Alahmadi, Jihan, Ph.D., December 2021 APPLIED MATHEMATICS

STANDARD AND RATIONAL GAUSS QUADRATURE RULES FOR THE APPROXIMATION OF MATRIX FUNCTIONALS (112 pages)

Dissertation Advisors: Lothar Reichel and Miroslav Pranić

In this thesis we develop efficient numerical methods for the approximation of matrix functionals of the form $F(A) := \boldsymbol{w}^T f(A)\boldsymbol{v}$, where A is a large symmetric or nonsymmetric matrix, $\boldsymbol{w}, \boldsymbol{v}$ are given vectors, and f is a function. Golub and Meurant describe a technique for computing upper and lower bounds for matrix functionals $F(A)$ based on the connection between the Lanczos process, orthogonal polynomials, and Gauss-type quadrature rules. Their technique considers the expression $F(A)$ as a Stieltjes integral. If the derivatives of the integrand f do not change sign on the convex hull of the support of the measure, then Gauss-type quadrature rules can be applied to compute upper and lower bounds for $F(A)$. However, when A is symmetric and derivatives of the integrand f change sign in the convex hull of the spectrum of A, or when the matrix A is nonsymmetric, then this approach is not guaranteed to yield upper and lower bounds.

We consider several extensions of the technique by Golub and Meurant for approximating matrix functions of the form $F(A)$. Let $A \in \mathbb{R}^{N\times N}$ be a large symmetric matrix. Our first extension is based on the use of pairs of Gauss, and suitable generalized Gauss–Radau or generalized Gauss–Lobatto rules that yield upper and lower bounds for $F(A)$ when some of the derivatives of f change sign on the convex hull of the support of the measure. We also describe new methods to evaluate these quadrature rules.

Our other extensions are concerned with the situation when the function f cannot be approximated accurately by a polynomial of small to moderate degree. Then Gauss-type rules may yield poor approximations of the functional $F(A)$. This situation occurs, for instance, when the function f has one or several singularities close to the support of the measure. This difficulty can be remedied by using rational Gauss rules. We discuss two approaches related to this case. First, we develop a technique to approximate matrix functionals of the form $F(A)$ with $A \in \mathbb{R}^{N\times N}$ a large nonsymmetric matrix when the function f has a singularity at the origin. We derive Gauss–Laurent quadrature rules that yield significantly more accurate approximations of $F(A)$ than Gauss-type rules with the same number of nodes. Also, we develop associated anti-Gauss–Laurent quadrature rules. Pairs of

Gauss–Laurent and anti-Gauss–Laurent rules can be applied to compute upper and lower bounds for $F(A)$ under suitable conditions. Second, we consider matrix functionals of the form $F(A)$, where $A \in \mathbb{R}^{N\times N}$ is a large symmetric positive definite matrix, and f is a Stieltjes function. In this case, upper and lower bounds can be computed using pairs of rational Gauss and Gauss–Radau rules or, under suitable conditions, pairs of rational Gauss and anti-Gauss rules that are determined by prescribed poles.

STANDARD AND RATIONAL GAUSS QUADRATURE RULES FOR THE APPROXIMATION OF MATRIX FUNCTIONALS

A dissertation submitted

to Kent State University

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

by

Jihan Alahmadi

December 2021

Dissertation written by

Jihan Alahmadi

B.S., Umm Al Qura University, 2005

M.S., Umm Al Qura University, 2011

Ph.D., Kent State University, 2021

Approved by

Lothar Reichel, Co-Chair, Doctoral Dissertation Committee

Miroslav Pranić, Co-Chair, Doctoral Dissertation Committee

Jing Li, Members, Doctoral Dissertation Committee

Jun Li

Mikhail Nesterenko

Arvind Bansal

Accepted by

Andrew M. Tonge, Chair, Department of Mathematical Sciences

Mandy Munro-Stasiuk, Interim Dean, College of Arts and Sciences

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisers, Professor Lothar Reichel and Professor Miroslav Pranić. I am very grateful to have been afforded the opportunity to work with you.

In particular, I would like to express my sincere gratitude to Professor Reichel for his constant assistance, encouragement, and guidance. I am greatly indebted to him for spending an incredible amount of time and untiring effort to discuss, proofread and correct all my works. I have learned so much from working with him. Thank you very much for being so patient and understanding with me.

I would like to thank Professor Pranić for sharing his valuable insights and knowledge with me. I want to take this opportunity to appreciate and thank him for always being available to answer all my questions. Thank you so much for the time and effort you have provided to help me.

I would like to thank Hessah Alqahtani who I had the privilege to work with.

Thank you to the faculty and staff at Mathematical Sciences Department at Kent State University for the assistance they have provided me.

Finally, I would like to thank my family for their boundless support and encouragement throughout my study.

CHAPTER 1

Introduction

The goal of this thesis is to study and propose efficient and inexpensive numerical methods to approximate the value of matrix functionals of the form $F(A) := \boldsymbol{w}^T f(A)\boldsymbol{v}$, where A is a large symmetric or nonsymmetric matrix, $\boldsymbol{w}, \boldsymbol{v}$ are given vectors, and f is a function. There are many applications that require the evaluation of the expression $F(A)$. For instance, consider the system of equations $A\boldsymbol{x} = \boldsymbol{b}$, where A is a large sparse symmetric positive definite matrix and $\boldsymbol{b}$ is a given vector. Let $\hat{\boldsymbol{x}}$ be an approximation to the exact solution, $\boldsymbol{x}$. Furthermore, let $\boldsymbol{r} = \boldsymbol{b} - A\hat{\boldsymbol{x}}$ be the residual vector. Then, the error $\boldsymbol{e} = \boldsymbol{x} - \hat{\boldsymbol{x}}$ is given by $\boldsymbol{e} = A^{-1}\boldsymbol{r}$. The square of the norm of the error is $\|\boldsymbol{e}\|_2^2 = \boldsymbol{r}^T A^{-2}\boldsymbol{r}$, which is of the form $F(A)$; see [42] for more details. Other applications arise in Tikhonov regularization and network analysis; see, e.g., [4, 11, 15, 17, 54].

In this research, we describe and explain the connection between large symmetric or nonsymmetric matrices, classical and rational Lanczos algorithms, orthogonal polynomials and rational functions, and quadrature rules.

Golub and collaborators developed elegant and powerful iterative methods for computing upper and lower bounds for $F(A)$ with $\boldsymbol{w} = \boldsymbol{v}$, utilizing the connection between the Lanczos procedure, orthogonal polynomials, and Gauss-type quadrature rules. However, in some cases, this approach is not guaranteed to provide upper and lower bounds. For instance, when A is symmetric and the derivatives of the integrand f change sign in the convex hull of the spectrum of A, or when the matrix A is nonsymmetric. Moreover, the quality of the upper and lower bounds obtained by applying a few steps of the standard Lanczos procedure to A depends on how well the function f can be approximated by a polynomial of small to moderate degree. However, when the integrand f has one or several singularities close to the spectrum of A, the approximation of $F(A)$ by a polynomial may not yield the desired accuracy with moderate computational effort. In this thesis, we introduce numerical methods that deal with these issues.

This thesis is organized as follows. Chapter 2 provides the necessary mathematical background, notation, definitions, and theorems that give the reader a good overview of the connections between orthogonal polynomials and rational functions, quadrature rules, and the Lanczos algorithm.

In Chapter 3 we discuss generalized Gauss-type rules which are Gauss-type quadrature formulas having nodes of arbitrary multiplicity at one or both end points of the interval of integration. Moreover, we present a new computational method for computing upper and lower bounds for $F(A)$, where $A \in \mathbb{R}^{N\times N}$ is a large symmetric matrix via pairs of Gauss, and generalized Gauss–Radau, or generalized Gauss–Lobatto rules. The advantage of these rules is that they give upper and lower bounds for $F(A)$ in some situations when pairs of Gauss and (standard) Gauss–Radau or Gauss–Lobatto rules are not guaranteed to furnish upper and lower bounds. Generalized Gauss–Radau rules are Gauss–Radau-type rules, in which the fixed node has multiplicity larger than one; similarly, generalized Gauss–Lobatto rules are Gauss–Lobatto-type rules, in which at least one of the fixed nodes has multiplicity larger than one. We evaluate these quadrature rules by modifying the tridiagonal matrix computed by a few steps of the Lanczos process.

When the integrand f has one or several singularities at or close to the spectrum of A, then Gauss-type quadrature rules with a small or moderate number of nodes will not yield accurate approximations of the expression $F(A)$. Then it may be useful to approximate f by rational functions with poles at or close to the singularities of f. In Chapters 4 and 5 , we present efficient methods for computing error bounds (or estimates of bounds) for matrix functions of the form $F(A)$ with a symmetric or nonsymmetric matrix A based on the rational Lanczos algorithm.

In Chapter 4 we assume that the singularity of the integrand f is at the origin. In this case, we use rational Gauss quadrature rules for the approximation of matrix functionals of the form $F(A)$ with $A \in \mathbb{R}^{N\times N}$ a large nonsymmetric matrix. These rules are exact for certain Laurent polynomials, which are polynomials in x and x^{-1}. They are referred to as Gauss–Laurent quadrature rules. These rules can provide much higher accuracy than (standard) Gauss-type rules with the same number of nodes. Gauss–Laurent quadrature rules can be determined by applying a few steps of the nonsymmetric rational Lanczos process to the matrix A with initial vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. We also define extended Krylov subspaces $\mathbb{K}_{\ell,m}$ corresponding to Laurent polynomials. These spaces require the evaluation of $m-1$ matrix-vector products with A and A^T, and require the solution of $\ell-1$ linear systems of equations with A and A^T. For many matrices, the evaluation of matrix-vector

products with A and A^T can be carried out faster than the solution of systems of equations with A and A^T. This suggests that it may be beneficial to choose m larger than ℓ in $\mathbb{K}_{\ell,m}$. In addition, we introduce the associated anti-Gauss–Laurent quadrature rules, allowing one to compute estimates of upper and lower bounds for the quadrature error in Gauss–Laurent rules.

Chapter 5 is concerned with the approximation of expressions of the form $F(A)$, where $A \in \mathbb{R}^{N\times N}$ is a large symmetric positive definite matrix, and f is a Stieltjes function. This class of function is defined in the complex plane except on the negative real axis. We therefore generalize the discussion of Chapter 4 to the case where Laurent polynomials are replaced by rational functions with prescribed poles on the negative real axis. In this chapter, we use rational Gauss rules that are presented in [56]. These rules are based on orthogonal rational functions that satisfy short recurrence relations and have a few prescribed poles. The number of terms in the recursion relations depends both on the number of distinct poles and their order. We also define the associated rational Gauss–Radau and rational anti-Gauss rules. Moreover, we describe how upper and lower bounds or estimates of upper and lower bounds for $F(A)$ can be computed by pairs of rational Gauss and Gauss–Radau quadrature rules, or by pairs of rational Gauss and anti-Gauss quadrature rules. The computation of rational anti-Gauss rules requires one more step of the symmetric rational Lanczos process. In order to reduce the computational cost, we introduce simplified rational anti-Gauss quadrature rules that require to determine the same number of steps of the symmetric rational Lanczos algorithm as the corresponding rational Gauss rule.

We conclude each chapter with numerical examples that illustrate the accuracy of the proposed methods with the existing methods. All computations were carried out using MATLAB R2017b on a 64-bit MacBook Pro personal computer with about 15 significant decimal digits.

CHAPTER 2

Background and Notation

In this chapter we provide basic concepts, properties, and algorithms that will be used throughout this thesis.

2.1 Orthogonal Polynomials and Rational Functions

We first define orthogonality with the aid of the definition of an inner product for functions of a real variable by using Riemann–Stieltjes integrals; see Golub and Meurant [29] and [64] for more discussions on Riemann–Stieltjes integrals.

Definition 1. *Let $[a,b] \subset \mathbb{R}$ be a finite interval. A partition $\wp$ of $[a,b]$ is a finite set of points of $[a,b]$ that satisfy*

$$a = x_0 \leq x_1 \leq \ldots \leq x_l = b.$$

The norm of $\wp$ is defined as

$$\|\wp\| = \max_{1 \leq i \leq l} \{x_i - x_{i-1}\}.$$

A sequence $c_i, i = 1, 2, \ldots, l$, of real numbers is said to be a choice sequence for $\wp$ if it satisfies

$$x_{i-1} \leq c_i \leq x_i, \quad \text{for} \quad i = 1, 2, \ldots, l.$$

Definition 2. *Let $[a,b] \subset \mathbb{R}$ be a finite interval, and let f and λ be real-valued functions defined on $[a,b]$. Further, let $\wp = \{x_0, x_1, \ldots, x_l\}$ be a partition of $[a,b]$ and let $c_i, i = 1, 2, \ldots, l$ be a choice sequence for $\wp$. Then the Riemann–Stieltjes sum of f and λ is defined as*

$$S((x_i)_{i=0,\ldots,l}, (c_i)_{i=1,\ldots,l}) = \sum_{i=1}^{l} f(c_i)(\lambda(x_i) - \lambda(x_{i-1})).$$

If there is a number $S \in \mathbb{R}$ such that for any $\varepsilon > 0$ there exists a $\delta > 0$ satisfying

$$\|S((x_i)_{i=0,\ldots,l}, (c_i)_{i=1,\ldots,l}) - S\| < \varepsilon,$$

whenever $\|\wp\| < \delta$*, then* S *is called the Riemann–Stieltjes integral of* f *with respect to* λ *on* $[a, b]$ *and is denoted by*

$$\mathcal{I}(f) := \int_a^b f(x) d\lambda(x). \tag{2.1}$$

The function f *is called the integrand and* λ *is a nondecreasing function on* $[a, b]$ *having finite limits at* $\pm\infty$ *if* $a = -\infty$ *and/or* $b = +\infty$ *and with infinitely many points of increase. The set of all points of increase of* λ *is called the support (or spectrum) of the measure* $d\lambda$*, and the smallest closed interval that contains these points is called the convex hull.*

If $\lambda(x) = x$ in (2.1), then the Riemann–Stieltjes integral reduces to the classical Riemann integral. In many cases, the Riemann–Stieltjes integral (2.1) can be written as

$$\int_a^b f(x) w(x) dx, \tag{2.2}$$

where w is a nonnegative integrable function on $[a, b]$, called a weight function. The integral (2.2) must be well defined if the interval $[a, b]$ is infinite. This can be achieved by requiring that all moments,

$$\mu_i = \int_a^b x^i w(x) dx, \quad i = 0, 1, 2, \ldots, \tag{2.3}$$

exist and be finite. See Golub and Meurant [29] for discussions on the existence of the Riemann–Stieltjes integral.

Let $\mathbb{P}$ denotes the space of all real polynomials and let $\mathbb{P}_d \subset \mathbb{P}$ be the space of all polynomials of degree at most d. A polynomial $p \in \mathbb{P}_d$ is said to be monic if its leading coefficient is equal to one, that is

$$p_d(x) = x^d + c_{d-1} x^{d-1} + c_{d-2} x^{d-2} + \ldots.$$

Definition 3. *We define an inner product associated with the measure* $d\lambda$ *as*

$$\langle f, g \rangle = \int_a^b f(x) g(x) d\lambda(x). \tag{2.4}$$

Let p *and* $q \in \mathbb{P}$*. The polynomials* p *and* q *are said to be orthogonal with respect to the inner product* (2.4) *if* $\langle p, q \rangle = 0$*. The norm of* p *is defined as*

$$\|p\| = \Big(\int_a^b p(x)^2 d\lambda(x) \Big)^{\frac{1}{2}}.$$

Definition 4. *A sequence of polynomials $p_i, i = 0, 1, 2, \ldots$, are called monic orthogonal polynomials with respect to the measure $d\lambda$, if they are monic and*

$$\langle p_i, p_j \rangle = 0, \quad \text{if} \quad i \neq j \quad i, j = 0, 1, 2, \ldots,$$
$$\|p_i\| > 0, \quad \text{for} \quad i = 0, 1, 2, \ldots.$$

By normalizing

$$\tilde{p_i} = p_i / \|p_i\|, i = 0, 1, 2, \ldots,$$

we obtain orthonormal polynomials, which satisfy

$$\langle \tilde{p_i}, \tilde{p_j} \rangle = \delta_{ij} := \begin{cases} 0, & \text{if} \quad i \neq j, \\ 1, & \text{if} \quad i = j. \end{cases}$$

Definition 5. *The inner product* (2.4) *is said to be positive-definite on $\mathbb{P}$ if $\|p\| > 0$ for all $p \in \mathbb{P}$, $p \neq 0$, and it is said to be positive-definite on $\mathbb{P}_d$ if $\|p\| > 0$ for any $p \in \mathbb{P}_d, p \neq 0$*

The following result provides a sufficient condition for guaranteeing the existence of orthogonal polynomials; see Gautschi [22] for more details.

Theorem 1. *If the inner product* (2.4) *is positive-definite on $\mathbb{P}$, then there exists a unique infinite sequence of monic orthogonal polynomials related to the measure $d\lambda$.*

Proof. *The polynomials $p_j(x)$ can be generated by applying the Gram–Schmidt process to the basis $1, x, x^2, \ldots$. Let $p_0(x) = 1$, and for $j = 1, 2, \ldots$ recursively generate*

$$p_j(x) = x^j - \sum_{i=0}^{j-1} c_i p_i(x), \quad c_i = \frac{\langle x^j, p_i \rangle}{\langle p_i, p_i \rangle}. \tag{2.5}$$

By the positive definiteness of the inner product, $\langle p_i, p_i \rangle > 0$. Thus, the polynomial $p_j(x)$ can be uniquely defined, and it is orthogonal to all polynomials $p_0(x), p_1(x), \ldots, p_{j-1}(x)$.

This condition is sufficient but not necessary for a polynomial to be exist. The inner product is enough to be different from zero, and it is not required to be positive. For instance, in Chapter 4, the inner product is defined with respect to a complex-valued measure. In this case, the sequence of orthogonal Laurent-polynomials can be generated, but the inner product is not positive definite.

The hypothesis of Theorem 1 is satisfied if λ has infinitely many points of increase. Assume now that λ has exactly $n \geq 1$ points of increase on the interval $[a, b]$. Then the Gram–Schmidt process

can be applied as long as $\langle p_i, p_i \rangle > 0$ in (2.5), that is, for $j \leq n$. The last generated polynomial $p_n(x)$ has norm zero. Thus, $p_n(x)$ vanishes on the finite support of λ. The Riemann-Stieltjes integral with respect to the measure $d\lambda$ where λ has finitely many points of increase plays a significant role in the approximation of matrix functions.

An important property of orthogonal polynomials is that they satisfy a three-term recurrence relation. More details of three-term recurrence relations can be found in Golub and Meurant [29].

Theorem 2. *Let $\{p_i\}_{i=0}^{\infty}$ be a sequence of monic orthogonal polynomials. Then there are two sequences of coefficients $\{\alpha_i\}_{i=0}^{\infty}$ and $\{\beta_i\}_{i=1}^{\infty}$ such that*

$$p_j(x) = (x - \alpha_{j-1})\, p_{j-1}(x) - \beta_{j-1} p_{j-2}(x), \quad j = 1, 2, \ldots,$$

$$p_{-1}(x) \equiv 0, \quad p_0(x) \equiv 1,$$

where

$$\alpha_{j-1} = \frac{\langle x p_{j-1}, p_{j-1} \rangle}{\langle p_{j-1}, p_{j-1} \rangle}, \quad j = 1, 2, \ldots,$$
$$\beta_{j-1} = \frac{\langle p_{j-1}, p_{j-1} \rangle}{\langle p_{j-2}, p_{j-2} \rangle}, \quad j = 2, 3, \ldots \ .$$

Since $p_{-1}(x) = 0$, the coefficient β_0 does not have to be defined.

Polynomials can be viewed as rational functions whose poles are located at infinity. By fixing a sequence of points $\{\zeta_1, \zeta_2, \ldots, \zeta_n\} \subset (\mathbb{C} \cup \{\infty\}) \backslash \{0\}$, where $\mathbb{C}$ denotes the complex plane, we define

$$\pi_n(x) := \prod_{j=1}^{n} (1 - x/\zeta_j).$$

Let Ω_n be the space of rational functions with poles at $\{\zeta_j\}_{j=1}^{n}$. Some of the poles can be repeated, and they can be fixed at infinity. A function f belongs to Ω_n if and only if it is of the form

$$f(x) = \frac{p_n(x)}{\pi_n(x)}, \quad p_n \in \mathbb{P}_n.$$

Now, define factors

$$Z_j(x) = \frac{x}{(1 - x/\zeta_j)}, \quad j = 1, 2, \ldots,$$

and products

$$\psi_0 \equiv 1, \quad \psi_j = Z_j(x) \psi_{j-1}, \quad j = 1, 2, \ldots \ .$$

Then, a basis for Ω_n is given by

$$\Omega_n := \text{span}\{\psi_0(x), \psi_1(x), \ldots, \psi_n(x)\},$$

Note that if $\zeta_j = \infty$ for all j, then

$$Z_j(x) = x, \quad \psi_j = x^j.$$

Thus, the space Ω_n reduces to the space of polynomials.

Assume that all moments μ_i defined by (2.3) exist and $\mu_0 = 1$. Then, the orthonormal rational functions $\{\phi_0, \phi_1, \ldots\}$ can be generated by applying the Gram–Schmidt process with respect to the inner product (2.4) and associated norm to the elements $\psi_j, j = 0, 1, \ldots$. Since $\mu_0 = 1$ and $\psi_0 = 1$, it follows that $\phi_0 = 1$.

Orthogonal rational functions with prescribed poles can satisfy short recurrence relations, similarly as orthogonal polynomials. The recursion relations are more complicated for orthogonal rational functions than for orthogonal polynomials, because they depend on the number of distinct poles and their ordering.

It is shown in [10] that the orthonormal rational functions ϕ_n with real poles, satisfy the following three-term recurrence relation

$$\phi_n(x) = \left(a_n Z_n + b_n \frac{Z_n}{Z_{n-1}}\right) \phi_{n-1}(x) + c_n \frac{Z_n}{Z_{n-2}} \phi_{n-2}(x), \quad \forall n \geq 1,$$

$$\phi_{-1}(x) \equiv 0, \quad \phi_0(x) \equiv 1,$$

for subtitle constants a_n, b_n and c_n. In the case of complex poles, the same relation holds, if ζ_{n-2} is replaced with $\bar{\zeta}_{n-2}$, where the bar denotes complex conjugation; see [69] for more details. In Chapters 4 and 5, we discuss orthogonal rational functions with fixed poles that satisfy short recurrence relations, analogous to the three-term recurrence relation for orthogonal polynomials.

2.2 Quadrature Rules

An n-point quadrature rule for the approximation of the integral (2.1) is of the form

$$\int_a^b f(x) d\lambda(x) = \sum_{j=1}^{n} w_j f(x_j) + \mathcal{E}_n(f), \tag{2.6}$$

where x_j and w_j are the nodes and weights of the quadrature rule, respectively. The sum in the right-hand side is an approximation of the integral on the left-hand side and the term $\mathcal{E}_n(f)$ is the quadrature error; see Golub and Meurant [29] and Gautschi [22] for discussions on quadrature rules.

Definition 6. *The quadrature rule* (2.6) *is said to have degree of exactness* d *if*

$$\mathcal{E}_n(p) = 0, \quad \forall p \in \mathbb{P}_d. \tag{2.7}$$

It is said to have precise degree of exactness d *if it has degree of exactness* d *but not* $d+1$*, that is if* (2.7) *holds but* $\mathcal{E}_n(p) \neq 0$ *for some* $p \in \mathbb{P}_{d+1}$*.*

It is well known that there exists a unique optimal choice of the nodes and weights that allow the quadrature rule (2.6) to integrate exactly all polynomials of degree at most $2n-1$. The resulting quadrature rule is called a Gauss quadrature rule. It was introduced by Gauss at the beginning of the nineteenth century.

2.2.1 Gauss-Type Quadrature Rules

In this subsection, we will briefly review some properties of Gauss-type quadrature rules; see, e.g., Golub and Meurant [29] or Gautschi [22] for more discussions on Gauss rules.

Definition 7. *The* n*-point Gauss quadrature rule is given by*

$$\mathcal{G}_n(f) := \sum_{i=1}^{n} w_i f(x_i), \tag{2.8}$$

and is characterized by the property that

$$\mathcal{I}(f) = \mathcal{G}_n(f), \quad \forall f \in \mathbb{P}_{2n-1}.$$

Assume that the integrand f *is* $2n$ *continuous derivatives on* $[a,b]$*. Then the error in the Gauss quadrature rule* (2.8) *can be expressed as*

$$\mathcal{E}_n(f) := (\mathcal{I} - \mathcal{G}_n)(f) = \frac{f^{(2n)}(x_G)}{(2n)!} \int_a^b \prod_{i=1}^{n} (x - x_i)^2 \, d\lambda(x), \tag{2.9}$$

for some $x_G \in (a,b)$*, where* $f^{(2n)}(x)$ *denotes the* $2n$*th derivative; see, e.g., [22, 29] for a proof.*

The following results illustrate the relation between orthogonal polynomials and Gauss quadrature. Complete proofs can be found in [66].

Theorem 3. *Let* $n \in \mathbb{N}$*. The nodes of the* n*-point Gauss quadrature rule* x_i *are distinct and the roots of the orthogonal polynomial* p_n *with respect to the measure* $d\lambda$*.*

Theorem 4. *The weights of the n-point Gauss quadrature rule w_i are the solution of the (nonsingular) system of equations*

$$\sum_{i=1}^{n} w_i p_k(x_i) = \begin{cases} 1, & \text{if } k = 0, \\ 0, & \text{if } k = 1, 2, \ldots, n-1. \end{cases}$$

The above two theorems present theoretical methods for the approximation of the Stieltjes integral (2.1) using Gauss quadrature rules (2.8). In fact, the nodes and weights of the Gauss rule can be more easily computed by other methods.

Let us assume that there is a family of orthonormal polynomials $p_0(x), p_1(x), \ldots$ with respect to λ:

$$\int_a^b p_i(x) p_j(x) d\lambda(x) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

These polynomials satisfy the three-term recurrence relation

$$\begin{gathered} \beta_j p_j(x) = (x - \alpha_{j-1})\, p_{j-1}(x) - \beta_{j-1} p_{j-2}(t), \quad j = 1, 2, \ldots, n, \\ p_{-1}(x) \equiv 0, \quad p_0(x) \equiv 1, \end{gathered} \tag{2.10}$$

if $\int_a^b d\lambda = 1$. The three-term recurrence relation can be written in a matrix form as

$$x\mathbf{P}(x) = T_n \mathbf{P}(x) + \beta_n p_n(x) e_n, \tag{2.11}$$

where $\mathbf{P}(x) = [p_0(x), p_1(x), \ldots, p_{n-1}(x)]^T$ and $e_n = [0, 0, \ldots, 0, 1]^T$. Moreover, the matrix

$$T_n := \begin{bmatrix} \alpha_0 & \beta_1 & & & 0 \\ \beta_1 & \alpha_1 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-2} & \alpha_{n-2} & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{2.12}$$

is symmetric and tridiagonal (i.e. a Jacobi matrix), and it is determined by the first $2n-1$ nontrivial recursion coefficients. Moreover, all the eigenvalues of T_n are real and simple, since $\beta_j \neq 0$ for $j = 1, 2, \ldots, n-1$.

Now, suppose that the x_i, $i = 1, 2, \ldots, n$, are the roots of $p_n(x)$; that is, $p_n(x_i) = 0$. Then, relation (2.11) reduces to

$$x_i \mathbf{P}(x_i) = T_n \mathbf{P}(x_i),$$

which shows that the x_i are eigenvalues of T_n, and $\mathbf{P}(x_i)$ are corresponding eigenvectors. The proof of the following theorem can be found in Golub and Meurant [29].

Theorem 5. *The eigenvalues of the Jacobi matrix* (2.12) *are the nodes* x_i *of the Gauss quadrature rule* (2.8) *and the weights* w_i *are the squares of the first elements of the normalized eigenvectors.*

In addition to Gauss rules, there are related quadrature rules having a fixed number of nodes. If there is a single node, x_0, which is fixed at an endpoint of the interval of integration, the corresponding $(n+1)$-point quadrature rule

$$\mathcal{G}_{n,1}(f) = w_0 f(x_0) + \sum_{i=1}^{n} w_i f(x_i), \tag{2.13}$$

is known as a *Gauss–Radau* quadrature rule with $-\infty < x_0 \leq a$ or $b \leq x_0 < \infty$ as a prescribed node. This rule is exact for all polynomials of degree at most $2n$. Moreover, if both of the endpoints, $-\infty < x_{0,1} \leq a$ and $b \leq x_{0,2} < \infty$, are prescribed nodes, then the $(n+2)$-point quadrature rule

$$\mathcal{G}_{n,1,1}(f) = w_0 f(x_{0,1}) + w_{n+1} f(x_{0,2}) + \sum_{i=1}^{n} w_i f(x_i), \tag{2.14}$$

is called a *Gauss–Lobatto* quadrature rule. It is exact for all polynomials of degree at most $2n+1$.

Similarly as for Gauss rules, we define reminder formulas for Gauss–Radau and Gauss–Lobatto rules. Let $\mathrm{int}(a, b, x_0)$ denote the convex hull of the set $\{a, b, x_0\}$. If the integrand f has $2n+1$ continuous derivatives in $\mathrm{int}(a, b, x_0)$. Then, the error in the Gauss–Radau quadrature rule (2.13) can be expressed as

$$\mathcal{E}_{n,1}(f) := (\mathcal{I} - \mathcal{G}_{n,1})(f) = \frac{f^{(2n+1)}(x_{GR})}{(2n+1)!} \int_a^b (x - x_0) \prod_{i=1}^{n} (x - x_i)^2 \, d\lambda(x), \tag{2.15}$$

for some $x_{GR} \in \mathrm{int}(a, b, x_0)$. Moreover, if the integrand f is $2n+2$ continuous derivatives in $\mathrm{int}(a, b, x_{0,1}, x_{0,2})$, the convex hull of the set $\{a, b, x_{0,1}, x_{0,2}\}$, then the error in the Gauss–Lobatto quadrature rule (2.14) can be expressed as

$$\mathcal{E}_{n,1,1}(f) := (\mathcal{I} - \mathcal{G}_{n,1,1})(f) = \frac{f^{(2n+2)}(x_{GL})}{(2n+2)!} \int_a^b (x - x_{0,1})(x - x_{0,2}) \prod_{i=1}^{n} (x - x_i)^2 \, d\lambda(x), \tag{2.16}$$

for some $x_{GL} \in \mathrm{int}(a, b, x_{0,1}, x_{0,2})$.

When the derivatives of f are of known constant sign, the error formulas (2.9), (2.15) and (2.16) are of known sign, and the quadrature rules (2.8) and (2.13) or (2.14) provide upper and lower bounds for (2.1). For instance, when $f^{(2n)}(x_G) > 0$ in (2.9), we have $\mathcal{E}_n > 0$ and, therefore,

$$\mathcal{G}_n(f) < \mathcal{I}(f).$$

If in addition, $f^{(2n+1)}(x_{GR}) < 0$ in (2.15), then an upper bound for $\mathcal{I}(f)$ can be determined by the application of an $(n+1)$-point Gauss–Radau quadrature rule with a fixed node at $x_0 \leq a$; see, e.g., [29] for more details.

In 1996, Laurie introduced the $(n+1)$-point anti-Gauss quadrature rule $\widetilde{\mathcal{G}}_{n+1}$ that gives an error of the same magnitude as the n-point Gauss quadrature rule and of opposite sign when applied to polynomials of degree at most $2n+1$, that is,

$$(\mathcal{I} - \widetilde{\mathcal{G}}_{n+1})(p) = -(\mathcal{I} - \mathcal{G}_n)(p), \quad \forall p \in \mathbb{P}_{2n+1}. \tag{2.17}$$

Therefore, for $f \in \mathbb{P}_{2n+1}$, pairs of the quadrature rules $\mathcal{G}_n$ and $\widetilde{\mathcal{G}}_{n+1}$ provide error bounds for the exact value of the integral (2.1). Moreover, an average of the Gauss and anti-Gauss formulae may yield more accurate results. The property (2.17) forms the basis for using pairs of Gauss and anti-Gauss rules to compute approximate bounds for $F(A)$. It was first applied to matrix functions in [12].

Gauss-type quadrature rules will only be an accurate approximation to the integral (2.1) if the integrand f can be well approximated by polynomials of small to moderate degree. However, when the integrand f has one or several singularities close to the interval of integration, Gauss-type quadrature rules (2.8), (2.13) and (2.14) will not provide accurate approximations of the expression (2.1). This difficulty can be remedied by using rational Gauss-type quadrature rules that are discussed in Chapters 4 and 5.

Similarly as orthogonal polynomials are related to Gauss quadrature rules, orthogonal rational functions are related to rational Gauss quadrature rules. The latter rules are exact for certain rational functions with prescribed poles.

We briefly review some properties of rational Gauss quadrature rules. Given an integer k with

$$0 \leq k \leq n,$$

we define the k-dimensional space of rational functions

$$\mathbb{Q}_k := \operatorname{span}\left\{ \tfrac{1}{(x-\zeta_j)^s}, s = 1, 2, \ldots, s_j; j = 1, 2, \ldots, \hat{d} \right\},$$
$$k := \textstyle\sum_{j=1}^{\hat{d}} s_j,$$

whose poles ζ_j are real or complex numbers. They are chosen to be at or close to singularities of the integrand f. The integers s_j are the multiplicities of the prescribed poles.

Consider a quadrature rule

$$\widehat{\mathcal{G}}_n(f) = \sum_{i=1}^{n} \widehat{w}_i f(\widehat{x}_i), \tag{2.18}$$

associated with the measure $d\lambda$. The aim is to construct a formula (2.18) such that it satisfies

$$\widehat{\mathcal{G}}_n(f) = \mathcal{I}(f), \quad \forall f \in \mathbb{S}_{2n}, \tag{2.19}$$

where

$$\mathbb{S}_{2n} := \mathbb{P}_{2n-k-1} \oplus \mathbb{Q}_k$$

is $2n$-dimensional linear space. The following theorem shows how to construct the desired quadrature rule. See Gautschi [22] for the proof and more discussions on rational Gauss rules.

Theorem 6. *Let*

$$\omega(x) = \prod_{j=1}^{d} (x - \zeta_j)^{s_j},$$

be a polynomial of exact degree k. Assume that the measure $d\lambda/\omega$ admits an n-point (standard) Gaussian quadrature formula

$$\int_a^b f(x) \frac{d\lambda(x)}{\omega(x)} = \sum_{i=1}^{n} w_i f(x_i), \quad \forall f \in \mathbb{P}_{2n-1},$$

having distinct nodes x_i contained in the support interval $[a, b]$ of $d\lambda$. Then,

$$\widehat{x}_i = x_i, \text{ and } \widehat{w}_i = \omega(x_i) w_i,$$

yields the desired n-point rational Gauss rule (2.18) *satisfying* (2.19).

Analogously to the standard Gauss rules, the zeros of orthogonal rational functions are nodes of rational Gauss rules. They can be computed as the eigenvalues of a symmetric banded matrix

defined by the recursion relations. The square of the first components of the associated eigenvectors of this matrix yield the weights of the rational Gauss quadrature rules; see, e.g., [42, 56] and Section 4.5 for more discussions.

2.3 The Lanczos Algorithm

In this section, we introduce the Lanczos algorithm for symmetric matrices as well as the nonsymmetric Lanczos algorithm. These algorithms are used to approximate eigenvalues of large sparse matrices, calculate quadrature formulas and estimate bilinear forms. Before we proceed, we define Krylov subspace upon which the Lanczos algorithm is based.

Definition 8. *Given a matrix $A \in \mathbb{R}^{N\times N}$ and initial vector $\boldsymbol{v} \in \mathbb{R}^N$, the nth Krylov subspace $\mathbb{K}_n(A, \boldsymbol{v})$ is spanned by n column vectors:*

$$\mathbb{K}_n(A, \boldsymbol{v}) = \operatorname{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{n-1}\boldsymbol{v}\}. \tag{2.20}$$

Krylov subspace methods are often used to compute an approximate solution to $A^{-1}\boldsymbol{v}$, $f(A)\boldsymbol{v}$, or $\boldsymbol{v}^T f(A)\boldsymbol{v}$, for a given large matrix A, and function f. The superscript ^{T} denotes transposition. Krylov subspaces have many properties that have proven their worth. A basis for a Krylov subspace can be constructed by evaluating matrix–vector products with A. This feature is useful when the matrix A is large and either sparse or structured, since then matrix–vector products with A can be carried out inexpensively. Moreover, each vector $\boldsymbol{x} \in \mathbb{K}_n$ can be expressed in the form $p(A)\boldsymbol{v}$, for some polynomial p of degree at most $n-1$. Conversely, if p is a polynomial of degree at most $n-1$, then $p(A)\boldsymbol{v} \in \mathbb{K}_n$. This yields an alternative characterization of $\mathbb{K}_n$:

$$\mathbb{K}_n(A, \boldsymbol{v}) = \{p(A)\boldsymbol{v},\ p \text{ a polynomial of degree} \leq n-1\}. \tag{2.21}$$

In the light of this observation, we may define an inner product on $\mathbb{P}_{n-1}$ via a given inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{K}_n$ as follows

$$\langle p, q\rangle := \langle p(A)\boldsymbol{v}, q(A)\boldsymbol{v}\rangle. \tag{2.22}$$

We now turn to the Lanczos algorithm which is the most important application of Krylov subspaces.

2.3.1 The Symmetric Lanczos Algorithm

Given $A \in \mathbb{R}^{N\times N}$ a symmetric matrix and an initial vector $\boldsymbol{v} \in \mathbb{R}^N$, the symmetric Lanczos algorithm is defined by the following recursion Algorithm 1

Algorithm 1 Symmetric Lanczos algorithm

1: **Input:** Integer n, matrix $A \in \mathbb{R}^{N\times N}$, vector $\boldsymbol{v} \in \mathbb{R}^N$.
2: **Output:** Orthogonal vectors $\{\boldsymbol{v}_i\}_{i=1}^n$
3: Initialize: $\boldsymbol{v}_0 := \boldsymbol{0}; \beta_0 = 0; \boldsymbol{v}_1 := \boldsymbol{v}/\|\boldsymbol{v}\|$;
4: **for** $i = 1, 2, \ldots, n$ **do**
5: $\quad \alpha_{i-1} := \boldsymbol{v}_i^T(A\boldsymbol{v}_i - \beta_{i-1}\boldsymbol{v}_{i-1})$;
6: $\quad \boldsymbol{u}_i := A\boldsymbol{v}_i - \alpha_{i-1}\boldsymbol{v}_i - \beta_{i-1}\boldsymbol{v}_{i-1}$;
7: $\quad \beta_i := \|\boldsymbol{u}_i\|$;
8: $\quad \boldsymbol{v}_{i+1} := \boldsymbol{u}_i/\beta_i$;
9: **end for**

The core of Algorithm 1 is the relation

$$\beta_i \boldsymbol{v}_{i+1} = (A - \alpha_{i-1})\boldsymbol{v}_i - \beta_{i-1}\boldsymbol{v}_{i-1}.$$

This three-term recurrence relation is reminiscent of the standard three-term recurrence relation of orthonormal polynomials (2.10). It also tells us that the vectors $\boldsymbol{v}_i$ form an orthonormal basis for $\mathbb{K}_n(A, \boldsymbol{v})$; i.e, they can be expressed as

$$\boldsymbol{v}_i = p_{i-1}(A)\boldsymbol{v}, \quad i = 1, \ldots, n,$$

for certain polynomials p_{i-1} of degree $i-1$. The orthogonality of the vectors $\boldsymbol{v}_i$ translates into the orthogonality of these polynomials with respect to the inner product (2.22). The Lanczos procedure is analogously to the Stieltjes algorithm for computing a sequence of orthogonal polynomials with respect to the inner product (2.22).

2.3.2 The Nonsymmetric Lanczos Algorithm

When the matrix A is not symmetric, the orthonormal vectors $\boldsymbol{v}_i$ cannot be generated by a short recurrence. To remedy this problem, Lanczos in 1950 introduced an algorithm for nonsymmetric matrices. Its aim is to construct two biorthogonal sequences of vectors that form bases for the

Krylov subspaces

$$\mathbb{K}_n(A, \boldsymbol{v}) = \text{span}\left\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{n-1}\boldsymbol{v}\right\},$$
$$\mathbb{K}_n\left(A^T, \boldsymbol{w}\right) = \text{span}\left\{\boldsymbol{w}, A^T\boldsymbol{w}, \ldots, \left(A^{n-1}\right)^T\boldsymbol{w}\right\},$$

where A^T denotes the transpose of A. Fortunately, these bases can be generated by short recurrences. The drawback is that the algorithm may break down. The standard Lanczos algorithm for nonsymmetric matrices requires the use of both the matrix A and of its transpose. It is defined as follows:

Algorithm 2 Nonsymmetric Lanczos algorithm

1: **Input:** Integer n, matrix $A \in \mathbb{R}^{N\times N}$, vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^N$ such that $\boldsymbol{w}^T\boldsymbol{v} = 1$.
2: **Output:** Biorthogonal vectors $\{\boldsymbol{v}_i\}_{i=0}^{n-1}, \{\boldsymbol{w}_i\}_{i=0}^{n-1}$.
3: Initialize: $\boldsymbol{v}_{-1} = \boldsymbol{w}_{-1} = \boldsymbol{0}; \boldsymbol{v}_0 = \boldsymbol{v}; \boldsymbol{w}_0 = \boldsymbol{w}/(\boldsymbol{w}^T\boldsymbol{v})$;
4: **for** $i = 1, 2, \ldots, n-1$ **do**
5: $\quad \alpha_{j-1} := \boldsymbol{w}_{i-1}^T(A\boldsymbol{v}_{i-1} - \gamma_{i-1}\boldsymbol{v}_{i-2})$;
6: $\quad \boldsymbol{r} := A\boldsymbol{v}_{i-1} - \alpha_{i-1}\boldsymbol{v}_{i-1} - \gamma_{i-1}\boldsymbol{v}_{i-2}$;
7: $\quad \boldsymbol{s} := A^T\boldsymbol{w}_{i-1} - \alpha_{i-1}\boldsymbol{w}_{i-1} - \beta_{i-1}\boldsymbol{w}_{i-2}$;
8: $\quad \beta_i := |\boldsymbol{r}^T\boldsymbol{s}|^{1/2}; \gamma_i := \boldsymbol{r}^T\boldsymbol{s}/\beta_i$;
9: $\quad \boldsymbol{v}_i := \boldsymbol{r}/\beta_i; \boldsymbol{w}_i := \boldsymbol{s}/\gamma_i$;
10: **end for**

Algorithm 2 may break down when some coefficient β_i vanishes. Breakdowns are classified into two kinds:

(i) If $\boldsymbol{r}^T\boldsymbol{s} = 0$ because one of the involved vectors is zero, this situation is called *lucky breakdown*. In this case, one may have to restart the algorithm with a pair of vectors that are biorthogonal to the invariant subspace already found, cf. [48].

(ii) If $\boldsymbol{r}^T\boldsymbol{s} = 0$ but $\boldsymbol{r} \neq \boldsymbol{0}$ and $\boldsymbol{s} \neq \boldsymbol{0}$, the algorithm is suffer from a *serious breakdown*. A way to deal with this problem is to apply a look-ahead technique.

The breakdowns of this algorithm have been discussed by many authors; see [32, 33, 34, 71]. For more details of the Lanczos process; see [22, 29, 62].

The symmetric rational Lanczos method reduces a large symmetric matrix A to a small one by both evaluating of matrix-vector products with A and solving linear systems of equations with

shifted matrices $(A - \zeta_j I)$, where ζ_j is a prescribed pole. Typically, these systems are solved by LU (or Cholesky) factorization. Therefore, the symmetric rational Lanczos method is attractive to use when the matrix A has a structure that allows the computation of its LU (or Cholesky) factorization for a reasonable computational cost.

Let $\zeta_i, i = 1, 2, \ldots, n-1$, be poles. A rational Krylov subspace is defined by

$$\mathbb{K}_n(A, \boldsymbol{v}) = \text{span}\{\boldsymbol{v}, (A - \zeta_1 I)^{-1}\boldsymbol{v}, (A - \zeta_2 I)^{-1}(A - \zeta_1 I)^{-1}\boldsymbol{v}, \ldots, \prod_{i=1}^{n-1}(A - \zeta_i I)^{-1}\boldsymbol{v}\}.$$

When all $\zeta_i = \infty$, the rational Krylov space reduces to the standard Krylov subspace (2.20). Rational Krylov subspaces with poles at zero and infinity only, form a special instance called extended Krylov subspaces. These subspaces are built by applying positive and negative powers of A to the vector $\boldsymbol{v}$, and they are related to Laurent polynomials.

Definition 9. *Let j, k be two nonnegative integers. Then the set of Laurent polynomials of numerator degree at most k and denominator degree at most j is defined as*

$$\mathcal{L}_{j,k} := \text{span}\left\{x^{-j}, x^{-j+1}, \ldots, 1, \ldots, x^{k-1}, x^k\right\}, \quad x \in \mathbb{R}\backslash\{0\}.$$

Obviously, if $j = 0$, then $\mathcal{L}_{0,k} = \mathbb{P}_k$.

An example of the extended Krylov subspace is given by:

$$\mathbb{K}_{n_l,n_r}(A, \boldsymbol{v}) = \text{span}\left\{A^{-n_l+1}\boldsymbol{v}, \ldots, A^{-1}\boldsymbol{v}, \boldsymbol{v}, A\boldsymbol{v}, A^2\boldsymbol{v}, \ldots, A^{n_r-1}\boldsymbol{v}\right\}.$$

The dimension of the subspace is equal to $n_r + n_l - 1$. Similarly to polynomial Krylov subspace (2.21), the extended Krylov subspace can be defined by using Laurent polynomials,

$$\mathbb{K}_{n_l,n_r}(A, \boldsymbol{v}) = \{\phi(A)\boldsymbol{v} : \phi \in \mathcal{L}_{n_l-1,n_r-1}\}.$$

These subspaces result in rational approximation to the integral (2.1). Generating an orthonormal basis for an extended Krylov subspace generally is more expensive than computing an orthonormal basis for a standard Krylov. However, rational Krylov subspaces of small dimensions may give higher accuracy than standard Krylov subspaces of the same dimensions. It also might be beneficial to allow the numerator degree of the orthonormal Laurent polynomials to grow faster than the denominator degree, because in the rational Lanczos method each increase in the numerator degree

requires the evaluation of a matrix-vector product with A, while each increase of the denominator degree demands the solution of a linear system of equations with A.

It is shown in [41] that the recursion formulas for this kind of rational Lanczos process can be short. They depend on the measure defined by the symmetric matrix A, the initial vector $\boldsymbol{v}$, and on how frequently the denominator degree is increased.

Also the recursion relations for the rational Lanczos process that is determined by several finite real or complex conjugate poles may be short. The number of terms in these recursion formulas depends on the number of distinct poles and their ordering. Recursion relations for this kind of orthonormal rational functions are introduced in [58].

CHAPTER 3

Generalized Gauss–Radau and Gauss–Lobatto Rules

3.1 Overview

Many functionals of a large symmetric matrix of interest in science and engineering can be expressed as a Stieltjes integral with a measure supported on the real axis. These functionals can be approximated by quadrature rules. Golub and Meurant proposed a technique for computing upper and lower error bounds for Stieltjes integrals with integrands whose derivatives do not change sign on the convex hull of the support of the measure. This technique is based on evaluating pairs of a Gauss quadrature rule and a suitably chosen Gauss–Radau or Gauss–Lobatto quadrature rule. However, when derivatives of the integrand change sign on the convex hull of the support of the measure, this technique is not guaranteed to give upper and lower error bounds for the functional. We describe an extension of the technique by Golub and Meurant that yields upper and lower error bounds for the functional in situations when only some derivatives of the integrand do not change sign on the convex hull of the support of the measure. This extension is based on the use of pairs of Gauss, and suitable generalized Gauss–Radau or Gauss–Lobatto rules. New methods to evaluate generalized Gauss–Radau and Gauss–Lobatto rules also are described.

3.2 Introduction

The main focus of this chapter is to evaluate matrix functionals of the form

$$F(A) := \boldsymbol{v}^T f(A)\boldsymbol{v}, \tag{3.1}$$

where $A \in \mathbb{R}^{N\times N}$ is a large symmetric matrix, $\boldsymbol{v} \in \mathbb{R}^N$, f is a function that is defined on the convex hull of the spectrum of A. For notational simplicity, we will assume that $\|\boldsymbol{v}\| = 1$. Here and throughout this thesis $\|\cdot\|$ denotes the Euclidean vector norm.

Golub and Meurant [28, 29] describe a technique for computing upper and lower error bounds for matrix functionals of the form (3.1) based on the connection between the Lanczos process,

orthogonal polynomials, and Gauss-type quadrature rules. Their technique considers the expression (3.1) as a Stieltjes integral with integrand f. This indicates that Gauss-type quadrature rules can be applied to compute approximations of (3.1). Assuming that derivatives of the integrand f do not change sign in the convex hull of the spectrum of A, Golub and Meurant [28, 29] observed that pairs of Gauss, and suitable Gauss–Radau or Gauss–Lobatto rules, provide upper and lower bounds for (3.1). This follows straightforwardly from the sign of the error formulas (2.9), (2.15), and (2.16).

When derivatives of the integrand f change sign in the convex hull of the spectrum of A, the technique developed by Golub and Meurant [28, 29] is not guaranteed to provide upper and lower error bounds for (3.1).

Example 3.2.1. Let $A \in \mathbb{R}^{200\times 200}$ be the symmetric Toeplitz matrix with first row $[2/3, 2/5, \ldots, 2/401]$. Its largest and smallest eigenvalues are given by $\lambda_{min} = 0.19175$ and $\lambda_{max} = 8.0626$, respectively. Consider the approximation of the functional

$$F(A) := \boldsymbol{v}^T \exp(-\frac{A}{4}) \sin(\frac{A}{4})\boldsymbol{v}. \tag{3.2}$$

and define the integrand

$$f(x) := \exp(-\frac{x}{4}) \sin(\frac{x}{4}). \tag{3.3}$$

Some derivatives of this integrand change sign on the interval $[\lambda_{min}, \lambda_{max}]$. We illustrate in Example 3.5.1 of Section 3.5 that pairs of Gauss and Gauss–Radau rules, or pairs of Gauss and Gauss–Lobatto rules, do not furnish upper and lower error bounds for (3.2).

We are interested in exploring whether the technique of Golub and Meurant can be extended to give upper and lower error bounds for (3.1) also in situations when some derivatives of the integrand f change sign in the convex hull of the spectrum of A. Specifically, we will show that pairs of Gauss rules and suitably chosen generalized Gauss–Radau or generalized Gauss–Lobatto rules give upper and lower error bounds for (3.1) in some situations when pairs of Gauss and (standard) Gauss–Radau or Gauss–Lobatto rules are not guaranteed to furnish upper and lower bounds.

Generalized Gauss–Radau rules are Gauss–Radau-type rules, in which the fixed node has multiplicity larger than one; similarly, generalized Gauss–Lobatto rules are Gauss–Lobatto-type rules, in which at least one of the fixed nodes has multiplicity larger than one. Generalized Gauss–Radau and Gauss–Lobatto rules have received considerable attention; see, e.g., [21, 23, 24, 26, 27, 55, 70].

Applications of these quadrature rules include the computation of spline approximations that reproduce as many consecutive moments of the integrand f as possible; see Gautschi [22, Section 3.3] for details.

This chapter is organized as follows: Section 3.3 reviews generalized Gauss–Radau quadrature rules, and describes a novel way to evaluate the quadrature rules. Generalized Gauss–Lobatto rules are considered in Section 3.4, and a few computed examples are presented in Section 3.5.

We conclude this section by discussing how the matrix functional (3.1) is related to a Stieltjes integral. The development follows Golub and Meurant [28, 29]. Introduce the spectral factorization

$$A = S\Lambda S^T, \qquad \Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_N], \tag{3.4}$$

with the eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$ and $S \in \mathbb{R}^{N\times N}$ an orthogonal matrix, whose columns are eigenvectors. Then we define

$$f(A) = Sf(\Lambda)S^T;$$

see, e.g., [30, 38] for discussions on the definition of matrix functions. Introduce the row vector $[\nu_1, \nu_2, \ldots, \nu_N] := \boldsymbol{v}^T S$. Then the functional (3.1) can be written as

$$F(A) = \boldsymbol{v}^T Sf(\Lambda)S^T\boldsymbol{v} = \sum_{j=1}^{N} f(\lambda_j)\nu_j^2. \tag{3.5}$$

The right-hand side can be expressed as a Stieltjes integral

$$\mathcal{I}(f) := \int_a^b f(x)d\lambda(x), \tag{3.6}$$

where the distribution function λ associated with the measure $d\lambda$ can be chosen to be piece-wise constant and defined by

$$\lambda(x) := \begin{cases} 0, & \text{if } \ x < a = \lambda_1, \\ \sum_{j=1}^{i} \nu_j^2, & \text{if } \ \lambda_i \leq x < \lambda_{i+1}, \quad i = 1, 2, \ldots, N-1, \\ \sum_{j=1}^{N} \nu_j^2, & \text{if } \ b = \lambda_N \leq x. \end{cases}$$

The m-point (standard) Gauss quadrature rule associated with the measure $d\lambda(x)$ is of the form

$$\mathcal{G}_m(f) := \sum_{i=1}^{m} w_i f(x_i),$$

and is characterized by the property that

$$\mathcal{I}(f) = \mathcal{G}_m(f), \qquad \forall f \in \mathbb{P}_{2m-1},$$

where $\mathbb{P}_{2m-1}$ denotes the set of polynomials of degree at most $2m-1$. The nodes x_i of the quadrature rule are distinct and known to be the zeros of an mth degree orthogonal polynomial with respect to the inner product

$$(f, g) := \mathcal{I}(fg). \tag{3.7}$$

We will approximate the integral (3.6), and therefore the functional (3.1), by Gauss-type quadrature rules. Under suitable conditions, the sign of the quadrature error can be inferred from the remainder terms of the quadrature rules used. While our discussion focuses on functionals of the form (3.1), a generalization to functionals $\boldsymbol{w}^T f(A)\boldsymbol{v}$ with $\boldsymbol{w} \in \mathbb{R}^N$ different from $\boldsymbol{v}$ is straightforward by using the identity

$$\boldsymbol{w}^T f(A)\boldsymbol{v} = \frac{1}{4}\left((\boldsymbol{w}+\boldsymbol{v})^T f(A)(\boldsymbol{w}+\boldsymbol{v}) - (\boldsymbol{w}-\boldsymbol{v})^T f(A)(\boldsymbol{w}-\boldsymbol{v})\right).$$

3.3 Generalized Gauss–Radau Formulas

This section considers generalized Gauss–Radau rules of the form

$$\mathcal{G}_{m,r}(f) = \sum_{i=1}^{m} w_i f(x_i) + \sum_{j=0}^{r-1} w_j^{(0)} f^{(j)}(x_0) \tag{3.8}$$

for approximating the integral (3.6), where the x_i, $1 \le i \le m$, are "free" distinct nodes in the open interval (a, b), and x_0 is a prescribed node of multiplicity $r \ge 2$ outside this interval. We assume that f and its required derivatives (see below) are defined in $\mathrm{int}(a, b, x_0)$, the convex hull of the set $\{a, b, x_0\}$ where $-\infty < x_0 \le a$ or $b \le x_0 < \infty$. Many properties of generalized Gauss–Radau rules are discussed in [21, 22, 23, 24, 26, 27, 55, 70]. Here we recall that the nodes $x_1, x_2, \ldots, x_m$ are the zeros of the mth degree orthogonal polynomial with respect to the measure $(x-a)^r\, d\lambda(x)$. The generalized Gauss–Radau quadrature rule satisfies

$$\mathcal{I}(f) = \mathcal{G}_{m,r}(f), \qquad \forall f \in \mathbb{P}_{2m+r-1}; \tag{3.9}$$

see, e.g., [22] for details.

When the integrand f is $2m+r$ times continuously differentiable in $\mathrm{int}(a,b,x_0)$, the error in the generalized Gauss–Radau quadrature rule (3.8) can be expressed as

$$\mathcal{E}_{m,r}(f) := (\mathcal{I}-\mathcal{G}_{m,r})(f) = \frac{f^{(2m+r)}(x_{\mathrm{GR}})}{(2m+r)!}\cdot\int_a^b (x-x_0)^r \prod_{i=1}^{m}(x-x_i)^2 d\lambda(x), \tag{3.10}$$

for some $x_{\mathrm{GR}} \in \mathrm{int}(a,b,x_0)$; see, e.g., [45] for a proof of (3.10). If the derivative $f^{(2m+r)}$ is of known constant sign in $\mathrm{int}(a,b,x_0)$, then we can tell the sign of $\mathcal{E}_{m,r}(f)$. For instance, when $f^{(2m+r)}(x) \leq 0$ for $x \in \mathrm{int}(a,b,x_0)$, and $x_0 = a$, the quadrature rule $\mathcal{G}_{m,r}(f)$ furnishes an upper bound for $\mathcal{I}(f)$.

Gautschi [21, 23, 24] describes several ways of computing the nodes and weights of generalized Gauss–Radau rules (3.8). We will describe a new approach to evaluate these quadrature rules that is convenient to use when the measure is implicitly defined by a sum (3.5). Our approach does not require the explicit evaluation of the nodes and weights.

Application of $m+r$ steps of the symmetric Lanczos process, which is described by Algorithm 1, to the matrix A with initial unit vector $\boldsymbol{v}$ gives the Lanczos decomposition

$$AV_{m+r} = V_{m+r}T_{m+r} + \beta_{m+r}\boldsymbol{v}_{m+r+1}\boldsymbol{e}_{m+r}^T, \tag{3.11}$$

where the matrix $V_{m+r} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m+r}] \in \mathbb{R}^{N\times(m+r)}$ and vector $\boldsymbol{v}_{m+r+1} \in \mathbb{R}^N$ satisfy $\boldsymbol{v}_1 = \boldsymbol{v}$, $V_{m+r}^T V_{m+r} = I_{m+r}$, $\|\boldsymbol{v}_{m+r+1}\| = 1$, and $V_{m+r}^T \boldsymbol{v}_{m+r+1} = \boldsymbol{0}$. Here and throughout this thesis $\boldsymbol{e}_j = [0,\ldots,0,1,0,\ldots,0]^T$ is the jth axis vector of suitable dimension, and I_j stands for the identity matrix of order j. Moreover, $\beta_{m+r} \in \mathbb{R}_+$ and the matrix

$$T_{m+r} := \begin{bmatrix} \alpha_0 & \beta_1 & & & \\ \beta_1 & \alpha_1 & \beta_2 & & \\ & \beta_2 & \alpha_2 & \ddots & \\ & & \ddots & \ddots & \beta_{m+r-1} \\ & & & \beta_{m+r-1} & \alpha_{m+r-1} \end{bmatrix} \in \mathbb{R}^{(m+r)\times(m+r)}.$$

is symmetric and tridiagonal. The Lanczos procedure is a discrete analogue of the Stieltjes procedure in the sense that the former is applied to a matrix and a vector; it requires the support of the measure to be a finite discrete point set. Of course, continuous analogues of the Lanczos procedure can be defined, in which case the matrix A is replaced by a symmetric operator. The Stieltjes procedure is described, e.g., by Gautschi [22] and the (discrete) Lanczos procedure is discussed by

Golub and Meurant [28, 29]. Typically, $1 < m + r \ll N$ in computations. We tacitly assume that $m + r$ is small enough so that the decomposition (3.11) with the stated properties exists. This is the generic situation. In the rare event that the Lanczos process breaks down before $m + r$ steps have been carried out, the computations simplify. We will not dwell on the ramification of breakdown.

The dominant computational effort required for the calculation of the decomposition (3.11) by the Lanczos process is the evaluation of $m + r$ matrix-vector products with the matrix A; see, e.g., [28, 29]. Each matrix-vector product evaluation with A requires $\mathcal{O}(cn)$ arithmetic floating-point operations (flops), where c is the average number of nonvanishing entries of A per row.

The relation (3.11) shows that the columns $\boldsymbol{v}_j$ of V_{m+r} can be expressed as

$$\boldsymbol{v}_j = p_{j-1}(A)\boldsymbol{v}, \qquad j = 1, 2, 3, \dots \ , \tag{3.12}$$

for certain polynomials $p_{j-1} \in \mathbb{P}_{j-1}$.

It follows from the orthonormality of the vectors $\boldsymbol{v}_j$ and (3.12) that

$$\begin{aligned}(p_{j-1}, p_{k-1}) &= \int_a^b p_{j-1}(x)p_{k-1}(x)d\lambda(x) = \boldsymbol{v}^T S p_{j-1}(\Lambda)p_{k-1}(\Lambda)S^T\boldsymbol{v} \\ &= \boldsymbol{v}^T p_{j-1}(A)p_{k-1}(A)\boldsymbol{v} = \boldsymbol{v}_j^T\boldsymbol{v}_k = \begin{cases} 0, & j \neq k, \\ 1, & j = k. \end{cases}\end{aligned}$$

Thus, the polynomials p_j are orthonormal with respect to the inner product (3.7).

The decomposition (3.11) defines a recurrence relation for the columns $\boldsymbol{v}_j$ of V_{m+r}, which, in view of (3.12), gives the following recurrence relation for the polynomials p_j,

$$\begin{aligned}\beta_1 p_1(x) &= (x - \alpha_0)p_0(x), \qquad p_0(x) = 1, \\ \beta_j p_j(x) &= (x - \alpha_{j-1})p_{j-1}(x) - \beta_{j-1}p_{j-2}(x), \qquad 2 \leq j \leq m + r,\end{aligned} \tag{3.13}$$

where

$$\alpha_{j-1} = (p_{j-1}, xp_{j-1}). \qquad j = 1, 2, \dots, m + r,$$

and the $\beta_j > 0$ are determined by the requirements $(p_j, p_j)^{1/2} = 1$ for all j.

Introduce the vector

$$\mathbf{p}(x) = \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{m+r-1}(x) \end{bmatrix}.$$

Then the recurrence relation (3.13) can be written in the form

$$x\,\mathbf{p}(x) = T_{m+r}\mathbf{p}(x) + \beta_{m+r}\,p_{m+r}(x)\boldsymbol{e}_{m+r}, \tag{3.14}$$

which shows that the eigenvalues of T_{m+r} are the zeros of the polynomial p_{m+r}. It can be shown that the $(m+r)$-node (standard) Gauss quadrature rule associated with the measure $d\lambda$ in (3.6) can be expressed as

$$\mathcal{G}_{m+r}(f) = \boldsymbol{e}_1^T f(T_{m+r})\boldsymbol{e}_1. \tag{3.15}$$

Here we have used the fact that the vector $\boldsymbol{v}$ in (3.5) is of unit norm; see [29] for details. Note that the Gauss rule (3.15) can be computed by evaluating the function f of the generally fairly small matrix T_{m+r}, without explicitly calculating the nodes and weights of the Gauss rule. Many algorithms for evaluating functions of a small to moderately-sized matrix are described and analyzed by Higham [38].

We now show how the generalized Gauss–Radau rule (3.8) can be evaluated without explicitly computing its nodes and weights. Let $\pi_0, \pi_1, \pi_2, \ldots$ be orthonormal polynomials with respect to the inner product

$$(f, g)_r = \int_a^b f(x)g(x)(x - x_0)^r d\lambda(x), \tag{3.16}$$

where the measure $d\lambda$ is the same as in (3.6). Thus,

$$(\pi_i, \pi_j)_r = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

and $\pi_i \in \mathbb{P}_i$. Define the polynomial $q_{m+r}(x) = \pi_m(x)(x - x_0)^r$. Then

$$xp_{m+r-1}(x) = \sum_{i=0}^{m+r-1} d_i p_i(x) + s_{m+r}q_{m+r}(x), \tag{3.17}$$

for suitable coefficients d_i and s_{m+r}. The orthonormality of the polynomials p_i with respect to the inner product (3.7) gives, for $i = 0, 1, \ldots, m+r-1$,

$$d_i = \int_a^b xp_{m+r-1}(x)p_i(x)d\lambda(x) - s_{m+r}\int_a^b \pi_m(x)p_i(x)(x - x_0)^r d\lambda(x).$$

Now using the orthogonality of the polynomials p_i with respect to the inner product (3.7) and the orthogonality of the polynomials π_j with respect to the inner product (3.16) shows that $d_i = 0$ for $0 \leq i < m$. It follows that (3.17) simplifies to

$$xp_{m+r-1}(x) = \sum_{i=m}^{m+r-1} d_i p_i(x) + s_{m+r}q_{m+r}(x). \tag{3.18}$$

We obtain analogously to (3.14) the relation

$$x\,\mathbf{p}(x) = T_{m+r}^R\,\mathbf{p}(x) + s_{m+r}q_{m+r}(x)\,\boldsymbol{e}_{m+r}, \tag{3.19}$$

where the matrix $T_{m+r}^R \in \mathbb{R}^{(m+r)\times(m+r)}$ is obtained from T_{m+r} by replacing the last row by the vector

$$[0,\ldots,0,d_m,d_{m+1},\ldots,d_{m+r-1}].$$

It follows from (3.14) that the nodes $x_0,x_1,\ldots,x_m$ of the quadrature rule (3.8) are eigenvalues of T_{m+r}^R, and the vectors $\mathbf{p}(x_i)$, $i=0,1,\ldots,m$, are corresponding eigenvectors. We will show below that the eigenvalue x_0 has algebraic multiplicity r and geometric multiplicity 1.

Let $p_i^{(j)}$ denote the jth derivative of the polynomial p_i. The nontrivial entries of the last row of T_{m+r}^R can be determined by solving the linear system of equations

$$x_0p_{m+r-1}^{(j)}(x_0) + jp_{m+r-1}^{(j-1)}(x_0) = \sum_{i=m}^{m+r-1} d_ip_i^{(j)}(x_0), \quad j=1,\ldots,r-1, \tag{3.20}$$

which is obtained by differentiating (3.18) and using the fact that $q_{m+r}^{(j)}(x_0)=0$ for $j=0,1,\ldots,r-1$.

We next verify that the Gauss–Radau rule (3.8) can be expressed as

$$\mathcal{G}_{m,r}(f) = \boldsymbol{e}_1^T f(T_{m+r}^R)\boldsymbol{e}_1. \tag{3.21}$$

This formula is analogous to (3.15). We show (3.21) by deriving the Jordan decomposition of the matrix T_{m+r}^R; see [61, Section 4] for details for more general situations. Differentiating equation (3.19) j times yields

$$x\,\mathbf{p}^{(j)}(x) + j\,\mathbf{p}^{(j-1)}(x) = T_{m+r}^R\,\mathbf{p}^{(j)}(x) + s_{m+r}\,q_{m+r}^{(j)}(x)\,\boldsymbol{e}_{m+r}, \quad j=1,2,\ldots,r-1,$$

where $\mathbf{p}^{(j)}(x)$ denotes j times component-wise differentiation of $\mathbf{p}(x)$ with respect to x. Dividing the right-hand side and left-hand side by $j!$ and setting $x=x_0$ gives

$$(T_{m+r}^R - x_0I)\,\frac{1}{j!}\,\mathbf{p}^{(j)}(x_0) = \frac{1}{(j-1)!}\mathbf{p}^{(j-1)}(x_0),$$

i.e., $\mathbf{p}^{(j)}(x_0)/(j!)$, $j=1,\ldots,r-1$, are principal (generalized eigen-) vectors of T_{m+r}^R. Introduce the $(m+r)\times(m+r)$ matrix

$$W = [\mathbf{p}(x_1),\ldots,\mathbf{p}(x_m),\mathbf{p}(x_0),\mathbf{p}^{(1)}(x_0),\ldots,\frac{1}{(r-1)!}\mathbf{p}^{(r-1)}(x_0)]. \tag{3.22}$$

We have derived the Jordan factorization

$$T_{m+r}^R W = W\Lambda, \qquad \Lambda = \begin{bmatrix} x_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & x_m & & & & & & \\ & & & x_0 & 1 & & & & \\ & & & & x_0 & 1 & & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & x_0 & 1 & \\ & & & & & & & x_0 & \end{bmatrix}. \tag{3.23}$$

Thus, the matrix $\Lambda \in \mathbb{R}^{(m+r)\times(m+r)}$ has a leading $m \times m$ diagonal block matrix and a trailing $r \times r$ Jordan block associated with the eigenvalue x_0.

Our proof of the representation (3.21) requires explicit formulas for the entries in the first column of W^{-1}. Introduce the matrix $U \in \mathbb{R}^{(m+r)\times(m+r)}$, whose ith row is $\boldsymbol{u}_i^T$, where

$$\begin{aligned} \boldsymbol{u}_i &= w_i \mathbf{p}(x_i), \qquad i = 1, 2, \ldots, m, \\ \boldsymbol{u}_{m+s} &= \sum_{u=s-1}^{r-1} u!\, w_u^{(0)}\, \mathbf{p}^{(u+1-s)}(x_0), \qquad s = 1, 2, \ldots, r. \end{aligned}$$

Denote the ith row of W by $\mathbf{a}_i^T = [a_1, a_2, \ldots, a_{m+r}]$, and the jth column of U by $\mathbf{b}_j = [b_1, b_2, \ldots, b_{m+r}]^T$. We will show that

$$\mathbf{a}_i^T \mathbf{b}_j = \mathcal{G}_{m+r}(p_{i-1} p_{j-1}).$$

Note that

$$a_k = p_{i-1}(x_k), \quad b_k = w_k p_{j-1}(x_k), \quad k = 1, 2, \ldots, m$$

and

$$a_{m+s} = \frac{1}{(s-1)!} p_{i-1}^{(s-1)}(x_0), \quad b_{m+s} = \sum_{u=s-1}^{r-1} u!\, w_u^{(0)}\, \frac{{p_{j-1}}^{(u+1-s)}(x_0)}{(u+1-s)!}, \quad s = 1, 2, \ldots, r.$$

It follows that

$$\begin{aligned}
\sum_{k=1}^{m+r} a_k b_k &= \sum_{k=1}^{m} w_k p_{i-1}(t_k) p_{j-1}(x_k) + \sum_{s=1}^{r} \sum_{u=s-1}^{r-1} p_{i-1}^{(s-1)}(x_0) \binom{u}{s-1} w_u^{(0)}\, p_{j-1}{}^{(u+1-s)}(x_0) \\
&= \sum_{k=1}^{m} w_k p_{i-1}(x_k) p_{j-1}(x_k) + \sum_{\ell=0}^{r-1} w_\ell^{(0)} \sum_{q=0}^{\ell} \binom{\ell}{q} p_{i-1}^{(q)}(x_0) p_{j-1}^{(\ell-q)}(x_0) \\
&= \sum_{k=1}^{m} w_k (p_{i-1}p_{j-1})(x_k) + \sum_{\ell=0}^{r-1} w_\ell^{(0)} (p_{i-1}p_{j-1})^{(\ell)}(x_0) \\
&= \mathcal{G}_{m,r}(p_{j-1}p_{i-1}). \qquad (3.24)
\end{aligned}$$

In view of (3.9), we have for $i + j - 2 \leq 2m + r - 1$ that

$$\mathcal{G}_{m,r}(p_{j-1}p_{i-1}) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

It now follows from (3.24) that the first $m + 1$ columns of the matrix U are the first $m + 1$ columns of W^{-1}. In particular,

$$W^{-1}\boldsymbol{e}_1 = [w_1, w_2, \ldots, w_m, w_0^{(0)}, \ldots, (r-1)!w_{r-1}^{(0)}]^T. \qquad (3.25)$$

We obtain from (3.22) that

$$W^T\boldsymbol{e}_1 = [1, 1, \ldots, 1, 0, \ldots, 0]^T. \qquad (3.26)$$

Finally, equations (3.23), (3.25), and (3.26) give

$$\boldsymbol{e}_1^T f(T_{m+r}^R)\boldsymbol{e}_1 = \boldsymbol{e}_1^T W f(\Lambda) W^{-1}\boldsymbol{e}_1 = \mathcal{G}_{m,r}(f),$$

which shows (3.21).

We conclude that the generalized Gauss–Radau rule $\mathcal{G}_{m,r}(f)$ can be evaluated by using either (3.8) or (3.21). Which one of these expressions is most convenient to compute depends on whether software for computing the integrand f at the small matrix T_{m+r}^R is available or easily can be written. This is the case, for instance, for the exponential function, logarithm, square root, and rational expressions. When the form (3.8) is used, the nodes and weights have to be evaluated. This can be done with software written by Gautschi [25].

We note that if the moments μ_i defined by (2.3) are explicitly known, which is the case for many classical positive measures on the real line, the modified moments

$$\nu_i = \int_a^b x^i (x - x_0)^r d\lambda(x), \qquad i = 0, 1, 2, \ldots \ ,$$

can be easily computed. Then the coefficients d_i in (3.20) can be evaluated without solving linear systems of equations.

Taking $f(x) = (x - x_0)^r P(x)$, where $P(x) \in \mathbb{P}_{2m-1}$, in (3.8) we verify that the first sum on the right-hand side in (3.8) is actually the (standard) m-point Gauss quadrature for the integral

$$\tilde{\mathcal{I}}(f) = \int_a^b f(x) d\tilde{\lambda}(x), \quad d\tilde{\lambda}(x) = (x - x_0)^r d\lambda(x).$$

Thus the quadrature $\mathcal{G}_{m,r} f$ can be written in the form

$$\mathcal{G}_{m,r}(f) = \tilde{m}_0 \boldsymbol{e}_1^T f(\tilde{T}_m) \boldsymbol{e}_1 + \sum_{j=0}^{r-1} w_j^{(0)} f^{(j)}(x_0),$$

where $\tilde{T}_m$ is the Jacobi matrix of dimension $m \times m$ associated with the modified positive measure $d\tilde{\lambda}(x)$, and $\tilde{m}_0 = \int_a^b d\tilde{\lambda}(x)$. This formula can be used for the computation of the generalized Gauss–Radau quadrature when the measure $d\lambda(x)$ is explicitly known, but not in the case when $d\lambda(x)$ is implicitly defined by the matrix A and the vector $\boldsymbol{v}$.

3.4 Generalized Gauss–Lobatto Formulas

This section discusses the application and computation of generalized Gauss–Lobatto rules

$$\mathcal{G}_{m,r,s}(f) = \sum_{j=0}^{r-1} w_j^{(0,1)} f^{(j)}(x_{0,1}) + \sum_{i=1}^{m} w_i f(x_i) + \sum_{j=0}^{s-1} w_j^{(0,2)} f^{(j)}(x_{0,2}) \tag{3.27}$$

for the approximation of the functional (3.1) or, equivalently, of the Stieltjes integral (3.6). Here the x_i, $1 \le i \le m$, are "free" distinct nodes in the open interval (a, b), $-\infty < x_{0,1} \le a$ is a prescribed node of multiplicity $r \ge 1$, and $b \le x_{0,2} < \infty$ is a prescribed node of multiplicity $s \ge 1$. We assume that $\max\{r, s\} \ge 2$ to avoid discussing "standard" Gauss–Lobatto rules. The nodes $x_1, x_2, \ldots, x_m$ are the zeros of the mth degree orthogonal polynomial π_m with respect to the modified measure

$$(x - x_{0,1})^r \, (x_{0,2} - x)^s \, d\lambda(x).$$

Many properties of generalized Gauss–Lobatto rules are discussed in [21, 22, 23, 24, 26, 27, 45, 55]. For instance, it is shown that

$$\mathcal{I}(f) = \mathcal{G}_{m,r,s}(f), \qquad \forall f \in \mathbb{P}_{2m+r+s-1}. \tag{3.28}$$

Moreover, let the integrand f be $2m + r + s$ times continuously differentiable in $\mathrm{int}(a, b, x_{0,1}, x_{0,2})$, the convex hull of the set $\{a, b, x_{0,1}, x_{0,2}\}$. Then analogously to (3.10), the error in the quadrature

rule (3.27) can be expressed as

$$\begin{aligned}\mathcal{E}_{m,r,s}(f) &:= (\mathcal{I} - \mathcal{G}_{m,r,s})(f) \\ &= \frac{f^{(2m+r+s)}(x_{\mathrm{GL}})}{(2m+r+s)!} \cdot \int_a^b (x - x_{0,1})^r (x - x_{0,2})^s \prod_{i=1}^m (x - x_i)^2 d\lambda(x),\end{aligned}$$

where $x_{\mathrm{GL}} \in \mathrm{int}(a, b, x_{0,1}, x_{0,2})$. If $f^{(2m+r+s)}$ is of constant sign in $\mathrm{int}(a, b, x_{0,1}, x_{0,2})$, then the sign of $\mathcal{E}_{m,r,s}(f)$ can be determined by choosing suitable multiplicities r and s.

We derive a formula analogous to (3.21) for the evaluation of $\mathcal{G}_{m,r,s}(f)$. Our derivation is similar to the one for (3.21). We therefore only provide an outline. Application of $m + r + s$ steps of the symmetric Lanczos process to the matrix A with initial unit vector $\boldsymbol{v}$ gives the Lanczos decomposition

$$AV_{m+r+s} = V_{m+r+s} T_{m+r+s} + \beta_{m+r+s} \boldsymbol{v}_{m+r+s+1} \boldsymbol{e}_{m+r+s}^T. \tag{3.29}$$

This decomposition is analogous to (3.11). Here we only note for future reference that the $(m + r + s) \times (m + r + s)$ matrix

$$T_{m+r+s} := \begin{bmatrix} \alpha_0 & \beta_1 & & & & \\ \beta_1 & \alpha_1 & \beta_2 & & & \\ & \beta_2 & \alpha_2 & \ddots & & \\ & & & & & \\ & & \ddots & \ddots & \beta_{m+r+s-1} & \\ & & & \beta_{m+r+s-1} & \alpha_{m+r+s-1} & \end{bmatrix}$$

is symmetric and tridiagonal; we assume that $m + r + s$ is small enough so that the decomposition (3.29) exists. Using (3.12) and defining

$$\mathbf{p}(x) = \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{m+r+s-1}(x) \end{bmatrix},$$

we can express (3.29) in the form

$$x\, \mathbf{p}(x) = T_{m+r+s}\, \mathbf{p}(x) + \beta_{m+r+s}\, p_{m+r+s}(x)\, \mathbf{e}_{m+r+s}.$$

Introduce the inner product

$$(f,g)_{r,s} = \int_a^b f(x)g(x)(x-x_{0,1})^r(x_{0,2}-x)^s d\lambda(x),$$

and let the polynomials $\pi_0, \pi_1, \pi_2, \ldots$ be orthonormal polynomials with respect to this inner product, i.e.,

$$(\pi_i, \pi_j)_{r,s} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

and $\pi_i \in \mathbb{P}_i$. Define the polynomial

$$q_{m+r+s}(x) = \pi_m(x)(x-x_{0,1})^r(x_{0,2}-x)^s.$$

Then

$$xp_{m+r+s-1}(x) = \sum_{i=0}^{m+r+s-1} d_i p_i(x) + \tau_{m+r+s} q_{m+r+s}(x), \tag{3.30}$$

for suitable coefficients d_i and τ_{m+r+s}. Using the orthogonality property of the p_i, we obtain

$$d_i = \int_a^b xp_{m+r+s-1}(x)p_i(x)d\lambda(x) - \tau_{m+r+s}\int_a^b \pi_m(x)p_i(x)(x-x_{0,1})^r(x_{0,2}-x)^s d\lambda(x),$$

for $i = 0, 1, \ldots, m+r+s-1$. Using the orthonormality properties of the polynomials p_i and π_m gives that $d_i = 0$ for $0 \leq i < m$. Thus, the relation (3.30) simplifies to

$$xp_{m+r+s-1}(x) = \sum_{i=m}^{m+r+s-1} d_i p_i(x) + \tau_{m+r+s} q_{m+r+s}(x). \tag{3.31}$$

The coefficients $d_m, d_{m+1}, \ldots, d_{m+r-1}$ can be determined by solving the linear system of equations

$$x_{0,1}p^{(j)}_{m+r+s-1}(x_{0,1}) + jp^{(j-1)}_{m+r+s-1}(x_{0,1}) = \sum_{i=m}^{m+r-1} d_i p_i^{(j)}(x_{0,1}), \quad j = 1, \ldots, r-1, \tag{3.32}$$

and the coefficients $d_{m+r}, d_{m+r+1}, \ldots, d_{m+r+s-1}$ are similarly obtained by solving the linear system of equations

$$x_{0,2}p^{(j)}_{m+r+s-1}(x_{0,2}) + jp^{(j-1)}_{m+r+s-1}(x_{0,2}) = \sum_{i=m+r}^{m+r+s-1} d_i p_i^{(j)}(x_{0,2}), \quad j = 1, \ldots, s-1. \tag{3.33}$$

We remark that the systems (3.32) and (3.33) are obtained from (3.31) by using the fact that $q^{(j)}_{m+r+s}(x_{0,1}) = 0$ for $j = 0, 1, \ldots, r-1$, and $q^{(j)}_{m+r+s}(x_{0,2}) = 0$ for $j = 0, 1, \ldots, s-1$.

Let the matrix $T^L_{m+r+s} \in \mathbb{R}^{(m+r+s)\times(m+r+s)}$ be determined from T_{m+r+s} by replacing the last row by

$$[0, \ldots, 0, d_m, d_{m+1}, \ldots, d_{m+r+s-1}].$$

This gives the relation

$$x\,\mathbf{p}(x) = T^L_{m+r+s}\,\mathbf{p}(x) + \tau_{m+r+s}\,q_{m+r+s}(x)\,\mathbf{e}_{m+r}. \tag{3.34}$$

It follows from this expression that the nodes $x_1, x_2, \ldots, x_m, x_{0,1}, x_{0,2}$ in the quadrature rule (3.27) are eigenvalues of T^L_{m+r+s}, and that $\mathbf{p}(x_1), \mathbf{p}(x_2), \ldots, \mathbf{p}(x_m), \mathbf{p}(x_{0,1}), \mathbf{p}(x_{0,2})$ are corresponding eigenvectors. Differentiation of (3.34) gives

$$x\,\mathbf{p}^{(j)}(x) + j\,\mathbf{p}^{(j-1)}(x) = T^L_{m+r}\,\mathbf{p}^{(j)}(x) + \tau_{m+r+s}\,q^{(j)}_{m+r}(x)\,\mathbf{e}_{m+r}, \quad j = 1, 2, \ldots, r-1.$$

Dividing the above equation by $j!$ and setting $x = x_{0,1}$ gives

$$(T^L_{m+r+s} - x_{0,1}I)\,\frac{1}{j!}\,\mathbf{p}^{(j)}(x_{0,1}) = \frac{1}{(j-1)!}\mathbf{p}^{(j-1)}(x_{0,1}), \quad j = 1, 2, \ldots, r-1.$$

Similarly, differentiating (3.34) component-wise and setting $x = x_{0,2}$ yields

$$(T^L_{m+r+s} - x_{0,2}I)\,\frac{1}{j!}\,\mathbf{p}^{(j)}(x_{0,2}) = \frac{1}{(j-1)!}\mathbf{p}^{(j-1)}(x_{0,2}), \quad j = 1, 2, \ldots, s-1.$$

Hence, $\mathbf{p}^{(j)}(x_{0,1})/(j!)$, $1 \le j < r$, and $\mathbf{p}^{(j)}(x_{0,2})/(j!)$, $1 \le j < s$, are principal vectors of T^L_{m+r+s} associated with the eigenvalues $x_{0,1}$ and $x_{0,2}$, respectively.

We are in a position to discuss the Jordan decomposition of T_{m+r+s}. Define the matrix

$$\begin{aligned} W = \Big[&\mathbf{p}(x_1), \ldots, \mathbf{p}(x_m), \mathbf{p}(x_{0,1}), \mathbf{p}^{(1)}(x_{0,1}), \ldots, \frac{1}{(r-1)!}\mathbf{p}^{(r-1)}(x_{0,1}), \\ &\mathbf{p}(x_{0,2}), \mathbf{p}^{(1)}(x_{0,2}), \ldots, \frac{1}{(s-1)!}\mathbf{p}^{(s-1)}(x_{0,2})\Big], \end{aligned} \tag{3.35}$$

and let

$$\Lambda = \begin{bmatrix} x_1 & & & & & & & & & & & & & & \\ & \ddots & & & & & & & & & & & & & \\ & & x_m & & & & & & & & & & & & \\ & & & x_{0,1} & 1 & & & & & & & & & & \\ & & & & x_{0,1} & 1 & & & & & & & & & \\ & & & & & \ddots & \ddots & & & & & & & & \\ & & & & & & x_{0,1} & 1 & & & & & & & \\ & & & & & & & x_{0,1} & 0 & & & & & & \\ & & & & & & & & x_{0,2} & 1 & & & & & \\ & & & & & & & & & x_{0,2} & 1 & & & & \\ & & & & & & & & & & \ddots & \ddots & & & \\ & & & & & & & & & & & x_{0,2} & 1 & & \\ & & & & & & & & & & & & x_{0,2} & & \end{bmatrix}.$$

Thus, the matrix $\Lambda \in \mathbb{R}^{(m+r+s)\times(m+r+s)}$ is bidiagonal with a leading $m \times m$ principal diagonal matrix, which is followed by a Jordan block associated with the eigenvalue $x_{0,1}$ of order r, and another Jordan block associated with the eigenvalue $x_{0,2}$ of order s. We have the Jordan factorization

$$T^L_{m+r+s}\, W = W\, \Lambda.$$

Similarly as at the end of Section 3.3, we need the first row of the matrix W and the first column of W^{-1} to define an expression for the quadrature rule (3.27) that does not require explicit knowledge of the nodes and weights. It follows from (3.35) that the first row of W is of the form

$$[1, \ldots, 1, 0, \ldots, 0, 1, 0, \ldots, 0];$$

the ones are in the positions where there is no derivative. To determine the first column of W^{-1}, we define the matrix U, whose rows are $\boldsymbol{u}_i^T$, $i = 1, 2, \ldots, m + r + s$, are defined as follows:

$$\begin{aligned}
\boldsymbol{u}_k &= w_k \mathbf{p}(x_k), \quad k = 1, 2, \ldots, m, \\
\boldsymbol{u}_{m+k} &= \sum_{u=k-1}^{r-1} u!\, w_u^{(0,1)}\, \frac{\mathbf{p}^{(u+1-k)}(x_{0,1})}{(u+1-k)!}, \quad k = 1, 2, \ldots, r, \\
\boldsymbol{u}_{m+r+k} &= \sum_{u=k-1}^{r-1} u!\, w_u^{(0,2)}\, \frac{\mathbf{p}^{(u+1-k)}(x_{0,2})}{(u+1-k)!}, \quad k = 1, 2, \ldots, s.
\end{aligned}$$

Denote the ith row of W by $\mathbf{a}_i^T = [a_1, \ldots, a_{m+r+s}]$, and the jth column of V by $\mathbf{b}_j = [b_1, \ldots, b_{m+r+s}]^T$. We will show that

$$\mathbf{a}_i^T \mathbf{b}_j = \mathcal{G}_{m,r,s}(p_{i-1} p_{j-1}).$$

Note that, in view of (3.28),

$$\mathcal{G}_{m,r,s}(p_{i-1} p_{j-1}) = \int_a^b p_{i-1}(x) p_{j-1}(x) d\lambda(x) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

for $i+j-2 \le 2m+r+s-1$. We have

$$\begin{aligned}
a_k &= p_{i-1}(x_k), \quad b_k = w_k p_{j-1}(x_k), \quad k = 1, 2, \ldots, m, \\
a_{m+k} &= \frac{1}{(k-1)!} p_{i-1}^{(k-1)}(x_{0,1}), \quad k = 1, 2, \ldots, r, \\
b_{m+k} &= \sum_{u=k-1}^{r-1} u!\, w_u^{(0,1)}\, \frac{p_{j-1}{}^{(u+1-k)}(x_{0,1})}{(u+1-k)!}, \quad k = 1, 2, \ldots, r, \\
a_{m+r+k} &= \frac{1}{(k-1)!} p_{i-1}^{(k-1)}(x_{0,2}), \quad k = 1, 2, \ldots, s, \\
b_{m+r+k} &= \sum_{u=k-1}^{s-1} u!\, w_u^{(0,2)}\, \frac{p_{j-1}{}^{(u+1-k)}(x_{0,2})}{(u+1-k)!}, \quad k = 1, 2, \ldots, s.
\end{aligned}$$

After some computations similar to those at the end of Section 3.3, we obtain

$$\begin{aligned}
\sum_{k=1}^{m+r+s} a_k b_k = &\sum_{k=1}^{m} w_k (p_{i-1}p_{j-1})(x_k) + \sum_{\ell=0}^{r-1} w_\ell^{(0,1)} (p_{i-1}p_{j-1})^{(\ell)}(x_{0,1}) \\
&+ \sum_{h=0}^{s-1} w_h^{(0,2)} (p_{i-1}p_{j-1})^{(h)}(x_{0,2}) = \mathcal{G}_{m+r+s}(p_{j-1}p_{i-1}).
\end{aligned}$$

It follows similarly as in Section 3.3 that

$$W^{-1}\mathbf{e}_1 = \left[w_1, w_2, \ldots, w_m, w_0^{(0,1)}, \ldots, (r-1)! w_{r-1}^{(0,1)}, w_0^{(0,2)}, \ldots, (s-1)! w_{s-1}^{(0,2)}\right]^T.$$

We finally obtain the desired representation of the quadrature rule,

$$\mathbf{e}_1^T f(T_{m+r+s}^L)\mathbf{e}_1 = \mathbf{e}_1^T W f(\Lambda) W^{-1}\mathbf{e}_1 = \mathcal{G}_{m,r,s}(f). \tag{3.36}$$

Similarly as at the end of Section 3.3, we conclude that the generalized Gauss–Lobatto rule $\mathcal{G}_{m,r,s}(f)$ can be evaluated by using either (3.27) or (3.36). Which one of these expressions is most conveinet to use depends on the integrand.

3.5 Computed Examples

In this section, we present three examples to illustrate the performance of the generalized Gauss–Radau and generalized Gauss–Lobatto quadrature rules. The examples show pairs of a Gauss rule and a generalized Gauss–Radau or generalized Gauss–Lobatto rule to provide upper and lower error bounds for the expression (3.1) in situations when pairs of Gauss rules and standard Gauss–Radau or standard Gauss–Lobatto rules do not.

Example 3.5.1. This example continues the discussion of Example 3.2.1. Thus, we would like to determine an approximation of the functional (3.2) with the matrix A and vector $\boldsymbol{v}$ defined as in Example 3.2.1. The exact value is $F(A) \approx 0.1183$.

We first consider the approximation of (3.2) by pairs of a Gauss rule and a standard or generalized Gauss–Radau rule with a fixed node $x_0 = \lambda_{min}$, and by pairs of a Gauss rule and standard or generalized Gauss–Lobatto rule with fixed nodes $x_{0,1} = \lambda_{min}$ and $x_{0,2} = \lambda_{max}$. We observe that the derivatives $f^{(2m+r)}$ and $f^{(2m+r+s)}$ of the integrand (3.3) change sign on the interval $[\lambda_{min}, \lambda_{max}]$ when $m = 2k$ and $r = s = 1$. This implies that pairs of the Gauss rule $\mathcal{G}_m(f)$ and the standard Gauss–Radau rule $\mathcal{G}_{m,1}(f)$ defined by (2.13), or pairs of the Gauss rule $\mathcal{G}_m(f)$ and the standard Gauss–Lobatto rule $\mathcal{G}_{m,1,1}(f)$ defined by (2.14), are not guaranteed to bracket the value $F(A)$. Indeed, for $m = 2$ we have $F(A) - \mathcal{G}_{m,1}(f) = -2.991 \cdot 10^{-5}$ and $F(A) - \mathcal{G}_{m,1,1}(f) = -1.021 \cdot 10^{-6}$. Table 1 shows that $F(A) - \mathcal{G}_m(f)$ also is negative. Thus, the value $F(A)$ is not bracketed by $\mathcal{G}_{m,1}(f)$ and $\mathcal{G}_{m,1,1}(f)$. We conclude that the technique described in [28, 29] for bounding $F(A)$ based on evaluating pairs of Gauss and (standard) Gauss–Radau or Gauss–Lobatto quadrature rules fails to yield upper and lower bounds for the expression (3.2). These quadrature rules therefore are not useful for assessing the errors in $\mathcal{G}_{m,1}(f)$ or $\mathcal{G}_{m,1,1}(f)$.

The derivatives $f^{(4\ell)}$ in (2.9), when ℓ is odd, are of negative sign in the interval $[\lambda_{min}, \lambda_{max}]$. This yields errors of negative sign and therefore the quadrature rule provides an upper bound for $\mathcal{I}(f)$. In addition, the derivatives $f^{(4\ell)}$, when ℓ is even, are of positive sign in the interval $[\lambda_{min}, \lambda_{max}]$. In this case, we have a positive error and the quadrature rule furnishes a lower bound for $\mathcal{I}(f)$. However, note that the derivatives $f^{(4\ell+1)}$ and $f^{(4\ell+2)}$ change sign in the interval $[\lambda_{min}, \lambda_{max}]$. Therefore, pairs of a Gauss rule and a (standard) Gauss–Radau or Gauss–Lobatto rule are not guaranteed to give upper and lower bounds for (3.2).

The above discussion suggests that pairs of suitable Gauss and generalized Gauss–Radau or generalized Gauss–Lobatto rules may be used to bracket (3.2). Let $r = 4$ and $x_0 = \lambda_{min}$ for the generalized Gauss–Radau rules $\mathcal{G}_{m,r}(f)$, and let $r = 2$, $s = 2$, $x_{0,1} = \lambda_{min}$, and $x_{0,2} = \lambda_{max}$, for the generalized Gauss–Lobatto rules $\mathcal{G}_{m,r,s}(f)$. Then pairs of Gauss rules and these generalized Gauss–Radau or generalized Gauss–Lobatto rules bracket (3.2). This is illustrated by Table 1.

Errors	$m = 2$	$m = 4$	$m = 6$
$F(A) - \mathcal{G}_m f$	$-1.900 \cdot 10^{-3}$	$2.112 \cdot 10^{-7}$	$-1.653 \cdot 10^{-13}$
$F(A) - \mathcal{G}_{m,r} f$	$3.312 \cdot 10^{-5}$	$-1.143 \cdot 10^{-10}$	$2.636 \cdot 10^{-16}$
$F(A) - \mathcal{G}_{m,r,s} f$	$1.050 \cdot 10^{-6}$	$-4.096 \cdot 10^{-11}$	$5.134 \cdot 10^{-16}$

Table 1: Example 3.5.1: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \exp(-\frac{A}{4}) \sin(\frac{A}{4})\boldsymbol{v}$, A a symmetric Toeplitz matrix, $r = 4$ in $\mathcal{G}_{m,r}$, and $r = s = 2$ in $\mathcal{G}_{m,r,s}$.

Example 3.5.2. We consider the approximation of the functional

$$F(A) := \boldsymbol{v}^T \exp(A)(\cos(A) - \sin(A))\boldsymbol{v}, \tag{3.37}$$

where $A = \frac{1}{6}(B + \frac{3\pi}{7} I) \in \mathbb{R}^{200\times 200}$ with $B \in \mathbb{R}^{200\times 200}$ a symmetric Toeplitz matrix with first row $[1, 1/2, \ldots, 1/200]$. The vector $\boldsymbol{v}$ has normally distributed entries with zero mean and is normalized to be of unit norm. The exact value is $F(A) \approx 0.7343$. In this example, the extreme eigenvalues of A are $\lambda_{min} = 0.28878$ and $\lambda_{max} = 1.7141$. Consider the integrand

$$f(x) := \exp(x)(\cos(x) - \sin(x)).$$

We compute approximations of (3.37) by pairs of Gauss rules and standard or generalized Gauss–Radau rules with a fixed node $x_0 = \lambda_{min}$, and by pairs of Gauss rules and standard or generalized Gauss–Lobatto rules with fixed nodes $x_{0,1} = \lambda_{min}$ and $x_{0,2} = \lambda_{max}$.

The derivatives $f^{(2m+r)}$ and $f^{(2m+r+s)}$ of the integrand change sign on the interval $[\lambda_{min}, \lambda_{max}]$ when $m = 2k+1$ and $r = s = 1$. This indicates that pairs of Gauss rules $\mathcal{G}_m(f)$ and standard Gauss–Radau rules $\mathcal{G}_{m,1}(f)$, or pairs of Gauss rules $\mathcal{G}_m(f)$ and standard Gauss–Lobatto rules $\mathcal{G}_{m,1,1}(f)$, are not guaranteed to bracket (3.37). For instance, we find for $m = 5$ that $F(A) - \mathcal{G}_{m,1}(f) = -6.452 \cdot 10^{-13}$ and $F(A) - \mathcal{G}_{m,1,1}(f) = -6.246 \cdot 10^{-11}$. Comparison with results of Tables 2 shows that the pairs of rules $\{\mathcal{G}_m(f), \mathcal{G}_{m,1}(f)\}$ and $\{\mathcal{G}_m(f), \mathcal{G}_{m,1,1}(f)\}$ do not bracket the value (3.37).

Note that the derivatives $f^{(4\ell+2)}$ are positive in the interval $[\lambda_{min}, \lambda_{max}]$ when ℓ is odd. This shows that the errors are positive, and then the quadrature rule yields a lower bound for $\mathcal{I}(f)$. Moreover, the derivatives $f^{(4\ell+2)}$ are negative in the interval $[\lambda_{min}, \lambda_{max}]$ when ℓ is even. Hence, we have negative errors and the quadrature rule yields an upper bound for $\mathcal{I}(f)$. We therefore can determine upper and lower bounds for (3.37) by suitable pairs of Gauss and generalized Gauss–Radau

or generalized Gauss–Lobatto rules. Let $r = 4$ and $x_0 = \lambda_{min}$ for the generalized Gauss–Radau rules $\mathcal{G}_{m,r}(f)$, and let $r = 2$, $s = 2$, $x_{0,1} = \lambda_{min}$, and $x_{0,2} = \lambda_{max}$ for the generalized Gauss–Lobatto rules $\mathcal{G}_{m,r,s}(f)$. Table 2 shows pairs of Gauss rules and these generalized Gauss–Radau or generalized Gauss–Lobatto rules to bracket (3.37).

Errors	$m = 3$	$m = 5$
$F(A) - \mathcal{G}_m f$	$3.862 \cdot 10^{-5}$	$-1.331 \cdot 10^{-10}$
$F(A) - \mathcal{G}_{m,r} f$	$-1.735 \cdot 10^{-8}$	$1.054 \cdot 10^{-14}$
$F(A) - \mathcal{G}_{m,r,s} f$	$-3.993 \cdot 10^{-9}$	$4.662 \cdot 10^{-15}$

Table 2: Example 3.5.2: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \exp(A)(\cos(A) - \sin(A))\boldsymbol{v}$, A a symmetric Toeplitz matrix, $r = 4$ in $\mathcal{G}_{m,r}$, and $r = s = 2$ in $\mathcal{G}_{m,r,s}$.

Example 3.5.3. We would like to compute an approximation of the functional

$$F(A) := \boldsymbol{v}^T \exp(-\frac{A}{7}) \cos(\frac{A}{7})\boldsymbol{v}, \tag{3.38}$$

where $A \in \mathbb{R}^{2114\times 2114}$ is the symmetric adjacency matrix for the Yeast network; see [44, 67]. This matrix is available at [6]. We let the vector $\boldsymbol{v}$ have normally distributed entries with zero mean and to be of unit norm. The extreme eigenvalues of A are $\lambda_{min} = -7.5159$ and $\lambda_{max} = 7.5412$. Introduce the integrand

$$f(x) := \exp(-\frac{x}{7}) \cos(\frac{x}{7}). \tag{3.39}$$

We consider the approximation of (3.38) by pairs of Gauss rules and standard or generalized Gauss–Radau rules with a fixed node $x_0 = \lambda_{min}$, and by pairs of Gauss rules and standard or generalized Gauss–Lobatto rules with fixed nodes $x_{0,1} = \lambda_{min}$ and $x_{0,2} = \lambda_{max}$. We observe that the derivatives $f^{(2m+r)}$ and $f^{(2m+r+s)}$ of the integrand (3.39) change sign on the interval $[\lambda_{min}, \lambda_{max}]$ when $m = 2k$ and $r = s = 1$. Therefore, pairs of Gauss rules $\mathcal{G}_m(f)$ and standard Gauss–Radau rules $\mathcal{G}_{m,1}(f)$, or pairs of Gauss rules $\mathcal{G}_m(f)$ and standard Gauss–Lobatto rules $\mathcal{G}_{m,1,1}(f)$, are not guaranteed to bracket the value (3.38). For instance, we obtain for $m = 4$ that $F(A) - \mathcal{G}_{m,1}(f) = 2.672 \cdot 10^{-8}$ and $F(A) - \mathcal{G}_{m,1,1}(f) = 1.499 \cdot 10^{-6}$. Comparison with results of Table 3 shows that the pairs of rules $\{\mathcal{G}_m(f), \mathcal{G}_{m,1}(f)\}$ and $\{\mathcal{G}_m(f), \mathcal{G}_{m,1,1}(f)\}$ do not bracket (3.38).

However, note that the derivatives $f^{(4\ell)}(x)$ are of a negative sign in the interval $[\lambda_{min}, \lambda_{max}]$ when ℓ is odd, and of a positive sign when ℓ is even. This observation allows us to compute upper and lower bounds for (3.38) by suitable pairs of Gauss and generalized Gauss–Radau or generalized Gauss–Lobatto rules. Let $r = 4$ and $x_0 = \lambda_{min}$ for the generalized Gauss–Radau rules $\mathcal{G}_{m,r}(f)$, and let $r = 2$, $s = 2$, $x_{0,1} = \lambda_{min}$, and $x_{0,2} = \lambda_{max}$ for the generalized Gauss–Lobatto rules $\mathcal{G}_{m,r,s}(f)$. Table 3 and shows that pairs of Gauss rules and these generalized Gauss–Radau or generalized Gauss–Lobatto rules bracket (3.38).

Errors	$m = 2$	$m = 4$	$m = 6$
$F(A) - \mathcal{G}_m f$	$-1.600 \cdot 10^{-3}$	$2.555 \cdot 10^{-7}$	$-1.083 \cdot 10^{-11}$
$F(A) - \mathcal{G}_{m,r} f$	$2.621 \cdot 10^{-5}$	$-7.180 \cdot 10^{-10}$	$6.106 \cdot 10^{-15}$
$F(A) - \mathcal{G}_{m,r,s} f$	$2.310 \cdot 10^{-6}$	$-4.266 \cdot 10^{-11}$	$2.220 \cdot 10^{-16}$

Table 3: Example 3.5.3: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \exp(\frac{-A}{7}) \cos(\frac{A}{7})\boldsymbol{v}$, A a symmetric adjacency matrix for the Yeast network, $r = 4$ in $\mathcal{G}_{m,r}$, and $r = s = 2$ in $\mathcal{G}_{m,r,s}$.

CHAPTER 4

Gauss–Laurent-Type Quadrature Rules

4.1 Overview

In this chapter, we derive Gauss–Laurent quadrature rules for the approximation of matrix functionals of the form $\boldsymbol{w}^T f(A)\boldsymbol{v}$, and also develop associated anti-Gauss–Laurent quadrature rules that allow us to estimate the quadrature error of the Gauss–Laurent rule. Computed examples illustrate the performance of the quadrature rules described.

4.2 Introduction

We are concerned with the approximation of matrix functionals of the form

$$F(A) := \boldsymbol{w}^T f(A)\boldsymbol{v} \tag{4.1}$$

by quadrature rules. Here $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^N$ with $\boldsymbol{v}^T\boldsymbol{w} = 1$, and $A \in \mathbb{R}^{N\times N}$ is a large nonsingular matrix, which may be nonsymmetric.

Assume for the moment that the matrix A has the spectral factorization

$$A = S\Lambda S^{-1}, \tag{4.2}$$

where $S \in \mathbb{C}^{N\times N}$ is nonsingular and $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_N] \in \mathbb{C}^{N\times N}$. We remark that the computation of the quadrature rules does not require this factorization. Substituting (4.2) into (4.1) gives

$$F(A) = \boldsymbol{w}^T S f(\Lambda) S^{-1}\boldsymbol{v} = \sum_{j=1}^{N} f(\lambda_j)\nu_j\nu_j', \tag{4.3}$$

where $[\nu_1, \nu_2, \ldots, \nu_N] := \boldsymbol{w}^T S$ and $[\nu_1', \nu_2', \ldots, \nu_N'] := (S^{-1}\boldsymbol{v})^T$. The right-hand side of (4.3) can be expressed as a Stieltjes integral

$$\mathcal{I}(f) := \int f(z)dw(z), \tag{4.4}$$

where dw is a complex-valued measure with support at the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ in the complex plane. It follows from $\boldsymbol{w}^T\boldsymbol{v} = 1$ that $\int dw(z) = 1$. A discussion on the situation when A does not have N linearly independent eigenvectors is provided by Pozza et al. [60, 61].

It is the purpose of the present chapter to derive Gauss–Laurent-type quadrature rules for the approximation of the integral (4.4) or, equivalently, of the functional (4.1). These rules are exact for certain Laurent polynomials, which are polynomials in z and $1/z$. Gauss–Laurent quadrature rules for the approximation of (4.1) can be computed by applying a few steps of the nonsymmetric rational Lanczos process to the matrix A with initial vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. Associated anti-Gauss-Laurent rules also are developed. The latter rules allow us to compute estimates for the quadrature error in Gauss–Laurent rules. Specifically, pairs of Gauss–Laurent and associated anti-Gauss–Laurent quadrature rules allow the computation of estimates of upper and lower bounds for the quadrature error in Gauss–Laurent rules. With this we mean that a pair of a Gauss–Laurent rule and an associated anti-Gauss–Laurent rule for many integrands f, matrices A, and vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, provide upper and lower bounds for the integral (4.4), and therefore of the functional (4.1). However, they do not provide upper and lower bounds for all integrands and it is difficult to assess a priori if the computed quantities are upper and lower bounds. We therefore refer to the computed quantities as estimates of upper and lower bounds.

Anti-Gauss rules for the estimation of the error in (standard) Gauss quadrature rules for the approximation of integrals with a nonnegative measure with support on (part of) the real axis were proposed in a seminal paper by Laurie [49]. An extension to the estimation of functionals of the form (4.1) by Gauss-type quadrature rules is described in [12]. Further extensions and modifications of Gauss and anti-Gauss rules are described in [1, 2, 16, 57]. However, none of these extensions and modifications are concerned with Gauss–Laurent and anti-Gauss–Laurent quadrature rules. The reason for our interest in Gauss–Laurent-type quadrature rules is that they may provide much higher accuracy than Gauss rules with the same number of nodes if the integrand has a singularity close to the support of the measure that determines the quadrature rules. Applications of Gauss–Laurent quadrature rules to the approximation of functionals (4.1) with a symmetric matrix A are described in [8, 40]. However, Gauss–Laurent quadrature rules and associated anti-Gauss–Laurent quadrature rules for the approximation of functionals (4.1) with a nonsymmetric matrix A have not been developed until now. We remark that the present chapter, as well as the references mentioned in this

paragraph, generalize and modify an approach described by Golub and Meurant [29] for computing upper and lower bounds for functionals (4.1) with a symmetric matrix $A \in \mathbb{R}^{N\times N}$ and an integrand f with derivatives that do not change sign on the convex hull of the spectrum of A.

This chapter is organized as follows. Section 4.3 reviews the approach described in [16] for approximating the functional (4.1) by first carrying out a few steps of the nonsymmetric Lanczos process, which is described by Algorithm 2, to the matrix A with initial vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, and then using the computed quantities to define a Gauss quadrature rule for the approximation of (4.4). Associated Krylov subspaces are defined. These spaces are determined by the matrix A, its transpose, and the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. Section 4.4 introduces extended Krylov subspaces, i.e., Krylov subspaces that are determined by the matrix A, its transpose, their inverses, as well as by the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. We remark that recursion formulas for extended Krylov subspaces that are determined by a symmetric matrix are discussed by Mach et al. [50] and recursion formulas for rational Krylov subspaces that are determined by a symmetric matrix A and inverses of shifted matrices, $(A-\sigma_j I)^{-1}$, for suitable scalars σ_j are considered by Mach et al. [51]. Applications and recursion formulas for rational Krylov subspaces of the latter kind also can be found in [39, 59]. Recently, Van Buggenhout et al. [68] discussed the recursion relations for biorthogonal bases for rational Krylov subspaces determined by A, A^T, as well as by inverses of shifted matrices $(A-\sigma_j I)^{-1}$ and $(A^T-\sigma_j' I)^{-1}$ for suitable scalars σ_j and σ_j'. Section 4.4 presents an alternate derivation of these recursion formulas for the case when $\sigma_j = \sigma_j' = 0$ for all j. Our derivation extends the approach described in [42] to nonsymmetric matrices. Section 4.5 discusses the application of the recursions of Section 4.4 to the computation of Gauss–Laurent and anti-Gauss–Laurent quadrature rules. The former rules are Gauss-type quadrature rules that are exact for specified positive and negative powers of z.

A nice introduction to rational Gauss rules is provided by Gautschi [22, Section 3.1.4]. More recent discussion of rational Gauss rules can be found in [14, 56]. Applications of rational Gauss quadrature to model reduction are described by Barkouki et al. [5] and Gallivan et al. [20].

4.3 Gauss Quadrature Rules

This section describes the application of the nonsymmetric Lanczos process to the nonsymmetric matrix $A \in \mathbb{R}^{N\times N}$ to compute Gauss quadrature rules for the approximation of the functional

(4.1). Further details and extensions can be found in [2, 12, 16]. Let the vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^N$ satisfy $\boldsymbol{w}^T\boldsymbol{v} = 1$. Then application of $1 \leq m \ll N$ steps of the nonsymmetric Lanczos process to A with initial vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ gives the Lanczos decompositions

$$\begin{aligned} AV_m &= V_m T_m + \gamma_m \boldsymbol{v}_m \boldsymbol{e}_m^T, \\ A^T W_m &= W_m T_m^T + \beta_m \boldsymbol{w}_m \boldsymbol{e}_m^T, \end{aligned} \tag{4.5}$$

where the matrices $V_m = [\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m-1}] \in \mathbb{R}^{N\times m}$ and $W_m = [\boldsymbol{w}_0, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-1}] \in \mathbb{R}^{N\times m}$ with $\boldsymbol{v}_0 := \boldsymbol{v}$ and $\boldsymbol{w}_0 := \boldsymbol{w}$ satisfy

$$W_m^T V_m = I_m, \tag{4.6}$$

and the columns of V_m and W_m form bases for the Krylov subspaces

$$\begin{aligned} \mathbb{K}_m(A, \boldsymbol{v}) &= \text{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}, \\ \mathbb{K}_m(A^T, \boldsymbol{w}) &= \text{span}\{\boldsymbol{w}, A^T\boldsymbol{w}, \ldots, (A^T)^{m-1}\boldsymbol{w}\}. \end{aligned} \tag{4.7}$$

Moreover, the vectors $\boldsymbol{v}_m, \boldsymbol{w}_m \in \mathbb{R}^N$ satisfy $V_m^T\boldsymbol{w}_m = \boldsymbol{0}$, $W_m^T\boldsymbol{v}_m = \boldsymbol{0}$, and $\boldsymbol{w}_m^T\boldsymbol{v}_m = 1$, and $\beta_m, \gamma_m \in \mathbb{R}_+$. The matrix

$$T_m := \begin{bmatrix} \alpha_0 & \gamma_1 & & & 0 \\ \beta_1 & \alpha_1 & \gamma_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{m-2} & \alpha_{m-2} & \gamma_{m-1} \\ 0 & & & \beta_{m-1} & \alpha_{m-1} \end{bmatrix} \in \mathbb{R}^{m\times m},$$

is nonsymmetric and tridiagonal. We assume that m is chosen small enough so that the decompositions (4.5) with the stated properties exist.

It follows from the recursion relations (4.5) that the jth columns of V_m and W_m can be expressed as

$$\boldsymbol{v}_{j-1} = p_{j-1}(A)\boldsymbol{v}, \quad \boldsymbol{w}_{j-1} = q_{j-1}(A^T)\boldsymbol{w}, \qquad j = 1, 2, \ldots, m, \tag{4.8}$$

where p_{j-1} and q_{j-1} are polynomials of degree $j-1$.

Introduce the bilinear form

$$(q, p) := (q(A^T)\boldsymbol{w})^T(p(A)\boldsymbol{v}) = \boldsymbol{w}^T S q(\Lambda) p(\Lambda) S^{-1}\boldsymbol{v} = \int q(z)p(z)dw(z), \tag{4.9}$$

where dw is the measure in (4.4). It follows from (4.6) that the families of polynomials $\{p_0, p_1, p_2, \dots\}$ and $\{q_0, q_1, q_2, \dots\}$ defined by (4.8) are biorthogonal with respect to the bilinear form (4.9). We have

$$(q_{k-1}, p_{j-1}) = (q_{k-1}(A^T)\boldsymbol{w})^T(p_{j-1}(A)\boldsymbol{v}) = \boldsymbol{w}_k^T\boldsymbol{v}_j = \begin{cases} 1 & k = j, \\ 0 & k \neq j, \end{cases}$$

where the last equality follows from (4.6). Using the orthogonality, we can show that

$$\mathcal{G}_m(f) := \boldsymbol{e}_1^T f(T_m)\boldsymbol{e}_1 \tag{4.10}$$

is a Gauss quadrature rule for the approximation of (4.1), i.e.,

$$\mathcal{G}_m(f) = \boldsymbol{w}^T f(A)\boldsymbol{v} \qquad \forall f \in \mathbb{P}_{2m-1},$$

where $\mathbb{P}_{2m-1}$ denotes the set of all polynomials of degree at most $2m-1$; see, e.g., [2, 12, 16] for proofs.

Assume for the moment that the matrix T_m has m distinct eigenvalues. Then substituting the spectral factorization of T_m into (4.10) shows that $\mathcal{G}_m(f)$ is a quadrature rule with m nodes, which may be complex-valued. The situation when T_m does not have a spectral factorization with m linearly independent eigenvectors is discussed by Pozza et al. [60, 61].

The application of a Gauss rule (4.10) to approximate the functional (4.1) is appropriate when the function f can be approximated well by a polynomial of small to moderate degree. However, if this is not the case, then Gauss rules (4.10) with a moderate number of nodes, m, may yield a poor approximation of the functional (4.1). It is sometimes possible to circumvent this difficulty by using rational Gauss rules. The following section discusses rational Gauss rules with one pole in the finite complex plane for the approximation of (4.1).

4.4 Recursion Relations for Extended Krylov Subspaces

When the function f in (4.1) has a singularity close to the support of the measure dw in (4.4), rational Gauss quadrature rules with a pole at or close to the singularity may yield approximations of (4.1) of higher accuracy than a Gauss rule (4.10) with the same number of nodes. This is illustrated in Section 4.6.

We will assume that the singularity of f close to the support of the measure is at the origin. Rational Gauss rules that are exact for as many positive and negative powers of z as possible are

commonly referred to as Gauss–Laurent quadrature rules. Similarly as Gauss rules are related to the Krylov subspaces (4.7), Gauss–Laurent quadrature rules are related to extended Krylov subspaces

$$\begin{aligned}\mathbb{K}_{\ell,m}(A,\boldsymbol{v}) &= \mathrm{span}\{A^{-\ell+1}\boldsymbol{v},\ldots,A^{-1}\boldsymbol{v},\boldsymbol{v},A\boldsymbol{v},\ldots,A^{m-1}\boldsymbol{v}\},\\ \mathbb{K}_{\ell,m}(A^T,\boldsymbol{w}) &= \mathrm{span}\{(A^T)^{-\ell+1}\boldsymbol{w},\ldots,(A^T)^{-1}\boldsymbol{w},\boldsymbol{w},A^T\boldsymbol{w},\ldots,(A^T)^{m-1}\boldsymbol{w}\}.\end{aligned} \tag{4.11}$$

Generically, the subspaces $\mathbb{K}_{\ell,m}(A,\boldsymbol{v})$ and $\mathbb{K}_{\ell,m}(A^T,\boldsymbol{w})$ are of dimension $m+\ell-1$; if $\ell=1$, then the spaces (4.11) simplify to the standard Krylov subspaces (4.7).

The computation of Gauss–Laurent quadrature rules for the approximation of (4.1) in the case when the matrix A is symmetric is discussed in [8, 40, 42], and several other applications of extended Krylov subspaces are described by Heyouni, Jbilou, Knizhnerman, and Simoncini [37, 47]. Our contribution differs from these works in that we use the pair of extended Krylov subspaces (4.11) and develop short recursion relations for biorthogonal bases. A different approach to the derivation of such recursion relations has recently been proposed by Van Buggenout et al. [68].

The remainder of this section discusses the generation of biorthogonal bases for pairs of nested Krylov subspaces

$$\begin{aligned}&\mathbb{K}_{1,i+1}(A,\boldsymbol{v})\subset\mathbb{K}_{2,2i+1}(A,\boldsymbol{v})\subset\ldots\subset\mathbb{K}_{m,mi+1}(A,\boldsymbol{v})\subset\mathbb{R}^N,\\ &\mathbb{K}_{1,i+1}(A^T,\boldsymbol{w})\subset\mathbb{K}_{2,2i+1}(A^T,\boldsymbol{w})\subset\ldots\subset\mathbb{K}_{m,mi+1}(A^T,\boldsymbol{w})\subset\mathbb{R}^N,\end{aligned} \tag{4.12}$$

where i is a positive integer. These recursions generalize those reported in [42] for a symmetric matrix A. Schweitzer [63] recently described recursion relations for the situation when $i=1$. Each increase in the denominator degree requires the solution of linear systems of equations with the matrices A and A^T, while each increase in the numerator degree demands the evaluation of matrix-vector products with the matrices A or A^T, which typically is cheaper than the solution of systems of equations. This makes it possible to compute rational Gauss–Laurent rules corresponding to $i>1$ faster than Gauss–Laurent rules with the same number of nodes corresponding to $i=1$. Illustrative examples are presented in Section 4.6. Computed examples for the situation when A is symmetric can be found in [41, 42].

4.4.1 Recursions Relations for Extended Krylov Subspaces

In this subsection, we will use biorthogonal sequences of Laurent polynomials to generate bases for the Krylov subspaces (4.12) corresponding to the orderings

$$\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^i\boldsymbol{v}, A^{-1}\boldsymbol{v}, A^{i+1}\boldsymbol{v}, \ldots, A^{2i}\boldsymbol{v}, A^{-2}\boldsymbol{v}, A^{2i+1}\boldsymbol{v}, \ldots \ ,$$
$$\boldsymbol{w}, A^T\boldsymbol{w}, \ldots, (A^T)^i\boldsymbol{w}, (A^T)^{-1}\boldsymbol{w}, (A^T)^{i+1}\boldsymbol{w}, \ldots, (A^T)^{2i}\boldsymbol{w}, (A^T)^{-2}\boldsymbol{w}, (A^T)^{2i+1}\boldsymbol{w}, \ldots \ ,$$

where the last powers of A and A^T are required to be positive.

Introduce the space of Laurent polynomials

$$\mathcal{L}_{m,im} := \operatorname{span}\{z^{-m}, z^{-m+1}, \ldots, 1, \ldots, z^{im-1}, z^{im}\} \qquad z \in \mathbb{R}\backslash\{0\}.$$

There are two sequences of monic biorthogonal Laurent polynomials

$$\begin{aligned} &\phi_0, \phi_1, \ldots \phi_i, \phi_{-1}, \phi_{i+1}, \ldots, \phi_{2i}, \phi_{-2}, \phi_{2i+1}, \ldots, \phi_{-m+1}, \ldots, \phi_{im}, \\ &\psi_0, \psi_1, \ldots \psi_i, \psi_{-1}, \psi_{i+1}, \ldots, \psi_{2i}, \psi_{-2}, \psi_{2i+1}, \ldots, \psi_{-m+1}, \ldots, \psi_{im}, \end{aligned} \tag{4.13}$$

of the forms

$$\phi_j(z) := \begin{cases} z^j + \displaystyle\sum_{\ell=-\lfloor (j-1)/i \rfloor}^{j-1} c_{j,\ell} z^\ell, & j = 1, 2, 3, \ldots \ , \\ z^j + \displaystyle\sum_{\ell=j+1}^{-ij} c_{j,\ell} z^\ell, & j = -1, -2, -3, \ldots \ , \end{cases}$$

and

$$\psi_k(z) := \begin{cases} z^k + \displaystyle\sum_{\ell=-\lfloor (k-1)/i \rfloor}^{k-1} d_{k,\ell} z^\ell, & k = 1, 2, 3, \ldots \ , \\ z^k + \displaystyle\sum_{\ell=k+1}^{-ik} d_{k,\ell} z^\ell, & k = -1, -2, -3, \ldots \ , \end{cases}$$

with $\phi_0(z) = \psi_0(z) = 1$. Thus,

$$(\phi_j, \psi_k) = 0, \qquad j \neq k,$$

where the bilinear form is given by (4.9). We assume here that i and m are small enough so that the Laurent polynomials (4.13) indeed form biorthogonal bases for the space $\mathcal{L}_{m,im}$.

The vectors

$$\boldsymbol{v}_j = \phi_j(A)\boldsymbol{v}, \qquad j = -m+1, -m+2, \ldots, im, \tag{4.14}$$

and

$$\boldsymbol{w}_j = \psi_j(A^T)\boldsymbol{w}, \qquad j = -m+1, -m+2, \ldots, im, \tag{4.15}$$

form biorthogonal bases for the extended Krylov subspaces $\mathbb{K}_{m,im+1}(A, \boldsymbol{v})$ and $\mathbb{K}_{m,im+1}(A^T, \boldsymbol{w})$, respectively, with $\boldsymbol{v}_0 = \boldsymbol{v}$ and $\boldsymbol{w}_0 = \boldsymbol{w}$. Hence, the determination of biorthogonal bases for these extended Krylov subspaces is equivalent to the generation of biorthogonal bases for the space $\mathcal{L}_{m-1,im}$ of Laurent polynomials.

Define for the nonsingular matrix $M \in \mathbb{R}^{N\times N}$ the bilinear form

$$[\boldsymbol{x}, \boldsymbol{y}]_M = \boldsymbol{x}^T M \boldsymbol{y}, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N,$$

which is needed in the following proposition. The proposition specifies some conditions that are required to compute the trailing and leading coefficients of $\{\phi_{im}, \psi_{im}\}$ and $\{\phi_{-m}, \psi_{-m}\}$.

Proposition 7. *Let the matrix A be such that*

$$[\boldsymbol{w}_{im}, \boldsymbol{v}_{im}]_{A^{-1}} \neq 0, \qquad [\boldsymbol{w}_{-m}, \boldsymbol{v}_{-m}]_A \neq 0.$$

Then the trailing coefficients $c_{im,-m+1}$, $d_{im,-m+1}$ of ϕ_{im}, ψ_{im}, respectively, and the leading coefficients $c_{-m,im}$, $d_{-m,im}$ of ϕ_{-m}, ψ_{-m}, respectively, are nonvanishing.

Proof. *We first show that the coefficient $c_{im,-m+1}$ is nonzero. Consider the Laurent polynomials $z^{-1}\phi_{im}$ and ψ_{im}. By the properties of the inner product* (4.9)*, we obtain*

$$(z^{-1}\phi_{im}, \psi_{im}) = [\boldsymbol{w}_{im}, \boldsymbol{v}_{im}]_{A^{-1}} \neq 0.$$

On the other hand,

$$(z^{-1}\phi_{im}, \psi_{im}) = (c_{im,-m+1}z^{-m} + \varphi, \psi_{im}),$$

where $\varphi \in \mathcal{L}_{m-1,im-1}$. Hence,

$$(z^{-1}\phi_{im}, \psi_{im}) = c_{im,-m+1}(z^{-m}, \psi_{im}).$$

It follows that $c_{im,-m+1} \neq 0$. In the same manner, we can show that $d_{im,-m+1} \neq 0$.

We now apply this argument again to show that $c_{-m,im}$ is a nonvanishing. Consider the Laurent polynomials $z\phi_{-m}$ and ψ_{-m}. Using the definition of the bilinear form (4.9)*, we have*

$$(z\phi_{-m}, \psi_{-m}) = [\boldsymbol{w}_{-m}, \boldsymbol{v}_{-m}]_A \neq 0.$$

Further,

$$(z\phi_{-m}, \psi_{-m}) = (c_{-m,im}z^{im+1} + \varphi, \psi_{-m}),$$

where $\varphi \in \mathcal{L}_{m-1,im}$. *Hence,*

$$(z\phi_{-m}, \psi_{-m}) = c_{-m,im}(z^{im+1}, \psi_{-m}),$$

and therefore $c_{-m,im} \neq 0$. *Analogously, we can show that* $d_{-m,im} \neq 0$.

Suppose that biorthogonal bases of Laurent polynomials

$$\{\phi_0, \phi_1, \ldots, \phi_i, \phi_{-1}, \phi_{i+1}, \ldots, \phi_{2i}, \phi_{-2}, \phi_{2i+1}, \ldots, \phi_{im}\}$$

$$\{\psi_0, \psi_1, \ldots, \psi_i, \psi_{-1}, \psi_{i+1}, \ldots, \psi_{2i}, \psi_{-2}, \psi_{2i+1}, \ldots, \psi_{im}\}$$

for $\mathcal{L}_{m-1,im}$ are available. The next subsections describe how to extend these bases to biorthogonal bases for the space $\mathcal{L}_{m,i(m+1)}$.

4.4.2 Computation of ϕ_{-m} and ψ_{-m}

The evaluation of ϕ_{-m} and ψ_{-m} correspond to determining biorthogonal bases for $\mathcal{L}_{m,im+j}$ for $j = 0$. Consider the Laurent polynomials

$$c_{im,-m+1}\phi_{-m}(z) - z^{-1}\phi_{im}(z), \quad d_{im,-m+1}\psi_{-m}(z) - z^{-1}\psi_{im}(z) \in \mathcal{L}_{m-1,im}. \tag{4.16}$$

By Proposition 7 the coefficients $c_{im,-m+1}$ and $d_{im,-m+1}$ of ϕ_{im} and ψ_{im}, respectively, are nonvanishing. Therefore,

$$\begin{aligned} c_{im,-m+1}\phi_{-m}(z) - z^{-1}\phi_{im}(z) &= -\sum_{k=-m+1}^{im} a_{im,k}\phi_k(z), \\ d_{im,-m+1}\psi_{-m}(z) - z^{-1}\psi_{im}(z) &= -\sum_{k=-m+1}^{im} b_{im,k}\psi_k(z), \end{aligned}$$

with the Fourier coefficients given by

$$a_{im,k} = \frac{(z^{-1}\phi_{im}, \psi_k)}{(\phi_k, \psi_k)}, \quad b_{im,k} = \frac{(z^{-1}\psi_{im}, \phi_k)}{(\phi_k, \psi_k)}, \quad k = -m+1, \ldots, im.$$

Since $\phi_{im}, \psi_{im} \perp \mathcal{L}_{m-1,im-1}$ and

$$z^{-1}\phi_k(z), z^{-1}\psi_k(z) \in \mathcal{L}_{m-1,im-1}, \qquad k = -m+2, \ldots, i(m-1),$$

it follows that the only nonvanishing Fourier coefficients are related to the previous sets of $i+1$ Laurent polynomials, $\{\phi_{-m+1},\ldots,\phi_{im}\}$ and $\{\psi_{-m+1},\ldots,\psi_{im}\}$. We therefore obtain

$$\begin{aligned} c_{im,-m+1}\phi_{-m}(z) &= z^{-1}\phi_{im}(z) - a_{im,im}\phi_{im}(z) - a_{im,im-1}\phi_{im-1}(z) - \ldots \\ &\quad -a_{im,i(m-1)+1}\phi_{i(m-1)+1}(z) - a_{im,-m+1}\phi_{-m+1}(z), \\ d_{im,-m+1}\psi_{-m}(z) &= z^{-1}\psi_{im}(z) - b_{im,im}\psi_{im}(z) - b_{im,im-1}\psi_{im-1}(z) - \ldots \\ &\quad -b_{im,i(m-1)+1}\psi_{i(m-1)+1}(z) - b_{im,-m+1}\psi_{-m+1}(z). \end{aligned}$$

This yields the $(i+2)$-term recursion formulas

$$\begin{aligned} \delta_{-m}\boldsymbol{v}_{-m} &= (A^{-1} - \zeta_{im,im}I_n)\boldsymbol{v}_{im} - \zeta_{im,im-1}\boldsymbol{v}_{im-1} - \ldots \\ &\quad -\zeta_{im,i(m-1)+1}\boldsymbol{v}_{i(m-1)+1} - \zeta_{im,-m+1}\boldsymbol{v}_{-m+1}, \\ \gamma_{-m}\boldsymbol{w}_{-m} &= ((A^T)^{-1} - \eta_{im,im}I_n)\boldsymbol{w}_{im} - \eta_{im,im-1}\boldsymbol{w}_{im-1} - \ldots \\ &\quad -\eta_{im,i(m-1)+1}\boldsymbol{w}_{i(m-1)+1} - \eta_{im,-m+1}\boldsymbol{w}_{-m+1} \end{aligned} \tag{4.17}$$

with $\zeta_{j,k} := \boldsymbol{w}_k^T A^{-1}\boldsymbol{v}_j$ and $\eta_{j,k} := \boldsymbol{v}_k^T(A^{-T})\boldsymbol{w}_j$.

4.4.3 Computation of ϕ_{im+1} and ψ_{im+1}

We determine biorthogonal bases for $\mathcal{L}_{m,im+j}$ for $j=1$. Regard the Laurent polynomials

$$c_{-m,im}\phi_{im+1}(z) - z\phi_{-m}(z), \quad d_{-m,im}\psi_{im+1}(z) - z\psi_{-m}(z) \in \mathcal{L}_{m,im}.$$

Analogously to the case $j=0$, we express the Laurent polynomials (4.16) in terms of their Fourier expansions with Fourier coefficients

$$a_{-m,k} = \frac{(z\phi_{-m},\psi_k)}{(\phi_k,\psi_k)}, \quad b_{-m,k} = \frac{(z\psi_{-m},\phi_k)}{(\phi_k,\psi_k)}, \qquad k = -m,\ldots,im.$$

Note that $\phi_{-m},\psi_{-m} \perp \mathcal{L}_{m-1,im}$ and

$$z\phi_k(z), z\psi_k(z) \in \mathcal{L}_{m-1,im}, \qquad k = -m+1,\ldots,im-1.$$

Therefore ϕ_{im+1} and ψ_{im+1} satisfy

$$\begin{aligned} c_{-m,im}\phi_{im+1}(z) &= z\phi_{-m}(z) - a_{-m,-m}\phi_{-m}(z) - a_{-m,im}\phi_{im}(z), \\ d_{-m,im}\psi_{im+1}(z) &= z\psi_{-m}(z) - b_{-m,-m}\psi_{-m}(z) - b_{-m,im}\psi_{im}(z). \end{aligned}$$

This gives the three-term recursion formulas

$$\begin{aligned} \delta_{im+1}\boldsymbol{v}_{im+1} &= (A-\alpha_{-m,-m}I_n)\boldsymbol{v}_{-m}-\alpha_{-m,im}\boldsymbol{v}_{im}, \\ \gamma_{im+1}\boldsymbol{w}_{im+1} &= (A^T-\beta_{-m,-m}I_n)\boldsymbol{w}_{-m}-\beta_{-m,im}\boldsymbol{w}_{im}, \end{aligned} \tag{4.18}$$

with $\alpha_{j,k} := \boldsymbol{w}_k^T A\boldsymbol{v}_j$ and $\beta_{j,k} := \boldsymbol{v}_k^T A^T\boldsymbol{w}_j$.

4.4.4 Computation of ϕ_{im+2} and ψ_{im+2}

We would like to determine biorthogonal bases for $\mathcal{L}_{m,im+2}$. Consider the functions

$$\phi_{im+2}(z)-z\phi_{im+1}(z), \quad \psi_{im+2}(z)-z\psi_{im+1}(z) \in \mathcal{L}_{m,im+1}.$$

The Fourier expansions of $\phi_{im+2}(z)-z\phi_{im+1}(z)$ has the coefficients

$$a_{im+1,k} = \frac{(z\phi_{im+1},\psi_k)}{(\phi_k,\psi_k)}, \qquad k=-m,\ldots,im+1,$$

and the Fourier coefficients of $\psi_{im+2}(z)-z\psi_{im+1}(z)$ are given by

$$b_{im+1,k} = \frac{(z\psi_{im+1},\phi_k)}{(\phi_k,\psi_k)}, \qquad k=-m,\ldots,im+1.$$

In view of that $\phi_{im+1},\psi_{im+1} \perp \mathcal{L}_{m,im}$ and

$$z\phi_k(z), z\psi_k(z) \in \mathcal{L}_{m,im}, \qquad k=-m+1,\ldots,im-1,$$

it follows that ϕ_{im+2} and ψ_{im+2} satisfy

$$\begin{aligned} \phi_{im+2}(z) &= z\phi_{im+1}(z)-a_{im+1,im+1}\phi_{im+1}(z) \\ &\quad -a_{im+1,-m}\phi_{-m}(z)-a_{im+1,im}\phi_{im}(z), \\ \psi_{im+2}(z) &= z\psi_{im+1}(z)-b_{im+1,im+1}\psi_{im+1}(z) \\ &\quad -b_{im+1,-m}\psi_{-m}(z)-b_{im+1,im}\psi_{im}(z), \end{aligned}$$

which yields the four-term recursion formula

$$\begin{aligned} \delta_{im+2}\boldsymbol{v}_{im+2} &= (A-\alpha_{im+1,im+1}I_n)\boldsymbol{v}_{im+1} \\ &\quad -\alpha_{im+1,-m}\boldsymbol{v}_{-m}-\alpha_{im+1,im}\boldsymbol{v}_{im}, \\ \gamma_{im+2}\boldsymbol{w}_{im+2} &= (A^T-\beta_{im+1,im+1}I_n)\boldsymbol{w}_{im+1} \\ &\quad -\beta_{im+1,-m}\boldsymbol{w}_{-m}-\beta_{im+1,im}\boldsymbol{w}_{im}. \end{aligned} \tag{4.19}$$

4.4.5 Computation of ϕ_{im+j} and ψ_{im+j} for $j = 3, 4, \ldots, i$

We describe how to determine the remaining biorthogonal bases elements for the subspace $\mathcal{L}_{m,im+j}$ for $3 \leq j \leq i$. They can be computed with the aid of the nonsymmetric Lanczos recursions. We have

$$\begin{aligned} \delta_{im+j}\boldsymbol{v}_{im+j} &= (A - \alpha_{im+j-1,im+j-1}I_n)\boldsymbol{v}_{im+j-1} \\ &\quad -\alpha_{im+j-1,im+j-2}\boldsymbol{v}_{im+j-2}, \\ \gamma_{im+j}\boldsymbol{w}_{im+j} &= (A^T - \beta_{im+j-1,im+j-1}I_n)\boldsymbol{w}_{im+j-1} \\ &\quad -\beta_{im+j-1,im+j-2}\boldsymbol{w}_{im+j-2}. \end{aligned} \tag{4.20}$$

This completes the computation of the biorthogonal bases for $\mathcal{L}_{m,i(m+1)}$.

4.4.6 Algorithm and Biorthogonal Projection

The following algorithm summarizes the computation of the biorthogonal bases for the extended Krylov subspaces $\mathbb{K}_{m+1,im+1}(A, \boldsymbol{v})$ and $\mathbb{K}_{m+1,im+1}(A^T, \boldsymbol{w})$. The algorithm is based on the recurrence relations for the biorthogonal bases for $\mathcal{L}_{m-1,im}$ derived in the previous subsections. Further details on the correspondence of the biorthogonal bases for $\mathbb{K}_{m+1,im+1}(A, \boldsymbol{v})$ and $\mathbb{K}_{m+1,im+1}(A^T, \boldsymbol{w})$, and for $\mathcal{L}_{m-1,im}$ can be found after the algorithm.

It is known that the nonsymmetric Lanczos algorithm may suffer from breakdown. This occurs when the inner products $\boldsymbol{r}^T\boldsymbol{s}$ or $\hat{\boldsymbol{r}}^T\hat{\boldsymbol{s}}$ in Algorithm 3 vanish, so that a coefficient δ_i or γ_i become zero. We will assume that m is small enough so that breakdown does not occur.

Algorithm 3 Biorthogonalization process for $\mathbb{K}_{m+1,im+1}(A,\boldsymbol{v})$ and $\mathbb{K}_{m+1,im+1}(A^T,\boldsymbol{w})$,Part 1.

Input: m, i, $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^N$ such that $(\boldsymbol{v},\boldsymbol{w})=1$, functions for evaluating matrix-vector products and solve linear systems of equations with $A, A^T\in\mathbb{R}^{N\times N}$;

Output: biorthogonal bases $\{\boldsymbol{v}_k\}_{k=-m}^{im}$ and $\{\boldsymbol{w}_k\}_{k=-m}^{im}$;

Initialization: $\boldsymbol{v}_{-1}:=\boldsymbol{0}$; $\boldsymbol{w}_{-1}:=\boldsymbol{0}$; $\boldsymbol{v}_0:=\boldsymbol{v}$; $\boldsymbol{w}_0:=\boldsymbol{w}/(\boldsymbol{w}^T\boldsymbol{v})$;

for $k=0,1,\ldots,m-1$ **do**

  $\boldsymbol{r}:=A\boldsymbol{v}_{-k}$;

  $\boldsymbol{s}:=A^T\boldsymbol{w}_{-k}$;

  $\alpha_{-k,ik}:=\boldsymbol{w}_{ik}^T\boldsymbol{r}$; $\boldsymbol{r}:=\boldsymbol{r}-\alpha_{-k,ik}\boldsymbol{v}_{ik}$

  $\beta_{-k,ik}:=\boldsymbol{v}_{ik}^T\boldsymbol{s}$; $\boldsymbol{s}:=\boldsymbol{s}-\beta_{-k,ik}\boldsymbol{w}_{ik}$;

  $\alpha_{-k,-k}:=\boldsymbol{w}_{-k}^T\boldsymbol{r}$; $\boldsymbol{r}:=\boldsymbol{r}-\alpha_{-k,-k}\boldsymbol{v}_{-k}$;

  $\boldsymbol{s}:=\boldsymbol{s}-\alpha_{-k,-k}\boldsymbol{w}_{-k}$;

  $\delta_{ik+1}:=|\boldsymbol{r}^T\boldsymbol{s}|^{1/2}$; $\boldsymbol{v}_{ik+1}:=\boldsymbol{r}/\delta_{ik+1}$;

  $\gamma_{ik+1}:=\boldsymbol{r}^T\boldsymbol{s}/\delta_{ik+1}$; $\boldsymbol{w}_{ik+1}:=\boldsymbol{s}/\gamma_{ik+1}$;

  $\boldsymbol{r}:=A\boldsymbol{v}_{ik+1}$;

  $\boldsymbol{s}:=A^T\boldsymbol{w}_{ik+1}$;

  $\alpha_{ik+1,ik}:=\boldsymbol{w}_{ik}^T\boldsymbol{r}$; $\boldsymbol{r}:=\boldsymbol{r}-\alpha_{ik+1,ik}\boldsymbol{v}_{ik}$;

  $\beta_{ik+1,ik}:=\boldsymbol{v}_{ik}^T\boldsymbol{s}$; $\boldsymbol{s}:=\boldsymbol{s}-\beta_{ik+1,ik}\boldsymbol{w}_{ik}$;

  $\alpha_{ik+1,-k}:=\boldsymbol{w}_{-k}^T\boldsymbol{r}$; $\boldsymbol{r}:=\boldsymbol{r}-\alpha_{ik+1,-k}\boldsymbol{v}_{-k}$;

  $\beta_{ik+1,-k}:=\boldsymbol{v}_{-k}^T\boldsymbol{s}$; $\boldsymbol{s}:=\boldsymbol{s}-\beta_{ik+1,-k}\boldsymbol{w}_{-k}$;

  $\alpha_{ik+1,ik+1}:=\boldsymbol{w}_{ik+1}^T\boldsymbol{r}$; $\boldsymbol{r}:=\boldsymbol{r}-\alpha_{ik+1,ik+1}\boldsymbol{v}_{ik+1}$;

  $\boldsymbol{s}:=\boldsymbol{s}-\alpha_{ik+1,ik+1}\boldsymbol{w}_{ik+1}$;

  $\delta_{ik+2}:=|\boldsymbol{r}^T\boldsymbol{s}|^{1/2}$; $\boldsymbol{v}_{ik+2}:=\boldsymbol{r}/\delta_{ik+2}$;

  $\gamma_{ik+2}:=\boldsymbol{r}^T\boldsymbol{s}/\delta_{ik+2}$; $\boldsymbol{w}_{ik+2}:=\boldsymbol{s}/\gamma_{ik+2}$;

  for $j=3,\ldots,i$ **do**

    $\boldsymbol{r}:=A\boldsymbol{v}_{ik+j-1}$;

    $\boldsymbol{s}:=A^T\boldsymbol{w}_{ik+j-1}$;

    $\alpha_{ik+j-1,ik+j-2}:=\boldsymbol{w}_{ik+j-2}^T\boldsymbol{r}$; $\boldsymbol{r}:=\boldsymbol{r}-\alpha_{ik+j-1,ik+j-2}\boldsymbol{v}_{ik+j-2}$;

    $\beta_{ik+j-1,ik+j-2}:=\boldsymbol{v}_{ik+j-2}^T\boldsymbol{s}$; $\boldsymbol{s}:=\boldsymbol{s}-\beta_{ik+j-1,ik+j-2}\boldsymbol{w}_{ik+j-2}$;

Algorithm 3 Biorthogonalization process for $\mathbb{K}_{m+1,im+1}(A,\boldsymbol{v})$ and $\mathbb{K}_{m+1,im+1}(A^T,\boldsymbol{w})$,Part 2.

$\alpha_{ik+j-1,ik+j-1} := \boldsymbol{w}_{ik+j-1}^T\boldsymbol{r}$; $\boldsymbol{r} := \boldsymbol{r} - \alpha_{ik+j-1,ik+j-1}\boldsymbol{v}_{ik+j-1}$;

$\boldsymbol{s} := \boldsymbol{s} - \alpha_{ik+j-1,ik+j-1}\boldsymbol{w}_{ik+j-1}$;

$\delta_{ik+j} := |\boldsymbol{r}^T\boldsymbol{s}|^{1/2}$; $\boldsymbol{v}_{ik+j} := \boldsymbol{r}/\delta_{ik+j}$;

$\gamma_{ik+j} := \boldsymbol{r}^T\boldsymbol{s}/\delta_{ik+j}$; $\boldsymbol{w}_{ik+j} := \boldsymbol{s}/\gamma_{ik+j}$;

end for

$\hat{\boldsymbol{r}} := A^{-1}\boldsymbol{v}_{i(k+1)}$;

$\hat{\boldsymbol{s}} := A^{-T}\boldsymbol{w}_{i(k+1)}$;

$\zeta_{i(k+1),-k} := \boldsymbol{w}_{-k}^T\hat{\boldsymbol{r}}$; $\hat{\boldsymbol{r}} := \hat{\boldsymbol{r}} - \zeta_{i(k+1),-k}\boldsymbol{v}_{-k}$;

$\eta_{i(k+1),-k} := \boldsymbol{v}_{-k}^T\hat{\boldsymbol{s}}$; $\hat{\boldsymbol{s}} := \hat{\boldsymbol{s}} - \eta_{i(k+1),-k}\boldsymbol{w}_{-k}$;

for $j = 0,\ldots,i-1$ **do**

$\zeta_{i(k+1),i(k+1)-j} := \boldsymbol{w}_{i(k+1)-j}^T\hat{\boldsymbol{r}}$; $\hat{\boldsymbol{r}} := \hat{\boldsymbol{r}} - \zeta_{i(k+1),i(k+1)-j}\boldsymbol{v}_{i(k+1)-j}$;

$\eta_{i(k+1),i(k+1)-j} := \boldsymbol{v}_{i(k+1)-j}^T\hat{\boldsymbol{s}}$; $\hat{\boldsymbol{s}} := \hat{\boldsymbol{s}} - \eta_{i(k+1),i(k+1)-j}\boldsymbol{w}_{i(k+1)-j}$;

end for

$\delta_{-(k+1)} := |\hat{\boldsymbol{r}}^T\hat{\boldsymbol{s}}|^{1/2}$; $\boldsymbol{v}_{-(k+1)} := \hat{\boldsymbol{r}}/\delta_{-(k+1)}$;

$\gamma_{-(k+1)} := \hat{\boldsymbol{r}}^T\hat{\boldsymbol{s}}/\delta_{-(k+1)}$;$\boldsymbol{w}_{-(k+1)} := \hat{\boldsymbol{s}}/\gamma_{-(k+1)}$;

end for

The biorthogonal bases for the subspaces $\mathbb{K}_{m,im+1}(A,\boldsymbol{v})$ and $\mathbb{K}_{m,im+1}(A^T,\boldsymbol{w})$ determine the matrices

$$\begin{aligned}
V_{m(i+1)+1} &= [\boldsymbol{v}_0,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_i,\boldsymbol{v}_{-1},\ldots,\boldsymbol{v}_{-m+1},\ldots,\boldsymbol{v}_{im},\boldsymbol{v}_{-m}] \in \mathbb{R}^{N\times m(i+1)+1},\\
V_{m(i+1)+2} &= [V_{m(i+1)+1},\boldsymbol{v}_{im+1}] \in \mathbb{R}^{N\times(m(i+1)+2)},\\
W_{m(i+1)+1} &= [\boldsymbol{w}_0,\boldsymbol{w}_1,\ldots,\boldsymbol{w}_i,\boldsymbol{w}_{-1},\ldots,\boldsymbol{w}_{-m+1},\ldots,\boldsymbol{w}_{im},\boldsymbol{w}_{-m}] \in \mathbb{R}^{N\times m(i+1)+1},\\
W_{m(i+1)+2} &= [W_{m(i+1)+1},\boldsymbol{w}_{im+1}] \in \mathbb{R}^{N\times(m(i+1)+2)}.
\end{aligned}$$

Equations (4.17), (4.18), (4.19), and (4.20) can be used to construct the matrix $\hat{H}_{m(i+1)+1} = [h_{j,k}] \in \mathbb{R}^{(m(i+1)+2)\times m(i+1)+1}$, such that

$$\begin{aligned}
AV_{m(i+1)+1} &= V_{m(i+1)+2}\hat{H}_{m(i+1)+1}\\
A^TW_{m(i+1)+1} &= W_{m(i+1)+2}\hat{H}_{m(i+1)+1}^T.
\end{aligned}$$

The leading submatrix $H_{m(i+1)+1} \in \mathbb{R}^{(m(i+1)+1)\times(m(i+1)+1)}$ of $\hat{H}_{m(i+1)+1}$ satisfies

$$H_{m(i+1)+1} = W_{m(i+1)+1}^TAV_{m(i+1)+1}. \tag{4.21}$$

This matrix is analogous to the matrix T_m in the nonsymmetric Lanczos decomposition (4.5). It is pentadiagonal and its non-zero entries can be computed column-wise for the columns $(i+1)k+j$, $0 \le j \le i$, $0 \le k \le m-1$. We examine the columns corresponding to different values of j.

4.4.6.1 The Case $j = 1$

The columns of $AV_{m(i+1)+1}$ and $A^T W_{m(i+1)+1}$ in this case correspond to $A\boldsymbol{v}_{-k}$ and $A^T\boldsymbol{w}_{-k}$, respectively. Equation (4.18) yields

$$\begin{aligned} A\boldsymbol{v}_{-k} &= \alpha_{-k,ik}\boldsymbol{v}_{ik} + \alpha_{-k,-k}\boldsymbol{v}_{-k} + \delta_{ik+1}\boldsymbol{v}_{ik+1}, \\ A^T\boldsymbol{w}_{-k} &= \beta_{-k,ik}\boldsymbol{w}_{ik} + \beta_{-k,-k}\boldsymbol{w}_{-k} + \gamma_{ik+1}\boldsymbol{w}_{ik+1}. \end{aligned}$$

Hence, the only nontrivial entries of the $((i+1)k+1)^{\text{th}}$ column of $H_{m(i+1)+1}$ are

$$\begin{aligned} &h_{(i+1)k,(i+1)k+1} = \alpha_{-k,ik}, \quad h_{(i+1)k+1,(i+1)k+1} = \alpha_{-k,-k}, \\ &h_{(i+1)k+2,(i+1)k+1} = \delta_{ik+1}. \end{aligned}$$

4.4.6.2 The Case $j = 2$

The columns of $AV_{m(i+1)+2}$ and $A^T W_{m(i+1)+2}$ in this case correspond to $A\boldsymbol{v}_{ik+1}$ and $A^T\boldsymbol{w}_{ik+1}$, respectively. Equation (4.19) gives

$$\begin{aligned} A\boldsymbol{v}_{ik+1} &= \alpha_{ik+1,ik}\boldsymbol{v}_{ik} + \alpha_{ik+1,-k}\boldsymbol{v}_{-k} + \alpha_{ik+1,ik+1}\boldsymbol{v}_{ik+1} + \delta_{ik+2}\boldsymbol{v}_{ik+2}. \\ A^T\boldsymbol{w}_{ik+1} &= \beta_{ik+1,ik}\boldsymbol{w}_{ik} + \beta_{ik+1,-k}\boldsymbol{w}_{-k} + \beta_{ik+1,ik+1}\boldsymbol{w}_{ik+1} + \gamma_{ik+2}\boldsymbol{w}_{ik+2}. \end{aligned}$$

It follows that the only nontrivial entries of the $((i+1)k+2)^{\text{th}}$ column of $H_{m(i+1)+1}$ are

$$\begin{aligned} &h_{(i+1)k,(i+1)k+2} = \alpha_{ik+1,ik}, && h_{(i+1)k+1,(i+1)k+2} = \alpha_{ik+1,-k}, \\ &h_{(i+1)k+2,(i+1)k+2} = \alpha_{ik+1,ik+1}, && h_{(i+1)k+3,(i+1)k+2} = \delta_{ik+2}, \end{aligned}$$

where

$$\begin{aligned} \alpha_{ik+1,-k} &= (A\boldsymbol{v}_{ik+1}, \boldsymbol{w}_{-k}) \\ &= (\boldsymbol{v}_{ik+1}, A^T\boldsymbol{w}_{-k}) \\ &= (\boldsymbol{v}_{ik+1}, \beta_{-k,ik}\boldsymbol{w}_{ik} + \beta_{-k,-k}\boldsymbol{w}_{-k} + \gamma_{ik+1}\boldsymbol{w}_{ik+1}) \\ &= \gamma_{ik+1}. \end{aligned}$$

4.4.6.3 The Cases $j = 3, 4, \ldots, i$

The columns of $AV_{m(i+1)+j}$ and $A^T W_{m(i+1)+j}$ in these cases correspond to $A\boldsymbol{v}_{ik+j-1}$ and $A^T\boldsymbol{w}_{ik+j-1}$, respectively. Equation (4.20) yields

$$\begin{aligned} A\boldsymbol{v}_{ik+j-1} &= \alpha_{ik+j-1,ik+j-2}\boldsymbol{v}_{ik+j-2} + \alpha_{ik+j-1,ik+j-1}\boldsymbol{v}_{ik+j-1} + \delta_{ik+j}\boldsymbol{v}_{ik+j}, \\ A^T\boldsymbol{w}_{ik+j-1} &= \beta_{ik+j-1,ik+j-2}\boldsymbol{w}_{ik+j-2} + \beta_{ik+j-1,ik+j-1}\boldsymbol{w}_{ik+j-1} + \gamma_{ik+j}\boldsymbol{w}_{ik+j}. \end{aligned}$$

The only nontrivial entries of the $((i+1)k+j)^{\text{th}}$ columns, for $j = 3, \ldots, i$, are

$$\begin{aligned} &h_{(i+1)k+j-1,(i+1)k+j} = \gamma_{ik+j-1}, \quad h_{(i+1)k+j,(i+1)k+j} = \alpha_{ik+j-1,ik+j-1}, \\ &h_{(i+1)k+1,(i+1)k+j} = \delta_{ik+j}. \end{aligned}$$

4.4.6.4 The Case $j = 0$

The $((i+1)k)^{\text{th}}$ columns of $AV_{m(i+1)+1}$ and $A^T W_{m(i+1)+1}$ correspond to $A\boldsymbol{v}_{ik}$ and $A^T\boldsymbol{w}_{ik}$, respectively. The expressions for $A\boldsymbol{v}_{ik}$ and $A^T\boldsymbol{w}_{ik}$ can be obtained by multiplying the first and second equations in (4.17) by A and A^T, respectively, and making the appropriate substitutions for $A\boldsymbol{v}_{-m+1}, \ldots, A\boldsymbol{v}_{-m}$ and $A^T\boldsymbol{w}_{-m+1}, \ldots, A^T\boldsymbol{w}_{-m}$. Then, we simplify the resulting expressions using the facts that i) $\zeta_{ik,ik} \neq 0$ and ii) $\boldsymbol{w}_{ik}^T A^{-1}\boldsymbol{v}_r = 0,\ r = -k+1, \ldots, 0, \ldots, ik-2$. Hence,

$$\begin{aligned} A\boldsymbol{v}_{ik} &= h_{(i+1)k-1,(i+1)k}\boldsymbol{v}_{ik-1} + h_{(i+1)k,(i+1)k}\boldsymbol{v}_{ik} + \\ &\quad h_{(i+1)k+1,(i+1)k}\boldsymbol{v}_{-k} + h_{(i+1)k+2,(i+1)k}\boldsymbol{v}_{ik+1}, \\ A^T\boldsymbol{w}_{ik} &= h^T_{(i+1)k-1,(i+1)k}\boldsymbol{w}_{ik-1} + h^T_{(i+1)k,(i+1)k}\boldsymbol{w}_{ik} + \\ &\quad h^T_{(i+1)k+1,(i+1)k}\boldsymbol{w}_{-k} + h^T_{(i+1)k+2,(i+1)k}\boldsymbol{v}_{ik+1}, \end{aligned}$$

where

$$h_{(i+1)k-1,(i+1)k} = \gamma_{ik},$$

$$h_{(i+1)k+1,(i+1)k} = \tfrac{-\delta_{-k}\alpha_{-k,-k}}{\zeta_{ik,ik}},$$

$$h_{(i+1)k+2,(i+1)k} = \tfrac{-\delta_{-k}\delta_{ik+1}}{\zeta_{ik,ik}}.$$

The diagonal element is given by

$$h_{(i+1)k,(i+1)k} = \frac{1 - \zeta_{ik,ik-1}\delta_{ik} - \alpha_{-k,ik}\delta_{-k}}{\zeta_{ik,ik}}.$$

Example 4.4.1. *Let $m = 3$ and $i = 3$. Then the matrix H_{13} is of the form*

$$\begin{bmatrix}
\alpha_{0,0} & \gamma_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta_1 & \alpha_{1,1} & \gamma_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \delta_2 & \alpha_{2,2} & \gamma_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \delta_3 & h_{4,4} & \alpha_{-1,3} & \alpha_{4,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & h_{5,4} & \alpha_{-1,-1} & \gamma_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & h_{6,4} & \delta_4 & \alpha_{4,4} & \gamma_5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \delta_5 & \alpha_{5,5} & \gamma_6 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \delta_6 & h_{8,8} & \alpha_{-2,6} & \alpha_{7,6} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{9,8} & \alpha_{-2,-2} & \gamma_7 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{10,8} & \delta_7 & \alpha_{7,7} & \gamma_8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_8 & \alpha_{8,8} & \gamma_9 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_9 & h_{12,12} & \alpha_{-3,9} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{13,12} & \alpha_{-3,-3}
\end{bmatrix}.$$

Let $G_{m(i+1)+1} \in \mathbb{R}^{(m(i+1)+1)\times(m(i+1)+1)}$ denote the projection of A^{-1} onto $\mathbb{K}_{m+1,im+1}(A, \boldsymbol{v})$ to $\mathbb{K}_{m+1,im+1}(A^T, \boldsymbol{w})$, that is

$$G_{m(i+1)+1} = W_{m(i+1)+1}^T A^{-1} V_{m(i+1)+1}. \tag{4.22}$$

The matrix $G_{m(i+1)+1}$ is a rank-one modification of $H_{m(i+1)+1}^{-1}$ and banded. Its non-vanishing entries form $(i+2)\times(i+2)$ blocks along the diagonal such that any two consecutive blocks overlap in one diagonal element; see [42] for a proof of this structure in the case when A is symmetric. This proof carries over to the present situation with obvious modifications.

4.5 Application to Rational Gauss Quadrature

This section discusses Gauss–Laurent quadrature rules for the approximation of functionals (4.1) based on quantities computed by Algorithm 3. Rational Gauss rules were first considered by Gonchar and López Lagomasino [31], and have subsequently received considerable attention; see, e.g., Gautschi [22, Section 3.1.4] as well as [14, 56] for discussions and references. An application of

Gauss–Laurent rules to the computation of upper and lower bounds for certain symmetric matrix functionals is described in [41, 42]. We consider the case $i \geq 1$ in (4.12).

Application of τ steps of Algorithm 3 to the matrix A with initial vector $\boldsymbol{v}$ and $\boldsymbol{w}$, such that $\boldsymbol{v}^T\boldsymbol{w} = 1$, yields the decompositions

$$AV_\tau = V_\tau H_\tau + (h_{\tau+1,\tau}\boldsymbol{v}_{-m} + h_{\tau+2,\tau}\boldsymbol{v}_{im+1})\boldsymbol{e}_\tau^T, \tag{4.23}$$

$$A^{-1}V_\tau = V_\tau G_\tau + \boldsymbol{v}_{-m}[g_{\tau+1,\tau-i}\boldsymbol{e}_{\tau-i}, \ldots, g_{\tau+1,\tau-1}\boldsymbol{e}_{\tau-1}, g_{\tau+1,\tau}\boldsymbol{e}_{\tau,}]^T, \tag{4.24}$$

$$A^TW_\tau = W_\tau H^T{}_\tau + (h_{\tau,\tau+1}\boldsymbol{w}_{-m} + h_{\tau,\tau+2}\boldsymbol{w}_{im+1})\boldsymbol{e}_\tau^T, \tag{4.25}$$

$$A^{-T}W_\tau = W_\tau G^T{}_\tau + \boldsymbol{w}_{-m}[g_{\tau-i,\tau+1}\boldsymbol{e}_{\tau-i}, \ldots, g_{\tau-1,\tau+1}\boldsymbol{e}_{\tau-1}, g_{\tau,\tau+1}\boldsymbol{e}_{\tau,}]^T,$$

where $\tau = m(i+1)$, and the columns of $V_\tau, W_\tau \in \mathbb{R}^{N\times m(i+1)}$ form biorthogonal bases for $\mathbb{K}_{m,im+1}(A, \boldsymbol{v})$ and $\mathbb{K}_{m,im+1}(A^T, \boldsymbol{w})$, respectively. The matrices $H_\tau, G_\tau \in \mathbb{R}^{m(i+1)\times m(i+1)}$ generally are nonsymmetric and pentadiagonal. The following example illustrates the structure of these matrices.

Example 4.5.1. *The matrices H_τ and G_τ in the decompositions* (4.23) *and* (4.24) *for $i = 3$, $m = 3$, and $\tau = 12$ may have nonvanishing entries in the positions marked by "$*$":*

$$H_{12} = \begin{bmatrix} * & * & & & & & & & & & & \\ * & * & * & & & & & & & & & \\ & * & * & * & & & & & & & & \\ & & * & * & * & * & & & & & & \\ & & & * & * & * & & & & & & \\ & & & * & * & * & * & & & & & \\ & & & & & * & * & * & & & & \\ & & & & & & * & * & * & * & & \\ & & & & & & & * & * & * & & \\ & & & & & & & * & * & * & * & \\ & & & & & & & & & * & * & * \\ & & & & & & & & & & * & * \end{bmatrix}, \qquad G_{12} = \begin{bmatrix} * & * & * & * & * & & & & & & & \\ * & * & * & * & * & & & & & & & \\ * & * & * & * & * & & & & & & & \\ * & * & * & * & * & & & & & & & \\ * & * & * & * & * & * & * & * & * & * & & \\ & & & & * & * & * & * & * & & & \\ & & & & * & * & * & * & * & & & \\ & & & & * & * & * & * & * & & & \\ & & & & * & * & * & * & * & * & * & * \\ & & & & & & & & * & * & * & * \\ & & & & & & & & * & * & * & * \\ & & & & & & & & * & * & * & * \end{bmatrix}.$$

We would like to establish that

$$\boldsymbol{w}^T f(A)\boldsymbol{v} = \boldsymbol{e}_1^T f(H_\tau)\boldsymbol{e}_1 \qquad \forall f \in \mathcal{L}_{2m-2,2im+1}, \tag{4.26}$$

where $A \in \mathbb{R}^{N\times N}$ is nonsingular and $\boldsymbol{w}, \boldsymbol{v} \in \mathbb{R}^N$ satisfy $\boldsymbol{w}^T\boldsymbol{v} = 1$. The right-hand side expression is a Gauss–Laurent quadrature rule for the approximation of the left-hand side. The quadrature rule on the right-hand side has τ nodes, which are the eigenvalues of H_τ. In order to show (4.26), we first need some auxiliary results on the properties of the matrices H_τ and G_τ. Analogous results for different spaces of Laurent polynomials have been shown by Schweitzer [63]. Related results for symmetric matrices can be found in [42].

Lemma 8. *Let $A \in \mathbb{R}^{N\times N}$ be nonsingular, let $\boldsymbol{w}, \boldsymbol{v} \in \mathbb{R}^N$ satisfy $\boldsymbol{w}^T\boldsymbol{v} = 1$, and let the matrices H_τ and G_τ be defined by* (4.21) *and* (4.22)*, respectively, with the matrices V_τ, W_τ computed by Algorithm 3. Assume that $\tau := m(i+1) > 1$. Then*

$$\boldsymbol{w}^T p(A)\boldsymbol{v} = \boldsymbol{e}_1^T p(H_\tau)\boldsymbol{e}_1, \qquad p \in \mathbb{P}_{2im+1}, \tag{4.27}$$

$$\boldsymbol{w}^T q(A^{-1})\boldsymbol{v} = \boldsymbol{e}_1^T q(G_\tau)\boldsymbol{e}_1, \qquad q \in \mathbb{P}_{2m-1}. \tag{4.28}$$

Proof. *We first show* (4.27) *and note that it suffices to show* (4.27) *for monomials $p(z) = z^j$, $j = 0, 1, \ldots, 2im+1$. For $j = 1$, we obtain from* (4.23)*, using $V_\tau\boldsymbol{e}_1 = \boldsymbol{v}$, that*

$$A\boldsymbol{v} = AV_\tau\boldsymbol{e}_1 = V_\tau H_\tau\boldsymbol{e}_1,$$

where the second term vanishes because $\boldsymbol{e}_\tau^T\boldsymbol{e}_1 = 0$. For increasing values of j, we obtain

$$A^j\boldsymbol{v} = A^jV_\tau\boldsymbol{e}_1 = V_\tau H_\tau^j\boldsymbol{e}_1 + \boldsymbol{z}_\tau\boldsymbol{e}_\tau^T H_\tau^{j-1}\boldsymbol{e}_1, \tag{4.29}$$

where

$$\boldsymbol{z}_\tau = h_{\tau+1,\tau}\boldsymbol{v}_{-m} + h_{\tau+2,\tau}\boldsymbol{v}_{im+1}.$$

Due to the structure of H_τ, the second term on the right-hand side of (4.29) *vanishes, and for $j \le im$ we get*

$$A^j\boldsymbol{v} = V_\tau H_\tau^j\boldsymbol{e}_1, \qquad j = 0, 1, \ldots, im. \tag{4.30}$$

Similarly, from (4.25) *we have*

$$(A^T)^j\boldsymbol{w} = W_\tau(H^T)_\tau^j\boldsymbol{e}_1, \qquad j = 0, 1, \ldots, im. \tag{4.31}$$

Combining (4.30) *and* (4.31) *gives*

$$\begin{aligned}\boldsymbol{w}^T A^{2im+1}\boldsymbol{v} &= ((A^T)^{im}\boldsymbol{w})^T A(A^{im}\boldsymbol{v}) \\ &= (W_\tau(H^T)_\tau^{im}\boldsymbol{e}_1)^T A(V_\tau H_\tau^{im}\boldsymbol{e}_1) \\ &= \boldsymbol{e}_1^T H_\tau^{2im+1}\boldsymbol{e}_1.\end{aligned}$$

The same conclusion can be drawn for lower powers of A. This shows (4.27)*. We can prove* (4.28) *in the same manner.*

Next we will show a relation between positive powers of G_τ and negative powers of H_τ.

Lemma 9. *Let the assumptions of Lemma 8 be satisfied, and assume that H_τ is nonsingular. Then*

$$\boldsymbol{e}_1^T p(G_\tau)\boldsymbol{e}_1 = \boldsymbol{e}_1^T p(H_\tau^{-1})\boldsymbol{e}_1, \qquad p \in \mathbb{P}_{2m-2}. \tag{4.32}$$

Proof. *Multiplying* (4.24) *by $W_\tau^T A$ from the left, we get*

$$I = H_\tau G_\tau + W_\tau^T A \boldsymbol{v}_{-m}[g_{\tau+1,\tau-i}\boldsymbol{e}_{\tau-i}, \ldots, g_{\tau+1,\tau-1}\boldsymbol{e}_{\tau-1}, g_{\tau+1,\tau}\boldsymbol{e}_\tau]^T.$$

This implies that

$$H_\tau G_\tau \boldsymbol{e}_1 = \boldsymbol{e}_1.$$

We obtain by induction that

$$H_\tau^j G_\tau^j \boldsymbol{e}_1 = H_\tau^{j-1}(I - W_\tau^T A \boldsymbol{v}_{-m}[g_{\tau+1,\tau-i}\boldsymbol{e}_{\tau-i}, \ldots, g_{\tau+1,\tau}\boldsymbol{e}_{\tau,}]^T) G_\tau^{j-1}\boldsymbol{e}_1$$

for $j = 0, 1, \ldots, m-1$. Observing that only the first $(2j-2)+i$ entries of $G_\tau^{j-1}\boldsymbol{e}_1$ may be nonzero, the above equation gives

$$H_\tau^j G_\tau^j \boldsymbol{e}_1 = H_\tau^{j-1} G_\tau^{j-1}\boldsymbol{e}_1 = \boldsymbol{e}_1, \quad j = 0, 1, \ldots, m-1.$$

Multiplying by H_τ^{-j} from the left shows that

$$G_\tau^j \boldsymbol{e}_1 = H_\tau^{-j}\boldsymbol{e}_1, \quad j = 0, 1, \ldots, m-1. \tag{4.33}$$

Similarly, we obtain

$$(G_\tau^T)^j \boldsymbol{e}_1 = (H_\tau^T)^{-j}\boldsymbol{e}_1, \quad j = 0, 1, \ldots, m-1. \tag{4.34}$$

Using (4.33) *and* (4.34) *together gives*

$$\boldsymbol{e}_1^T G_\tau^{2m-2}\boldsymbol{e}_1 = \boldsymbol{e}_1^T H_\tau^{-(2m-2)}\boldsymbol{e}_1.$$

A similar argument holds for lower powers of G_τ. This shows (4.32).

We are now in a position to show that (4.26) holds.

Theorem 10. *The right-hand side of* (4.26) *is a Gauss–Laurent quadrature rule for the expression on the left-hand side.*

Proof. *Let $f \in \mathcal{L}_{2m-2,2im+1}$. Then $f(A) = p(A) + q(A^{-1})$ for some polynomials $p \in \mathbb{P}_{2im+1}$ and $q \in \mathbb{P}_{2m-2}$. We obtain from Lemma 8 that*

$$\boldsymbol{w}^T f(A)\boldsymbol{v} = \boldsymbol{w}^T p(A)\boldsymbol{v} + \boldsymbol{w}^T q(A^{-1})\boldsymbol{v} = \boldsymbol{e}_1^T p(H_\tau)\boldsymbol{e}_1 + \boldsymbol{e}_1^T q(G_\tau)\boldsymbol{e}_1.$$

Applying Lemma 9 gives

$$\boldsymbol{w}^T f(A)\boldsymbol{v} = \boldsymbol{e}_1^T p(H_\tau)\boldsymbol{e}_1 + \boldsymbol{e}_1^T q(H_\tau^{-1})\boldsymbol{e}_1 = \boldsymbol{e}_1^T f(H_\tau)\boldsymbol{e}_1.$$

This shows (4.26).

It is shown in [42] that for suitable integrands, appropriate pairs of Gauss–Laurent and Gauss–Laurent–Radau quadrature rules can be applied to determine upper and lower bounds for the functional (4.1) when the matrix A is symmetric. However, this approach is not guaranteed to furnish upper and lower bounds when the matrix A is nonsymmetric. We will show that in this situation, estimates of error bounds can be determined by evaluating appropriate pairs of Gauss–Laurent and anti-Gauss–Laurent quadrature rules.

Laurie [49] introduced the (standard) $(m+1)$-point anti-Gauss quadrature rule that gives an error of the same magnitude and of opposite sign as the (standard) m-point Gauss quadrature rule. The evaluation of the standard $(m+1)$-point anti-Gauss quadrature rule requires the computation of $m+1$ steps of the (standard) Lanczos process; see, e.g., [1, 2, 12] for details. We will show that anti-Gauss–Laurent rules can be computed in an analogous fashion.

Let

$$\tilde{\mathcal{G}}_{\tau+1}(f) := \sum_{j=1}^{\tau+1} f(\tilde{\lambda}_j)\tilde{\nu}_j\tilde{\nu}_j'$$

be the $(\tau+1)$-point anti-Gauss–Laurent rule associated with the measure dw in (4.4). This rule is determined by the requirement that

$$(\mathcal{I} - \tilde{\mathcal{G}}_{\tau+1})(f) = -(\mathcal{I} - \mathcal{G}_\tau)(f) \qquad \forall f \in \mathcal{L}_{2m-2,2im+3}, \tag{4.35}$$

where $\mathcal{G}_\tau(f)$ is characterized by

$$\mathcal{G}_\tau(f) = \mathcal{I}(f), \qquad \forall f \in \mathcal{L}_{2m-2,2im+1}. \tag{4.36}$$

Relation (4.35) shows that $\tilde{\mathcal{G}}_{\tau+1}$ is the $(\tau+1)$-point Gauss–Laurent rule for the functional

$$\mathcal{J}(f) := (2\mathcal{I} - \mathcal{G}_\tau)(f).$$

Introduce, analogously to (4.14) and (4.15), the vectors

$$\begin{aligned} \tilde{\boldsymbol{v}}_j &= \tilde{\phi}_j(A)\boldsymbol{v}, \\ \tilde{\boldsymbol{w}}_j &= \tilde{\psi}_j(A^T)\boldsymbol{w}, \end{aligned} \qquad j = -m+1, -m+2, \ldots, im+1,$$

where $\tilde{\phi}_j$ and $\tilde{\psi}_j$ are two families of biorthogonal Laurent polynomials with respect to the bilinear form

$$\{p, q\} := \mathcal{J}(pq),$$

i.e., $\{\tilde{\phi}_i, \tilde{\psi}_j\} = 0$ for all $i \neq j$ and $\{\tilde{\phi}_j, \tilde{\psi}_j\} = 1$ for all j. The biorthogonal bases $\tilde{V}_{\tau+1} = [\tilde{\boldsymbol{v}}_j]_{j=-m+1}^{im+1}$ and $\tilde{W}_{\tau+1} = [\tilde{\boldsymbol{w}}_j]_{j=-m+1}^{im+1} \in \mathbb{R}^{N\times(\tau+1)}$ for the extended Krylov subspaces $\mathbb{K}_{m,im+2}(A, \boldsymbol{v})$ and $\mathbb{K}_{m,im+2}(A^T, \boldsymbol{w})$, respectively, with $\tilde{\boldsymbol{v}}_0 = \boldsymbol{v}$ and $\tilde{\boldsymbol{w}}_0 = \boldsymbol{w}$ satisfy the decompositions

$$\begin{aligned} A\tilde{V}_{\tau+1} &= \tilde{V}_{\tau+1}\tilde{H}_{\tau+1} + \tilde{h}_{\tau+2,\tau+1}\tilde{\boldsymbol{v}}_{im+2}\boldsymbol{e}_{\tau+1}^T, \\ A^T\tilde{W}_{\tau+1} &= \tilde{W}_{\tau+1}\tilde{H}_{\tau+1}^T + \tilde{h}_{\tau+1,\tau+2}\tilde{\boldsymbol{w}}_{im+2}\boldsymbol{e}_{\tau+1}^T, \end{aligned}$$

where $\tau = m(i+1)$ and the matrix $\tilde{H}_{\tau+1} = [\tilde{h}_{j,k}] \in \mathbb{R}^{(\tau+1)\times(\tau+1)}$ is a nonsymmetric and pentadigonal. It follows from (4.35) and (4.36) that

$$\{\phi, \psi\} = [\phi, \psi] = \mathcal{I}(\phi\psi), \qquad \forall \phi\psi \in \mathcal{L}_{2m-2,2im+1}.$$

These equalities show that

$$\tilde{h}_{j,k} = h_{j,k}, \qquad j, k = 1, 2, \ldots, \tau$$

Therefore, $\tilde{\phi}_j = \phi_j$ and $\tilde{\psi}_j = \psi_j$ for $j = -m+1, -m+2, \ldots, im$.

It follows from the structure of H_τ and relations (4.23) and (4.25), in view of (4.14) and (4.15), that the Laurent polynomials

$$\begin{aligned} \breve{\phi}_\tau(z) &= h_{\tau+1,\tau}\phi_{-m}(z) + h_{\tau+2,\tau}\phi_{im+1}(z), \\ \breve{\psi}_\tau(z) &= h_{\tau,\tau+1}\psi_{-m}(z) + h_{\tau,\tau+2}\psi_{im+1}(z), \end{aligned}$$

for $i = 1$ can be computed with four-term recursion formulas

$$\begin{aligned} \breve{\phi}_\tau(z) &= (z - h_{\tau,\tau})\phi_{im}(z) - h_{\tau-1,\tau}\phi_{-m+1}(z) - h_{\tau-2,\tau}\phi_{im-1}(z), \\ \breve{\psi}_\tau(z) &= (z - h_{\tau,\tau})\psi_{im}(z) - h_{\tau,\tau-1}\psi_{-m+1}(z) - h_{\tau,\tau-2}\psi_{im-1}(z), \end{aligned}$$

which can be written as

$$\begin{aligned}\breve{\phi}_\tau(z) &= (z-\tilde{h}_{\tau,\tau})\tilde{\phi}_{im}(z) - \tilde{h}_{\tau-1,\tau}\tilde{\phi}_{-m+1}(z) - \tilde{h}_{\tau-2,\tau}\tilde{\phi}_{im-1}(z),\\ \breve{\psi}_\tau(z) &= (z-\tilde{h}_{\tau,\tau})\tilde{\psi}_{im}(z) - \tilde{h}_{\tau,\tau-1}\tilde{\psi}_{-m+1}(z) - \tilde{h}_{\tau,\tau-2}\tilde{\psi}_{im-1}(z).\end{aligned}$$

For $i > 1$, $\breve{\phi}_\tau$ and $\breve{\psi}_\tau$ can be determined with three-term recursion formulas

$$\begin{aligned}\breve{\phi}_\tau(z) &= (z-h_{\tau,\tau})\phi_{im}(z) - h_{\tau-1,\tau}\phi_{im-1}(z),\\ \breve{\psi}_\tau(z) &= (z-h_{\tau,\tau})\psi_{im}(z) - h_{\tau,\tau-1}\psi_{im-1}(z),\end{aligned}$$

which also can be written as

$$\begin{aligned}\breve{\phi}_\tau(z) &= (z-\tilde{h}_{\tau,\tau})\tilde{\phi}_{im}(z) - \tilde{h}_{\tau-1,\tau}\tilde{\phi}_{im-1}(z),\\ \breve{\psi}_\tau(z) &= (z-\tilde{h}_{\tau,\tau})\tilde{\psi}_{im}(z) - \tilde{h}_{\tau,\tau-1}\tilde{\psi}_{im-1}(z).\end{aligned}$$

For all $i \geq 1$, we can determine Laurent polynomials $\breve{\phi}_{\tau+1}$ and $\breve{\psi}_{\tau+1} \in \mathcal{L}_{m-1,im+2}$ that are biorthogonal to $\mathcal{L}_{m-1,im+1}$ with three-term recursion formulas

$$\begin{aligned}\breve{\phi}_{\tau+1}(z) &= (z-\alpha)\breve{\phi}_\tau(z) - \tilde{\gamma}_\tau\phi_{im}(z),\\ \breve{\psi}_{\tau+1}(z) &= (z-\alpha)\breve{\psi}_\tau(z) - \tilde{\delta}_\tau\psi_{im}(z).\end{aligned}$$

where

$$\tilde{\delta}_\tau\tilde{\gamma}_\tau = \{\breve{\phi}_\tau, \breve{\psi}_\tau\} = 2\mathcal{I}(\breve{\phi}_\tau\breve{\psi}_\tau) - \mathcal{G}_\tau(\breve{\phi}_\tau\breve{\psi}_\tau) = 2\mathcal{I}(\breve{\phi}_\tau\breve{\psi}_\tau) = 2[\breve{\phi}_\tau, \breve{\psi}_\tau] = 2\delta_\tau\gamma_\tau.$$

We may choose $\tilde{\delta}_\tau = \sqrt{2}\delta_\tau$ and $\tilde{\gamma}_\tau = \sqrt{2}\gamma_\tau$. It follows that the nonsymmetric pentadigonal matrix associated with the anti-Gauss–Laurent rule $\tilde{\mathcal{G}}_{\tau+1}$ is given by

$$\tilde{H}_{\tau+1} = \begin{bmatrix} H_\tau & \sqrt{2}\gamma_\tau \\ \sqrt{2}\delta_\tau & \alpha \end{bmatrix} \in \mathbb{R}^{(\tau+1)\times(\tau+1)},$$

where the last diagonal coefficient can be determined by from $\alpha = (z\breve{\phi}_\tau, \breve{\psi}_\tau)$. This coefficient can be evaluated by carrying out one additional "standard step" of Algorithm 3 that uses the three-term recurrence relation, i.e., we evaluate

$$\begin{aligned}\boldsymbol{r} &:= A\boldsymbol{v}_{im} - h_{\tau,\tau}\boldsymbol{v}_{im} - \gamma_{im}\boldsymbol{v}_{im-1};\\ \boldsymbol{s} &:= A^T\boldsymbol{w}_{im} - h_{\tau,\tau}\boldsymbol{w}_{im} - \delta_{im}\boldsymbol{w}_{im-1};\end{aligned}$$

and then compute

$$\delta_{im+1} := |\boldsymbol{r}^T \boldsymbol{s}|^{1/2}; \quad \gamma_{im+1} := \boldsymbol{r}^T \boldsymbol{s}/\delta_{im+1}; \quad \boldsymbol{v}_{im+1} := \boldsymbol{r}/\delta_{im+1}; \quad \boldsymbol{w}_{im+1} := \boldsymbol{s}/\gamma_{im+1}.$$

Finally,

$$h_{\tau+1,\tau+1} = \alpha := \boldsymbol{w}_{im+1}^T A \boldsymbol{v}_{im+1};$$

Analogously to formula (4.26), the anti-Gauss–Laurent quadrature rule can be evaluated according to

$$\tilde{\mathcal{G}}_{\tau+1}(f) = e_1^T f(\tilde{H}_{\tau+1}) e_1 \qquad \forall f \in \mathcal{L}_{2m-2,2im+3}.$$

We are now in a position to provide sufficient conditions for $\mathcal{G}_\tau(f)$ and $\tilde{\mathcal{G}}_{\tau+1}(f)$ to bracket $\mathcal{I}(f)$. Assume that we can carry out N steps of the Algorithm 3 without breakdown. This yields biorthogonal bases $\{\boldsymbol{v}_j, \boldsymbol{w}_j\}_{j=0}^{N-1}$ of $\mathbb{R}^N$ and associated sequences of Laurent polynomials $\{\phi_j, \psi_j\}_{j=0}^{N-1}$ defined by (4.14) and (4.15) that satisfy (4.9).

Theorem 11. *Consider the expansion of the integrand*

$$f(z) = \sum_{j=0}^{N-1} \sigma_j \phi_j(z), \quad z \in \lambda(A), \tag{4.37}$$

in terms of the Laurent polynomials ϕ_j, and assume that the coefficients σ_j in (4.37) *are such that*

$$\left| \sum_{j=2im+2}^{2im+3} \sigma_j \mathcal{G}_\tau(\phi_j) \right| \geq \max \left\{ \left| \sum_{j=2im+4}^{N-1} \sigma_j \mathcal{G}_\tau(\phi_j) \right|, \left| \sum_{j=2im+4}^{N-1} \sigma_j \tilde{\mathcal{G}}_{\tau+1}(\phi_j) \right| \right\}. \tag{4.38}$$

Then the quadrature rules $\mathcal{G}_\tau(f)$ and $\tilde{\mathcal{G}}_{\tau+1}(f)$ bracket $\mathcal{I}(f)$.

Proof. *Since*

$$\mathcal{I}(f) = \sigma_0 \mathcal{I}(\phi_0), \qquad \mathcal{I}(\phi_j) = 0, \quad \forall j > 0$$

we have, in view of (4.35) *and* (4.36)*, that*

$$\begin{aligned} \mathcal{G}_\tau(f) = \sum_{j=0}^{N-1} \sigma_j \mathcal{G}_\tau(\phi_j) &= \sum_{j=0}^{2im+1} \sigma_j \mathcal{G}_\tau(\phi_j) + \sum_{j=2im+2}^{N-1} \sigma_j \mathcal{G}_\tau(\phi_j) \\ &= \mathcal{I}(f) + \sigma_{2im+2} \mathcal{G}_\tau(\phi_{2im+2}) + \sigma_{2im+3} \mathcal{G}_\tau(\phi_{2im+3}) + \sum_{j=2im+4}^{N-1} \sigma_j \mathcal{G}_\tau(\phi_j). \end{aligned} \tag{4.39}$$

$$\tilde{\mathcal{G}}_{\tau+1}(f) = \sum_{j=0}^{N-1} \sigma_j \tilde{\mathcal{G}}_{\tau+1}(\phi_j) = \sum_{j=0}^{2im+3} \sigma_j (2\mathcal{I} - \mathcal{G}_\tau)(\phi_j) + \sum_{j=2im+4}^{N-1} \sigma_j \tilde{\mathcal{G}}_{\tau+1}(\phi_j) \tag{4.40}$$

$$= \mathcal{I}(f) - \sigma_{2im+2}\mathcal{G}_\tau(\phi_{2im+2}) - \sigma_{2im+3}\mathcal{G}_\tau(\phi_{2im+3}) + \sum_{j=2im+4}^{N-1} \sigma_j \tilde{\mathcal{G}}_{\tau+1}(\phi_j).$$

combining (4.39) *and* (4.40) *show* (4.38).

Theorem 11 shows that if the coefficients σ_j decay sufficiently rapidly with increasing j, then rational Gauss and anti-Gauss rules provide quadrature errors that are of opposite sign and of roughly the same magnitude. The following example illustrates the structure of the matrix $\tilde{H}_{\tau+1}$ for the cases $i = 1, 2, 3$.

Example 4.5.2. *The matrix $\tilde{H}_{\tau+1}$ may have nonvanishing entries in the positions marked by "$*$":*

$$\tilde{H}_7 = \begin{bmatrix} * & * & & & & & \\ * & * & * & * & & & \\ & * & * & * & & & \\ & * & * & * & * & * & \\ & & & * & * & * & \\ & & & * & * & * & * \\ & & & & & * & * \end{bmatrix} \qquad \text{for} \quad i = 1, m = 3, \tau = 6$$

$$\tilde{H}_7 = \begin{bmatrix} * & * & & & & & \\ * & * & * & & & & \\ & * & * & * & * & & \\ & & * & * & * & & \\ & & * & * & * & * & \\ & & & & * & * & * \\ & & & & & * & * \end{bmatrix} \qquad \text{for} \quad i = 2, m = 2, \tau = 6$$

$$\tilde{H}_{13} = \begin{bmatrix} * & * & & & & & & & & & & & \\ * & * & * & & & & & & & & & & \\ & * & * & * & & & & & & & & & \\ & & * & * & * & * & & & & & & & \\ & & & * & * & * & & & & & & & \\ & & & * & * & * & * & & & & & & \\ & & & & & * & * & * & & & & & \\ & & & & & * & * & * & * & & & & \\ & & & & & & & * & * & * & & & \\ & & & & & & & & * & * & * & & \\ & & & & & & & & & * & * & * & \\ & & & & & & & & & & * & * & \end{bmatrix} \qquad \text{for} \quad i = 3, m = 3, \tau = 12$$

4.6 Computed Examples

In this section, we illustrate the performance of the Gauss–Laurent and associated anti-Gauss–Laurent rules when applied to several functionals (4.1).

The purpose of these examples is to compare the performance of the standard Gauss and Gauss–Laurent rules for the case $i = 3$. The last example illustrates the performance of these quadrature rules for $i = 1, 2$, and 3. Also, we show that pairs of Gauss–Laurent and anti-Gauss–Laurent quadrature rules provide upper and lower bounds for certain functionals (4.1). We compare the approximations obtained by the quadrature rules and the values computed by explicitly evaluating the functional (4.1). We choose $\tau = 0 \mod 4$ and $\tau = 0 \mod 3$ to ensure that the matrix H_τ defined by (4.23) is of the appropriate dimensions for $i = 1, 2, 3$. The table column headings $\boldsymbol{e}_1^T f(T)\boldsymbol{e}_1$, $\boldsymbol{e}_1^T f(H)\boldsymbol{e}_1$, and $\boldsymbol{e}_1^T f(\tilde{H})\boldsymbol{e}_1$ refer to standard Gauss, Gauss–Laurent and anti-Gauss–Laurent quadrature rules, respectively.

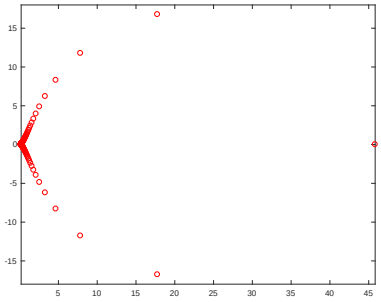

Figure 1: Example 4.6.1: Spectrum of matrix A in $\mathbb{C}$. The eigenvalues are marked by "o". The horizontal axis shows the real parts of the eigenvalues, the vertical axis the imaginary parts.

Example 4.6.1. We would like to determine approximations of the functional

$$F(A) := \boldsymbol{w}^T \exp(-A) A^{-1/2} \boldsymbol{v},$$

where $A \in \mathbb{R}^{200\times 200}$ is a real nonsymmetric Toeplitz matrix with first row and column $[1, 1/2, \ldots, 1/200]$ and $[1, 1, \ldots, 1]^T$, respectively. The vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ have normally distributed random entries with zero mean and are scaled so that $\boldsymbol{w}^T\boldsymbol{v} = 1$. Figure 1 shows the eigenvalues of A. The largest magnitude is real-valued and about 45.8; the eigenvalues with the largest imaginary parts (in magnitude) are approximately $17.8 \pm 16.8i$, where $i = \sqrt{-1}$ is the imaginary unit. The eigenvalue of smallest magnitude is real and about 0.195.

We evaluate (4.1) as $\boldsymbol{w}^T \exp(-A)A^{-1/2}\boldsymbol{v}$, where the vector $A^{-1/2}\boldsymbol{v}$ is calculated by first computing the matrix square root and then solving a linear system of equations. The exact value of $F(A)$ is approximately 0.9990. We report this value to allow a reader to estimate the relative approximation error from Table 4.

The Gauss–Laurent rule is evaluated as

$$\boldsymbol{e}_1^T \exp(-H_\tau)G_\tau^{1/2}\boldsymbol{e}_1,$$

where $G_\tau^{1/2}\boldsymbol{e}_1$ is the first column of the inverse of the square root of the matrix H. It is determined by first computing the matrix square root and then solving a linear system of equations. The exponential is computed with the MATLAB function `expm`. The standard Gauss rule

$$\boldsymbol{e}_1^T \exp(-T_\tau)T_\tau^{-1/2}\boldsymbol{e}_1$$

is determined by first computing the matrix square root of T_τ and then solving a linear system of equations for the vector $T_\tau^{-1/2}\boldsymbol{e}_1$.

Columns 2 and 3 of Table 4 display the errors in approximations determined by standard Gauss and Gauss–Laurent rules for $i = 3$. We observe that the Gauss–Laurent rules yield higher accuracy than the standard Gauss rules when using the same number, τ, of quadrature nodes. Columns 3 and 4 of Table 4 show the errors in approximations obtained by Gauss–Laurent and associated anti-Gauss–Laurent rules to have opposite sign and to be of about the same magnitude for each value of τ. Therefore, the average rules

$$\frac{1}{2}(\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1 + \boldsymbol{e}_1^T f(\tilde{H}_{i=3})\boldsymbol{e}_1) \tag{4.41}$$

for the different τ-values determine more accurate approximations of $F(A)$ than the corresponding Gauss–Laurent rules. In applications, we use pairs of Gauss–Laurent and associated anti-Gauss–Laurent rules to determine estimates of upper and lower bounds for the functional $F(A)$, and use the averages rule as an approximation of $F(A)$.

Example 4.6.2. This example determines approximations of the functional

$$F(A) := \boldsymbol{w}^T \ln(A)\boldsymbol{v},$$

with A the same matrix as in Example 4.6.1. We let $\boldsymbol{v} = [1, 1, \ldots, 1]^T \in \mathbb{R}^{200}$ and $\boldsymbol{w} = [1/200, 1/200, \ldots, 1/200]^T \in \mathbb{R}^{200}$ so that $\boldsymbol{w}^T\boldsymbol{v} = 1$. The exact value of $F(A)$ is approximately

τ	$\boldsymbol{e}_1^T f(T)\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=3})\boldsymbol{e}_1$
12	$1.08 \cdot 10^{-1}$	$9.69 \cdot 10^{-4}$	$-8.69 \cdot 10^{-4}$
16	$3.15 \cdot 10^{-2}$	$2.69 \cdot 10^{-5}$	$-2.26 \cdot 10^{-5}$
20	$8.80 \cdot 10^{-3}$	$1.28 \cdot 10^{-6}$	$-1.48 \cdot 10^{-6}$

Table 4: Example 4.6.1: Errors for computed approximations of $F(A) = \boldsymbol{w}^T \exp(-A)A^{-1/2}\boldsymbol{v}$ with A a nonsymmetric Toeplitz matrix.

$2.924 \cdot 10^{-4}$. Columns 2 and 3 of Table 5 show the difference between the exact value and the approximations determined by the standard Gauss and Gauss–Laurent rules for $i = 3$ and the same number of nodes, τ. It can be seen that the quadrature error for the Gauss–Laurent rules is the smallest for all values of τ. Column 4 of Table 5 displays the errors achieved with the anti-Gauss–Laurent rules. We observe that the errors of these quadrature rules are of opposite sign and of about the same magnitude as the error in the corresponding Gauss–Laurent rules. Similarly as above, this indicates that the average rules (4.41) are more accurate than the corresponding Gauss–Laurent and anti-Gauss–Laurent rules.

τ	$\boldsymbol{e}_1^T f(T)\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=3})\boldsymbol{e}_1$
8	$-1.10 \cdot 10^{-3}$	$-1.78 \cdot 10^{-5}$	$1.75 \cdot 10^{-5}$
12	$-9.71 \cdot 10^{-5}$	$-4.61 \cdot 10^{-8}$	$4.36 \cdot 10^{-8}$
16	$-7.58 \cdot 10^{-6}$	$-1.19 \cdot 10^{-10}$	$1.14 \cdot 10^{-10}$

Table 5: Example 4.6.2: Errors for computed approximations of $F(A) = \boldsymbol{w}^T \ln(A)\boldsymbol{v}$ with A a nonsymmetric Toeplitz matrix.

Example 4.6.3. In this example, we approximate the value

$$F(A) := \boldsymbol{w}^T A^{-1/2}\boldsymbol{v},$$

where $A \in \mathbb{R}^{1000\times 1000}$ is the nonsymmetric tridiagonal Toeplitz matrix $[-1, 2, 1]$. The vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ have normally distributed random entries with zero mean; they are scaled so that $\boldsymbol{w}^T\boldsymbol{v} = 1$. The eigenvalues of A all have real part 2 and their imaginary parts are zeros of a Chebyshev polynomial of the first kind of degree 1000 for the interval $[-2, 2]$. The exact value $F(A)$ is approximately

0.6201. Columns 2 and 3 of Table 6 display the errors in approximations obtained by the standard Gauss and Gauss–Laurent rules for $i = 3$. We find that Gauss–Laurent rules give significantly smaller approximations errors than the standard Gauss rules. Columns 4 of Table 6 shows the Gauss–Laurent and anti-Gauss–Laurent rules to bracket the exact value. This implies that the average (4.41) will be quite accurate.

τ	$\boldsymbol{e}_1^T f(T)\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=3})\boldsymbol{e}_1$
8	$1.22 \cdot 10^{-7}$	$3.17 \cdot 10^{-10}$	$-3.13 \cdot 10^{-10}$
12	$9.41 \cdot 10^{-11}$	$1.22 \cdot 10^{-15}$	$-1.99 \cdot 10^{-15}$

Table 6: Example 4.6.3: Errors for computed approximations of $F(A) = \boldsymbol{w}^T A^{-1/2}\boldsymbol{v}$ with A a nonsymmetric tridiagonal Toeplitz matrix.

Example 4.6.4. In this example, we determine approximations of the functional

$$F(A) := \boldsymbol{w}^T(A^5 + A^{-6})\boldsymbol{v},$$

where $A \in \mathbb{R}^{1000\times 1000}$ is the same matrix as in Example 4.6.3, $\boldsymbol{v} = [1, 1, \ldots, 1]^T$, and $\boldsymbol{w} = [1, 0, \ldots, 0]^T$. Thus, $\boldsymbol{w}^T\boldsymbol{v} = 1$. The value of $F(A)$ is approximately $7.340 \cdot 10^1$. Columns 2 and 3 of Table 7 display the errors in approximations determine by the standard Gauss and Gauss–Laurent rules for $i = 3$. Column 4 of Table 7 shows the approximations determine anti-Gauss-Laurent rules and illustrates that Gauss–Laurent and anti-Gauss–Laurent rules bracket the exact value.

τ	$\boldsymbol{e}_1^T f(T)\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=3})\boldsymbol{e}_1$
8	$1.60 \cdot 10^{-5}$	$5.77 \cdot 10^{-7}$	$-5.77 \cdot 10^{-7}$
12	$8.75 \cdot 10^{-8}$	$1.08 \cdot 10^{-11}$	$-1.09 \cdot 10^{-11}$

Table 7: Example 4.6.4: Errors for computed approximations of $F(A) = \boldsymbol{w}^T(A^5 + A^{-6})\boldsymbol{v}$ with A a nonsymmetric tridiagonal Toeplitz matrix.

Example 4.6.5. In our last example, we approximate the value

$$F(A) := \boldsymbol{w}^T \ln(A)\boldsymbol{v},$$

where $\boldsymbol{v} = [1, 1, \ldots, 1]^T$, $\boldsymbol{w} = [1, 0, \ldots, 0]^T$, and the matrix A is obtained by discretizing the differential operator $-\Delta + \rho_1 \frac{\partial}{\partial x} + \rho_2 \frac{\partial}{\partial y}$. Here Δ denotes the two-dimensional Laplacian, which is

discretized on the unit square by the standard 5-point stencil on a uniform mesh with grid size $h = \frac{1}{41}$. The discretization error is $O(h^2)$ as $h \searrow 0$. The partial first derivatives are discretized by the standard symmetric 2-point stencil with discretization error $O(h^2)$. Dirichlet boundary conditions are imposed. The coefficients ρ_i are defined below. This gives a nonsymmetric matrix $A \in \mathbb{R}^{1600\times 1600}$ that can be represented as follows

$$A := -\frac{1}{h^2}(I_{40} \otimes C_1 + C_2 \otimes I_{40}),$$

where

$$C_i = \begin{bmatrix} -2 & 1-\rho_i\frac{h}{2} & 0 & \cdots & 0 \\ 1+\rho_i\frac{h}{2} & -2 & 1-\rho_i\frac{h}{2} & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & 1+\rho_i\frac{h}{2} & -2 & 1-\rho_i\frac{h}{2} \\ 0 & \cdots & 0 & 1+\rho_i\frac{h}{2} & -2 \end{bmatrix} \in \mathbb{R}^{40\times 40}.$$

see, e.g., [53, 63]. The convection coefficients ρ_i are chosen such that the Péclet numbers $Pe_i = \frac{\rho_i h}{2}$ are equal to $Pe_1 = 0.2$ and $Pe_2 = 0.1$, respectively. All eigenvalues of A are real and positive; the extreme eigenvalues are $\lambda_1 = 1.04 \cdot 10^2$ and $\lambda_{1600} = 1.33 \cdot 10^4$.

Table 8 displays the difference between the exact value, $F(A) \approx 8.019$, and some approximations determined by the standard Gauss and Gauss–Laurent quadrature rules for $i = 1, 2, 3$. Since $\tau = m(i+1)$, only certain combinations of i and m give quadrature rules with τ nodes. The entrries '$-$' mark combinations of m and i that do not correspond to quadrature rules with τ nodes. We note that Gauss–Laurent rules give the most accurate approximations of $F(A)$. Furthermore, the results achieved with Gauss–Laurent rules are fairly insensitive to the choice of $i \geq 1$. Therefore, it might be beneficial to use a value of i larger than one and in this manner reduce the computational cost. The Tables 9, 10, and 11 show the Gauss–Laurent and associated anti-Gauss–Laurent quadrature rules to give errors of about the same magnitude and of opposite sign.

τ	$\boldsymbol{e}_1^T f(T)\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=1})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=2})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1$
6	$-3.40 \cdot 10^{-3}$	–	$-4.47 \cdot 10^{-4}$	–
8	$-1.10 \cdot 10^{-3}$	$-1.84 \cdot 10^{-5}$	–	$-9.11 \cdot 10^{-5}$
12	$-1.56 \cdot 10^{-4}$	$-9.59 \cdot 10^{-8}$	$-3.40 \cdot 10^{-7}$	$-1.08 \cdot 10^{-6}$
15	$-4.16 \cdot 10^{-5}$	–	$-8.66 \cdot 10^{-9}$	–
16	$-2.72 \cdot 10^{-5}$	$-3.50 \cdot 10^{-10}$	–	$-1.33 \cdot 10^{-8}$

Table 8: Example 4.6.5: Errors for computed approximations of $F(A) = \boldsymbol{w}^T \ln(A)\boldsymbol{v}$ for $i = 1, 2, 3$ when A is a discretization of a differential operator.

τ	$\boldsymbol{e}_1^T f(H_{i=1})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=1})\boldsymbol{e}_1$
8	$-1.84 \cdot 10^{-5}$	$1.82 \cdot 10^{-5}$
12	$-9.59 \cdot 10^{-8}$	$9.55 \cdot 10^{-8}$
16	$-3.50 \cdot 10^{-10}$	$3.49 \cdot 10^{-10}$

Table 9: Example 4.6.5: Errors for computed approximations of $F(A) = \boldsymbol{w}^T \ln(A)\boldsymbol{v}$ for $i = 1$, when A is a discretization of a differential operator.

τ	$\boldsymbol{e}_1^T f(H_{i=2})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=2})\boldsymbol{e}_1$
6	$-4.47 \cdot 10^{-4}$	$4.42 \cdot 10^{-4}$
12	$-3.40 \cdot 10^{-7}$	$3.39 \cdot 10^{-7}$
15	$-8.66 \cdot 10^{-9}$	$8.67 \cdot 10^{-9}$

Table 10: Example 4.6.5: Errors for computed approximations of $F(A) = \boldsymbol{w}^T \ln(A)\boldsymbol{v}$ for $i = 2$, when A is a discretization of a differential operator.

τ	$\boldsymbol{e}_1^T f(H_{i=3})\boldsymbol{e}_1$	$\boldsymbol{e}_1^T f(\tilde{H}_{i=3})\boldsymbol{e}_1$
8	$-9.11 \cdot 10^{-5}$	$9.06 \cdot 10^{-5}$
12	$-1.08 \cdot 10^{-6}$	$1.08 \cdot 10^{-6}$
16	$-1.33 \cdot 10^{-8}$	$1.34 \cdot 10^{-8}$

Table 11: Example 4.6.5: Errors for computed approximations of $F(A) = \boldsymbol{w}^T \ln(A)\boldsymbol{v}$ for $i = 3$, when A is a discretization of a differential operator.

CHAPTER 5

Rational Gauss-Type Quadrature Rules

5.1 Overview

This chapter is concerned with computing approximations of matrix functionals of the form $F(A) := \boldsymbol{v}^T f(A)\boldsymbol{v}$ with the aid of rational Gauss quadrature rules. Associated rational Gauss–Radau and rational anti-Gauss rules are developed. Pairs of rational Gauss and rational Gauss–Radau quadrature rules, or pairs of rational Gauss and rational anti-Gauss quadrature rules, can be used to determine upper and lower bounds, or approximate upper and lower bounds, for $F(A)$. The application of rational Gauss rules, instead of standard Gauss rules, is beneficial when the function f has singularities close to the spectrum of A.

5.2 Introduction

We are interested in computing approximations of matrix functionals of the form

$$F(A) := \boldsymbol{v}^T f(A)\boldsymbol{v} \tag{5.1}$$

by quadrature rules, where $A \in \mathbb{R}^{N\times N}$ is a large symmetric positive definite matrix, the vector $\boldsymbol{v} \in \mathbb{R}^N \backslash \{\boldsymbol{0}\}$ is of unit Euclidean norm, and f is a Stieltjes function, i.e., f has the representation

$$f(z) = \int_0^\infty \frac{1}{t+z}\, d\mu(t), \quad z \in \mathbb{C} \setminus (-\infty, 0], \tag{5.2}$$

where the nonnegative measure $d\mu$ is such $f(z)$ is well defined; see, e.g., [7, 9, 18, 19, 36, 46, 52] for discussions and illustrations of Stieltjes functions. Examples include

$$f(z) = z^{-a} = \frac{\sin(a\pi)}{\pi} \int_0^\infty \frac{1}{t+z}\, d\mu(t), \quad \text{with} \quad d\mu(t) = t^{-a}dt, \quad a \in (0,1), \tag{5.3}$$

$$f(z) = \frac{\log(1+z)}{z} = \int_0^\infty \frac{1}{t+z}\, d\mu(t), \quad \text{with} \quad d\mu(t) = \begin{cases} 0, & 0 \le t \le 1, \\ t^{-1}dt, & t > 1. \end{cases}$$

Let the matrix A have the spectral factorization (3.4) where the eigenvalues λ_i of A are ordered according to $0 < \lambda_1 \le \cdots \le \lambda_N$. The spectral factorization is used in the derivation of the

quadrature rules of this chapter, but application of these rules does not require the computation of this factorization.

Using the factorization (3.4), the expression (5.1) with f defined by (5.2) can be written as

$$\begin{aligned} \boldsymbol{v}^T f(A)\boldsymbol{v} = \boldsymbol{v}^T S f(\Lambda) S^T \boldsymbol{v} &= \int_0^\infty \boldsymbol{v}^T S(tI+\Lambda)^{-1} S^T \boldsymbol{v}\, d\mu(t) \\ &= \int_0^\infty \sum_{i=1}^N (t+\lambda_i)^{-1} \nu_i^2 d\mu(t) \\ &= \int_0^\infty \int_0^\infty (t+y)^{-1} d\nu(y) d\mu(t) \\ &= \int_0^\infty f(y) d\nu(y) =: \mathcal{I}(f), \end{aligned} \tag{5.4}$$

where the vector $\boldsymbol{v}^T S$ is defined in Chapter 3, and the nonnegative measure $d\nu(y)$ has support at the eigenvalues λ_i of A. The associated distribution function $\nu(y)$ can be chosen to be a nondecreasing piece-wise constant function with jumps at the eigenvalues λ_i.

The approximation of expressions $f(A)\boldsymbol{v}$ when f is a Stieltjes function, as well as the computation of error bounds or estimates, has received considerable attention; see, e.g., [7, 18, 19, 52] and references therein. Frommer and Schweitzer [18, 19] approximate $f(A)$ by a polynomial, while Massei and Robol [52] use a rational function. A difficulty with polynomial approximants is that their degree may have to be large to yield desired accuracy of the computed approximation of (5.1). This is illustrated in Section 5.6.

Our contribution differs from the works [18, 19, 52] in that we use rational Gauss quadrature rules for the approximation of (5.1). These rules were introduced in [56] and are based on orthogonal rational functions that satisfy short recurrence relations and have prescribed poles. Rational Gauss quadrature rules associated with the measure $d\nu$ in (5.4) can be computed by applying a few steps of a rational Lanczos process to the matrix A with initial vector $\boldsymbol{v}$. These rules are exact for certain rational functions with prescribed poles. This chapter defines rational Gauss–Radau quadrature rules, and shows how pairs of rational Gauss and Gauss–Radau rules can be applied to determine upper and lower bounds for the functional (5.1). These bounds are analogues of bounds provided by pairs of standard Gauss and Gauss–Radau rules that have been described by Golub and Meurant [29]. The evaluation of these bounds requires that bounds for extreme eigenvalues of A be available.

We also define rational anti-Gauss quadrature rules and simplified rational anti-Gauss rules. Pairs of rational Gauss and rational anti-Gauss rules or pairs of rational Gauss and simplified

rational anti-Gauss rules also provide upper and lower bounds for the functional (5.1) under suitable conditions, and do not require knowledge of extreme eigenvalues of A. Rational anti-Gauss rules provide an extension to rational Gauss quadrature of the anti-Gauss rules associated with (standard) Gauss quadrature rules that were introduced by Laurie [49].

This chapter is organized as follows. Section 5.3 reviews how (standard) Gauss quadrature rules for the approximation of (5.4) can be determined by carrying out a few steps of the Lanczos process applied to the symmetric matrix A with initial vector $\boldsymbol{v}$. Section 5.4 discusses available results on recursion formulas for orthonormal bases for rational Krylov subspaces associated with orthogonal rational functions with several fixed poles. The recursion formulas are determined by the vector $\boldsymbol{v}$, the symmetric matrix A, and by shifted matrices $(A-\alpha_i I)^{-1}$. In our computations, the α_i are negative poles of the rational functions that determine the rational Gauss, rational Gauss–Radau, rational anti-Gauss, and simplified rational anti-Gauss quadrature rules used to approximate (5.1). It also is possible to let some poles appear in complex conjugate pairs. Recursion relations for rational orthogonal functions with poles α_i are reviewed in Section 5.4, and their application to rational Gauss, rational Gauss–Radau, rational anti-Gauss, and simplified rational anti-Gauss quadrature rules is discussed in Section 5.5. Section 5.6 presents a few computed examples.

5.3 Gauss Quadrature Rules

This section reviews the application of the symmetric Lanczos process to the symmetric positive definite matrix $A \in \mathbb{R}^{N\times N}$ to evaluate Gauss quadrature rules for the approximation of the functional (5.1); see, e.g., Golub and Meurant [29] for further details. The application of $1 \leq m \ll N$ steps of the symmetric Lanczos process to A with initial unit vector $\boldsymbol{v}$ gives the Lanczos decomposition

$$AV_m = V_m T_m + \beta_m \boldsymbol{v}_{m+1}\boldsymbol{e}_m^T, \tag{5.5}$$

where the matrix $V_m = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m] \in \mathbb{R}^{N\times m}$, with $\boldsymbol{v}_1 := \boldsymbol{v}$, has orthonormal columns that form a basis for the Krylov subspace

$$\mathbb{K}_m(A, \boldsymbol{v}) = \text{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}.$$

Moreover, the unit vector $\boldsymbol{v}_{m+1} \in \mathbb{R}^N$ satisfies $V_m^T \boldsymbol{v}_{m+1} = \boldsymbol{0}$, and $T_m \in \mathbb{R}^{m\times m}$ is a symmetric positive definite tridiagonal matrix; the scalar β_m is nonnegative. We assume that $1 \leq m \ll N$ is

chosen small enough so that the decomposition (5.5) with the stated properties exists. This is the generic situation.

It follows from the recursion relation (5.5) for the columns of V_m that the jth column can be expressed as

$$\boldsymbol{v}_j = p_{j-1}(A)\boldsymbol{v}, \quad j = 1, 2, \ldots, m,$$

where p_{j-1} is a polynomial of degree $j-1$. Golub and Meurant [29] show that the expression

$$\mathcal{G}_m(f) := \boldsymbol{e}_1^T f(T_m)\boldsymbol{e}_1 \tag{5.6}$$

is the m-point Gauss quadrature rule for the approximation of (5.1), i.e.,

$$\mathcal{G}_m(f) = \boldsymbol{v}^T f(A)\boldsymbol{v}, \qquad \forall f \in \mathbb{P}_{2m},$$

where $\mathbb{P}_{2m}$ denotes the space of all polynomials of degree at most $2m-1$, i.e., a space of dimension $2m$. This Gauss rule is associated with the bilinear form

$$\langle f, g\rangle := (f(A)\boldsymbol{v})^T(g(A)\boldsymbol{v}), \tag{5.7}$$

which is an inner product for polynomials f and g of sufficiently low degree. This can be seen by substituting the spectral factorization (3.4) into (5.7). Substituting the spectral factorization of the matrix T_m into the right-hand side of (5.6) shows that $\mathcal{G}_m(f)$ indeed is a quadrature rule with m nodes.

When the integrand f has one or several singularities close to the support of the measure $d\nu$ in (5.4), Gauss rules (5.6) with a moderate number of nodes, m, may yield poor approximations of the functional (5.1). This difficulty can be remedied by using rational Gauss rules. They were first discussed by Gonchar and López Lagomasino [31], and have subsequently received considerable attention; see, e.g., [8, 22, 40, 42].

We will use the rational Gauss quadrature rules described in [56] and define associated rational Gauss–Radau rules, as well as rational anti-Gauss and simplified anti-Gauss rules. The computation of the rational Gauss rules described in [56] is based on the observation in [58] that a sequence of certain orthogonal rational functions satisfy short recursion relations, i.e., the number of terms in the recursion relations can be bounded independently of the number of orthogonal rational functions in the sequence. We note that different sequences of orthogonal rational functions that satisfy a

three-term recursion relation have been developed by Deckers and Bultheel [13]. These sequences also can be used to construct rational Gauss rules; see [14]. We use the rational Gauss quadrature rules described in [56], because their computation requires fewer linear systems of equations with matrices that are determined from A to be solved than the approach in [14].

5.4 Recursion Relations for Rational Krylov Subspaces

The first part of this section reviews results in [58] on recursion relations for certain orthogonal rational functions. The number of terms in the recursion relations depends on the number of distinct poles of the rational functions and on the ordering of certain elementary rational basis functions. The recursion relations are applied in a rational Lanczos process, which is described in Subsection 5.4.1.

Introduce linear spaces of rational functions with finite poles,

$$\mathbb{Q}_{i,k_i} = \text{span}\left\{ \frac{1}{(y-\alpha_i)^j} : \ j = 1, 2, \ldots, k_i \right\}, \quad i = 1, 2, \ldots, \ell, \tag{5.8}$$

where the α_i are real distinct poles of multiplicity $k_i, i = 1, 2, \ldots, \ell$. They are assumed to lie outside the convex hull of the support of the measure $d\nu$. In the applications of this chapter, they will live on the negative real axis.

Now let $\alpha_{\ell+2i}$, $i = 1, 2, \ldots, \hat{\ell}$, denote distinct complex conjugate poles with nonvanishing imaginary part, and assume that each pole and its complex conjugate are adjacent, i.e., $\alpha_{\ell+2i} = \bar{\alpha}_{\ell+2i-1}$, where the bar denotes complex conjugation. Since we are interested in integrating functions f, such that $f(y) \in \mathbb{R}$ for $y > 0$, we replace each pair of rational functions

$$\frac{1}{(y-\alpha_{\ell+2i-1})^j}, \quad \frac{1}{(y-\bar{\alpha}_{\ell+2i-1})^j}$$

by a pair

$$\frac{1}{(y^2+\beta_i y+\gamma_i)^j}, \quad \frac{y}{(y^2+\beta_i y+\gamma_i)^j},$$

where the coefficients $\beta_i, \gamma_i \in \mathbb{R}$ are defined by $y^2+\beta_i y+\gamma_i = (y-\alpha_{\ell+2i-1})(y-\bar{\alpha}_{\ell+2i-1})$. Analogously to (5.8), we define the spaces

$$\mathbb{W}_{i,2s_i} = \text{span}\left\{ \frac{1}{(y^2+\beta_i y+\gamma_i)^j}, \frac{y}{(y^2+\beta_i y+\gamma_i)^j} : \ j = 1, 2, \ldots, s_i \right\}, \quad i = 1, 2, \ldots, \hat{\ell},$$

where s_i is the multiplicity of the complex conjugate poles.

Let

$$k = \sum_{i=1}^{\ell} k_i, \quad s = \sum_{i=1}^{\hat{\ell}} s_i, \tag{5.9}$$

and introduce the $(m+1)$-dimensional linear space

$$\mathbb{S}_{m+1} := \mathbb{P}_{m+1-k-2s} \oplus \mathbb{Q}_{1,k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,k_\ell} \oplus \mathbb{W}_{1,2s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},2s_{\hat{\ell}}}, \tag{5.10}$$

where we assume that the k_i and s_i are chosen so that $0 \le k + 2s \le m-1$. Then the space (5.10) contains linear functions. Let

$$\Psi_{m+1} = \{\psi_0, \psi_1, \ldots, \psi_m\} \tag{5.11}$$

denote an elementary basis for the space $\mathbb{S}_{m+1}$, i.e., $\psi_0(y) = 1$ and each basis function $\psi_i(y)$, for $i = 1, 2, \ldots, m$, is one of the functions

$$y^j, \ \frac{1}{(y-\alpha_i)^j}, \ \frac{1}{(y^2+\beta_i y+\gamma_i)^j}, \ \frac{y}{(y^2+\beta_i y+\gamma_i)^j}$$

for some positive integers i and j. The notation $\psi_s \prec \psi_t$ indicates that the basis function ψ_s comes before ψ_t. We say that the ordering of the basis functions (5.11) is *natural* if $\psi_0(y) = 1$ and the remaining functions ψ_j, $j = 1, 2, \ldots, m$, satisfy:

1. $y^j \prec y^{j+1}$ for all integers $j > 0$,
2. $\frac{1}{(y-\alpha_i)^j} \prec \frac{1}{(y-\alpha_i)^{j+1}}$ for all integers $j > 0$ and every real pole α_i,
3. $\frac{1}{(y^2+\beta_i y+\gamma_i)^j} \prec \frac{y}{(y^2+\beta_i y+\gamma_i)^j} \prec \frac{1}{(y^2+\beta_i y+\gamma_i)^{j+1}}$ for all positive integers j and every pair $\{\beta_i, \gamma_i\}$,
4. if $\psi_j(y) = \frac{1}{(y^2+\beta_j y+\gamma_j)^p}$, then $\psi_{j+1}(y) = \frac{y}{(y^2+\beta_j y+\gamma_j)^p}$.

It is shown in [58] that a Stieltjes procedure for orthogonalizing elementary basis functions with respect to some inner product [1] on $\mathbb{S}_{m+1}$ only requires short recursion relations when the basis functions are in natural order. We mention three types of recursion relations from [58] that are needed in this chapter. We start with recursions for $y\phi_r$:

$$y\phi_r(y) = \sum_{j=-n_1}^{n_2} c_{r,r+j}\phi_{r+j}(y), \quad r = 0, 1, \ldots \ , \tag{5.12}$$

[1] We can also orthogonalize the elementary basis with respect to a bilinear form $[f, g] = \mathcal{L}(fg)$, where $\mathcal{L}$ is a linear functional on $\mathbb{S}_{m+1}$, as long as $\mathcal{L}(\phi_j^2) \neq 0$.

where $r - n_1$ is the largest integer smaller than r such that ψ_{r-n_1} is a monomial if there is such a monomial (otherwise $r - n_1 = 0$), and $r + n_2$ is the smallest integer larger than r such that ψ_{r+n_2} is a monomial.

In order to introduce a new real pole α_i we need the following recursions:

$$\frac{1}{y - \alpha_i}\phi_r(y) = \sum_{j=-n_3}^{n_4} c_{r,r+j}^{(i)} \phi_{r+j}(y), \quad r = 0, 1, 2, \dots \ , \tag{5.13}$$

where $r - n_3$ is the largest integer smaller than r such that ψ_{r-n_3} is a rational function with a pole at α_i if there is such a rational function (otherwise $r - n_3 = 0$), and $r + n_4$ is the smallest integer larger than r such that ψ_{r+n_4} is a rational function with a pole at α_i.

To introduce a new pair of complex conjugate poles, we use the the following formulas with $p = 0, 1$:

$$\frac{y^p}{y^2 + \beta_j y + \gamma_j}\phi_r(y) = \sum_{i=-n_5}^{n_6} c_{r,r+i}^{(j)} \phi_{r+i}(y), \quad r = 0, 1, 2, \dots \ , \tag{5.14}$$

where $r - n_5$ is the largest integer smaller than r such that $\psi_{r-n_5}(y) = (y^2 + \beta_j y + \gamma_j)^s$ if there is such a rational function (otherwise $r - n_5 = 0$), and $r + n_6$ is the smallest integer larger than r such that $\psi_{r+n_6}(y) = y(y^2 + \beta_j y + \gamma_j)^s$.

Define the vector of orthonormal rational functions,

$$\Phi_m(y) := [\phi_0(y), \phi_1(y), \dots, \phi_{m-1}(y)].$$

If ψ_m is a monomial and $m - d$ is the largest integer smaller than m such that ψ_{m-d} is a monomial, then the recursion formulas (5.12) can be written in the matrix form

$$y\Phi_m(y) = H_m\Phi_m(y) + \sum_{j=1}^{d} h_{m-j,m}\phi_m(y)\boldsymbol{e}_{m+1-j}. \tag{5.15}$$

The matrix H_m has the following block-diagonal structure: it has $m - k - 2s - 1$ square blocks along the diagonal such that any two consecutive blocks overlap in one diagonal element. The jth block of H_m is of dimension $r \times r$, where $r - 2$ is the number of rational functions between consecutive monomials y^{j-1} and y^j. More precisely, the jth block of H_m is the submatrix $H_m\,(r_1 : r_2, r_1 : r_2)$ (in MATLAB notation) with the entries h_{ij} for $r_1 \le i, j \le r_2$, where $\psi_{r_1}(y) = y^{j-1}$ and $\psi_{r_2}(y) = y^j$. Also, $r = r_2 - r_1 + 1$. The non-zero entries of $H_m = [h_{ij}]_{i,j=0}^{m-1}$ are recursion coefficients for the functions $\phi_i, i = 0, 1, \dots, m - 1$. They satisfy

$$h_{i,j} = (y\phi_i(y), \phi_j(y)) = (y\phi_j(y), \phi_i(y)) = h_{j,i}, \tag{5.16}$$

which shows that the matrix H_m depends only on the first m elementary basis functions ψ_j, $j = 0, \ldots, m-1$, and does not depend on ψ_m. In this chapter we will always assume that ψ_m (or ψ_{m+1} if we deal with H_{m+1}) is a monomial only for one reason: the zeros of ϕ_m are eigenvalues of H_m.

The following example illustrates the structure of the matrix H_m.

Example 5.4.1. Let $\alpha_i, \beta_j, \gamma_j$ be defined as described in the beginning of this section and consider the elementary basis

$$\left\{1, \frac{1}{y-\alpha_1}, y, \frac{1}{(y-\alpha_1)^2}, y^2, \ldots, \frac{1}{y-\alpha_\ell}, y^{k-k_\ell+1}, \ldots, \frac{1}{(y-\alpha_\ell)^{k_\ell}}, y^k, \frac{1}{y^2+\beta_1 y+\gamma_1}, \right.$$
$$\frac{y}{y^2+\beta_1 y+\gamma_1}, y^{k+1}, \ldots, \frac{1}{(y^2+\beta_1 y+\gamma_1)^{s_1}}, \frac{y}{(y^2+\beta_1 y+\gamma_1)^{s_1}}, y^{k+s_1}, \frac{1}{y^2+\beta_2 y+\gamma_2}, $$
$$\left. \frac{y}{y^2+\beta_2 y+\gamma_2}, y^{k+s_1+1}, \ldots, \frac{1}{(y^2+\beta_{\hat{\ell}} y+\gamma_{\hat{\ell}})^{s_{\hat{\ell}}}}, \frac{y}{(y^2+\beta_{\hat{\ell}} y+\gamma_{\hat{\ell}})^{s_{\hat{\ell}}}}, y^{k+s} \right\}.$$

This basis together with the function $\psi_m(y) = y^{k+s+1}$, where k and s are defined by (5.9), and $m = 1 + 2k + 3s$, satisfy the requirements of natural ordering and make up the space $\mathbb{S}_{m+1}$. The matrix H_m in this case has k 3×3 blocks and s trailing 4×4 blocks along the diagonal:

$$H_m = \begin{bmatrix} * & * & * & & & & & & \\ * & * & * & & & & & & \\ * & * & * & * & * & & & & \\ & & * & * & * & & & & \\ & & * & * & * & & & & \\ & & & & \ddots & & & & \\ & & & & * & * & * & * & \\ & & & & * & * & * & * & \\ & & & & * & * & * & * & \\ & & & & * & * & * & * & * & * & * \\ & & & & & & & * & * & * & * \\ & & & & & & & * & * & * & * \\ & & & & & & & * & * & * & * \end{bmatrix}.$$

Matrix entries that may be nonvanishing are marked by "$*$". If we place one more rational function between two consecutive monomials y^{j-1} and y^j in the elementary basis, then the size of the jth block of H_m increases by one. Similarly, removing one rational function between two consecutive monomials would decrease the size of the corresponding block by one.

□

5.4.1 The Rational Lanczos Algorithm

In the rest of this chapter we focus on functionals $\mathcal{I}$ defined by (5.1) when A is a symmetric positive definite matrix, and $\boldsymbol{v}$ is a normalized vector. If the functions ϕ_j, $j = 0, \ldots, m$, are orthonormal

with respect to the bilinear form

$$(f, g) := \mathcal{I}(fg) = \boldsymbol{v}^T f(A)g(A)\boldsymbol{v},$$

then the vectors

$$\{\boldsymbol{v}_0 = \phi_0(A)\boldsymbol{v}, \ldots, \boldsymbol{v}_j = \phi_j(A)\boldsymbol{v}\} \tag{5.17}$$

form an orthonormal basis for the rational Krylov subspace

$$\mathbb{K}_{j+1}(A, \boldsymbol{v}) = \text{span}\{\psi_0(A)\boldsymbol{v}, \psi_1(A)\boldsymbol{v}, \ldots, \psi_j(A)\boldsymbol{v}\}, \tag{5.18}$$

for $j = 0, 1, \ldots, m$. Indeed,

$$\boldsymbol{v}_j^T \boldsymbol{v}_i = (\phi_j(A)\boldsymbol{v})^T(\phi_i(A)\boldsymbol{v}) = \boldsymbol{v}^T\phi_j(A)\phi_i(A)\boldsymbol{v} = (\phi_j, \phi_i).$$

The vectors $\boldsymbol{v}_j$ satisfy the same recursion relations as rational functions ϕ_j and can be constructed by the rational Lanczos process. This process is analogous to the Stieltjes-type procedure [58, Algorithm 3.1] for computing an orthonormal basis for the space $\mathbb{S}_{m+1}$ which is based on the recursion relations (5.12), (5.13) and (5.14).

The implementation of the rational Lanczos process requires the solution of linear systems of equations with matrices of the forms $A - \alpha_i I$ and $A^2 + \beta_i A + \gamma_i I$, where $\alpha_i, \beta_i, \gamma_i$ are suitable real scalars. The norm $\|\cdot\|$ in Algorithm 4 denotes the Euclidean vector norm.

Algorithm 4 Orthonormalization process, Part 1.

Input: $\boldsymbol{v} \in \mathbb{R}^N \backslash \{\boldsymbol{0}\}$, and functions for evaluating matrix-vector products with A and for solving linear systems of equations with matrices of the form $A - \alpha_\ell I$ and $A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I$. Thus, we do not explicitly form the matrices A^j, $(A - \alpha_\ell I)^{-j}$, and $(A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-j}$.

Output: Orthonormal basis $\{\boldsymbol{v}_r\}_{r=0}^m$.

Initialization: $\boldsymbol{v}_0 := \boldsymbol{v}/\|\boldsymbol{v}\|; r := 1;$

while $r \leq m$ **do**

  if $\psi_r = A^j$ for some $j \in \mathbb{N}$ **then**

    $\boldsymbol{u} := A\boldsymbol{v}_{r-1};$

    for $i = \hat{r} : r - 1$ **do**

      $c_{r-1,i} := \boldsymbol{v}_i^T \boldsymbol{u};\ \boldsymbol{u} := \boldsymbol{u} - c_{r-1,i}\boldsymbol{v}_i;$

    end for

Algorithm 4 Orthonormalization process, Part 2.

$\delta_r := \|\boldsymbol{u}\|$; $\boldsymbol{v}_r := \boldsymbol{u}/\delta_r$;

else if $\psi_r = (A - \alpha_\ell I)^{-j}$ for some $j \in \mathbb{N}$ **then**

$\boldsymbol{u} := (A - \alpha_\ell I)^{-1}\boldsymbol{v}_{r-1}$;

for $i = \hat{r} : r-1$ **do**

$c_{r-1,i} := \boldsymbol{v}_i^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - c_{r-1,i}\boldsymbol{v}_i$;

end for

$\delta_r := \|\boldsymbol{u}\|$; $\boldsymbol{v}_r := \boldsymbol{u}/\delta_r$;

else if $\psi_r = (A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-j}$ for some $j \in \mathbb{N}$ **then**

$\boldsymbol{u} := (A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-1}\boldsymbol{v}_{r-1}$;

$\boldsymbol{w} := A(A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-1}\boldsymbol{v}_{r-1}$;

for $i = \hat{r} : r-1$ **do**

$c_{r-1,i} := \boldsymbol{v}_i^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - c_{r-1,i}\boldsymbol{v}_i$;

$d_{r-1,i} := \boldsymbol{v}_i^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - d_{r-1,i}\boldsymbol{v}_i$;

end for

$\delta_r := \|\boldsymbol{u}\|$; $\boldsymbol{v}_r := \boldsymbol{u}/\delta_r$;

$\eta := \boldsymbol{v}_r^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \eta\boldsymbol{v}_r$;

$\delta'_r := \|\boldsymbol{w}\|$; $\boldsymbol{v}_{r+1} := \boldsymbol{w}/\delta'_r$;

end if

$r := r + 1$;

with the exception of $r := r + 2$ when pairs of complex conjugate

poles are evaluated

end while

We assume in Algorithm 4 that the basis (5.11) satisfies conditions 1-4 of natural ordering. The value of $\hat{r}$ is such that $\psi_{\hat{r}}$ is a basis function having the same pole(s) as ψ_r but of order one less (of order two less if $\hat{r} = r + 1$) if there are such rational basis function; otherwise $\hat{r} = 0$. Performing m steps of Algorithm 4 yields an orthonormal basis $\{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ for the rational Krylov subspace $\mathbb{K}_{m+1}(A, \boldsymbol{v})$. The matrix

$$V_m = [\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m-1}] \in \mathbb{R}^{N\times m}, \quad \text{with } \boldsymbol{v}_0 = \boldsymbol{v},$$

and the symmetric matrix

$$H_m = [h_{i,j}]_{i,j=0}^{m-1} \in \mathbb{R}^{m\times m}, \quad h_{i,j} = \mathcal{I}(y\phi_i\phi_j) = \boldsymbol{v}_i^T A \boldsymbol{v}_j,$$

satisfy the decomposition

$$AV_m = V_m H_m + \sum_{j=1}^{d} h_{m-j,m} \boldsymbol{v}_m \boldsymbol{e}_{m+1-j}^T, \tag{5.19}$$

where $\boldsymbol{v}_m \in \mathbb{R}^N$ is such that $V_m^T \boldsymbol{v}_m = \boldsymbol{0}$. The orthogonal projection of A onto the rational Krylov subspace (5.18) is given by

$$H_m = V_m^T A V_m.$$

We tacitly assume that m is small enough so that the decomposition (5.19) with the stated properties exists. This is the generic situation. Algorithm 4 breaks down before m steps have been carried out if $\delta_r = 0$ or $\delta_r^{'} = 0$. Ramification of breakdown is out of scope of this chapter.

5.5 Application to Rational Gauss Quadrature

This section discusses rational Gauss quadrature rules for the approximation of functionals (5.1) when f is a Stieltjes function (5.2) and A is a symmetric positive definite matrix. Then

$$F(A) = \boldsymbol{v}^T f(A) \boldsymbol{v} = \int_0^\infty \boldsymbol{v}^T (tI + A)^{-1} \boldsymbol{v} \, d\mu(t) = \mathcal{I}(f). \tag{5.20}$$

It is shown in [56] that the expression

$$\widehat{\mathcal{G}}_m(f) := \boldsymbol{e}_1^T f(H_m) \boldsymbol{e}_1$$

is a rational Gauss quadrature rule for the approximation of the expression (5.20). It is characterized by the property

$$\widehat{\mathcal{G}}_m(f) = \mathcal{I}(f), \quad \forall f \in \mathbb{S}_{2m}, \tag{5.21}$$

where

$$\mathbb{S}_{2m} := \mathbb{P}_{2m-2k-4s} \oplus \mathbb{Q}_{1,2k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,2k_\ell} \oplus \mathbb{W}_{1,4s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},4s_{\hat{\ell}}} = \frac{\mathbb{P}_{2m}}{w^2}$$

denotes a $2m$-dimensional linear space of certain rational functions and

$$w(y) = \prod_{i=1}^{\ell} (y - \alpha_i)^{k_i} \prod_{j=1}^{\hat{\ell}} \left(y^2 + \beta_j y + \gamma_j\right)^{s_j} \tag{5.22}$$

is a polynomial of degree $k+2s$ with k and s defined by (5.9). The scalars $\alpha_i, \beta_j, \gamma_j$ are determined as described in the beginning of Section 5.4. In view of (5.4), this rational Gauss quadrature rule can be written as

$$\widehat{\mathcal{G}}_m(f) = \int_0^\infty \boldsymbol{e}_1^T (tI + H_m)^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(H_m) \boldsymbol{e}_1, \tag{5.23}$$

The orthonormal rational function ϕ_m has m distinct zeros $\{y_i\}_{i=1}^m$ that lie in the convex hull of the support of the measure $d\nu$, and they are the eigenvalues of H_m in (5.15); see [56, Theorem 2.5]. Recall that we assume ψ_m to be a monomial. Thus

$$\phi_m(y) = c_m \frac{\prod_{i=1}^m (y - y_i)}{w(y)} \in \mathbb{S}_{m+1}, \tag{5.24}$$

where c_m is a constant.

The remainder term for rational Gauss rules (5.23) can be derived by considering rational Hermite interpolation. The following result is shown similarly as [42, Theorem 5.4]. The Lagrange fundamental functions associated with the function (5.24) are defined by

$$l_i(y) := \frac{\phi_m(y)}{\phi_m'(y_i)(y - y_i)}, \quad i = 1, 2, \ldots, m,$$

and satisfy

$$l_i(y_j) = \begin{cases} 1 & \text{if} \quad j = i, \\ 0 & \text{if} \quad j \neq i. \end{cases}$$

Introduce the rational Hermite interpolation function

$$\hat{L}(y) := \sum_{i=1}^m \left(\hat{l}_i(y) f(y_i) + \tilde{l}_i(y) f'(y_i) \right), \tag{5.25}$$

where

$$\hat{l}_i(y) = \left[1 - 2(y - y_i) l_i'(y_i)\right] l_i^2(y), \quad \tilde{l}_i(y) = (y - y_i) l_i^2(y).$$

Then the function $\hat{L}(y)$ satisfies the interpolation conditions

$$\hat{L}(y_i) = f(y_i), \quad \hat{L}'(y_i) = f'(y_i), \quad i = 1, 2, \ldots, m.$$

Theorem 12. *Let $\hat{L}$ be the rational Hermite function* (5.25) *determined by the interpolation nodes $y_1 < y_2 < \ldots < y_m$, which are the m zeros of $\phi_m \in \mathbb{S}_{m+1}$. Assume that f is $2m$ times continuously differentiable in the interval between the nodes y_1 and y_m. Then for some scalar $c = c(y)$ depending on y in this interval, we have*

$$f(y) = \hat{L}(y) + \frac{d^{2m}}{dt^{2m}} \left(w(t)^2 f(t) \right)_{t=c} \frac{\phi_m^2(y)}{c_m^2 (2m)!}. \tag{5.26}$$

Proof. *If $y = y_i$ for some $i = 1, 2, \ldots, m$, then*

$$f(y) = \hat{L}(y),$$

in (5.26)*. Thus, the error vanishes. The result holds for an arbitrary constant c. Now, assume that $y \neq y_i$ for all i. In this case, we consider*

$$g(t) = w^2(t)(f(t) - \hat{L}(t)) - w^2(y)(f(y) - \hat{L}(y)) \prod_{i=1}^{m} \frac{(t - y_i)^2}{(y - y_i)^2},$$

where $w^2(x)\hat{L}(x)$ is a polynomial of degree $2m - 1$. The function g is $2m$ times continuously differentiable and has $2m + 1$ zeros. By Rolle's theorem, there exists a scalar $c = c(y)$ depending on y in the interval between the nodes y_1 and y_m, for which

$$\begin{aligned} 0 &= \frac{d^{2m}}{dt^{2m}}(g(t))_{t=c} \\ &= \frac{d^{2m}}{dt^{2m}}\left(w(t)^2 f(t)\right)_{t=c} - w^2(y)(f(y) - \hat{L}(y)) \frac{d^{2m}}{dt^{2m}} \left(\prod_{i=1}^{m} \frac{(t - y_i)^2}{(y - y_i)^2} \right)_{t=c} \\ &= \frac{d^{2m}}{dt^{2m}}\left(w(t)^2 f(t)\right)_{t=c} - w^2(y)(f(y) - \hat{L}(y)) \frac{(2m)!}{\prod_{i=1}^{m}(y - y_i)^2}. \end{aligned}$$

Rearranging terms, and dividing by w^2 yields (5.26)*.*

Theorem 12 can be used to construct an error term for the m-point rational Gauss quadrature rule $\widehat{\mathcal{G}}_m$. The following results can be shown similarly as [42, Corollary 5.5].

Corollary 13. *Assume that f is $2m$ times continuously differentiable in the convex hull of the spectrum of A. Then, the reminder term for the rational Gauss rule* (5.23) *can be expressed as*

$$\mathcal{E}_m(f) := \mathcal{I}(f) - \widehat{\mathcal{G}}_m(f) = \frac{d^{2m}}{dy^{2m}}(w^2(y)f(y))_{y=c} \cdot \frac{1}{c_m^2(2m)!} \int_0^\infty \frac{\prod_{j=1}^{m}(y - y_j)^2}{w^2(y)} d\nu(y) \tag{5.27}$$

for some scalar c in an interval that contains the spectrum of A.

Proof. *Theorem 12 shows that the rational Hermite interpolation function $\hat{L}$ lives in $\mathbb{S}_{2m}$. It now follows from* (5.21) *that*

$$\widehat{\mathcal{G}}_m(\hat{L}) = \mathcal{I}(\hat{L}),$$

and we obtain

$$(\mathcal{I} - \widehat{\mathcal{G}}_m)(f) = \mathcal{I}(f - \hat{L}) = \frac{d^{2m}}{dt^{2m}}\left(w(t)^2 f(t)\right)_{t=c} \frac{\mathcal{I}(\phi_m^2(y))}{c_m^2(2m)!}.$$

This shows (5.27)*.*

Assume that f satisfies the conditions of Corollary 13, and that

$$\frac{d^{2m}}{dy^{2m}}\left(w^2(y)f(y)\right) \geq 0, \tag{5.28}$$

in some open interval containing the spectrum of A. Then Corollary 13 gives a lower bound for $\mathcal{I}(f)$. We have

$$\widehat{\mathcal{G}}_m(f) \leq \mathcal{I}(f).$$

5.5.1 Rational Gauss–Radau Quadrature Rules

This subsection discusses the computation of rational Gauss–Radau rules and error bounds that can be determined with these rules. This approach of bracketing (5.1) is a rational analogue of the technique advocated by Golub and Meurant [29] for computing upper and lower bounds for (5.1) by evaluating pairs of (standard) Gauss and Gauss–Radau quadrature rules.

The $(m+1)$-point rational Gauss–Radau quadrature rule with a prescribed node θ can be expressed as

$$\widehat{\mathcal{R}}_{m+1}^{\theta}(f) = \int_0^{\infty} \boldsymbol{e}_1^T (tI + H_{m+1}^{\theta})^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(H_{m+1}^{\theta}) \boldsymbol{e}_1 \tag{5.29}$$

for a suitable matrix $H_{m+1}^{\theta} \in \mathbb{R}^{(m+1)\times(m+1)}$. This quadrature rule is characterized by the property

$$\widehat{\mathcal{R}}_{m+1}^{\theta}(f) = \mathcal{I}(f), \quad \forall f \in \mathbb{S}_{2m+1},$$

where

$$\mathbb{S}_{2m+1} := \mathbb{P}_{2m+1-2k-4s} \oplus \mathbb{Q}_{1,2k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,2k_\ell} \oplus \mathbb{W}_{1,4s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},4s_{\hat{\ell}}} = \frac{\mathbb{P}_{2m+1}}{w^2}.$$

The standard $(m+1)$-point Gauss–Radau rule for the approximation of (5.1) can be determined by modifying the tridiagonal matrix T_{m+1} in (5.5). Analogously, we will show that the rational Gauss–Radau rule (5.29) can be determined by modifying the last diagonal entry of the matrix H_{m+1}.

Note that in our definition of the rational $(m+1)$-point Gauss–Radau rule we assume exactness on the space $\frac{\mathbb{P}_{2m+1}}{w^2}$ with the same w as in definition of $\widehat{\mathcal{G}}_m$. In other words, we do not introduce a new finite pole. Thus, we can introduce the finite poles in the same order as in the construction of $\widehat{\mathcal{G}}_m$: take ψ_m to be a monomial, and determine the matrix H_{m+1} such that

$$y\Phi_{m+1}(y) = H_{m+1}\Phi_{m+1}(y) + \varphi_{m+1}(y)\boldsymbol{e}_{m+1}, \tag{5.30}$$

and

$$\boldsymbol{e}_1^T f(H_{m+1})\boldsymbol{e}_1 = \mathcal{I}(f), \quad \forall f \in \frac{\mathbb{P}_{2m+2}}{w^2}.$$

We will show that the rational Gauss–Radau rule (5.29) with a fixed node $\theta \le \lambda_1$ can be determined by replacing the last diagonal entry $h_{m,m}$ of H_{m+1} by $h^\theta_{m,m}$ so that the resulting matrix

$$H^\theta_{m+1} = \begin{bmatrix} H_m & \boldsymbol{w}_m \\ \boldsymbol{w}_m^T & h^\theta_{m,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \quad \boldsymbol{w}_m = [0,\dots,0,h_{m,m-d},\dots,h_{m,m-1}]^T \in \mathbb{R}^m,$$

has an eigenvalue at $\theta \le \lambda_1$. Only d trailing entries of the vector $\boldsymbol{w}_m$ might be nonvanishing, where d is the same as in (5.15). We use $\varphi_{m+1} = h_{m,m+1}\phi_{m+1}$ instead of ϕ_{m+1} for simplicity. The last equation in (5.30) can be written in the form

$$\varphi_{m+1}(y) = (y - h_{m,m})\phi_m(y) - \sum_{i=m-d}^{m-1} h_{m,i}\phi_i(y).$$

First we show that

$$\boldsymbol{e}_1^T f(H^\theta_{m+1})\boldsymbol{e}_1 = \mathcal{I}(f), \quad \forall f \in \frac{\mathbb{P}_{2m+1}}{w^2}, \tag{5.31}$$

for any $h^\theta_{m,m}$. Replacing $h_{m,m}$ in the matrix H_{m+1} by $h_{m,m} + c$, we obtain the recursion coefficients for the rational functions $\phi_0, \phi_1, \dots, \phi_m, \tilde{\varphi}_{m+1}$, where

$$\tilde{\varphi}_{m+1} = \varphi_{m+1} - c\phi_m.$$

We see that $\tilde{\varphi}_{m+1}$ is orthogonal to the space $\frac{\mathbb{P}_m}{w} = \operatorname{span}\{\phi_0,\dots,\phi_{m-1}\}$ with respect to the bilinear form defined by the integral $\mathcal{I}$. Then we can construct the new functional $\tilde{\mathcal{I}}$ on the space $\frac{\mathbb{P}_{2m+2}}{w^2}$ in the following way: $\tilde{\mathcal{I}}(f) = \mathcal{I}(f)$ for $f \in \frac{\mathbb{P}_{2m+1}}{w^2}$, and $\tilde{\mathcal{I}}$ on $\frac{\mathbb{P}_{2m+2}}{w^2} \setminus \frac{\mathbb{P}_{2m+1}}{w^2}$ is such that $\tilde{\mathcal{I}}(\tilde{\varphi}_{m+1}\phi_m) = 0$. Thus $\phi_0, \phi_1, \dots, \phi_m, \tilde{\varphi}_{m+1}$ is the sequence of orthogonal rational functions with respect to $\tilde{\mathcal{I}}$, and

$$\boldsymbol{e}_1^T f(H^\theta_{m+1})\boldsymbol{e}_1 = \tilde{\mathcal{I}}(f), \quad \forall f \in \frac{\mathbb{P}_{2m+2}}{w^2}.$$

The formula (5.31) follows from the fact that $\tilde{\mathcal{I}}(f) = \mathcal{I}(f)$ for $f \in \frac{\mathbb{P}_{2m+1}}{w^2}$.

We finish the construction of the rational Gauss–Radau quadrature by choosing c so that $\tilde{\varphi}_{m+1}(\theta) = 0$. Thus, we get

$$c = \frac{\varphi_{m+1}(\theta)}{\phi_m(\theta)}.$$

The rational function $\tilde{\varphi}_{m+1}$ can be written in the form

$$\tilde{\varphi}_{m+1}(y) = a\frac{(y-\theta)q_m(y)}{w(y)},$$

where a is a constant and q_m is a polynomial of degree m. Since $\tilde{\varphi}_{m+1}$ is orthogonal to the space $\frac{\mathbb{P}_m}{w}$, we see that $\frac{q_m}{w}$ is orthogonal to the space $\frac{\mathbb{P}_m}{w}$ with respect to the integral $I(f) = \mathcal{I}((y-\theta)f)$ with nonnegative measure $(y-\theta)d\nu(y)$. Thus, we conclude that the rational Gauss-Radau quadrature $\widehat{\mathcal{G}}_{m+1}$ has a node at θ and m distinct nodes in the open interval that contains the spectrum of A.

Similarly, we may define rational Gauss–Radau rules with a fixed node $\theta \geq \lambda_N$. The following result is shown in the same manner as [42, Theorem 7.1].

Theorem 14. *Assume that f is $2m+1$ times continuously differentiable in the convex hull of the spectrum of A and θ. Then the remainder term for the rational Gauss–Radau rule* (5.29) *can be written as*

$$\begin{aligned}\mathcal{E}_{m+1}(f) :=& \mathcal{I}(f) - \widehat{\mathcal{R}}^{\theta}_{m+1}(f) \\ =& \frac{d^{2m+1}}{dy^{2m+1}}(w^2(y)f(y))_{y=c} \cdot \frac{1}{c_m^2(2m+1)!}\int_0^{\infty}(y-\theta)\frac{\prod_{j=1}^m (y-y_j)^2}{w^2(y)}d\nu(y),\end{aligned} \tag{5.32}$$

where the scalar c lives in an interval that contains the spectrum of A and θ.

Proof. *Let the function ϕ_m in Theorem 12 have an additional node θ and assume that the rational Hermite interpolation function $\hat{L}^{\theta} \in \mathbb{S}_{2m+1}$ also interpolates f at $y_1, y_2, \ldots, y_m$ and θ. Then by modifying the error term* (5.26) *and applying Corollary 13, we obtain* (5.32).

Assume that f satisfies the conditions of Theorem 14, and that

$$\frac{d^{2m+1}}{dy^{2m+1}}\left(w^2(y)f(y)\right) \geq 0 \tag{5.33}$$

in some open interval containing the spectrum of A and θ. Since

$$\begin{aligned}\int_0^{\infty}(y-\lambda_1)\frac{\prod_{j=1}^m (y-y_j)^2}{w^2(y)}d\nu(y) \geq 0, \\ \int_0^{\infty}(y-\lambda_N)\frac{\prod_{j=1}^m (y-y_j)^2}{w^2(y)}d\nu(y) \leq 0,\end{aligned} \tag{5.34}$$

the remainder terms for the rational Gauss–Radau rule with a prescribed node $\theta \leq \lambda_1$ or $\theta \geq \lambda_N$ are of opposite sign. It follows that

$$\widehat{\mathcal{R}}^{\lambda_1}_{m+1}(f) \leq \mathcal{I}(f) \leq \widehat{\mathcal{R}}^{\lambda_N}_{m+1}(f). \tag{5.35}$$

Analogously to (5.35), when the derivative $(w^2 f)^{(2m)}$ in (5.27) is of constant sign in an interval that contains the spectrum of A, and the derivative $(w^2 f)^{(2m+1)}$ in (5.32) is of constant sign in an

interval that contains the spectrum of A and θ, and the Radau point $\theta \leq \lambda_1$ or $\theta \geq \lambda_N$ is suitably chosen, the values $\widehat{\mathcal{G}}_m(f)$ and $\widehat{\mathcal{R}}_{m+1}^{\theta}(f)$ bracket (3.1). We tacitly assume here that θ is chosen so that $\widehat{\mathcal{R}}_{m+1}^{\theta}(f)$ can be evaluated.

It is known that for every Stieltjes functions f, we have for all nonnegative integers k and ℓ,

$$(-1)^k \frac{d^{k+\ell}}{dx^{k+\ell}}(x^\ell f(x)) \geq 0, \quad x > 0; \tag{5.36}$$

see, e.g., Sokal [65, Theorem 1] and references therein. Setting $\ell = 0$ shows that Stieltjes functions are completely monotonic. Formula (5.36) demonstrates that for certain simple polynomials (5.22), such as $w(x) = x$, the properties (5.28) and (5.33) hold. Then pairs of rational Gauss and rational Gauss–Radau rules, or pairs of rational Gauss–Radau rules, can be used to bracket (5.1); this is a consequence of Corollary 13 and Theorem 14. However, the bracketing cannot be guaranteed for all polynomials (5.22). Note that the computation of a Gauss–Radau rule requires knowledge of the location of the largest or smallest eigenvalue of A in order to allocate the Radau point.

5.5.2 Rational Anti-Gauss Quadrature Rules

When the derivatives $(w^2 f)^{(2m)}$ or $(w^2 f)^{(2m+1)}$ change sign on the convex hull of the spectrum of A, pairs of rational Gauss and rational Gauss–Radau quadrature rules are not guaranteed to bracket (5.1). In this case, estimates of upper and lower bounds for (5.1) can be determined by evaluating appropriate pairs of rational Gauss and anti-Gauss quadrature rules. An advantage of this approach is that the sign of derivatives of $(w^2 f)$ are allowed to change in an interval that contains the spectrum of A. Moreover, knowledge of the location of the largest or smallest eigenvalues of A is not required.

In this subsection, we will show that rational anti-Gauss rules can be computed analogously to the standard $(m+1)$-point anti-Gauss quadrature rule.

The $(m+1)$-point rational anti-Gauss rule $\widetilde{\mathcal{G}}_{m+1}$ associated with the functional $\mathcal{I}$ defined by (5.4) is determined by the requirement that

$$(\mathcal{I} - \widetilde{\mathcal{G}}_{m+1})(f) = -(\mathcal{I} - \widehat{\mathcal{G}}_m)(f), \quad \forall f \in \mathbb{S}_{2m+2}, \tag{5.37}$$

where

$$\mathbb{S}_{2m+2} := \mathbb{P}_{2m+2-2k-4s} \oplus \mathbb{Q}_{1,2k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,2k_\ell} \oplus \mathbb{W}_{1,4s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},4s_{\hat{\ell}}} = \frac{\mathbb{P}_{2m+2}}{w^2}.$$

We will show that $\widetilde{\mathcal{G}}_{m+1}(f)$ can be expressed as

$$\widetilde{\mathcal{G}}_{m+1}(f) = \int_0^\infty \boldsymbol{e}_1^T (tI + \widetilde{H}_{m+1})^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(\widetilde{H}_{m+1}) \boldsymbol{e}_1 \tag{5.38}$$

for a suitable matrix $\widetilde{H}_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$.

The relation (5.37) shows that $\widetilde{\mathcal{G}}_{m+1}$ is the $(m+1)$-point rational Gauss rule for the functional

$$\mathcal{J}(f) := (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f).$$

Introducing the poles in the same way as in construction of the rational Gauss-Radau rule gives the matrix $\widetilde{H}_{m+1}$ having the same block-diagonal structure as H_{m+1}.

Define, analogously to (5.17), the vectors

$$\tilde{\boldsymbol{v}}_j = \tilde{\phi}_j(A)\boldsymbol{v}, \quad j = 0, 1, \ldots, m+1,$$

where the $\tilde{\phi}_j$ are orthonormal rational functions with respect to the bilinear form

$$\{f, g\} := \mathcal{J}(fg), \tag{5.39}$$

i.e, $\{\tilde{\phi}_i, \tilde{\phi}_j\} = 0$ for $i \neq j$, and $\{\tilde{\phi}_j, \tilde{\phi}_j\} = 1$ for all j. These orthonormal functions satisfy recurrence relations of the form

$$y\tilde{\Phi}_{m+1}(y) = \widetilde{H}_{m+1}\tilde{\Phi}_{m+1}(y) + \tilde{\varphi}_{m+1}(y)\boldsymbol{e}_{m+1}, \tag{5.40}$$

where $\tilde{\Phi}_{m+1}(y) = [\tilde{\phi}_0, \tilde{\phi}_2, \ldots, \tilde{\phi}_m]^T$, and the remainder term $\tilde{\varphi}_{m+1}(y)\boldsymbol{e}_{m+1}$ in (5.40) is of the form $\tilde{h}_{m,m+1}\tilde{\phi}_{m+1}(y)\boldsymbol{e}_{m+1}$ since the elementary basis functions ψ_m and ψ_{m+1} are monomials. Analogously to (5.16), the matrix $\widetilde{H}_{m+1} = [\tilde{h}_{ij}]_{i,j=0}^m$ is determined by the coefficients of the recursion relation that express the bilinear form (5.39),

$$\tilde{h}_{i,j} = \{y\tilde{\phi}_i, \tilde{\phi}_j\} = \{y\tilde{\phi}_j, \tilde{\phi}_i\} = \tilde{h}_{j,i}. \tag{5.41}$$

It follows from (5.21) and (5.37) that for rational functions f and g such that $fg \in \mathbb{S}_{2m}$, we have

$$\{f, g\} = (f, g) = \mathcal{I}(fg).$$

These equalities show that

$$\tilde{h}_{i,j} = h_{i,j}, \quad i, j = 0, 1, \ldots, m-1.$$

Therefore, $\tilde{\phi}_j = \phi_j$ for $0 \le j < m$, and

$$\tilde{H}_{m+1} = \begin{bmatrix} H_m & \tilde{\boldsymbol{w}}_m \\ \tilde{\boldsymbol{w}}_m^T & \tilde{h}_{m,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \quad \tilde{\boldsymbol{w}}_m = \begin{bmatrix} 0, \ldots, 0, \tilde{h}_{m-d,m}, \ldots, \tilde{h}_{m-1,m} \end{bmatrix}^T \in \mathbb{R}^m.$$

Further, we get

$$h_{m-1,m}\phi_m(y) = (y - h_{m-1,m-1})\phi_{m-1}(y) - \sum_{j=m-d}^{m-2} h_{m-1,j}\phi_j(y) = \tilde{h}_{m-1,m}\tilde{\phi}_m(y).$$

We use the previous equality, (5.41), and the fact that $\widehat{\mathcal{G}}_m(\phi_m g) = 0$ for any function g to show that $\phi_m = \sqrt{2}\tilde{\phi}_m$. Indeed,

$$\tilde{h}_{m-1,m} = \mathcal{J}(y\tilde{\phi}_{m-1}\tilde{\phi}_m) = \frac{h_{m-1,m}}{\tilde{h}_{m-1,m}}\mathcal{J}(y\phi_{m-1}\phi_m) = \frac{h_{m-1,m}}{\tilde{h}_{m-1,m}}\, 2\,\mathcal{I}(y\phi_{m-1}\phi_m) = \frac{2h_{m-1,m}^2}{\tilde{h}_{m-1,m}}.$$

The same relation holds for other entries (if any) of $\boldsymbol{w}_m$:

$$\tilde{h}_{m-j,m} = \mathcal{J}(y\tilde{\phi}_{m-j}\tilde{\phi}_m) = \frac{1}{\sqrt{2}}\, 2\,\mathcal{I}(y\phi_{m-j}\phi_m) = \sqrt{2}h_{m-j,m}, \quad j = d, \ldots, 2.$$

Finally,

$$\tilde{h}_{m,m} = \mathcal{J}(y\tilde{\phi}_m\tilde{\phi}_m) = 2\mathcal{I}(y\frac{\phi_m}{\sqrt{2}}\frac{\phi_m}{\sqrt{2}}) = h_{m,m}.$$

Therefore, the matrix $\tilde{H}_{m+1}$ associated with the rational anti-Gauss rule $\tilde{\mathcal{G}}_{m+1}$ is given by

$$\widetilde{H}_{m+1} = \begin{bmatrix} H_m & \sqrt{2}\boldsymbol{w}_m \\ \sqrt{2}\boldsymbol{w}_m^T & h_{m,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \tag{5.42}$$

Analogously to formula (5.23), the rational anti-Gauss quadrature rule can be evaluated according to (5.38).

We are now in a position to provide some sufficient conditions for $\widehat{\mathcal{G}}_m(f)$ and $\widetilde{\mathcal{G}}_{m+1}(f)$ to bracket $\mathcal{I}(f)$. Assume that we can carry out N steps of the Algorithm 4 without breakdown. This yields an orthonormal basis $\{\boldsymbol{v}_j\}_{j=0}^{N-1}$ of $\mathbb{R}^N$, and an associated sequence of orthonormal rational function $\{\phi_j\}_{j=0}^{N-1}$ determined by (5.17).

Theorem 15. *Consider the expansion of the integrand*

$$f(y) = \sum_{j=0}^{N-1} \omega_j\phi_j(y), \quad y \in \lambda(A), \tag{5.43}$$

in terms of the rational function ϕ_j *determined by* (5.17)*, and assume that the coefficients* ω_j *in* (5.43) *are such that*

$$\left|\sum_{j=2m}^{2m+1}\omega_j\widehat{\mathcal{G}}_m(\phi_j)\right| \geq \max\left\{\left|\sum_{j=2m+2}^{N-1}\omega_j\widehat{\mathcal{G}}_m(\phi_j)\right|, \left|\sum_{j=2m+2}^{N-1}\omega_j\widetilde{\mathcal{G}}_{m+1}(\phi_j)\right|\right\}. \tag{5.44}$$

Then the quadrature rules $\widehat{\mathcal{G}}_m(f)$ *and* $\widetilde{\mathcal{G}}_{m+1}(f)$ *bracket* $\mathcal{I}(f)$*.*

Proof. *Since*

$$\mathcal{I}(f) = \omega_0\mathcal{I}(\phi_0), \qquad \mathcal{I}(\phi_j) = 0, \quad \forall j > 0,$$

we have, in view of (5.21) *and* (5.37)*, that*

$$\begin{aligned}\widehat{\mathcal{G}}_m(f) = \sum_{j=0}^{N-1}\omega_j\widehat{\mathcal{G}}_m(\phi_j) &= \sum_{j=0}^{2m-1}\omega_j\widehat{\mathcal{G}}_m(\phi_j) + \sum_{j=2m}^{N-1}\omega_j\widehat{\mathcal{G}}_m(\phi_j)\\ &= \mathcal{I}(f) + \omega_{2m}\widehat{\mathcal{G}}_m(\phi_{2m}) + \omega_{2m+1}\widehat{\mathcal{G}}_m(\phi_{2m+1}) + \sum_{j=2m+2}^{N-1}\omega_j\widehat{\mathcal{G}}_m(\phi_j)\end{aligned} \tag{5.45}$$

and

$$\begin{aligned}\widetilde{\mathcal{G}}_{m+1}(f) = \sum_{j=0}^{N-1}\omega_j\widetilde{\mathcal{G}}_{m+1}(\phi_j) &= \sum_{j=0}^{2m+1}\omega_j(2\mathcal{I} - \widehat{\mathcal{G}}_m)(\phi_j) + \sum_{j=2m+2}^{N-1}\omega_j\widetilde{\mathcal{G}}_{m+1}(\phi_j)\\ &= \mathcal{I}(f) - \omega_{2m}\widehat{\mathcal{G}}_m(\phi_{2m}) - \omega_{2m+1}\widehat{\mathcal{G}}_m(\phi_{2m+1}) + \sum_{j=2m+2}^{N-1}\omega_j\widetilde{\mathcal{G}}_{m+1}(\phi_j).\end{aligned} \tag{5.46}$$

Combining (5.45) *and* (5.46) *shows* (5.44)*.*

Theorem 15 shows that if the coefficients ω_j decay sufficiently rapidly with increasing index j, then rational Gauss and anti-Gauss rules provide quadrature errors that are of opposite sign and of roughly the same magnitude. Since it is difficult to verify for a given expression (5.1) whether the conditions of the theorem hold, we say that pairs of rational Gauss and rational anti-Gauss quadrature rules provide estimates of upper and lower bounds for (5.1).

It is natural to consider the average quadrature rule

$$\mathcal{A}_{2m+1}(f) := \frac{1}{2}(\widehat{\mathcal{G}}_m + \widetilde{\mathcal{G}}_{m+1})(f). \tag{5.47}$$

It follows from (5.45) and (5.46) that

$$\mathcal{A}_{2m+1}(f) = \mathcal{I}(f) + \sum_{j=2m+2}^{N-1}\omega_j\mathcal{A}_{2m+1}(\phi_j).$$

This shows that

$$\mathcal{A}_{2m+1}(f) = \mathcal{I}(f), \quad f \in \mathbb{S}_{2m+2}.$$

This rule is an extension to the average rule defined by Laurie [49].

The computation of the matrix H_m that determines the rational rule $\widehat{\mathcal{G}}_m(f)$ requires that m steps of Algorithm 4 be carried out, while the calculation of the matrix $\widetilde{H}_{m+1}$ that defines the associated rational anti-Gauss rules $\widetilde{\mathcal{G}}_m(f)$ demands $m+1$ steps of Algorithm 4. The last step of the algorithm determines the last diagonal entry entry, $h_{m,m}$, of (5.42). We can reduce the number of steps by replacing this entry by an arbitrary scalar, $\breve{h}$. This defines the matrix $\breve{H}_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$. We refer to the quadrature rule so obtained,

$$\breve{\mathcal{G}}_{m+1}(f) = \int_0^\infty \boldsymbol{e}_1^T (tI + \breve{H}_{m+1})^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(\breve{H}_{m+1}) \boldsymbol{e}_1, \tag{5.48}$$

as a simplified rational anti-Gauss rule. This rule is an extension to rational Gauss quadrature of the simplified anti-Gauss rules associated with (standard) Gauss rules discussed in [1]. In the computed examples of Section 5.6, we found the choice $\breve{h} = h_{m-1,m-1}$, where $h_{m-1,m-1}$ is the last diagonal element of the matrix H_m to yield good results.

The following result provides sufficient conditions for the quadrature rules $\widehat{\mathcal{G}}_m(f)$ and $\breve{\mathcal{G}}_{m+1}(f)$ to bracket (5.1), and holds for an arbitrary scalar $\breve{h}$.

Theorem 16. *The simplified rational anti-Gauss rule* (5.48) *satisfies*

$$\breve{\mathcal{G}}_{m+1}(f) = \mathcal{I}(f), \qquad \forall f \in \mathbb{S}_{2m}, \tag{5.49}$$

$$\breve{\mathcal{G}}_{m+1}(f) = (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f), \qquad \forall f \in \mathbb{S}_{2m+1}. \tag{5.50}$$

Consider the expansion (5.43) *with the rational function* ϕ_j *determined by* (5.17)*, and assume that the coefficients* ω_j *in* (5.43) *are such that*

$$\left|\omega_{2m}\widehat{\mathcal{G}}_m(\phi_{2m})\right| \geq \max\left\{ \left| \sum_{j=2m+1}^{N-1} \omega_j \widehat{\mathcal{G}}_m(\phi_j) \right|, \left| \sum_{j=2m+1}^{N-1} \omega_j \breve{\mathcal{G}}_{m+1}(\phi_j) \right| \right\}. \tag{5.51}$$

Then the quadrature rules $\widehat{\mathcal{G}}_m(f)$ *and* $\breve{\mathcal{G}}_{m+1}(f)$ *bracket* (5.1)*.*

Proof. *The rational anti-Gauss rule* $\widetilde{\mathcal{G}}_{m+1}(f)$ *satisfies* $\widetilde{\mathcal{G}}_{m+1}(f) = (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f)$ *for all rational functions in* $\mathbb{S}_{2m+2}$*. This rule is determined by the matrix* $\widetilde{H}_{m+1}$*, while the simplified rational*

anti-Gauss rule, $\breve{\mathcal{G}}_{m+1}(f)$, is defined by the matrix $\breve{H}_{m+1}$. These matrices have all the same entries except the last diagonal element. In the same way as in Subsection 5.5.1 we conclude that the simplified rational anti-Gauss rule satisfies

$$\breve{\mathcal{G}}_{m+1}(f) = \widetilde{\mathcal{G}}_{m+1}(f) = (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f), \qquad \forall f \in \mathbb{S}_{2m+1}.$$

This shows (5.50). *Property* (5.49) *follows from* (5.50) *since $\widehat{\mathcal{G}}_m(f) = \mathcal{I}(f)$ for $f \in \mathbb{S}_{2m}$. Property* (5.51) *can be shown similarly as Theorem 15.*

Similarly to (5.47), we define the average quadrature rule

$$\breve{\mathcal{A}}_{2m+1}(f) := \frac{1}{2}(\widehat{\mathcal{G}}_m + \breve{\mathcal{G}}_{m+1})(f).$$

It follows from Theorem 16 that this rule satisfies

$$\breve{\mathcal{A}}_{2m+1}(f) = \mathcal{I}(f), \quad f \in \mathbb{S}_{2m+1}.$$

5.6 Computed Examples

In this section, we illustrate the performance of the rational Gauss rules when applied to Stieltjes matrix functions of a symmetric matrices.

The examples of this section compare the performance of standard Gauss and rational Gauss rules. We also illustrate in Example 5.6.4 that rational Gauss rules (5.23) with several distinct poles may give higher accuracy than rational Gauss rules with a single pole at the origin with the same number of nodes.

In all examples, when $m \geq k + 2s$, where $k + 2s$ is the degree of $w(y)$ defined by (5.22), we observed that the derivative $(w^2 f)^{(2m)}$ in (5.27) is positive in an interval that contains the spectrum of A, and the derivative $(w^2 f)^{(2m+1)}$ in (5.32) is negative in an interval that contains the spectrum A and θ. It follows from (5.34) that pairs of rational Gauss and rational Gauss–Radau rules with a fixed node at $\theta \geq \lambda_N$ do not bracket $F(A)$, while pairs of rational Gauss and rational Gauss–Radau rules with a fixed node at $\theta \leq \lambda_1$ give lower and upper bounds for $F(A)$, respectively. We also will illustrate that error bounds for certain functionals (5.1) can be computed by pairs of rational Gauss-Radau rules, and estimates of upper and lower bounds can be determined by pairs of rational Gauss and simplified rational anti-Gauss quadrature rules. To determine the quadrature error we explicitly evaluate the functionals (5.1). This limits the size of the matrices A considered. To

compute the quadrature rules (5.38) and (5.48), we require the elementary basis function ψ_{m-1} to be a monomial. In all examples, we let the elementary basis functions ψ_m and ψ_{m+1} be monomials.

Example 5.6.1. Consider the Stieltjes function (5.3) with $a = 1/2$. We would like to approximate the functional

$$F(A) := \boldsymbol{v}^T A^{-1/2} \boldsymbol{v}, \tag{5.52}$$

where $A \in \mathbb{R}^{1000\times 1000}$ is a symmetric Toeplitz matrix with first row $[1, 1/2, \ldots, 1/1000]$, and $\boldsymbol{v} = [1/\sqrt{1000}, \ldots, 1/\sqrt{1000}]^T \in \mathbb{R}^{1000}$. The smallest eigenvalue of A is $\lambda_1 = 0.3863$ and the largest one is $\lambda_{1000} = 12.1259$. The value of $F(A)$ is approximately 0.2897. Approximations of (5.52) determined by standard and rational Gauss rules, rational Gauss–Radau rules, rational anti-Gauss and simplified rational anti-Gauss quadrature rules are presented. The computations require the solution of linear system of equations with the symmetric positive definite Toeplitz matrices $A - \alpha_i I$, where the α_i are poles. We remark that fast algorithms for the solution of systems of equations with this kind of matrix are available; see, e.g., [3].

Consider the rational Krylov subspace

$$\begin{aligned}\mathbb{K}_m(A, \boldsymbol{v}) = \operatorname{span}\Big\{ & \boldsymbol{v}, A\boldsymbol{v}, (A-\alpha_1 I)^{-1}\boldsymbol{v}, A^2\boldsymbol{v}, \ldots, (A-\alpha_1 I)^{-k_1}\boldsymbol{v}, A^{k_1+1}\boldsymbol{v}, \\ & (A-\alpha_2 I)^{-1}\boldsymbol{v}, A^{k_1+2}\boldsymbol{v}, \ldots, (A-\alpha_2 I)^{-k_2}\boldsymbol{v}, A^{k_1+k_2+1}\boldsymbol{v}, \ldots, \\ & (A-\alpha_\ell I)^{-1}\boldsymbol{v}, \ldots\ , (A-\alpha_\ell I)^{-k_\ell}\boldsymbol{v}, A^{k+1}\boldsymbol{v}\Big\},\end{aligned} \tag{5.53}$$

where k is determined by (5.9). The Stieltjes function (5.3) is defined in the complex plane except for on the interval $(-\infty, 0]$. It therefore is natural to allocate poles on this interval. We consider rational Krylov subspaces $\mathbb{K}_m(A, \boldsymbol{v})$ with poles allocated on $(-\infty, 0]$ in three different ways:

(i) $\mathbb{K}_6(A, \boldsymbol{v})$ is determined by a simple pole, $\alpha_1 = -1/2$ of multiplicity two.

(ii) $\mathbb{K}_8(A, \boldsymbol{v})$ is determined by two distinct poles, $\alpha_1 = -0.4310$ of multiplicity two and $\alpha_2 = -0.9024$ of multiplicity one. These poles are the zeros of the Chebyshev polynomial of the first kind of degree two for the interval $[-1, -1/3]$.

(iii) $\mathbb{K}_{10}(A, \boldsymbol{v})$ is determined by four equidistant poles $\alpha_i \in \{0, -1/2, -1, -3/2\}$ of multiplicity one.

The poles in $\mathbb{K}_6(A, \boldsymbol{v})$ are the zeros of the polynomial $w(y)$ defined by (5.22). We have

$$w(y) = y^2 + y + \frac{1}{4}.$$

In the same manner, we can define $w(y)$ associated with $\mathbb{K}_8(A, \boldsymbol{v})$ and $\mathbb{K}_{10}(A, \boldsymbol{v})$. We evaluate (5.52) as $\boldsymbol{v}^T A^{-1/2} \boldsymbol{v}$, where the vector $A^{-1/2}\boldsymbol{v}$ is calculated by first computing the matrix square root and then solving a linear system of equations. The standard Gauss rule

$$\boldsymbol{e}_1^T T_m^{-1/2} \boldsymbol{e}_1$$

requires the computation of m steps of the standard Lanczos process. The rational Gauss rule is evaluated as

$$\boldsymbol{e}_1^T H_m^{-1/2} \boldsymbol{e}_1,$$

where $H_m^{-1/2}\boldsymbol{e}_1$ is determined by first computing the matrix square root and then solving a linear system of equations. Analogously to the rational Gauss, the simplified rational anti-Gauss, and Gauss-Radau rules with a fixed node $\theta \leq \lambda_1$ or $\theta \geq \lambda_{1000}$ can be computed by carrying out m steps of the Algorithm 4, while the rational anti-Gauss rule is determined by $m + 1$ steps of Algorithm 4.

Columns 2 and 3 of Table 12 display the errors in approximations determined by standard and rational Gauss rules. We observe that rational Gauss rules yield higher accuracy than the standard Gauss rules. Column 4 of Table 12 displays the errors achieved with the rational Gauss–Radau rules. We chose the Radau point $\theta = 0.3$. A comparison with the errors in column 3 shows that pairs of rational Gauss, $\widehat{\mathcal{G}}_m(f)$, and rational Gauss–Radau rules, $\widehat{\mathcal{R}}_m^{0.3}(f)$, provide lower and upper bounds for (5.52), respectively.

Columns 5 and 7 of Table 12 display the errors in approximations obtained by rational anti-Gauss and simplified rational anti-Gauss rules, respectively. It can be seen that the errors of these quadrature rules are of opposite sign and of about the same magnitude as the errors in the corresponding rational Gauss rules. In this example, we chose the last diagonal entry of the matrix $\breve{H}_{m+1}$ that determines the simplified rational anti-Gauss rules to be $\breve{h} = h_{m-1,m-1}$. We also observed that the choice of $\breve{h} = (h_{m-1,m-1} + h_{m-2,m-2})/2$ yields similar results. For instance, we found for $m = 8$ that

$$F(A) - \breve{\mathcal{G}}_{m+1}(f) = -9.21 \cdot 10^{-11}.$$

This illustrates that the results achieved with simplified rational anti-Gauss rules are fairly insensitive to the choice of $\breve{h}$. Therefore, simplified rational anti-Gauss rules can be used to reduce the computational cost. Table 12 also shows that the pairs of rules $\{\widehat{\mathcal{G}}_m, \widetilde{\mathcal{G}}_{m+1}\}$ and $\{\widehat{\mathcal{G}}_m, \breve{\mathcal{G}}_{m+1}\}$ yield tighter error bounds than the pairs of $\{\widehat{\mathcal{G}}_m, \widehat{\mathcal{R}}^{\theta}_{m+1}\}$.

Columns 6 and 8 of Table 12 show the errors in computed approximations determined by the average rules associated with rational Gauss and anti-Gauss rules, and rational Gauss and simplified rational anti-Gauss rules, respectively. These quadrature rules yield more accurate approximations of (5.52) than the corresponding rational Gauss rules.

Table 13 displays the errors in the computed rational Gauss–Radau quadrature rules with fixed nodes at $\theta = 0.3$ and $\theta = 13$. The table illustrates that pairs of rational Gauss–Radau rules provide upper and lower bounds for (5.52), we have

$$\widehat{\mathcal{R}}^{0.3}_{m+1}(f) \geq F(A) \geq \widehat{\mathcal{R}}^{13}_{m+1}(f), \quad \forall m.$$

m	$F(A) - \mathcal{G}_m(f)$	$F(A) - \widehat{\mathcal{G}}_m(f)$	$F(A) - \widehat{\mathcal{R}}^{\theta}_{m+1}(f)$	$F(A) - \widetilde{\mathcal{G}}_{m+1}(f)$	$F(A) - \mathcal{A}_{2m+1}(f)$	$F(A) - \breve{\mathcal{G}}_{m+1}(f)$	$F(A) - \breve{\mathcal{A}}_{2m+1}(f)$
6	$5.79 \cdot 10^{-7}$	$2.75 \cdot 10^{-9}$	$-6.09 \cdot 10^{-9}$	$-2.86 \cdot 10^{-9}$	$-5.57 \cdot 10^{-11}$	$-2.38 \cdot 10^{-9}$	$1.85 \cdot 10^{-10}$
8	$7.28 \cdot 10^{-8}$	$3.95 \cdot 10^{-11}$	$-1.16 \cdot 10^{-10}$	$-4.10 \cdot 10^{-11}$	$-7.65 \cdot 10^{-13}$	$-3.45 \cdot 10^{-11}$	$2.48 \cdot 10^{-12}$
10	$9.20 \cdot 10^{-9}$	$5.46 \cdot 10^{-14}$	$-2.23 \cdot 10^{-13}$	$-5.71 \cdot 10^{-14}$	$-1.22 \cdot 10^{-15}$	$-4.99 \cdot 10^{-14}$	$2.38 \cdot 10^{-15}$

Table 12: Example 5.6.1: Errors for computed approximations of $F(A) = \boldsymbol{v}^T A^{-1/2} \boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau node is fixed at $\theta = 0.3$

m	$F(A) - \widehat{\mathcal{R}}^{0.3}_{m+1}(f)$	$F(A) - \widehat{\mathcal{R}}^{13}_{m+1}(f)$
6	$-6.09 \cdot 10^{-9}$	$2.21 \cdot 10^{-9}$
8	$-1.16 \cdot 10^{-10}$	$3.32 \cdot 10^{-11}$
10	$-2.23 \cdot 10^{-13}$	$4.61 \cdot 10^{-14}$

Table 13: Example 5.6.1: Errors for computed approximations of $F(A) = \boldsymbol{v}^T A^{-1/2} \boldsymbol{v}$ by rational Gauss–Radau rules with A a symmetric Toeplitz matrix. The Radau nodes are $\theta = 0.3$ and $\theta = 13$.

Example 5.6.2. This example determines an approximation of the functional

$$F(A) := \boldsymbol{v}^T \log(A + I) A^{-1} \boldsymbol{v},$$

where $A \in \mathbb{R}^{1000 \times 1000}$ is symmetric Toeplitz matrix with first row $[3, 3/2, \ldots, 3/1000]$. The vector $\boldsymbol{v} \in \mathbb{R}^{1000}$ and the rational Krylov subspaces $\mathbb{K}_6(A, \boldsymbol{v})$ and $\mathbb{K}_8(A, \boldsymbol{v})$ are defined to be the

same as in Example 5.6.1. The subspace $\mathbb{K}_{10}(A, \boldsymbol{v})$ is determined by four equidistant poles $\alpha_i \in \{0, -1/4, -1/2, -1\}$ of multiplicity one. The smallest eigenvalue of A is $\lambda_1 = 1.1589$ and the largest one is $\lambda_{1000} = 36.3776$. Consider the Stieltjes function

$$f(y) = \frac{\log(1+y)}{y} = \int_1^\infty \frac{t^{-1}}{t+y}\, dt.$$

The value of $F(A)$ is approximately 0.1009.

Columns 2 and 3 of Table 14 show the difference between the exact value and the approximations determined by the standard and rational Gauss rules. We note that the quadrature error for the rational Gauss rules is the smallest for all values of m. Column 4 of Table 14 displays the errors in approximations obtained by rational Gauss–Radau rules with a fixed node at $\theta = 1.1$. The table illustrates that pairs of rational Gauss rule, $\widehat{\mathcal{G}}_m(f)$, and associated Gauss–Radau rules, $\widehat{\mathcal{R}}^{1.1}_{m+1}(f)$, bracket the exact value.

Columns 3 and 5 of Table 14 show the errors in the rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and in rational anti-Gauss rules, $\widetilde{\mathcal{G}}_{m+1}(f)$, to have opposite sign and be of about the same magnitude. Similarly, Columns 3 and 7 of Table 14 show the errors in rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and in simplified rational anti-Gauss rules, $\breve{\mathcal{G}}_{m+1}(f)$, with $\breve{h} = (h_{m-1,m-1} + h_{m-2,m-2})/2$ to be of opposite sign and of about the same magnitude. Columns 6 and 8 of Table 14 illustrate that the average rules yield the best approximations of $F(A)$.

Table 15 displays the errors in approximations obtained by rational Gauss–Radau quadrature rules. The table illustrates that

$$\widehat{\mathcal{R}}^{1.1}_{m+1}(f) \geq F(A) \geq \widehat{\mathcal{R}}^{37}_{m+1}(f), \quad \forall m.$$

m	$F(A) - \mathcal{G}_m(f)$	$F(A) - \widehat{\mathcal{G}}_m(f)$	$F(A) - \widehat{\mathcal{R}}^{\theta}_{m+1}(f)$	$F(A) - \widetilde{\mathcal{G}}_{m+1}(f)$	$F(A) - \mathcal{A}_{2m+1}(f)$	$F(A) - \breve{\mathcal{G}}_{m+1}(f)$	$F(A) - \breve{\mathcal{A}}_{2m+1}(f)$
6	$9.65 \cdot 10^{-8}$	$1.88 \cdot 10^{-9}$	$-7.92 \cdot 10^{-9}$	$-1.91 \cdot 10^{-9}$	$-1.57 \cdot 10^{-11}$	$-3.13 \cdot 10^{-9}$	$-6.25 \cdot 10^{-10}$
8	$5.93 \cdot 10^{-9}$	$1.32 \cdot 10^{-11}$	$-3.98 \cdot 10^{-11}$	$-1.33 \cdot 10^{-11}$	$-8.45 \cdot 10^{-14}$	$-2.01 \cdot 10^{-11}$	$-3.44 \cdot 10^{-12}$
10	$3.56 \cdot 10^{-10}$	$1.99 \cdot 10^{-13}$	$-5.21 \cdot 10^{-13}$	$-2.01 \cdot 10^{-13}$	$-1.05 \cdot 10^{-15}$	$-2.97 \cdot 10^{-13}$	$-4.87 \cdot 10^{-14}$

Table 14: Example 5.6.2: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \log(A+I)A^{-1}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau node is $\theta = 1.1$.

Example 5.6.3. Consider the Stieltjes function

$$f(y) = \frac{1}{\log(1+y)} = \int_1^\infty \frac{1}{(y+t)} \frac{dt}{\log^2(y+t)}.$$

m	$F(A)-\widehat{\mathcal{R}}_{m+1}^{1.1}(f)$	$F(A)-\widehat{\mathcal{R}}_{m+1}^{37}(f)$
6	$-7.92\cdot 10^{-9}$	$1.23\cdot 10^{-9}$
8	$-3.98\cdot 10^{-11}$	$8.60\cdot 10^{-12}$
10	$-5.21\cdot 10^{-13}$	$1.31\cdot 10^{-13}$

Table 15: Example 5.6.2: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \log(A+I)A^{-1}\boldsymbol{v}$ by rational Gauss–Radau rules with A a symmetric Toeplitz matrix. The Radau nodes are fixed at $\theta = 1.1$ and $\theta = 37$.

We would like to approximate the functional

$$F(A) := \boldsymbol{v}^T(\log(I+A))^{-1}\boldsymbol{v},$$

where the matrix $A \in \mathbb{R}^{1000\times 1000}$ is the same as in Example 5.6.2. The vector $\boldsymbol{v} \in \mathbb{R}^{1000}$ has normally distributed random entries with zero mean and is normalized to be of unit norm. The value of $F(A)$ is approximately 0.9472. We consider the rational Krylov subspace (5.53) with poles α_i in $(-\infty, 0]$ allocated in three different ways:

(i) $\mathbb{K}_8(A, \boldsymbol{v})$ is determined by two equidistant poles, $\alpha_1 = 0$ of multiplicity two and $\alpha_2 = -1/2$ of multiplicity one.

(ii) $\mathbb{K}_{10}(A, \boldsymbol{v})$ is determined by four equidistant poles $\alpha_i \in \{0, -1, -2, -3\}$ of multiplicity one.

(iii) $\mathbb{K}_{14}(A, \boldsymbol{v})$ is determined by three equidistant poles $\alpha_i \in \{0, -1/4, -1/2\}$ of multiplicity two.

Columns 2 and 3 of Table 16 report the errors in approximations determined by the standard and rational Gauss rules. Columns 4 and 5 of Table 16 show the approximations determined by rational Gauss–Radau and simplified rational anti-Gauss rules. The table illustrates that pairs of rational Gauss and Gauss–Radau rules with fixed node at $\theta = 1.1$, or pairs of rational Gauss and simplified rational anti-Gauss rules with $\breve{h} = h_{m-1,m-1}$, bracket the exact value. Column 6 of Table 16 illustrates that the average rules can yield much higher accuracy than rational Gauss and anti-Gauss rules. Table 17 shows that the values determined by rational Gauss–Radau quadrature rules bracket $F(A)$.

□

m	$F(A)-\mathcal{G}_m(f)$	$F(A)-\widehat{\mathcal{G}}_m(f)$	$F(A)-\widehat{\mathcal{R}}^{\theta}_{m+1}(f)$	$F(A)-\breve{\mathcal{G}}_{m+1}(f)$	$F(A)-\breve{\mathcal{A}}_{2m+1}(f)$
8	$1.10\cdot 10^{-3}$	$1.81\cdot 10^{-8}$	$-1.11\cdot 10^{-7}$	$-1.36\cdot 10^{-8}$	$2.24\cdot 10^{-9}$
10	$1.81\cdot 10^{-4}$	$2.18\cdot 10^{-12}$	$-4.13\cdot 10^{-12}$	$-1.80\cdot 10^{-12}$	$1.93\cdot 10^{-13}$
14	$5.84\cdot 10^{-6}$	$3.60\cdot 10^{-14}$	$-1.70\cdot 10^{-13}$	$-3.91\cdot 10^{-14}$	$-1.55\cdot 10^{-15}$

Table 16: Example 5.6.3: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\log(I+A))^{-1}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau node is $\theta = 1.1$.

m	$F(A)-\widehat{\mathcal{R}}^{1.1}_{m+1}(f)$	$F(A)-\widehat{\mathcal{R}}^{37}_{m+1}(f)$
8	$-1.11\cdot 10^{-7}$	$1.14\cdot 10^{-8}$
10	$-4.13\cdot 10^{-12}$	$1.24\cdot 10^{-12}$
14	$-1.70\cdot 10^{-13}$	$2.29\cdot 10^{-14}$

Table 17: Example 5.6.3: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\log(I+A))^{-1}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau nodes are $\theta = 1.1$ and $\theta = 37$.

Example 5.6.4. In our last example, we compute an approximate of

$$F(A) := \boldsymbol{v}^T(\pi(I+\sqrt{A})^{-1})\boldsymbol{v}, \tag{5.54}$$

where the matrix A is obtained from the discretization of the self-adjoint differential operator $L(u) = \frac{1}{10}u_{xx} + u_{yy}$ in the unit square. Each derivative is approximated by the standard three-point stencil with 40 equally spaced interior nodes in each space dimension. Homogeneous boundary conditions are used. This gives a symmetric positive definite matrix $A \in \mathbb{R}^{1600\times 1600}$. The initial vector $\boldsymbol{v}$ is given by $\boldsymbol{v} = \boldsymbol{e}_1 \in \mathbb{R}^{1600}$. The extreme eigenvalues of A are $\lambda_1 = 0.0646$ and $\lambda_{1600} = 43.9354$. Define the Stieltjes function

$$f(y) = \frac{\pi}{1+\sqrt{y}} = \int_0^\infty \frac{1}{(y+t)}\left(\frac{\sqrt{t}}{1+t}\right)dt,$$

and consider the subspace (5.53) with a single pole, $\alpha_1 = -1/2$, of high multiplicity.

We compare the performance of the methods of this chapter with rational Gauss rules that are presented in [42]. The latter rule is exact for Laurent polynomials, which are rational functions, whose only finite pole is at the origin, and it is known as a Gauss–Laurent quadrature rule. We will denote these rules by $\widehat{\mathcal{G}}^{\mathcal{L}}_m(f)$; they are described in [42]. Algorithm 4 requires the solution of linear systems of equations with the matrix $(A-\alpha_1 I)$, where $\alpha_1 = -1/2$. An algorithm for computing an

approximation of (5.54) by Gauss–Laurent rules is presented in [41]. The computation of this rule requires the solution of linear systems of equations with the matrix A.

Columns 2 and 3 of Table 18 display the difference between the exact value, $F(A) \approx 0.5983$, and the approximations obtained by rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and Gauss–Laurent rules, $\widehat{\mathcal{G}}_m^{\mathcal{L}}(f)$. We find that rational Gauss rules associated with the rational Krylov subspace (5.53) give higher accuracy than Gauss–Laurent rules for all values of m. Columns 3, 4, and 5 of Table 18 show the pairs $\{\widehat{\mathcal{G}}_m(f), \widehat{\mathcal{R}}_{m+1}^{0.05}(f)\}$ and $\{\widehat{\mathcal{G}}_m(f), \breve{\mathcal{G}}_{m+1}(f)\}$ to bracket $F(A)$. The average rules, which are displayed in column 6 of Table 18, are seen to be quite accurate.

Table 19 displays the errors in approximations obtained by the rational Gauss–Radau rules. The table illustrates that

$$\widehat{\mathcal{R}}_{m+1}^{0.05}(f) \geq F(A) \geq \widehat{\mathcal{R}}_{m+1}^{45}(f), \quad \forall m.$$

m	$F(A) - \widehat{\mathcal{G}}_m^{\mathcal{L}}(f)$	$F(A) - \widehat{\mathcal{G}}_m(f)$	$F(A) - \widehat{\mathcal{R}}_{m+1}^{\theta}(f)$	$F(A) - \breve{\mathcal{G}}_{m+1}(f)$	$F(A) - \breve{\mathcal{A}}_{2m+1}(f)$
8	$1.70 \cdot 10^{-5}$	$3.85 \cdot 10^{-7}$	$-1.99 \cdot 10^{-6}$	$-3.90 \cdot 10^{-7}$	$-2.82 \cdot 10^{-9}$
10	$3.17 \cdot 10^{-6}$	$2.28 \cdot 10^{-8}$	$-1.24 \cdot 10^{-7}$	$-2.33 \cdot 10^{-8}$	$-2.75 \cdot 10^{-10}$
14	$9.77 \cdot 10^{-8}$	$1.09 \cdot 10^{-10}$	$-4.67 \cdot 10^{-10}$	$-1.13 \cdot 10^{-10}$	$-1.83 \cdot 10^{-12}$

Table 18: Example 5.6.4: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\pi(I + \sqrt{A})^{-1})\boldsymbol{v}$ when A is a discretization of a differential operator. The Radau node is fixed at $\theta = 0.05$

m	$F(A) - \widehat{\mathcal{R}}_{m+1}^{0.05}(f)$	$F(A) - \widehat{\mathcal{R}}_{m+1}^{45}(f)$
8	$-1.99 \cdot 10^{-6}$	$2.51 \cdot 10^{-7}$
10	$-1.24 \cdot 10^{-7}$	$1.51 \cdot 10^{-8}$
14	$-4.67 \cdot 10^{-10}$	$8.33 \cdot 10^{-11}$

Table 19: Example 5.6.4: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\pi(I + \sqrt{A})^{-1})\boldsymbol{v}$ when A is a discretization of a differential operator. The Radau nodes are $\theta = 0.05$ and $\theta = 45$

CHAPTER 6

Conclusions

In this chapter we draw conclusions about the methods that we proposed in this thesis to approximate expression of the form $\boldsymbol{w}^T f(A)\boldsymbol{v}$, where $\boldsymbol{v}, \boldsymbol{w}$ are given vectors and A is a large symmetric or nonsymmetric matrix.

6.1 Conclusion for Gauss-Type Quadrature Rules of Chapter 3

Golub and Meurant [28, 29] described a technique for computing upper and lower error bounds for a Stieltjes integral by evaluating pairs of Gauss, and suitable Gauss–Radau or Gauss–Lobatto quadrature rules. This technique is not guaranteed to furnish upper and lower error bounds when derivatives of the integrand f change sign on the convex hull of spectrum of A. This chapter extends the technique by Golub and Meurant by using pairs of Gauss, and suitable generalized Gauss–Radau or generalized Gauss–Lobatto rules, to determine upper and lower error bounds for Stieltjes integrals with an integrand f, some of whose derivatives change sign on the convex hull of the support of the measure. New methods for evaluating generalized Gauss–Radau and Gauss–Lobatto rules are described. Computed examples illustrate the benefit of using these quadrature rules.

6.2 Conclusion for Gauss–Laurent-Type Quadrature Rules of Chapter 4

It is known that Gauss–Laurent quadrature rules associated with a real nonnegative measure with support on the real axis are determined by symmetric pentadiagonal matrices. This chapter extends the methods described in [42] to complex-valued measures with support in the complex plane. We investigate the structure of the matrices for Gauss–Laurent and associated anti-Gauss–Laurent quadrature rules and discuss properties of these quadrature rules. Computed examples show that Gauss–Laurent rules may give higher accuracy than standard Gauss rules with the same number of nodes. Moreover, they illustrate that pairs of Gauss–Laurent and anti-Gauss–Laurent rules provide upper and lower bounds for certain matrix functionals.

6.3 Conclusion for Rational Gauss-Type Quadrature Rules of Chapter 5

This chapter discusses the approximation of the expression $\boldsymbol{v}^T f(A)\boldsymbol{v}$, where A is a symmetric positive definite matrix and f is a Stieltjes function, by rational Gauss rules with preselected poles. Associated rational Gauss–Radau and anti-Gauss rules are introduced. Computed examples show that when the integrand f has singularities close to the spectrum of A, rational Gauss rules with poles at or close to these singularities give higher accuracy than standard Gauss rules and Gauss–Laurent rules with a pole at the origin with the same number of nodes. The examples also illustrate that pairs of rational Gauss and Gauss–Radau rules, or pairs of rational Gauss and rational anti-Gauss rules, or simplified rational anti-Gauss rules, provide error bounds, or estimates of bounds.

BIBLIOGRAPHY

[1] H. Alqahtani and L. Reichel, *Simplified anti-Gauss quadrature rules with applications in linear algebra*, Numer. Algorithms, 77 (2018), pp. 577–602.

[2] H. Alqahtani and L. Reichel, *Generalized block anti-Gauss quadrature rules*, Numer. Math., 143 (2019), pp. 605–648.

[3] G. S. Ammar and W. B. Gragg, *Superfast solution of real positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 61–76.

[4] Z. Bai, M. Fahey, and G. H. Golub, *Some large scale matrix computation problems*, J. Comput. Appl. Math., 74 (1996), pp. 71–89.

[5] H. Barkouki, A. H. Bentbib, and K. Jbilou, *A matrix rational Lanczos method for model reduction in large-scale first- and second-order dynamical systems*, Numer. Linear Algebra Appl., 24 (2017), Art. e2077.

[6] V. Batagelj and A. Mrvar, *Pajek data sets*, (2006). Available at `http://vlado.fmf.uni-lj.si/pub/networks/data/`

[7] B. Beckermann and L. Reichel, *Error estimation and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.

[8] A. H. Bentbib, M. El Ghomari, C. Jagels, K. Jbilou, and L. Reichel, *The extended global Lanczos method for matrix function approximation*, Electron. Trans. Numer. Anal., 50 (2018), pp. 144–163.

[9] C. Berg, *Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity*, in Positive Definite Functions. From Schoenberg to Space-Time Challenges, S. Mateu and E. Porcu, eds., Dept. of Mathematics, University Jaume I, Castellón de la Plana, Spain, 2008.

[10] A. Bultheel, P. González Vera, E. Hendriksen, and O. Njåstad, *Orthogonal Rational Functions*, Cambridge University Press, Cambridge, UK, 1999.

[11] D. Calvetti, P. C. Hansen, and L. Reichel, *L-curve curvature bounds via Lanczos bidiagonalization*, Electron. Trans. Numer. Anal., 14 (2002), pp. 20–35.

[12] D. Calvetti, L. Reichel, and F. Sgallari, *Application of anti-Gauss quadrature rules in linear algebra*, in Applications and Computation of Orthogonal Polynomials, W. Gautschi, G. H. Golub, and G. Opfer, eds., Birkhäuser, Basel, 1999, pp. 41–56.

[13] K. Deckers and A. Bultheel, *Recurrence and asymptotics for orthonormal rational functions on an interval*, IMA J. Numer. Anal., 29 (2009), pp. 1–23.

[14] K. Deckers and A. Bultheel, *The existence and construction of rational Gauss-type quadrature rules*, Appl. Math. Comput., 218 (2012), pp. 10299–10320.

[15] C. Fenu, D. Martin, L. Reichel, and G. Rodriguez, *Network analysis via partial spectral factorization and Gauss quadrature*, SIAM J. Sci. Comput., 35 (2013), pp. A2046–A2068.

[16] C. Fenu, D. Martin, L. Reichel, and G. Rodriguez, *Block Gauss and anti-Gauss quadrature with application to networks*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1655–1684.

[17] P. Fika, M. Mitrouli, and P. Roupa, *Estimates for the bilinear form $x^T A^{-1} y$ applications to linear algebra problems*, Electron. Trans. Numer. Anal., 43 (2014), pp. 70–89.

[18] A. Frommer, *Monotone convergence of the Lanczos approximations to matrix functions of Hermitian matrices*, Electron. Trans. Numer. Anal., 35 (2009), pp. 118–128.

[19] A. Frommer and M. Schweitzer, *Error bounds and estimates for Krylov subspace approximations of Stieltjes matrix functions*, BIT Numer. Math., 56 (2016), pp. 865–892.

[20] K. Gallivan, E. Grimme, and P. Van Dooren, *A rational Lanczos algorithm for model reduction*, Numer. Algorithms, 12 (1996), pp. 33–66.

[21] W. Gautschi, *High-order Gauss–Lobatto formulae*, Numer. Algorithms, 25 (2000), pp. 213–222.

[22] W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, Oxford, 2004.

[23] W. Gautschi, *Generalized Gauss–Radau and Gauss–Lobatto formulae*, BIT Numer. Math., 44 (2004), pp. 711–720.

[24] W. Gautschi, *High-order generalized Gauss–Radau and Gauss–Lobatto formulae for Jacobi and Laguerre weight functions*, Numer. Algorithms, 51 (2009), pp. 143–149.

[25] W. Gautschi, MATLAB codes for computing nodes and weights of quadrature rules described in [24] are available at
`https://www.cs.purdue.edu/archives/2002/wxg/codes/HOGGRL.html`

[26] W. Gautschi and S. Li, *The remainder term for analytic functions of Gauss–Radau and Gauss–Lobatto rules with multiple end points*, J. Comput. Appl. Math., 33 (1990), pp. 315–329.

[27] W. Gautschi and S. Li, *Gauss–Radau and Gauss–Lobatto rules with double end points*, J. Comput. Appl. Math., 34 (1991), pp. 343–360.

[28] G. H. Golub and G. Meurant, *Matrices, moments and quadrature*, in Numerical Analysis 1993, eds. D. F. Griffiths and G. A. Watson, Longman, Essex, England, 1994, pp. 105–156.

[29] G. H. Golub and G. Meurant, *Matrices, Moments and Quadrature with Applications*, Princeton University Press, Princeton, 2010.

[30] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 2013.

[31] A. A. Gonchar and G. López Lagomasino, *On Markov's theorem for multipoint Padé approximants*, Math. USSR Sb., 34 (1978), pp. 449–459.

[32] M. H. Gutknecht, *The unsymmetric Lanczos algorithms and their relations to Padé approximation, continued fractions and the QD algorithm*, Proceedings of the Copper Mountain Conference on Iterative Methods, vol. 2, 1990.

[33] M. H. Gutknecht, *A completed theory of the unsymmetric Lanczos process and related algorithms*, Part I, SIAM Journal on Matrix Analysis and Applications 13 (1992), no. 2, 594–639.

[34] M. H. Gutknecht, *A completed theory of the unsymmetric Lanczos process and related algorithms.* Part II, SIAM Journal on Matrix Analysis and Applications 15 (1994), no. 1, 15–58.

[35] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM-Mitt., 36 (2013), pp. 8–31.

[36] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 2, Wiley, New York, 1977.

[37] M. Heyouni and K. Jbilou, *An extended block Arnoldi algorithm for large-scale solutions of the continuous-time algebraic Riccati equation*, Electron. Trans. Numer. Anal., 33 (2009), pp. 53–62.

[38] N. J. Higham, *Functions of Matrices*, SIAM, Philadelphia, 2008.

[39] C. Jagels, T. Mach, L. Reichel, and R. Vandebril, *Convergence rates for inverse-free rational approximation of matrix functions*, Linear Algebra Appl., 510 (2016), pp. 291–310.

[40] C. Jagels, K. Jbilou, and L. Reichel, *The extended global Lanczos method, Gauss–Radau quadrature, and matrix function approximation*, J. Comput. Appl. Math., 381 (2021), Art. 113027.

[41] C. Jagels and L. Reichel, *Recursion relations for the extended Krylov subspace method*, Linear Algebra Appl., 434 (2011), pp. 1716–1732.

[42] C. Jagels and L. Reichel, *The structure of matrices in rational Gauss quadrature*, Math. Comp., 82 (2013), pp. 2035–2060.

[43] C. Jagels and L. Reichel, *The extended Krylov subspace method and orthogonal Laurent polynomials*, Linear Algebra Appl., 431 (2009), pp. 441–458.

[44] H. Jeong, S. Mason, A.-L. Barabási, and Z. N. Oltvai, *Lethality and centrality of protein networks*, Nature, 411 (2001), pp. 41–42.

[45] H. Joulak and B. Beckermann, *On Gautschi's conjecture for generalized Gauss–Radau and Gauss–Lobatto formulae*, J. Comput. Appl. Math., 233 (2009), pp. 768–774.

[46] G. Keady, N. Khajohnsaksumeth, and B. Wiwatanapataphee, *On functions and inverses, both positive, decreasing and convex: and Stieltjes functions*, Cog. Math. Stat., 5 (2018), Art. ID 1477543.

[47] L. Knizhnerman and V. Simoncini, *A new investigation of the extended Krylov subspace method for matrix function evaluations*, Numer. Linear Algebra Appl., 17 (2010), pp. 615–638.

[48] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, United States Governm. Press Office, 1950.

[49] D. P. Laurie, *Anti-Gaussian quadrature formulas*, Math. Comp., 65 (1996), pp. 739–747.

[50] T. Mach, M. S. Pranić, and R. Vandebril, *Computing approximate extended Krylov subspaces without explicit inversion*, Electron. Trans. Numer. Anal., 40 (2013), pp. 414–435.

[51] T. Mach, M. S. Pranić, and R. Vandebril, *Computing approximate (block) rational Krylov subspaces without explicit inversion with extensions to symmetric matrices*, Electron. Trans. Numer. Anal., 43 (2014), pp. 100–124.

[52] S. Massei and L. Robol, *Rational Krylov for Stieltjes matrix functions: Convergence and pole selection*, BIT Numer. Math., 61 (2021), pp. 237–273.

[53] I. Moret and P. Novati, *RD-rational approximations of the matrix exponential*, BIT, 44 (2004), pp. 595–615.

[54] S. Morigi, L. Reichel, and F. Sgallari, *An iterative Lavrentiev regularization method*, BIT Numer. Math., 46 (2006), pp. 589–606.

[55] G. Petrova, *Generalized Gauss–Radau and Gauss–Lobatto formulas with Jacobi weight functions*, BIT Numer. Math., 57 (2017), pp. 191–206.

[56] M. S. Pranić and L. Reichel, *Rational Gauss quadrature*, SIAM J. Numer. Anal., 52 (2014), pp. 832–851.

[57] M. S. Pranić and L. Reichel, *Generalized anti-Gauss quadrature rules*, J. Comput. Appl. Math., 284 (2015), pp. 235–243.

[58] M. S. Pranić and L. Reichel, *Recurrence relations for orthogonal rational functions*, Numer. Math., 123 (2013), pp. 629–642.

[59] M. Pranić, L. Reichel, G. Rodriguez, Z. Wang, and X. Yu, *A rational Arnoldi process with applications*, Numer. Linear Algebra Appl., 23 (2016), pp. 1007–1022.

[60] S. Pozza, M. S. Pranić, and Z. Strakoš, *Gauss quadrature for quasi-definite linear functionals*, IMA J. Numer. Anal., 37 (2017), pp. 1468–1495.

[61] S. Pozza, M. S. Pranić, and Z. Strakoš, *The Lanczos algorithm and complex Gauss quadrature*, Electron. Trans. Numer. Anal., 50 (2018), pp. 1–19.

[62] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.

[63] M. Schweitzer, *A two-sided short-recurrence extended Krylov subspace method for nonsymmetric matrices and its relation to rational moment matching*, Numer. Algorithms, 76 (2016), pp. 1–31.

[64] M. Schweitzer, *Restarting and error estimation in polynomial and extended Krylov subspace methods for the approximation of matrix functions*, PhD. Thesis, Bergische Universität Wuppertal, Wuppertal, 2016.

[65] A. D. Sokol, *Real-variables characterization of generalized Stieltjes functions*, Expo. Math., 28 (2010), pp. 179–185.

[66] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer, 2013

[67] S. Sun, L. Ling, N. Zhang, G. Li, and R. Chen, *Topological structure analysis of the protein-protein interaction network in budding yeast*, Nucleic Acids Research, 31 (2003), pp. 2443–2450.

[68] N. Van Buggenhout, M. Van Barel, and R. Vandebril, *Biorthogonal rational Krylov subspace methods*, Electron. Trans. Numer. Anal., 51 (2019), pp. 451–468.

[69] J. Van Deun, A. Bultheel, *Orthogonal rational functions on an interval and quadrature*, Technical Report TW322, Department of Computer Science, KULeuven, March 2001.

[70] B. D. Welfert, *The weights of classical Gauss–Radau quadratures with double end-point*, Appl. Numer. Math., 60 (2010), pp. 574–586.

[71] Q. Ye, *A breakdown-free variation of the nonsymmetric Lanczos algorithms*, Mathematics of Computation 62 (1994), no. 205, 179–207