# LINEAR CONVERGENCE
# IN THE SHIFTED QR ALGORITHM

STEVE BATTERSON AND DAVID DAY

ABSTRACT. Global and asymptotic convergence properties for the QR algorithm with Francis double shift are established for certain orthogonal similarity classes of $4 \times 4$ real matrices. It is shown that in each of the classes every unreduced Hessenberg matrix will decouple and that the rate of decoupling is almost always linear. The effect of the EISPACK exceptional shift strategy is shown to be negligible.

## 1. INTRODUCTION AND SUMMARY

A standard approach to approximating the eigenvalues of an $n \times n$ real nonsymmetric matrix $B$ is as follows (see [3, §7.5]). First $B$ is reduced by orthogonal similarity transformations to an upper Hessenberg matrix $H^0$. The QR algorithm with Francis double shift (QRF) is then employed to produce a sequence $\{H^i\}$ of orthogonally similar Hessenberg matrices. While in practice the above procedure usually produces a rapid decoupling, there is little theoretical foundation for its overall speed and reliability. In [8, 9], conditions are identified which assure asymptotic quadratic and cubic decoupling rates. Examples of unreduced Hessenberg matrices which are fixed under QRF are known (see [7, p. 379] for an example) but believed to be unstable under perturbation. We prove that there exist orthogonal similarity classes of matrices for which global convergence is assured but the asymptotic rate of decoupling is only linear.

**Theorem 1.1.** *Suppose $\lambda \in \mathbf{R}$ and $H^0$ is an unreduced Hessenberg matrix in the orthogonal similarity class of*

$$\begin{pmatrix} \lambda & -1 & 1 & 0 \\ 1 & \lambda & 0 & 1 \\ 0 & 0 & \lambda & -1 \\ 0 & 0 & 1 & \lambda \end{pmatrix}.$$

*Then the QRF sequence $\{H^k\}$ will decouple ($H^k_{32} \to 0$). Furthermore, there exists a matrix $H$ such that*

*(1) if there does not exist a diagonal matrix $D$ with $D_{ii} = \pm 1$ such that $H^0 = DHD$, then the rate of decoupling is linear;*

---

(2) *if there exists a diagonal matrix $D$ with $D_{ii} = \pm 1$ such that $H^0 = DHD$, then exact decoupling is immediate $(H^1_{32} = 0)$.*

QRF is used in the HQR2 module in EISPACK [6], the prevailing algorithm for computing all the eigenvalues and all the eigenvectors of a real Hessenberg matrix. In HQR2, if QRF has not produced a decoupling by the tenth iteration, the *exceptional shift* is employed on the eleventh iteration. QRF is then resumed until the twenty-first iteration, at which the exceptional shift is repeated. The remaining iterations are QRF. The motivation of the exceptional shift is as follows. Suppose one begins with a matrix for which QRF has difficulty. The exceptional shift produces an equivalent problem for which QRF will hopefully be more successful. In [1] it is shown that there exist matrices which, in exact arithmetic, will not decouple under HQR2; however, the known examples are extremely unstable. A corollary of the above theorem is that linear convergence persists even after exceptional shifting. Moreover, the linearity is robust under perturbation *within the orthogonal similarity class.*

## 2. Background

In this section we define the iterative algorithm discussed in the paper. These algorithms may be viewed as functions on a space of candidates (e.g., Hessenberg matrices).

Suppose $H^i$ is an unreduced Hessenberg matrix. Let $\chi(\cdot)$ denote the characteristic polynomial of the lower right $2 \times 2$ submatrix of $H^i$. The QRF iteration is given by $H^{i+1} = F(H^i) = Q^t H^i Q$, where $Q$ comes from the QR factorization of $\chi(H^i)$.

We now define the EISPACK exceptional shift for an unreduced Hessenberg matrix $H^i$. Let $\beta = |H^i_{n,n-1}| + |H^i_{n-1,n-2}|$, and define the polynomial $\omega(t) = t^2 - 1.5\beta t + \beta^2$. Then $H^{i+1} = E(H^i) = Q^t(H^i - H^i_{nn}I)Q$, where $Q$ is from the QR factorization of $\omega(H^i - H^i_{nn}I)$. Note that $F$ preserves the orthogonal similarity class and that $E$ maps into the orthogonal similarity class of the shifted matrix $H^i - H^i_{nn}I$.

Fix an orthogonal similarity class and canonial (Schur) form representative, $A$. By the implicit Q theorem, each unreduced Hessenberg matrix in the class is essentially determined by the first column of the orthogonal similarity class from the Schur form. While QRF tracks a sequence of Hessenberg matrices, one can define an analogous iteration which tracks the sequence of corresponding unit vectors. We call this iteration PQRF, and it is discussed in depth in [1, 2].

To define PQRF, let $u^i$ be a unit vector which does not lie in an invariant subspace, and consider the Krylov matrix $K = [u^i | Au^i | \cdots | A^{n-1}u^i]$. Let $q_j$ denote the $j$th column of the QR factorization of $K$. Define $\chi(\cdot)$ to be the characteristic polynomial of the $2 \times 2$ matrix $[q_{n-1} | q_n]^t A [q_{n-1} | q_n]$. The PQRF iteration is given by $u^{i+1} = G(u^i) = \chi(A)u^i / \|\chi(A)u^i\|$.

The relation between PQRF and QRF is formalized by a surjection from PQRF vectors onto the corresponding unreduced Hessenberg matrices in the orthogonal similarity class. The map is defined by $\Gamma(u) = Q^t A Q$, where $Q$ comes from the QR factorization of $K$. It can be shown that $\Gamma$ maps the PQRF sequence for $u$ to the QRF sequence for $\Gamma(u)$, and if the PQRF sequence for $u$ converges to an invariant subspace, then the QRF sequence for $\Gamma(u)$ will decouple. Thus, one can utilize PQRF to analyze the convergence of QRF

within a particular orthogonal similarity class. One advantage of PQRF is the exploitation of Schur forms.

We establish linear convergence of PQRF and QRF in the orthogonal similarity class of

$$A = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The roots of the characteristic polynomial are $\pm i$, and the only proper invariant subspace is the plane spanned by the first two coordinate vectors. Thus for PQRF to succeed, the iteration sequence must converge to this plane.

## 3. THE PQRF ALGORITHM

The PQRF algorithm starts by computing the QR factorization of

$$K = [u|Au|A^2u|A^3u],$$

where $u$ is a unit vector.

We factor $K$, using the Gram-Schmidt algorithm. Let $u = (w, x, y, z)^t$ and $\varepsilon = \sqrt{y^2 + z^2} > 0$. The Gram-Schmidt algorithm applied to $K$ produces $q_1 = u$, $q_2' = Au - \langle u, Au \rangle u$, $q_2 = q_2'/\|q_2'\|$, and so on, where $q_i$ is the $i$th column of $Q$. Let

$$u^\perp = \begin{pmatrix} -x \\ w \\ -z \\ y \end{pmatrix}, \quad v = \begin{pmatrix} y \\ z \\ 0 \\ 0 \end{pmatrix}, \quad v^\perp = \begin{pmatrix} -z \\ y \\ 0 \\ 0 \end{pmatrix}$$

and

$$a = wy + xz, \qquad b = wz - xy.$$

Then for our choice of $A$, we have $Au = u^\perp + v$, $a^2 + b^2 = \varepsilon^2 - \varepsilon^4$, $a = \langle u, Au \rangle$, and

$$\langle q_2', q_2' \rangle = \langle Au, Au \rangle - a^2 = (1 + b)^2 + \varepsilon^4.$$

Using the ordered basis $\mathscr{B} = \{u, u^\perp, v, v^\perp\}$ simplifies the calculations. For any $p \in \mathbf{R}^4$, let $[p]_\mathscr{B}$ denote the coordinates of $p$ with respect to $\mathscr{B}$. For any $p$ and $q$ in $\mathbf{R}^4$,

$$\langle p, q \rangle = \langle [p]_\mathscr{B}, M[q]_\mathscr{B} \rangle,$$

where

$$M = \begin{pmatrix} 1 & 0 & a & -b \\ 0 & 1 & b & a \\ a & b & \varepsilon^2 & 0 \\ -b & a & 0 & \varepsilon^2 \end{pmatrix}.$$

The vector $q_2' = (A - aI)u$ has the coordinates $[q_2']_\mathscr{B} = (-a, 1, 1, 0)^t$ with respect to the new basis. Let $r = \|q_2'\|^2$. We have $r = 1 + 2b + \varepsilon^2 - a^2 = (1+b)^2 + \varepsilon^4$. Now before continuing Gram-Schmidt, note that $A^2u = -u + 2v^\perp$. Since $u \in \text{span}\{u, Au\}$, the vector $q_3$ is determined by orthogonalizing $v^\perp$ against $\text{span}\{u, Au\}$,

$$\langle q_2', v^\perp \rangle = \langle [q_2']_\mathscr{B}, Me_4 \rangle = a(b + 1).$$

By the Gram-Schmidt formula,

$$[q_3']_\mathscr{B} = (b, 0, 0, 1)^t - \frac{a(b+1)}{r}(-a, 1, 1, 0)^t.$$

After multiplying by $M$ and simplifying, we obtain

$$M[q_2']_\mathscr{B} = \begin{pmatrix} 0 \\ 1+b \\ r-1-b \\ a(1+b) \end{pmatrix}, \qquad M[q_3']_\mathscr{B} = \frac{\varepsilon^4}{r} \begin{pmatrix} 0 \\ a \\ -a \\ r+a^2 \end{pmatrix}.$$

Hence, $\|q_3'\|^2 = \langle [q_3']_\mathscr{B}, M[q_3']_\mathscr{B} \rangle = \frac{1}{r}\varepsilon^4(r+a^2)$.

The next step is to compute $q_4$ by orthogonalizing $v$ against $u$, $q_2$, and $q_3$:

$$\langle q_2', v \rangle = b + \varepsilon^2 - a^2, \qquad \langle q_3', v \rangle = -\frac{a}{r}\varepsilon^4.$$

By the Gram-Schmidt formula,

$$[q_4']_\mathscr{B} = (-a, 0, 1, 0)^t - \frac{b+\varepsilon^2-a^2}{r}[q_2']_\mathscr{B} + \frac{a}{r+a^2}[q_3']_\mathscr{B}.$$

After algebraic simplification,

$$M[q_4']_\mathscr{B} = \frac{\varepsilon^4}{r+a^2}(0, -1, 1, 0)^t, \qquad \|q_4'\|^2 = \frac{\varepsilon^4}{r+a^2}.$$

This completes the explicit computation of $Q$.

We now determine explicit formulae for $G(u)$. Let $u'$ denote $G(u)$ prior to normalization,

$$u' = (d-1)u - t(u^\perp + v) + 2v^\perp,$$

where $t$ and $d$ are respectively the trace and determinant of $[q_3|q_4]^t A[q_3|q_4]$. Thus,

$$\|u'\|^2 = 4\varepsilon^2 + (d-1)^2 + t^2(r+a^2) - 2a(d+1)t - 4b(d-1).$$

Let $\tilde{\varepsilon}$ denote the norm of the last two coordinates of the representation of $G(u)$ with respect to the standard basis,

$$\tilde{\varepsilon}^2 = \varepsilon^2((d-1)^2 + t^2)/\|u'\|^2.$$

In the new basis,

$$[A]_\mathscr{B} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

$$[A]_\mathscr{B}[q_2']_\mathscr{B} = -\begin{pmatrix} 1 \\ a \\ a \\ -2 \end{pmatrix}, \qquad [A]_\mathscr{B}[q_3']_\mathscr{B} = \begin{pmatrix} 0 \\ b \\ b-1 \\ 0 \end{pmatrix} + a\frac{b+1}{r}\begin{pmatrix} 1 \\ a \\ a \\ -2 \end{pmatrix},$$

$$[A]_\mathscr{B}[q_4']_\mathscr{B} = \begin{pmatrix} 0 \\ -a \\ -a \\ 1 \end{pmatrix} + \frac{b+\varepsilon^2}{r+a^2}\begin{pmatrix} 1 \\ a \\ a \\ -2 \end{pmatrix} + \frac{a}{r+a^2}\begin{pmatrix} 0 \\ b \\ b-1 \\ 0 \end{pmatrix}.$$

Using the formula

$$\langle p, Aq \rangle = \langle M[p]_{\mathscr{B}}, [A]_{\mathscr{B}}[q]_{\mathscr{B}} \rangle,$$

we have

$$[q_3'|q_4']^t A[q_3'|q_4'] = \begin{pmatrix} \frac{1}{r^2}\varepsilon^4 a(r - 2(1+b)(r+a^2)) & \frac{1}{r}\varepsilon^4(1 - \varepsilon^2 + \frac{1}{r+a^2}a^2) \\ -\frac{1}{r+a^2}\varepsilon^4 & -\frac{1}{(r+a^2)^2}a\varepsilon^4 \end{pmatrix}.$$

Now we obtain closed form expressions for $t$ and $d$:

$$t = \frac{\langle q_3', Aq_3' \rangle}{\|q_3'\|^2} + \frac{\langle q_4', Aq_4' \rangle}{\|q_4'\|^2} = -2a\frac{1+b}{r},$$

$$d = \frac{\langle q_3', Aq_3' \rangle\langle q_4', Aq_4' \rangle - \langle q_3', Aq_4' \rangle\langle q_4', Aq_3' \rangle}{\|q_3'\|^2\|q_4'\|^2} = 1 + \frac{2}{r}(1 + b - r).$$

Substituting these values into the expression for $\|u'\|^2$ and reducing, yields

$$\|u'\|^2 = \frac{4\varepsilon^2}{r^2}((1+b)^2(4(1+b) - 3\varepsilon^2) + r(r+a^2)\varepsilon^2).$$

Similarly,

$$(d-1)^2 + t^2 = \frac{4\varepsilon^2}{r^2}((1+b)^2 + (b^2 - 1 + \varepsilon^4)\varepsilon^2)$$

and

$$\frac{\tilde{\varepsilon}^2}{\varepsilon^2} = \frac{(1+b)^2 + (b^2 - 1 + \varepsilon^4)\varepsilon^2}{(1+b)^2(4(1+b) - 3\varepsilon^2) + r(r+a^2)\varepsilon^2}.$$

Since $r + a^2 = 1 + 2b + \varepsilon^2$, $\tilde{\varepsilon}$ is a function of $\varepsilon$ and $b$.

### 3.1. The convergence of PQRF.
Recall that decoupling under QRF corresponds to convergence of the PQRF sequence to an invariant subspace of $A$. For PQRF to map a vector to this target space, we must have $t = 0$ and $d = 1$.

**Lemma 3.1.** *We have $t = 0$ and $d = 1$ if and only if $\varepsilon = 1/\sqrt{2}$ and $b = -\frac{1}{2}$.*

*Proof.* Since $r$ and $1 + b$ are strictly positive, $t = 0$ if and only if $a = 0$. Furthermore, $d = 1$ and $a = 0$ if and only if $r = 1 + b$ and $a = 0$ if and only if $\varepsilon = 1/\sqrt{2}$ and $b = -\frac{1}{2}$. □

Lemma 3.1 characterizes the set of vectors which hit the target after a single iteration. One can show that these are the unit vectors of the form $(w, x, x, -w)^t$. The following lemma shows that all other vectors converge to the target.

**Lemma 3.2.** *There holds $\tilde{\varepsilon}^2/\varepsilon^2 < \frac{1}{2}$.*

*Proof.* We must show that for $-\varepsilon\sqrt{1 - \varepsilon^2} \le b \le \varepsilon\sqrt{1 - \varepsilon^2}$ and $0 < \varepsilon \le 1$,

$$\frac{(1+b)^2 + (b^2 - 1 + \varepsilon^4)\varepsilon^2}{(1+b)^2(4(1+b) - 3\varepsilon^2) + r(r+a^2)\varepsilon^2} < \frac{1}{2}.$$

The term $4(1 + b) - 3\varepsilon^2 \ge 4(1 - \sqrt{\varepsilon^2 - \varepsilon^4}) - 3\varepsilon^2 \ge 0$ with equality holding if and only if $\varepsilon^2 = \frac{4}{5}$. We have $r(r+a^2) = ((1+b)^2 + \varepsilon^4)(1 + 2b + \varepsilon^2)$. Therefore,

the denominator above is strictly positive, and we can cross-multiply. It suffices to show that the following expression is positive:

$$(1 + b)^2(4(1 + b) - 3\varepsilon^2) + ((1 + b)^2 + \varepsilon^4)(1 + 2b + \varepsilon^2)\varepsilon^2$$
$$- 2(1 + b)^2 - 2(b^2 - 1 + \varepsilon^4)\varepsilon^2$$
$$= \varepsilon^8 + (2b - 1)\varepsilon^6 + (1 + b)^2\varepsilon^4 + 2b(b^2 - 1)\varepsilon^2 + 2(1 + b)^2(2b + 1).$$

From the domain restriction it follows that $-\frac{1}{2} \le b \le \frac{1}{2}$. Note that the $\varepsilon^8$, $\varepsilon^4$, and constant terms are nonnegative.

We consider two cases for the above polynomial.

*Case* 1: $0 \le b \le \frac{1}{2}$. Here,

$$\varepsilon^8 + (2b - 1)\varepsilon^6 + (1 + b)^2\varepsilon^4 + 2b(b^2 - 1)\varepsilon^2 + 2(1 + b)^2(2b + 1)$$
$$\ge \varepsilon^8 - \varepsilon^6 + \varepsilon^4 - \tfrac{3}{4}\varepsilon^2 + 2 > \varepsilon^8 + \varepsilon^4 + \tfrac{1}{8} > 0.$$

*Case* 2: $-\frac{1}{2} \le b \le 0$. Note that if $\varepsilon = 1$, then $b = 0$, which was dealt with in Case 1. Thus, we may assume $0 < \varepsilon < 1$.

Evaluated at $b = -\frac{1}{2}$,

$$\varepsilon^8 + (2b - 1)\varepsilon^6 + (1 + b)^2\varepsilon^4 + 2b(b^2 - 1)\varepsilon^2 + 2(1 + b)^2(2b + 1)$$
$$= \varepsilon^8 - 2\varepsilon^6 + \tfrac{1}{4}\varepsilon^4 + \tfrac{3}{4}\varepsilon^2 > \varepsilon^8 - 2\varepsilon^6 + \varepsilon^4 = \varepsilon^4(\varepsilon^2 - 1)^2 \ge 0.$$

Take the partial derivative with respect to $b$ of the original polynomial:

$$\frac{\partial}{\partial b}(\varepsilon^8 + (2b - 1)\varepsilon^6 + (1 + b)^2\varepsilon^4 + 2b(b^2 - 1)\varepsilon^2 + 2(1 + b)^2(2b - 1))$$
$$= 2\varepsilon^6 + 2(1 + b)\varepsilon^4 - 2(3b^2 - 1)\varepsilon^2 + 4(1 + b)(3b + 2)$$
$$\ge 2\varepsilon^6 + \varepsilon^4 - 2\varepsilon^2 + 1 = 2\varepsilon^6 + (\varepsilon^2 - 1)^2 > 0.$$

The function is positive at any point on the line $b = -\frac{1}{2}$. Increases in $b$ up to 0 produce increases in the function. □

The following theorem is obtained by combining Lemmas 1 and 2 and noting that only the set from Lemma 1 may reach the target prematurely. In the event that the target is not reached on the first iteration, each subsequent iterate will have a value of $\varepsilon < 1/\sqrt{2}$. Thus exact decoupling will never occur after the first iteration.

**Theorem 3.1.** *The PQRF algorithm with*

$$A = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

*converges in zero, one, or an infinite number of iterations to $\varepsilon = 0$. PQRF converges immediately on $\{u: \varepsilon = 0\}$ and in one iteration on $\{u: \varepsilon = 1/\sqrt{2}$ and $b = -\frac{1}{2}\}$. Both sets are one-dimensional curves.*

**Theorem 3.2.** *The PQRF algorithm converges linearly in the case of*

$$A = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

*Proof*. We have

$$\frac{\tilde{\varepsilon}^2}{\varepsilon^2} = \frac{(1+b)^2 + (b^2 - 1 + \varepsilon^4)\varepsilon^2}{(1+b)^2(4(1+b) - 3\varepsilon^2) + r(r+a^2)\varepsilon^2} = \frac{1 + O(\varepsilon^2)}{4(1+b) + O(\varepsilon^2)}.$$

Recall that $b^2 \le a^2 + b^2 = \varepsilon^2 - \varepsilon^4 \le \varepsilon^2$. Thus $\tilde{\varepsilon}^2/\varepsilon^2 = \frac{1}{4} + O(\varepsilon)$, and

$$\tilde{\varepsilon} = \tfrac{1}{2}\varepsilon + O(\varepsilon^2). \quad \square$$

## 3.2. The convergence of QRF.

**Lemma 3.3.** *We have* $(\Gamma(u^i))_{32} = \frac{1}{r}2\varepsilon^2\sqrt{r+a^2}$, *where* $r = (1+b)^2 + \varepsilon^4$.

*Proof*. Suppose $\varepsilon > 0$. Then

$$\langle q_3', Aq_2' \rangle = \langle M[q_3']_{\mathscr{B}}, [A]_{\mathscr{B}}[q_2']_{\mathscr{B}} \rangle = \frac{1}{r}2\varepsilon^4(r+a^2),$$

$$(\Gamma(u^i))_{32} = \left\langle \frac{q_3'}{\|q_3'\|}, A\frac{q_2'}{\|q_2'\|} \right\rangle = \frac{1}{r}2\varepsilon^2\sqrt{r+a^2}. \quad \square$$

**Corollary 3.1.** *One has* $(\Gamma(u^i))_{32} = 0$ *if and only if* $\varepsilon = 0$.

Suppose $H^0$ is an unreduced Hessenberg matrix in the orthogonal similarity class of

$$\begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Produce $u^0$ such that $\Gamma(u^0) = H^0$. Decoupling of the QRF sequence follows from Theorem 3.1.

If the QRF sequence starting from $H^0$ were to decouple in the $(2, 1)$ or $(4, 3)$ positions, this would imply that $H^0$ had a real eigenvalue. Since the roots of the characteristic polynomial are $\pm i$, decoupling must occur in the $(3, 2)$ position.

Suppose $Q$ comes from the QR factorization of the Krylov matrix corresponding to $u^i$. By the previous lemma,

$$H_{32}^i = (\Gamma(u^i))_{32} = \langle q_3, Aq_2 \rangle = \frac{1}{r}2\varepsilon^2\sqrt{r+a^2},$$

where $r = 1 + O(\varepsilon)$. Therefore, $H_{32}^{i+1}/H_{32}^i = \frac{1}{4} + O(\varepsilon)$.

## 4. PROOF OF THEOREM 1.1 AND LINEAR CONVERGENCE OF HQR2

We now complete the proof of Theorem 1.1. The matrix $M$ and the vectors $[q_1]_{\mathscr{B}}, \ldots, [q_4]_{\mathscr{B}}$ are functions of $a$, $b$, and $\varepsilon$ only, and $[A]_{\mathscr{B}}$ is constant. Since $[Q^t AQ]_{ij} = \langle M[q_i]_{\mathscr{B}}, [A]_{\mathscr{B}}[q_j]_{\mathscr{B}} \rangle$, we have that $\Gamma(u) = Q^t AQ$ is a function of $a$, $b$, and $\varepsilon$ only. Since $Q$ is determined up to multiplication on the right by a diagonal matrix $D$ such that $D_{ii} = \pm 1$, the matrices that decouple after one QRF step are of the form $D^t\Gamma(u)D$, where $u$ is any unit vector such that $a = 0$, $b = -\frac{1}{2}$, and $\varepsilon^2 = \frac{1}{2}$. This completes the proof of Theorem 1.1 for $A$.

A Hessenberg matrix orthogonally similar to $A(\lambda) = A + \lambda I$ has the form $H^0 + \lambda I$, where $H^0$ is orthogonally similar to $A$. If $s_1$ and $s_2$ are the roots of

the characteristic polynomial of the lower right $2 \times 2$ submatrix of $H^0$, then $s_1 + \lambda$ and $s_2 + \lambda$ are the shifts for $H^0 + \lambda I$. Shifting by $s_1 + \lambda$ and $s_2 + \lambda$ nullifies the $\lambda I$ for the QR factorization and produces the QRF iterate $H^1 + \lambda I$. Thus, the argument for $A(\lambda)$ follows from that for $A$.

Now consider applying the exceptional shift $E$ to an unreduced Hessenberg matrix $H$ whose Schur form is $A(\lambda)$. The result is a Hessenberg matrix $E(H)$ with Schur form $A(\mu)$ for some $\mu \in \mathbf{R}$. Almost every unreduced Hessenberg matrix in this orthogonal similarity class converges linearly under QRF.

**Corollary 4.1.** *Suppose* $\lambda \in \mathbf{R}$ *and* $H^0$ *is an unreduced Hessenberg matrix in the orthogonal similarity class of*

$$\begin{pmatrix} \lambda & -1 & 1 & 0 \\ 1 & \lambda & 0 & 1 \\ 0 & 0 & \lambda & -1 \\ 0 & 0 & 1 & \lambda \end{pmatrix}.$$

*Then the HQR2 sequence* $\{H^k\}$ *will decouple. Furthermore, there exists a set* $\mathscr{H}$ *of measure* 0 *such that*

(1) *if* $H^0 \notin \mathscr{H}$, *then the rate of decoupling is linear;*
(2) *if* $H^0 \in \mathscr{H}$, *then exact decoupling occurs in a finite number of iterations.*

*Note.* By the HQR2 sequence we mean the algorithm obtained by applying the exceptional shift at steps 11 and 21 with Francis shifts employed for all other iterations.

*Proof of Corollary* 4.1. If decoupling has not occurred by the 21st iteration, then Theorem 1.1 guarantees that it will subsequently occur. We concern ourselves with $\{H | E(H) \text{ or } F(E(H)) \text{ is reduced}\}$, the set of matrices for which the exceptional shift effects an exact decoupling. To show this set is meager, we first develop the exceptional shift analogue of PQRF.

Recall from §2 that $\beta = |H_{n,n-1}| + |H_{n-1,n-2}|$, $\omega(t) = t^2 - 1.5\beta t + \beta^2$, and $E(H) = Q^t(H - H_{nn}I)Q$, where $Q$ is from the QR factorization of $\omega(H^i - H_{nn}^i I)$. As with PQRF, we will define a map of unit vectors. If $u$ does not lie in an invariant subspace of $A(\lambda)$, define

$$D(u) = \frac{\omega(A(\lambda) - H_{nn}I)(u)}{\|\omega(A(\lambda) - H_{nn}I)(u)\|},$$

where $H = \Gamma(u)$.

Details about the relationship between $D$, $E$, and $\Gamma$ are available in [1, 2]. From that analysis it follows that $E(H)$ is reduced if and only if $D(u)$ lies in an invariant subspace. Furthermore, $F(E(H))$ is reduced if and only if $D(u)$ is in the set of Lemma 3.1.

For $D(u)$ to lie in an invariant subspace, the last two coordinates of $u$ must be annihilated by the matrix $\omega(A(\lambda) - H_{44}I)(u)$. The bottom two rows of this matrix have the form

$$\begin{bmatrix} 0 & 0 & \omega\left(\begin{bmatrix} \lambda - H_{44} & -1 \\ 1 & \lambda - H_{44} \end{bmatrix}\right) \end{bmatrix}.$$

The last two coordinates will be annihilated when $\omega(t)$ is the characteristic polynomial of

$$\begin{pmatrix} \lambda - H_{44} & -1 \\ 1 & \lambda - H_{44} \end{pmatrix}.$$

This occurs when $\frac{3}{2}\beta = 2(\lambda - H_{44})$ and $\beta^2 = (\lambda - H_{44})^2 + 1$. It follows that $D(u)$ lies in an invariant subspace only if $\beta^2 = \frac{16}{7}$. From the calculations of the previous sections,

$$\beta = |H_{32}| + |H_{43}| = \frac{2\varepsilon^2(r + a^2)^{3/2} + r^{3/2}}{r(r + a^2)}.$$

Since $\varepsilon$, $r$, and $a$ are each real analytic functions of

$$u = \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix}$$

defined on the hemisphere $z > 0$, $\beta^2$ is an analytic function on this variety. The level set $\beta^2 = \frac{16}{7}$ is a subanalytic set and thus has a triangulation [4]. The function $\beta^2$ is not a constant function, since its values approach 1 as $\varepsilon$ goes to 0. Consequently, the triangulation contains no 3-simplices, and the level set is a measure 0 collection of smaller simplices.

We have shown that $\{u \mid D(u)$ is in an invariant subspace$\}$ has measure 0. This is the set corresponding to $\{H \mid E(H)$ is reduced$\}$. Actually, we are concerned about $\{H \mid EF^{10}(H)$ is reduced$\}$, which corresponds to $\{u \mid DG^{10}(u)$ is in an invariant subspace$\}$. That this set is meager follows from composing $\beta^2$ with $G^{10}$ and applying the previous argument.

If $D(u)$ is in the set of Lemma 3.1, then the sum of the squares of its last two coordinates is $\frac{1}{2}$. Once again, this is the level set of a nonconstant analytic function, and the argument proceeds as in the previous paragraphs.

The set $\mathscr{H}$ consists of the image under $\Gamma$ of a measure-zero set of vectors. Since $\Gamma$ depends only on $b$ and $\varepsilon$, each 2-simplex is mapped to a one-dimensional set which has measure zero in the space of matrices.   $\square$

## 5. NUMERICAL EXAMPLE

We applied the EISPACK subroutine HQR2 to

$$H^0 = \Gamma(u^0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 & 1 \\ 2 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

for $u^0 = (0, 0, 0, 1)^t$, and monitored the $(3, 2)$ element of $H$ until it decoupled. This is displayed in column two of the table. This is the only way for HQR2 to converge when the eigenvalues of $H^0$ are all complex. Column three shows the ratios of consecutive $(3, 2)$ elements. This ratio converges to $\frac{1}{4}$ and persists after the exceptional shifts at iterations 11 and 21. Decoupling occurs when $H_{32}^i$ reaches the double-precision machine epsilon. HQR2 computes the first two eigenvalues of $H^0$ in 34 iterations.

We also computed the corresponding sequence of PQRF vectors. The norm of the last two components of each vector is given in column four of the table. Column five shows the ratios of these norms. As predicted, these ratios converge to $\frac{1}{2}$.

| Iteration | $H_{32}$ | ratio | $\epsilon$ | ratio |
|-----------|----------|-------|-----------|-------|
| 1 | 0.28D+00 | 0.2000000 | 0.45D+00 | 0.4472136 |
| 2 | 0.57D-01 | 0.2032374 | 0.18D+00 | 0.4120817 |
| 3 | 0.13D-01 | 0.2279239 | 0.84D-01 | 0.4572662 |
| 4 | 0.31D-02 | 0.2395767 | 0.40D-01 | 0.4795176 |
| 5 | 0.77D-03 | 0.2449622 | 0.20D-01 | 0.4900170 |
| 6 | 0.19D-03 | 0.2475265 | 0.98D-02 | 0.4950764 |
| 7 | 0.47D-04 | 0.2487748 | 0.49D-02 | 0.4975555 |
| 8 | 0.12D-04 | 0.2493903 | 0.24D-02 | 0.4987821 |
| 9 | 0.29D-05 | 0.2496959 | 0.12D-02 | 0.4993922 |
| 10 | 0.74D-06 | 0.2498481 | 0.61D-03 | 0.4996964 |
| 11 | -0.73D-06 | -0.9987853 | 0.61D-03 | 0.9993923 |
| 12 | -0.18D-06 | 0.2499241 | 0.30D-03 | 0.4998483 |
| 13 | -0.46D-07 | 0.2499620 | 0.15D-03 | 0.4999242 |
| 14 | -0.11D-07 | 0.2499810 | 0.76D-04 | 0.4999621 |
| 15 | -0.29D-08 | 0.2499905 | 0.38D-04 | 0.4999811 |
| 16 | -0.72D-09 | 0.2499952 | 0.19D-04 | 0.4999905 |
| 17 | -0.18D-09 | 0.2499976 | 0.95D-05 | 0.4999953 |
| 18 | -0.45D-10 | 0.2499988 | 0.47D-05 | 0.4999976 |
| 19 | -0.11D-10 | 0.2499993 | 0.24D-05 | 0.4999988 |
| 20 | -0.28D-11 | 0.2499996 | 0.12D-05 | 0.4999994 |
| 21 | 0.28D-11 | -0.9999976 | 0.12D-05 | 0.9999988 |
| 22 | 0.70D-12 | 0.2499994 | 0.59D-06 | 0.4999997 |
| 23 | 0.18D-12 | 0.2499983 | 0.30D-06 | 0.4999999 |
| 24 | 0.44D-13 | 0.2499938 | 0.15D-06 | 0.4999999 |
| 25 | 0.11D-13 | 0.2499755 | 0.74D-07 | 0.5000000 |
| 26 | 0.27D-14 | 0.2499021 | 0.37D-07 | 0.5000000 |
| 27 | 0.68D-15 | 0.2496088 | 0.19D-07 | 0.5000000 |
| 28 | 0.17D-15 | 0.2484401 | 0.93D-08 | 0.5000000 |
| 29 | 0.41D-16 | 0.2438378 | 0.46D-08 | 0.5000000 |
| 30 | 0.94D-17 | 0.2265230 | 0.23D-08 | 0.5000000 |
| 31 | 0.16D-17 | 0.1715232 | 0.12D-08 | 0.5000000 |
| 32 | 0.11D-18 | 0.682D-01 | 0.58D-09 | 0.4999999 |
| 33 | 0.68D-21 | 0.623D-02 | 0.29D-09 | 0.4999999 |
| 34 | 0.27D-25 | 0.397D-04 | 0.14D-09 | 0.4999998 |

## BIBLIOGRAPHY

1. S. Batterson, *Convergence of the shifted QR algorithm on* $3 \times 3$ *normal matrices*, Numer. Math. **58** (1990), 341–352.
2. ____, *The dynamics of eigenvalue computation*, preprint.
3. G. H. Golub and C. F. Van Loan, *Matrix computations*, 2nd ed., Johns Hopkins Univ. Press, Baltimore, 1989.

4. H. Hironaka, *Triangulations of algebraic sets*, Proc. Sympos. Pure Math., vol. 29, Amer. Math. Soc., Providence, RI, 1975, pp. 165–185.

5. R. S. Martin, G. Peters, and J. H. Wilkinson, *The QR algorithm for real Hessenberg matrices*, Numer. Math. **14** (1970), 219–231.

6. B. T. Smith, J. M. Boyle, J. J. Dongarra, B. S. Garbow, Y. Ikebe, V. C. Klema, and C. B. Moler, *Matrix eigensystem routines—Eispack guide*, Lecture Notes in Comput. Sci., vol. 6, Springer-Verlag, 1974.

7. G. W. Stewart, *Introduction to matrix computations*, Academic Press, New York, 1973.

8. R. van de Geijn, *Deferred shifting schemes for parallel QR methods*, preprint.

9. D. S. Watkins and L. Elsner, *Convergence of algorithms of decomposition type for the eigenvalue problem*, Linear Algebra Appl. **143** (1991), 19–47.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT BERKELEY, BERKELEY, CALIFORNIA 94720

*Current address*, S. Batterson: Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia 30322

*E-mail address*, S. Batterson: sb@mathes.emory.edu

*E-mail address*, D. Day: dday@math.berkeley.edu