An Efficient QR Algorithm for a Hessenberg Submatrix of a Unitary Matrix

Gregory S. Ammar[1], William B. Gragg[2], and Chunyang He[3]

[1] Department of Mathematical Sciences, Northern Illinois University, DeKalb, IL 60115 USA. ammar@math.niu.edu
[2] Department of Mathematics, Naval Postgraduate School, Monterey, CA 93943 USA. bill_gragg@yahoo.com
[3] Department of Mathematics and Statistics, University of Missouri–Kansas City, 5100 Rockhill Road, Kansas City, MO 64110–2499 USA. chunyang@sbcglobal.net

Dedicated to Clyde Martin on the occasion of his sixtieth birthday.

Summary. We describe an efficient procedure for implementing the Hessenberg QR algorithm on a class of matrices that we refer to as subunitary matrices. This class includes the set of Szegő-Hessenberg matrices, whose characteristic polynomials are Szegő polynomials, i.e., polynomials orthogonal with respect to a measure on the unit circle in the complex plane. Computing the zeros of Szegő polynomials is important in time series analysis and in the design and implementation of digital filters. For example, these zeros are the poles of autoregressive filters.

1 Introduction

A real-valued bounded nondecreasing function $\mu(t)$ on the interval $[-\pi, \pi]$ with infinitely many points of increase defines an inner product according to

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(\lambda)} g(\lambda) d\mu(t), \qquad \lambda = \mathrm{e}^{\mathrm{i}t}, \tag{1}$$

where i denotes the imaginary unit and the bar denotes complex conjugation. There is then a unique family $\{\psi_k(\lambda)\}_{k=0}^{\infty}$ of monic polynomials such that each $\psi_k(\lambda)$ has degree k and such that

$$\langle \psi_j(\lambda), \psi_k(\lambda) \rangle = \begin{cases} 0 & \text{if } j \neq k \ , \\ \delta_k^2 > 0 & \text{if } j = k \ . \end{cases}$$

The polynomials $\{\psi_k\}$ are said to be *orthogonal on the unit circle*, and we call them *Szegő polynomials*.

The monic Szegő polynomials satisfy the recurrence relation

$$\psi_{k+1}(\lambda) = \lambda\psi_k(\lambda) + \gamma_{k+1}\tilde{\psi}_k(\lambda), \quad k = 0, 1, 2, \ldots, \tag{2}$$

where $\psi_0(\lambda) = 1$, and $\tilde{\psi}_k(\lambda) = \lambda^k \bar{\psi}_k(1/\lambda)$ is the polynomial obtained by reversing and conjugating the power basis coefficients of ψ_k. The recurrence coefficients $\gamma_{k+1} \in \mathbb{C}$ are given by

$$\gamma_{k+1} = -\langle 1, \lambda\psi_k(\lambda)\rangle / \delta_k^2, \tag{3}$$

and the squared norm of ψ_{k+1} is recursively given by

$$\delta_{k+1}^2 = \delta_k^2(1 - |\gamma_{k+1}|^2) \ ,$$

where $\delta_0^2 = \langle 1, 1\rangle$. See, for example, [22, Chapter 11] or [16].

The Szegő recursion coefficients γ_j are known as *Schur parameters*. Since the distribution function $\mu(t)$ has infinitely many points of increase, $|\gamma_j| < 1$ for all j, and the zeros of each $\psi_j(\lambda)$ lie in the open unit disk $|\lambda| < 1$. On the other hand, if $\mu(t)$ has only n points of increase, then $|\gamma_j| < 1$ for $j = 1, 2, \ldots, n-1$ and $|\gamma_n| = 1$. In this case the orthogonal polynomials $\{\psi_j(\lambda)\}_{j=1}^{n-1}$ are defined only up to degree $n-1$, and (2) formally defines the monic polynomial ψ_n whose norm $\delta_n = 0$. In this case, the zeros of $\psi_n(\lambda)$ are pairwise distinct, of unit modulus, and equal to $\{e^{\mathrm{i}t_j}\}$, where $\{t_j\}$ are the points of increase of $\mu(t)$. See [22, Chapter 11] or [16].

Szegő polynomials arise in applications such as signal processing and time series analysis because of their connection with stationary time series. In these applications the Szegő polynomials are sometimes referred to as *backward predictor polynomials* or *Levinson polynomials*, and the Schur parameters are better known as *reflection coefficients* or *partial correlation coefficients*.

The moments associated with $\mu(t)$,

$$\mu_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-\mathrm{i}jt} d\mu(t), \qquad j = 0, \pm 1, \pm 2, \ldots \ , \tag{4}$$

form a positive definite Toeplitz matrix $M_{n+1} = [\mu_{j-k}]_{j,k=0}^n$. If we use these moments to compute the inner products of polynomials in (3) to obtain the Schur parameters, and use these to construct the power basis coefficients of $\psi_k(\lambda)$, $k = 0, 1, \ldots n$, with the Szegő recursion (2), the resulting algorithm is known as the *Levinson-Durbin algorithm* for solving the Yule-Walker equations, which is fundamental among fast algorithms for solving Toeplitz systems of equations. In fact, a variety of efficient Toeplitz solvers generate the Szegő polynomials and/or the Schur parameters. See, for example, [3, 12].

The Szegő polynomials $\{\psi_j\}_{j=0}^n$ are determined by the Schur parameters $\{\gamma_j\}_{j=1}^n$, and problems involving the Szegő polynomials can often be re-cast in terms of the Schur parameters. In particular, the Szegő polynomials can be identified as the characteristic polynomials of the leading principal submatrices of an upper Hessenberg matrix H determined by the Schur parameters as follows.

Given n complex parameters $\{\gamma_j\}_{j=1}^n$ with $|\gamma_j| \le 1$, define the *complementary parameters* $\sigma_j = (1 - |\gamma_j|^2)^{1/2}$. For $j = 1, \dots, n-1$, define the unitary transformation of order n in the $(j, j+1)$ coordinate plane

$$G_j(\gamma_j) = \begin{bmatrix} I_{j-1} & & \\ & \begin{matrix} -\gamma_j & \sigma_j \\ \sigma_j & \bar{\gamma}_j \end{matrix} & \\ & & I_{n-j-1} \end{bmatrix} \in \mathbb{C}^{n \times n}, \qquad 1 \le j < n, \tag{5}$$

where I_k denotes the $k \times k$ identity matrix. Also define a truncated matrix

$$\tilde{G}_n(\gamma_n) = \begin{bmatrix} I_{n-1} & \\ & -\gamma_n \end{bmatrix}. \tag{6}$$

Then we call the upper Hessenberg matrix

$$H_n = G_1(\gamma_1) G_2(\gamma_2) \cdots G_{n-1}(\gamma_{n-1}) \tilde{G}_n(\gamma_n) \tag{7}$$

the *Szegő-Hessenberg matrix* determined by the Schur parameters $\{\gamma_j\}_{j=1}^n$, and we write $H_n = H(\gamma_1, \gamma_2, \dots \gamma_n)$. Although H_n is mathematically determined by the Schur parameters, we retain the complementary parameters in computational procedures to avoid numerical instability, as σ_j cannot be accurately computed from γ_j when the latter has magnitude close to one. The Szegő-Hessenberg matrix H_n is then represented by *Schur parameter pairs* $\{(\gamma_j, \sigma_j)\}_{j=1}^n$. The last complementary parameter σ_n, which is not needed in H_n, is included for notational convenience.

The leading principal submatrix of order k of H_n is

$$H_k = H(\gamma_1, \dots, \gamma_k).$$

Let $\psi_k(\lambda) = \det(\lambda I - H_k)$ denote the characteristic polynomial of H_k. It is shown in [12] that these polynomials satisfy the recurrence relation (2) with $\psi_0(\lambda) = 1$. Consequently, these characteristic polynomials are the monic Szegő polynomials determined by the Schur parameters $\{\gamma_j\}_{j=1}^n$. See [2] for a short induction proof.

In the special case that $|\gamma_n| = 1$, the Szegő-Hessenberg matrix $H(\gamma_1, \dots \gamma_n)$ is a *unitary* Hessenberg matrix, and there are a variety of efficient algorithms for computing its eigenvalues. These include the unitary Hessenberg QR algorithm [11], divide-and-conquer methods [5, 6, 14, 15, 17], a matrix pencil approach [7], and a Sturm-sequence type method [8]. The existence of these algorithms stems from the fact that unitary Hessenberg matrices are invariant under unitary similarity transformations, so that the computations can be performed on the Schur parameters that determine the intermediate matrices, rather than on the matrix elements explicitly.

However, the set of Szegő-Hessenberg matrices with $|\gamma_n| < 1$ is not invariant under unitary similarity transformations. In particular, a general Szegő-Hessenberg matrix $H_n = H(\gamma_1, \ldots, \gamma_n)$ satisfies

$$H_n^H H_n = I - \sigma_n^2 e_n e_n^H \ , \tag{8}$$

where the superscript H denotes the conjugate transpose of a matrix and e_n denotes the nth column of I_n. Since $\sigma_n > 0$, this property is preserved by the unitary similarity transformation $Q^H H_n Q$ only if $Q^H e_n$ is proportional to e_n. The development of efficient algorithms for general Szegő-Hessenberg matrices is therefore more complicated than that of unitary Hessenberg matrices.

One approach to efficiently compute the eigenvalues of a Szegő-Hessenberg matrix is a continuation method given in [2]. In order to find the eigenvalues of $H_n = H(\gamma_1, \ldots, \gamma_n)$ with $|\gamma_n| < 1$, we first find the eigenvalues of a unitary Hessenberg matrix $H_n' = H(\gamma_1, \ldots, \gamma_{n-1}, \gamma_n')$, where $|\gamma_n'| = 1$, using any of the established $O(n^2)$ methods. A continuation method is then applied to track the path of each eigenvalue, beginning with those of H_n' on the unit circle, and ending with those of H_n, as the last Schur parameter is varied from γ_n' to γ_n. This results in an $O(n^2)$ algorithm that also lends itself well to parallel computation.

Another approach is presented in [9], in which H_n is viewed as being in a larger class of matrices, called *fellow matrices*, defined as additive rank-one perturbations of unitary Hessenberg matrices. While fellow matrices are not invariant under the QR iteration, each QR iteration adds $O(n)$ additional parameters to the representation of the matrix. One can contain the number of parameters required by using a periodically restarted QR iteration, which leads to an efficient, $O(n^2)$, algorithm for computing the eigenvalues of fellow matrices. See [9] for details.

In this paper we present an efficient implementation of the QR algorithm on another class of matrices that include the Szegő-Hessenberg matrices. In view of (8), we consider matrices A that have the property that

$$A^H A = I_n - uu^H \quad \text{with } \|u\| \leq 1.$$

We will refer to such a matrix as a *subunitary* matrix. This class of matrices includes the Szegő-Hessenberg matrices, and moreover, is invariant under unitary similarity transformations. Our goal is to present an efficient implementation of the QR algorithm on the set of subunitary Hessenberg matrices.

In [1] it is shown that Szegő-Hessenberg matrices provide an alternative to companion matrices for finding the zeros of a general polynomial from its power basis coefficients. In particular, any polynomial, after a suitable change of variable, can be identified as a Szegő polynomial. Experiments presented in [1] indicate computing the zeros of a polynomial by applying the QR algorithm to an associated Szegő-Hessenberg matrix often yields more accurate results than the traditional use of the QR algorithm on companion matrices. The development of efficient algorithms for Szegő-Hessenberg eigenproblems will therefore have a direct impact on the problem of computing the zeros of a general polynomial.

In Section 2 we summarize the mechanics of the bulge chasing procedure for performing one step of the implicit Hessenberg QR algorithm. Some basic properties of subunitary matrices are presented in Section 3, where we see that A is a subunitary matrix if and only if it is the leading principal submatrix of a unitary matrix of size one larger. In Section 4 we show that a subunitary Hessenberg matrix is represented by approximately $4n$ real parameters, and describe how the QR iteration can be efficiently performed on a subunitary Hessenberg matrix implicitly in terms of the underlying parameters.

2 Overview of the Hessenberg QR algorithm

In finding the eigenvalues of a matrix using the QR algorithm, the matrix is first transformed by a unitary similarity transformation to upper Hessenberg form. The QR algorithm then iteratively generates a sequence of upper Hessenberg matrices by unitary similarity transformations. Implicit implementations of the Hessenberg QR algorithm can be viewed in terms of a *bulge chasing procedure*, which is a general computational procedure for performing a similarity transformation on a Hessenberg matrix to obtain another Hessenberg matrix. This is possible by virtue of the fact that a unitary matrix U such that U^*AU is an upper Hessenberg matrix is essentially determined by its first column. In order to establish notation, we now summarize the mechanics of the bulge chasing procedure that underlies the implicitly shifted QR algorithm. See, for example, [10, 18, 19, 25] for background on bulge-chasing procedures and their application in the implicitly shifted QR algorithm.

Let A be an upper Hessenberg matrix. One step of the Hessenberg QR algorithm with a single shift $\mu \in \mathbb{C}$ applied to A results in a new Hessenberg matrix A', given by

$$A' := RQ + \mu I_n = Q^H AQ$$

where

$$A - \mu I_n =: QR$$

is the QR factorization of the shifted matrix. The transformation Q is essentially determined by its first column, and since A is in upper Hessenberg form, so is Q. Let Q_1 be a unitary transformation in the $(1, 2)$ coordinate plane with the same first column as that of Q, and set $A_0 := A$. The bulge-chasing procedure now proceeds as follows.

The matrix $A_1 := Q_1^H A_0$ is also an upper Hessenberg matrix, and completing the similarity transformation yields the matrix $K_1 := Q_1^H A_0 Q_1$. The matrix K_1 would be a Hessenberg matrix if its $(3, 1)$ element were nonzero. This element is the *bulge*. A unitary transformation Q_2 in the $(2, 3)$ plane is then chosen to to annihilate the bulge in K_1 by left multiplication, so that $Q_2^H K_1 = A_2$ is in Hessenberg form. After multiplying on the right by Q_2 to complete the similarity transformation, the matrix $K_2 = A_2 Q_2$ has a bulge in the (4,2) position. The bulge in K_2 is annihilated by left multiplication by a unitary transformation Q_3 in the $(3,4)$-plane, and the process continues until the bulge is 'chased' diagonally down the matrix until we obtain the Hessenberg matrix $K_{n-1} = Q_{n-1}^H K_{n-2} Q_{n-1}$, which is unitarily similar to the initial Hessenberg matrix A. Finally, a diagonal unitary similarity transformation, equal to the identity matrix except possibly for its (n,n) entry, is performed to make the $(n, n-1)$ entry nonnegative, resulting in $A' = \tilde{Q}_n^H K_{n-1} \tilde{Q}_n^H$. Then $A' = Q^H A Q$, where $Q = Q_1 Q_2 \cdots Q_{n-1} \tilde{Q}_n$, is the result of a single bulge-chasing sweep on the original matrix A.

The pattern of nonzero elements in the intermediate matrices is displayed in Figure 1 for $n = 5$.

$$A_1 = Q_1^H A = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \bullet & \times & \times & \times \\ & & \bullet & \times & \times \\ & & & \bullet & \times \end{bmatrix} \qquad K_1 = Q_1^H A_0 Q_1 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ + & \times & \times & \times & \times \\ & & \bullet & \times & \times \\ & & & \bullet & \times \end{bmatrix}$$

$$K_2 = Q_2^H K_1 Q_2 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \bullet & \times & \times & \times & \times \\ \oplus & \times & \times & \times & \times \\ & + & \times & \times & \times \\ & & & \bullet & \times \end{bmatrix} \qquad K_3 = Q_3^H K_2 Q_3 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \bullet & \times & \times & \times & \times \\ & \bullet & \times & \times & \times \\ & \oplus & \times & \times & \times \\ & & + & \times & \times \end{bmatrix}$$

$$K_4 = Q_4^H K_3 Q_4 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \bullet & \times & \times & \times & \times \\ & \bullet & \times & \times & \times \\ & & \bullet & \times & \times \\ & & \oplus & \times & \times \end{bmatrix} \qquad A' = \tilde{Q}_5^H K_4 \tilde{Q}_5 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \bullet & \times & \times & \times & \times \\ & \bullet & \times & \times & \times \\ & & \bullet & \times & \times \\ & & & \bullet & \times \end{bmatrix}$$

Fig. 1. Matrix profiles during one bulge chasing sweep. Entries $\times$ represent complex numbers, $\bullet$ represent nonnegative numbers, $+$ represents a fill-in, and $\oplus$ represents a zero element introduced after annihilating a fill-in element

The matrix A' is uniquely determined by A and Q_1 provided that every nonnegative subdiagonal element σ_j' $(j = 1, 2, \ldots, n-1)$ of A' (represented by the $\bullet$ symbols in Figure 1) is positive. If the subdiagonal element σ_j' of A' vanishes, then the procedure terminates early, and the eigenproblem for A' deflates into two smaller eigenproblems.

For general Hessenberg matrices A, a bulge-chasing step requires $O(n^2)$ floating-point operations (flops).

3 Subunitary matrices

We will say that $A \in \mathbb{C}^{n \times n}$ is a *subunitary matrix* if $A^H A = I - uu^H$ for some vector u such that $\|u\|_2 \leq 1$. We refer to u as the *departure vector* of the subunitary matrix A, and to $\|u\|_2$ as its *departure norm.*

Many elementary properties of subunitary matrices are easily derived. For example, the proof of the following proposition is immediate.

Proposition 1. *Let A be a subunitary matrix with $\|u\| = \nu$. Then every eigenvalue λ of A is contained in the annulus $\sqrt{1-\nu^2} \leq \lambda \leq 1$. Moreover, A has $n-1$ singular values equal to one, and one singular value equal to $(1-\nu^2)^{1/2}$.*

The next proposition gives some support to the choice of the appellation *subunitary*.

Proposition 2. *The $n \times n$ matrix A is a subunitary matrix if and only if it is a submatrix of a unitary matrix of order $n+1$.*

Proof. There is no loss of generality to prove this result for A the leading principal submatrix of a unitary matrix B, since this can be enforced after performing row and column permutations on B. Let B be an $(n+1)\times(n+1)$ matrix, and write B in partitioned form,

$$B = \begin{bmatrix} A & v \\ u^H & \beta \end{bmatrix}, \tag{9}$$

where A is an $n \times n$ matrix, $u, v \in \mathbb{C}^n$ are column vectors, and $\beta \in \mathbb{C}$. Then

$$B^H B = \begin{bmatrix} A^H A + uu^H & A^H v + u\beta \\ v^H A + \bar{\beta} u^H & v^H v + |\beta|^2 \end{bmatrix}.$$

Thus, if B is unitary, then $A^H A = I - uu^H$, where $\|u\| \leq 1$.

Conversely, assume that A is a subunitary matrix with departure vector u. If $\|u\| < 1$, then A is nonsingular, and the equation $A^H x = -u$ has a unique solution x. Set $\beta = (1 + \|x\|^2)^{-1/2}$ and $v = \beta x$. Then the matrix B given by (9) is unitary. If $\|u\| = 1$, then A is singular with a one-dimensional nullspace spanned by u. Let v be such that $A^H v = 0$ and $\|v\| = 1$, and set $\beta = 0$ to obtain a unitary matrix B whose leading principal submatrix is A. □

Proposition 3. *Let A be a subunitary matrix with departure vector u, and let Q be any unitary matrix. Then:*

1. *QA is a subunitary matrix with the same departure vector u.*
2. *AQ is a subunitary matrix with departure vector $Q^H u$.*

It follows immediately that set of subunitary matrices of departure norm ν is invariant under unitary equivalence transformations and unitary similarity transformations. In particular, the set is invariant under the QR iteration. It also follows that for any factorization of the form $A = QX$, where Q is unitary, the matrix X is also subunitary with the same departure vector as that of A. In particular, in the QR factorization of a subunitary matrix A, the upper triangular factor R is a subunitary matrix with the same departure vector u.

Proposition 4. *Let R be a subunitary upper triangular matrix with positive diagonal elements. Then R is uniquely determined by its departure vector u. Moreover, the entries of $R = [\rho_{jk}]_{j,k=1}^{n}$ are explicitly given by*

$$\rho_{jk} = \begin{cases} \dfrac{-\upsilon_j \overline{\upsilon}_k}{\kappa_{j-1}\kappa_j} & \text{for } j < k, \\ \kappa_j / \kappa_{j-1} & \text{for } j = k, \\ 0 & \text{for } j > k, \end{cases} \tag{10}$$

where $u = [\upsilon_j]_{j=1}^{n}$ and

$$\kappa_j^2 = 1 - \sum_{k=1}^{j} |\upsilon_k|^2 = \kappa_{j+1}^2 + |\upsilon_{j+1}|^2$$

for $j = 0, \ldots, n-1$, with $\kappa_n^2 := 1 - \|u\|_2^2$.

Proof. Since R has full rank, its departure vector u has norm strictly less than one. From $R^H R = I - uu^H$, we see that R is the unique Cholesky factor of the positive definite matrix $I - uu^H$. The formulas (10) follow from the Cholesky factorization algorithm. □

If $\|u\| = 1$ and $\upsilon_n \neq 0$ (i.e., $\kappa_n = 0$ and $\kappa_{n-1} > 0$), the above formulas for the entries of R remain valid. In this case R is unique up to a unimodular scaling of its last column.

4 Efficient QR iteration on subunitary Hessenberg matrices

Let $\mathcal{H}_0$ denote the set of unitary upper Hessenberg matrices with nonnegative subdiagonal elements, and let $\mathcal{H}_1$ denote the set of subunitary upper Hessenberg matrices with nonnegative subdiagonal elements. Then in the QR factorization of $A \in \mathcal{H}_1$,

$$A = HR \, ,$$

we have that $H \in \mathcal{H}_0$, and therefore H has a unique Schur parametrization $H = H(\gamma_1, \ldots, \gamma_n)$. This combined with Proposition 4 yields the following result.

Theorem 1. *Any subunitary upper Hessenberg matrix $A \in \mathbb{C}^{n \times n}$ with nonnegative subdiagonal elements and departure vector u with $\|u\| < 1$ is uniquely represented by $4n - 1$ real parameters. In particular, $A = HR$, where $H = H(\gamma_1, \ldots, \gamma_n) \in \mathcal{H}_0$ (with $|\gamma_n| = 1$), and where R is the subunitary upper triangular matrix given in Proposition 4 whose departure vector $u \in \mathbb{C}^n$ is the same as that of A.*

The parametrization of $\mathcal{H}_1$ given in Theorem 1 now allows us to approach the problem of efficiently implementing the QR iteration on $\mathcal{H}_1$. The key is to perform the QR iteration on the parameterized product $A = HR$, keeping the parameterized form of each factor intact during the iteration.

Let $A =: HR$ be the QR factorization of $A \in \mathcal{H}_1$, where

$$H = H(\gamma_1, \ldots, \gamma_n) \in \mathcal{H}_0,$$

and where R is the upper triangular matrix with positive diagonal elements satisfying $R^H R = A^H A = I - uu^H$. In one step of the QR algorithm on A with shift $\tau \in \mathbb{C}$, the QR factorization of the shifted matrix,

$$A - \tau I =: QU,$$

defines the unitary Hessenberg matrix Q that produces the result of the QR step $A' = Q^H AQ$. We seek to compute the QR factorization $A' = H'R' \in \mathcal{H}_1$. The Hessenberg-triangular product representation is maintained through the introduction of a unitary matrix Z such that

$$A' = Q^H AQ = (Q^H HZ)(Z^H RQ) =: H'R'.$$

We will therefore implement the implicitly shifted QR step on $A \in \mathcal{H}_1$ keeping the parameterized factors H and R separate, as in implementations of the QZ algorithm for matrix pencils (see, e.g., [10]).

Let us now consider the individual steps of an implicit QR step on $A = HR \in \mathcal{H}_1$. Let Q_1 denote the initial transformation in the $(1, 2)$ coordinate plane that implicitly defines Q and A'. Write $H = G_1 G_2 \cdots \tilde{G}_n$, where each $G_j = G_j(\gamma_j, \sigma_j)$. In addition to the matrices G_j, throughout the following discussion, Q_j, T_j, Z_j will denote unitary transformations in the $(j, j+1)$-plane (for $j < n$), and $\tilde{Q}_n$, $\tilde{T}_n$, $\tilde{Z}_n$ will denote unitary matrices that differ from the identity matrix only in the (n, n) entry. We consider the special structure of intermediate matrices generated during the QR step on $A = HR$ as described in Section 2.

The initial bulge in the (3,1) position of the matrix

$$K_1 = Q_1^H AQ_1 = Q_1^H HRQ_1$$

arises from the bulge in the (2,1) position of RQ_1. Choose a unitary transformation Z_1 to annihilate this bulge, so that $Z_1^H RQ_1 =: R_1$ is upper triangular with positive diagonal entries. Then $K_1 = (Q_1^H HZ_1)R_1$, and the bulge resides in the Hessenberg factor of the product. Since Z_j commutes with G_k whenever $|j - k| > 1$, we have

$$\begin{aligned} K_1 &= Q_1^H G_1 G_2 \cdots \tilde{G}_n Z_1 R_1 \\ &= (T_1 G_2 Z_1) G_3 G_4 \cdots \tilde{G}_n R_1 \ , \end{aligned}$$

where $T_1 := Q_1^H G_1$. The bulge in K_1 now arises from the $(3, 1)$ entry of the matrix $W_1 = T_1 G_2 Z_1$. Note that W_1 differs from the identity matrix only in its 3×3 leading principal submatrix. Also, R_1 is the subunitary upper triangular matrix determined by its departure vector $u_1 = Q_1^H u$.

A unitary transformation Q_2 is now chosen so that $Q_2^H W_1$ is in upper Hessenberg form, and then choose $G_1' = G_1(\gamma_1', \sigma_1')$ so that the (2,1) entry of $G_1'^H Q_2^H W_1$ is annihilated, and note that since W_1 is unitary, this last matrix is a unitary matrix T_2 which differs from the identity matrix only in the $(2,3)$ principal submatrix. In this way, we obtain unitary plane transformations Q_2, G_1', and T_2 such that

$$W_1 = T_1 G_2 Z_1 =: Q_2 G_1' T_2$$

and

$$K_1 = Q_2 G_1' T_2 G_3 G_4 \cdots \tilde{G}_n R_1 \ .$$

The similarity transformation defined by Q_2 on K_1 then yields the matrix K_2 with bulge in the (4,2) position,

$$\begin{aligned} K_2 = Q_2^H K_1 Q_2 &= G_1' T_2 G_3 G_4 \cdots \tilde{G}_n R_1 Q_2 \\ &= G_1' T_2 G_3 G_4 \cdots \tilde{G}_n Z_2 (Z_2^H R_1 Q_2) \\ &= G_1' (T_2 G_3 Z_2) G_4 \cdots \tilde{G}_n R_2 \\ &= G_1' (Q_3 G_2' T_3) G_4 \cdots \tilde{G}_n R_2 \\ &= Q_3 (G_1' G_2') T_3 G_4 \cdots \tilde{G}_n R_2 \ , \end{aligned}$$

where Z_2 is chosen so that $Z_2^H R_1 Q_2^H =: R_2$ is the upper triangular subunitary matrix with departure vector $u_2 = Q_2^H u_1$, and the identification

$$T_2 G_3 Z_2 =: W_2 =: Q_3 G_2' T_3$$

is made as above, using the matrix W_2 that differs from the identity matrix only in the 3×3 principal submatrix from rows and columns $(2,3,4)$. Now $K_3 = Q_3^H K_2 Q_3$ has a bulge in the (5,3) position, and the process continues until we obtain the upper Hessenberg matrix

$$K_{n-1} = \tilde{Q}_n G_1' G_2' \cdots G_{n-1}' \tilde{T}_n R_{n-1}$$

and

$$\begin{aligned} A' = \tilde{Q}_n^H K_{n-1} \tilde{Q}_n &= G_1' G_2' \cdots G_{n-1}' (\tilde{T}_n \tilde{Q}_n)(\tilde{Q}_n^H R_{n-1} \tilde{Q}_n) \\ &= G_1' G_2' \cdots G_{n-1}' \tilde{G}_n' R_n \\ &= H(\gamma_1', \ldots, \gamma_n') R_n \\ &= H' R' \ . \end{aligned}$$

Thus, the transition from $A = HR$ to $A' = H'R'$ can be achieved by keeping the Hessenberg factors and upper triangular factors separate. Moreover, individual operations can be performed on 2×2 and 3×3 matrices. This leads to the following algorithm for performing one QR step on a subunitary Hessenberg matrix using $O(n)$ flops.

Algorithm 1.
Input: Schur parameter pairs $\{\gamma_j, \sigma_j\}_{j=1}^n$ of $H = H(\gamma_1, \ldots, \gamma_n) \in \mathcal{H}_0$, departure vector $u = [\upsilon_j] \in \mathbb{C}^n$ of a subunitary upper triangular matrix R, with $\|u\| = \nu < 1$, and initial unitary transformation Q_1 in the $(1, 2)$ coordinate plane.

Output: Schur parameter pairs $\{\gamma'_j, \sigma'_j\}_{j=1}^n$ of $H' = H(\gamma'_1, \ldots, \gamma'_n) \in \mathcal{H}_0$, departure vector $u' \in \mathbb{C}^n$ of the subunitary upper triangular matrix R', with $\|u'\| = \nu$, such that $A' = H'R'$ is the QR factorization of the subunitary upper Hessenberg matrix $A' = Q^H AQ = Q^H HZZ^H RQ$ that results from one bulge chasing sweep applied to the initial subunitary Hessenberg matrix $A = HR$ with initial transformation Q_1.

Set $\kappa_n := (1 - \|u\|_2^2)^{1/2}$.
Set $\kappa_j := (\kappa_{j+1}^2 + |\upsilon_j|^2)^{1/2}$, $j = n-1, n-2, \ldots, 0$.
Form the 2×2 matrix $T_1 = Q_1^H G_1$, where $G_1 = G_1(\gamma_1, \sigma_1)$.
For $j = 1, 2, \ldots n-1$

Form the 2×2 matrix $X = \begin{bmatrix} \rho_{jj} & \rho_{j,j+1} \\ 0 & \rho_{j+1,j+1} \end{bmatrix}$ according to the formulas (10).

Overwrite $XQ_j =: X =: \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix}$.

Set $\rho'_{jj} = (|\xi_{11}|^2 + |\xi_{21}|^2)^{1/2}$, $\mu_j := -\xi_{11}/\rho'_{jj}$, $\kappa_j = \xi_{21}/\rho'_{jj}$.

Update $\begin{bmatrix} \upsilon_j \\ \upsilon_{j+1} \end{bmatrix} := Q_j^H \begin{bmatrix} \upsilon_j \\ \upsilon_{j+1} \end{bmatrix}$ and $\kappa_j := \rho'_{jj}\kappa_{j-1}$.

Form $Z_j = G_j(\mu_j, \kappa_j)$ and $G_{j+1} = G_{j+1}(\gamma_{j+1}, \sigma_{j+1})$.

Form the 3×3 matrix $W = T_j G_{j+1} Z_j$.

Set $\sigma'_j = (|\omega_{21}|^2 + \omega_{31}^2)^{1/2}$, $\alpha_{j+1} = -\omega_{21}/\sigma'_j$, $\beta_{j+1} = \omega_{31}/\sigma'_j$.

Form $Q_{j+1} = \begin{bmatrix} -\alpha_{j+1} & \beta_{j+1} \\ \beta_{j+1} & \bar{\alpha}_{j+1} \end{bmatrix} = G_{j+1}(\alpha_{j+1}, \beta_{j+1})$.

Overwrite $Q_{j+1}^H W =: W$. Set $\gamma'_j = -\omega_{11}$.

Form $G'_j = G_j(\gamma'_j, \sigma'_j)$, and overwrite $(G'_j)^H W =: W$.

Set $T_{j+1} = \begin{bmatrix} \omega_{22} & \omega_{23} \\ \omega_{32} & \omega_{33} \end{bmatrix}$.

End (for j).

Form the 2×2 matrix $W = T_{n-1}\tilde{G}_n Z_{n-1}$.

Set $\sigma'_{n-1} = |\omega_{21}|$, $\alpha_n = -\omega_{21}/\sigma'_j$.

Form $\tilde{Q}_n = \begin{bmatrix} 1 & 0 \\ 0 & -\alpha_n \end{bmatrix}$.

Overwrite $\tilde{Q}_n^H W =: W$, and set $\gamma'_j = -\omega_{11}$.

Form $G'_{n-1} = G_{n-1}(\gamma'_{n-1}, \sigma'_{n-1})$, and overwrite $(G'_{n-1})^H W =: W$.

Set $\tilde{T}_n = \begin{bmatrix} 1 & 0 \\ 0 & \omega_{22} \end{bmatrix}$.

Update $\upsilon_n = -\bar{\alpha}_n \upsilon_n$.

Set $\tilde{T}_n \tilde{Q}_n =: \tilde{G}'_n =: \begin{bmatrix} 1 & 0 \\ 0 & -\gamma'_n \end{bmatrix}$.

Set $u' := u$.

end algorithm.

5 Concluding Remarks

We considered some basic aspects of subunitary matrices, and showed that the QR algorithm can be implemented on subunitary Hessenberg matrices using $O(n)$ flops per iteration. The resulting subunitary Hessenberg QR (SUHQR) algorithm can be regarded as a generalization of the general idea behind the unitary Hessenberg QR (UHQR) algorithm. In particular, UHQR operates on the Schur parameters of intermediate unitary Hessenberg matrices to implicitly perform unitary similarity transformations on the initial matrix, while SUHQR implicitly performs unitary equivalence transformations on both a unitary Hessenberg matrix and a subunitary upper triangular matrix, using the Schur parameters of the former and the departure vector of the latter.

The algorithm outlined above represents only a beginning for the study of QR iterations on $\mathcal{H}_1$. A next step is to further streamline the algorithm by implementing it directly on the parameters that determine the intermediate matrices, rather than explicitly on the entries of these matrices. This will result in an implementation more closely related to the original implementation of the UHQR algorithm given in [11]. In fact, several different implementations of the UHQR algorithm have been developed [23,24], and because the SUHQR algorithm is essentially a generalization of UHQR, there will be corresponding variants on the SUHQR algorithm as well.

A numerical stability analysis of our algorithm remains open. We make no claim on this matter here. In [13] it is shown that the original version of the UHQR algorithm is not numerically stable. A source of instability is identified there, and a modified UHQR algorithm is proposed to avoid the instability. Numerical experiments in [13] confirm the improved stability of this modified UHQR algorithm. A detailed error analysis of UHQR algorithms has been performed by Stewart [20,21], where additional potential instabilities are identified and resolved with provably numerical stable implementations of the UHQR algorithm. Certainly this recent work on stabilizing the UHQR algorithm will be relevant for SUHQR as well.

Acknowledgement. GA began his study of the QR algorithm during his work toward his doctorate under the direction of Clyde Martin. In [4]

it is shown that the basic QR algorithm can be interpreted in terms of a linear action on the full flag manifold. This is in analogy with the study of control-theoretic matrix Riccati equations in terms of linear actions on Grassmannians, which was initiated by Hermann and Martin. GA takes this opportunity to thank Clyde Martin for introducing him to the QR algorithm, which continues to provide for interesting and fruitful study.

References

1. G. Ammar, D. Calvetti, W. Gragg, and L. Reichel. Polynomial zerofinders based on Szegő polynomials. *J. Comput. Appl. Math.*, 127:1–16, 2001.
2. G. S. Ammar, D. Calvetti, and L. Reichel. Continuation methods for the computation of zeros of Szegő polynomials. *Lin. Alg. Appl.*, 249:125–155, 1996.
3. G. S. Ammar and W. B. Gragg. The generalized Schur algorithm for the superfast solution of Toeplitz systems. In J. Gilewicz, M. Pindor, and W. Siemaszko, editors, *Rational Approximation and its Applications in Mathematics and Physics*, number 1237 in Lecture Notes in Mathematics, pages 315–330. Springer-Verlag, Berlin, 1987.
4. G. S. Ammar and C. F. Martin. The geometry of matrix eigenvalue methods. *Acta Applicandae Math.*, 5:239–278, 1986.
5. G. S. Ammar, L. Reichel, and D. C. Sorensen. An implementation of a divide and conquer algorithm for the unitary eigenproblem. *ACM Trans. Math. Software*, 18:292–307, 1992.
6. G. S. Ammar, L. Reichel, and D. C. Sorensen. Algorithm 730: An implementation of a divide and conquer algorithm for the unitary eigenproblem. *ACM Trans. Math. Software*, 20:161, 1994. (Mathematical software.).
7. A. Bunse-Gerstner and L. Elsner. Schur parameter pencils for the solution of the unitary eigenproblem. *Lin. Alg. Appl.*, 154–156:741–778, 1991.
8. A. Bunse-Gerstner and C. He. On the Sturm sequence of polynomials for unitary Hessenberg matrices. *SIAM J. Matrix Anal. Appl.*, 16(4):1043–1055, October 1995.
9. D. Calvetti, S.-M. Kim, and L. Reichel. The restarted QR-algorithm for eigenvalue computation of structured matrices. *J. Comput. Appl. Math.*, 149:415–422, 2002.
10. G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, third edition, 1996.
11. W. B. Gragg. The QR algorithm for unitary Hessenberg matrices. *J. Comput. Appl. Math.*, 16:1–8, 1986.
12. W. B. Gragg. Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle. *J. Comput. Appl. Math.*, 46:183–198, 1993. Originally published (in Russian) in: E.S. Nikolaev, Ed., *Numerical Methods in Linear Algebra,* Moscow Univ. Press, 1982, pp. 16–32.
13. W. B. Gragg. Stabilization of the uhqr algorithm. In Z. Chen, Y. Li, C. Micchelli, and Y. Xu, editors, *Advances in Computational Mathematics: Proceedings of the Guangzhou International Symposium*, pages 139–154, New York, 1999. Marcel Dekker.

14. W. B. Gragg and L. Reichel. A divide and conquer method for the unitary eigenproblem. In M. T. Heath, editor, *Hypercube Multiprocessors 1987*, Philadelphia, 1987. SIAM.
15. W. B. Gragg and L. Reichel. A divide and conquer method for unitary and orthogonal eigenproblems. *Numer. Math.*, 57:695–718, 1990.
16. U. Grenander and G. Szegő. *Toeplitz Forms and Their Applications.* Chelsea, New York, second edition, 1984.
17. M. Gu, R. Guzzo, X.-B. Chi, and X.-Q. Cao. A stable divide and conquer algorithm for the unitary eigenproblem. *SIAM J. Matrix Anal. Appl.*, 25(2):385–404, 2003.
18. B. N. Parlett. *The Symmetric Eigenvalue Problem.* Society for Industrial and Applied Mathematics, Philadelphia, 1998.
19. G. W. Stewart. *Introduction to Matrix Computations.* Academic Press, New York, 1973.
20. M. Stewart. Stability properties of several variants of the unitary Hessenberg QR algorithm. In V. Olshevsky, editor, *Structured Matrices in Mathematics, Engineering, and Computer Science II (Boulder, CO, 1999)*, volume 281 of *Contemporary Math.*, pages 57–72, Providence, RI, 2001. Amer. Math. Soc.
21. M. Stewart. An error analysis of a unitary Hessenberg QR algorithm. Preprint, 2004.
22. G. Szegő. *Orthogonal Polynomials.* American Math. Soc., Providence, RI, fourth edition, 1975.
23. T.-L. Wang and W. B. Gragg. Convergence of the shifted QR algorithm for unitary Hessenberg matrices. *Math. Comp.*, 71:1473–1496, 2002.
24. T.-L. Wang and W. B. Gragg. Convergence of the unitary QR algorithm with a unimodular Wilkinson shift. *Math. Comp.*, 72:375–385, 2003.
25. D. S. Watkins. *Fundamentals of Matrix Computations.* John Wiley and Sons, New York, 1991.