MATHEMATICS OF COMPUTATION
Volume 71, Number 240, Pages 1473–1496
S 0025-5718(01)01387-4
Article electronically published on November 30, 2001

# CONVERGENCE OF THE SHIFTED *QR* ALGORITHM FOR UNITARY HESSENBERG MATRICES

TAI-LIN WANG AND WILLIAM B. GRAGG

*Dedicated to the memory of James H. Wilkinson*

ABSTRACT. This paper shows that for unitary Hessenberg matrices the *QR* algorithm, with (an exceptional initial-value modification of) the Wilkinson shift, gives global convergence; moreover, the asymptotic rate of convergence is at least cubic, higher than that which can be shown to be quadratic only for Hermitian tridiagonal matrices, under no further assumption. A general mixed shift strategy with global convergence and cubic rates is also presented.

## 1. INTRODUCTION

The *QR* algorithm has been known as a standard method for computing the eigenvalues of a dense matrix [5, 17, 1, 13, 6]. One remarkable feature in the development of *QR* is Wilkinson's discovery that the algorithm, when incorporated with his efficient shift strategy, gives fast convergence for all real symmetric tridiagonal matrices [18, 8]. In this paper we extend Wilkinson's famous results to the unitary case. We show that for unitary Hessenberg matrices the *QR* algorithm, with an exceptional initial-value modification of the Wilkinson shift, gives *global* convergence in exact arithmetic. The proof is based on the Schur parameterization of unitary Hessenberg matrices [6] and a residual estimation for the shifted *QR* decomposition of these matrices. Furthermore, we show that the asymptotic rate of convergence with the Wilkinson shift is at least *cubic* in the unitary case. A general mixed shift strategy with global convergence and cubic rates is also included for reference. A special case of this general strategy (with parameter $\theta = \sqrt{2}$) was shown to have global convergence by Eberlein and Huang [4], in which the rate of convergence was claimed to be only quadratic. The analysis we consider here is purely theoretical. Numerical testing and experiments are prepared in a later paper.

We adhere to the following notational conventions: upper case letters for matrices, lower case Latin letters for column vectors (except for $i, j, k$, and $n$, which are used as indices), and lower case Greek letters for scalars. The conjugate transpose of a vector $a$ and of a matrix $A$ is denoted by $a^*$ and $A^*$, respectively, while the

Received by the editor March 9, 1999 and, in revised form, January 6, 2000 and November 7, 2000.

2000 *Mathematics Subject Classification.* Primary 65F15, 15A18.

*Key words and phrases.* *QR* algorithm, shift strategy, unitary Hessenberg matrices.

The research of the first author was supported by the Center for Computational Sciences at the University of Kentucky.

The research of the second author was supported in part by the National Science Foundation under grant DMS-8704196.

conjugate of a complex number $\alpha$ is denoted by $\bar{\alpha}$. We use the Euclidean norm $\|a\| := \|a\|_2$ for vectors and the spectral norm $\|A\| := \|A\|_2$ for matrices. Throughout, $A \in \mathbf{C}^{n\times n}$ will represent an upper Hessenberg matrix of order $n$ with entries $\alpha_{jk} := e_j^* A e_k$, $1 \le j \le k \le n$, in the upper triangular section and *positive* elements $\beta_k := e_{k+1}^* A e_k$, $1 \le k < n$, on the subdiagonal, where $e_k$ denotes the $k$th column vector of an identity matrix with appropriate dimension. The leading principal submatrices of $A$ will be expressed by $A_j := E_j^* A E_j \in \mathbf{C}^{j\times j}$, $1 \le j \le n$, where $E_j := [e_1, e_2, \dots, e_j] \in \mathbf{C}^{n\times j}$. Same structure and similar symbols apply to $\hat{A}$ and $A^{(k)}$, which will be defined in later sections. Iteration indices are usually indicated by superscript $k$ in parentheses. In case $A$ (or $\hat{A}$ or $A^{(k)}$) is *unitary* we use letter $U$ (or $\hat{U}$ or $U^{(k)}$) to represent the matrix. We let $\lambda(A)$ be the set of the eigenvalues of $A$. To avoid triviality we assume that $n$, the order of $A$, is at least 3.

## 2. Schur parameterization

It is straightforward to see that every unitary (upper) Hessenberg matrix $U \in \mathbf{C}^{n\times n}$ with *positive* subdiagonal elements $\{\beta_k\}_{k=1}^{n-1}$ can be uniquely factorized into a product of $n$ elementary unitary matrices [4, 6]:

$$U = U(\alpha_1, \alpha_2, \dots, \alpha_n) =: G_1(\alpha_1)G_2(\alpha_2)\cdots G_n(\alpha_n), \tag{2.1}$$

where

$$\begin{aligned} G_k(\alpha_k) &:= \operatorname{diag}\left(I_{k-1}, \begin{bmatrix} -\alpha_k & \beta_k \\ \beta_k & \bar{\alpha}_k \end{bmatrix}, I_{n-k-1}\right), \quad |\alpha_k|^2 + \beta_k^2 = 1, \ 1 \le k < n; \\ G_n(\alpha_n) &:= \operatorname{diag}\left(I_{n-1}, -\alpha_n\right), \quad |\alpha_n| = 1. \end{aligned}$$

Here, $I_j$ denotes the $j \times j$ identity matrix, and $\alpha_1, \alpha_2, \dots, \alpha_n$ are called the *Schur parameters* of $U$ [6]. These parameters can be determined from the top row and the subdiagonal of $U$:

$$\begin{aligned} \alpha_1 &= -e_1^* U e_1, \\ \alpha_k &= -e_1^* U e_k / \beta_1\beta_2\cdots\beta_{k-1}, \ 2 \le k \le n. \end{aligned}$$

To see this, we can multiply out the product $G_1(\alpha_1)G_2(\alpha_2)\cdots G_n(\alpha_n)$ and obtain

$$U = \begin{bmatrix} -\bar{\alpha}_0\alpha_1 & -\bar{\alpha}_0\beta_1\alpha_2 & \cdots & -\bar{\alpha}_0\beta_1\beta_2\cdots\beta_{k-1}\alpha_k & \cdots & -\bar{\alpha}_0\beta_1\beta_2\cdots\beta_{n-1}\alpha_n \\ \beta_1 & -\bar{\alpha}_1\alpha_2 & \cdots & -\bar{\alpha}_1\beta_2\cdots\beta_{k-1}\alpha_k & \cdots & -\bar{\alpha}_1\beta_2\cdots\beta_{n-1}\alpha_n \\ & \beta_2 & \ddots & \vdots & & \vdots \\ & & \ddots & -\bar{\alpha}_{k-1}\alpha_k & \cdots & -\bar{\alpha}_{k-1}\beta_k\cdots\beta_{n-1}\alpha_n \\ & & & \beta_k & \ddots & \vdots \\ & & & & \ddots & -\bar{\alpha}_{n-1}\alpha_n \end{bmatrix}, \tag{2.2}$$

where $\alpha_0 := 1$. We refer to the representation (2.1) as the *Schur parametric form* of $U$.

## 3. The shifted QR algorithm

Given an upper Hessenberg matrix $A \in \mathbf{C}^{n\times n}$ and a shift parameter $\lambda \in \mathbf{C}$, we form the $QR$ factorization of

$$A - \lambda I =: QR, \tag{3.1}$$

with $Q$ unitary and $R$ upper triangular with nonnegative diagonal elements. Here $Q$ is obtained by performing the Gram-Schmidt orthonormalizing process on the columns of $A - \lambda I$ from left to right, and hence is also upper Hessenberg. This factorization is unique if $\lambda \notin \lambda(A)$, and this is the case we usually assume hereafter. From $Q$ we define $\hat{A}$, the *QR transform* of $A$, by setting

$$\hat{A} := Q^*AQ = RQ + \lambda I. \tag{3.2}$$

The shifted $QR$ algorithm iterates the $QR$ transformation $A \to \hat{A}$, with an appropriate shift $\lambda$ chosen at each step:

$$\begin{aligned}
&A^{(1)} := A\\
&\text{for } k = 1, 2, 3, \dots\\
&A^{(k)} - \lambda^{(k)} I \ =: Q^{(k)}R^{(k)};\\
&A^{(k+1)} := R^{(k)}Q^{(k)} + \lambda^{(k)} I.
\end{aligned}$$

It is well known that the Hessenberg structure of $A^{(k)}$ and $Q^{(k)}$ is preserved and that all the $A^{(k)}$ are unitarily similar to each other. The efficiency of this algorithm depends critically on the choice of the shift sequence $\lambda^{(k)}$. In particular, if $\lambda^{(k)}$ could be chosen close to an isolated eigenvalue of $A$, then $\beta_{n-1}^{(k)}$ (the last subdiagonal element of $A^{(k)}$) would eventually decrease rapidly. As soon as $\beta_{n-1}^{(k)}$ becomes negligible to working precision, $\alpha_{nn}^{(k)}$ (the last diagonal element of $A^{(k)}$) may be taken as a computed eigenvalue; we can then delete the last row and column, and proceed with a matrix of lower order [1, 17, 18, 13]. In fact if any of the subdiagonal elements of $A^{(k)}$ vanishes, then the eigenproblem splits into that for two or more smaller Hessenberg matrices. An upper Hessenberg matrix is said to be *unreduced* if its subdiagonal elements are all nonzero [11]. For the convenience of theoretical analysis, there is no loss of generality in assuming that all the $\beta_j$ of $A$ are *positive* and hence by the following lemma that, if $\lambda^{(k)} \notin \lambda(A)$, all the $\beta_j^{(k)}$ of $A^{(k)}$ are positive, $1 \le j < n$.

**Lemma 1** (Basic relations in $QR$). *Let $\hat{A}$ be the QR transform of $A$ with shift $\lambda$. Then, for $1 \le k < n$,*

(a)

$$\beta_k = \sigma_k \rho_k, \qquad \hat{\beta}_k = \sigma_k \rho_{k+1},$$

*where $\beta_k, \hat{\beta}_k$, and $\sigma_k$ are, respectively, the subdiagonal elements of $A, \hat{A}$, and $Q$, and $\rho_k$ are the nonnegative diagonal elements of $R$;*

(b)

$$\rho_j \ge 0,\ \beta_k > 0 \text{ and } \lambda \notin \lambda(A) \Longrightarrow \rho_j > 0,\ \sigma_k > 0,\ \hat{\beta}_k > 0,\ 1 \le j \le n;$$

*hence*

$$\lambda \in \lambda(A) \iff \rho_n = 0 \iff \hat{\beta}_{n-1} = 0 \Longrightarrow \hat{\alpha}_{nn} = \lambda;$$

(c)

$$\hat{\beta}_k\hat{\beta}_{k+1}\cdots\hat{\beta}_{n-1} \le \beta_{k+1}\beta_{k+2}\cdots\beta_{n-1}\rho_n\,;$$

*in particular,*

$$\begin{array}{lrcl} \text{(i)} & \hat{\beta}_{n-1} & \le & \rho_n, \\ \text{(ii)} & \hat{\beta}_{n-2}\hat{\beta}_{n-1} & \le & \beta_{n-1}\rho_n, \\ \text{(iii)} & \hat{\beta}_{n-2}\hat{\beta}_{n-1}^2 & \le & \beta_{n-1}\rho_n^2. \end{array}$$

*Proof.* (a) This is straightforward by equating the corresponding subdiagonal elements on each side of the matrix equations (3.1) and (3.2), respectively.

(b) The implications are direct consequences of the relations stated in (a) and the fact that $\lambda \notin \lambda(A) \iff \rho_1\rho_2\cdots\rho_n > 0$, where $\{\rho_j\}_{j=1}^n$ are chosen nonnegative in the factorization $A - \lambda I = QR$, and that

$$\hat{\alpha}_{nn} - \lambda = e_n^*(\hat{A} - \lambda I)e_n = e_n^* RQe_n = \rho_n e_n^* Qe_n.$$

(c) From (a) we have

$$\begin{aligned} \hat{\beta}_k\hat{\beta}_{k+1}\cdots\hat{\beta}_{n-1} &= \sigma_k\rho_{k+1}\sigma_{k+1}\rho_{k+2}\cdots\sigma_{n-1}\rho_n \\ &= \sigma_k\beta_{k+1}\beta_{k+2}\cdots\beta_{n-1}\rho_n \\ &\le \beta_{k+1}\beta_{k+2}\cdots\beta_{n-1}\rho_n, \end{aligned}$$

since $0 < \sigma_k \le 1$. Setting $k = n-1$ and $k = n-2$ we get (i) and (ii), respectively. Multiplying (ii) with (i) side by side we obtain (iii). □

3.1. **Shift strategies.** To achieve rapid convergence (for the definition of convergence see Section 5) it is essential to incorporate an efficient strategy into the algorithm. We consider, in each step of the *QR* transformation from $A$ to $\hat{A}$, the following choices of the shift parameter $\lambda$ [18, 4, 13, 10]:

1) The Rayleigh shift (R-shift).
   $\lambda := \alpha_{nn}$, the last diagonal element of $A$.
2) The Wilkinson shift (W-shift).
   $\lambda$ is taken as an eigenvalue of $\begin{bmatrix} \alpha_{n-1,n-1} & \alpha_{n-1,n} \\ \beta_{n-1} & \alpha_{nn} \end{bmatrix}$, the trailing $2\times 2$ submatrix of $A$, which is closer to $\alpha_{nn}$; that is, we choose $\lambda$ to satisfy
   (i) $\delta(\lambda) := (\lambda - \alpha_{n-1,n-1})(\lambda - \alpha_{nn}) - \alpha_{n-1,n}\beta_{n-1} = 0$
   (ii) $|\lambda - \alpha_{nn}| \le \sqrt{|\alpha_{n-1,n}|\beta_{n-1}} \le |\lambda - \alpha_{n-1,n-1}|$.
3) The mixed shift (M-shift).
   $\lambda$ is taken as the R-shift if $\theta\beta_{n-2} \ge \beta_{n-1}$, the W-shift if $\theta\beta_{n-2} < \beta_{n-1}$, where $\theta$ is a positive parameter to be determined.

In case $A$ is unitary, it can be written in the Schur parametric form $A =: U(\alpha_1, \alpha_2, \ldots, \alpha_n)$, as expressed by (2.2). Hence the R-shift has the form

$$\lambda = -\bar{\alpha}_{n-1}\alpha_n,$$

and the W-shift is chosen as an eigenvalue $\lambda$ of

$$\begin{bmatrix} -\bar{\alpha}_{n-2}\alpha_{n-1} & -\bar{\alpha}_{n-2}\beta_{n-1}\alpha_n \\ \beta_{n-1} & -\bar{\alpha}_{n-1}\alpha_n \end{bmatrix},$$

which satisfies the following characteristic relations:

$$\begin{aligned} \delta(\lambda) &:= (\lambda + \bar{\alpha}_{n-2}\alpha_{n-1})(\lambda + \bar{\alpha}_{n-1}\alpha_n) + \bar{\alpha}_{n-2}\beta_{n-1}^2\alpha_n \\ &= \lambda^2 + (\bar{\alpha}_{n-2}\alpha_{n-1} + \bar{\alpha}_{n-1}\alpha_n)\lambda + \bar{\alpha}_{n-2}\alpha_n = 0, \end{aligned} \tag{3.3}$$

$$|\lambda + \bar{\alpha}_{n-1}\alpha_n| \le \sqrt{|\alpha_{n-2}|}\,\beta_{n-1} \le |\lambda + \bar{\alpha}_{n-2}\alpha_{n-1}|. \tag{3.4}$$

3.2. **Initial-value modification of the shifts.** Since unitary matrices (and scalar multiples of them) remain invariant under the *basic QR* algorithm (i.e., $\lambda^{(k)} \equiv 0$ for all $k$), it is therefore essential that *nonzero* shifts be taken in the unitary *QR* algorithm. Consider the simple unitary matrix

$$U = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

as an example. If we follow the conventional definition for either R- or W-shift to determine $\lambda$, then $\lambda = 0$ and the *unshifted QR* transformation produces no change of the matrix at all; that is, $\hat{U} = U$. This is indeed the case for any unitary Hessenberg matrices with $\beta_{n-1} = 1$ when the R-shift is applied, and with $\beta_{n-2} = \beta_{n-1} = 1$ when the W-shift is applied. (A thorough investigation of the invariant Hessenberg form under the *QR* algorithm with Francis' double shift was given by Parlett [11].) To avoid any such invariant cycling we make the following modification for the shift in the *QR* transformation [4]:

> *For unitary matrices in case the shift $\lambda$ following the usual rule is null, we take it to be any nonzero number with modulus equal to unity; for definiteness, we choose $\lambda = 1$. The* R-, W-, *and* M-*shift with such modification will be denoted as the* R$'$-, W$'$-, *and* M$'$*-shift, respectively.*

Note that throughout the entire *QR* iteration this modification on the shift sequence $\lambda^{(k)}$, if necessary, has *only* to be made *at the very initial step*, $k = 1$; subsequent values of $\lambda^{(k)}$ will *never* be null again (in fact, $|\lambda^{(k)}| \to 1$ eventually when $\beta_{n-1}^{(k)} \to 0$). The detailed analysis is given in Sections 6, 7, and 8.

## 4. THE *QR* FACTORIZATION OF $A - \lambda I$

In this section our attention is focused on the *QR* factorization of a shifted Hessenberg matrix $A - \lambda I$. Useful expressions and formulas that are crucial to the proof of global convergence will be derived.

4.1. **Characteristic polynomials.** First we express the characteristic polynomials of the leading principal submatrices $A_k$ of $A$ in terms of entries of the conformal sections of the factor matrices $Q$ and $R$. Partition $A - \lambda I = QR$ as

$$\left[\begin{array}{c|c} A_k - \lambda I_k & X \\ \hline \beta_k e_1 e_k^* & X \end{array}\right] = \left[\begin{array}{c|c} Q_k & X \\ \hline \sigma_k e_1 e_k^* & X \end{array}\right] \left[\begin{array}{c|c} R_k & X \\ \hline O & X \end{array}\right], \quad 1 \le k \le n,$$

where $A_k - \lambda I_k$, $Q_k$, and $R_k$ are square submatrices of size $k$, and the $X$'s are irrelevant submatrices of appropriate sizes. Clearly, from the upper triangularity of $R$,

$$A_k - \lambda I_k = Q_k R_k. \tag{4.1}$$

Since $Q$ is unitary Hessenberg with positive subdiagonal elements $\{\sigma_j\}_{j=1}^{n-1}$, by (2.1) it can be written in the product form $Q = G_1 G_2 \cdots G_n$, where $G_j = G_j(\gamma_j)$, $|\gamma_j|^2 + \sigma_j^2 = 1$, $1 \le j < n$; $G_n = G_n(\gamma_n)$, $|\gamma_n| = 1$. Each $Q_k := E_k^* Q E_k$ can then be expressed correspondingly in the product form

$$Q_k = \bar{G}_1 \bar{G}_2 \cdots \bar{G}_k,$$

where $\bar{G}_j := E_k^* G_j E_k$ is the $k$th leading section of $G_j$, $1 \le j \le k$. Note that while $G_j \in \mathbf{C}^{n\times n}$, $\bar{G}_j \in \mathbf{C}^{k\times k}$. Since $\det \bar{G}_j = -1$ for $1 \le j < k$ and $\det \bar{G}_k = -\gamma_k$, we see that

$$\det Q_k = (-1)^k \gamma_k$$

and from (4.1) we obtain the following formula:

$$\chi_k = \chi_k(\lambda) := \det(\lambda I_k - A_k) = \rho_1 \rho_2 \cdots \rho_k \gamma_k, \ 1 \le k \le n. \tag{4.2}$$

Here $\chi_k$ is the (monic) characteristic polynomial of $A_k$, $\rho_1, \rho_2, \ldots, \rho_k$ are the leading $k$ diagonal elements of $R$, and $\gamma_k$ is the $k$th Schur parameter of $Q$. Putting (4.2) in modulus form, we have

$$|\chi_k| = \rho_1 \rho_2 \cdots \rho_k |\gamma_k| = \rho_1 \rho_2 \cdots \rho_k \sqrt{1 - \sigma_k^2}, \ 1 \le k \le n, \tag{4.3}$$

where $\sigma_n := 0$. Observe that $|\chi_n| = \rho_1 \rho_2 \cdots \rho_n$, since $|\gamma_n| = 1$.

4.2. **Recurrence relations.** Next we introduce a specific form (see (4.9)) of the Szegö recurrence relations, which play a fundamental role in the convergence proof of the unitary $QR$ iteration.

We begin with the relations that the $\chi_k$ satisfy for an upper Hessenberg $A$ [17, p. 411]:

$$\begin{aligned} &\chi_0 := 1, \ \chi_1 = \lambda - \alpha_{11}; \\ &\chi_k = (\lambda - \alpha_{kk}) \chi_{k-1} - \beta_{k-1} \sum_{j=1}^{k-1} \chi_{j-1} \beta_j \beta_{j+1} \cdots \beta_{k-2} \alpha_{jk}, \ 1 < k \le n. \end{aligned} \tag{4.4}$$

If, in addition, $A$ is unitary, then the Schur parameterization of $A := U(\alpha_1, \alpha_2, \ldots, \alpha_n)$ gives (see (2.2))

$$\alpha_{jk} := e_j^* A e_k = -\bar{\alpha}_{j-1} \beta_j \beta_{j+1} \cdots \beta_{k-1} \alpha_k, \ 1 \le j \le k \le n, \ \alpha_0 := 1,$$

and, using this expression in (4.4), we get

$$\begin{aligned} \chi_k &= \lambda \chi_{k-1} - \sum_{j=1}^{k} \chi_{j-1} \beta_j \beta_{j+1} \cdots \beta_{k-1} \alpha_{jk} \\ &= \lambda \chi_{k-1} + \alpha_k \sum_{j=1}^{k} \bar{\alpha}_{j-1} \chi_{j-1} (\beta_j \beta_{j+1} \cdots \beta_{k-1})^2 \\ &=: \lambda \chi_{k-1} + \alpha_k \tilde{\chi}_{k-1}, \end{aligned} \tag{4.5}$$

where we have assigned $\tilde{\chi}_{k-1}$. Hence

$$
\begin{aligned}
\tilde{\chi}_k &:= \sum_{j=1}^{k+1} \bar{\alpha}_{j-1}\chi_{j-1}(\beta_j\beta_{j+1}\cdots\beta_k)^2 \\
&= \bar{\alpha}_k\chi_k + \beta_k^2 \sum_{j=1}^{k} \bar{\alpha}_{j-1}\chi_{j-1}(\beta_j\beta_{j+1}\cdots\beta_{k-1})^2 \\
&= \bar{\alpha}_k(\lambda\chi_{k-1} + \alpha_k\tilde{\chi}_{k-1}) + \beta_k^2\tilde{\chi}_{k-1}, \text{ from (4.5),} \\
&= \tilde{\chi}_{k-1} + \bar{\alpha}_k\lambda\chi_{k-1}, \text{ since } |\alpha_k|^2 + \beta_k^2 = 1. \qquad (4.6)
\end{aligned}
$$

Summarizing (4.5) and (4.6), we get the Szegö recurrence relations for the unitary Hessenberg matrix $U(\alpha_1, \alpha_2, \ldots, \alpha_n)$:

$$
\begin{aligned}
&\chi_0 := 1, \ \tilde{\chi}_0 := 1; \\
&\chi_k = \lambda\chi_{k-1} + \alpha_k\tilde{\chi}_{k-1}, \ \tilde{\chi}_k = \tilde{\chi}_{k-1} + \bar{\alpha}_k\lambda\chi_{k-1}, \ 1 \le k \le n.
\end{aligned} \qquad (4.7)
$$

The characteristic polynomials $\{\chi_k(\lambda)\}_{k=0}^n$ of the successive leading principal submatrices of $U$ are also called the (monic) Szegö polynomials associated with the Schur parameters $\{\alpha_k\}_{k=1}^n$ of $U$ [6]. It can be shown by induction that the auxiliary polynomials $\{\tilde{\chi}_k(\lambda)\}_{k=0}^n$ in (4.7) satisfy $\tilde{\chi}_k(\lambda) = \lambda^k\chi_k^c(1/\lambda)$, with the superscript $c$ denoting complex conjugation of the coefficients of a polynomial. Note that $\alpha_k = \chi_k(0)$. For further applications of these polynomials, see [7].

We may put $\tilde{\chi}_k$ in product form similar to that obtained in (4.2) for $\chi_k$ [6]:

$$\tilde{\chi}_k =: \rho_1\rho_2\cdots\rho_k\tilde{\gamma}_k, \ 1 \le k \le n. \qquad (4.8)$$

Then, since $\rho_1\rho_2\cdots\rho_{n-1} > 0$, (4.7) can be replaced by

$$
\begin{aligned}
&\gamma_0 := 1, \ \tilde{\gamma}_0 := 1; \\
&\rho_k\gamma_k = \lambda\gamma_{k-1} + \alpha_k\tilde{\gamma}_{k-1}, \ \rho_k\tilde{\gamma}_k = \tilde{\gamma}_{k-1} + \bar{\alpha}_k\lambda\gamma_{k-1}, \ 1 \le k \le n.
\end{aligned} \qquad (4.9)
$$

Note that

$$|\gamma_k|^2 + \sigma_k^2 = 1, \ |\tilde{\gamma}_k|^2 + |\lambda|^2\sigma_k^2 = 1, \ 1 \le k \le n, \qquad (4.10)$$

where $\sigma_n := 0$. To see why $|\tilde{\gamma}_k|^2 + |\lambda|^2\sigma_k^2 = 1$, we can make use of the two recurrence formulas presented in (4.9) (subtracting the square of the modulus form of one from that of the other), together with $|\alpha_k|^2 + \beta_k^2 = 1$ and $\beta_k = \sigma_k\rho_k$, to obtain

$$|\tilde{\gamma}_k|^2 - |\gamma_k|^2 = \sigma_k^2(|\tilde{\gamma}_{k-1}|^2 - |\lambda|^2|\gamma_{k-1}|^2), \ 1 \le k \le n,$$

from which an inductive argument, with the use of $|\gamma_k|^2 + \sigma_k^2 = 1$, proves the second identity in (4.10).

4.3. **A residual estimate.** In the Hermitian tridiagonal case a constructive proof for the global convergence of the shifted $QR$ algorithm was obtained by exploiting the connection between $QR$ and inverse iteration [3, 8, 13]. We generalize this approach and derive a residual bound for *normal* Hessenberg matrices.

Let $A$ be normal Hessenberg. Take the conjugate transpose of $A - \lambda I = QR$ and postmultiply by $Q$ to get

$$(A - \lambda I)^*Q = R^*.$$

Equating the last column on each side gives

$$(A - \lambda I)^*q_n = \rho_n e_n,$$

where $q_n := Qe_n$ and $\rho_n := e_n^* R e_n \geq 0$. Since $A$ is normal, $\|(A-\lambda I)q_n\| = \|(A-\lambda I)^* q_n\| = \rho_n$, and $\langle \lambda, q_n \rangle$ is an eigenpair of $A$ if and only if $\rho_n = 0$. Assume $\lambda \notin \lambda(A)$ and hence $\rho_n > 0$. Put $x := q_n/\rho_n$ so that

$$(A-\lambda I)^* x = e_n. \tag{4.11}$$

Partition $A - \lambda I$ and let

$$B := [B_1 \;\; B_2] := A - \lambda I = \left[\begin{array}{c|c} A_{n-2} - \lambda I_{n-2} & X \\ \hline \beta_{n-2} e_1 e_{n-2}^* & \tilde{A}_2 - \lambda I_2 \end{array}\right], \tag{4.12}$$

where

$$B_2 \;\; := \;\; \begin{bmatrix} X \\ \tilde{A}_2 - \lambda I_2 \end{bmatrix} \in \mathbf{C}^{n\times 2}, \; X \in \mathbf{C}^{(n-2)\times 2},$$

and

$$\tilde{A}_2 \;\; := \;\; \begin{bmatrix} \alpha_{n-1,n-1} & \alpha_{n-1,n} \\ \beta_{n-1} & \alpha_{nn} \end{bmatrix} \in \mathbf{C}^{2\times 2}. \tag{4.13}$$

We look for an upper bound on $\rho_n$ (or a lower bound on $\|x\| = 1/\rho_n$) by considering only the last two equations of (4.11), that is,

$$B_2^* x = e_2, \tag{4.14}$$

and calculating the norm of the "minimal solution" $\hat{x}$ of this underdetermined system in the sense that $\|\hat{x}\| \leq \|x\|$ for all the solutions $x$ of (4.14). This can easily be done by forming the $QR$ factorization of $B_2$. Let

$$B_2 =: Q_2 \begin{bmatrix} R_2 \\ O \end{bmatrix}$$

where $Q_2 \in \mathbf{C}^{n\times n}$ is unitary and $R_2 \in \mathbf{C}^{2\times 2}$ is upper triangular with positive diagonal elements. Then (4.14) is equivalent to

$$[R_2^* \;\; O] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = e_2,$$

where $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} := y := Q_2^* x$, and $y_1 \in \mathbf{C}^2$, $y_2 \in \mathbf{C}^{n-2}$. Since

$$\|x\|^2 = \|Q_2^* x\|^2 = \|y\|^2 = \|y_1\|^2 + \|y_2\|^2 \geq \|y_1\|^2,$$

the length of the minimal solution $\hat{x}$ is obtained by setting $y_2 = 0$ and computing $\|y_1\|$ from $y_1$, which is the unique solution of the triangular system

$$R_2^* y_1 = e_2. \tag{4.15}$$

Putting

$$R_2 := \begin{bmatrix} \rho_{11} & \rho_{12} \\ 0 & \rho_{22} \end{bmatrix}$$

in (4.15), we get $y_1 = [0, 1/\rho_{22}]^T$ and

$$\|y_1\|^2 = \frac{1}{\rho_{22}^2} = \|\hat{x}\|^2 \leq \|x\|^2 = \frac{1}{\rho_n^2}, \tag{4.16}$$

where, from the triangularity of $R_2$ and the unitariness of $Q_2$,

$$\rho_{22}^2 = \frac{\det(R_2^* R_2)}{e_1^* R_2^* R_2 e_1} = \frac{\det(B_2^* B_2)}{e_1^* B_2^* B_2 e_1}. \tag{4.17}$$

Hence combining (4.17) with (4.16) we have

$$\rho_n^2 \leq \frac{\det(B_2^* B_2)}{e_1^* B_2^* B_2 e_1}. \tag{4.18}$$

Next, we seek a simpler formula to compute $B_2^* B_2 \in \mathbf{C}^{2\times 2}$ in (4.18) by taking advantage of the normality of $A$. We claim

$$B_2^* B_2 = \beta_{n-2}^2 e_1 e_1^* + (\tilde{A}_2 - \lambda I_2)(\tilde{A}_2 - \lambda I_2)^*, \tag{4.19}$$

where $\tilde{A}_2$ is the lower right $2 \times 2$ submatrix of $A$, as was defined in (4.13). Since $B := [B_1 \;\, B_2] := A - \lambda I$ is normal, formula (4.19) is readily obtained by equating the lower right blocks of $B^*B$ and $BB^*$ (cf. (4.12)):

$$B^*B = \begin{bmatrix} B_1^* \\ B_2^* \end{bmatrix} [B_1 \;\, B_2] = \begin{bmatrix} X & X \\ X & B_2^* B_2 \end{bmatrix},$$

$$\begin{aligned} BB^* &= \left[\begin{array}{c|c} A_{n-2} - \lambda I_{n-2} & X \\ \hline \beta_{n-2} e_1 e_{n-2}^* & \tilde{A}_2 - \lambda I_2 \end{array}\right] \left[\begin{array}{c|c} (A_{n-2} - \lambda I_{n-2})^* & \beta_{n-2} e_{n-2} e_1^* \\ \hline X & (\tilde{A}_2 - \lambda I_2)^* \end{array}\right] \\ &= \begin{bmatrix} X & X \\ X & D \end{bmatrix}, \end{aligned}$$

where $D := \beta_{n-2}^2 e_1 e_1^* + (\tilde{A}_2 - \lambda I_2)(\tilde{A}_2 - \lambda I_2)^* \in \mathbf{C}^{2\times 2}$ and the $X$'s are irrelevant submatrices of appropriate sizes. From (4.18) and (4.19) we obtain an upper bound for $\rho_n^2$ expressed in terms of $\lambda$, $\beta_{n-2}$, and the entries of $\tilde{A}_2$ in the following lemma.

**Lemma 2** (A residual bound). *Given a normal Hessenberg matrix $A \in \mathbf{C}^{n\times n}$ with positive subdiagonal elements $\{\beta_k\}_{k=1}^{n-1}$, let $\rho_n$ be the last diagonal element of the upper triangular matrix $R$ in the QR factorization of $A - \lambda I$, where $\lambda \in \mathbf{C}$ is the shift, and let $q_n$ be the last column of the unitary matrix $Q$. Then*

$$\|(A - \lambda I)q_n\|^2 = \rho_n^2 \leq \frac{\beta_{n-2}^2 \beta_{n-1}^2 + \beta_{n-2}^2 |\alpha_{nn} - \lambda|^2 + |\delta(\lambda)|^2}{\beta_{n-2}^2 + |\alpha_{n-1,n-1} - \lambda|^2 + |\alpha_{n-1,n}|^2}, \tag{4.20}$$

*where* $\delta(\lambda) := \det(\tilde{A}_2 - \lambda I_2) = (\alpha_{n-1,n-1} - \lambda)(\alpha_{nn} - \lambda) - \alpha_{n-1,n}\beta_{n-1}$.

*Remark.* The bound we obtain in (4.20) is in fact a least upper bound, which can be attained in extreme cases. For example, let

$$A = \begin{bmatrix} \alpha + \sqrt{\beta^2 - \delta^2} & 0 & 0 & -\delta \\ \delta & \alpha & 0 & \sqrt{\beta^2 - \delta^2} \\ 0 & \beta & \alpha & 0 \\ 0 & 0 & \beta & \alpha \end{bmatrix}, \tag{4.21}$$

where $\alpha \in \mathbf{C}$, $\beta \geq \delta > 0$ (cf. [9, p. 132]). Clearly $A$ is normal (Hessenberg). If we choose $\lambda = \alpha$ (which is both the R- and W-shift), then it is easy to see that

$$Q = \frac{1}{\beta}(A - \alpha I), \;\; R = \beta I,$$

and that equality holds in this case: $\|(A - \lambda I)q_n\| = \rho_n = \beta$. Actually, with shift $\lambda = \alpha$, matrix $A$ is *invariant* under the *QR* transformation defined by (3.1) and (3.2).

## 5. Properties related to convergence

We say the *QR* algorithm is convergent if the last subdiagonal element $\beta_{n-1}^{(k)}$ of $A^{(k)}$ in the iterating process converges to zero; in other words, the last row of $A^{(k)}$ tends to a limit form $\lambda_n^{(k)} e_n^*$, where $\lambda_n^{(k)} \in \lambda(A)$ [3, 13]. As we shall see, convergence of $\beta_{n-1}^{(k)}$ depends on how we choose the shift sequence $\lambda^{(k)}$, and the selection of an efficient shift strategy is of crucial importance to the implementation of the algorithm. The lemmas and theorem presented in this section will be applied repeatedly later when we discuss convergence of the algorithm with the various shift strategies. The next lemma is a modification of [18, Lemma 1] used for tridiagonal *QR*.

**Lemma 3** (Boundedness property). *With any of the shift strategies mentioned before, all elements $\alpha_{ij}^{(k)}, \beta_j^{(k)}$ of $A^{(k)}$ are bounded by $\|A\|$, the spectral norm of $A$, and all elements derived from the QR transformation of $A^{(k)}$ (and hence all elements of $R^{(k)}$) are bounded by $2\|A\|$ for all $k$, the iteration index. Consequently, $\chi_j^{(k)}(\lambda^{(k)}) := \det(\lambda^{(k)} I_j - A_j^{(k)})$ and $\delta_j^{(k)} := \rho_1^{(k)} \rho_2^{(k)} \cdots \rho_j^{(k)}$ are bounded for all $k$ and $1 \le j \le n$.*

*Proof.* It is clear that $|\alpha_{ij}^{(k)}|, \beta_j^{(k)} \le \|A^{(k)}\| = \|A\|$, likewise for the shift $\lambda^{(k)}$ because it is chosen as an eigenvalue of a submatrix of $A^{(k)}$. Let $\rho_{ij}^{(k)}$ be the $i$th component of the $j$th column vector $r_j^{(k)}$ in $R^{(k)}$, and $\rho_{jj}^{(k)} =: \rho_j^{(k)}$. Then

$$\begin{aligned} |\rho_{ij}^{(k)}| &\le \|r_j^{(k)}\| \le \|R^{(k)}\| = \|Q^{(k)} R^{(k)}\| = \|A^{(k)} - \lambda^{(k)} I\| \\ &\le \|A^{(k)}\| + \|\lambda^{(k)} I\| \le 2\|A\|. \end{aligned}$$

Therefore $\delta_j^{(k)}$ are bounded and $|\chi_j^{(k)}| \le \delta_j^{(k)}$, by (4.3). $\square$

To estimate the rate of convergence for $\beta_{n-1}^{(k)} \to 0$, we need a relation between $\hat{\beta}_{n-1}$ and $\beta_{n-1}$. Jiang and Zhang [10, Lemma 2] proposed a relation for real symmetric tridiagonal matrices. In the next lemma we extend the relation to Hessenberg matrices; at this stage normality of the matrix $A$ is *not* required.

**Lemma 4** (Relations for the subdiagonal elements). *Let $\hat{A}$ be the QR transform of $A$ with shift $\lambda$. Then*

(a) $\delta_k^2 = \beta_k^2 \delta_{k-1}^2 + |\chi_k(\lambda)|^2 = \sum_{j=0}^{k} (|\chi_j(\lambda)| \beta_{j+1} \beta_{j+2} \cdots \beta_k)^2,$
*where* $\delta_k := \rho_1 \rho_2 \cdots \rho_k, \ 1 \le k \le n, \ \delta_0 := 1, \ \beta_n := 0,$

(b) $\hat{\beta}_{n-1} = \left[ \frac{\delta_{n-2} |\chi_n(\lambda)|}{\beta_{n-1}^2 \delta_{n-2}^2 + |\chi_{n-1}(\lambda)|^2} \right] \beta_{n-1}.$

*Proof.*

$$\begin{aligned} \text{(a) } \delta_k^2 &= \sigma_k^2 \delta_k^2 + |\gamma_k|^2 \delta_k^2, \text{ because } \sigma_k^2 + |\gamma_k|^2 = 1, \\ &= \sigma_k^2 (\rho_k^2 \delta_{k-1}^2) + |\gamma_k|^2 \delta_k^2, \text{ from the definition of } \delta_k, \\ &= \beta_k^2 \delta_{k-1}^2 + |\chi_k(\lambda)|^2, \text{ since } \beta_k = \sigma_k \rho_k \text{ and } \chi_k(\lambda) = \rho_1 \rho_2 \cdots \rho_k \gamma_k. \end{aligned}$$

By induction on $\delta_j^2$, $j = k-1, k-2, \ldots, 1$, we get, with $\delta_0 := 1 = \chi_0(\lambda)$,

$$\delta_k^2 = \sum_{j=0}^{k} (|\chi_j(\lambda)| \beta_{j+1} \beta_{j+2} \cdots \beta_k)^2.$$

$$
\begin{aligned}
\text{(b) } \hat{\beta}_{n-1} &= \sigma_{n-1}\rho_n = \left(\frac{\rho_n}{\rho_{n-1}}\right)\beta_{n-1}, \text{ from Lemma 1(a),} \\
&= \left(\frac{\delta_{n-2}\delta_n}{\delta_{n-1}^2}\right)\beta_{n-1}, \text{ from the definition of } \delta_k \text{ given in (a),} \\
&= \left(\frac{\delta_{n-2}|\chi_n(\lambda)|}{\delta_{n-1}^2}\right)\beta_{n-1}, \text{ since } |\chi_n(\lambda)| = \delta_n \text{ from (4.3),} \\
&= \left[\frac{\delta_{n-2}|\chi_n(\lambda)|}{\beta_{n-1}^2\delta_{n-2}^2 + |\chi_{n-1}(\lambda)|^2}\right]\beta_{n-1}, \text{ by applying (a) to } \delta_{n-1}^2. \quad \square
\end{aligned}
$$

**Lemma 5** (Basic facts from normality). *Let $A$ be normal. Then*

(a) *the eigenvalues of $A$ are mutually distinct and*

(b) $|\chi_n - (\lambda - \alpha_{nn})\chi_{n-1}| \le \beta_{n-1}^2\delta_{n-2}$.

*Proof.* (a) It is clear that the eigenvalues of an *unreduced* Hessenberg matrix have unit geometric multiplicity. Since normal matrices are unitarily diagonalizable, the eigenvalues of an unreduced normal Hessenberg matrix can only be mutually distinct.

(b) From the recurrence relation for $\chi_k$ in (4.4), we have

$$
\begin{aligned}
&|\chi_k - (\lambda - \alpha_{kk})\chi_{k-1}| \\
&\quad = \beta_{k-1}\left|\sum_{j=1}^{k-1}\chi_{j-1}\beta_j\beta_{j+1}\cdots\beta_{k-2}\alpha_{jk}\right| \\
&\quad \le \beta_{k-1}\left(\sum_{j=1}^{k-1}|\chi_{j-1}\beta_j\beta_{j+1}\cdots\beta_{k-2}|^2\right)^{1/2}\left(\sum_{j=1}^{k-1}|\alpha_{jk}|^2\right)^{1/2}, \\
&\qquad \text{by applying the Cauchy-Schwarz inequality,} \\
&\quad = \beta_{k-1}\delta_{k-2}\left(\sum_{j=1}^{k-1}|\alpha_{jk}|^2\right)^{1/2}, \text{ from Lemma 4(a) after shifting the index } j.
\end{aligned}
$$

Now let $k = n$; since $A$ is normal, the last factor $(\sum_{j=1}^{n-1}|\alpha_{jn}|^2)^{1/2} = \beta_{n-1}$ and the inequality is proved. $\square$

**Theorem 6** (Convergence properties for the $QR$ iteration). *Assume $A$ is normal. Let $A^{(k)}$ be the unreduced $QR$ iterates of $A$ with either the* R*-shift or the* W*-shift, $\lambda^{(k)}$. If $\beta_{n-1}^{(k)} \to 0$, then :*

(a) *$\lambda^{(k)} \to \lambda_n$ for some $\lambda_n \in \lambda(A)$;*

(b) *$|\chi_{n-1}^{(k)}(\lambda^{(k)})| \ge \sigma^{n-1} + O(\varepsilon)$, where $\sigma := \min_{j\ne k}\{|\lambda_j - \lambda_k| : \lambda_j, \lambda_k \in \lambda(A)\} > 0$ and $\varepsilon$ is an arbitrarily small number*;

(c) *$\rho_n^{(k)} \to 0$, and $\{\rho_j^{(k)}\}_{j=1}^{n-1}$ are bounded away from zero.*

*Proof.* (a) If $\beta_{n-1}^{(k)} \to 0$, then, according to Hessenberg structure of $A^{(k)}$, the last diagonal element $\alpha_{nn}^{(k)}$ clusters to an eigenvalue of $A$, say $\lambda_n^{(k)}$, which may depend on $k$; that is,

$$
\beta_{n-1}^{(k)} \to 0 \Longrightarrow |\alpha_{nn}^{(k)} - \lambda_n^{(k)}| \to 0 \text{ for some } \lambda_n^{(k)} \in \lambda(A).
$$

From the $QR$ transformation: $A - \lambda I = QR$, $\hat{A} - \lambda I = RQ$ we infer, with the normality of $A$, that

$$\begin{aligned}
&(\hat{\beta}_{n-1}^2 + |\hat{\alpha}_{nn} - \lambda|^2)^{1/2} \\
&\quad = \|e_n^*(\hat{A} - \lambda I)\| = \|e_n^* RQ\| = \|e_n^* R\| = \rho_n \\
&\quad \leq \|Re_n\| = \|QRe_n\| = \|(A - \lambda I)e_n\| = \|(A - \lambda I)^* e_n\| \\
&\quad = (\beta_{n-1}^2 + |\alpha_{nn} - \lambda|^2)^{1/2}. \qquad (5.1)
\end{aligned}$$

Hence $|\alpha_{nn}^{(k+1)} - \lambda^{(k)}| \to 0$, as $\beta_{n-1}^{(k)} \to 0$ and

$$|\alpha_{nn}^{(k)} - \lambda^{(k)}| \begin{cases} = 0 & \text{with the R-shift} \\ \leq \sqrt{|\alpha_{n-1,n}^{(k)}|\beta_{n-1}^{(k)}} \to 0 & \text{with the W-shift.} \end{cases} \qquad (5.2)$$

Therefore, $|\alpha_{nn}^{(k+1)} - \alpha_{nn}^{(k)}| \to 0$ and, since the eigenvalues of $A$ are mutually distinct (Lemma 5(a)), the sequence $\alpha_{nn}^{(k)}$ converges to a *fixed* eigenvalue $\lambda_n$ of $A$. So does $\lambda^{(k)}$, the shift sequence, by (5.2); that is, $\lambda^{(k)} \to \lambda_n$ for some $\lambda_n \in \lambda(A)$.

(b) Let $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$ be the remaining distinct eigenvalues in any order so that for $\lambda \in \mathbf{C}$

$$\chi_n(\lambda) = (\lambda - \lambda_n) \prod_{j=1}^{n-1} (\lambda - \lambda_j). \qquad (5.3)$$

Let $\psi^{(k)} \leftrightarrow \phi^{(k)}$ denote $|\psi^{(k)} - \phi^{(k)}| \to 0$ as $k \to \infty$. From Lemma 5(b) there follows, as $\beta_{n-1}^{(k)} \to 0$,

$$\begin{aligned}
\chi_n(\lambda^{(k)}) &\leftrightarrow (\lambda^{(k)} - \alpha_{nn}^{(k)})\chi_{n-1}^{(k)}(\lambda^{(k)}) \\
&\leftrightarrow (\lambda^{(k)} - \lambda_n)\chi_{n-1}^{(k)}(\lambda^{(k)}), \qquad (5.4)
\end{aligned}$$

since $\alpha_{nn}^{(k)} \to \lambda_n$, as proved in (a), and $|\chi_{n-1}^{(k)}(\lambda^{(k)})|$ is bounded for all $k$, by Lemma 3. Comparing (5.4) with (5.3) we have, as $\beta_{n-1}^{(k)} \to 0$,

$$\chi_{n-1}^{(k)}(\lambda^{(k)}) \leftrightarrow \prod_{j=1}^{n-1} (\lambda^{(k)} - \lambda_j).$$

Since $\lambda^{(k)} \to \lambda_n$ as $\beta_{n-1}^{(k)} \to 0$, this gives, for a sufficiently small $\varepsilon$,

$$|\chi_{n-1}^{(k)}(\lambda^{(k)})| \geq \sigma^{n-1} + O(\varepsilon), \qquad (5.5)$$

where $\sigma := \min_{j \neq k}\{|\lambda_j - \lambda_k| : \lambda_j, \lambda_k \in \lambda(A)\} > 0$.

(c) This is a direct consequence of (a), (b) and Lemma 3: As $\beta_{n-1}^{(k)} \to 0$, $\lambda^{(k)} \to \lambda_n \in \lambda(A)$ and hence

$$|\chi_n(\lambda^{(k)})| = \rho_1^{(k)} \rho_2^{(k)} \cdots \rho_{n-1}^{(k)} \rho_n^{(k)} \to 0,$$

but from Lemma 4(a)

$$\begin{aligned}
\rho_1^{(k)} \rho_2^{(k)} \cdots \rho_{n-1}^{(k)} =: \delta_{n-1}^{(k)} &= \sqrt{\beta_{n-1}^{(k)2} \delta_{n-2}^{(k)2} + |\chi_{n-1}^{(k)}(\lambda^{(k)})|^2}, \\
&\geq |\chi_{n-1}^{(k)}(\lambda^{(k)})| \geq \sigma^{n-1} + O(\varepsilon), \text{ by (5.5).} \quad \square
\end{aligned}$$

Finally in this section, we examine, for unitary Hessenberg matrices with $\beta_{n-1} = 1$, how $\hat{\beta}_{n-1}$ can change after one $QR$ step if a nonzero shift is taken. Eberlein and Huang [4, Lemma 1] showed that $\hat{\beta}_{n-1} < 1$ by a plane-rotation argument; we prove this instead through a constructive inequality which is useful for the analysis of numerical decrease of $\hat{\beta}_{n-1}$ from unity.

**Lemma 7** (Nonzero shift in the extreme case). *Let $\hat{U}$ be the $QR$ transform of $U$ with any nonzero shift $\lambda$, $0 < |\lambda| \le 1$. Assume $\beta_{n-1} = 1$. Then*

$$\hat{\beta}_{n-1} \le \frac{|\lambda|^2\sqrt{1-\sigma_{n-2}^2} + \sqrt{1-|\lambda|^2\sigma_{n-2}^2}}{1+|\lambda|^2(1-\sigma_{n-2}^2)} < 1, \tag{5.6}$$

*where $\sigma_{n-2} := e_{n-1}^* Q e_{n-2}$.*

*Proof.* From the recurrence relations (4.7) we have

$$\begin{aligned} \chi_n &= \lambda\chi_{n-1} + \alpha_n\tilde{\chi}_{n-1} \\ &= \lambda(\lambda\chi_{n-2} + \alpha_{n-1}\tilde{\chi}_{n-2}) + \alpha_n(\tilde{\chi}_{n-2} + \bar{\alpha}_{n-1}\lambda\chi_{n-2}) \\ &= \lambda^2\chi_{n-2} + \alpha_n\tilde{\chi}_{n-2}, \text{ since } \beta_{n-1} = 1 \Longleftrightarrow \alpha_{n-1} = 0, \end{aligned}$$

and hence, after taking modulus on each side and eliminating the common factor $\rho_1\rho_2\cdots\rho_{n-2}$,

$$\rho_{n-1}\rho_n = |\lambda^2\gamma_{n-2} + \alpha_n\tilde{\gamma}_{n-2}|, \tag{5.7}$$

because $\chi_k = \rho_1\rho_2\cdots\rho_k\gamma_k$, $|\gamma_n| = 1$, and $\tilde{\chi}_k = \rho_1\rho_2\cdots\rho_k\tilde{\gamma}_k$. Similarly, working on

$$\chi_{n-1} = \lambda\chi_{n-2} + \alpha_{n-1}\tilde{\chi}_{n-2} = \lambda\chi_{n-2}$$

gives, with the modulus on each side squared,

$$\rho_{n-1}^2|\gamma_{n-1}|^2 = |\lambda|^2|\gamma_{n-2}|^2.$$

Combining this relation with the identity

$$\rho_{n-1}^2\sigma_{n-1}^2 = \beta_{n-1}^2$$

side by side (note that $|\gamma_{n-1}|^2 + \sigma_{n-1}^2 = 1$ and $\beta_{n-1} = 1$) we get

$$\rho_{n-1}^2 = 1 + |\lambda|^2|\gamma_{n-2}|^2. \tag{5.8}$$

From (5.7) and (5.8), together with the fact that $\sigma_{n-1} = \frac{\beta_{n-1}}{\rho_{n-1}} = \frac{1}{\rho_{n-1}}$, we obtain

$$\hat{\beta}_{n-1} = \sigma_{n-1}\rho_n = \frac{\rho_{n-1}\rho_n}{\rho_{n-1}^2} = \frac{|\lambda^2\gamma_{n-2} + \alpha_n\tilde{\gamma}_{n-2}|}{1+|\lambda|^2|\gamma_{n-2}|^2}.$$

Therefore,

$$\hat{\beta}_{n-1} \le \frac{|\lambda|^2|\gamma_{n-2}| + |\tilde{\gamma}_{n-2}|}{1+|\lambda|^2|\gamma_{n-2}|^2} = \frac{|\lambda|^2\sqrt{1-\sigma_{n-2}^2} + \sqrt{1-|\lambda|^2\sigma_{n-2}^2}}{1+|\lambda|^2(1-\sigma_{n-2}^2)}$$

by applying the triangle inequality and the identities (4.10) in tandem. To show that the right-hand side of the above inequality is less than 1, we square both sides and obtain, after some calculations,

$$\hat{\beta}_{n-1}^2 \le 1 - |\lambda|^2\left[\frac{1-\sqrt{1-\sigma_{n-2}^2}\sqrt{1-|\lambda|^2\sigma_{n-2}^2}}{1+|\lambda|^2(1-\sigma_{n-2}^2)}\right]^2 < 1,$$

since $0 < |\lambda| \le 1$ and $0 < \sigma_{n-2} \le 1$. $\square$

*Remark.* To obtain the result (5.6) of Lemma 7, we have used the triangle inequality. By varying $\alpha_n$ in (5.7) we see that, for a given *fixed* shift, say $\lambda = 1$, the bound can be attained. (However, in the *usual* shift strategies, $\lambda$ will depend on $\alpha_n$.) Let us examine, with $\lambda = 1$ in (5.6), how close to unity the right-hand side of

$$\hat{\beta}_{n-1} \leq \frac{2\sqrt{1-\sigma_{n-2}^2}}{2-\sigma_{n-2}^2} \tag{5.9}$$

could conceivably be. Asymptotically, as $\sigma_{n-2} \to 0$,

$$\frac{2\sqrt{1-\sigma_{n-2}^2}}{2-\sigma_{n-2}^2} \sim 1 - \frac{1}{8}\sigma_{n-2}^4 - \frac{1}{8}\sigma_{n-2}^6 - \cdots .$$

Now, for $|\lambda| \leq 1$, $\beta_{n-2} = \sigma_{n-2}\rho_{n-2} \leq 2\sigma_{n-2}$ (see Lemma 3). Hence the decrease

$$1 - \hat{\beta}_{n-1} \geq \frac{1}{128}\beta_{n-2}^4 \tag{5.10}$$

cannot be arbitrarily small. On the other hand, if equality in (5.10) were (nearly) possible and $\beta_{n-2} = \varepsilon$ is small, a tiny perturbation of order $\beta_{n-2}^4$ could easily not be detected by the computer. Consequently, the *numerical* value of $\hat{\beta}_{n-1}$ could stay at 1 with the nonzero special shift $\lambda = 1$. However, if the W-shift were used in this exceptional case (i.e., $\beta_{n-1} = 1, \beta_{n-2} = \varepsilon$), $\hat{\beta}_{n-1}$ would drop to a small value $\leq \varepsilon$ in one *QR* step (using the inequality $\hat{\beta}_{n-1} \leq \beta_{n-2}$ given later in Lemma 11(d)).

## 6. Convergence with the Rayleigh shift

There is an intimate connection between the Rayleigh quotient iteration (RQI) and the *QR* algorithm with the Rayleigh shift, and the convergence properties of RQI can be translated (with a proper selection of the initial vector) into statements about *QR* with the R-shift [13, pp.144-148]. Extensive analyses and rigorous proofs about convergence of RQI were given by Ostrowski, Kahan, and Parlett; for details and further references, see [12, 13]. In the *QR* language, results were given by Wilkinson [18] for symmetric tridiagonal matrices and by Eberlein and Huang [4] for unitary Hessenberg matrices. In fact, more general (but weaker) convergence results for normal Hessenberg matrices, closely related to the minimal residual property of the Rayleigh quotient, were given in Buurema's thesis [2]; see also the results obtained by Watkins and Elsner [16] using nested subspace iteration analysis. Here we derive some of these properties through the use of recurrence formulas, and summarize them for comparison with those from the W-shift, which is analyzed in the next section. These results will again be used in the final section to establish the convergence of *QR* with a mixed shift strategy.

For normal matrices, the monotonicity property $\hat{\beta}_{n-1} \leq \beta_{n-1}$ with the R-shift is well known [12], and can be readily seen from (5.1). In the unitary case (with the R$'$-shift), we derive a *strict* inequality between $\hat{\beta}_{n-1}$ and $\beta_{n-1}$ (and thus, in exact arithmetic, a stationary state of $\beta_{n-1}$ will not occur), and from which global convergence of either $\beta_{n-1}$ or $\beta_{n-2}$ to zero is a direct consequence.

**Lemma 8** (Monotonicity property). *Let $\hat{U}$ be the QR transform of $U$ with the* R$'$*-shift $\lambda$. Then $\hat{\beta}_{n-1} < \beta_{n-1}$; more precisely,*

(a) $\hat{\beta}_{n-1} = \sqrt{1-(1-\beta_{n-1}^2)\sigma_{n-2}^2}\;\sigma_{n-1}^2\beta_{n-1} < \beta_{n-1}$ *if* $\beta_{n-1} < 1$,

(b) $\hat{\beta}_{n-1} \leq \frac{2\sqrt{1-\sigma_{n-2}^2}}{2-\sigma_{n-2}^2} < 1$ *if* $\beta_{n-1} = 1$.

*Proof.* (a) If $\beta_{n-1} < 1$, then $\lambda = -\bar{\alpha}_{n-1}\alpha_n \neq 0$ and

$$\begin{aligned} \chi_n &= \lambda\chi_{n-1} + \alpha_n\tilde{\chi}_{n-1} \\ &= \lambda(\lambda\chi_{n-2} + \alpha_{n-1}\tilde{\chi}_{n-2}) + \alpha_n(\tilde{\chi}_{n-2} + \bar{\alpha}_{n-1}\lambda\chi_{n-2}) \\ &= \alpha_n(1 - |\alpha_{n-1}|^2)\tilde{\chi}_{n-2} \\ &= \alpha_n\beta_{n-1}^2\tilde{\chi}_{n-2}. \end{aligned}$$

Hence

$$|\chi_n| = \beta_{n-1}^2|\tilde{\chi}_{n-2}|,$$

and

$$\rho_n = |\tilde{\gamma}_{n-2}|\sigma_{n-1}\beta_{n-1}$$

after using (4.2),(4.8) and Lemma 1(a). Therefore,

$$\begin{aligned} \hat{\beta}_{n-1} &= \sigma_{n-1}\rho_n = |\tilde{\gamma}_{n-2}|\sigma_{n-1}^2\beta_{n-1} \\ &= \sqrt{1-(1-\beta_{n-1}^2)\sigma_{n-2}^2}\;\sigma_{n-1}^2\beta_{n-1}, \\ &\qquad \text{since } |\tilde{\gamma}_{n-2}|^2+|\lambda|^2\sigma_{n-2}^2 = 1 \text{ and } |\lambda| = |\alpha_{n-1}|, \\ &< \beta_{n-1}, \text{ because } 0 < \beta_{n-1} < 1 \text{ and } 0 < \sigma_k \leq 1. \end{aligned} \tag{6.1}$$

(b) If $\beta_{n-1} = 1$, then $\lambda = 1$ by definition of the R$'$-shift and the result is just a special case of (5.6). □

The result given in the following theorem is already known [4, Lemma 2], and here we prove it in a different way by using formulas derived from the recurrence relations.

**Theorem 9** (Global convergence). *Let $U^{(k)}$ be the QR iterates of $U$ with the exclusive use of the R$'$-shift. Then either $\beta_{n-1}^{(k)} \to 0$ or $\beta_{n-2}^{(k)} \to 0$.*

*Proof.* By Lemma 8 the sequence $\beta_{n-1}^{(k)}$ decreases monotonically and thus tends to a limit $\delta$, say. If $\delta = 0$, then $\beta_{n-1}^{(k)} \to 0$. If $\delta > 0$, then, from Lemma 8(a),

$$\frac{\beta_{n-1}^{(k+1)}}{\beta_{n-1}^{(k)}} = \sqrt{1-(1-\beta_{n-1}^{(k)2})\sigma_{n-2}^{(k)2}}\;\sigma_{n-1}^{(k)2}$$

and the left-hand side tends to unity since $\beta_{n-1}^{(k)} \to \delta > 0$. Hence $\sigma_{n-1}^{(k)} \to 1$, $\sigma_{n-2}^{(k)} \to 0$ on the right-hand side, and therefore

$$\beta_{n-2}^{(k)} = \rho_{n-2}^{(k)}\sigma_{n-2}^{(k)} \to 0$$

because $\rho_{n-2}^{(k)}$ is bounded, by Lemma 3. □

In practice, convergence of $\beta_{n-2}^{(k)} \to 0$ (while $\beta_{n-1}^{(k)} \to \delta > 0$) is exceedingly slow as compared to that of $\beta_{n-1}^{(k)} \to 0$ which, if it occurs, has a cubic rate as we provide a simple proof in the next theorem. Note that a proof for the more general case, namely, that for normal matrices (indeed, for matrices with properties weaker than normality) the *QR* iteration with the *generalized* Rayleigh-quotient shift has cubic rates if it converges, was given by Watkins and Elsner [16] using subspace iteration technique.

**Theorem 10** (Local convergence). *Assume $A$ is normal. Let $\hat{A}$ be the QR transform of $A$ with the* R*-shift $\lambda$. If $\beta_{n-1} \to 0$, then $\hat{\beta}_{n-1} = O(\beta_{n-1}^3)$, that is, the asymptotic rate of convergence is cubic.*

*Proof.*

$$\begin{aligned}
\hat{\beta}_{n-1} &= \left[\frac{\delta_{n-2}|\chi_n(\lambda)|}{\beta_{n-1}^2\delta_{n-2}^2 + |\chi_{n-1}(\lambda)|^2}\right]\beta_{n-1}, \text{ from Lemma 4(b)},\\
&\le \left[\frac{\delta_{n-2}^2}{\beta_{n-1}^2\delta_{n-2}^2 + |\chi_{n-1}(\lambda)|^2}\right]\beta_{n-1}^3,\\
&\quad \text{since } |\chi_n(\lambda)| \le \beta_{n-1}^2\delta_{n-2} \text{ by Lemma 5(b)},\\
&= O(\beta_{n-1}^3) \text{ as } \beta_{n-1} \to 0, \text{ from Lemma 3 and Theorem 6(b)}. \quad \square
\end{aligned}$$

## 7. Convergence with the Wilkinson shift

We now arrive at the main results of this paper: global convergence (Theorem 12) and local convergence (Theorem 13) of the *QR* iteration with the (modified) Wilkinson shift for unitary Hessenberg matrices [14]. We begin with a technical lemma in which a constructive analysis for the decrease of $\beta_{n-2}\beta_{n-1}^2$ in one *QR* step is given, through the use of the basic relations stated in Lemma 1 and the residual bound (4.20) obtained in Lemma 2. This approach was used by Parlett [13, Chapter 8] in the Hermitian tridiagonal case; see also [3, Lemma 7.4].

**Lemma 11** (One-step changes and relations). *Let $\hat{U}$ be the QR transform of $U$ with the* W*-shift $\lambda$. Then :*

(a) $\hat{\beta}_{n-1} \le \rho_n \le \sqrt{1+\sqrt{1-\beta_{n-2}^2}}\,\beta_{n-1} < \sqrt{2}\,\beta_{n-1}$;

(b) $\hat{\beta}_{n-1}^2 \le \rho_n^2 \le \omega(\beta_{n-2},\beta_{n-1})\beta_{n-2}\beta_{n-1} \le \beta_{n-2}\beta_{n-1}$,
*where*

$$\omega(\beta_{n-2},\beta_{n-1}) := \frac{\beta_{n-2}\beta_{n-1} + \beta_{n-2}\sqrt{1-\beta_{n-2}^2}\,\beta_{n-1}}{\beta_{n-2}^2 + \sqrt{1-\beta_{n-2}^2}\beta_{n-1}^2 + (1-\beta_{n-2}^2)\beta_{n-1}^2},$$

*and*
(i)

$$\begin{aligned}
0 &< \omega(\beta_{n-2},\beta_{n-1})\\
&\le \left\{\begin{array}{ll}
\min\left[\frac{1}{\sqrt{2-\beta_{n-1}^2}}, \beta_{n-2}\right] & \textit{if } \beta_{n-2} > \frac{\sqrt{3}}{2}\\
\min\left[\frac{1}{\sqrt{2-\beta_{n-1}^2}}, \frac{1}{2}\sqrt{1+\frac{1}{\sqrt{1-\beta_{n-2}^2}}}\right] & \textit{if } \beta_{n-2} \le \frac{\sqrt{3}}{2}
\end{array}\right\} \le 1,
\end{aligned}$$

(ii) $\omega(\beta_{n-2},\beta_{n-1}) \to 1$ *if and only if* $\beta_{n-1} \to 1$ *and* $\beta_{n-2} \to 1$;

(c) $\hat{\beta}_{n-2}\hat{\beta}_{n-1}^2 \le \omega(\beta_{n-2},\beta_{n-1})\beta_{n-2}\beta_{n-1}^2 \le \beta_{n-2}\beta_{n-1}^2$;

(d) *if* $\beta_{n-1} = 1$, *then* $\hat{\beta}_{n-1} \le \beta_{n-2}$;

(e) *if* $\beta_{n-2} = 1$, *then the* W*-shift degenerates to the* R*-shift, i.e.,* $\lambda = -\bar{\alpha}_{n-1}\alpha_n$;

(f) $\chi_n(\lambda) = \beta_{n-2}^2(\alpha_{n-1}\lambda + \alpha_n)\tilde{\chi}_{n-3}(\lambda)$.

*Proof.* First of all, (b) is readily obtained from the basic relation (c)(i) of Lemma 1, the residual bound (4.20) derived in Lemma 2, the characteristic relations (3.3), (3.4) for the W-shift, and the Schur parametric form of $U$ given by (2.2); properties

(i) and (ii) of $\omega(\beta_{n-2}, \beta_{n-1})$ are derived in the Appendix. (c) is from (b) and Lemma 1(c)(iii). From (b) we have

$$\rho_n^2 \leq \omega(\beta_{n-2}, \beta_{n-1})\beta_{n-2}\beta_{n-1} \leq (1 + \sqrt{1 - \beta_{n-2}^2})\beta_{n-1}^2 < 2\beta_{n-1}^2;$$

this gives (a). Setting $\beta_{n-1} = 1$ in (b), we obtain, with $\omega(\beta_{n-2}, 1) = \beta_{n-2}$,

$$\hat{\beta}_{n-1}^2 \leq \rho_n^2 \leq \beta_{n-2}^2;$$

this gives (d). Since $|\alpha_{n-2}| = \sqrt{1 - \beta_{n-2}^2} = 0$ if $\beta_{n-2} = 1$, (e) is trivial from (3.4). Finally, to prove (f), substitutions of the recurrence relations from (4.7) give

$$\begin{aligned}
\chi_n &= \lambda\chi_{n-1} + \alpha_n\tilde{\chi}_{n-1} \\
&= \lambda(\lambda + \bar{\alpha}_{n-1}\alpha_n)\chi_{n-2} + (\alpha_{n-1}\lambda + \alpha_n)\tilde{\chi}_{n-2} \\
&= \lambda[\lambda^2 + (\bar{\alpha}_{n-2}\alpha_{n-1} + \bar{\alpha}_{n-1}\alpha_n)\lambda + \bar{\alpha}_{n-2}\alpha_n]\chi_{n-3} \\
&\quad + [\alpha_{n-2}\lambda^2 + (\alpha_{n-2}\bar{\alpha}_{n-1}\alpha_n + \alpha_{n-1})\lambda + \alpha_n]\tilde{\chi}_{n-3} \\
&= \beta_{n-2}^2(\alpha_{n-1}\lambda + \alpha_n)\tilde{\chi}_{n-3},
\end{aligned}$$

where, for the last equality to hold, we apply (3.3), the characteristic equation for the W-shift. □

**Theorem 12** (Glocal convergence). *Let $U^{(k)}$ be the QR iterates of $U$ with the $\mathrm{W}'$-shift used exclusively. Then $\beta_{n-1}^{(k+1)3}$ can be majorized by $\sqrt{2}\ \beta_{n-2}^{(k)}\beta_{n-1}^{(k)2}$ which is monotonically convergent to zero with a substantially decreasing ratio $\omega(\beta_{n-2}^{(k)}, \beta_{n-1}^{(k)})$, and hence $\beta_{n-1}^{(k)} \to 0$.*

*Proof.* If, and only if, $\beta_{n-2} = 1$ and $\beta_{n-1} = 1$ in the very beginning, the unit shift $\lambda = 1$ is applied. (Only in this exceptional case does the W-shift become null.) Then, by Lemma 7, $\hat{\beta}_{n-1} < 1$ in one *QR* step. Therefore, with the $\mathrm{W}'$-shift, there is no loss of generality in assuming that the starting value of $\beta_{n-1}^{(k)}$ is less than unity, say $\beta_{n-1}^{(1)} := \beta < 1$. This implies, from Lemma 11(b)(i), that

$$\omega(\beta_{n-2}^{(1)}, \beta_{n-1}^{(1)}) \leq \frac{1}{\sqrt{2 - \beta^2}} < 1. \tag{7.1}$$

At each *QR* step $U \to \hat{U}$,

$$\hat{\beta}_{n-2}\hat{\beta}_{n-1}^2 \leq \omega(\beta_{n-2}, \beta_{n-1})\beta_{n-2}\beta_{n-1}^2 \leq \beta_{n-2}\beta_{n-1}^2 \tag{7.2}$$

by Lemma 11(c). So $\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2}$ form a bounded monotonically decreasing sequence which has a limit, say $\delta$. We claim $\delta = 0$. For if $\delta > 0$, then $\omega(\beta_{n-2}^{(k)}, \beta_{n-1}^{(k)}) \to 1$ as $\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2} \to \delta > 0$. From Lemma 11(b)(ii) this implies that $\beta_{n-1}^{(k)} \to 1$ and $\beta_{n-2}^{(k)} \to 1$; hence $\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2} \to 1$. But from properties (7.1) and (7.2), for $k \geq 2$,

$$\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2} \leq \left[\prod_{j=1}^{k-1} \omega(\beta_{n-2}^{(j)}, \beta_{n-1}^{(j)})\right] \beta_{n-2}^{(1)}\beta_{n-1}^{(1)2} \leq \frac{\beta^2}{\sqrt{2 - \beta^2}} < \beta^2,$$

a fixed number which is strictly less than unity, a contradiction. Therefore,

$$\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2} \searrow 0.$$

Since from Lemma 11(a) and (b)

$$\hat{\beta}_{n-1}^3 < \sqrt{2}\ \beta_{n-2}\beta_{n-1}^2,$$

at each step $\beta_{n-1}^{(k+1)3}$ is dominated by $\sqrt{2}\,\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2}$ which converges monotonically to zero with a ratio $\omega(\beta_{n-2}^{(k)},\beta_{n-1}^{(k)}) \le \frac{1}{\sqrt{2-\beta_{n-1}^{(k)2}}}$. □

We now examine the asymptotic behavior of $\beta_{n-1}^{(k)}$ as it converges to zero with the use of the W-shift. The iteration index $k$ is usually suppressed and, to represent $|\psi^{(k)} - \phi^{(k)}| \to 0$ as $k \to \infty$, we use the notation $\psi \leftrightarrow \phi$.

**Theorem 13** (Local convergence). *Let $\hat{U}$ be the QR transform of $U$ with the W$'$-shift $\lambda$. Then, as $\beta_{n-1} \to 0$,*

(a) $|\chi_n(\lambda)| \leftrightarrow \beta_{n-2}^2\beta_{n-1}^2|\tilde{\chi}_{n-3}(\lambda)| = O(\beta_{n-2}^2\beta_{n-1}^2)$,

(b) $\hat{\beta}_{n-1} = O(\beta_{n-2}^2\beta_{n-1}^3)$, *that is, the rate of convergence is cubic in $\beta_{n-1}$.*

*Proof.* (a) We claim, as $\beta_{n-1} \to 0$, the following asymptotic relations:

(i) $\lambda \leftrightarrow -\bar{\alpha}_{n-1}\alpha_n$

(ii) $\alpha_{n-1}\lambda + \alpha_n \leftrightarrow \alpha_n\beta_{n-1}^2$.

While (ii) comes from (i) directly with some simple calculations (note that $|\alpha_{n-1}|^2 + \beta_{n-1}^2 = 1$ and $|\alpha_n| = 1$), (i) comes from a characteristic property of the W-shift given in (3.4): $|\lambda + \bar{\alpha}_{n-1}\alpha_n| \le \sqrt{|\alpha_{n-2}|}\,\beta_{n-1}$. Therefore,

$$\begin{aligned} \chi_n(\lambda) &= \beta_{n-2}^2(\alpha_{n-1}\lambda + \alpha_n)\tilde{\chi}_{n-3}(\lambda), \text{ from Lemma 11(f)}, \\ &\leftrightarrow \alpha_n\beta_{n-2}^2\beta_{n-1}^2\tilde{\chi}_{n-3}(\lambda), \text{ by (ii)}, \end{aligned}$$

and

$$|\chi_n(\lambda)| \leftrightarrow \beta_{n-2}^2\beta_{n-1}^2|\tilde{\chi}_{n-3}(\lambda)| = O(\beta_{n-2}^2\beta_{n-1}^2) \text{ as } \beta_{n-1} \to 0,$$

since $|\tilde{\chi}_{n-3}(\lambda)|$ is bounded.

(b)

$$\begin{aligned} \hat{\beta}_{n-1} &= \left[\frac{\delta_{n-2}|\chi_n(\lambda)|}{\beta_{n-1}^2\delta_{n-2}^2 + |\chi_{n-1}(\lambda)|^2}\right]\beta_{n-1}, \text{ from Lemma 4(b)}, \\ &\leftrightarrow \left[\frac{\delta_{n-2}|\tilde{\chi}_{n-3}(\lambda)|}{|\chi_{n-1}(\lambda)|^2}\right]\beta_{n-2}^2\beta_{n-1}^3 \text{ as } \beta_{n-1} \to 0, \text{ by (a)}, \\ &= O(\beta_{n-2}^2\beta_{n-1}^3), \text{ from Lemma 3 and Theorem 6(b)}. \quad \square \end{aligned}$$

*Remark.* Let us take a closer look at why, with the Wilkinson shift, the asymptotic rate of convergence for unitary Hessenberg matrices is cubic. (That for Hermitian tridiagonal matrices can only be shown to be quadratic [18].) From the characteristic relations (3.4) and (3.3) for the W-shift $\lambda$, we have

$$|\lambda + \bar{\alpha}_{n-1}\alpha_n| \le \sqrt{|\alpha_{n-2}|}\,\beta_{n-1} \tag{7.3}$$

$$|\lambda + \bar{\alpha}_{n-2}\alpha_{n-1}|\,|\lambda + \bar{\alpha}_{n-1}\alpha_n| = |\alpha_{n-2}|\beta_{n-1}^2. \tag{7.4}$$

We see directly from (7.3) that, as $\beta_{n-1} \to 0$, $|\lambda + \bar{\alpha}_{n-1}\alpha_n| = O(\beta_{n-1})$ *at least*; however, in the unitary case, the factor $|\lambda + \bar{\alpha}_{n-2}\alpha_{n-1}|$ in (7.4) is always bounded away from zero (without any further assumption like $\beta_{n-2} \to 0$, as is usually made in the tridiagonal case [8, 13], in order to guarantee cubic convergence of $\beta_{n-1}$, see the analysis given below); consequently $|\lambda + \bar{\alpha}_{n-1}\alpha_n| = O(\beta_{n-1}^2)$ as $\beta_{n-1} \to 0$ by (7.4); this is equivalent to $|\alpha_{n-1}\lambda + \alpha_n| = O(\beta_{n-1}^2)$ as $\beta_{n-1} \to 0$, shown by (ii) in the proof of Theorem 13(a).

We analyze this in more detail. Recall that, as $\beta_{n-1} \to 0$, the shift $\lambda \to \lambda_n$ for some fixed eigenvalue $\lambda_n \in \lambda(U)$ (Theorem 6(a)), and so

$$|\lambda| \to 1 \text{ and } |\alpha_{n-1}| \to 1 \text{ as } \beta_{n-1} \to 0 \tag{7.5}$$

because $|\lambda_n| = 1$ and $|\alpha_{n-1}|^2 + \beta_{n-1}^2 = 1$.

Now examine the behavior of $|\lambda + \bar{\alpha}_{n-2}\alpha_{n-1}|$ as $\beta_{n-1} \to 0$ by checking the $\beta_{n-2}^{(k)}$ sequence. If there is a subsequence $\beta_{n-2}^{(j)} \to 0$, then for this subsequence, $-\bar{\alpha}_{n-2}^{(j)}\alpha_{n-1}^{(j)} \leftrightarrow \lambda_{n-1}^{(j)}$ for some $\lambda_{n-1}^{(j)} \in \lambda(U)$ distinct from $\lambda_n$ anyway (see (2.2) and Lemma 5(a)). Hence

$$|\lambda^{(j)} + \bar{\alpha}_{n-2}^{(j)}\alpha_{n-1}^{(j)}| \leftrightarrow |\lambda_n - \lambda_{n-1}^{(j)}| \geq \sigma > 0. \tag{7.6}$$

For the rest of the sequence, $\beta_{n-2}^{(k)} > \varepsilon$ ( $\Longleftrightarrow |\alpha_{n-2}^{(k)}| < \sqrt{1-\varepsilon^2}$ ) for some $\varepsilon$, where $0 < \varepsilon < 1$. Then

$$|\lambda^{(k)} + \bar{\alpha}_{n-2}^{(k)}\alpha_{n-1}^{(k)}| \geq |\lambda^{(k)}| - |\alpha_{n-2}^{(k)}|\, |\alpha_{n-1}^{(k)}| \geq 1 - \sqrt{1-\varepsilon^2} > 0 \tag{7.7}$$

eventually, by (7.5). We conclude from (7.6) and (7.7) that, as $\beta_{n-1} \to 0$,

$$|\lambda + \bar{\alpha}_{n-2}\alpha_{n-1}| \geq \delta > 0 \text{ for some } \delta < \min\{\sigma,\ 1-\sqrt{1-\varepsilon^2}\},$$

that is, $|\lambda + \bar{\alpha}_{n-2}\alpha_{n-1}|$ is always bounded below from zero.

## 8. Convergence with the mixed shift

For theoretical interest we propose a general mixed shift strategy, with which the *QR* iteration has global convergence and cubic rates at least, and of which the modified Wilkinson shift (in Section 7) and the Eberlein-Huang shift [4] can be viewed as special cases.

**Theorem 14** (Global convergence). *Let $\theta$ be a real number, $0 \leq \theta < \infty$. Let $U^{(k)}$ be the QR iterates with the following shift strategy:*

$$\begin{cases} \text{if } \theta\beta_{n-2}^{(k)} \geq \beta_{n-1}^{(k)}, & \text{use the R}'\text{-shift,} \\ \text{if } \theta\beta_{n-2}^{(k)} < \beta_{n-1}^{(k)}, & \text{use the W}'\text{-shift.} \end{cases} \tag{8.1}$$

*Then $\beta_{n-1}^{(k)} \to 0$; in particular,*

(a) *if $\theta = 0$, then the W$'$-shift is used exclusively,*

(b) *if $0 \leq \theta < 1$, then $\beta_{n-1}^{(k)}$ can be majorized by a sequence which is monotonically convergent to zero,*

(c) *if $\theta \geq 1$, then $\beta_{n-1}^{(k+1)} < \beta_{n-1}^{(k)}$.*

*Proof.* Similarly with the argument given in the proof of Theorem 12, we may assume the starting $\beta_{n-1}^{(1)} =: \beta < 1$ in all cases, and thereafter the R$'$- and W$'$-shifts are the same as the R- and W-shifts. The implication of (a) is trivial from (8.1), because $0 < \beta_{n-1}^{(k)}$ is always assumed and convergence of $\beta_{n-1}^{(k)}$ with the W$'$-shift was proved in Theorem 12.

(b) Consider one step of *QR*: Since

$$\begin{aligned} \hat{\beta}_{n-1}^2 &\leq \rho_n^2 \\ &\leq \begin{cases} \beta_{n-1}^2 \leq \theta\beta_{n-2}\beta_{n-1} & \text{if the R-shift is used (by (6.1) and (8.1))} \\ \omega(\beta_{n-2}, \beta_{n-1})\beta_{n-2}\beta_{n-1} & \text{if the W-shift is used (from Lemma 11(b)),} \end{cases} \end{aligned}$$

we have

$$\begin{aligned}\hat{\beta}_{n-2}\hat{\beta}_{n-1}^2 &\le \beta_{n-1}\rho_n^2 \text{ (from Lemma 1(c) (iii))}\\ &\le \max\{\theta,\ \omega(\beta_{n-2},\beta_{n-1})\}\beta_{n-2}\beta_{n-1}^2,\end{aligned}$$

where $0 \le \theta < 1,\ 0 < \omega(\beta_{n-2},\beta_{n-1}) < 1$, and

$$\hat{\beta}_{n-1}^3 < \begin{cases} \beta_{n-1}^3 \le \theta\beta_{n-2}\beta_{n-1}^2 & \text{if the R-shift is used (Lemma 8 and (8.1))}\\ \sqrt{2}\,\beta_{n-2}\beta_{n-1}^2 & \text{if the W-shift is used (Lemma 11(a) and (b)).}\end{cases}$$

Therefore,

$$\beta_{n-1}^{(k+1)3} < \sqrt{2}\,\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2}$$

and

$$\beta_{n-2}^{(k)}\beta_{n-1}^{(k)2} \searrow 0,$$

by the same argument as stated in the proof of Theorem 12.

(c) We show, for $\theta \ge 1$, $\beta_{n-1}^{(k+1)} < \beta_{n-1}^{(k)}$ and $\beta_{n-1}^{(k)} \searrow 0$.

Clearly $\hat{\beta}_{n-1} < \beta_{n-1}$ if the R-shift is used. For the W-shift

$$\begin{aligned}\hat{\beta}_{n-1}^2 &\le \omega(\beta_{n-2},\beta_{n-1})\beta_{n-2}\beta_{n-1} \text{ (from Lemma 11(b))}\\ &< \frac{1}{\theta}\omega(\beta_{n-2},\beta_{n-1})\beta_{n-1}^2 \text{ (using (8.1))}\\ &< \beta_{n-1}^2 \text{ (since } \theta \ge 1 \text{ and } 0 < \omega(\beta_{n-2},\beta_{n-1}) < 1).\end{aligned} \tag{8.2}$$

Hence the monotonic decreasing of $\beta_{n-1}^{(k)}$ holds.

To further conclude that $\beta_{n-1}^{(k)} \searrow 0$, we consider two situations (which are mutually exclusive):

(i) The W-shift is applied *infinitely* many times. In this case

$$\beta_{n-1}^{(k+1)2} < \begin{cases} \beta_{n-1}^{(k)2} & \text{if the R-shift is used}\\ \frac{1}{\theta\sqrt{2-\beta^2}}\beta_{n-1}^{(k)2} & \text{if the W-shift is used,}\end{cases}$$

where $\frac{1}{\theta\sqrt{2-\beta^2}}$ is a fixed number less than unity, because from (8.2) we know

$$\begin{aligned}\frac{1}{\theta}\omega(\beta_{n-2}^{(k)},\beta_{n-1}^{(k)}) &\le \frac{1}{\theta\sqrt{2-\beta_{n-1}^{(k)2}}} \text{ (by Lemma 11(b)(i))}\\ &\le \frac{1}{\theta\sqrt{2-\beta^2}} \text{ (since } \beta_{n-1}^{(k)} \text{ is decreasing and } \beta := \beta_{n-1}^{(1)})\\ &< 1 \text{ (since } \theta \ge 1 \text{ and } \beta < 1).\end{aligned}$$

Therefore, the entire sequence $\beta_{n-1}^{(k)} \searrow 0$.

(ii) The R-shift is applied *ultimately.* Then we know, from (8.1), that $\theta\beta_{n-2}^{(k)} \ge \beta_{n-1}^{(k)}$ eventually holds. If $\beta_{n-1}^{(k)} \searrow \delta > 0$, then, from (the proof of) Theorem 9, $\beta_{n-2}^{(k)} \to 0$ which by (8.1) implies $\beta_{n-1}^{(k)} \to 0$, a contradiction. □

Though not important in practice, the sequence $\beta_{n-1}^{(k)}$ *itself* might not decrease monotonically for $0 \le \theta < 1$ in the above theorem. To guarantee a monotonic decreasing of $\beta_{n-1}^{(k)}$ and a constructive convergence analysis, we could choose a function $\theta(\beta_{n-2},\beta_{n-1})$, instead of a fixed number $\theta$, in the mixed shift strategy such that $\omega(\beta_{n-2},\beta_{n-1}) \le \theta(\beta_{n-2},\beta_{n-1}) \le 1$; we further present the following

**Corollary 15** (Global convergence). *Let*

$$\theta(\beta_{n-2}, \beta_{n-1}) := \min\{\phi(\beta_{n-1}), \psi(\beta_{n-2})\}, \tag{8.3}$$

*where*

$$\phi(\beta_{n-1}) := \frac{1}{\sqrt{2-\beta_{n-1}^2}}$$

*and*

$$\psi(\beta_{n-2}) := \begin{cases} \beta_{n-2} & \text{if } \beta_{n-2} > \frac{\sqrt{3}}{2} \\ \frac{1}{2}\sqrt{1+\frac{1}{\sqrt{1-\beta_{n-2}^2}}} & \text{if } \beta_{n-2} \le \frac{\sqrt{3}}{2}. \end{cases}$$

*Let $U^{(k)}$ be the QR iterates with the following shift strategy:*

$$\begin{cases} \text{if } \theta^{(k)}\beta_{n-2}^{(k)} \ge \beta_{n-1}^{(k)}, & \text{use the R}'\text{-shift}, \\ \text{if } \theta^{(k)}\beta_{n-2}^{(k)} < \beta_{n-1}^{(k)}, & \text{use the W-shift}, \end{cases} \tag{8.4}$$

*where*

$$\theta^{(k)} := \theta(\beta_{n-2}^{(k)}, \beta_{n-1}^{(k)}), \quad \frac{1}{\sqrt{2}} < \theta^{(k)} \le 1.$$

*Then $\beta_{n-1}^{(k)}$ is monotonically decreasing to zero.*

*Proof.* Similar as part (b) of the proof given to the preceding theorem, with either shift (note that $\omega(\beta_{n-2}, \beta_{n-1}) \le \theta(\beta_{n-2}, \beta_{n-1}) \le 1$)

$$\hat{\beta}_{n-1}^2 \le \rho_n^2 \le \min\{\beta_{n-1}^2, \theta(\beta_{n-2}, \beta_{n-1})\beta_{n-2}\beta_{n-1}\},$$

and we have

$$\hat{\beta}_{n-1}^3 = \hat{\beta}_{n-1}\hat{\beta}_{n-1}^2 \le \hat{\beta}_{n-1}\rho_n^2 \le \theta(\beta_{n-2}, \beta_{n-1})\beta_{n-2}\beta_{n-1}^2 \to 0,$$

because (from Lemma 1(c) (iii) and Lemma 11(b),(c))

$$\hat{\beta}_{n-2}\hat{\beta}_{n-1}^2 \le \beta_{n-1}\rho_n^2 \le \theta(\beta_{n-2}, \beta_{n-1})\beta_{n-2}\beta_{n-1}^2$$

which implies, similarly as (7.2) did in the proof of Theorem 12, that $\beta_{n-2}\beta_{n-1}^2 \to 0$. □

*Remark.* Eberlein and Huang [4] proposed a mixed shift strategy as follows:

(i) If $\beta_{n-1} = 1$, choose $|\lambda| = 1$. (Initial-value modification)

(ii) If $\sqrt{2}\beta_{n-2} \le \beta_{n-1}$, use the W-shift.

(iii) If neither of the above holds, use the R-shift.

From the *modified* version of the shifts defined in Section 3 (i.e., R-shift $\equiv 1$ instead of 0 when $\beta_{n-1} = 1$, and W-shift $\equiv 1$ instead of 0 when $\beta_{n-2} = \beta_{n-1} = 1$) this shift strategy can essentially be considered (with (i) included in (iii)) as a special case of the general shift strategy given in Theorem 14 with parameter $\theta = \sqrt{2}$.

In numerical computation, Eberlein-Huang's shift strategy may have the following drawback: If $\beta_{n-1} = 1$ and $\beta_{n-2}$ is very small, then, from the Remark following Lemma 7, there is a possibility that the shift under (i) may employ no decrease of $\beta_{n-1}$ from unity at all, on a digital computer with finite precision arithmetic. We do not worry about this if strategy (8.1) with $0 \le \theta < 1$ (or strategy (8.4)) is applied, because then the W-shift is used and in one step $\hat{\beta}_{n-1}$ becomes very small, by Lemma 11(d). In other words, one should always use the W-shift, instead of the R-shift or a nonzero shift, in case $\beta_{n-2}$ is much smaller than $\beta_{n-1}$, even when $\beta_{n-1} = 1$.

Finally we show, for all the mixed shift strategies considered in this section, that the rate of convergence is at least cubic, rather than just quadratic, as was claimed in [4, p.104] in the special case $\theta = \sqrt{2}$.

**Theorem 16** (Local convergence). *Let $\hat{U}$ be the QR transform of $U$ with the following shift strategy:*

$$\begin{cases} \text{if } \theta\beta_{n-2} \geq \beta_{n-1}, & \text{use the } \mathrm{R}'\text{-shift,} \\ \text{if } \theta\beta_{n-2} < \beta_{n-1}, & \text{use the } \mathrm{W}'\text{-shift,} \end{cases}$$

*where $\theta$ is either a fixed nonnegative real number or $\theta = \theta(\beta_{n-2}, \beta_{n-1})$ as defined by (8.3). Then, as $\beta_{n-1} \to 0$, $\hat{\beta}_{n-1} = O(\beta_{n-1}^3)$ at least.*

*Proof.* For the R-shift, $\hat{\beta}_{n-1} = O(\beta_{n-1}^3)$ by Theorem 10; for the W-shift, $\hat{\beta}_{n-1} = O(\beta_{n-2}^2\beta_{n-1}^3) = O(\beta_{n-1}^5)$ by Theorem 13 and $\theta\beta_{n-2} < \beta_{n-1}$ in this case. □

## Appendix. Basic properties of $\omega(\beta_{n-2}, \beta_{n-1})$

In this section we do not follow the notational conventions used in the main text; here, all the quantities dealt with, whether represented by Greek letters or not, are *positive real numbers less than or equal to one.* In Lemma 11(b) we have, with $\beta_{n-1} =: x$ and $\beta_{n-2} =: y$,

$$\begin{aligned} \omega(\beta_{n-2}, \beta_{n-1}) &:= \frac{\beta_{n-2}\beta_{n-1} + \beta_{n-2}\sqrt{1-\beta_{n-2}^2}\,\beta_{n-1}}{\beta_{n-2}^2 + \sqrt{1-\beta_{n-2}^2}\,\beta_{n-1}^2 + (1-\beta_{n-2}^2)\beta_{n-1}^2} \\ &= \frac{xy + xy\sqrt{1-y^2}}{y^2 + x^2\sqrt{1-y^2} + x^2(1-y^2)} \\ &= \frac{xy}{1-(1-x^2)\sqrt{1-y^2}} \\ &=: f(x,y). \end{aligned}$$

Properties of $f(x,y)$ are given in the following

**Lemma.** *Let $f(x,y) = \frac{xy}{1-(1-x^2)\sqrt{1-y^2}}$, $0 < x \leq 1$, $0 < y \leq 1$. Then*

(a)

$$\begin{aligned} f(x,y) &\leq \min\{\phi(x), \psi(y)\} \\ &= \begin{cases} \min\left\{\frac{1}{\sqrt{2-x^2}}, y\right\} & \text{if } y > \frac{\sqrt{3}}{2} \\ \min\left\{\frac{1}{\sqrt{2-x^2}}, \frac{1}{2}\sqrt{1+\frac{1}{\sqrt{1-y^2}}}\right\} & \text{if } y \leq \frac{\sqrt{3}}{2}, \end{cases} \end{aligned}$$

*where*

$$\phi(x) := \frac{1}{\sqrt{2-x^2}}$$

*and*

$$\psi(y) := \begin{cases} y & \text{if } y > \frac{\sqrt{3}}{2} \\ \frac{1}{2}\sqrt{1+\frac{1}{\sqrt{1-y^2}}} & \text{if } y \leq \frac{\sqrt{3}}{2} \end{cases}$$

*are increasing functions of $x$ and $y$, respectively, and*

$$\frac{1}{\sqrt{2}} < \phi(x) \le 1, \; \frac{1}{\sqrt{2}} < \psi(y) \le 1;$$

(b) $0 < f(x,y) \le 1$; $f(x,y) \to 1$ *if and only if* $x \to 1$ *and* $y \to 1$.

*Proof.* For each fixed $x$, $0 < x \le 1$, $f(x,y)$ has a global maximum at $y = x\sqrt{2-x^2}$. Hence

$$f(x,y) \le f(x, x\sqrt{2-x^2}) = \frac{1}{\sqrt{2-x^2}} =: \phi(x). \tag{A.1}$$

For each fixed $y$, $0 < y \le 1$, $f(x,y)$ has a global maximum at

$$x = \begin{cases} 1 & \text{if } y > \frac{\sqrt{3}}{2} \\ \sqrt{\frac{1}{\sqrt{1-y^2}} - 1} & \text{if } y \le \frac{\sqrt{3}}{2}. \end{cases}$$

Hence

$$f(x,y) \le \begin{cases} f(1,y) & = y & \text{if } y > \frac{\sqrt{3}}{2} \\ f(\sqrt{\frac{1}{\sqrt{1-y^2}} - 1}, y) & = \frac{1}{2}\sqrt{1 + \frac{1}{\sqrt{1-y^2}}} & \text{if } y \le \frac{\sqrt{3}}{2} \end{cases} \tag{A.2}$$
$$=: \psi(y).$$

Combining (A.1) and (A.2) we have (a). It is clear that $\phi(x)$ and $\psi(y)$ are increasing functions with values in the interval $(\frac{1}{\sqrt{2}}, 1]$.

(b) is immediate from (a): Note that $f(x,y)$ is continuous at $(1,1)$ and $f(1,1) = 1$; as $f(x,y) \to 1$,

$$f(x,y) \le \min\left\{\frac{1}{\sqrt{2-x^2}}, y\right\} \le 1$$

implies that $x \to 1$ and $y \to 1$. □

## Acknowledgments

The authors would like to thank the referee for providing helpful suggestions, which led to improvements in the presentation of this paper. The work of the first author is based on part of his Ph.D. thesis [14]. A preliminary form of this research appeared as [15].

## References

1. H. Bowdler, R. S. Martin, C. Reinsch, and J. H. Wilkinson, *The QR and QL algorithms for symmetric matrices*, Numer. Math. **11** (1968), 293-306.
2. H. J. Buurema, *A geometric proof of convergence for the QR method*, doctoral dissertation, Rijksuniversiteit Te Groningen, Groningen, The Netherlands, 1970. MR **52:**4597
3. T. J. Dekker and J. F. Traub, *The shifted QR algorithm for Hermitian matrices*, Linear Algebra Appl. **4** (1971), 137-154. MR **44:**1207
4. P. J. Eberlein and C. P. Huang, *Global convergence of the QR algorithm for unitary matrices with some results for normal matrices*, SIAM J. Numer. Anal. **12** (1975), 97-104. MR **50:**8948
5. J. G. F. Francis, *The QR transformation*. I, II, Comput. J. **4** (1961-1962), 265-271, 332-345. MR **25:**744
6. W. B. Gragg, *The QR algorithm for unitary Hessenberg matrices*, J. Comput. Appl. Math. **16** (1986), 1-8.
7. 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** (1993), 183-198. MR **94e:**65046

W. Hoffmann and B. N. Parlett, *A new proof of global convergence for the tridiagonal QL algorithm*, SIAM J. Numer. Anal. **15** (1978), 929-937. MR **80a:**65075
9. C. P. Huang, *On the convergence of the QR algorithm with origin shifts for normal matrices*, IMA J. Numer. Anal. **1** (1981), 127–133. MR **87d:**65130
10. E. Jiang and Z. Zhang, *A new shift of the QL algorithm for irreducible symmetric tridiagonal matrices*, Linear Algebra Appl. **65** (1985), 261-272. MR **86g:**65082
11. B. N. Parlett, *Singular and invariant matrices under the QR transformation*, Math. Comp. **20** (1966), 611-615. MR **35:**3870
12. B. N. Parlett, *The Rayleigh quotient iteration and some generalizations for nonnormal matrices*, Math. Comp. **28** (1974), 679-693. MR **53:**9615
13. B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980. MR **81j:**65063
14. T.-L. Wang, *Convergence of the QR algorithm with origin shifts for real symmetric tridiagonal and unitary Hessenberg matrices*, Ph.D. thesis, University of Kentucky, Lexington, Kentucky, 1988.
15. T.-L. Wang and W. B. Gragg, *Convergence of the shifted QR algorithm for unitary Hessenberg matrices*, Technical Report NPS-53-90-007, Naval Postgraduate School, Monterey, California, 1990.
16. D.S. Watkins and L. Elsner, *Convergence of algorithms of decomposition type for the eigenvalue problem*, Linear Algebra Appl. **143** (1991), 19-47. MR **91m:**65114
17. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
18. J. H. Wilkinson, *Global convergence of tridiagonal QR algorithm with origin shifts*, Linear Algebra Appl. **1** (1968), 409-420. MR **38:**2938

Department of Mathematical Sciences, National Chengchi University, Taipei, Taiwan, Republic of China
*E-mail address*: `wang@math.nccu.edu.tw`

Department of Mathematics, Naval Postgraduate School, Monterey, California 93943
*E-mail address*: `gragg@nps.navy.mil`