Global Convergence of the Basic QR Algorithm on Hessenberg Matrices*

By Beresford Parlett

0. Introduction. The QR algorithm was developed by Francis (1960) to find the eigenvalues (or roots) of real or complex matrices. We shall consider it here in the context of exact arithmetic.

Sufficient conditions for convergence, listed in order of increasing generality have been given by Francis [1], Kublanovskaja [3], Parlett [4], and Wilkinson [8]. It seems that necessary and sufficient conditions would be very complicated for a general matrix.

One of the many merits of Francis' paper was the observation that the Hessenberg form ($a_{ij} = 0$, $i > j + 1$) is invariant under the QR transformation and the algorithm is usually applied to Hessenberg matrices which are unreduced, that is $a_{ij} \neq 0$, $i = j + 1$. The properties of this form combine with those of the algorithm in such a way that a complete convergence theory can be stated quite simply. The aim is to produce a sequence of unitarily similar matrices whose limit is upper triangular.

Elementwise convergence to a particular triangular matrix is not necessary for determining eigenvalues; block triangular form with 1×1 and 2×2 blocks on the diagonal is sufficient.

Definition. A sequence $\{H^{(s)} = (h_{(ij)}^{(s)}),\ s = 1, 2, \cdots\}$ of $n \times n$ Hessenberg matrices is said to *"converge"* whenever $h_{j+1,j}^{(s)} h_{j,j-1}^{(s)} \to 0$, for each $j = 2, \cdots, n - 1$.

THEOREM 1. *The basic QR algorithm applied to an unreduced Hessenberg matrix H produces a sequence of Hessenberg matrices which "converges" if, and only if, among each set of H's eigenvalues with equal magnitude, there are at most two of even and two of odd multiplicity.*

This is a special case, tailored to computer programs, of the main theorem. In general let $\omega_1 > \omega_2 > \cdots > \omega_r > 0$ be the distinct nonzero magnitudes occurring among the roots of H. Of the roots of magnitude ω_i let $p(i)$ have even multiplicities

$$m_1^{\ i} \geqq m_2^{\ i} \geqq \cdots \geqq m_{p(i)}^{i} > m_{p(i)+1}^{i} \equiv 0\,,$$

and let $q(i)$ have odd multiplicities,

$$n_1^{\ i} \geqq n_2^{\ i} \geqq \cdots \geqq n_{q(i)}^{i} > n_{q(i)+1}^{i} \equiv 0\,.$$

MAIN THEOREM. *Let $H^{(s)}$ be the sth term of the basic QR sequence derived from an unreduced Hessenberg matrix H. If zero is a root of multiplicity m, then the last m rows of $H^{(s)}$ are null for $s > m$ and they and the last m columns are discarded from $H^{(s)}$. As $s \to \infty$, $H^{(s)}$ becomes block triangular, $(H_{ij}^{(s)})$, and the spectrum of $H_{ii}^{(s)}$ con-*

Received May 24, 1966. Revised January 23, 1967 and February 19, 1968.

* This work was begun while the author was a summer visitor at the Argonne National Laboratory, Argonne, Illinois. It was completed under Office of Naval Research contract NONR 3656 (23) with the Computing Center of the University of California, Berkeley.

verges to the set of eigenvalues with magnitude ω_i. *Further* $H_{ii}^{(s)}$ *itself tends to block triangular form. There emerge* $m_j^i - m_{j+1}^i$ *unreduced diagonal blocks of order* j $[j = 1, \cdots, p(i)]$, *the union of whose spectra converges to the eigenvalues of even multiplicity. Similarly, there emerge* $n_j^i - n_{j+1}^i$ *unreduced diagonal blocks of order* j $[j = 1, \cdots, q(i)]$, *the union of whose spectra converges to the eigenvalues of odd multiplicity.*

Theorem 1 follows because if any $p(i)$ or $q(i)$ exceeds 2, then there will be a principal submatrix of order greater than 2, none of whose subdiagonal elements converge to zero. Conversely, if $p(i) \leqq 2$, $q(i) \leqq 2$ for all i then $H^{(s)}$ reduces to block triangular form with 1×1 and 2×2 diagonal blocks.

The position of the unreduced blocks depends on how the m_j^i, n_k^i interlace when ordered monotonically.

The rate of convergence is very slow (like s^{-1}). This is not a disaster, because in Francis' program the basic algorithm is used only until at least one of the roots of the bottom 2×2 submatrix "settles down" to 1 binary bit (that is, to within 50%). Then the extended algorithm is applied to hasten convergence. Theorem 1 ensures that when the hypotheses hold, this test will be passed. Of more importance, the test will not be passed only if there are too many roots of equal modulus. This modulus is easily calculated from the determinant of the unreduced submatrix. The only problem is to decide early when the test will not be passed.

Note that convergence is certain when the roots are real.

A preliminary report of these results appeared as *Necessary and Sufficient Conditions for Convergence of the QR Algorithm on Hessenberg Matrices*, Proc. of the ACM National Meeting, Los Angeles, Calif., 1966, Thompson, Washington, D. C., 1966, pp. 13–16.

1. The Algorithm, its Essential Convergence and Known Properties. We shall assume that the reader has some familiarity with the QR algorithm of J. G. F. Francis. For expositions of it, see [1], [5], or [9, Chapter 8]. Here we shall give a brief outline of the algorithm and those convergence properties which are already known.

From any given square matrix A_1 the algorithm generates a sequence $\{A_s\}$ of matrices unitarily congruent to A_1. Under certain mild conditions, it is known that, as $s \to \infty$, A_s tends to a form which is essentially triangular; namely a block triangular matrix whose diagonal blocks have orders one or two. When A_1, and hence each A_s, is real, complex eigenvalues will be found from real two-rowed principal submatrices.

The Factorization. Any square matrix A can be expressed as the product, QR, of a unitary matrix Q and a right triangular matrix R. When A is real, Q can be taken orthogonal. It is customary to normalize the factorization by requiring that the diagonal of R have nonnegative elements. When A is nonsingular Q and R are unique and will be denoted by $Q(A)$ (or Q_A) and $R(A)$ (or R_A) respectively.

Without the normalization, Q and R are unique only to within a diagonal unitary factor. Thus for any diagonal unitary matrix D we have $A = (Q_A\bar{D})(DR_A) = QR$.

The Basic Algorithm. Given a nonsingular matrix A_1, the algorithm is given by the rule

$$\text{(1.1)} \qquad \text{for } s = 1, 2, \cdots \begin{cases} \text{factor } A_s \text{ into } Q_s R_s\,, \\ \text{form } R_s Q_s \text{ and call it } A_{s+1}\,. \end{cases}$$

It follows from (1.1) that

$$\text{(1.2)} \qquad \begin{aligned} A_{s+1} &= P_s^* A_1 P_s, \ (M^* \text{ is the conjugate transpose of } M)\,, \\ P_s &= Q(A_1^s) = Q_1 \cdots Q_s\,, \end{aligned}$$

and so the convergence of A_s as $s \to \infty$ depends on the unitary factor of A_1^s.

In practice we are interested in a less stringent property which Wilkinson calls *essential* convergence, namely the convergence of A_s to within a diagonal unitary congruence. Thus if there is a sequence of diagonal unitary matrices D_s such that $P_s\bar{D}_s$ converges then we say that P_s, A_s, and the algorithm all converge essentially. We shall extend the usage further in the real case by allowing D_s to be orthogonal and block diagonal with blocks of order 1 and 2.

Convergence. The fundamental result given in [1], [4], [8] is that when the eigenvalues of A_1 have distinct moduli, then $\{A_s\}$ converges essentially to upper triangular form. Wilkinson showed that, under a certain assumption, if there is only one eigenvalue (of any multiplicity) of a given modulus, then the algorithm converges essentially.

Hessenberg Form. It is a useful fact that any matrix may be put into upper Hessenberg form H ($h_{ij} = 0$, $i > j + 1$) by a finite sequence of similarity transformations [9, Chapter 5]. Indeed this form can be achieved in several ways (including orthogonal congruences). It was one of the many merits of Francis' article that it recognised the invariance of the Hessenberg form under the QR transformation.

The importance of the reduction of the given matrix to this form is not just the arithmetic economy in transforming Hessenberg matrices as against full ones, but the clever devices which Francis was able to use in calculating the transformed matrix. Moreover, we shall show in later sections that the QR algorithm has strong convergence properties when applied to Hessenberg matrices.

Definition. An $n \times n$ Hessenberg matrix H is *unreduced* if $h_{i,i-1} \neq 0$, $i = 2, \cdots, n$.

We recall that a matrix is called *derogatory* if the eigenspace of any eigenvalue has dimension greater than 1.

LEMMA. *An unreduced Hessenberg matrix is not derogatory.*

Proof. The minor of the $(1, n)$ element of $H - zI$ is nonzero and independent of z. Thus the null space of $H - zI$ has dimension ≤ 1 for all z. Q.E.D.

Is it possible that the basic algorithm might fail to resolve an unreduced Hessenberg submatrix of order greater than 2? The answer is yes, but we shall see that this can *only* happen in cases which are easily remedied by the extension of the basic algorithm introduced by Francis.

2. A Particular Case. The permutation matrix

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

has spectrum $\{1, \omega, \omega^2\}$, $\omega = \exp(2\pi i/3)$. Like any unitary matrix it is invariant under the QR transformation. Moreover it can be shown that no unreduced 3×3 Hessenberg matrix with 3 equimodular simple eigenvalues yields a convergent QR sequence. Consequently it is surprising to discover that the spectrum alone does not determine convergence but that the multiplicities of the eigenvalues do play a role.

The proof of the main theorem is somewhat involved and in this section we analyze a 4×4 example which exhibits the crucial aspects of the general case. Consider any real unreduced 4×4 Hessenberg matrix H_1 with the same spectrum as P above. The Jordan form of $H_1 = Y^{-1}JY$ is given by

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & \omega & 0 \\ 0 & 0 & 0 & \omega^2 \end{bmatrix}.$$

Since $H_{s+1} = P_s^* H_1 P_s$, $P_s = Q(H_1^s)$, $s = 1, 2, \cdots$ we must investigate the unitary factor of H_1^s. By Theorem 2, Section 3, Y has a triangular decomposition $Y = L_Y U_Y$ with

$$L_Y = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ l_{31} & l_{32} & 1 & \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix}.$$

See [9, Chapter 4] for a discussion of triangular (LU) factorization, L unit lower, U upper triangular.

Thus

$$H_1^s = Y^{-1}J^sY = Y^{-1}J^sL_YU_Y.$$

Following an idea of Wilkinson, we wish to manipulate the factors of H_1^s into the form (unitary) (upper triangular). Although J^sL_Y is unit lower triangular it is unbounded in s and this obstructs the analysis. However, a suitable permutation of the rows yields a matrix with an LU decomposition with L bounded as $s \to \infty$. Our problem is to find a permutation matrix B, independent of s, such that $BJ^sL_Y = L_sU_s$ with L_s bounded as $s \to \infty$.

On writing out J^sL_Y in extenso, we see that row 2 should become row 1 of BJ^sL_Y. On checking all 2-rowed minors in the first two columns, we find that row 1 should become row 4. Let $B = (e_4, e_1, e_2, e_3)$ where $I = (e_1, e_2, e_3, e_4)$. Then

$$BJ^sL_Y = \begin{bmatrix} s + l_{21} & 1 & 0 & 0 \\ \omega^s l_{31} & \omega^s l_{32} & \omega^s & 0 \\ \omega^{2s} l_{41} & \omega^{2s} l_{42} & \omega^{2s} l_{43} & \omega^{2s} \\ 1 & 0 & 0 & 0 \end{bmatrix} = L_sU_s,$$

where the order of magnitude of the elements of L_s is given below:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ s^{-1} & 1 & 0 & 0 \\ s^{-1} & 1 & 1 & 0 \\ s^{-1} & s^{-1} & s^{-1} & 1 \end{bmatrix}.$$

The (3, 2) element is

$$\omega^s\left(\frac{l_{42}(s+l_{21})-l_{41}}{l_{32}(s+l_{21})-l_{31}}\right)\sim\left(\frac{l_{42}}{l_{32}}\right)\omega^s, \quad \text{as } s\to\infty .$$

In a previous paper [7] we showed that l_{42}, l_{32} and their analogues in the general case cannot vanish because H_1 is an unreduced Hessenberg matrix.

The major problem in the general case is to determine the matrix B.

The surprising point here is that the two rows corresponding to the double eigenvalue 1 have been separated as far as possible.

Returning to the factorization of H_1^s we find

$$\begin{aligned} H_1^s &= Y^{-1}B^{-1}(BJ^sL_Y)U_Y \\ &= Y^{-1}B^{-1}L_sU_sU_Y \\ &= \bar{Q}\bar{R}L_sU_sU_Y, \qquad \text{defining } \bar{Q}\bar{R} = Y^{-1}B^{-1} \\ &= (\bar{Q}Q_s)(R_sU_sU_Y), \qquad \text{defining } Q_sR_s = \bar{R}L_s . \end{aligned}$$

Since $\bar{R}$ is nonsingular upper triangular the strictly lower triangular part of Q_s is determined by L_s (this is proved in Lemma 3.1). Thus as $s\to\infty$

$$Q_s\sim\begin{bmatrix}1&0&0&0\\0&x&x&0\\0&x&x&0\\0&0&0&1\end{bmatrix},$$

where the x represent elements which do not converge. By the uniqueness of the QR factorization $P_s = Q(H_1^s) = \bar{Q}Q_s$, essentially and $H_{s+1} = Q_s^*\bar{Q}^*H_1\bar{Q}Q_s$.

Thus, as $s\to\infty$,

$$Q_sH_{s+1}Q_s^*\to(RBY)(Y^{-1}JY)(Y^{-1}B^{-1}R^{-1}) = R(BJB^{-1})R^{-1},$$

and

$$BJB^{-1} = \begin{bmatrix}1&0&0&1\\0&\omega&0&0\\0&0&\omega^2&0\\0&0&0&1\end{bmatrix}.$$

3. The General Case. Any square matrix $A = (a_{ij})$ (with real or complex elements) may be taken into upper Hessenberg form H by a similarity transformation (see Section 2). Some subdiagonal elements $h_{j+1,j}$ may be zero. By partitioning with respect to these elements we may write H as a block upper triangular matrix (H_{ij}) where each H_{ii} is a Hessenberg matrix with *nonzero* subdiagonal elements. We will call such Hessenberg matrices *unreduced.* Typically, we might have

$$H = \begin{pmatrix}H_{11}&H_{12}&H_{13}\\0&H_{22}&H_{23}\\0&0&H_{33}\end{pmatrix}. \tag{3.1}$$

The QR transformation acts independently on each H_{ii}. Indeed if $H_{ii} = Q_iR_i$ and $H = QR$ then Q is the direct sum of the Q_i and the diagonal blocks of RQ are just R_iQ_i. Thus it suffices to restrict attention to *unreduced* Hessenberg matrices.

Suppose such a matrix is singular. We showed in an earlier paper [6] that each single QR transformation annihilates the bottom row. The algorithm then proceeds on the submatrix obtained by omitting the last row and column. After a finite number of steps all the zero eigenvalues will have been found. Thus it remains to consider *nonsingular unreduced Hessenberg* matrices.

Nonsingular unreduced Hessenberg matrices. Let H be such a matrix. Then H is nonderogatory and so has only one Jordan block to each eigenvalue.

Subsequent analysis will be simplified if we write the Jordan submatrix corresponding to a nonlinear elementary divisor in the slightly unconventional form, illustrated below for a cubic divisor:

$$\begin{pmatrix} \lambda & 0 & 0 \\ |\lambda| & \lambda & 0 \\ 0 & |\lambda| & \lambda \end{pmatrix} \equiv \begin{pmatrix} \theta & 0 & 0 \\ 1 & \theta & 0 \\ 0 & 1 & \theta \end{pmatrix} \omega , \quad |\lambda| = \omega > 0 .$$

Geometrically, this amounts to a nonstandard selection of principal vectors to span the cyclic subspace associated with λ. There is no loss of generality. Let

$$\omega_1 > \omega_2 > \cdots > \omega_r > 0 \tag{3.3}$$

be the distinct magnitudes occurring among the roots of H. We shall write the Jordan canonical form of H as JW $(=WJ)$ where

$$W = \omega_1 I_1 \oplus \cdots \oplus \omega_r I_r , \qquad J = J_{11} \oplus \cdots \oplus J_{rr} , \tag{3.4}$$

and each J_{ii} is a direct sum of the Jordan blocks of the arguments of the eigenvalues of magnitude ω_i.

We wish to study the sequence $\{H_s, \ s = 1, 2, \cdots\}$ obtained by applying the basic QR algorithm to $H_1 = H$. This depends (see Section 1) on the unitary factor P_s of H^s and we shall follow Wilkinson's idea of exhibiting P_s explicitly by manipulating the canonical factorization of H^s. We begin with an essential result proved in [7].

THEOREM 2. *Let the Jordan decomposition of H be $H = Y^{-1}JWY$. Then Y permits a triangular decomposition without interchanges, $Y = L_Y U_Y$, L_Y unit lower triangular, U_Y upper triangular.*

Our modification of the Jordan form does not invalidate this result since it corresponds to a premultiplication of Y by a positive diagonal matrix. So

$$\begin{aligned} H^s &= XJ^sW^sY , && \text{where } X \equiv Y^{-1} , \\ &= XJ^sW^sL_YU_Y , && \text{by Theorem 2} , \\ &= XMW^sU_Y , && \text{where } M(s) = J^sW^sL_YW^{-s} . \end{aligned} \tag{3.5}$$

One of the principal results of the next three sections is that there is a fixed permutation matrix B such that as $s \to \infty$ BM permits a triangular decomposition $BM = L_sU_s$ with L_s bounded. Then, for large enough s,

$$\begin{aligned} H^s &= XB^*L_sU_sW^sU_Y , \\ &= \bar{Q}\bar{R}L_sU_sW^sU_Y, \text{ where } \bar{Q}\bar{R} \text{ is the } Q - R \text{ factorization of } XB^* , \end{aligned} \tag{3.6}$$

$$= (\bar{Q}\tilde{Q}_s)(\tilde{R}_sU_sW^sU_Y), \text{ where } \tilde{Q}_s\tilde{R}_s \text{ is the } Q - R \text{ factorization of } \bar{R}L_s . \tag{3.7}$$

This is a unitary-triangular factorization of H^s. Hence, (see Section 1) $P_s = \bar{Q}\tilde{Q}_s$, essentially, and, to within a similarity by a diagonal unitary matrix,

$$H_{s+1} = P_s^*HP_s = \tilde{Q}_s^*(\bar{Q}^*H\bar{Q})\tilde{Q}_s = \tilde{Q}_s^*(\bar{R}BJWB^{-1}\bar{R}^{-1})\tilde{Q}_s . \tag{3.8}$$

The usefulness of this analysis depends on the following observation. Let the matrices $L_s = (L_{ij})$ and $\tilde{Q}_s = (Q_{ij})$ of (3.6) and (3.7) be partitioned conformably in any manner which makes the diagonal blocks square.

LEMMA 3.1. *As* $s \to \infty$, $Q_{ij} \to 0$ *for all* i, j $(i \neq j)$ *if, and only if,* $L_{ij} \to 0$ *for all* i, j $(i > j)$.

Proof. L_s is unit lower triangular and bounded as $s \to \infty$. Hence L_s^{-1} has the same properties. Moreover, since $\tilde{Q}_s$ is unitary and $\tilde{Q}_s\tilde{R}_s = \bar{R}L_s$, it follows that $\det \tilde{R}_s = \det \bar{R} = |\det X| > 0$. For any unitarily invariant norm, we then have

$$\|\tilde{R}_s\| = \|\tilde{Q}_s\tilde{R}_s\| \leqq \|\bar{R}\| \, \|L_s\| \leqq \gamma\|\bar{R}\| ,$$

$$\|\tilde{R}_s^{-1}\| = \|\tilde{R}_s^{-1}Q_s^*\| \leqq \|L_s^{-1}\| \, \|\bar{R}^{-1}\| \leqq \gamma\|\bar{R}^{-1}\| ,$$

where γ is a bound on $||L_s||$ and $||L_s^{-1}||$. Now partitioning all the matrices conformably with L_s and $\tilde{Q}_s$ we have, for $i > j$,

$$Q_{ij} = \sum_{\mu \geqq i} \sum_{\nu \leqq j} \bar{R}_{i\mu}L_{\mu\nu}(\tilde{R}_s^{-1})_{\nu j} , \qquad L_{ij} = \sum_{\mu \geqq i} \sum_{\nu \leqq j} (\bar{R}^{-1})_{i\mu}Q_{\mu\nu}\tilde{R}_{\nu j} .$$

Hence, as $s \to \infty$,

$$Q_{ij} \to 0 , \quad \text{all } i, j \ (i > j) \text{ if, and only if, } L_{ij} \to 0 , \quad \text{all } i, j \ (i > j) .$$

Equating corresponding blocks in the equations $\tilde{Q}_s^*\tilde{Q}_s = \tilde{Q}_s\tilde{Q}_s^* = I$ yields the lemma. □

To see that $\tilde{Q}_s$, and therefore L_s, determines the block triangular structure to which H_{s+1} tends we consider (3.8) and use the corollary of Lemma 6.2, which states that

$$BJWB^{-1} \textit{ is upper triangular} .$$

These results will be proved in the following sections. Thus the matrix $(\bar{R}BJWB^{-1}\bar{R}^{-1})$ of (3.8) is upper triangular and the block triangular form of H_s is completely determined by $\tilde{Q}_s$, and therefore by L_s, as $s \to \infty$. The purpose of this section was to show that it suffices to consider the matrix L_s, the bounded lower triangular factor of some permutation of $J^sW^sL_YW^{-s}$.

4. Eigenvalues of Different Magnitudes. The matrix $M(s) = J^s(W^sL_YW^{-s})$ is a product of two lower triangular matrices and, in general, J^s is unbounded as $s \to \infty$. Now partition M into blocks, one block for all eigenvalues with a common magnitude. Then

$$L_Y = (L_{ij}) , \qquad (i, j = 1, \cdots, r) , \tag{4.1}$$

where the partition conforms with (3.4). Then for $i > j$, as $s \to \infty$, by (3.3),

$$M_{ij} = J_{ii}^sL_{ij}(\omega_i/\omega_j)^s \to 0 , \tag{4.2}$$

since $s^m(\omega_i/\omega_j)^s \to 0$ for any fixed m.

Thus M tends to block diagonal form. However, each diagonal block $M_{ii}(s) = J_{ii}^sL_{ii}$ $(i = 1, \cdots, r)$ is unbounded as $s \to \infty$ except in the trivial case when J_{ii}

is diagonal. We seek fixed permutation matrices B_i such that for $i = 1, \cdots, r$,

(4.3) $$B_i M_{ii} = \tilde{L}_i \tilde{U}_i, \quad \tilde{L}_i(s) \text{ bounded as } s \to \infty .$$

We define

(4.4) $$B = B_1 \oplus \cdots \oplus B_r$$

and then the matrix L_s of Section 3 tends to $\tilde{L}_1 \oplus \cdots \oplus \tilde{L}_r$ as $s \to \infty$.

It remains to study $J_{ii}^s L_{ii}$ and $\tilde{L}_i$. It happens that L_{ii} depends on *all* the eigenvalues with magnitude greater than or equal to ω_i. See [7] or Lemma 6.3. Thus the main theorem cannot be established by considering only matrices H all of whose eigenvalues have equal magnitude.

In Section 6 we examine in detail a typical diagonal block $M_{ii} = J_{ii}^s L_{ii}$ and there we drop the index i.

5. Triangular Factorization. If A is a matrix and $\alpha = (\alpha_1, \cdots, \alpha_j), \beta = (\beta_1, \cdots, \beta_j)$ are multi-indices let A_β^α or $A(\alpha; \beta)$ denote the submatrix of A lying in rows $\alpha_1, \cdots, \alpha_j$ and columns $\beta_1, \cdots, \beta_j$. Let det $[A_\beta^\alpha]$ be its minor. We shall need

(5.1) A complex $n \times n$ matrix A permits a triangular factorization $A = LU$ if and only if the first $n - 1$ leading principal minors do not vanish. See [2, p. 11].

(5.2) If $L = (l_{ij})$ in (5.1), then

$$l_{ij} = \det [A(1, \cdots, j-1, i; 1, \cdots, j)] / \det [A(1, \cdots, j; 1, \cdots, j)] .$$

See [2, p. 11].

(5.3) If det $[A] \neq 0$ there is a permutation matrix Π such that $\Pi A = LU$ and the elements of L are bounded by 1 in magnitude. See [9, Chapter 1].

The arguments which yield (5.3) also show

(5.4) If the elements of A are polynomials in one variable and det $[A] \not\equiv 0$ then there is a permutation matrix B such that $BA = LU$ and the elements of L are rational functions with the degree of the numerator bounded by the degree of the denominator. Thus the elements of L are bounded in a neighborhood of infinity of the variable.

Formula (5.2) shows that the permuted matrices ΠA, BA in (5.3), (5.4) are such that for each $j = 1, \cdots, n-1$,

(5.5) det $[BA(1, \cdots, j; 1, \cdots, j)]$ is maximal among all

$$\det [BA(1, \cdots, j-1, i; 1, \cdots, j)] , \qquad i \geqq j .$$

In (5.3) the ordering is by absolute value, in (5.4) it is by (polynomial) degree.

In the case of (5.4) let s be the variable. If the degrees of the minors of A are independent of s, then B may also be chosen independent of s. We shall study the matrix $M(s)$ of (6.1). Nonzero minors of J^s are products of nonzero minors of the $J_i^s = (\theta_i I_i + N_i)^s$, $|\theta_i| = 1$. Thus the coefficients of the powers of s do depend on the θ_i^s. So to prove that B is independent of s we must prove that the degrees of certain minors of M are constant.

6. Eigenvalues of Equal Magnitude. In Section 4 we reduced the problem of finding those subdiagonal elements of H_s which tend to zero as $s \to \infty$ to the study of the bounded triangular factorization of matrices of the form

$$M = M(s) = J^s L \,, \tag{6.1}$$

where

$$\begin{aligned} J &= J_1 \oplus \cdots \oplus J_t \,, \\ J_i &= \theta_i I_i + N_i \,, \quad |\theta_i| = 1 \,, \quad I_i = (e_1, \cdots, e_{m_i}) \,, \\ N_i &= (e_2, \cdots, e_{m_i}, 0) \,, \qquad (i = 1, \cdots, t) \,. \end{aligned} \tag{6.2}$$

L = the principal submatrix of L_Y (see Theorem 2) corresponding to eigenvalues with some common modulus ω and arguments (or signa) $\theta_1, \cdots, \theta_t$.

For any matrix A and natural number ν let A_ν, $A_{\bar\nu}$ denote the matrices formed from the first ν columns of A and last ν columns of A respectively. Let A^ν, $A^{\bar\nu}$ be similarly defined for the rows.

We order the minors of M and J^s by their degrees as polynomials in s. Let $\sum x_\nu$ denote summation of the x_ν, $\nu = 1, \cdots, t$.

We begin with a crucial but simple result.

LEMMA 6.1. *The maximum $\nu \times \nu$ minor of J_i^s is* $\det [(J_i^s)_\nu^{\bar\nu}]$, *is unique, and has degree $\nu(m_i - \nu)$.*

Proof.

$$J_i^s = \sum_{\sigma=0}^{m_i} \binom{s}{\sigma} \theta_i^{s-\sigma} N_i^\sigma \,, \quad \text{where} \binom{s}{\sigma} \text{ is a binomial coefficient} \,.$$

Consider the minor in rows $\alpha = (\alpha_1, \cdots, \alpha_\nu)$ and columns $\beta = (\beta_1, \cdots, \beta_\nu)$. Replace

$$\binom{s}{\sigma}$$

by $s^\sigma/\sigma!$ and observe that the determinant is homogeneous in s of degree $|\alpha| - |\beta|$, where $|\alpha| = \alpha_1 + \cdots + \alpha_\nu$, and in θ_i of degree $s - |\alpha| + |\beta|$. The coefficient of these terms is of Vandermonde type and is nonzero.

The degree is maximized only when $\alpha = (m_i - \nu + 1, \cdots, m_i)$, $\beta = (1, \cdots, \nu)$, and the coefficient of $s^{\nu(m_i - \nu)}$ is given by

$$\xi_i(\nu) = \theta_i^{s-\nu(m_i-\nu)} (m_i - \nu + 1)!!(\nu - 1)!!/(m_i - 1)!! \,, \tag{6.3}$$

where $k!! = k!(k-1)! \cdots 2!$. □

Let $Z_j(k) = \{\alpha = (\alpha_1, \cdots, \alpha_j) \colon 1 \leqq \alpha_1 < \cdots < \alpha_j \leqq k,\ \alpha_i \text{ integers}\}$ and $Z_0(k) = \varnothing$, the empty set. The order of M and J is $|m| = \sum m_i$. The indices of any set of j rows of M could be designated by a multi-index in $Z_j(|m|)$. However, J is block diagonal and it is more convenient to designate row indices by multi-multi-indices as follows. Let $\beta = (\beta_1, \cdots, \beta_t)$, $\beta_i \in Z_{\mu_i}(m_i)$ and $|\beta| = \sum \mu_i$. Then M^β is the submatrix of M obtained by taking the rows of M with indices

$$\beta_{jk} + \sum_{\nu=1}^{j-1} m_\nu \,, \quad k = 1, \cdots, \mu_j, j = 1, \cdots, t \,. \tag{6.4}$$

If $\mu_j = 0$, then β_j does not exist and no rows of M^β involve the submatrix J_j.

Example. Let $m = (4, 3, 3)$, $\beta = ((3, 4), (2, 3), (3))$,

$$|\beta| = \mu_1 + \mu_2 + \mu_3 = 2 + 2 + 1 \,, \qquad M = J^s L \,,$$

$M_5{}^{\beta}$ is as indicated below.

$$
\begin{array}{c}
\\ \\ \beta_1 \begin{cases} \to \\ \to \end{cases} \\ \\ \beta_2 \begin{cases} \to \\ \to \end{cases} \\ \\ \\ \beta_3 \to
\end{array}
\left[\begin{array}{cccc|ccc|ccc}
x & & & & & & & & & \\
x & x & & & & & & & & \\
x & x & x & & & 0 & & & 0 & \\
x & x & x & x & & & & & & \\
\hline
 & & & & x & & & & & \\
 & 0 & & & x & x & & & 0 & \\
 & & & & x & x & x & & & \\
\hline
 & & & & & & & x & & \\
 & & & & & 0 & & x & x & \\
 & & & & & & & x & x & x
\end{array}\right]
\left[\begin{array}{ccccc}
x & & & & \\
x & x & & 0 & \\
x & x & x & & \\
x & x & x & x & \\
x & x & x & x & x \\
x & x & x & x & x \\
x & x & x & x & x \\
x & x & x & x & x \\
x & x & x & x & x \\
x & x & x & x & x
\end{array}\right].
$$

Note that μ_i is the number of rows of $J_i{}^s$ involved in the minor.

We recall that if ν is a natural number, A_ν designates the first ν columns of A. If p is a polynomial, deg p denotes its degree.

LEMMA 6.2 *Let* $\alpha = (\alpha_1, \cdots, \alpha_t)$, $\alpha_i \in Z_{\mu_i}(m_i)$. *Then*

$$\deg \det [M^{\alpha}_{|\alpha|}] \leqq \sum \mu_i(m_i - \mu_i)$$

and equality occurs if, and only if, $\alpha = \bar{\mu} = (\bar{\mu}_1, \cdots, \bar{\mu}_t)$, *where* $\bar{\mu}_i = (m_i - \mu_i + 1, \cdots, m_i)$.

Proof. Let $j = |\alpha| = \Sigma\mu_i$. By the theorem of corresponding minors (the Binet-Cauchy theorem) [2, p. 14],

$$\det [M_j{}^{\alpha}] = \Sigma \det [(J^s)_{\beta}{}^{\alpha}] \det [L_j{}^{\beta}] \qquad \text{over all } \beta \in Z_j(|m|)\ .$$

Since J is block diagonal, nonzero terms occur only when β yields exactly μ_i indices corresponding to columns of J_i for each i. Such β may be rewritten $\beta = (\beta_1, \cdots, \beta_t)$, $\beta_i \in Z_{\mu_i}(m_i)$ by using (6.4) and then

$$\det [(J^s)_{\beta}{}^{\alpha}] = \prod_{i=1}^{t} \det [(J_i{}^s)^{\alpha_i}_{\beta_i}]\ , \qquad (\det [A^{\alpha_i}_{\beta_i}] = 1 \text{ if } \mu_i = 0)\ .$$

By Lemma 6.2

$$\deg \det [(J^s)_{\beta}{}^{\alpha}] \leqq \sum \mu_i(m_i - \mu_i)$$

with equality if and only if $\alpha_i = (m_i - \mu_i + 1, \cdots, m_i)$, $\beta_i = (1, \cdots, \mu_i)$. Denote this maximal minor by $\alpha = \bar{\mu}$, $\beta = \mu$. The coefficient of this minor in the expansion of $\det [M_j{}^{\bar{\mu}}]$ is $\det [L_j{}^{\mu}]$. By Lemma 6.3 $\det [L_j{}^{\mu}] \neq 0$. The lemma follows. □

We are now in a position to reorder the rows of M to get a new matrix BM with triangular factorization $BM = \tilde{L}\tilde{U}$ where $\tilde{L}$ is bounded as $s \to \infty$. Let row π_i of M be row i of BM. Let $\pi(j)$ denote the indices $\pi_1, \pi_2, \cdots, \pi_j$ arranged in increasing order. By (5.5), for $j = 1, \cdots, t$,

$$(6.2) \qquad \deg \det [M_j{}^{\pi(j)}] \geqq \deg \det [M_j{}^{\pi(j-1),\pi_i}]\ , \qquad i > j\ .$$

COROLLARY. BJB^{-1} *is upper triangular.*

Proof. We say that two indices i, j in $\{1, \cdots, |m|\}$ belong to the same block if row i and row j of J lie in the same principal submatrix J_k of J. Let i and j $(i > j)$ be indices such that π_i, π_j are in the same block. Suppose $\pi_i > \pi_j$. Then, by Lemma 6.1, the degree of $\det [M_j{}^{\pi(j)}]$ is less than the degree of the minor obtained by replacing row π_j by row π_i. This contradicts the maximality condition

(6.2). Hence $\pi_i < \pi_j$. But $J_{\pi_i,\pi_j} = (BJB^{-1})_{ij}$ and the only nonzero off-diagonal elements of J have $\pi_i = \pi_j + 1$ and π_i, π_j in the same block. It follows that $i \leqq j$ for the nonzero elements. □

At this stage we have proved that, as $s \to \infty$, the block structure to which H_s tends is the same as the block structure to which L_s tends. We have seen that L_s tends to block diagonal form, one block to all eigenvalues with the same magnitude. We have not yet determined the structure to which each diagonal block $\tilde{L}_i$ tends ($i = 1, \cdots, r$). See (4.4) for the definition of $\tilde{L}_i$.

The matrix L_s is not to be confused with the matrix L of Lemma 6.3 which is a principle submatrix of L_Y; see (6.2) and Section 4.

LEMMA 6.3. det $[L_j^\mu] \neq 0$.

Proof. Let $V = V(\lambda_1, \cdots, \lambda_1, \lambda_2, \cdots, \lambda_2, \lambda_3, \cdots, \lambda_r)$ be the confluent Vandermonde matrix associated with the eigenvalues λ_i of H. For example,

$$V(\lambda_1, \lambda_1, \lambda_1, \lambda_2, \lambda_2) = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \lambda_1^4 \\ 0 & 1 & 2\lambda_1 & 3\lambda_1^2 & 4\lambda_1^3 \\ 0 & 0 & 1 & 3\lambda_1 & 6\lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 & \lambda_2^4 \\ 0 & 1 & 2\lambda_2 & 3\lambda_2^2 & 4\lambda_2^3 \end{bmatrix}.$$

Let $V = L_V U_V$ be the triangular factorization of V.

In [7] we proved that $L_Y = L_V$ and obtained explicit formulas for the elements of L_V. The principle submatrix L of L_Y corresponds to all the eigenvalues of modulus ω_k (say), see (3.3). Let us relabel these eigenvalues $\eta_1, \cdots, \eta_t$ and their multiplicities $m_1, \cdots, m_t$. Let π be the monic polynomial whose zeros are all the eigenvalues λ_i (counting multiplicity) with $|\lambda_i| > \omega_k$, while $p_1, \cdots, p_t$ are defined by $p_1(z) = 1$,

$$p_j(z) = \prod_{i=1}^{j-1} (z - \eta_i)^{m_i}.$$

Let $L = (L^{ij})$ be partitioned conformably with the Jordan blocks of the η_i ($i = 1, \cdots, t$). Then $L^{ij} = (l^{ij}_{\alpha\beta})$ is of order $m_i \times m_j$ and in [7] we proved

$$\begin{aligned} l^{ij}_{\alpha\beta} &= (\pi p_i)^{(\alpha-\beta)}(\eta_i)/(\alpha - \beta)!(\pi p_i)(\eta_i), \quad i = j, \alpha \geqq \beta, \\ &= (d/d\eta_i)^{\alpha-1}[(\eta_i - \eta_j)^{\beta-1}(\pi p_j)(\eta_i)]/(\alpha - 1)!(\pi p_j)(\eta_j), \quad i > j. \end{aligned}$$

Here $(\pi p)(z) = \pi(z)p(z)$. By omitting the π in these expressions we obtain the elements $\hat{l}^{ij}_{\alpha\beta}$ of the lower triangular factor $\hat{L}$ of the Vandermonde matrix $\hat{V} = \hat{V}(\eta_1^{m_1}, \cdots, \eta_t^{m_t})$ associated with the eigenvalues of magnitude ω_k. By using Leibnitz' rule for differentiating πp, we find that

$$l^{ij}_{\alpha\beta} = \left(\frac{\pi(\eta_i)}{\pi(\eta_j)}\right) \sum_{\nu=0}^{\alpha-1} \frac{1}{\nu!} \frac{\pi^{(\nu)}(\eta_i)}{\pi(\eta_i)} \hat{l}^{ij}_{\alpha-\nu,\beta}, \quad i \geqq j.$$

To use this result, we define unit lower triangular matrices T_i by

$$T_i = \sum_{\nu=0}^{m_i-1} \pi^{(\nu)}(\eta_i)/\nu!\pi(\eta_i)N_i^\nu, \quad N_i = (e_2, e_3, \cdots, e_{m_i}, 0),$$

and

$$T = T_1 \oplus \cdots \oplus T_t, \qquad D = \pi(\eta_1)I_1 \oplus \cdots \oplus \pi(\eta_t)I_t.$$

Then $L = DT\hat{L}D^{-1}$ and, since D is diagonal and T block unit lower triangular

$$\det [L_j{}^\mu] = \det [D_\mu{}^\mu] \det [T_\mu{}^\mu] \det [\hat{L}_j{}^\mu]/\det [D_j{}^j] .$$

Now we observe that $\hat{V}_j{}^\mu$ is itself the Vandermonde matrix associated with $\eta_1{}^{\mu_1}, \cdots, \eta_t{}^{\mu_t}$ and so

$$0 \neq \prod_{\alpha>\beta} (\eta_\alpha - \eta_\beta)^{\mu_\alpha \mu_\beta} = \det \hat{V}_j{}^\mu = \det [\hat{L}_j{}^\mu] \det [\hat{U}_j{}^j] ,$$

the last part following from the triangular structure of $\hat{U}$. Since $\hat{U}$ is nonsingular $\det [\hat{L}_j{}^\mu] \neq 0$ and the lemma follows. □

7. Block Structure of $\tilde{L}$ as $s \to \infty$. We consider a typical block $M_{ii} = J_{ii}L_{ii}$ (see (4.2)) and drop the subscript i. With the aid of Lemma 6.2 we can describe the permutation matrix B and the bounded unit lower triangular factor $\tilde{L}$ in $BM = \tilde{L}\tilde{U}$.

To determine B it suffices to describe the permutation π characterized by condition (6.2).

Observe first that, by (5.2), for $i > j$, as $s \to \infty$,

$$\deg \det [M_j^{\pi(j)}] > \deg \det [M_j^{\pi(j-1),\pi_i}] \text{ implies } \tilde{L}(i;j) \to 0 , \tag{7.1}$$

and

$$\deg \det [M_j^{\pi(j)}] = \deg \det [M_j^{\pi(j-1),\pi_i}] \text{ implies } \tilde{L}(i;j) \not\to 0 . \tag{7.2}$$

Recall that $\pi(j)$ denotes the indices $\pi_1, \cdots, \pi_j$ arranged in increasing order.

We now imagine that the rows of M are taken one by one in some order $(\pi_1, \pi_2, \cdots)$ and placed in natural order $(1, 2, \cdots)$ in BM. At the ith step, we ask which of the remaining rows of M should be chosen as the ith row of BM. The process may be described by a variable index $\mu = (\mu_1, \cdots, \mu_t)$. Initially $\mu = (0, \cdots, 0)$. When a row is taken from block ν (the rows of J which lie in J_ν) the index μ_ν is increased by 1. This simplification is possible because of

LEMMA 7.1. *The rows of each block in M are taken in decreasing order. Thus μ indicates that the last μ_ν rows of block ν have been chosen and the first $m_\nu - \mu_\nu$ remain $(\nu = 1, \cdots, t)$.*

Proof. By Lemma 6.2, π_1 must be the last index in a block of maximal order. As induction hypothesis suppose that at step $|\mu|$ the last μ_ν rows from block ν have been assigned to BM. By criterion (6.2) the next index chosen must make $\det [M_{|\mu|+1}^{\pi(|\mu|+1)}]$ maximal among all other possible choices. By Lemma 6.2 $\pi_{|\mu|+1}$ must be the last remaining index of one of the blocks. By the principle of finite induction the lemma holds for $|\mu| = 1, \cdots, |m|$. □

This proof shows that at each step there are at most t possibilities for the next row. Let

$$d(\mu) = \sum \mu_i(m_i - \mu_i) = \sum (m_i/2)^2 - \sum (m_i/2 - \mu_i)^2 . \tag{7.3}$$

Then we seek the maximal value, $\delta(|\mu| + 1)$, among

$$\begin{aligned} &d(\mu_1 + 1, \mu_2, \cdots, \mu_t) , \\ &d(\mu_1, \mu_2 + 1, \mu_3, \cdots, \mu_t) , \\ &\quad \cdot \quad \cdot \quad \cdot \\ &d(\mu_1, \cdots, \mu_{t-1}, \mu_t + 1) . \end{aligned} \tag{7.4}$$

From the second term in (7.3) we obtain immediately

LEMMA 7.2. $\delta(|\mu| + 1) = d(\mu) + \max_i (m_i - 2\mu_i - 1)$.

This implicitly describes the permutation π and the matrix B. At step $|\mu|$ we increase one of the μ_i ($<m_i$) which satisfy

$$m_i - 2\mu_i = \max_\nu (m_\nu - 2\mu_\nu) \equiv \epsilon(|\mu|) , \tag{7.5}$$

where the maximum is over all ν with $\mu_\nu < m_\nu$. If, at step $|\mu|$, there is an r-fold choice of μ_i which achieve the maximum, then $r - 1$ steps later, at step $|\mu| + r - 1$, there will be a unique μ_i achieving the maximum because

$$\epsilon(|\mu|) = \epsilon(|\mu| + 1) = \cdots = \epsilon(|\mu| + r - 1) > \epsilon(|\mu| + r) . \tag{7.6}$$

There is a unique choice for $\pi_{|\mu|+1}$ if, and only if, $\epsilon(|\mu|) > \epsilon(|\mu| + 1)$. Hence, by (7.1), as $s \to \infty$

$$\tilde{L}(|\mu| + i; |\mu| + r) \to 0 , \qquad i > r , \tag{7.7}$$

and, by (7.2),

$$\tilde{L}(|\mu| + i; |\mu| + j) \nrightarrow 0 , \qquad i, j = 1, \cdots, r , \qquad i > j . \tag{7.8}$$

We thus see that if, at step $|\mu| - 1$, there is only one possibility for $\pi_{|\mu|}$ and, at step $|\mu|$, an r-fold choice for $\pi_{|\mu|+1}$, then the $r \times r$ principal submatrix of $\tilde{L}$ in rows $|\mu| + 1, \cdots, |\mu| + r$ has no subdiagonal elements which vanish at $s = \infty$.

We now observe that, at any step, if $\epsilon(|\mu|) = \max_\nu (m_\nu - 2\mu_\nu)$ is even then no i with m_i odd could achieve it, and vice versa. *Thus an r-fold choice occurs only among r blocks whose orders have the same parity.*

Consequently we relabel the multiplicities so that

$$e_1 = m_1 \geqq e_2 = m_2 \geqq \cdots \geqq e_p = m_p > e_{p+1} = 0$$

are even and

$$f_1 = m_{p+1} \geqq f_2 = m_{p+2} \geqq \cdots \geqq f_q = m_{p+q} > f_{q+1} = 0$$

are odd, and $p + q = t$.

LEMMA 7.3. *For each $i = 1, \cdots, p$, there are $e_i - e_{i+1}$ steps when a unique choice for $\pi_{|\mu|-1}$ is followed by an i-fold choice for $\pi_{|\mu|}$ among the i largest blocks of even order. A similar result holds for the odd case.*

Proof. The selection begins with all $\mu_\nu = 0$. Select any i among $\{1, \cdots, p\}$. We ask when μ_i increases to 1. By (7.5) $\mu_i, \cdots, \mu_p$ remain zero while at least one $e_j - 2\mu_j > e_i$, $j < i$. On the other hand while $\mu_i = 0$ we cannot have $e_j - 2\mu_j < e_i$ for any $j < i$. Thus at some stage $e_j - 2\mu_j = e_i$, $j < i$. This situation obtains until the odd blocks satisfy $m_\nu - 2\mu_\nu < e_i$, $\nu > p$. Thus when μ is such that $\epsilon(|\mu| - 1) > e_i$, $\epsilon(|\mu|) = e_i$ an i-fold choice occurs; any one of $\mu_1, \cdots, \mu_i$ may be increased.

By the same reasoning there will be an i-fold choice, following a unique choice, at each increase of μ_i until $e_i - 2\mu_i = e_{i+1}$ at which step an $(i + 1)$-fold choice occurs. This yields $\frac{1}{2}(e_i - e_{i+1})$ occurrences of an i-fold choice.

However, $d(\mu) = d(m - \mu)$, $m = (m_1, \cdots, m_t)$, and so the selection process is symmetric about $\frac{1}{2}m$. In detail, we ask when μ_i increases to $\frac{1}{2}(e_i + e_{i+1})$. By (7.5) again there must be some stage at which

$$\begin{aligned} e_\nu - 2\mu_\nu &= -e_{i+1}, \qquad \nu \leqq i, \\ \mu_\nu &= e_i, \qquad \nu = i+1, \cdots, p, \\ m_\nu - 2\mu_\nu &< -e_{i+1}, \qquad \nu > p, \\ \epsilon(|\mu| - 1) &> -e_{i+1}. \end{aligned}$$

Again any of $\mu_1, \cdots, \mu_i$ is eligible for an increase. By the same reasoning there will be such a choice at each increase of μ_i until $e_i - 2\mu_i = -e_i$, at which point an $(i - 1)$-fold choice appears. This yields another $\frac{1}{2}(e_i - e_{i+1})$ occurrences and proves the lemma. □

Since

$$(BJB^{-1})_{ii} = J_{\pi_i,\pi_i} = \omega\theta_{\pi_i}$$

it follows that the eigenvalues of the diagonal blocks of H_s whose subdiagonal elements fail to converge to zero do tend to eigenvalues whose multiplicities have the same parity.

We have not determined the exact positions of these blocks. These follow readily from (7.5) once we know the interlacing of the multiplicities $e_1, \cdots, e_p$ and $f_1, \cdots, f_q$ when they are ordered monotonically. The details are left to the interested reader.

8. An Example. Consider a 10×10 unreduced Hessenberg matrix, necessarily complex, with four distinct eigenvalues $\theta_1, \theta_2, \theta_3, \theta_4$ of modulus 1. Let their multiplicities be $m_1 = 4$, $m_2 = 3$, $m_3 = 2$, $m_1 = 1$. Then

$$(8.1)\quad J = \begin{bmatrix} \theta_1 & & & \\ 1 & \theta_1 & & \\ & 1 & \theta_1 & \\ & & 1 & \theta_1 \end{bmatrix} \oplus \begin{pmatrix} \theta_2 & & \\ 1 & \theta_2 & \\ & 1 & \theta_2 \end{pmatrix} \oplus \begin{pmatrix} \theta_3 & \\ 1 & \theta_3 \end{pmatrix} \oplus \theta_4 .$$

We give below a table showing $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$ and $\epsilon(|\mu|) = \max\,(m_\nu - 2\mu_\nu)$, $\mu_\nu < m_\nu$.

(8.2)

$\lvert\mu\rvert$	μ_1	μ_2	μ_3	μ_4	$\epsilon(\lvert\mu\rvert)$
0	0	0	0	0	4
1	1	0	0	0	3
2	1	1	0	0	2
3	2	1	0	0	2
4	2	1	1	0	1
5	2	2	1	0	1
6	2	2	1	1	0
7	3	2	1	1	0
8	3	2	2	1	−1
9	3	3	2	1	−2
10	4	3	2	1	·

Consequently one choice for B is given by

$$B^* = (e_4, e_7, e_3, e_9, e_6, e_{10}, e_2, e_8, e_5, e_1)$$

and

$$(8.3)\quad BJB^{-1} = \begin{bmatrix} \lambda_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & \lambda_2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ & & \lambda_1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ & & * & \lambda_3 & 0 & 0 & 0 & 1 & 0 & 0 \\ & & & & \lambda_2 & 0 & 0 & 0 & 1 & 0 \\ & & & & * & \lambda_4 & 0 & 0 & 0 & 0 \\ & & & & & & \lambda_1 & 0 & 0 & 1 \\ & & & & & & * & \lambda_3 & 0 & 0 \\ & & & & & & & & \lambda_2 & 0 \\ & & & & & & & & & \lambda_1 \end{bmatrix}.$$

Here the asterisk indicates that although BJB^{-1} is triangular, the matrix $\tilde{Q}_s$ tends to diagonal form except for 2×2 principal submatrices in rows 3, 4 and 5, 6 and 7, 8. Since $H_{s+1} = \tilde{Q}_s(\bar{R}BJB^{-1}\bar{R}^{-1})\tilde{Q}_s^*$ it follows that $h_{43}^{(s)}$, $h_{65}^{(s)}$ and $h_{87}^{(s)}$ are the only subdiagonal elements which fail to converge to zero as $s \to \infty$. This does not contradict the convergence of the algorithm in our use of the word.

University of California
Department of Computer Science
Berkeley, California 94720

1. J. G. F. Francis, (a) "The QR transformation: a unitary analogue to the LR transformation. I," *Comput. J.*, v. 4, 1961/62, pp. 265–271. MR **23** #B3143. (b) "The QR transformation. II," *Comput. J.*, v. 4, 1961/62, pp. 332–345. MR **25** #744.

2. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964. MR **30** #5475.

3. B. H. Kublanovskaja, "On some algorithms for the solution of the complete problem of proper values," *J. Comput. Math. and Math. Phys.*, v. 1, 1961, pp. 555–570.

4. B. N. Parlett, (a) "Convergence of the QR algorithm," *Numer. Math.*, v. 7, 1965, pp. 187–193. MR **31** #872.

Correction: *Numer. Math.*, v. 10, 1967, pp. 163–164. MR **35** #5129.

5. B. N. Parlett, (b) "The LU and QR transformations," in *Mathematical Methods for Digital Computers*, Vol. II, Wiley, New York, 1967, Chapter 5.

6. B. N. Parlett, (c) "Singular and invariant matrices under the QR algorithm," *Math. Comp.*, v. 20, 1966, pp. 611–615. MR **35** #3870.

B. N. Parlett, (d) "Canonical decomposition of Hessenberg matrices," *Math. Comp.*, v. 21, 1967, pp. 223–227.

8. J. H. Wilkinson, "Convergence of the LR, QR, and related algorithms," *Comput. J.*, v. 8, 1965, pp. 77–84. MR **32** #590.

9. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. MR **32** #1894.