NAVAL POSTGRADUATE SCHOOL
Monterey, California

A DIVIDE AND CONQUER METHOD
FOR UNITARY AND ORTHOGONAL EIGENPROBLEMS

William Gragg

L. Reichel

February 1989

Approved for public release; distribution unlimited
Prepared for: Naval Postgraduate School and the
                National Science Foundation, Washington
                D.C.   20550

NAVAL POSTGRADUATE SCHOOL
Department of Mathematics

Rear Admiral R. C. Austin
Superintendent

Harrison Shull
Provost

Prepared by:

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b RESTRICTIVE MARKINGS			
2a SECURITY CLASSIFICATION AUTHORITY			3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited			
2b DECLASSIFICATION/DOWNGRADING SCHEDULE						
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NPS-53-89-007			5 MONITORING ORGANIZATION REPORT NUMBER(S) NPS-53-89-007			
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School	6b OFFICE SYMBOL (If applicable) 53		7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School and the National Science Foundation			
6c ADDRESS (City, State, and ZIP Code) Monterey, CA 93943			7b ADDRESS (City, State, and ZIP Code) Monterey, CA 93943 and Washington, D.C. 20550			
8a NAME OF FUNDING/SPONSORING ORGANIZATION Naval Postgraduate School	8b OFFICE SYMBOL (If applicable) 53		9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER O&MN, Direct funding			
8c ADDRESS (City, State, and ZIP Code) Monterey, CA 93943			10 SOURCE OF FUNDING NUMBERS			
			PROGRAM ELEMENT NO	PROJECT NO	TASK NO.	WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

A DIVIDE AND CONQUER METHOD FOR UNITARY AND ORTHOGONAL EIGENPROBLEMS

12 PERSONAL AUTHOR(S)
William Gragg and L. Reichel

13a TYPE OF REPORT Technical Report	13b TIME COVERED FROM 1 Oct 88 TO 30 Jan 89	14 DATE OF REPORT (Year, Month, Day) 1989 February 22	15 PAGE COUNT 37

16 SUPPLEMENTARY NOTATION

COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	unitary eigenproblem, orthogonal eigenproblem, divide and
			conquer, parallel algorithm, Pisarenko frequencies, Gauss-Szegö quadrature.

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Let $H \in C^{n \times n}$ be a unitary upper Hessenberg matrix whose eigenvalues, and possibly also eigenvectors, are to be determined. We describe how this eigenproblem can be solved by a divide and conquer method, in which the matrix H is split into two smaller unitary right Hessenberg matrices H_1 and H_2 by a rank-one modification of H. The eigenproblems for H_1 and H_2 can be solved independently, and the solutions of these smaller eigenproblems define a rational function, whose zeros on the unit circle are the eigenvalues of H. The eigenvectors of H can be determined from the eigenvalues of H and the eigenvectors of H_1 and H_2. The outlined splitting of unitary upper Hessenberg matrices into smaller such matrices is carried out recursively. This gives rise to a divide and conquer method that is suitable for implementation on a parallel computer.

When $H \in R^{n \times n}$ is orthogonal, the divide and conquer scheme simplifies and is described separately. Our interest in the orthogonal eigenproblem stems from applications in signal processing. Numerical examples for the orthogonal eigenproblem conclude the paper.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a NAME OF RESPONSIBLE INDIVIDUAL William Gragg	22b TELEPHONE (Include Area Code) (408) 646-2194	22c OFFICE SYMBOL 53Gr

DD FORM 1473, 84 MAR   83 APR edition may be used until exhausted   SECURITY CLASSIFICATION OF THIS PAGE
All other editions are obsolete   UNCLASSIFIED

A Divide and Conquer Method for Unitary and Orthogonal Eigenproblems[*]

W.B. Gragg[**]
Naval Postgraduate School
Department of Mathematics
Monterey, CA 93943, USA

L. Reichel[**]
Bergen Scientific Centre
Allegaten 36
N-5007 Bergen, Norway

Abstract: Let $H \in \mathbb{C}^{n \times n}$ be a unitary upper Hessenberg matrix whose eigenvalues, and possibly also eigenvectors, are to be determined. We describe how this eigenproblem can be solved by a divide and conquer method, in which the matrix H is split into two smaller unitary right Hessenberg matrices H_1 and H_2 by a rank-one modification of H. The eigenproblems for H_1 and H_2 can be solved independently, and the solutions of these smaller eigenproblems define a rational function, whose zeros on the unit circle are the eigenvalues of H. The eigenvectors of H can be determined from the eigenvalues of H and the eigenvectors of H_1 and H_2. The outlined splitting of unitary upper Hessenberg matrices into smaller such matrices is carried out recursively. This gives rise to a divide and conquer method that is suitable for implementation on a parallel computer.

When $H \in \mathbb{R}^{n \times n}$ is orthogonal, the divide and conquer scheme simplifies and is described separately. Our interest in the orthogonal eigenproblem stems from applications in signal processing. Numerical examples for the orthogonal eigenproblem conclude the paper.

Subject classification: AMS(MOS): 65F15; CR: 1.3

Key Words: unitary eigenproblem, orthogonal eigenproblem, divide and conquer, parallel algorithm, Pisarenko frequencies, Gauss-Szegö quadrature.

1

1. Introduction

Divide and conquer (DC) methods have been developed for the symmetric eigenproblem, see Cuppen [Cu], Dongarra and Sorenson [DS], and Krishnakumar and Morf [KM], and have for these problems been shown to be efficient on parallel computers and competitive on single processor machines [DS],[Cu],[KM]. The DC method has also been applied successfully to the computation of singular values by Jessup and Sorensen [JS]. In the present paper we describe a DC method for the unitary eigenproblem, and we also discuss the simplifications that arise for real orthogonal matrices.

Let $H \in \mathbb{C}^{n \times n}$ be unitary. Then H is unitarily similar to a upper Hessenberg matrix with real-valued non-negative subdiagonal elements. If a subdiagonal element vanishes, then the eigenproblem splits into eigenproblems for smaller upper Hessenberg matrices. We therefore may assume that H is a upper Hessenberg matrix with <u>positive</u> subdiagonal elements. Then all eigenvalues of H are simple. It is easily seen that H can be written as a product of n Givens reflectors $G_j \in \mathbb{C}^{n \times n}$,

$$H = H(\gamma_1, \gamma_2, \ldots, \gamma_n) := G_1 G_2 \ldots G_{n-1} G_n, \tag{1.1}$$

where, for $1 \leq k < n$,

$$G_k := \begin{bmatrix} I_{k-1} & & & \\ & -\gamma_k & \sigma_k & \\ & \sigma_k & \bar{\gamma}_k & \\ & & & I_{n-k-1} \end{bmatrix}, \quad \gamma_k \in \mathbb{C}, \ \sigma_k \in \mathbb{R}, \ \sigma_k > 0, \quad |\gamma_k|^2 + \sigma_k^2 = 1, \tag{1.2a}$$

and

$$G_n := \begin{bmatrix} I_{n-1} & \\ & -\gamma_n \end{bmatrix}; \quad \gamma_n \in \mathbb{C}, \ |\gamma_n| = 1 \ . \tag{1.2b}$$

2

Here I_j denotes the $j \times j$ identity matrix. The γ_j, $1 \leq j \leq n$, are the so-called <u>Schur</u> <u>parameters</u> of H, and $\bar{\gamma}_j$ denotes the complex conjugate of γ_j. The σ_j, $1 \leq j \leq n$, are said to be <u>complementary</u> <u>parameters</u> of H, and are the subdiagonal elements of H.

The DC method described uses the product representation (1.1) of H, the so-called <u>Schur</u> <u>parametric</u> <u>form</u> of H. An application of particular interest to us is the computation of Pisarenko frequency estimates for a random stationary stochastic process, see below. In this application H is defined by its Schur parametric form. The determination of Gaussian quadrature rules on the unit circle, so-called Gauss-Szegö quarature rules, also gives rise to unitary (or real orthogonal) matrices in Schur parametric form, see Section 5. When the Schur parameters are not explicitly known, they can be computed from

$$\gamma_1 = -H_{11}$$

$$\gamma_j = -(G_{j-1}^H \ G_{j-2}^H \ \cdots \ G_2^H \ G_1^H \ H)_{jj}, \quad j = 2,3,\ldots,n,$$

where G_k^H denotes the conjugate transpose of G_k, and M_{jj} denotes the element (j,j) of a matrix $M \in \mathbb{C}^{n \times n}$.

The DC method is most easily described for $H \in \mathbb{R}^{n \times n}$ orthogonal. Then

$$H = G_1 \ G_2 \ \cdots \ G_{s-1} \ G_s \ G_{s+1} \ \cdots \ G_n =: \begin{bmatrix} H_1 & \\ & I_{n-s} \end{bmatrix} G_s \begin{bmatrix} I_s & \\ & H_2 \end{bmatrix}, \qquad (1.3)$$

where $H_1 \in \mathbb{R}^{s \times s}$ and $H_2 \in \mathbb{R}^{(n-s) \times (n-s)}$ are orthogonal. The Givens reflector $G_s \in \mathbb{R}^{n \times n}$, $s < n$, can be written as a Householder transformation

$$G_s = I - 2ww^H \qquad (1.4)$$

where:

3

$$w := e_s \omega_s + e_{s+1} \omega_{s+1} \in \mathbb{R}^n , \tag{1.5}$$

$$\omega_s := 2^{-1/2} (1 + \gamma_s)^{1/2} , \tag{1.6a}$$

$$\omega_{s+1} := -2^{-1/2} (1 - \gamma_s)^{1/2} . \tag{1.6b}$$

Throughout this paper e_j denotes the jth column of an identity matrix of appropriate order. By (1.3)-(1.4), H is orthogonally similar to

$$H' := \begin{bmatrix} H_1 & \\ & H_2 \end{bmatrix} (I - 2ww^H) =: \tilde{H} - 2\tilde{H}ww^H . \tag{1.7}$$

This is one step of the DC method for orthogonal matrices: The eigenvalues, and if so desired the eigenvectors, of H_1 and H_2 are computed first. H' is a rank-one modification of $\tilde{H}$, and the eigenvalues of $\tilde{H}$ are computed as the zeros on the unit circle of a rational function, whose poles are the eigenvalues of H'.

Section 2 describes the DC method for unitary matrices. In Section 3 we show some results on the orthogonality of the eigenvectors and on the location of the eigenvalues. These results are analogous to bounds presented by Dongarra and Sorensen [DS] for the DC method for symmetric matrices. Section 4 discusses simplifications that can be made when H is real and orthogonal, and also considers some computational details. Computed examples for the orthogonal eigenvalue problem are presented in Section 5.

An outline of a unitary DC method with convergence results for the root-finder has previously been presented in [GR]. The splitting into subproblems is done differently in the present paper. A related DC method is described by Arbenz and Golub [AG]. Cybenko [Cy] reduces the orthogonal eigenproblem to an eigenproblem for a symmetric tridiagonal

4

matrix. The orthogonal eigenproblem is in [AGR1] solved by solving
singular value problems for certain bidiagonal matrices, and a QR
algorithm for unitary matrices is presented in [Gr1]. A comparison with
respect to accuracy and speed of these methods still remains to be done.
Here we only note that DC methods are suitable for implementation on
parallel computers, see [DS], [KM], and Section 2. Some of the schemes
mentioned transform the orthogonal eigenvalue problem to a symmetric one.
The real eigenvalues of the latter problem are then mapped to the unit
circle to yield the eigenvalues of the orthogonal eigenproblem. The
mapping from the interval to the unit circle may be sensitive to
perturbations of arguments near the end points of the interval, and it may
therefore be difficult to determine eigenvalues close to ± 1 accurately
with these schemes.

Pisarenko [Pi] proposed a method for decomposing a random stationary
stochastic process $\{x_m\}_{m=0}^{\infty}$, $x_m \in \mathbb{R}$, into a sum of harmonics in white
noise, i.e.,

$$x_m = \sum_{\ell=1}^{p} \alpha_\ell \, \cos(m\phi_\ell + \theta_\ell) + y_m, \quad m \geq 0, \tag{1.8}$$

where the θ_ℓ are arbitrary phase shifts and $\{y_m\}_{m=0}^{\infty}$ is a zero mean white
noise process with variance σ^2. The ϕ_ℓ are called <u>Pisarenko frequency</u>
<u>estimates</u>. Assume for simplicity that p is the number of distinct
harmonics in the 'signal' $\{x_m\}_{m=0}^{\infty}$ is known, and that $0 < \phi_\ell < \pi$ for
$1 \leq \ell \leq p$. Then the ϕ_ℓ can be determined as follows. Form the
$(2p+1) \times (2p+1)$ Toeplitz covariance matrix M for the signal $\{x_m\}_{m=0}^{\infty}$, and
compute its least eigenvalue λ_{min}. Then $\lambda_{min} = \sigma^2$, see [Pi]. Let $\{\gamma_j\}_{j=1}^{2p}$ be
the Schur parameters associated with the Toeplitz matrix $M-\lambda_{min}I$. They can

5

be determined from the Szegö recursions (Levinson's algorithm), see e.g. [AGR2]. From our assumptions it follows that $M-\lambda_{min}I$ is singular, but leading principal submatrices not identical with $M-\lambda_{min}I$ are not. Therefore, $-1 < \gamma_j < 1$ for $1 \leq j < 2p$, and $\gamma_{2p} \in \{-1,1\}$. By (1.1)-(1.2) it follows that the Schur parameters $\{\gamma_j\}_{j=1}^{2p}$ define an orthogonal matrix $H \in \mathbb{R}^{2p \times 2p}$ with distinct eigenvalues $\{\lambda_j\}_{j=1}^{2p}$. Enumerate the eigenvalues so that those with $Im(\lambda_j) \geq 0$ have smaller index than the eigenvalues with $Im'(\lambda_j) < 0$. Then the Pisarenko frequency estimates are given by

$$\phi_j := \arg(\lambda_j) , \quad 1 \leq j \leq p .$$

The coefficients α_j of (1.8) are two times the weights belonging to the Gauss-Szegö quadrature rule with abscissas λ_j, $1 \leq j \leq p$. For details see [AGR2], where also references to related work can be found. The unitary DC method yields the Gauss-Szegö weights with no extra computational effort when computing the eigenvalues λ_j. Gauss-Szegö quadrature is discussed in Example 5.1 of Section 5.

2. The unitary eigenproblem

In this section we describe a divide and conquer method for unitary right Hessenberg matrices with positive subdiagonal elements. First we need to generalize the splitting (1.3)-(1.7) to Givens reflectors with complex-valued Schur parameters. This is accomplished by noting that the G_s defined by (1.2a) are diagonally unitarily equivalent with a real Householder transformation. Introduce

$$\gamma_s' := \begin{cases} \gamma_s/|\gamma_s| \,, & \gamma_s \neq 0 \\ 1 \,, & \gamma_s = 0 \,. \end{cases}$$

Then $|\gamma_s'| = 1$, and for G_s defined by (1.2a) we obtain

$$\begin{bmatrix} I_{s+1} & & \\ & \bar{\gamma}_s' & \\ & & I_{n-s} \end{bmatrix} G_s \begin{bmatrix} I_s & & \\ & \gamma_s' & \\ & & I_{n-s-1} \end{bmatrix}$$

$$= \begin{bmatrix} I_{s-1} & & & \\ & -|\gamma_s| & \sigma_s & \\ & \sigma_s & |\gamma_s| & \\ & & & I_{n-s-1} \end{bmatrix} \tag{2.1}$$

where, similarly to (1.5)-(1.6), $w = e_s \omega_s + e_{s+1} \omega_{s+1} \in \mathbb{R}^n$ and

$$\omega_s := 2^{-1/2} (1 + |\gamma_s|)^{1/2} \,, \tag{2.2a}$$

$$\omega_{s+1} := -2^{-1/2} (1 - |\gamma_s|)^{1/2} \,. \tag{2.2b}$$

Substitution of (2.1) into (1.1) yields

$$H = \begin{bmatrix} H_1 & \\ & I_{n-s} \end{bmatrix} (I - 2ww^H) \begin{bmatrix} I_s & \\ & H_2 \end{bmatrix} \tag{2.3}$$

with

$$H_1 := H(\gamma_1, \gamma_2, \ldots, \gamma_{s-1}, -\gamma_s'),$$

$$H_2 := H(\bar{\gamma}_s' \gamma_{s+1}, \bar{\gamma}_s' \gamma_{s+2}, \ldots, \bar{\gamma}_s' \gamma_n).$$

A unitary similarity transform of (2.3) yields, analogously with (1.7),

$$H' := \begin{bmatrix} H_1 & \\ & H_2 \end{bmatrix} (I - 2ww^H) =: \tilde{H} - 2\tilde{H}ww^H. \tag{2.4}$$

Let

$$H_k = W_k \Lambda_k W_k^H, \quad k = 1, 2, \tag{2.5}$$

be spectral resolutions, i.e., the W_k are unitary and the Λ_k are diagonal. Then $\tilde{H}$ has the spectral resolution $\tilde{H} = \tilde{W} \tilde{\Lambda} \tilde{W}^H$, where

$$\tilde{W} := \begin{bmatrix} W_1 & \\ & W_2 \end{bmatrix}, \tilde{\Lambda} := \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \tag{2.6}$$

with $|\lambda_k| = 1$ for $1 \leq k \leq n$.

We are in a position to describe how the spectrum of H can be obtained from $\tilde{\Lambda}$, the last row of W_1 and the first row of W_2. Introduce the characteristic polynomial

$$\begin{aligned} \chi(\lambda) &:= \det(\lambda I - H) = \det(\lambda I - H') = \det(\lambda I - \tilde{H} + 2\tilde{H}ww^H) \\ &= \det(\lambda I - \tilde{H}) \det(I + 2(\lambda I - \tilde{H})^{-1}\tilde{H}ww^H) \\ &= \psi(\lambda)(1 + 2w^H(\lambda I - \tilde{H})^{-1}\tilde{H}w) \\ &= \psi(\lambda)(1 + 2w^H \tilde{W}(\lambda I - \tilde{\Lambda})^{-1}\tilde{\Lambda}\tilde{W}^H w), \end{aligned}$$

8

where Π' is defined by (2.4), $\tilde{W}$ and $\tilde{\Lambda}$ by (2.6), w by (2.2) and $\psi(\lambda) :=$ $\det(\lambda[-\tilde{\Pi}])$. Let $z = [\zeta]_{j=1}^{n}$ be given by

$$z := \tilde{W}^H w = \begin{bmatrix} W_1^H e_s \omega_s \\ W_2^H e_1 \omega_{s+1} \end{bmatrix} , \tag{2.7}$$

and define the <u>spectral function</u>

$$\phi(\lambda) := \frac{\chi(\lambda)}{\psi(\lambda)} = 1 + 2z^H(\lambda I - \tilde{\Lambda})^{-1}\tilde{\Lambda}z = 1 + 2\sum_{j=1}^{n} |\zeta_j|^2 \frac{\lambda_j}{\lambda - \lambda_j}$$

$$= \sum_{j=1}^{n} |\zeta_j|^2 \frac{\lambda + \lambda_j}{\lambda - \lambda_j} , \tag{2.8}$$

where we have used that $z^H z = 1$. Let

$$\theta_j := \arg(\lambda_j) , \quad \theta := \arg(\lambda) , \quad 0 \le \theta_j, \theta < 2\pi .$$

Then, with $i := \sqrt{-1}$,

$$\phi(\lambda) = i \sum_{j=1}^{n} |\zeta_j|^2 \cot\left(\frac{\theta_j - \theta}{2}\right) =: i\Phi(\theta) , \tag{2.9}$$

$$\Phi'(\theta) = \frac{1}{2} \sum_{j=1}^{n} |\zeta_j|^2 / \sin^2\left(\frac{\theta_j - \theta}{2}\right) \ge \frac{1}{2} z^H z = \frac{1}{2} . \tag{2.10}$$

We may assume that the θ_j are distinct and that all $\zeta_j \neq 0$, because otherwise we can make these conditions hold by deflation, see below. Let $\theta_j' \in [0, 2\pi[$, $1 \le j \le n$, denote the zeros of $\Phi(\theta)$. Then the eigenvalues of H' and of Π are given by $\lambda_j' := \exp(i\theta_j')$, $1 \le j \le n$. The sets $\{\theta_j\}_{j=1}^{n}$ and $\{\theta_j'\}_{j=1}^{n}$ strictly interlace.

We describe a rootfinder for $\Phi(\theta)$. By the inequality (2.10), the zeros of $\Phi(\theta)$ can be determined accurately. We may assume that

9

$0 < \theta_1 < \theta_2 < \dots < \theta_n < 2\pi$ and that $\theta^{(0)}$, our initial approximation of a zero of $\Phi(\theta)$, satisfies $\theta_n - 2\pi < \theta^{(0)} < \theta_1$. By the strict interlacing of the sets $\{\theta_j\}_{j=1}^n$ and $\{\theta_j'\}_{j=1}^n$, $\Phi(\theta)$ has precisely one zero, denoted θ_1', in the open interval $]\theta_n - 2\pi, \theta_1[$. Assume for the moment that

$$\Phi(\theta^{(0)}) < 0 , \tag{2.11}$$

and introduce

$$\hat{\Phi}(\theta) := \rho + \sigma \, \cot\left(\frac{\theta_1 - \theta}{2}\right) . \tag{2.12}$$

The coefficients ρ and σ are determined by osculatory interpolation, i.e.,

$$\hat{\Phi}(\theta^{(0)}) = \Phi(\theta^{(0)}) , \quad \hat{\Phi}'(\theta^{(0)}) = \Phi'(\theta^{(0)}) , \tag{2.13}$$

which yields

$$\begin{cases} \rho = \Phi(\theta^{(0)}) - \Phi'(\theta^{(0)}) \, \sin(\theta_1 - \theta^{(0)}) , \\ \sigma = 2\Phi'(\theta^{(0)}) \, \sin^2\left(\frac{\theta_1 - \theta^{(0)}}{2}\right) . \end{cases}$$

The zero $\theta^{(1)}$ of $\hat{\Phi}(\theta)$ in $]\theta_n - 2\pi, \theta_1[$ is our next approximation of θ_1'. New approximations $\theta^{(j+1)}$ of θ_1' are computed from $\theta^{(j)}$, $j \geq 1$, in a similar fashion. The sequence $\{\theta^{(j)}\}_{j=0}^\infty$ satisfies $\theta^{(j)} \leq \theta_1'$ for $j \geq 0$, and converges monotonically and quadratically to θ_1' as j increases, see [GR] for a proof.

If instead of (2.11) we have

$$\Phi(\theta^{(0)}) > 0 , \tag{2.14}$$

then we replace (2.12) by

$$\hat{\Phi}(\theta) = \rho + \sigma \, \cot\left(\frac{\theta_n - \theta}{2}\right) \tag{2.15}$$

10

in (2.13). This defines ρ and σ. The zero $\theta^{(1)}$ of $\hat{\Phi}(\theta)$ in the open

interval $]\theta_n - 2\pi, \theta_1[$ is our next approximation of θ_1'. New approximations

$\theta^{(j+1)}$ of θ_1' are computed from $\theta^{(j)}$ for $j \geq 1$ in a similar fashion. The

sequence $\{\theta^{(j)}\}_{j=0}^{\infty}$ satisfies $\theta^{(j)} \geq \theta_1'$ for $j \geq 0$, and converges monotonically

and quadratically to θ_1', see [GR]. In the implementation used to generate

the computed examples of Section 5, the iterations are carried out until

$\Phi(\theta^{(J+1)}) \geq \Phi(\theta^{(i)} < \Phi(\theta^{(j-1)})$. The value $\theta^{(j)}$ is accepted as an approximate

root of $\Phi(\theta) = 0$.

From

$$\Lambda := \mathrm{diag}[e^{i\theta_1'}, e^{i\theta_2'}, \ldots, e^{i\theta_n'}] \qquad (2.16)$$

and the spectral resolutions (2.5) of H_1 and H_2, we can now compute the

spectral resolution of H:

$$H = W\Lambda W^H , \qquad W^H W = I . \qquad (2.17)$$

By (2.3)-(2.4), we can for some vector $\tilde{u} \in \mathbb{C}^n$ express H as

$$H = \tilde{H} - 2u\tilde{u}^H , \qquad u := \begin{bmatrix} H_1 e_s \omega_s \\ e_1 \omega_{s+1} \end{bmatrix} \in \mathbb{C}^n .$$

Let $\lambda := \exp(i\theta')$ and $v \in \mathbb{C}^n$ form an eigenpair of H, i.e. $Hv = v\lambda$. Then

$$(\tilde{H} - 2u\tilde{u}^H)v = v\lambda ,$$

or, equivalently,

$$(\tilde{H} - I\lambda)v = u\alpha , \qquad \alpha := 2\tilde{u}^H v .$$

This shows that any normalized eigenvector v of H associated with the

eigenvalue λ is a normalization of

11

$$v' := (\tilde{H} - I\lambda)^{-1}u = \begin{bmatrix} (H_1 - I\lambda)^{-1}H_1 e_s \omega_s \\ (H_2 - I\lambda)^{-1}e_1 \omega_{s+1} \end{bmatrix}$$

$$= \begin{bmatrix} W_1(I - \Lambda_1^H \lambda)^{-1}W_1^H e_s \omega_s \\ W_2(\Lambda_2 - I\lambda)^{-1}W_2^H e_s \omega_{s+1} \end{bmatrix} , \qquad (2.18)$$

where W_k and Λ_k are given by (2.5). Let $\| \ \|_2$ denote the Euclidean vector and matrix norms. From $\|W_1\|_2 = \|W_2\|_2 = \|\Lambda_1\|_2 = 1$ and (2.6), (2.7), (2.10), (2.18) it follows that

$$\delta(\lambda) := \|v'\|_2 = \|(\tilde{\Lambda} - I\lambda)^{-1} \ x^2 \ \|_2$$

$$= (\sum_{j=1}^{n} \frac{|\zeta_j|^2}{|\lambda_j - \lambda|^2})^{1/2} = (\tfrac{1}{2}\Phi'(\theta))^{1/2} \geq \tfrac{1}{2} . \qquad (2.19)$$

We choose

$$v_\lambda = \begin{bmatrix} W_1(I - \Lambda_1^H \lambda)^{-1}W_1^H e_s \omega_s \\ W_2(\Lambda_2 - I\lambda)^{-1}W_2^H e_1 \omega_{s+1} \end{bmatrix} /\delta(\lambda) , \qquad (2.20)$$

and note that the lower bound (2.19) for $\delta(\lambda)$ indicates that severe cancellation of significant digits does not take place in the computation of v_λ by (2.20).

By (2.7), we only require the last row of W_1 and the first row of W_2 (as well as $\tilde{\Lambda}$) in order to determine the spectral function (2.8). Hence, if we do not desire the eigenvectors, then it suffices to determine the first and last rows of W in order to be able to compute the spectrum of larger problems. We call the triplet $\{\Lambda, e_1^H W, e_n^H W\}$ the _partial_ _spectral_ _resolution_ of H. The first and last elements of v_λ can easily be determined from

$$\begin{cases} e_1^H v_\lambda = e_1^H W_1(I - \Lambda_1^H \lambda)^{-1}W_1^H e_s \omega_s/\delta(\lambda) , \\ e_n^H v_\lambda = e_{n-s}^H W_2(\Lambda_2 - I\lambda)^{-1}W_2^H e_1 \omega_{s+1}/\delta(\lambda) . \end{cases} \qquad (2.21)$$

We may assume that the columns of W are such that all components of the vector $W^H e_1$ are real and positive. Then $e_k^H W^H e_1$ is the weight corresponding to the node $\exp(i\theta_k')$ in the Gauss–Szegö quadrature rule with nodes $\{\exp(i\theta_j')\}_{j=1}^n$, see [Gr2] and Example 5.1.

We assumed above all components ζ_ℓ of z to be non-vanishing and all eigenvalues λ_j of $\tilde{H}$ to be distinct. These conditions can be made to hold true by <u>deflation</u>. Our discussion follows Dongarra and Sorensen [DS]. First assume that ζ_ℓ vanishes. By (2.4)–(2.7),

$$\tilde{W}^H H' \tilde{W} = \tilde{\Lambda} - 2\tilde{W}^H \tilde{H} w w^H \tilde{W} = \tilde{\Lambda}(I - 2zz^H) \; , \tag{2.22}$$

and from $e_\ell^H z = 0$ it follows that

$$\tilde{\Lambda}(I - 2zz^H)e_\ell = \tilde{\Lambda} e_\ell = \lambda_\ell e_\ell, \qquad \lambda_\ell := e_\ell^H \tilde{\Lambda} e_\ell \; . \tag{2.23}$$

Substituting (2.23) into (2.22) yields

$$H'\tilde{W} e_\ell = \tilde{W} e_\ell \lambda_\ell \; ,$$

and therefore

$$H \begin{bmatrix} W_1 & \\ & W_2 \Lambda_2{}^H \end{bmatrix} e_\ell = \begin{bmatrix} W_1 & \\ & W_2 \Lambda_2{}^H \end{bmatrix} e_\ell \lambda_\ell \; .$$

Thus if $\zeta_\ell = 0$, then we can determine an eigenpair of H without explicitly computing a zero of $\Phi(\theta)$ and without using (2.20).

For a general $z \in \mathbb{C}^n$, with $z^H z = 1$, we obtain

$$\|\tilde{\Lambda}(I - 2zz^H)e_\ell\|_2 = 2|\zeta_\ell| \; ,$$

and we accept $\{\lambda_\ell, e_\ell\}$ as an eigenpair of $\tilde{\Lambda}$ if

13

$$2|\zeta_\ell| \le \epsilon_1 \tag{2.24}$$

for some small constant ϵ_1.

Assume that $z = [\zeta_j]_{j=1}^n$ with $2|\zeta_j| > \epsilon_1$ for all j. We may now be able to deflate due to close eigenvalues. Let $\tilde{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$ with $\lambda_1 \approx \lambda_2$, and choose the Givens reflector

$$G = \begin{bmatrix} -\gamma & \sigma & \\ \sigma & \bar{\gamma} & \\ & & I_{n-2} \end{bmatrix} \in \mathbb{C}^{n \times n} \, , \tag{2.25}$$

so that

$$Gz = [(|\zeta_1|^2 + |\zeta_2|^2)^{1/2}, 0, \zeta_3, \zeta_4, \ldots, \zeta_n]^\top \, ,$$

i.e.

$$\begin{cases} \sigma := |\zeta_2|/(|\zeta_1|^2 + |\zeta_2|^2)^{1/2} \, , \\ \gamma := -\bar{\zeta}_1 \dfrac{\zeta_2}{|\zeta_2|}/(|\zeta_1|^2 + |\zeta_2|^2)^{1/2} \, . \end{cases}$$

We accept $\{\gamma_1 \sigma^2 + \gamma_2 |\gamma|^2, G^H e_2\}$ as an (approximate) eigenpair of $\tilde{\Lambda}(I - 2zz^H)$ if

$$|\gamma\sigma(\lambda_1 - \lambda_2)| \le \epsilon_2, \tag{2.26}$$

for some small constant e_2, because

$$\|\tilde{\Lambda}(I-2zz^H)G^H e_2 - (\lambda_1 \sigma^2 + \gamma_2|\gamma|^2)G^H e_2\|_2 = |\gamma\sigma(\gamma_1 - \gamma_2)| \, .$$

If $\gamma_1 = \gamma_2$ then we have determined an eigenpair exactly. In case $\lambda_1 \ne \lambda_2$ we note that $|\gamma\sigma(\lambda_1 - \lambda_2)| \le \frac{1}{2}|\lambda_1 - \lambda_2|$, and, moreover, if $|\gamma| \approx 0$ or $|\gamma| \approx 1$ then $|\gamma\sigma(\lambda_1 - \lambda_2)| \ll |\lambda_1 - \lambda_2|$. Hence, inequality (2.26) may be satisfied, even if $|\lambda_1 - \lambda_2| > \epsilon_2$. Assume that (2.26) is valid. Then $\tilde{\Lambda}$ is replaced by

14

$$\hat{\Lambda} := \operatorname{diag}[\lambda_1(\gamma^2) + \lambda_2\sigma^2, \lambda_3, \lambda_4, \ldots, \lambda_n] \in \mathbb{C}^{(n-1)\times(n-1)}$$

and if $\hat{\Lambda}$ has close eigenvalues, then deflation is repeated.

The unitary DC method can be used in two ways. One approach is to divide the original eigenproblem, as well as subproblems so obtained, until only trivial eigenproblems of orders two and one remain. These small eigenproblems are solved analytically. From the solutions of small eigenproblems, the solutions of eigenproblems of larger size are computed, and this step is repeated until the solution of the original eigenproblem has been determined. This approach is used in the numerical examples of Section 5.

An alternative approach is to use the DC technique to generate just a few subeigenproblems, each of which can be solved independently by some other numerical scheme, such as the unitary GR method [Gr1], or the scheme in [AGR1], in case the matrix is real orthogonal.

We conclude this section with some bounds of the computational complexity of the unitary DC method. Assume that $H \in \mathbb{C}^{n\times n}$ is given in Schur parametric form (1.1) with positive subdiagonal elements σ_j. Let $n = 2^\ell$ for some positive integer ℓ, and subdivide the given eigenproblem until $\frac{n}{2}$ eigenproblems for 2×2 matrics are obtained. The latter eigenproblems are solved analytically. We assume that the number of iterations required by the rootfinder for $\Phi(\theta)$ can be bounded independently of n.

Let first n independent processors be available. The reduction of the original eigenproblem for H to $\frac{n}{2}$ eigenproblems for 2×2 matrices can be carried out in $t_1 := O(\log_2\frac{n}{2})$ time steps. This computation only requires the determination of the Schur parameters for the unitary matrices of the smaller eigenproblems, see (2.3). Let the Schur parameters for all $\frac{n}{2}$ unitary 2×2 matrices be known. The spectral resolution of all these

15

matrices can be computed in $t_2 := O(1)$ time steps. Assume that the partial spectral resolutions (2.21) of all $2^{\ell-j+1}$ unitary $2^{j-1} \times 2^{j-1}$ matrices are known for some $j \in [2,\ell]$. In order to compute the partial spectral resolution of all $2^{\ell-j}$ unitary $2^j \times 2^j$ matrices, we have to compute 2^j zeros of each of the $2^{\ell-j}$ functions $\Phi(\theta)$, see (2.8). Hence, a total number of n zeros have to be computed, and we use one processor to determine each one. Each function $\Phi(\theta)$ has 2^j terms, and can therefore be evaluated in $O(2^j)$ time steps for each value of θ. Hence, we can determine all eigenvalues of all $2^{\ell-j}$ unitary $2^j \times 2^j$ matrices in $t_3^{(j)} := O(2^j)$ time steps. For each eigenvalue we compute the first and last elements of the corresponding eigenvector from (2.21). The first and last element of one eigenvector can be determined by one processor in $O(2^j)$ time steps. These computations have to be carried out for n eigenvectors by n processors and therefore require $t_4^{(j)} = O(2^j)$ time steps. Hence, the number of time steps required to determine the partial spectral resolution of H by n processors is

$$t_1 + t_2 + \sum_{j=2}^{\ell} t_3^{(j)} + \sum_{j=2}^{\ell} t_4^{(j)} = O(n) \quad . \tag{2.27}$$

Now let n^2 independent processors be available, and assume that the partial spectral resolutions of all $2^{\ell-j+1}$ unitary $2^{j-1} \times 2^{j-1}$ matrices are known for some $j \in [2,\ell]$. We have now n processors available for each evaluation of each of the $2^{\ell-j}$ functions $\Phi(\theta)$. Each of these functions $\Phi(\theta)$ has 2^j terms and can for each value of θ be evaluated in $O(\log_2 2^j)$ time steps. Hence, we can compute all eigenvalues of all $2^{\ell-j}$ unitary $2^j \times 2^j$ matrices in $\tilde{t}_3^{(j)} := O(\log_2 2^j)$ time steps. The first and last elements of each eigenvector of each of the $2^{\ell-j}$ unitary $2^j \times 2^j$ matrices can be

16

determined in $\tilde{t}_4^{(j)} := O(\log_2 2^j)$ time steps, by using n processors to compute each sum with $2^j \leq n$ terms. The initial determination of the $\frac{n}{2}$ unitary 2 × 2 matrices and their spectral resolutions cannot be sped up essentially by using more than n processors, and requires $O(t_1) + O(t_2)$ time steps. Hence, the number of time steps required in order to compute the partial spectral resolution of H by n^2 processors is

$$O(t_1) + O(t_2) + \sum_{j=2}^{\ell} \tilde{t}_3^{(j)} + \sum_{j=2}^{\ell} \tilde{t}_4^{(j)} = O(\log_2 n) + \sum_{j=2}^{\ell} O(j) = O(\log_2^2 n) . \qquad (2.28)$$

The time complexities (2.27)-(2.28) suggest that the unitary DC method presented could be attractive for use in real-time signal processing applications.

3. Some properties of the unitary DC method

We show some properties of the eigenvectors of H and zeros of $\Phi(\theta)$. Analogous results have previously been obtained by Dongarra and Sorensen [DS, Lemmas 4.2, 4.6 and 4.7] for the DC method for the eigenproblem for symmetric tridiagonal matrices. The following formulas are used in several of the proofs:

$$\phi(\lambda) = 1 + 2 \sum_{j=1}^{n} |\zeta_j|^2 \frac{\lambda_j}{\lambda - \lambda_j} , \tag{3.1}$$

$$\phi'(\lambda) = -2 \sum_{j=1}^{n} |\zeta_j|^2 \frac{\lambda_j}{(\lambda - \lambda_j)^2} , \tag{3.2}$$

$$\Phi'(\theta) = \phi'(\lambda)\lambda . \tag{3.3}$$

<u>Lemma 3.1</u> Let $\lambda_1 \mu \in \mathbb{C}$, $|\lambda| = |\mu| = 1$ and $\lambda \neq \mu$. Assume that $\lambda_1 \mu \notin \{\lambda_j\}_{j=1}^{n}$, where $\tilde{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Let v_λ and v_μ be defined by (2.20). Then

$$|v_\lambda^H v_\mu| = |\phi'(\lambda)\phi'(\mu)|^{-1/2} \left| \frac{\phi(\lambda) - \phi(\mu)}{\lambda - \mu} \right| \tag{3.4}$$

$$= \left| \Phi'(\theta_\lambda)\Phi'(\theta_\mu) \right|^{-1/2} \left| \frac{\Phi(\theta_\lambda) - \Phi(\theta_\mu)}{e^{i\theta_\lambda} - e^{i\theta_\mu}} \right| ,$$

where $\lambda = e^{i\theta_\lambda}$, $\mu = e^{i\theta_\mu}$, $0 \leq \theta_\lambda, \theta_\mu < 2\pi$. In particular, if λ and μ are distinct eigenvalues of H, then $\phi(\lambda) = \phi(\mu) = 0$, and therefore $v_\lambda^H v_\mu = 0$.

Proof. By (2.20), (2.6) and (2.7),

$$v_\lambda = \begin{bmatrix} W_1 & \\ & W_2 \end{bmatrix} \begin{bmatrix} \Lambda_1 & \\ & I_{n-s} \end{bmatrix} (\tilde{\Lambda} - I\lambda)^{-1} z / \delta(\lambda) , \tag{3.5}$$

and therefore

18

$$v_\lambda{}^H v_\mu = \frac{z^H}{\delta(\lambda)}(\tilde\Lambda - I\bar\lambda)^{-1}(\tilde\Lambda - I\mu)^{-1}\frac{z}{\delta(\mu)}$$

$$= (\delta(\lambda)\delta(\mu))^{-1}\sum_{j=1}^{n}\frac{|\zeta_j|^2}{(\bar\lambda_j - \bar\lambda)(\lambda_j - \mu)} \quad . \tag{3.6}$$

Now

$$\frac{1}{(\bar\lambda_j-\bar\lambda)(\lambda_j-\mu)} = -\frac{\lambda\lambda_j}{(\lambda_j-\lambda)(\lambda_j-\mu)} = -\frac{\lambda}{\lambda-\mu}\Big(\frac{\lambda_j}{\lambda_j-\lambda} - \frac{\lambda_j}{\lambda_j-\mu}\Big) \quad . \tag{3.7}$$

Substituting (3.7) into (3.6) yields

$$v_\lambda{}^H v_\mu = (2\delta(\lambda)\delta(\mu))^{-1}\frac{\lambda}{\lambda+\mu}\Big(2\sum_{j=1}^{n}\frac{|\zeta_j|^2\lambda_j}{\lambda-\lambda_j} - 2\sum_{j=1}^{n}\frac{|\zeta_j|^2\lambda_j}{\mu-\lambda_j}\Big) \quad ,$$

and by (2.19), (3.1) and (3.3),

$$v_\lambda{}^H v_\mu = (\phi'(\lambda)\phi(\mu)\lambda\mu)^{-1/2}\lambda\,\frac{\phi(\lambda) - \phi(\mu)}{\lambda - \mu}$$

$$= (\Phi'(\theta_\lambda)\Phi'(\theta_\mu))^{-1/2}\,ie^{i\theta_\lambda}\,\frac{\Phi(\theta_\lambda) - \Phi(\theta_\mu)}{e^{i\theta_\lambda} - e^{i\theta_\mu}} \quad .$$

This shows (3.4).   □

The denominator $|\lambda-\mu|$ in (3.4) suggests that it may be numerically difficult to obtain orthogonal eigenvectors when the associated eigenvalues are very close. The following lemma sheds some light on this situation, and shows that due to deflation the roots of $\Phi(\theta)$ are bounded away from each other.

<u>Lemma 3.2</u> Let $\lambda_j = e^{i\theta_j}$, $1 \leq j \leq n$, be the eigenvalues of $\tilde\Lambda$, and let $z = [\zeta_j]_{j=1}^{n}$ be defined by (2.7). Let ϵ_1 be an arbitrary but fixed positive

constant and assume that the λ_j are pairwise distinct and that $2|\zeta_j| > \epsilon_1$ for all j. These conditions can be made valid by deflation. Assume that the λ_j are sorted so that $0 \leq \theta_1 < \theta_2 < \ldots < \theta_n < 2\pi$, and let $\theta_{n+1} := 2\pi + \theta_1$. Let $\theta \in [0, 2\pi[$ be a zero of $\Phi(\theta)$, and let k be such that $\theta_k < \theta < \theta_{k+1}$. Then

$$\theta - \theta_k \geq \tfrac{1}{2}(\theta_{k+1} - \theta_k) \Rightarrow \theta_{k+1} - \theta \geq \min\{\tfrac{\epsilon_1^2}{16}(\theta_{k+1} - \theta_k), \tfrac{2\pi}{3}\} \, , \tag{3.8}$$

$$\theta_{k+1} - \theta \geq \tfrac{1}{2}(\theta_{k+1} - \theta_k) \Rightarrow \theta - \theta_k \geq \min\{\tfrac{\epsilon_1^2}{16}(\theta_{k+1} - \theta_k), \tfrac{2\pi}{3}\} \, . \tag{3.9}$$

Proof. Introduce the index sets

$$I_1 := \{j: \theta < \theta_j + 2\pi\ell \leq \theta + \pi, \text{ for some } \ell \in \mathbb{Z}, \ 1 \leq j \leq n\} \, ,$$

$$I_2 := \{j: \theta - \pi < \theta_j + 2\pi\ell < \theta, \text{ for some } \ell \in \mathbb{Z}, \ 1 \leq j \leq n\} \, .$$

Then $I_1 \cap I_2 = \emptyset$ and $I_1 \cup I_2 = \{1, 2, \ldots, n\}$. Further

$$i \, \frac{\lambda + \lambda_j}{\lambda - \lambda_j} = \cot\left(\frac{\theta - \theta_j}{2}\right) \begin{cases} \leq 0 \, , & j \in I_1, \\ > 0 \, , & j \in I_2. \end{cases}$$

In particular, $k \in I_2$ and, provided that $k < n$, $k+1 \in I_1$. If $k = n$ then $1 \in I_1$. Moreover,

$$\begin{cases} \cot\left(\frac{\theta - \theta_{k+1}}{2}\right) \leq \cot\left(\frac{\theta - \theta_j}{2}\right) \, , & \forall j \in I_1 \\ \\ \cot\left(\frac{\theta - \theta_k}{2}\right) \geq \cot\left(\frac{\theta - \theta_j}{2}\right) \, , & \forall j \in I_2 \end{cases} \tag{3.10}$$

From $\Phi(\theta) = 0$, it follows that

$$-\sum_{j \in I_1} |\zeta_j|^2 \cot\left(\frac{\theta - \theta_j}{2}\right) = \sum_{j \in I_2} |\zeta_j|^2 \cot\left(\frac{\theta - \theta_j}{2}\right) \, . \tag{3.11}$$

20

By (3.10), (3.11) and $z^H z = 1$ we obtain, provided that $k < n$,

$$- |\zeta_{k+1}|^2 \cot\left(\frac{\theta - \theta_{k+1}}{2}\right) \le \cot\left(\frac{\theta - \theta_k}{2}\right) , \qquad (3.12)$$

or, equivalently,

$$|\zeta_{k+1}|^2 \tan\left(\frac{\theta - \theta_k}{2}\right) \le \tan\left(\frac{\theta_{k+1} - \theta}{2}\right) . \qquad (3.13)$$

If $k = n$ then we define $\zeta_{n+1} := \zeta_n$ and (3.12)-(3.13) remain valid.

Now assume that

$$\theta - \theta_k \ge \frac{1}{2}(\theta_{k+1} - \theta_k) . \qquad (3.14)$$

We wish to determine a lower bound for $\theta_{k+1} - \theta$. From $\tan \frac{x}{2} \le x$ for $0 \le x \le \frac{2\pi}{3}$ it follows that if $0 \le \theta_{k+1} - \theta \le \frac{2\pi}{3}$, then

$$\tan\left(\frac{\theta_{k+1} - \theta}{2}\right) \le \theta_{k+1} - \theta . \qquad (3.15)$$

Substituting (3.14)-(3.15) into (3.13) yields

$$|\zeta_{k+1}|^2 \tan(\tfrac{1}{4}(\theta_{k+1} - \theta_k)) \le \theta_{k+1} - \theta ,$$

and from $\tan(\frac{1}{4}(\theta_{k+1} - \theta_k)) \ge \frac{1}{4}(\theta_{k+1} - \theta_k)$, we obtain

$$|\zeta_{k+1}|^2 \tfrac{1}{4}(\theta_{k+1} - \theta_k) \le \theta_{k+1} - \theta . \qquad (3.16)$$

Finally, substituting $|\zeta_{k+1}| \ge \frac{\epsilon_1}{2}$ into (3.16) yields (3.8).

In order to show (3.9), we note that from (3.10)-(3.11) and $z^H z = 1$ it follows that

$$- \cot\left(\frac{\theta - \theta_{k+1}}{2}\right) \ge |\zeta_k|^2 \cot\left(\frac{\theta - \theta_k}{2}\right)$$

or, equivalently,

$$\tan\left(\frac{\theta - \theta_k}{2}\right) \ge |\zeta_k|^2 \tan\left(\frac{\theta_{k+1} - \theta}{2}\right) , \qquad (3.17)$$

21

which corresponds to (3.13). We now assume that

$$\theta_{k+1} - \theta \geq \frac{1}{2}(\theta_{k+1} - \theta_k) \ . \tag{3.18}$$

We would like to determine a lower bound for $\theta - \theta_k$. Similarly as in the derivation of (3.15) we obtain that if $0 \leq \theta - \theta_k \leq \frac{2\pi}{3}$, then

$$\tan\left(\frac{\theta - \theta_k}{2}\right) \leq \theta - \theta_k \ . \tag{3.19}$$

From (3.17)-(3.19) and $\tan(\frac{1}{2}(\theta_{k+1}-\theta)) \geq \frac{1}{2}(\theta_{k+1}-\theta)$, we obtain

$$\theta - \theta_k \geq |\zeta_k|^2 \frac{1}{4}(\theta_{k+1} - \theta_k) \ . \tag{3.20}$$

Finally, substituting $|\zeta_k| \geq \frac{\epsilon_1}{2}$ into (3.20) yields (3.9).   □

Our final lemma shows that the computed eigenvectors are close to orthogonal if the zeros of $\Phi(\theta)$ are evaluated with sufficient accuracy.

<u>Lemma 3.3</u>.   Let $\tilde{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$, and let $\hat{\lambda}$, $\hat{\mu}$ be computed approximations of the distinct roots λ, μ of ϕ. Introduce the relative errors α_k, β_k of $\lambda_k - \hat{\lambda}$ and $\lambda_k - \hat{\mu}$, respectively, i.e.

$$\begin{cases} \lambda_k - \hat{\lambda} = (\lambda_k - \lambda)(1 + \alpha_k) \\ \lambda_k - \hat{\mu} = (\lambda_k - \mu)(1 + \beta_k) \ . \end{cases} \qquad k = 1, 2, \ldots, n \ . \tag{3.21}$$

Assume that for some constant $0 \leq \epsilon < 1$, $|\alpha_k| \leq \epsilon$ and $|\beta_k| \leq \epsilon$ for all k, and that $|\hat{\lambda}| = |\hat{\mu}| = 1$. Then

$$|v_{\hat{\lambda}}^H v_{\hat{\mu}}| = |v_\lambda^H C v_\mu| \leq \epsilon(2 + \epsilon)\left(\frac{1 + \epsilon}{1 - \epsilon}\right)^2 \ ,$$

where $C = \text{diag}[\rho_1, \rho_2, \ldots, \rho_n]$ with

22

$$\rho_k := -\frac{\bar{a}_k + \beta_k + \bar{a}_k \beta_k}{(1 + \bar{a}_k)(1 + \beta_k)}\left(\frac{\delta(\lambda)\delta(\mu)}{\delta(\hat{\lambda})\delta(\hat{\mu})}\right) .$$

(3.22)

For $\eta := e^{i\theta_\eta}, 0 \le \theta_\eta < 2\pi$, we define $\delta(\eta) := (\frac{1}{2}\phi'(\eta)\eta)^{1/2} = \left(\frac{1}{2}\Phi'(\theta_\eta)\right)^{1/2}$.

Proof. We first note that since $\hat{\lambda}$, $\hat{\mu}$ are computed by determining zeros of $\Phi(\theta)$, $0 \le \theta < 2\pi$, the requirement $|\hat{\lambda}| = |\hat{\mu}| = 1$ is satisfied. Analogously to (3.6) we obtain

$$v_{\hat{\lambda}}^H v_{\hat{\mu}} = (\delta(\hat{\lambda})\delta(\hat{\mu}))^{-1}\sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\hat{\lambda}})(\lambda_k - \hat{\mu})}$$

$$= (\delta(\hat{\lambda})\delta(\hat{\mu}))^{-1}\sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\lambda})(\lambda_k - \mu)(1 + \bar{a}_k)(1 + \beta_k)} ,$$

where the last equality follows from (3.21). Now

$$0 = v_\lambda^H v_\mu = (\delta(\lambda)\delta(\mu))^{-1}\sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\lambda})(\lambda_k - \mu)}$$

and (2.19) imply that

$$\sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\lambda})(\lambda_k - \mu)} = 0 .$$

Therefore

$$v_{\hat{\lambda}}^H v_{\hat{\mu}} = (\delta(\hat{\lambda})\delta(\hat{\mu}))^{-1}\left(\sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\lambda})(\lambda_k - \mu)(1 + \bar{a}_k)(1 + \beta_k)}\right.$$

$$\left. - \sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\lambda})(\lambda_k - \mu)}\right)$$

$$= (\delta(\hat{\lambda})\delta(\hat{\mu}))^{-1}\sum_{k=1}^{n}\frac{|\zeta_k|^2}{(\bar{\lambda}_k - \bar{\lambda})(\lambda_k - \mu)}\left(\frac{1}{(1 + \bar{a}_k)(1 + \beta_k)} - 1\right) ,$$

(3.23)

23

which shows that $v_{\hat{\lambda}}{}^H v_{\hat{\mu}} = v_\lambda{}^H C v_\mu$ with C defined by (3.22). From (2.19), (3.3), (3.2) and (3.21) it follows that

$$
\begin{aligned}
\left(\frac{\delta(\lambda)}{\delta(\hat{\lambda})}\right)^2 &= \frac{\phi'(\lambda)\lambda}{\phi'(\hat{\lambda})\hat{\lambda}} \\
&= \frac{\displaystyle\sum_{k=1}^{n} |\zeta_k|^2 \frac{(-\lambda_k)\lambda}{(\lambda-\lambda_k)^2}}{\displaystyle\sum_{k=1}^{n} \left(|\zeta_k|^2 \frac{(-\lambda_k)\lambda}{(\lambda-\lambda_k)^2} \cdot \frac{\hat{\lambda}/\lambda}{(1+\alpha_k)^2}\right)} \quad .
\end{aligned}
\tag{3.24}
$$

From $(-\lambda_k\lambda)/(\lambda-\lambda_k)^2 > 0$ it follows that $\hat{\lambda}/(\lambda(1+\alpha_k)^2) > 0$ and therefore

$$
\frac{\hat{\lambda}/\lambda}{(1 + \alpha_k)^2} \geq (1 + |\alpha_k|)^{-2} \geq (1 + \epsilon)^{-2} \quad .
\tag{3.25}
$$

Substituting (3.25) into (3.24) yields $\delta(\lambda)/\delta(\hat{\lambda}) \leq 1 + \epsilon$, and similarly one can show that $\delta(\mu)/\delta(\hat{\mu}) \leq 1 + \epsilon$. Hence, by (3.22),

$$
|\rho_k| \leq \frac{\epsilon + \epsilon + \epsilon^2}{(1 - \epsilon)^2} (1 + \epsilon)^2 = \epsilon(2 + \epsilon)\left(\frac{1 + \epsilon}{1 - \epsilon}\right)^2 \quad .
\tag{3.26}
$$

Finally,

$$
|v_{\hat{\lambda}}{}^H v_{\hat{\mu}}| = |v_\lambda{}^H C v_\mu| \leq \|v_\lambda\|_2 \|C\|_2 \|v_\mu\|_2 = \|C\|_2 = \max_{1 \leq k \leq n} |\rho_k| \quad ,
$$

and the desired bound now follows from (3.26). $\qquad \Box$

4. The orthogonal eigenproblem

The computational work required for the real orthogonal eigenproblem is smaller than for the unitary one. This section discusses these differences, and considers some details of our implementation of a DC scheme for the real orthogonal eigenproblem. Our computer program is for the case when $H \in \mathbb{R}^{n \times n}$ with $n = 2^{\ell}$, where ℓ is a positive integer, and we assume in this section that n is of this form. Many of our comments are valid for more general values of n, also.

We first note that the subdivision of the eigenproblem for H into smaller eigenproblems, as described by (1.3)-(1.7), does not require any computational work. Subdivision yields the block-diagonal matrix

$$\hat{H} := G_1 G_3 G_5 \ldots G_{n-5} G_{n-3} G_{n-1} G_n \ , \tag{4.1}$$

and we obtain simple formulas for the eigenpairs of each 2x2 block on the diagonal as follows. Let

$$G := \begin{bmatrix} -\gamma & \sigma \\ \sigma & \gamma \end{bmatrix} \in \mathbb{R}^{2 \times 2} \ , \quad -1 < \gamma < 1, \ \sigma > 0, \ \gamma^2 + \sigma^2 = 1 \ . \tag{4.2}$$

Since G is real, symmetric, orthogonal and has distinct eigenvalues $\{\lambda_1, \lambda_2\}$, we have $\lambda_1 = 1$ and $\lambda_2 = -1$. Let $x_1 = [\xi_1, \xi_2]^T$ be an eigenvector of unit length. Then we can choose

$$\begin{cases} \xi_1 = \sigma 2^{-1/2} (1 + \gamma)^{-1/2} \\ \xi_2 = 2^{-1/2} (1 + \gamma)^{1/2}, \end{cases} \tag{4.3}$$

and from $\sigma = (1 - \gamma^2)^{1/2}$ it follows that

$$\begin{cases} \xi_1 = 2^{-1/2} (1 - \gamma)^{1/2} \\ \xi_2 = \sigma 2^{-1/2} (1 - \gamma)^{-1/2} \ . \end{cases} \tag{4.4}$$

25

Cancellation of significant digits is avoided by using (4.3) if $\gamma \geq 0$ and (4.4) otherwise. An eigenvector associated with $\lambda_2 = -1$ is given by $x_2 := [\xi_2, -\xi_1]^T$.

If $\gamma_n = -1$ then $G_n = I_n$, and the eigenpairs of $G_{n-1}G_n$ are those of G_{n-1}. If $\gamma_\eta = 1$ we need to determine the eigenpairs of

$$G' := \begin{bmatrix} -\gamma_{n-1} & -\sigma_{n-1} \\ \sigma_{n-1} & -\gamma_{n-1} \end{bmatrix} .$$

We find that the eigenvalue $\lambda_1 = -\gamma_{n-1} + i\sigma_{n-1}$ of G' has an associated eigenvector $x_1 := [2^{-1/2}, -2^{-1/2}]^T$, and the eigenvalue $\lambda_2 = -\gamma_{n-1} - i\sigma_{n-1}$ has an associated eigenvector $x_2 := [2^{-1/2}, 2^{-1/2}]^T$.

Note that since the eigenvalues of G given by (4.2) are $\lambda = \pm 1$, independent of $-1 < \gamma < 1$, deflation takes place numerous times during the computations.

We turn to the computation of the Householder transformation (1.4). In oder to avoid cancellation of significant digits, we compute $\{\omega_s, \omega_{s+1}\}$ given by (1.6) as follows. If $\gamma_s \geq 0$, then we use (1.6a) and replace (1.6b) by

$$\omega_{s+1} := -\sigma_s 2^{-1/2}(1 + \gamma_s)^{-1/2} . \qquad (1.6b')$$

In case $\gamma_s < 0$, we use (1.6b) and replace (1.6a) by

$$\omega_s := \sigma_s 2^{-1/2}(1 - \gamma_s)^{-1/2} . \qquad (1.6a')$$

Due to H having real-valued elements, the eigenvalues and eigenvectors of H occur in complex conjugate pairs. Therefore only zeros of $\Phi(\theta)$ for $0 \leq \theta \leq \pi$ have to be computed. Moreover,

$$\Phi(\theta) = \sum_{j=1}^{n} |\zeta_j|^2 \cot\left(\frac{\theta_j - \theta}{2}\right)$$

can be simplified. Assume that $0 < \theta_k < \pi$ for some $k < n$ and let $\theta_{k+1} = 2\pi - \theta_k$. Then $|\zeta_k| = |\zeta_{k+1}|$, and we obtain

$$|\zeta_k|^2 \cot\left(\frac{\theta_k - \theta}{2}\right) + |\zeta_{k+1}|^2 \cot\left(\frac{\theta_{k+1} - \theta}{2}\right) = |\zeta_k|^2\left(\cot\left(\frac{\theta_k - \theta}{2}\right) - \cot\left(\frac{\theta_k + \theta}{2}\right)\right)$$

$$= |\zeta_k|^2 \frac{2 \sin \theta}{\cos \theta - \cos \theta_k} = |\zeta_k|^2 \frac{\sin \theta}{\sin\left(\frac{\theta_k + \theta}{2}\right)\sin\left(\frac{\theta_k - \theta}{2}\right)} \qquad (4.5)$$

We use the right hand side of (4.5) in the computations. If $\theta_k = 0$ then we need to evaluate $\cot(-\frac{\theta}{2})$ as well as $\cot\left(\frac{\pi - \theta}{2}\right) = \tan \frac{\theta}{2}$.

The contribution from (4.5) to $\Phi'(\theta)$ is

$$2|\zeta_k|^2 \frac{d}{d\theta}\left(\frac{\sin \theta}{\cos \theta - \cos \theta_k}\right) = 2|\zeta_k|^2 \frac{1 - \cos \theta \cos \theta_k}{(\cos \theta - \cos \theta_k)^2} \qquad (4.6)$$

The stable evaluation of the right hand side of (4.6) can be accomplished as described in Table 4.1.

Conditions	Evaluate
$cc_k \leq 0$	$\frac{1}{4}(s_+^{-2} + s_-^{-2})$
$cc_k > 0$ and $\frac{c}{s_+ s_-} > 0$	$\frac{c}{2s_+ s_-} + \left(\frac{s}{2s_+ s_-}\right)^2$
$cc_k > 0$ and $\frac{c}{s_+ s_-} < 0$	$-\frac{c_k}{2s_+ s_-} + \left(\frac{s_k}{2s_+ s_-}\right)^2$

Table 4.1:   Stable evaluation of $(1 - cc_k)/(c - c_k)^2$, where
$c := \cos \theta$, $c_k := \cos \theta_k$, $s := \sin \theta$, $s_k := \sin \theta_k$, $s_+ := \sin\left(\frac{\theta_k + \theta}{2}\right)$,
$s_- := \sin\left(\frac{\theta_k - \theta}{2}\right)$.

The interlacing of the zeros of $\phi(\lambda)$ with the $\{\lambda_k\}_{k=1}^n$ implies that it easily can be determined whether $\theta = 0$ or $\theta = \pi$ are zeros of $\Phi(\theta)$. Let

27

the θ_k, $1 \leq k \leq n$, be ordered so that $0 \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_p \leq \pi < \theta_{p+1} \leq \ldots \leq \theta_n < 2\pi$. Since the $\lambda_k = \exp(i\theta_k)$ appear in complex conjugate pairs, we obtain

$$\begin{cases} \theta_1 > 0 & \Rightarrow & \Phi(0) = 0 \ , \\ \theta_p < \pi & \Rightarrow & \Phi(\pi) = 0 \ . \end{cases}$$

Finally, we consider the computation of eigenvectors v_λ defined by (2.20). Let

$$W_1 =: [w_1^{(1)}, w_2^{(1)}, \ldots, w_s^{(1)}], \qquad w_j^{(1)} \in \mathbb{C}^s \ ,$$

$$W_2 =: [w_1^{(2)}, w_2^{(2)}, \ldots, w_{n-s}^{(2)}], \qquad w_j^{(2)} \in \mathbb{C}^{n-s} \ ,$$

$$\Lambda_1 =: \mathrm{diag}[\exp(i\theta_1^{(1)}), \exp(i\theta_2^{(1)}), \ldots, \exp(i\theta_s^{(1)})], \qquad 0 \leq \theta_j^{(1)} < 2\pi \ ,$$

$$\Lambda_2 =: \mathrm{diag}[\exp(i\theta_1^{(2)}), \exp(i\theta_2^{(2)}), \ldots, \exp(i\theta_{n-s}^{(2)})], \qquad 0 \leq \theta_j^{(2)} < 2\pi \ ,$$

$$W_1^H e_s \omega_s =: [\zeta_1^{(1)}, \zeta_2^{(1)}, \ldots, \zeta_s^{(1)}]^\top \ ,$$

$$W_2^H e_1 \omega_{s+1} =: [\zeta_1^{(2)}, \zeta_2^{(2)}, \ldots, \zeta_{n-s}^{(2)}]^\top \ ,$$

and $\lambda =: \exp(i\theta)$, $0 \leq \theta < 2\pi$. Then

$$W_1(I - \Lambda_1^H \lambda)^{-1} W_1^H e_s \omega_s = \sum_{j=1}^{s} (1 - \exp(i(\theta - \theta_j^{(1)})))^{-1} \zeta_j^{(1)} w_j^{(1)}$$

$$= \sum_{\theta_j^{(1)} \in \{0, \pi\}} (1 - \exp(i(\theta - \theta_j^{(1)})))^{-1} \zeta_j^{(1)} w_j^{(1)}$$

$$+ \sum_{0 < \theta_j^{(1)} < \pi} [(1 - \exp(i(\theta - \theta_j^{(1)})))^{-1} \zeta_j^{(1)} w_j^{(1)} + (1 - \exp(i(\theta + \theta_j^{(1)})))^{-1} \overline{\zeta}_j^{(1)} \overline{w}_j^{(1)}]$$

$$= \frac{1}{2} \sum_{\theta_j^{(1)}=0} (1+i\cot(\tfrac{\theta}{2}))\zeta_j^{(1)}w_j^{(1)} + \frac{1}{2} \sum_{\theta_j^{(1)}=\pi} (1-i\tan(\tfrac{\theta}{2}))\zeta_j^{(1)}w_j^{(1)}$$

$$+ \sum_{0<\theta_j^{(1)}<\pi} \left[(\mathrm{Re}(\zeta_j^{(1)}w_j^{(1)})) + \frac{\sin\theta_j^{(1)}}{2\,\sin\left(\frac{\theta_j^{(1)}+\theta}{2}\right)\sin\left(\frac{\theta_j^{(1)}-\theta}{2}\right)}\,\mathrm{Im}(\zeta_j^{(1)}\zeta_j^{(1)}) \right] \qquad (4.7)$$

$$- \frac{i}{2}\sin\theta \sum_{0<\theta_j^{(1)}<\pi} \left(\sin\left(\frac{\theta_j^{(1)}+\theta}{2}\right)\sin\left(\frac{\theta_j^{(1)}-\theta}{2}\right) \right)^{-1} \mathrm{Re}(\zeta_j^{(1)}w_j^{(1)}) \;.$$

We may assume that close eigenvalues have been eliminated from $\Lambda^{(1)}$ and $\Lambda^{(2)}$ by deflation, and that therefore the $\theta_j^{(1)}$ and $\theta_j^{(2)}$ are distinct. Hence, the sums over $\theta_j^{(1)} = 0$ and $\theta_j^{(1)} = \pi$ contain at most one term each.

Analogously to (4.7) we obtain

$$W_2(\Lambda_2-I\lambda)^{-1}W_2{}^H e_1\omega_{s+1} = \sum_{j=1}^{n-s} (\exp(i\theta_j^{(2)})-\exp(i\theta))^{-1}\zeta_j^{(2)}w_j^{(2)}$$

$$= \frac{1}{2} \sum_{\theta_j^{(2)}=0} (1+i\cot(\tfrac{\theta}{2}))\zeta_j^{(2)}w_j^{(2)} + \frac{1}{2} \sum_{\theta_j^{(2)}=\pi} (-1+i\tan(\tfrac{\theta}{2}))\zeta_j^{(2)}w_j^{(2)}$$

$$+ \frac{1}{2} \sum_{0<\theta_j^{(2)}<\pi} \left(\frac{\sin\left(\frac{\theta+\theta_j^{(2)}}{2}\right)}{\sin\left(\frac{\theta-\theta_j^{(2)}}{2}\right)} + \frac{\sin\left(\frac{\theta-\theta_j^{(2)}}{2}\right)}{\sin\left(\frac{\theta+\theta_j^{(2)}}{2}\right)} \right) \mathrm{Re}(\zeta_j^{(2)}w_j^{(2)}) \qquad (4.8)$$

$$- \frac{1}{2}e^{-i\theta} \sum_{0<\theta_j^{(2)}<\pi} \sin\theta_j^{(2)} \left(\sin\left(\frac{\theta+\theta_j^{(2)}}{2}\right)\sin\left(\frac{\theta-\theta_j^{(2)}}{2}\right) \right)^{-1} \mathrm{Im}(\zeta_j^{(2)}w_j^{(2)})$$

$$+ \frac{i}{2}\sin\theta \sum_{0<\theta_j^{(2)}<\pi} \cos\theta_j^{(2)} \left(\sin\left(\frac{\theta+\theta_j^{(2)}}{2}\right)\sin\left(\frac{\theta-\theta_j^{(2)}}{2}\right) \right)^{-1} \mathrm{Re}(\zeta_j^{(2)}w_j^{(2)}) \;.$$

The evaluation of $\delta(\lambda)$ defined by (2.19) can also be simplified. We have

$$\delta(\lambda) = \left(\sum_{j=1}^{s} \frac{|\zeta_j^{(1)}|^2}{|\lambda - \exp(i\theta_j^{(1)})|^2} + \sum_{j=1}^{n-s} \frac{|\zeta_j^{(2)}|^2}{|\lambda - \exp(i\theta_j^{(2)})|^2} \right)^{1/2} ,$$

where, e.g.,

$$\sum_{j=1}^{s} \frac{|\zeta_j^{(1)}|^2}{|\lambda - \exp(i\theta_j^{(1)})|^2} = \frac{1}{4} \sum_{\theta_j^{(1)}=0} |\zeta_j^{(1)}|^2 / \sin^2(\tfrac{\theta}{2}) + \frac{1}{4} \sum_{\theta_j^{(1)}=\pi} |\zeta_j^{(1)}|^2 / \cos^2(\tfrac{\theta}{2})$$

$$+ \frac{1}{4} \sum_{0 < \theta_j^{(1)} < \pi} |\zeta_j^{(1)}|^2 \left[\left(\sin\left(\tfrac{\theta - \theta_j^{(1)}}{2} \right) \right)^{-2} + \left(\sin^2\left(\tfrac{\theta + \theta_j^{(1)}}{2} \right) \right)^{-2} \right] .$$

(4.9)

The simplifications of this section for the orthogonal eigenproblem have been implemented in a Pascal program. Several other mathematically equivalent forms of (4.7)-(4.9) could also be used. We have tried to find formulas that avoid unnecessary loss of significant digits.

30

5. Numerical examples

We report results of some computed examples with an experimental program for the orthogonal eigenproblem. The program is written in Turbo Pascal 4.0 and was run on an IBM PC AT computer with unit roundoff $u = 2^{-39} \approx 2 \cdot 10^{-12}$. Our code implements the formulas of Section 4. Generally very accurate answers are obtained. Lemma 3.2 indicates, however, that a zero $\tilde{\theta}$ of $\Phi(\theta)$ may be very close to a singular point θ_j of $\Phi(\theta)$ and by Lemma 3.3 the difference $\tilde{\theta} - \theta_j$ has to be computed to high _relative_ accuracy in order to yield nearly orthogonal eigenvectors. Example 5.2 below shows that, indeed, $\tilde{\theta} - \theta_j$ can be extremely tiny and that loss of accuracy in both eigenvectors and eigenvalues may result. This loss of accuracy could be reduced, e.g., by representing $\tilde{\theta}$ and θ_j in higher precision arithmetic.

In this section $\Lambda \in \mathbb{C}^{n \times x}$ denotes the diagonal matrix with the computed eigenvalues of $H \in \mathbb{R}^{n \times n}$ as entries, and $W \in \mathbb{C}^{n \times n}$ is the matrix with the computed eigenvectors. We evaluate the residual errors $\|HW - W\Lambda\|_\infty$ and $\|W^H W - I\|_\infty$, where $\| \ \|_\infty$ denotes the uniform matrix norm.

Example 5.1. This example discusses the application of the unitary and orthogonal eigenproblems to the construction of Gauss–Szegö quadrature rules. Consider the inner product on the unit circle

$$<f,g> = \int_{|\lambda| = 1} f(\lambda) \ \overline{g(\lambda)} \ d\alpha(\lambda) \ , \tag{5.1}$$

with a positive measure $d\alpha(\lambda)$. Let ψ_k, $0 \le k < n$, be monic orthogonal polynomials with respect to (5.1). They satisfy a recurrence relation

$$\begin{cases} \psi_0(\lambda) := 1 & (5.1a) \\ \psi_k(\lambda) := \lambda\psi_{k-1}(\lambda) + \bar{\gamma}_k\tilde{\psi}_{k-1}(\lambda), \ 1 \le k < n, & (5.2b) \end{cases}$$

for some parameters $\gamma_k \in \mathbb{C}$ such that $|\gamma_k| < 1$ for $1 \le k < n$. Here $\tilde{\psi}_{k-1}(\lambda) := \lambda^{k-1}\bar{\psi}_{k-1}(1/\lambda)$ is the "reversed polynomial." Let $\gamma_n \in \mathbb{C}$ be an arbitrary complex number of unit magnitude, and define ψ_n by (5.2b) with $k := n$. Writing the recursions (5.2) (for $1 \le k \le n$) in matrix form, yields the unitary matrix

$$H = G_1 G_2 \ldots G_{n-1} G_n, \qquad (5.3)$$

whose eigenvalues $\{\lambda_k\}_{k=1}^n$ are the zeros of ψ_n. Here G_k is defined by γ_k according to (1.2) for $1 \le k \le n$. Hence, the parameters $\{\gamma_k\}^n$ are the Schur parmaeters for H. Let $H = W\Lambda W^H$ be a spectral resolution, and define the weights $\rho_k := |e_1^T W e_k|^2$ for $1 \le k \le n$. Then

$$\int_{|\lambda|=1} f(\lambda)\,d\alpha(\lambda) = \sum_{k=1}^n \rho_k f(\lambda_k) + \epsilon_n(f)$$

is a Gauss–Szegö quadrature rule with respect to the measure $d\alpha(\lambda)$, because the error $\epsilon_n(f)$ vanishes when f is any trigonometric polynomial of degree less than n. See [Gr2] for details. The computed examples illustrate the case when all Schur parameters γ_k are real valued and H therefore is real orthogonal.

A particularly simple example is $\gamma_k := 0$, $1 \le k < n$, and $\gamma_n := -1$. Then $\psi_k(\lambda) = \lambda^k$, $0 \le k < n$, and $\psi_n(\lambda) = \lambda^n - 1$, and therefore

$$\begin{cases} \lambda_k = \exp(2\pi i (k-1)/n), \\ \rho_k = 1/n. \end{cases} \qquad 1 \le k \le n. \qquad (5.4)$$

These Schur parameters have been used for Table 5.1. In the table "# defl. close e.v." stands for number of deflations due to close eigenvalues, and "# defl. small $|\zeta_k|$" is short for number of deflations due to components ζ_k of z of small magnitude. Two eigenvalues are considered close if (2.26) is satisfied for $\epsilon_2 := 1 \cdot 10^{-5}$, and $|\zeta_k|$ is regarded small if (2.24) is valid for $\epsilon_1 := 1 \cdot 10^{-5}$. These values of ϵ_1 and ϵ_2 are used in all computed examples of this section.

n	$\|HW-W\Lambda\|_\infty$	$\|W^H W - I\|_\infty$	# defl. close e.v.	# defl. small $|\zeta_k|$
4	$4.6 \cdot 10^{-12}$	$1.5 \cdot 10^{-11}$	2	0
8	$6.4 \cdot 10^{-12}$	$3.0 \cdot 10^{-11}$	8	0
16	$1.4 \cdot 10^{-11}$	$5.4 \cdot 10^{-11}$	24	0
32	$2.9 \cdot 10^{-11}$	$1.8 \cdot 10^{-10}$	64	0
64	$3.9 \cdot 10^{-11}$	$3.4 \cdot 10^{-10}$	160	0

Table 5.1: $\gamma_k := 0$, $1 \leq k < n$; $\gamma_n := -1$

For $\gamma_k := 0$, $1 \leq k < n$, and $\gamma_n := 1$, we obtain the polynomials $\psi_k(\lambda) = \lambda^k$, $0 \leq k < n$, and $\psi_n(\lambda) := \lambda^n + 1$. Hence, the eigenvalues are $\lambda_k = \exp(i\pi(2k-1)/n)$, $1 \leq k \leq n$, and the Gauss-Szegö weights ρ_k are the same as in (5.4). Table 5.2 shows computations for the present Schur parameters, and differs from Table 5.1 mainly in that fewer deflations take place.

n	$\|HW-W\Lambda\|_\infty$	$\|W^H W - I\|_\infty$	# defl. close e.v.	# defl. small $|\zeta_k|$
4	$7.8 \cdot 10^{-12}$	$1.6 \cdot 10^{-11}$	0	0
8	$1.7 \cdot 10^{-11}$	$4.2 \cdot 10^{-11}$	2	0
16	$3.1 \cdot 10^{-11}$	$1.6 \cdot 10^{-10}$	10	0
32	$4.1 \cdot 10^{-11}$	$3.5 \cdot 10^{-10}$	34	0
64	$5.6 \cdot 10^{-11}$	$7.5 \cdot 10^{-10}$	98	0

Table 5.2: $\gamma_k := 0$, $1 \leq k < n$; $\gamma_n := 1$

33

In Tables 5.3-5.4 we have chosen $\gamma_k := 0.8$, $1 \leq k < n$. This makes the λ_k gather in the left half plane. For the examples of Table 5.3 we have $\max\limits_{\lambda_k \neq 1}$ Re $\lambda_k < -\frac{1}{4}$. For the examples of Table 5.4 we obtain max Re $\lambda_k \leq -\frac{1}{4}$.

n	$\|HW-W\Lambda\|_\infty$	$\|W^H W-I\|_\infty$	# defl. close e.v.	# defl. small $|\zeta_k|$
4	$2.7 \cdot 10^{-12}$	$1.6 \cdot 10^{-11}$	2	0
8	$5.5 \cdot 10^{-11}$	$1.8 \cdot 10^{-10}$	8	0
16	$5.2 \cdot 10^{-11}$	$3.2 \cdot 10^{-10}$	24	0
32	$3.2 \cdot 10^{-8}$	$9.3 \cdot 10^{-8}$	63	1
64	$3.2 \cdot 10^{-8}$	$1.6 \cdot 10^{-7}$	157	3

Table 5.3: $\gamma_k := 0.8$, $1 \leq k < n$; $\gamma_n := -1$

n	$\|HW-W\Lambda\|_\infty$	$\|W^H W-I\|_\infty$	# defl. close e.v.	# defl. small $|\zeta_k|$
4	$4.8 \cdot 10^{-11}$	$1.7 \cdot 10^{-10}$	0	0
8	$9.4 \cdot 10^{-11}$	$5.5 \cdot 10^{-10}$	2	0
16	$4.8 \cdot 10^{-10}$	$6.6 \cdot 10^{-9}$	10	0
32	$6.3 \cdot 10^{-10}$	$2.3 \cdot 10^{-8}$	34	0
64	$4.2 \cdot 10^{-8}$	$1.9 \cdot 10^{-7}$	97	1

Table 5.4: $\gamma_k := 0.8$, $1 \leq k < n$; $\gamma_n := 1$

In the last computed quadrature rules of this example we let the γ_k, $1 \leq k < n$, be uniformly distributed in the open interval $]-1,1[$, and let γ_n be -1 or 1 with probability $\frac{1}{2}$ each. The γ_k are determined with the random number generator of Pascal. Table 5.5 shows the result of 30 eigenproblems so generated. The maximum, average and minimum in Table 5.5 are over all 30 eigenproblems.

	$\|HW-W\Lambda\|_\infty$	$\|W^H W-I\|_\infty$	# defl. close e.v.	# defl. small $|\zeta_k|$
max	$7.2 \cdot 10^{-7}$	$2.5 \cdot 10^{-6}$	30	0
average	$5.8 \cdot 10^{-8}$	$2.3 \cdot 10^{-7}$	26.5	0
min	$2.9 \cdot 10^{-9}$	$1.5 \cdot 10^{-8}$	22	0

Table 5.5:   Uniformly distributed $\gamma_k \in\,]-1,1[$, $1 \le k < n$; uniformly distributed $\gamma_n \in \{-1,1\}$.   Max, average and min are over 30 eigenproblems with n := 32

The numerical experiments of Table 5.5 indicate that for many choices of Schur parameters γ_k, the magnitudes $|\zeta_k|$ are not sufficiently small to give rise to frequent deflations.   This behavior has also been observed in many other computed experiments.   In contrast, massive deflation in DC methods for symmetric tridiagonal matrices often is caused by small components of the vector correspnding to $z = [\zeta_k]_{k=1}^n$.   □

Example 5.2.   This example suggests that it might not be possible to increase the small lower bound for $\min_j |\theta-\theta_j|$ of Lemma 3.2 significantly. The Schur parameters for Table 5.6 are obtained by reversing the sign of the γ_k, $1 \le k < n$, of Table 5.4.

n	$\min |\theta_k| \atop 1 \le k \le n$	$\|HW-W\Lambda\|_\infty$	$\|W^H W-I\|_\infty$	# defl. close e.v.	# defl. small $|\zeta_k|$
4	$6.6 \cdot 10^{-2}$	$6.9 \cdot 10^{-11}$	$3.1 \cdot 10^{-10}$	0	0
8	$8.1 \cdot 10^{-4}$	$7.2 \cdot 10^{-10}$	$2.5 \cdot 10^{-8}$	2	0
16	0^*	$1.2 \cdot 10^{-7}$	$3.1 \cdot 10^{-8}$	10	0
32	0^*	$7.2 \cdot 10^{-7}$	$1.9 \cdot 10^{-6}$	34	2
64	0^*	$7.2 \cdot 10^{-7}$	$2.6 \cdot 10^{-6}$	97	5

Table 5.6: $\gamma_k := -0.8$, $1 \le k < n$; $\gamma_n := 1$.   *The matrix has numerically the eigenvalue $\lambda = 1$ of multiplicity two.

35

Because $\sigma_k = 0.6 > 0$, $1 \leq k < n$, the matrix H has distinct eigenvalues mathematically. Numerically two eigenvalues are so close that they are not distinguished with our present choice of $\epsilon_2 = 1 \cdot 10^{-5}$. A smaller value of ϵ_2, such as $\epsilon_2 = 1 \cdot 10^{-6}$, gave in some numerical experiments larger residual errors $\|HW - W\Lambda\|_\infty$ or $\|W^H W - I\|_\infty$. $\square$

Acknowledgment

References

[AGR1]   Ammar, G.S., Gragg, W.B., and Reichel, L.: On the eigenproblem
           for orthogonal matrices. In Proc. 25th IEEE Conference on
           Decision and Control, Athens, Greece, 1986, pp. 1963-1966.

[AGR2]   Ammar, G.S., Gragg, W.B., and Reichel, L.: Determination of
           Pisarenko frequency estimates as eigenvalues of an orthogonal
           matrix. In SPIE vol. 826, Advanced Algorithms and Architectures
           for Signal Processing II, 1987, pp. 143-145.

[AG]   Arbenz, P., and Golub, G.H.: On the spectral decomposition of
           Hermitian matrices modified by low rank perturbations with
           applications. SIAM J. Matrix Anal. Appl. 9, 40-58 (1988).

[Cu]   Cuppen, J.J.M.: A divide and conquer method for the symmetric
           tridiagonal eigenproblem. Numer. Math. 36, 177-195 (1981).

[Cy]   Cybenko, G.: Computing Pisarenko frequency estimates. In Proc.
           1984 Conference on Information Systems and Sciences, Princeton
           University, 1984, pp. 587-591.

[DS]   Dongarra, J.J., and Sorensen, D.C.: A fully parallel algorithm
           for the symmetric eigenvalue problem. SIAM J. Sci. Stat. Comput.
           8, s139-s154 (1987).

[Gr1]   Gragg, W.B.: The QR algorithm for unitary Hessenberg matrices.
           J. Comput. Appl. Math. 16, 1-8 (1986).

[Gr2]   Gragg, W.B.: Positive definite Toeplitz matrices, the Arnoldi
           process for isometric operators and Gaussian quadrature on the
           unit circle (in Russian). In Numerical Methods in Linear
           Algebra, ed. E.S. Nikolaev, Moscow University Press, 1982.

[GR]   Gragg, W.B., and Reichel, L.: A divide and conquer algorithm
           for the unitary eigenproblem. In Hypercube Multiprocessors,
           1987, ed. M.T. Heath, SIAM, Philadelphia, 1987, pp. 639-647.

[JS]   Jessup, E.R., and Sorensen, D.C.: A parallel algorithm for
           computing the singular value decomposition of a matrix. Report
           ANL/MCS-TM-102 Math. Comp. Sci. Div., Argonne National
           Laboratory, 1987.

[KM]   Krishnakumar, A.S., and Morf, M.: Eigenvalues of a symmetric
           tridiagonal matrix: a divide-and-conquer approach. Numer. Math.
           48, 349-368 (1986).

[Pi]   Pisarenko, V.F.: The retrieval of harmonics from a covariance
           function. Geophys. J. R. Astr. Soc. 33, 347-366 (1973).

37

DISTRIBUTION LIST

DIRECTOR   (2)
DEFENSE TECH. INFORMATION
   CENTER, CAMERON STATION
ALEXANDRIA, VA 22314

DIRECTOR OF RESEARCH ADMIN.
CODE 012
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

LIBRARY   (2)
CODE 0142
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

DEPARTMENT OF MATHEMATICS
CODE 53
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

CENTER FOR NAVAL ANALYSES
4401 FORD AVENUE
ALEXANDRIA, VA 22302-0268

PROFESSOR WILLIAM GRAGG   (15)
CODE 53Gr
DEPARTMENT OF MATHEMATICS
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

NATIONAL SCIENCE FOUNDATION
WASHINGTON, D.C. 20550