I returned to Belgium and started my PhD, to my surprise I had two supervisors. My second supervisor was Raf, who I met shortly before I started. From the start it was a pleasure to work with Raf. I consider myself lucky to have been supervised by Marc and Raf these past years. Thank you for the freedom I had to choose my own path and for the guidance I received when this path got a bit too rough. I am grateful for the ideas you shared with me, the interesting and frequent meetings we had and the opportunities you provided me.

The quality of this thesis improved thanks to the jury members who careful read the text, provided me with valuable comments and asked me challenging question on the preliminary defense. So, Paola Boito, Karl Meerbergen, Wim Michiels and Walter Van Assche, thank you for being a part of my jury. Thanks to Patrick Wollants for chairing my defenses. Lothar Reichel, thank you for being part of my jury as well and agreeing to host me when I wanted to do a research visit. Unfortunately this visit never happened, but getting the opportunity already means a lot.

I did have the opportunity to attend many conferences. My first conference, M2A19 in Morocco, was a nice experience and I am glad that Dries attended the conference as well. It was nice to share it with a colleague and a friend. The following conferences brought me to every corner of the world, ILAS in Brazil, ETNA25 in Italy with Daan and NASC in China where I accompanied Raf. These conferences broadened my view on my own topic and introduced me to new topics. The last conference I attended, SIAM LA, was held in the virtual world, and it was nice to see some familiar faces,

Leonardo Robol, Stefano Pozza and Thomas Mach, again after more than a year.

When I was not at a conference, I was often at the office. It was a pleasure to share an office with you, Elien, Paul, Hannes, Haik and Sahar. When it was not allowed to go to the office, I did miss the conversations during our tea breaks. I appreciate that some colleagues did regularly visit the office once it was allowed again, it was nice to have these brave souls, Simon T., Pieter and Tom, accompanying me in an empty building and sharing lunches together again. The office was a good working environment thanks to many other colleagues, Daan, Andreas, Emil, Bert, Pieterjan, Philippe, Koen, Luca, Francesco, Nick, Ward, Ignace, Rob, Thijs, Simon D., Adrian, Marcus, Wouter, Nina, Jonas, Jonathan and Tim.

Once I left the office, I would usually return home. During my co-housing I could always rely, in good and not so good times, on my housemates whom I treasure a lot. Bri, Gil, Jana and Clea, the time we spend living together in Heverlee was intense, fun and valuable. Laure and Nathalie, sharing a house with you was a pleasure, thank you for being so welcoming during my short time in Sint-Gillis.

After some weeks of writing this text I needed a change of location. Most chapters are written at home or in the office, but some of them are written elsewhere. For this reason I dedicate Chapter 3 and 4 to Barcelona, thank you Clea, Marta and Marina for hosting me, I had a wonderful time. Chapter 5 is dedicated to Rebecq where we could enjoy breaks in the sun in a big, wild garden. When the commute between my office and Brussels became too much on a long day of writing, I had a place to sleep close by thanks to the hospitality of Camille, Cooper, Bri and Gil. So I dedicate chapter 6 to the PDS-laan. And I would like to thank Tom for proofreading a part of this text.

Under some external pressure in 2020 everyone picked up hiking as a hobby and had to choose a single companion to share these hikes with. I am happy that my hiking companion was Mathilde, whom I got to know better during these hikes and with whom I found a lot of support. There are still many friends that I would like to thank, for their friendship, openness and kindness, thank you Fre, Kaatje, Sam, Schoofs, Enya, Bas, Eva, Joost, Lize, Seppe, Simon, Helena, Klaas, Annelies, Koen en Jolle.

Mama en papa, dank jullie om mij te steunen doorheen mijn opleiding en doctoraat, zowel met praktische zaken als een luisterend oor te bieden. Ook bedankt om mij, wanneer ik te veel klaagde, er aan te herinneren dat ik er zelf voor had gekozen om een doctoraat te doen. Femke, Lobke en Sil, dank jullie voor de tijd samen en de steun. Zichem is altijd een plaats geweest waar ik even tot rust kon komen. Ook bij ma en pa, waar ik altijd terecht kon voor een bezoekje en een praatje. En als laatste wil ik Lily bedanken, om te tonen dat er altijd mag gedanst worden, waar je ook bent en ook al ben je de enige die danst.

Niel

# Abstract

Polynomials are a powerful tool to approximate functions. If the function of interest does not resemble a polynomial, rational function based methods might be more appropriate.

The theory of polynomials is well established in the field of classical analysis. Sequences of polynomials with orthogonality properties are preferred for computations in finite precision. Many effective numerical methods for polynomials originate from numerical linear algebra. A connection between orthogonal polynomials and Krylov subspaces allows to translate theory from classical analysis to numerical linear algebra and to apply numerical procedures from numerical linear algebra to problems in classical analysis.

For rational functions similar connections to rational Krylov subspaces remain unexploited. Here we develop the necessary theory to identify these connections, relate rational functions to structured matrices and develop numerical procedures based on structured matrices for problems involving rational functions.

Rational Krylov subspaces are chosen as the starting point of our exposition. All structured matrix pencils that construct orthogonal and biorthogonal bases for these spaces are derived. In particular, a tridiagonal matrix pencil is shown to suffice for the construction of biorthogonal bases. This allows us to design an efficient Lanczos like iteration for rational Krylov subspaces.

Generating orthogonal and biorthogonal vectors is related to the factorization of the associated Gram matrix. The displacement structure of Gram matrices related to rational Krylov spaces is studied and the ones exhibiting low displacement rank are classified.

The (bi)orthogonal basis vectors for rational Krylov subspaces can be related to rational functions (bi)orthogonal with respect to a discrete (linear functional) inner product. The specific form of these discrete inner products and linear functionals is derived, these are weighted sums of function evaluations. For these specific (linear

functionals) inner products, (bi)orthogonal rational functions can be directly related to rational Krylov subspaces and the corresponding structured pencils.

These connections are used to propose numerical procedures based on Krylov subspaces and structured matrices for problems involving rational functions. In particular, the computation of a sequence of (bi)orthogonal rational functions is reformulated as an inverse eigenvalue problem. The latter problem is a problem in linear algebra and techniques from numerical linear algebra can be applied to solve this problem.

# Beknopte samenvatting

Veeltermbenadering werkt goed voor het benaderen van functies. Als de functie, waarin we geïnteresseerd zijn, niet lijkt op een veelterm, dan zijn methodes gebaseerd op rationale functies mogelijks beter geschikt.

Klassieke analyse bevat een uitgebreide theorie over veeltermen. Rijen van veeltermen met orthogonaliteitseigenschappen genieten de voorkeur om berekeningen uit te voeren in eindige precisie. Veel effectieve numerieke methoden voor veeltermen zijn onstaan uit onderzoek in numerieke lineaire algebra. Een verband tussen orthogonale veeltermen en Krylov deelruimten laat toe om de theorie van klassieke analyse te vertalen naar numerieke lineaire algebra en om numerieke procedures van numerieke lineaire algebra toe te passen op problemen uit klassieke analyse.

Voor rationale functies zijn gelijkaardige verbanden met rationale Krylov deelruimten nog niet uitgebuit. In deze tekst ontwikkelen we de nodige theorie om deze verbanden bloot te leggen, brengen we rationale functies in verband met gestructureerde matrices en ontwikkelen we numerieke procedures gebaseerd op gestructureerde matrices om problemen geformuleerd met rationale functies op te lossen.

Rationale Krylov deelruimtes zijn gekozen als startpunt voor deze tekst. Alle gestructureerde matrix pencils die orthogonale en biorthogonale basissen genereren voor deze deelruimtes worden afgeleid. We bewijzen dat een tridiagonaal matrix pencil volstaat om biorthogonale basissen op te stellen. Dit laat toe om een efficiënte veralgemening van de Lanczos iteratie te ontwikkelen voor rationale Krylov deelruimtes.

Het opstellen van orthogonale en biorthogonale vectoren is nauw verwant aan het factoriseren van de geassocieerde Gram matrix. De structuur van zulke matrices, gerelateerd aan rationale Krylov deelruimtes, onder goed gekozen verplaatsingsoperatoren wordt bestudeerd. De matrices die hieronder een lage rang hebben worden geclassificeerd.

De (bi)orthogonale basisvectoren voor rationale Krylov deelruimtes kunnen in verband gebracht worden met rationale functies die (bi)orthogonaal zijn ten opzichte van

een discreet (lineaire functionaal) inwendig product. We bepalen ook de specifieke vorm van deze discrete inwendige producten en lineaire functionalen. Voor deze specifieke (lineaire functionalen) inwendige producten kunnen we een direct verband tussen (bi)orthogonale rationale functies en rationale Krylov deelruimtes en bijhorende gestructureerde pencils leggen.

Uitgaande van deze verbanden stellen we numerieke procedures voor, gebaseerd op Krylov deelruimten en gestructureerde matrices, om problemen geformuleerd voor rationale functies op te lossen. In het bijzonder herformuleren we het berekenen van een rij (bi)orthogonale rationale functies als een invers eigenwaardeprobleem. Een invers eigenwaardeprobleem is een probleem uit lineaire algebra en technieken uit numerieke lineaire algebra kunnen toegepast worden om dit probleem op te lossen.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Identifying connections between different topics in mathematics allows us to translate theoretical results and numerical procedures from one topic to another topic. The topics in this text are generalizations of Krylov subspaces from numerical linear algebra and orthogonal polynomials from classical analysis. The interplay between numerical linear algebra and classical analysis is fruitful in both directions.

Many numerical procedures in mathematics make use of the theory and methods developed in numerical linear algebra. With the advent of computers capable of storing program instructions, stored-program computers, in the late 1940s, interest in error analysis describing the cumulative effects of rounding errors in finite precision arithmetic increased significantly. Matrix theory has proven to be a natural tool to study the effects of error propagation. Research in this area has led to the development of effective numerical procedures for problems in linear algebra and in other fields of mathematics [184]. For example solving least squares problems, constructing sequences of orthogonal polynomials and computing the nodes and weights of quadrature rules [89].

Classical analysis is an older discipline, the study of orthogonal polynomials arose from continued fractions in the late 19th century [4]. A vast theory on orthogonal polynomials is available, some standard introductions to this topic are [43,110,154,157]. This theory can be applied or translated to other disciplines in mathematics, such as numerical analysis, approximation theory and numerical linear algebra. Examples are the computation of nodes and weights for Gauss quadrature rules, error estimates for linear algebraic systems and convergence theory for Krylov subspace methods.

The specific topics handled in this text are described for the classical case in Section 1.1. In Section 1.2 two generalizations that will be studied are described and justified.

## 1.1 The interplay - classical case

The now classical connections between the Lanczos iteration for Krylov subspaces and polynomials orthogonal with respect to an inner product are sketched. For a more complete discussion we refer to the literature [28, 82, 89, 124].
A Krylov subspace for a matrix $A \in \mathbb{C}^{m\times m}$ and vector $v \in \mathbb{C}^m$ is

$$\mathcal{K}_k(A, v) = \text{span}\{v, Av, \ldots, A^{k-1}v\}.$$

For a Hermitian matrix $A^H = A$, the Hermitian Lanczos iteration computes an orthogonal basis for this space via a three term recurrence relation. This three term recurrence relation also generates a sequence of orthogonal polynomials with respect to a discrete inner product. The recurrence coefficients can be grouped into a Jacobi matrix (a Hermitian tridiagonal matrix), which represents the projection of $A$ onto the Krylov subspace. Roots of orthogonal polynomials correspond to the eigenvalues of the associated Jacobi matrix. This Jacobi matrix matches the first $2k-1$ moments of $A$, these moments are $v^H A^i v$, $i = 0, 1, \ldots, 2k-2$. Moments also play a key role in Padé approximation, approximating the power series of a function of interest up to its first $2k-1$ terms, and in the theory of quadrature. A Gauss quadrature rule with $k$ nodes matches $2k-1$ moments of the integrand of interest according to a given measure. The $k$ nodes of a Gauss quadrature rule correspond to the roots of a polynomial of degree $k$ orthogonal with respect to an inner product based on this measure. And thus the nodes of a Gauss quadrature rule with respect to certain discrete inner products correspond to the eigenvalues of a Jacobi matrix generated by the Hermitian Lanczos iteration applied to Krylov subspaces generated by a specific matrix $A$ and vector $v$. Which brings us back at Krylov subspaces.
Some specific examples of the above connections are:

- The Golub-Welsch algorithm to compute Gauss quadrature nodes and weights by solving an eigenvalue problem for a Jacobi matrix [91].
- The construction of orthogonal polynomials, given an inner product, by computing the related Jacobi matrix with techniques from numerical linear algebra [92].
- Convergence results for Krylov subspace methods by applying potential theory developed for orthogonal polynomials [14, 100, 119, 120].
- Computing a Padé approximation by applying the Lanczos iteration [93].

## 1.2 The interplay - Generalizations

This text generalizes these connections in two ways. The first is a generalization from an inner product to a linear functional (indefinite inner product). Orthogonality with respect to a linear functional will, in this thesis, be interpreted as biorthogonality. The numerical linear algebra techniques for polynomials orthogonal with respect to an inner product make use of unitary transformations to manipulate the related matrix. Unitary transformations of a matrix do not cause a large growth in the Euclidean or Frobenius norm going from the original to the transformed matrix in finite precision. The crucial factor for numerical stability of a numerical procedure manipulating a matrix is the control of growth in the size of the entries of this matrix [184]. Hence, unitary transformations are inherently stable.
For linear functionals it is, in general, no longer possible to use unitary transformations. The transformations that will be used are not inherently stable, which makes the development of numerically stable procedures more difficult. However, such biorthogonal procedures lead to more efficient algorithms and in some applications a linear functional must be used. Justification of the development of methods using nonunitary transformations are provided below for several different topics.
The second is a generalization from polynomials to rational functions with prescribed poles. Polynomials are a powerful tool to approximate functions. However, if the function to be approximated is not 'polynomial-like', the approximation might converge slowly. For example, if we want to approximate an integral of a function that has a singularity just outside the region of interest. In this case, a rational function with poles close to this singularity is a more natural candidate for approximation than a polynomial [54].

### 1.2.1 Krylov subspace methods

In Krylov subspace methods, the interest in biorthogonal procedures arises from their potential gain in efficiency. An orthogonal basis for a Krylov subspace $\mathcal{K}_k(A, v)$, with $A^H \neq A$, satisfies a recurrence relation governed by a Hessenberg matrix. This implies that the computation of basis vectors requires orthogonalization with respect to all previously generated basis vectors. If $A$ is very large, then keeping all basis vectors in memory might become prohibitively expensive. Usually one is interested in the matrix containing the recurrence coefficients and not in the basis vectors. A more efficient procedure is obtained if a pair of biorthogonal bases for specific Krylov subspaces is computed. One basis for $\mathcal{K}_k(A, v)$ and another, with $A^H$ the Hermitian conjugate of $A$ and some vector $w$, for

$$\mathcal{K}_k(A^H, w) = \text{span}\{w, A^H w, \ldots, (A^H)^{k-1} w\}.$$

Such a pair of biorthogonal bases can be constructed via a three term recurrence relation. This leads to a tridiagonal matrix and requires only 3 basis vectors to be kept in memory. The tridiagonal matrix represents the oblique projection onto $\mathcal{K}_k(A, v)$. Hence, using biorthogonal bases leads to a more efficient procedure, but developing a biorthogonal algorithm with limited amplification of rounding errors is challenging. A popular biorthogonal Krylov subspace method for solving large systems of linear equations is BiCGSTAB [168].
Krylov subspaces implicitly use polynomials, this can be seen by noting that $x \in \mathcal{K}_k(A, v)$ can be written as $x = p_{k-1}(A)v$, for some polynomial $p_{k-1}$ of degree $k-1$. Eigenvalue methods based on Krylov subspaces typically converge first to extreme eigenvalues of $A$. If, instead of a polynomial, a rational function underlies the subspace, i.e., $x = r_{k-1}(A)v$, the poles of this rational function can be chosen to focus convergence of eigenvalue methods on specific regions in the complex plane [143]. This idea has proven to be especially powerful for the approximation of matrix functions [58, 138]. Other applications can be found in [133] and references therein.

### 1.2.2 Gauss quadrature

Quadrature rules for linear functionals arise in quantum physics when trying to compute the *scattering amplitude* $w^\top A^{-1} v$. One approach to approximate the scattering amplitude is to use biorthogonal Krylov subspace methods, another is by complex Gaussian quadrature [150]. In the approximation of highly oscillatory integrals Gauss quadrature rules lead to *kissing polynomials* [42]. These polynomials are orthogonal with respect to a linear functional. The theory of kissing polynomials is well-developed in classical analysis, however, stable numerical procedures to compute with these polynomials are still lacking.

Gauss quadrature rules are exact for a space of polynomials up to a certain degree. Rational Gauss quadrature rules are exact for a space spanned by rational functions with prescribed poles. If the integrand of the integral of interest can be better approximated by a rational function, e.g., it has singularities outside the region of interest and close to its boundary, then rational Gauss quadrature rules are more appropriate than standard Gauss quadrature rules [54, 81, 139]. To obtain upper and lower bounds for functionals of the form $v^\top f(A)v$, for some function $f$, rational functions provide a powerful alternative to polynomials [127].

### 1.2.3 Model order reduction

In system theory, a linear, time invariant dynamical system is represented by state space equations

$$\dot{x}(t) = Ax(t) + bu(t)$$

$$y(t) = c^H x(t)$$

with a matrix $A$ and vectors $b, c$. The functions $x, y, u$ vary in time $t$ and represent the state, output and input of the system, respectively. If the matrix $A$ is very large, it is interesting to approximate the system, determined by $(A, b, c)$, by a smaller system $(\hat{A}, \hat{b}, \hat{c})$. A popular approach to construct the smaller system is by moment matching. That is, the reduced order model $(\hat{A}, \hat{b}, \hat{c})$ is constructed such that the first $l$ moments of both systems correspond

$$c^H A^i b = \hat{c}^H \hat{A}^i \hat{b}, \quad \text{for } i = 0, 1, \ldots l.$$

This is related to the approximation of the $l$ first terms in the power series expansion of the associated transfer function. The biorthogonal Lanczos iteration can be applied to obtain such a reduced order model [93].

Methods based on polynomial Krylov subspaces will focus on one frequency in the construction of the reduced order model. This corresponds to approximating a power series expansion of the original transfer function around a single frequency. Approximation around multiple frequencies with one reduced model can lead to a smaller model and thus to a more efficient representation of the original model. For details we refer to [2, 73].

## 1.3 Outline

An outline of this text is provided here. Our starting point is numerical linear algebra, more precisely Krylov subspaces and structured matrices. The goal is to obtain efficient and numerically stable procedures to solve problems in classical analysis.

Chapter 2 is a preliminary chapter and contains the basic notions that will be used. Two specific vector spaces are introduced and the notion of orthogonality and biorthogonality is defined. The relation between (bi)orthogonality and structured matrices is discussed shortly.

Chapter 3 studies the structured matrices that arise from (polynomial) Krylov subspaces. The recurrence relations to generate bases for these subspaces, how these

recurrence relations can be represented by a matrix and which properties this matrix possesses are the topics of this chapter. For orthogonal bases the *core factorization* of the matrix of recurrence coefficients is introduced, which represents the matrix as a product of essentially $2 \times 2$ unitary matrices multiplied with an upper triangular matrix. For biorthogonal bases, a nonunitary analogue of the core factorization is introduced, which is called the *eliminator factorization*.

Chapter 4 studies the structured matrices that arise from rational Krylov subspaces. The core and eliminator factorizations are used to derive the structures of these. This leads to a framework that allows to describe all the possible recurrence relations for (bi)orthogonal bases. One of the possible recurrence relations is a three term recurrence, which implies an efficient procedure to construct the bases. The recurrence coefficients can be captured in a tridiagonal matrix pencil. The results of this chapter are published in [165].

In Chapter 5 a meaningful generalization of moments to the context of rational Krylov subspaces is introduced. The connection between rational Krylov subspaces, moment matching and multi-point Padé approximation is discussed. The displacement structure of Gram matrices associated with rational Krylov subspaces are studied and suitable displacement operators are proposed. Gram matrices related to pairs of Krylov subspaces, one generated with a matrix and the other with the complex conjugate of the same matrix, are shown to have displacement rank at most two.

Chapter 6 relates polynomials to structured matrices via the recurrence relation that generates these polynomials. The connection between (bi)orthogonal vectors in polynomial Krylov subspaces and inner products or linear functionals on the space of polynomials is identified. Properties shared by polynomials and the related structured matrices are stated here. Problems formulated for polynomials can, by the theory of this chapter, be reformulated as problems for certain structured matrices.

Chapter 7 develops the theory that allows the reformulation of problems involving rational functions into problems for structured matrix pencils. Rational functions orthogonal with respect to a specific inner product or linear functional induced by (bi)orthogonal bases for rational Krylov subspaces are described. Results from the literature on orthogonal rational functions relevant to the topics discussed in this text are listed.

In Chapter 8 two numerical procedures are proposed, based on the theory developed in the aforementioned chapters. One procedure provides a Lanczos like iteration

constructing biorthogonal bases for rational Krylov subspaces using a three term recurrence relation. This is based on [165]. The second procedure is a Levinson procedure applied to a specific pair of rational Krylov subspaces. This procedure is published in [166]. Numerical tests proving the validity of both procedures are included.

Chapter 9 employs the connections between rational Krylov subspaces and rational functions to reformulate a problem in classical analysis into a problem for structured matrix pencils. This problem is the computation of a sequence of orthogonal rational functions for a given inner product or a pair of biorthogonal rational functions for a linear functional. The reformulated problem is a structured inverse eigenvalue problem and the remainder of this chapter is dedicated to developing numerical procedures to solves, update and downdate inverse eigenvalue problems. Some results in this chapter appeared in [167], and the downdating procedures are based on an article in preparation [164].

Appendix A contains some proofs that are omitted from the main text.

Appendix B contains the derivation of the rational Lanczos iteration.

The dependencies among the chapters are shown in Figure 1.1. For Chapter 2 reading can be restricted to Section 2.1.2, which clarifies the terms orthogonal vectors and biorthogonal vectors. Dashed lines indicate only a weak dependency, full lines indicate a strong dependency on another chapter. The left side of the figure uses Krylov subspaces as a starting point, the right side uses orthogonal polynomials and the center combines both topics.

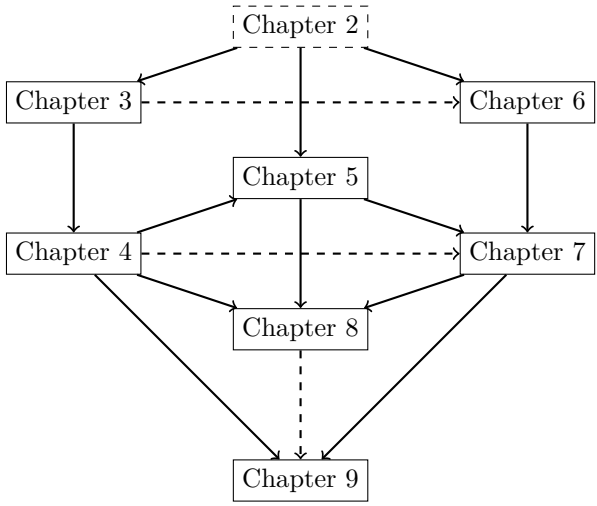



Figure 1.1: Scheme showing principal dependencies between different chapters.

# Chapter 2

# Orthogonal and biorthogonal vectors

The study of orthogonality leads naturally to structured matrices and matrix factorizations. The arising structured matrices and their relation to (bi)orthogonal vectors are the main subject of this thesis. Section 2.1 introduces some basic concepts that are useful later on, such as vector spaces, and formally defines an inner product and linear functionals used to characterize orthogonality. These formal definitions are given to avoid confusion, since there are different conventions in the literature. The connection between (bi)orthogonal vectors and Gram matrices is elaborated on in Section 2.2. Gram matrices can exhibit displacement structure, which is defined in Section 2.3 together with other structured matrices. Section 2.4 provides the remaining concepts required, such as the eigenvalue decomposition, function of a matrix and grade of a vector. The contents of this chapter are mostly based on the book by Horn and Johnson [105] and the concepts from functional analysis can be found in standard texts on this topic [158].

## 2.1 Vectors

Vectors with some properties, being orthogonal or biorthogonal, are the main objects of study in this thesis. Orthogonality is a well-known concept, biorthogonality is not as well-known. Biorthogonality refers to two sets of vectors that are orthogonal with respect to each other for some linear functional. Section 2.1.1 introduces the vector spaces appearing in this text and Section 2.1.2 defines (bi)orthogonal vectors and

bases. The discussion on biorthogonality is restricted to the vector spaces introduces in the former section.

### 2.1.1 Vector spaces and bases

Vector spaces (or linear space) over some field $\mathbb{F}$ are denoted by $\mathcal{V}$. In this text the only field that is considered is the field of complex numbers, i.e., $\mathbb{F} = \mathbb{C}$.
We will, essentially, work with two vector spaces: the space of $m$-tuples of complex numbers $\mathbb{C}^m$ and the space of complex polynomials $\mathcal{P}$.
A *basis* for $\mathcal{V}$ is a linearly independent subset $B$ which spans the whole vector space $\mathcal{V}$. The *span* of a subset, defined by the columns of the matrix $B = \begin{bmatrix} b_0 & b_1 & \dots & b_{k-1} \end{bmatrix}$, of a vector space $\mathcal{V}$ is the set $\text{span}\{b_0, b_1, \dots, b_{k-1}\} := \{\alpha_0 b_0 + \alpha_1 b_1 + \cdots + \alpha_{k-1} b_{k-1} | \alpha_i \in \mathbb{C}, i = 1, 2, \dots, k-1\}$. The span of a subset composed of the columns of a matrix $B$ will be denoted as $\text{span}\{B\} = \text{span}\{b_0, b_1, \dots, b_{k-1}\}$.
If $\text{span}\{B\} = \mathcal{V}$, then $B$ is said to span the vector space $\mathcal{V}$. Bases are not unique, and we will argue below that some bases are more appropriate than others, depending on the situation. Let (the columns of) $B$ form a basis for some vector space $\mathcal{V}$, then each element of $\mathcal{V}$ has a unique representation in $B$. An example of a basis for $\mathcal{P}$ is the *monomial basis* $\begin{bmatrix} 1 & z & z^2 & \dots \end{bmatrix}$.
A subset $\mathcal{S}$ of $\mathcal{V}$ is a *subspace* of $\mathcal{V}$, i.e., $\mathcal{S} \subseteq \mathcal{V}$, whenever $\mathcal{S}$ itself is a vector space over $\mathbb{C}$. For example, the space of polynomials up to degree $k$, denoted by $\mathcal{P}_k$, is the span of $\begin{bmatrix} 1 & z & z^2 & \dots & z^k \end{bmatrix}$ and is a subspace of $\mathcal{P}$.

### 2.1.2 Orthogonality and biorthogonality

Orthogonality relies on an inner product and biorthogonality relies on a linear functional. These are introduced below together with the notion of (bi)orthogonal vectors. Bases formed with (bi)orthogonal vectors are called (bi)orthogonal bases and are the main topic of this manuscript. We will see that these have several nice properties.

#### Inner product

An inner product, denoted as $\langle ., . \rangle$, is defined below. It is a Hermitian positive-definite sesquilinear form.

**Definition 2.1** (Sesquilinear form)**.** *Let $\mathcal{V}, \mathcal{W}$ be vector spaces over $\mathbb{C}$. A map $\mathcal{V} \times \mathcal{W} \to \mathbb{C} : (x, y) \mapsto \langle x, y \rangle$ is called a sesquilinear form if $\langle ., . \rangle$ is linear in the first*

*and anti-linear in the second variable:*

$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle \text{ for all } x, y \in \mathcal{V}, z \in \mathcal{W}, \quad \alpha, \beta \in \mathbb{C};$$

$$\langle x, \alpha y + \beta z\rangle = \bar{\alpha}\langle x, y\rangle + \bar{\beta}\langle x, z\rangle \text{ for all } x \in \mathcal{V}, y, z \in \mathcal{W}, \quad \alpha, \beta \in \mathbb{C}.$$

**Definition 2.2** (Hermitian form)**.** *A Hermitian form is a sesquilinear form with $\mathcal{V} = \mathcal{W}$ that is Hermitian:*

$$\langle y, x\rangle = \overline{\langle x, y\rangle}.$$

**Definition 2.3** (Inner product as positive-definite Hermitian form)**.** *An inner product is a positive-definite Hermitian form*

$$\langle x, x\rangle > 0 \text{ for all } x \in \mathcal{V}\backslash\{0\},$$

*where the superscript $H$ denotes the Hermitian conjugate.*

For $\mathbb{C}^m$, the *Euclidean inner product* is denoted by $\langle ., .\rangle_E$, that is, for $x, y \in \mathbb{C}^m$ we have

$$\langle x, y\rangle_E := y^H x.$$

For $\mathcal{P}$, let $p, q \in \mathcal{P}$ be some complex polynomials. An inner product can be defined as $\int p(z)\overline{q(z)}d\mu(z)$, where the bar denotes the *complex conjugate* and $\mu$ is a finite Borel measure on $\mathbb{C}$ with all the associated moments finite. For details see [154]. We consider inner products of the form $\int_\gamma p(z)\overline{q(z)}\alpha(z)|dz|$ on $\mathcal{P}$, with $\alpha(z)$ a positive weight function, $\gamma$ an arc in the complex plane and where $|dz|$ denotes the arc length [150]. Our focus will be on discrete inner products on the subspace $\mathcal{P}_{m-1}$, for *nodes* $z_i \in \mathbb{C}$ and *weights* $\alpha_i > 0$, $i = 1, \ldots, m$,

$$\langle p, q\rangle_m := \sum_{i=1}^{m} \alpha_i p(z_i)\overline{q(z_i)}, \quad z_i \in \mathbb{C},$$

with $p, q \in \mathcal{P}_{m-1}$.

### Orthogonal vectors

Let $\delta_{ij}$ denote the *Kronecker delta*. A set of vectors $\{v_i\}_i$ in a vector space $\mathcal{V}$ satisfying

$$\langle v_i, v_j\rangle = \eta_i \delta_{ij}, \quad \eta_i \in \mathbb{R}$$

is said to be *orthogonal* with respect to the inner product $\langle ., .\rangle$. If $\eta_i = 1$ for all $i$, the set of vectors is called *orthonormal*.
An *orthogonal basis* $V$ is a basis for a vector space $\mathcal{V}$ formed by vectors $\{v_i\}_i$ that are

orthogonal with respect to an inner product.
Given a set of linearly independent vectors $\{b_i\}_{i=0}^{k-1}$, $b_i \in \mathcal{V}$ and an inner product $\langle .,. \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$. Then the Gram-Schmidt orthogonalization procedure [122] can be applied to obtain an orthogonal set of vectors $\{v_i\}_{i=0}^{k-1}$ such that $\text{span}\{v_0, \ldots, v_l\} = \text{span}\{b_0, \ldots, b_l\}$ for $l = 0, 1, \ldots, k-1$. In finite precision the modified Gram-Schmidt procedure should be used and some reorthogonalization might be necessary [24, 75, 126, 146].

### Linear functional

Biorthogonality can be defined with respect to a *bilinear form* [35] or with respect to a *linear functional* [29]. In this manuscript the latter is used. A linear functional $\mathcal{L}$ is a linear mapping of some vector space $\mathcal{V}$ into the field $\mathbb{C}$, i.e., $\mathcal{L}\{.\} : \mathcal{V} \to \mathbb{C}$.

For $\mathcal{P}_l$, polynomials up to degree $l$, the linear functional $\mathcal{L}$ used in this text is defined by fixing its value for certain polynomials to a scalar from a given sequence of so-called *moments*. The sequence is $\{m_i\}_{i=0}^{2k-1}$, $m_i \in \mathbb{C}$, and the linear functional is defined by

$$\mathcal{L}\{z^i\} = m_i, \quad i = 0, 1, \ldots, 2k-1.$$

The scalars $m_i$ are called the *(classical) moments* of the linear functional if the monomial basis is used. If another basis is used, these are called *modified moments*.

**Remark 2.1.** *Consider two distinct subspaces of the space of m-tuples, $\mathcal{S}_1 \subset \mathbb{C}^m$ and $\mathcal{S}_2 \subset \mathbb{C}^m$ and the Euclidean inner product $\langle .,. \rangle_E$. For these subspaces the definition $\langle .,. \rangle_E : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathbb{C}$ is meaningful, however it is only an inner product on $\mathcal{S}_1 \cap \mathbb{S}_2$. Suppose we have two sets of vectors $\{v_1, \ldots, v_k\}$, $v_i \in \mathcal{S}_1$ and $\{w_1, \ldots, w_k\}$, $w_i \in \mathcal{S}_2$ such that $v_j$ and $w_j$ are related in a meaningful way. Then, applying the Euclidean inner product to these vectors does not guarantee that $\langle v_j, w_j \rangle_E > 0$. Therefore, when working with distinct subspaces of $\mathbb{C}^m$ we interpret the Euclidean inner product sesquilinear form or, equivalently, a linear functional. We will, however, keep using the name Euclidean inner product for this case.*

### Biorthogonal vectors

A pair of sets of polynomials $\{p_i\}_i$ and $\{q_j\}_j$, $p_i, q_j \in \mathcal{P}$, is said to be *biorthogonal* with respect to $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ if they satisfy

$$\mathcal{L}\{p_i q_j\} \begin{cases} = 0, & \text{if } i \neq j \\ \neq 0, & \text{if } i = j \end{cases}.$$

They are said to be *biorthonormal* if it is a biorthogonal pair and $\mathcal{L}\{p_i q_i\} = 1$.
A pair of distinct sets of vectors $\{x_i\}_i$ and $\{y_j\}_j$, with $x_i, y_j \in \mathbb{C}^m$, satisfying

$$\langle x_i, y_j \rangle_E = \delta_{ij},$$

is said to be biorthogonal. It is not orthogonal, because the property $\langle x_i, y_i \rangle_E > 0$ cannot be guaranteed for $x_i \neq y_i$.
A pair of bases $V = \begin{bmatrix} v_0 & v_1 & \dots & v_{k-1} \end{bmatrix}$, $W = \begin{bmatrix} w_0 & w_1 & \dots & w_{k-1} \end{bmatrix}$ for some vector spaces $\mathcal{V}, \mathcal{W}$ is called biorthogonal if the sets $\{v_i\}_{i=0}^{k-1}$ and $\{w_i\}_{i=0}^{k-1}$ are biorthogonal. A pair of biorthogonal sets of vectors can be generated from a pair of linearly independent vector sets by the two-sided Gram-Schmidt procedure [135]. However these do not always exist, a condition for their existence is given in the following section.

## 2.2 Gram matrices

Gram matrices are an important tool in the study of (bi)orthogonality. A Gram matrix $M$ can be defined using an inner product or linear functional. Consider an inner product $\langle ., . \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ and a set of linear independent vectors $\{b_i\}_{i=0}^{k-1}$ in $\mathcal{V}$, then the associated Gram matrix $M_k \in \mathbb{C}^{k \times k}$ is

$$M_k := \begin{bmatrix} \langle b_0, b_0 \rangle & \langle b_1, b_0 \rangle & \dots & \langle b_{k-1}, b_0 \rangle \\ \langle b_0, b_1 \rangle & \langle b_1, b_1 \rangle & \dots & \langle b_{k-1}, b_1 \rangle \\ \vdots & \vdots & & \vdots \\ \langle b_0, b_{k-1} \rangle & \langle b_1, b_{k-1} \rangle & \dots & \langle b_{k-1}, b_{k-1} \rangle \end{bmatrix}. \tag{2.1}$$

This matrix is Hermitian positive-definite thanks to the properties of an inner product. For the linear functionals in this text a Gram matrix can be defined in the following way. Consider the Euclidean inner product $\langle ., . \rangle_E$ and two sets of linearly independent vectors $\{x_i\}_{i=1}^k$ and $\{y_j\}_{j=1}^k$, with $x_i, y_i \in \mathbb{C}^m$ for $i = 1, \dots, k$. Then the associated Gram matrix is

$$M_k := \begin{bmatrix} \langle x_0, y_0 \rangle_E & \langle x_1, y_0 \rangle_E & \dots & \langle x_{k-1}, y_0 \rangle_E \\ \langle x_0, y_1 \rangle_E & \langle x_1, y_1 \rangle_E & \dots & \langle x_{k-1}, y_1 \rangle_E \\ \vdots & \vdots & & \vdots \\ \langle x_0, y_{k-1} \rangle_E & \langle x_1, y_{k-1} \rangle_E & \dots & \langle x_{k-1}, y_{k-1} \rangle_E \end{bmatrix}, \tag{2.2}$$

which is, in general, an indefinite matrix.
For the linear functional $\mathcal{L}\{.\} : \mathcal{P} \to \mathbb{C}$ and two sets of linearly independent polynomials

$\{p_i\}_{i=0}^{k-1}$ and $\{q_j\}_{j=0}^{k-1}$, with $p_i, q_i \in \mathcal{P}_i$, the associated Gram matrix is

$$M_k := \begin{bmatrix} \mathcal{L}\{p_0q_0\} & \mathcal{L}\{p_1q_0\} & \dots & \mathcal{L}\{p_{k-1}q_0\} \\ \mathcal{L}\{p_0q_1\} & \mathcal{L}\{p_1q_1\} & \dots & \mathcal{L}\{p_{k-1}q_1\} \\ \vdots & \vdots & & \vdots \\ \mathcal{L}\{p_0q_{k-1}\} & \mathcal{L}\{p_1q_{k-1}\} & \dots & \mathcal{L}\{p_{k-1}q_{k-1}\} \end{bmatrix}. \tag{2.3}$$

Note that at least $2k-2$ moments are required to define $\mathcal{L}$ for this Gram matrix.

### LR factorization

The LR factorization of a Gram matrix gives rise to (bi)orthogonal vectors. First the LR factorization must be introduced. Let $a_{i,j}$ denote the entry on row $i$ and column $j$ of a matrix $A \in \mathbb{C}^{m\times m}$.
A *lower triangular matrix* $L \in \mathbb{C}^{m\times m}$ satisfies $l_{i,j} = 0$ for $i < j$ and an *upper triangular matrix* $R \in \mathbb{C}^{m\times m}$ satisfies $r_{i,j} = 0$ for $i > j$.
The LR factorization of $A \in \mathbb{C}^{m\times m}$ is

$$A = LR,$$

where $L \in \mathbb{C}^{m\times m}$ is a nonsingular lower triangular and $R \in \mathbb{C}^{m\times m}$ is a nonsingular upper triangular matrix.
The LR factorization does not always exist. Lemma 2.1 provides a necessary and sufficient condition for the existence of this factorization. The *principal leading submatrix* of size $i$ of some matrix $A \in \mathbb{C}^{m\times m}$ is denoted by $A^{(i)} \in \mathbb{C}^{i\times i}$.

**Lemma 2.1** (Existence of LR factorization [105])**.** *Consider $A \in \mathbb{C}^{m\times m}$ and suppose rank$(A) = m$. If $\det\left(A^{(i)}\right) \neq 0$, $i = 1, \dots, m$, then the LR factorzation of $A$ exists.*

A matrix $A$ of which all principal leading submatrices are nonsingular is called *strongly nonsingular* or *quasi-definite*. The LR factorization is not unique, by normalization it can be made unique. Normalization is done by using a *diagonal matrix* $D \in \mathbb{C}^{m\times m}$, with diagonal elements $d_i$, $i = 1, 2, \dots, m$, denoted as $D = \text{diag}(\{d_i\}_i)$. A unit (upper) lower triangular matrix is a (upper) lower triangular matrix with all diagonal elements equal to 1.
The LDR factorization of $A \in \mathbb{C}^{m\times m}$ is

$$A = LDR$$

with $L \in \mathbb{C}^{m\times m}$ a unit lower triangular matrix, $R \in \mathbb{C}^{m\times m}$ a unit upper triangular matrix and $D \in \mathbb{C}^{m\times m}$ a nonsingular diagonal matrix. The LDR factorization of a strongly nonsingular matrix is unique. The LR factorization of a Hermitian positive-definite matrix $A$ can be written as $A = R^H R$, with $R$ an upper triangular matrix with nonnegative diagonal entries. This factorization is called the *Cholesky decomposition*.

Lemma 2.2 states that sets can be orthogonalized using a factorization of the associated Gram matrix $M$.

**Lemma 2.2** (Biorthonormal vectors via Gram matrix factorization [51])**.** *Let* $X = \begin{bmatrix} x_0 & x_1 & \dots & x_{k-1} \end{bmatrix}$ *and* $Y = \begin{bmatrix} y_0 & y_1 & \dots & y_{k-1} \end{bmatrix}$ *denote matrices containing linearly independent vectors in some vector spaces* $\mathcal{V}$ *and* $\mathcal{W}$*, respectively. Let* $M_k \in \mathbb{C}^{k\times k}$ *be the associated Gram matrix generated by an inner product or a linear functional. Suppose* $M_k$ *is strongly nonsingular and its LR factorization is* $M_k = L_k R_k$*. Then these factors allow the construction of biorthonormal sets of vectors* $\{v_i\}_i, \{w_j\}_j$ *satisfying, for* $l = 0, 1, \dots, k-1$*,*

$$span\{v_0, \dots, v_l\} = span\{x_0, \dots, x_l\}, \quad span\{w_0, \dots, w_l\} = span\{y_0, \dots, y_l\}.$$

The LR factorization can be used as follows:

- For $M_k$ in (2.1), the Cholesky factorization is $M_k = R_k^H R_k$ and
$$\begin{bmatrix} v_0 & v_1 & \dots & v_{k-1} \end{bmatrix} := \begin{bmatrix} b_0 & b_1 & \dots & b_{k-1} \end{bmatrix} R_k^{-1}$$
forms an orthonormal set of vectors with respect to $\langle ., . \rangle$.
- For $M_k$ in (2.2), the LR factorization is $M_k = L_k R_k$ and
$$\begin{bmatrix} v_0 & v_1 & \dots & v_{k-1} \end{bmatrix} := \begin{bmatrix} x_0 & x_1 & \dots & x_{k-1} \end{bmatrix} R_k^{-1},$$
$$\begin{bmatrix} w_0 & w_1 & \dots & w_{k-1} \end{bmatrix} := \begin{bmatrix} y_0 & y_1 & \dots & y_{k-1} \end{bmatrix} L_k^{-H}$$
form a pair of biorthonormal sets of vectors with respect to $\langle ., . \rangle_E$.
- For $M_k$ in (2.3), the LR factorization is $M_k = L_k R_k$ and
$$\begin{bmatrix} v_0 & v_1 & \dots & v_{k-1} \end{bmatrix} := \begin{bmatrix} p_0 & p_1 & \dots & p_{k-1} \end{bmatrix} R_k^{-1},$$
$$\begin{bmatrix} w_0 & w_1 & \dots & w_{k-1} \end{bmatrix} := \begin{bmatrix} q_0 & q_1 & \dots & q_{k-1} \end{bmatrix} L_k^{-\top}$$
form a pair of biorthonormal sets of vectors with respect to $\mathcal{L}\{.\} : \mathcal{P} \to \mathbb{C}$.

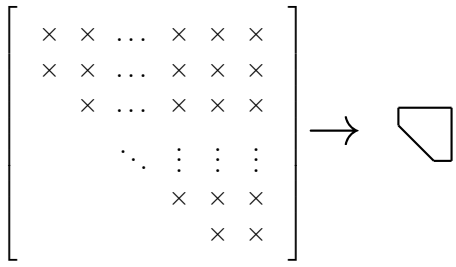
In finite precision arithmetic the above method to generate (bi)orthonormal vectors has two issues:

1. The Gram matrix tends to be ill-conditioned.
2. The computation of the LR factorization without pivoting is numerically unstable [183].

For the Hermitian positive definite Gram matrix $M_k$ in (2.1), an alternative procedure can be applied. Let $B_k \in \mathbb{C}^{m\times k}$, $m \geq k$, satisfy $M_k = B_k^H B_k$. Then the QR factorization of $B_k = Q_k R_k$, with $Q_k \in \mathbb{C}^{m\times k}$ having orthonormal columns and $R_k \in \mathbb{C}^{k\times k}$ upper triangular, generates the Cholesky factor $R_k$. That is $M_k = B_k^H B_k = R_k^H Q_k^H Q_k R_k = R_k^H R_k$.
Using the QR factorization is an improvement over the Cholesky decomposition, however the matrix $B_k$ also tends to be ill-conditioned in the context of Krylov subspaces.

## 2.3 Structured matrices

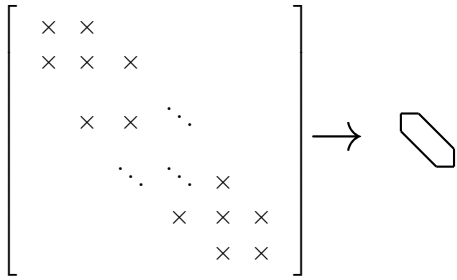
We study alternative methods to generate (bi)orthogonal vectors. These techniques rely heavily on structured matrices. Matrices can exhibit several kinds of structure. The structure of tridiagonal and Hessenberg matrices is determined by sparsity, i.e., certain entries are zero. Generic nonzero elements in a matrix are denoted by $\times$. A *Hessenberg matrix* $H_k \in \mathbb{C}^{k\times k}$ is a matrix with zeros below its first subdiagonal, i.e., $h_{i,j} = 0$ for $i > j + 1$. A Hessenberg matrix will be represented pictographically as



The *leftshift matrix* $Z_k \in \mathbb{C}^{k\times k}$ is a Hessenberg matrix with ones on its subdiagonal and zeros elsewhere

$$Z_k := \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix}.$$

A *tridiagonal matrix* $T_k \in \mathbb{C}^{k\times k}$ satisfies $t_{i,j} = 0$ for $|i-j| > 1$ and its pictograph is



Matrices with *low rank structure* are matrices which have submatrices that are of low rank. For example, the inverse of a nonsingular Hessenberg matrix has in its lower triangular part low rank structure. More precisely, any submatrix that can be taken in its lower triangular part has rank equal to 1. For details concerning rank structure we refer to the literature [61, 66–68, 106, 173, 174, 176].

Another type of structure, especially important for Gram matrices, is *displacement structure*.

**Definition 2.4** (Displacement structure [134])**.** *The displacement structure of a matrix $M$ is the image $L(M)$ of an appropriate linear* displacement operator *$L$ applied to a matrix $M$. The operator $L$ is appropriate when it reveals the structure of $M$. Two types of linear operators $L$ can be used, for a fixed pair of operator matrices $\{A, B\}$,*

- *Sylvester type, $L = \nabla_{A,B}$*

$$L(M) = \nabla_{A,B}(M) = AM - MB.$$

- *Stein type, $L = \Delta_{A,B}$*

$$L(M) = \Delta_{A,B}(M) = M - AMB.$$

For further details, we refer to [116, 134]. We will use the Sylvester type and we will refer to $A$ and $B$ as displacement operators. The *displacement rank* of $M$ with respect to the displacement operators $A, B$ is defined as the rank of the resulting matrix $L(M)$.

## 2.4 Matrix theory

This section groups all remaining concepts from matrix theory that have not been introduced above.

### Eigenvalue decomposition

A scalar $\lambda \in \mathbb{C}$ is called an *eigenvalue* of a matrix $A \in \mathbb{C}^{m\times m}$ if

$$Ax = \lambda x$$

is satisfied for a nonzero vector $x \in \mathbb{C}^m$, called the right *eigenvector* corresponding to $\lambda$. A left eigenvector of $A$ is a nonzero vector $y \in \mathbb{C}^{1\times m}$ that satisfies $yA = \lambda y$.
Let $I$ denote the identity matrix, its size will be clear from the context. An eigenvalue can also be characterized as a value $\lambda$ such that $\det(A - \lambda I) = 0$ is satisfied. The polynomial $p(z) = \det(A - zI)$ is called the *characteristic polynomial*.
The set of all eigenvalues of $A$ is called the *spectrum* of $A$ and is equal to

$$\sigma(A) = \{\lambda \in \mathbb{C} | \det(A - \lambda I) = 0\}.$$

Suppose that $A \in \mathbb{C}^{m\times m}$ has $m$ linearly independent eigenvectors $x_i$, with corresponding eigenvalue $\lambda_i$, then its *eigenvalue decomposition* is

$$A = X\Lambda X^{-1},$$

where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$ and $X := \begin{bmatrix} x_1 & \ldots & x_m \end{bmatrix}$. Such a matrix is called *diagonalizable*.
A *matrix pencil* consisting of matrices $A, B \in \mathbb{C}^{m\times m}$ is denoted as $(A, B) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$. The *generalized eigenvalues* of a matrix pencil $(A, B)$ are the scalars belonging to the spectrum of the pencil

$$\sigma(A, B) = \{\lambda = \alpha/\beta \in \overline{\mathbb{C}} | \det(\beta A - \alpha B) = 0\}, \text{ with } \overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}.$$

### Grade of a vector

For eigenvectors $x$ of a matrix $A \in \mathbb{C}^{m\times m}$ the vectors $x$ and $Ax$ are linearly dependent. For an arbitrary vector $v \in \mathbb{C}^m$ (which is not an eigenvector) this does not hold. However, there does exist a relationship between $v, Av, \ldots, A^g v$, i.e., for some nonzero coefficients $\alpha_i$

$$(\alpha_0 I + \alpha_1 A + \cdots + \alpha_{g-1} A^{g-1} + A^g)v = 0.$$

Let $g$ be the smallest integer for which this relation is satisfied. The polynomial $p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_{g-1} z^{g-1} + z^g$ is the *minimum polynomial* of $v$ with respect to

$A$. The degree $g$ of the minimum polynomial $p(z)$ is called the *grade* of $v$ with respect to $A$. Sometimes the degree of $v$ is denoted by $g_v$, if multiple vectors are considered at once. The grade will play an important role in the classification of breakdowns in the procedures discussed throughout this manuscript.
A *companion matrix* $Z$ for a monic polynomial $p(z) = z^m + \alpha_{m-1}z^{m-1} + \alpha_{m-2}z^{m-2} + \cdots + \alpha_1 z + \alpha_0$, with $\alpha_i \in \mathbb{C}$ is defined as

$$Z := \begin{bmatrix} 0 & & & & -\alpha_0 \\ 1 & 0 & & & -\alpha_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -\alpha_{m-2} \\ & & & 1 & -\alpha_{m-1} \end{bmatrix}.$$

The roots of the polynomial $p(z)$ correspond to the eigenvalues of the associated companion matrix $Z$, i.e., $p(\lambda) = 0$ if $\lambda \in \sigma(C)$.

### Function of a matrix

Let $f$ be a function defined on the spectrum of $A \in \mathbb{C}^{m\times m}$ and suppose this matrix is diagonalizable, i.e., $A = X\Lambda X^{-1}$, with $\Lambda = \mathrm{diag}(\{\lambda_i\}_i)$. Then

$$f(A) := Xf(\Lambda)X^{-1} = X\mathrm{diag}(\{f(\lambda_i)\}_i)X^{-1}.$$

See [104] for details.

# Chapter 3

# Structured matrices in polynomial Krylov subspaces

Computing eigenvalues of very large matrices and solving systems of equations are probably the most known uses of Krylov subspaces. If matrix-vector multiplication can be performed cheaply, then Krylov subspaces are a powerful tool to approximate the large matrix. Aside from the practical use for very large matrices, Krylov subspaces are a useful theoretical tool since they form a link between matrix theory and other disciplines or topics in mathematics, e.g., orthogonal polynomials, moment problems, quadrature and Padé approximation. These links allow translating some problems in these disciplines to problems in linear algebra. Numerical linear algebra offers many well-studied and efficient algorithms to solve these problems on a finite precision machine. Often the most effective methods for finite precision computation, in any of these disciplines, are methods from numerical linear algebra [70, 78]. The development and study of efficient algorithms starts with research on the matrix structures appearing in the problem to be solved. Studying the structure of matrices appearing in Krylov subspace computations allows to exploit this structure for efficiency, by using less memory and/or less operations, and to study the stability of the algorithm in detail. This chapter is dedicated to the matrix structures that are linked to Krylov subspaces. Polynomial Krylov subspaces and a suitable inner product on these spaces are introduced in Section 3.1. The following sections discuss different bases for polynomial Krylov subspaces, their theoretical and numerical properties and how to generate them. Section 3.2 discusses the most straightforward basis, a *Krylov basis*. This basis is, in general, not suited for numerical computation and two alternative types of bases are discussed with better numerical properties, orthogonal bases in Section 3.3, which possess inherent stability properties, and biorthogonal bases in Section 3.4, which

inherently lead to efficient numerical procedures. Important special cases, e.g., where biorthogonal bases reduce to an orthogonal basis, are discussed in Section 3.5. A conclusion is formulated in Section 3.6.

## 3.1 Polynomial Krylov subspaces

The origin of polynomial Krylov subspaces can be traced back to Krylov [118], who used them to compute the characteristic polynomial explicitly. The form in which they are still used today has its origin in several papers from the 1950s [3, 102, 121]. Polynomial Krylov subspaces $\mathcal{K}(A, v)$ for a given matrix $A \in \mathbb{C}^{m\times m}$ are subspaces formed by consecutive powers of $A$ multiplied by some vector $v \in \mathbb{C}^m$, called the *starting vector*,

$$\mathcal{K}(A, v) = \text{span}\{v, Av, A^2v, \dots\}.$$

The vectors $\{A^i v\}_i$ spanning a Krylov subspace become linearly dependent when the grade $g$ of the starting vector $v$ with respect to $A$ is reached. Krylov subspaces spanned by fewer than $g$ vectors are meaningful as well and will be the main object of study throughout this manuscript. Definition 3.1 defines polynomial Krylov subspaces.

**Definition 3.1** (Polynomial Krylov subspace [118])**.** *A polynomial Krylov subspace of dimension $k \leq g$ of a matrix $A \in \mathbb{C}^{m\times m}$ with starting vector $v \in \mathbb{C}^m$ is*

$$\mathcal{K}_k(A, v) = \textit{span}\{v, Av, A^2v, \dots, A^{k-1}v\}.$$

Krylov subspaces are nested, which is advantageous when computing (implicitly) with them.

**Property 3.1** (Nestedness of Krylov subpaces)**.** *Consider $\mathcal{K}_k(A, v)$, $k = 1, 2, \dots$, then*

$$\mathcal{K}_k \subseteq \mathcal{K}_{k+1}.$$

When $A\mathcal{K}_k(A, v) = \mathcal{K}_k(A, v)$, then $\mathcal{K}_{k+1}(A, v) = \mathcal{K}(A, v)$, and we have found an invariant subspace of $A$. This occurs when $k = g$, and thus the grade $g$ can be equivalently characterized as the value at which $A\mathcal{K}_g(A, v) = \mathcal{K}_g(A, v)$. Finding an invariant subspace is interesting in, e.g., the QR algorithm, where at the discovery of an invariant subspace, the problem can be decomposed into two smaller problems [183, p.525]. Note that the QR algorithm applied to Hessenberg matrices is intimately connected to Krylov subspaces [181].
Invariant subspaces contain a lot of information about $A$, however finding invariant subspaces can be very costly, especially for large $A$. Even if $\mathcal{K}_k(A, v)$ is not an invariant subspace, it still contains some information about $A$. In order to be able to compute with Krylov subspaces, a basis must be constructed. Choosing a different basis will

alter the representation of the underlying linear operator on the Krylov subspace. Bases which possess some orthogonality property are the most interesting. The notion of orthogonality requires a suitable inner product, this is the Euclidean inner product on $\mathbb{C}^m$

$$\mathbb{C}^m \times \mathbb{C}^m \to \mathbb{C} : (x, y) \mapsto \langle x, y \rangle_E := y^H x. \tag{3.1}$$

Now, for subspaces $\mathcal{K}_k(A, v)$ of the vector space $\mathbb{C}^m$ endowed with the Euclidean inner product $\langle ., . \rangle_E$ (3.1), three bases are discussed. A Krylov basis possesses no orthogonality properties, an orthogonal basis consists of vectors that are orthogonal to one another and biorthogonal bases consists of a pair of bases whose vectors are orthogonal to the vectors in the other basis with which they form the pair.

## 3.2 Krylov basis

The most straightforward basis for $\mathcal{K}_k(A, v)$ is the *Krylov basis*, which consists of the vectors $v, Av, \ldots A^{k-1}v$. The *Krylov matrix* $B_k$ is composed of these vectors, see Definition 3.2.

**Definition 3.2** (Krylov matrix)**.** *The matrix $B_k \in \mathbb{C}^{m \times k}$, $k \leq g \leq m$, associated with a Krylov subspace $\mathcal{K}_k(A, v)$ is called the Krylov matrix (of $\mathcal{K}_k$) if*

$$B_k = \begin{bmatrix} | & | & & | \\ v & Av & \ldots & A^{k-1}v \\ | & | & & | \end{bmatrix}.$$

An invariant subspace is found when $\text{rank}(B_k) < k$, i.e., the columns of $B_k$ become linearly dependent. Theorem 3.1 relates the rank of the Krylov matrix to the eigenvalues of a nondefective matrix $A$ and to the starting vector $v$. The grade $g$ plays an important role.

**Theorem 3.1** (Linear independence of Krylov matrix)**.** *Let $v \in \mathbb{C}^m$ be the starting vector and $A \in \mathbb{C}^{m \times m}$ be a nondefective matrix, i.e., it has the eigenvalue decomposition $A = X \Lambda X^{-1}$, where $\Lambda = diag(\lambda_1, \ldots, \lambda_m)$. Consider $\mathcal{K}_k(A, v)$ and corresponding Krylov matrix $B_k \in \mathbb{C}^{m \times k}$. Then rank$(B_k) = \min\{k, g\}$, with $g$ the grade of $A$ with respect to $v$. Moreover $g \leq \min\{c, d\}$ for $c$ the number of nonzero elements in $X^{-1}v$ and $d$ the number of distinct eigenvalues.*

*Proof.* A similar derivation appeared in [63] for normal, square matrices, the following derivation is a generalization.

Substitute $A = X\Lambda X^{-1}$ in $B_k$ and set $w := X^{-1}v$

$$B_k = \begin{bmatrix} | & | & & | \\ v & Av & \dots & A^{k-1}v \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ v & X\Lambda X^{-1}v & \dots & (X\Lambda X^{-1})^{k-1}v \\ | & | & & | \end{bmatrix}$$

$$= X \begin{bmatrix} w_1 & \lambda_1 w_1 & \dots & \lambda_1^{k-1} w_1 \\ w_2 & \lambda_2 w_2 & \dots & \lambda_2^{k-1} w_2 \\ \vdots & \vdots & & \vdots \\ w_m & \lambda_m w_m & \dots & \lambda_m^{k-1} w_1 \end{bmatrix}$$

$$= X \underbrace{\begin{bmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_m \end{bmatrix}}_{=:W} \underbrace{\begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^{k-1} \end{bmatrix}}_{=:V_k}.$$

Hence, $\operatorname{rank}(V_k) = \min\{k, d\}$. Clearly $\operatorname{rank}(W) = c$, since it is a diagonal matrix with elements $w_i$. Then

$$\operatorname{rank}(B_k) = \operatorname{rank}(XWV_k) = \operatorname{rank}(WV_k) = \min\{k, g\}$$
$$\leq \min\{\operatorname{rank}(W), \operatorname{rank}(V_k)\} = \min\{c, d\}.$$

To conclude, $\operatorname{rank}(B_k) = \min\{g, k\} \leq \min\{c, d, k\}$. □

Example 3.1 illustrates Theorem 3.1.

**Example 3.1.** *Consider a diagonal matrix $A$ with some multiple eigenvalues and a starting vector $v$ with a component along the direction of every eigenvector*

$$A = \begin{bmatrix} \imath & & & & & & \\ & 5 & & & & & \\ & & \imath & & & & \\ & & & 1 & & & \\ & & & & 2 & & \\ & & & & & 5 \end{bmatrix}, \qquad v = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

*The resulting Krylov matrix corresponding to $\mathcal{K}(A, v)$ is*

$$B = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & \imath & \dots & \imath^6 \\ 1 & 5 & \dots & 5^6 \\ 1 & \imath & \dots & \imath^6 \\ 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^6 \\ 1 & 5 & \dots & 5^6 \end{bmatrix}$$

*and has rank 4, equal to the number of distinct eigenvalues d.*
*Consider another starting vector which does not have a component along the direction of an eigenvector belonging to a simple eigenvalue*

$$\tilde{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

*Then the resulting Krylov matrix*

$$B = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & \imath & \dots & \imath^6 \\ 1 & 5 & \dots & 5^6 \\ 1 & \imath & \dots & \imath^6 \\ 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^6 \\ 1 & 5 & \dots & 5^6 \end{bmatrix} = \begin{bmatrix} 1 & \imath & \dots & \imath^6 \\ 1 & 5 & \dots & 5^6 \\ 1 & \imath & \dots & \imath^6 \\ 0 & 0 & \dots & 0 \\ 1 & 2 & \dots & 2^6 \\ 1 & 5 & \dots & 5^6 \end{bmatrix}$$

*has rank equal to 3, i.e., the case* $g < \min\{c, d\}$.

**Corollary 3.1.** *If A has distinct eigenvalues and v contains a component along the direction of every eigenvector* $x_i$ *in X, i.e.,* $v = \sum_{i=1}^{m} \alpha_i x_i$, *with* $\alpha_i \neq 0$ *for all i. Then, by Theorem 3.1,* $rank(B_k) = k$ *for* $k = 1, 2, \dots, m$ *and* $\mathcal{K}_m(A, v) = \mathbb{C}^m$.

Theorem 3.1 provides an expression for the rank of the Krylov matrix and thus for the maximum dimension of the Krylov subspace. To see what happens when the maximal rank, equal to $g$, is reached and why this is a good situation, we take a look at the recurrence relations underlying the Krylov matrix in Section 3.2.1.

### 3.2.1 Recurrence relation

The recurrence relation underlying the Krylov matrix is very simple and gives, when *breakdown* of the recurrence relation occurs, the companion matrix of the minimal polynomial of $A$. A breakdown of the recurrence relation takes place when the dimension of $\mathcal{K}_k(A, v)$ reaches the grade $g$, i.e., $k = g$. The recurrence relation is provided in Lemma 3.1 and Corollary 3.2 shows the connection to the companion matrix of the minimal polynomial of $A$ in case of a breakdown.

**Lemma 3.1** (Krylov matrix recurrence relation)**.** *Let $B_k$ be the Krylov matrix corresponding to $\mathcal{K}_k(A, v)$, $k < g$, then the following recurrence relation is satisfied*

$$AB_k = B_k Z_k + A^k v e_k^\top, \quad Z_k := \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix} \in \mathbb{C}^{k\times k}. \tag{3.2}$$

*Proof.* Straightforward. □

Equivalently, with $B_{k+1}$ the Krylov basis for $\mathcal{K}_{k+1}(A, v)$, the relation (3.2) can be written as

$$AB_k = B_{k+1}\underline{Z}_k, \quad \underline{Z}_k := \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \\ & & & & 1 \end{bmatrix} \in \mathbb{C}^{(k+1)\times k}.$$

The line underneath $Z_k$ indicates that the matrix is of size $(k + 1) \times k$ and implies that the term $A^k v e_k^\top$ is included in the basis.

**Corollary 3.2** (Companion matrix)**.** *Let $B_g$ denote the Krylov matrix for $\mathcal{K}_g(A, v)$, with $g$ the grade of $v$ with respect to $A$, then the following recurrence relation is satisfied*

$$AB_g = B_g Z, \quad Z := \begin{bmatrix} 0 & & & & \alpha_0 \\ 1 & 0 & & & \alpha_1 \\ & 1 & 0 & & \alpha_2 \\ & & \ddots & \ddots & \vdots \\ & & & 1 & 0 & \alpha_{g-2} \\ & & & & 1 & \alpha_{g-1} \end{bmatrix}, \tag{3.3}$$

*with $\alpha_0, \alpha_1, \ldots, \alpha_{g-1}$, the coefficients of the minimal polynomial of $v$ with respect to $A$ in the monomial basis.*

*Proof.* Since $\mathcal{K}_g(A, v)$ is invariant under multiplication by $A$, vector $A^g v \in \mathcal{K}_g(A, v)$ and can therefore be represented by

$$A^g v = \sum_{i=0}^{g-1} \alpha_i A^i v = B_g \begin{bmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_{g-1} \end{bmatrix}^\top.$$

□

The companion matrix $Z$ from Corollary 3.2 represents the operator $A$ restricted to the invariant subspace $\mathcal{K}_g(A, v)$. Hence, the eigenvalues of $Z$ are exact eigenvalues of $A$.
The *recurrence matrix* $Z_k$ contains little information about the matrix $A$ until there is a breakdown at $k = g$. To obtain a matrix $Z_k$, $k < g$, which contains more information, we can replace its last column by some projection of $A^k v$ onto $\mathcal{K}_k(A, v)$. This can be achieved by the Moore-Penrose pseudo-inverse $B_k^\dagger$ [90] of the Krylov matrix $B_k$. Since $B_k$ has linearly independent columns for $k < g$, $B_k^\dagger$ is the left inverse of $B_k$. Therefore we have, from Lemma 3.1, $B_k^\dagger A B_k = Z_k + B_k^\dagger A^k v e_k^\top$. The matrix of interest is then

$$\widehat{Z}_k := Z_k + B_k^\dagger A^k v e_k^\top. \tag{3.4}$$

Hence, we have constructed a smaller matrix $\widehat{Z}_k \in \mathbb{C}^{k\times k}$ representing $A \in \mathbb{C}^{m\times m}$ in the subspace $\mathcal{K}_k(A, v)$. The pseudo-inverse can be interpreted as a combination of two consecutive projectors [20]. A more detailed discussion on the pseudo-inverse in relation to Krylov subspaces and minimization properties is available in the thesis of Güttel [98, Chapter 3].

### 3.2.2 Recurrence relation - inverse companion matrix

An alternative recurrence relation is provided. It seems to be merely a curiosity, however, it will prove to be useful in studying matrix structures of recurrence matrices for biorthogonal bases. The recurrence relation is stated in Lemma 3.2, and Corollary 3.3 shows the recurrence matrix when an invariant subspace is reached.

**Lemma 3.2** (Left Krylov matrix recurrence relation)**.** *Let $B_k$ be the Krylov matrix corresponding to $\mathcal{K}_k(A, v)$, $k < g$, with $g$ the grade of $v$ with respect to $A$. Then the following recurrence relation is satisfied*

$$AB_k\widetilde{Z}_k + ve_1^\top = B_k, \quad \widetilde{Z}_k = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} \in \mathbb{C}^{k\times k}. \tag{3.5}$$

*Proof.* Straightforward. □

**Corollary 3.3** (Inverse companion matrix)**.** *Assume that $A$ is nonsingular and that $k = g$ in Lemma 3.2. Then the recurrence relation*

$$AB_g\widetilde{Z} = B_g$$

*is satisfied, where $\widetilde{Z}$ is the inverse of the companion matrix of the minimal polynomial $p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_{g-1} z^{g-1} + z^g$ of $v$ with respect to $A$.*

*Proof.* The assumption that $A$ is nonsingular guarantees that the companion matrix $Z$ satisfying $AK_g = K_g Z$ is nonsingular. Then $AK_g Z^{-1} = K_g$ and clearly $\widetilde{Z} = Z^{-1}$. □

The inverse companion matrix $\widetilde{Z}$ has a particular rank structure in its lower triangular part. It is a so-called *semiseparable matrix*: any (rectangular) submatrix that can be taken in the lower triangular part, including the diagonal, has rank $\leq 1$. For more information on the specific rank structure exhibited by $\widetilde{Z}$, we refer to literature [68,176]. We will study this rank structure via a particular factorization, the *core factorization*, introduced in the section on orthogonal bases for $\mathcal{K}_k(A, v)$, Section 3.3. The discussion is postponed since the orthogonal basis allows for a more natural study of the core factorization.

Again, the Moore-Penrose pseudo-inverse can be used to obtain intermediate matrices which contain some information of $A$. From the recurrence relation in Lemma 3.2 and under the assumption that $A$ is nonsingular it follows that

$$(AB_k)^\dagger (AB_k)\widetilde{Z}_k + (AB_k)^\dagger v e_1^\top = (AB_k)^\dagger B_k$$

$$\widetilde{Z}_k + (AB_k)^\dagger v e_1^\top = (AB_k)^\dagger B_k$$

where we used that $B_k$ has linearly independent columns and $A$ is nonsingular such that the equality $(AB_k)^\dagger (AB_k) = I$ is valid. The matrix of interest then is

$$\widetilde{Z}_k + (AB_k)^\dagger v e_1^\top. \tag{3.6}$$

Note that this matrix has the same lower triangular rank structure as $\widetilde{Z}$.

### 3.2.3 Numerical procedure

Mathematically Equation (3.4) and Equation (3.6) provide a projection of $A$ onto a Krylov subspace. However recurrence relation (3.5) requires the pseudo-inverse of $AB_k$, while it might be too expensive to compute the inverse of $A$. Recurrence relation (3.3) needs only the pseudo-inverse of $B_k$. The columns of $B_k$ are the vectors which are used in the power method and, as is well-known, the vectors in the power sequence $v, Av, A^2v, \ldots$ start to align with the eigenvector belonging to the dominant eigenvalue [90]. This can lead to numerical linear dependence, and will certainly lead to a larger condition number as $k$ increases [11, 90]. For more details we refer to the paper by Beckermann [11], where the condition number for Krylov matrices for Hermitian matrices $A$ is shown to increase exponentially in some cases. This suggests

that one should try to find a numerically better behaved basis: an orthogonal basis and biorthogonal bases.

## 3.3 Orthogonal basis

Orthogonal bases are more suited than Krylov bases for computation on a finite precision machine. A nested orthogonal basis $Q_k = \begin{bmatrix} q_0 & q_1 & \dots & q_{k-1} \end{bmatrix} \in \mathbb{C}^{m \times k}$ for $\mathcal{K}_k(A, v)$ spans the nested subspaces

$$\text{span}\{Q_i\} := \text{rank}\{q_0, q_1 \dots, q_{i-1}\} = \mathcal{K}_i(A, v)$$

and satisfies the orthogonality property

$$q_i \perp \mathcal{K}_i(A, v), \quad i = 2, 3, \dots k-1.$$

An orthonormal basis is an orthogonal basis with the normalization

$$\langle q_i, q_i \rangle_E = 1, \quad i = 0, 1, \dots, k-1$$

and for this case the orthogonality condition can be compactly denoted as

$$Q_k^H Q_k = I.$$

Since a Krylov matrix $B_k$, with $k < g$, has linearly independent columns, the Gram-Schmidt orthogonalization procedure can be applied to these columns. The result is a matrix with orthonormal columns. The matrix version of the Gram-Schmidt procedure is the QR-factorization, so we have

$$B_k = Q_k R_k, \quad \text{with } Q_k^H Q_k = I \text{ and } R_k \text{ upper triangular.}$$

Using the QR factorization of $B_k$ the recurrence relation for nested orthonormal bases of Krylov subspaces is derived in Section 3.3.1. The recurrence matrix contains all recurrence coefficients and will exhibit a particular structure, namely Hessenberg structure. Section 3.3.2 introduces a factorization using *core transformations* to factor this Hessenberg matrix. This factorization allows us to obtain an alternative recurrence relation in Section 3.3.3, which is the equivalent of the recurrence relation in Section 3.2.2. The structure of the recurrence matrix in this alternative recurrence relation is useful to study the matrix structures in Section 3.4.

### 3.3.1 Recurrence relation

The recurrence relation for an orthonormal basis $Q_k$ for $\mathcal{K}_k(A, v)$ is provided in Lemma 3.3. Using the QR factorization of the Krylov basis it can be derived starting from the Krylov recurrence relation (3.2).

**Lemma 3.3** (Orthonormal Krylov basis recurrence relation)**.** *Consider a nested orthonormal basis $Q_{k+1} \in \mathbb{C}^{m\times(k+1)}$ for $\mathcal{K}_{k+1}(A, v)$, $k < g$. Then a Hessenberg matrix $\underline{H}_k \in \mathbb{C}^{(k+1)\times k}$ exists such that*

$$AQ_k = Q_{k+1}\underline{H}_k = Q_k H_k + h_{k+1,k} q_k e_k^\top, \tag{3.7}$$

*where $q_k = Q_{k+1}e_{k+1}$.*

*Proof.* This proof is inspired by proofs from the literature [90, 169]. First note that a nested orthonormal basis for $\mathcal{K}_k(A, v)$ must satisfy $\widetilde{Q}_k = B_k\widetilde{R}_k$, for some nonsingular upper triangular matrix $\widetilde{R}_k$. By the uniqueness of the QR decomposition, this is (essentially) the same orthogonal matrix as obtained by the QR decomposition of $B_k = Q_k R_k$, in other words $\widetilde{Q}_k = Q_k D$, for a unitary diagonal matrix $D$. Substitute the QR decomposition into (3.2)

$$\begin{aligned} AB_k &= B_k C_k + A^k v e_k^\top \\ AQ_k R_k &= Q_k R_k C_k + A^k v e_k^\top \\ AQ_k &= Q_k \underbrace{R_k C_k R_k^{-1}}_{\hat{H}_k} + A^k v e_k^\top R_k^{-1} = Q_k \hat{H}_k + \frac{1}{r_{kk}} A^k v e_k^\top \end{aligned}$$

The term $A^k v = B_{k+1}e_{k+1}$ can be written in terms of the orthonormal basis spanned by columns of $Q_{k+1} = B_{k+1}R_{k+1}^{-1}$: let

$$R_{k+1} = \begin{bmatrix} R_k & \tilde{r} \\ 0 & r_{k+1,k+1} \end{bmatrix}$$

then

$$A^k v = Q_{k+1}R_{k+1}e_{k+1} = Q_k\tilde{r} + r_{k+1,k+1}q_k.$$

Then substitution in the above recurrence relation results in

$$AQ_k = Q_k\hat{H}_k + \frac{1}{r_{kk}}\left(Q_k\tilde{r} + r_{k+1,k+1}q_k e_k^\top\right) = Q_k H_k + \frac{r_{k+1,k+1}}{r_{kk}} q_k e_k^\top,$$

where $H_k := \hat{H}_k + \frac{1}{r_{kk}}Q_k\tilde{r}$. □

The orthonormal basis $Q_k$ for $\mathcal{K}_k(A, v)$ can be obtained by applying the Gram-Schmidt orthogonalization procedure to the Krylov matrix $B_k$ or, equivalently, computing its QR decomposition. The associated recurrence matrix $H_k$ has Hessenberg structure. Since these imply the construction of the ill-conditioned Krylov matrix $B_k$, they are not suited for numerical computation. A better method to construct $Q_k$ and $H_k$ in

finite precision is the Arnoldi iteration [3] discussed in Section 3.3.4. The Arnoldi iteration is in essence the Gram-Schmidt procedure applied to well-chosen vectors. Projection of $A$ onto $\mathcal{K}_k(A, v)$ with the orthogonal basis $Q_k$ does not need to be explicitly computed if the recurrence matrix $H_k$ is available, since $Q_k^H Q_k = I$, $Q_k^H q_k = 0$ and therefore

$$Q_k^H A Q_k = Q_k^H Q_k H_k + Q_k^H q_k e_k^\top = H_k. \tag{3.8}$$

That is, the recurrence matrix $H_k$ is the orthogonal projection of $A$ onto $\mathcal{K}_k(A, v)$.

#### Normal matrices

The Hessenberg matrix $H_k$ in the orthogonal projection (3.8) inherits some properties of the matrix $A$:

- if $A^H = A$, then $H_k^H = H_k$, i.e., a Jacobi matrix,
- if $A^H A = I$ and $k = m$, then $H_m^H H_m = I$, i.e., a unitary Hessenberg matrix [36, 155].

### 3.3.2 Core factorization

A natural tool to study the structures of the matrices appearing in this manuscript is the core factorization. The core factorization is a QR factorization where the unitary matrix $Q$ is represented as a product of core transformations.

**Definition 3.3** (Core transformations)**.** *A core transformation matrix $C_i \in \mathbb{C}^{m\times m}$ is a unitary matrix of the form*

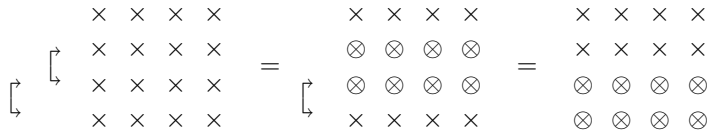
$$C_i = \begin{bmatrix} I_{i-1} & & & \\ & \times & \times & \\ & \times & \times & \\ & & & I_{m-i-1} \end{bmatrix}, \tag{3.9}$$

*where $I_k$ denotes the identity matrix of size $k \times k$.*

Core transformations are essentially $2 \times 2$ matrices, since their only active part is a $2 \times 2$ diagonal block. The parameter $i$ in $C_i$ indicates where, on the diagonal, the active block appears, and $\mathfrak{C}_i$ denotes the class of all these core transformations.

The pictographic notation for $C_i \in \mathfrak{C}_i$ is $\begin{smallmatrix}\cdot\\ \cdot\end{smallmatrix}$, with the top arrow pointing to row $i$ and the bottom arrow pointing to row $i + 1$. Example 3.2 shows the action of core transformations and illustrates how the pictographic notation is used.

**Example 3.2.** *Consider $A \in \mathbb{C}^{4\times 4}$ and core transformations $C_i \in \mathfrak{C}_i$. The action of core transformations will be visualized, an $\times$ indicates a generic nonzero element and $\otimes$ an element altered by a core transformation. Consider premultiplication of $A$ with $C_3$ and $C_2$, i.e., $C_3C_2A$, pictographically this is*



*For postmultiplication $AC_3C_2$ we have*

As Example 3.2 illustrates, premultiplication with $C_i$ only affects the $i$th and $(i+1)$st rows of a matrix and postmultiplication the $i$th and $(i+1)$st columns.
The unitary matrix of the QR-factorization of a Hessenberg matrix can be represented in terms of core transformations. Moreover, the structure in the lower triangular part of a Hessenberg matrix, the nonzero subdiagonal, is reflected in the core transformations. Lemma 3.4 states this more formally.

**Lemma 3.4** (Core factorization of Hessenberg matrices)**.** *Consider a Hessenberg matrix $H_k \in \mathbb{C}^{k\times k}$, the unitary matrix in the QR factorization of $H_k$ can be expressed as $H_k = C_1C_2 \dots C_{k-1}R$, where $C_i \in \mathfrak{C}_i$.*

The product of core transformations $\prod_i C_i$ with strictly increasing parameter $i$ and an upper triangular $R$ constitute the core factorization of a Hessenberg matrix. A visualization of this core factorization and resulting Hessenberg matrix is shown in Figure 3.1.

The product of core transformations contains all information on the lower triangular structure of the resulting matrix. We will refer to the product of core transformations with strictly increasing parameter as a *descending pattern*, this name arises from the

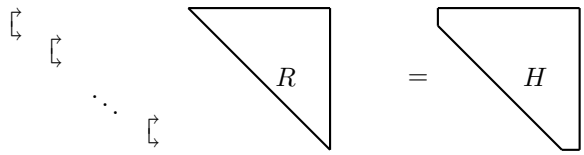
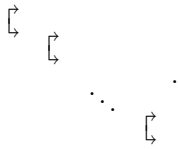



Figure 3.1: Core factorization $\left(\prod_{i=1}^{k-1} C_i\right) R$ of a Hessenberg matrix $H \in \mathbb{C}^{k\times k}$

visual representation

$$C_1 C_2 \dots C_{k-1} =$$



Note that for $k \leq g$ the Hessenberg matrix $H_k$ appearing in the (3.7) is a *proper* (also called unreduced) Hessenberg matrix, i.e., a Hessenberg matrix with only nonzero elements on its subdiagonal. The core factorization of proper Hessenberg matrices consists of a descending pattern of nontrivial core transformations, i.e., $H = \left(\prod_{i=1}^{k-1} C_i\right) R$, with $C_i \neq I$.
Next we consider the core factorization of the leftshift matrix appearing in the Krylov recurrence relation. Afterwards, in Section 3.2.2, the core factorization is used to derive an alternative representation of the recurrence matrix for an orthogonal basis.

### Leftshift matrix

The recurrence matrices associated with Krylov bases can also be factorized using core transformations. Consider $Z_k$ from Lemma 3.1 and let $\dot{C} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then

$$Z_k = \begin{bmatrix} \dot{C} & \\ & I_{k-2} \end{bmatrix} \begin{bmatrix} 1 & & \\ & \dot{C} & \\ & & I_{k-3} \end{bmatrix} \cdots \begin{bmatrix} I_{k-3} & & \\ & \dot{C} & \\ & & 1 \end{bmatrix} \begin{bmatrix} I_{k-2} & & \\ & 0 & 0 \\ & 1 & 0 \end{bmatrix}.$$

The last matrix in the product is not a core transformation, if we consider $\widehat{Z}_k$ from (3.4), then the last matrix is also a core transformation.

### 3.3.3 Alternative recurrence relation - inverse Hessenberg matrix

The core factorization facilitates the study of structures in matrices. For example the fact that a Hessenberg matrix corresponds to a descending pattern. Starting from the recurrence relation (3.7) with a Hessenberg recurrence matrix we derive an alternative form of the recurrence matrix with an interesting structure. This alternative form is obtained by moving the recurrence matrix to the left-hand side of the recurrence relation. To do so, manipulations of the core transformations are applied. The *transfer through property* formulated in Lemma 3.5 will allow a characterization of the structure of the alternative recurrence matrix.

**Lemma 3.5** (Transfer through property [172])**.** *A pattern (of core transformations) can be transferred through a nonsingular upper triangular matrix $R$ without altering the pattern. That is, for an upper triangular $R \in \mathbb{C}^{m\times m}$ and $C_i \in \mathfrak{C}_i \subset \mathbb{C}^{m\times m}$ there exist an upper triangular $\widetilde{R} \in \mathbb{C}^{m\times m}$ $\widetilde{C}_i \in \mathfrak{C}_i \subset \mathbb{C}^{m\times m}$ such that*

$$\left(\prod_{\sigma(i)} C_i\right) R = \widetilde{R} \left(\prod_{\sigma(i)} \widetilde{C}_i\right),$$

*where $\sigma(i)$ is some permutation of the sequence $i = 1, 2, \ldots, m-1$.*

Example 3.3 clarifies the statement in Lemma 3.5.

**Example 3.3.** *Consider core transformations $C_i, \widetilde{C}_i \in \mathfrak{C}_i$ and nonsingular upper triangular matrices $R, \widetilde{R}$. Then $C_1C_3C_2C_4R = \widetilde{R}\widetilde{C}_1\widetilde{C}_3\widetilde{C}_2\widetilde{C}_4$. The matrices involved will generally change (its elements), but the pattern, i.e., the mutual ordering of the core transformations (and therefore the structure of the resulting matrix) remains the same. This is shown in Figure 3.2.*

Figure 3.2: Illustration of transfer through property of the pattern $C_1C_3C_2C_4$.

Lemma 3.6 states an obvious result which is used in Lemma 3.7.

**Lemma 3.6** (Inverse of core transformations)**.** *The class of core transformations $\mathfrak{C}_i$ is closed under inversion. More specifically, let $C_i \in \mathfrak{C}_i$ then $C_i^{-1} = C_i^H \in \mathfrak{C}_i$.*

*Proof.* Follows from the fact that $C_i$ is unitary. □

**Lemma 3.7** (Inverse Hessenberg recurrence matrix)**.** *Consider a nested orthogonal basis $Q_{k+1} \in \mathbb{C}^{m\times(k+1)}$ for $\mathcal{K}_{k+1}(A, v)$, $k < g$. The matrix $Q_k$ satisfies*

$$AQ_k H_k^{inv} = Q_{k+1} \begin{bmatrix} & I_k & \\ \times & \dots & \times \end{bmatrix}, \tag{3.10}$$

*where $H_k^{inv} \in \mathbb{C}^{k\times k}$ has a core factorization with an ascending pattern, under the assumption that it is nonsingular.*

*Proof.* Consider the recurrence relation from Lemma 3.3, with $H_k$ in its core factorization and postmultiply and write $H_k^{-1}$ in its factored form:

$$AQ_k = Q_k H_k + h_{k+1,k} q_k e_k^\top = Q_k \left( \prod_{i=1}^{k-1} C_i \right) R_H + h_{k+1,k} q_k e_k^\top$$

$$AQ_k R_H^{-1} = Q_k \left( \prod_{i=1}^{k-1} C_i \right) + (e_k^\top R_H^{-1} e_k) h_{k+1,k} q_k e_k^\top$$

$$AQ_k R_H^{-1} C_{k-1}^H \dots C_2^H C_1^H = Q_k + (e_k^\top R_H^{-1} e_k) h_{k+1,k} q_k e_k^\top C_{k-1}^H \dots C_2^H C_1^H$$

$$AQ_k \underbrace{R_H^{-1} \left( \prod_{i=1}^{k-1} C_i \right)^H}_{=:H_k^{\mathrm{inv}}} = Q_k + E,$$

where $E := (e_k^\top R_H^{-1} e_k) h_{k+1,k} q_k e_k^\top C_{k-1}^H \dots C_2^H C_1^H$. Using Lemma 3.5 the core factorization of $H_k^{\mathrm{inv}}$ is manipulated to arrive at the usual form, see Figure 3.3. To complete the proof note that

$$Q_k + q_k \begin{bmatrix} \times & \dots & \times \end{bmatrix} = \begin{bmatrix} Q_k & q_k \end{bmatrix} \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ \times & \dots & \times \end{bmatrix}.$$

□

In the proof of Lemma 3.7 the structure of $H_k^{\mathrm{inv}}$ is derived without the need to explicitly compute the elements nor to rely on the nullity theorem. The nullity theorem links structured rank blocks in some nonsingular matrix $A$ to structured rank blocks in $A^{-1}$.

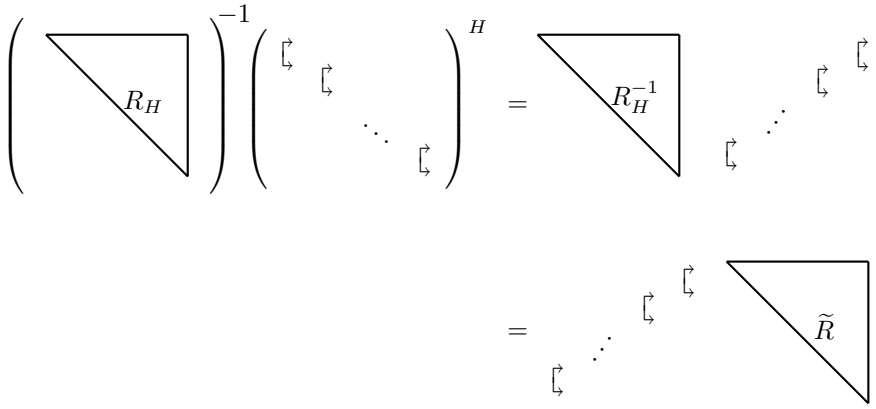
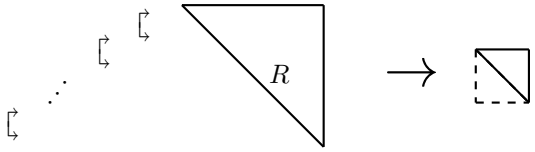



Figure 3.3: Core factorization manipulation.

For details on the nullity theorem, see the paper by Vandebril and Van Barel [173] and references therein. An alternative proof for Lemma 3.7 can be formulated which follows the proof of Lemma 3.3. Starting from the alternative recurrence relation for the Krylov basis provided in Lemma 3.2 and using the QR factorization of the Krylov matrix.
The lower triangular structure of the inverse Hessenberg recurrence matrix (3.10) will be important later, Definition 3.4 formalizes this structure and introduces a pictographic notation.

**Definition 3.4** (Inverse Hessenberg structure)**.** *A matrix $H_k$ with inverse Hessenberg structure, or short, inverse Hessenberg matrix is a matrix that allows a core factorization with a strictly ascending pattern of core transformations, i.e., $H_k = C_{k-1} \dots C_2 C_1 R$. A shorthand notation will be used: The full lines enclose elements*




*without rank structure and dashed lines enclose submatrices which exhibit low rank structure.*

Inverse Hessenberg matrices are also called Hessenberg-like matrices in the literature [175]. Hessenberg and inverse Hessenberg matrices both belong to the class of generalized Hessenberg matrices. These enjoy many interesting properties, such

as the product of a generalized Hessenberg and a nonsingular upper triangular matrix is again a generalized Hessenberg matrix and the class of generalized Hessenberg matrices is closed under inversion. For a full discussion on generalized Hessenberg matrices we refer to the literature [61, 68].
Note the essential difference between the recurrence matrix in Lemma 3.3, where a column is appended to the already computed recurrence matrix, and Lemma 3.7, where all the elements of the recurrence matrix change at each expansion of the Krylov subspace.

### 3.3.4 Arnoldi iteration

The Arnoldi iteration computes an orthogonal basis for $\mathcal{K}_k(A, v)$ and simultaneously provides the orthogonal projection of $A$ onto $\mathcal{K}_k(A, v)$. Essentially the Arnoldi iteration applies the Gram-Schmidt orthonormalization procedure to a well chosen representative for $A^k v$ in $\mathcal{K}_{k+1}(A, v)$ constructed from $A$ and a vector in the subspace $\mathcal{K}_k(A, v) = \operatorname{span}\{v, Av, \ldots, A^{k-1}v\}$ for which an orthogonal basis $Q_k$ is already available. For polynomials $p_i$ of exact degree $i$ we have

$$\mathcal{K}_k(A, v) = \operatorname{span}\{v, Av, \ldots, A^{k-1}v\} = \operatorname{span}\{p_0(A)v, p_1(A)v, \ldots, p_{k-1}(A)v\},$$

and $\mathcal{K}_k(A, v) \cup \{Ap_{k-1}(A)v\} = \mathcal{K}_{k+1}(A, v)$. Hence, any vector of the form $Ap_{k-1}(A)v$ is a suitable candidate to expand $\mathcal{K}_k(A, v)$ to $\mathcal{K}_{k+1}(A, v)$.
This is the idea behind a *continuation vector*. A continuation vector $p_{k-1}(A)v \in \mathcal{K}_k(A, v)$, $p_{k-1} \in \mathcal{P}_{k-1}$ of exact degree, is represented in some basis $B$ for $\mathcal{K}_k(A, v)$. For some $\tau_k \in \mathbb{C}^k$, with $e_k^\top \tau_k \neq 0$, the continuation vector is $B\tau_k = p_{k-1}(A)v$.
The orthogonal basis $Q_k$ is a well-conditioned basis for $\mathcal{K}_k$, so it is natural to take $B = Q_k$, i.e., $p_{k-1}(A)v = Q_k\tau_k$. In the Arnoldi iteration the orthonormal basis $Q_k$ for $\mathcal{K}_k(A, v)$ and continuation vector $\tau_k = e_k$ is chosen, thus $p_{k-1}(A)v = Q_k e_k = q_{k-1}$. Orthonormalizing $Ap_{k-1}(A)v = Aq_{k-1}$ provides $q_k$, thereby obtaining the orthonormal basis $Q_{k+1} = \begin{bmatrix} Q_k & q_k \end{bmatrix}$ for $\mathcal{K}_{k+1}(A, v)$. Algorithm 1 contains the Arnoldi iteration.

There is no need to explicitly project $A$ onto $\mathcal{K}_k(A, v)$ via the formula $Q_k^H A Q_k = H_k$, since the recurrence matrix $H_k$ is obtained as a by-product of the orthonormalization performed by the Arnoldi iteration.

The recurrence relation constructing an orthogonal basis for Krylov subspaces is, in general, a long recurrence relation. It requires all basis vectors to be saved in memory and at each step the new vector must be orthogonalized to all of them. This could become troublesome for very large matrices, especially when convergence is slow. Biorthogonal bases provide the means to alleviate this drawback. These will lead to short (three term) recurrence relations, which implies that a fixed number of basis vectors are needed to expand $\mathcal{K}_k$. There are however several complications in the use

**Algorithm 1** Arnoldi iteration [3]

**Input:** $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$, integer $k < g$
**Output:** Orthogonal $Q_{k+1} \in \mathbb{C}^{m\times k+1}$, Hessenberg matrix $\underline{H}_k \in \mathbb{C}^{(k+1)\times k}$ such that $AQ_k = Q_{k+1}\underline{H}_k$.
**procedure** ARNOLDI_ITERATION($A, v, k$)
    $q_0 = v/\|v\|$
    **for** $i = 1, 2, \ldots, k$ **do**
        $q_i = Aq_{i-1}$
        **for** $j = 1, 2, \ldots, i$ **do** ▷ Orthogonalization
            $h_{j,i} = \langle q_i, q_j\rangle_E$
            $q_i = q_i - h_{j,i}q_j$
        **end for**
        $h_{i+1,i} = \|q_i\|$
        $q_i = q_i/h_{i+1,i}$ ▷ Normalization
    **end for**
**end procedure**

of biorthogonal bases. These complications and the derivation of the short recurrence relation is the topic of the next section.

## 3.4 Biorthogonal bases

Short recurrence relations generating bases for polynomial Krylov subspaces require two bases: one forms a basis for a Krylov subspace generated with $A \in \mathbb{C}^{m\times m}$ and the other for a Krylov subspace generated with $A^H \in \mathbb{C}^{m\times m}$. If these bases are orthogonal with respect to each other, i.e., biorthogonal, then a pair of three term recurrence relations suffices to construct them.
Formally, we have two Krylov subspaces $\mathcal{K}_k(A, v)$ and $\mathcal{K}_k(A^H, w)$, where the matrices are obviously related and the starting vectors satisfy $\langle v, w\rangle_E \neq 0$. For these subspaces two bases $V_k, W_k \in \mathbb{C}^{m\times k}$ are constructed which are nested,

$$\operatorname{span}\{V_i\} := \operatorname{span}\{v_0, v_1, \ldots, v_{i-1}\} = \mathcal{K}_i(A, v) \quad \text{and}$$

$$\operatorname{span}\{W_i\} := \operatorname{span}\{w_0, w_1, \ldots, w_{i-1}\} = \mathcal{K}_i(A^H, w), \text{ for } i = 1, \ldots, k,$$

and satisfy the biorthonormality condition $\langle v_i, w_j\rangle_E = \delta_{ij}$, in matrix notation

$$W_k^H V_k = I.$$

The orthogonality can be expressed in terms of Krylov subspaces

$$v_i \perp \mathcal{K}_i(A^H, w) \text{ and } w_i \perp \mathcal{K}_i(A, v), \quad i = 1, 2, \ldots k-1.$$

Section 3.4.1 discusses breakdowns in the biorthogonalization of two sequences of vectors. A pair of three term recurrence relations for bases of Krylov subspaces is derived in Section 3.4.2, which can be equivalently represented by a tridiagonal matrix. Tridiagonal matrices allow for a factorization resembling the core factorization, but now using nonunitary triangular matrices, called *eliminators*, which are introduced in Section 3.4.3. The non-Hermitian Lanczos iteration, which constructs biorthogonal bases and the corresponding tridiagonal recurrence matrix is the topic of Section 3.4.4.

## 3.4.1 Breakdown

The situation for biorthogonalization procedures is more complicated than for orthogonalization, due to the possibility of a *serious breakdown*. This term is introduced below. Biorthonormal bases for two sequences of linearly independent vectors can be constructed via the two-sided Gram-Schmidt procedure. This procedure is applied to the columns of the Krylov matrices $B_k^V$ and $B_k^W$ related to $\mathcal{K}_k(A, v)$ and $\mathcal{K}_k(A^H, w)$, respectively,

$$B_k^V := \begin{bmatrix} v & Av & \dots & A^{k-1}v \end{bmatrix},$$

$$B_k^W := \begin{bmatrix} w & A^H w & \dots & (A^H)^{k-1}w \end{bmatrix}.$$

The resulting matrices $V_k = B_k^V R_k^V$ and $W_k = B_k^W R_k^W$, with nonsingular upper triangular matrices $R_k^V, R_k^W$, have biorthonormal columns, i.e., $W_k^H V_k = I$. Two types of breakdowns can occur while executing the two-sided Gram-Schmidt procedure. A lucky breakdown signals the discovery of an invariant subspace and can now occur for both spaces, $\mathcal{K}_k(A, v)$ and/or $\mathcal{K}_k(A^H, w)$. A new type of breakdown, a *serious breakdown*, is a breakdown where no invariant subspace is found.

### Lucky breakdown

A lucky breakdown is good news since an invariant subspace has been found. Two Krylov subspaces are used with respect to (possibly) distinct starting vectors $v, w$, therefore both the grade $g_v$ of $v$ with respect to $A$ and the grade $g_w$ of $w$ with respect to $A^H$ must be taken into account. If any of the two sequences forms an invariant subspace, i.e., $k = \min\{g_v, g_w\}$, then the procedure breaks down.
Suppose, without loss of generality, that $k = g_v$ and $g_w > g_v$, then $A\mathcal{K}_{g_v}(A, v) = \mathcal{K}_{g_v}(A, v)$ and $v_{g_v}$ will vanish. This implies that $\langle v_{g_v}, w_{g_v} \rangle_E = 0$ and the normalization cannot be achieved, thus the process breaks down. Example 3.4 illustrates such a lucky breakdown.

**Example 3.4.** *Consider* $A = diag(1, 2, 3, 4)$, $v = \begin{bmatrix} 1, 1, 0, 0 \end{bmatrix}^\top$ *and* $w = \begin{bmatrix} 1, 1, 1, 1 \end{bmatrix}^\top$. *Clearly,* $g_v = 2$ *and* $g_w = 4$. *When generating the sequences of biorthogonal*

*vectors for $\mathcal{K}(A, v)$ and $\mathcal{K}(A^H, w)$ via the two-sided Gram-Schmidt procedure, without normalization, we obtain:*

$$\{v_0, v_1, v_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -0.5 \\ 0.5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

$$\{w_0, w_1, w_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -0.5 \\ 0.5 \\ 1.5 \\ 2.5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 6 \end{bmatrix} \right\}.$$

*The vectors still satisfy $v_2 \perp span\{w_0, w_1\}$, $w_2 \perp span\{v_0, v_1\}$. However they do not satisfy $\langle v_2, w_2 \rangle_E \neq 0$ and the process breaks down, we cannot generate an appropriate $w_3$.*
*In case of orthonormalization, the vectors would only differ by multiplicative scaling until arriving at $v_2$. Since $v_2 = 0$, and therefore $\langle v_2, w_2 \rangle_E = 0$, the vectors $v_2$ and $w_2$ cannot be normalized.*

#### Serious breakdown

If $\langle v_j, w_j \rangle_E = 0$ for some $j$ and $v_j \neq 0$, $w_j \neq 0$, then we have encountered a *serious breakdown*. Serious because we have not found an invariant subspace, but the biorthonormalization procedure breaks down. Such breakdowns can occur because for $v_j \neq w_j$, the inner product $\langle v_j, w_j \rangle_E$ must not be positive. When $v_j = w_j$, then $\langle v_j, v_j \rangle_E \geq 0$, by the properties of an inner product, and positive for $v_j \neq 0$. Further remarks on breakdowns specific for the nonhermitian Lanczos iteration are provided in Section 3.4.4.

### 3.4.2 Three term recurrence relation

A pair of three term recurrence relations suffices to generate the nested biorthogonal bases $V_k := \begin{bmatrix} v_0 & v_1 & \dots & v_{k-1} \end{bmatrix}$ and $W_k := \begin{bmatrix} w_0 & w_1 & \dots & w_{k-1} \end{bmatrix}$. The pair of short recurrence relations imply more efficient procedures compared to the long recurrence relation for constructing an orthogonal basis. A derivation of the recurrence relations starting from the vectors in the sets $\{v_i\}, \{w_i\}$ and using the property of the inner product, $\langle Ax, y \rangle_E = \langle x, A^H y \rangle_E$ for any $x, y \in \mathbb{C}^m$, is often used. In Lemma 3.8 a derivation using matrix theory is provided, which leads to a shorter proof.

**Lemma 3.8** (Biorthogonal Krylov bases recurrence relations)**.** *Let $A \in \mathbb{C}^{m \times m}$, $v, w \in \mathbb{C}^m$, with $\langle v, w \rangle_E \neq 0$. Consider the Krylov subspaces $\mathcal{K}_k(A, v), \mathcal{K}_k(A^H, w)$,*

*with $k < \min\{g_v, g_w\}$. Then biorthonormal nested bases $V_k, W_k \in \mathbb{C}^{m\times k}$ for these subspaces satisfy the recurrence relations*

$$AV_k = V_k T_k + t_{k+1,k} v_k e_k^\top,$$

$$A^H W_k = W_k T_k^H + \bar{t}_{k,k+1} w_k e_k^\top,$$

*where $T_k \in \mathbb{C}^{k\times k}$ is a tridiagonal matrix.*

*Proof.* Since the biorthonormality of the bases $V_k, W_k$ with respect to the Euclidean inner product implies $W_k^H V_k = I$ and $v_k \perp \operatorname{span}\{W_k\}$, $w_k \perp \operatorname{span}\{V_k\}$, we have

$$W_k^H A V_k = W_k^H V_k T_k = T_k,$$

$$V_k^H A^H W_k = V_k^H W_k T_k^H = T_k^H.$$

The columns of $V_k$ form a nested basis for $\mathcal{K}_k(A, v)$, so there exists a nonsingular upper triangular matrix $R_k$ such that $V_k = B_k^V R_k$, where $B_k^V$ is the Krylov basis. By substituting this into the Krylov recurrence relation, one easily finds that $T_k$ must be a upper Hessenberg matrix. A similar argument is valid for $W_k$ and thus $T_k^H$ is upper Hessenberg and therefore $T_k$ is lower Hessenberg. Only a tridiagonal matrix is simultaneously upper-and lower Hessenberg. Hence, $T_k$ has tridiagonal structure. □

This lemma shows that biorthonormalizing a pair of bases for Krylov subspaces generated by $A$ and $A^H$ leads to structure in both the upper-and lower triangular part of the recurrence matrix $T_k$.
Next we provide a proof which uses the associated Gram matrix to relate the upper and lower triangular part of $T_k$ more explicitly to the lower triangular structure of the recurrence matrices appearing in the Krylov recurrence relations for $\mathcal{K}(A, v)$ and $\mathcal{K}(A^H, w)$. This proof makes use of the recurrence relation with the inverse Hessenberg matrix from Lemma 3.2. The idea of the alternative proof can be generalized to biorthogonal bases for rational Krylov subspace, discussed in Chapter 4.

*Proof: Alternative proof for Lemma 3.8.* The proof is provided for the case $k = \min\{g_v, g_w\}$. Assume without loss of generality that $k = g_w \le g_v$. Consider orthogonal bases $Q_k^V, Q_k^W$ for $\mathcal{K}_k(A, v), \mathcal{K}_k(A^H, w)$ in the form given in Lemma 3.3 and Lemma 3.7, respectively:

$$AQ_k^V = Q_k^V H_k^V + h_{k+1,k}^V q_{k+1}^V e_k^\top,$$

$$A^H Q_k^W H_k^W = Q_k^W + E.$$

Since $k = g_w$, $E = 0$. Let $M_k := (Q_k^W)^H Q_k^V$ denote the associated Gram matrix and assume that $M_k$ is strongly nonsingular. Factorize $M_k = L_k R_k$, with $L_k$ lower and

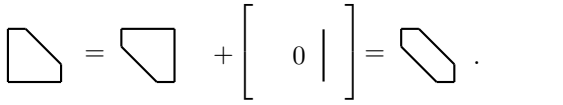
$R_k$ upper triangular. Form the new bases $V_k := Q_k^V R^{-1}$, $W_k := Q_k^W L^{-H}$. By Lemma 2.2 $V_k, W_k$ form nested biorthonormal bases for $\mathcal{K}_k(A, v), \mathcal{K}_k(A^H, w)$, respectively. Substitute this into the recurrence relations

$$\begin{cases} AQ_k^V & = Q_k^V H_k^V + h_{k+1,k}^V q_{k+1}^V e_k^\top \\ A^H Q_k^W H_k^W & = Q_k^W \end{cases}$$

$$\begin{cases} AV_k & = V_k R_k H_k^V R_k^{-1} + \frac{h_{k+1,k}^V}{r_{k+1,k+1}} q_{k+1}^V e_k^\top \\ A^H W_k L_k^H H_k^W L_k^{-H} & = W_k \end{cases}$$

$$\begin{cases} W_k^H AV_k & = R_k H_k^V R_k^{-1} + \frac{h_{k+1,k}^V}{r_{k+1,k+1}} W_k^H q_{k+1}^V e_k^\top \\ W_k^H AV_k & = L_k (H_k^W)^{-H} L_k^{-1} \end{cases}$$

Set $T_k := W_k^H AV_k$, then.

$T_k =$  .

For $k < \min\{g_v, g_w\}$ a similar derivation can be followed, a rank argument on the upper and lower triangular part of $T_k$ then results in $AV_k = V_k T_k + t_{k+1,k} v_{k+1} e_k^\top$ and $A^H W_k = W_k T_k^H + \bar{t}_{k,k+1} w_{k+1} e_k^\top$. □

### 3.4.3 Eliminator factorization

A nonunitary analogue to the core factorization is based on eliminators. Eliminators are essentially $2 \times 2$ matrices of upper or lower triangular form. Definition 3.5 introduces these eliminators together with a pictographical notation.

**Definition 3.5** (Eliminators)**.** *An upper eliminator $E_i^u \in \mathfrak{E}_i^u \subset \mathbb{C}^{m \times m}$ is an upper triangular matrix of the form*

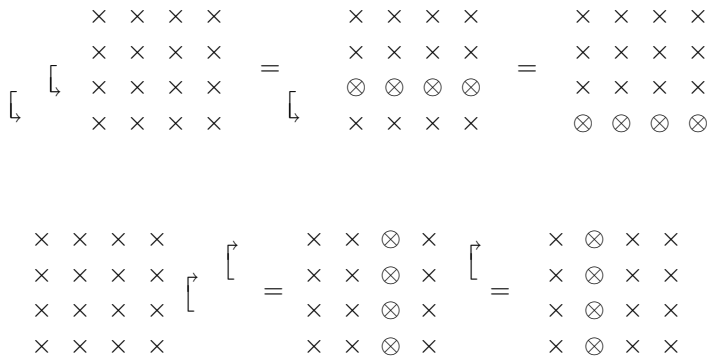
$$E_i^u = \begin{bmatrix} I_{i-1} & & & \\ & 1 & \times & \\ & & 1 & \\ & & & I_{m-i-1} \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \vec{} \end{bmatrix} \tag{3.11}$$

*and a lower eliminator $E_i^l \in \mathfrak{E}_i^l \subset \mathbb{C}^{m \times m}$ has lower triangular form*

$$E_i^l = \begin{bmatrix} I_{i-1} & & & \\ & 1 & & \\ & \times & 1 & \\ & & & I_{m-i-1} \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \\ \hookrightarrow \end{bmatrix} . \tag{3.12}$$

This definition of eliminators requires only one parameter, naturally the elements 1 can be replaced by parameters as well. That additional freedom might be useful in developing numerical procedures, however, for theoretical purposes there is no harm to fix them to 1. For clarity and ease of notation, we will do so. An obvious, but important, property of lower (upper) eliminators is that multiplication only influences one row or column and preserves upper (lower) rank structures. Example 3.5 illustrates the effect of eliminators and shows how the notation is used.

**Example 3.5.** *For some matrix $A \in \mathbb{C}^{m\times m}$, the rows altered by performing the premultiplication $E_3^l E_2^l M$ is and an example of postmultiplication is $M E_2^u E_1^u$*



Lemma 3.9 states that a strongly nonsingular tridiagonal matrix can be decomposed using these eliminators. The eliminator factorization is, in fact, a particular LR factorization, where $L$ and $R$ are represented as a product of eliminators.

**Lemma 3.9** (Eliminator factorization of tridiagonal matrix)**.** *Let $T \in \mathbb{C}^{m\times m}$ be a strongly nonsingular tridiagonal matrix. Then $T$ can be factorized as $T = \left(\prod_{i=1}^{m-1} E_i^l\right) D \left(\prod_{i=m-1}^{1} E_i^u\right)$, where $D$ is a diagonal matrix.*

*Proof.* Consider $T \in \mathbb{C}^{3\times 3}$ and set $\tilde{t}_{22} := t_{22} - \frac{t_{21}t_{12}}{t_{11}}$, then the factorization is

$$\begin{aligned}
\begin{bmatrix} t_{11} & t_{12} & \\ t_{21} & t_{22} & t_{23} \\ & t_{32} & t_{33} \end{bmatrix} &= \begin{bmatrix} 1 & & \\ \frac{t_{21}}{t_{11}} & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} t_{11} & & \\ & \tilde{t}_{22} & t_{23} \\ & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} 1 & \frac{t_{12}}{t_{11}} & \\ & 1 & \\ & & 1 \end{bmatrix} \\
&= \begin{bmatrix} \\ \hookrightarrow \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & \frac{t_{32}}{\tilde{t}_{22}} & 1 \end{bmatrix} \begin{bmatrix} t_{11} & & \\ & \tilde{t}_{22} & \\ & & \tilde{t}_{33} \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \frac{t_{23}}{\tilde{t}_{22}} \\ & & 1 \end{bmatrix} \begin{bmatrix} \\ \rightarrow \end{bmatrix} \\
&= \begin{bmatrix} \\ \hookrightarrow \end{bmatrix} \begin{bmatrix} \\ \hookrightarrow \end{bmatrix} \begin{bmatrix} t_{11} & & \\ & \tilde{t}_{22} & \\ & & \tilde{t}_{33} \end{bmatrix} \begin{bmatrix} \\ \rightarrow \end{bmatrix} \begin{bmatrix} \\ \rightarrow \end{bmatrix},
\end{aligned}$$

where $\tilde{t}_{33} := t_{33} - \frac{t_{32}t_{23}}{\tilde{t}_{22}}$. Already factorized principal leading submatrices are not affected by later eliminators. Hence, this process is immediately valid for tridiagonal matrices of arbitrary size. The diagonal terms $\tilde{t}_{i,i}$ can never vanish, which corresponds to $T$ being strongly nonsingular. The strongly nonsingular requirement follows from the fact that the eliminator factorization is an LR factorization. □

If the tridiagonal matrix is not strongly nonsingular, then Lemma 3.9 holds until the first singular principal leading submatrix is reached. Just as for the core factorization, the eliminator factorization can be used to study the structure of matrices.

### 3.4.4 Nonhermitian Lanczos iteration

Lanczos [121] proposed two iterations to compute eigenvalues (which he called latent roots) of matrices. An iteration for non-Hermitian matrices, which is discussed at present and an iteration for Hermitian matrices, which is discussed in Section 3.5. The Lanczos iteration is an instance of the two-sided Gram-Schmidt procedure applied to two (nested) sets of linearly independent vectors in the Krylov subspaces $\mathcal{K}(A, v)$ and $\mathcal{K}(A^H, w)$. These particular choices of Krylov subspaces, i.e., generated with $A$ and $A^H$, lead to short recurrence relations, as is shown in Section 3.4.2. Since we are generating spaces there is flexibility to choose a candidate vector for expansion, i.e., suitable vectors $\hat{v}, \hat{w}$ such that $\mathcal{K}_{k+1}(A, v) = \mathcal{K}_k(A, v) \cup \{\hat{v}\}$ and $\mathcal{K}_{k+1}(A^H, w) = \mathcal{K}_k(A^H, w) \cup \{\hat{w}\}$. Similarly to the Arnoldi iteration these candidate vectors will be represented as $\hat{v} = AV_k t_k^V$ and $\hat{w} = A^H W_k t_k^W$, where $V_k, W_k$ is a pair of biorthonormal bases for $\mathcal{K}_k(A, v)$ and $\mathcal{K}_k(A^H, w)$, respectively. The continuation vectors are chosen to be $t_k^V = t_k^W = e_k$ and the resulting algorithm is provided in Algorithm 2.

**Algorithm 2** Lanczos iteration [121, 148]

1: **Input:** $A \in \mathbb{C}^{m\times m}$, $v, w \in \mathbb{C}^m$, integer $k < \min\{g_v, g_w\}$
2: **Output:** Biorthogonal $W_{k+1}, V_{k+1} \in \mathbb{C}^{m\times k+1}$, tridiagonal matrix $\underline{T}_k \in \mathbb{C}^{(k+1)\times k}$ such that $W_k^H A V_k = T_k$.
3: **procedure** LANCZOS_ITERATION($A, v, w, k$)
4: $\eta := \langle v, w\rangle_E$
5: $v_0 := v/|\eta|^{1/2}$ & $w_0 := w/(\eta/|\eta|^{1/2})$
6: $v_{-1} := 0$ & $w_{-1} := 0$
7: **for** $i = 1, 2, \ldots, k$ **do**
8: $t_{i,i} := \langle Av_{i-1}, w_{i-1}\rangle_E$
9: $\hat{v}_i := Av_{i-1} - t_{i,i}v_{i-1} - t_{i-1,i}v_{i-2}$
10: $\hat{w}_i := A^H w_{i-1}\bar{t}_{i,i}w_{i-1} - t_{i,i-1}w_{i-2}$
11: $t_{i+1,i} := |\langle \hat{v}_i, \hat{w}_i\rangle_E|^{1/2}$
12: $t_{i,i+1} := \langle \hat{v}_i, \hat{w}_i\rangle_E / t_{i+1,i}$
13: $v_i := \hat{v}_i / t_{i+1,i}$
14: $w_i := \hat{w}_i / \bar{t}_{i,i+1}$
15: **end for**
16: **end procedure**

Note that the normalization, Step 11 and Step 12 in Algorithm 2, can be done in any way as long as $t_{i+1,i}t_{i,i+1} = \langle \hat{v}_i, \hat{w}_i\rangle_E$.

### Breakdowns and look ahead

Serious breakdowns in the non-Hermitian Lanczos procedure are well studied and we refer to the literature for the consequences of serious breakdowns and possible remedies [97]. Serious breakdowns in the Lanczos iteration can be classified as curable or incurable serious breakdowns. A curable breakdown can be resolved by using a look ahead procedure to 'jump over' the vectors that have $\langle v_j, w_j\rangle_E = 0$. An incurable breakdown is a serious breakdown that cannot be cured. At the occurrence of an incurable breakdown, the eigenvalues of the generated tridiagonal recurrence matrix are exact eigenvalues of the matrix $A$.

## 3.5 Special cases

Two important special cases occur when the matrix $A \in \mathbb{C}^{m\times m}$ is unitary $A^H A = I$ or Hermitian $A^H = A$. For a unitary matrix the *isometric Arnoldi iteration* [100,155,188] is developed, which we will not discuss in detail.
For Hermitian matrices the Arnoldi and Lanczos iteration reduce to the same iteration, the Hermitian Lanczos iteration. This implies a three term recurrence relation which generates an orthogonal basis. So it is as efficient as the biorthogonal procedure while still providing an orthogonal result, and thus is inherently more stable than a biorthogonal result. First we show that an orthonormal basis for $\mathcal{K}(A, v)$, with $A = A^H$, satisfies a three term recurrence relation and second the equivalent statement that a pair of biorthonormal bases for $\mathcal{K}(A, v)$ and $\mathcal{K}(A^H, w)$ reduces to a single orthonormal basis if $A^H = A$ and $v = w$. The resulting recurrence matrix is a Hermitian tridiagonal matrix and this is reflected in its eliminator factorization. Then the Hermitian Lanczos iteration is provided in its most used form, the form introduced by Paige [132].

### Three term recurrence relation

Lemma 3.10 states that for Hermitian matrices $A^H = A$ there exists a three term recurrence relation that generates an orthonormal basis for $\mathcal{K}(A, v)$. Two proofs are provided. The first starts from an orthonormal basis and shows that the Hessenberg matrix reduces to a Hermitian tridiagonal matrix. The second starts from biorthonormal bases and shows that the pair of three term recurrence relations reduces to a single three term recurrence relation. The second proof uses the same idea as a proof that will be used in the next chapter to show that for rational Krylov subspaces and a Hermitian matrix there also exist short recurrence relations for an orthonormal basis.

**Lemma 3.10** (Short orthogonal Krylov basis recurrence relation)**.** *Let $A \in \mathbb{C}^{m\times m}$ be Hermitian, i.e., $A^H = A$ and $v \in \mathbb{C}^m$. Consider $\mathcal{K}_k(A, v)$, with $k < g$. Then a nested orthonormal basis $Q_k \in \mathbb{C}^{m\times k}$ for $\mathcal{K}(A, v)$ satisfies the relation*

$$AQ_k = Q_k T_k + t_{k+1,k} q_k e_k^\top,$$

*where $T_k \in \mathbb{C}^{k\times k}$ is a Hermitian tridiagonal matrix and $q_k \in \mathcal{K}_{k+1}(A, v)$, $q_k \perp \mathcal{K}_k A, v$, $\|q_k\| = 1$.*

*Proof.* Use the recurrence relation, $AQ_k = Q_k H_k + h_{k+1,k} q_k e_k^\top$, from Lemma 3.3, and using $A^H = A$ and $Q_k^H q_k = 0$, premultiplication with $Q_k^H$ results in

$$Q_k^H A Q_k = H_k = (Q_k^H A Q_k)^H = H_k^H.$$

Hence the Hessenberg matrix $H_k$ must be Hermitian, making it a tridiagonal matrix $T_k$. □

*Alternative proof.* Start from the pair of recurrence relations from Lemma 3.8 and show that it reduces to a single recurrence relation when $v = w$. Since $\mathcal{K}_k(A^H, w) = \mathcal{K}_k(A, v)$, the biorthonormal bases $V_k, W_k$ now span the same space, for any $k$. By induction it can be shown that $V_k = W_k$ for any $k$, since the basisvectors $v_i, w_i \in \mathcal{K}_{i+1}(A, v)$ are uniquely determined by the orthogonality conditions $v_i, w_i \perp \mathcal{K}_i(A, v)$ and normalization. Thus, by using the appropriate normalization, $W_k = V_k$ and

$$A^H W_k = W_k T_k^H + \bar{t}_{k,k+1} w_{k+1} e_k^\top$$

$$\Leftrightarrow A V_k = V_k T_k^H + \bar{t}_{k,k+1} v_{k+1} e_k^\top = V_k T_k + t_{k+1,k} v_{k+1} e_k^\top$$

and thus the tridiagonal matrix $T_k$ is Hermitian and generates a single orthogonal nested basis $V_k$. □

The off-diagonal elements of $T_k$ can be chosen to be positive and then it is called a Jacobi matrix. The Hermitian Lanczos iteration is well studied [130, 132] and the properties of the resulting recurrence matrix, the Jacobi matrix are also extensively studied. Therefore we will not discuss this method further.

### Eliminator factorization

The eliminator factorization of a Hermitian tridiagonal matrix is discussed in Lemma 3.11. This is related to the reduction of the LR factorization to Cholesky factorization for Hermitian positive definite matrices.

**Lemma 3.11** (Eliminator factorization of tridiagonal matrix)**.** *Let $T \in \mathbb{C}^{k\times k}$ be a strongly nonsingular Hermitian tridiagonal matrix. Then $T$ can be factorized as $T = R^H D R$, where $R = \left(\prod_{i=k-1}^{1} E_i^u\right)$ is upper triangular and $D$ a real diagonal matrix.*

*Proof.* Lemma 3.9 states that the decomposition $T = LDR$ exists, where $R = \left(\prod_{i=k-1}^{1} E_i^u\right)$ and $L = \left(\prod_{i=1}^{k-1} E_i^l\right)$ are unit upper and lower triangular. From $T^H = T$ we have $R^H D^H L^H = LDR$ and from the uniqueness of the LDR factorization $L = R^H$ and $D = D^H \in \mathbb{R}^{k\times k}$. □

### Hermitian Lanczos iteration

The implementation of the Hermitian Lanczos iteration that is most often used, introduced and analyzed by Paige [132], is given in Algorithm 3.

**Algorithm 3** Hermitian Lanczos iteration [132]

**Input:** $A \in \mathbb{C}^{m\times m}$, $A^H = A$, $v \in \mathbb{C}^m$, integer $k < g$
**Output:** Orthogonal $Q_{k+1} \in \mathbb{C}^{m\times k+1}$, tridiagonal matrix $\underline{T}_k \in \mathbb{C}^{(k+1)\times k}$ such that $AQ_k = Q_{k+1}\underline{T}_k$.
**procedure** ARNOLDI_ITERATION($A, v, k$)
    $q_0 := v/\|v\|$
    $t_{1,1} := q_0^H A q_0$
    $\tilde{q}_1 := Aq_0 - t_{1,1}q_0$
    **for** $i = 1, 2, \ldots, k$ **do**
        $t_{i,i-1} := \|\tilde{q}_i\|_2$, $t_{i-1,i} := \bar{t}_{i,i-1}$
        $q_i = \tilde{q}_i/t_{i,i-1}$
        $u_i = Aq_i - t_{i,i-1}q_{i-1}$
        $t_{i,i} = q_i^H u_i$
        $\tilde{q}_{i+1} = u_i - t_{i,i}q_i$
    **end for**
**end procedure**

For a full discussion on this implementation and its finite precision behavior we refer to the dissertation of Paige [132] and for a recent survey to the book by Meurant [130].

## 3.6 Conclusion

Three types of bases for polynomial Krylov subspaces and their recurrence relations are discussed. The Krylov basis is easy to study thanks to its simple structure and provides clear conditions for breakdowns in terms of the rank of certain matrices. However due to ill-conditioning this basis is unsuited for numerical computations. The orthogonal basis is expected to deliver the most numerically stable results and is characterized by a Hessenberg recurrence matrix, i.e., a long recurrence relation. To alleviate the high cost associated with the long recurrence relation, biorthogonal bases can be used. Biorthogonal bases, for a specific pair of Krylov subspaces, lead to a pair of three term recurrence relations. These do not possess the inherent stability of orthogonal bases. This should be taken into account when developing biorthogonal procedures.

# Chapter 4

# Structured matrices in rational Krylov subspaces

Polynomial Krylov subspaces provide subspaces which contain mainly information related to the best separated eigenvalues, typically the extreme eigenvalues. For more precise statements on convergence of Krylov subspace methods, there are results for solving linear systems for non-Hermitian matrices [125] and based on potential theory for Hermitian matrices [120] and unitary matrices [100]. To obtain subspaces which represent other, chosen eigenvalues several generalizations have been introduced:

- polynomial Krylov subspaces generated with a shift-invert version of the matrix $A$ with shift $s \in \mathbb{C}$, will generate a space representing the dominant eigenvalues of $(A - sI)^{-1}$ well. That is, those eigenvalues close to the shift $s$.
- rational Krylov subspaces, which generalize the previous idea further, by allowing the shift $s$ to change for every new vector added to the subspace being constructed.

Ruhe [143] introduced rational Krylov subspaces together with an Arnoldi-type iteration to construct an orthogonal basis for these subspaces. Ruhe argued that, apart from polynomials, rational functions with prescribed poles are a suitable candidate for numerical computation thanks to the simplicity and efficiency with which they can be manipulated.
Rational Krylov subspaces are introduced in Section 4.1 and we comment on some different conventions used in the literature. Polynomial Krylov subspaces can be regarded as a special case of rational Krylov subspaces, so why do we study them separately? First, polynomial Krylov subspaces have an intimate connection to orthogonal polynomials [43, 157], whereas rational Krylov subspaces are connected

to orthogonal rational functions. Second, because of the nature of a rational Krylov subspace, systems of equations have to be solved. Hence, they are not suited for solving systems of equations, which is an important application of polynomial Krylov subspaces (e.g., GMRES [149], CG [102], BiCGSTAB [168]). With inexact solves for the shift-inversion, rational Krylov subspaces can be used to solve systems of equations. An equivalent interpretation is to view them as a preconditioned polynomial Krylov method where the preconditioner is allowed to change at each iteration, in literature referred to as flexible GMRES [147]. For rational Krylov subspaces three types of bases and their corresponding recurrence relations are discussed.

Krylov bases are discussed in Section 4.2, which are a useful theoretical tool, but fail to guarantee a practical basis for numerical computation. The analysis of Krylov bases provides insight into breakdowns and rank structures appearing in the recurrence matrix, the matrix containing the recurrence coefficients, and the Gram matrix associated with rational Krylov subspaces. For practical computation orthogonal bases are preferred, where the rational Arnoldi iteration [143] is the method of choice to generate them. Section 4.3 discusses these orthogonal bases and their relation to Krylov bases. For rational Krylov subspaces generated with nonnormal matrices the orthogonal bases must be generated with a long recurrence relation. To obtain short recurrence relations we consider biorthogonal bases in Section 4.4. All possible representations of the recurrence relations for biorthogonal bases are derived from the recurrence matrices appearing for orthogonal bases. This derivation is based on the factorization of the associated Gram matrix. Instead of using a single matrix to represent the recurrence relation, a matrix pencil can be used. Matrix pencils representing recurrence coefficients are the natural object to study short recurrence relations for rational Krylov subspaces. In particular, a tridiagonal pencil is derived for rational Krylov subspaces. This is a novel result. Section 4.5 identifies special cases appearing in the literature of these results on matrix and pencil structures in biorthogonal rational Krylov bases.

The main results presented in this chapter are published in [165].

## 4.1 Rational Krylov subspaces

Rational Krylov subspaces are introduced in Definition 4.1. They are constructed by multiplication by $A$ and/or shift inverted versions of $A$, i.e., $(A-\xi_i I)^{-1}$. The shifts do not have to be known a priori, only at the introduction of a new vector to the subspace.

**Definition 4.1** (Rational Krylov subspace: practical form)**.** *Let $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$ and $\Xi = \{\xi_i\}_{i=1}^k$, with $\xi_i \in \overline{\mathbb{C}}$. A rational Krylov subspace with poles $\xi_i = \frac{\nu_i}{\mu_i}$, shifts $\frac{\eta_i}{\rho_i} \in \mathbb{C}$ and continuation vectors $t_i \in \mathcal{K}_i(A, v; \Xi)$ is defined as*

$$\mathcal{K}_k(A, v; \Xi) := span\left\{v, \frac{\rho_1 A - \eta_1 I}{\mu_1 A - \nu_1 I} t_1, \frac{\rho_2 A - \eta_2 I}{\mu_2 A - \nu_2 I} t_2, \ldots, \frac{\rho_{k-1} A - \eta_{k-1} I}{\mu_{k-1} A - \nu_{k-1} I} t_{k-1}\right\}.$$

The parameters $\rho_i, \eta_i, \mu_i, \nu_i$ must satisfy $\frac{\eta_i}{\rho_i} \neq \frac{\nu_i}{\mu_i}$, $\rho_i, \mu_i \neq 0$ and the continuation vectors $t_i \in \mathcal{K}_i(A, v; \Xi)$, $i = 0, 1, \ldots, k-1$, must be chosen such that $\dim(\mathcal{K}_k(A, v; \Xi)) = k$ for all $k \leq g$, where $g$ denotes the grade of $v$ with respect to $A$. Definition 4.1 allows us to deal with finite and infinite poles simultaneously, for an infinite pole the choice $\rho_k \neq 0$ and $\mu_k = 0$ is appropriate.
Intuitively, in eigenvalue computation, the pole $\xi_i = \frac{\nu_i}{\mu_i}$ increases the contribution of eigenvalues near $\xi_i$, while the shift $\frac{\eta_i}{\rho_i}$ suppresses the contribution near this quantity. Moreover there is flexibility in the choice of the continuation vector $t_i$ and the choice of the shift, determined by $\eta_i, \rho_i$. Lemma 4.1 states several properties of rational Krylov subspaces.

**Lemma 4.1** (Properties of rational Krylov subspaces [98])**.** *For a rational Krylov subspace $\mathcal{K}_k(A, v; \Xi)$ and the polynomial $\pi_{k-1}(z) := \prod_{i=1, \xi_i \neq \infty}^{k-1} (\mu_i z - \nu_i)$ of degree $\leq k-1$, with $\xi_i = \frac{\nu_i}{\mu_i}$ the following statements hold*

- $\mathcal{K}_k(A, v; \Xi) = \mathcal{K}_k(A, \pi_{k-1}^{-1}(A)v)$
- $v \in \mathcal{K}_k(A, v; \Xi)$
- $\dim(\mathcal{K}_k(A, v; \Xi)) = \min\{k, g\}$*, where $g$ is the usual grade of $v$ with respect to $A$*
- *the spaces $\mathcal{K}_k(A, v; \Xi)$ and $\mathcal{P}_{k-1}/\pi_{k-1}$ are isomorph, where $\mathcal{P}_{k-1}/\pi_{k-1} = span\{1/\pi_{k-1}(z), z/\pi_{k-1}(z), \ldots, z^{k-1}/\pi_{k-1}(z)\}$.*

Since the representation in Definition 4.1 complicates notation tremendously and we are foremost interested in theoretical properties, we will mostly use the form introduced in Definition 4.2. The only parameters are the poles $\xi_i$, and $\xi_i = \infty$ is dealt with by using a *selection vector* to keep track of finite and infinite poles.

**Definition 4.2** (Rational Krylov subspace (RKS)). *Let $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$ and $\Xi = \{\xi_i\}_{i=1}^{k-1}$, with $\xi_i \in \overline{\mathbb{C}}$. Consider a* selection vector *$s$ indicating when a pole is finite or infinite, $s_i := \begin{cases} 1 & \text{if } \xi_i = \infty \\ 0 & \text{else} \end{cases}$, and indices $k_i^+ := \sum_{j=1}^{i} s_j$, $k_i^- := i - k_i^+$. The corresponding rational Krylov subspace is*

$$\mathcal{K}_k(A, v; \Xi) = span\{v, \psi_1(A)v, \psi_2(A)v, \ldots, \psi_{k-1}(A)v\} \tag{4.1}$$

*where $\psi_i(z) := \begin{cases} z^{k_i^+} & \text{if } s_i = 1 \\ \phi_{k^-}(z) = (z-\xi_i)^{-1}\phi_{k^- -1}(z) & \text{else} \end{cases}$ and $\phi_0(z) := 1$.*

The selection vector still complicates the notation, but it is required in some statements in this section. Some results will be stated only for finite poles, the omission of infinite poles avoids complicated expressions that are merely a form of bookkeeping. Examples will then be provided which show that the inclusion of infinite poles does not change anything essential.
To illustrate the rational Krylov subspaces in the form proposed by Definition 4.2, two examples, Example 4.1 and Example 4.2, are provided.

**Example 4.1.** *Let $\Xi = \{\xi_1, \xi_2, \xi_3, \infty, \xi_5, \xi_6\}$. Then the bookkeeping is given in Table 4.1.*

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $s_i$ | 0 | 0 | 0 | 1 | 0 | 0 |
| $k_i^+$ | 0 | 0 | 0 | 1 | 1 | 1 |
| $k_i^-$ | 1 | 2 | 3 | 3 | 4 | 5 |

Table 4.1: Bookkeeping for a rational Krylov subspace.

*The corresponding RKS is*

$$\mathcal{K}_7(A, v) = span\{v, \phi_1(A)v, \phi_2(A)v, \phi_3(A)v, Av, \phi_4(A)v, \phi_5(A)v\},$$

*where*

$$\phi_1(z) = (z-\xi_1)^{-1},$$

$$\phi_2(z) = (z-\xi_2)^{-1}(z-\xi_1)^{-1},$$

$$\phi_3(z) = (z-\xi_3)^{-1}(z-\xi_2)^{-1}(z-\xi_1)^{-1},$$

$$\phi_4(z) = (z-\xi_5)^{-1}(z-\xi_3)^{-1}(z-\xi_2)^{-1}(z-\xi_1)^{-1},$$

$$\phi_5(z) = (z-\xi_6)^{-1}(z-\xi_5)^{-1}(z-\xi_3)^{-1}(z-\xi_2)^{-1}(z-\xi_1)^{-1}.$$

**Example 4.2.** *Let* $\Xi = \{\infty, \xi_2, \xi_3, \infty, \infty, \infty\}$. *Then the bookkeeping is given in Table 4.2.*

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $s_i$ | 1 | 0 | 0 | 1 | 1 | 1 |
| $k_i^+$ | 1 | 1 | 1 | 2 | 3 | 4 |
| $k_i^-$ | 0 | 1 | 2 | 2 | 2 | 2 |

Table 4.2: Bookkeeping for a rational Krylov subspace.

*The corresponding RKS is*

$$\mathcal{K}_7(A, v) = span\{v, Av, \phi_1(A)v, \phi_2(A)v, A^2v, A^3v, A^4v\},$$

*where*

$$\phi_1(z) = (z - \xi_2)^{-1},$$

$$\phi_2(z) = (z - \xi_3)^{-1}(z - \xi_2)^{-1}.$$

In order to be able to compute with these rational Krylov subspaces a basis is required. The following three sections describe possible bases to use: the Krylov basis, an orthogonal basis and biorthogonal bases.

## 4.2 Rational Krylov basis

The (rational) Krylov basis $\{v, \psi_1(A)v, \ldots, \psi_{k-1}(A)v\}$ associated with a rational Krylov subspace $\mathcal{K}_k(A, v; \Xi)$ forms the *rational Krylov matrix*

$$B_k^\Xi := \left[v, \psi_1(A)v, \psi_2(A)v, \ldots, \psi_{k-1}(A)v\right]. \tag{4.2}$$

Theorem 4.1 provides an expression for the rank of $B_k^\Xi$ in terms of the grade $g$ of $v$ with respect to $A$. The proof of the following theorem is made analogously to the proof of Theorem 3.1 for the polynomial case without explicitly relying on the connection between polynomial and rational Krylov subspaces. Fasino [63] provided this proof for a normal matrix $A$ and square rational Krylov matrices where only finite poles are allowed. Here, the result is given for a diagonalizable matrix $A$ and rectangular rational Krylov matrices, allowing any pole in $\overline{\mathbb{C}}$. The key idea in the proof is to extract a Vandermonde matrix, two illustrative examples of this procedure are given in Example 4.3 and Example 4.4.

**Theorem 4.1** (Rank of a rational Krylov matrix)**.** *Let $v \in \mathbb{C}^m$ and $A \in \mathbb{C}^{m\times m}$ be a diagonalizable matrix, i.e., $A = X\Lambda X^{-1}$, where $\Lambda = diag(\lambda_1, \ldots, \lambda_m)$ exists. Consider $\mathcal{K}_k(A, v; \Xi)$, with a given set of poles $\Xi = \{\xi_i\}_{i=1}^{k-1}$, $\xi_i \in \overline{\mathbb{C}}$, and corresponding rational Krylov matrix $B_k^{\Xi} \in \mathbb{C}^{m\times k}$. Then rank$(B_k^{\Xi}) = \min\{k, g\}$, with $g$ the grade of $v$ with respect to $A$. Moreover $g \leq \min\{c, d\}$ for $c$ the number of nonzero elements in $X^{-1}v$ and $d$ the number of distinct eigenvalues.*

*Proof.* A Vandermonde matrix is present in the rational Krylov matrix, as in the polynomial case, however it takes some effort to reveal it. First a preliminary decomposition isolating the rational functions $\psi_i(z)$

$$
\begin{aligned}
B_k^{\Xi} &= \begin{bmatrix} | & | & & | \\ v & \psi_1(A)v & \ldots & \psi_{k-1}(A)v \\ | & | & & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & & | \\ XX^{-1}v & \psi_1(X\Lambda X^{-1})v & \ldots & \psi_{k-1}(X\Lambda X^{-1})v \\ | & | & & | \end{bmatrix} \\
&= X \begin{bmatrix} | & | & & | \\ X^{-1}v & \psi_1(\Lambda)X^{-1}v & \ldots & \psi_{k-1}(\Lambda)X^{-1}v \\ | & | & & | \end{bmatrix} \\
&\quad \text{Set } w_i := (X^{-1}v)e_i \\
&= \underbrace{X \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & w_m \end{bmatrix}}_{=:W} \underbrace{\begin{bmatrix} 1 & \psi_1(\lambda_1) & \ldots & \psi_{k-1}(\lambda_1) \\ \vdots & & \vdots & \\ 1 & \psi_1(\lambda_m) & \ldots & \psi_{k-1}(\lambda_m) \end{bmatrix}}_{=:F}.
\end{aligned}
$$

Now we focus on the matrix $F$, which hides a Vandermonde matrix. Construct a permutation matrix $P$ that orders the rational functions in $F$ by decreasing degree, i.e.

$$\begin{bmatrix} 1 & \psi_1(z) & \ldots & \psi_{k-1}(z) \end{bmatrix} P = \begin{bmatrix} \tilde{\psi}_{k-1}(z) & \tilde{\psi}_{k-2}(z) & \ldots & \tilde{\psi}_0(z) \end{bmatrix}$$

with $\deg(\tilde{\psi}_i) > \deg(\tilde{\psi}_{i-1})$. The degree of a rational function is taken to be the degree of the numerator minus the degree of the denominator. The matrix $P$ is unique.

Set $\pi_i(z) = \prod_{l=i+1}^{k}(z - \hat{\xi}_l)$, with $\begin{cases} \hat{\xi}_i = \xi_i, & \text{if } \xi_i \neq \infty \\ \hat{\xi}_i = 0, & \text{if } \xi_i = \infty \end{cases}$. For the products $\pi_i(z)\phi_{k^-}(z)$, $i = 0, 1, \ldots, k-1$, there exists a unit lower triangular matrix $L$ such that

$$\begin{bmatrix} \tilde{\psi}_{k-1}(z) & \ldots & \tilde{\psi}_0(z) \end{bmatrix} = \begin{bmatrix} \pi_0(z)\phi_{k^-}(z) & \ldots & \pi_{k-1}(z)\phi_{k^-}(z) \end{bmatrix} L.$$

This allows the extraction of the common term $\phi_{k^-}(z)$

$$\begin{bmatrix}\pi_0(z)\phi_{k^-}(z) & \dots & \pi_{k-1}(z)\phi_{k^-}(z)\end{bmatrix} = \phi_{k^-}(z)\begin{bmatrix}\pi_0(z) & \pi_1(z) & \dots & \pi_{k-1}(z)\end{bmatrix}.$$

The monic polynomials $\pi_i(z)$ have exact degree $\deg(\pi_i) = k - i - 1$. Hence, there exists a unit lower triangular matrix $\hat{L}$ such that

$$\begin{bmatrix}\pi_0(z) & \pi_1(z) & \dots & \pi_{k-1}(z)\end{bmatrix} = \begin{bmatrix}z^{k-1} & \dots & z & 1\end{bmatrix}\hat{L}.$$

To summarize, in matrix notation:

$$\begin{aligned}
F &= FPP^\top \\
&= \begin{bmatrix}\tilde{\psi}_{k-1}(\lambda_1) & \dots & \tilde{\psi}_1(\lambda_1) & \tilde{\psi}_0(\lambda_1)\\ \vdots & & \vdots & \vdots \\ \tilde{\psi}_{k-1}(\lambda_m) & \dots & \tilde{\psi}_1(\lambda_m) & \tilde{\psi}_0(\lambda_m)\end{bmatrix} P^\top \\
&= \underbrace{\begin{bmatrix}\phi_{k^-}(\lambda_1) & & \\ & \ddots & \\ & & \phi_{k^-}(\lambda_m)\end{bmatrix}}_{=:\Phi}\begin{bmatrix}\pi_0(\lambda_1) & \dots & \pi_{k-2}(\lambda_1) & \pi_{k-1}(\lambda_1)\\ \vdots & & \vdots & \vdots \\ \pi_0(\lambda_m) & \dots & \pi_{k-2}(\lambda_m) & \pi_{k-1}(\lambda_m)\end{bmatrix} LP^\top \\
&= \Phi\underbrace{\begin{bmatrix}\lambda_1^{k-1} & \dots & \lambda_1 & 1\\ \vdots & & \vdots & \vdots \\ \lambda_m^{k-1} & \dots & \lambda_m & 1\end{bmatrix}}_{=:V}\hat{L}LP^\top
\end{aligned}$$

By construction the matrix $\Phi$ is nonsingular (poles cannot coincide with eigenvalues), permutation matrices are nonsingular and the product $\hat{L}L$ is unit lower triangular and thus, nonsingular. Hence, $\operatorname{rank}(F) = \operatorname{rank}(V)$. Then $\operatorname{rank}(B_k^\Xi) = \operatorname{rank}(XWF) = \operatorname{rank}(WF) = \operatorname{rank}(WV)$. □

The relation between Vandermonde and (rational) Krylov matrices allows to relate their respective condition numbers [11]. In Chapter 5 the relationship between a rational Krylov matrix and a Vandermonde matrix will be used to study its displacement rank and that of the related Gram matrix.

**Example 4.3.** *Denote by $\xi_i$ finite poles and let $\Xi = \{\infty, \xi_2, \xi_3, \infty\}$. The RKS is*

$$\begin{aligned}
\mathcal{K}_5(A, v; \Xi) &= \mathit{span}\{v, \psi_1(A)v, \psi_2(A)v, \psi_3(A)v, \psi_4(A)v\} \\
&= \mathit{span}\{v, Av, \phi_1(A)v, \phi_2(A)v, A^2v\},
\end{aligned}$$

*where*

$$\phi_1(z) = (z-\xi_2)^{-1},$$

$$\phi_2(z) = (z-\xi_3)^{-1}(z-\xi_2)^{-1}.$$

*With the corresponding*

$$F = \begin{bmatrix} 1 & \psi_1(\lambda_1) & \dots & \psi_4(\lambda_1) \\ \vdots & \vdots & & \vdots \\ 1 & \psi_1(\lambda_m) & \dots & \psi_4(\lambda_m) \end{bmatrix} = \begin{bmatrix} 1 & \lambda_1 & \phi_1(\lambda_1) & \phi_2(\lambda_1) & \lambda_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_m & \phi_1(\lambda_m) & \phi_2(\lambda_m) & \lambda_m^2 \end{bmatrix}.$$

*Table 4.3 shows the parameters keeping track of positive and negative powers.*

| $i$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $s_i$ | 1 | 0 | 0 | 1 |
| $k_i^+$ | 1 | 1 | 1 | 2 |
| $k_i^-$ | 0 | 1 | 2 | 2 |

Table 4.3: Bookkeeping for a rational Krylov subspace.

*The polynomials* $\pi_i(z) = \prod_{l=i+1}^{k}(z-\hat{\xi}_l)$ *and the products, denoted by* $\tilde{\psi}_i$*, allow the extraction of the common factor* $\phi_2(z)$ *are*

$$\pi_0(z) = z(z-\xi_2)(z-\xi_3)z, \qquad \tilde{\psi}_4 := \pi_0(z)\phi_2(z) = z^2,$$

$$\pi_1(z) = (z-\xi_2)(z-\xi_3)z, \qquad \tilde{\psi}_3 := \pi_1(z)\phi_2(z) = z,$$

$$\pi_2(z) = (z-\xi_3)z, \qquad \tilde{\psi}_2 := \pi_2(z)\phi_2(z) = \frac{z}{z-\xi_2},$$

$$\pi_3(z) = z, \qquad \tilde{\psi}_1 := \pi_3(z)\phi_2(z) = \frac{z}{(z-\xi_3)(z-\xi_2)},$$

$$\pi_4(z) = 1, \qquad \tilde{\psi}_0 := \pi_4(z)\phi_2(z) = \frac{1}{(z-\xi_3)(z-\xi_2)}.$$

*First a reordering by degree is performed using the permutation matrix* $P$

$$FP = F\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} z^2 & z & 1 & \phi_1(z) & \phi_2(z) \end{bmatrix}$$

*Clearly a lower triangular $L$ can be found such that $FP = \widetilde{\Psi}L$, with $\widetilde{\Psi} = \begin{bmatrix}\tilde{\psi}_4 & \tilde{\psi}_3 & \dots & \tilde{\psi}_0\end{bmatrix}$,*

$$FP = \begin{bmatrix} z^2 & z & \frac{z}{z-\xi_2} & \frac{z}{(z-\xi_3)(z-\xi_2)} & \frac{1}{(z-\xi_3)(z-\xi_2)} \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\xi_2 & 1 & 0 \\ 0 & 0 & \xi_3\xi_2 & -\xi_3 & 1 \end{bmatrix}}_{=:L}.$$

*Since $\pi_i(z)$ is a monic polynomial of exact degree $i$ it allows a representation in the monomial basis*

$$\begin{bmatrix}\pi_0 & \pi_1 & \dots & \pi_4\end{bmatrix} = \begin{bmatrix} z^4 & \dots & z & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & & & & \\ -\xi_3-\xi_2 & 1 & & & \\ \xi_3\xi_2 & -\xi_3-\xi_2 & 1 & & \\ & \xi_3\xi_2 & -\xi_3 & 1 & \\ & & & & 1 \end{bmatrix}}_{=:\hat{L}}.$$

*Hence,*

$$F = FPP^\top = \begin{bmatrix} \phi_2(\lambda_1) & & \\ & \ddots & \\ & & \phi_2(\lambda_m) \end{bmatrix} \begin{bmatrix} \lambda_1^4 & \dots & \lambda_1 & 1 \\ \vdots & & \vdots & \vdots \\ \lambda_1^4 & \dots & \lambda_m & 1 \end{bmatrix} \hat{L}LP^\top.$$

**Example 4.4.** *Denote by $\xi_i$ finite poles and let $\Xi = \{\xi_1, \xi_2, \xi_3, \infty, \xi_5, \xi_6\}$. The RKS is*

$$\mathcal{K}_7(A, v; \Xi) = span\{v, \phi_1(A)v, \phi_2(A)v, \phi_3(A)v, Av, \phi_4(A)v, \phi_5(A)v\},$$

*where*

$$\begin{aligned} \phi_1(z) &= (z-\xi_1)^{-1}, \\ \phi_2(z) &= (z-\xi_2)^{-1}(z-\xi_1)^{-1}, \\ \phi_3(z) &= (z-\xi_3)^{-1}(z-\xi_2)^{-1}(z-\xi_1)^{-1}, \\ \phi_4(z) &= (z-\xi_5)^{-1}(z-\xi_3)^{-1}(z-\xi_2)^{-1}(z-\xi_1)^{-1}, \\ \phi_5(z) &= (z-\xi_6)^{-1}(z-\xi_5)^{-1}(z-\xi_3)^{-1}(z-\xi_2)^{-1}(z-\xi_1)^{-1}. \end{aligned}$$

*The corresponding parameters are shown in Table 4.4.*

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $s_i$ | 0 | 0 | 0 | 1 | 0 | 0 |
| $k_i^+$ | 0 | 0 | 0 | 1 | 0 | 0 |
| $k_i^-$ | 1 | 2 | 3 | 3 | 4 | 5 |

Table 4.4: Bookkeeping for a rational Krylov subspace.

*Using $\pi_i(z)$ we can extract the common factor $\phi_5(z)$*

$$\pi_0(z) = (z-\xi_6)(z-\xi_5)z(z-\xi_3)(z-\xi_2)(z-\xi_1), \quad \tilde{\psi}_6 := \pi_0(z)\phi_5(z) = z,$$

$$\pi_1(z) = (z-\xi_6)(z-\xi_5)z(z-\xi_3)(z-\xi_2), \quad \tilde{\psi}_5 := \pi_1(z)\phi_5(z) = \frac{z}{z-\xi_1},$$

$$\pi_2(z) = (z-\xi_6)(z-\xi_5)z(z-\xi_3), \quad \tilde{\psi}_4 := \pi_2(z)\phi_5(z) = \frac{z}{(z-\xi_2)(z-\xi_1)},$$

$$\pi_3(z) = (z-\xi_6)(z-\xi_5)z, \quad \tilde{\psi}_3 := \pi_3(z)\phi_5(z) = \frac{z}{(z-\xi_3)(z-\xi_2)(z-\xi_1)},$$

$$\pi_4(z) = (z-\xi_6)(z-\xi_5), \quad \tilde{\psi}_2 := \pi_4(z)\phi_5(z) = \frac{1}{(z-\xi_3)(z-\xi_2)(z-\xi_1)}$$

$$\pi_5(z) = z-\xi_6, \quad \tilde{\psi}_1 := \pi_5(z)\phi_5(z) = \frac{1}{(z-\xi_5)(z-\xi_3)(z-\xi_2)(z-\xi_1)}$$

$$\pi_6(z) = 1, \quad \tilde{\psi}_0 := \pi_6(z)\phi_5(z) = \frac{1}{(z-\xi_6)(z-\xi_5)(z-\xi_3)(z-\xi_2)(z-\xi_1)}.$$

*Then for*

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } L = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & -\xi_1 & 1 & & & & \\ & \xi_2\xi_1 & -\xi_2 & 1 & & & \\ & -\xi_3\xi_2\xi_1 & \xi_3\xi_2 & -\xi_3 & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix}$$

*we have*

$$FP = \begin{bmatrix} \tilde{\psi}_6(z) & \tilde{\psi}_5(z) & \dots & \tilde{\psi}_0(z) \end{bmatrix} L.$$

*Then, let column $i$ of $\hat{L}$, $\hat{l}_i$, be the representation of $\pi_{6-i}(z)$ in the monomial basis, i.e., $\pi_{6-i}(z) = \begin{bmatrix}1 & z & \dots & z^6\end{bmatrix}\hat{l}_i$ and we obtain*

$$F = \begin{bmatrix}\phi_5(\lambda_1) & & \\ & \ddots & \\ & & \phi_5(\lambda_m)\end{bmatrix}\begin{bmatrix}\lambda_1^6 & \dots & \lambda_1 & 1 \\ \vdots & & \vdots & \vdots \\ \lambda_m^6 & \dots & \lambda_m & 1\end{bmatrix}\hat{L}LP^\top.$$

### 4.2.1 Recurrence relations

The rational Krylov matrix $B_k^\Xi$ satisfies a recurrence relation. This relation can be described by a recurrence matrix or it can be represented as a *recurrence pencil*. A recurrence pencil is a pair of matrices which contain recurrence coefficients, so a generalization of a matrix to a pencil in the usual sense. Matrix pencils are not unique, therefore several different representations exist. Three representations are important in this manuscript: a Hessenberg pencil provides an elegant and intuitive representation, a single recurrence matrix follows from projection onto the Krylov subspace and an inverse Hessenberg pencil is useful for the analysis of structures of recurrence pencils of biorthogonal bases.
Only the Hessenberg pencil is discussed in detail, its derivation is natural in the Krylov basis. The single recurrence matrix and inverse Hessenberg pencil are more tedious to derive. Their discussion is postponed to Section 4.3 on orthogonal bases, where the core factorization allows elegant characterization and definition.

#### Hessenberg pencil

A structural constraint is set on the recurrence pencil. The structure of the matrices appearing in the recurrence pencil must both adhere to Hessenberg structure, i.e., they must form a *Hessenberg pencil*. Then the poles of the related RKS can be easily retrieved from this representation. Lemma 4.2 elaborates on this representation for finite poles. For brevity the infinite poles are excluded, including infinite poles is not a problem but it requires a lot of bookkeeping. Example 4.5 provides an example with both finite and infinite poles.

**Lemma 4.2** (Rational Krylov recurrence relation: Hessenberg pencil)**.** *Let $B_{k+1}^\Xi \in \mathbb{C}^{m\times(k+1)}$, $k < g$, be the rational Krylov matrix corresponding to $\mathcal{K}_{k+1}(A, v; \Xi)$ and poles satisfying $\xi_i \in \mathbb{C}\backslash\sigma(A)$ for all $i$. Then $B_{k+1}^\Xi$ satisfies the following recurrence*

*relation for the Hessenberg pencil* $(\underline{C}_k, \underline{F}_k) \in \mathbb{C}^{(k+1)\times k} \times \mathbb{C}^{(k+1)\times k}$

$$AB_k^{\Xi} F_k + Ab_k e_k^\top = B_k^{\Xi} C_k + b_k e_k^\top$$

$$AB_{k+1}^{\Xi} \underline{F}_k = B_{k+1}^{\Xi} \underline{C}_k$$

$$AB_{k+1}^{\Xi} \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & \ddots & \\ & & \ddots & 0 \\ & & & 1 \end{bmatrix} = B_{k+1}^{\Xi} \begin{bmatrix} 1 & & & \\ \xi_1 & 1 & & \\ & \xi_2 & \ddots & \\ & & \ddots & 1 \\ & & & \xi_k \end{bmatrix},$$

*and, clearly,* $\underline{C}_k = I + \underline{F}_k \begin{bmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_k \end{bmatrix}$.

*Proof.* For only finite poles the rational Krylov matrix is given by $B_{k+1}^{\Xi} = \begin{bmatrix} v & \phi_1(A)v & \dots & \phi_{k-1}(A)v & \phi_k(A)v \end{bmatrix}$. The statement follows immediately from the equality $z\phi_i(z) = \phi_{i-1}(z) + \xi_i \phi_i(z)$. □

Example 4.5 illustrates that including infinite poles will not change anything essential to the statement of Lemma 4.2. That is, the matrices have Hessenberg structure and the ratios of the subdiagonal elements reveal the poles. In the context of poles, if $\alpha \neq 0$ then $\frac{\alpha}{0} = \infty$, i.e., an infinite poles.

**Example 4.5.** $\mathcal{K}(A, v; \{\xi_1, \xi_2, \xi_3, \infty, \xi_5, \xi_6\})$ *and the corresponding Krylov basis* $B_7 = \begin{bmatrix} v & \phi_1(A)v & \phi_2(A)v & \phi_3(A)v & Av & \phi_4(A)v & \phi_5(A)v \end{bmatrix}$

$$AB_7 \begin{bmatrix} 0 & & & 1 & & & \\ 1 & 0 & & & & & \\ & 1 & 0 & & & & \\ & & 1 & 0 & & & \\ & & & 0 & 0 & & \\ & & & & 1 & 0 & \\ & & & & & 1 & \end{bmatrix} = B_7 \begin{bmatrix} 1 & & & & & & \\ \xi_1 & 1 & & & & & \\ & \xi_2 & 1 & & & & \\ & & \xi_3 & 0 & 1 & & \\ & & & 1 & 0 & & \\ & & & & & \xi_4 & 1 \\ & & & & & & \xi_5 \end{bmatrix}.$$

*Subdiagonal ratios are* $\{\frac{\xi_1}{1}, \frac{\xi_2}{1}, \frac{\xi_3}{1}, \frac{1}{0}, \frac{\xi_4}{1}, \frac{\xi_5}{1}\}$.

By including infinite poles the elegant connection between $\underline{C}_k$ and $\underline{F}_k$ has, however, been lost for some part. A connection can be obtained by changing $I$ by a permutation matrix of Hessenberg form, but this is not explored further.

**Corollary 4.1.** *If* $\xi_k = \infty$*, then the recurrence relation* $AB_k^{\Xi} F_k + Ab_k e_k^\top = B_k^{\Xi} C_k + \xi_k b_k e_k^\top$ *from Lemma 4.2 becomes* $AB_k^{\Xi} F_k = B_k^{\Xi} C_k + b_k e_k^\top$.

### Rational Hessenberg matrix

A single matrix representation, corresponding to the Krylov subspace from Example 4.5, is provided in Example 4.6.

**Example 4.6.** *Consider* $\mathcal{K}(A, v; \Xi)$*, with* $\hat{\Xi} = \{\xi_1, \xi_2, \xi_3, \infty, \xi_5, \xi_6\}$*, and its Krylov basis* $B_7^{\hat{\Xi}} = \begin{bmatrix} v & \phi_1(A)v & \phi_2(A)v & \phi_3(A)v & Av & \phi_4(A)v & \phi_5(A)v \end{bmatrix}$*. Then the selection vector* $s = \{0, 0, 0, 1, 0, 0\}$ *and*

$$AB_7^{\hat{\Xi}} I = B_7^{\hat{\Xi}} Z$$

$$AB_7^{\hat{\Xi}} = B_7^{\hat{\Xi}} \begin{bmatrix} 0 & 1 & & & 0 & & \\ & \xi_1 & 1 & & 0 & & \\ & & \xi_2 & 1 & 0 & & \\ & & & \xi_3 & 0 & 1 & \\ 1 & 0 & 0 & 0 & 0 & 0 & \\ & & & & 0 & \xi_5 & 1 \\ & & & & 0 & & \xi_6 \end{bmatrix}.$$

*Clearly, the selection vector, i.e., occurrences of* $\xi_i = \infty$*, determines the lower triangular structure.*

The resulting matrix is a *rational Hessenberg matrix*. In Section 4.3.4 an elegant derivation of this structure will be provided starting from the core factorization of the Hessenberg pencil. The discussion on core factorization is postponed, since the representation in an orthogonal basis provides a more natural starting point.

## 4.3 Orthogonal basis

Three types of recurrence relations for an orthogonal basis for a RKS are discussed. The analysis of the structure of the associated recurrence matrices or recurrence pencils relies on manipulating their core factorization. Ruhe [143] introduced rational Krylov subspaces and a corresponding Arnoldi-type iteration [144] to generate an orthogonal basis. This Arnoldi-type iteration is commonly referred to as the *rational Arnoldi iteration* and uses a Hessenberg pencil to represent the recurrence relation. Section 4.3.1 derives this Hessenberg recurrence pencil. Afterwards the core factorization of Hessenberg pencils is elaborated upon in Section 4.3.2. The core factorization allows for the derivation of another recurrence pencil, an inverse Hessenberg pencil, which is the subject of Section 4.3.3. The final representation is a single matrix representation discussed in Section 4.3.4. To conclude, the rational Arnoldi iteration is introduced in Section 4.3.5.

### 4.3.1 Hessenberg pencil

An orthonormal basis $Q_k \in \mathbb{C}^{m\times k}$ for $\mathcal{K}_k(A, v; \Xi)$ can be obtained by orthonormalizing the Krylov basis or, equivalently, by computing the QR decomposition of the Krylov matrix $B_k^{\Xi}$. The recurrence pencil describing the generation of an orthonormal basis can be represented as a Hessenberg pencil. This is stated in Theorem 4.2. This theorem is stated for finite poles, since it uses Lemma 4.2 in its proof. In Section 4.3.5, Theorem 4.4, it is shown that Theorem 4.2 holds for $\xi_i \in \overline{\mathbb{C}}\backslash\sigma(A)$.

**Theorem 4.2** (Orthogonal rational Krylov basis: Hessenberg pencil)**.** *Consider a nested orthonormal basis $Q_{k+1} \in \mathbb{C}^{m\times k}$ for $\mathcal{K}_{k+1}(A, v; \Xi)$, $k < g$, where $\xi_i \in \mathbb{C}\backslash\sigma(A)$. Then a Hessenberg pencil $(\underline{H}_k, \underline{K}_k) \in \mathbb{C}^{(k+1)\times k} \times \mathbb{C}^{(k+1)\times k}$ exists which satisfies*

$$AQ_{k+1}\underline{K}_k = Q_{k+1}\underline{H}_k \text{ or, equivalently,}$$

$$AQ_kK_k + k_{k+1,k}Aq_ke_k^\top = Q_kH_k + h_{k+1,k}q_ke_k^\top,$$

*where $q_k = Q_{k+1}e_{k+1}$ and $\frac{h_{i+1,i}}{k_{i+1,i}} = \xi_i$, $i = 1, 2, \ldots, k$.*

*Proof.* Consider the Krylov recurrence relation

$$AB_{k+1}^{\Xi}\underline{C}_k = B_{k+1}^{\Xi}\underline{F}_k$$

and the QR decomposition $B_{k+1}^{\Xi} = Q_{k+1}R_{k+1}$. Then substituting the QR decomposition into the recurrence relation reveals the recurrence pencil for an orthonormal basis

$$AQ_{k+1}\underbrace{R_{k+1}\underline{C}_k}_{=:\underline{K}_k} = Q_{k+1}\underbrace{R_{k+1}\underline{F}_k}_{=:\underline{H}_k}.$$

Since multiplication with an upper triangular matrix preserves the lower triangular structure, $\underline{H}_k$ and $\underline{K}_k$ are Hessenberg matrices. Writing $Q_{k+1} = \begin{bmatrix} Q_k & q_k \end{bmatrix}$ and some basic manipulation of the recurrence relation shows that

$$AQ_kK_k + r_{k+1,k+1}Aq_ke_k^\top = Q_kH_k + \xi_kr_{k+1,k+1}q_ke_k^\top.$$

Note that $k_{k+1,k} = r_{k+1,k+1}$ and $h_{k+1,k} = \xi_kr_{k+1,k+1}$, which completes the proof. □

**Corollary 4.2.** *If $\xi_k = \infty$, then*

- *the recurrence relation $AQ_kK_k + k_{k+1,k}Aq_ke_k^\top = Q_kH_k + h_{k+1,k}q_ke_k^\top$ from Theorem 4.2 becomes $AQ_kK_k = Q_kH_k + h_{k+1,k}q_ke_k^\top$.*
- *the orthogonal projection of $A$ onto the space spanned by $Q_k$ is given by $(H_k, K_k)$.*

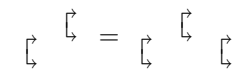
### 4.3.2 Core factorization

The core factorization of a Hessenberg pencil $(H, K) \in \mathbb{C}^{k\times k} \times \mathbb{C}^{k\times k}$ is the pencil of the core factorizations of the two Hessenberg matrices $H, K$. By Lemma 3.4, the core factorization of $H$ is $H = Q_H R_H$, with $Q_H = C_1 C_2 \dots C_{k-1}$ and of $K$ is $K = Q_K R_K$, with $Q_K = \widetilde{C}_1 \widetilde{C}_2 \dots \widetilde{C}_{k-1}$, with $C_i, \widetilde{C}_i \in \mathfrak{C}_i$. Thus, the core factorization of $(H, K)$ is $(C_1 C_2 \dots C_{k-1} R_H, \widetilde{C}_1 \widetilde{C}_2 \dots \widetilde{C}_{k-1} R_K)$, with $C_i, \widetilde{C}_i \in \mathfrak{C}_i$.

**Definition 4.3** (Proper Hessenberg pencil)**.** *A Hessenberg pencil $(H, K)$ is said to be proper if $h_{i+1,i} k_{i+1,i} \neq 0$, i.e., no subdiagonal elements of $H, K$ are simultaneously zero.*

The core factorization is a powerful tool to analyze structures of matrices thanks to the operations that can be defined between core transformations. Lemma 4.3 provides such an operation, called a *turnover*.

**Lemma 4.3** (Turnover lemma [176], Lemma 9.38)**.** *Consider the product of three core transformations $C_{i-1} C_i \hat{C}_{i-1}$, where $C_{i-1}, \hat{C}_{i-1} \in \mathfrak{C}_{i-1}$ and $C_i \in \mathfrak{C}_i$. Then there exist $\Gamma_{i-1} \in \mathfrak{C}_{i-1}$ and $\Gamma_i, \hat{\Gamma}_i \in \mathfrak{C}_i$ such that*



$$C_{i-1} C_i \hat{C}_{i-1} = \Gamma_i \Gamma_{i-1} \hat{\Gamma}_i.$$

The turnover operation is used in Section 4.3.3 and Section 4.3.4 to derive alternative representations of the recurrence coefficients.

### 4.3.3 Inverse Hessenberg pencil

The recurrence relation for an orthonormal basis for rational Krylov subspaces can be represented by a recurrence pencil with inverse Hessenberg structure. Theorem 4.3 states the result for poles in $\overline{\mathbb{C}}$, starting from Theorem 4.4. This representation of the recurrence coefficients is paramount to deriving the structures of recurrence pencils for biorthogonal bases.

**Theorem 4.3** (Orthonormal basis for RKS: inverse Hessenberg pencil)**.** *Let $Q_{k+1} \in \mathbb{C}^{m\times(k+1)}$ be an orthonormal nested basis for $\mathcal{K}_{k+1}(A, v; \Xi)$, $k < g$. Then an inverse Hessenberg pencil $(H_k^{inv}, K_k^{inv}) \in \mathbb{C}^{k\times k} \times \mathbb{C}^{k\times k}$ exists for which the recurrence relation*

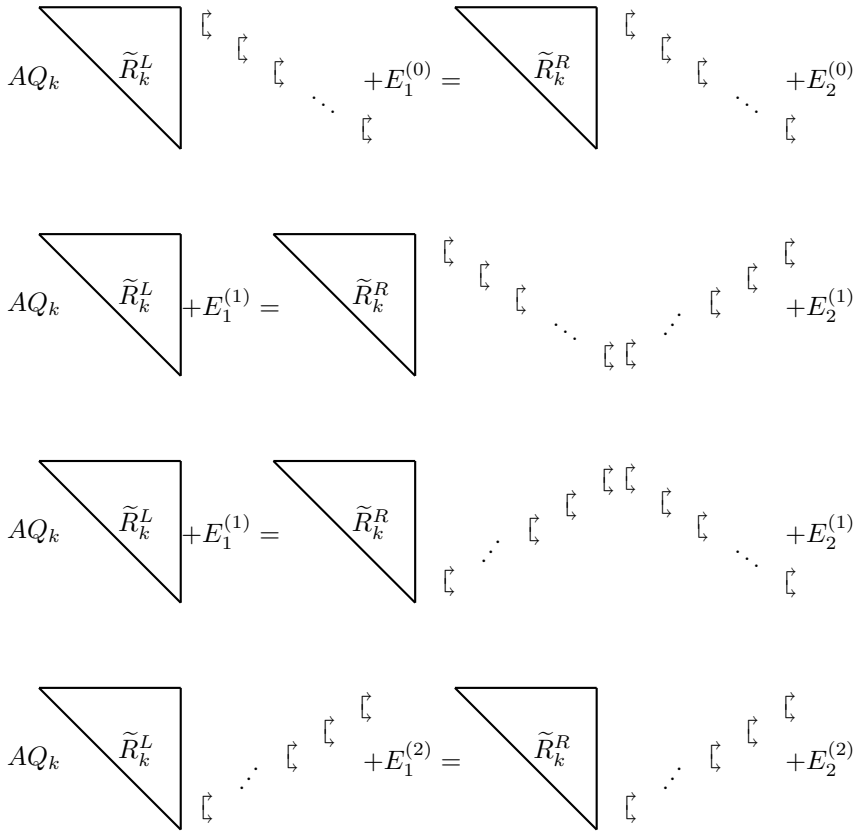
$$A Q_k K_k^{inv} + E_1 = Q_k H_k^{inv} + E_2,$$

*with $rank(E_1) = rank(E_2) = 1$, holds.*

*Proof.* A Hessenberg pencil $(H_k, K_k)$ follows from Theorem 4.4 and satisfies

$$AQ_kK_k + k_{k+1,k}Aq_ke_k^\top = Q_kH_k + h_{k+1,k}q_ke_k^\top.$$

Consider the core factorizations $H_k = \prod_{i=1}^{k-1} C_i R_k^R$ and $K_k = \prod_{i=1}^{k-1} \widetilde{C}_i R_k^L$. Now substitute the core transformations into the recurrence relation, use Lemma 3.5, basic manipulations of the core transformations and Lemma 4.3 to obtain




The terms $E_1^{(i)}$ and $E_2^{(i)}$, $i = 0, 1, 2$, have at most rank 1. Use Lemma 3.5 again and we obtain, for upper triangular $\hat{R}_k^R$ and $\hat{R}_k^L$, the pencil $(\prod_{i=k-1}^{i} \dot{C}_i \hat{R}_k^R, \prod_{i=k-1}^{i} \hat{C}_i \hat{R}_k^L)$, which has inverse Hessenberg structure and $\operatorname{rank}(\hat{R}_k^R) = \operatorname{rank}(R_k^R)$, $\operatorname{rank}(\hat{R}_k^L) = \operatorname{rank}(R_k^L)$ since all multiplications involved nonsingular matrices. □

### 4.3.4 Recurrence matrix

A single matrix representation $Z_k$ of the recurrence matrix can be derived from the Hessenberg recurrence pencil. The general core factorization of this single matrix representation is shown in Figure 4.1. The recurrence matrix for orthogonal bases of

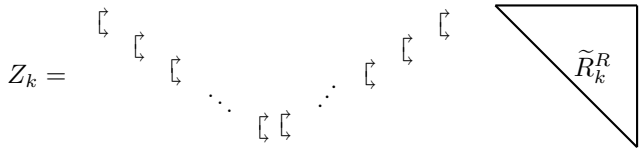
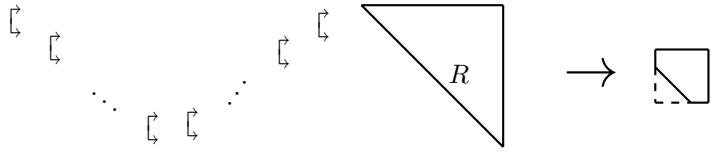



Figure 4.1: Core factorization of a recurrence matrix for an orthonormal basis of $\mathcal{K}(A, v; \Xi)$

rational Krylov subspaces will be referred to as having *rational Hessenberg structure*, see Definition 4.4 where also a pictographic notation is introduced.

**Definition 4.4** (Rational Hessenberg structure)**.** *A matrix $Z_k$ has rational Hessenberg structure, or equivalently, $Z_k$ is a rational Hessenberg matrix, if it has a core factorization consisting of a descending pattern multiplied by an ascending pattern (pre-or postmultiplication, by Lemma 4.3). A pictographic notation will be used:*




The pictographic notation in Definition 4.4 highlights the structure of $Z_k$, full lines enclose elements without rank structure and dashed lines enclose submatrices which exhibit low rank structure. Example 4.7 illustrates this notation.
Before providing this example, the following discussion is given for completeness. The properties discussed are not essential in the remainder of this manuscript. Rational Hessenberg matrices can be decomposed using an extended Hessenberg matrix. These arise from extended Krylov subspaces. *Extended Krylov subspaces* are rational Krylov subspaces which only allow $\xi_i = 0$ or $\xi_i = \infty$. Core factorizations of recurrence matrices for orthogonal bases of extended Krylov subspaces consist of a single pattern, which has descending and ascending subpatterns. This follows by noting that if $\xi_i = 0$ or $\xi_i = \infty$, a core transformation appears only in one of the two matrices in the recurrence pencil, the other core transformation must be trivial.

**Definition 4.5** (Extended Hessenberg structure)**.** *A matrix $Z_k \in \mathbb{C}^{k\times k}$ that has the core factorization $\left(\prod_{\sigma(i)} C_i\right) R_k$, $C_i \in \mathfrak{C}_i$ and $R_k$ upper triangular for some permutation $\sigma$ of $\{1, 2, \ldots, k-1\}$ is said to have extended Hessenberg structure.*

A rational Hessenberg matrix can be decomposed as the sum of an extended Hessenberg matrix plus a diagonal matrix containing the poles. Our interest is mainly in the pencil representation, so we do not go into detail on this decomposition, for more details we refer to the literature [39].

**Example 4.7.** *Consider the poles $\hat{\Xi} = \{\infty, 0, 0, \infty, \infty, \infty\}$ which generate the extended Krylov subspace*

$$\mathcal{K}_7(A, v; \hat{\Xi}) = span\{v, Av, A^{-1}v, A^{-2}v, A^2v, A^3v, A^4v\}.$$

*The recurrence matrix $\hat{Z}$ for an orthogonal basis of $\mathcal{K}_7(A, v; \hat{\Xi})$ has the core factorization $\hat{Z} = C_1C_4C_5C_6C_3C_2R$, $C_i \in \mathfrak{C}_i$. The structure of $Z$ is shown in Figure 4.2,*
*On the diagonal of $\hat{Z}$, submatrices with Hessenberg or inverse Hessenberg structure*

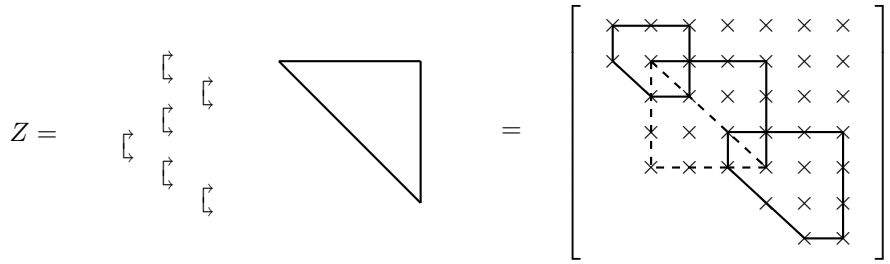



Figure 4.2: Recurrence matrix $\hat{Z} = C_1C_4C_5C_6C_3C_2R$ of an orthonormal basis for $\mathcal{K}_7(A, v; \hat{\Xi})$.

*appear, they are indicated by the lines in the figure.*
*If the finite poles $\xi_2 \neq 0, \xi_3 \neq 0$ and $\dot{\Xi} = \{\infty, \xi_2 = \frac{\mu_2}{\nu_2}, \xi_3 = \frac{\mu_3}{\nu_3}, \infty, \infty, \infty\}$ then*

$$\mathcal{K}_7(A, v; \dot{\Xi}) = span\{v, Av, \phi_1(A)v, \phi_2(z)v, A^2v, A^3v, A^4v\},$$

*with $\phi_1(z) = (\nu_2 z - \mu_2)^{-1}$ and $\phi_2(z) = (\nu_3 z - \mu_3)^{-1}\phi_1(z)$. The recurrence matrix $\dot{Z} = \dot{C}_1\dot{C}_2\dot{C}_3\dot{C}_4\dot{C}_5\dot{C}_6\widetilde{C}_3\widetilde{C}_2\hat{R}$ is a rational Hessenberg matrix. A rational Hessenberg matrix can be decomposed in an extended Hessenberg matrix plus a diagonal, as is shown in Figure 4.3. Note that in this figure the second subblock indicated by a contour changes from a rational Hessenberg block to a inverse Hessenberg block by the extraction of the poles $\xi_2$ and $\xi_3$.*

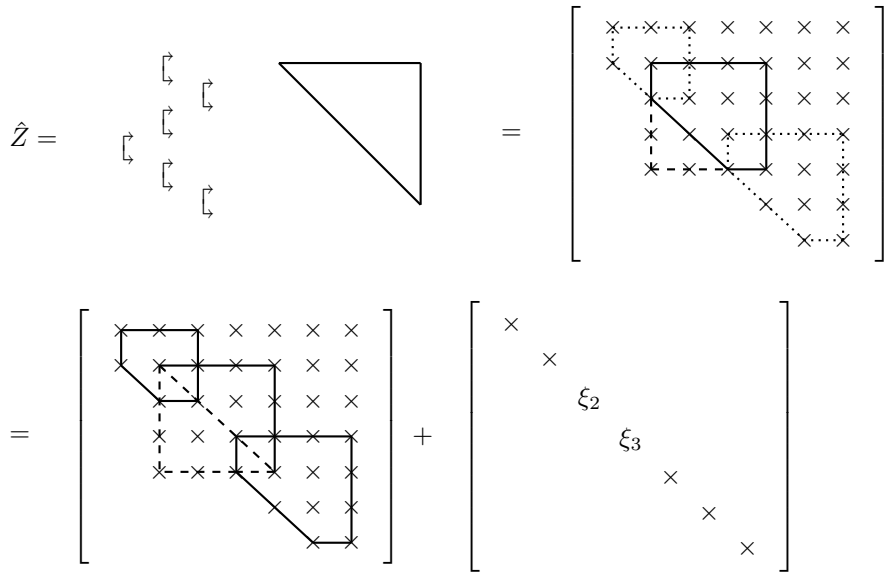



Figure 4.3: Recurrence matrix $\dot{Z} = \dot{C}_1\dot{C}_2\dot{C}_3\dot{C}_4\dot{C}_5\dot{C}_6\widetilde{C}_3\widetilde{C}_2\hat{R}$ of an orthonormal basis for $\mathcal{K}_7(A, v; \dot{\Xi})$.

### 4.3.5 Rational Arnoldi iteration

The idea is the same as for the polynomial Arnoldi iteration. The expansion is done by multiplying a suitable vector by $A$ or $(A - \xi_i I)^{-1}$. Any vector $t_k$ in the space $\mathcal{K}_k(A, v; \Xi_{k-1})$, which satisfies $(\nu_k A - \mu_k I)^{-1}(\rho_k A - \eta_k I)t_k \in \mathcal{K}_{k+1}(A, v, \Xi_k)$ for suitable parameters $\nu_k, \mu_k, \rho_k, \eta_k$, is a candidate continuation vector. Mathematically all suitable choices of $t_k$ will be equivalent. However, numerical stability of a rational Krylov subspace method will depend on an appropriate choice and representation of the continuation vector [18, 144]. This vector will be represented in some basis $\mathcal{K}_k(A, v; \Xi_{k-1})$ and its coefficients in that basis is denoted by the vector $\tau_k$. After the expansion, Gram-Schmidt orthogonalization is applied to $(\nu_n A - \mu_n I)^{-1}(\rho_n A - \eta_n I)t_k$, which orthogonalizes with respect to $\mathcal{K}_k(A, v; \Xi_{k-1})$.
Suppose an orthonormal basis $Q_k$ is available for $\mathcal{K}_k(A, v; \Xi_{k-1})$. The rational Arnoldi iteration [144] chooses $t_k = Q_k e_k = q_{k-1}$, i.e., a representation in the orthonormal basis $Q_k$ and $\tau_k = e_k$. Theorem 4.4 shows that a Hessenberg recurrence pencil with specific subdiagonal elements represents an orthonormal basis for any RKS $\mathcal{K}(A, v; \Xi)$, with $\xi_i \in \overline{\mathbb{C}}$.

**Theorem 4.4** (Rational Arnoldi iteration [144])**.** *Consider $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$ and a set of poles $\Xi_k = \{\xi_1, \xi_2, \ldots, \xi_k\} \in \overline{\mathbb{C}}$. Let $Q_{k+1} \in \mathbb{C}^{m\times(k+1)}$ be an orthonormal*

*nested basis for the rational Krylov subspace $\mathcal{K}_{k+1}(A, v; \Xi_k)$, $k < g$, with $g$ the grade of $v$ with respect to $A$. Then $Q_{k+1}$ satisfies*

$$AQ_{k+1}\underline{K}_k = Q_{k+1}\underline{H}_k$$

*for Hessenberg matrices $\underline{H}_k, \underline{K}_k \in \mathbb{C}^{(k+1)\times k}$, with $\frac{(h_k)_{i+1,i}}{(\underline{k}_k)_{i+1,i}} = \xi_i$, $i = 1, 2 \ldots, k$.*

*Proof.* Expansion of $\mathcal{K}_l(A, v; \Xi_{l-1})$ is performed by $(\nu_l A - \mu_l I)^{-1}(\rho_l A - \eta_l I)Q_l e_l = (\nu_l A - \mu_l I)^{-1}(\rho_l A - \eta_l I)q_l$, with $\frac{\mu_l}{\nu_l} = \xi_l$ and parameters such that $\mathcal{K}_l(A, v; \Xi_{l-1}) \cup \{(\nu_l A - \mu_l I)^{-1}(\rho_l A - \eta_l I)q_l\} = \mathcal{K}_{l+1}(A, v; \Xi_l)$. Orthonormalization with respect to $\mathcal{K}_l(A, v; \Xi_{l-1})$, using the basis $Q_l$, provides the coefficients $h_{i,l}$, $i = 1, 2, \ldots, l+1$,

$$h_{l+1,l}q_{l+1} = (\nu_l A - \mu_l I)^{-1}(\rho_l A - \eta_l I)q_l - h_{1l}q_1 - \cdots - h_{ll}q_l. \tag{4.3}$$

Rewriting (4.3) reveals the $l$th column of matrices $\underline{H}_l$ and $\underline{K}_l$

$$(\nu_l A - \mu_l I)h_{l+1,l}q_{l+1} = (\rho_l A - \eta_l I)q_l - (\nu_l A - \mu_l I)\sum_{i=1}^{l} h_{il}q_i$$

$$\nu_l A h_{l+1,l}q_{l+1} + \nu_l A\sum_{i=1}^{l} h_{il}q_i - \rho_l A v_l = -\eta_l q_l + \mu_l \sum_{i=1}^{l} h_{il}q_i + \mu_l h_{l+1,l}q_{l+1}$$

$$A\left((\nu_l \sum_{i=1}^{l+1} h_{il}q_i) - \rho_l q_l\right) = \mu_l \left(\sum_{i=1}^{l+1} h_{il}q_i\right) - \eta_l q_l$$

$$A\nu_l \begin{bmatrix} q_1 & \ldots & q_l & q_{l+1} \end{bmatrix} \begin{bmatrix} h_{1l} \\ \vdots \\ h_{ll} - \rho_l/\nu_l \\ h_{l+1,l} \end{bmatrix} = \mu_l \begin{bmatrix} q_1 & \ldots & q_l & q_{l+1} \end{bmatrix} \begin{bmatrix} h_{1l} \\ \vdots \\ h_{ll} - \eta_l/\mu_l \\ h_{l+1,l} \end{bmatrix}.$$

The last equation reveals that the subdiagonal element ratio is $\frac{\mu_l h_{l+1,l}}{\nu_l h_{l+1,l}} = \frac{\mu_l}{\nu_l} = \xi_l$. Since this is valid for $l \leq k$, the proof is complete. □

Algorithm 4 contains the procedure following from the Hessenberg recurrence pencil in Theorem 4.4. Because of the Hessenberg structure the algorithm relies on an infinite recurrence relation.

Other choices of continuation vectors have been studied by Berljafa and Güttel. They explored optimal and near-optimal continuation pairs (a continuation vector $t_k$ and shift parameters $\rho_k, \eta_k$) [16] and suitable choices of continuation pairs for a parallel rational Arnoldi iteration [19]. For state-of-the-art methods for rational Krylov subspaces we refer to `rktoolbox` [17].

**Algorithm 4** Rational Arnoldi iteration [143]

**Input:** $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$, integer $k < g$, $\Xi_{k-1} = \{\xi_1, \xi_2, \ldots, \xi_{k-1}\}$
**Output:** Orthogonal $Q_{k+1} \in \mathbb{C}^{m\times k+1}$, Hessenberg matrices $\underline{H}_k, \underline{K}_k \in \mathbb{C}^{(k+1)\times k}$ such that $AQ_{k+1}\underline{K}_k = Q_{k+1}\underline{H}_k$.
**procedure** ARNOLDI_ITERATION($A, v, k, \Xi$)
    $q_0 = v/\|v\|$
    **for** $i = 1, 2, \ldots, k$ **do**
        $q_i = (\nu_k A - \mu_k I)^{-1}(\rho_k A - \eta_k I)q_{i-1}$ ▷ with $\frac{\mu_k}{\nu_k} = \xi_k \in \overline{\mathbb{C}}$
        **for** $j = 1, 2, \ldots, i$ **do** ▷ Orthogonalization
            $h_{j,i} = \langle q_i, q_j \rangle_E$
            $q_i = q_i - h_{j,i}q_j$
        **end for**
        $h_{i+1,i} = \|q_i\|$
        $q_i = q_i/h_{i+1,i}$ ▷ Normalization
    **end for**
    $\underline{K}_k = \mathrm{diag}(\nu_1, \ldots, \nu_k)\underline{H}_k - \mathrm{diag}(\rho_1, \ldots, \rho_k)$
    $\underline{H}_k = \mathrm{diag}(\mu_1, \ldots, \mu_k)\underline{H}_k - \mathrm{diag}(\eta_1, \ldots, \eta_k)$
**end procedure**

### Orthogonal projection

The pencil $(H_k, K_k)$ only represents the orthogonal projection of $A$ onto $\mathcal{K}_k(A, v; \Xi)$ if $\xi_k = \infty$, which follows from Corollary 4.2.
If $\xi_k \neq \infty$, then the pencil $(\underline{H}_k, \underline{K}_k)$ can still be used to obtain an orthogonal projection, without explicitly computing $Q_k^H A Q_k$ [16].

## 4.4 Biorthogonal bases

To obtain short recurrence relations for bases of rational Krylov subspaces, two related Krylov subspaces must be generated, one with $A$ and one with $A^H$,

$$\mathcal{K}_k(A, v; \Xi) \text{ and } \mathcal{K}_k(A^H, w; \Theta).$$

The starting vectors must satisfy $\langle v, w \rangle_E \neq 0$. Then biorthonormal nested bases $V_k, W_k \in \mathbb{C}^{m \times k}$ for these spaces exists, i.e., for $i = 1, 2, \ldots, k$

$$\operatorname{span}\{V_i\} = \operatorname{span}\{v_0, v_1, \ldots, v_{i-1}\} = \mathcal{K}_i(A, v; \Xi),$$

$$\operatorname{span}\{W_i\} = \operatorname{span}\{w_0, w_1, \ldots, w_{i-1}\} = \mathcal{K}_i(A^H, w; \Theta),$$

$$W_k^H V_k = I$$

The pair of recurrence relations generating these bases are of the form

$$AV_{k+1}\underline{K}_k = V_{k+1}\underline{H}_k,$$

$$A^H W_{k+1}\underline{\widetilde{K}}_k = W_{k+1}\underline{\widetilde{H}}_k.$$

The goal of this section is to show that the matrices $\underline{H}_k, \underline{K}_k, \underline{\widetilde{H}}_k, \underline{\widetilde{K}}_k$ can be chosen to be tridiagonal matrices, implying short recurrence relations for the biorthonormal bases. First, several possible structures for these matrices are studied in Section 4.4.1 Then, Section 4.4.2 discusses the three term recurrence relation generating biorthogonal bases in detail. The eliminator factorization of tridiagonal pencils is the topic of Section 4.4.3. All the results in this section are new and are published in our paper [165], with the exception of Theorem 4.5, which has been proven before for extended Krylov subspaces [128] and is generalized here to all rational Krylov subspaces.

### 4.4.1 Recurrence relations for biorthonormal bases

The recurrence matrix of the pair of recurrence relations for biorthonormal bases is discussed first. Afterwards, recurrence pencils are derived. This result is novel and generalizes many special cases appearing in the literature.
Throughout this section only full decompositions will be used in the proofs, i.e., recurrence matrices and pencils for bases that span $\mathcal{K}_m(A, v; \Xi) = \mathbb{C}^m$ and $\mathcal{K}_m(A^H, w; \Theta) = \mathbb{C}^m$. This is done for simplicity of the exposition; the proofs simplify for full decompositions.

### Recurrence matrix

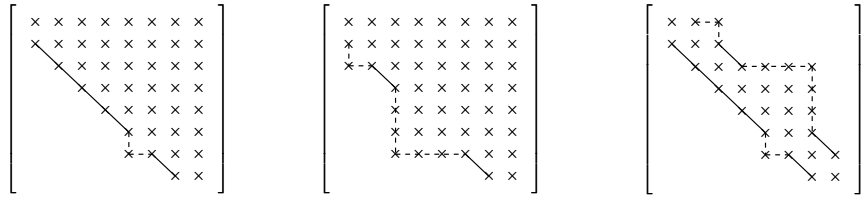
A recurrence matrix for biorthogonal bases has upper and lower triangular rank structures determined by the related recurrence matrices for orthogonal bases. This is illustrated by Example 4.8 and formalized in Theorem 4.5.

**Example 4.8.** *Consider* $A \in \mathbb{C}^{8\times 8}$ *and rational Krylov subspaces with* $\hat{\Xi} = \{\infty, \infty, \infty, \infty, 0, \infty, 0\}$ *and* $\hat{\Theta} = \{0, \infty, 0, 0, 0, \infty, \infty\}$,

$$\mathcal{K}_8(A, v; \hat{\Xi}) = span\{v, Av, A^2v, A^3v, A^4v, A^{-1}v, A^5v, A^{-2}v\},$$

$$\mathcal{K}_8(A^H, w; \hat{\Theta}) = span\{w, (A^H)^{-1}w, A^Hw, (A^H)^{-2}w, (A^H)^{-3}w, (A^H)^{-4}w,$$
$$(A^H)^2w, (A^H)^3w\}.$$

*The structure of the recurrence matrices* $Z_8^V, Z_8^W$ *for orthogonal bases,* $Q_8^V$ *for* $\mathcal{K}_8(A, v; \hat{\Xi})$ *and* $Q_8^W$ *for* $\mathcal{K}_8(A^H, w; \hat{\Theta})$ *is discussed in Section 4.3.4 and shown in Figure 4.4. This figure also shows the recurrence matrix* $Z_8$ *for generating a basis* $V_8$*, for* $\mathcal{K}_8(A, v; \hat{\Xi})$ *that is orthogonal to* $\mathcal{K}_8(A^H, w; \hat{\Theta})$*. The lower and upper triangular structure of* $Z_8$ *corresponds to the lower triangular structure of* $Z_8^V$ *and* $Z_8^W$*, respectively. Black lines are added to highlight the structure.*



(a) $AQ_8^V = Q_8^V Z_8^V$ (b) $A^H Q_8^W = Q_8^W Z_8^W$ (c) $AV_8 = Z_8 V_8$

Figure 4.4: Structure of recurrence matrices or orthogonal and biorthogonal bases for $\mathcal{K}_8(A, v; \hat{\Xi})$ and $\mathcal{K}_8(A^H, w; \hat{\Theta})$ from Example 4.8.

The result, provided in Theorem 4.5, is valid for any configuration of poles in both subspaces. This generalizes results which restrict the allowed poles [41, 112, 152, 179].

**Theorem 4.5** (Recurrence matrix for biorthonormal basis for a rational Krylov subspace)**.** *Consider* $A \in \mathbb{C}^{m\times m}$*,* $v, w \in \mathbb{C}^m$*, with* $\langle v, w\rangle_E \neq 0$ *and* $\Xi, \Theta$*, with* $\xi_i \in \overline{\mathbb{C}}\backslash\sigma(A)$ *and* $\theta_i \in \overline{\mathbb{C}}\backslash\sigma(A^H)$*. Then, under the assumption that no breakdowns occur, nested biorthogonal bases* $V, W$ *for* $\mathcal{K}(A, v; \Xi)$ *and* $\mathcal{K}(A^H, w; \Theta)$*, respectively, satisfy*

$$AV = VZ, \quad A^HW = WZ^H$$

*where the lower triangular structure of $Z$, $Z^H$ is equivalent to the lower triangular structure of the orthogonal recurrence matrix related to $\mathcal{K}(A, v; \Xi)$, $\mathcal{K}(A^H, w; \Theta)$, respectively.*

*Proof.* Consider the matrices $Z^V$ and $Z^W$

$$Z^V = (Q^V)^H A Q^V, \quad Z^W = (Q^W)^H A^H Q^W,$$

where $Q^V$ and $Q^W$ are orthogonal bases for the rational Krylov subspaces $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$, respectively, with $\langle v, w \rangle_E \neq 0$. Use Lemma 2.2 for the Gram matrix $M = (Q^W)^H Q V$:

$$M = (Q^W)^H Q^V = LR$$

$$(\underbrace{Q^W L^{-H}}_{=:W})^H \underbrace{Q^V R^{-1}}_{=:V} = I$$

$$W^H V = I.$$

This decomposition exists under the assumption that $M$ is strongly nonsingular, which is implied by the assumption that no breakdown occurs. The structure of $Z$ can be derived as follows. First consider

$$AV = VZ$$

$$A \underbrace{VR}_{Q^V} = \underbrace{VR}_{Q^V} R^{-1} Z R$$

$$AQ^V = Q^V \underbrace{R^{-1} Z R}_{Z^V}$$

which provides the equality

$$Z = R Z^V R^{-1}. \tag{4.4}$$

Second consider the relations

$$A^H W = W Z^H$$

$$A^H \underbrace{WL^H}_{Q^W} = \underbrace{WL^H}_{Q^W} L^{-H} Z^H L^H$$

$$A^H Q^W = Q^W \underbrace{L^{-H} Z^H L^H}_{Z^W}$$

which provides the equality

$$Z^H = L^H Z^W L^{-H}. \tag{4.5}$$

Multiplication with an upper triangular matrix preserves the structure in the lower triangular part. Hence, from (4.4) it follows that the lower triangular structure of $Z$ is the same as the lower triangular structure of $Z^V$. The upper triangular structure of $Z$ is the same as the lower triangular structure of $Z^W$. This follows from (4.5). □

Note that when all poles are chosen to be infinite, the recurrence matrix is a tridiagonal matrix, which agrees with the results from Section 3.4.

### Recurrence pencil

The recurrence pencil for biorthogonal bases also inherits the structure of recurrence pencils for orthogonal bases. For an orthogonal basis we studied Hessenberg or inverse Hessenberg pencil. Based on these, four important forms in which the recurrence pencil for biorthogonal bases can appear are derived. By *inverted structure*, we mean that the pattern of the core factorization is changed from ascending to descending.

**Lemma 4.4** (Recurrence pencil structure for biorthogonal bases of rational Krylov subspaces)**.** *Consider $A \in \mathbb{C}^{m\times m}$, $v, w \in \mathbb{C}^m$, with $\langle v, w\rangle_E \neq 0$ and $\Xi, \Theta$, with $\xi_i \in \overline{\mathbb{C}}\backslash\sigma(A)$ and $\theta_i \in \overline{\mathbb{C}}\backslash\sigma(A^H)$. Let $Q^V, Q^W \in \mathbb{C}^{m\times m}$ be orthogonal nested bases for $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$, respectively, and corresponding pencils $(H^V, K^V)$, $(H^W, K^W)$ satisfying*

$$AQ^V H^V = Q^V K^V, \qquad A^H Q^W H^W = Q^W K^W.$$

*Then, under the assumption that no breakdown occurs, biorthonormal nested bases $V, W$ for $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$, respectively, satisfy*

$$AVK = VH,$$

$$A^H W\tilde{K} = W\tilde{H},$$

*where the recurrence pencils can be chosen such that*

- *$H$ has the lower triangular structure of $H^V$ and the upper triangular structure equals the inverted upper triangular structure of $K^W$,*
- *$K$ has the same lower triangular structure as $K^V$ and the upper triangular structure equals the inverted lower triangular structure of $H^W$,*
- *$\tilde{H}$ has the lower triangular structure of $H^W$ and the upper triangular structure equals the inverted upper triangular structure of $K^V$,*
- *$\tilde{K}$ has the same lower triangular structure as $K^W$ and the upper triangular structure equals the inverted lower triangular structure of $H^V$.*

*Proof.* From the orthogonal bases $Q^V$ and $Q^W$, the biorthonormal bases $V$ and $W$ can be constructed via Lemma 2.2, i.e., $V := Q^V R^{-1}$ and $W := Q^W L^{-H}$, with $M = (Q^W)^H Q^V = LR$. Substituting the expressions for $V, W$ in the recurrence relations for $Q^V, Q^W$ provides

$$\begin{cases} AQ^V K^V = Q^V H^V \\ A^H Q^W K^W = Q^W H^W \end{cases}$$

$$\Leftrightarrow \begin{cases} AVRK^V = VRH^V \\ A^H W L^H K^W = W L^H H^W \end{cases}$$

$$\Leftrightarrow \begin{cases} W^H AVRK^V = RH^V \\ V^H A^H W L^H K^W = L^H H^W \end{cases}.$$

Taking the Hermitian conjugate of the second equation and rewriting it reveals the connection between the matrices at play

$$\begin{cases} W^H AVRK^V = RH^V \\ W^H AVL^{-1}(H^W)^{-H} = L^{-1}(K^W)^{-H} \end{cases}.$$

Since these expressions are only unique up to right multiplication with a nonsingular matrix $B$, we get

$$RK^V B = L^{-1}(H^W)^{-H}$$

$$RH^V B = L^{-1}(K^W)^{-H}.$$

To obtain a particular choice for the structure of $H$ and $K$ it suffices to represent $B$ in its $RL$-decomposition $B = R_B L_B$ (assuming it exists), where $R_B$ is an upper triangular matrix and $L_B$ a lower triangular matrix,

$$\begin{cases} RK^V B = L^{-1}(H^W)^{-H} \\ RH^V B = L^{-1}(K^W)^{-H} \end{cases}$$

$$\Leftrightarrow \begin{cases} RK^V R_B L_B = L^{-1}(H^W)^{-H} \\ RH^V R_B L_B = L^{-1}(K^W)^{-H} \end{cases}$$

$$\Leftrightarrow \begin{cases} RK^V R_B = L^{-1}(H^W)^{-H} L_B^{-1} =: K \\ RH^V R_B = L^{-1}(K^W)^{-H} L_B^{-1} =: H \end{cases}.$$

For the remainder of this proof $H$ and $K$ are defined as in the last equation. Since $R$ and $R_B$ are upper triangular matrices, they preserve the structure in the lower

triangular part. This means that $K$ and $K^V$ have the same lower triangular structure and so do $H$ and $H^V$. On the other hand $K$ shares its upper triangular structure with $(H^W)^{-H}$ and $H$ with $(K^W)^{-H}$, since $L$ and $L_B$ are lower triangular matrices. An analogous derivation holds for $\tilde{H}, \tilde{K}$. □

The principal result of this section is Theorem 4.6, which states that a tridiagonal recurrence pencil exists describing the construction of biorthogonal bases for rational Krylov subspaces.

**Theorem 4.6** (Tridiagonal pencil for biorthogonal rational Krylov subspaces)**.** *The recurrence pencil* $(H, K)$ *for biorthonormal nested bases* $V, W$ *of* $\mathcal{K}(A, v; \Xi), \mathcal{K}(A^H, w; \Theta)$*, as in Theorem 4.4, can be chosen to be tridiagonal pencils* $(T, S)$ *and* $(\tilde{T}, \tilde{S})$*.*

*Proof.* If $(H_V, K_V)$ is chosen to be a Hessenberg pencil (Theorem 4.4) and $(H_W, K_W)$ to be an inverse Hessenberg pencil (Theorem 4.3), then Lemma 4.4 guarantees that $(T, S)$ has tridiagonal structure. For $(\tilde{T}, \tilde{S})$ the same argument holds by relying on the fact that an inverse Hessenberg pencil representation for $(H_V, K_V)$ can be chosen as well, which is equivalent in the sense that it describes the same recurrence relation. Hence, both representations are valid at the same time. And for $(H_W, K_W)$, the equivalent Hessenberg pencil representation must be chosen. □

The tridiagonal recurrence pencil is only one instance of the possible representations, of which we will distinguish the four principal representations.

### Classification of structures

A classification of the structures of the recurrence matrix and pencils is shown in Table 4.5. The recurrence matrix structure follows from Theorem 4.5. The four recurrence pencils follow from Lemma 4.4 by choosing different representations for the recurrence pencils of the related orthogonal bases that appear in Section 4.3.1 or Section 4.3.3.

A tridiagonal pencil is discussed above and will be elaborated upon further in this manuscript. A *semiseparable pencil* is a pencil of semiseparable matrices. Details on semiseparable matrices can be found in the literature [173–176]. The remaining two recurrence pencils classified in the table are hybrids between tridiagonal and semiseparable. All the matrices appearing in the recurrence pencils belong to the class of *basic matrices*, introduced by Fiedler [66].
The recurrence matrix for non-Hermitian $A$ and extended Krylov subspaces was first studied in the context of inverse eigenvalue problems [128], it is called an *extended tridiagonal matrix*. If $A$ is Hermitian, the recurrence matrix representation, $Z$, is a semiseparable-plus-diagonal matrix [63, Theorem 1] and the connection between

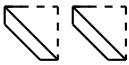
| $W^H AV = Z$ | $W^H AVK = H$ | |
|---|---|---|
|  |  tridiagonal pencil |  semiseparable pencil |
| |  |  |

Table 4.5: Summary of matrix structures appearing for the recurrence coefficients for biorthogonal bases of rational Krylov subspaces, with the biorthonormal bases $V$ and $W$.

a Hermitian semiseparable-plus-diagonal matrix and a tridiagonal pencil has been exploited by Fasino and Gemignani to solve inverse eigenvalue problems [65].

### 4.4.2 Three term recurrence relation

The tridiagonal recurrence pencil from Theorem 4.6 implies that a pair of short recurrence relations exist which generate biorthogonal bases for rational Krylov subspaces. We are interested in short recurrence relations since they can lead to much cheaper algorithms for rational Krylov subspace methods. And, as discussed in Chapter 7, the tridiagonal pencil is related to biorthogonal rational functions, so efficient matrix theoretical algorithms for rational functions can be developed. In order to develop numerical procedures the location of the poles of the rational Krylov subspaces in the pencil must be determined. This knowledge will allow us to develop a Lanczos-type iteration.

#### Location of poles

A tridiagonal recurrence pencil for rational Krylov subspaces reveals the poles of these spaces on the sub-and superdiagonal of the pencil. Lemma 4.5 and Lemma 4.6 state the exact location of the poles in the pencil.

**Lemma 4.5** (Poles on subdiagonal of tridiagonal pencil)**.** *Let $(T, S)$ be the tridiagonal recurrence pencil from Theorem 4.6. The ratio of the subdiagonal elements of $(T, S)$ reveals the poles $\Xi = \{\xi_1, \xi_2, \ldots, \xi_{m-1}\}$ of $\mathcal{K}(A, v; \Xi)$*

$$\frac{t_{i+1,i}}{s_{i+1,i}} = \xi_i, \qquad i = 1, 2, \ldots, m-1. \tag{4.6}$$

*Proof.* Use the variables from the proof of Lemma 4.4. Consider the Hessenberg pencil $(H^V, K^V)$, then, from Theorem 4.4, it follows that the ratio of the subdiagonal elements of $(H^V, K^V)$ equals the poles

$$\frac{h^V_{i+1,i}}{k^V_{i+1,i}} = \xi_i, \quad i = 1, 2, \ldots, m-1.$$

Since, for upper triangular matrices $R$ and $R_B$ we have

$$\frac{t_{i+1,i}}{s_{i+1,i}} = \frac{r_{i+1,i+1} h^V_{i+1,i} (r_B)_{ii}}{r_{i+1,i+1} k^V_{i+1,i} (r_B)_{ii}} = \frac{h^V_{i+1,i}}{k^V_{i+1,i}}, \quad i = 1, 2, \ldots, m-1,$$

the statement follows. In the first equality we used a result stated in the proof of Theorem 4.4. □

**Lemma 4.6** (Poles on superdiagonal of tridiagonal pencil)**.** *Let $(T, S)$ be the tridiagonal pencil from Theorem 4.6. The ratio of the superdiagonal elements of $(T, S)$ reveals the (complex conjugate of the) poles $\Theta = \{\theta_1, \theta_2, \ldots, \theta_{m-2}\}$ of $\mathcal{K}(A^H, w; \Theta)$*

$$\frac{t_{i,i+1}}{s_{i,i+1}} = \bar{\theta}_{i-1}, \qquad i = 2, 3, \ldots, m-1. \tag{4.7}$$

*Proof.* Use the variables from Lemma 4.4 and Theorem 4.6. Another tridiagonal pencil $(\widetilde{T}, \widetilde{S})$ exists for which

$$A^H W \widetilde{S} = W \widetilde{T}. \tag{4.8}$$

Hence, from Lemma 4.5 we know that the ratios of the subdiagonals of $(T, S)$ and $(\widetilde{T}, \widetilde{S})$ reveal the poles of $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$, respectively.
Starting from the tridiagonal pencil $(T, S)$ satisfying $AVS = VT$ and Equation (4.8) we can relate the matrix pencils as follows

$$\begin{cases} W^H A V = T S^{-1} \\ V^H A^H W = \widetilde{T} \widetilde{S}^{-1} \end{cases} \Rightarrow \begin{cases} W^H A V = T S^{-1} \\ W^H A V = \widetilde{S}^{-H} \widetilde{T}^H \end{cases}$$

concluding that $TS^{-1} = \widetilde{S}^{-H} \widetilde{T}^H$. Rewriting this equation as $\widetilde{S}^H T = \widetilde{T}^H S$ leads to two pentadiagonal matrices. Let us assign a variable to each off-diagonal element,

diagonal elements are denoted by $\times$, because these are not relevant for the proof

$$\begin{bmatrix} \times & \tilde{\sigma}_1 & & & \\ \tilde{s}_1 & \times & \tilde{\sigma}_2 & & \\ & \tilde{s}_2 & \times & \ddots & \\ & & \ddots & \ddots & \tilde{\sigma}_{n-1} \\ & & & \tilde{s}_{n-1} & \times \end{bmatrix}^H \begin{bmatrix} \times & \tau_1 & & & \\ t_1 & \times & \tau_2 & & \\ & t_2 & \times & \ddots & \\ & & \ddots & \ddots & \tau_{n-1} \\ & & & t_{n-1} & \times \end{bmatrix} =$$

$$\begin{bmatrix} \times & \tilde{s}_1 & & & \\ \tilde{t}_1 & \times & \tilde{s}_2 & & \\ & \tilde{t}_2 & \times & \ddots & \\ & & \ddots & \ddots & \tilde{s}_{n-1} \\ & & & \tilde{t}_{n-1} & \times \end{bmatrix}^H \begin{bmatrix} \times & \sigma_1 & & & \\ s_1 & \times & \sigma_2 & & \\ & s_2 & \times & \ddots & \\ & & \ddots & \ddots & \sigma_{n-1} \\ & & & s_{n-1} & \times \end{bmatrix}.$$

Equate the second superdiagonals and second subdiagonals of both pentadiagonal matrices

$$\begin{cases} t_i \tilde{\sigma}_{i+1}^H = s_i \tilde{\tau}_{i+1}^H \\ \tau_{i+1} \tilde{s}_i^H = \sigma_{i+1} \tilde{t}_i^H \end{cases}, \qquad i = 1, \ldots, m-2,$$

$$\Rightarrow \begin{cases} \xi_i = t_i / s_i = \tilde{\tau}_{i+1}^H / \tilde{\sigma}_{i+1}^H \\ \psi_i = \tilde{t}_i / \tilde{s}_i = \tau_{i+1}^H / \sigma_{i+1}^H \end{cases}, \qquad i = 1, \ldots, m-2,$$

where the last equality uses Lemma 4.5. □

Note that Lemma 4.6 allows for freedom in the choice of $t_{1,2}$ and $s_{1,2}$, since there is no relation to the given poles. The results from Theorem 4.6, Lemma 4.5 and Lemma 4.6 can be used to construct a pair of three term recurrence relations generating biorthogonal bases for rational Krylov subspaces. This is a Lanczos-type iteration and is derived in Chapter 8.

### 4.4.3 Eliminator factorization

The tridiagonal matrices $T, S$ allow for an eliminator factorization as in Lemma 3.9. The eliminator factorization of the tridiagonal pencil $(T, S)$ equals the pencil of the eliminator factorization of $T$ and the eliminator factorization of $S$. Manipulation of the eliminators appearing in the tridiagonal pencil $(T, S)$ allows us to obtain all the recurrence pencils from Table 4.5. The necessary operations to derive the other recurrence pencils are provided here. These are operations working directly with

eliminators. First note that $\mathfrak{E}_k^l$ is closed under multiplication with $E_k^l \in \mathfrak{E}_k^l$ and under inversion, i.e., for $E_k^l, \hat{E}_k^l \in \mathfrak{E}_k^l$ we have $E_k^l \hat{E}_k^l \in \mathfrak{E}_k^l$ and $\left(E_k^l\right)^{-1} \in \mathfrak{E}_k^l$. The same is valid for $\mathfrak{E}_k^u$.
Lemma 4.7 show how an upper and lower eliminator can switch places by forming the product and factorizing again. An analogue to the turnover of core transformations is provided for eliminators in Lemma 4.8. And finally, the passing through of an eliminator, through a diagonal matrix, is provided in Lemma 4.9.

**Lemma 4.7** (LDR/RDL factorization of a $2 \times 2$ matrix [128])**.** *A strongly nonsingular $2 \times 2$ matrix, that is also strongly nonsingular for all its trailing principal submatrices, has an LDR and RDL factorization consisting of eliminators*

$$\begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix} = \begin{bmatrix} \times & \\ & \times \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix} = \quad .$$

**Lemma 4.8** (Turnover lemma: eliminators [128])**.** *Consider the product of three lower eliminators $E_{i-1}^l E_i^l \hat{E}_{i-1}^l$, where $E_{i-1}^l, \hat{E}_{i-1}^l \in \mathfrak{E}_{i-1}^l$ and $E_i^l \in \mathfrak{E}_i^l$. Then there exist $\epsilon_{i-1}^l \in \mathfrak{E}_{i-1}^l$ and $\epsilon_i^l, \hat{\epsilon}_i^l \in \mathfrak{E}_i^l$ such that*

$$= \quad .$$

$$E_{i-1}^l E_i^l \hat{E}_{i-1}^l = \epsilon_i^l \epsilon_{i-1}^l \hat{\epsilon}_i^l.$$

*The same is valid for upper eliminators $E_{i-1}^u, \hat{E}_{i-1}^u \in \mathfrak{E}_{i-1}^u$, $E_i^u \in \mathfrak{E}_i^u$, $\epsilon_{i-1}^u \in \mathfrak{E}_{i-1}^u$ and $\epsilon_i^u, \hat{\epsilon}_i^u \in \mathfrak{E}_i^u$,*

$$= \quad .$$

$$E_{i-1}^u E_i^u \hat{E}_{i-1}^u = \epsilon_i^u \epsilon_{i-1}^u \hat{\epsilon}_i^u.$$

**Lemma 4.9** (Passing through [128])**.** *Let $D \in \mathbb{C}^{m \times m}$ be a nonsingular diagonal matrix, then the following equalities hold for $E_i^l, \hat{E}_i^l \in \mathfrak{E}_i^l$ and $E_i^u, \hat{E}_i^u \in \mathfrak{E}_i^u$*

$$E_i^l D = D \hat{E}_i^l, \qquad D = D \quad ,$$

$$E_i^u D = D \hat{E}_i^u, \qquad D = D \quad .$$

*Proof.* Since the eliminator only influences a $2 \times 2$ submatrix of $D$, the proof is stated for a $2 \times 2$ matrix and follows immediately from basic linear algebra manipulations:

$$\begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} = \begin{bmatrix} 1 & \\ l & 1 \end{bmatrix} \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} = \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 1 & \\ l d_1/d_2 & 1 \end{bmatrix} = \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \quad .$$

The element $d_2$ is nonzero by the nonsingularity of $D$. Proof for upper eliminators is the same. □

These three lemmas allow for the derivation of all the recurrence pencils and the recurrence matrix appearing in biorthogonal recurrence relations for a rational Krylov subspace. This is omitted here, since the structures have already been derived in the above section.

## 4.5 Special cases

Two important special cases are discussed in detail: the case when $A$ is a Hermitian matrix and the case when it is a unitary matrix. These cases are discussed and they are put into the biorthogonal framework developed throughout this chapter. Rational Krylov subspaces for Hermitian matrices recently received some attention thanks to the underlying short recurrence relations, which provide procedures that are more efficient than Arnoldi-type iterations for certain applications [133]. The tridiagonal recurrence pencil for a Hermitian matrix is discussed in Section 4.5.1. For unitary matrices, Section 4.5.2 discusses a pencil representation for the (Hessenberg) recurrence matrix for polynomial Krylov subspaces. This result is an important example when discussing short recurrence relations in Chapter 5.

### 4.5.1 Hermitian matrix

For a Hermitian matrix $A = A^H$ and real poles, $\Xi \in \mathbb{R}$, an interesting result originating from the study of orthogonal rational functions [53] states that a tridiagonal recurrence relation exists for rational Krylov subspaces $\mathcal{K}(A, v; \Xi)$. A matrix theoretic proof is given by Güttel [98]. Theorem 4.7 states the result for real poles and in Theorem 4.8 we show that this result fits in the biorthogonal framework and that it also holds for complex poles.

**Theorem 4.7** (Hermitian rational Lanczos [53, 98])**.** *Let $A \in \mathbb{C}^{m\times m}$ be Hermitian, i.e., $A^H = A$, $v \in \mathbb{C}^m$ and $\Xi = \{\xi_i\}_{i=1}^{m-1}$ a set of real poles, $\xi_i \in \mathbb{R}$, for all $i$. Then the recurrence relation generating an orthogonal basis $Q_{k+1}$ for $\mathcal{K}_k(A, v)$ is satisfied for a tridiagonal pencil $(\underline{T}_k, \underline{S}_k) \in \mathbb{C}^{(k+1)\times k}$*

$$AQ_{k+1}\underline{S}_k = Q_{k+1}\underline{T}_k.$$

*Proof.* See Appendix A.1 □

A slightly more general result than Theorem 4.7 can be obtained by using a continuation vector, instead of the last $v_j$ to expand the RKS. However this does not change anything to the proof, only the identity matrix $\underline{I}_k$ must be replaced by an upper triangular matrix. Theorem 4.8 starts from a tridiagonal pencil and shows that for any choice of poles $\xi_i \in \overline{\mathbb{C}}$ and $A^H = A$, $v = w$, the biorthonormal bases $V, W$ reduce to a single orthogonal basis $Q$. None of the tridiagonal matrices in the resulting pencil will be Hermitian, this can only be obtained for real poles. This result is stronger than the one stated in Theorem 4.7, dropping the need to restrict to real poles. A similar result is stated in the paper by Deckers and Bultheel [54], who arrive at this result by means of classical analysis. We provide a matrix theoretical approach.

**Theorem 4.8** (Hermitian rational Lanczos: biorthogonal interpretation)**.** *Let $A^H = A$, $v, w \in \mathbb{C}^m$, $\Xi \subset \overline{\mathbb{C}}$ and consider a nested orthonormal basis $Q \in \mathbb{C}^{m\times m}$ for $\mathcal{K}(A, v; \Xi)$. Then $Q$ satisfies*

$$AQS = QT$$

*for tridiagonal matrices $T, S \in \mathbb{C}^{m\times m}$.*

*Proof.* Consider biorthonormal nested bases $V, W$ for $\mathcal{K}(A, v; \Xi)$, $\mathcal{K}(A^H, w; \Xi)$, $A^H = A$ and $w = v$, respectively. From Theorem 4.6 it follows that

$$AVS = VT,$$

$$AW\widetilde{S} = W\widetilde{T},$$

with $T, S, \widetilde{T}, \widetilde{S}$ tridiagonal. Lemma 4.5 and Lemma 4.6 state that the two pencils have the same ratios on the sub-and superdiagonal, $\frac{T_{i+1,i}}{S_{i+1,i}} = \xi_i$, $\frac{\widetilde{T}_{i+1,i}}{\widetilde{S}_{i+1,i}} = \xi_i$, for $i = 1, 2, \ldots, m-1$ $\frac{T_{i,i+1}}{S_{i,i+1}} = \bar{\xi}_{i-1}$, $\frac{\widetilde{T}_{i,i+1}}{\widetilde{S}_{i,i+1}} = \bar{\xi}_{i-1}$, for $i = 2, 3, \ldots, m-1$. Let $B$ be the Krylov basis for $\mathcal{K}(A, v; \Xi) = \mathcal{K}(A^H, w, \Xi)$, then the Cholesky factorization of the corresponding Gram matrix is

$$M = B^H B = R^H R.$$

And $V = BR^{-1}$, $W = BR^{-1}$, therefore $V = W = Q$ and $Q^H Q = I$, i.e., an orthonormal nested basis for $\mathcal{K}(A, v; \Xi)$. Hence,

$$AQS = QT,$$

$$AQ\widetilde{S} = Q\widetilde{T},$$

and the sub-and superdiagonal elements can be altered, while preserving the ratios revealing the poles, such that $\widetilde{T} = T$ and $S = \widetilde{S}$. □

### 4.5.2 Unitary matrices

The recurrence matrix $H$ for an orthonormal basis for $\mathcal{K}_m(U, v)$ for a unitary matrix $U \in \mathbb{C}^{m\times m}$ is unitary. This is stated in Section 3.3.1. Here another approach is taken to illustrate the use of the biorthogonal framework once more and to derive a bidiagonal pencil representation that reveals a short recurrence relation. Theorem 4.9 elaborates on the structure of the recurrence matrix.

**Theorem 4.9** (Recurrence matrix for orthogonal basis of polynomial Krylov subspace: unitary matrix)**.** *Consider a unitary matrix $U \in \mathbb{C}^{m\times m}$, and the polynomial Krylov subspace $\mathcal{K}(U, v)$. Let $V$ form an orthonormal basis for $\mathcal{K}_m(U, v)$. Then, under the assumption that no breakdown occurs, the projection $Z = V^H UV$ has Hessenberg structure below and inverse Hessenberg structure above the diagonal.*

*Proof.* Consider a unitary matrix $U$, $U^{-1} = U^H$ and rational (more precisely extended) Krylov subspaces

$$\mathcal{K}(U, v; \hat{\Xi} = \{\infty, \infty, \dots\}),$$

$$\mathcal{K}(U^H, v; \hat{\Theta} = \{0, 0, \dots\}),$$

with respective orthonormal bases $Q^V$ and $Q^W$. Since $U^{-1} = U^H$, $\mathcal{K}(U, v; \hat{\Xi}) = \mathcal{K}(U^H, v; \hat{\Theta}) = \mathcal{K}(U, v)$ and therefore $Q^V = Q^W =: Q$ implying $(Q^V)^H Q^w = Q^H Q = I$. Hence, they are simultaneously orthogonal and biorthogonal bases.
Using the knowledge from Section 4.3.4 it is clear that the structure of

$$Z^V = (Q^V)^H A Q^V,$$

$$Z^W = (Q^W)^H A^H Q^W,$$

is Hessenberg and inverse Hessenberg, respectively. Theorem 4.5 then states that $Z = (Q^W)^H A Q^V = Q^H A Q$ has Hessenberg structure in its lower triangular part and inverse Hessenberg structure in its upper triangular part. □

The pencil analogue to Theorem 4.9 can be derived from Theorem 4.4. Theorem 4.10 reveals that a two coupled two-term recurrence relations generate an orthonormal basis for $\mathcal{K}(U, v)$. This is equivalent to the recurrence relation for Szegö polynomials [157], which is a classical result in classical analysis. The relation between structured matrices and orthogonal polynomials is discussed in Chapter 6.

**Theorem 4.10** (Recurrence pencil for orthogonal basis of polynomial Krylov subspace: unitary matrix)**.** *Consider a unitary matrix $U \in \mathbb{C}^{m\times m}$ and $\mathcal{K}(U, v)$. Let $V$ form an orthonormal basis for $\mathcal{K}(U, v)$. Then, under the assumption that no breakdown occurs, the equation $V^H UVK = H$ is satisfied for a proper lower-bidiagonal and upper-bidiagonal pencil $(H, K)$.*

*Proof.* Consider a unitary matrix $U$, $U^{-1} = U^H$ and the same subspaces $\mathcal{K}(U, v; \hat{\Xi})$ and $\mathcal{K}(U^H, v; \hat{\Theta})$ as in the proof of Theorem 4.9, with respective orthogonal bases $QV$ and $Q^W$, note that $QV = Q^W =: V$.
The pencil representation of orthogonal projections onto these subspaces are the following

$$(Q^V)^H U Q^V K^V = H^V,$$

$$(Q^W)^H U^H Q^W K^W = H^W.$$

For $(H^V, K^V)$, consider the standard case: $K^V$ is upper triangular and $H^V$ is of Hessenberg form. For $(H^W, K^W)$, choose $H^W$ to be of inverse Hessenberg form and $K^W$ to be upper triangular. Then following from Theorem 4.4 the structure of $(H, K)$ is a lower bidiagonal and upper bidiagonal pencil. □

Theorem 4.10 together with Theorem 4.9 shows that a unitary Hessenberg matrix $Z$ can be factorized as the product of a lower-bidiagonal matrix $H$ and the inverse of an upper-bidiagonal matrix $K$ [176].

## 4.6 Conclusion

Three types of bases for rational Krylov subspaces are discussed. Rational Krylov bases provide theoretical tools to analyze structure. Most notably, we show that the rational Krylov basis can be represented as a product of matrices involving a Vandermonde matrix. Orthogonal bases can be constructed in a numerically stable way. Two important recurrence relations for orthogonal bases are derived in terms of their core factorization. The possible structures of the recurrence matrices and pencils generating biorthogonal bases are derived using the relation between recurrence pencils for orthogonal bases and those for biorthogonal bases. This derivation makes use of the core factorization and leads to a framework able to classify all structured matrices arising from rational Krylov subspaces. As a main result we obtain that a tridiagonal recurrence pencil suffices to generate biorthogonal bases.

# Chapter 5

# Gram matrices

Moments of a matrix $A \in \mathbb{C}^{m\times m}$ with respect to $v, w \in \mathbb{C}^m$,

$$m_j := w^H A^j v, \quad j = 0, 1, \ldots, \tag{5.1}$$

can be used to construct approximations to $A$. This is the idea behind moment matching: an approximation $\hat{A} \in \mathbb{C}^{k\times k}$ is constructed such that its first moments coincide with those of $A$. The moment matching properties of projections onto polynomial Krylov subspaces are well known. These can be obtained by using a functional approach [29, 30, 124] initiated by Vorobyev [177] or by studying Gram matrices. This chapter discusses the latter. Gram matrices contain moments as their entries and might exhibit displacement structure [115]. Gram matrices with displacement structure allow efficient algorithms, e.g., compute their LR factorization by Schur reduction [117]. The LR factorization of a Gram matrix is related to generating (bi)orthogonal vectors. In fact, studying the displacement structure can reveal short recurrence relations for these (bi)orthogonal vectors derived directly from the Gram matrix by the Levinson procedures [176]. For other fast and superfast algorithms and connections to polynomial computation, we refer to the literature [21, 22, 64, 88, 116, 117, 134, 140]. These results are well known for polynomial Krylov subspaces, i.e., the relation between the Lanczos iteration, tridiagonal matrices, Hankel matrices, partial realization and Padé approximation [35, 74, 93, 135, 179]. Similar results are known for Szegö polynomials, unitary Hessenberg matrices and Toeplitz matrices [111, 116].

This chapter elaborates on the connection between rational Krylov subspaces, Gram matrices with a particular displacement structure and multi-point Padé approximation (rational interpolation).

Section 5.1 introduces a generalization of moments, suitable for rational Krylov subspaces, via their appearance in multi-point Padé approximation [73]. These

moments allow us to construct Gram matrices for rational Krylov subspaces in Section 5.2. In this section the displacement structure of all matrices, including Gram matrices, arising from rational Krylov subspaces is analyzed. Section 5.3 elaborates on the connection between the factorization of Gram matrices and how this generates (bi)orthogonal vectors. A short note on the Levinson procedure is provided, which can be used to derive short recurrence relations. This is done in Chapter 8. Most of the results discussed here are known. The general form of the Gram matrices appearing in rational Krylov subspaces and the corresponding displacement structure, see Section 5.2.3, is new.

## 5.1 Moments matching

The $j$th moment $m_j$ of a matrix $A \in \mathbb{C}^{m\times m}$ with respect to $v, w \in \mathbb{C}^m$ contains information about the matrix $A$. Thus it can be used to construct an approximation to $A$. The moment matching technique constructs an approximation $\hat{A}$ such that the first $2k-1$ (or $k$) moments of $\hat{A}$ match those of $A$, i.e.,

$$\hat{w}^H \hat{A}^j \hat{v} = w^H A^j v, \quad j = 0, 1, \ldots, 2k-2, \quad (\text{ or } j = 0, 1, \ldots, k-1).$$

Moment matching for moments of the form (5.1) has an intimate connection to polynomial Krylov subspaces [124]. Moments related to rational Krylov subspaces require a generalization of (5.1). These are introduced by discussing their appearance in system theory.
In system theory moment matching is used to obtain a reduced order model [2]. Consider the state space equations

$$E\dot{x}(t) = Ax(t) + bu(t)$$

$$y(t) = c^H x(t)$$

with $E, A \in \mathbb{C}^{m\times m}$ and $b, c \in \mathbb{C}^m$. A reduced order model is computed, i.e., matrices $\hat{A}, \hat{E} \in \mathbb{C}^{k\times k}$ and vectors $\hat{b}, \hat{c} \in \mathbb{C}^k$, such that

$$\hat{E}\dot{x}(t) = \hat{A}x(t) + \hat{b}u(t)$$

$$y(t) = \hat{c}^H x(t)$$

is a good approximation. The starting vectors for Krylov subspace methods are chosen to be $v = b$, the input vector, and $w = c$, the output vector of the system. The quality of the approximation will be indicated by the transfer functions. How well does the transfer function of the reduced system correspond to the transfer function of the original system? The transfer functions $g(\xi)$ and $\hat{g}(\xi)$ of the original and reduced

system, respectively, are obtained by applying the Laplace transformation [73],

$$g(\xi) := \frac{y(\xi)}{u(\xi)} = c^H(\xi E - A)^{-1}b,$$

$$\hat{g}(\xi) := \frac{\hat{y}(\xi)}{\hat{u}(\xi)} = \hat{c}^H(\xi \hat{E} - \hat{A})^{-1}\hat{b}.$$

The approximate transfer function $\hat{g}(\xi)$ should match the first coefficients of the power series expansion of $g(\xi)$. Partial realization [2] uses the power series approximation around a single frequency $\xi$, typically $\xi = 0$ and for $m_j = -c^H(A^{-1}E)^j A^{-1}b$ the expansion is $g(\xi) = m_0 + m_1\xi + \frac{\xi^2}{2!}m_2 + \frac{\xi^3}{3!}m_3 + \cdots$ or $s = \infty$ and for $m_{-j} = c^H(E^{-1}A)^{(j-1)}E^{-1}b$ we have $g(\xi) = m_{-1}\xi^{-1} + m_{-2}\xi^{-2} + m_{-3}\xi^{-3} + \cdots$. An effective numerical approach to compute the reduced order model is the Lanczos iteration, which avoids explicit computation of the moments [93, 95].

Approximation around a single frequency can lead to slow convergence far away from this frequency, requiring a reduced model of large size to obtain a good approximation. A possible approach to avoid this limitation is to approximate around multiple frequencies. This approach is proposed by Gallivan et al. [73] and corresponds to multi-point Padé approximation and to projections onto a rational Krylov subspace. In terms of the transfer functions, their approach is to construct $\hat{g}(\xi)$ such that it matches coefficients of multiple power series expansions of $g(\xi)$ around different frequencies $\{\xi_1, \xi_2, \ldots, \xi_i\}$. A rational generalization of the Lanczos iteration is a good candidate to construct multi-point Padé approximations. A suitable rational Lanczos iteration is proposed in Chapter 8 and the difference with the approach of Gallivan and coauthors [73] is discussed.

Now we introduce the moments that arise from this approach and that will be used throughout this manuscript. The power series expansion around $\hat{\imath}$ frequencies, corresponding to rational Krylov subspaces, is

$$g(\xi) = \sum_{i=1}^{\hat{\imath}} \sum_{j=0}^{\infty} m_j(\xi_i) \frac{(\xi - \xi_i)^j}{j!}, \quad m_j(\xi_i) = -c^H(\xi_i E - A)^{-j} E^j (\xi_i E - A)^{-1} b.$$

Therefore, moments of the form

$$m_j(\xi) = -c^H(A - \xi I)^{-j-1}b \tag{5.2}$$

occur naturally in model order reduction.

These moments (5.2) appear in the Gram matrices associated with rational Krylov subspaces. The Gram matrix $M_k \in \mathbb{C}^{k\times k}$ for a single subspace $\mathcal{K}_k(A, v; \Xi)$ is

$$M_k = (B_k^{\Xi})^H B_k^{\Xi},$$

with the Krylov basis $B_k^{\Xi}$, $\text{span}\{B_k^{\Xi}\} = \mathcal{K}_k(A, v; \Xi)$. In general this matrix contains $k$ moments, and if $A^H = A$ it contains $2k - 1$ moments.

For the pair of subspaces $\mathcal{K}_k(A, v; \Xi)$ and $\mathcal{K}_k(A^H, w; \Theta)$, the Gram matrix is

$$M_k = (B_k^{\Theta})^H B_k^{\Xi},$$

with Krylov bases $B_k^{\Xi}$ and $B_k^{\Theta}$, $\text{span}\{B_k^{\Xi}\} = \mathcal{K}_k(A, v; \Xi)$ and $\text{span}\{B_k^{\Theta}\} = \mathcal{K}_k(A^H, w; \Theta)$. This Gram matrix always contains $2k - 1$ moments.

## 5.2 Displacement structure

Gram matrices arise in many applications [116, 134], and often have a certain displacement structure. For matrices with low displacement rank specialized algorithms can be developed: fast algorithms with complexity $\mathcal{O}(m^2)$, and superfast algorithms, with complexity $\mathcal{O}(m(\log m)^2)$ [99, 116, 134]. Krylov matrices and rational Krylov matrices have low displacement rank. Gram matrices arising from Krylov subspaces are composed of the product of two Krylov matrices and might exhibit some displacement structure as well. The analysis of the displacement structure of Gram matrices is the main topic of this section.
The following discussion relies heavily on two well-known matrices with displacement structure, the Vandermonde and Hankel matrices. Section 5.2.1 introduces these and shows how they are related to each other. Polynomial Krylov matrices are intimately related to these matrices. The relations are discussed in Section 5.2.2. Section 5.2.3 generalizes the approach used for polynomial Krylov to Gram matrices for rational Krylov subspaces. A procedure to construct suitable displacement operators for any Gram matrix, arising from rational Krylov subspaces, is described.

### 5.2.1 Basics

Two 'classical' matrices with displacement structure are central to the analysis of displacement structure in Krylov matrices. These are the Vandermonde and Hankel matrix. A Vandermonde matrix $V_k \in \mathbb{C}^{m \times k}$ generated by $\{\lambda_i\}_{i=1}^m$ is

$$V_k = \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^{k-1} \end{bmatrix}.$$

A Vandermonde matrix exhibits low displacement rank, see Lemma 5.1.

**Lemma 5.1** (Displacement rank: Vandermonde matrix [134])**.** *A Vandermonde matrix*

$$V_k = \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^{k-1} \end{bmatrix}$$

*has displacement rank smaller than or equal to 1.*

*Proof.* The only difficulty is to find suitable displacement operators. These are the diagonal matrix $\Lambda = \operatorname{diag}(\{\lambda_i\}_{i=1}^m) \in \mathbb{C}^{m\times m}$ and leftshift matrix $Z_k \in \mathbb{C}^{k\times k}$. Then

$$\Lambda V_k - V_k Z_k = \begin{bmatrix} & \lambda_1^k \\ 0 & \vdots \\ & \lambda_m^k \end{bmatrix},$$

which has rank 1. □

A Hankel matrix $H_k \in \mathbb{C}^{k\times k}$ has constant anti-diagonals

$$H_k = \begin{bmatrix} h_0 & h_1 & h_2 & \dots & h_{k-1} \\ h_1 & h_2 & h_3 & \dots & h_k \\ h_2 & h_3 & h_4 & \dots & h_{k+1} \\ \vdots & \vdots & \vdots & \iddots & \vdots \\ h_{k-1} & h_k & h_{k+1} & \dots & h_{2k-2} \end{bmatrix}.$$

The Hankel matrices of interest here are formed by Vandermonde matrices, for some diagonal matrix $D = \operatorname{diag}(d_1, \dots, d_m)$,

$$H_k = V_k^\top D V_k.$$

This relation is useful to link properties of both matrices, such as condition number, to each other [107, 159]. Hankel matrices have low displacement rank, as stated in Lemma 5.2.

**Lemma 5.2** (Displacement rank: Hankel matrix [134])**.** *A Hankel matrix*

$$H_k = \begin{bmatrix} h_0 & h_1 & h_2 & \dots & h_{k-1} \\ h_1 & h_2 & h_3 & \dots & h_k \\ h_2 & h_3 & h_4 & \dots & h_{k+1} \\ \vdots & \vdots & \vdots & \iddots & \vdots \\ h_{k-1} & h_k & h_{k+1} & \dots & h_{2k-2} \end{bmatrix}$$

*has displacement rank smaller than or equal to 2.*

*Proof.* Consider the leftshift matrix $Z_k \in \mathbb{C}^{k\times k}$, then

$$Z_k^H H_k - H_k Z_k = \begin{bmatrix} & & & h_k \\ & 0 & & \vdots \\ & & & h_{2k-2} \\ -h_k & \dots & -h_{2k-2} & 0 \end{bmatrix},$$

which has rank at most 2. □

The displacement structure of $H_k = V_k^\top D V_k$ can be explained by the displacement rank of $V_k$. Consider the relations that exist for $DV_k$ and $V_k^\top$ for the leftshift matrix $Z_k \in \mathbb{C}^{m\times m}$ and $\Lambda := \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$,

$$\Lambda D V_k - D V_k Z_k = \begin{bmatrix} & d_1\lambda_1^k \\ 0 & \vdots \\ & d_m\lambda_m^k \end{bmatrix}, \quad Z_k^\top V_k^\top - V_k^\top \Lambda^\top = \begin{bmatrix} & 0 & \\ -\lambda_1^k & \ldots & -\lambda_m^k \end{bmatrix}.$$

Use the second relation, postmultiply by $DV_k$, and use the first relation

$$Z_k^\top V_k^\top D V_k - V_k^\top \Lambda^\top D V_k = \begin{bmatrix} & 0 & \\ -\lambda_1^k & \ldots & -\lambda_m^k \end{bmatrix} D V_k$$

$$Z_k^\top H_k - V_k^\top D \Lambda V_k = \begin{bmatrix} & 0 & \\ -\sum_i d_i\lambda_i^k & \ldots & -\sum_i d_i\lambda_i^{2k-1} \end{bmatrix}$$

$$Z_k^\top H_k - H_k Z_k = \begin{bmatrix} & & \sum_i d_i\lambda_i^k \\ & 0 & \vdots \\ & & \sum_i d_i\lambda_i^{2k-2} \\ -\sum_i d_i\lambda_i^k & \ldots & -\sum_i d_i\lambda_i^{2k-2} \quad 0 \end{bmatrix}.$$

### 5.2.2 Polynomial Krylov subspaces

Corollary 5.1 states that a Krylov matrix $B_k$ forming the Krylov basis for $\mathcal{K}_k(A, v)$ with a diagonalizable matrix $A$ has low displacement rank. This follows from its relation to a Vandermonde matrix.

**Corollary 5.1.** *Consider the Krylov matrix $B_k \in \mathbb{C}^{m\times k}$ of a Krylov subspace $\mathcal{K}_k(A, v)$ for a diagonalizable matrix $A \in \mathbb{C}^{m\times m}$. Then $B_k$ has displacement rank less than or equal to one.*

*Proof.* Start from Lemma 5.1 and adjust the displacement matrices, based on the factorization of $B_k$ used in the proof of Theorem 3.1. Consider the eigendecomposition $A = X\Lambda X^{-1}$ and, for $\alpha_i := e_i^\top(X^{-1}v)$, $B_k = X \underbrace{\begin{bmatrix}\alpha_1 & & \\ & \ddots & \\ & & \alpha_m\end{bmatrix}}_{=:W} \underbrace{\begin{bmatrix}1 & \lambda_1 & \dots & \lambda_1^{k-1}\\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^{k-1}\end{bmatrix}}_{=:V_k}$.

Then

$$XW\Lambda V_k = X\Lambda X^{-1}XWV_k = AB_k$$

and

$$AB_k - B_kZ_k = XW(\Lambda V_k - V_kZ_k) = XW\begin{bmatrix}\lambda_1^k \\ \vdots \\ \lambda_m^k\end{bmatrix} e_k^\top,$$

which has rank equal to 1. When $k = g$, the grade of $v$ with respect to $A$, then it follows from Corollary 3.2 that $AB_g - B_gC$ has rank 0, with $C$ the companion matrix. □

Note that Corollary 5.1 describes a property that follows immediately from the Krylov recurrence relation in Lemma 3.1.

#### Orthogonal basis

Since a Krylov matrix $B_k$ is related to a Vandermonde matrix, the associated Gram matrix $M_k = B_k^H B_k$ might be related to a Hankel matrix. Lemma 5.3 states that the Gram matrix $M_k$ for a Krylov subspace has Hankel structure when the space is generated for a Hermitian matrix $A^H = A$.

**Lemma 5.3** (Gram matrix: single polynomial Krylov subspace)**.** *The Gram matrix $M_k$ for $\mathcal{K}_k(A, v)$ is of the form $H_k = V_k^\top DV_k$ if and only if $A^H = A$.*

*Proof.* Let $A = X\Lambda X^{-1}$, and $B_k = \left[v, Av, \ldots, A^{k-1}v\right] = XWV_k$ as in the proof of Corollary 5.1, then

$$M_k = B_k^H B_k = V_k^H W^H X^H X W V_k.$$

For a Hermitian matrix A one has $X^H X = I$ and $D = W^H W \in \mathbb{R}^{m\times m}$ is diagonal and positive semidefinite, thus

$$M_k = V_k^H D V_k.$$

Hence, $M_k = V_k^\top D V_k$ if and only if $V_k^\top = V_k^H$ or, equivalently, $\bar{\lambda}_i = \lambda_i$. □

Let us look at the moments appearing in the Gram matrix $M_k$ for a Krylov subspace $\mathcal{K}(A, v)$. The inner product on Krylov subspaces is the Euclidean inner product $\langle ., . \rangle_E$. For $A^H = A$ the property $\langle Ax, y\rangle_E = \langle x, Ay\rangle_E$ is satisfied. Then $M_k$ is the Hankel matrix built with the first $2k-1$ moments of $A$, $m_j = v^H A^j v = \langle A^j v, v\rangle_E$,

$$M_k = \begin{bmatrix} \langle v,v\rangle_E & \langle Av,v\rangle_E & \langle A^2v,v\rangle_E & \ldots & \langle A^{k-1}v,v\rangle_E \\ \langle v,Av\rangle_E & \langle Av,Av\rangle_E & \langle A^2v,Av\rangle_E & \ldots & \langle A^{k-1}v,Av\rangle_E \\ \langle v,A^2v\rangle_E & \langle Av,A^2v\rangle_E & \langle A^2v,A^2v\rangle_E & \ldots & \langle A^{k-1}v,A^2v\rangle_E \\ \vdots & \vdots & \vdots & & \vdots \\ \langle v,A^{k-1}v\rangle_E & \langle Av,A^{k-1}v\rangle_E & \langle A^2v,A^{k-1}v\rangle_E & \ldots & \langle A^{k-1}v,A^{k-1}v\rangle_E \end{bmatrix}$$

$$= \begin{bmatrix} \langle v,v\rangle_E & \langle Av,v\rangle_E & \langle A^2v,v\rangle_E & \ldots & \langle A^{k-1}v,v\rangle_E \\ \langle Av,v\rangle_E & \langle A^2v,v\rangle_E & \langle A^3v,v\rangle_E & \ldots & \langle A^{k}v,v\rangle_E \\ \langle A^2v,v\rangle_E & \langle A^3v,v\rangle_E & \langle A^4v,v\rangle_E & \ldots & \langle A^{k+1}v,v\rangle_E \\ \vdots & \vdots & \vdots & & \vdots \\ \langle A^{k-1}v,v\rangle_E & \langle A^kv,v\rangle_E & \langle A^{k+1}v,v\rangle_E & \ldots & \langle A^{2k-2}v,v\rangle_E \end{bmatrix}.$$

For nonnormal matrices $A^H A \neq AA^H$, only the first $k$ moments of the matrix $A$ appear in $M_k$,

$$M_k = \begin{bmatrix} \langle v,v\rangle_E & \langle Av,v\rangle_E & \ldots & \langle A^{k-1}v,v\rangle_E \\ \langle v,Av\rangle_E & \langle Av,Av\rangle_E & \ldots & \langle A^{k-1}v,Av\rangle_E \\ \vdots & \vdots & & \vdots \\ \langle v,A^{k-1}v\rangle_E & \langle Av,A^{k-1}v\rangle_E & \ldots & \langle A^{k-1}v,A^{k-1}v\rangle_E \end{bmatrix}$$

$$= \begin{bmatrix} \langle v,v\rangle_E & \langle Av,v\rangle_E & \ldots & \langle A^{k-1}v,v\rangle_E \\ \langle A^Hv,v\rangle_E & \langle A^HAv,v\rangle_E & \ldots & \langle A^HA^kv,v\rangle_E \\ \vdots & \vdots & & \vdots \\ \langle (A^H)^{k-1}v,v\rangle_E & \langle (A^H)^{k-1}Av,v\rangle_E & \ldots & \langle (A^H)^{k-1}A^{k-1}v,v\rangle_E \end{bmatrix}.$$

The terms in $M_k$, where $A^H$ appears in the first term of the inner product, do not have a connection to moment matching of the matrix $A$ [124, p.153].

For normal matrices $A^H A = AA^H$, or equivalently $A^H = p_l(A)$ for some polynomial $p_l$ of degree $l$, the Gram matrix has a displacement rank at most $l$. This is a result on short recurrence relations by Faber and Manteuffel [62].

### Biorthogonal bases

Two Krylov matrices, $B_k$ and $\tilde{B}_k$, for $\mathcal{K}(A, v)$ and $\mathcal{K}(A^H, w)$, respectively, lead to a Gram matrix with Hankel structure $H_k = V_k^\top D V_k$. Lemma 5.4 formalizes this statement.

**Lemma 5.4** (Gram matrix: pair of polynomial Krylov subspaces)**.** *Let $A$ be a diagonalizable matrix. Consider Krylov subspaces $\mathcal{K}(A, v)$ and $\mathcal{K}(A^H, w)$, with $\langle v, w\rangle_E \neq 0$ and corresponding Krylov matrices $B_k$ and $\tilde{B}_k$. Then the associated Gram matrix $M_k = \tilde{B}_k^H B_k$ is a Hankel matrix of the form $V_k^\top D V_k$, with diagonal matrix $D$.*

*Proof.* Write out the columns of the Krylov matrices, and use $\tilde{V}_k = \overline{V}_k$, then

$$M_k = \tilde{B}_k^H B_k = \tilde{V}_k^H \tilde{W}^H X^{-1} X W V_k = V_k^\top D V_k,$$

with $D = \tilde{W}^H W$ diagonal and indefinite. □

The Gram matrix $M_k = \tilde{B}_k^H B_k$ contains the first $2k-1$ moments of the matrix $A$, $m_j = w^H A^j v = \langle Av, w\rangle_E$,

$$M_k = \begin{bmatrix} \langle v, w\rangle_E & \langle Av, w\rangle_E & \dots & \langle A^{k-1}v, w\rangle_E \\ \langle v, A^H w\rangle_E & \langle Av, A^H w\rangle_E & \dots & \langle A^{k-1}v, A^H w\rangle_E \\ \vdots & \vdots & & \vdots \\ \langle v, (A^H)^{k-1} w\rangle_E & \langle Av, (A^H)^{k-1} w\rangle_E & \dots & \langle A^{k-1}v, (A^H)^{k-1} w\rangle_E \end{bmatrix}$$

$$= \begin{bmatrix} \langle v, w\rangle_E & \langle Av, w\rangle_E & \langle A^2 v, w\rangle_E & \dots & \langle A^{k-1}v, w\rangle_E \\ \langle Av, w\rangle_E & \langle A^2 v, w\rangle_E & \langle A^3 v, w\rangle_E & \dots & \langle A^k v, w\rangle_E \\ \langle A^2 v, w\rangle_E & \langle A^3 v, w\rangle_E & \langle A^4 v, w\rangle_E & \dots & \langle A^{k+1} v, w\rangle_E \\ \vdots & \vdots & \vdots & & \vdots \\ \langle A^{k-1} v, w\rangle_E & \langle A^k v, w\rangle_E & \langle A^{k+1} v, w\rangle_E & \dots & \langle A^{2k-2} v, w\rangle_E \end{bmatrix},$$

thanks to the fact that the second Krylov subspace is constructed using $A^H$. Whereas in the orthogonal case, we needed $A^H = A$ to obtain this property.
The low displacement rank of $M_k$ for any diagonalizable matrix $A$ explains the existence of a pair of short recurrence relations for biorthogonal bases for the Krylov subspaces $\mathcal{K}_k(A, v)$ and $\mathcal{K}_k(A^H, w)$.

### 5.2.3 Rational Krylov subspaces

A rational Krylov matrix $B_k^{\Xi} \in \mathbb{C}^{m\times k}$, related to the rational Krylov subspace $\mathcal{K}(A, v; \Xi)$ for a diagonalizable $A \in \mathbb{C}^{m\times m}$, is also related to a Vandermonde matrix $V_k \in \mathbb{C}^{m\times k}$. The relation between $B_k^{\Xi}$ and $V_k$ is given by

$$B_k^{\Xi} = XW\Phi V_k J\hat{L}LP^{\top}, \tag{5.3}$$

with the variables that appeared in the proof of Theorem 4.1:

- $X$ the eigenvector matrix of $A$, i.e., $A = X\Lambda X^{-1}$ with $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$,
- $W = \operatorname{diag}(\{\alpha_i\}_i)$, where $\alpha_i := e_i^{\top}(X^{-1}v)$, $i = 1, 2, \ldots, m$,
- $\Phi = \operatorname{diag}(\{\phi_{k^-}(\lambda_i)\})$ and nonsingular, where $\phi_{k^-}(z)$ is the denominator of highest degree appearing in $\mathcal{K}_k(A, v; \Xi)$ (see Definition 4.2),
- $J = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$,
- $\hat{L}$ and $L$ are unit lower triangular matrices,
- $P$ is a permutation matrix which orders the columns of $B_k^{\Xi}$ by decreasing degree.

The low displacement rank of rational Krylov matrices is stated in Corollary 5.2.

**Corollary 5.2.** *Consider the rational Krylov matrix $B_k^{\Xi}$ of a rational Krylov subspace $\mathcal{K}_k(A, v; \Xi)$ for a diagonalizable matrix $A \in \mathbb{C}^{m\times m}$. Then $B_k^{\Xi}$ has displacement rank less than or equal to one.*

*Proof.* Start from Lemma 5.1 and adjust the displacement matrices, using the decomposition (5.3). Consider the eigendecomposition $A = X\Lambda X^{-1}$ and, for $\alpha_i := e_i^{\top}(X^{-1}v)$, $B_k^{\Xi} = X \underbrace{\begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_m \end{bmatrix}}_{=:W} \Phi \underbrace{\begin{bmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \ldots & \lambda_m^{k-1} \end{bmatrix}}_{=:V_k} J\hat{L}LP$. Since $S := J\hat{L}LP$ is nonsingular, we have

$$XW\Phi(\Lambda V_k - V_k Z_k)S = AXW\Phi V_k - XW\Phi V_k S\tilde{Z}_k,$$

where $\tilde{Z}_k = S^{-1}Z_k S$. Thus the displacement matrices $A$ and $\tilde{Z}_k$ reveal the displacement rank, which is smaller than or equal to 1 since $\operatorname{rank}(\Lambda V_k - V_k Z_k) \leq 1$. □

Note that the $\tilde{Z}_k$ in the proof of Corollary 5.2 is the single matrix representation for the Krylov recurrence relation, i.e., a rational Hessenberg matrix.

### Orthogonal basis

A Gram matrix related to a single rational Krylov subspace $\mathcal{K}(A, v; \Xi)$ with $A^H = A$ has low displacement rank, see Lemma 5.5.

**Lemma 5.5** (Gram matrix: single rational Krylov subspace)**.** *Let $B_k^\Xi \in \mathbb{C}^{m\times k}$ be the Krylov matrix of $\mathcal{K}(A, v; \Xi)$ for a matrix $A \in \mathbb{C}^{m\times m}$. The associated Gram matrix $M_k = (B_k^\Xi)^H B_k^\Xi$ has displacement rank less than or equal to 2 if and only if $A^H = A$.*

*Proof.* The proof of Lemma 5.3 holds when $D = \Phi^H W^H W \Phi$ is chosen, and therefore

$$\begin{aligned} M_k = (B_k^\Xi)^H B_k^\Xi &= PL^H \hat{L}^H J V_k^H \Phi^H W^H X^H X W \Phi V_k J \hat{L} L P^\top \\ &= PL^H \hat{L}^H J V_k^\top \underbrace{\Phi^H W^H W \Phi}_{=:D} V_k J \hat{L} L P^\top \\ &= PL^H \hat{L}^H J H_k J \hat{L} L P^\top. \end{aligned}$$

Thus, $M_k$ has displacement rank less than or equal to 2 since $H_k$ has displacement rank less than or equal to 2. □

The Gram matrix for $\mathcal{K}(A, v; \Xi)$, with $A^H = A$, hides a Hankel matrix of the form $V_k^\top D V_k$. The diagonal matrix $D$ now includes the matrix $\Phi$. This points to a rational modification of the underlying inner product [108, 192]. Chapter 7 discusses the rational modification of inner products.

### Biorthogonal bases

For the biorthogonal setting, i.e., a Gram matrix $M_k = (B_k^\Theta)^H B_k^\Xi$, with $B_k^\Xi$ a nested basis for $\mathcal{K}_k(A, v; \Xi)$ and $B_k^\Theta$ for $\mathcal{K}_k(A^H, w; \Theta)$, always has low displacement rank. In Lemma 5.6 the poles $\Xi$ and $\Theta$ can be chosen completely independent from each other.

**Lemma 5.6** (Gram matrix: pair of rational Krylov subspaces)**.** *Let $A \in \mathbb{C}^{m\times m}$ be a diagonalizable matrix. Consider $\mathcal{K}_k(A, v; \Xi)$ and $\mathcal{K}_k(A^H, w; \Theta)$, with $\langle v, w\rangle_E \neq 0$ and corresponding Krylov matrices $B_k^\Xi, B_k^\Theta \in \mathbb{C}^{m\times k}$. Then the associated Gram matrix $M_k = (B_k^\Theta)^H B_k^\Xi$ has displacement rank less than or equal to 2.*

*Proof.* The decompositions as in Equation (5.3) will be used, where $\bar{V}_k$ is the complex conjugate of $V_k$:

$$B_k^{\Xi} = XW\Phi V_k J \underbrace{\hat{L}L}_{=:L^{\Xi}} P^{\top},$$

$$B_k^{\Theta} = X^{-H}\widetilde{W}\widetilde{\Phi}\bar{V}_k J \underbrace{\hat{L}L}_{=:L^{\Theta}} \widetilde{P}^{\top}.$$

The Gram matrix is

$$M_k = (B_k^{\Theta})^H B_k^{\Xi} = \widetilde{P}(L^{\Theta})^H J V_k^{\top} \widetilde{\Phi}^H \widetilde{W}^H \underbrace{X^{-1}X}_{=I} W\Phi V_k J L^{\Xi} P^{\top}$$

and setting $D := \widetilde{\Phi}^H \widetilde{W}^H W\Phi$ reveals the Hankel matrix $H_k = V_k^{\top} D V_k$ in the expression

$$M_k = \widetilde{P}(L^{\Theta})^H J H_k J L^{\Xi} P^{\top}.$$

Starting from the displacement operators for $H_k$, leftshift $Z_k$ such that $Z_k^H H_k - H_k Z_k$ has rank 2, the displacement operators $Z_k^{\Theta}$, $Z_k^{\Xi}$ for $M_k$ are determined:

$$\begin{aligned}
&\widetilde{P}(L^{\Theta})^H J(Z_k^H H_k - H_k Z_k) J L^{\Xi} P^{\top} \\
&= \widetilde{P}(L^{\Theta})^H J Z_k^H H_k J L^{\Xi} P^{\top} - \widetilde{P}(L^{\Theta})^H J H_k Z_k J L^{\Xi} P^{\top} \\
&= (Z_k^{\Theta})^H \widetilde{P}(L^{\Theta})^H J H_k J L^{\Xi} P^{\top} - \widetilde{P}(L^{\Theta})^H J H_k J L^{\Xi} P^{\top} Z_k^{\Xi}
\end{aligned}$$

Hence, $M_k$ has displacement rank less than or equal to 2 for the displacement operators

$$Z_k^{\Xi} = P(L^{\Xi})^{-1} J Z J L^{\Xi} P^{\top}, \qquad Z_k^{\Theta} = \widetilde{P}(L^{\Theta})^{-1} J Z J L^{\Theta} \widetilde{P}^{\top}.$$

These exist since $L^{\Xi}$ and $L^{\Theta}$ are unit lower triangular matrices and $P, \widetilde{P}$ permutation matrices. □

This result suggests short (three term) recurrence relations for biorthogonal bases for rational Krylov subspaces, i.e., exactly the result obtained in Section 4.4.2.

### Toeplitz matrix

The result from Lemma 5.6 allows us to derive a well known result for Toeplitz matrices. Displacement rank was first used for Toeplitz matrices [115], and later generalized to, among others, Hankel, Vandermonde, and Cauchy matrices [134]. Toeplitz matrices arise naturally in many applications [116] and also as the Gram matrix for a polynomial

Krylov subspace $\mathcal{K}(U,v)$, with a unitary matrix $U \in \mathbb{C}^{m\times m}$. Consider $B_k$, the Krylov matrix for $\mathcal{K}_k(U,v)$, and its Gram matrix

$$M_k = B_k^H B_k = \begin{bmatrix} \langle v,v\rangle_E & \langle Uv,v\rangle_E & \dots & \langle U^{k-1}v,v\rangle_E \\ \langle v,Uv\rangle_E & \langle Uv,Uv\rangle_E & \dots & \langle U^{k-1}v,Uv\rangle_E \\ \vdots & \vdots & \ddots & \vdots \\ \langle v,U^{k-1}v\rangle_E & \langle Uv,U^{k-1}v\rangle_E & \dots & \langle U^{k-1}v,U^{k-1}v\rangle_E \end{bmatrix}$$

$$= \begin{bmatrix} \langle v,v\rangle_E & \langle Uv,v\rangle_E & \langle U^2v,v\rangle_E & \dots & \langle U^{k-1}v,v\rangle_E \\ \langle U^{-1}v,v\rangle_E & \langle v,v\rangle_E & \langle Uv,v\rangle_E & \dots & \langle U^{k-2}v,v\rangle_E \\ \langle U^{-2}v,v\rangle_E & \langle U^{-1}v,v\rangle_E & \langle v,v\rangle_E & \dots & \langle U^{k-3}v,v\rangle_E \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle U^{-(k-1)}v,v\rangle_E & \langle U^{-(k-2)}v,v\rangle_E & \langle U^{-(k-3)}v,v\rangle_E & \dots & \langle v,v\rangle_E \end{bmatrix}$$

$$= \begin{bmatrix} t_0 & t_1 & t_2 & \dots & t_{k-1} \\ t_{-1} & t_0 & t_1 & \dots & t_{k-2} \\ t_{-2} & t_{-1} & t_0 & \dots & t_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{-(k-1)} & t_{-(k-2)} & t_{-(k-3)} & \dots & t_0 \end{bmatrix} =: T_k.$$

A unitary matrix is normal, $U^HU = I = UU^H$, however, the polynomial Krylov interpretation does not allow us to conclude that $T_k$ has displacement rank 2 or even low displacement rank. However, the isometric Arnoldi provides coupled two-term recurrence relations which construct an orthogonal basis for $\mathcal{K}_k(U,v)$ [111, 188]. This suggests that $T_k$ does have low displacement rank, namely,

$$\operatorname{rank}\left(JZ_k^H JT_k - T_k Z_k\right) \le 2. \tag{5.4}$$

The reason is that restricting to polynomial Krylov subspaces corresponds to restricting the recurrence relation to be of the form related to $A^H = p(A)$, i.e.,

$$Av_k = \sum_{i=k-l}^{k+1} c_i v_i,$$

where $l+1$ is the amount of terms in the recurrence relation.
A recurrence relation corresponding to rational Krylov subspaces is of the form

$$\sum_{i=k-\tilde{l}}^{k+1} \tilde{c}_i Av_i = \sum_{i=k-l}^{k+1} c_i v_i.$$

This is related to $A^H = r_l(A)$, where the rational function $r_l(z) = \frac{p_l(z)}{q_l(z)}$, with $p_l, q_l$ polynomials of at most degree $l$ [123]. For the above example, $U^H = r_1(U)$, with

$r_1(z) = z^{-1}$ and thus short recurrence relations exist [188].
The rational framework allows to obtain the result (5.4). The Krylov basis for a unitary matrix is interpreted as a pair of rational (extended) Krylov subspaces, which allows to derive the displacement structure of $T_k$. Since $U^H = U^{-1}$, we have $\mathcal{K}_k(U, v) = \mathcal{K}_k(U^{-1}, v) = \mathcal{K}_k(U, v; \{0, 0, \dots\})$. Consider $B_k^\Xi$, the Krylov matrix for $\mathcal{K}_k(U, v; \{\infty, \infty, \dots\})$ and $B_k^\Theta$ for $\mathcal{K}_k(U, v; \{0, 0, \dots\})$ and their decompositions, with $U = X\Lambda X^H$ and $U^H = U^{-1} = X\Lambda^{-1}X^H$,

$$B_k^\Xi = \begin{bmatrix} v & Uv & \dots & U^{k-1}v \end{bmatrix} = XWV_k$$

$$B_k^\Theta = \begin{bmatrix} v & U^{-1}v & \dots & U^{-(k-1)}v \end{bmatrix} = XW\Phi\bar{V}_k J$$

with $W = \mathrm{diag}(\{X^H v e_i\}_{i=1}^m)$ and $\Phi = \mathrm{diag}(\{\phi_{k^-}(\lambda_i)\}_{i=1}^m)$ for $\phi_{k^-}(z) = z^{-(k-1)}$. The Gram matrix $M_k = (B_k^\Theta)^H B_k^\Xi$ is the Toeplitz matrix $T_k$ from before and

$$T_k = (B_k^\Theta)^H B_k^\Xi = JV_k^\top \underbrace{\Phi W^H W}_{=:D} V_k = JH_k$$

reveals the well known relation between a Toeplitz and Hankel matrix and allows to conclude that $T_k$ has the same displacement rank as $H_k$. Displacement operators for $T_k$ follow easily from this relation

$$J(Z_k^H H_k - H_k Z_k) = JZ_k^H JJH_k - JH_k Z_k = JZ_k^H JT_k - T_k Z_k.$$

### Construct permutation matrix

The construction of the permutation matrix $P$ in (5.3) from the poles $\Xi$ is straightforward. Algorithm 5 provides the procedure.

This procedure is illustrated by Example 5.1.

**Example 5.1.** *Consider $m \geq 8$, $v = \begin{bmatrix} \alpha_1 & \dots & \alpha_m \end{bmatrix}^\top \in \mathbb{C}^m$ and $\Lambda = diag(\lambda_1, \dots, \lambda_m) \in \mathbb{C}^{m\times m}$. Let $\hat{\Xi} = \{0, \infty, \infty, 0, \infty, 0, 0\}$. The Krylov matrix is*

$$B_k^\Xi = \begin{bmatrix} v & \Lambda^{-1}v & \Lambda v & \Lambda^2 v & \Lambda^{-2}v & \Lambda^3 v & \Lambda^{-3}v & \Lambda^{-4}v \end{bmatrix}.$$

*The matrix $\Phi = \Lambda^{-4}$ and $P \in \mathbb{C}^{8\times 8}$, as constructed by Algorithm 5 is*

$$P^\top = \begin{bmatrix} e_8 & e_7 & e_5 & e_2 & e_1 & e_3 & e_4 & e_6 \end{bmatrix}.$$

*And, indeed, this forms the decomposition from* (5.3)*:*

$$B_k^\Xi = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_m \end{bmatrix} \Lambda^{-4} \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^7 \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^7 \end{bmatrix} P^\top.$$

**Algorithm 5** Construct permutation matrix

**Input:** $\Xi = \{\xi_i\}_{i=1}^{k-1}$
**Output:** $P \in \mathbb{C}^{k\times k}$ appearing in (5.3) associated with given $\Xi$
**procedure** ConstructPermutationMatrix($\Xi$)
    $P := 1$
    **for** $i = 1, 2, \ldots, k-1$ **do**
        $P = \begin{bmatrix} P \\ \mathbf{0} \end{bmatrix}$ ▷ Embed $P$ in $\mathbb{C}^{i+1}$ by appending a zero row
        **if** $\xi_i == \infty$ **then**
            $P = \begin{bmatrix} P & e_{i+1} \end{bmatrix}$
        **else**
            $P = \begin{bmatrix} e_{i+1} & P \end{bmatrix}$
        **end if**
    **end for**
    $P = P^\top$
**end procedure**

## 5.3 Generating (bi)orthogonal vectors

Orthonormal and biorthonormal nested bases for rational Krylov subspaces can be generated using the associated Gram matrix. Lemma 2.2 states that computing the LR factorization of a Gram matrix corresponds to computing the recurrence coefficients of biorthonormal vectors. In finite precision two issues are encountered with this procedure:

1. The LR factorization without pivoting is numerically unstable. Pivoting is not allowed, since this would destroy the nestedness of the bases formed by the vectors.
2. The Gram matrix with classical moments is ill-conditioned.

The numerical instability of computing the LR factorization, the first issue, can be dealt with straightforwardly for Hermitian positive definite Gram matrices. For a Hermitian positive definite Gram matrix $M_k \in \mathbb{C}^{k\times k}$, the Cholesky decomposition $M_k = R_k^H R_k$ can be used. The Cholesky factor $R_k$ corresponds to the triangular factor in the QR decomposition of $B_k^\Xi = Q_k R_k$, if $M_k = (B_k^\Xi)^H B_k^\Xi$, with rational Krylov matrix $B_k^\Xi \in \mathbb{C}^{m\times k}$. The QR decomposition can be computed stably, but the ill-conditioning of the rational Krylov matrix remains an issue. However, rational Krylov subspace methods avoid the explicit construction of this matrix and generate the orthonormal vectors, that is, the matrix $Q_k$ and their representation as a recurrence pencil. The recurrence pencil representation is preferred over the representation in a

basis by the matrix $R$.
For Gram matrices that do not allow a Cholesky decomposition, methods based on Krylov subspaces, will improve the stability and avoid explicit construction of the ill-conditioned moment or Krylov matrices. Chapter 9 elaborates on methods for the construction of (bi)orthonormal vectors based on Krylov subspaces.
The conditioning of the Gram matrix, the second issue, also can be improved by using modified moments. An updating strategy is proposed in Section 5.3.1. This technique uses modified moments. The updating idea is important in Chapter 9. Section 5.3.2 discusses Levinson procedures, which uses low displacement rank of Gram matrices to obtain efficient algorithms for the (implicit) computation of their LR factorization. Levinson procedures are discussed further in Chapter 8, where they are used to obtain short recurrence relations for (bi)orthogonal vectors.

### 5.3.1 Updating Gram matrix factorization

For simplicity we discuss the updating procedure for Krylov subspaces generated by a diagonal matrix $A = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$. As we will see in Chapter 6 and Chapter 7, this is an important case. Consider starting vectors $v = \begin{bmatrix}\alpha_1 & \ldots & \alpha_m\end{bmatrix}^\top$, $w = \begin{bmatrix}\beta_1 & \ldots & \beta_m\end{bmatrix}^\top$ and the rational Krylov matrices

$$B_k^\Xi = \begin{bmatrix} v & \psi_1^\Xi(\Lambda)v & \ldots & \psi_{k-1}^\Xi(\Lambda)v \end{bmatrix} = \begin{bmatrix} \alpha_1 & \psi_1^\Xi(\lambda_1)\alpha_1 & \ldots & \psi_{k-1}^\Xi(\lambda_1)\alpha_1 \\ \vdots & \vdots & & \vdots \\ \alpha_m & \psi_1^\Xi(\lambda_m)\alpha_m & \ldots & \psi_{k-1}^\Xi(\lambda_m)\alpha_m \end{bmatrix},$$

such that $\mathrm{span}\{B_k^\Xi\} = \mathcal{K}(\Lambda, v; \Xi)$ and

$$\begin{aligned} B_k^\Theta &= \begin{bmatrix} w & \psi_1^\Theta(\Lambda^H)w & \ldots & \psi_{k-1}^\Theta(\Lambda^H)w \end{bmatrix} \\ &= \begin{bmatrix} \beta_1 & \psi_1^\Theta(\bar{\lambda}_1)\beta_1 & \ldots & \psi_{k-1}^\Theta(\bar{\lambda}_1)\beta_1 \\ \vdots & \vdots & & \vdots \\ \beta_m & \psi_1^\Theta(\bar{\lambda}_m)\beta_m & \ldots & \psi_{k-1}^\Theta(\bar{\lambda}_m)\beta_m \end{bmatrix}, \end{aligned}$$

with $\mathrm{span}\{B_k^\Theta\} = \mathcal{K}(\Lambda^H, w; \Theta)$. Suppose that for a Gram matrix $M_k = (B_k^\Xi)^H B_k^\Xi$ or $M_k = (B_k^\Theta)^H B_k^\Xi$ the LR factorization is available, i.e., $M_k = L_k R_k$. We are interested in the LR factorization of $\tilde{M}_k$, the Gram matrix for Krylov subspaces generated with $\tilde{\Lambda} = \begin{bmatrix}\Lambda & \\ & \lambda_{m+1}\end{bmatrix}$, $\tilde{v} = \begin{bmatrix} v \\ \alpha_{m+1}\end{bmatrix}$ and, possibly, $\tilde{w} = \begin{bmatrix} w \\ \beta_{m+1}\end{bmatrix}$. This Gram matrix satisfies $\tilde{M}_k = M_k + \hat{M}_k$, with $\mathrm{rank}(\hat{M}_k) = 1$. The updating procedure reuses the factorization of $M_k$ to compute the factorization of $\tilde{M}_k$. The factorization of $M_k = L_k R_k$ can be reused by noting that

$$\tilde{M}_k = M_k + \hat{M}_k \Leftrightarrow \tilde{L}_k \tilde{R}_k = L_k R_k + \hat{M}_k \Leftrightarrow L_k^{-1} \tilde{L}_k \tilde{R}_k R_k^{-1} = I + L_k^{-1} \hat{M}_k R_k^{-1}.$$

The only factorization that must be computed is $I + L_k^{-1}\hat{M}_k R_k^{-1} = \hat{L}_k\hat{R}_k$ and the solution is obtained by $\tilde{L}_k = L_k\hat{L}_k$ and $\tilde{R}_k = \hat{R}_k R_k$.

### Single space

Consider a Gram matrix corresponding to a single rational Krylov subspace $\mathcal{K}_k(\Lambda, v; \Xi)$:

$$M_k = (B_k^{\Xi})^H B_k^{\Xi}$$

$$= \begin{bmatrix} \langle v, v\rangle_E & \langle \psi_1^{\Xi}(\Lambda)v, v\rangle_E & \dots & \langle \psi_{k-1}^{\Xi}(\Lambda)v, v\rangle_E \\ \langle v, \psi_1^{\Xi}(\Lambda)v\rangle_E & \langle \psi_1^{\Xi}(\Lambda)v, \psi_1^{\Xi}(\Lambda)v\rangle_E & \dots & \langle \psi_{k-1}^{\Xi}(\Lambda)v, \psi_1^{\Xi}(\Lambda)v\rangle_E \\ \vdots & \vdots & & \vdots \\ \langle v, \psi_{k-1}^{\Xi}(\Lambda)v\rangle_E & \langle \psi_1^{\Xi}(\Lambda)v, \psi_{k-1}^{\Xi}(\Lambda)v\rangle_E & \dots & \langle \psi_{k-1}^{\Xi}(\Lambda)v, \psi_{k-1}^{\Xi}(\Lambda)v\rangle_E \end{bmatrix}.$$

The Gram matrix $\tilde{M}_k = (\tilde{B}_k^{\Xi})^H \tilde{B}_k^{\Xi}$ corresponding to $\mathcal{K}_k(\tilde{\Lambda}, \tilde{v}; \Xi)$ with basis $\tilde{B}_k^{\Xi} = \begin{bmatrix} v & \psi_1^{\Xi}(\tilde{\Lambda})\tilde{v} & \dots & \psi_{k-1}^{\Xi}(\tilde{\Lambda})\tilde{v} = \end{bmatrix}$ is equal to

$$\tilde{M}_k = M_k + |\alpha_{m+1}|^2 \begin{bmatrix} 1 \\ \bar{\psi}_1^{\Xi}(\bar{\lambda}_{m+1}) \\ \vdots \\ \bar{\psi}_{k-1}^{\Xi}(\bar{\lambda}_{m+1}) \end{bmatrix} \begin{bmatrix} 1 & \psi_1^{\Xi}(\lambda_{m+1}) & \dots & \psi_{k-1}^{\Xi}(\lambda_{m+1}) \end{bmatrix}.$$

The Cholesky factorization of

$$I + |\alpha_{m+1}|^2 R_k^{-H} \begin{bmatrix} 1 \\ \bar{\psi}_1^{\Xi}(\bar{\lambda}_{m+1}) \\ \vdots \\ \bar{\psi}_{k-1}^{\Xi}(\bar{\lambda}_{m+1}) \end{bmatrix} \begin{bmatrix} 1 & \psi_1^{\Xi}(\lambda_{m+1}) & \dots & \psi_{k-1}^{\Xi}(\lambda_{m+1}) \end{bmatrix} R_k^{-1}$$

must be computed. In this equation we see that this in fact uses modified moments, using the orthonormal basis for $\mathcal{K}(\Lambda, v; \Xi)$. We will not discuss the implementation and numerical properties further. Details can be found in the paper by Gill et al., [86]. Alternatively, the QR decomposition can be used. If the QR decomposition of $B_k^{\Xi}$ is available,

$$B_k^{\Xi} = \begin{bmatrix} \alpha_1 & \psi_1^{\Xi}(\lambda_1)\alpha_1 & \dots & \psi_{k-1}^{\Xi}(\lambda_1)\alpha_1 \\ \vdots & \vdots & & \vdots \\ \alpha_m & \psi_1^{\Xi}(\lambda_m)\alpha_m & \dots & \psi_{k-1}^{\Xi}(\lambda_m)\alpha_m \end{bmatrix} = Q_k R_k,$$

then using the embedded matrix $\hat{Q}_k = \begin{bmatrix} Q_k & \\ & 1 \end{bmatrix}$, we get

$$\hat{Q}_k^H \tilde{B}_k^{\Xi} = \begin{bmatrix} & R_k & \\ \alpha_{m+1} & \psi_1^{\Xi}(\lambda_{m+1})\alpha_{m+1} & \dots & \psi_{k-1}^{\Xi}(\lambda_{m+1})\alpha_{m+1} \end{bmatrix} = \begin{bmatrix} R_k \\ x \end{bmatrix} =: \hat{R}_k,$$

with $x = Q_k^H \begin{bmatrix} \alpha_{m+1} & \psi_1^{\Xi}(\lambda_{m+1})\alpha_{m+1} & \dots & \psi_{k-1}^{\Xi}(\lambda_{m+1})\alpha_{m+1} \end{bmatrix}$. The matrix $\hat{R}_k$ deviates from an upper triangular matrix only by its last row. Premultiplication with a sequence of plane rotations can restore the upper triangular structure efficiently and leads to a QR decomposition of $\tilde{M}_k$. Details can be found in the literature [25, 50, 86].

### Pair of spaces

A Gram matrix corresponding to a pair of rational Krylov subspaces $\mathcal{K}(\Lambda, v; \Xi)$ and $\mathcal{K}(\Lambda^H, w; \Theta)$ has the form

$$M_k = (B_k^{\Theta})^H B_k^{\Xi} = \begin{bmatrix} \langle v, w\rangle_E & \dots & \langle \psi_{k-1}^{\Xi}(\Lambda)v, w\rangle_E \\ \langle v, \psi_1^{\Theta}(\Lambda^H)w\rangle_E & \dots & \langle \psi_{k-1}^{\Xi}(\Lambda)v, \psi_1^{\Theta}(\Lambda^H)w\rangle_E \\ \vdots & & \vdots \\ \langle v, \psi_{k-1}^{\Theta}(\Lambda^H)w\rangle_E & \dots & \langle \psi_{k-1}^{\Xi}(\Lambda)v, \psi_{k-1}^{\Theta}(\Lambda^H)w\rangle_E \end{bmatrix}.$$

Suppose its LR factorization is available, $M_k = L_k R_k$. The updated Gram matrix is

$$\begin{aligned} \tilde{M}_k &= (\tilde{B}_k^{\Theta})^H \tilde{B}_k^{\Xi} \\ &= M_k + \bar{\beta}_{m+1}\alpha_{m+1} \begin{bmatrix} 1 \\ \bar{\psi}_1^{\Theta}(\lambda_{m+1}) \\ \vdots \\ \bar{\psi}_{k-1}^{\Theta}(\lambda_{m+1}) \end{bmatrix} \begin{bmatrix} 1 & \psi_1^{\Xi}(\lambda_{m+1}) & \dots & \psi_{k-1}^{\Xi}(\lambda_{m+1}) \end{bmatrix}, \end{aligned}$$

and only the LR factoriation of

$$I + \bar{\beta}_{m+1}\alpha_{m+1} L_k^{-1} \begin{bmatrix} 1 \\ \bar{\psi}_1^{\Theta}(\lambda_{m+1}) \\ \vdots \\ \bar{\psi}_{k-1}^{\Theta}(\lambda_{m+1}) \end{bmatrix} \begin{bmatrix} 1 & \psi_1^{\Xi}(\lambda_{m+1}) & \dots & \psi_{k-1}^{\Xi}(\lambda_{m+1}) \end{bmatrix} R_k^{-1}$$

must be computed to obtain the LR factorization of $\tilde{M}_k$. Details can be found in the paper by Bennet [15].

## 5.3.2 Levinson procedure

The LR factorization of Gram matrices with low displacement rank can be efficiently computed by Levinson procedures by studying its displacement structure. This results in short recurrence relations for (bi)orthogonal vectors related to the considered Gram matrix. The derivation of short recurrence relations is postponed to Chapter 8, since

having results from Chapter 6 and Chapter 7 facilitates the description of the Levinson procedures. Now, only the idea behind these procedures is sketched.
Set $M_k = \dot{B}_k^H B_k$, with $B_k, \dot{B}_k \in \mathbb{C}^{m\times k}$ which can represent a basis for a single Krylov subspace $B_k = \dot{B}_k = B_k^{\Xi}$ or for a pair of Krylov subspace $B_k = B_k^{\Xi}$, generated with some matrix $A$, and $\dot{B}_k = B_k^{\Theta}$, generated with $A^H$.
Suppose $M_k$ has low displacement rank. A Levinson procedure assumes (implicit) knowledge of the LR factorization of $M_k = L_k R_k$. From these factors we can obtain the (bi)orthonormal vectors $\{v_i\}_{i=1}^k$ and $\{w_i\}_{i=1}^k$. Note that

$$M_k v_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \times \end{bmatrix} \text{ and } w_k^H M_k = \begin{bmatrix} 0 & \cdots & 0 & \times \end{bmatrix}. \tag{5.5}$$

An efficient way to obtain the next vectors in the (bi)orthonormal sequence, $v_{k+1}, w_{k+1}$, is derived. The Gram matrix is

$$M_{k+1} = \dot{B}_{k+1}^H B_{k+1} = \begin{bmatrix} M_k & m \\ \hat{m}^H & \mu \end{bmatrix},$$

where $m, \hat{m} \in \mathbb{C}^k$ and $\mu \in \mathbb{C}$. Studying the structure of the multiplication of $M_{k+1}$ with $\dot{v}_k := \begin{bmatrix} 0 \\ v_k \end{bmatrix}$, an embedding of $v_k$ in $\mathbb{C}^{k+1}$, allows the derivation of recurrence relations. Another possible embedding is $\begin{bmatrix} v_k \\ 0 \end{bmatrix}$. If $M_{k+1}$ has displacement rank 2, we expect no more than 3 nonzero elements appearing in the vector $M_{k+1}\dot{v}_k$. With the embedding of the available vectors $\{v_i\}_{i=1}^k$, a linear combination $v_{k+1}$ is sought such that $M_{k+1}v_{k+1} = \eta e_{k+1}$, for some nonzero constant $\eta$. For details on the Levinson procedure for Toeplitz and Hankel matrices we refer the survey paper by Heinig and Rost [99] and references therein.

## 5.4 Conclusion

A meaningful generalization of moments is described which is related to rational Krylov subspaces. The displacement structure of Gram matrices for rational Krylov subspaces is studied and suitable displacement operators are proposed. This reveals the low displacement rank of Gram matrices; certain Gram matrices have displacement rank at most 2. We show that these are Gram matrices arising from normal matrices and a single rational Krylov subspace or from a diagonalizable matrix and a pair of rational Krylov subspaces. Displacement rank equal to 2 suggests the existence of three term recurrence relations for the construction of (bi)orthonormal nested bases for rational Krylov subspaces. Two procedures to generate (bi)orthonormal nested bases from a Gram matrix are discussed.

# Chapter 6

# Orthogonal polynomials

The interplay between Krylov subspaces (matrix theory) and orthogonal polynomials (classical analysis) is fruitful in both directions. Classical analysis is an older discipline and many theoretical results are known for orthogonal polynomials (OPs) [43, 110, 154, 157]. Potential theory for orthogonal polynomials can be used to study convergence results for Krylov subspace methods [100, 119, 120].
Matrix theory, numerical linear algebra, provides very effective procedures to solve linear algebraic problems and provides a natural framework to study numerical stability of algorithms and condition of problems. For numerical computation involving polynomials, many of the most effective methods rely on solving an equivalent (or approximate/discretized) problem in matrix theory [23, 70, 77, 78, 91, 134]. Figure 6.1 shows the idea of solving functional problems (problems for functions, i.e., polynomials in the current case) with linear algebraic methods schematically.

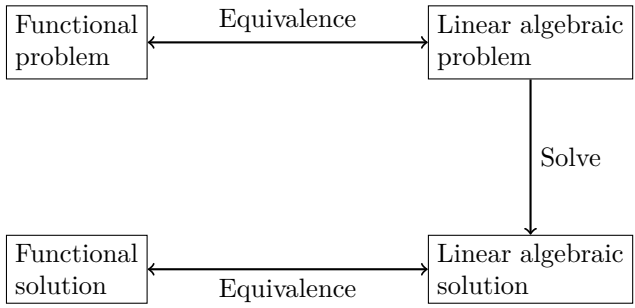



Figure 6.1: Solving functional problems with numerical linear algebra techniques.

Some examples are the Golub-Welsch algorithm for the computation of weights and nodes for Gaussian quadrature [91], the construction of orthogonal polynomials [78] and computation of roots of polynomials [5, 23].
First, general polynomials and their relation to structured matrices are discussed in Section 6.1. This allows us to represent polynomials as structured matrices, i.e., represent a functional solution as an equivalent algebraic solution. Orthogonal polynomials enjoy many interesting, theoretical and numerical, properties. Polynomials orthogonal to discrete inner products are related to certain Krylov subspaces, Section 6.2 introduces OPs and identifies the relation to Krylov subspace. Polynomials orthogonal with respect a more general linear functional are formal orthogonal polynomials [56], we interpret these as biorthogonal polynomials (biOPs). These always satisfy a short recurrence relation but sacrifice some of the attractive properties of OPs, e.g., the associated linear functional can be indefinite. Section 6.3 describes biorthogonal polynomials and the relation to biorthonormal bases for Krylov subspaces. These results allow the formulation of problems for (bi)orthogonal polynomials as problems for structured matrices, most notably, inverse eigenvalue problems, which are the subject of Chapter 9.

## 6.1 Polynomials and structured matrices

The connection between polynomials and vectors in (polynomial) Krylov subspaces arises from the manner in which a Krylov subspace is constructed, multiplying consecutive powers of $A \in \mathbb{C}^{m\times m}$ with a starting vector $v \in \mathbb{C}^m$

$$\mathcal{K}_k(A, v) = \text{span}\{v, Av, A^2v, \ldots, A^{k-1}v\}.$$

A vector $x \in \mathcal{K}_k(A, v)$ can be written in terms of a polynomial $p_{k-1}$ of degree at most $k-1$,

$$x = p_{k-1}(A)v, \quad p_{k-1} \in \mathcal{P}_{k-1},$$

where $\mathcal{P}_{k-1}$ denotes the space of polynomials up to degree $k-1$. When no breakdown occurs, i.e., $\dim(\mathcal{K}_l(A, v)) = l$, $l = 1, 2, \ldots k$, the recurrence relations for nested Krylov, orthogonal and biorthogonal bases of $\mathcal{K}_k(A, v)$, discussed in Chapter 3, lead to proper Hessenberg matrices. The sequence of polynomials $\{p_i(z)\}_{i=0}^{k-1}$, $p_i \in \mathcal{P}_i$ and satisfying $\mathcal{K}_l(A, v) = \text{span}\{p_0(A)v, \ldots, p_{l-1}(A)v\}$, forms a triangle family of polynomials. Section 6.1.1 elaborates on the relation between triangle families and proper Hessenberg matrices, they are in fact equivalent. Evaluation of polynomials can be reformulated as a problem involving a Hessenberg matrix, Section 6.1.2 proposes two matrix theoretical formulations of polynomial evaluation. These Hessenberg matrices related to triangle families are in fact congenial matrices. Section 6.1.3 provides a short remark on such matrices and their relation to Krylov subspaces.

### 6.1.1 Triangle family

Definition 6.1 introduces triangle families of polynomials whose relation to proper Hessenberg matrices is elaborated on below.

**Definition 6.1** (Triangle family of polynomials [49])**.** *A triangle family of polynomials* $\{p_0, p_1, \ldots, p_k\}$ *is a sequence*

$$\begin{aligned} p_0(z) &= \gamma_{1,0} \\ p_1(z) &= \gamma_{1,1} + \gamma_{2,1} z \\ p_2(z) &= \gamma_{1,2} + \gamma_{2,2} z + \gamma_{3,2} z^2 \\ &\cdots \\ p_k(z) &= \gamma_{1,k} + \gamma_{2,k} z + \gamma_{3,k} z^2 + \cdots + \gamma_{k+1,k} z^k \end{aligned}$$

*where* $\gamma_{i+1,i} \neq 0$ *for all* $i$*.*

Triangle families of polynomials have two important properties stated in Property 6.1.

**Property 6.1** (Properties of triangle family of polynomials [49])**.** *A triangle family of polynomials* $\{p_0, \ldots, p_k\}$ *satisfies the following two properties*

- *Every triangle family of* $k+1$ *members is a basis for* $\mathcal{P}_k$
- *An* $l+1$*-term recurrence relation* $z p_{l-1}(z) = \sum_{i=0}^{l} h_{i+1,l} p_i(z)$*, with* $h_{l+1,l} \neq 0$*, holds for* $l = 0, 1, \ldots, k-1$*.*

The second property implies that a proper Hessenberg matrix $H_k = \left[h_{i,j}\right]_{i,j=1}^{k} \in \mathbb{C}^{k \times k}$ is the recurrence matrix for the relation

$$z P_k = P_k H_k + h_{k+1,k} p_k(z) e_k, \tag{6.1}$$

with $P_k := \begin{bmatrix} p_0(z) & p_1(z) & \ldots & p_{k-1}(z) \end{bmatrix}$, where $\{p_i\}$ forms a triangular family of polynomials. This recurrence relation links the roots of polynomials with eigenvalues of the recurrence matrix. Lemma 6.1 states this link formally.

**Lemma 6.1** (Correspondence roots of polynomial and eigenvalues of recurrence matrix [151])**.** *Let* $\{p_i\}_{i=0}^{k-1}$ *be a triangular family and* $H_k \in \mathbb{C}^{k \times k}$ *a proper Hessenberg matrix such that* (6.1) *holds. Then the roots of* $p_k(z)$ *and the eigenvalues of* $H_k$ *coincide.*

Property 6.1 and Lemma 6.1 suggest that recurrence matrices for a nested basis of a polynomial Krylov subspace can be used to study polynomials. In fact, knowing $H_k$ and $p_0$ corresponds to knowing $\{p_0, \ldots, p_{k-1}\}$, it is the matrix representation of the polynomial sequence. In the sequel the polynomials must be evaluated, e.g., to perform error analysis, and the preferred way is to use the matrix $H_k$. Section 6.1.2 describes how a polynomial can be evaluated when the recurrence matrix (the linear algebraic solution) is available.

## 6.1.2 Evaluate polynomials

A triangular family of polynomials $\{p_0, \ldots, p_k\}$ with corresponding recurrence matrix $H_k$ satisfying (6.1) can be evaluated at a point $z = z^\star$ by solving a linear system of equations with $H_k$ or by computing eigenvalues of $H_k$. These two procedures are described here and for both the condition of the problem of polynomial evaluation is provided.

### Solve linear system

Polynomial evaluation can be formulated as an upper triangular linear system of equations, which can be solved by forward substitution. Proposition 6.1 provides the details.

**Proposition 6.1** (Polynomial evaluation - system solve)**.** *Consider a triangle family of polynomials $\{p_i\}_{i=0}^k$ and corresponding recurrence matrix $H_k \in \mathbb{C}^{k\times k}$ satisfying* (6.1)*. Then evaluating $p_i(z)$, $i = 0, 1, \ldots, k-1$ at $z = z^\star \in \mathbb{C}$ is equivalent to solving the system of equations*

$$x^\top \begin{bmatrix} 1 & \\ 0 & \\ \vdots & (H_k - z^\star I)I^{k\times(k-1)} \\ 0 & \\ 0 & \end{bmatrix} = \begin{bmatrix} p_0 & 0 & \ldots & 0 \end{bmatrix}$$

*for $x \in \mathbb{C}^k$ and where $I^{k\times(k-1)}$ is the $k \times (k-1)$ leading principal submatrix of the identity matrix. Equivalent in the sense that $x_l := x^\top e_l = p_{l-1}(z^\star)$.*

*Proof.* The equivalence follows directly from the fact that $p_0(z)$ is a known constant and from recurrence relation (6.1). Clearly the statement holds for $x_1$,

$$x^\top \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} p_0 & 0 & \cdots & 0 \end{bmatrix} \Leftrightarrow x_1 = p_0.$$

For polynomials $p_l$, $1 < l < k$, the statement follows from the recurrence relation and using commutativity of polynomials under multiplication,

$$\begin{aligned} & z \begin{bmatrix} p_0 & p_1 & \cdots p_{k-1} \end{bmatrix} = \begin{bmatrix} p_0 & p_1 & \cdots p_{k-1} \end{bmatrix} H_k + h_{k+1,k} p_k e_k^\top \\ \Leftrightarrow & \begin{bmatrix} p_0 & p_1 & \cdots p_{k-1} \end{bmatrix} (zI - H_k) = h_{k+1,k} p_k e_k^\top. \end{aligned}$$

Using the first $k-1$ elements (dropping the $k$th element and thus the contribution of $h_{k+1,k}p_k$) in the above equation allows to show the equivalence $p_{l-1}(z^\star) = x_l$, $1 < l < k$. □

Proposition 6.1 is the matrix notation of evaluating the recurrence relation (6.1). The matrix notation allows to straightforwardly compute the condition of the problem of polynomial evaluation. The condition is given by $\kappa(G)$, with

$$G := \begin{bmatrix} 1 & \\ 0 & \\ \vdots & (H_n - z^\star I) I^{n \times (n-1)} \\ 0 & \\ 0 & \end{bmatrix}.$$

The matrix $G$ depends on both $H_n$ and $z^\star$. A structured condition number would be more appropriate. This is subject to future research.

### Eigenvalue decomposition

The eigenvalues of the recurrence matrix $H_l \in \mathbb{C}^{l \times l}$, $1 < l \leq k$, appearing in (6.1) correspond to the roots of the polynomial $p_l$. Thanks to the availability of effective eigenvalue solvers, the eigenvalues of $H_l$, and thus the roots of $p_l(z)$, can be computed efficiently and numerically stable. This allows us to write $p_l$ into the factored form $p_l(z) = \eta_l \prod_{i=1}^{l} (z - z_i)$ and evaluate $p_l(z)$ using this form. Proposition 6.2 formalizes this idea.

**Proposition 6.2** (Polynomial evaluation - eigenvalue decomposition)**.** *Consider a triangle family of polynomials* $\{p_i\}_{i=0}^{k}$ *and corresponding recurrence matrix* $H_k \in \mathbb{C}^{k\times k}$*, which satisfies* (6.1)*. Then, for* $l \leq k$*,*

$$p_l(z) = \eta_l \prod_{i=1}^{l} (z - z_i^{(l)}),$$

*where* $\{z_i^{(l)}\}_{i=1}^{l}$ *are the eigenvalues of the principal submatrix* $H_k^{(l)} = H_l$ *of size* $l \times l$ *of* $H_k$ *and* $\eta_l$ *is determined by normalization.*

*Proof.* The equivalence between the eigenvalues of $H_k^{(l)}$, the $l \times l$ principal submatrix of $H_k$, and the roots of $p_l$ follows from Lemma 6.1. Hence, $p_l(z) = \eta_l \hat{p}_l(z)$, with $\hat{p}_l(z) := \prod_{i=1}^{l}(z - z_i^{(l)})$. For monic polynomials, $\eta_l = 1$ and for orthonormal polynomials we have

$$\begin{aligned} \langle p_l, p_l \rangle &= |\eta_l|^2 \langle \hat{p}_l, \hat{p}_l \rangle \\ 1 &= |\eta_l|^2 \langle \hat{p}_l, \hat{p}_l \rangle \\ |\eta_l|^2 &= (\langle \hat{p}_l, \hat{p}_l \rangle)^{-1}. \end{aligned}$$

So the expression for $p_l(z)$ is completely determined. □

For the condition of this evaluation procedure we note that multiplication of the factors is a well-conditioned problem. Hence, the condition number of the eigenvalue computation is assumed to dominate the condition of the problem. This condition number is stated in Proposition 6.3. Only simple eigenvalues are considered, since recurrence matrices originating from orthogonal and biorthogonal Krylov subspace methods cannot have multiple eigenvalues.

**Proposition 6.3** (Condition of a simple eigenvalue [90, p. 344])**.** *Let* $\lambda$ *be a simple eigenvalue of* $A \in \mathbb{C}^{m\times m}$ *and let* $x, y$ *be vectors such that* $Ax = \lambda x$ *and* $y^H A = \lambda y^H$ *with* $\|x\|_2 = \|y\|_2 = 1$*. The condition of the eigenvalue* $\lambda$ *is given by*

$$s(\lambda) = |y^H x|^{-1}. \tag{6.2}$$

In practice, recurrence matrices originating from biorthogonal Krylov subspace methods, including the Hermitian Lanczos iteration, might have multiple eigenvalues. These might or might not correspond to actual multiple eigenvalues, e.g., a distinct eigenvalue might be found multiple times due to the loss of orthogonality and round-off errors, especially in the absence of reorthogonalization [46, 47].

### 6.1.3 Congenial matrix

Congenial matrices [7] represent a polynomial as a matrix in a certain basis. The polynomial can be studied by studying its congenial matrix. The class of congenial matrices include the companion matrix and, in fact, all recurrence matrices for a nested basis of polynomial Krylov subspaces. The difference among these congenial matrices is the basis with respect to which the polynomial is represented. The companion matrix uses the monomial basis and corresponds to the recurrence matrix for the Krylov recurrence relation. In the following sections the polynomial represented in an orthogonal basis and in biorthogonal bases is elaborated on. These choices allow for an efficient computation of the elements of the recurrence matrix.

## 6.2 Orthogonal polynomials

Just as for Krylov subspaces, it is interesting to consider an orthogonal basis for the space of polynomials. A sequence of polynomials $\{p_l\}_l$ is said to be orthogonal with respect to some inner product $\langle ., . \rangle : \mathcal{P} \times \mathcal{P} \to \mathbb{C}$ if $p_l \in \mathcal{P}_l$, $\deg(p_l) = l$ and

$$\langle p_i, p_j \rangle \begin{cases} = 0, & \text{if } i \neq j \\ > 0, & \text{if } i = j \end{cases} .$$

An inner product for complex valued functions $f$, $g$ can be defined as

$$\langle f, g \rangle := \int_\gamma f(z)\overline{g(z)}\alpha(z)|dz|,$$

where $\alpha(z)$ is a positive weight function and $\gamma$ an arc [157]. Since $z = ai + b$ is a complex variable, by $|dz|$ is meant the arc length, i.e., $|dz| = \sqrt{da^2 + db^2}$ [150]. Our interest is in discrete inner products, introduced in Section 6.2.1. Section 6.2.2 about the recurrence matrix for OPs, shows that the recurrence coefficients can be computed by inner products. In some cases the orthogonal basis vectors for Krylov subspaces give rise to OPs with respect to a discrete inner product, this is the topic of Section 6.2.3. To conclude, Section 6.2.4 provides a short note on the Gram matrix associated with OPs.

### 6.2.1 Discrete inner product

The discrete inner products considered in this thesis are:

$$\langle f, g \rangle_m := \sum_{i=1}^{m} \alpha_i f(z_i)\overline{g(z_i)}, \quad z_i \in \mathbb{C},$$

with $m$ distinct nodes $\{z_i\}_{i=1}^m$ and positive weights $\{\alpha_i\}_{i=1}^m$. If we are interested in polynomials orthogonal with respect to a continuous inner product, then it must first be discretized. One way to discretize an inner product is via its first $2m-1$ moments

$$g_i = \int_\gamma z^i \alpha(z)|dz|, \quad i = 0, 1, \ldots, 2m-2,$$

which are finite. These moments can be used to define $\langle ., . \rangle_m$. For numerical computation it is advised to use modified moments [78]

$$\tilde{g}_i = \int_\gamma q_k(z)\alpha(z)|dz|, \quad i = 0, 1, \ldots, 2m-2,$$

where $q_k$ is some polynomial of exact degree $k$. A usual choice is to use a sequence of polynomials orthogonal with respect to another inner product over the same arc $\gamma$.
Another possible discretization is applying a suitable quadrature rule to the integral $\int_\gamma f(z)\overline{g(z)}\alpha(z)|dz| = \sum_{i=1}^m f(z_i)\overline{g(z_i)}\alpha(z_i) + R_m$, with $R_m$ the error term [69, 77, 83, 150].
Good choices to discretize a continuous inner product depend on the specific problem. We assume that the discretization has been done and a discrete inner product $\langle ., . \rangle_m$ is available.
Note that $\langle ., . \rangle_m$ with $m$ distinct nodes $\{z_i\}$ is only an inner product on the space $\mathcal{P}_{m-1}$. The positive-definiteness of $\langle p, p \rangle_m$, $p \in \mathcal{P}_m$ does not hold. Let $p = \prod_{i=1}^m (z - z_i)$, then $p(z_i) = 0$ for all $i$ and therefore $\langle p, p \rangle_m = 0$ for a nonzero polynomial. Hence, $\langle ., . \rangle_m$ is an indefinite Hermitian form on $\mathcal{P}$.

### 6.2.2 Recurrence matrix

A sequence of orthogonal polynomials $\{p_i\}$ is obviously a triangle family of polynomials and satisfies recurrence relation (6.1). The advantage of the recurrence relation for orthogonal polynomials is that the recurrence coefficients can be computed by inner products, this is stated in Lemma 6.2.

**Lemma 6.2** (Recurrence matrix for orthogonal polynomials)**.** *Let $\langle ., . \rangle : \mathcal{P} \times \mathcal{P} \to \mathbb{C}$ be an inner product and let $\{p_i\}_{i=0}^k$ be the set of orthogonal polynomials with respect to $\langle ., . \rangle$. Then the Hessenberg matrix $H_k = [h_{i,j}]_{i,j=1}^k$ in (6.1) is composed of the elements*

$$h_{i,j} = \frac{\langle z p_{j-1}, p_{i-1} \rangle}{\langle p_{i-1}, p_{i-1} \rangle}.$$

*Proof.* The statement follows immediately from the recurrence relation (6.1) and properties of orthogonal polynomials, consider

$$\begin{bmatrix} zp_0 & zp_1 & \cdots & zp_{k-1} \end{bmatrix} = \begin{bmatrix} p_0 & p_1 & \cdots & p_{k-1} \end{bmatrix} H_k + h_{k+1,k} p_k e_k^\top,$$

and apply the inner product with $p_i(z)$, $i < k$, and use the orthogonality property, then, for $e_{i+1}$ the $(i+1)$st canonical unit vector,

$$\begin{bmatrix} \frac{\langle zp_0, p_i\rangle}{\langle p_i, p_i\rangle} & \frac{\langle zp_1, p_i\rangle}{\langle p_i, p_i\rangle} & \dots & \frac{\langle zp_{k-1}, p_i\rangle}{\langle p_i, p_i\rangle} \end{bmatrix} = e_{i+1}^\top H_k = \begin{bmatrix} h_{i,0} & \dots & h_{i,k-1} \end{bmatrix}.$$

□

Orthonormal polynomials are orthogonal polynomials satisfying $\langle p_i, p_i\rangle = 1$.
The Stieltjes procedure generates a sequence of orthogonal polynomials $\{p_i\}_i$ for a given inner product $\langle ., .\rangle$ and given $p_0$. For a discrete inner product $\langle f, g\rangle_m = \sum_{i=1}^m \alpha_i f(z_i)\overline{g(z_i)}$ the Stieltjes procedure is given in Algorithm 6. It computes alternatingly recurrence coefficients for the next polynomial in the sequence of orthogonal polynomials (Step 8) and evaluates this polynomial at the nodes of the inner product (Step 9). For some discrete inner products $\langle ., .\rangle_m$, this procedure is quite stable [69, 77, 79].

**Algorithm 6** Stieltjes procedure [77]

1: **Input:** Nodes $\{z_i\}$, weights $\{\alpha_i\}$ determining $\langle f, g\rangle_m = \sum_{i=1}^m \alpha_i f(z_i)\overline{g(z_i)}$, and $p_0 = \sum_{i=1}^m \alpha_i$, and integer $k < m$
2: Set $Z = \mathrm{diag}(\{z_i\})$ and $W = \mathrm{diag}(\{\alpha_i\})$
3: **Output:** Recurrence matrix $H_k \in \mathbb{C}^{k\times k}$ and $Q_k \in \mathbb{C}^{m\times k}$, with $Q_k e_i = \begin{bmatrix} p_{i-1}(z_1) & \dots p_{i-1}(z_m) \end{bmatrix}^\top$, with $Q_k^H W Q_k = D$, diagonal.
4: **procedure** Stieltjes procedure($p_0, Z, W$)
5: $\rho_0 = \begin{bmatrix} p_0 & \dots & p_0 \end{bmatrix}$ and $\tilde{\rho}_0 = Z\rho_0$
6: **for** $i = 1, 2, \dots, k-1$ **do**
7: **for** $j = 1, 2, \dots, i$ **do**
8: $h_{j,i} = \langle zp_{i-1}(z), p_{j-1}\rangle_m = \rho_{j-1}^H W \tilde{\rho}_i$
9: $\tilde{\rho}_i = \tilde{\rho}_i - h_{j,i}\rho_{j-1}$
10: **end for**
11: $\rho_i = h_{i+1,i}\tilde{\rho}_i$ ▷ $h_{i+1,i}$ is used for normalization
12: $\tilde{\rho}_{i+1} = Z\rho_i$
13: **end for**
14: **end procedure**

Section 6.2.3 shows that the Stieltjes procedure corresponds to the Arnoldi iteration in Algorithm 1 for a matrix $A = Z = \mathrm{diag}(\{z_i\})$ and starting vector $v = \begin{bmatrix} \sqrt{\alpha_1} & \dots & \sqrt{\alpha_m} \end{bmatrix}^\top$. The Stieltjes procedure can be sensitive to round-off errors. The advised numerical procedure is by solving an inverse eigenvalue problem. This is the subject of Chapter 9.

### 6.2.3 Inner product induced by Krylov basis

The similarity between the Arnoldi iteration, computing an orthonormal basis for Krylov subspaces, and the Stieltjes procedure, computing orthogonal polynomials for a given inner product, gives rise to the following question. When does the recurrence matrix for a nested orthogonal basis for $\mathcal{K}_k(A, v)$ contain recurrence coefficients for a sequence of orthogonal polynomials $\{p_i\}_{i=0}^{k-1}$?
Lemma 6.3 relates vectors in Krylov subspaces with polynomials and states how the inner products correspond.

**Lemma 6.3** (Krylov induced inner product)**.** *Consider a normal matrix $A \in \mathbb{C}^{m\times m}$ and $v \in \mathbb{C}^m$ of grade $g$. Let $\{b_i\}_{i=0}^{k-1}$ form a nested basis for $\mathcal{K}_k(A, v)$, and let the polynomials $p_i \in \mathcal{P}_i$ be such that $b_i = p_i(A)v$. Then the Euclidean inner product $\langle ., . \rangle_E$ on $\mathcal{K}_k(A, v)$ induces an inner product of the form*

$$\langle p_i, p_j \rangle_g = \sum_{k=1}^{g} \alpha_k p_i(z_k)\overline{p_j(z_k)}, \quad \alpha_k > 0,$$

*on the space of polynomials $\mathcal{P}_g$.*

*Proof.* A diagonalizable $A$ has an eigenvalue decomposition $A = X\Lambda X^{-1}$, with $\Lambda = \operatorname{diag}(\{\lambda_i\})$ and for a normal matrix $X^{-1} = X^H$. Substitute the eigenvalue decomposition of $A$ and use $X^H X = I$ to obtain

$$\begin{aligned}
\langle b_i, b_j \rangle_E &= \langle p_i(A)v, p_j(A)v \rangle_E \\
&= v^H \bar{p}_j(A^H) p_i(A) v \\
&= v^H X^{-H} \bar{p}_j(\Lambda^H) X^H X p_i(\Lambda) X^{-1} v \\
&= v^H X^{-H} \bar{p}_j(\Lambda^H) p_i(\Lambda) \underbrace{X^{-1} v}_{=:c} \\
&= c^H \begin{bmatrix} \bar{p}_j(\bar{\lambda}_1) & & \\ & \ddots & \\ & & \bar{p}_j(\bar{\lambda}_m) \end{bmatrix} \begin{bmatrix} p_i(\lambda_1) & & \\ & \ddots & \\ & & p_i(\lambda_m) \end{bmatrix} c \\
&= \sum_{k=1}^{m} |e_k^\top c| \overline{p_j(\lambda_k)} p_i(\lambda_k).
\end{aligned}$$

Denote by $\{\tilde{z}_i\}_{i=1}^{d}$ the set of $d \leq m$ distinct eigenvalues of $A$. The weight $\tilde{\alpha}_i = \sum_{i=l,\lambda_l=\tilde{z}_i}^{m} |e_i^\top c|^2 \geq 0$ corresponds to $\tilde{z}_i$. Let $\{\alpha_i\}_{i=1}^{g}$ be the weights in $\{\tilde{\alpha}_i\}_{i=1}^{d}$ which

are nonzero and let $z_i$ be the node corresponding to $\alpha_i$. Then

$$\langle p_i, p_j \rangle_g := \sum_{k=1}^{g} \alpha_k \overline{g(z_k)} f(z_k)$$

is a discrete inner product on $\mathcal{P}_g \times \mathcal{P}_g$. □

In case the matrix $A$ is diagonalizable, but nonnormal, the term $X^H X$ will not cancel. It might be possible to interpret this term as a (nondiagonal) weight matrix, but this is not explored further.
Theorem 6.1 reveals the connection between orthonormal vectors for Krylov subspaces and OPs.

**Theorem 6.1** (Krylov induced orthogonal polynomials)**.** *Consider a normal matrix $A \in \mathbb{C}$ and $v \in \mathbb{C}^m$ with grade $g$. Let $\begin{bmatrix} q_0 & q_1 & \dots & q_{k-1} \end{bmatrix} = Q_k \in \mathbb{C}^{m \times k}$ form a nested orthonormal basis for $\mathcal{K}_k(A, v)$,$k < g$. Let polynomials $p_i \in \mathcal{P}_i$ satisfy $q_i = p_i(A)v$. Then the sequence of polynomials $\{p_i\}_{i=0}^{k-1}$ consists of orthonormal polynomials with respect to an inner product of the form*

$$\langle p_i, p_j \rangle_g = \sum_{k=1}^{g} \alpha_k p_i(z_k) \overline{p_j(z_k)}.$$

*Proof.* The basis vectors satisfy $q_i = p_i(A)v$, with $p_i \in \mathcal{P}_i$ of exact degree $i$. By Lemma 6.3, the sequence of polynomials $\{p_i\}_{i=1}^{g}$ is orthonormal with respect to the inner product $\langle p_i, p_j \rangle_g$, since

$$\delta_{ij} = \langle q_i, q_j \rangle_E = \langle p_i(A)v, p_j(A)v \rangle_E = \sum_{k=1}^{g} \alpha_k p_i(z_k) \overline{p_j(z_k)} = \langle p_i, p_j \rangle_g = \delta_{ij}.$$

□

Theorem 6.1 allows a functional problem to be translated into a linear algebra problem, shown in Figure 6.2. The linear algebra problem is in fact an inverse eigenvalue problem. Choose a diagonal matrix with (distinct) nodes $Z = \operatorname{diag}(\{z_i\}_{i=1}^{g})$ and starting vector with weights $v = \begin{bmatrix} \sqrt{\alpha_1} & \dots & \sqrt{\alpha_g} \end{bmatrix}$. Then an orthonormal basis for $\mathcal{K}(Z, v)$ corresponds to OPs with respect to $\langle p_i, p_j \rangle_g = \sum_{k=1}^{g} \alpha_k p_i(z_k) \overline{p_j(z_k)}$.

### 6.2.4 Gram matrix

The Gram matrix of a Krylov basis for a normal matrix and the Gram matrix for monomials with respect to a discrete inner product correspond. Corollary 6.1 follows immediately from Lemma 6.3, since $m_{i,j} := \langle A^j v, A^i v \rangle_E = \langle z^j, z^i \rangle_g$.

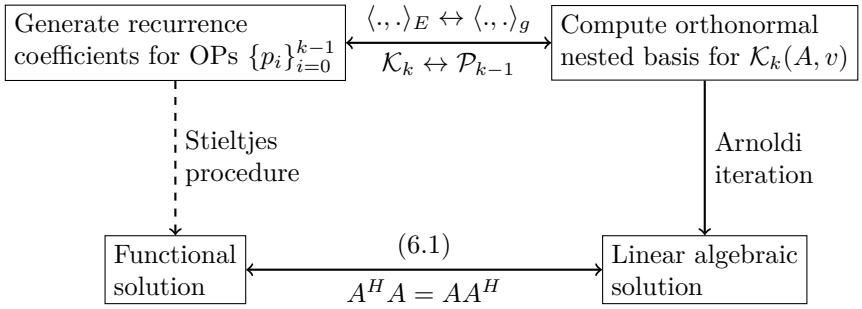



Figure 6.2: Scheme for generating recurrence coefficients for OPs.

**Corollary 6.1.** *Let* $B_k = \begin{bmatrix} v & Av & \dots & A^{k-1}v \end{bmatrix}$ *form a nested basis for* $\mathcal{K}_k(A, v)$*, with* $A^H A = A A^H$ *and* $k < g$*, with* $g$ *the grade of* $v$ *with respect to* $A$*. The corresponding Gram matrix*

$$M_k = B_k^H B_k = \begin{bmatrix} \langle v, v \rangle_E & \langle Av, v \rangle_E & \dots & \langle A^{k-1}v, v \rangle_E \\ \langle v, Av \rangle_E & \langle Av, Av \rangle_E & \dots & \langle A^{k-1}v, Av \rangle_E \\ \vdots & \vdots & & \vdots \\ \langle v, A^{k-1}v \rangle_E & \langle Av, A^{k-1}v \rangle_E & \dots & \langle A^{k-1}v, A^{k-1}v \rangle_E \end{bmatrix}$$

*equals the Gram matrix for monomials generated by* $\langle .,. \rangle_g$*, as defined in Lemma 6.3,*

$$\begin{bmatrix} \langle 1, 1 \rangle_g & \langle z, 1 \rangle_g & \dots & \langle z^{k-1}, 1 \rangle_g \\ \langle 1, z \rangle_g & \langle z, z \rangle_g & \dots & \langle z^{k-1}, z \rangle_g \\ \vdots & \vdots & & \vdots \\ \langle 1, z^{k-1} \rangle_g & \langle z, z^{k-1} \rangle_g & \dots & \langle z^{k-1}, z^{k-1} \rangle_g \end{bmatrix} = M_k.$$

This corollary implies that theoretical results and numerical procedures for Gram matrices of Krylov subspaces are also valid or applicable to the Gram matrices of polynomials and vice versa. Results for Gram matrices can be found in Chapter 5.
A common representation for orthogonal polynomials is the determinantal formula in Property 6.2.

**Property 6.2** (Determinantal formula)**.** *The determinant, with moments* $m_{i,j} = \langle z^j, z^i \rangle_g$*,*

$$p_k(z) = \det \left( \begin{bmatrix} m_{0,0} & m_{0,1} & \dots & m_{0,k} \\ m_{1,0} & m_{1,1} & \dots & m_{1,k} \\ \vdots & \vdots & & \vdots \\ m_{k-1,0} & m_{k-1,1} & \dots & m_{k-1,k} \\ 1 & z & \dots & z^k \end{bmatrix} \right)$$

*is an orthogonal polynomial with respect to* $\langle .,. \rangle_g$*, i.e.,* $\langle z^j, p_k(z) \rangle_g = 0$ *for* $j < k$*.*

## 6.3 Biorthogonal polynomials

Polynomials orthogonal with respect to a linear functional are called *formal orthogonal polynomials* [32, 56]. We will use the term biorthogonal polynomials, which stresses that there are in fact two sequences of polynomials. This term is better suited for the discussion on biorthogonal rational functions in Chapter 7 .
The dual of the vector space of polynomials $\mathcal{P}$ is the space of linear functionals, with a linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$. For the vector space of column vectors $\mathbb{C}^m$, the dual space is formed by row vectors $\mathbb{C}^{1\times m}$.
The linear functional $\mathcal{L}$ on $\mathcal{P}$ is defined, here, by fixing its value for the monomials

$$\mathcal{L}\{z^i\} := m_i, \quad i = 0, 1, \ldots,$$

where $m_i$ are given and will be called the moments (associated with $\mathcal{L}$). A biorthogonal polynomial $p_{k-1}$ is a polynomial of exact degree $k-1$ such that

$$\mathcal{L}\{z^i p_{k-1}\} = 0, \quad i = 0, 1, \ldots, k-2.$$

The Gram matrix induced by $\{m_i\}$ is of Hankel form and leads to a three term recurrence relation for biorthogonal polynomials, called Lanczos polynomials, which are the subject of Section 6.3.1. Section 6.3.2 discusses the recurrence matrix and Section 6.3.3 the relation to the Lanczos iteration. A justification for the generalization from inner product to linear functionals is given in Section 6.3.4.

### 6.3.1 Lanczos polynomials

Lanczos polynomials are the polynomials representing the biorthonormal basis vectors generated by the Lanczos iteration. The Lanczos iteration generates a pair of biorthonormal sequences $\{v_i\}_{i=0}^{k-1}$, $\{w_i\}_{i=0}^{k-1}$ which satisfy, for $l = 0, 1, \ldots, k-1$,

$$\operatorname{span}\{v_0, v_1, \ldots, v_l\} = \mathcal{K}_l(A, v) := \operatorname{span}\{v, Av, \ldots, A^{l-1}v\},$$

$$\operatorname{span}\{w_0, w_1, \ldots, w_l\} = \mathcal{K}_l(A^H, w) := \operatorname{span}\{w, A^H w, \ldots, (A^H)^{l-1}w\},$$

and the biorthonormality condition

$$\langle v_i, w_j\rangle_E = \delta_{ij}.$$

The basisvector $v_l$, $0 \le l < k$, is an element of $\mathcal{K}_l(A, v)$, i.e., $v_l = p_{l-1}(A)v$ for $\alpha_0 + \alpha_1 z + \cdots + \alpha_{l-1}z^{l-1} =: p_{l-1} \in \mathcal{P}_{l-1}$ of exact degree. The orthogonality conditions $v_l \perp \mathcal{K}_{l-1}(A^H, w)$ require the first $2l-2$ moments $m_i$, $i = 0, 1, \ldots, 2l-3$. The orthogonality conditions $\langle v_l, (A^H)^i w\rangle_E = 0$ for $i = 0, 1, \ldots, l-2$, can be written as a

system of $l-1$ equations with $l$ unknowns

$$\begin{cases} w^H A^0 p_{l-1}(A)v & = 0 \\ w^H A^1 p_{l-1}(A)v & = 0 \\ & \vdots \\ w^H A^{2l-3} p_{l-1}(A)v & = 0 \end{cases} \Leftrightarrow \begin{bmatrix} w^H v & \dots & w^H A^{l-1} v \\ w^H A v & \dots & w^H A^l v \\ \vdots & & \vdots \\ w^H A^{l-2} v & \dots & w^H A^{2l-3} v \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{l-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

A normalization condition provides an additional equation. Here we choose $w^H A^{l-1} p_{l-1}(A)v = \eta$ to get the square system

$$M_l a := \begin{bmatrix} w^H v & w^H A v & \dots & w^H A^{l-1} v \\ w^H A v & w^H A^2 v & \dots & w^H A^l v \\ \vdots & \vdots & & \vdots \\ w^H A^{l-2} v & w^H A^{l-1} v & \dots & w^H A^{2l-3} v \\ w^H A^{l-1} v & w^H A^l v & \dots & w^H A^{2l-2} v \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{l-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \eta \end{bmatrix}$$

and $M_l$ is a Gram matrix with moments $\mu_{i,j} = \mathcal{L}\{z^{i+j}\} = c_{i+j} = w^H A^{i+j} v$ and thus, a Hankel matrix.

From $\mathcal{L}\{z^i p_{l-1}\} = 0$, for $0 \le i < l-1$, it follows that $p_{l-1}(z)$ is orthogonal to $\mathcal{P}_{l-2}$ with respect to the linear functional $\mathcal{L}$ defined by the $2l-1$ moments $\mathcal{L}\{z^i\} := m_i = w^H A^i v$, for $i = 0, 1, \dots, 2l-2$.

The other sequence $\{w_i\}$ in the pair is characterized by the polynomial sequence $\{q_i\}$, $q_i \in \mathcal{P}_i$ satisfying $w_l = q_{l-1}(A^H)w$ and $\langle w_l, A^i v\rangle_E = 0$, $i = 0, 1, \dots, l-2$. The induced system of equations for $q_{l-1}(z) = \beta_0 + \beta_1 z + \cdots + \beta_{l-1} z$ is

$$M_k^H b := \begin{bmatrix} v^H w & v^H A^H w & \dots & v^H (A^H)^{l-1} w \\ v^H A^H w & v^H (A^H)^2 w & \dots & v^H (A^H)^l w \\ \vdots & \vdots & & \vdots \\ v^H (A^H)^{l-2} w & v^H (A^H)^{l-1} w & \dots & v^H (A^H)^{2l-3} w \\ v^H (A^H)^{l-1} w & v^H (A^H)^l w & \dots & v^H (A^H)^{2l-2} w \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{l-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \nu \end{bmatrix}.$$

To reveal the relationship between $p_{l-1}(z)$ and $q_{l-1}(z)$, take a look at the Gram matrix formulation. Orthonormality, with normalization $\langle p_{l-1}(A)v, q_{l-1}(A^H)w\rangle_E = 1$, requires

$$M_k a = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \eta \end{bmatrix}, \qquad M_k^H b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \nu \end{bmatrix} \quad \text{and} \quad b^H M_k a = \bar{\nu}\eta = 1.$$

Group all coefficients of the pair of sequences of Lanczos polynomials $\{p_i\}_{i=0}^{k-1}$ and $\{q_i\}_{i=0}^{k-1}$ into the columns of $R_k^V$ and $R_k^W$, respectively. These are upper triangular matrices by construction and represent the LR factorization of the Gram matrix $M_k$,

i.e., $M_k = (R_k^W)^{-H}(R_k^V)^{-1}$. From the symmetry of $M_k$, i.e., $M_k^\top = M_k$, we have that $M_k = (R_k^W)^{-H}(R_k^V)^{-1} = (R_k^V)^{-\top}(\overline{R_k^W})^{-1}$ and by the uniqueness of the LR factorization, and for appropriate normalization, $R_k^W = \overline{R_k^V}$. Hence, $q_i(z) = \bar{p}_i(z)$ for all $i$, they only differ from each other by complex conjugation of their coefficients. This explains the commonly used name formal orthogonal polynomial instead of biorthogonal polynomials.
The Lanczos polynomials exist up to degree $k$ if and only if $M_k$ is quasi-definite. If they exist, then they are unique up to scaling [10]. If $M_k$ becomes singular, a notion of block orthogonality can be introduced to attempt to skip over the singular principal submatrix [10, 72]. A quasi-definite Hankel matrix implies a three term recurrence relation and the relation between the coefficients implies that the recurrence matrix is tridiagonal and symmetric.

### 6.3.2 Recurrence matrix

The Gram matrices for the Lanczos polynomials are Hankel matrices, therefore the same three term recurrence relation underlies these polynomials and the polynomial Lanczos iteration. The recurrence relation also can be obtained by starting from the two-sided Gram-Schmidt procedure, using orthogonality of the polynomials and $\mathcal{L}\{(zp(z))\, q(z)\} = \mathcal{L}\{p(z)\,(zq(z))\}$.

**Lemma 6.4** (Recurrence matrix for Lanczos polynomials)**.** *A biorthonormal pair of sequences of polynomials* $\{p_i\}_{i=0}^{k-1}$, $\{q_i\}_{i=0}^{k-1}$ *with respect to a quasi-definite linear functional* $\mathcal{L} : \mathcal{P}_{2k-1} \to \mathbb{C}$ *satisfies the pair of recurrence relations*

$$zp_i(z) = p_i(z)T_i + t_{i+1,i}p_{i+1}(z)$$

$$zq_i(z) = q_i(z)\overline{T}_i + \bar{t}_{i+1,i}q_{i+1}(z)$$

*for a tridiagonal matrix* $T_i \in \mathbb{C}^{k\times k}$.

The recurrence matrix $T_k \in \mathbb{C}^{k\times k}$

$$T_k := \begin{bmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_k \\ & & \beta_k & \alpha_k \end{bmatrix}$$

is called *complex symmetric* because $T_k = T_k^\top$. This is to stress that it is in general not Hermitian, which requires $T_k = T_k^\top$ and $T_k \in \mathbb{R}^{k\times k}$. The determinant of the recurrence matrix satisfies the three term recurrence relation represented by this matrix.

**Property 6.3** (Recurrence relation for recurrence matrix)**.** *The determinant of the recurrence matrix $T_k \in \mathbb{C}^{k\times k}$ from Lemma 6.4 satisfies a three term recurrence relation*

$$\det(T_1) = \alpha_1, \qquad \det(T_2) = \alpha_1\alpha_2 - \beta_2^2$$

$$\det(T_k) = \alpha_k \det(T_{k-1}) - \beta_2^2 \det(T_{k-2}).$$

*Proof.* Expand the determinant along last column and last row. □

The correspondence of the roots of $p_l(z)$ to the eigenvalues of $T_l$ is proved by Draux [57].

### 6.3.3 Relation to Lanczos iteration

To obtain a relation to the Lanczos iteration, a linear functional $\mathcal{L}$ can be defined using moments $w^H A^i v$ arising from Krylov subspaces. Lemma 6.5 provides an explicit form for the linear functional related to polynomial Krylov subspaces.

**Lemma 6.5** (Krylov induced linear functional)**.** *Consider Krylov subspaces $\mathcal{K}_k(A, v)$, $\mathcal{K}_k(A^H, w)$ for a diagonalizable matrix $A \in \mathbb{C}^{m\times m}$ and $v, w \in \mathbb{C}^m$. Then $\langle ., . \rangle_E$ on $\mathcal{K}_k(A, v) \times \mathcal{K}_k(A^H, w)$ induces the linear functional $\mathcal{L} : \mathcal{P}_{2k-1} \to \mathbb{C}$ with moments*

$$m_i = w^H A^i v.$$

*More precisely the linear functional is of the form $\mathcal{L}\{z^i\} = \sum_{k=1}^{g} \alpha_k z_k^i$, with $g$ the amount of distinct eigenvalues $\{z_k\}_{k=0}^{m}$ and corresponding nonzero weights $\alpha_k \in \mathbb{C}$.*

*Proof.* A diagonalizable $A$ has an eigenvalue decomposition $A = X\Lambda X^{-1}$, with $\Lambda = \operatorname{diag}(\{\lambda_i\})$. The linear functional can be defined by assigning values to its moments

$$\langle A^i v, (A^H)^j w\rangle_E = w^H A^{i+j} v = \underbrace{w^H X^{-1}}_{=:c_w} \Lambda^{i+j} \underbrace{Xv}_{c_v} = \sum_{k=1}^{m} (e_k^\top c_w)^H (e_k^\top c_v) \lambda_k^{i+j}.$$

Denote by $\{\tilde{z}_i\}_{i=1}^{d}$ the set of $d \leq m$ distinct eigenvalues of $A$. The weight $\tilde{\alpha}_i = \sum_{i=l,\lambda_l=z_i}^{m} (e_i^\top c_w)(e_i^\top c_v)$ corresponds to $\tilde{z}_i$ and can vanish. Let $\{\alpha_i\}_{i=1}^{g}$ be the weights in $\{\tilde{\alpha}_i\}_{i=1}^{d}$ which are nonzero and let $z_i$ be the node corresponding to $\alpha_i$. Then

$$m_i := \sum_{k=1}^{g} \alpha_k z_k^i, \quad i = 0, 1, \ldots, 2k-2$$

are the moments which define the linear functional $\mathcal{L}$. Note that $g$ is not necessarily equal the grade of $v$ or $w$, since weights of multiple eigenvalues may cancel each other. □

The connection between the Lanczos iteration and the construction of biorthogonal polynomials suggests that a normalization of the biorthonormal bases $V_k, W_k$ exists such that the tridiagonal matrix from Lemma 3.8 is complex symmetric.

**Corollary 6.2** (Biorthogonal Krylov bases recurrence relations-complex symmetric)**.** *Let $A \in \mathbb{C}^{m\times m}$, $v, w \in \mathbb{C}^m$. Consider the Krylov subspaces $\mathcal{K}_k(A, v), \mathcal{K}_k(A^H, w)$, with $k < \min\{g_v, g_w\}$. Then there exist, under the assumption that no breakdown occurs, biorthonormal nested bases $V_k, W_k \in \mathbb{C}^{m\times k}$ for these subspaces such that*

$$AV_k = V_k T_k + t_{k+1,k} v_k e_k^\top$$

$$A^H W_k = W_k \overline{T}_k + \bar{t}_{k+1,k} w_k e_k^\top,$$

*where $T_k \in \mathbb{C}^{k\times k}$ is a complex symmetric tridiagonal matrix.*

For more properties of the complex symmetric tridiagonal recurrence matrix and connections to biorthogonal polynomials we refer to the paper by Beckermann [12]. Throughout this manuscript the linear functional will be assumed to be quasi-definite, which corresponds to the no-breakdown assumption for the Lanczos iteration. A recent paper [136] discusses indefinite linear functionals and a look-ahead strategy to deal with breakdowns.

### 6.3.4 Applications

Computing the recurrence matrix $T_k$ corresponds to computing the recurrence coefficients of sequences of biorthogonal polynomials. Figure 6.3 shows a scheme for computing recurrence coefficients for biorthogonal polynomials via reformulation as a linear algebraic problem.

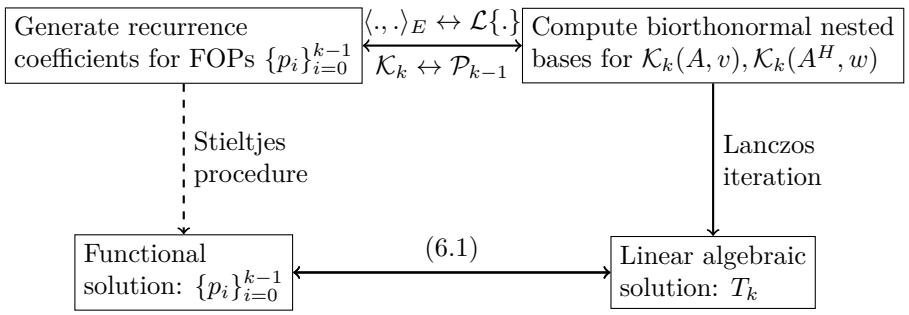



Figure 6.3: Scheme for generating recurrence coefficients for biorthogonal polynomials.

Applications of polynomials orthogonal with respect to an inner product and connections to problems in physics are well known. When the inner product is replaced by a (indefinite) linear functional, the applications are much less known. So a short justification of the study of biorthogonal polynomials is in order. An overview of applications for biorthogonal polynomials is provided:

- In model order reduction, see Chapter 5, moments of the form $w^H A^{-j-1} v$ occur naturally in the power series expansion of the transfer function [74, 136, 156].
- Sequence transformation methods, to accelerate the convergence of sequences [29, 57].
- Gaussian quadrature in the complex plane [137, 150] and quadrature of highly oscillatory integrals [42] make use of biorthogonal polynomials.

## 6.4 Conclusion

The connection between polynomials and structured matrices is identified. Triangle families of polynomials have a proper Hessenberg matrix as a recurrence matrix. The roots of the polynomials correspond to the eigenvalues of this matrix. Sequences of orthogonal and biorthogonal polynomials are triangle families and the recurrence matrices are shown to have Hessenberg and tridiagonal structure. For specific discrete inner products, the recurrence matrix for (bi)orthogonal polynomials corresponds to the recurrence matrix for (bi)orthogonal basis vectors for Krylov subspaces. The recurrence matrix allows us to represent a sequence of polynomials in the language of linear algebra.

# Chapter 7

# Orthogonal rational functions

Rational functions with prescribed poles are especially interesting for the approximation of certain functions in regions of the complex plane, see [138] and references therein. In numerical quadrature, if the function of interest is not well approximated by a polynomial, e.g, the function has singularities close to the region of interest, then rational Gauss quadrature rules [54, 81] can be used. These are Gauss quadrature rules which are exact for certain rational functions with prescribed poles.
The literature on orthogonal and biorthogonal rational functions is quite extensive: the monograph by Bultheel et al. [34] is a nice introduction into the field of orthogonal rational functions. The focus is often on inner products or linear functionals on the real line or the unit circle. There are some notable results for more general inner products and linear functionals. These are our primary interest, since the results in Chapter 4 are also of a general nature. The real line and unit circle are considered to be special cases of this more general theory. The study of these special cases in their own right is however still very important.
Numerical procedures to efficiently and stably solve problems involving (bi)orthogonal rational functions can be developed with a structured matrix approach. The goal of this chapter is to uncover the relation between (bi)orthogonal vectors in rational Krylov subspaces and (bi)orthogonal rational functions. This enables us to reformulate problems defined for rational functions as problems in linear algebra and solve them using linear algebraic techniques. For example, nodes for rational Gauss quadrature rules correspond to the roots of orthogonal rational functions and these correspond to eigenvalues of the associated recurrence pencil. Thus, Golub-Welsch type algorithms can be developed [54, 114, 139], which rely on solving a generalized eigenvalue problem for this recurrence pencil.
First a vector space of rational functions with prescribed poles is introduced in Section 7.1. Orthogonal rational functions which are related to rational Krylov subspaces

are introduced by deriving a suitable discrete inner product for rational functions in Section 7.2. Section 7.3 discusses orthogonality with respect to a linear functional, leading to biorthogonal rational functions. These can be represented by a tridiagonal pencil and are therefore related to biorthogonal rational Krylov subspaces. Section 7.4 summarizes notable results from the literature on ORF which are directly related to the topics in this manuscript. This summary is purposely kept short to show only the essence, references are provided which contain the details.

## 7.1 Rational functions and structured matrices

A rational function $r(z) \in \mathcal{R}$ is the ratio of two polynomials,

$$r(z) = \frac{p(z)}{q(z)}, \quad p(z), q(z) \in \mathcal{P}.$$

Our interest is in rational functions with prescribed poles. Given a set of poles $\Xi = \{\xi_1, \xi_2, \ldots, \xi_k\}$, with $\xi_i \in \overline{\mathbb{C}}$,

$$r(z) = \frac{p(z)}{\pi(z)}, \quad p(z) \in \mathcal{P} \text{ and } \pi(z) = \prod_{\substack{i=1 \\ \xi_i \neq \infty}}^{k} (z - \xi_i). \tag{7.1}$$

The space formed by these rational functions with prescribed poles will be denoted by $\mathcal{R}^{\Xi}$, which is a short notation for $\mathcal{P}/\pi(z)$. The finite dimensional space $\mathcal{R}_k^{\Xi}$ is defined as $\text{span}\{1, r_1(z), \ldots, r_k(z)\}$, with,

$$r_l(z) = \frac{p_l(z)}{\pi_l(z;\Xi)}, \quad \pi_l(z;\Xi) := \prod_{\substack{i=1 \\ \xi_i \neq \infty}}^{l} (z - \xi_i), \quad p_l \in \mathcal{P}_l, \quad \deg(p_l) = l. \tag{7.2}$$

The set of poles $\Xi$ for $\mathcal{R}_k^{\Xi}$ must contain (at least) $k$ poles $\xi_i \in \overline{\mathbb{C}}$. If $\Xi$ contains more than $k$ poles, the first $k$ are used for the space $\mathcal{R}_k^{\Xi}$, conform with the definition in (7.2).
If $\Xi$ contains exactly $k$ poles and the considered space is of dimension $k+1$, then the order of the poles does not matter, as stated by Property 7.1. This property is useful below to discuss moments arising from rational function spaces.

**Property 7.1** (Equivalence of rational functions spaces [170])**.** *Consider a set of $k$ poles $\Xi = \{\xi_1, \ldots, \xi_k\}$ and the rational function space $\mathcal{R}_k^{\Xi}$. For a fixed value of $k$ and $\widetilde{\Xi}$ any permutation of the poles in $\Xi$, the space $\mathcal{R}_k^{\widetilde{\Xi}} = \mathcal{R}_k^{\Xi}$.*

Of course, $\mathcal{R}_i^{\Xi}$ and $\mathcal{R}_i^{\widetilde{\Xi}}$, for $i < k$ and $\Xi$ containing $k$ poles, are not necessarily the same. Since nestedness, stated in Property 7.2, is paramount to obtain a connection to rational Krylov subspaces, the order of the poles cannot be arbitrarily changed.

**Property 7.2** (Nestedness of rational function spaces)**.** *For a fixed set $\Xi$ of $m$ poles, the spaces $\mathcal{R}_k^{\Xi}$, $k = 0, 1, \ldots, m$ are strictly nested, i.e.,*

$$\mathcal{R}_k^{\Xi} \subset \mathcal{R}_{k+1}^{\Xi}, \quad k = 0, 1, \ldots, m-1.$$

A nested basis $\{r_0, r_1, \ldots, r_{l-1}\}$ for the space $\mathcal{R}_{l-1}^{\Xi}$ generates a rational Krylov subspace as follows

$$\operatorname{span}\{r_0(A)v, r_1(A)v, \ldots, r_{l-1}(A)v\} = \mathcal{K}_l(A, v; \Xi), \quad l = 1, 2, \ldots, g, \tag{7.3}$$

with $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$ and $g$ the grade of $v$ with respect to $A$.
Such a nested basis satisfies a recurrence relation with a proper Hessenberg recurrence pencil $(H_l, K_l)$, let $\boldsymbol{r}_l := \begin{bmatrix} r_0 & r_1 & \ldots & r_{l-1} \end{bmatrix}$ and $r_l \in \mathcal{R}_l^{\Xi} \backslash \mathcal{R}_{l-1}^{\Xi}$, then

$$z\boldsymbol{r}_l K_l + k_{l+1,l} z r_l(z) = \boldsymbol{r}_l H_l + h_{l+1,l} r_l(z). \tag{7.4}$$

Hence, the pencil $(H_l, K_l)$ represents the sequence of rational functions $\{r_i\}_{i=0}^{l-1}$. Note that this is not the only possible representation, Chapter 4.1 discusses other representations for rational Krylov subspaces.
Only the Hessenberg recurrence pencil will be considered in the sequel. Section 7.1.1 describes how the Hessenberg recurrence pencil can be used to evaluate the sequence of rational functions.

### 7.1.1 Rational function evaluation

A sequence of rational functions $\{r_i(z)\}_{i=0}^{l}$ forming a nested basis for $\mathcal{R}_{l-1}^{\Xi}$ can be evaluated at $z = z^\star \notin \Xi$ using its recurrence pencil $(H_l, K_l)$ appearing in (7.4). Two possible procedures are discussed, one based on the solution of a linear system of equations and one based on solving a generalized eigenvalue problem. The interest in such procedures for rational function evaluation originates from the availability of efficient and stable numerical methods to solve these problems.

#### Solve linear system

The linear system to be solved is in fact a matrix notation of the recurrence relation (7.4) and can be solved by forward substitution. Proposition 7.1 provides the system to be solved.

**Proposition 7.1** (Rational function evaluation - system solve)**.** *Consider a sequence of rational functions* $\{r_i\}_{i=0}^{l-1}$ *with poles* $\Xi = \{\xi\}_{i=1}^{l-1}$ *with recurrence pencil* $(H_l, K_l) \in \mathbb{C}^{l\times l}$ (7.4)*. Then evaluating* $r_i(z)$*,* $i = 0, 1, \ldots, l-1$*, at* $z = z^\star \in \mathbb{C}\backslash\Xi$ *is equivalent to solving the system of equations*

$$x^\top \begin{bmatrix} 1 & \\ 0 & \\ \vdots & (H_n - z^\star K_n)I^{n\times(n-1)} \\ 0 & \\ 0 & \end{bmatrix} = \begin{bmatrix} r_0 & 0 & \ldots & 0 \end{bmatrix}$$

*for* $x \in \mathbb{C}^l$ *and where* $I^{l\times(l-1)}$ *is the* $l \times (l-1)$ *leading principal submatrix of the identity matrix. Equivalent in the sense that* $x_l := xe_l = r_{l-1}(z^\star)$*,* $1 < i < l$*.*

*Proof.* Analogous to the proof of Proposition 6.1. □

The condition is given by $\kappa(G)$, where $G := \begin{bmatrix} e_1 & (H_l - z^\star K_l)I^{l\times(l-1)} \end{bmatrix}$. This depends on both $(H_l, K_l)$ and $z^\star$. Note that the polynomial form, given in Proposition 6.1 is a special case, i.e., $K_l = I_l$.

### Eigenvalue decomposition

The roots of $r_l(z)$, appearing in recurrence relation (7.4), correspond to the generalized eigenvalues of $(H_l, K_l)$ if there are no multiple eigenvalues. In other words, if $(H_l, K_l)$ has $l$ distinct eigenvalues, then these eigenvalues are the $l$ roots of the rational function $r_l(z)$. Then the factored form of $r_l(z)$ can be obtained, i.e., $r_l(z) = \eta_l \dfrac{\prod_{i=1}^{l}(z-z_i)}{\prod_{\substack{j=1 \\ \xi_j\neq\infty}}^{l}(z-\xi_j)}$, with $\{z_i\}_{i=1}^{l}$ the roots and $\Xi = \{\xi_j\}_{j=1}^{l}$ the poles of $r_l(z)$. This is stated in Proposition 7.2.

**Proposition 7.2** (Rational function evaluation - eigenvalue decomposition)**.** *Consider a sequence of rational functions* $\{r_i\}_{i=0}^{l-1}$ *with poles* $\Xi = \{\xi\}_{i=1}^{l-1}$ *and recurrence pencil* $(H_l, K_l) \in \mathbb{C}^{l\times l}$ *satisfying* (7.4)*. If, for all* $j$*, the* $j \times j$ *principal leading subpencil* $(H_l^{(j)}, K_l^{(j)})$ *has simple generalized eigenvalues* $\{z_i^{(j)}\}$*, with* $z_i^{(j)} \neq \xi_k$ $i = 1, \ldots, j$ *and* $k = 1, 2, \ldots, l-1$*, then*

$$r_j(z) = \eta_j \frac{\prod_{i=1}^{j}(z - z_i^{(j)})}{\prod_{\substack{k=1 \\ \xi_k\neq\infty}}^{l}(z-\xi_k)}, \qquad \textit{where } \eta_j \textit{ is determined by normalization.}$$

*Proof.* Analogous to the proof of Proposition 6.2. □

Proposition 7.2 requires that none of the roots overlap with any of the poles in $\Xi$. In some cases this can be guaranteed. For orthogonal rational functions for an inner product, all roots are located in the convex hull of the support of the measure associated with this inner product [154]. In the sequel the support of the measure is some arc $\gamma$ in the complex plane. Hence, choosing poles outside this region guarantees that the requirement is satisfied.

The condition number for generalized eigenvalues $(\alpha, \beta)$ such that $\beta H_n x = \alpha K_n x$ can be defined using the (scale-dependent) chordal metric [90, p.396] or for $\lambda$ such that $H_n x = \lambda K_n x$, following Higham and Higham [103].

## 7.2 Orthogonal rational functions

Orthogonal rational functions are the obvious generalization of orthogonal polynomials. The vector space in which the functions of interest live is now $\mathcal{R}^{\Xi}$ instead of $\mathcal{P}$. A sequence of rational functions $\{r_i\}_i$, $r_i \in \mathcal{R}_i^{\Xi}$, is said to be orthogonal with respect to an inner product $\langle ., . \rangle : \mathcal{R}^{\Xi} \times \mathcal{R}^{\Xi} \to \mathbb{C}$ if $r_i \in \mathcal{R}_i^{\Xi}$, $r_i \notin \mathcal{R}_{i-1}^{\Xi}$ and

$$\langle r_k, r_j \rangle \begin{cases} = 0, & \text{if } k \neq j \\ > 0, & \text{if } k = j \end{cases} .$$

A suitable inner product defined for complex valued functions $f$, $g$ is very similar to the inner product used for OPs in Section 6.2. However, an additional restriction must be imposed, namely that the poles $\Xi$ cannot lie on support of the measure $\gamma$

$$\langle f, g \rangle := \int_{\gamma} f(z)\overline{g(z)}\alpha(z)|dz|,$$

where $\alpha(z)$ is a positive weight function, $\gamma$ an arc in the complex plane and $\Xi \cap \gamma = \emptyset$. In order to develop linear algebraic procedures for ORFs, in particular a procedure for the computation of their recurrence coefficients, a connection with rational Krylov subspaces is identified. Section 7.2.1 introduces the discrete inner product on $\mathcal{R}_l^{\Xi}$, whose relation to the Euclidean inner product on $\mathcal{K}_l(A, v; \Xi)$ is stated in Section 7.2.2. A procedure to generate ORFs starting from OPs is shortly described in Section 7.2.3.

### 7.2.1 Discrete inner product

A discrete inner product suitable for $\mathcal{R}_l^{\Xi}$ is

$$\langle f, g\rangle_m := \sum_{i=1}^{m} f(z_i)\overline{g(z_i)}\alpha(z_i), \quad z_i \in \mathbb{C} \tag{7.5}$$

with $m \geq l$ distinct nodes $\{z_i\}_{i=1}^m$, a positive weight function $\alpha(z)$ and poles that do not coincide with any of the nodes, i.e., $\Xi \cap \{z_i\}_{i=1}^m = \emptyset$. Discretization of a continuous inner product can be performed by using the moments $g_{i,j}$ associated with the poles $\Xi = \{\xi_1, \ldots, \xi_{l-1}\}$, if these moments are available.
These moments are defined to correspond to those appearing in Chapter 5 in the context of multi-point Padé approximation and in the literature on rational functions [170]. Let $\hat{\Xi} = \{\hat{\xi}_1, \ldots, \hat{\xi}_{\hat{l}}\}$ denote the set of distinct elements in $\Xi = \{\xi_1, \ldots, \xi_{l-1}\}$, that is, the set of distinct poles. And let $l_i$, for $i = 1, \ldots, \hat{l}$, denote the number of occurrences $l_i$ of $\hat{\xi}_i$ in $\Xi$. Then the *classical* moments associated with $\Xi$ are the zeroth moment $g_{0,0} := \int_\gamma \alpha(z)|dz|$ and

$$g_{i,j} = \begin{cases} \int_\gamma (z - \hat{\xi}_i)^{-j}\alpha(z)|dz|, & \text{if } \hat{\xi}_i \neq \infty \\ \int_\gamma z^j \alpha(z)|dz|, & \text{if } \hat{\xi}_i = \infty \end{cases}, \quad j = 1, \ldots, 2l_i, \quad i = 1, \ldots, \hat{l},$$

for a total of $2l - 1$ moments. An illustrative example is provided in Example 7.1.

**Example 7.1** (Example of moments for rational functions)**.** *Consider the set of poles* $\Xi = \{\xi_1, \infty, \xi_2, \xi_1, \infty\}$*. Then* $\hat{\Xi} = \{\xi_1, \infty, \xi_2\}$ *and* $\{n_1, n_2, n_3\} = \{2, 2, 1\}$*. The moments are*

$$g_{1,j} = \int_\gamma (z - \xi_1)^{-j}\alpha(z)|dz|, \quad j = 1, 2, 3, 4,$$

$$g_{2,j} = \int_\gamma z^j \alpha(z)|dz|, \quad j = 1, 2, 3, 4,$$

$$g_{3,j} = \int_\gamma (z - \xi_2)^{-j}\alpha(z)|dz|, \quad j = 1, 2.$$

*Together with the zeroth moment, these are* $2l - 1 = 11$ *moments.*

Modified moments also can be used and are preferred over the classical moments, since the map from the classical moments to recurrence coefficients is ill-conditioned. The map from modified moments to recurrence coefficients is better conditioned. The difficulty now is to obtain the modified moments accurately. For details about the conditioning of these maps and for a procedure to obtain modified moments via

interpolatory quadrature we refer to the paper by Van Deun and Bultheel [170]. Once a discrete inner product (7.5) is obtained, there is a connection to the Euclidean inner product for Krylov subspaces. The next section clarifies this connection.

### 7.2.2 Krylov induced inner product

The relation between the Euclidean inner product on rational Krylov subspaces $\mathcal{K}_l(A, v; \Xi)$, for $A \in \mathbb{C}^{m\times m}$ and of an inner product on the space $\mathcal{R}_{l-1}^{\Xi}$ is derived. Lemma 7.1 states that for a normal matrix $A$ the induced inner product is the weighted sum of function evaluations at the eigenvalues of $A$. A similar result is derived by Güttel [98, p.43].

**Lemma 7.1** (Krylov induced inner product)**.** *Consider a normal matrix $A \in \mathbb{C}^{m\times m}$ and $v \in \mathbb{C}^m$ of grade $g$. Let $\{b_i\}_{i=0}^{l-1}$ form a nested basis for $\mathcal{K}_l(A, v; \Xi)$, and let rational functions $r_i \in \mathcal{R}_i^{\Xi}$ be such that $b_i = r_i(A)v$ and $\Xi \cap \sigma(A) = \emptyset$. Then the Euclidean inner product $\langle ., . \rangle_E$ on $\mathcal{K}_l(A, v; \Xi)$ induces an inner product of the form*

$$\langle r_i, r_j \rangle_g = \sum_{k=1}^{g} \alpha_k \overline{r_j(z_k)} r_i(z_k), \quad \alpha_k > 0,$$

*on the space of rational functions $\mathcal{P}_{l-1}/\pi$ with $\pi(z) := \prod_{\substack{i=1 \\ \xi_i \neq \infty}}^{l-1} (z - \xi_i)$. If $g \geq l$, then $\mathcal{P}_{l-1}/\pi = \mathcal{R}_{l-1}^{\Xi}$.*

*Proof.* Similar to the proof of Lemma 6.3. □

Theorem 7.1 states that orthogonal vectors in rational Krylov subspaces correspond to ORFs, which provides the means to generate recurrence coefficients for a sequence of orthogonal rational functions via (rational) Krylov subspace methods.

**Theorem 7.1** (Krylov induced orthogonal rational functions)**.** *Consider a normal matrix $A \in \mathbb{C}^{m\times m}$ and $v \in \mathbb{C}^m$ with grade $g$. Let $\begin{bmatrix} q_0 & q_1 & \dots & q_{l-1} \end{bmatrix} = Q_l \in \mathbb{C}^{m\times l}$ form a nested orthonormal basis for $\mathcal{K}_l(A, v; \Xi)$, $l < g$. Let rational functions $r_i \in \mathcal{R}_i^{\Xi}$ satisfy $q_i = r_i(A)v$. Then the sequence of rational functions $\{r_i\}_{i=0}^{l-1}$ consists of orthonormal rational functions with respect to an inner product of the form*

$$\langle r_i, r_j \rangle_g = \sum_{k=1}^{g} \alpha_k \overline{r_j(r_k)} r_i(z_k).$$

*Proof.* Follows from Lemma 7.1, $\delta_{i,j} = \langle q_i, q_j \rangle_E = \langle r_i(A)v, r_j(A)v \rangle_E = \langle r_i, r_j \rangle_g$. □

A numerical procedure applying the result in this theorem to compute the recurrence coefficients for ORFs is proposed in Chapter 9. The following section discusses an alternative procedure.

### 7.2.3 Rational modification of a measure

A high level description of a procedure for the generation of recurrence coefficients of ORFs is provided. Details can be found in the literature. The procedure is proposed by Gautschi [76, 80, 81] and see also López Lagomasino et al. [127]. Consider the rational function $r(z) = \frac{p(z)}{\pi(z)} \in \mathcal{P}/\pi$, with a prescribed polynomial $\pi(z)$ satisfying $\pi(z) > 0$ on the support of some given measure $\mu$. Then a sequence of ORFs $\{r_i\}_i$, $r_i \in \mathcal{P}_i/\pi$, for the inner product $\langle r, s\rangle_\mu = \int_\gamma r(z)\overline{s(z)}d\mu(z)$ can be generated by

1. computing the recurrence coefficients of a sequence of OPs $\{p_i\}$ with respect to $\langle ., .\rangle_\mu$;
2. applying a procedure that implements the generalized Christoffel transformation [76, 190], which modifies the recurrence coefficients of $\{p_i\}$ to be orthogonal with respect to the measure $\tilde{\mu} := \frac{\mu}{|\pi(z)|^2}$;
3. the resulting recurrence coefficients generate the requested ORFs $\{r_i\}_i$.

We will not pursue this idea further, the procedure based on an inverse eigenvalue problem (IEP) for structured matrices provides a simpler relation between weights in the inner product and the measure required for the problem formulation. Such procedures, based on IEPs, are among the most effective and stable methods for the generation of recurrence coefficients of orthogonal functions [80, 141]. Chapter 9 proposes an IEP for ORFs with respect to discrete inner products (7.5) and corresponding procedures to solve them.

### 7.2.4 Gram matrix

The Gram matrix for rational functions for a discrete inner product (7.5) corresponds, by Lemma 7.1, to the Gram matrix for rational Krylov subspaces generated by appropriate $A$ and $v$. If the functions in $\boldsymbol{r}_l := \begin{bmatrix} r_0 & r_1 & \dots & r_{l-1} \end{bmatrix}$ form a nested basis for $\mathcal{R}_{l-1}^{\Xi}$ and $B_l^{\Xi}$ a nested basis for $\mathcal{K}_l(A, v; \Xi)$, and $\langle ., .\rangle_g$ is the inner product on $\mathcal{R}_{l-1}^{\Xi}$ induced by the Euclidean inner product on $\mathcal{K}_l(A, v; \Xi)$, then the Gram matrices are the same

$$(B_l^{\Xi})^H B_l^{\Xi} = \left[\langle r_i, r_j\rangle_g\right]_{j,i=1}^{l-1}.$$

A detailed description of the structure of Gram matrices for a rational Krylov subspace can be found in Section 5.2.3.
Orthogonal rational functions also allow a determinantal expression [34, Theorem 2.2.4].

## 7.3 Biorthogonal rational functions

Biorthogonality for rational functions is characterized by a linear functional $\mathcal{L}$. Biorthogonal rational functions are two sequences of rational functions $\{r_l\}_l$ and $\{s_k\}_k$ that satisfy, for some given sets of poles $\Xi$ and $\Theta$,

$$r_l \in \mathcal{R}_l^{\Xi} \quad \text{and} \quad s_k \in \mathcal{R}_k^{\Theta}$$

and

$$\mathcal{L}\{r_i(z)s_j(z)\} \begin{cases} = 0, & \text{if } i \neq j \\ \neq 0, & \text{if } i = j \end{cases}.$$

We will define a linear functional acting on the finite space $\mathcal{R}_{l-1}^{\Xi} \cdot \mathcal{R}_{l-1}^{\Theta}$, which is short for $\mathcal{R}_{2l-1}^{\{\Xi \cup \Theta\}}$. In the remainder of this chapter we assume that there are exactly $l-1$ poles in $\Xi$ and $\Theta$, then this notation makes sense. A sequence of moments can be used to define a linear functional, Section 7.3.1 provides the details. Once it is clear how a linear functional can be defined, Section 7.3.2 shows that a linear functional can be constructed which relates biorthogonal rational functions to biorthogonal vectors in rational Krylov subspaces.

### 7.3.1 Linear functional

A linear functional $\mathcal{L} : \mathcal{R}^{\Xi} \cdot \mathcal{R}^{\Theta} \to \mathbb{C}$ will be defined by fixing its values at certain elementary rational functions. These elementary rational functions are positive powers of $z$ or $(z-\xi)^{-1}$, for any $\xi \in \mathbb{C}$. Exactly which elementary basic functions are relevant depends on the choice of poles $\Xi$ and $\Theta$. Two cases are distinguished: the case $\Xi = \Theta$, where the two spaces have the same poles of the same multiplicity, and the case where none of the poles in $\Xi$ appear in $\Theta$. When only a subset of the poles in $\Xi$ corresponds to those in $\Theta$, these two cases can be combined.
A sequence of $2l-1$ moments $\{g_{i,j}\}$ must be available to be able to define a linear functional on $\mathcal{R}_{l-1}^{\Xi} \cdot \mathcal{R}_{l-1}^{\Theta}$.

**Case:** $\Xi = \Theta$

The poles $\Xi$ are reordered to obtain $\widetilde{\Xi}$ such that all equal poles are grouped together

$$\{\xi_1, \xi_2, \ldots, \xi_{l-1}\} \to \{\underbrace{\tilde{\xi}_1, \ldots, \tilde{\xi}_1}_{l_1}, \underbrace{\tilde{\xi}_2, \ldots, \tilde{\xi}_2}_{l_2}, \ldots, \underbrace{\tilde{\xi}_{\hat{l}}, \ldots, \tilde{\xi}_{\hat{l}}}_{l_{\hat{l}}}\},$$

with $\sum_{i=1}^{\hat{l}} l_i = l - 1$. The elementary rational functions are

$$1, \underbrace{\frac{1}{z-\tilde{\xi}_1}, \frac{1}{(z-\tilde{\xi}_1)^2}, \ldots, \frac{1}{(z-\tilde{\xi}_1)^{2l_1}}}_{l_1}, \ldots, \underbrace{\frac{1}{z-\tilde{\xi}_{\hat{l}}}, \frac{1}{(z-\tilde{\xi}_{\hat{l}})^2}, \ldots, \frac{1}{(z-\tilde{\xi}_{\hat{l}})^{2l_{\hat{l}}}}}_{l_{\hat{l}}}.$$

The linear functional $\mathcal{L}$ is defined by assigning a given value $g_{i,j}$ to $\mathcal{L}\{\frac{1}{(z-\tilde{\xi}_i)^j}\}$, i.e.,

$$\begin{cases} \mathcal{L}\{1\}, \\ \mathcal{L}\{\frac{1}{z-\tilde{\xi}_1}\}, \mathcal{L}\{\frac{1}{(z-\tilde{\xi}_1)^2}\}, \ldots, \mathcal{L}\{\frac{1}{(z-\tilde{\xi}_1)^{2l_1}}\}, \\ \mathcal{L}\{\frac{1}{z-\tilde{\xi}_2}\}, \mathcal{L}\{\frac{1}{(z-\tilde{\xi}_2)^2}\}, \ldots, \mathcal{L}\{\frac{1}{(z-\tilde{\xi}_2)^{2l_2}}\}, \\ \quad \vdots \\ \mathcal{L}\{\frac{1}{z-\tilde{\xi}_{\hat{l}}}\}, \mathcal{L}\{\frac{1}{(z-\tilde{\xi}_{\hat{l}})^2}\}, \ldots, \mathcal{L}\{\frac{1}{(z-\tilde{\xi}_{\hat{l}})^{2l_{\hat{l}}}}\} \end{cases} := \begin{cases} g_{0,0}, \\ g_{1,1}, g_{1,2}, \ldots, g_{1,2l_1}, \\ g_{2,1}, g_{2,2}, \ldots, g_{2,2l_2}, \\ \quad \vdots \\ g_{\hat{l},1}, g_{\hat{l},2}, \ldots, g_{\hat{l},2l_{\hat{l}}} \end{cases}.$$

**Case:** $\Xi \cap \Theta = \emptyset$

Consider the sets of poles $\Xi$ and $\Theta$ satisfying $\xi_i \neq \theta_j$ for all $i, j$. That is, the two sets have no poles in common. Reorder both sets to group equal poles, obtaining $\widetilde{\Xi}$ and $\widetilde{\Theta}$

$$\{\xi_1, \xi_2, \ldots, \xi_{l-1}\} \to \{\underbrace{\tilde{\xi}_1, \ldots, \tilde{\xi}_1}_{l_1}, \underbrace{\tilde{\xi}_2, \ldots, \tilde{\xi}_2}_{l_2}, \ldots, \underbrace{\tilde{\xi}_{\hat{l}}, \ldots, \tilde{\xi}_{\hat{l}}}_{l_{\hat{l}}}\},$$

$$\{\theta_1, \theta_2, \ldots, \theta_{k-1}\} \to \{\underbrace{\tilde{\theta}_1, \ldots, \tilde{\theta}_1}_{k_1}, \underbrace{\tilde{\theta}_2, \ldots, \tilde{\theta}_2}_{k_2}, \ldots, \underbrace{\tilde{\theta}_{\hat{k}}, \ldots, \tilde{\theta}_{\hat{k}}}_{k_{\hat{k}}}\},$$

with $\sum_{i=1}^{\hat{l}} l_i = l - 1$ and $\sum_{i=1}^{\hat{k}} k_i = l - 1$. The basis vectors are

$$1,$$

$$\frac{1}{z-\tilde{\xi}_1}, \ldots, \frac{1}{(z-\tilde{\xi}_1)^{l_1}}, \frac{1}{z-\tilde{\xi}_2}, \ldots, \frac{1}{(z-\tilde{\xi}_2)^{l_2}}, \ldots, \frac{1}{z-\tilde{\xi}_{\hat{l}}}, \ldots, \frac{1}{(z-\tilde{\xi}_{\hat{l}})^{l_{\hat{l}}}}$$

$$\frac{1}{z-\tilde{\theta}_1}, \ldots, \frac{1}{(z-\tilde{\theta}_1)^{k_1}}, \frac{1}{z-\tilde{\theta}_2}, \ldots, \frac{1}{(z-\tilde{\theta}_2)^{k_2}}, \ldots \frac{1}{z-\tilde{\theta}_{\hat{k}}}, \ldots, \frac{1}{(z-\tilde{\theta}_{\hat{k}})^{k_{\hat{k}}}}.$$

The linear functional $\mathcal{L}$ can be defined by assigning values from the sequence $\{g_{i,j}\}$ to the evaluation of $\mathcal{L}$ in these basisvectors.

### 7.3.2 Lanczos rational functions

The connection between the Euclidean inner product on rational Krylov subspaces and certain linear functionals on rational function spaces allows the construction of biorthogonal rational functions via structured matrix procedures. The biorthogonal rational functions that can be generated in this manner will be called *Lanczos rational functions*, in analogy with Lanczos polynomials. Appropriate linear functionals on rational function spaces are defined in Lemma 7.2 by using moments arising from rational Krylov subspaces.

**Lemma 7.2** (rational Krylov induced linear functional)**.** *Consider rational Krylov subspaces $\mathcal{K}_l(A, v; \Xi)$, $\mathcal{K}_l(A^H, w; \Theta)$ for a diagonalizable matrix $A \in \mathbb{C}^{m\times m}$ and $v, w \in \mathbb{C}^m$. Then $\langle ., . \rangle_E$ on $\mathcal{K}_l(A, v; \Xi) \times \mathcal{K}_l(A^H, w; \Theta)$ induces a linear functional $\mathcal{L}_g : \mathcal{R}^{\Xi}_{l-1} \cdot \mathcal{R}^{\Theta}_{l-1} \to \mathbb{C}$ of the form $\mathcal{L}_g\{t(z)\} = \sum_{k=1}^{g} \alpha_k t(z_k)$, with $g$ the amount of distinct eigenvalues of $A$ with nonzero weights $\alpha_k \in \mathbb{C}$ and $t(z) \in \mathcal{R}^{\Xi}_{l-1} \cdot \mathcal{R}^{\Theta}_{l-1}$.*

*Proof.* Consider $x \in \mathcal{K}_l(A, v; \Xi)$ and $y \in \mathcal{K}_l(A^H, w; \Theta)$ and rational functions $r, s$ such that $r(A)v = x$ and $\bar{s}(A^H)w = y$. Clearly $r \in \mathcal{R}^{\Xi}_{l-1}$ and $s \in \mathcal{R}^{\Theta}_{l-1}$. Then, for $A = X\Lambda X^{-1}$ with $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$,

$$\langle x, y\rangle_E = w^H s(A)r(A)v = \underbrace{w^H X}_{=:c_w^H} s(\Lambda)r(\Lambda)\underbrace{X^{-1}v}_{=:c_v} = \sum_{k=1}^{m} (e_k^\top c_w)^H (e_k^\top c_v) r(z_k)s(z_k).$$

Denote by $\{\tilde{z}_i\}_{i=1}^{d}$ the set of $d \leq m$ distinct eigenvalues of $A$. The weight $\tilde{\alpha}_k := \sum_{i=l,\lambda_l=z_i}^{m} (e_k^\top c_w)^H (e_k^\top c_v)$ corresponds to $\tilde{z}_i$ and can vanish. Let $\{\alpha_i\}_{i=1}^{g}$ be the weights in $\{\tilde{\alpha}_i\}_{i=1}^{d}$ which are nonzero and let $z_i$ be the node corresponding to $\alpha_i$. Then the linear functional for $t(z) \in \mathcal{R}^{\Xi}_{k-1} \cdot \mathcal{R}^{\Theta}_{k-1}$ has the form $\mathcal{L}_g\{t(z)\} = \sum_{k=1}^{g} \alpha_k t(z_k)$. □

The linear functional from Lemma 7.2 is characterized by its moments $g_{0,0} = w^H v$ and

$$g_{i,j} = m_{j-1}(\xi_i), \text{ for } i, j \geq 1,$$

with $m_j(\xi)$ as defined in (5.2). Our main interest in this functional, however, is that we can determine the nodes $\{z_i\}_i$ and weights $\{e_i^\top \bar{c}_w, e_i^\top c_v\}_i$ which are used in the linear functional. For more details about the moments we remark that the Gram matrix generated by $\mathcal{L}$ for $\mathcal{R}^{\Xi}_{l-1} \cdot \mathcal{R}^{\Theta}_{l-1}$ can be studied by relating it to the Gram matrix for $\mathcal{K}_l(A, v; \Xi) \times \mathcal{K}_l(A^H, w; \Theta)$ and this is discussed in Section 5.2.3.

### Recurrence pencil

Lanczos rational functions, which are biorthogonal rational functions, satisfy a recurrence relation governed by a tridiagonal recurrence pencil. This result follows from combining Theorem 7.2, relating biorthogonal vectors in RKS to Lanczos rational functions, with Theorem 4.6, proving that a tridiagonal recurrence pencil underlies the biorthogonal Krylov vectors.

**Theorem 7.2** (Krylov induced biorthogonal rational functions)**.** *Consider a diagonalizable matrix $A \in \mathbb{C}^{m\times m}$ and $v \in \mathbb{C}^m$ with grade $g$. Let $\begin{bmatrix} v_0 & v_1 & \dots & v_{l-1} \end{bmatrix} = V_l \in \mathbb{C}^{m\times l}$ and $\begin{bmatrix} w_0 & w_1 & \dots & w_{l-1} \end{bmatrix} = W_l \in \mathbb{C}^{m\times l}$ form nested biorthonormal bases for $\mathcal{K}_l(A, v; \Xi)$ and $\mathcal{K}_l(A^H, w; \Theta)$. Let rational functions $r_i \in \mathcal{R}_i^{\Xi}$, $s_i \in \mathcal{R}_i^{\Theta}$ satisfy $v_i = r_i(A)v$, $w_i = s_i(A^H)w$. Then the pair of sequences of rational functions $\{r_i\}_{i=0}^{l-1}$ and $\{s_i\}_{i=0}^{l-1}$ are biorthonormal with respect to the linear functional*

$$\mathcal{L}_g\{t(z)\} = \sum_{k=1}^{g} \alpha_k t(z_k).$$

*Proof.* It follows from Lemma 7.2 that

$$\delta_{i,j} = \langle v_i, w_j \rangle_E = \langle r_i(A)v, s_j(A^H)w \rangle_E = \mathcal{L}_g\{r_i(z)s_j(z)\}.$$

□

This connection is exploited in Chapter 9, where linear algebraic techniques for the generation of recurrence coefficients of Lanczos rational functions are developed.

## 7.4 Connections to ORF literature

The literature on orthogonal rational functions is too vast to be summarized in a single section. Only some notable results which are immediately applicable to the discussion in this manuscript are summarized. In the ORF literature, often rational Krylov subspaces are generated with $A$ and $A^\top$ and the bilinear form $\langle x, y\rangle_\top = y^\top x$, $x, y \in \mathbb{C}^m$ instead of our choice $A$, $A^H$ and $\langle x, y\rangle_E = y^H x$. The results we obtained above remain, with minor modifications, valid for $A$, $A^\top$ and $\langle x, y\rangle_\top$. Appendix A.2 contains modified proofs of the principal theorems of Chapter 3 and Chapter 4 on matrix structures in Krylov subspaces. Tridiagonal recurrence pencils underlying (bi)orthogonal rational functions are discussed in Section 7.4.1. Gauss quadrature rules use the roots of OPs as nodes, and rational Gauss quadrature uses the roots of ORFs. Section 7.4.2 provides some references to results linking these roots to eigenvalue problems for structured matrices. There are also some connections to spectral theory, which are discussed in Section 7.4.3.

### 7.4.1 Tridiagonal pencil

The connection between a tridiagonal recurrence pencil and biorthogonal rational functions with finite poles seems to be first explicitly noted by Zhedanov [192]. Zhedanov shows that $\mathrm{R_{II}}$-polynomials arise as the eigenvectors from a generalized eigenvalue problem for a tridiagonal pencil. These $\mathrm{R_{II}}$-polynomials satisfy the recurrence relation

$$P_{n+1}(z) + \rho_n(v_n - z)P_n(z) + u_n(z - a_n)(z - b_n)P_{n-1}(z) = 0, \tag{7.6}$$

with initial conditions

$$P_{-1}(z) \equiv 0, \qquad P_0(z) \equiv 1.$$

Assume that the following restrictions are satisfied

$$P_n(a_k) \neq 0, \qquad P_n(b_k) \neq 0 \qquad u_n \neq 0, \qquad \text{for all} \quad n, k. \tag{7.7}$$

For brevity we introduce the following notations:

$$A_0 = B_0 = 1, \qquad A_n(z) = \prod_{k=1}^{n}(z - a_k), \qquad B_n(z) = \prod_{k=1}^{n}(z - b_k).$$

The related generalized eigenvalue problem is $T\psi = \lambda S\psi$, where the components $\psi_i$ of $\psi$ are rational functions in $\lambda$. The relation between the pencil $(T, S)$ and the

recurrence relation (7.6) is:

$$\rho_n = s_{n,n}, \quad v_n = \frac{t_{n,n}}{s_{n,n}}, \qquad n = 1, 2, \ldots, m,$$

$$u_n = s_{n,n+1}s_{n+1,n}, \quad b_n = \frac{t_{n,n+1}}{s_{n,n+1}}, \quad a_n = \frac{t_{n+1,n}}{s_{n+1,n}}, \qquad n = 1, 2, \ldots, m-1.$$

For a specific $R_{II}$-polynomial sequence Ismail and Sri Ranga [109] showed that the zeros are simple and lie on the real line, which means that all the roots of these polynomials can be obtained by solving a related generalized eigenvalue problem.
A connection between some Geronimus transformations of complex Jacobi matrices for biOPs and tridiagonal pencils for $R_{II}$-polynomials is discussed in [6].

#### Note on the history

Ismail and Masson [108] introduced $R_{II}$-polynomials via related continued fractions. Earlier Hendriksen and Njåstad [101] introduced them in relation to multi-point Padé approximation. A more recent paper on tridiagonal pencils, continued fractions and biorthogonal rational functions [13] includes the connection to multi-point Padé approximation, a Favard type theorem and an explicit expression for the linear functional accompanying this Favard theorem.

### 7.4.2 Rational Gauss quadrature

Gauss quadrature rules that are exact for rational functions with prescribed poles are called *rational Gauss quadrature rules*. These are interesting if the function of interest does not behave like a polynomial, for example it has singularities outside of the region of interest and close to its boundary. The use of rational Gauss quadrature is illustrated in [54, 81]. One approach is to start from a sequence of OPs and apply rational transformations on the associated measure to obtain a sequence of ORFs [76, 160]. Another approach uses the structure of the recurrence matrix or pencil of orthogonal rational functions to generate a sequence of ORFs directly [54, 55, 113, 127, 139]. These are Golub-Welsch type algorithms [91].
For quadrature it is interesting that the weights are positive and the nodes are inside the convex hull of the measure. Therefore, mainly orthogonal rational functions (with respect to an inner product) are considered in literature on rational Gauss quadrature. For a tridiagonal recurrence pencil representing ORFs it is proved [33, 54, 55] that, under some mild conditions on the poles, the roots of the ORFs and eigenvalues of a GEVP for a tridiagonal pencil correspond.

### 7.4.3 Spectral theory

In spectral theory, the recurrence matrices and pencils that are discussed in this thesis appear as well. The CMV decomposition [41] is a well-known example, which is a pentadiagonal matrix or sparse tridiagonal pencil, the relation to rational Krylov subspaces can be found in the paper by Watkins [179]. See also the survey paper by Simon [153].
So-called *generalized moment problem matrices* (GMP matrices) have been receiving some attention the past years [59, 187]. These matrices correspond to the matrices in Theorem 4.5. Eichinger and collaborators [59] use GMP matrices to study asymptotic behavior of orthogonal rational functions. In order to do so, the GMP matrix must be a finite band matrix, which restricts the choice of possible poles for the underlying rational functions. Since the tridiagonal pencil representation always has finite band structure for any choice of poles, this might be an interesting representation to study instead of the single matrix representation. This is future research.

### 7.4.4 Conclusion

Literature on orthogonal rational functions, rational Gauss quadrature and spectral theory for GMP and CMV matrices contain many theoretical results. Some of these results are directly applicable to the structured matrices and pencils studied here. For numerical computation, for example constructing (bi)orthogonal rational functions, the numerical linear algebraic approach taken in this manuscript can lead to powerful algorithms. Algorithms are proposed in Chapter 8 and Chapter 9.

## 7.5 Conclusion

Sequences of orthogonal and biorthogonal rational functions are shown to be related to Hessenberg and tridiagonal pencils, respectively. For certain discrete inner products and linear functionals these sequences correspond to orthogonal and biorthogonal bases for rational Krylov subspaces. The general form of such inner products and linear functionals is derived, they are weighted sums of function evaluations. These relations allow us to represent rational functions as matrices. Related results appearing in the literature on orthogonal rational functions is summarized.

# Chapter 8

# Biorthogonal Methods

Biorthogonal methods are less popular in numerical linear algebra than their orthogonal counterparts. Contrary to orthogonal methods, biorthogonal methods often do not possess the inherent stability that is natural to orthogonal transformations used in orthogonal methods. Sometimes biorthogonal methods are required because it is dictated by the problem that is posed, e.g., biorthogonal polynomials appear in quadrature [32, 42, 136, 137, 150], non-Hermitian operators appear in quantum mechanics [9, 48, 185, 186] and the approximation of the time-ordered exponential is attempted by using a variant of the non-Hermitian Lanczos iteration [87].
If we do have a choice to use orthogonal or biorthogonal methods, it can still be worth considering biorthogonal methods. Biorthogonal methods often lead to more efficient procedures, which allow to solve larger problems than their orthogonal counterpart. For example, generating biorthonormal bases for Krylov subspaces uses a three term recurrence relation, implying that only six basis vectors must be kept in memory at a time, since orthogonalization must only be done with respect to two previous basis vectors. For orthonormal bases, in general, all basis vectors must be kept in memory. The fast deterioration of the biorthonormality and the possibility of serious breakdowns make a simple method generating biorthonormal bases unsuited for numerical computation. With the use of reorthogonalization procedures and look-ahead strategies, some of the efficiency is lost, but the numerical behavior of biorthogonal Krylov subspaces methods is improved. Thus, biorthonormal methods are interesting as a basis for numerical algorithms, however a lot of research is often needed to develop biorthogonal algorithms that are robust and efficient. One example is BiCGSTAB [168], which solves systems of equations for general matrices using the non-Hermitian Lanczos iteration.
Biorthogonal methods are also an interesting theoretical tool. By studying a general biorthogonal framework, it is possible to obtain results for special cases. This is

illustrated in Chapter 5, where a Toeplitz matrix is interpreted as a Gram matrix related to a pair of rational Krylov subspaces.
In this chapter the potential efficiency gain and the power of the general framework are illustrated. Section 8.1 is based on [165] and Section 8.2 on [166]. In Section 8.1 the rational Lanczos iteration, which follows from the results in Chapter 4, is discussed and some numerical testing is done as a proof of concept. Section 8.2 derives short recurrence relations for biorthonormal bases spanning a specific pair of rational Krylov subspaces. Two approaches to derive these recurrence relations are used, one based on orthogonality properties of the Euclidean inner product and the structure of the recurrence matrix and one based on a Levinson procedure. The specific pair of rational Krylov subspaces is related to the CMV decomposition. The CMV decomposition received a lot of attention since its introduction into the literature on spectral theory [41, 153, 179].

## 8.1 Rational Lanczos iteration

An efficient procedure to compute biorthonormal nested bases for rational Krylov subspaces

$$\mathcal{K}(A, v; \Xi) \text{ and } \mathcal{K}(A^H, w; \Theta)$$

is suggested by Theorem 4.6 and Lemma 5.6. If no breakdowns occur, these bases, $V_k, W_k \in \mathbb{C}^{m\times k}$, span the rational Krylov subspaces

$$\text{span}\{V_k\} = \text{span}\{v_0, v_1 \ldots, v_{k-1}\} = \mathcal{K}_k(A, v; \Xi)$$

$$\text{span}\{W_k\} = \text{span}\{w_0, w_1 \ldots, w_{k-1}\} = \mathcal{K}_k(A^H, w; \Theta)$$

and are biorthonormal

$$W_k^H V_k = I.$$

They are generated by the recurrence relation

$$AV_{k+1}\underline{S}_k = V_{k+1}\underline{T}_k,$$

$$A^H W_{n+1}\underline{\widetilde{T}}_n = W_{n+1}\underline{\widetilde{S}}_n,$$

with tridiagonal pencils $(\underline{T}_k, \underline{S}_k), (\widetilde{\underline{T}}_k, \widetilde{\underline{S}}_k) \in \mathbb{C}^{(k+1)\times k} \times \mathbb{C}^{(k+1)\times k}$. The matrices in these pencils are

$$\underline{T}_k =: \begin{bmatrix} d_1 & a_2 & & \\ b_2 & d_2 & \ddots & \\ & \ddots & \ddots & a_k \\ & & b_k & d_k \\ & & & b_{k+1} \end{bmatrix}, \qquad \underline{S}_k =: \begin{bmatrix} c_1 & u_2 & & \\ l_2 & c_2 & \ddots & \\ & \ddots & \ddots & u_k \\ & & l_k & c_k \\ & & & l_{k+1} \end{bmatrix},$$

$$\widetilde{\underline{T}}_k =: \begin{bmatrix} \gamma_1 & \mu_2 & & \\ \lambda_2 & \gamma_2 & \ddots & \\ & \ddots & \ddots & \mu_k \\ & & \lambda_k & \gamma_k \\ & & & \lambda_{k+1} \end{bmatrix}, \qquad \widetilde{\underline{S}}_k =: \begin{bmatrix} \delta_1 & \alpha_2 & & \\ \beta_2 & \delta_2 & \ddots & \\ & \ddots & \ddots & \alpha_k \\ & & \beta_k & \delta_k \\ & & & \beta_{k+1}, \end{bmatrix}$$

with the poles appearing as the sub-and superdiagonal ratios

$$\frac{b_{i+1}}{l_{i+1}} = \xi_i, \qquad \frac{\beta_{i+1}}{\lambda_{i+1}} = \theta_i, \qquad i = 1, 2, \ldots, k,$$

$$\frac{a_{i+1}}{u_{i+1}} = \overline{\theta}_{i-1}, \qquad \frac{\alpha_{i+1}}{\mu_{i+1}} = \overline{\xi}_{i-1}, \qquad i = 1, 2, \ldots, k-1.$$

The elements appearing in the recurrence pencils follow from these ratios, biorthogonality conditions on the columns of $V_k, W_k$ and from normalization. The expressions for these elements can be found in Appendix B.1, the derivation is lengthy and technical and therefore it is not included here. An implementation is available online [162].
The following short discussion is a proof of concept, it verifies that the proposed iteration is valid. The validity of the rational Lanczos iteration is verified by applying it to solve an eigenvalue problem. Its behavior is then compared qualitatively with the polynomial Lanzcos iteration and with rational Krylov subspace methods. To talk about the numerical behavior, three metrics are introduced. The biorthonormality of these bases is quantified by the *biorthonormality error*

$$\|W_k^H V_k - I\|_2.$$

The *recursion error* quantifies the accuracy of the recurrence pencil,

$$\|W_{k+1}^H A V_{k+1} \underline{S}_k - \underline{T}_k\|_2.$$

A Ritz plot is used to visualize the accuracy of the Ritz values. The Ritz values $\theta^{(k)}$ of $(T_k, S_k)$, obtained by removing the last row of $(\underline{T}_k, \underline{S}_k)$ and computing its eigenvalues, are compared with the eigenvalues $\lambda$ of $A$. Ritz plots visualize how close the $k$ Ritz

values $\theta_i^{(k)}$, $1 \leq i \leq k$, are to the closest eigenvalue $\lambda_i := \min_\lambda |\theta_i^{(k)} - \lambda|$, for increasing $k$. The colors show how accurate the approximation is:

$$\text{red: } \|\theta_i^{(k)} - \lambda_i\|_2 < 10^{-8},$$

$$\text{yellow: } \|\theta_i^{(k)} - \lambda_i\|_2 < 10^{-5},$$

$$\text{green: } \|\theta_i^{(k)} - \lambda_i\|_2 < 10^{-2},$$

$$\text{blue: } \|\theta_i^{(k)} - \lambda_i\|_2 \geq 10^{-2}.$$

Consider a random $50 \times 50$ upper triangular matrix with eigenvalues $\lambda_i = i$, $i = 1, 2, \ldots, 50$. Krylov subspaces $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$ are built using $v = w$ and $\Xi = \Theta := \{0, 24.1, 0, 24.1, \ldots\}$. Figure 8.1 shows the metrics and Figure 8.2 shows the Ritz plot. The Ritz plot clearly shows that convergence is concentrated around the chosen poles 0 and 24.1. This is the expected behavior, the convergence of rational Krylov subspace methods can be focused on certain parts of the spectrum [143].

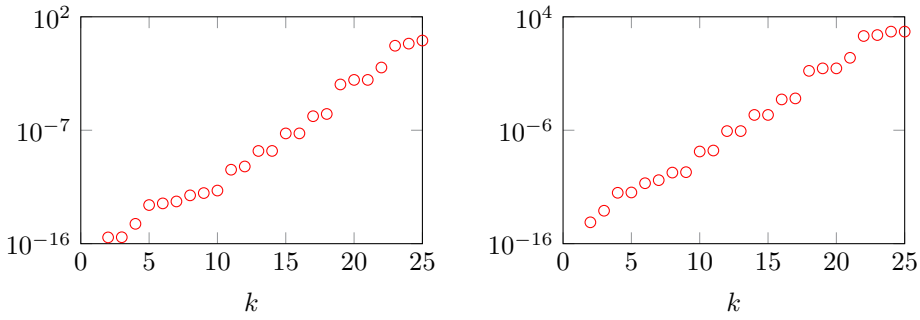



Figure 8.1: Biorthonormality and recurrence error, respectively left and right, for the tridiagonal recurrence pencil generated by the rational Lanczos iteration, with $\Xi = \Theta = \{0, 24.1, 0, 24.1, \ldots\}$.

To show the connection between convergence of eigenvalues and biorthonormality, we choose a pole closer to an eigenvalue. This leads to faster convergence to this eigenvalue and thus to faster loss of biorthonormality [132]. The poles chosen now are $\Xi = \Theta := \{0, 24 + 10^{-5}, 0, 24 + 10^{-5}, \ldots\}$. Figure 8.4 shows that the eigenvalue $\lambda = 24$ is found in fewer iterations than for the poles chosen above. The faster loss of biorthonormality is apparent from Figure 8.3. It also shows that the quality of the recurrence pencil is related to the biorthonormality of the bases.

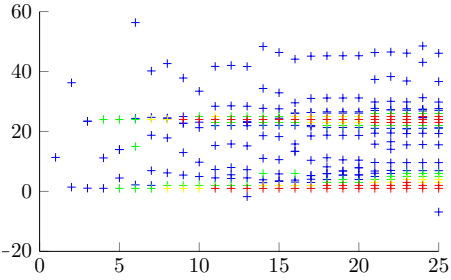



Figure 8.2: Ritz plot for the tridiagonal recurrence pencil generated by the rational Lanczos iteration, with $\Xi = \Theta = \{0, 24.1, 0, 24.1, \dots\}$. X-axis shows the size of the pencil.

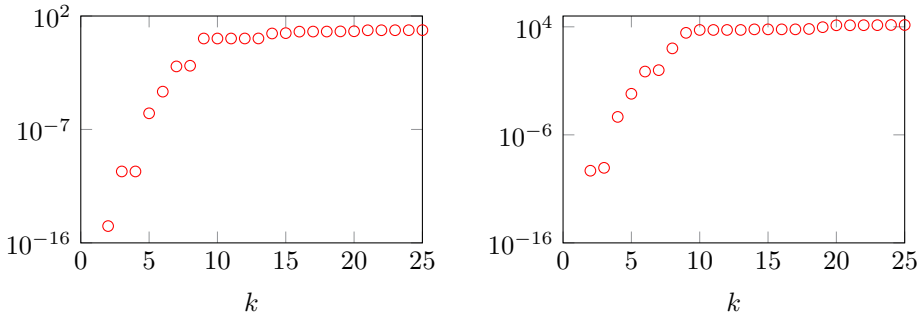



Figure 8.3: Biorthonormality and recurrence error, respectively left and right, for the tridiagonal recurrence pencil generated by the rational Lanczos iteration, with $\Xi = \Theta = \{0, 24 + 10^{-5}, 0, 24 + 10^{-5}, \dots\}$.

Note that we did not consider the case $\Xi \cap \Theta = \emptyset$, since the behavior of such choices is not comparable with any existing Lanczos-type iterations and is subject to future research. We conclude that the novel rational Lanczos iteration exhibits the expected behavior, i.e., comparable to that of known iterations. Hence, its validity is substantiated.

**Remark 8.1** (The rational Lanczos iteration of Grimme)**.** *In the context of model order reduction, Grimme and collaborators [73, 74, 94], discussed the use of an iteration resembling the rational Lanczos iteration proposed here. This is the iteration related to moment matching and multi-point Padé approximation described in Chapter 5. Their method generates biorthonormal bases for rational Krylov subspaces with equal poles, i.e., for $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Xi)$, and is for this choice of poles very similar to our rational Lanczos iteration. But not the same, the recurrence pencil that they propose*

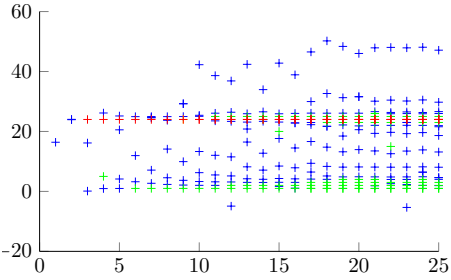



Figure 8.4: Ritz plot for the tridiagonal recurrence pencil generated by the rational Lanczos iteration, with $\Xi = \Theta = \{0, 24 + 10^{-5}, 0, 24 + 10^{-5}, \dots\}$. X-axis shows the size of the pencil.

*differs from a tridiagonal pencil. The structure of their recurrence pencil depends on the chosen poles. Every time the pole changes, i.e., $\xi_i \neq \xi_{i-1}$, the tridiagonal structure of the pencil is distorted by nonzeros in column $i+1$ and $i+2$ in the upper triangular part of the matrices in the pencil. These nonzeros appear up to row $j$, with $j$ such that $\xi_i = \xi_j$, $j < i$. Such a peak implies that more basis vectors must be kept in memory than for the tridiagonal pencil that we compute by using our rational Lanczos iteration.*

**Remark 8.2** (Moment matching)**.** *Some of the results obtained by Grimme and collaborators are valid here. Most notably, they proved the moment matching property of projections onto rational Krylov subspaces [73, Theorem 3]. This property is valid for our approach, since it only makes use of the oblique projector $V_k W_k^H$ onto rational Krylov subspaces with equal poles and does not depend on the structure of the recurrence pencil.*

### 8.1.1 Conclusion

Based on the recurrence pencil results from Chapter 4 a rational Lanczos iteration is proposed which generates biorthonormal bases via a three term recurrence relation. The proposed rational Lanczos iteration is valid, its behavior resembles that of the non-Hermitian Lanczos iteration and that of other rational Krylov subspaces methods. Further analysis of its finite precision behavior is required to develop robust algorithms based on this iteration. The requirements of such an algorithm will depend on the specific problem considered. One possible avenue is to reconsider the rational Lanczos iteration for model order reduction [73], but now starting from a more solid theoretical background linking the bases to structured matrices.

## 8.2 Nonunitary CMV decomposition

For $A \in \mathbb{C}^{m\times m}$ and $v, w \in \mathbb{C}^m$, $\langle v, w\rangle_E \neq 0$ consider the spaces associated with $\Xi := \{\infty, 0, \infty, 0, \dots\}$ and $\Theta := \{0, \infty, 0, \infty, \dots\}$

$$\mathcal{K}(A, v; \Xi) = \operatorname{span}\{v, Av, A^{-1}v, A^2v, A^{-2}v, \dots\}, \tag{8.1}$$

$$\mathcal{K}(A^H, w; \Theta) = \operatorname{span}\{w, A^{-H}w, A^Hw, A^{-2H}w, A^{2H}w, \dots\}. \tag{8.2}$$

A Levinson procedure is used to derive short recurrence relations for biorthonormal bases spanning these spaces. These are in fact extended Krylov subspaces [58] and the related rational functions are Laurent polynomials

$$\mathcal{R}^{\Xi} = \operatorname{span}\{1, z, z^{-1}, z^2, z^{-2}, \dots\},$$

$$\mathcal{R}^{\Theta} = \operatorname{span}\{1, z^{-1}, z, z^{-2}, z^2, \dots\}.$$

If $A$ is a unitary matrix, these subspaces are connected to the CMV-decomposition [41, 153]. Watkins [179] showed that, for a unitary matrix and $v = w$, an orthonormal basis can be constructed by short recurrence relations. This orthonormal basis spans $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$ simultaneously. He also discussed the link to orthogonal Laurent polynomials, quadrature formulas, Szegő polynomials and Toeplitz matrices. Here the restriction to unitary matrices is dropped. The connection between extended Krylov subspaces and Laurent polynomials implies short recurrence relations for biorthogonal Laurent polynomials [191].
The latter is shown by making use of the Gram matrix related to the pair of extended Krylov subspaces. The link to biorthogonal Szegő polynomials [8] and quadrature rules is not discussed. Section 8.2.1 uses orthogonality properties to derive short recurrence relations to construct biorthogonal bases for the pair of extended Krylov subspaces. A sparse, factored matrix representation of these recurrence relations is given by the recurrence matrix or pencil. The recurrence pencil is also derived in Section 8.2.2 by using a Levinson-type procedure. The link to the paper of Watkins [179] and the CMV-decomposition [41] is elaborated on in Section 8.2.3. Section 8.2.4 discusses the numerical properties of the proposed recurrence relations. This discussion is limited to a proof of concept, it verifies that the recurrence relations are valid and provides an indication of their stability.

### 8.2.1 Short recurrence relations

Consider a nonunitary, nonsingular matrix $A \in \mathbb{C}^{m\times m}$, vectors $v, w \in \mathbb{C}^m$, $\langle v, w\rangle_E \neq 0$, and the pair of subspaces $\mathcal{K}(A, v; \Xi)$ (8.1) and $\mathcal{K}(A^H, w; \Theta)$ (8.2). The goal is to construct biorthonormal bases $V_l, W_l \in \mathbb{C}^{m\times l}$ for the finite extended Krylov subspaces

$\mathcal{K}_l(A, v; \Xi)$ and $\mathcal{K}_l(A^H, w; \Theta)$, respectively. These subspaces are, for a nonnegative integer $k$:
For odd $l = 2k + 1$,

$$\mathcal{K}_l(A, v; \Xi) = \text{span}\{v, Av, A^{-1}v, A^2v, A^{-2}v, \ldots, A^kv, A^{-k}v\},$$

$$\mathcal{K}_l(A^H, w; \Theta) = \text{span}\{w, A^{-H}w, A^Hw, A^{-2H}w, A^{2H}w, \ldots, A^{-kH}w, A^{kH}w\}.$$

For even $l = 2k + 2$,

$$\mathcal{K}_l(A, v; \Xi) = \text{span}\{v, Av, A^{-1}v, A^2v, A^{-2}v, \ldots, A^kv, A^{-k}v, A^{k+1}v\},$$

$$\mathcal{K}_l(A^H, w; \Theta) = \text{span}\{w, A^{-H}w, A^Hw, \ldots, A^{-kH}w, A^{kH}w, A^{(-k-1)H}w\}.$$

The nested biorthonormal bases satisfy

$$\text{span}\{V_i\} = \text{span}\{v_0, v_1, \ldots, v_{i-1}\} = \mathcal{K}_i(A, v; \Xi),$$

$$\text{span}\{W_i\} = \text{span}\{w_0, w_1, \ldots, w_{i-1}\} = \mathcal{K}_i(A^H, w; \Theta), \text{ for } i = 1, \ldots, l,$$

$$W_l^H V_l = I.$$

The bases $V_l, W_l$ are obtained by computing the basis vectors $v_l$ and $w_l$ such that $\text{span}\{V_{l-1}, v_{l-1}\} = \mathcal{K}_l(A, v; \Xi)$, $\text{span}\{W_{l-1}, w_{l-1}\} = \mathcal{K}_l(A^H, w; \Theta)$ and $v_{l-1} \perp \text{span}\{W_{l-1}\} = \mathcal{K}_{l-1}(A^H, w; \Theta)$, $w_{l-1} \perp \text{span}\{V_{l-1}\} = \mathcal{K}_{l-1}(A, v; \Xi)$.
To compute $v_{l-1}$, any vector $x \in \mathcal{K}_l(A, v; \Xi) \backslash \mathcal{K}_{l-1}(A, v; \Xi)$ is a valid candidate. And for $w_{l-1}$, any vector $y \in \mathcal{K}_l(A^H, w; \Theta) \backslash \mathcal{K}_{l-1}(A^H, w; \Theta)$. The choice of candidate will influence the length of the recurrence relation generating the biorthonormal basis vectors. We will look at two such choices, one leading to a 4-term recurrence relation and another leading to a coupled 2-term recurrence relation.
For simplicity we assume that no breakdowns occur, neither lucky nor serious. This no-breakdown assumption implies that $\langle v_{l-1}, w_{l-1} \rangle_E \neq 0$ and $\mathcal{K}_{l-1}(A, v; \Xi) \subset \mathcal{K}_l(A, v; \Xi)$, $\mathcal{K}_{l-1}(A^H, w; \Theta) \subset \mathcal{K}_l(A^H, w; \Theta)$.

### Four term recurrence relation

A short (four term) recurrence relation is derived for the biorthonormal bases $V_l$ and $W_l$. The property $\langle x, y \rangle_E = \langle Ax, A^{-H}y \rangle_E$ and the orthogonality properties of $V_l$ and $W_l$ are the key to obtain the short recurrence relations in Theorem 8.1. The candidates chosen here are

$$x = \begin{cases} Av_{2k-1}, & \text{if } l \text{ is even} \\ A^{-1}v_{2k}, & \text{if } l \text{ is odd} \end{cases} \quad \text{and} \quad y = \begin{cases} A^{-H}w_{2k-1}, & \text{if } l \text{ is even} \\ A^Hw_{2k}, & \text{if } l \text{ is odd} \end{cases}.$$

**Theorem 8.1.** *Let $A \in \mathbb{C}^{m\times m}$ be a nonsingular matrix and $v, w \in \mathbb{C}^m$, then biorthogonal bases $V_l \in \mathbb{C}^{m\times l}$ and $W_l \in \mathbb{C}^{m\times l}$ can be constructed by four term recurrence relations.*
*For $l = 2k+2$, $k \geq 0$, with $v_{-1} = v_0$, $v_{-2} = 0$ and $\alpha_{-1,0} = 0$,*

$$\eta_{2k+1,2k} v_{2k+1} = Av_{2k-1} - \alpha_{2k-2,2k} v_{2k-2} - \alpha_{2k-1,2k} v_{2k-1} - \alpha_{2k,2k} v_{2k}, \tag{8.3}$$

$$\nu_{2k+1,2k} w_{2k+1} = A^{-H} w_{2k-1} - \beta_{2k-2,2k} w_{2k-2} - \beta_{2k-1,2k} w_{2k-1} - \beta_{2k,2k} w_{2k}, \tag{8.4}$$

*where $\alpha_{i,2k} = \langle Av_{2k-1}, w_i\rangle_E$ and $\beta_{i,2k} = \langle A^{-H} w_{2k-1}, v_i\rangle_E$.*
*For $l = 2k+3$, $\quad k \geq 0$, with $v_{-1} = 0$,*

$$\eta_{2k+2,2k+1} v_{2k+2} = A^{-1} v_{2k} - \alpha_{2k-1,2k+1} v_{2k-1} - \alpha_{2k,2k+1} v_{2k} - \alpha_{2k+1,2k+1} v_{2k+1}, \tag{8.5}$$

$$\nu_{2k+2,2k+1} w_{2k+2} = A^{H} w_{2k} - \beta_{2k-1,2k+1} w_{2k-1} - \beta_{2k,2k+1} w_{2k} - \beta_{2k+1,2k+2} w_{2k+1}, \tag{8.6}$$

*where $\alpha_{i,2k+1} = \langle A^{-1} v_{2k}, w_i\rangle_E$ and $\beta_{i,2k+1} = \langle A^{H} w_{2k}, v_i\rangle_E$. Normalization is done by choosing $\eta_{l,l-1}$ and $\nu_{l,l-1}$ such that $\langle v_{l-1}, w_{l-1}\rangle_E = 1$. This is assumed to always be possible, i.e., the assumption that breakdowns do not occur.*

*Proof.* We will prove the recurrence relations (8.3) and (8.5). The proof of (8.4) and (8.6) is analogous. Assume, without loss of generality, that $l = 2k+1$. The next basis vector $v_{2k+1}$ must be constructed such that it expands $\mathcal{K}_{2k+1}(A, v; \Xi)$ to $\mathcal{K}_{2k+2}(A, v; \Xi)$, i.e., $v_{k+1}$ must contain a component along the direction $A^{k+1}v$. And it must be orthogonal to $\mathcal{K}_{2k+1}(A^H, w; \Theta)$. Consider $v_{2k-1}$, with properties

$$\begin{aligned} v_{2k-1} &\in \operatorname{span}\{v, Av, A^{-1}v, A^2v, A^{-2}v, \ldots, A^{k-1}v, A^{-k+1}v, A^k v\} = \mathcal{K}_{2k}(A, v; \Xi) \\ &\perp \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(-k+1)H} w, A^{(k-1)H} w\} = \mathcal{K}_{2k-1}(A^H, w; \Theta). \end{aligned}$$

Multiplication with $A$ results in

$$Av_{2k-1} \in \operatorname{span}\{Av, A^2v, v, A^3v, A^{-1}v \ldots, A^k v, A^{-k+2}v, A^{k+1}v\},$$

which shows that $Av_{2k-1}$ has a component along the required direction $A^{k+1}v$. And by the no-breakdown assumption, this vector will be linearly independent of $\mathcal{K}_{2k+1}(A, v; \Xi)$.
Using $\langle x, y\rangle_E = \langle Ax, A^{-H} y\rangle_E$, we obtain

$$\begin{aligned} Av_{2k-1} &\perp A^{-H} \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(-k+1)H} w, A^{(k-1)H} w\} \\ &\perp \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(k-2)H} w, A^{(-k+1)H} w, A^{-kH} w\}. \end{aligned}$$

Hence, vector $Av_{2k-1}$ is orthogonal with respect to $\mathcal{K}_{2k-2}(A^H, w; \Theta)$. It remains to orthogonalize with respect to $w_{2k-2}, w_{2k-1}$ and $w_{2k}$ in order to satisfy the orthogonality condition $v_{2k+1} \perp \mathcal{K}_{2k+1}(A^H, w; \Theta)$. Thus, (8.3) is proven, since $\alpha_{i,2k}$ is chosen such that it eliminates the aforementioned directions from $Av_{2k-1}$. Similar reasoning can be applied to construct $v_{2k+2}$. Consider $v_{2k}$, with properties

$$\begin{aligned} v_{2k} &\in \operatorname{span}\{v, Av, A^{-1}v, \ldots, A^{k-1}v, A^{-k+1}v, A^k v, A^{-k}v\} \\ &\perp \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(-k+1)H}w, A^{(k-1)H}w, A^{-kH}w\}. \end{aligned}$$

Multiplication with $A^{-1}$ and using $\langle x, y\rangle_E = \langle A^{-1}x, A^H y\rangle_E$ provides

$$\begin{aligned} A^{-1}v_{2k} &\in \operatorname{span}\{A^{-1}v, v, A^{-2}v, \ldots, A^{k-2}v, A^{-k}v, A^{k-1}v, A^{-k-1}v\} \\ &\perp \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(-k+1)H}w, A^{(k-1)H}w, A^{-kH}w\}. \end{aligned}$$

Variables $\alpha_{i,2k+1}$ are chosen such that they eliminate the directions such that $v_{2k+2} \perp \mathcal{K}_{2k+1}(A^H, w; \Theta)$. Thus proving (8.5). □

Next, in Theorem 8.2, the matrix of recurrence coefficients is given, which has pentadiagonal structure. This result also follows from Theorem 4.5.

**Theorem 8.2.** *Consider a nonsingular matrix $A \in \mathbb{C}^{m\times m}$, $v \in \mathbb{C}^m$ and basis $V_l \in \mathbb{C}^{m\times l}$ spanning $\mathcal{K}_l(A, v; \Xi)$. The matrix of recurrence coefficients $\underline{Z}_l \in \mathbb{C}^{(l+1)\times l}$ satisfying*

$$AV_l = V_{l+1}\underline{Z}_l,$$

*has pentadiagonal structure. More precisely, for $h_{i,2k} = \langle Av_{2k}, w_i\rangle_E$ and, $\alpha_{i,2k} = \langle Av_{2k-1}, w_i\rangle_E$ and $\eta_{2k+1,2k}$ as a normalizing constant, i.e., as in Theorem 8.1, the matrix $\underline{Z}_l$ is the $(l+1)\times l$ principal leading submatrix of*

$$Z = \begin{bmatrix} h_{0,0} & \alpha_{0,2} & h_{0,2} & & & \\ h_{1,0} & \alpha_{1,2} & h_{1,2} & & & \\ & \alpha_{2,2} & h_{2,2} & \alpha_{2,4} & h_{2,4} & \\ & \eta_{3,2} & h_{3,2} & \alpha_{3,4} & h_{3,4} & \\ & & & \alpha_{4,4} & h_{4,4} & \\ & & & \eta_{5,4} & h_{5,4} & \\ & & & & & \ddots \end{bmatrix}. \tag{8.7}$$

*Proof.* Consider the recurrence relation following immediately from Equation (8.3),

$$Av_{2k-1} = \alpha_{2k-2,2k}v_{2k-2} + \alpha_{2k-1,2k}v_{2k-1} + \alpha_{2k,2k}v_{2k} + \eta_{2k+1,2k}v_{2k+1}.$$

This relation forms the even columns of $\underline{Z}_l$. To obtain a recurrence relation for $Av_{2k}$, i.e., the odd columns of $\underline{Z}_l$, look at the space in which this vector lives

$$\begin{aligned} v_{2k} &\in \operatorname{span}\{v, Av, A^{-1}v, \ldots, A^k v, A^{-k}v\} \\ Av_{2k} &\in \operatorname{span}\{Av, A^2 v, v, \ldots, A^k v, A^{-k+2}v, A^{k+1}v, A^{-k+1}v\} \\ &\subseteq \operatorname{span}\{v, Av, A^{-1}v, \ldots, A^{-k+1}v, A^k v, A^{-k}v, A^{k+1}v\} \\ &= \operatorname{span}\{v_0, \ldots, v_{2k}, v_{2k+1}\}. \end{aligned}$$

Hence, $Av_{2k} = \sum_{i=0}^{2k+1} h_{i,2k} v_i$. A short recurrence relation is obtained by looking at orthogonality properties

$$\begin{aligned} v_{2k} &\perp \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(-k+1)H}w, A^{(k-1)H}w, A^{-kH}w\} \\ Av_{2k} &\perp A^{-H}\operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(-k+1)H}w, A^{(k-1)H}w, A^{-kH}w\} \\ &\perp \operatorname{span}\{w, A^{-H}w, A^H w, \ldots, A^{(k-2)H}w, A^{(-k+1)H}w, A^{-kH}w, A^{(-k-1)H}w\} \\ &\perp \operatorname{span}\{w_0, \ldots, w_{2k-4}, w_{2k-3}\}. \end{aligned}$$

Thus the short recurrence relation, with $h_{i,2k} = \langle Av_{2k}, w_i \rangle_E$, is

$$Av_{2k} = h_{2k-2,2k} v_{2k-2} + h_{2k-1,2k} v_{2k-1} + h_{2k,2k} v_{2k} + h_{2k+1,2k} v_{2k+1}.$$

□

The four term recurrence relation contains some redundant information. This is suggested by the similarity of the coefficients $\alpha_{i,l}$ and $\beta_{i,l}$ occurring in Theorem 8.1 and verified by the low rank structure exhibited by the matrix of recurrence coefficients $\underline{Z}_l$. Example 8.1 illustrates the low rank structure of $\underline{Z}_l$.

**Example 8.1.** *Following the notation of Theorem 8.2, let $l = 7$. Then the matrix of recurrence coefficients*

$$\underline{Z}_7 = \begin{bmatrix} \times & \times & \star & & & & & \\ \times & \times & \star & & & & & \\ & \star & \times & \times & \star & & & \\ & \star & \times & \times & \star & & & \\ & & & \star & \times & \times & \star & \\ & & & \star & \times & \times & \star & \\ & & & & & \star & \times & \\ & & & & & \star & \times & \end{bmatrix},$$

*where $\times$ and $\star$ denote a generic nonzero element, exhibits some low rank structure. Namely, the pairs of nonzero elements represented as $\star$ equal their neighboring elements $\times$ multiplied with the same factor. Or in other words, every submatrix*

$$\begin{bmatrix} \times & \star \\ \times & \star \end{bmatrix} \text{ or } \begin{bmatrix} \star & \times \\ \star & \times \end{bmatrix}$$

*has rank equal to 1.*

### Two term recurrence relation

The low rank structure in the matrix of recurrence coefficients $\underline{Z}_l$ from Theorem 8.2 implies that a shorter (two term) recurrence relation can be derived for the bases $V_l$ and $W_l$. To do so, the candidates for expansion are chosen from auxiliary subspaces. These auxiliary subspaces are

$$\mathcal{K}(A, v; \Theta) = \text{span}\{v, A^{-1}v, Av, A^{-2}v, A^2v, \dots\},$$

$$\mathcal{K}(A^H, w; \Xi) = \text{span}\{w, A^H w, A^{-H}w, A^{2H}w, A^{-2H}w, \dots\}.$$

Thus we have four subspaces: $\mathcal{K}(A, v; \Xi)$, $\mathcal{K}(A, v; \Theta)$, $\mathcal{K}(A^H, w; \Theta)$ and $\mathcal{K}(A^H, w; \Xi)$. Note that Property 7.1 states that following spaces are equal $\mathcal{K}_{2k+1}(A, v; \Xi) = \mathcal{K}_{2k+1}(A, v; \Theta)$ and $\mathcal{K}_{2k+1}(A^H, w; \Theta) = \mathcal{K}_{2k+1}(A^H, w; \Xi)$.
The candidate vectors are chosen from the vectors which form biorthonormal bases for the auxiliary spaces, i.e., from $\tilde{V}_l, \tilde{W}_l$ satisfying

$$\text{span}\{\tilde{V}_i\} = \text{span}\{\tilde{v}_0, \tilde{v}_1, \dots, \tilde{v}_{i-1}\} = \mathcal{K}_l(A, v; \Theta),$$

$$\text{span}\{\tilde{W}_i\} = \text{span}\{\tilde{w}_0, \tilde{w}_1, \dots, \tilde{w}_{l-1}\} = \mathcal{K}_l(A^H, w; \Xi), \text{ for } i = 1, \dots, l-1,$$

$$\widetilde{W}_l^H \widetilde{V}_l = I.$$

The recurrence relations are given in Theorem 8.3.

**Theorem 8.3.** *Let $A \in \mathbb{C}^{m\times m}$ be a nonsingular matrix and $v, w \in \mathbb{C}^m$, $\langle v, w\rangle_E \neq 0$, then biorthonormal bases $V_l$, $W_l$, $\widetilde{V}_l$ and $\widetilde{W}_l$ can be constructed by pairs of two term recurrence relations.*
*For $l = 2k+2, \quad k \geq 0$,*

$$\eta_{2k} v_{2k+1} = A\tilde{v}_{2k} - \gamma_{2k} v_{2k}, \tag{8.8}$$

$$\tilde{\eta}_{2k} \tilde{v}_{2k+1} = A^{-1} v_{2k} - \tilde{\gamma}_{2k} \tilde{v}_{2k}, \tag{8.9}$$

$$\nu_{2k} w_{2k+1} = A^{-H} \tilde{w}_{2k} - \bar{\tilde{\gamma}}_{2k} w_{2k}, \tag{8.10}$$

$$\tilde{\nu}_{2k} \tilde{w}_{2k+1} = A^H w_{2k} - \bar{\gamma}_{2k} \tilde{w}_{2k}, \tag{8.11}$$

*where* $\gamma_{2k} = \langle A\tilde{v}_{2k}, w_{2k}\rangle_E$ *and* $\tilde{\gamma}_{2k} = \langle A^{-1}v_{2k}, \tilde{w}_{2k}\rangle_E$.
*For* $l = 2k+3, \quad k \geq 0$,

$$\eta_{2k+1}v_{2k+2} = \tilde{v}_{2k+1} - \gamma_{2k+1}v_{2k+1}, \tag{8.12}$$

$$\tilde{\eta}_{2k+1}\tilde{v}_{2k+2} = v_{2k+1} - \tilde{\gamma}_{2k+1}\tilde{v}_{2k+1}, \tag{8.13}$$

$$\nu_{2k+1}w_{2k+2} = \tilde{w}_{2k+1} - \bar{\tilde{\gamma}}_{2k+1}w_{2k+1}, \tag{8.14}$$

$$\tilde{\nu}_{2k+1}\tilde{w}_{2k+2} = w_{2k+1} - \tilde{\gamma}_{2k+1}\tilde{w}_{2k+1}, \tag{8.15}$$

*where* $\gamma_{2k+1} = \langle \tilde{v}_{2k+1}, w_{2k+1}\rangle_E$ *and* $\tilde{\gamma}_{2k+1} = \langle v_{2k+1}, \tilde{w}_{2k+1}\rangle_E$. *Normalization coefficients* $\eta_i$, $\nu_i$, $\tilde{\eta}_i$ *and* $\tilde{\nu}_i$ *are chosen such that* $\langle v_i, w_i\rangle_E = 1$ *and* $\langle \tilde{v}_i, \tilde{w}_i\rangle_E = 1$. *This is assumed to be possible under the no-breakdown assumption.*

*Proof.* The proof is given here for (8.8), (8.9), (8.12) and (8.13). For the remaining recurrence relations the proof is analogous. Assume, without loss of generality, $l = 2k+1$. The next basis vector $v_{2k+1}$ must be constructed such that it expands $\mathcal{K}_{2k+1}(A, v; \Xi)$ to $\mathcal{K}_{2k+2}(A, v; \Xi)$, i.e., introduce a component along the direction $A^{k+1}v$. And it must be orthogonal to $\mathcal{K}_{2k+1}(A^H, w; \Theta)$. Consider

$$\begin{aligned}\tilde{v}_{2k} &\in \operatorname{span}\{v, A^{-1}v, Av, \ldots, A^{-k}v, A^k v\} = \mathcal{K}_{2k+1}(A, v; \Theta)\\ &\perp \operatorname{span}\{w, A^H w, A^{-H}w, \ldots, A^{(-k+1)H}w, A^{kH}w\} = \mathcal{K}_{2k}(A^H, w; \Xi).\end{aligned}$$

Multiplication with $A$ results in

$$A\tilde{v}_{2k} \in \operatorname{span}\{Av, v, A^2 v, \ldots, A^{-k+1}v, A^{k+1}v\} \tag{8.16}$$

$$\perp A^{-H}\operatorname{span}\{w, A^H w, A^{-H}w, \ldots, A^{(-k+1)H}w, A^{kH}w\}$$

$$\perp \operatorname{span}\{w, A^H w, A^{-H}w, \ldots, A^{(k-1)H}w, A^{-kH}w\} = \mathcal{K}_{2k}(A^H, w; \Theta). \tag{8.17}$$

The required component $A^{k+1}v$ is present (8.16) and orthogonality is satisfied with respect to $\mathcal{K}_{2k}(A^H, w; \Theta)$ (8.17). Note that $A\tilde{v}_{2k}$ is orthogonal to a larger subspace of $\mathcal{K}_{2k+1}(A^H, w; \Xi)$ than $Av_{2k-1}$, this is the key observation to explain the shorter recurrence relation. Orthogonalization only remains to be done to eliminate components along $w_{2k}$, thus obtaining (8.8). The same derivation can be done for $\tilde{v}_{2k+1}$, which must expand $\mathcal{K}_{2k+1}(A, v; \Theta)$ to $\mathcal{K}_{2k+2}(A, v; \Theta)$ and must be orthogonal to $\mathcal{K}_{2k+1}(A^H, w; \Xi)$. Thereby proving (8.9). For $v_{2k+2}$, consider $\tilde{v}_{2k+1}$

$$\tilde{v}_{2k+1} \in \operatorname{span}\{v, A^{-1}v, Av, \ldots, A^{-k}v, A^k v, A^{-k-1}v\} \tag{8.18}$$

$$\perp \operatorname{span}\{w, A^H w, A^{-H}w, \ldots, A^{kH}w, A^{-kH}w\} = \mathcal{K}_{2k+1}(A^H, w; \Theta). \tag{8.19}$$

The component along $A^{-k-1}v$ is present in $\tilde{v}_{2k+1}$ (8.18). So it only remains to enforce the orthogonality conditions, orthogonalize along $w_{2k+1}$ to obtain (8.12). Similarly for $\tilde{v}_{2k+2}$, to obtain (8.13). □

From these recurrence relations, a matrix pencil representation of the matrix of recurrence coefficients $\underline{Z}_l$ from Theorem 8.2 can be derived. This result is given in Theorem 8.4. This representation reveals that $\underline{Z}_l$ can be represented by a product of essentially $2 \times 2$ matrices. This allows for an efficient way to store and manipulate this matrix on a computer.

**Theorem 8.4.** *Consider a nonsingular matrix $A \in \mathbb{C}^{m\times m}$, $v, w \in \mathbb{C}^m$ and basis $V_l \in \mathbb{C}^{m\times l}$ spanning $\mathcal{K}_l(A, v; \Xi)$. The recurrence pencil $(\underline{T}_l, \underline{S}_l) \in \mathbb{C}^{(l+1)\times l} \times \mathbb{C}^{(l+1)\times l}$ satisfying*

$$AV_{l+1}\underline{S}_l = V_{l+1}\underline{T}_l,$$

*can be represented by a sparse tridiagonal pencil. More precisely, for the same coefficients as used in Theorem 8.3, $\underline{T}_l, \underline{S}_l$ are, respectively, the $(l+1) \times l$ principal leading submatrices of*

$$T = \begin{bmatrix} \gamma_0 & \tilde{\eta}_0^{-1}(1-\gamma_0\tilde{\gamma}_0) & & & & & \\ \eta_0 & -\tilde{\eta}_0^{-1}\eta_0\tilde{\gamma}_0 & & & & & \\ & & \gamma_2 & \tilde{\eta}_2^{-1}(1-\gamma_2\tilde{\gamma}_2) & & & \\ & & \eta_2 & -\tilde{\eta}_2^{-1}\eta_2\tilde{\gamma}_2 & & & \\ & & & & \gamma_4 & & \\ & & & & \eta_4 & & \\ & & & & & \ddots & \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & & & & & & \\ & \gamma_1 & \tilde{\eta}_1^{-1}(1-\gamma_1\tilde{\gamma}_1) & & & & \\ & \eta_1 & -\tilde{\eta}_1^{-1}\eta_1\tilde{\gamma}_1 & & & & \\ & & & \gamma_3 & \tilde{\eta}_3^{-1}(1-\gamma_3\tilde{\gamma}_3) & & \\ & & & \eta_3 & -\tilde{\eta}_3^{-1}\eta_3\tilde{\gamma}_3 & & \\ & & & & & \ddots & \end{bmatrix}.$$

*Proof.* Rewrite the pairs of recurrence relations (8.8), (8.9), (8.12) and (8.13) in matrix notation to obtain

$$A\begin{bmatrix}\tilde{v}_{2k} & \tilde{v}_{2k+1}\end{bmatrix} = \begin{bmatrix}v_{2k} & v_{2k+1}\end{bmatrix}\begin{bmatrix}\gamma_{2k} & \tilde{\eta}_{2k}^{-1}(1-\gamma_{2k}\tilde{\gamma}_{2k}) \\ \eta_{2k} & -\tilde{\eta}_{2k}^{-1}\eta_{2k}\tilde{\gamma}_{2k}\end{bmatrix}, \tag{8.20}$$

$$\begin{bmatrix}\tilde{v}_{2k+1} & \tilde{v}_{2k+2}\end{bmatrix} = \begin{bmatrix}v_{2k+1} & v_{2k+2}\end{bmatrix}\begin{bmatrix}\gamma_{2k+1} & \tilde{\eta}_{2k+1}^{-1}(1-\gamma_{2k+1}\tilde{\gamma}_{2k+1}) \\ \eta_{2k+1} & -\tilde{\eta}_{2k+1}^{-1}\eta_{2k+1}\tilde{\gamma}_{2k+1}\end{bmatrix}. \tag{8.21}$$

The proof consists of substituting (8.21) into (8.20). Substitution is done as follows, consider

$$A\begin{bmatrix}\tilde{v}_{2k} & \tilde{v}_{2k+1} & \tilde{v}_{2k+2} & \tilde{v}_{2k+3}\end{bmatrix} = \begin{bmatrix}v_{2k} & v_{2k+1} & v_{2k+2} & v_{2k+3}\end{bmatrix}\begin{bmatrix}D_{2k} & \\ & D_{2k+2}\end{bmatrix}, \tag{8.22}$$

where $D_i = \begin{bmatrix}\gamma_i & \tilde{\eta}_i^{-1}(1-\gamma_i\tilde{\gamma}_i)\\ \eta_i & -\tilde{\eta}_i^{-1}\eta_i\tilde{\gamma}_i\end{bmatrix}$. Then (8.22) can be used to obtain

$$A\begin{bmatrix}\tilde{v}_{2k} & v_{2k+1} & v_{2k+2} & \tilde{v}_{2k+3}\end{bmatrix}\begin{bmatrix}1 & & & \\ & D_{2k+1} & & \\ & & & 1\end{bmatrix}$$
$$= \begin{bmatrix}v_{2k} & v_{2k+1} & v_{2k+2} & v_{2k+3}\end{bmatrix}\begin{bmatrix}D_{2k} & \\ & D_{2k+2}\end{bmatrix}.$$

Repeating this procedure for $v_i$, $\tilde{v}_i$, $i = 0, 1, \ldots, l$ proves the statement. □

Note that the structure of the recurrence pencil appearing in Theorem 8.4 follows from Theorem 4.6, Lemma 4.5 and Lemma 4.6.

### 8.2.2 Levinson procedure

The two term recurrence relation can also be derived starting from the Gram matrix arising from the subspaces $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^H, w; \Theta)$. The derivation here will follow a Levinson procedure. Such procedures make use of the isomorphism between the vector space of $(i+1)$-tuples and of polynomials of degree $i$. In the case studied here, the connection between vectors $v_i \in \mathcal{K}_i(A, v; \Xi)$ and Laurent polynomials $a_i \in \mathcal{R}_i^{\Xi}(z)$ is used. Similar to the definition of $\mathcal{K}_l(A, v; \Xi)$ and $\mathcal{K}_l(A^H, w; \Theta)$, we define, for a nonnegative integer $k$:
For even $l = 2k$,

$$\mathcal{R}_l^{\Xi}(z) = \operatorname{span}\{1, z, z^{-1}, z^2, z^{-2}, \ldots, z^k, z^{-k}\},$$

$$\mathcal{R}_l^{\Theta}(z) = \operatorname{span}\{1, z^{-1}, z, z^{-2}, z^2, \ldots, z^{-k}, z^k\}.$$

For odd $l = 2k+1$,

$$\mathcal{R}_l^{\Xi}(z) = \operatorname{span}\{1, z, z^{-1}, z^2, z^{-2}, \ldots, z^k, z^{-k}, z^{k+1}\},$$

$$\mathcal{R}_l^{\Theta}(z) = \operatorname{span}\{1, z^{-1}, z, z^{-2}, z^2, \ldots, z^{-k}, z^k, z^{-k-1}\}.$$

A vector $v_i \in \mathcal{K}_{i+1}(A, v; \Xi)$ can be written in terms of a corresponding Laurent polynomial $a_i(z) \in \mathcal{R}_i^{\Xi}(z)$, i.e., $v_i = a_i(A)v$. Similarly, $w_i = b_i(A^H)w$, with $w_i \in \mathcal{K}(A^H, w; \Theta)$ and $b_i(z) \in \mathcal{R}_i^{\Theta}(z)$. By Lemma 7.2 we know that a linear functional $\mathcal{L}_g\{.\}$ exists for which $\{a_i(z), b_i(z)\}_i$ forms a biorthonormal pair of sequences, i.e., $\mathcal{L}_g\{a_i(z)b_j(z)\} = \delta_{ij}$.
The bases for the spaces $\mathcal{R}^{\Xi}$ and $\mathcal{R}^{\Theta}$ are, respectively,

$$B^{\Xi} = \begin{bmatrix} 1 & z & z^{-1} & z^2 & z^{-2} & \dots \end{bmatrix},$$

$$B^{\Theta} = \begin{bmatrix} 1 & z^{-1} & z & z^{-2} & z^2 & \dots \end{bmatrix}.$$

Consider the Gram matrix $M$,

$$M = \begin{bmatrix} m_0 & m_1 & m_{-1} & m_2 & m_{-2} & \\ m_{-1} & m_0 & m_{-2} & m_1 & m_{-3} & \\ m_1 & m_2 & m_0 & m_3 & m_{-1} & \dots \\ m_{-2} & m_{-1} & m_{-3} & m_0 & m_{-4} & \\ m_2 & m_3 & m_1 & m_4 & m_0 & \\ & & \vdots & & & \ddots \end{bmatrix}, \tag{8.23}$$

with $m_i := \mathcal{L}_g\{z^i\} = w^H A^i v = \langle A^i v, w \rangle_E$. Denote by $M_l$ the $(l+1) \times (l+1)$ principal leading submatrix of $M$. The sequence of pairs $\{a_i(z), b_i(z)\}_{i=0}^{l}$ is the biorthonormal Laurent polynomial sequence following from the factorization of $M_l$. That is, let $a_i, b_i \in \mathbb{C}^{i+1}$ denote the coefficients of $a_i(z)$, $b_i(z)$ in the basis $B_{i+1}^{\Xi}, B_{i+1}^{\Theta} \in \mathbb{C}^{m \times (i+1)}$, respectively. Then $M_i a_i = \begin{bmatrix} 0 & \dots & 0 & \times \end{bmatrix}^{\top}$ and $b_i^{\top} M_i = \begin{bmatrix} 0 & \dots & 0 & \times \end{bmatrix}$. These are matrix representations of the orthogonality conditions [51, p.44] imposed on the biorthonormal Laurent polynomials.
The goal is to obtain short recurrence relations to construct the sequence $\{a_i(z), b_i(z)\}_i$ by analyzing the displacement structure of $M$. The coupled short recurrence relation is given in Theorem 8.5. First we sketch the main ideas and introduce the tools required to prove this theorem.
As is illustrated above, a good choice of candidate to expand the biorthonormal sequence is paramount to obtaining short recurrence relations. To this end, an auxiliary sequence of biorthonormal Laurent polynomials $\{\tilde{a}_i(z), \tilde{b}_i(z)\}_{i=0}^{l}$ is introduced, which follows from the factorization of the Gram matrix $\tilde{M}_l = M_l^{\top}$. Let $\tilde{a}_i, \tilde{b}_i \in \mathbb{C}^{i+1}$ denote their coefficients in the bases $B_{i+1}^{\Theta}, B_{i+1}^{\Xi} \in \mathbb{C}^{m \times (i+1)}$, respectively. Then $\tilde{M}_i \tilde{a}_i = \begin{bmatrix} 0 & \dots & 0 & \times \end{bmatrix}^{\top}$ and $\tilde{b}_i^{\top} \tilde{M}_i = \begin{bmatrix} 0 & \dots & 0 & \times \end{bmatrix}$.
Furthermore, a matrix $S_{2k+1} \in \mathbb{C}^{(2k+2) \times (2k+2)}$ allowing to switch between bases $B_{2k+1}^{\Xi}$

and $B^{\Theta}_{2k+1}$ for $\mathcal{R}^{\Xi}_{2k+1} = \mathcal{R}^{\Theta}_{2k+1}$ is the permutation matrix

$$S_{2k+2} := \begin{bmatrix} 1 & & & & & & & \\ & 0 & 1 & & & & & \\ & 1 & 0 & & & & & \\ & & & 0 & 1 & & & \\ & & & 1 & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & 1 \\ & & & & & & 1 & 0 \end{bmatrix}.$$

This matrix satisfies $B^{\Xi}_{2k+1} = B^{\Theta}_{2k+1} S_{2k+1}$ and $S^2_{2k+1} = I$.
Matrices which represent multiplication with $z$ in $B^{\Xi}_{2k+1}$ and with $z^{-1}$ in $B^{\Theta}_{2k+1}$ are the last required components for the proof. These representations can be obtained by using the techniques from Chapter 5. Consider the decomposition (5.3) of $B^{\Xi}_{2k+1}$, with the Vandermonde matrix $V_i := \begin{bmatrix} 1 & z & \dots & z^{i-1} \end{bmatrix}$,

$$B^{\Xi}_{2k+1} = z^{-k} V_{2k+1} \left( P^{\Xi}_{2k+1} \right)^{\top}.$$

The permutation matrix $P^{\Xi}_{2k+1}$ can be generated by Algorithm 5 and is, with canonical unit vectors $e_i \in \mathbb{C}^{2k+1}$,

$$P^{\Xi}_{2k+1} = \begin{bmatrix} e_{2k+1} & \dots & e_5 & e_3 & e_1 & e_2 & e_4 & \dots & e_{2k} \end{bmatrix}.$$

From this decomposition and $zV_i = V_i Z_i + z^i e_i^{\top}$, with $Z_i \in \mathbb{C}^{i \times i}$ the leftshift matrix, the representation for the multiplication with $z$ can be obtained

$$\begin{aligned} zB^{\Xi}_{2k+1} &= z^{-k} z V_{2k+1} \left( P^{\Xi}_{2k+1} \right)^{\top} \\ &= z^{-k} V_{2k+1} Z_{k+1} \left( P^{\Xi}_{2k+1} \right)^{\top} + z^{-k} z^{2k+1} e^{\top}_{2k+1} \left( P^{\Xi}_{2k+1} \right)^{\top} \\ &= \underbrace{z^{-k} V_{2k+1} \left( P^{\Xi}_{2k+1} \right)^{\top}}_{=: B^{\Xi}_{2k+1}} \underbrace{P^{\Xi}_{2k+1} Z_{2k+1} \left( P^{\Xi}_{2k+1} \right)^{\top}}_{=: Z^{\Xi}_{2k+1}} + z^{k+1} e^{\top}_{2k+1} \\ &= B^{\Xi}_{2k+1} Z^{\Xi}_{2k+1} + z^{k+1} e^{\top}_{2k+1}. \end{aligned}$$

The matrix $Z^{\Xi}_{2k+1}$ is the $(2k+1) \times (2k+1)$ leading principal submatrix of

$$Z = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \\ 1 & 0 & & 0 & 0 & 0 & \\ & 0 & & 0 & 1 & 0 & \\ & 1 & & 0 & & 0 & \\ & & & 0 & & 0 & \\ & & & 1 & & 0 & \\ & & & & & & \ddots \end{bmatrix}.$$

For $z^{-1}B^{\Theta}_{2k+1}$ a similar derivation results in $z^{-1}B^{\Theta}_{2k+1} = B^{\Theta}_{2k+1}Z^{\Theta}_{2k+1} + z^{-2k-1}e^{\top}_{2k+1}$, with $Z^{\Theta}_{2k+1} = \left(Z^{\Xi}_{2k+1}\right)^{\top}$.

**Theorem 8.5.** *Consider the Laurent polynomials $a_i(z), \tilde{b}_i(z) \in \mathcal{R}^{\Xi}_i(z)$ and $\tilde{a}_i(z), b_i(z) \in \mathcal{R}^{\Theta}_i(z)$. If they are constructed via*

$$\begin{aligned}
\eta_{2k}a_{2k+1}(z) &= z\tilde{a}_{2k}(z) - \gamma_{2k}a_{2k}(z),\\
\tilde{\eta}_{2k}\tilde{a}_{2k+1}(z) &= z^{-1}a_{2k}(z) - \tilde{\gamma}_{2k}\tilde{a}_{2k}(z),\\
\nu_{2k}b_{2k+1}(z) &= z^{-1}\tilde{b}_{2k}(z) - \bar{\tilde{\gamma}}_{2k}b_{2k}(z),\\
\tilde{\nu}_{2k}\tilde{b}_{2k+1}(z) &= zb_{2k}(z) - \bar{\gamma}_{2k}\tilde{b}_{2k}(z),\\
\eta_{2k+1}a_{2k+2}(z) &= \tilde{a}_{2k+1}(z) - \gamma_{2k+1}a_{2k+1}(z),\\
\tilde{\eta}_{2k+1}\tilde{a}_{2k+2}(z) &= a_{2k+1}(z) - \tilde{\gamma}_{2k+1}\tilde{a}_{2k+1}(z),\\
\nu_{2k+1}b_{2k+2}(z) &= \tilde{b}_{2k+1}(z) - \bar{\gamma}_{2k+1}b_{2k+1}(z),\\
\tilde{\nu}_{2k+1}\tilde{b}_{2k+2}(z) &= b_{2k+1}(z) - \bar{\tilde{\gamma}}_{2k+1}\tilde{b}_{2k+1}(z),
\end{aligned}$$

*with coefficients $\gamma_{2k} = \{z\tilde{a}_{2k}(z), b_{2k}(z)\}$, $\tilde{\gamma}_{2k} = \mathcal{L}_g\{z^{-1}a_{2k}(z), \tilde{b}_{2k}(z)\}$, $\gamma_{2k+1} = \mathcal{L}_g\{\tilde{a}_{2k+1}(z), b_{2k+1}(z)\}$ and $\tilde{\gamma}_{2k+1} = \mathcal{L}_g\{a_{2k+1}(z), \tilde{b}_{2k+1}(z)\}$, then the sequences $\{a_i(z)\}_i$ and $\{b_i(z)\}_i$ form biorthonormal Laurent polynomials. That is, if $\eta_i, \nu_i$ are chosen such that $\mathcal{L}_g\{a_i(z), b_i(z)\} = 1$, then $\mathcal{L}_g\{a_i(z), b_j(z)\} = \delta_{ij}$, biorthogonal with respect to the linear functional from Lemma 7.2. Similarly for $\{\tilde{a}_i(z)\}_i$ and $\{\tilde{b}_i(z)\}_i$.*

*Proof.* The proof is given by induction and follows a Levinson procedure. Consider the Gram matrix $M_0 := \begin{bmatrix} m_0 \end{bmatrix}$, then $a_0 := m_0^{-1/2}$ and $b_0 := \overline{m}_0^{-1/2}$ satisfy the biorthonormality condition $\mathcal{L}_g\{a_0, b_0\} = 1$. Assume that we possess $a_{2k+1}(z), \tilde{a}_{2k+1}(z)$ with coefficients $a_n, \tilde{a}_n \in \mathbb{C}^{n+1}$ with respect to bases $B^{\Xi}_{2k+1}$ and $B^{\Theta}_{2k+1}$, respectively, satisfying

$$M_{2k+1}a_{2k+1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \times \end{bmatrix} \quad \text{and} \quad \tilde{M}_{2k+1}\tilde{a}_{2k+1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \times \end{bmatrix}.$$

From the coefficients of $a_{2k+1}(z)$ and $\tilde{a}_{2k+1}(z)$ we want to compute the coefficients for $a_{2k+2}(z)$ and $\tilde{a}_{2k+2}(z)$. These can be obtained from the linear combinations of the coefficients of $a_{2k+1}(z), \tilde{a}_{2k+1}(z), z^{-1}a_{2k+1}(z)$ and $z\tilde{a}_{2k+2}(z)$, which follows from the

derivation below. Let $a_{2k+2} \in \mathbb{C}^{2k+2}$ represent the coefficients of $a_{2k+2}(z)$ in the basis $B^{\Xi}_{2k+2}$ for $\mathcal{R}^{\Xi}_{2k+1}$, then it must satisfy

$$M_{2k+2}a_{2k+2} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \times \end{bmatrix}.$$

Candidates for the expansion from $\mathcal{R}^{\Xi}_{2k}$ to $\mathcal{R}^{\Xi}_{2k+1}$ are $za_{2k+1}(z)$ and $z\tilde{a}_{2k+1}(z)$. From the displacement structure of $M_{2k+2}$ it follows that these result in

$$M_{2k+2}Z^{\Xi}_{k+2}\begin{bmatrix} a_{2k+1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \times \\ 0 \\ \times \\ 0 \end{bmatrix} \quad \text{and} \quad M_{2k+2}Z^{\Xi}_{k+2}\begin{bmatrix} S_{2k+1}\tilde{a}_{2k+1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \tilde{\alpha}_n \\ \tilde{\tau}_n \end{bmatrix}.$$

Since the embedding of $a_{2k+1}(z)$ in $\mathcal{R}^{\Xi}_{2k+2}$ leads to

$$M_{2k+2}\begin{bmatrix} a_{2k+1} \\ 0 \end{bmatrix} = \begin{bmatrix} & & & m_{k+1} \\ & M_{2k+2} & & m_k \\ & & & \vdots \\ m_{-k-1} & m_{-k} & \dots & m_0 \end{bmatrix}\begin{bmatrix} a_{2k+1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \tau_n \\ \alpha_n \end{bmatrix},$$

the candidate leading to the shortest recurrence relation is $z\tilde{a}_{2k+1}(z)$. It is easy to verify that $\eta_{2k+1}a_{2k+2} = Z^{\Xi}_{2k+2}\begin{bmatrix} S_{2k+1}\tilde{a}_{2k+1} & 0 \end{bmatrix}^{\top} - \gamma_{2k+1}\begin{bmatrix} a_{2k+1} & 0 \end{bmatrix}^{\top}$, with $\gamma_{2k+1} = \frac{\tilde{\alpha}_{2k+1}}{\tau_{2k+1}}$, will satisfy the orthogonality condition. Similarly for $\tilde{a}_{n+1}(z)$, the recurrence relation is $\tilde{\eta}_{2k+1}\tilde{a}_{2k+2}(z) = z^{-1}a_{2k+1}(z) - \tilde{\gamma}_{2k+1}\tilde{a}_{2k+1}(z)$, with $\tilde{\gamma}_{2k+1} = \frac{\alpha_{2k+1}}{\tilde{\tau}_{2k+1}}$. Hence, the recurrence relations for odd indices has been shown.

Now, we consider even indices, i.e., $2k+2$. The spaces are not the same $\mathcal{R}^{\Xi}_{2k+2} \neq \mathcal{R}^{\Theta}_{2k+2}$. Moreover, the term needed to expand $\mathcal{R}^{\Xi}_{2k+2}$ to $\mathcal{R}^{\Xi}_{2k+3}$ appears in the space $\mathcal{R}^{\Theta}_{2k+2}$, that is, $z_{-k-1} \in \mathcal{R}^{\Theta}_{2k+2}$. The biorthonormal Laurent polynomial can thus be used to expand $\mathcal{R}^{\Xi}_{2k+2}$ and leads to a short recurrence relation since

$$M_{2k+3}\begin{bmatrix} a_{2k+2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \tau_{2k+2} \\ \alpha_{2k+2} \end{bmatrix} \quad \text{and} \quad M_{2k+3}S_{2k+3}\begin{bmatrix} \tilde{a}_{2k+2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \tilde{\alpha}_{2k+2} \\ \tilde{\tau}_{2k+2} \end{bmatrix}.$$

Note that the use of the matrix $S_{2k+3}$ is appropriate, since the embedding $\begin{bmatrix}\tilde{a}_{2k+2} & 0\end{bmatrix}$ of $\tilde{a}_{2k+2}(z)$ into $\mathcal{R}^{\Theta}_{2k+3}$ is an element of $\mathcal{R}^{\Theta}_{2k+3} = \mathcal{R}^{\Xi}_{2k+3}$. Then, $\eta_{2k+2}a_{2k+3} = S_{2k+3}\begin{bmatrix}\tilde{a}_{2k+3} & 0\end{bmatrix}^{\top} - \gamma_{2k+3}\begin{bmatrix}a_{2k+3} & 0\end{bmatrix}^{\top}$, $\gamma_{2k+3} = \frac{\tilde{\alpha}_{2k+3}}{\tau_{2k+3}}$ construct the next biorthonormal Laurent polynomial $a_{2k+3}(z)$. Similarly $\tilde{a}_{2k+3}(z)$ can be shown to satisfy $\tilde{\eta}_{2k+2}a_{2k+3}(z) = \tilde{a}_{2k+2}(z) - \tilde{\gamma}_{2k+2}a_{2k+2}(z)$, $\tilde{\gamma}_{2k+2} = \frac{\alpha_{2k+2}}{\tilde{\tau}_{2k+2}}$.
The proof of the recurrence relations for $b_i(z), \tilde{b}_i(z)$ can be done in a similar way. □

### 8.2.3 Connection to the CMV decomposition

To retrieve the results reported by Watkins [179], it suffices to consider the same starting vector for all spaces $v = w$ and a unitary matrix $U$. The following corollaries summarize the results for the CMV decomposition which are given above for general matrices.
For the CMV decomposition, the two pairs of four term recurrence relations from Theorem 8.1 collapse into one pair of recurrence relations. The corresponding recurrence matrix is unitary. These results are given in Corollary 8.1 and 8.2, respectively.

**Corollary 8.1.** *Consider a unitary matrix $U \in \mathbb{C}^{m\times m}$ and $h \in \mathbb{C}^m$. Then, with normalization of $\eta_i$ such that $\langle v_i, v_i\rangle_E = 1$,*

$$\eta_{2k+1,2k}v_{2k+1} = Uv_{2k-1} - \alpha_{2k-2,2k}v_{2k-2} - \alpha_{2k-1,2k}v_{2k-1} - \alpha_{2k,2k}v_{2k},$$

$$\eta_{2k+2,2k+1}v_{2k+2} = U^H v_{2k} - \alpha_{2k-1,2k+1}v_{2k-1} - \alpha_{2k,2k+1}v_{2k} - \alpha_{2k+1,2k+1}v_{2k+1}.$$

**Corollary 8.2.** *Consider a unitary matrix $U \in \mathbb{C}^{m\times m}$, $h \in \mathbb{C}^m$ and an orthogonal basis $V_l \in \mathbb{C}^{m\times l}$ for $\mathcal{K}_l(U, v; \Xi)$. Then the orthogonal projection of $U$ onto $\mathcal{K}_l(U, v; \Xi)$,*

$$V_l^H U V_l = Z_l, \tag{8.24}$$

*is a unitary matrix $Z_l \in \mathbb{C}^{l\times l}$ with pentadiagonal structure.*

For the two term recurrence relations, given in Theorem 8.3, the four pairs of recurrence relations collapse into two pairs. The coefficients also simplify, since $\tilde{\gamma} = \bar{\gamma}$. The result is stated in Corollary 8.3. The related recurrence pencil consists of two unitary tridiagonal matrices, Corollary 8.4 states this result.

**Corollary 8.3.** *Consider a unitary matrix $U \in \mathbb{C}^{m\times m}$ and $h \in \mathbb{C}^m$. Then, with normalization of $\eta_i$ and $\tilde{\eta}_i$ such that $\langle v_i, v_i\rangle_E = 1$ and $\langle \tilde{v}_i, \tilde{v}_i\rangle_E = 1$, the recurrence relations* (8.8) - (8.15) *become*

$$\eta_{2k}v_{2k+1} = U\tilde{v}_{2k} - \gamma_{2k}v_{2k},$$

$$\tilde{\eta}_{2k}\tilde{v}_{2k+1} = U^H v_{2k} - \tilde{\gamma}_{2k}\tilde{v}_{2k},$$

*and*

$$\eta_{2k+1} v_{2k+2} = \tilde{v}_{2k+1} - \gamma_{2k+1} v_{2k+1},$$

$$\tilde{\eta}_{2k+1} \tilde{v}_{2k+2} = v_{2k+1} - \tilde{\gamma}_{2k+1} \tilde{v}_{2k+1}.$$

*The vectors $v_i, \tilde{v}_i$, $i = 0, 1, \ldots, l$, form orthogonal bases for $\mathcal{K}_l(U, h)$ and $\mathcal{S}_l(U, h)$, respectively.*

**Corollary 8.4.** *Consider a unitary matrix $U \in \mathbb{C}^{m\times m}$, a vector $h \in \mathbb{C}^m$ and an orthogonal basis $V_l \in \mathbb{C}^{m\times(l+1)}$ for $\mathcal{K}_l(U, v; \Xi)$. The matrix pencil of recurrence coefficients $(\underline{T}_l, \underline{S}_l)$, with $\underline{T}_l, \underline{S}_l \in \mathbb{C}^{(l+1)\times l}$ satisfying*

$$AV_{l+1}\underline{S}_l = V_{l+1}\underline{T}_l,$$

*can be represented by two unitary tridiagonal matrices with the same sparsity structure as in Theorem 8.4.*

The recurrence relations for the Laurent polynomials simplify in a similar manner, stated in Corollary 8.5.

**Corollary 8.5.** *Consider Laurent polynomials $a_i(z) \in \mathcal{R}_i^{\Xi}(z)$ and $\tilde{a}_i(z) \in \mathcal{R}_i^{\Theta}(z)$. If these are constructed as*

$$\eta_{2k} a_{2k+1}(z) = z\tilde{a}_{2k}(z) - \gamma_{2k} a_{2k}(z),$$

$$\tilde{\eta}_{2k} \tilde{a}_{2k+1}(z) = z^{-1} a_{2k}(z) - \tilde{\gamma}_{2k} \tilde{a}_{2k}(z),$$

$$\eta_{2k+1} a_{2k+2}(z) = \tilde{a}_{2k+1}(z) - \gamma_{2k+1} a_{2k+1}(z),$$

$$\tilde{\eta}_{2k+1} \tilde{a}_{2k+2}(z) = a_{2k+1}(z) - \tilde{\gamma}_{2k+1} \tilde{a}_{2k+1}(z),$$

*where $\gamma_{2k} = \mathcal{L}_g\{z\tilde{a}_{2k}(z), a_{2k}(z)\}$, $\gamma_{2k+1} = \mathcal{L}_g\{\tilde{a}_{2k+1}(z), a_{2k+1}(z)\}$ , $\tilde{\gamma}_i = \bar{\gamma}_i$ and $\eta_i$, $\tilde{\eta}_i$ are chosen such that $\langle v_i, v_i\rangle_E = 1$, $\langle \tilde{v}_i, \tilde{v}_i\rangle_E = 1$, respectively. Then the sequence $\{a_i(z)\}_i$ is a sequence of orthogonal Laurent polynomials. That is, $\langle a_i(z), a_j(z)\rangle_g = \delta_{ij}$, with respect to the inner product from Lemma 7.1. Similarly for the sequence $\{\tilde{a}_i(z)\}_i$.*

### 8.2.4 Numerical behavior

The numerical behavior of the proposed four and two term recurrence relations from Theorem 8.1 and Theorem 8.3 is analyzed. The matrices $A \in \mathbb{C}^{m\times m}$ used are restricted to well-conditioned normal matrices. Testing the recurrence relations with these matrices will shed some light on the stability of the methods. We are interested in the biorthogonality of the generated bases and the quality of the recurrence matrix or pencil. The generated bases $V_k, W_k \in \mathbb{C}^{m\times k}$ span the Krylov subspaces $\mathcal{K}_k(A, v; \Xi)$ and $\mathcal{K}_k(A^H, v; \Theta)$, respectively, where $\Xi = \{\infty, 0, \infty, 0, \dots\}$ and $\Theta = \{0, \infty, 0, \infty, \dots\}$. The biorthonormality of these bases is quantified by the *biorthonormality error*

$$\|W_k^H V_k - I\|_2,$$

The *recursion errors* are, for the recurrence matrix from Theorem 8.2,

$$\frac{\|AV_k - V_{k+1}\underline{Z}_{k+1}\|_2}{\|A\|_2}$$

and, for the recurrence pencil from Theorem 8.4,

$$\frac{\|AV_kS_k - V_kT_k\|_2}{\|A\|_2}.$$

Two other metrics are useful in order to interpret the results. These are the condition number of the matrix or pencil and the *growth factor*. The growth factor is defined, for the four and two term recurrence relation respectively, as

$$\rho := \frac{\max\left(|Z_k|\right)}{\max\left(|A|\right)} \text{ and } \rho := \frac{\max\left(\max(|T_k|, |S_k|)\right)}{\max\left(|A|\right)}.$$

The growth factor provides a metric for stability. Throughout this section a modest size of matrices is chosen, a dimension of $m = 200$. This can be done without loss of generality, larger matrices (with the same properties) do inherently cause larger numerical errors, however the numerical behaviour remains similar.

#### Unitary matrices

Consider the case corresponding to the CMV decomposition, $U \in \mathbb{C}^{m\times m}$ is a unitary matrix and the starting vectors are equal $v = w \in \mathbb{C}^m$. Let $m = 200$ and $v$ be a random vector. Figure 8.5 shows the biorthonormality and recursion error, both errors are low. When the dimension of the subspaces $k$ approaches the dimension of the matrix $m$, biorthonormality is lost more rapidly since the Ritz values (i.e., eigenvalues of the recurrence matrix or pencil) start to accurately approximate the eigenvalues of $U$. Furthermore, the matrices, $Z_k, S_k$ and $T_k$ are close to unitary. Hence, for a unitary matrix and $v = w$, the CMV case is retrieved.

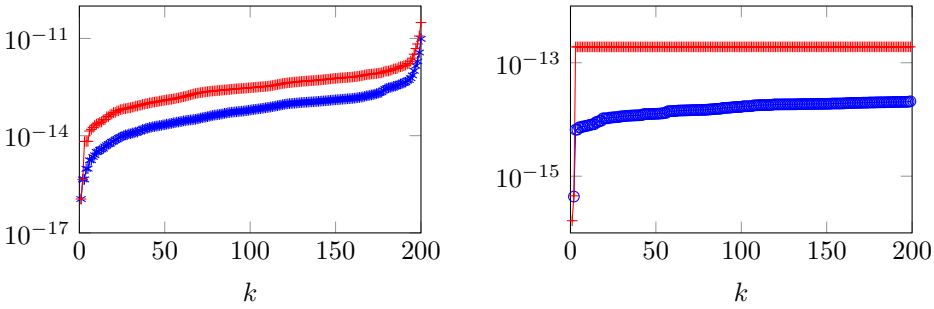



Figure 8.5: Biorthogonality and recursion error, respectively left and right, for a unitary matrix $U \in \mathbb{C}^{200\times 200}$ and $v = w$. Two term recurrence relation in blue and four term recurrence relation in red.

### Scaled and shifted unitary matrices

Consider scaled and shifted unitary matrices, i.e., for a unitary matrix $U \in \mathbb{C}^{m\times m}$, starting vectors $v = w$ and scalars $\alpha, \lambda \in \mathbb{R}$, $A = (\alpha U + \lambda I) \in \mathbb{C}^{m\times m}$. Figure 8.6 shows the biorthonormality and recursion error for $\alpha = 1.1$ and $\lambda = 0$. Both errors are still small, however compared to the unitary case, shown in Figure 8.5, they are larger. The recursion error increases steadily from $k = 55$ and $k = 120$, for the two and four term recurrence relations, respectively. This can be (at least partially) explained by the increasing growth factor $\rho$, which is approximately $\rho(k) = \alpha^k$. Figure 8.7 shows the errors for the choice $\alpha = 1$ and $\lambda = 0.1$. The two term recurrence relation performs better than the four term recurrence relation, especially in terms of the recursion error.

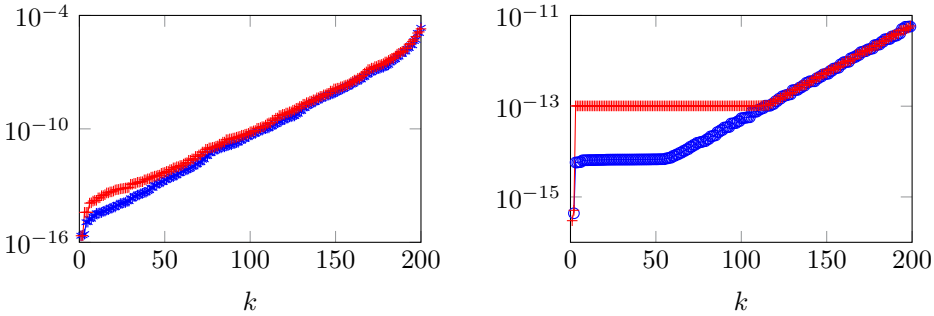



Figure 8.6: Biorthonormality and recursion error, respectively left and right, for a scaled unitary matrix $\alpha U \in \mathbb{C}^{200\times 200}$, with $\alpha = 1.1$ and $v = w$. Two term recurrence relation in blue and four term recurrence relation in red.

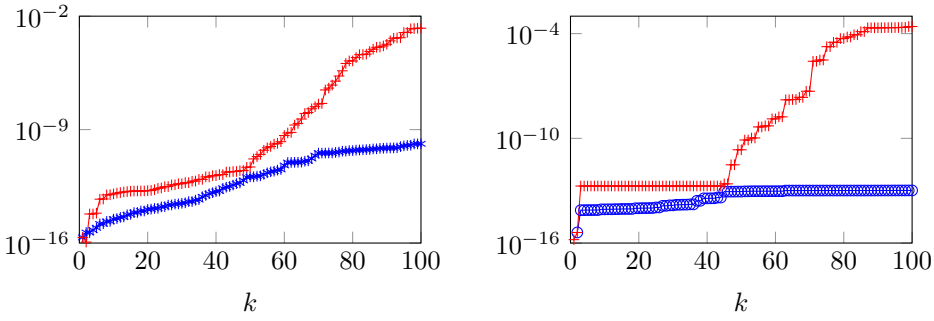



Figure 8.7: Biorthonormality and recursion error, respectively left and right, for a shifted unitary matrix $(U + \lambda I) \in \mathbb{C}^{200\times 200}$, with $\lambda = 0.1$ and $v = w$. Two term recurrence relation in blue and four term recurrence relation in red.

### General normal matrices

Consider a normal matrix $A \in \mathbb{C}^{m\times m}$ with condition number $\kappa(A) = 10$ and starting vectors $v, w \in \mathbb{C}^m$, not necessarily equal. The errors are shown in Figure 8.8. The two and four term recurrence relations have similar behavior. Biorthonormality is lost quickly and there is no Ritz value which approximates an eigenvalue accurately. Hence, this loss of biorthonormality is due to the error propagation in the recurrence relations themselves, the two term recurrence relation preserves biorthonormality better as $k$ increases. Figure 8.9 shows the magnitude of the entries in the Gram matrix $W_{20}^H V_{20} - I$. There is clearly a pattern visible, however further research into the numerical properties of the recurrence relation must be done in order to interpret this.

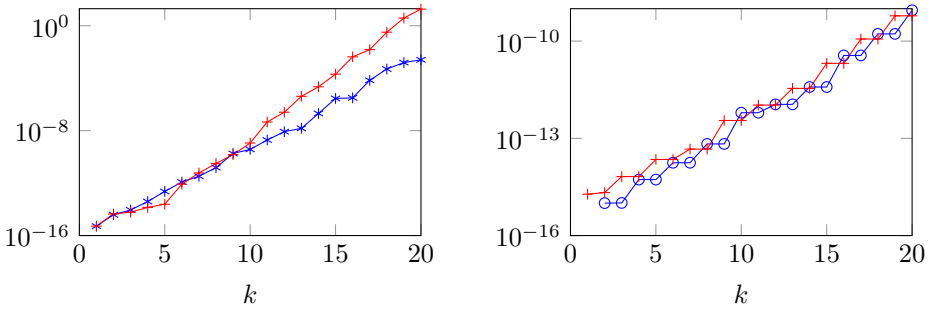



Figure 8.8: Biorthonormality and recursion error, respectively left and right, for a normal matrix $A \in \mathbb{C}^{200\times 200}$, with condition number $\kappa(A) = 10$ and $v \neq w$. Two term recurrence relation in blue and four term recurrence relation in red.

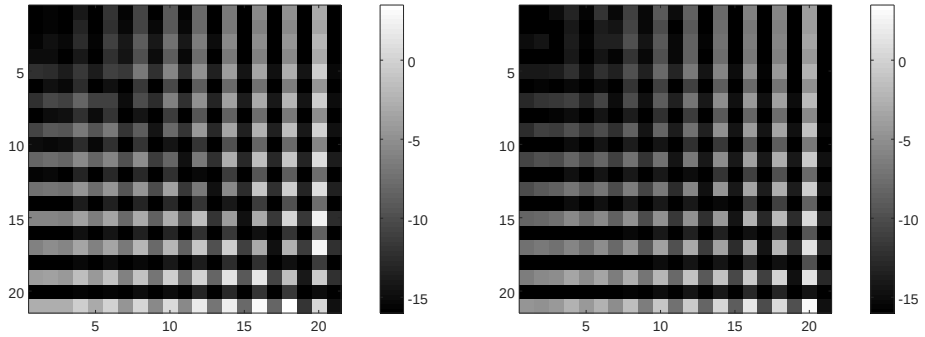



Figure 8.9: Magnitude (base 10) of the elements appearing in the Gram matrix $W_{20}^H V_{20} - I$ for the two and four term recurrence relation on the left and right, respectively. Bases $V, W$ are constructed for a normal matrix $A$ with $\kappa(A) = 10$ and $v \neq w$.

### 8.2.5 Conclusion

The study of structures in matrices allows to derive short recurrence relations. Low rank structure in recurrence matrices indicate that a certain redundancy appears in the underlying recurrence relation and a shorter (coupled) recurrence relation can be derived. The displacement structure of the related Gram matrix can also be used to derive short recurrence relations. The coupled two term recurrence relations are shown to outperform the four term recurrence relation in terms of the biorthonormality of the generated bases and the quality of the recurrence pencil.

# Chapter 9

# Inverse eigenvalue problems

This chapter is dedicated to numerical procedures for the computation of sequences of polynomials and rational functions with prescribed poles that are orthogonal with respect to a discrete inner product or linear functional. Discrete inner products of the form

$$\langle f, g \rangle_m = \sum_{i=1}^{m} |\alpha_i|^2 f(z_i)\overline{g(z_i)} \tag{9.1}$$

and discrete linear functionals of the form

$$\mathcal{L}\{fg\}_m = \sum_{i=1}^{m} \bar{\beta}_i \alpha_i f(z_i) g(z_i) \tag{9.2}$$

are considered. Three mathematically equivalent formulations are discussed: computing the recurrence coefficients of a sequence of (bi)ORFs, computing the LR factorization of a Gram matrix containing the moments of a sequence of (nonorthogonal) rational functions and constructing a structured pencil from given spectral data, that is, an inverse eigenvalue problem (IEP). These three formulations are discussed in Section 9.1. The IEPs related to (bi)ORFs are identified, these impose Hessenberg and tridiagonal structure on the pencil. An overview of some solution strategies for these three formulations is provided. The remainder of this chapter focuses on solution strategies for IEPs formulated for (bi)OPs and (bi)ORFs. Since the rational Arnoldi and Lanczos iteration constructs a Hessenberg and tridiagonal recurrence pencil, respectively these are suited to solve the IEPs. Section 9.2 provides the details. The connection between (bi)ORFs and (tridiagonal) Hessenberg pencils can also be used directly, without explicit use of Krylov subspaces. Section 9.3 and Section 9.4 propose updating strategies for the IEPs related to (bi)OPs and (bi)ORFs. Updating procedures assume the availability of a sequence of (bi)ORFs with respect to $\langle ., . \rangle_m$

(or $\mathcal{L}\{.\}_m$) and compute, starting from this sequence, the sequence for $\langle ., . \rangle_{m+1} = \langle ., . \rangle_m + w_{m+1} f(z_{m+1}) \overline{g(z_{m+1})}$ (or $\mathcal{L}\{.\}_{m+1} = \mathcal{L}\{.\}_m + w_{m+1} f(z_{m+1}) g(z_{m+1})$). For OPs we propose a novel parallelizable updating procedure. For ORFs, the pencil approach is new and so are the proposed procedures, these results are published in [167]. Section 9.5 contains numerical experiments for the proposed procedures.
Downdating procedures compute a sequence of (bi)ORFs with respect to $\langle ., . \rangle_m$ (or $\mathcal{L}\{.\}_m$) when the sequence with respect to $\langle ., . \rangle_{m+1} = \langle ., . \rangle_m + \alpha_{m+1} f(z_{m+1}) g(z_{m+1})$ (or $\mathcal{L}\{.\}_{m+1} = \mathcal{L}\{.\}_m + \bar{\beta}_{m+1} \alpha_{m+1} f(z_{m+1}) g(z_{m+1})$) is available. The problem of downdating is more difficult than updating, due to its inherent ill-conditioning [25, 60]. Downdating procedures for discrete inner products are proposed in Section 9.6. The results for downdating procedures will be published in a paper that is in preparation [164].

## 9.1 Problem reformulation

The main problem is that of constructing rational functions with prescribed poles orthogonal with respect to a discrete inner product (9.1) or linear functional (9.2). A sequence of (bi)ORFs can be represented in a given basis of rational functions or it can be represented by its recurrence pencil. The first representation depends strongly on the choice of a good basis to represent the sequence. The second representation uses the (bi)ORFs themselves as the basis, moreover, it does so in an implicit manner, see Chapter 6 and Chapter 7 for details of how a recurrence pencil represents a sequence of rational functions.
Three mathematically equivalent formulations of the main problem are introduced in this section. A functional and structured matrix problem require a recurrence pencil, these are introduced in Section 9.1.1 and Section 9.1.3, respectively. Section 9.1.2 introduces the Gram matrix formulation, based on Lemma 2.2. This formulation leads to a solution represented in the basis used to generate the Gram matrix. The equivalence between the three formulations follows largely from the discussion in above chapters, the IEP formulation is derived in detail in Section 9.1.3. Section 9.1.4 shortly recapitulates the connections between the three formulations.

### 9.1.1 Functional problem

Problem 9.1 introduces the orthogonal case, i.e., orthonormality with respect to $\langle ., . \rangle_m$.

**Problem 9.1** (Generate ORFs). *Given nodes* $\{z_i\}_{i=1}^m$, $z_i \in \mathbb{C}$ *and weights* $\{\alpha_i\}_{i=1}^m$, $\alpha_i \in \mathbb{C}$, *determining the inner product* $\langle f, g \rangle_m := \sum_{i=1}^m |\alpha_i|^2 f(z_i) \overline{g(z_i)}$. *Let* $\Xi = \{\xi_i\}_{i=1}^{m-1}$ *be a set of given poles* $\xi_i \in \overline{\mathbb{C}}$ *and let* $r_0 \equiv \sqrt{\sum_{i=1}^m |\alpha_i|^2}$. *Compute a sequence* $\{r_k\}_{k=0}^{m-1}$ *of rational functions* $r_k \in \mathcal{R}_k^\Xi \backslash \mathcal{R}_{k-1}^\Xi$ *which are orthonormal with respect to*

$\langle .,. \rangle_m$, *i.e.,*

$$\langle r_i, r_j \rangle_m \begin{cases} = 0, & \text{if } i \neq j \\ = 1, & \text{if } i = j \end{cases}.$$

Problem 9.2 replaces the inner product by a linear functional $\mathcal{L}_m\{.\}$, thus leading to biorthonormal rational functions.

**Problem 9.2** (Generate biORFs)**.** *Given nodes* $\{z_i\}_{i=1}^m$, $z_i \in \mathbb{C}$ *and weights* $\{\alpha_i\}_{i=1}^m, \{\beta_i\}_{i=1}^m$, $\alpha_i, \beta_i \in \mathbb{C}$, *which determine a linear functional* $\mathcal{L}_m\{f,g\} := \sum_{i=1}^m \beta_i \alpha_i f(z_i) g(z_i)$. *Let* $\Xi = \{\xi_i\}_{i=1}^{m-1}$, $\Theta = \{\theta_i\}_{i=1}^{m-1}$ *be sets of given poles* $\xi_i, \theta_i \in \overline{\mathbb{C}}$. *And let* $r_0, s_0$ *be given constants such that* $\mathcal{L}_m\{r_0 s_0\} = 1$. *Compute a pair of sequences* $\{r_k\}_{k=0}^{m-1}$, $\{s_k\}_{k=0}^{m-1}$ *of rational functions* $r_k \in \mathcal{R}_k^{\Xi} \backslash \mathcal{R}_{k-1}^{\Xi}$, $s_k \in \mathcal{R}_k^{\Theta} \backslash \mathcal{R}_{k-1}^{\Theta}$ *which are biorthonormal with respect to* $\mathcal{L}_m$, *i.e.,*

$$\mathcal{L}_m\{r_i s_j\} \begin{cases} = 0, & \text{if } i \neq j \\ = 1, & \text{if } i = j \end{cases}.$$

When all poles are chosen to be infinite $\xi_i = \infty$, Problem 9.1 reduces to generating OPs and Problem 9.2 to generating biOPs. The polynomial problems will be treated separately in the sequel and not just as a special case of the rational function problem. Stieltjes-like procedures [70, 79–81, 83, 84, 138, 140] can be used to solve Problem 9.1 and Problem 9.2. The discretized Stieltjes procedure for OPs is provided in Algorithm 6. This procedure [78] computes alternately recurrence coefficients for $p_i$, the $i$th orthogonal polynomial, and the value of $p_i$ for the given (discrete) inner product, which can then be used to compute the recurrence coefficients of $p_{i+1}$ and so on.
For ORFs this Stieltjes procedure can be combined with a modification of the inner product by division with a polynomial [80, 81, 83]. Or, a Stieltjes-like procedure can be developed based on recurrence relations for ORFs [138]. This idea is not pursued further. Since, based on observations and results for the classical OP case, it is expected that methods based on IEPs will provide a more accurate solution at the same or lower cost.
These observations are summarized now, with the necessary references. For discrete inner products with $m$ nodes on the real line and $i$ approaching $m$, the Stieltjes procedure for OPs may suffer accuracy loss due to the underlying three term recurrence relation exhibiting *pseudostability* [69, 79, 80, 83]. This can lead to large errors on the computed recurrence coefficients. Gautschi mentioned, after analyzing the effect of pseudostability on the Stieltjes procedure, that the method by Gragg and Harrod [92] does not suffer from the pseudostability of the underlying three term recurrence relation. Their method solves an IEP with unitary similarity transformations on a diagonal matrix, for details see Section 9.1.3. Also Reichel [140] noted that this method achieves the same accuracy as the Stieltjes procedure and better accuracy as $i$ approaches $m$. While these observations are made for OPs, we expect that for

ORFs the behavior is similar. That is, methods based on IEPs will lead to the most numerically stable methods.

### 9.1.2 Gram matrix problem

Orthogonal rational functions $\{r_i\}_{i=0}^{m-1}$ for $\langle .,. \rangle_m$ are the rational functions for which the associated Gram matrix is the identity matrix

$$\left[\langle r_i, r_j \rangle_m\right]_{j,i=0}^{m-1} = I.$$

Starting from any sequence of rational functions $\{\hat{r}_i\}_{i=0}^{m-1}$ satisfying $\hat{r}_i \in \mathcal{R}_i^{\Xi} \backslash \mathcal{R}_{i-1}^{\Xi}$, a sequence of ORFs can be generated. The factorization of the Gram matrix for the sequence $\{\hat{r}_i\}_i$ generated by $\langle .,. \rangle_m$ or $\mathcal{L}_m\{.\}$ provides recurrence coefficients for the ORFs in the basis $\hat{\boldsymbol{r}} := \begin{bmatrix}\hat{r}_0 & \hat{r}_1 & \dots & \hat{r}_{m-1}\end{bmatrix}$. For an inner product the Gram matrix is Hermitian positive definite and, therefore, a Cholesky decomposition can be used. The factor $R$, more precisely $\bar{R}^{-1}$, obtained by solving Problem 9.3 contains in its columns the coefficients for the ORFs $\{r_i\}_i$ in the basis $\hat{\boldsymbol{r}}$.

**Problem 9.3** (Generate ORFs: Gram matrix)**.** *Given any sequence of rational functions $\{\hat{r}_i\}_{i=0}^{m-1}$, with $\hat{r}_i \in \mathcal{R}_i^{\Xi} \backslash \mathcal{R}_{i-1}^{\Xi}$, and the corresponding moments $\hat{\mu}_{j,i} = \langle \hat{r}_i, \hat{r}_j \rangle_m$. Let $\hat{M} = \left[\hat{\mu}_{j,i}\right]_{j,i=0}^{m-1}$ be the associated Gram matrix. Compute the Cholesky factorization of $\hat{M} = R^H R$.*

Linear functionals do not guarantee a positive definite Gram matrix. Assume the given discrete linear functional is quasi-definite, then Problem 9.4 has a solution.

**Problem 9.4** (Generate biORFs: Gram matrix)**.** *Given any pair of sequences of rational functions $\{\hat{r}_i\}_{i=0}^{m-1}$ and $\{\hat{s}_i\}_{i=0}^{m-1}$, with $\hat{r}_i \in \mathcal{R}_i^{\Xi} \backslash \mathcal{R}_{i-1}^{\Xi}$ and $\hat{s}_i \in \mathcal{R}_i^{\Theta} \backslash \mathcal{R}_{i-1}^{\Theta}$ Consider the corresponding moments $\hat{\mu}_{j,i} = \mathcal{L}_m\{\hat{r}_i \hat{s}_j\}$ and the Gram matrix, which is assumed to be strongly nonsingular, $\hat{M} = \left[\hat{\mu}_{j,i}\right]_{j,i=0}^{m-1}$. Compute the LR factorization of $\hat{M} = LR$.*

The computation of the LR factorization is numerically unstable since pivoting strategies are not allowed.
For the orthogonal formulation, an alternative problem can be formulated, avoiding the need to construct a Gram matrix. This alternative problem is formulated in terms of the QR factorization of the Vandermonde-like matrix

$$\begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_m \end{bmatrix} \begin{bmatrix} \hat{r}_0(z_1) & \hat{r}_1(z_1) & \dots & \hat{r}_{m-1}(z_1) \\ \hat{r}_0(z_2) & \hat{r}_1(z_2) & \dots & \hat{r}_{m-1}(z_2) \\ \vdots & \vdots & & \vdots \\ \hat{r}_0(z_m) & \hat{r}_1(z_m) & \dots & \hat{r}_{m-1}(z_m) \end{bmatrix}.$$

For details we refer to Chapter 5.

### 9.1.3 Structured matrix problem

The following problems are formulated by relying on results discussed in Chapter 6 and Chapter 7, more precisely, the connection between (bi)ORFs and (bi)orthonormal bases for rational Krylov subspaces. This connection is provided in Lemma 7.1 and Lemma 7.2 for orthogonal and biorthogonal rational functions, respectively. When these lemmas are applied to a diagonal matrix $A = \Lambda = \operatorname{diag}(\{z_i\}_i)$, the recurrence relation for (bi)orthonormal bases in matrix notation can also be interpreted as a generalized eigenvalue decomposition (GEVD) for the recurrence pencil. Details are provided below. Now we sketch the general idea of how this interpretation leads to an IEP.
The interpretation of the matrix form of the recurrence relations for ORFs as a GEVD of a structured pencil allows a reverse argument. Namely, a pencil with a specific structure and certain spectral data is a recurrence pencil for (bi)ORFs with respect to some discrete inner product or linear functional. Imposing spectral data and searching for a pencil with a certain structure, denoted as belonging to a class of matrix pencils $\mathcal{N}$, is a structured inverse eigenvalue problem. Problem 9.5 provides the definition of a structured IEP [27, 44].

**Problem 9.5** (Structured IEP-matrix formulation)**.** *Given a diagonal matrix* $\Lambda = diag(z_1, \ldots, z_m) \in \mathbb{C}^{m\times m}$*,* $z_i \neq z_j$ *for* $i \neq j$*, and vectors* $v, w \in \mathbb{C}^m$*. Find a pencil* $(B, C) \in \mathcal{N}$ *and* $V \in \mathbb{C}^{m\times m}$*, where* $Ve_1 = \frac{v}{\nu}$ *and* $V^{-H}e_1 = \frac{w}{\eta}$*, with* $\bar{\eta}\nu = \langle v, w\rangle_E = w^H v \neq 0$*, such that*

$$V^{-1}\Lambda V C = B. \tag{9.3}$$

The derivation and the formulation of four structured IEPs related to (bi)OPs and (bi)ORFs are provided below. The structure that is imposed on the pencil $(B, C)$ in Problem 9.5 determines the function class and type of orthogonality. A short overview of relevant pencil structures in IEPs is given together with some references where they appear:

- OPs: $\mathcal{N}_1 = \{\text{Hessenberg matrices}\}$, Problem 9.7.
- OPs on $\mathbb{R}$: $\mathcal{N}_2 = \{\text{Jacobi matrices}\}$, see literature [52, 92].
- OPs on $\mathbb{T}$, the unit circle: $\mathcal{N}_3 = \{\text{unitary Hessenberg matrices}\}$, see literature [1, 142].
- BiOPs: $\mathcal{N}_4 = \{\text{tridiagonal matrices}\}$, Problem 9.9.
- ORFs: $\mathcal{N}_5 = \{\text{Hessenberg pencils}\}$, Problem 9.6.

- ORFs: $\mathcal{N}_6 = \{\text{Semi-separable plus diagonal}\}$, see literature [161].
- ORFs: $\mathcal{N}_7 = \{\text{extended Hessenberg matrix}\}$, see literature [128].
- BiORFs: $\mathcal{N}_8 = \{\text{tridiagonal pencils}\}$, Problem 9.8.
- BiORFs: $\mathcal{N}_9 = \{\text{extended tridiagonal matrix}\}$, see literature [128].

As can be seen in the overview, the formulation for (bi)ORFs as a matrix IEP, $\mathcal{N}_6$, $\mathcal{N}_7$ and $\mathcal{N}_9$, is discussed in literature already. The formulation as a pencil IEP, $\mathcal{N}_5$ and $\mathcal{N}_8$ is novel and so are the procedures that will be proposed. The pencil formulation provides more flexibility than the single matrix representations, and the expectation is that this flexibility can be used to obtain procedures with better numerical stability properties.
First IEPs for rational functions orthogonal with respect to an inner product are introduced and afterwards for rational functions biorthonormal with respect to a linear functional. Afterwards an introduction into two possible solution strategies is given, one uses Krylov subspaces methods to generate (bi)orthonormal bases and the other uses an updating procedure.

### IEP for ORFs

The correspondence between recurrence relations for ORFs and a GEVD for a structured pencil is clarified. This allows the formulation of Problem 9.6. Corollary 9.1 relates ORFs with a discrete inner product to an orthonormal basis for a rational Krylov subspace.

**Corollary 9.1** (Recurrence pencil equivalence: ORFs)**.** *A sequence of rational functions* $\{r_i\}_{i=0}^{m-1}$ *in* $\mathcal{R}_{m-1}^{\Xi}$ *orthonormal with respect to*

$$\langle f, g\rangle_m = \sum_{i=1}^{m} |\alpha_i|^2 f(z_i)\overline{g(z_i)}$$

*is generated by the same recurrence coefficients as the nested orthonormal basis* $\{q_i\}_{i=0}^{m-1}$ *for* $\mathcal{K}_m(\Lambda, v; \Xi)$*, where* $\Lambda = diag(\{z_i\}_{i=1}^{m})$*,* $v = \begin{bmatrix}\alpha_1 & \dots & \alpha_m\end{bmatrix}^\top$.

*Proof.* Follows from Lemma 7.1 and Theorem 7.1. □

This corollary implies that a Hessenberg pencil $(H, K) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ satisfying

$$\Lambda Q K = Q H, \tag{9.4}$$

with $\Lambda = \text{diag}(\{z_i\}_{i=1}^m)$, $Qe_1 = v/\|v\|$, $\text{span}\{Q\} = \mathcal{K}_m(\Lambda, v; \Xi)$ and $Q^H Q = I$, also satisfies

$$z\boldsymbol{r}K = \boldsymbol{r}H, \tag{9.5}$$

with $\boldsymbol{r} = \begin{bmatrix} r_0 & \dots & r_{m-1} \end{bmatrix}$ forming a sequence of ORFs in $\mathcal{R}_{m-1}^{\Xi}$ for the discrete inner product $\langle f, g \rangle_m = \sum_{i=1}^m |\alpha_i|^2 f(z_i)\overline{g(z_i)}$.
From the equivalence between $(H, K)$ in (9.4) and in (9.5) it is easy to derive the following equality relating $Q$, the orthonormal basis for $\mathcal{K}(\Lambda, v; \Xi)$, to $\boldsymbol{r}$, the sequence of rational functions in $\mathcal{R}^{\Xi}$ orthonormal to $\langle ., . \rangle_m$,

$$Q = \begin{bmatrix} \alpha_1 r_0(z_1) & \alpha_1 r_1(z_1) & \dots & \alpha_1 r_{m-1}(z_1) \\ \alpha_2 r_0(z_2) & \alpha_2 r_1(z_2) & \dots & \alpha_2 r_{m-1}(z_2) \\ \vdots & \vdots & & \vdots \\ \alpha_m r_0(z_m) & \alpha_m r_1(z_m) & \dots & \alpha_m r_{m-1}(z_m) \end{bmatrix}.$$

The correspondence of orthonormality for $\{q_i\}_{i=0}^{m-1}$ with respect to $\langle ., . \rangle_E$ to the orthonormality of $\{r_i\}_{i=0}^{m-1}$ with respect to $\langle ., . \rangle_m$ follows from the equality of their respective Gram matrices. Both Gram matrices are equal to $Q^H Q$,

$$I = Q^H Q = \begin{bmatrix} \sum_{i=1}^m |\alpha_i|^2 r_0(z_i)\overline{r_0(z_i)} & \dots & \sum_{i=1}^m |\alpha_i|^2 r_0(z_i)\overline{r_{m-1}(z_i)} \\ \sum_{i=1}^m |\alpha_i|^2 r_1(z_i)\overline{r_0(z_i)} & \dots & \sum_{i=1}^m |\alpha_i|^2 r_1(z_i)\overline{r_{m-1}(z_i)} \\ \vdots & & \vdots \\ \sum_{i=1}^m |\alpha_i|^2 r_{m-1}(z_i)\overline{r_0(z_i)} & \dots & \sum_{i=1}^m |\alpha_i|^2 r_{m-1}(z_i)\overline{r_{m-1}(z_i)} \end{bmatrix}$$

$$= \begin{bmatrix} \langle r_0, r_0 \rangle_m & \dots & \langle r_0, r_{m-1} \rangle_m \\ \langle r_1, r_0 \rangle_m & \dots & \langle r_1, r_{m-1} \rangle_m \\ \vdots & & \vdots \\ \langle r_{m-1}, r_0 \rangle_m & \dots & \langle r_{m-1}, r_{m-1} \rangle_m \end{bmatrix} = I.$$

The key observation that must be made is that Equation (9.4) represents the recurrence relation for the orthonormal basis vectors for $\mathcal{K}_m(\Lambda, v; \Xi)$ and simultaneously it represents the generalized eigenvalue decomposition of the pencil $(H, K)$, with eigenvalues $\{z_i\}_i$ and eigenvectors the columns of $Q^H$.
From (9.5) follows that the eigenvalues $\{z_i\}$ correspond to the nodes of $\langle ., . \rangle_m$. The weight $|\alpha_i|^2$, more precisely $\alpha_i$, appears as the first element of the normalized eigenvector corresponding to $z_i$, this follows from $Q^H e_i = \bar{\alpha}_i \overline{\boldsymbol{r}(z_i)}$. Now that the location of the nodes and weights of $\langle ., . \rangle_m$ in the GEVD of the recurrence pencil is known, an IEP can be formulated for this pencil. Problem 9.6 introduces the Hessenberg pencil IEP.

**Problem 9.6** (Hessenberg pencil inverse eigenvalue problem (HPIEP) $\mathcal{N}_5$)**.** *Given a diagonal matrix $\Lambda = diag(z_1, \dots, z_m)$ of distinct nodes $z_i \in \mathbb{C}$, a vector of weights*

*$v \in \mathbb{C}^m$ and a set of poles $\Xi = \{\xi_i\}_{i=1}^{m-1}$, $\xi_i \in \overline{\mathbb{C}}$. Construct a Hessenberg pencil $(H_m, K_m) \in \mathbb{C}^{m\times m}$, with $\frac{h_{i+1,i}}{k_{i+1,i}} = \xi_i$, $i = 1, 2, \ldots, m-1$, and a unitary matrix $Q_m \in \mathbb{C}^{m\times m}$ such that*

$$Q_m^H \Lambda Q_m K_m = H_m \quad \textit{and} \quad Q_m e_1 = \frac{v}{\|v\|}. \tag{9.6}$$

The uniqueness of the solution to Problem 9.6 follows from the implicit Q theorem for Hessenberg pencils [40, Theorem 5.1], which states a one-to-one correspondence between proper Hessenberg pencils and rational Krylov subspaces.
The polynomial case is retrieved by choosing $\xi_i = \infty$ for all $i$. We formulate the related IEP in Problem 9.7, because it is an important special case for which many results and procedures exist and for which we will propose a new parallelizable approach in Section 9.3.2.

**Problem 9.7** (Hessenberg inverse eigenvalue problem (HIEP) $\mathcal{N}_1$)**.** *Given a diagonal matrix $\Lambda = diag(z_1, \ldots, z_m)$ of distinct nodes $z_i \in \mathbb{C}$ and a vector of weights $v \in \mathbb{C}^m$. Construct a Hessenberg matrix $H_m \in \mathbb{C}^{m\times m}$ and a unitary matrix $Q_m \in \mathbb{C}^{m\times m}$ such that*

$$Q_m^H \Lambda Q_m = H_m \quad \textit{and} \quad Q_m e_1 = \frac{v}{\|v\|}. \tag{9.7}$$

### IEP for biORFs

For discrete linear functionals the relation between biORFs and a pair of biorthonormal bases for Krylov subspaces is provided in Corollary 9.2.

**Corollary 9.2** (Recurrence pencil equivalence: biORFs)**.** *A pair of sequences of rational functions $\{r_i\}_{i=0}^{m-1}$ and $\{s_i\}_{i=0}^{m-1}$ in $\mathcal{R}_{m-1}^{\Xi}$ and $\mathcal{R}_{m-1}^{\Theta}$ biorthonormonal with respect to*

$$\mathcal{L}_m\{fg\} = \sum_{i=1}^{m} \bar{\beta}_i \alpha_i f(z_i) g(z_i)$$

*is generated by the same recurrence coefficients as the pair of nested biorthonormal bases $\{v_i\}_{i=0}^{m-1}$ and $\{w_i\}_{i=0}^{m-1}$ for $\mathcal{K}_m(\Lambda, v; \Xi)$ and $\mathcal{K}_m(\Lambda^H, w; \Theta)$, respectively, with $\Lambda = diag(\{z_i\}_{i=1}^{m-1})$ and $v = \begin{bmatrix}\alpha_1 & \ldots & \alpha_m\end{bmatrix}^\top$, $w = \begin{bmatrix}\beta_1 & \ldots & \beta_m\end{bmatrix}^\top$.*

*Proof.* Follows from Lemma 7.2 and Theorem 7.2. □

A similar argument as for the IEPs above holds. By Theorem 4.6, we have a tridiagonal pencil $(T, S)$ and biorthonormal bases $V, W \in \mathbb{C}^{m\times m}$ for $\mathcal{K}(\Lambda, v; \Xi)$ and $\mathcal{K}(\Lambda^H, w; \Theta)$ satisfying

$$\Lambda V S = V T$$

and simultaneously, for biORFs $\boldsymbol{r} := \begin{bmatrix} r_0 & \dots & r_{m-1} \end{bmatrix}$,

$$\Lambda \boldsymbol{r} S = \boldsymbol{r} T.$$

And we have

$$V = \begin{bmatrix} \alpha_1 r_0(z_1) & \alpha_1 r_1(z_1) & \dots & \alpha_1 r_{m-1}(z_1) \\ \alpha_2 r_0(z_2) & \alpha_2 r_1(z_2) & \dots & \alpha_2 r_{m-1}(z_2) \\ \dots & \vdots & & \vdots \\ \alpha_m r_0(z_m) & \alpha_m r_1(z_m) & \dots & \alpha_m r_{m-1}(z_m) \end{bmatrix}$$

$$W = \begin{bmatrix} \beta_1 \overline{s}_0(\bar{z}_1) & \beta_1 \overline{s}_1(\bar{z}_1) & \dots & \beta_1 \overline{s}_{m-1}(\bar{z}_1) \\ \beta_2 \overline{s}_0(\bar{z}_2) & \beta_2 \overline{s}_1(\bar{z}_2) & \dots & \beta_2 \overline{s}_{m-1}(\bar{z}_2) \\ \dots & \vdots & & \vdots \\ \beta_m \overline{s}_0(\bar{z}_m) & \beta_m \overline{s}_1(\bar{z}_m) & \dots & \beta_m \overline{s}_{m-1}(\bar{z}_m) \end{bmatrix}.$$

Noting that $W^H V = I$ implies $W^H = V^{-1}$ and therefore

$$W^H \Lambda V S = T \Leftrightarrow V^{-1} \Lambda V S = T,$$

that is, a GEVD of the recurrence pencil $(T, S)$. This holds in both directions. Thus an IEP can be formulated, Problem 9.8 provides an IEP corresponding to biORFs.

**Problem 9.8** (Tridiagonal pencil inverse eigenvalue problem (TPIEP) $\mathcal{N}_8$)**.** *Given a diagonal matrix $\Lambda = diag(z_1, \dots, z_m)$ of distinct nodes $z_i \in \mathbb{C}$, two vectors containing weights $v, w \in \mathbb{C}^m$, $\langle v, w \rangle \neq 0$, and sets of poles $\Xi = \{\xi_i\}_{i=1}^{m-1}$ and $\Theta = \{\theta_i\}_{i=1}^{m-2}$, $\xi_i, \theta_i \in \overline{\mathbb{C}}$. Construct a tridiagonal pencil $(T_m, S_m) \in \mathbb{C}^{m \times m}$, with $\frac{t_{i,i+1}}{s_{i,i+1}} = \xi_i$, $i = 1, 2, \dots, m-1$ and $\frac{t_{i+1,i}}{s_{i+1,i}} = \bar{\theta}_{i-1}$, $i = 2, 3, \dots, m-1$ and matrices $V_m, W_m \in \mathbb{C}^{m \times m}$ such that*

$$W_m^H \Lambda V_m S_m = T_m \quad \text{and} \quad V_m e_1 = \frac{v}{\nu}, \quad W_m e_1 = \frac{w}{\eta}, \tag{9.8}$$

*where $W_m^H V_m = I$ and $w^H v = \nu \bar{\eta} \neq 0$.*

Thus a tridiagonal pencil $(T, S)$, where both the subdiagonal and superdiagonal elements should satisfy some restrictions on their ratios, must be constructed. These ratio restrictions guarantee that the corresponding rational Krylov subspaces have the appropriate poles [165]. To obtain the other sequence $\{s_i\}$, biorthonormal to $\{r_i\}$, an additional tridiagonal pencil $(\widetilde{T}, \widetilde{S})$ must be constructed such that $V^H \Lambda^H W \widetilde{S} = \widetilde{T}$. For biOPs a matrix IEP can be formulated, this is a tridiagonal IEP and is stated in Problem 9.9.

**Problem 9.9** (Tridiagonal inverse eigenvalue problem (TIEP) $\mathcal{N}_4$)**.** *Given a diagonal matrix $\Lambda = diag(z_1, \dots, z_m)$ of distinct nodes $z_i \in \mathbb{C}$ and two vectors containing weights*

$v, w \in \mathbb{C}^m$, $\langle v, w \rangle \neq 0$. *Construct a tridiagonal matrix* $T_m \in \mathbb{C}^{m \times m}$ *and matrices* $V_m, W_m \in \mathbb{C}^{m \times m}$ *such that*

$$W_m^H \Lambda V_m S_m = T_m \quad \textit{and} \quad V_m e_1 = \frac{v}{\nu}, \quad W_m e_1 = \frac{w}{\eta}, \tag{9.9}$$

*where* $W_m^H V_m = I$ *and* $w^H v = \nu \bar{\eta} \neq 0$.

The tridiagonal IEP has appeared in the literature, for a tridiagonal matrix with additional structure, called a *pseudo-Jacobi matrix*. Pseudo-Jacobi matrices arise in quantum mechanics [9, 48, 185, 186], the study of nonlinear Toda lattices [96] and for certain Sturm-Liouville operators [189]. For details on the structure of a pseudo-Jacobi matrix we refer to [186].

### Arnoldi/Lanczos-type procedures

The relationship stated in Corollary 9.1 and Corollary 9.2 between bases for Krylov spaces and ORFs suggests that Arnoldi and Lanczos-type iterations are suitable solution procedures. Applying Arnoldi/Lanczos-type procedures for a diagonal matrix $\Lambda = \text{diag}(z_1, \ldots, z_m)$ and starting vectors $v, w \in \mathbb{C}^m$ results in (bi)orthonormal bases $V, W \in \mathbb{C}^{m \times m}$ and an associated recurrence pencil $(H, K)$. This pencil generates (bi)ORFs with respect to an inner product or linear functional with nodes $\{z_i\}_{i=1}^m$ and weights $\{\alpha_i\}_{i=1}^m$ and $\{\beta_i\}_{i=1}^m$. Section 9.2 elaborates on Krylov subspace methods for the solution of IEPs. For the IEPs involving long recurrence relations, corresponding to $\mathcal{N}_1, \mathcal{N}_5$ these procedures are expected to perform well if reorthogonalization is used [50]. For IEPs with underlying short term recurrence relations, $\mathcal{N}_2, \mathcal{N}_4$ and $\mathcal{N}_8$, Lanczos-type procedures are used. The biorthonormality of the generated bases deteriorates fast with growing size, reorthogonalization can improve this. Alternatively, updating procedures work directly with the structure of the recurrence pencils and avoid the cost of reorthogonalization.

### Updating procedures

The relationship between Krylov subspaces and rational functions provides a connection between sequences of (bi)ORFs and certain structured matrices, the recurrence pencils for bases of the Krylov spaces. Hence, it is also possible to operate on the recurrence pencils immediately without the explicit use of Krylov subspaces. The one-to-one correspondence of a Hessenberg pencil to a rational Krylov subspace follows from the rational implicit Q theorem [40].
The proposed updating procedures are of this kind, they work immediately on the structure of the recurrence pencil. Representing a sequence of rational functions by

their recurrence pencil is discussed in Chapter 7.
Section 9.3 and Section 9.4 propose updating procedures for HPIEPs and TPIEPs, respectively. Such a procedure starts from, e.g., a Hessenberg pencil $(H_m, K_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ with eigenvalues $\{\lambda_i\}_{i=1}^{m}$ and $\{\alpha_i\}_{i=0}^{m}$ the first components of the eigenvectors. Reusing this solution, the solution $(H_{m+1}, K_{m+1}) \in \mathbb{C}^{(m+1)\times(m+1)} \times \mathbb{C}^{(m+1)\times(m+1)}$ for $\{\lambda_i\}_{i=1}^{m+1}$ and $\{\alpha_i\}_{i=0}^{m+1}$, i.e., after adding a node-weight pair $(\lambda_{m+1}, \alpha_{m+1})$ to the HPIEP can be computed. For TPIEPs the idea is the same, only requiring an additional weight.
The introduction of the new node-weight pair, details follow in Section 9.3 and Section 9.4, will disrupt the structure of the recurrence pencil. Restoring the structure of this pencil to Hessenberg or tridiagonal form corresponds to computing the solution $(H_{m+1}, K_{m+1})$ to the updated problem.

An updating procedure can be used to solve an IEP of size $m > 1$ without possessing a (nontrivial) solution to a smaller IEP. The idea is explained by using a HIEP, Problem 9.7, as an example. For any IEP, the same principle applies.
The updating procedure for HIEPs can be initiated with the trivial $1 \times 1$ solution $H, Q \in \mathbb{C}^{1\times 1}$ to the HIEP with $\Lambda := z_1 \in \mathbb{C}$ and $v := 1$, i.e., $H = z_1$ and $Q = 1$. The associated inner product is $\langle f, g\rangle_1 = f(z_1)\overline{g(z_1)}$. For $k = 2, 3, \ldots, m$, a solution for $\langle ., .\rangle_k$ can be obtained by updating the solution for $\langle ., .\rangle_{k-1}$.
The order in which the node-weight pairs are added can be chosen and will influence the accuracy of the finite precision solution.

### 9.1.4 Equivalence between problems

The solutions to Problem 9.1, Problem 9.3 and Problem 9.6 are mathematically equivalent in the sense that the generated sequences of orthonormal rational functions are the same. Lemma 9.1 formalizes this equivalence and summarizes the main connections.

**Lemma 9.1.** *Consider the vector space $\mathcal{R}_{m-1}^{\Xi}$ and a discrete inner product on this space $\langle ., .\rangle_m$. Let $\{r_i\}, \{\hat{r}_i\}$ and $\{\tilde{r}_i\}$ denote the ORFs obtained by solving Problem 9.1, Problem 9.3 and Problem 9.6, respectively. Then these are essentially the same, i.e., $r_i = \hat{r}_i = \tilde{r}_i$ for all $i$, up to multiplication with a constant of modulus 1.*

*Proof.* A sequence of ORFs $\{r_i\}_{i=0}^{k}$, $k < m$, with respect to $\langle ., .\rangle_m$ is unique. Thus $\{r_i\}$ solving Problem 9.1 is unique. The Gram matrix $M$ associated with $\langle ., .\rangle_m$ and $\mathcal{R}_k^{\Xi}$ is Hermitian positive definite, thus $M = R^H R$, and contains in $R$, more precisely $\bar{R}^{-1}$, the coefficients of $\{\hat{r}_i\}$, which form ORFs for $\langle ., .\rangle_m$ By uniqueness of the ORFs and of the Cholesky decomposition, $r_i = \hat{r}_i$ for all $i$.
For the solution $\{\tilde{r}_i\}$, the connection to Krylov subspaces is used, the Gram matrix

$M$ of ORFs with respect to $\langle ., . \rangle_m$ in the monomial basis and the Gram matrix $M = B_m^H B_m$ associated with the Krylov matrix $B_m$ for $\mathcal{K}_m(\Lambda, v; \Xi)$ ($\Lambda$ and $v$ as in Corollary 9.1) are the same. The Hessenberg recurrence pencil for a nested orthonormal basis $Q_m$ for $\mathcal{K}_m(\Lambda, v; \Xi)$ is unique and essentially equal to the QR decomposition of $B_m$, i.e., $B_m = Q_m \tilde{R}_m$. From orthonormality we have $I = Q_m^H Q_m = \tilde{R}_m^{-H} B_m^H B_m \tilde{R}_m^{-1}$ and thus $M = \tilde{R}_m^H \tilde{R}_m$ which is a Cholesky decomposition and since this is unique, we have $\tilde{r}_i = \hat{r}_i$. □

A similar result for biorthonormonal rational functions solving Problem 9.2, Problem 9.4 and Problem 9.8 is formulated in Lemma 9.2. The biORFs solving these three problems are essentially the same, under the condition that no breakdowns occur, i.e., the linear functional is quasi-definite or, equivalently, the Gram matrix is strongly nonsingular.

**Lemma 9.2.** *Consider the vector spaces $\mathcal{R}_{m-1}^{\Xi}$ and $\mathcal{R}_{m-1}^{\Theta}$ and a discrete quasi-definite linear functional $\mathcal{L}_m$ on the space $\mathcal{R}_{m-1}^{\Xi} \cdot \mathcal{R}_{m-1}^{\Theta}$. Let $\{(r_i, s_i)\}$, $\{(\hat{r}_i, \hat{s}_i)\}$ and $\{(\tilde{r}_i, \tilde{s}_i)\}$ denote the pairs of biORFs obtained by solving Problem 9.2, Problem 9.4 and Problem 9.8, respectively. Then these are essentially the same, i.e., $r_i = \hat{r}_i = \tilde{r}_i$ and $s_i = \hat{s}_i = \tilde{s}_i$ for all $i$, up to normalization.*

*Proof.* The proof is analogous to the one in Lemma 9.1, thus we only provide the necessary components. A sequence of biORFs exists and is unique if the associated Gram matrix (appearing in the determinantal expression) is strongly nonsingular [10, 31]. The LDR factorization of a matrix is unique [105]. The relation between the LDR factorization of the Gram matrix and the recurrence pencil for biorthonormal vectors for Krylov subspaces is clear from the exposition in Chapter 4. □

## 9.2 Krylov based procedures

The solution of the structured IEPs in Section 9.1.3 can be obtained by running an Arnoldi or Lanczos-type iteration for Krylov subspaces generated by suitably chosen matrix and starting vector(s). Given a discrete inner product or linear functional, Corollary 9.1 or Corollary 9.2, respectively, specify how to choose the matrix and starting vectors. The usual Arnoldi and Lanczos iteration solve Problem 9.7 and Problem 9.9, respectively. The discussion of these classical cases is limited to an overview with some references to the literature. Problem 9.6 and Problem 9.8 can be solved with the rational Arnoldi iteration, see Algorithm 4, and the rational Lanczos iteration introduced in Chapter 8. The formulation of the IEPs in terms of a pencil is new and is therefore discussed in more detail. Section 9.2.1 is dedicated to IEPs related to inner products and Section 9.2.2 to IEPs related to linear functionals.

### 9.2.1 Orthogonal

Inverse eigenvalue problems related to polynomials and rational functions orthogonal with respect to an inner product are the classes $\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3, \mathcal{N}_5, \mathcal{N}_6$ and $\mathcal{N}_7$.

#### Polynomials

A Hessenberg IEP, corresponding to $\mathcal{N}_1$, can be solved by the Arnoldi iteration, Algorithm 1. The resulting Hessenberg recurrence matrix generates polynomials orthogonal with respect to some inner product with nodes $\{z_i\}$ anywhere in the complex plane, $z_i \in \mathbb{C}$.
If all the nodes are located on the real line $z_i \in \mathbb{R}$, then $\Lambda^H = \Lambda$ and a Jacobi IEP, $\mathcal{N}_2$, can be formulated. The Hermitian Lanczos iteration [45] solves the Jacobi IEP and leads to OPs on the real line. If all nodes lie on the unit circle $z_i \in \mathbb{T}$, then $\Lambda^H = \Lambda^{-1}$ and a unitary Hessenberg IEP, $\mathcal{N}_3$, is appropriate. The recommended solution procedure uses the Schur form (core factorization) of the unitary Hessenberg [142]. The resulting OPs are Szegő polynomials [157].

#### Rational functions

The Hessenberg pencil formulation of an IEP, $\mathcal{N}_5$, corresponding to ORFs is new and was published in [167]. In [128, 161] the IEPs formulated for ORFs use a single recurrence matrix, $\mathcal{N}_6$ and $\mathcal{N}_7$.
The rational Arnoldi iteration, given in Algorithm 4, computes a nested orthonormal basis for $\mathcal{K}_k(\Lambda, v; \Xi)$ and corresponding Hessenberg pencil. Theorem 9.1 provides details on how the rational Arnoldi procedure can be used to solve Problem 9.6.

**Theorem 9.1.** *Let the unitary matrix $Q_m \in \mathbb{C}^{m\times m}$ and Hessenberg matrices $H_m, K_m \in \mathbb{C}^{m\times m}$ be obtained by applying the rational Arnoldi iteration for $\mathcal{K}_m(\Lambda, v; \Xi)$, with $\Lambda$, $v$ and $\Xi$ as in Problem 9.6. Then these matrices solve Problem 9.6.*

*Proof.* The matrix $Q_m$ is unitary since $\mathcal{K}_m(\Lambda, v; \Xi) = \mathbb{C}^m$ by the conditions on $\Lambda$ and $v$. Thus $Q_m^{-1}\Lambda Q_m K_m = H_m$, i.e., $\sigma(H_m, K_m) = \sigma(\Lambda)$. The remainder of the proof follows immediately from the properties of the recurrence pencil obtained by the rational Arnoldi iteration. □

The obtained pencil $(H_m, K_m)$ is the Hessenberg recurrence pencil representing the sequence of ORFs in $\mathcal{R}^{\Xi}$ for $\langle ., . \rangle_m$. For a state of the art implementation of the rational Arnoldi iteration we refer to `rktoolbox` [17].
If all nodes are on the real line, a tridiagonal pencil IEP can be formulated for ORFs

and the Hermitian Lanczos iteration [98, 133] is suited to solve this problem. More details follow in Section 9.2.2.

## 9.2.2 Biorthogonal

IEPs related to polynomials or rational functions biorthogonal with respect to a linear functional, $\mathcal{N}_4, \mathcal{N}_8$ and $\mathcal{N}_9$, must be solved with biorthogonal procedures. Hence, Lanczos-type iterations are appropriate. In finite precision Lanczos-type iterations suffer from loss of biorthogonality of the generated bases due to the underlying short recurrence relations. Thus, some form of reorthogonalization might be necessary to obtain satisfactory results.

### Polynomials

For biorthogonal polynomials a tridiagonal IEP, $\mathcal{N}_4$, can be formulated and it can be solved by the Lanczos iteration, given in Algorithm 2. For nodes $z_i$ on the real line and positive real weights $\alpha_i > 0$, the linear functional is an inner product and the tridiagonal IEP reduces to the Jacobi IEP. Just as the Lanczos iteration reduces to the Hermitian Lanczos iteration for $A^H = A$ and $v = w$.

### Rational functions

To solve the TPIEP, $\mathcal{N}_8$, the rational Lanczos iteration described in Appendix B.1 can be used. We made our `Matlab` code implementing the rational Lanczos iteration, proposed in our paper [165], available online [162]. Theorem 9.2 formalizes the connection between the solution of the rational Lanczos iteration and the solution to the TPIEP.

**Theorem 9.2.** *Let the matrices $V_m, W_m \in \mathbb{C}^{m\times m}$ be the bases and $T_m, S_m \in \mathbb{C}^{m\times m}$ the matrices of recurrences coefficients obtained by applying the rational Lanczos iteration for $\mathcal{K}_m(\Lambda, v; \Xi)$ and $\mathcal{K}_m(\Lambda^H, w; \Theta)$, with $\Lambda$, $v, w$ and $\Xi, \Theta$ as in Problem 9.8. Then these matrices solve Problem 9.8, under the assumption that the iteration does not break down.*

*Proof.* Since $\mathcal{K}_m(\Lambda, v; \Xi) = \mathcal{K}_m(\Lambda^H, w; \Theta) = \mathbb{C}^m$, under the no-breakdown assumption, and $W_m^H V_m = I$, we have $W_m = V_m^{-H}$. Thus $V_m^{-1}\Lambda V_m S_m = T_m$, i.e., $\sigma(T_m, S_m) = \sigma(\Lambda)$. The remainder of the proof follows immediately from the rational Lanczos iteration. With appropriate normalization $Ve_1 = \frac{v}{\nu}$ and $We_1 = \frac{w}{\eta}$, with $\nu\bar{\eta} = w^H v$. □

If all nodes $\{z_i\}$ defining the linear functional $\mathcal{L}_m$ are on the real line, $z_i \in \mathbb{R}$ for all $i$, and the associated weights $\alpha_i > 0$, then $\Lambda = \mathrm{diag}(\{z_i\})$ is a Hermitian matrix and the linear functional is in fact an inner product. By Theorem 4.8 there exists a tridiagonal pencil which generates a sequences of ORFs. Fasino and Gemignani [65] also noted the connection between a real semi-separable plus diagonal IEP and a tridiagonal pencil IEP.
The Hessenberg pencil computed with the rational Arnoldi iteration, however, does not reduce to a tridiagonal pencil when all $z_i \in \mathbb{R}$. So in order to exploit the short recurrence relation, the biorthogonal method, i.e., the rational Lanczos iteration, is a more natural starting point. The rational Lanczos iteration will generate a biorthonormal pair of bases that reduces to a single orthonormal basis. This is, of course, mathematically equivalent to using the Hermitian rational Lanczos iteration. For numerical implementation it is advised to develop a method specifically for the Hermitian case.

## 9.3 Updating procedures - inner product

The connection between ORFs and certain structured pencils, a Hessenberg pencil with a particular ratio satisfied by its subdiagonal elements, can be exploited directly. Updating procedures do not explicitly use Krylov subspaces, they operate directly on the structure of the pencil. For IEPs with orthonormal bases the updating procedure uses unitary similarity transformations with *plane rotations*. Plane rotations are essentially $2 \times 2$ unitary matrices $P_i$ with parameters $a, b \in \mathbb{C}$, $|a|^2 + |b|^2 = 1$,

$$P_i := \begin{bmatrix} I_{i-1} & & & \\ & \bar{a} & & -\bar{b} \\ & & I_{m-i} & \\ & b & & a \end{bmatrix}. \tag{9.10}$$

The class of plane rotations $P_i$ is denoted by $\mathfrak{P}_i$.
A node-weight pair must be introduced, which is performed by embedding of the available solution and by multiplication with a plane rotation. This will perturb the recurrence pencil, it will no longer be a Hessenberg matrix (or Hessenberg pencil). Then the Hessenberg (pencil) structure is restored using a sequence of plane rotations. The origin of this idea can be traced back to Rutishauser [145]. Rutishauser used a similar procedure with plane rotations to express the product of two J-fractions as a J-fraction, this is a continued fraction related to Jacobi matrices. In the same conference proceedings Wilkinson [182] published a paper on the error analysis of operations with plane rotations. Gragg and Harrod were the first to apply Rutishauser's procedure to solve a Jacobi IEP [92]. Their procedure constructs the Jacobi matrix in a numerically stable manner, whereas the Lanczos iteration suffers from instability. Following the

success of this method, the same idea has been applied to solve many other IEPs, see e.g., [27, 45, 161].
An updating procedure for the Hessenberg IEP is discussed in Section 9.3.1. In Section 9.3.2 a novel, parallelizable variant of this updating procedure is proposed. Section 9.3.3 proposes a novel, numerically stable updating procedure for Hessenberg pencil IEPs. This novel procedure is published in [167].

### 9.3.1 Hessenberg IEP

Suppose a solution to Problem 9.7 of size $m$ is available, this is a proper Hessenberg matrix $H_m \in \mathbb{C}^{m\times m}$ and a unitary matrix $Q_m \in \mathbb{C}^{m\times m}$ satisfying

$$H_m = Q_m^H \Lambda Q_m, \quad \Lambda = \operatorname{diag}(\{z_i\}_{i=1}^m), \quad Q_m e_1 = v/\|v\|.$$

Given a new weight $\alpha_{m+1} \neq 0$ and node $z_{m+1} \notin \{z_i\}_{i=1}^m$, the updating problem consists of computing the Hessenberg matrix $H_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ and unitary matrix $Q_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ satisfying

$$H_{m+1} = Q_{m+1}^H \underbrace{\begin{bmatrix}\Lambda & \\ & z_{m+1}\end{bmatrix}}_{=:\tilde{\Lambda}} Q_{m+1}, \quad Q_{m+1}e_1 = \tilde{v}/\|\tilde{v}\|, \text{ with } \tilde{v} = \begin{bmatrix} v & \alpha_{m+1}\end{bmatrix}^\top.$$

The first step in the updating procedure is to embed the available solution of size $m$ into matrices of size $m+1$. The embedded matrices are denoted by $\hat{H}, \hat{Q} \in \mathbb{C}^{(m+1)\times(m+1)}$:

$$\hat{H} = \begin{bmatrix} H_m & \\ & z_{m+1}\end{bmatrix}, \quad \hat{Q} = \begin{bmatrix} Q_m & \\ & 1\end{bmatrix},$$

$\hat{H}$ is a Hessenberg matrix, $\hat{Q}^H\hat{Q} = I$ and $\hat{Q}^H\tilde{\Lambda}\hat{Q} = \hat{H}$. These matrices form the solution to a HIEP, but not the correct one, since $\hat{Q}e_1 = \begin{bmatrix} v & 0\end{bmatrix}^\top/\|v\| \neq \tilde{v}/\|\tilde{v}\|$. So, the first column of $\hat{Q}$ must be altered, which corresponds to introducing the weight $\alpha_{m+1}$ into the associated inner product.
An important note is that the matrices $Q_m$ and $\hat{Q}$ need not be known at any point in the procedure, only the Hessenberg matrix and the weights must be saved. In the following exposition, $Q_m$ and $\hat{Q}$ are used solely for clarity. The two steps of the updating procedure are now discussed in detail and afterwards an algorithm is proposed, implementing this procedure.

**Introduce weight** $\alpha_{m+1}$

The first column of the unitary matrix $Q_{m+1}$ must correspond to the normalized weight vector $\tilde{v}/\|\tilde{v}\|$ , i.e., $Q_{m+1}e_1 = \tilde{v}/\|\tilde{v}\|$. The plane rotation $P_1 \in \mathfrak{P}_1 \subset \mathbb{C}^{(m+1)\times(m+1)}$ enforces this condition. The value for $a$ and $b$ follow from $|a|^2 + |b|^2 = 1$ and

$$\hat{Q}P_1^H e_1 = \begin{bmatrix} | & \times & \dots & \times & 0 \\ \frac{v}{\|v\|} & \vdots & & \vdots & \vdots \\ | & \times & \dots & \times & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} a & & & \bar{b} \\ & & I_{m-1} & \\ -b & & & \bar{a} \end{bmatrix} e_1 = \begin{bmatrix} \frac{\bar{a}v}{\|v\|} \\ -b \end{bmatrix} = \tilde{v}/\|\tilde{v}\|.$$

The values are

$$\begin{cases} a = \frac{\sqrt{\|v\|^2 - \|v\|^2|\alpha_{m+1}|^2 + |\alpha_{m+1}|^2}}{\sqrt{\|v\|^2 + |\alpha_{m+1}|^2}} \\ b = -\frac{\alpha_{m+1}\|v\|}{\sqrt{\|v\|^2 + |\alpha_{m+1}|^2}} \end{cases} \tag{9.11}$$

and if $\|v\|^2 = 1$, then the expressions simplify to

$$\begin{cases} a = \frac{1}{\sqrt{1+|\alpha_{m+1}|^2}} \\ b = -\frac{v_{m+1}}{\sqrt{1+|\alpha_{m+1}|^2}} \end{cases} .$$

This step of the updating procedure is stable since multiplication with plane rotations is numerically stable [183, p.133]. The eigenvalue decomposition changes since the eigenvectors are altered by $P_1$,

$$\hat{Q}^H \tilde{\Lambda} \hat{Q} = \hat{H} \rightarrow P_1 \hat{Q}^H \tilde{\Lambda} \hat{Q} P_1^H = P_1 \hat{H} P_1^H.$$

The new eigenvectors are orthonormal, which follows from $(\hat{Q}P_1^H)^H(\hat{Q}P_1^H) = P_1 P_1^H = I$. And the eigenvalues are unaltered since a similarity transformation is performed. However, the matrix $\dot{H} := P_1 \hat{H} P_1^H$ is no longer a Hessenberg matrix. Figure 9.1 shows the structure of $\dot{H}$, which has Hessenberg structure except for the last row.

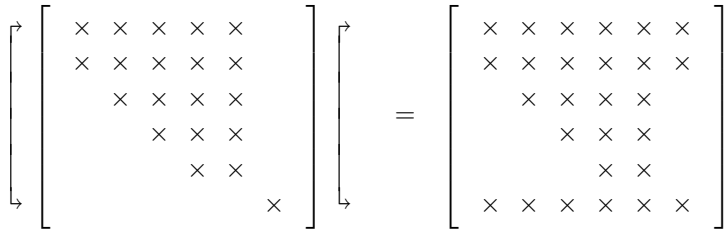


Figure 9.1: Structure of $P_1 \hat{H} P_1^H = \dot{H}$.

### Restore structure

The matrix $\dot{H} =: \dot{H}^{[1]}$ only differs from a Hessenberg matrix by its last row. Using unitary similarity transformations with plane rotations the elements in the last row can be eliminated one by one, thereby restoring the Hessenberg structure. These transformations should preserve the first column of $\hat{Q}P_1^H$. Without loss of generality we assume that the $k-1$ first elements in the last row have been eliminated by the product of plane rotations $\prod_{i=k}^{2} P_i = P_k P_{k-1} \dots P_2$. At step $k$ the matrix is $\dot{H}^{[k]} := \left(\prod_{i=k}^{1} P_i\right) \dot{H} \left(\prod_{i=k}^{1} P_i\right)^H$, the superscript $[k]$ reveals the step in the structure restoring process. Next $\dot{h}_{m+1,k}^{[k]}$ must be eliminated using $P_{k+1} \in \mathfrak{P}_{k+1}$ with parameters $a_{k+1}$ and $b_{k+1}$. The element $\dot{h}_{m+1,k}^{[k]}$ is eliminated by a sum with the element $\dot{h}_{k+1,k}^{[k]}$ which has not been altered by any of the previous transformations, thus $\dot{h}_{k+1,k}^{[k]} = h_{k+1,k}$. The parameters $a_{k+1}, b_{k+1}$ are determined such that

$$\begin{bmatrix} \bar{a}_{k+1} & -\bar{b}_{k+1} \\ b_{k+1} & a_{k+1} \end{bmatrix} \begin{bmatrix} h_{k+1,k} \\ \dot{h}_{m+1,k}^{[k]} \end{bmatrix} = \begin{bmatrix} \sqrt{|h_{i+1,i}|^2 + |\dot{h}_{m+1,k}^{[k]}|^2} \\ 0 \end{bmatrix},$$

which leads to

$$\begin{cases} a_{k+1} = \sqrt{\frac{|h_{k+1,k}|^2}{|h_{k+1,k}|^2 + |\dot{h}_{m+1,k}^{[k]}|^2}} \\ b_{k+1} = -\frac{a_{k+1}\dot{h}_{m+1,k}^{[k]}}{h_{k+1,k}} \end{cases}.$$

These parameters always exist because $H_m$ is a proper Hessenberg matrix, i.e., $h_{k+1,k} \neq 0$ for $k = 1, \dots, m-1$.

Figure 9.2 shows, for $m = 6$, the structure of $\dot{H}^{[k]}$ throughout the updating procedure. As illustrated in this figure, the Hessenberg structure is recovered after $m-2$ similarity transformations with plane rotations. Once the Hessenberg structure is restored, the resulting matrix $\dot{H}^{[m]} \in \mathbb{C}^{(m+1)\times(m+1)}$ is the solution to the updated IEP.
Similarity transformations with plane rotations are numerically stable [183, p.140]. Hence, restoring the Hessenberg structure is performed in a numerically stable manner.

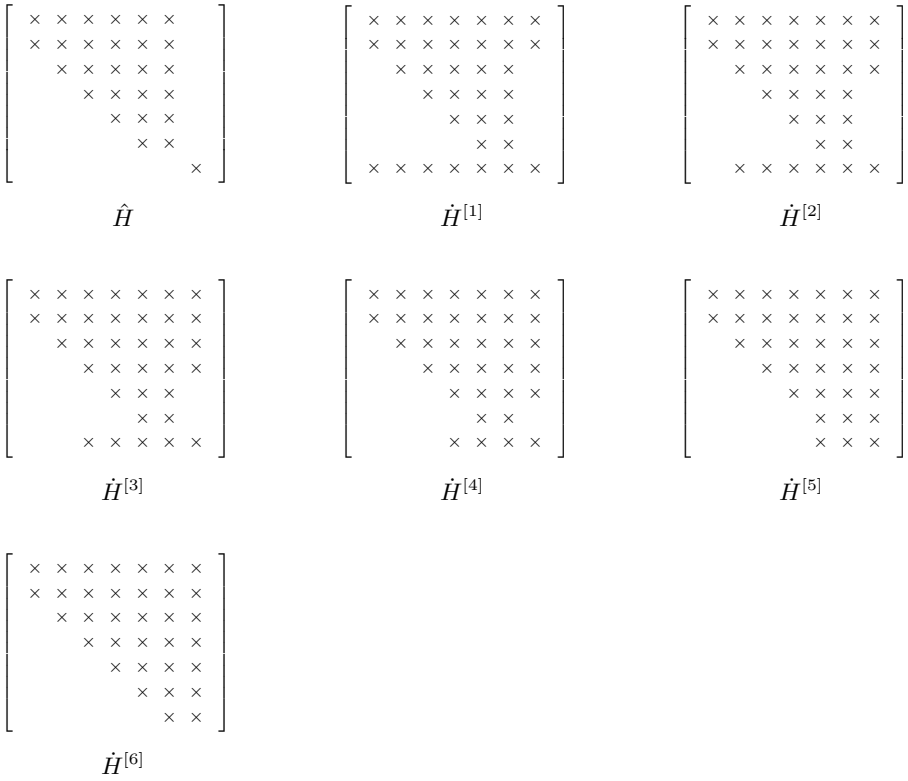



Figure 9.2: Updating strategy restoring Hessenberg structure.

## Algorithm

By combining the above steps a numerically stable algorithm to compute the solution of an updated HIEP is obtained. For simplicity we restrict to the case where $z_{m+1}$ is not in the spectrum of the available solution $H_m$, $z_{m+1} \notin \sigma(H_m)$. Algorithm 7 provides the implementation of the updating procedure for $z_{m+1} \notin \sigma(H_m)$. In this algorithm the notation $P(x, y, i)$ is used, which is a short notation for $P_i \in \mathfrak{P}_i$ with active parameters $a = x$ and $b = y$. Note that in Step 9 of Algorithm 7 the eliminated element $\dot{h}_{m+1,k}$ is set explicitly to zero.
If the nodes $z_i$ are all on the real line, then the updating procedure constructs a Jacobi matrix, this follows from $\Lambda = \Lambda^H$ and $Q^H Q = I$. Then Algorithm 7 corresponds (mathematically) to the procedure proposed by Gragg and Harrod [92]. Figure 9.3 shows the updating procedure for a Jacobi matrix, i.e., $z_i \in \mathbb{R}$. If $z_{m+1}$ does coincide with an eigenvalue of $H_m$, $z_{m+1} \in \sigma(H_m)$, say $z_{m+1} = z_k$, then the size of the problem

**Algorithm 7** Updating procedure HIEP

**Input:** Proper Hessenberg matrix $H_m \in \mathbb{C}^{m\times m}$, node $z_{m+1} \in \mathbb{C}$, $z_{m+1} \notin \sigma(H_m)$ and weight $\alpha_{m+1} > 0$.
**Output:** Proper Hessenberg matrix $H_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$, such that $\sigma(H_{m+1}) = \sigma(H_m) \cup z_{m+1}$.
**procedure** UPDATEHIEP($H_m, z_{m+1}, \alpha_{m+1}$)
    $P := P \begin{pmatrix} \frac{1}{\sqrt{1+|v_{m+1}|^2}}, \frac{\alpha_{m+1}}{\sqrt{1+|\alpha_{m+1}|^2}}, 1 \end{pmatrix}$
    $\dot{H} := P^H \begin{bmatrix} H_m & \\ & z_{m+1} \end{bmatrix} P$
    **for** $k = 1, \ldots, m-1$ **do**
        $P := P \begin{pmatrix} \sqrt{\frac{|h_{k+1,k}|^2}{|h_{k+1,k}|^2+|\dot{h}^{[k]}_{m+1,k}|^2}}, -\frac{a_{k+1}\dot{h}^{[k]}_{m+1,k}}{h_{k+1,k}}, k+1 \end{pmatrix}$
        $\dot{H} := P^H \dot{H} P$
        Set $\dot{h}_{m+1,k} := 0$
    **end for**
    $H_{m+1} := \dot{H}$
**end procedure**

remains the same, but the weight $w_k$ in the inner product $\langle ., . \rangle_m$ associated with $z_k$ will be altered by the updating procedure. With minor modifications the updating procedure can then be used as a downdating procedure [60], i.e., to remove a node-weight pair from an inner product. But this does not lead to a numerically stable downdating procedure so this idea will not be pursued further. Downdating procedures are discussed in Section 9.6.

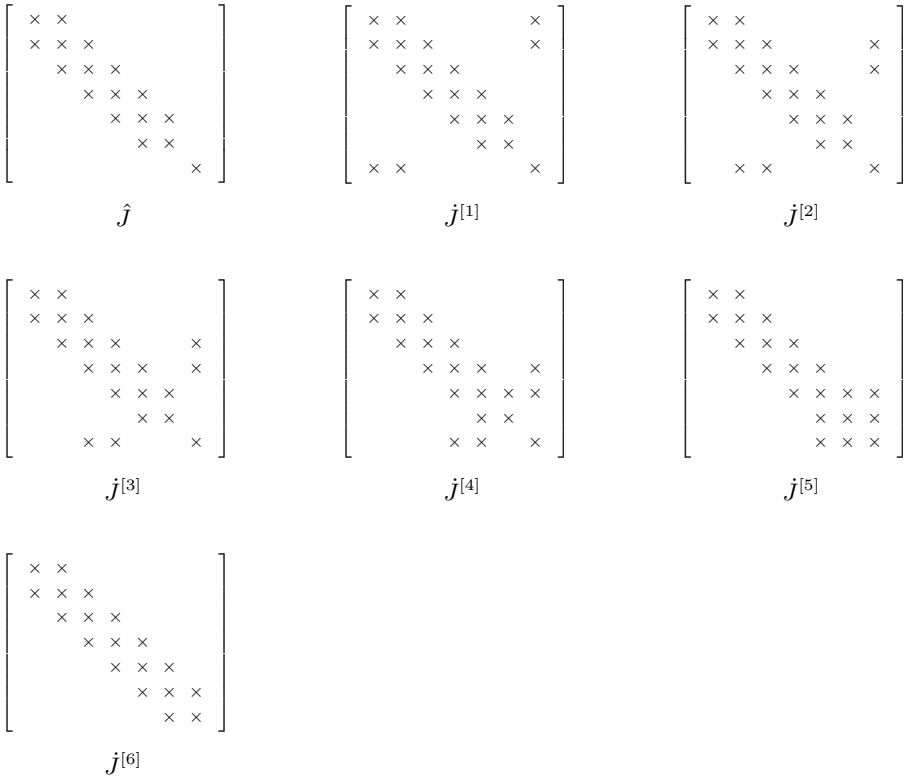



Figure 9.3: Updating strategy restoring tridiagonal (Jacobi) structure.

### 9.3.2 Hessenberg IEP - Recursive updating

A new updating procedure is proposed which allows for a parallel implementation, it will be called *recursive updating procedure* (RUP). The idea is as follows. Instead of adding one node $z_{m+1}$ to an available solution $H_m \in \mathbb{C}^{m\times m}$ and updating to obtain $H_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$, RUP starts from two available solutions $H_l \in \mathbb{C}^{l\times l}, \breve{H}_k \in \mathbb{C}^{k\times k}$ and merges them to obtain the solution $\tilde{H}_{k+l} \in \mathbb{C}^{(k+l)\times(k+l)}$.
More formally, consider $H_l, Q_l \in \mathbb{C}^{l\times l}$, the solution to a HIEP with

$$\Lambda = \mathrm{diag}(\{z\}_{i=1}^{l}), \quad \hat{Q}_l e_1 = \frac{v}{\|v\|_2},$$

where $z_i \neq z_j$ if $i \neq j$ and $e_i^\top v \neq 0$ for all $i$. Consider also, the solution $\breve{H}_k, \breve{Q}_l \in \mathbb{C}^{k\times k}$

$$\breve{\Lambda} = \mathrm{diag}(\{\breve{z}_{i=1}^{k}\}), \breve{Q}_k e_1 = \frac{\breve{v}}{\|\breve{v}\|_2},$$

with $\breve{z}_i \neq \breve{z}_j$ if $i \neq j$ and $e_i^\top \breve{v} \neq 0$ for all $i$.
The updated IEP of size $m = k + l$ consists of

$$\tilde{\Lambda} := \begin{bmatrix} \Lambda & \\ & \breve{\Lambda} \end{bmatrix}, \quad \tilde{v} := \begin{bmatrix} v \\ \breve{v} \end{bmatrix},$$

the corresponding solution is denoted by $\tilde{H}_m, \tilde{Q}_m \in \mathbb{C}^{m\times m}$. Assume, for simplicity that $\tilde{\Lambda}$ has distinct elements.
The matrices $Q_l \in \mathbb{C}^{l\times l}$ and $\breve{Q}_k \in \mathbb{C}^{k\times k}$ are used in the following discussion, but they are not required, only the norms of the weight vectors $v, \breve{v}$ are needed. The idea is the same as above: embed the matrices, introduce the correct weights and restore Hessenberg structure. The embedded matrices are

$$\hat{H} := \begin{bmatrix} H_l & \\ & \breve{H}_k \end{bmatrix}, \quad \hat{Q} := \begin{bmatrix} Q_l & \\ & \breve{Q}_k \end{bmatrix},$$

with $\hat{H}$ a Hessenberg matrix, $\hat{Q}^H\hat{Q} = I$ and $\hat{Q}^H\tilde{\Lambda}\hat{Q} = \hat{H}$. Clearly $\hat{Q}e_1 \neq \frac{\tilde{v}}{\|\tilde{v}\|_2}$, which can be enforced by a unitary transformation.

### Introduce weights

The weights appearing in the first column of $\breve{Q}_k$ must appear in the first column of the solution $\tilde{Q}_m$. This is accomplished by $\breve{P}_1 := P_1 \oplus I_{k-1}$, $P_1 \in \mathfrak{P}_1 \subset \mathbb{C}^{(l+1)\times(l+1)}$,

$$\breve{P}_1 = \begin{bmatrix} \bar{a} & & -\bar{b} & \\ & I_{l-1} & & \\ b & & a & \\ & & & I_{k-1} \end{bmatrix}, \quad \text{with } \begin{cases} a = \frac{\|v\|_2}{\sqrt{\|v\|_2+\|\breve{v}\|_2}} \\ b = -\frac{\|\breve{v}\|_2}{\sqrt{\|v\|_2+\|\breve{v}\|_2}} \end{cases}.$$

The matrix $\dot{Q} := \hat{Q}\breve{P}_1^H$ satisfies $\dot{Q}e_1 = \frac{\tilde{v}}{\|\tilde{v}\|_2}$. The eigenvalue decomposition changes

$$\hat{Q}^H\tilde{\Lambda}\hat{Q} = \hat{H} \to \breve{P}_1\hat{Q}^H\tilde{\Lambda}\hat{Q}\breve{P}_1^H = \breve{P}_1\hat{H}\breve{P}_1^H.$$

Set $\dot{H} := \breve{P}_1\hat{H}\breve{P}_1^H$, this matrix is no longer a Hessenberg matrix, its structure is shown in Figure 9.4. Note that in the first column of $\dot{H}$ there are two elements that must be annihilated.

### Restore structure

The matrix $\dot{H} =: \dot{H}^{[1]}$ can be brought to Hessenberg form by unitary similarity transformations. As before, plane rotations can be used or Householder reflectors [105, 117].

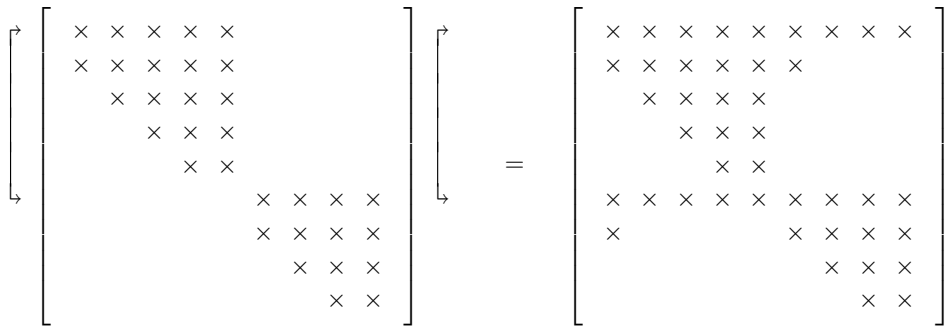


Figure 9.4: Structure of $\dot{H} = \check{P}_1 \hat{H} \check{P}_1^H$.

**Definition 9.1** ((Complex) Householder reflectors [117, Appendix B.1])**.** *Consider the vector $x \in \mathbb{C}^n$ and $\alpha = \pm \|x\|_2 e^{\imath \arg(x_1)} \in \mathbb{C}$, with $x_1 = xe_1$. The Householder reflector is the unitary, Hermitian matrix*

$$F = I - 2\frac{yy^H}{y^H y}, \quad \textit{with } y = x + \alpha e_1$$

*transforming $x$ into a multiple of $e_1$, i.e., $Fx = -\alpha e_1$.*

An algorithm implementing this is provided in Algorithm 8.

**Algorithm 8** Compute Householder reflector

**Input:** $x \in \mathbb{C}^n$.
**Output:** $F \in \mathbb{C}^{n \times n}$ such that $Fx = -\alpha e_1$, with $\alpha = \|x\|_2 e^{\imath \arg(xe_1)} \in \mathbb{C}$.
**procedure** HouseHolderReflector($x$)
    $\alpha := \|x\|_2 e^{\imath \arg(xe_1)}$
    $y := x + \alpha e_1$
    $F := I - 2\frac{yy^H}{y^H y}$
**end procedure**

If there are few nonzero elements in each column to be eliminated, then plane rotations are more efficient than Householder reflectors. The structure restoring process here will lead to an increase in the amount of nonzero elements per column. A preview of the structure restoring process is provided in Figure 9.5.

Hence, Householder reflectors become a viable alternative to plane rotations [85]. For ease of notation, the procedure for the Householder reflectors will be described. In

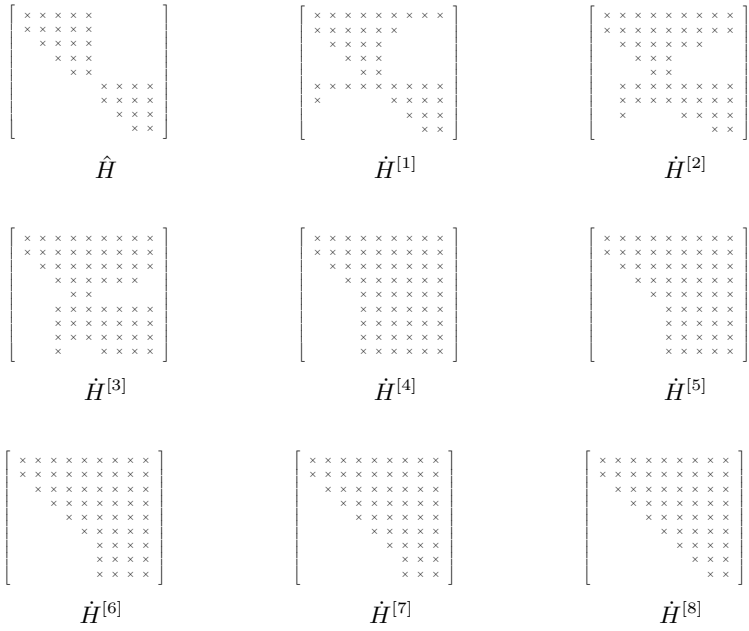



Figure 9.5: Recursive updating strategy for HIEP

the structure restoring process a Householder reflector $F$ is constructed such that $Fx = \alpha\|x\|_2 e_1$, with $|\alpha| = 1$. The vector $x$ is constructed from the subdiagonal element, used for elimination, and all nonzero elements to be eliminated. Throughout the updating, two cases can be distinguished:

- For column $i$, with $1 \leq i < l$, $n := \min\{k, i+1\}$ elements must be annihilated in this column. The relevant elements are grouped in the vector $x := \begin{bmatrix} h_{i+1,i} \\ \dot{h}^{[i]}_{l+1,i} \\ \dot{h}^{[i]}_{l+2,i} \\ \vdots \\ \dot{h}^{[i]}_{l+n,i} \end{bmatrix}$.
  Note that the first element of $x$ is an element of the original matrix $H_l$, and therefore nonzero.
- For column $i$, with $i \geq l$, $n := \min\{k, l+i\}$ and $n-1$ elements must be annihilated in this column. The relevant elements are grouped in the vector

$$x := \begin{bmatrix} \dot{h}^{[i]}_{i+1,i} \\ \dot{h}^{[i]}_{i+2,i} \\ \vdots \\ \dot{h}^{[i]}_{i+n,i} \end{bmatrix}.$$

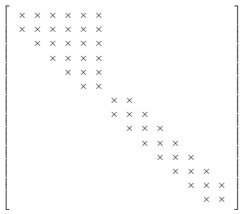
For $x \in \mathbb{C}^n$ a Householder reflector $\dot{F} \in \mathbb{C}^{n \times n}$ is constructed. Then $\dot{F}$ is correctly embedded in a matrix $F \in \mathbb{C}^{m \times m}$, such that $FH$ only influences the rows used to construct $x$. Details are provided in the algorithm.

### Algorithm

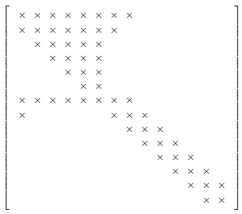
Algorithm 9 provides the recursive updating procedure for HIEPs. The Householder reflectors can be replaced by any sequence of plane transformations that reduces $x$ to $\alpha e_1$. Details for this approach are omitted, our code implementing Algorithm 9 can be found online [163]. Figure 9.6 shows the structure of the matrix throughout the updating procedure, on this figure $k > l$, whereas Figure 9.5 shows the case $l > k$.

**Algorithm 9** Recursive updating procedure

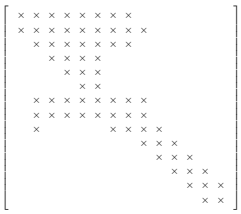
**Input:** Proper, simple Hessenberg matrices $H_l \in \mathbb{C}^{l\times l}, \breve{H}_k \in \mathbb{C}^{k\times k}$, with $\sigma(H_l) \cap \sigma(\breve{H}_k) = \emptyset$ and $\eta, \nu$ the normalization of the weight vectors, i.e., $\eta = \|v\|_2$ and $\nu = \|\breve{v}\|_2$.
**Output:** $\tilde{H} \in \mathbb{C}^{(k+l)\times(k+l)}$, with $\sigma(H) = \sigma(H_l) \cup \sigma(\breve{H}_k)$.
**procedure** RUP$(H_l, \breve{H}_k, \eta, \nu)$
    $H := \begin{bmatrix} H_l & \\ & \breve{H}_k \end{bmatrix}$
    $\dot{C} :=$ CoreTransformation$(\eta, \nu)$
    $C := I$, $C([1, l+1], [1, l+1]) := \dot{C}$
    $H := CHC^H$ ▷ Introduce weights
    **for** $i = 1, 2, \ldots, l-1$ **do** ▷ Columns 1 through $l-1$
        $n := \min\{k, i+1\}$
        $x := \begin{bmatrix} h_{i+1,i} & h_{l+1,i} & h_{l+2,i} & h_{l+n,i} \end{bmatrix}^\top$
        $\dot{F} :=$ HouseHolderReflector$(x)$
        $F := I$, $F([i+1, l+1 : l : n], i+1) := \dot{F}([1 : n+1], 1)$, $F([i+1, l+1 : l+n], [l+1 : l+n]) := \dot{F}([1 : n+1], [2 : n+1])$
        $H = FHF^H$
        $H([l+1, l+n], i) = 0$
    **end for**
    $n := l + k$
    **for** $i = l, l+1, \ldots, l+k-2$ **do** ▷ Columns $l$ through $k+l-2$
        **if** $l + i + 1 \leq l + k$ **then**
            $n := l + i + 1$
        **end if**
        $x := \begin{bmatrix} h_{i+1,i} & h_{i+2,i} & \ldots & h_{n,i} \end{bmatrix}^\top$
        $\dot{F} :=$ HouseHolderReflector$(x)$
        $F := I$, $F([i+1, n], [i+1, n]) := \dot{F}$
        $H = FHF^H$
        $H([i+2 : n], i) = 0$
    **end for**
    $\tilde{H} := H$
**end procedure**



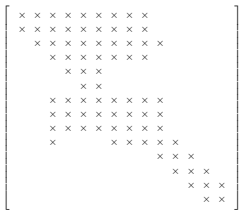
$\hat{H}$



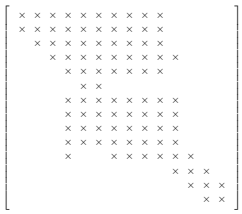
$\dot{H}^{[1]}$



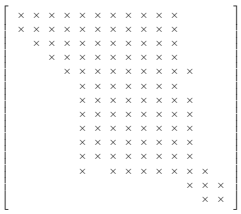
$\dot{H}^{[2]}$



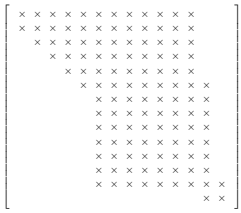
$\dot{H}^{[3]}$



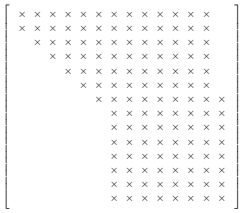
$\dot{H}^{[4]}$



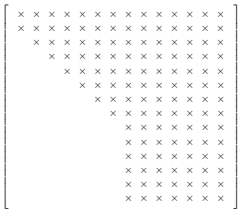
$\dot{H}^{[5]}$



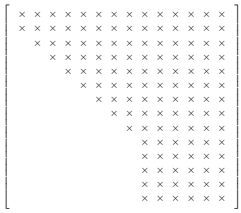
$\dot{H}^{[6]}$



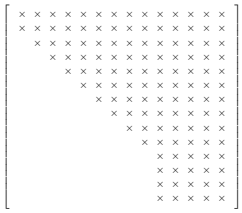
$\dot{H}^{[7]}$



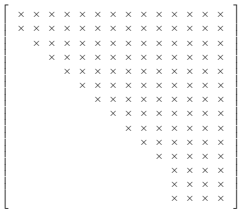
$\dot{H}^{[8]}$



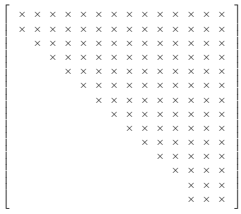
$\dot{H}^{[9]}$



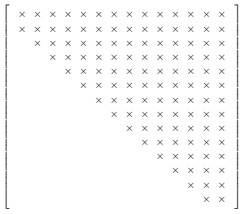
$\dot{H}^{[10]}$



$\dot{H}^{[11]}$



$\dot{H}^{[12]}$



$\dot{H}^{[13]}$

Figure 9.6: Hessenberg recursive updating, trailing submatrix larger than leading submatrix with Householder

### 9.3.3 Hessenberg pencil IEP

An updating procedure for a HPIEP, formulated in Problem 9.6, is similar in spirit to the updating discussed in Section 9.3.1, with the essential difference that a pencil must be manipulated. This especially influences the step which restores the structure of the recurrence pencil.
The procedure starts from a unitary matrix $Q_m \in \mathbb{C}^{m\times m}$ and a proper Hessenberg pencil $(H_m, K_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$, which form the solution to a HPIEP. That is, the following equalities are satisfied

$$Q_m^H \Lambda Q_m K_m = H_m, \quad Q_m e_1 = \frac{v}{\|v\|}, \quad \frac{h_{i+1,i}}{k_{i+1,i}} = \xi_i,\ i = 1, \ldots, m-1. \tag{9.12}$$

From this available solution it constructs the Hessenberg pencil $(H_{m+1}, K_{m+1}) \in \mathbb{C}^{(m+1)\times(m+1)} \times \mathbb{C}^{(m+1)\times(m+1)}$ and orthonormal basis $Q_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ which solve the updated HPIEP. The HPIEP is updated with the node-weight pair $z_{m+1}, \alpha_{m+1}$ and pole $\xi_m$. That is, the updated solution must satisfy

$$Q_{m+1}^H \underbrace{\begin{bmatrix} \Lambda & \\ & z_{m+1} \end{bmatrix}}_{=:\tilde{\Lambda}} Q_{m+1} K_{m+1} = H_{m+1}, \quad \frac{h_{i+1,i}}{k_{i+1,i}} = \xi_i,\ i = 1, \ldots, m-1, m,$$

$$Q_{m+1} e_1 = \frac{\tilde{v}}{\|\tilde{v}\|}, \quad \text{with } \tilde{v} := \begin{bmatrix} v & \alpha_{m+1} \end{bmatrix}^\top.$$

First, the $m$-dimensional solution is embedded in matrices of size $m+1$, while preserving some key properties. The embedded matrices $\hat{Q}, \hat{H}, \hat{K} \in \mathbb{C}^{(m+1)\times(m+1)}$ should satisfy $\hat{Q}^H \hat{Q} = I$, $\hat{Q}^H \tilde{\Lambda} \hat{Q} \hat{K} = \hat{H}$, and $\hat{H}, \hat{K}$ must be Hessenberg matrices:

$$\widehat{Q} := \begin{bmatrix} Q & \\ & 1 \end{bmatrix}, \quad \widehat{H} := \begin{bmatrix} H & \\ & \hat{h} \end{bmatrix}, \quad \widehat{K} := \begin{bmatrix} K & \\ & \hat{k} \end{bmatrix}, \text{ with } z_{m+1}\hat{k} = \hat{h}. \tag{9.13}$$

Since the procedure proposed in the sequel is new, a formal statement in Theorem 9.3 describes the form of the solution. The update requires only $2m$ plane rotations to update an available solution of size $m$ to size $m+1$.

**Theorem 9.3.** *Let $Q_m \in \mathbb{C}^{m\times m}$, $(H_m, K_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ be the solution to Problem 9.6 of size $m$. And let $\widehat{Q}$, $\widehat{H}$ and $\widehat{K}$ denote the embedded matrices from* (9.13)*. Then there exist $P_i, \dot{P}_i \in \mathfrak{P}_i$, $i = 1, \ldots, m$, such that*

$$H_{m+1} = \prod_{l=m}^{1} P_l \widehat{H} \prod_{k=1}^{m} \dot{P}_k, \quad K_{m+1} = \prod_{l=m}^{1} P_l \widehat{K} \prod_{k=1}^{m} \dot{P}_k, \quad \widetilde{Q} = Q_{m+1} \prod_{l=1}^{m} P_l^H$$

*solve a HPIEP of size $m+1$, which is obtained by adding a node $z_{m+1}$, weight $\alpha_{m+1}$ and pole $\xi_m$ to the original IEP.*

The remainder of the section is dedicated to proving results that can be combined to obtain this theorem. The proofs of these results are constructive in the sense that they describe how the plane rotations in Theorem 9.3 must be computed. Remarks on stability are included outside of the proofs, which clarify how to compute the plane rotations in the most numerically stable way. Afterwards these results are combined to propose a numerically stable algorithm that implements the updating procedure.

### Introducing weight $\alpha_{m+1}$

The introduction of the weight in the first column of $\hat{Q}$ is the same as described for the HIEP updating. Postmultiplication with a plane rotation $P_1^H \in \mathfrak{P}_1 \subset \mathbb{C}^{(m+1)\times(m+1)}$ with parameters (9.11) has the desired effect. The generalized eigenvalue decomposition changes under this change of eigenvector basis,

$$\hat{Q}^H \tilde{\Lambda} \hat{Q} \hat{K} = \hat{H} \rightarrow P_1 \hat{Q}^H \tilde{\Lambda} \hat{Q} P_1^H P_1 \hat{K} = P_1 \hat{H}.$$

The structure of the pencil is perturbed in the last row of both matrices:

$$P_1\hat{H} = \begin{bmatrix} \times & \times & \dots & \times & \times & \times \\ \times & \times & \dots & \times & \times & 0 \\ & \times & \dots & \times & \times & 0 \\ & & \ddots & \vdots & \vdots & \vdots \\ & & & \times & \times & 0 \\ \times & \times & \dots & \times & \times & \times \end{bmatrix}, \qquad P_1\hat{K} = \begin{bmatrix} \times & \times & \dots & \times & \times & \times \\ \times & \times & \dots & \times & \times & 0 \\ & \times & \dots & \times & \times & 0 \\ & & \ddots & \vdots & \vdots & \vdots \\ & & & \times & \times & 0 \\ \times & \times & \dots & \times & \times & \times \end{bmatrix}.$$

### Restore Hessenberg pencil structure

The following step is to restore the Hessenberg pencil structure using unitary similarity transforms on the pencil $(P_1\hat{H}, P_1\hat{K}) = (\dot{H}^{[1]}, \dot{K}^{[1]})$. The elements in the last row of both matrices appearing in the pencil must be eliminated simultaneously. Lemma 9.3 provides the details and stresses an important property, namely, the poles present in the pencil are preserved under transformations that restore the structure. The poles must be preserved since these characterize the connection between the pencil and ORFs with these poles.

**Lemma 9.3.** *Consider the embedded pencil* $(\widehat{H}, \widehat{K}) \in \mathbb{C}^{(m+1)\times(m+1)} \times \mathbb{C}^{(m+1)\times(m+1)}$ (9.13) *and* $P_1$ *with parameters* (9.11)*. There exist* $P_l \in \mathfrak{P}_l$*,* $l = 2, 3, \dots, m$ *and* $\dot{P}_k \in \mathfrak{P}_k$*,* $k = 1, 2, \dots, m-1$ *such that for* $X := \prod_{l=m}^{1} P_l$ *and* $\dot{X} := \prod_{j=1}^{m-1} \dot{P}_j$*,*

$$\dot{H}^{[m]} := X\widehat{H}\dot{X} \text{ and } \dot{K}^{[m]} := X\widehat{K}\dot{X}$$

*are Hessenberg matrices with* $\frac{\dot{h}_{i+1,i}}{\dot{k}_{i+1,i}} = \frac{\hat{h}_{i+1,i}}{\hat{k}_{i+1,i}}$*,* $i = 1, \dots, m-1$.

*Proof.* Starting from $(\dot{H}^{[1]}, \dot{K}^{[1]})$, the elements of the last row are annihilated using plane rotations. Assume, without loss of generality, that the Hessenberg structure is restored up to column $i$.
Set $X^{[i+1]} := \prod_{l=i+1}^{1} P_l$ and $\dot{X}^{[i]} := \prod_{j=i}^{1} P_j$, then the first $i$ columns of $\dot{H}^{[i+1]} := X^{[i+1]}\hat{H}\dot{X}^{[i]}$ and $\dot{K}^{[i+1]} := X^{[i+1]}\widehat{K}\dot{X}^{[i]}$ have Hessenberg structure. Let $\gamma := \dot{h}^{[i+1]}_{m+1,i+1}$, $\alpha := \dot{k}^{[i+1]}_{m+1,i+1}$, $\delta := \dot{h}^{[i+1]}_{i+2,i+1}$, $\beta := \dot{k}^{[i+1]}_{i+2,i+1}$, $\eta := \dot{h}^{[i+1]}_{m+1,m+1}$ and $\epsilon := \dot{k}^{[i+1]}_{m+1,m+1}$, then

$$\dot{H}^{[i+1]} =: \begin{bmatrix} \widetilde{H}^{(i+1)\times i} & \times & M_H & \times \\ & \delta e_1 & B_H & \mathbf{0} \\ & \gamma & \times^\top & \eta \end{bmatrix}, \dot{K}^{[i+1]} =: \begin{bmatrix} \widetilde{K}^{(i+1)\times i} & \times & M_K & \times \\ & \beta e_1 & B_K & \mathbf{0} \\ & \alpha & \times^\top & \epsilon \end{bmatrix}. \tag{9.14}$$

In the matrices from (9.14):

- $\widetilde{H}^{(i+1)\times i}$, $\widetilde{K}^{(i+1)\times i}$ are Hessenberg matrices of size $(i+1)\times i$. These will not be altered in subsequent steps of the structure restoring process.
- $M_H, M_K \in \mathbb{C}^{(i+1)\times(m-i-1)}$ are (generically) full matrices.
- $B_H, B_K \in \mathbb{C}^{(m-i-1)\times(m-i-1)}$ are Hessenberg matrices.
- the zero vectors, denoted by $\mathbf{0}$, and vectors containing generic nonzero elements, denoted by $\times$, are assumed to be of appropriate size.

The elements $\alpha, \gamma$ must be eliminated, this is achieved in two steps. Since plane rotations are used, the relevant elements can be isolated in an equivalent $2\times 2$ problem, i.e. find parameters $a, b$ and $c, d$ appearing in $P_{i+2}$ and $\dot{P}_{i+1}$, respectively, such that

$$P^b_{i+2}H^b\dot{P}^b_{i+1} := \begin{bmatrix} \bar{a} & -\bar{b} \\ b & a \end{bmatrix}\begin{bmatrix} \delta & 0 \\ \gamma & \eta \end{bmatrix}\begin{bmatrix} \bar{c} & -\bar{d} \\ d & c \end{bmatrix} = \begin{bmatrix} h & \times \\ 0 & \times \end{bmatrix},$$

$$P^b_{i+2}K^b\dot{P}^b_{i+1} := \begin{bmatrix} \bar{a} & -\bar{b} \\ b & a \end{bmatrix}\begin{bmatrix} \beta & 0 \\ \alpha & \epsilon \end{bmatrix}\begin{bmatrix} \bar{c} & -\bar{d} \\ d & c \end{bmatrix} = \begin{bmatrix} k & \times \\ 0 & \times \end{bmatrix}$$

with $\frac{h}{k} = \frac{\delta}{\beta} = \xi_{i+1}$ and the superscript $b$, for *block*. First $\dot{P}^b_{i+1}$ is constructed such that $M\dot{P}^b_{i+1}e_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}^\top$ with

$$M := \delta K^b - \beta H^b = \begin{bmatrix} 0 & 0 \\ \times & \times \end{bmatrix}.$$

This results in $\operatorname{rank}\left(\begin{bmatrix} H^b\dot{P}^b_{i+1}e_1 & K^b\dot{P}^b_{i+1}e_1 \end{bmatrix}\right) = 1$, i.e., the first column of $H^b\dot{P}^b_{i+1}$ and the first column $K^b\dot{P}^b_{i+1}$ are colinear (also called parallel). Second, $P^b_{i+2}$ is chosen to

make $H^b\dot{P}^b_{i+1}$ (or $K^b\dot{P}^b_{i+1}$) upper triangular, i.e., $P^b_{i+2}H^b\dot{P}^b_{i+1}e_1 = \begin{bmatrix} k & 0 \end{bmatrix}^\top$. Thanks to the colinearity, the same plane rotation $P^b_{i+2}$ results in $P^b_{i+2}H^b\dot{P}^b_{i+1} = \begin{bmatrix} h & 0 \end{bmatrix}^\top$. For some nonzero constant $u$, $h = u\delta$ and $k = u\beta$ and therefore the ratio of subdiagonal elements $\frac{h}{k} = \frac{u\delta}{u\beta} = \xi_{i+1}$ is preserved. By a simple rank argument on the $2 \times 2$ matrices involved it can be shown that the elements $e_2^\top P^b_{i+2}H^b\dot{P}^b_{i+1}e_2$ and $e_2^\top P^b_{i+2}K^b\dot{P}^b_{i+1}e_2$ are nonzero. Hence, this process can be repeated until $i = m-2$. If $\delta = 0$ (or $\beta = 0$, never both), then $\operatorname{rank}(H^b) = 1$ (or $\operatorname{rank}(K^b) = 1$) and the rank argument no longer holds, however straightforward computation shows that, in this case, $e_2^\top P^b_{i+2}H^b\dot{P}^b_{i+1}e_1 \neq 0$ (or $e_2^\top P^b_{i+2}K^b\dot{P}^b_{i+1}e_1 \neq 0$). □

Note that $\dot{P}_l$, $l = 1, 2, \ldots, m-1$ can be replaced by nonsingular matrices in the above proof. However unitary matrices are preferred for numerical computations.
Lemma 9.3 guarantees that a Hessenberg pencil can be obtained via unitary similarity transformations. Figure 9.7 shows the structures occurring in the updating procedure for one of the matrices in the pencil, for the other matrix the occurring structures are the same.

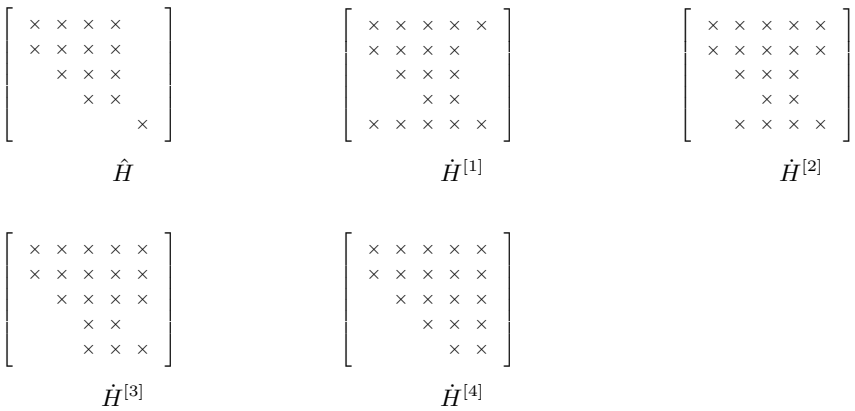



Figure 9.7: Updating strategy restoring Hessenberg pencil structure.

Thus we have a Hessenberg pencil $(\dot{H}, \dot{K})$ with $\sigma(\dot{H}, \dot{K}) = \{z_i\}_{i=1}^{m+1}$. However the ratio of their last subdiagonal elements will not necessarily be equal to the new pole $\xi_m$.

### Introduce new pole $\xi_m$

The pole $\xi_m$ will be introduced by postmultiplying $\dot{H}^{[m]}$ and $\dot{K}^{[m]}$ with a plane rotation in $\mathfrak{P}_m$. Postmultiplication of $\dot{H}^{[m]}$ and $\dot{K}^{[m]}$ with a nonsingular matrix preserves the eigenvalues of the pencil $(\dot{H}^{[m]}, \dot{K})^{[m]}$. Lemma 9.4 provides a formal statement.

**Lemma 9.4.** *Consider $\xi_m \in \overline{\mathbb{C}}$ and the matrices from Lemma 9.3,*

$$\dot{H}^{[m]} = \prod_{l=m}^{1} P_l \widehat{H} \prod_{k=1}^{m-1} \dot{P}_k, \qquad \dot{K}^{[m]} = \prod_{l=m}^{1} P_l \widehat{K} \prod_{k=1}^{m-1} \dot{P}_k. \tag{9.15}$$

*Then $\dot{P}_m \in \mathfrak{P}_m$, with parameters $c, d$, exists such that, for $\tilde{H} := \dot{H}^{[m]}\dot{P}_m$ and $\tilde{K} := \dot{K}^{[m]}\dot{P}_m$, the ratio $\frac{\tilde{h}_{m+1,m}}{\tilde{k}_{m+1,m}} = \xi_m$.*

*Proof.* The matrix $\dot{P}_m = \begin{bmatrix} I_{m-1} & & \\ & \bar{c} & -\bar{d} \\ & d & c \end{bmatrix}$ acts only on the last two columns of $\dot{H}^{[m]}$ and $\dot{K}^{[m]}$ and its parameters $c, d$ can be determined from the equations

$$\begin{cases} c\bar{c} + d\bar{d} = 1 \\ \frac{\bar{c}\dot{h}^{[m]}_{m+1,m} + d\dot{h}^{[m]}_{m+1,m+1}}{\bar{c}\dot{k}^{[m]}_{m+1,m} + d\dot{k}^{[m]}_{m+1,m+1}} = \xi \in \mathbb{C}. \end{cases}$$

For $\xi = \infty$, the second expression becomes $\bar{c}\dot{k}^{[m]}_{m+1,m} + d\dot{k}^{[m]}_{m+1,m+1} = 0$. □

Introducing the pole is performed by multiplication with a plane rotation, which is numerically stable [183, p.133].

### Algorithm

A numerically stable algorithm can be developed which computes $(\tilde{H}, \tilde{K})$ (and optionally $\tilde{Q}$) appearing in Theorem 9.3. Introducing the new weight and new pole is performed by plane rotations and is numerically stable. For the structure restoring process, a pencil is manipulated and in the algorithm the numerically computed annihilated elements will be set explicitly to zero. We must verify whether this is numerically stable, i.e., verify that the annihilated element is small enough so that setting it to zero does not compromise numerical stability.
The error analysis for this has, in fact, been performed in the context of computing deflating subspaces [38, 171]. These results are applicable here, we repeat the criterion leading to the most numerically stable procedure. this criterion dictates whether to use

$\dot{H}^{[i]}\dot{P}_i$ or $\dot{K}^{[i]}\dot{P}_i$ to compute $P_{i+1}$ in the proof of Lemma 9.3. The most numerically stable implementation [38] is obtained by, using the notation of the proof of Lemma 9.3,

$$\text{compute } P_{i+1} \text{ from } \begin{cases} \prod_{l=i}^{1} P_l \widehat{H} \prod_{k=1}^{i} \dot{P}_k, & \text{if } \frac{\eta}{\epsilon} < \frac{\delta}{\alpha} \\ \prod_{l=i}^{1} P_l \widehat{K} \prod_{k=1}^{i} \dot{P}_k, & \text{else} \end{cases} . \tag{9.16}$$

Algorithm 10 provides all the details, criterion (9.16) appears on Step 10 of this algorithm.

**Algorithm 10** Updating procedure HPIEP

1: **Input:** Proper Hessenberg pencil $(H_m, K_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$, node $z_{m+1} \in \mathbb{C}$, $z_{m+1} \notin \sigma(H_m)$ and weight $\alpha_{m+1} \neq 0$.
2: **Output:** Proper Hessenberg pencil $(H_{m+1}, K_{m+1}) \in \mathbb{C}^{(m+1)\times(m+1)} \times \mathbb{C}^{(m+1)\times(m+1)}$, such that $\sigma(H_{m+1}) = \sigma(H_m) \cup z_{m+1}$.
3: Assumption: $\|v\| = 1$.
4: **procedure** UPDATEHPIEP($H_m, K_m, z_{m+1}, \alpha_{m+1}, \xi_m$)
5: $P := P\left(\frac{1}{\sqrt{1+|\alpha_{m+1}|^2}}, -\frac{\alpha_{m+1}}{\sqrt{1+|\alpha_{m+1}|^2}}, 1\right)$
6: $H := P^H \begin{bmatrix} H_m & \\ & z_{m+1} \end{bmatrix} P$, $K := P^H \begin{bmatrix} K_m & \\ & 1 \end{bmatrix} P$
7: **for** $i = 1, \ldots, m-1$ **do**
8: $z = -\frac{h_{i+1,i}k_{m+1,m+1} - k_{i+1,i}h_{m+1,m+1}}{h_{i+1,i}k_{m+1,i} - k_{i+1,i}h_{m+1,i}}$
9: $\dot{P} := P\left(-\bar{z}\left(\sqrt{(1+|z|^2)}\right)^{-1}, \sqrt{1+|z|^2}^{-1}, i\right)$
10: **if** $\frac{h_{m+1,m+1}}{k_{m+1,m+1}} < \frac{h_{i+1,i}}{k_{i+1,i}}$ **then**
11: $H = H\dot{P}, \quad K = K\dot{P}$
12: $a := \sqrt{\frac{|h_{i+1,i}|^2}{|h_{i+1,i}|^2} + |h_{m+1,i}|^2}$
13: $P := P(a, -ah_{m+1,i}/h_{i+1,i}, i+1)$
14: **else**
15: $H := H\dot{P}, \quad K := K\dot{P}$
16: $a := \sqrt{\frac{|k_{i+1,i}|^2}{|k_{i+1,i}|^2} + |k_{m+1,i}|^2}$
17: $P := P(a, -ak_{m+1,i}/k_{i+1,i}, i+1)$
18: **end if**
19: $H := PH, \quad K := PK$
20: Set $h_{m+1,i} := 0$ and $k_{m+1,i} := 0$
21: **end for**
22: $H_{m+1} := H$ and $K_{m+1} := K$
23: **end procedure**

The elements in the last rows of the pencil, which $P_{i+1}$ should eliminate, are set explicitly to zero in Step 18 of the algorithm.
Since the annihilated elements in the last rows are set explicitly to zero, the resulting structure is exactly a Hessenberg pencil and therefore corresponds exactly to a structured matrix pencil containing recurrence coefficients of rational functions with prescribed poles.

## 9.4 Updating procedures - linear functional

Updating procedures for IEPs related to biOPs and biORFs can, in general, not make use of unitary similarity transformations. Instead of unitary plane rotations, nonsingular *eliminators* will be used, which are essentially $2 \times 2$ triangular matrices. Let $\mathfrak{L}_i$ denote the class of lower triangular eliminators and $\mathfrak{R}_i$ the class of upper triangular eliminators. The class $\mathfrak{L}_i \subset \mathbb{C}^{(m+1)\times(m+1)}$ and $\mathfrak{R}_i \subset \mathbb{C}^{(m+1)\times(m+1)}$ are composed, respectively, of matrices of the form

$$L_i := \begin{bmatrix} I_{i-1} & & & \\ & 1 & & \\ & & I_{m-i} & \\ & l_i & & 1 \end{bmatrix} \text{ and } R_i := \begin{bmatrix} I_{i-1} & & & \\ & 1 & & r_i \\ & & I_{m-i} & \\ & & & 1 \end{bmatrix}, \tag{9.17}$$

with parameters $l_i, r_i \in \mathbb{C}$. Breakdowns can occur in the updating procedures proposed in this section. For simplicity of the exposition we assume that no breakdowns occur, some remarks will be given regarding breakdowns.
Property 9.1 is important in the following discussion.

**Property 9.1** (Inverse of eliminators)**.** *If $L_i \in \mathfrak{L}_i$, then $L_i^{-1} \in \mathfrak{L}_i$ and if $R_i \in \mathfrak{R}_i$, then $R_i^{-1} \in \mathfrak{R}_i$.*

An updating procedure for TIEPs is discussed in Section 9.4.1, showing how eliminators can restore structure in the upper and lower triangular part of a matrix. Then this is generalized to TPIEPs in Section 9.4.2.

### 9.4.1 Tridiagonal IEP

An updating procedure for Problem 9.9, the (non-Hermitian) tridiagonal IEP is proposed here. This procedure could provide an alternative for or complement the Lanczos algorithms that are developed for pseudo-Jacobi matrices [9,185,186]. Pseudo-Jacobi matrices have additional structure that must be taken into account. We will propose an updating procedure for general tridiagonal matrices, the special case of pseudo-Jacobi matrices is not discussed.

A solution to a TIEP of size $m$ is available, a tridiagonal matrix $T_m \in \mathbb{C}^{m\times m}$ and pair of biorthonormal matrices $V_m, W_m \in \mathbb{C}^{m\times m}$ satisfying

$$T_m = W_m^H \Lambda V_m, \quad V_m e_1 = v/\eta, \quad W_m e_1 = w/\nu,$$

with $\Lambda = \operatorname{diag}(\{z_i\}_{i=1}^m)$ and $w^H v = \bar{\nu}\eta \neq 0$.
Given new nonzero weights $\alpha_{m+1}, \beta_{m+1}$ and a node $z_{m+1} \notin \{z_i\}_{i=1}^m$, compute the tridiagonal matrix $T_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ and the biorthonormal pair $V_{m+1}, W_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ satisfying

$$T_{m+1} = W_{m+1}^H \underbrace{\begin{bmatrix}\Lambda & \\ & z_{m+1}\end{bmatrix}}_{=:\tilde{\Lambda}} V_{m+1}, \quad V_{m+1}e_1 = \tilde{v}/\eta, \quad W_{m+1}e_1 = \tilde{w}/\nu$$

with $\tilde{v} = \begin{bmatrix} v & \alpha_{m+1}\end{bmatrix}^\top$ and $\tilde{w} = \begin{bmatrix} w & \beta_{m+1}\end{bmatrix}^\top$. These matrices are embedded in $\mathbb{C}^{(m+1)\times(m+1)}$, preserving tridiagonality for $\hat{T}$, biorthonormality for $\hat{V}, \hat{W}$ and satisfying the equation $\hat{W}^H \tilde{Z} \hat{V} = \hat{T}$:

$$\hat{T} = \begin{bmatrix} T_m & \\ & z_{m+1}\end{bmatrix}, \quad \hat{V} = \begin{bmatrix} V_m & \\ & 1\end{bmatrix}, \quad \hat{W} = \begin{bmatrix} W_m & \\ & 1\end{bmatrix}.$$

Similarly as the updating procedure for HIEPs, first the new weights are introduced in the first columns of the biorthonormal pair $\hat{V}, \hat{W}$. This leads to a perturbation of the structure of the recurrence matrix. The tridiagonal structure is then restored with similarity transformations using eliminators.

### Introduce weights $\alpha_{m+1}, \beta_{m+1}$

The first columns of the biorthonormal pair $V_{m+1}, W_{m+1}$ must contain the normalized weight vectors $\tilde{v}/\tilde{\eta}, \tilde{w}/\tilde{\nu}$, respectively, where $\langle \tilde{v}/\tilde{\eta}, \tilde{w}/\tilde{\nu}\rangle_E = 1$. Eliminators $L_1 \in \mathfrak{L}_1$ and $R_1 \in \mathfrak{R}_1$ are constructed such that they introduce the new weights in the first columns

$$\hat{V} L_1 R_1 e_1 = \tilde{v}/\tilde{\eta} \quad \text{and} \quad \hat{W}(L_1)^{-H}(R_1)^{-H} e_1 = \tilde{w}/\tilde{\nu}.$$

Set $\dot{V} := \hat{V} L_1 R_1$ and $\dot{W} := \hat{W}(L_1)^{-H}(R_1)^{-H}$. The above expressions preserve biorthonormality, $\dot{W}^H \dot{V} = I$. The parameters $l_1, r_1$ in $L_1, R_1$, introducing the weights, are

$$\begin{cases} l_1 = \frac{\alpha_{m+1}}{\eta} \\ r_1 = -\frac{\bar{\beta}_{m+1}}{\bar{\nu} + l_1 \bar{\beta}_{m+1}} = \frac{\bar{\beta}_{m+1}}{\eta\bar{\nu} + \bar{\beta}_{m+1}\alpha_{m+1}} \end{cases}. \tag{9.18}$$

The resulting first columns are

$$\hat{V} L_1 R_1 e_1 = \begin{bmatrix} | \\ \frac{v}{\eta} \\ | \\ l_1 \end{bmatrix}, \qquad \hat{W}(L_1)^{-H}(R_1)^{-H} e_1 = \begin{bmatrix} | \\ \left(1 + \bar{r}_1 \bar{l}_1\right) \frac{w}{\nu} \\ | \\ -\bar{r}_1 \end{bmatrix}.$$

For numerical computation it can be interesting to be able to scale the resulting first columns, this is possible by multiplying with $L_1 D_1 R_1$, where $D_1$ differs from the identity matrix only in its first and last element. Details are provided in Appendix B.2.
The change in basis implies that now the relevant matrix is $\dot{T} := R_1^{-1} L_1^{-1} \hat{T} L_1 R_1$, with structure shown in Figure 9.8.

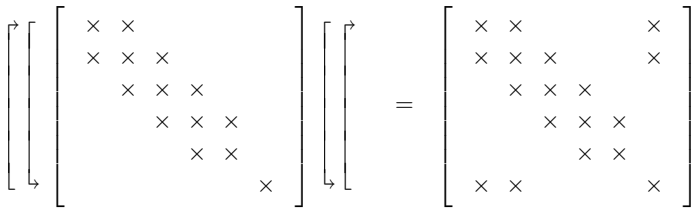


Figure 9.8: Structure of $R_1^{-1} L_1^{-1} \hat{T} L_1 R_1 = \dot{T}$.

### Restore structure

The matrix representation $\dot{T}$ of $\tilde{\Lambda}$ in the new basis is not a tridiagonal matrix. To restore the tridiagonal structure eliminators are constructed to eliminate the nonzero elements in the last row and last column deviating from the tridiagonal structure. To preserve the first columns of $\dot{V}$ and $\dot{W}$, the first elements are eliminated with $L_2 \in \mathfrak{L}_2$ and $R_2 \in \mathfrak{R}_2$. Eliminators $L_{i+1} \in \mathfrak{L}_{i+1}$ and $R_{i+1} \in \mathfrak{R}_{i+1}$, $i = 1, \ldots, m-1$, can be constructed to subsequently eliminate the elements $\dot{t}^{[i]}_{m+1,i}$ and $\dot{t}^{[i]}_{i,m+1}$ in $\dot{T}^{[i]} := R_i^{-1} L_i^{-1} \dot{T}^{[i-1]} L_i R_i$, $\dot{T}^{[1]} := \dot{T}$. The parameters are

$$\begin{cases} l_{i+1} = -\frac{\dot{t}^{[i]}_{m+1,i}}{\dot{t}^{[i]}_{i+1,i}} \\ r_{i+1} = -\frac{\dot{t}^{[i]}_{i,m+1}}{\dot{t}^{[i]}_{i,i+1} + l_{i+1} \dot{t}^{[i]}_{i,m+1}} \end{cases}$$

This procedure can break down, namely the denominator in the expression for $r_{i+1}$ can vanish. If this denominator becomes very small, then a numerical breakdown might occur. Under the assumption that no breakdown occurs, the product $X = \prod_{k=1}^{m} L_k R_k$ restores the tridiagonal structure and provides the solution to the updated TIEP

$$X^{-1} \hat{T} X = T_{m+1}.$$

The resulting biorthonormal bases are $V_{m+1} = \hat{V} X$ and $W_{m+1} = \hat{W} X^{-H}$, but these are not required in the updating procedure.
Figure 9.9 shows $\dot{T}^{[i]}$ during the updating process.

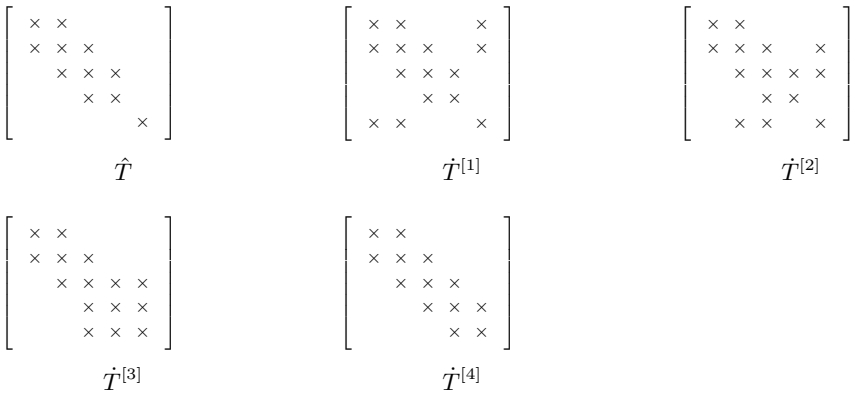


Figure 9.9: Updating strategy restoring tridiagonal structure.

### Algorithm

Algorithm 11 describes the updating procedure for nodes $z_{m+1} \notin \sigma(T_m)$. Assuming no breakdowns occur this algorithm results in a solution to an updated TIEP. In the algorithm the notation $L(l,i)$ and $R(r,i)$ is used, which is short for $L_i \in \mathfrak{L}_i$ and $R_i \in \mathfrak{R}_i$ with parameters $l_i = l$ and $r_i = r$, respectively.

**Algorithm 11** Updating procedure TIEP

**Input:** Proper tridiagonal matrix $T_m \in \mathbb{C}^{m\times m}$, node $z_{m+1} \in \mathbb{C}$, $z_{m+1} \notin \sigma(T_m)$ and weights $\alpha_{m+1}, \beta_{m+1} \neq 0$.
**Output:** Assuming no breakdowns, proper tridiagonal matrix $T_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$
**procedure** UPDATETIEP($T_m, z_{m+1}, \alpha_{m+1}, \beta_{m+1}$)
    $L := L\left(\frac{\alpha_{m+1}}{\eta}, 1\right)$, $R := R\left(-\frac{\eta\bar{\beta}_{m+1}}{\bar{\nu}\eta+\bar{\beta}_{m+1}\alpha_{m+1}}, 1\right)$
    $\dot{T} := R^{-1}L^{-1}\begin{bmatrix} T_m & \\ & z_{m+1} \end{bmatrix} LR$
    **for** $k = 1, \ldots, m-1$ **do**
        $L := L\left(-\dot{t}_{m+1,k}/t_{k+1,k}, k+1\right)$
        $R := R\left(\dot{t}_{k,m+1}t_{k+1,k}/\left(\dot{t}_{k,k+1}t_{k+1,k} + \dot{t}_{m+1,k}\dot{t}_{k,m+1}\right), k+1\right)$
        $\dot{T} := R^{-1}L^{-1}\dot{T}LR$
    **end for**
    $T_{m+1} := \dot{T}$
**end procedure**

### 9.4.2 Tridiagonal pencil IEP

Using eliminators the solution to a TPIEP can also be updated. Suppose a solution to the TPIEP with the matrix of nodes $\Lambda = \operatorname{diag}(\{z_i\}_{i=1}^m)$, weights $v = \begin{bmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_m\end{bmatrix}^\top$, $w = \begin{bmatrix}\beta_2 & \beta_2 & \dots & \beta_m\end{bmatrix}^\top$ and poles $\Xi = \{\xi_i\}_{i=1}^{m-1}$, $\Theta = \{\theta_i\}_{i=1}^{m-2}$ is available. This solution consists of biorthonormal matrices $V_m, W_m \in \mathbb{C}^{m\times m}$ and a tridiagonal pencil $(T_m, S_m) \in \mathbb{C}^{m\times m}$ satisfying, for $w^H v = \bar{\eta}\nu \neq 0$,

$$W_m^H \Lambda V_m S_m = T_m, \quad V_m e_1 = \frac{v}{\eta}, \quad W_m e_1 = \frac{w}{\nu}. \tag{9.19}$$

Next, a node $z_{m+1}$, weights $\alpha_{m+1}, \beta_{m+1}$ and poles $\xi_m, \theta_{m-1}$ are added to the original problem. The new problem consists of nodes $\widetilde{\Lambda} = \operatorname{diag}(\{z_i\}_{i=1}^{m+1})$, weights $\tilde{v} = \begin{bmatrix}v & \alpha_{m+1}\end{bmatrix}^\top$, $\tilde{w} = \begin{bmatrix}w & \beta_{m+1}\end{bmatrix}^\top$ and poles $\widetilde{\Xi} = \{\xi_i\}_{i=1}^{m}$, $\widetilde{\theta} = \{\theta_i\}_{i=1}^{m-1}$. The biorthonormal matrices $V_m, W_m \in \mathbb{C}^{m\times m}$ are embedded while preserving their biorthonormality

$$\widehat{V} := \begin{bmatrix} V_m & \\ & 1 \end{bmatrix}, \qquad \widehat{W} := \begin{bmatrix} W_m & \\ & 1 \end{bmatrix}. \tag{9.20}$$

And for $\hat{s}_{m+1}$ and $\hat{t}_{m+1}$ satisfying $z_{m+1}\hat{s}_{m+1} = \hat{t}_{m+1}$, define $\widehat{T}$ and $\widehat{S}$ as

$$\widetilde{\Lambda}\widehat{V} \underbrace{\begin{bmatrix} S_m & \\ & \hat{s}_{m+1} \end{bmatrix}}_{=:\widehat{S}} = \widehat{V} \underbrace{\begin{bmatrix} T_m & \\ & \hat{t}_{m+1} \end{bmatrix}}_{=:\widehat{T}}. \tag{9.21}$$

The bases $V, W$ do not have to be available in order to execute the updating procedure. Theorem 9.4 states how a solution can be efficiently obtained from the embedded matrices (9.20), (9.21) via similarity transformations with eliminators.

**Theorem 9.4.** *Let $V_m, W_m \in \mathbb{C}^{m\times m}$, $(T_m, S_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ be the solution to Problem 9.8 of size $m$ and let $\widehat{V}, \widehat{W}, \widehat{T}, \widehat{S}$ denote the corresponding embedded matrices* (9.20), (9.21). *Then there exist a nonsingular matrix $C$, eliminators $L_i \in \mathfrak{L}_i$, $R_i \in \mathfrak{R}_i$, $i = 1, 2, \dots, m$ and $\dot{L}_j \in \mathfrak{L}_j$, $\dot{R}_j \in \mathfrak{R}_j$, $j = 1, 2, \dots, m$ such that*

$$\widetilde{T} = \left(\prod_{k=m}^{1} R_k \dot{L}_k\right) \widehat{T} C L_1 \left(\prod_{k=2}^{m} \dot{R}_k L_k\right), \quad \widetilde{S} = \left(\prod_{k=m}^{1} R_k \dot{L}_k\right) \widehat{S} C L_1 \left(\prod_{k=2}^{m} \dot{R}_k L_k\right),$$

$$\widetilde{V} = \widehat{V}\left(\prod_{k=1}^{m} \dot{L}_k^{-1} R_k^{-1}\right) \quad \text{and} \quad \widetilde{W} = \widehat{W}\left(\prod_{k=1}^{m} \dot{L}_k^{H} R_k^{H}\right),$$

*solve the IEP of size $m+1$. This IEP is obtained by adding a node $z_{m+1} \notin \sigma(T_m, S_m)$, weights $\alpha_{m+1}, \beta_{m+1}$ and poles $\xi_m, \theta_{m-1}$ to the original IEP of dimension $m$.*

The remainder of this section is dedicated to providing the components required to prove this theorem. Three steps compose the updating procedure: introduction of the new weights, which will perturb the structure of the pencil, restoring the structure and introducing the new poles. The proofs provided in this section are constructive, they provide the necessary components for an algorithm implementing the updating procedure. We published our implementation of the algorithm online [163].

### Introduce weights $\alpha_{m+1}, \beta_{m+1}$

The new weights $\alpha_{m+1}, \beta_{m+1}$ can be introduced as described for the TIEP. Let $L_1 \in \mathfrak{L}_1$, $R_1 \in \mathfrak{R}_1$ be such that

$$\hat{V} L_1^{-1} R_1^{-1} e_1 = \tilde{v}/\tilde{\eta} \text{ and } \hat{W} L_1^{-H} R_1^{-H} e_1 = \tilde{w}/\tilde{\nu}.$$

The new bases lead to the pencil $(R_1 \dot{L}_1 \widehat{T}, R_1 \dot{L}_1 \widehat{S})$, which is no longer a tridiagonal pencil

$$\widetilde{\Lambda} \widehat{V} \dot{L}_1^{-1} R_1^{-1} \left[\begin{array}{ccc|c} & \breve{S} & & \times \\ & & & 0 \\ & & & \mathbf{0} \\ \hline \times & \times & \mathbf{0}^\top & \times \end{array}\right] = \widehat{V} \dot{L}_1^{-1} R_1^{-1} \left[\begin{array}{ccc|c} & \breve{T} & & \times \\ & & & 0 \\ & & & \mathbf{0} \\ \hline \times & \times & \mathbf{0}^\top & \times \end{array}\right], \tag{9.22}$$

with $\breve{T}, \breve{S}$ tridiagonal matrices of size $m$ which, respectively, differ from $T, S$ only in their first row.

### Restore tridiagonal pencil structure

Using eliminators, the tridiagonal structure of the pencil will be restored, Lemma 9.5 provides the details.

**Lemma 9.5.** *Let $(\widehat{T}, \widehat{S})$ denote the embedded pencil* (9.21) *and $R_1$, $\dot{L}_1$ the matrices from* (9.18)*. Then, under the assumption that no breakdown occurs, there exist a nonsingular upper triangular matrix $U$, eliminators $L_i \in \mathfrak{L}_i$, $R_i \in \mathfrak{R}_i$, for $i = 1, 2, \ldots, m-1$ and $\dot{L}_j \in \mathfrak{L}_j$, $\dot{R}_j \in \mathfrak{R}_j$, $j = 2, 3, \ldots, m$ such that*

$$\dot{T}^{[m]} := \dot{L}_m \left( \prod_{k=1}^{m-1} R_k \dot{L}_k \right) \widehat{T} U L_1 \left( \prod_{k=2}^{m-1} \dot{R}_k L_k \right) \dot{R}_m,$$

$$\dot{S}^{[m]} := \dot{L}_m \left( \prod_{k=1}^{m-1} R_k \dot{L}_k \right) \widehat{S} U L_1 \left( \prod_{k=2}^{m-1} \dot{R}_k L_k \right) \dot{R}_m \tag{9.23}$$

*have tridiagonal structure and the same poles as $(\widehat{T}, \widehat{S})$.*

*Proof.* The proof is by induction. The first step, $i = 1$, differs from the general iteration ($i \geq 2$). Consider the matrices $\dot{L}_1, R_1$ from (9.18), and denote their parameter by $\dot{l}_1$, $r_1$, respectively. Then the pencil which must be reduced to tridiagonal form is, for some $\dot{t}^{[1]}, \dot{s}^{[1]}$ satisfying $\dot{t}^{[1]} z_{m+1} = \dot{s}^{[1]}$,

$$\dot{T}^{[1]} := R_1 \dot{L}_1 \widehat{T} = \left[\begin{array}{c|cc|c} (1+\dot{l}_1 r_1)t_{11} & (1+\dot{l}_1 r_1)t_{12} & \mathbf{0} & b_1 \dot{t}^{[1]} \\ \hline t_{21} & & & 0 \\ \mathbf{0} & \multicolumn{2}{c|}{T^{(m-1)}} & \mathbf{0} \\ & & & \\ \hline \dot{l}_1 t_{11} & \dot{l}_1 t_{12} & \mathbf{0}^\top & \dot{t}^{[1]}_{m+1,m+1} \end{array}\right],$$

$$\dot{S}^{[1]} := R_1 \dot{L}_1 \widehat{S} = \left[\begin{array}{c|cc|c} (1+\dot{l}_1 r_1)s_{11} & (1+\dot{l}_1 r_1)s_{12} & \mathbf{0} & r_1 \dot{s}^{[1]} \\ \hline s_{21} & & & 0 \\ \mathbf{0} & \multicolumn{2}{c|}{S^{(m-1)}} & \mathbf{0} \\ & & & \\ \hline \dot{l}_1 s_{11} & \dot{l}_1 s_{12} & \mathbf{0}^\top & \dot{s}^{[1]}_{m+1,m+1} \end{array}\right],$$

where $T^{(m-1)}$ and $S^{(m-1)}$ denote, respectively, the principal trailing submatrix of size $(m-1) \times (m-1)$ of $T$ and $S$ from (9.19). Note in the above equation that the matrices are very similar, therefore, we will only explicitly write down $\dot{T}^{[i]}$ and omit $\dot{S}^{[i]}$. The annihilation of the first element in the last row and last column is performed in two steps. The first step creates colinearity between relevant elements in the pencil and the second step eliminates the first elements in the last row (and last column) simultaneously.

First, create colinearity, i.e., find suitable $L_1 \in \mathfrak{L}_1$ with parameter $l_1$ and $U := \begin{bmatrix} 1 & u & \\ & 1 & \\ & & I_{m-1} \end{bmatrix}$ which act on $\dot{T}^{[1]}$ as follows

$$\dot{T}^{[1]} U L_1 = \left[\begin{array}{c|cc|c} \tilde{t}_{11} & \tilde{t}_{12} & \mathbf{0} & r_1 \dot{t}^{[1]} \\ \hline \tilde{t}_{21} & & & 0 \\ \mathbf{0} & \multicolumn{2}{c|}{T^{(m-1)} + u t_{21} e_1 e_1^\top} & \mathbf{0} \\ & & & \\ \hline \dot{l}_1 t_{11} + l_1 \dot{t}^{[1]} & \dot{l}_1 (t_{12} + u t_{11}) & \mathbf{0}^\top & \dot{t}^{[1]}_{m+1,m+1} \end{array}\right],$$

where $\tilde{t}_{11} := (1+\dot{l}_1 r_1)t_{11} + l_1 r_1 \dot{t}^{[1]}_{m+1,m+1}$, $\tilde{t}_{12} := (1+\dot{l}_1 r_1)(t_{12}+ut_{11})$, $\tilde{t}_{21} := t_{21}$. The resulting matrix should satisfy the colinearity conditions

$$\operatorname{rank}\left(\begin{bmatrix} \tilde{t}_{12} & r_1\dot{t}^{[1]}_{m+1,m+1} \\ \tilde{s}_{12} & r_1\dot{s}^{[1]}_{m+1,m+1}\end{bmatrix}\right) = 1,$$

$$\operatorname{rank}\left(\begin{bmatrix} \tilde{t}_{21} & \tilde{s}_{21} \\ \dot{l}_1 t_{11} + a_1\dot{t}^{[1]}_{m+1,m+1} & \dot{l}_1 s_{11} + l_1\dot{s}^{[1]}_{m+1,m+1}\end{bmatrix}\right) = 1,$$

this can be achieved by choosing appropriate $u$ and $l_1$. Next, $\dot{L}_2 \in \mathfrak{L}_2$ and $\dot{R}_2 \in \mathfrak{R}_2$ eliminate these colinear elements in the last row and last column, i.e.,

$$\dot{L}_2\dot{T}[1]UL_1\dot{R}_2 = \left[\begin{array}{cc|cc|c} \tilde{t}_{11} & \tilde{t}_{12} & 0 & \mathbf{0}^\top & 0 \\ \hline \tilde{t}_{21} & & & & \dot{t}^{[2]}_{2,m+1} \\ 0 & \multicolumn{2}{c|}{T^{(m-1)} + u\tilde{t}_{21}e_1e_1^\top} & & \dot{b}_2 t_{32} \\ \mathbf{0} & & & & \mathbf{0} \\ \hline 0 & \dot{t}^{[2]}_{m+1,2} & \dot{l}_2 t_{23} & \mathbf{0}^\top & \dot{t}^{[2]}_{m+1,m+1} \end{array}\right] =: \dot{T}^{[2]},$$

where $\dot{t}^{[2]}_{m+1,2} := \dot{l}_1(t_{12}+ut_{11}) + \dot{l}_2(\dot{t}^{[1]}_{22} + u\tilde{t}_{21})$, $\dot{t}^{[2]}_{2,m+1} := \dot{r}_1(\dot{t}^{[1]}_{22} + u\tilde{t}_{21})$ and $\dot{t}^{[2]}_{m+1,m+1} := (1+\dot{l}_2 b_1)\dot{t}^{[1]} + \dot{r}_2\dot{l}_1(t_{12}+ut_{11}) + \dot{r}_2\dot{l}_2(\dot{t}^{[1]}_{22} + ut_{21})$. By the colinearity, the same holds for $\dot{S}^{[2]}$. This shows that the initial step can be performed using the matrices $U, L_1, \dot{L}_2$ and $\dot{R}_2$. Note that this $U$ is only required in the first step, in the subsequent steps it will be replaced by a matrix in $\mathfrak{R}_i$. Under the induction hypothesis, we have

$$\dot{T}^{[i]} := \dot{L}_i R_{i-1}\dot{T}^{[i-1]}L_{i-1}\dot{R}_i = \left[\begin{array}{cc|cc|c} \widetilde{T}^{(i-1)} & \tilde{t}_{i-1,i}e_i & & & \mathbf{0} \\ \hline \tilde{t}_{i,i-1}e_i^\top & & & & \dot{t}^{[i]}_{i,m+1} \\ & & \multicolumn{2}{c|}{T^{(m-i+1)}} & \dot{r}_i t_{i+1,i} \\ & & & & \mathbf{0} \\ \hline \mathbf{0} & \dot{t}^{[i]}_{m+1,i} & \dot{l}_i t_{i,i+1} & \mathbf{0}^\top & \dot{t}^{[i]}_{m+1,m+1} \end{array}\right],$$

where $T^{(m-i+1)} \in \mathbb{C}^{(m-i+1)\times(m-i+1)}$, the principal trailing submatrix of $T$ (9.19) and $\widetilde{T}^{(i-1)}$ is the $(i-1)\times(i-1)$ leading principal submatrix of the solution $\widetilde{T}$ to the TPIEP. All the action takes place in the $(m-i+1)\times(m-i+1)$ principal trailing submatrix of $\dot{T}^{[i]}$. For the proof to hold, $\dot{T}^{[i+1]} = \dot{L}_{i+1}R_i\dot{T}^{[i]}L_i\dot{R}_{i+1}$ and $\dot{S}^{[i+1]} = \dot{L}_{i+1}R_i\dot{S}^{[i]}L_i\dot{R}_{i+1}$ must create zeros on positions $(m+1,i)$ and $(i,m+1)$. The first step enforcing colinearity, determining $L_i \in \mathfrak{L}_i$ and $R_i \in \mathfrak{R}_i$, can be elegantly formulated using

$$M_l := s_{i+1,i}\begin{bmatrix} t_{i+1,i} & \dot{r}_i t_{i+1,i} \\ \dot{t}^{[i]}_{m+1,i} & \dot{t}^{[i]}_{m+1,m+1}\end{bmatrix} - t_{i+1,i}\begin{bmatrix} s_{i+1,i} & \dot{r}_i s_{i+1,i} \\ \dot{s}^{[i]}_{m+1,i} & \dot{s}^{[1]}_{m+1,m+1}\end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \times & \times\end{bmatrix},$$

$$M_r := s_{i,i+1}\begin{bmatrix} t_{i,i+1} & \dot{t}^{[i]}_{i,m+1} \\ \dot{l}_i t_{i,i+1} & \dot{t}^{[i]}_{m+1,m+1}\end{bmatrix} - t_{i+1,i}\begin{bmatrix} s_{i,i+1} & \dot{s}^{[i]}_{i,m+1} \\ \dot{l}_i s_{i,i+1} & \dot{s}^{[1]}_{m+1,m+1}\end{bmatrix} = \begin{bmatrix} 0 & \times \\ 0 & \times\end{bmatrix}.$$

They are constructed such that, isolating their active part in $L_i^b$ and $R_i^b$, they create rank one matrices

$$M_l L_i^b = \begin{bmatrix} 0 & 0 \\ 0 & \times \end{bmatrix} \text{ and } R_i^b M_r = \begin{bmatrix} 0 & 0 \\ 0 & \times \end{bmatrix}.$$

Clearly, appropriate $\dot{L}_{i+1} \in \mathfrak{L}_{i+1}$ and $\dot{R}_{i+1} \in \mathfrak{R}_{i+1}$ can be found which will eliminate the $i$th elements in the last row and column of $\dot{T}^{[i+1]}$ and $\dot{S}^{[i+1]}$ such that

$$\dot{T}^{[i+1]} := \dot{L}_{i+1} R_i \dot{T}^{[i]} L_i \dot{R}_{i+1}$$

$$= \left[\begin{array}{c|c|c} \widetilde{T}^{(i)} & \tilde{t}_{i,i+1} e_{i+1} & \mathbf{0} \\ \hline \tilde{t}_{i+1,i} e_{i+1}^\top & & \dot{t}_{i+1,m+1}^{[i+1]} \\ & T^{(m-i)} & \dot{r}_{i+1} t_{i+2,i+1} \\ & & \mathbf{0} \\ \hline \mathbf{0} & \dot{t}_{m+1,i+1}^{[i+1]} \quad \dot{l}_{i+1} t_{i+1,i+2} \quad \mathbf{0}^\top & \dot{t}_{m+1,m+1}^{[i+1]} \end{array}\right],$$

with $\dot{t}_{m+1,i+1}^{[i+1]} = \dot{l}_i t_{i,i+1} + \dot{l}_{i+1} t_{i+1,i+1}$, $\dot{t}^{[i+1]} = \hat{t}_{m+1,m+1}^{[i]} + \dot{l}_{i+1} \dot{r}_i t_{i+1,i} + \dot{r}_{i+1} \dot{r}_i t_{i,i+1}$ and $\dot{t}_{i+1,m+1}^{[i+1]} = \dot{r}_i t_{i+1,i} + \dot{r}_{i+1} t_{i+1,i+1}$. Furthermore, $\tilde{t}_{i,i} = t_{i,i} + r_i \dot{t}_{m+1,i}^{[i]} + l_i \dot{t}_{i,m+1}^{[i]}$, $\tilde{t}_{i,i+1} = (1 + \dot{l}_i r_i) t_{i,i+1}$ and $\tilde{t}_{i+1,i} = (1 + \dot{r}_i l_i) t_{i+1,i}$. The same holds for $\dot{S}^{[i+1]}$ thanks to the colinearity. Hence, the poles are preserved, $\tilde{t}_{i+1,i}/\tilde{s}_{i+1,i} = t_{i+1,i}/s_{i+1,i} = \xi_i$, $i = 1, 2, \ldots, m-1$ and $\tilde{t}_{i,i+1}/\tilde{s}_{i,i+1} = t_{i,i+1}/s_{i,i+1} = \bar{\psi}_{i-1}$, $i = 2, 3, \ldots, m-1$.
This can be continued for $i < m$, thereby obtaining at $i = m-1$ a tridiagonal pencil $(\dot{T}^{[m]}, \dot{S}^{[m]})$. Whenever $\dot{t}^{[i]}$ (or $\dot{s}^{[i]}$) vanishes for some $i$, the above transformations cannot be determined, i.e., a breakdown occurs. This situation is excluded by the assumption that no breakdowns occur. □

The structure restoring process for one of the matrices from the pencil is shown in Figure 9.10, the structure for both matrices in the pencil is at each step the same. Lemma 9.5 shows that a tridiagonal pencil $(\dot{T}^{[m]}, \dot{S}^{[m]})$ with eigenvalues $z_i$ can be constructed while preserving its poles, i.e., $\frac{\dot{t}_{i+1,i}^{[m]}}{\dot{s}_{i+1,i}^{[m]}} = \xi_i$, $i = 1, 2, \ldots, m-1$ and $\frac{\dot{t}_{i,i+1}^{[m]}}{\dot{s}_{i,i+1}^{[m]}} = \bar{\psi}_{i-1}$, $i = 2, 3, \ldots, m-1$. However, in general $\frac{\dot{t}_{m+1,m}^{[m]}}{\dot{s}_{m+1,m}^{[m]}} \neq \xi_m$ and $\frac{\dot{t}_{m,m+1}^{[m]}}{\dot{s}_{m,m+1}^{[m]}} \neq \bar{\psi}_{m-1}$, which are the poles belonging to the rational functions that are added to the spaces.

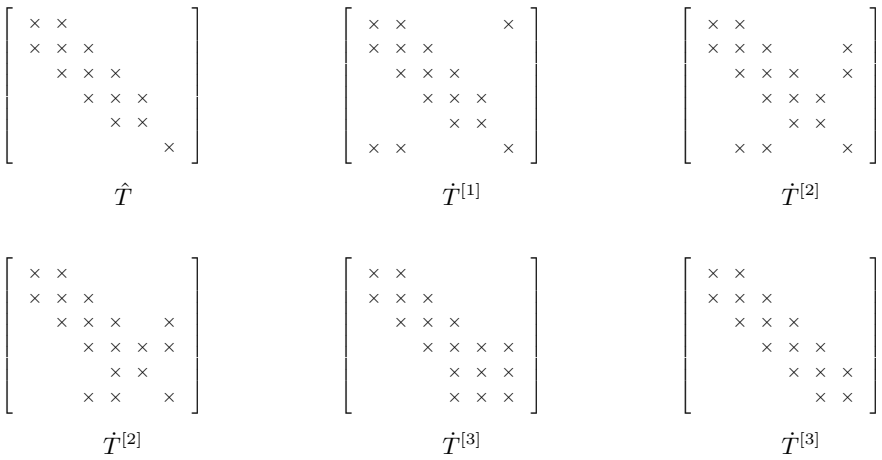



Figure 9.10: Updating strategy restoring tridiagonal pencil structure.

**Introduce poles** $\xi_m, \theta_{m-1}$

The new poles must be introduced in the pencil without disturbing the tridiagonal structure, Lemma 9.6 provides the details.

**Lemma 9.6.** *Consider a tridiagonal pencil* $(\dot{T}, \dot{S})$*, where* $\dot{t}_{m+1,m+1}, \dot{s}_{m+1,m+1} \neq 0$*. Premultiplication with* $R_m \in \mathfrak{R}_m$ *and postmultiplication with* $L_m \in \mathfrak{L}_m$ *suffice to alter, respectively, the ratio of elements* $(m, m+1)$ *and* $(m+1, m)$ *of* $(\dot{T}, \dot{S})$ *to any values* $\theta_{m-1} \in \bar{\mathbb{C}}$ *and* $\xi_m \in \bar{\mathbb{C}}$*.*

*Proof.* Only the $2 \times 2$ trailing principal submatrices of $(\dot{T}, \dot{S})$ are altered by the transformation $(R_m \dot{T} L_m, R_m \dot{S} L_m)$, denoted compactly by

$$\begin{bmatrix} 1 & r_m \\ & 1 \end{bmatrix} \begin{bmatrix} \dot{t}_{m,m} & \dot{t}_{m,m+1} \\ \dot{t}_{m+1,m} & \dot{t}_{m+1,m+1} \end{bmatrix} \begin{bmatrix} 1 & \\ l_m & 1 \end{bmatrix} = \begin{bmatrix} \times & \alpha \\ \mu & \times \end{bmatrix}$$

$$\begin{bmatrix} 1 & r_m \\ & 1 \end{bmatrix} \begin{bmatrix} \dot{s}_{m,m} & \dot{s}_{m,m+1} \\ \dot{s}_{m+1,m} & \dot{s}_{m+1,m+1} \end{bmatrix} \begin{bmatrix} 1 & \\ l_m & 1 \end{bmatrix} = \begin{bmatrix} \times & \beta \\ \nu & \times \end{bmatrix}.$$

The formation of the product on the left-hand side shows that, under the assumption $\dot{t}_{m+1,m+1} \neq 0$ and $\dot{s}_{m+1,m+1} \neq 0$, we can create any ratios $\alpha/\beta, \mu/\nu \in \bar{\mathbb{C}}$. □

## 9.5 Numerical experiments

The proposed solution strategies, based on Krylov subspace methods and on the updating procedures, are analyzed numerically. Consider the diagonal matrix of distinct nodes $\Lambda \in \mathbb{C}^{m\times m}$, a weight vector $v \in \mathbb{C}^m$ and a set of poles $\Xi = \{\xi_i\}_{i=1}^{m-1}$, with $\xi_i \in \mathbb{C}\backslash\{0\}$ (the exclusion of $\xi_i = 0$ and $\xi_i = \infty$ is done solely for simplicity of notation). In case of biorthogonality, an extra weight vector $w \in \mathbb{C}^m$ and set of poles $\Theta$ is provided. Both solution strategies compute a solution to Problem 9.5, i.e., a pencil $(B, C)$ such that

$$W^H \Lambda V C = B,$$

with $W^H V = I$, $We_1 = \frac{w}{\nu}$, $Ve_1 = \frac{v}{\eta}$ and $(B, C)$ adhering to either Hessenberg or tridiagonal structure. When computing this in finite precision some errors will arise and these are measured. The biorthogonality of the formed bases $V, W$ is measured by

$$\mathrm{err_o} := \|W^H V - I\|_2,$$

where $I$ is a unit matrix of appropriate size and if we consider an orthogonal basis $Q$, then $V = W = Q$. The accuracy of the recurrence relation, consisting of recurrence matrices $(B, C)$ and basis $V$, is measured by

$$\mathrm{err_r} := \frac{\|\Lambda V C - V B\|_2}{\max\left(\|\Lambda V C\|_2, \|V B\|_2\right)}.$$

The elements $(B, C)$ represent recurrence coefficients of sequences of biorthogonal rational functions, $\{r_i\}_{i=0}^{m-1}$ and $\{s_i\}_{i=0}^{m-1}$, or a single sequence of orthogonal rational functions $\{r_i\}_{i=0}^{m-1}$. The orthogonality of these functions is checked by constructing their Gram matrix, which should equal the unit matrix. We get for the orthogonal case with inner product $\langle ., . \rangle_m$:

$$\mathrm{err_f} := \left\| \left[ \langle r_i, r_j \rangle_m \right]_{i,j=0}^{m-1} - I \right\|_2$$

and for the biorthogonal case with linear functional $\mathcal{L}_m\{.\}$:

$$\mathrm{err_f} := \left\| \left[ \mathcal{L}_m\{r_i s_j\} \right]_{i,j=0}^{m-1} - I \right\|_2.$$

The evaluation of the rational functions $\{r_i\}$ and $\{s_i\}$ using its recurrence pencil $(B, C)$ is discussed in Chapter 7.

The final error metric quantifies the accuracy of the poles. The poles of the computed pencil $(B, C)$ are compared to the given poles $\xi_i$,

$$\mathrm{err_p} = \max_{1\le i\le m-1} \left\{ \frac{\left| \frac{B(i+1,i)}{C(i+1,i)} - \xi_i \right|}{|\xi_i|} \right\}.$$

For the TPIEP, also the superdiagonal ratios reveal poles and must be taken into account, which is taken to be the maximum of the above metric and the following

$$\mathrm{err}_{\mathrm{p}} = \max_{2 \le i \le m-1} \left\{ \frac{\left| \frac{B(i,i+1)}{C(i,i+1)} - \bar{\psi}_{i-1} \right|}{|\psi_{i-1}|} \right\}.$$

Throughout this section all weights are chosen to be equal to the value 1.
Throughout the following discussion, it is important to be aware of the essential difference between both solution procedures: the updating procedure starts from an already known solution to construct the next solution, whereas the Krylov procedure must start over every time the problem changes. Therefore, the updating procedure is much more efficient in situations where the solution to a related problem is available. On the other hand, the Krylov procedure possesses all information about the whole problem, which typically leads to a more accurate solution.

### Hessenberg pencil

Two experiments are discussed. The first uses equidistant nodes on the unit circle and highlights the numerical stability of the proposed updating procedure. The second illustrates the influence of the given nodes on the accuracy of the numerical solution by choosing two nodes close to each other.
The first experiment uses equidistant nodes on the unit circle. Since updating always adds one node and keeps all others fixed, we cannot have equidistant nodes at each step. The node is then added at the largest distance from all nodes already generated, exactly in the middle of two adjacent nodes. This order of adding nodes is chosen because it is a good order for the updating procedure, the order has little effect on the final solution (for the same nodes) but strongly influences the intermediate behavior. The poles $\Xi$ are chosen equidistant on a circle of radius 1.5. The result for problem sizes $m = 3, 18, \ldots, 393$ is shown in Figure 9.11. The three metrics for the matrix solution, $\mathrm{err}_{\mathrm{o}}$, $\mathrm{err}_{\mathrm{r}}$ and $\mathrm{err}_{\mathrm{p}}$, show very good accuracy for both procedures, with the Krylov procedure performing slightly better, which can be attributed to the benefit of solving the complete problem every time. The metric for the orthogonality of the rational functions, $\mathrm{err}_{\mathrm{f}}$, shows that the updating procedure performs much better than the Krylov procedure.

To explain this, we must look at the condition number $\kappa(B, C)$ for the pencil $(B, C)$ obtained by both procedures, shown in Table 9.1. This table shows that the condition of the system of equations is much larger for the pencil obtained by the Krylov procedure. The updating procedure performs unitary similarity transformations, therefore if nodes are located on the unit circle then the pencil consists of unitary matrices and this leads to much better conditioning than the Krylov procedure, which does not generate unitary matrices. Note that the pencil $(B, C)$ as a whole is unitary in both cases. The

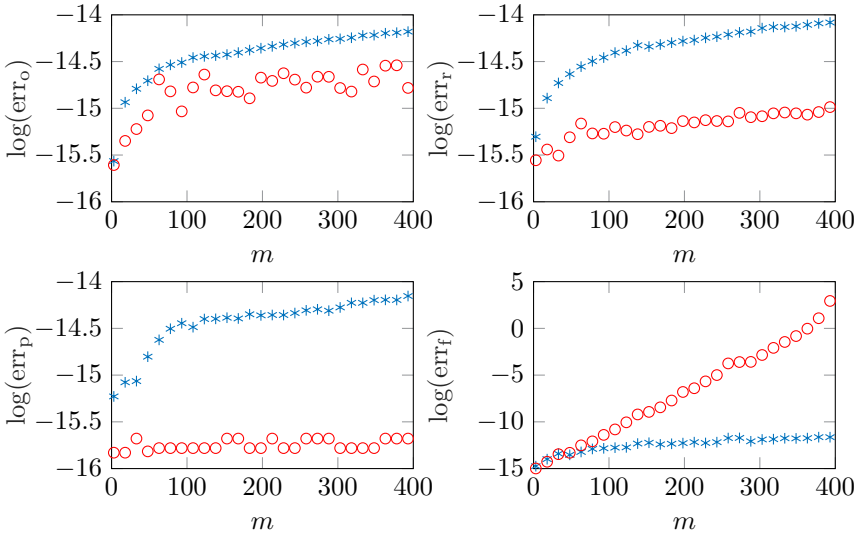



Figure 9.11: HPIEP with nodes on the unit circle and poles on a circle with radius 1.5. Error metrics for Krylov '∘' and updating '*' procedure in log scale for problem size $m$.

| $m$ | 10 | 100 | 200 | 300 | 400 |
|---|---|---|---|---|---|
| Update | 1.9e01 | 2.3e02 | 4.8e02 | 1.4e03 | 9.1e03 |
| Krylov | 2.9e01 | 2.8e05 | 4.0e09 | 3.3e13 | 7.7e17 |

Table 9.1: Condition number $\kappa(B, C)$ for the pencil $(B, C)$ obtained by the updating and Krylov procedure for the first experiment, with poles on a circle of radius 1.5, for problem size $m$.

condition number $\kappa(B, C)$ of the solution obtained by the Krylov procedure depends on the choice of poles. We repeat the above experiment with poles on a circle with radius 3. Table 9.2 shows a much smaller $\kappa(B, C)$ for the Krylov solution. This illustrates the dependence of $\kappa(B, C)$ of the Krylov solution on the prescribed poles. The updating solution is much less influenced by the choice of poles.

The second experiment shows how the numerical solution of a HPIEP depends on the location of the given nodes. The nodes are chosen as above, up to the following change, the $m_p$th node is chosen on the circle and close to the $(m_p - 1)$th node. That is, for $m \geq m_p$ nodes, we have $m - 1$ equidistant on the circle as above, and a node close to one of these nodes, the distance between these two nodes is given by the angle $\omega$. For small $\omega$ this leads to an underlying discrete inner product with $m$ nodes that

| $m$ | 10 | 100 | 200 | 300 | 400 |
|---|---|---|---|---|---|
| Update | 2.0e01 | 2.2e02 | 4.4e02 | 1.4e03 | 9.8e02 |
| Krylov | 2.4e01 | 3.2e02 | 6.1e02 | 1.8e03 | 1.4e03 |

Table 9.2: Condition number $\kappa(B, C)$ for the pencil $(B, C)$ obtained by the updating and Krylov procedure for the first experiment, with poles on a circle of radius 3, and problem size $m$.

is very close to an inner product with $m-1$ nodes. Therefore, the $m$th orthogonal rational function will become closer to numerical linear dependence, as $\omega$ gets smaller, with a deterioration of the orthonormality of the generated rational functions as a consequence. Figure 9.12 shows the results for $m_p = 50$, $\omega = 10^{-6}$ and equidistant poles on a circle of radius 3. The metrics $\mathrm{err_o}$ and $\mathrm{err_r}$ behave nicely, as with the first experiment. The error on the poles obtained by the updating procedure, $\mathrm{err_p}$, makes a jump when a value $m > m_p$ is reached and stagnates thereafter. This jump is caused by the small value obtained for the last element in the Hessenberg pencil, this element is the inner product of $r_{m_p-1}(z)$ with $r_{m_p-2}(z)$ which is small due to the similarity of the inner product with $m_p$ and $m_p - 1$ nodes. And this element is used to introduce the new pole, but due to the difference in order of magnitude (about the size of $\omega$), loss of accuracy (about 5-6 digits) is expected. As predicted, the orthonormality, measured by $\mathrm{err_f}$, of the complete set of the $m$ orthogonal rational functions deteriorates for $m > m_p$. However, if we look at the $m-1$ first orthogonal rational functions, the situation is much better. This means that the loss of orthogonality, as expected by the closeness of an inner product of $m$ and $m-1$ nodes, is isolated in the $m$th ORF. Hence, the first $m-1$ ORFs are still accurately computed.

### Tridiagonal pencil

The tridiagonal pencil is interesting because of the underlying short recurrence relation. If this can be combined with an inner product, which has preferred numerical properties over a general linear functional, then it will lead to an efficient procedure to generate ORFs in a stable manner. This scenario occurs for nodes on the real line and poles $\xi_i = \bar{\theta}_i$. The first experiment chooses Chebyshev nodes, obtained by projecting the equidistant nodes on the unit circle onto the real line. The second experiment serves as a proof of concept, the nodes are chosen equidistant on a thin ellipse, a choice between the unit circle and Chebyshev nodes, where Chebyshev nodes would be the limit case (such an ellipse of height zero). For both experiments all weights are equal to 1, i.e., $v = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^\top$ and $w = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^\top$.
The metrics for the first experiment, with Chebyshev nodes on the interval $[-1, 1]$ and equidistant poles on a circle of radius 3, are provided in Figure 9.13. As expected from using nonunitary similarity transformations, the procedure is no longer numerically

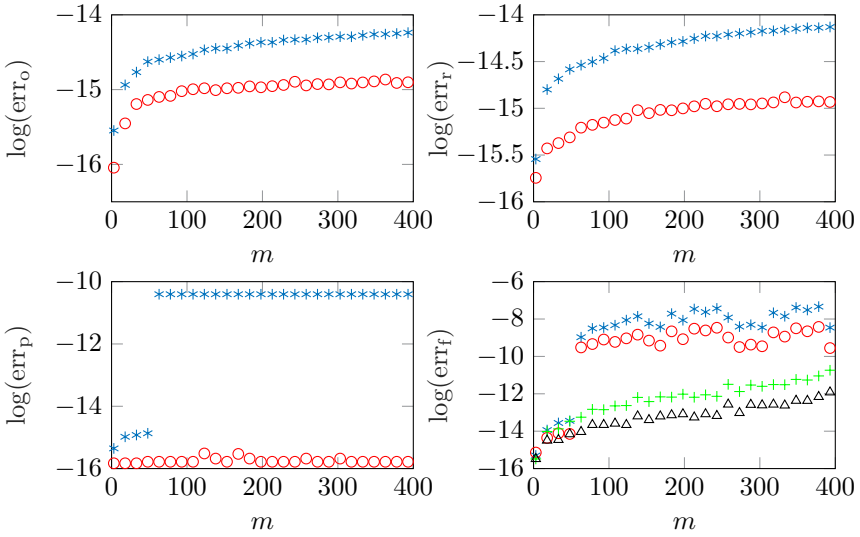



Figure 9.12: HPIEP with nodes on the unit circle, perturbation of $\omega = 10^{-6}$ at $m = 50$ and poles on a circle with radius 3. Error metrics for Krylov '$\circ$' and updating '$*$' procedure in log scale for problem size $m$. The metric $\text{err}_\text{f}$ for the $m-1$ first ORFs is indicated by '$+$' for Krylov and '$\triangle$' for updating.

stable, we see a steady deterioration of all metrics. The errors are however still relatively small, especially for $\text{err}_\text{f}$. Table 9.3 shows this metric for solutions obtained by the solution procedures for the HPIEP and TPIEP, which are equivalent except for the imposed structure on the pencil. The Hessenberg pencil achieves only one significant digit more than the tridiagonal pencil, which makes the procedures based on the biorthogonal formulation competitive thanks to the efficiency gained by the underlying short recurrence relation.

| m | 18 | 93 | 198 | 288 |
|---|---|---|---|---|
| Update TP | e-12.9 | e-10.6 | e-10.1 | e-9.8 |
| Krylov TP | e-13.2 | e-12.1 | e-11.5 | e-11.8 |
| Update HP | e-13.5 | e-11.5 | e-10.6 | e-10.4 |
| Krylov HP | e-13.6 | e-12 | e-12 | e-11.9 |

Table 9.3: Metric $\text{err}_\text{f}$ for solution of the IEP with Chebyshev nodes in $[-1, 1]$ and poles equidistant on circle with radius 3. Solutions are obtained by solving the HPIEP with updating and Krylov procedure and by solving the equivalent TPIEP with its respective updating and Krylov procedure.

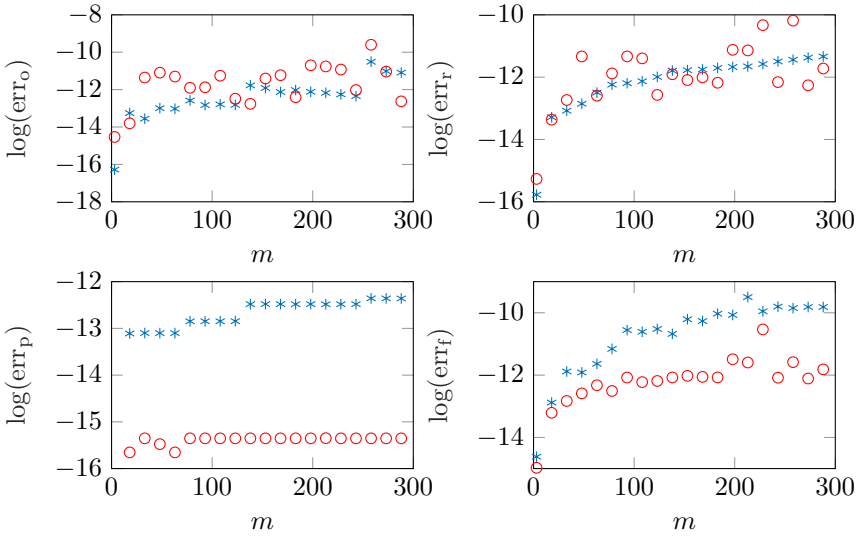



Figure 9.13: TPIEP with Chebyshev nodes on the interval $[-1, 1]$ and equidistant poles on a circle with radius 3. Error metrics for Krylov '$\circ$' and updating '*' procedure in log scale for problem size $m$.

The second experiment uses equidistant nodes on a thin ellipse $x^2 + (y/0.01)^2 = 1$, obtained by compressing the unit circle $x^2 + y^2 = 1$ in height. The poles $\Xi$ are chosen equidistant on a circle of radius 3 and $\Theta$ on a circle of radius 4. Results of this experiment are shown in Figure 9.14. The updating procedure performs well despite its biorthogonal nature. The solution obtained by the Krylov procedure deteriorates fast, this can be explained by the fact that it is a general purpose Lanczos-like recurrence relation, whereas the updating procedure is designed for this specific problem. For poles which are equal, the situation improves, as is shown by the metrics in Figure 9.15, where $\Xi = \Theta$.

These results show the potential of the biorthogonal procedures, especially for special cases such as discussed in the first experiment.

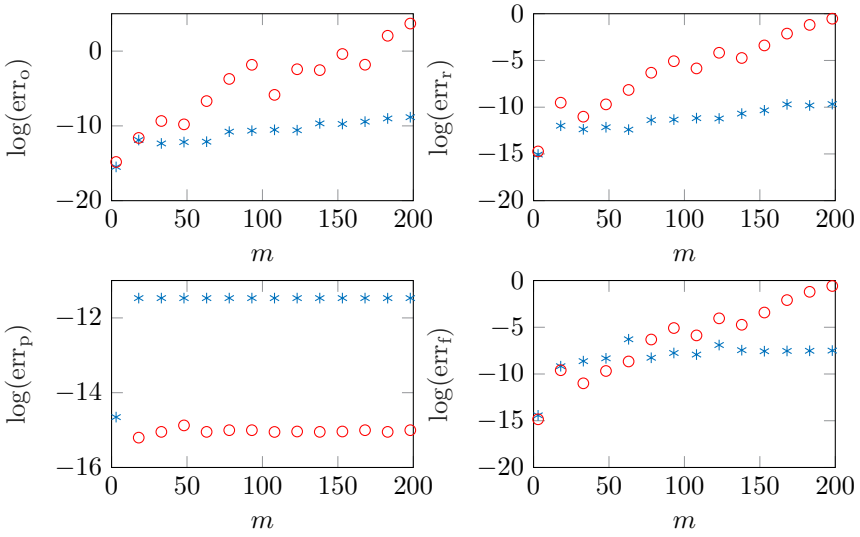



Figure 9.14: TPIEP with equidistant nodes on an ellipse $x^2 + (y/0.01)^2 = 1$ and poles $\Xi$ and $\Theta$ on a circle with radius 3 and 4, respectively. Error metrics for Krylov '$\circ$' and updating '*' procedure in log scale for problem size $m$.

## 9.6 Downdating procedures - inner product

Downdating procedures are procedures that remove an eigenvalue from an available solution to some IEP. This corresponds to removing a node-weight pair from the underlying inner product. Downdating an IEP is closely related to downdating matrix factorizations, which is known to be a much more difficult problem than updating [60]. For least squares problems a downdating procedure is proposed by Björck et al. [25]. Their procedure requires the full QR decomposition of the matrix of interest to be available. Downdating a Cholesky decomposition is discussed in [25, 26].
The downdating procedures proposed here only require the recurrence matrix or pencil and the node of the inner product to be removed. Downdating is discussed in more detail for the HIEP and HPIEP. The matrices arising from solving a HIEP are normal proper Hessenberg matrices with simple spectrum and the pencils from the HPIEP are normal, proper Hessenberg pencils with simple spectrum. These properties facilitate the discussion of the procedures proposed for downdating.
The downdating procedures extract an eigenvalue from a matrix or pencil. A single step of the RQ algorithm applied to a matrix shifted by an eigenvalue allows deflation of this eigenvalue. Section 9.6.1 discusses the RQ algorithm in more detail and addresses the difficulties arising from its computation in finite precision. Based on this exposition

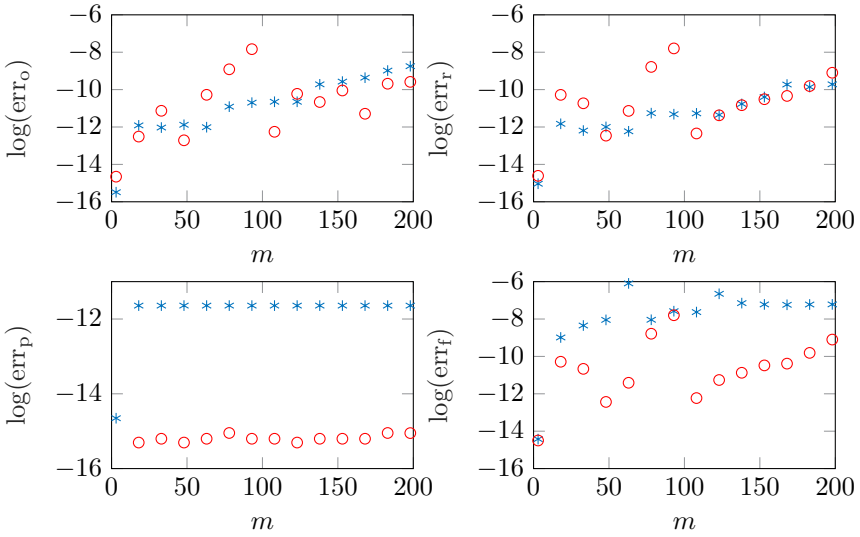



Figure 9.15: TPIEP with equidistant nodes on an ellipse $x^2 + (y/0.01)^2 = 1$ and poles $\Xi = \Theta$ on a circle with radius 3. Error metrics for Krylov '$\circ$' and updating '$*$' procedure in log scale for problem size $m$.

Section 9.6.2 proposes two procedures to downdate the HIEP. Section 9.6.3 performs numerical experiments for the proposed procedures. The generalization of the RQ algorithm to matrix pencils is the backward RQZ algorithm, which is the subject of Section 9.6.4. A single step of the backward RQZ algorithm applied to a pencil shifted by one of its eigenvalues allows deflation of this eigenvalue. Section 9.6.5 uses this property to downdate a Hessenberg pencil, two new procedures are proposed. A publication with these results is in preparation [164].

### 9.6.1 RQ algorithm

The QR algorithm is an algorithm for the computation of eigenvalues of a matrix $A \in \mathbb{C}^{m\times m}$. Usually, a preliminary step reduces $A$ to a proper Hessenberg matrix, if the Hessenberg matrix is not proper, it can be split into two (or more) smaller problems for proper Hessenberg matrices. The QR algorithm is thus applied to proper Hessenberg matrices. In our context the solution to a HIEP is already a proper Hessenberg matrix, the QR algorithm can be applied immediately.
A single step of the QR algorithm with shift consists of computing the QR

decomposition of the proper Hessenberg matrix $H$ shifted by $\lambda I$ for some $\lambda \in \mathbb{C}$

$$H - \lambda I = QR$$

and then recombining these factors to obtain the Hessenberg matrix

$$\hat{H} = RQ + \lambda I.$$

This performs a unitary similarity transformation on the matrix $H$, since $\hat{H} = RQ + \lambda I = Q^H(H - \lambda I)Q + \lambda I = Q^H HQ$. Note that $Q$ is a unitary Hessenberg matrix. For details on the QR algorithm and the connection to Krylov subspaces, we refer to [178].
The RQ algorithm applies the same idea, but now the RQ decomposition of the shifted matrix is used,

$$H - \lambda I = R\tilde{Q}^H,$$

the unitary factor in the RQ decomposition is denoted by a complex conjugated matrix $\tilde{Q}^H$, this is merely convention and useful in the following discussion. The unitary factor $\tilde{Q}^H$ is a unitary Hessenberg matrix. The recombination is

$$\hat{H} = \tilde{Q}^H R + \lambda I$$

and $\hat{H} = \tilde{Q}^H H \tilde{Q}$.
The theoretical and numerical behavior of the RQ step will be described in more detail below. Some known results are provided together with the proof if it contributes to the understanding of the downdating procedure in finite arithmetic.

### Perfect shift

When a proper, simple Hessenberg matrix $H \in \mathbb{C}^{m\times m}$ is shifted by one of its eigenvalues, it will be called a *perfect shift*. In exact arithmetic a single RQ step with perfect shift, $\lambda \in \sigma(H)$, allows to deflate $\lambda$ from $\hat{H} = \tilde{Q}^H H \tilde{Q}$. More precisely $\hat{H}e_1 = \lambda e_1$, where $\lambda \in \sigma(H)$. Lemma 9.7 leads to a result on invariant subspaces required to explain the deflation of the perfect shift.

**Lemma 9.7** (RQ step with perfect shift [178])**.** *Let $H \in \mathbb{C}^{m\times m}$ be a proper, simple Hessenberg matrix and $\lambda \in \sigma(H)$. Then the triangular factor $R$ in its shifted RQ decomposition $H - \lambda I = R\tilde{Q}^H$ satisfies $r_{1,1} := e_1^\top R e_1 = 0$ .*

*Proof.* The shifted Hessenberg matrix $H - \lambda I$ is singular. Hence, the upper triangular matrix $R \in \mathbb{C}^{m\times m}$ from the RQ decomposition

$$H - \lambda I = R\tilde{Q}^H$$

must also be singular and therefore a diagonal element $r_{i,i}$ must satisfy $r_{i,i} = 0$ for some $i$. Since $H - \lambda I$ is Hessenberg, its last $m-1$ rows are linearly independent, from which it follows that these rows of $R$ are also linearly independent and therefore $r_{1,1} = 0$. □

From Lemma 9.7 it follows that a single RQ step with perfect shift leads to the decomposition

$$H - \lambda I = R\tilde{Q}^H = \begin{bmatrix} 0 & r_{1,2} & r_{1,3} & \dots & r_{1,m} \\ & r_{2,2} & r_{2,3} & \dots & r_{2,m} \\ & & r_{3,3} & \dots & r_{3,m} \\ & & & \ddots & \vdots \\ & & & & r_{m,m} \end{bmatrix} \underbrace{\begin{bmatrix} \tilde{q}_1^H \\ \tilde{q}_2^H \\ \vdots \\ \tilde{q}_m^H \end{bmatrix}}_{=: \begin{bmatrix} \tilde{q}_1^H \\ \dot{Q}^H \end{bmatrix}} \tag{9.24}$$

with $\dot{Q} \in \mathbb{C}^{(m-1)\times m}$, $\dot{Q}\dot{Q}^H = I$ and $\dot{Q}\tilde{q}_1 = \mathbf{0}$. Therefore $H\tilde{q}_1 = \lambda\tilde{q}_1$, i.e., $\tilde{q}_1$ is the eigenvector corresponding to $\lambda$. The eigenvector $\tilde{q}_1$ forms an invariant subspace under $H$ of dimension 1. The discovery of an invariant subspace allows deflation.

### Deflation

Once an invariant subspace of $H$ is found, $\hat{H}$ will have a zero on its subdiagonal. This allows deflation of the eigenvalue problem, $\hat{H}$ is split into two matrices at the location of the zero. Theorem 9.5 provides the details for a normal Hessenberg matrix.

**Theorem 9.5** (Isolate eigenvalue using perfect shift RQ [178])**.** *Consider a proper, simple, normal Hessenberg matrix $H \in \mathbb{C}^{m\times m}$ and one of its eigenvalues $\lambda$. Let $R$ be an upper triangular and $\tilde{Q}^H$ a unitary matrix such that $H - \lambda I = R\tilde{Q}^H$. Then for a proper Hessenberg matrix $\tilde{H} \in \mathbb{C}^{(m-1)\times(m-1)}$,*

$$\hat{H} = \tilde{Q}^H R + \lambda I = \tilde{Q}^H H \tilde{Q} = \begin{bmatrix} \lambda & \mathbf{0} \\ \mathbf{0} & \tilde{H} \end{bmatrix}.$$

*Proof.* Let $X, \Lambda$ be the factors of the eigenvalue decomposition of the normal matrix $H$, i.e., $H = X\Lambda X^H$, $X^H X = I$ and $\Lambda = \mathrm{diag}(\{\lambda_i\}_{i=1}^m)$. From Lemma 9.7, more precisely (9.24), we have that $\tilde{q}_1 = x_k$ for some $k$. Now, by relying on the orthogonality of the

eigenvectors we obtain

$$\tilde{q}_1^H X = \tilde{q}_1^H \begin{bmatrix} | & & | & | & | & & | \\ x_1 & \dots & x_{k-1} & x_k & x_{k+1} & \dots & x_m \\ | & & | & | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{bmatrix} = e_k^\top,$$

and

$$\dot{Q}^H X = \dot{Q}^H \begin{bmatrix} | & & | & | & | & & | \\ x_1 & \dots & x_{k-1} & x_k & x_{k+1} & \dots & x_m \\ | & & | & | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & & | & & | & & | \\ \times & \dots & \times & \mathbf{0} & \times & \dots & \times \\ | & & | & & | & & | \end{bmatrix}$$

This isolates the eigenvalue $\lambda$ corresponding to the eigenvector $\tilde{q}_1$

$$\tilde{Q}^H R + \lambda I = \tilde{Q}^H H \tilde{Q} = \tilde{Q}^H X \Lambda X^H \tilde{Q} = \begin{bmatrix} \lambda & \mathbf{0} \\ \mathbf{0} & \tilde{H} \end{bmatrix},$$

where $\tilde{H} \in \mathbb{C}^{(m-1)\times(m-1)}$ is a proper Hessenberg matrix with spectrum $\sigma(\tilde{H}) = \sigma(H)\backslash\{\lambda\}$. □

The isolation of the perfect shift $\lambda$ in $\hat{H} = \tilde{Q}^H H \tilde{Q}$ is the key to downdating solutions to HIEPs. Before proposing downdating procedures based on the perfect shift RQ step, the computation of $\hat{H}$ in finite precision is discussed.

### Numerical computation

In finite precision, blurring can occur, which obstructs deflation [129, 180]. The computed $\hat{H} = \tilde{Q}^H H \tilde{Q}$ might have a first column that differs significantly from $\lambda e_1$, depending on the condition number of the eigenvalue $\lambda$. The perfect shift RQ step described above can be implemented in three different ways [129], these are stated in Theorem 9.6. Mathematically they are equivalent, but numerically the choice matters.

**Theorem 9.6** (Equivalencies for RQ step [129])**.** *Let $H \in \mathbb{C}^{m\times m}$ be a proper Hessenberg matrix and $\lambda \in \sigma(H)$. Then the following hold:*

1. *$H$ has a normalized eigenvector $x$ corresponding to $\lambda$*

$$Hx = \lambda x, \quad \|x\|_2 = 1$$

*which is unique up to unimodular scaling and $e_m^\top x \neq 0$.*

2. *An essentially unique sequence of* $\{C_i\}_{i=1}^{m-1}$, $C_i \in \mathfrak{C}_i$ *forming the matrix*

$$\tilde{Q}^H := C_1 C_2 \dots C_{m-1}$$

*exists that transforms the pair* $(H, x)$ *to a similar one*

$$(\hat{H}, \hat{x}) := (\tilde{Q}^H H \tilde{Q}, \tilde{Q}^H x)$$

*with*

$$\hat{x} = \alpha e_1, \quad |\alpha| = 1, \quad \hat{H} e_1 = \lambda e_1 \text{ and } \hat{H} \text{ is a Hessenberg matrix.}$$

3. *The Hessenberg matrix* $H - \lambda I$ *has the RQ decomposition*

$$H - \lambda I = R \tilde{Q}^H,$$

*where* $e_1^\top R e_1 = 0$ *and* $\tilde{Q}^H$ *is essentially the same matrix as the one transforming* $x$ *to* $\hat{x}$.

Ammar et al. [1] proposed to compute the matrix $\tilde{Q}^H$ by one step of the shifted RQ algorithm, corresponding to 3 in Theorem 9.6. They studied the case where the solution to a unitary Hessenberg matrix IEP is downdated, but the idea applies to the general HIEP as well. This method will be referred to as the *matrix method*. Mastronardi and Van Dooren [129] proposed a method based on 2 in Theorem 9.6. Their procedure leads to a more accurate isolation of the given eigenvalue, on the condition that the eigenvector is computed with sufficient accuracy. This procedure will be called the *eigenvector method*.
Assume that the given eigenvalue $\lambda$ is such that the smallest singular value of $H - \lambda I$ equals $\epsilon_{\text{mach}} \|H - \lambda I\|_2$, denoted as

$$s_{\min}(H - \lambda I) = \epsilon_{\text{mach}} \|H - \lambda I\|_2. \tag{9.25}$$

If this is not the case, the given $\lambda$ is not a suitable candidate to perform a perfect shift RQ step. If the condition is satisfied, then an eigenvector $\dot{x}$ can be computed which guarantees

$$\begin{aligned} &\tilde{Q}^H (H + \Delta_H) \tilde{Q} = \hat{H}, \quad \|\Delta_H\|_2 \le c \epsilon_{\text{mach}} \|H\|_F, \quad \Delta_H \text{ is a Hessenberg matrix,} \\ &\tilde{Q}^H (\dot{x} + \Delta_{\dot{x}}) = e_1, \quad \|\Delta_{\dot{x}}\|_2 \le c \epsilon_{\text{mach}}, \quad \text{with } c \text{ a constant of order 1,} \end{aligned} \tag{9.26}$$

with $\tilde{Q}^H := C_1 C_2 \dots C_{m-1}$, where $C_i \in \mathfrak{C}_i$ is exactly unitary and eliminates the element $e_{i+1}^\top \prod_{k=i+1}^{m-1} C_k \dot{x}$. This implies that $(\lambda, \dot{x})$ is an exact eigenpair of $\hat{H}$, after setting $\hat{h}_{1,1} = \lambda$ and $\hat{h}_{2,1} = 0$, up to a Hessenberg structured perturbation $\Delta_{\hat{H}}$

satisfying $\|\Delta_{\hat{H}}\|_F \leq c\epsilon_{\text{mach}}\|H\|_F$.
Such an eigenvector $\dot{x}$ can be computed as described below. Details about the error analysis can be found in [129]. To obtain a suitable eigenvector $\dot{x}$, a diagonal scaling is applied to balance the elements in the available approximate eigenvector $x$, an inverse iteration with the scaled vector $Dx$ is applied and the obtained vector is scaled back. This procedure is described in Algorithm 12.

**Algorithm 12** Compute eigenvector for eigenvector method [129]

**Input:** An approximate eigenpair $(\lambda, x)$, $x \in \mathbb{C}^m$ of Hessenberg matrix $H \in \mathbb{C}^{m\times m}$, which satisfies $s_{\min}(H-\lambda I) = \epsilon_{\text{mach}}\|H-\lambda I\|_2$, with $s_{\min}(A)$ the smallest singular value of matrix $A$.
**Output:** An approximate eigenvector $\dot{x} \in \mathbb{C}^m$ of $H$ satisfying (9.26).
**procedure** ACCURATEEIGENVECTOR($H, \lambda, x$)
    $d := \max\left(\min\left[\max_{i\leq m-2}\left\{|x_i/x_{m-1}|^{1/m-i-1}\right\}, \max_{i\leq m-2}\left\{|x_i/x_m|^{1/m-i}\right\}\right], 1\right)$
    Round $d$ to the nearest power of 2
    $D := \text{diag}(1, d, d^2, \ldots, d^{m-1})$
    $H_D := DHD^{-1}$, $x_D = \frac{Dx}{\|Dx\|_2}$
    One step of inverse iteration: $X_D = (H_D - \lambda I)\hat{x}_D$
    $\hat{x}_D = \frac{\hat{x}_D}{\|\hat{x}_D\|_2}$
    $\dot{x} := \frac{D^{-1}\hat{x}_D}{\|D^{-1}\hat{x}_D\|_2}$ ▷ Scale eigenvector back
**end procedure**

The matrix method and eigenvector method for downdating solutions to HIEPs are proposed in the next section.

## 9.6.2 Downdating HIEP

The downdating problem for HIEPs is introduced and it is shown that a RQ step with perfect shift provides a solution to this problem. Suppose $H_m \in \mathbb{C}^{m\times m}$ is the solution to a HIEP, the downdated HIEP removes the node-weight pair $z^\star, w^\star$. Let $z^\star = z_k$ then the downdated HIEP is Problem 9.7 with

$$\tilde{\Lambda} = \begin{bmatrix} z_1 & & & & & & \\ & \ddots & & & & & \\ & & z_{k-1} & & & & \\ & & & z_{k+1} & & & \\ & & & & \ddots & & \\ & & & & & z_m \end{bmatrix}, \quad \tilde{v} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{k-1} \\ \alpha_{k+1} \\ \vdots \\ \alpha_m \end{bmatrix}.$$

The available solution consists of the proper Hessenberg matrix $H_m \in \mathbb{C}^{m\times m}$ with simple eigenvalues and $Q_m \in \mathbb{C}^{m\times m}$ satisfying $Q_m^H Q_m = I$, $Q_m e_1 = \frac{v}{\|v\|}$. The last property will be used in the equivalent form: $Q_m^H v = \|v\| e_1$.
The solution of the downdated HIEP, with node $z^\star = z_k \in \sigma(H_m)$ and weight $\alpha^\star = \alpha_k$, consists of $H_{m-1} \in \mathbb{C}^{(m-1)\times(m-1)}$ and $Q_{m-1} \in \mathbb{C}^{(m-1)\times(m-1)}$ satisfying

$$\sigma(H_{m-1}) = \sigma(H_m)\backslash\{z^\star\}, \quad Q_{m-1}^H Q_{m-1} = I, \quad Q_{m-1} e_1 = \frac{\tilde{v}}{\|\tilde{v}\|}, \tag{9.27}$$

where $\tilde{v} = \begin{bmatrix}\alpha_1 & \dots & \alpha_{k-1} & \alpha_{k+1} & \dots & \alpha_m\end{bmatrix}^\top$.

### Mathematical solution

The solution (9.27) will be obtained by computing $\tilde{Q} \in \mathbb{C}^{m\times m}$ satisfying

$$\tilde{Q}^H H_m \tilde{Q} = \begin{bmatrix} z^\star & \\ & H_{m-1}\end{bmatrix}, \quad \tilde{Q}^H \tilde{Q} = I, \tag{9.28}$$

$$\tilde{Q}^H Q_m^H v = \hat{\alpha} e_1 + \eta e_2 = \begin{bmatrix}\hat{\alpha}\\ \eta \\ 0 \\ \vdots \\ 0\end{bmatrix}, \quad \text{with } |\hat{\alpha}| = |\alpha^\star|, \quad |\eta| = \|\tilde{v}\|_2. \tag{9.29}$$

The solution $H_{m-1} \in \mathbb{C}^{(m-1)\times(m-1)}$ is obtained by deflating $z^\star$ from $\tilde{Q}^H H_m \tilde{Q}$. Clearly, since $z^\star$ is known and is an eigenvalue of $H_m$, the RQ step with perfect shift provides the solution to this downdated HIEP. The conditions (9.28) are satisfied by the matrices in the RQ step, see Theorem 9.5. Condition (9.29) reveals why a RQ step is used instead of a QR step.
The matrix $\tilde{Q}^H$ is a unitary Hessenberg matrix if it is computed as $H - z^\star I = R\tilde{Q}^H$, and this results in

$$\tilde{Q}^H Q_m^H v = \tilde{Q}^H \begin{bmatrix}\|v\|\\ 0 \\ 0 \\ \vdots \\ 0\end{bmatrix} = \begin{bmatrix}\hat{\alpha}\\ \eta \\ 0 \\ \vdots \\ 0\end{bmatrix},$$

with $\hat{\alpha}, \eta$ as in (9.29). If, on the other hand, the QR decomposition is used $H - z^\star I = QR$, then $Q$ is a unitary Hessenberg matrix and $Q^H Q_m^H v = \begin{bmatrix}\times & \times & \dots & \times\end{bmatrix}^\top$, i.e., a vector full of generic nonzeros. Thus the QR decomposition does not (immediately) lead to an appropriate solution to the downdated HIEP.
Next the matrix and eigenvector method, discussed in Section 9.6.1, are introduced for solving the downdated HIEP.

### Numerical procedures

Algorithm 13 provides the matrix method, since usually the QR algorithm is available, but not the RQ algorithm, the algorithm uses matrix transformations such that the QR algorithm can be used to apply an RQ step. Algorithm 14 is an auxiliary procedure

**Algorithm 13** Downdating procedure: matrix method

1: **Input:** Proper Hessenberg matrix $H_m \in \mathbb{C}^{m\times m}$, node $z^\star \in \sigma(H_m)$.
2: **Output:** Proper Hessenberg matrix $H_{m-1} \in \mathbb{C}^{(m-1)\times(m-1)}$, such that $\sigma(H_{m-1}) = \sigma(H_m)\backslash z^\star$.
3: **procedure** DOWNDATEHIEP_MATRIX($H_m, z^\star$)
4: $J = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{C}^{m\times m}$
5: Compute QR decomposition of $J(H - z^\star I)^\top J =: QR$
6: $Q := JQ^\top J$, $R := JR^\top J$
7: $H := QR + z^\star I$
8: $\begin{bmatrix} \times & \\ & H_{m-1} \end{bmatrix} := H$
9: **end procedure**

used in the eigenvector method proposed in Algorithm 15.

**Algorithm 14** Compute core transformation

1: **Input:** $a, b \in \mathbb{C}$.
2: **Output:** $C \in \mathbb{C}^{2\times 2}$ such that $C \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \hat{a} \\ 0 \end{bmatrix}$, with $|\hat{a}| = \sqrt{|a|^2 + |b|^2}$.
3: **procedure** CORETRANSFORMATION($a, b$)
4: $\eta := |a| + |b|\sqrt{|\frac{a}{|a|+|b|}|^2 + |\frac{b}{|a|+|b|}|^2}$
5: $c := \frac{|a|}{\eta}$
6: $s = \frac{a\bar{b}}{|a|\eta}$
7: **end procedure**

For the actual implementation of Algorithm 15, the condition on the eigenvector $x$, specified in [129, Theorem 4.2], is checked. If the given eigenvector $x$ satisfies the condition, no scaling is needed and the core transformations forming $\tilde{Q}^H$ reduce the eigenvector $x$, i.e, Step 4 is skipped and $\dot{x} := x$. If the condition for $x$ is not satisfied, Step 4 will compute a suitable eigenvector $\dot{x}$.

**Algorithm 15** Downdating procedure: eigenvector method

**Input:** Proper Hessenberg matrix $H_m \in \mathbb{C}^{m\times m}$, node $z^\star \in \sigma(H_m)$, approximate eigenvector $x \in \mathbb{C}^m$ corresponding to $z^\star$.
**Output:** Proper Hessenberg matrix $H_{m-1} \in \mathbb{C}^{(m-1)\times(m-1)}$, such that $\sigma(H_{m-1}) = \sigma(H_m)\backslash z^\star$.
**procedure** DOWNDATEHIEP_EIGENVECTOR($H_m, z^\star, x$)
    $\dot{x} = \text{accurateEigevector}(H_m, z^\star, x)$
    $H := H_m$
    **for** $i = m : -1 : 2$ **do**
        $\hat{C} := \text{CoreTransformation}(\dot{x}_{i-1}, \dot{x}_i)$
        $C := I_{i-2} \oplus \hat{C} \oplus I_{m-i}$
        $\dot{x} = C\dot{x}$
        $H = CHC^H$
    **end for**
    $\begin{bmatrix} \times & \\ & H_{m-1} \end{bmatrix} := H$
**end procedure**

### 9.6.3 Numerical experiments

Numerical experiments are performed for the matrix and eigenvector method. The setup is as follows. A solution $H_m, Q_m \in \mathbb{C}^{m\times m}$ is available to a HIEP, with weights $\{\alpha_i\}_{i=1}^m$ all equal to 1 and some nodes $\{z_i\}_{i=1}^m$. The nodes are imposed by the experiments below. One by one nodes are downdated from $H_m$, resulting in $H_{m-l}$, $l = 1, 2, \ldots, m/2$, we will downdate half of the nodes. The error metrics $\text{err}_\text{o}$, $\text{err}_\text{r}$ and $\text{err}_\text{f}$ from Section 9.5 are used. For the current situation, $(B, C) = (H_{m-l}, I)$, $V = W = Q_{m-l}$ and $r_i = p_i \in \mathcal{P}_i$. An additional metric, which quantifies how well the first column of $Q_{m-l}$ represents the weight vector $\tilde{v} := \begin{bmatrix}1 & \ldots & 1\end{bmatrix}^\top \in \mathbb{C}^{m-l}$, is used

$$\text{err}_\text{w} := \|\|\tilde{v}\|_2 Qe_1 - \tilde{v}\|_2.$$

Three choices of nodes are discussed: nodes equidistant on the unit circle, Chebyshev nodes and random nodes inside of the unit circle. The former two are the same as used in Section 9.5, details can be found there.
The initial solution $H_m, Q_m$ is computed via the Arnoldi iteration. The eigenvector method requires the eigenvector corresponding to the node that will be downdated. This eigenvector is computed by a QR step with perfect shift and iterative refinement is applied until a relaxed version of condition (9.25) is satisfied:

$$s_{\min}(H - \lambda I) = 100\epsilon_{\text{mach}}\|H - \lambda I\|_2.$$

A thorough analysis of the effect of this relaxation is future research.

### Nodes on the unit circle

For $m$ nodes equidistant on the unit circle, $m/2$ nodes are downdated in the reverse order as explained for updating in Section 9.5. Figure 9.16 shows the downdating of these nodes as performed by the matrix and eigenvector method. The errors for the

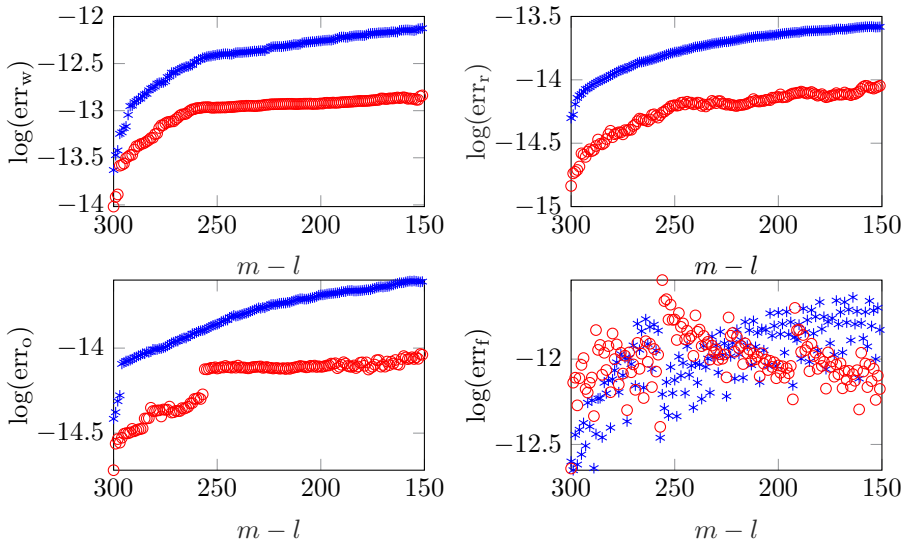



Figure 9.16: Error metrics for downdating with $m = 300$ nodes equidistant on the unit circle. Blue '$\star$' corresponds to the matrix method and red '$\circ$' to the eigenvector method.

matrix method increase more quickly than for the eigenvector method. The eigenvector method, as it is implemented for this experiment, is more expensive than the matrix method.

### Chebyshev nodes

Figure 9.17 shows the downdating of a Jacobi matrix $H_m$ for Chebyshev nodes on the interval $[-1, 1]$. The downdating is performed in the reverse order as described for the Chebyshev nodes for updating in Section 9.5. The error $\mathrm{err_f}$ is significantly larger for this experiment than for the first experiment. This can be explained by the condition of polynomial evaluation for the matrices $H_{m-l}$ in the nodes.

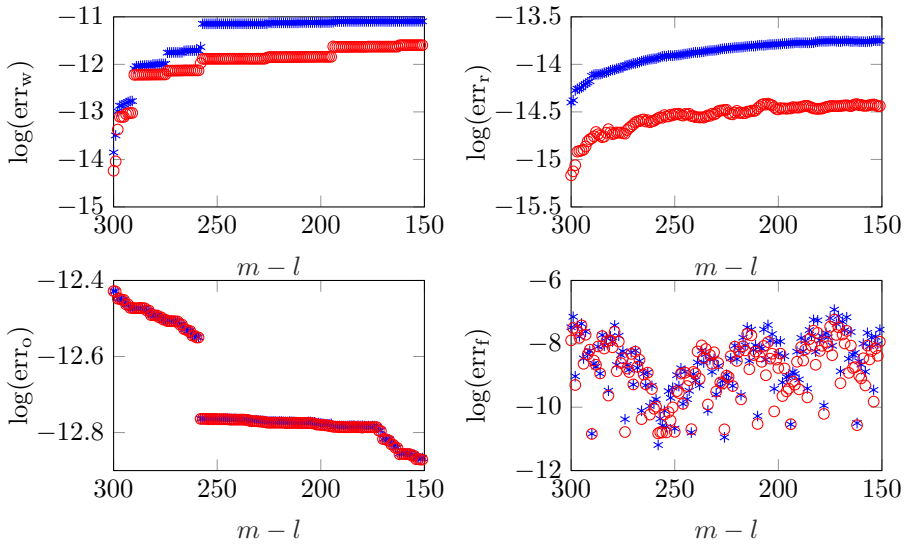



Figure 9.17: Error metrics for downdating with $m = 300$ Chebyshev nodes. Blue '$\star$' corresponds to the matrix method and red '$\circ$' to the eigenvector method.

### Complex plane

The final experiment tests a more general case, $m$ nodes are generated in the complex plane, lying inside the unit circle. Then $m/2$ nodes are downdated, in no particular order. Figure 9.18 shows the metrics. Again the eigenvector method performs slightly better. The conditioning of the corresponding polynomial evaluation is very large, we have omitted the metric $\mathrm{err_f}$ for this reason.

### Conclusion

The eigenvector method performs better than the matrix method. The current implementation of the eigenvector method is more expensive than the matrix method. Both methods show promising results for these simple experiments, but a more thorough study to their finite precision behavior is required. This is future research.

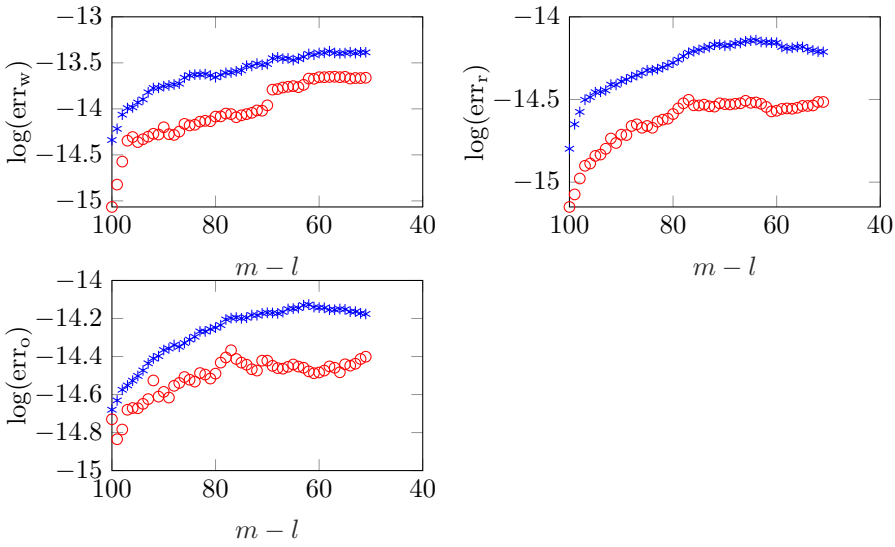



Figure 9.18: Error metrics for downdating with $m = 100$ nodes randomly chosen inside the unit circle. Blue '$\star$' corresponds to the matrix method and red '$\circ$' to the eigenvector method.

### 9.6.4 RQZ algorithm

The RQZ algorithm is a generalization of the QR algorithm that operates on Hessenberg pencils. This suggests that it might be suited to downdate HPIEPs. The (backward) RQZ step [40] performs nested subspace iterations and consists of essentially 2 operations acting on Hessenberg pencils, *pole swapping* and *pole changing*. These operations are discussed and the backward RQZ step with perfect shift is stated.
Some terminology is introduced to facilitate the discussion of the pole swapping and changing operations that are discussed below. The pair of subdiagonal elements $(h_{i+1,i}, k_{i+1,i})$ of a Hessenberg pencil $(H, K) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ will be referred to as *pole position* $i$. The poles $\Xi = \{\xi_j\}_{j=1}^{m-1}$ are distributed, in factored form $\xi_j = \frac{a_j}{b_j}$ over the pole positions. If $\xi_j = \frac{a_j}{b_j}$ appears on pole position $l$, then $h_{l+1,l} = a_j$ and $k_{l+1,l} = b_j$. Note that this is a break from the usual convention in this thesis that $\xi_i = \frac{h_{i+1,i}}{k_{i+1,i}}$, i.e., that $\xi_i$ always appears on pole position $i$.

### Pole swapping

A pole swap is the interchanging of poles on two neighboring pole positions, i.e., $\xi_i$ appearing on pole position $i$ and $\xi_{i+1}$ on position $i+1$ are interchanged such that $\xi_i$ now appears on pole position $i+1$ and $\xi_{i+1}$ on position $i$. Lemma 9.8 states that swapping poles can be performed by core transformations.

**Lemma 9.8** (Pole swapping [18, 40])**.** *Consider the proper Hessenberg pencil with poles $\xi_1 = \frac{h_{1,1}}{k_{1,1}}, \xi_2 = \frac{h_{2,2}}{k_{2,2}} \in \overline{\mathbb{C}}$ on pole position 1 and 2, respectively,*

$$H = \begin{bmatrix} h_{1,1} & h_{1,2} \\ & h_{2,2} \end{bmatrix}, \quad K = \begin{bmatrix} k_{1,1} & k_{1,2} \\ & k_{2,2} \end{bmatrix}.$$

*Then core transformations $C, \dot{C} \in \mathbb{C}^{2\times 2}$ can be constructed such that*

$$C^H H \dot{C} = C^H \begin{bmatrix} h_{1,1} & h_{1,2} \\ & h_{2,2} \end{bmatrix} \dot{C} = \begin{bmatrix} \tilde{h}_{1,1} & \tilde{h}_{1,2} \\ & \tilde{h}_{2,2} \end{bmatrix},$$

$$C^H K \dot{C} = C^H \begin{bmatrix} k_{1,1} & k_{1,2} \\ & k_{2,2} \end{bmatrix} \dot{C} = \begin{bmatrix} \tilde{k}_{1,1} & \tilde{k}_{1,2} \\ & \tilde{k}_{2,2} \end{bmatrix},$$

*where $\frac{\tilde{h}_{2,2}}{\tilde{k}_{2,2}} = \xi_1$ and $\frac{\tilde{h}_{1,1}}{\tilde{k}_{1,1}} = \xi_2$, i.e., $\xi_1$ appears on pole position 2 and $\xi_2$ appears on pole position 1.*

*Proof.* The procedure consists of three steps

1. Compute $M := k_{2,2}\hat{H} - h_{2,2}\hat{K}$

$$M = \begin{bmatrix} k_{2,2}h_{1,1} - h_{2,2}k_{1,1} & k_{2,2}h_{1,2} - h_{2,2}k_{1,2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \times & \times \\ 0 & 0 \end{bmatrix}.$$

2. Compute $\dot{C}$ such that $M\dot{C} = \begin{bmatrix} 0 & \times \\ 0 & 0 \end{bmatrix}$ which implies that

$$\operatorname{rank}\left(\begin{bmatrix} \hat{H}\dot{C}e_1 & \hat{K}\dot{C}e_1 \end{bmatrix}\right) = 1,$$

   i.e., the first column of $\hat{H}\dot{C}$ and $\hat{K}\dot{C}$ are colinear.

3. Compute $C$ such that $C^H\hat{H}\dot{C} = \begin{bmatrix} \times \\ 0 \end{bmatrix}$, from which immediately follows that $C^H\hat{K}\dot{C} = \begin{bmatrix} \times \\ 0 \end{bmatrix}$, i.e., simultaneous elimination of the element on position $(2,1)$.

The resulting matrices $\tilde{H} = C^H H \dot{C}$ and $\tilde{K} = C^H K \dot{C}$ satisfy the stated property. □

Note that this procedure corresponds to the structure restoring procedure that is described in the proof of Lemma 9.3 in the context of updating a HPIEP. In order to apply Lemma 9.8 to a Hessenberg pencil $(H, K) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ of size $m > 2$, plane rotations $C_i, \dot{C}_i \in \mathfrak{C}_i$ can be constructed with as its parameters those elements appearing in $C, \dot{C}$ obtained from Lemma 9.8 applied to the subpencil $\left(\begin{bmatrix} h_{i,i-1} & h_{i,i} \\ & h_{i+1,i} \end{bmatrix}, \begin{bmatrix} k_{i,i-1} & k_{i,i} \\ & k_{i+1,i} \end{bmatrix}\right)$. Let $\xi = \frac{h_{i,i-1}}{k_{i,i-1}}$ appear on pole position $i-1$ and $\theta = \frac{h_{i+1,i}}{k_{i+1,i}}$ on pole position $i$. Then the pencil $(C_i^H H \dot{C}_i, C_i^H K \dot{C}_i)$ will have $\theta$ on pole position $i-1$ and $\xi$ on pole position $i$. All other poles in the pencil $(H, K)$ remain unaltered by the core transformations $C_i, \dot{C}_i$.

### Pole changing

Changing a pole is possible on the first pole position or the last, i.e., pole position 1 or $m-1$ for a pencil of size $m$. Lemma 9.9 states that changing a pole can be performed by a unitary similarity transformation.

**Lemma 9.9** (Changing pole - first pole position [18,40])**.** *Let $(H, K) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ be a proper Hessenberg pencil with poles $\Xi = \{\xi_1, \xi_2, \xi_3, \ldots, \xi_{m-1}\}$, where $\xi_i \in \bar{\mathbb{C}}$ appears on pole position i, i.e., $\xi_1 = \frac{h_{2,1}}{k_{2,1}}$. Let $\hat{\xi} = \frac{\hat{a}}{\hat{b}} \notin \sigma(H, K)$ be a given pole. Consider the core transformation $C_1 \in \mathfrak{C}_1$ which, for an arbitrary nonzero constant $\gamma$ and $x := \gamma(\hat{b}H - \hat{a}K)(k_{2,1}H - h_{2,1}K)^{-1}e_1$, satisfies*

$$C_1^H x = \alpha e_1.$$

*Premultiplication with $C_1$ introduces the pole $\hat{\xi}$ in pole position 1, removing $\xi_1$ in the process. That is, $(\hat{H}, \hat{K}) := (C_1^H H, C_1^H K)$ is a proper Hessenberg pencil with poles $\{\hat{\xi}, \xi_2, \xi_3, \ldots, \xi_{m-1}\}$.*

*Proof.* Since $H$ and $K$ are both Hessenberg matrices, $He_1 = h_{1,1}e_1 + h_{2,1}e_2$ and $Ke_1 = k_{1,1}e_1 + k_{2,1}e_2$. Therefore it is not required to compute any inverses since $(k_{2,1}H - h_{2,1}K)^{-1}e_1 = \frac{1}{k_{2,1}h_{1,1} - h_{2,1}k_{1,1}}e_1$. Thus

$$x = \frac{\hat{\gamma}}{k_{2,1}h_{1,1} - h_{2,1}k_{1,1}} \begin{bmatrix} \hat{b}h_{1,1} - \hat{a}k_{1,1} \\ \hat{b}h_{2,1} - \hat{a}k_{2,1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and clearly a core transformation $C_1 := \begin{bmatrix} \bar{a} & -\bar{b} & \\ b & a & \\ & & I_{m-2} \end{bmatrix}$ can be constructed such that $C_1^H x = \alpha e_1$. The core transformation is determined by $|a|^2 + |b|^2 = 1$ and $-b(\hat{b}h_{11} + \bar{a}k_{11}) + \bar{a}(\hat{b}k_{11} + \bar{a}h_{11}) = 0$. The latter equation can be rewritten as

$$\begin{aligned} -b(\hat{b}h_{11} + \bar{a}k_{11}) + \bar{a}(\hat{b}k_{11} + \bar{a}h_{11}) &= 0 \\ \Leftrightarrow \hat{b}(-bh_{11} + \bar{a}h_{2,1}) - \hat{a}(-bk_{11} + \bar{a}k_{21}) &= 0 \\ \Leftrightarrow \frac{-bh_{11} + \bar{a}h_{2,1}}{-bk_{11} + \bar{a}k_{2,1}} &\frac{\hat{a}}{\hat{b}}. \end{aligned}$$

The statement follows by noting that $\hat{h} = -bh_{11} + \bar{a}h_{2,1} = \hat{a}$ and $\hat{k} = -bk_{11} + \bar{a}k_{2,1} = \hat{b}$. □

The pole on the last pole position can also be changed, as stated in Lemma 9.10.

**Lemma 9.10** (Changing pole - last pole position [18, 40])**.** *Let $(H, K) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ be a proper Hessenberg pencil with poles $\Xi = \{\xi_1, \ldots, \xi_{m-2}, \xi_{m-1}\}$, where $\xi_i \in \bar{\mathbb{C}}$ appears on pole position $i$, i.e., $\xi_{m-1} = \frac{h_{m,m-1}}{k_{m,m-1}}$. Let $\hat{\xi} = \frac{\hat{a}}{\hat{b}} \notin \sigma(H, K)$ be a given pole. Consider the core transformation $\dot{C}_{m-1} \in \mathfrak{C}_{m-1}$ which, for an arbitrary nonzero constant $\gamma$ and $x^\top := \gamma e_m^\top (k_{m,m-1}H - h_{m,m-1}K)^{-1}(\hat{b}H - \hat{a}K)$, satisfies*

$$x^\top \dot{C}_{m-1} = \alpha e_m$$

*Postmultiplication with $C_1$ introduces the pole $\hat{\xi}$ in pole position $m-1$, removing $\xi_{m-1}$ in the process. That is, $(\hat{H}, \hat{K}) := (H\dot{C}_{m-1}, K\dot{C}_{m-1})$ is a proper Hessenberg pencil with poles $\{\xi_1, \ldots, \xi_{m-3}, \xi_{m-2}, \hat{\xi}\}$.*

*Proof.* Similar to the proof of Lemma 9.9. □

### Perfect shift RQZ

A backward RQZ step combines the pole changing and pole swapping operation to perform nested subspace iteration [40]. Consider a proper Hessenberg pencil $(H, K) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ with poles $\Xi = \{\xi_1, \ldots, \xi_{m-1}\}$, where $\xi_i$ appears on pole position $i$. For a given shift $\lambda = \frac{a}{b} \in \mathbb{C}$ a backward RQZ step consists of the following three steps

1. Introduce the shift $\lambda$ into the last pole position, removing $\xi_{m-1}$ in the process, see Lemma 9.10.

2. Swap poles, via the procedure in Lemma 9.8, until $\lambda$ appears on the first pole position.

3. Remove the shift $\lambda$, now on pole position 1, by reintroducing the removed pole $\xi_{m-1}$ by applying Lemma 9.9.

The backward RQZ step is described in Algorithm 16.

**Algorithm 16** Backward RQZ step

**Input:** Proper Hessenberg pencil $(H_m, K_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ with poles $\{\xi_i\}_{i=1}^{m-1}$, $\xi_i := \frac{h_{i+1,i}}{k_{i+1,i}}$ and a shift $\lambda = \frac{a}{b} \in \overline{\mathbb{C}}$.
**Output:** Hessenberg pencil $(\hat{H}, \hat{K}) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$, unitarily similar to $(H_m, K_m)$, with $\xi_i := \frac{\hat{h}_{i+2,i+1}}{\hat{k}_{i+2,i+1}}$ for $i = 1, 2, \ldots, m-2$.
**procedure** RQZ($H_m, K_m, \lambda$)
    $H := H_m$, $K := K_m$
    $\hat{C} :=$ CoreTransformation$(bh_{m,m} - ak_{m,m}, bh_{m-1,m} - ak_{m-1,m})$
    $\dot{C} := I_{m-2} \oplus \hat{C}$
    $H = H\dot{C}$, $K = K\dot{C}$ ▷ Introduces shift
    **for** $i = m-1 : -1 : 2$ **do** ▷ Swap shift to top
        $M := k_{i+1,i}H - h_{i+1,i}K$
        $\hat{C} :=$ CoreTransformation$(m_{1,2}, m_{1,1})$
        $\dot{C} := I_{i-2} \oplus \hat{C} \oplus I_{m-i}$
        **if** $\frac{h_{i+1,i}}{k_{i+1,i}} < \frac{h_{i,i-1}}{k_{i,i-1}}$ **then**
            $H = H\dot{C}, \quad K = K\dot{C}$
            $x := He_{i-1}$
        **else**
            $H = H\dot{C}, \quad K = K\dot{C}$
            $x := Ke_{i-1}$
        **end if**
        $\hat{C} :=$ CoreTransformation$(e_i^\top x, e_{i+1}^\top x)$
        $C := I_{i-1} \oplus \hat{C} \oplus I_{m-i-1}$
        $H = CH$, $K = CK$
    **end for**
    $\hat{C} :=$ CoreTransformation$(\tilde{k}h_{1,1} - \tilde{h}k_{1,1}, \tilde{k}h_{2,1} - \tilde{h}k_{2,1})$, with $\frac{\tilde{h}}{\tilde{k}} = \xi_{m-1}$
    $C := \hat{C} \oplus I_{m-2}$
    $H = CH$, $K = CK$ ▷ Remove shift
    $\hat{H} := H$, $\hat{K} := K$
**end procedure**

If the backward RQZ step is applied with a perfect shift, then deflation occurs in the first column of the resulting pencil. Theorem 9.7 states this result formally.

**Theorem 9.7** (Deflation for perfect shift backward RQZ [37, Theorem 3.8.1])**.** *Let $(H, K)$ be a $m \times m$ proper Hessenberg pencil with poles $\Xi = \{\xi_1, \xi_2, \ldots, \xi_{m-1}\}$, where the poles are not necessarily distinct from the eigenvalues. Let $\sigma \in \bar{\mathbb{C}}$ be an eigenvalue of $(H, K)$ with $\lambda \notin \Xi$. Then a backward RQZ step with perfect shift $\lambda$ deflates the first column.*

*Proof.* See [37], where it is given for the forward RQZ step. □

It should be noted that in finite precision blurring can obstruct the deflation, just as in the RQ step with perfect shift for Hessenberg matrices.

### 9.6.5 Downdating HPIEP

A backward RQZ step with perfect shift is suitable for downdating the solution to a HPIEP. This follows from Theorem 9.7. The solution $(H_m, K_m) \in \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times m}$ and $Q_m \in \mathbb{C}^{m\times m}$ to a HPIEP will be downdated. Suppose this HPIEP is formulated with

$$\Lambda = \operatorname{diag}(z_1, \ldots, z_m), \quad v = \begin{bmatrix}\alpha_1 & \ldots \alpha_m\end{bmatrix}^\top, \quad \Xi = \{\xi_1, \ldots, \xi_{m-1}\}.$$

The downdated HPIEP is obtained by removing a node-weight pair $z^\star, \alpha^\star$ and a pole $\xi^\star$ from the HPIEP formulation. Let $z^\star = z_k$, then the associated weight is $\alpha_k = \alpha^\star$. The pole $\xi^\star$ that will be removed can be chosen independently of the node-weight pair, let $\xi^\star = \xi_i$. Then the downdated HPIEP is the HPIEP for

$$\tilde{\Lambda} = \begin{bmatrix} z_1 & & & & & \\ & \ddots & & & & \\ & & z_{k-1} & & & \\ & & & z_{k+1} & & \\ & & & & \ddots & \\ & & & & & z_m \end{bmatrix}, \quad \tilde{v} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{k-1} \\ \alpha_{k+1} \\ \vdots \\ \alpha_m \end{bmatrix}^\top, \quad \tilde{\Xi} = \Xi \backslash \{\xi_i\}.$$

The downdating procedure consists of a preliminary step, bringing $\xi^\star$ into the last pole position, and a backward RQZ step with shift equal to $z^\star$. Since the shift

$\lambda = z^\star = z_k \in \sigma(H_m, K_m)$, by Theorem 9.7, the resulting pencil $(\widetilde{H}, \widetilde{K})$ has structure

$$\left( \begin{bmatrix} h & & & & & & \\ & \times & \times & \times & \dots & \times & \times \\ & \times & \times & \times & \dots & \times & \times \\ & & \ddots & \vdots & & \vdots & \vdots \\ & & & \times & \dots & \times & \times \\ & & & & \ddots & \vdots & \vdots \\ & & & & & \times & \times \end{bmatrix}, \begin{bmatrix} k & & & & & & \\ & \times & \times & \times & \dots & \times & \times \\ & \times & \times & \times & \dots & \times & \times \\ & & \ddots & \vdots & & \vdots & \vdots \\ & & & \times & \dots & \times & \times \\ & & & & \ddots & \vdots & \vdots \\ & & & & & \times & \times \end{bmatrix} \right),$$

where $\frac{h}{k} = z^\star$, the node to be downdated, and the set of poles of the $(m-1) \times (m-1)$ principal trailing submatrix equals $\widetilde{\Xi} = \Xi \backslash \xi^\star$. This structure implies that the eigenvalue $z^\star$ can be deflated and that the pole $\xi^\star$ is removed from the denominator of the rational functions. This procedure is in fact a generalization of the matrix method for downdating HIEPs to downdating HPIEPs.

A generalization of the eigenvector method can also be devised. In this case, the left and right eigenvectors $x_l^\star, x_r^\star$ of the pencil $(H_m, K_m)$ corresponding to an eigenvalue $z^\star$ are required. The two matrices, formed by a product of core transformations, reducing these eigenvectors to a multiple of the canonical unit vector, are mathematically the same as those obtained by the RQZ step with perfect shift. Developing this method further is future research.

## 9.7 Conclusion

Numerical procedures based on structured matrix theory are proposed to solve the problem of generating (bi)orthogonal rational functions or polynomials. The relation of (bi)ORFs to certain structured matrices or pencils allows the development of new procedures to solve this problem. The rational Arnoldi iteration can be applied to generate orthogonal rational functions and the novel rational Lanczos iteration to generate biorthogonal rational functions. Several new updating procedures are proposed, a parallelizable variant for orthogonal polynomials, numerically stable updating procedures for orthogonal rational functions based on the related Hessenberg pencil and an updating procedure for biorthogonal rational functions. For downdating of a sequence of orthogonal polynomials, two approaches are discussed. One approach, the eigenvector method, is new in the sense that it has not been applied in this context yet. The first steps are taken to generalize the downdating procedures to orthogonal rational functions.

# Chapter 10

# Conclusion and future research

In this manuscript a general biorthogonal framework is described to classify all the structured matrices and pencils appearing in the study of rational Krylov subspaces. These structured pencils contain the recurrence coefficients of (bi)orthogonal bases spanning rational Krylov subspaces. We showed that a tridiagonal recurrence pencil suffices to represent biorthogonal bases. This leads to the most efficient procedure to compute such bases. This procedure is introduced here and is called the *rational Lanczos iteration*.

The Gram matrices that appear for rational Krylov subspace exhibit certain displacement structure. We have shown that for biorthogonal rational Krylov subspaces the associated Gram matrix always has displacement rank at most 2. The displacement operators which reveal the displacement structure are derived and a procedure to construct such operators is described. Using these displacement operators, a Levinson procedure is used to derive a coupled short recurrence relation to generate biorthogonal bases for a specific pair of extended Krylov subspaces.

These structured matrices and pencils not only appear in the context of Krylov subspace methods, they are also recurrence pencils for sequences of (bi)orthogonal rational functions with prescribed poles. We identified which (bi)orthogonal rational functions are related to (bi)orthogonal bases for rational Krylov subspaces. Using this relation, procedures for rational Krylov subspaces can be applied to rational functions. Most notably, the rational Lanczos iteration is suitable to compute the recurrence coefficients of biorthogonal rational functions for a given linear functional.

The relation between structured recurrence pencils and (bi)orthogonal rational functions can also be exploited directly. We developed several updating procedures based on structured matrices to compute the recurrence coefficients of (bi)orthogonal

rational functions, if the corresponding inner product or linear functional is given. A parallelizable variant is derived for orthogonal polynomials. A numerically stable procedure is designed for orthogonal rational functions. And a procedure to generate biorthogonal rational functions is proposed.

Two procedures to compute the perfect shift QR step are applied to the problem of downdating orthogonal polynomials. One of these procedures is generalized for the application to downdating orthogonal rational functions. And the first step to generalize the other procedure to the downdating of ORFs is taken as well.

## 10.1 Contributions

An overview of the main contributions of this manuscript is provided below.

Chapter 3 contains mostly known results, three results are new:

- Lemma 3.7 provides a recurrence relation for orthogonal basis vectors of polynomial Krylov subspaces in terms of a rank structured matrix.
- An alternative proof is given for the structure of the recurrence matrix for biorthogonal basis vectors of polynomial Krylov subspaces, stated in Lemma 3.8.
- Theorem 3.1 generalizes a known result from normal matrices to diagonalizable matrices.

Chapter 4 generalizes some proofs appearing in literature and introduces the general framework classifying all structured matrices and pencils appearing for biorthogonal bases of rational Krylov subspaces. The contents of this chapter are published [165].

- Theorem 4.1 generalizes a known result from normal matrices to diagonalizable matrices and the proof provides a new decomposition of rational Krylov bases.
- Section 4.4 proves the structure of recurrence matrices and pencils which generate biorthogonal bases for rational Krylov subspaces. It also provides a short overview of notable special cases of these results appearing in the literature.

Chapter 5 contains a new result on the displacement structure of Gram matrices appearing in rational Krylov subspaces.

- Section 5.2 studies the displacement operators for Gram matrices arising from rational Krylov subspaces using the decomposition of rational Krylov bases. This leads to constructive proofs, which allows the construction of these displacement operators.

- Lemma 5.6 is a new result on the displacement structure of Gram matrices arising from rational Krylov subspaces. This lemma also implies the existence of three term recurrence relations for the construction of biorthogonal bases.

Chapter 6 and Chapter 7 identify the inner products and linear functionals for polynomials and rational functions that are related to the Euclidean inner product on (rational) Krylov subspaces.

- Lemma 6.5 is a new result for biorthogonal polynomials linking a linear functional to the Euclidean inner product on polynomial Krylov subspaces.
- Lemma 7.2 is a new result for biorthogonal rational functions, linking a linear functional to the Euclidean inner product on rational Krylov subspaces.

Chapter 8 contains the results of two published papers [165, 166].

- Section 8.1 describes the rational Lanczos iteration, whose derivation can be found in Appendix B.1. Its validity is tested by performing some numerical experiments.
- Section 8.2.1 derives new short recurrence relations for biorthogonal bases for CMV-like extended Krylov subspaces.
- Section 8.2.2 provides a new Levinson procedure for Gram matrices arising from CMV-like extended Krylov subspaces.
- Section 8.2.3 shows how the CMV decomposition can be retrieved as a special case when the matrix generating the extended Krylov subspaces is unitary.

Chapter 9 contains several new results. The Krylov and updating procedures are published [167]. The recursive updating procedure for polynomials is based on an article in preparation [164], as well as the downdating procedures.

- Problem 9.6 and Problem 9.8 are new formulations for the problem of computing (bi)orthogonal rational functions.
- Theorem 9.2 shows that the rational Lanczos iteration is suitable to construct biorthogonal rational functions.
- Section 9.3.2 introduces a solution procedure based on unitary similarity transformations that merges two Hessenberg matrices for representing polynomials orthogonal to some inner product to obtain a Hessenberg matrix for polynomials orthogonal with respect to the sum of these inner products.

- Section 9.3.3 proposes a procedure to update solutions to Hessenberg pencil inverse eigenvalue problems. This can be used to generate orthogonal rational functions.
- Section 9.4.2 proposes a procedure to update solutions to tridiagonal pencil inverse eigenvalue problems. This can be used to generate biorthogonal rational functions.
- Section 9.6 applies results from structured matrix error analysis to the problem of downdating polynomials and generalizes the idea to rational functions.

## 10.2 Future research

This manuscript answered some questions and generated new questions and possible directions for research. We list here some possible topics for future research.

- The proposed rational Lanczos iteration is a general purpose algorithm. A proper error analysis of this iteration is required to make it suitable for numerical computation. Restriction to specific problems can facilitate this analysis. For example, in moment matching for model order reduction the poles in both Krylov spaces are chosen to be equal and this facilitates the study of the iteration. An in-depth analysis of rational Lanczos iterations for state space equations for general matrices is a possible avenue. Another example is the tridiagonal pencil inverse eigenvalue problem. The Krylov subspaces are generated for a simple diagonal matrix. Preliminary numerical tests show that the rational Lanczos iteration performs well, considering that it has not been adapted to this specific case. An adaptation of the iteration to this case and corresponding error analysis is necessary to develop a robust and efficient algorithm to generate biorthogonal rational functions.
- The identification of the structure of Gram matrices appearing in rational Krylov subspaces allow to develop Levinson procedures for these Gram matrices. A specific instance is provided in this thesis. Does the knowledge of the displacement operators lead to a general Levinson procedure that can be used for any Gram matrix arising from rational Krylov subspaces?
- In numerical analysis, quadrature rules for highly oscillatory integrals use *kissing polynomials*. These polynomials are orthogonal with respect to a specific linear functional, i.e., they are biorthogonal polynomials. A stable procedure to generate these polynomials has yet to be developed. The updating procedure based on a tridiagonal inverse eigenvalue problem might provide a good starting point.

- The pseudo-Jacobi inverse eigenvalue problem is a specific instance of the tridiagonal IEP. Current procedures make use of the non-Hermitian Lanczos iteration. An adaptation of the updating procedure for tridiagonal IEPs might provide a more stable alternative to Lanczos-based procedures.

- Throughout this manuscript the connections between the Euclidean inner product and inner products or linear functionals for polynomials or rational functions are derived for diagonalizable matrices. If the matrix is not diagonalizable, its Jordan form will lead to an inner product or linear functional containing derivatives of these polynomials or rational functions. This is related to Sobolev orthogonality. Is it possible to develop algorithms, similar to those proposed here, for this case?

- The matrix method for downdating procedures is generalized from the polynomial case to the rational function case. The eigenvector method can also be generalized from the polynomial case to the rational function case. The necessary theory and error analysis must still be developed.

- In least squares problems sometimes new data becomes available and old data can be forgotten. If we possess a solution to a least squares problem, new data can be added efficiently by an updating procedure and old data can be removed efficiently by downdating. Removing old data and adding new data consecutively leads to a so-called sliding window scheme. Can the up-and downdating procedures in this manuscript be combined to develop a numerically stable sliding window scheme?

- Block Krylov methods are related to vector orthogonal functions. Can similar relationships, as presented in this manuscript, be derived and exploited for the block case?

- In this thesis the poles for the rational Krylov subspaces or rational functions are assumed to be given. A proper choice of poles depends on the application and is paramount to developing effective procedures. Two applications are solving least squares problems involving rational functions using the techniques presented here and model order reduction using the rational Lanczos iteration. Can these be combined effectively with the AAA algorithm [131] to choose the poles of the rational functions?

# Appendix A

# A collection of proofs

This appendix contains some alternative proofs, which are very similar to proofs already included in the manuscript. And proofs which are too technical and lengthy to be included in the manuscript itself.

## A.1 Proof Hermitial RatLan

The proof to Theorem 4.7 is given here.

*Proof.* From the rational Arnoldi iteration, in a simplified form, we have

$$h_{k+1,k}q_k = (A - \xi_k I)^{-1} A q_{k-1} - \sum_{i=0}^{k-1} h_{i,k} q_i,$$

which expresses $h_{k+1,k}q_k$ in the basis $\{(A-\xi_k I)^{-1}Aq_{k-1}, q_{k-1}, \ldots, q_0\}$ for the subspace $\mathcal{K}_{k+1}(A, v; \Xi)$. In matrix notation, for $Q_{k+1} := \begin{bmatrix} q_0 & \ldots q_k \end{bmatrix}$ and $\underline{I}_k = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$,

$$AQ_{k+1}\left(\underline{H}_k - \underline{I}_k\right) = Q_{k+1}\left(\underline{H}_k \begin{bmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_k \end{bmatrix}\right).$$

Express $\tilde{h}_{k+1,k}q_k$ in another basis $\{(A-\xi_k I)^{-1}Aq_{k-1}, \tilde{q}_{k-1}, \ldots, \tilde{q}_0\}$, where $\tilde{q}_i := (A-\xi_k)^{-1}(A-\xi_{i-1})q_i$ and $\xi_0 \neq 0$,

$$\begin{aligned}\tilde{h}_{k+1,k}q_k &= (A-\xi_k I)^{-1}Aq_{k-1} - \sum_{i=0}^{k-1}\tilde{h}_{i,k}\tilde{q}_i \\ &= (A-\xi_k I)^{-1}Aq_{k-1} - \sum_{i=0}^{k-1}\tilde{h}_{i,k}(A-\xi_k)^{-1}(A-\xi_{i-1})q_i.\end{aligned}$$

Now write it in matrix notation, use

$$\begin{aligned}(A-\xi_k I)\tilde{h}_{k+1,k}q_k &= Aq_{k-1} - (A-\xi_k I)\sum_{i=0}^{k-1}\tilde{h}_{i,k}(A-\xi_k)^{-1}(A-\xi_{i-1})q_i \\ &= Aq_{k-1} - \sum_{i=0}^{k-1}\tilde{h}_{i,k}(A-\xi_{i-1})q_i\end{aligned}$$

$$\Leftrightarrow AQ_{k+1}\begin{bmatrix}\tilde{h}_{1,k}\\ \vdots \\ \tilde{h}_{k-1,k}\\ \tilde{h}_{k,k}-1\\ \tilde{h}_{k+1,k}\end{bmatrix} = Q_{k+1}\begin{bmatrix}\xi_0\tilde{h}_{1,k}\\ \vdots \\ \xi_{k-2}\tilde{h}_{k-1,k}\\ \xi_{k-1}\tilde{h}_{k,k}-1\\ \xi_k\tilde{h}_{k+1,k}\end{bmatrix}$$

to obtain

$$AQ_{k+1}\left(\underline{\widetilde{H}}_k - \underline{I}_k\right) = Q_{k+1}\begin{bmatrix}\xi_0 & & \\ & \ddots & \\ & & \xi_k\end{bmatrix}\underline{\widetilde{H}}_k.$$

From this expression it is possible to show that the Hessenberg matrix $\widetilde{H}_k$ is Hermitian and, thus, tridiagonal. Assume that $\widetilde{H}_k$ is nonsingular, postmultiply by $\widetilde{H}_k^{-1}$,

$$Q_{k+1}^H A Q_{k+1} \left( \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ x_1 & \dots & x_k \end{bmatrix} - \begin{bmatrix} & \widetilde{H}_k^{-1} & \\ 0 & \dots & 0 \end{bmatrix} \right) = D_\xi \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ x_1 & \dots & x_k \end{bmatrix}$$

$$\left(Q_{k+1}^H A Q_{k+1} - D_\xi\right) \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ x_1 & \dots & x_k \end{bmatrix} = Q_{k+1}^H A Q_{k+1} \begin{bmatrix} & \widetilde{H}_k^{-1} & \\ 0 & \dots & 0 \end{bmatrix}$$

$$\underbrace{Q_{k+1}^H A^{-1} Q_{k+1} \left(Q_{k+1}^H A Q_{k+1} - D_\xi\right)}_{=:B} \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ x_1 & \dots & x_k \end{bmatrix} = \begin{bmatrix} & \widetilde{H}_k^{-1} & \\ 0 & \dots & 0 \end{bmatrix}.$$

Clearly $B = B^H$, since the product commutes and is Hermitian. From the equations

$$\begin{cases} B \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ x_1 & \dots & x_k \end{bmatrix} = \begin{bmatrix} & \widetilde{H}_m^{-1} & \\ 0 & \dots & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & & & \bar{x}_1 \\ & \ddots & & \vdots \\ & & 1 & \bar{x}_k \end{bmatrix} B = \begin{bmatrix} & & 0 \\ & \widetilde{H}_m^{-H} & \vdots \\ & & 0 \end{bmatrix} \end{cases} \tag{A.1}$$

we can conclude that $\widetilde{H}_k^{-1} = \widetilde{H}_k^{-H}$ and therefore $\widetilde{H} = \widetilde{H}^H$. The details for this equality follow below and conclude the proof. Let

$$B := \begin{bmatrix} b_{1,1} & \bar{b}_{2,1} & \dots & \bar{b}_{k,1} & \bar{b}_{k+1,1} \\ b_{2,1} & b_{2,2} & \dots & \bar{b}_{k,2} & \bar{b}_{k+1,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{k,1} & b_{k,2} & \dots & b_{k,k} & \bar{b}_{k+1,k} \\ b_{k+1,1} & b_{k+1,2} & \dots & b_{k+1,k} & b_{k+1,k+1} \end{bmatrix}$$

and write out (A.1)

$$\left\{ \begin{bmatrix} b_{1,1}+x_1\bar{b}_{k+1,1} & \bar{b}_{2,1}+x_2\bar{b}_{k+1,1} & \dots & \bar{b}_{k,1}+x_k\bar{b}_{k+1,1} \\ b_{2,1}+x_1\bar{b}_{k+1,2} & b_{2,2}+x_2\bar{b}_{k+1,2} & \dots & \bar{b}_{k,2}+x_k\bar{b}_{k+1,2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k,1}+x_1\bar{b}_{k+1,k} & b_{k,2}+x_2\bar{b}_{k+1,k} & \dots & b_{k,k}+x_k\bar{b}_{k+1,k} \\ b_{k+1,1}+x_1b_{k+1,k+1} & b_{k+1,2}+x_2\bar{b}_{k+1,k+1} & \dots & b_{k+1,k}+x_kb_{k+1,k+1} \end{bmatrix} \right.$$
$$\left. \begin{bmatrix} b_{1,1}+\bar{x}_1b_{k+1,1} & \dots & \bar{b}_{k,1}+\bar{x}_1b_{k+1,k} & \bar{b}_{k+1,1}+\bar{x}_1b_{k+1,k+1} \\ b_{2,1}+\bar{x}_2b_{k+1,1} & \dots & \bar{b}_{k,2}+\bar{x}_2b_{k+1,k} & \bar{b}_{k,2}+\bar{x}_2b_{k+1,k+1} \\ \vdots & \ddots & \vdots & \vdots \\ b_{k,1}+\bar{x}_kb_{k+1,1} & \dots & b_{k,k}+\bar{x}_kb_{k+1,k} & \bar{b}_{k+1,k}+\bar{x}_kb_{k+1,k+1} \end{bmatrix} \right\}$$
$$= \left\{ \begin{bmatrix} & \widetilde{H}_m^{-1} & \\ 0 & \dots & 0 \end{bmatrix} \right. \quad .$$
$$\left. \begin{bmatrix} & 0 \\ \widetilde{H}_m^{-H} & \vdots \\ & 0 \end{bmatrix} \right\}$$

Note that $b_{k+1,i} = -x_i b_{k+1,k+1}$ and $\bar{b}_{k+1,i} = -\bar{x}_i b_{k+1,k+1}$, which we substitute in the second equation

$$\widetilde{H}_m^{-H} = \begin{bmatrix} b_{1,1}+\bar{x}_1b_{k+1,1} & \bar{b}_{2,1}+\bar{x}_2b_{k+1,2} & \dots & \bar{b}_{k,1}+\bar{x}_1b_{k+1,k} \\ b_{2,1}+\bar{x}_2b_{k+1,1} & b_{2,2}+\bar{x}_2b_{k+1,2} & \dots & \bar{b}_{k,2}+\bar{x}_2b_{k+1,k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k,1}+\bar{x}_kb_{k+1,1} & b_{k,2}+\bar{x}_kb_{k+1,2} & \dots & b_{k,k}+\bar{x}_kb_{k+1,k} \end{bmatrix}$$
$$= \begin{bmatrix} b_{1,1}-\bar{x}_1x_1b_{k+1,k+1} & \bar{b}_{2,1}-\bar{x}_2x_2b_{k+1,k+1} & \dots & \bar{b}_{k,1}-\bar{x}_1x_kb_{k+1,k+1} \\ b_{2,1}-\bar{x}_2x_1b_{k+1,k+1} & b_{2,2}-\bar{x}_2x_2b_{k+1,k+1} & \dots & \bar{b}_{k,2}-\bar{x}_2x_kb_{k+1,k+1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k,1}-\bar{x}_kx_1b_{k+1,k+1} & b_{k,2}-\bar{x}_kx_2b_{k+1,k+1} & \dots & b_{k,k}-\bar{x}_kx_kb_{k+1,k+1} \end{bmatrix}$$
$$= \begin{bmatrix} b_{1,1}+x_1\bar{b}_{k+1,1} & \bar{b}_{2,1}+x_2\bar{b}_{k+1,2} & \dots & \bar{b}_{k,1}+x_k\bar{b}_{k+1,2} \\ b_{2,1}+x_1\bar{b}_{k+1,2} & b_{2,2}+x_2\bar{b}_{k+1,2} & \dots & \bar{b}_{k,2}+x_k\bar{b}_{k+1,2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k,1}+x_1\bar{b}_{k+1,k} & b_{k,2}+x_2\bar{b}_{k+1,k} & \dots & b_{k,k}+x_k\bar{b}_{k+1,k} \end{bmatrix} = \widetilde{H}_m^{-1}.$$

□

## A.2 Bilinear form for Krylov subspaces

In literature on biorthogonal rational functions, instead of generating Krylov subspaces with $A$, $A^H$ and using the Euclidean inner product $\langle x, y\rangle_E = y^H x$, sometimes $A$, $A^\top$ and $\langle x, y\rangle_\top = y^\top x$ is used. Note that $\langle ., .\rangle_\top$ is a Hermitian bilinear form.
We provide two main results, which are sufficient to prove all other statements in Chapter 3 and Chapter 4 for this alternative choice.

### Biorthonormal bases for polynomial Krylov subspaces

The polynomial Krylov subspaces are, for $A \in \mathbb{C}^{m\times m}$ and $v, w \in \mathbb{C}^m$,

$$\mathcal{K}_k(A, v) \text{ and } \mathcal{K}_k(A^\top, w).$$

Biorthonormal nested bases $V_k, W_k \in \mathbb{C}^{m\times k}$ for these spaces satisfy

$$\operatorname{span}\{V_i\} = \operatorname{span}\{v_0, \ldots, v_{i-1}\} = \mathcal{K}_i(A, v)$$

$$\operatorname{span}\{W_i\} = \operatorname{span}\{w_0, \ldots, w_{i-1}\} = \mathcal{K}_i(A^\top, w), \text{ for } i = 1, \ldots, k,$$

and $W_k^\top V_k = I$. These biorthonormal vectors satisfy a three term recurrence relation, this result is stated in Lemma A.1. This lemma is a variant on Lemma 3.8.

**Lemma A.1** (Biorthogonal Krylov bases recurrence relations - for $\langle ., .\rangle_\top$)**.** *Let $A \in \mathbb{C}^{m\times m}$, $v, w \in \mathbb{C}^m$, with $\langle v, w\rangle_\top \neq 0$. Consider the Krylov subspaces $\mathcal{K}_k(A, v), \mathcal{K}_k(A^\top, w)$, with $k < \min\{g_v, g_w\}$. Then biorthonormal nested bases $V_k, W_k \in \mathbb{C}^{m\times k}$ for these subspaces satisfy the recurrence relations*

$$AV_k = V_k T_k + t_{k+1,k} v_k e_k^\top$$

$$A^H W_k = W_k T_k^\top + t_{k,k+1} w_k e_k^\top,$$

*where $T_k \in \mathbb{C}^{k\times k}$ is a tridiagonal matrix.*

*Proof.* Since the biorthonormality of the bases $V_k, W_k$ with respect to the inner product $\langle ., .\rangle_\top$ implies $W_k^\top V_k = I$ and $v_k \perp \operatorname{span}\{W_k\}$, $w_k \perp \operatorname{span}\{V_k\}$, we have

$$W_k^\top A V_k = W_k^\top V_k T_k = T_k,$$

$$V_k^\top A^H W_k = V_k^\top W_k T_k^\top = T_k^\top.$$

The columns of $V_k$ form a nested basis for $\mathcal{K}_k(A, v)$, so there exists a nonsingular upper triangular matrix $R_k$ such that $V_k = B_k^V R_k$, where $B_k^V$ is the Krylov basis. By

substituting this into the Krylov recurrence relation, one easily finds that $T_k$ must be a upper Hessenberg matrix. A similar argument is valid for $W_k$ and thus $T_k^\top$ is upper Hessenberg and therefore $T_k$ is lower Hessenberg. Only a tridiagonal matrix is simultaneously upper-and lower Hessenberg. Hence, $T_k$ has tridiagonal structure. □

If $A = A^\top$, that is, $A$ is complex symmetric, and $v = w$ then a basis $V_k$ orthogonal with respect to $\langle ., . \rangle_\top$ is obtained, $V_k^\top V_k = I$ and $T_k \top = T_k$ [71].

### Biorthonormal bases for rational Krylov subspaces

Consider rational Krylov subspaces

$$\mathcal{K}_k(A, v; \Xi) \text{ and } \mathcal{K}_k(A^\top, w; \Theta).$$

The starting vectors must satisfy $\langle v, w \rangle_\top \neq 0$. Then biorthonormal nested bases $V_k, W_k \in \mathbb{C}^{m \times k}$ for these spaces exists, i.e., for $i = 1, 2, \ldots, k$

$$\operatorname{span}\{V_i\} = \operatorname{span}\{v_0, v_1, \ldots, v_{i-1}\} = \mathcal{K}_i(A, v; \Xi),$$

$$\operatorname{span}\{W_i\} = \operatorname{span}\{w_0, w_1, \ldots, w_{i-1}\} = \mathcal{K}_i(A^\top, w; \Theta),$$

$$W_k^\top V_k = I$$

The pair of recurrence relations generating these bases is of the form

$$AV_{k+1}\underline{K}_k = V_{k+1}\underline{H}_k,$$

$$A^\top W_{k+1}\underline{\widetilde{K}}_k = W_{k+1}\underline{\widetilde{H}}_k.$$

We will show that $\underline{H}_k, \underline{K}_k, \underline{\widetilde{H}}_k, \underline{\widetilde{K}}_k$ can be chosen to be tridiagonal matrices. The following lemma, Lemma A.2, is a variant of Lemma 4.4.

**Lemma A.2** (Recurrence pencil structure for biorthogonal bases of rational Krylov subspaces- for $\langle ., . \rangle_\top$)**.** *Consider $A \in \mathbb{C}^{m \times m}$, $v, w \in \mathbb{C}^m$, with $\langle v, w \rangle_\top \neq 0$ and $\Xi, \Theta$, with $\xi_i, \theta_i \in \overline{\mathbb{C}} \backslash \sigma(A)$. Let $Q^V, Q^W \in \mathbb{C}^{m \times m}$ be orthogonal nested bases for $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^\top, w; \Theta)$, respectively, and corresponding pencils $(H^V, K^V)$, $(H^W, K^W)$ satisfying*

$$AQ^V H^V = Q^V K^V, \qquad A^H Q^W H^W = Q^W K^W.$$

*Then, under the assumption that no breakdown occurs, biorthonormal nested bases $V, W$ for $\mathcal{K}(A, v; \Xi)$ and $\mathcal{K}(A^\top, w; \Theta)$, respectively, satisfy*

$$AVK = VH,$$

$$A^\top W\tilde{K} = W\tilde{H},$$

*where the recurrence pencils can be chosen such that*

- *$H$ has the lower triangular structure of $H^V$ and the upper triangular structure equals the inverted upper triangular structure of $K^W$,*
- *$K$ has the same lower triangular structure as $K^V$ and the upper triangular structure equals the inverted lower triangular structure of $H^W$,*
- *$\tilde{H}$ has the lower triangular structure of $H^W$ and the upper triangular structure equals the inverted upper triangular structure of $K^V$,*
- *$\tilde{K}$ has the same lower triangular structure as $K^W$ and the upper triangular structure equals the inverted lower triangular structure of $H^V$.*

*Proof.* From the orthogonal bases $Q^V$ and $Q^W$, the biorthonormal bases $V$ and $W$ can be constructed via Lemma 2.2, i.e., $V := Q^V R^{-1}$ and $W := Q^W L^{-H}$, with $M = (Q^W)^H Q^V = LR$. Substituting the expressions for $V, W$ in the recurrence relations for $Q^V, Q^W$ provides

$$\begin{cases} AQ^V K^V = Q^V H^V \\ A^\top Q^W K^W = Q^W H^W \end{cases} \Leftrightarrow \begin{cases} AVRK^V = VRH^V \\ A^\top W L^H K^W = W L^\top H^W \end{cases}$$

$$\Leftrightarrow \begin{cases} W^\top AVRK^V = RH^V \\ V^\top A^\top W L^\top K^W = L^\top H^W \end{cases}.$$

Taking the transpose of the second equation and rewrite it:

$$\begin{cases} W^\top AVRK^V = RH^V \\ W^\top AVL^{-1}(H^W)^{-\top} = L^{-1}(K^W)^{-\top} \end{cases}.$$

Since these expressions are only unique up to right multiplication with a nonsingular matrix $B$, we get

$$RK^V B = L^{-1}(H^W)^{-\top}, \quad RH^V B = L^{-1}(K^W)^{-\top}.$$

To obtain a particular choice for the structure of $H$ and $K$ it suffices to represent $B$ in its $RL$-decomposition $B = R_B L_B$ (assuming it exists), where $R_B$ is an upper triangular matrix and $L_B$ a lower triangular matrix,

$$\begin{cases} RK^V B = L^{-1}(H^W)^{-\top} \\ RH^V B = L^{-1}(K^W)^{-\top} \end{cases} \Leftrightarrow \begin{cases} RK^V R_B L_B = L^{-1}(H^W)^{-\top} \\ RH^V R_B L_B = L^{-1}(K^W)^{-\top} \end{cases}$$

$$\Leftrightarrow \begin{cases} RK^V R_B = L^{-1}(H^W)^{-\top} L_B^{-1} =: K \\ RH^V R_B = L^{-1}(K^W)^{-\top} L_B^{-1} =: H \end{cases}.$$

And the same reasoning as in Lemma 4.4 finishes the proof. □

# Appendix B

# Biorthogonal procedures

Section B.1 provides the derivation of the rational Lanczos iteration. That is, expressions for the entries of the tridiagonal recurrence pencil are derived. In Section B.2 a remark on the scaling of introduced weight vectors for inverse eigenvalue problems is given.

## B.1 Rational Lanczos iteration

The rational Lanczos iteration proposed in Section 8.1 is derived here. Consider two rational Krylov subspaces, generated by $A \in \mathbb{C}^{m\times m}$ and its complex conjugate $A^H$,

$$\mathcal{K}(A, v; \Xi) \text{ and } \mathcal{K}(A^H, w; \Theta),$$

with starting vectors $v, w \in \mathbb{C}^m$ satisfying $\langle v, w\rangle_E \neq 0$ and two sets of poles $\Xi = \{\xi_i\}_i$, with $\xi \in \overline{\mathbb{C}}$ and $\Theta = \{\theta_i\}_i$, with $\theta_i \in \overline{\mathbb{C}}$.
Let $\hat{g}$ denote the index of a breakdown, this can be a lucky breakdown or a serious breakdown. The goal is to construct bases $V_k, W_k \in \mathbb{C}^{m\times k}$ such that, for any $k \leq \hat{g}$, their columns span the rational Krylov subspaces

$$\text{span}\{V_k\} = \text{span}\{v_0, v_1 \ldots, v_{k-1}\} = \mathcal{K}_k(A, v; \Xi)$$

$$\text{span}\{W_k\} = \text{span}\{w_0, w_1 \ldots, w_{k-1}\} = \mathcal{K}_k(A^H, w; \Theta)$$

and are biorthogonal

$$W_k^H V_k = D,$$

where $D \in \mathbb{C}^{n\times n}$ is a diagonal matrix whose elements depend on the chosen normalization.

By Theorem 4.6 (and also Lemma 5.6), we know that two tridiagonal pencils $(\underline{T}_k, \underline{S}_k), (\widetilde{\underline{T}}_k, \widetilde{\underline{S}}_k) \in \mathbb{C}^{(k+1)\times k}$ exist such that

$$AV_{k+1}\underline{S}_k = V_{k+1}\underline{T}_k \text{ and } A^H W_{n+1}\widetilde{\underline{S}}_n = W_{n+1}\widetilde{\underline{T}}_n.$$

We assign variables to the elements of the recurrence pencils:

$$\underline{T}_k =: \begin{bmatrix} d_1 & a_2 & & \\ b_2 & d_2 & \ddots & \\ & \ddots & \ddots & a_k \\ & & b_k & d_k \\ & & & b_{k+1} \end{bmatrix} \qquad \underline{S}_k =: \begin{bmatrix} c_1 & u_2 & & \\ l_2 & c_2 & \ddots & \\ & \ddots & \ddots & u_k \\ & & l_k & c_k \\ & & & l_{k+1} \end{bmatrix}$$

$$\widetilde{\underline{T}}_k =: \begin{bmatrix} \delta_1 & \alpha_2 & & \\ \beta_2 & \delta_2 & \ddots & \\ & \ddots & \ddots & \alpha_k \\ & & \beta_k & \delta_k \\ & & & \beta_{k+1} \end{bmatrix} \qquad \widetilde{\underline{S}}_k =: \begin{bmatrix} \gamma_1 & \mu_2 & & \\ \lambda_2 & \gamma_2 & \ddots & \\ & \ddots & \ddots & \mu_k \\ & & \lambda_k & \gamma_k \\ & & & \lambda_{k+1} \end{bmatrix}$$

Additional structural properties are provided by Lemma 4.5 and Lemma 4.6, namely the ratios of the sub-and superdiagonal elements are determined by the given poles

$$\frac{b_{i+1}}{l_{i+1}} = \xi_i, \qquad \frac{\beta_{i+1}}{\lambda_{i+1}} = \theta_i, \qquad i = 1, 2, \ldots, k,$$

$$\frac{a_{i+1}}{u_{i+1}} = \bar{\theta}_{i-1}, \qquad \frac{\alpha_{i+1}}{\mu_{i+1}} = \bar{\xi}_{i-1}, \qquad i = 1, 2, \ldots, k-1.$$

### B.1.1 Initialize

For $i = 1$, the equations are different than the others, therefore these will be discussed first. The recurrence relations follow from the first columns of the recurrence pencil

$$\begin{cases} c_1 A v_1 + l_2 A v_2 & = d_1 v_1 + b_2 v_2 \\ \delta_1 A^H w_1 + \beta_2 A^H w_2 & = \gamma_1 w_1 + \lambda_2 w_2 \end{cases}$$

$$\Leftrightarrow \begin{cases} v_2 & = -\underbrace{(l_2 A - b_2 I)^{-1}}_{=:B}(c_1 A - d_1 I)v_1 \\ w_2 & = -\underbrace{(\lambda_2 A - \beta_2 I)^{-1}}_{=:F}(\gamma_1 A - \delta_1 I)w_1 \end{cases}.$$

The matrices $B$ and $F$ can be computed, since $A$ and the poles are given. Now, let us obtain expressions for the coefficients $c_1, d_1, \gamma_1, \delta_1$.
Consider the biorthogonality conditions

$$\langle v_2, w_1\rangle_E = 0, \quad \langle w_2, v_1\rangle_E = 0.$$

Using these conditions, expressions for the ratios $\frac{d_1}{c_1}$ and $\frac{\delta_1}{\gamma_1}$ are obtained:

$$\begin{cases} \langle v_2, w_1\rangle_E & = \langle -B(c_1A - d_1I)v_2, w_1\rangle_E = 0 \\ \langle w_2, v_1\rangle_E & = \langle -F(\gamma_1 A^H - \delta_1 I)w_1, v_1\rangle_E = 0 \end{cases} \Leftrightarrow \begin{cases} \frac{d_1}{c_1} & = \frac{\langle BAv_1, w_1\rangle_E}{\langle Bv_1, w_1\rangle_E} \\ \frac{\delta_1}{\gamma_1} & = \frac{\langle FA^H w_1, v_1\rangle_E}{\langle Fw_1, v_1\rangle_E} \end{cases}$$

Only the ratios are specified, to make the coefficients unique, a normalization must be chosen.

**Example B.1.** *Biorthonormal bases, i.e., normalization such that* $\langle v_2, w_2\rangle_E = 1$. *Let*

$$\hat{v}_2 := v_2/c_1 = -B(A - \frac{d_1}{c_1}I)v_1, \quad \hat{w}_2 := w_2/\delta_1 = -F(A^H - \frac{\delta_1}{\gamma_1}I)w_1.$$

*Then the normalization leads to*

$$\langle v_2, w_2\rangle_E = 1 \Leftrightarrow \langle \hat{v}_2, \hat{w}_2\rangle_E = \frac{1}{c_1\bar{\gamma}_1} \Leftrightarrow c_1\bar{\gamma}_1 = \frac{1}{\langle \hat{v}_2, \hat{w}_2\rangle_E}.$$

*Still, there is freedom in the choice of* $c_1$ *and* $\gamma_1$.

### B.1.2 Iterate

Now the coefficients for $1 < i \leq k < \hat{g}$ are derived. From the $i$th column of the recurrence pencils we get the pair of equations

$$\begin{cases} u_iAv_{i-1} + c_iAv_i + l_{i+1}Av_{i+1} & = a_iv_{i-1} + d_iv_i + b_{i+1}v_{i+1} \\ \mu_iA^Hw_{i-1} + \gamma_iA^Hw_i + \lambda_{i+1}A^Hw_{i+1} & = \alpha_iw_{i-1} + \delta_iw_i + \beta_{i+1}w_{i+1} \end{cases}$$

$$\Leftrightarrow \begin{cases} v_{i+1} & = -B(u_iA - a_iI)v_{i-1} - B(c_iA - d_iI)v_i \\ w_{i+1} & = -F(\mu_iA^H - \alpha_iI)w_{i-1} - F(\gamma_iA^H - \delta_iI)w_i \end{cases},$$

with $B = (l_{i+1}A - b_{i+1}I)^{-1}$, $F = (\lambda_{i+1}A^H - \beta_{i+1}I)^{-1}$. Expressions for the coefficients $u_i, a_i, c_i, d_i, \alpha_i, \mu_i, \delta_i, \gamma_i$ can be obtained by taking the structure, orthogonality conditions and normalization into account.

### Poles

First, the poles must appear in the recurrence pencil

$$\frac{a_i}{u_i} = \bar{\theta}_{i-2}, \quad \frac{\alpha_i}{\mu_i} = \bar{\xi}_{i-2}.$$

Note the special cases

- If $\xi_{i-2} = \infty$, then $\mu_i = 0$.
- If $\xi_{i-2} = 0$, then $\alpha_i = 0$.
- If $\theta_{i-2} = \infty$, then $u_i = 0$.
- If $\theta_{i-2} = 0$, then $a_i = 0$.

These will be important for normalization, the variable used for normalization depends on the pole chosen. A distinction between finite and infinite poles must be made.

### Orthogonality

Second, the orthogonality conditions dictate

$$\langle v_{i+1}, w_{i-1}\rangle_E = 0, \quad \langle v_{i+1}, w_i\rangle_E = 0, \quad \langle w_{i+1}, v_{i-1}\rangle_E = 0 \text{ and } \langle w_{i+1}, v_i\rangle_E = 0.$$

The first two lead to

$$\begin{aligned}
&\langle v_{i+1}, w_i\rangle \\
&= \langle -B(u_iA - a_iI)v_{i-1} - B(c_iA - d_iI)v_i, w_i\rangle_E \\
&= -u_i \underbrace{\langle BAv_{i-1}, w_i\rangle_E}_{x} + a_i \underbrace{\langle Bv_{i-1}, w_i\rangle_E}_{y} - c_i \underbrace{\langle BAv_i, w_i\rangle_E}_{z} + d_i \underbrace{\langle Bv_i, w_i\rangle_E}_{q} \\
&= u_ix + a_iy - c_iz + d_iq = 0
\end{aligned}$$

and

$$\begin{aligned}
&\langle v_{i+1}, w_{i-1}\rangle_E \\
&= \langle -B(u_iA - a_iI)v_{i-1} - B(c_iA - d_iI)v_i, w_{i-1}\rangle_E \\
&= -u_i \underbrace{\langle BAv_{i-1}, w_{i-1}\rangle_E}_{\tilde{x}} + a_i \underbrace{\langle Bv_{i-1}, w_{i-1}\rangle_E}_{\tilde{y}} - c_i \underbrace{\langle BAv_i, w_{i-1}\rangle_E}_{\tilde{z}} + d_i \underbrace{\langle Bv_i, w_{i-1}\rangle_E}_{\tilde{q}} \\
&= u_i\tilde{x} + a_i\tilde{y} - c_i\tilde{z} + d_i\tilde{q} = 0
\end{aligned}$$

To summarize, for $v_{i+1}$ we have three equations for four unknowns

$$\begin{cases} \frac{a_i}{u_i} = \bar{\theta}_{i-2} \\ -u_i x + a_i y - c_i z + d_i q = 0 \\ -u_i \tilde{x} + a_i \tilde{y} - c_i \tilde{z} + d_i \tilde{q} = 0 \end{cases} . \tag{B.1}$$

Some choices of poles lead to a reduction in parameters:

- If $\xi_i = \infty$, then $B = I$ and $y = \tilde{q} = 0$.
- If $\xi_i = 0$, then $B = A^{-1}$ and $x = \tilde{z} = 0$.

The fourth equation is obtained by normalization. In order to normalize, the other sequence is required.
Using the latter two orthogonality conditions, a similar derivation is performed for $w_{i+1}$:

$$\begin{aligned} &\langle w_{i+1}, v_i \rangle_E \\ &= -\mu_i \underbrace{\langle FA^H w_{i-1}, v_i \rangle_E}_{\chi} + \alpha_i \underbrace{\langle F w_{i-1}, v_i \rangle_E}_{\tau} - \gamma_i \underbrace{\langle FA^H w_i, v_i \rangle_E}_{\eta} + \delta_i \underbrace{\langle F w_i, v_i \rangle}_{\rho} \\ &= -\mu_i \chi + \alpha_i \tau - \gamma_i \eta + \delta_i \rho = 0 \end{aligned}$$

$$\begin{aligned} &\langle w_{i+1}, v_{i-1} \rangle_E \\ &= -\mu_i \underbrace{\langle FA^H w_{i-1}, v_{i-1} \rangle_E}_{\tilde{\chi}} + \alpha_i \underbrace{\langle F w_{i-1}, v_{i-1} \rangle_E}_{\tilde{\tau}} - \gamma_i \underbrace{\langle FA^H w_i, v_{i-1} \rangle_E}_{\tilde{\eta}} + \delta_i \underbrace{\langle F w_i, v_{i-1} \rangle_E}_{\tilde{\rho}} \\ &= -\mu_i \tilde{\chi} + \alpha_i \tilde{\tau} - \gamma_i \tilde{\eta} + \delta_i \tilde{\rho} = 0. \end{aligned}$$

To summarize, for $w_{i+1}$ we have

$$\begin{cases} \frac{\alpha_i}{\mu_i} = \bar{\xi}_{i-2} \\ -\mu_i \chi + \alpha_i \tau - \gamma_i \eta + \delta_i \rho = 0 \\ -\mu_i \tilde{\chi} + \alpha_i \tilde{\tau} - \gamma_i \tilde{\eta} + \delta_i \tilde{\rho} = 0 \end{cases} \tag{B.2}$$

And for poles at zero or infinity:

- If $\theta_i = \infty$, then $F = I$ and $\tau = \tilde{\rho} = 0$.
- If $\theta_i = 0$, then $F = A^{-H}$ and $\chi = \tilde{\eta} = 0$.

### Normalization

Third, normalization will provide the fourth equation for (B.1) and (B.2), such that the four unknowns are determined uniquely. We will define the nonnormalized vectors $\hat{v}_{i+1}$ and $\hat{w}_{i+1}$, which can be computed at this point. We must make distinction between $\theta_{i-2} \neq \infty$ and $\theta_{i-2} = \infty$.

- If $\theta_{i-2} = \infty$, then $\beta_{i-2} = 0$ and the expressions become

$$\begin{cases} u_i & = 0 \\ \frac{d_i}{a_i} & = \frac{y\tilde{z}-\tilde{y}z}{\tilde{q}z-q\tilde{z}} \\ \frac{c_i}{a_i} & = \frac{y+\frac{d_i}{a_i}q}{z} \end{cases}.$$

  The nonnormalized vector is

$$\hat{v}_{i+1} := \frac{1}{a_i} v_{i+1} = Bv_{i-1} - B(\frac{c_i}{a_i}A - \frac{d_i}{a_i}I)v_i$$

  and the normalizing variable is $a_i$.

- If $\theta_{i-2} \neq \infty$, then the expressions become

$$\begin{cases} \frac{a_i}{u_i} & = \bar{\theta}_{i-2} \\ \frac{d_i}{u_i} & = \frac{\bar{\theta}_{i-2}(y\tilde{z}-\tilde{y}z)-(\tilde{x}z-x\tilde{z})}{\tilde{q}z-q\tilde{z}} \\ \frac{c_i}{u_i} & = \frac{\bar{\theta}_{i-2}y-x+\frac{d_i}{u_i}q}{z} \end{cases}.$$

  The nonnormalized vector is

$$\hat{v}_{i+1} := \frac{1}{u_i} v_{i+1} = -B(A - \bar{\theta}_{i-2}I)v_{i-1} - B(\frac{c_i}{u_i}A - \frac{d_i}{u_i}I)v_i.$$

  and the normalizing variable is $u_i$.

Similarly for $\hat{w}_{i+1}$, a distinction is made.

- If $\xi_{i-2} = \infty$, then $l_{i-2} = 0$ and the expressions become

$$\begin{cases} \mu_i & = 0 \\ \frac{\delta_i}{\alpha_i} & = \frac{\tau\tilde{\eta}-\tilde{\tau}\eta}{\tilde{\rho}\eta-\rho\tilde{\eta}} \\ \frac{\gamma_i}{\alpha_i} & = \frac{\tau+\frac{\delta_i}{\alpha_i}\rho}{\eta} \end{cases}.$$

The nonnormalized vector is

$$\hat{w}_{i+1} := \frac{1}{\alpha_i} w_{i+1} = F w_{i-1} - F(\frac{\gamma_i}{\alpha_i} A^H - \frac{\delta_i}{\alpha_i} I) w_i$$

and the normalizing variable is $\alpha_i$.

- If $\xi_{i-2} \neq \infty$, then the expressions become

$$\begin{cases} \frac{\alpha_i}{\mu_i} & = \bar{\xi}_{i-2} \\ \frac{\delta_i}{\mu_i} & = \frac{\bar{\xi}_{i-2}(\tau\tilde{\eta} - \tilde{\tau}\eta) - (\bar{\chi}\eta - \chi\tilde{\eta})}{\tilde{\rho}\eta - \rho\tilde{\eta}} \\ \frac{\gamma_i}{\mu_i} & = \frac{\bar{\xi}_{i-2}\tau - \chi + \frac{\delta_i}{\alpha_i}\rho}{\eta} \end{cases}.$$

The nonnormalized vector is

$$\hat{w}_{i+1} := \frac{1}{\mu_i} w_{i+1} = -F(A^H - \bar{\xi}_{i-2} I) w_{i-1} - F(\frac{\gamma_i}{\mu_i} A^H - \frac{\delta_i}{\mu_i} I) w_i$$

and normalization is performed by $\mu_i$.

The normalization constants can be determined by using the nonnormalized vectors $\hat{v}_{i+1}$ and $\hat{w}_{i+1}$.

### B.1.3 Termination

If a lucky breakdown occurs, i.e., $k = \min\{g_v, g_w\}$, then an invariant subspace is found. Assume without loss of generality that $k = g_v < g_w$. The restriction of $A$ to this subspace is then given by the pencil $(T_k, S_k)$. The recurrence relation becomes

$$AV_k S_k = V_k T_k$$

$$AV_k \begin{bmatrix} c_1 & u_2 & & \\ l_2 & c_2 & \ddots & \\ & \ddots & \ddots & u_k \\ & & l_k & c_k \end{bmatrix} = V_k \begin{bmatrix} d_1 & a_2 & & \\ b_2 & d_2 & \ddots & \\ & \ddots & \ddots & a_k \\ & & b_k & d_k \end{bmatrix},$$

i.e., there is no new basis vector introduced, since it is linearly dependent on the others $v_i, i = 1, 2, \ldots, k$. The columns of the basis $V_k$ span the invariant subspace.
We will derive expressions for $u_k, c_k, a_k, d_k$. Consider the equation

$$v_{k+1} = -B(u_k A - a_k I) v_{k-1} - B(c_k A - d_k I) v_k,$$

which must be equal to zero.
The orthogonality conditions are

$$\langle v_{k+1}, w_k\rangle_E = 0, \quad \langle v_{k+1}, w_{k-1}\rangle_E = 0.$$

These lead to the expressions:

$$\langle v_{k+1}, w_k\rangle_E = c_k\langle Av_k, w_k\rangle_E + u_k\langle Av_{k-1}, w_k\rangle_E - d_k\langle v_k, w_k\rangle_E = 0,$$

$$\langle v_{k+1}, w_{k-1}\rangle_E = c_k\langle Av_k, w_{k-1}\rangle_E + u_k\langle Av_{k-1}, w_{k-1}\rangle_E - a_k\langle v_{k-1}, w_{k-1}\rangle_E = 0.$$

And, distinguishing finite and infinite poles:

- If $\theta_{m-2} = \infty$, then
$$\begin{cases} \frac{d_k}{a_k} &= \frac{c_k}{a_k}\frac{\langle Av_k, w_k\rangle_E}{\langle v_k, w_k\rangle_E} \\ \frac{c_k}{a_k} &= \frac{\langle v_{m-1}, w_{m-1}\rangle_E}{\langle Av_k, w_{m-1}\rangle_E} \end{cases}$$
- If $\theta_{k-2} \neq \infty$, then
$$\begin{cases} \frac{d_k}{u_k} &= \frac{c_k}{u_k}\frac{\langle Av_k, w_k\rangle_E}{\langle v_k, w_k\rangle_E} + \frac{\bar{\theta}_{m-2}\langle Av_{m-1}, w_k\rangle_E}{\langle v_k, w_k\rangle_E} \\ \frac{c_k}{u_k} &= \frac{\bar{\theta}_{m-2}\langle v_{k-1}, w_{k-1}\rangle_E - \langle A_{k-1}, w_{k-1}\rangle_E}{\langle Av_k, w_{k-1}\rangle_E} \end{cases}.$$

The parameters $a_k$ and $u_k$ can still be chosen, we set $u_k = 1$ if $\theta_{m-2} = \infty$ and $a_k = 1$ if $\theta_{m-2} \neq \infty$.

## B.2 Scaling in Biorthogonal IEP

In Section 9.4.1 it was mentioned that in numerical computation with biorthonormal vectors the mutual scaling is important. It might be interesting to keep the difference in their order of magnitude as small as possible. Using the notation of Section 9.4.1, eliminators $L_1 \in \mathfrak{L}_1$ and $R_1 \in \mathfrak{R}_1$ are constructed such that they introduce the new weights in the first columns of $\hat{V}, \hat{W} \in \mathbb{C}(m+1) \times (m+1)$. For scaling purposes, a diagonal matrix $D_1 \in \mathbb{C}(m+1) \times (m+1)$ is introduced,

$$D_1 = \begin{bmatrix} d_1 & & \\ & I_{m-1} & \\ & & d_{m+1} \end{bmatrix}.$$

This freedom originates from the fact that in normalization, the value of $\langle \tilde{v}, \tilde{w}\rangle_E$ can be distributed over $\tilde{\eta}$ and $\tilde{\nu}$ in any way as long as $\langle \frac{\tilde{v}}{\tilde{\eta}}, \frac{\tilde{w}}{\tilde{\nu}u}\rangle_E = 1$. The equations are now

$$\hat{V}L_1D_1R_1e_1 = \left[\tilde{v}/\tilde{\eta}\right] \quad \text{and} \quad \hat{W}(L_1)^{-H}D_1^{-H}(R_1)^{-H}e_1 = \left[\tilde{w}/\tilde{\nu}\right].$$

The parameters $l_1, r_1$ in $L_1, R_1$ are

$$\begin{cases} l_1 = \frac{v_{m+1}}{\eta} \\ r_1 = -\frac{d_2}{d_1} \frac{\bar{w}_{m+1}}{\bar{\nu} + l_1 \bar{w}_{m+1}} \end{cases} . \tag{B.3}$$

The resulting first columns are now scaled

$$\hat{V} L_1 D_1 R_1 e_1 = \begin{bmatrix} | \\ d_1 \frac{v}{\eta} \\ | \\ d_1 l_1 \end{bmatrix} = \begin{bmatrix} | \\ d_1 \frac{v}{\eta} \\ | \\ d_1 \frac{v_{m+1}}{\eta} \end{bmatrix}$$

$$\hat{W} (L_1)^{-H} (D_1)^{-H} (R_1)^{-H} e_1 = \begin{bmatrix} | \\ \left( \frac{1}{\bar{d}_1} + \frac{\bar{r}_1 \bar{l}_1}{\bar{d}_2} \right) \frac{w}{\nu} \\ | \\ -\frac{\bar{r}_1}{\bar{d}_2} \end{bmatrix} = \begin{bmatrix} | \\ \frac{\bar{\eta}\nu}{\bar{d}_1 (\bar{\eta}\nu + \bar{v}_{m+1} w_{m+1})} \frac{w}{\nu} \\ | \\ \frac{\bar{\eta}\nu}{\bar{d}_1 (\bar{\eta}\nu + \bar{v}_{m+1} w_{m+1})} \frac{w_{m+1}}{\nu} \end{bmatrix} .$$

# Bibliography


[1] Ammar, G., Gragg, W. B., and Reichel, L. Downdating of Szegő polynomials and data-fitting applications. *Linear Algebra and its Applications 172* (1992), 315–336.

[2] Antoulas, A. C. *Approximation of large-scale dynamical systems.* Advances in design and control. Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[3] Arnoldi, W. E. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of Applied Mathematics 9*, 1 (1951), 17–29.

[4] Askey, R., and Ismail, M. *Recurrence relations, continued fractions and orthogonal polynomials*, vol. 300. American Mathematical Soc., 1984.

[5] Aurentz, J. L., Mach, T., Vandebril, R., and Watkins, D. S. Fast and backward stable computation of roots of polynomials. *SIAM Journal on Matrix Analysis and Applications 36*, 3 (2015), 942–973.

[6] Bailey, R., and Derevyagin, M. Complex Jacobi matrices generated by Darboux transformations. *arXiv preprint arXiv:2107.09824* (2021).

[7] Barnett, S. Congenial matrices. *Linear Algebra and its Applications 41* (1981), 277–298.

[8] Baxter, G. Polynomials defined by a difference system. *Journal of Mathematical Analysis and Applications 2*, 2 (1961), 223 – 263.

[9] Bebiano, N., da Providência, J., Nishiyama, S., and da Providência, J. P. A quantum system with a non-Hermitian Hamiltonian. *Journal of Mathematical Physics 61*, 082106 (2020).

[10] Beckermann, B. The stable computation of formal orthogonal polynomials. *Numerical Algorithms 11*, 1 (1996), 1–23.

[11] Beckermann, B. The condition number of real Vandermonde, Krylov and positive definite Hankel matrices. *Numerische Mathematik 85*, 4 (2000), 553–577.

[12] Beckermann, B. Complex Jacobi matrices. *Journal of Computational and Applied Mathematics 127*, 1 (2001), 17–65. Numerical Analysis 2000. Vol. V: Quadrature and Orthogonal Polynomials.

[13] Beckermann, B., Derevyagin, M., and Zhedanov, A. The linear pencil approach to rational interpolation. *Journal of Approximation Theory 162*, 6 (2010), 1322–1346.

[14] Beckermann, B., Güttel, S., and Vandebril, R. On the convergence of rational Ritz values. *SIAM Journal on Matrix Analysis and Applications 31*, 4 (2010), 1740–1774.

[15] Bennett, J. M. Triangular factors of modified matrices. *Numerische Mathematik 7*, 3 (1965), 217–221.

[16] Berljafa, M. *Rational Krylov Decompositions: Theory and Applications.* PhD thesis, The University of Manchester, 2017.

[17] Berljafa, M., Elsworth, S., and Güttel, S. A rational Krylov toolbox for MATLAB. `http://rktoolbox.org`, 2014.

[18] Berljafa, M., and Güttel, S. Generalized rational Krylov decompositions with an application to rational approximation. *SIAM Journal on Matrix Analysis and Applications 36*, 2 (2015), 894–916.

[19] Berljafa, M., and Güttel, S. Parallelization of the rational Arnoldi algorithm. *SIAM Journal on Scientific Computing 39*, 5 (2017), S197–S221.

[20] Beutler, F. J. The operator theory of the pseudo-inverse I. Bounded operators. *Journal of Mathematical Analysis and Applications 10*, 3 (1965), 451–470.

[21] Bini, D. A., and Böttcher, A. Polynomial factorization through Toeplitz matrix computations. *Linear Algebra and its Applications 366* (2003), 25 – 37. Special issue on Structured Matrices: Analysis, Algorithms and Applications.

[22] Bini, D. A., Gemignani, L., and Meini, B. Computations with infinite Toeplitz matrices and polynomials. *Linear Algebra and its Applications 343-344* (2002), 21 – 61. Special Issue on Structured and Infinite Systems of Linear equations.

[23] Bini, D. A., Gemignani, L., and Pan, V. Y. Inverse power and Durand-Kerner iterations for univariate polynomial root-finding. *Computers & Mathematics with Applications 47*, 2 (2004), 447–459.

[24] Björck, Å. Numerics of Gram-Schmidt orthogonalization. *Linear Algebra and its Applications 197-198* (1994), 297–316.

[25] Björck, r., Park, H., and Eldén, L. Accurate downdating of least squares solutions. *SIAM Journal on Matrix Analysis and Applications 15*, 2 (1994), 549–568.

[26] Bojanczyk, A. W., Brent, R. P., Van Dooren, P., and de Hoog, F. R. A note on downdating the Cholesky factorization. *SIAM journal on scientific and statistical computing 8*, 3 (1987), 210–221.

[27] Boley, D., and Golub, G. H. A survey of matrix inverse eigenvalue problems. *Inverse Problems 3*, 4 (1987), 595–622.

[28] Brezinski, C. *Padé-type approximation and general orthogonal polynomials*. Springer, 1980.

[29] Brezinski, C. *Biorthogonality and its Applications to Numerical Analysis*. CRC Press, 1992.

[30] Brezinski, C. The methods of Vorobyev and Lanczos. *Linear Algebra and its Applications 234* (1996), 21–41.

[31] Brezinski, C., and Redivo-Zaglia, M. Breakdowns in the computation of orthogonal polynomials. In *Nonlinear Numerical Methods and Rational Approximation II* (Antwerp, Belgium, 1994), A. Cuyt, Ed., vol. 296, Springer, pp. 49–59.

[32] Bultheel, A., Cuyt, A., Van Assche, W., Van Barel, M., and Verdonk, B. Generalizations of orthogonal polynomials. *Journal of Computational and Applied Mathematics 179*, 1 (2005), 57–95. Proceedings of the Conference on Orthogonal Functions and Related Topics, Held in Honor of Olav Njåstad.

[33] Bultheel, A., González-Vera, P., Hendriksen, E., and Njåstad, O. Orthogonal rational functions and tridiagonal matrices. *Journal of Computational and Applied Mathematics 153*, 1 (2003), 89–97. Proceedings of the 6th International Symposium on Orthogonal Polynomials, Special Functions and their Applications, Rome, Italy, 18-22 June 2001.

[34] Bultheel, A., González-Vera, P., Hendriksen, E., and Njôstad, O. *Orthogonal rational functions*, vol. 5. Cambridge University Press; Cambridge, 1999.

[35] Bultheel, A., and Van Barel, M. *Linear algebra, rational approximation and orthogonal polynomials*, vol. 6. Elsevier, 1997.

[36] Bunse-Gerstner, A., and Fassbender, H. Error bounds in the isometric Arnoldi process. *Journal of Computational and Applied Mathematics 86*, 1 (1997), 53–72.

[37] Camps, D. *Pole Swapping Methods for the Eigenvalue Problem - Rational QR Algorithms*. PhD thesis, KU Leuven, 2019.

[38] Camps, D., Mach, T., Vandebril, R., and Watkins, D. S. On pole-swapping algorithms for the eigenvalue problem. *arXiv preprint arXiv:1906.08672* (2020).

[39] Camps, D., Meerbergen, K., and Vandebril, R. An implicit filter for rational Krylov using core transformations. *Linear Algebra and its Applications 561* (2019), 113 – 140.

[40] Camps, D., Meerbergen, K., and Vandebril, R. A Rational QZ method. *SIAM Journal on Matrix Analysis and Applications 40*, 3 (2019), 943–972.

[41] Cantero, M. J., Moral, L., and Velázquez, L. Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra and its Applications 362* (2003), 29 – 56.

[42] Celsus, A. F., Deaño, A., Huybrechs, D., and Iserles, A. The kissing polynomials and their Hankel determinants. *arXiv preprint arXiv:1504.07297* (2021).

[43] Chihara, T. S. *An introduction to orthogonal polynomials*. Mathematics and its Applications. Gordon and Breach, New York, NY, 1978. This book has also been published by Dover in 2011.

[44] Chu, M. T. Inverse eigenvalue problems. *SIAM Review 40*, 1 (1998), 1–39.

[45] Chu, M. T., and Golub, G. H. Structured inverse eigenvalue problems. *Acta Numerica 11* (2002), 1–71.

[46] Cullum, J., Kerner, W., and Willoughby, R. A generalized nonsymmetric Lanczos procedure. *Computer Physics Communications 53*, 1-3 (1989), 19–48.

[47] Cullum, J., and Willoughby, R. A. Computing eigenvalues of very large symmetric matrices—An implementation of a Lanczos algorithm with no reorthogonalization. *Journal of Computational Physics 44*, 2 (1981), 329–358.

[48] da Providência, J., Bebiano, N., and da Providencia, J. P. Non-hermitian Hamiltonians with real spectrum in quantum mechanics. *Brazilian Journal of Physics 41*, 1 (2011), 78.

[49] Dahlquist, G., and Björck, r. *Numerical Methods in Scientific Computing, Volume I*. Society for Industrial and Applied Mathematics, 2008.

[50] Daniel, J. W., Gragg, W. B., Kaufman, L., and Stewart, G. W. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Mathematics of Computation 30*, 136 (1976), 772–795.

[51] Davis, P. J. *Interpolation and approximation.* Dover books on advanced mathematics. Dover, New York, 1975.

[52] de Boor, C., and Golub, G. H. The numerically stable reconstruction of a Jacobi matrix from spectral data. *Linear Algebra and its Applications 21*, 3 (1978), 245–260.

[53] Deckers, K., and Bultheel, A. Orthogonal rational functions with complex poles: The Favard theorem. *Journal of Mathematical Analysis and Applications 356*, 2 (2009), 764–768.

[54] Deckers, K., and Bultheel, A. The existence and construction of rational Gauss-type quadrature rules. *Applied Mathematics and Computation 218*, 20 (2012), 10299–10320.

[55] Deckers, K., Bultheel, A., and Van Deun, J. A generalized eigenvalue problem for quasi-orthogonal rational functions. *Numerische Mathematik 117*, 3 (2011), 463–506.

[56] Draux, A. *Polynômes orthogonaux formels : applications.* Lecture notes in mathematics 974. Springer, Berlin, 1983.

[57] Draux, A. Formal orthogonal polynomials revisited. Applications. *Numerical Algorithms 11* (1996), 143–158.

[58] Druskin, V., and Knizhnerman, L. Extended Krylov subspaces: approximation of the matrix square root and related functions. *SIAM Journal on Matrix Analysis and Applications 19*, 3 (1998), 755–771.

[59] Eichinger, B., Lukić, M., and Young, G. Orthogonal rational functions with real poles, root asymptotics, and GMP matrices. *arXiv preprint arXiv:2008.11884* (2020).

[60] Elhay, S., Golub, G. H., and Kautsky, J. Updating and downdating of orthogonal polynomials with data fitting applications. *SIAM Journal on Matrix Analysis and Applications 12*, 2 (1991), 327–353.

[61] Elsner, L. A note on generalized Hessenberg matrices. *Linear Algebra and its Applications 409* (2005), 147–152. Special Issue in honor of Pauline van den Driessche.

[62] Faber, V., and Manteuffel, T. Necessary and sufficient conditions for the existence of a conjugate gradient method. *SIAM Journal on Numerical Analysis 21*, 2 (1984), 352–362.

[63] Fasino, D. Rational Krylov matrices and QR steps on Hermitian diagonal-plus-semiseparable matrices. *Numerical Linear Algebra with Applications 12*, 8 (2005), 743–754.

[64] Fasino, D., and Gemignani, L. A Lanczos-type algorithm for the QR factorization of regular Cauchy matrices. *Numerical Linear Algebra with Applications 9*, 4 (2002), 305–319.

[65] Fasino, D., and Gemignani, L. Direct and inverse eigenvalue problems for diagonal-plus-semiseparable matrices. *Numerical Algorithms 34* (2003), 313–324.

[66] Fiedler, M. Basic matrices. *Linear Algebra and its Applications 373* (2003), 143–151. Combinatorial Matrix Theory Conference (Pohang, 2002).

[67] Fiedler, M. Complementary basic matrices. *Linear Algebra and its Applications 384* (2004), 199–206.

[68] Fiedler, M., and Vavřín, Z. Generalized Hessenberg matrices. *Linear Algebra and its Applications 380* (2004), 95–105.

[69] Fischer, H.-J. On generating orthogonal polynomials for discrete measures. *Zeitschrift für Analysis und ihre Anwendungen 17*, 1 (1998), 183–205.

[70] Forsythe, G. E. Generation and use of orthogonal polynomials for data-fitting with a digital computer. *Journal of the Society for Industrial and Applied Mathematics 5*, 2 (1957), 74–88.

[71] Freund, R. W. Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices. *SIAM Journal on Scientific and Statistical Computing 13*, 1 (1992), 425–448.

[72] Freund, R. W., and Zha, H. Formally biorthogonal polynomials and a look-ahead Levinson algorithm for general Toeplitz systems. *Linear Algebra and its Applications 188-189* (1993), 255–303.

[73] Gallivan, K., Grimme, E. J., and Van Dooren, P. A rational Lanczos algorithm for model reduction. *Numerical Algorithms 12*, 1 (1996), 33–63.

[74] Gallivan, K., Grimme, Eric, J., and Van Dooren, P. Padé approximation of large-scale dynamic systems with Lanczos methods. In *Proceedings of 1994 33rd IEEE Conference on Decision and Control* (1994), vol. 1, IEEE, pp. 443–448.

[75] Gander, W. Algorithms for the QR decomposition. Tech. Rep. 02, Angewandte Mathematik, ETH, 1980.

[76] Gautschi, W. An algorithmic implementation of the generalized Christoffel theorem. In *Numerical integration* (Oberwolfach, 1982), vol. 57 of *ISNM*, Birkhäuser, Basel, pp. 89–106.

[77] Gautschi, W. On generating orthogonal polynomials. *SIAM Journal on Scientific and Statistical Computing 3*, 3 (1982), 289–317.

[78] Gautschi, W. Orthogonal polynomials—constructive theory and applications. *Journal of Computational and Applied Mathematics 12-13* (1985), 61–76.

[79] Gautschi, W. Is the recurrence relation for orthogonal polynomials always stable? *BIT Numerical Mathematics 33*, 2 (1993), 277–284.

[80] Gautschi, W. Orthogonal polynomials: applications and computation. *Acta Numerica 5* (1996), 45–119.

[81] Gautschi, W. The use of rational functions in numerical quadrature. *Journal of Computational and Applied Mathematics 133*, 1 (2001), 111–126. 5th Int. Symp. on Orthogonal Polynomials, Special Functions and their Applications.

[82] Gautschi, W. The interplay between classical analysis and (numerical) linear algebra - a tribute to Gene H. Golub. *Electron. Trans. Numer. Anal. 13* (2002), 119–147.

[83] Gautschi, W. *Orthogonal polynomials computation and approximation.* Numerical mathematics and scientific computation. Oxford university press, 2004.

[84] Gautschi, W. Orthogonal polynomials (in Matlab). *Journal of Computational and Applied Mathematics 178*, 1 (2005), 215–234. Proceedings of the Seventh International Symposium on Orthogonal Polynomials,Special Functions and Applications.

[85] George, A., and Liu, J. W. Householder reflections versus Givens rotations in sparse orthogonal decomposition. *Linear Algebra and its Applications 88-89* (1987), 223–238.

[86] Gill, P. E., Golub, G. H., Murray, W., and Saunders, M. A. Methods for modifying matrix factorizations. *Mathematics of Computation 28*, 126 (1974), 505–535.

[87] Giscard, P.-L., and Pozza, S. Lanczos-like method for the time-ordered exponential. *arXiv preprint arXiv:1909.03437* (2020).

[88] Gohberg, I., Kailath, T., and Olshevsky, V. Fast Gaussian elimination with partial pivoting for matrices with displacement structure. *Mathematics of computation 64*, 212 (1995), 1557–1576.

[89] Golub, G. H., and Meurant, G. *Matrices, Moments and Quadrature with Applications.* Princeton University Press, 2009.

[90] Golub, G. H., and Van Loan, C. F. *Matrix computations*, third ed. John Hopkins University Press, 1996.

[91] Golub, G. H., and Welsch, J. H. Calculation of Gauss quadrature rules. *Mathematics of Computation 23*, 106 (1969), 221–230.

[92] Gragg, W. B., and Harrod, W. J. The numerically stable reconstruction of Jacobi matrices from spectral data. *Numerische Mathematik 44* (1984), 317–335.

[93] Gragg, W. B., and Lindquist, A. On the partial realization problem. *Linear Algebra and its Applications 50* (1983), 277–319.

[94] Grimme, E. J. *Krylov projection methods for model reduction.* PhD thesis, University of Illinois at Urbana Champaign, 1997.

[95] Grimme, E. J., Sorensen, D. C., and Van Dooren, P. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numerical Algorithms 12* (1996), 1–31.

[96] Guseinov, G. S. An inverse spectral problem for complex Jacobi matrices. *Communications in Nonlinear Science and Numerical Simulation 15*, 4 (2010), 840–851.

[97] Gutknecht, M. H. Lanczos-type solvers for nonsymmetric linear systems of equations. *Acta Numerica 6* (1997), 271–397.

[98] Güttel, S. *Rational Krylov Methods for Operator Functions.* PhD thesis, Technische Universität Bergakademie Freiberg, Germany, 2010. Dissertation available as MIMS Eprint 2017.39.

[99] Heinig, G., and Rost, K. Fast algorithms for Toeplitz and Hankel matrices. *Linear Algebra and its Applications 435*, 1 (2011), 1–59.

[100] Helsen, S., Kuijlaars, A. B., and Van Barel, M. Convergence of the isometric Arnoldi process. *SIAM Journal on Matrix Analysis and Applications 26*, 3 (2005), 782–809.

[101] Hendriksen, E., and Njåstad, O. A Favard theorem for rational functions. *Journal of Mathematical Analysis and Applications 142*, 2 (1989), 508–520.

[102] Hestenes, M. R., and Stiefel, E. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards 49*, 6 (1952), 409–436.

[103] Higham, D. J., and Higham, N. J. Structured backward error and condition of generalized eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications 20*, 2 (1998), 493–512.

[104] Higham, N. J. *Functions of matrices: theory and computation*. SIAM, 2008.

[105] Horn, R. A., and Johnson, C. R. *Matrix analysis*. Cambridge University press, Cambridge, 1985.

[106] Ikebe, Y. On inverses of Hessenberg matrices. *Linear Algebra and its Applications 24* (1979), 93–97.

[107] Inglese, G. A relation between the Hilbert and Vandermonde matrices. *Calcolo 27*, 1 (1990), 73–79.

[108] Ismail, M., and Masson, D. Generalized orthogonality and continued fractions. *Journal of Approximation Theory 83*, 1 (1995), 1 – 40.

[109] Ismail, M., and Sri Ranga, A. RII type recurrence, generalized eigenvalue problem and orthogonal polynomials on the unit circle. *Linear Algebra and its Applications 562* (2019), 63–90.

[110] Ismail, M. E. H. *Classical and quantum orthogonal polynomials in one variable*. Encyclopedia of mathematics and its applications 98. Cambridge university press, 2005.

[111] Jagels, C., and Reichel, L. On the construction of Szegő polynomials. *Journal of Computational and Applied Mathematics 46*, 1 (1993), 241–254.

[112] Jagels, C., and Reichel, L. Recursion relations for the extended Krylov subspace method. *Linear Algebra and its Applications 434*, 7 (2011), 1716–1732.

[113] Jagels, C., and Reichel, L. The structure of matrices in rational Gauss quadrature. *Mathematics of computation 82*, 284 (2013), 2035–2060.

[114] Jagels, C., and Reichel, L. On the computation of Gauss quadrature rules for measures with a monomial denominator. *Journal of Computational and Applied Mathematics 286* (2015), 102–113.

[115] Kailath, T., Kung, S.-Y., and Morf, M. Displacement ranks of matrices and linear equations. *Journal of Mathematical Analysis and Applications 68*, 2 (1979), 395–407.

[116] Kailath, T., and Sayed, A. H. Displacement structure: Theory and applications. *SIAM Review 37*, 3 (1995), 297–386.

[117] Kailath, T., and Sayed, A. H. *Fast reliable algorithms for matrices with structure*. SIAM, Philadelphia (Pa.), 1999.

[118] Krylov, A. N. On the numerical solution of the equation by which the frequency of small oscillations is dertermined in technical problems. *Izvestiya Akad. Nauk. SSSR, Ser. Fiz.-Mat. 4* (1931), 491–539. In Russian.

[119] Kuijlaars, A. B. Which eigenvalues are found by the Lanczos method? *SIAM Journal on Matrix Analysis and Applications 22*, 1 (2000), 306–321.

[120] Kuijlaars, A. B. J. Convergence analysis of Krylov subspace iterations with methods from potential theory. *SIAM Review 48*, 1 (2006), 3–40.

[121] Lanczos, C. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards 45*, 4 (1950), 255–282.

[122] Leon, S. J., Björck, Å., and Gander, W. Gram-Schmidt orthogonalization: 100 years and more. *Numerical Linear Algebra with Applications 20*, 3 (2013), 492–532.

[123] Liesen, J. When is the adjoint of a matrix a low degree rational function in the matrix? *SIAM Journal on Matrix Analysis and Applications 29*, 4 (2008), 1171–1180.

[124] Liesen, J., and Strakoš, Z. *Krylov Subspace Methods: Principles and Analysis*. Oxford University Press, 2013.

[125] Liesen, J., and Tichý, P. Convergence analysis of Krylov subspace methods. *GAMM-Mitteilungen 27*, 2 (2004), 153–173.

[126] Longley, J. W. Modified Gram-Schmidt process vs. classical Gram-Schmidt. *Communications in Statistics - Simulation and Computation 10*, 5 (1981), 517–527.

[127] López Lagomasino, G., Reichel, L., and Wunderlich, L. Matrices, moments, and rational quadrature. *Linear Algebra and its Applications 429*, 10 (2008), 2540–2554. Special Issue in honor of Richard S. Varga.

[128] Mach, T., Van Barel, M., and Vandebril, R. Inverse eigenvalue problems for extended Hessenberg and extended tridiagonal matrices. *Journal of Computational and Applied Mathematics 272* (2014), 377–398.

[129] Mastronardi, N., and Van Dooren, P. The $QR$ steps with perfect shifts. *SIAM Journal on Matrix Analysis and Applications 39*, 4 (2018), 1591–1615.

[130] Meurant, G. *The Lanczos and Conjugate Gradient Algorithms*. Society for Industrial and Applied Mathematics, 2006.

[131] Nakatsukasa, Y., Sète, O., and Trefethen, L. N. The AAA algorithm for rational approximation. *SIAM Journal on Scientific Computing 40*, 3 (2018), A1494–A1522.

[132] Paige, C. C. *The computation of eigenvalues and eigenvectors of very large sparse matrices*. PhD thesis, University of London, 1971.

[133] Palitta, D., Pozza, S., and Simoncini, V. The short-term rational Lanczos method and applications. *arXiv preprint arXiv:2103.04054* (2021).

[134] Pan, V. Y. *Structured Matrices and Polynomials: unified superfast algorithms.* Birkhäuser, Boston, MA, 2001.

[135] Parlett, B. N. Reduction to tridiagonal form and minimal realizations. *SIAM Journal on Matrix Analysis and Applications 13*, 2 (1992), 567–593.

[136] Pozza, S., and Pranić, M. The Gauss quadrature for general linear functionals, Lanczos algorithm, and minimal partial realization. *Numerical Algorithms* (2021). Published online: https://doi.org/10.1007/s11075-020-01052-y.

[137] Pozza, S., Pranić, M. S., and Strakoš, Z. The Lanczos algorithm and complex Gauss quadrature. *Electron. Trans. Numer. Anal. 50* (2018), 1–19.

[138] Pranić, M. S., and Reichel, L. Recurrence relations for orthogonal rational functions. *Numerische Mathematik 123* (2013), 629–642.

[139] Pranić, M. S., and Reichel, L. Rational Gauss quadrature. *SIAM Journal on Numerical Analysis 52*, 2 (2014), 832–851.

[140] Reichel, L. Fast QR decomposition of Vandermonde-like matrices and polynomial least squares approximation. *SIAM Journal on Matrix Analysis and Applications 12*, 3 (1991), 552–564.

[141] Reichel, L. Construction of polynomials that are orthogonal with respect to a discrete bilinear form. *Advances in Computational Mathematics 1*, 2 (1993), 241–258.

[142] Reichel, L., Ammar, G., and Gragg, W. Discrete least squares approximation by trigonometric polynomials. *Mathematics of Computation 57* (1991), 273–289.

[143] Ruhe, A. Rational Krylov sequence methods for eigenvalue computation. *Linear Algebra and its Applications 58* (1984), 391–405.

[144] Ruhe, A. Rational Krylov algorithms for nonsymmetric eigenvalue problems. II. Matrix pairs. *Linear algebra and its Applications 197* (1994), 283–295.

[145] Rutishauser, H. On Jacobi rotation patterns. In *Experimental Arithmetic, High Speed Computing and Mathematics. Proceedings of Symposia in Applied Mathematics* (Providence, 1963), vol. 15, American Mathematical Society, pp. 241–258.

[146] Rutishauser, H. *Description of Algol 60.* Handbook for automatic computation 1a. Springer, Berlin, 1967.

[147] Saad, Y. A flexible inner-outer preconditioned GMRES algorithm. *SIAM Journal on Scientific Computing 14*, 2 (1993), 461–469.

[148] Saad, Y. *Iterative Methods for Sparse Linear Systems*. SIAM, 2003.

[149] Saad, Y., and Schultz, M. H. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing 7*, 3 (1986), 856–869.

[150] Saylor, P. E., and Smolarski, D. C. Why Gaussian quadrature in the complex plane? *Numerical Algorithms 26* (2001), 251–280.

[151] Saylor, P. E., and Smolarski, S.J., D. C. Computing the roots of complex orthogonal and kernel polynomials. *SIAM Journal on Scientific and Statistical Computing 9*, 1 (1988), 1–13.

[152] Schweitzer, M. A two-sided short-recurrence extended Krylov subspace method for nonsymmetric matrices and its relation to rational moment matching. *Numerical Algorithms 76*, 1 (2017), 1–31.

[153] Simon, B. CMV matrices: Five years after. *Journal of Computational and Applied Mathematics 208*, 1 (2007), 120–154.

[154] Stahl, H., and Totik, V. *General orthogonal polynomials*. No. 43 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992.

[155] Stewart, M. A generalized isometric Arnoldi algorithm. *Linear Algebra and its Applications 423*, 2 (2007), 183–208.

[156] Strakoš, Z., and Tichý, P. On efficient numerical approximation of the bilinear form $c^*a^{-1}b$. *SIAM Journal on Scientific Computing 33*, 2 (2011), 565–587.

[157] Szegő, G. *Orthogonal Polynomials*, 4th ed. American Mathematical Society, Providence, RI, 1975.

[158] Taylor, A. E. *Introduction to functional analysis*, 6th print. ed. Wiley, New York (N.Y.), 1967.

[159] Tyrtyshnikov, E. E. How bad are Hankel matrices? *Numerische Mathematik 67*, 2 (1994), 261–269.

[160] Van Assche, W., and Vanherwegen, I. Quadrature formulas based on rational interpolation. *Mathematics of Computation 61*, 204 (1993), 765–783.

[161] Van Barel, M., Fasino, D., Gemignani, L., and Mastronardi, N. Orthogonal rational functions and structured matrices. *SIAM Journal on Matrix Analysis and Applications 26*, 3 (2005), 810–829.

[162] Van Buggenhout, N. RationalLanczos v0.1.0. `https://github.com/nielvb/RationalLanczos`, 2019.

[163] Van Buggenhout, N. ResearchCode. `https://github.com/nielvb/ResearchCode`, 2021.

[164] Van Buggenhout, N., Van Barel, M., and Vandebril, R. Downdating procedures for orthogonal polynomials and rational functions. In preparation.

[165] Van Buggenhout, N., Van Barel, M., and Vandebril, R. Biorthogonal rational Krylov subspace methods. *Electron. Trans. Numer. Anal. 51* (2019), 451–468.

[166] Van Buggenhout, N., Van Barel, M., and Vandebril, R. Non-unitary CMV-decomposition. *Special Matrices 8*, 1 (2020), 144 – 159.

[167] Van Buggenhout, N., Van Barel, M., and Vandebril, R. Generation of orthogonal rational functions by procedures for structured matrices. *Numerical Algorithms* (2021). Published online: https://doi.org/10.1007/s11075-021-01125-6.

[168] van der Vorst, H. A. BiCGSTAB a fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems. *SIAM Journal on Statistical Computing 13* (1992), 631–644.

[169] Van der Vorst, H. A. *Iterative Krylov Methods for Large Linear Systems*. Cambridge University Press, 2003.

[170] Van Deun, J., and Bultheel, A. Modified moments and orthogonal rational functions. *Applied Numerical Analysis & Computational Mathematics 1*, 2 (2004), 455–468.

[171] Van Dooren, P. A generalized eigenvalue approach for solving Riccati equations. *SIAM Journal on Scientific and Statistical Computing 2*, 2 (1981), 121–135.

[172] Vandebril, R. Chasing bulges or rotations? A metamorphosis of the QR-algorithm. *SIAM Journal on Matrix Analysis and Applications 32*, 1 (2011), 217–247.

[173] Vandebril, R., and Van Barel, M. A note on the nullity theorem. *Journal of Computational and Applied Mathematics 189*, 1 (2006), 179–190. Proceedings of The 11th International Congress on Computational and Applied Mathematics.

[174] Vandebril, R., Van Barel, M., Golub, G. H., and Mastronardi, N. A bibliography on semiseparable matrices. *Calcolo 42*, 3 (2005), 249–270.

[175] Vandebril, R., Van Barel, M., and Mastronardi, N. An implicit Q theorem for Hessenberg-like matrices. *Mediterranean Journal of Mathematics 2*, 3 (2005), 259–275.

[176] Vandebril, R., Van Barel, M., and Mastronardi, N. *Matrix Computations and Semiseparable Matrices: Linear Systems*, vol. 1. John Hopkins University Press, 2007.

[177] Vorobyev, Y. V. *Method of Moments in Applied Mathematics*. Gordon and Breach Science Publishers, New York, 1965. Translated from Russian by Bernard Seckler.

[178] Watkins, D. S. Understanding the QR algorithm. *SIAM Review 24*, 4 (1982), 427–440.

[179] Watkins, D. S. Some perspectives on the eigenvalue problem. *SIAM Review 35*, 3 (1993), 430–471.

[180] Watkins, D. S. The transmission of shifts and shift blurring in the QR algorithm. *Linear Algebra and its Applications 241-243* (1996), 877–896. Proceedings of the Fourth Conference of the International Linear Algebra Society.

[181] Watkins, D. S. The QR algorithm revisited. *SIAM Review 50*, 1 (2008), 133–145.

[182] Wilkinson, J. H. Plane rotations in floating-point arithmetic. In *Experimental Arithmetic, High Speed Computing and Mathematics. Proceedings of Symposia in Applied Mathematics* (Providence, 1963), vol. 15, American Mathematical Society, pp. 185–198.

[183] Wilkinson, J. H. *The algebraic eigenvalue problem*. Monographs on numerical analysis. Clarendon, Oxford, 1965.

[184] Wilkinson, J. H. Modern error analysis. *SIAM Review 13*, 4 (1971), 548–568.

[185] Xu, W.-R., Bebiano, N., and Chen, G.-L. Generalized inverse spectral problem for pseudo-Jacobi matrices with mixed eigendata. *Inverse Problems in Science and Engineering 27*, 6 (2019), 773–789.

[186] Xu, W.-R., Bebiano, N., and Chen, G.-L. An inverse eigenvalue problem for pseudo-Jacobi matrices. *Applied Mathematics and Computation 346* (2019), 423–435.

[187] Yuditskii, P. Killip–Simon problem and Jacobi flow on GMP matrices. *Advances in Mathematics 323* (2018), 811–865.

[188] Zha, H., and Zhang, Z. The Arnoldi process, short recursions, and displacement ranks. *Linear Algebra and its Applications 249*, 1 (1996), 169 – 188.

[189] Zhang, L., and Ao, J.-j. Inverse spectral problem for Sturm-Liouville operator with coupled eigenparameter-dependent boundary conditions of Atkinson type. *Inverse Problems in Science and Engineering 27*, 12 (2019), 1689–1702.

[190] Zhedanov, A. Rational spectral transformations and orthogonal polynomials. *Journal of Computational and Applied Mathematics 85*, 1 (1997), 67–86.

[191] Zhedanov, A. The "classical" Laurent biorthogonal polynomials. *Journal of Computational and Applied Mathematics 98*, 1 (1998), 121 – 147.

[192] Zhedanov, A. Biorthogonal rational functions and the generalized eigenvalue problem. *Journal of Approximation Theory 101*, 2 (1999), 303 – 329.

# Curriculum Vitae

## Studies

**2018-2021**
PhD in Engineering Science: Computer Science.
Department of Computer Science, KU Leuven, Belgium.

**2015-2017**
Master of Science in Mathematical Engineering.
KU Leuven, Belgium.

**2012-2015**
Bachelor of Science in Engineering Science: Computer Science.
KU Leuven, Belgium.

## Teaching

**2015-2016**
Teaching assistent for *Analyse II*.

**2016-2017**
Teaching assistent for *Analyse III*.

**2018-2021**
Teaching assistant for *Computergesteund probleemoplossen voor natuurkundigen*.

**2018-2021**
Teaching assistant for *Probleemoplossen en ontwerpen*.

**2018-2019**: Daily supervisor for Master thesis on *Inexact rational Krylov subspace methods for large systems of equations* by Steel, T., supervised by Vandebril, R. and Van Barel, M.

# List of publications

**Articles in international reviewed papers**

- Van Buggenhout, N., Van Barel, M., and Vandebril, R., Biorthogonal rational Krylov subspace methods, Electronic Transactions on Numerical Analysis, vol. 51 (2019), pp. 451-468. Published.
- Van Buggenhout, N., Van Barel, M., and Vandebril, R., Non-unitary CMV-decomposition, Special Matrices, vol. 8, no. 1, (2020), pp. 144-159. Published.
- Van Buggenhout, N., Van Barel, M., and Vandebril, R., Generation of orthogonal rational functions by procedures for structured matrices, Numerical Algorithms, (2021). Published online.

**Articles in preparation**

- Van Buggenhout, N., Van Barel, M., and Vandebril, R., Downdating procedures for orthogonal polynomials and rational functions. In preparation.

**Presentations at international conferences**

- Van Buggenhout, N., Van Barel, M., and Vandebril, R., Generating Orthogonal Rational Functions by Solving Structured Inverse Eigenvalue Problems, SIAM Conference on Applied Linear Algebra (LA21).
- Van Buggenhout, N., Vandebril, R., and Van Barel, M. (2019). Matrix Structures Related to Rational Krylov Subspaces. Presented at the The Seventh International Conference on Numerical Algebra and Scientific Computing (NASC 2019), Nanjing, China, 19 Oct 2019 - 23 Oct 2019.
- Van Buggenhout, N., Van Barel, M., and Vandebril, R. (2019). Matrix Structures in Rational Krylov Subspace Methods. Presented at the 22nd Conference of the

International Linear Algebra Society (ILAS), Rio de Janeiro, Brazil, 08 Jul 2019 - 12 Jul 2019.

- Van Buggenhout, N., Van Barel, M., and Vandebril, R. (2019). Biorthogonal Rational Krylov Subspace Methods. Presented at the Recent Advances in Scientific Computation (ETNA25), Santa Margherita di Pula, Italy, 27 May 2019 - 29 May 2019.
- Van Buggenhout, N., Vandebril, R., and Van Barel, M. (2019). Biorthogonal Rational Krylov Subspace Methods. Presented at the The International Conference Mathematical Modeling with Applications (M2A19), Mohammed V University, Rabat, Morocco, 01 Apr 2019 - 04 Apr 2019.
- Van Buggenhout, N., Van Barel, M., and Vandebril, R. (2018). Biorthogonal Extended Lanczos Iteration. Presented at the forty-third Woudschoten Conference, Zeist, The Netherlands, 03 Oct 2018 - 05 Oct 2018. (Poster)

FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF COMPUTER SCIENCE
NUMA
Celestijnenlaan 200A box 2402
B-3001 Leuven
niel.vanbuggenhout@kuleuven.be