An Analysis of the *HR* Algorithm for Computing the Eigenvalues of a Matrix

A. Bunse-Gerstner
Fakultät für Mathematik
Universität Bielefeld
Postfach 8640
4800 Bielefeld 1, West Germany

Submitted by Richard A. Brualdi

ABSTRACT

The *HR* algorithm, a method of computing the eigenvalues of a matrix, is presented. It is based on the fact that almost every complex square matrix A can be decomposed into a product $A=HR$ of a so-called pseudo-Hermitian matrix H and an upper triangular matrix R. This algorithm is easily seen to be a generalization of the well-known QR algorithm. It is shown how it is related to the power method and inverse iteration, and for special matrices the connection between the LR and HR algorithms is indicated.

1. INTRODUCTION

A certain class of algorithms for the computation of the eigenvalues of a nonsingular square matrix A is based on the fact that almost every matrix can be decomposed into a product of two matrices of a special form, usually one of them being upper triangular. An algorithm of this kind can be described in the following way: Let G be a subset of the nonsingular square matrices and T be a subset of the nonsingular upper triangular matrices. Starting with the first iterate

$$A_1=A,$$

at the ith step the decomposition

$$A_i=S_iR_i,\qquad S_i\in G,\quad R_i\in T,$$

52 Vanderbilt Ave., New York, NY 10017 0024-3795/81/010155+19$02.50

is computed. Then the next iterate is

$$A_{i+1}=R_iS_i.$$

Among other possibilities of choosing G, Della-Dora [5] mentioned the group O_J of the pseudo-orthogonal matrices which are nonsingular real matrices H with the property $H^TJH=J$ for a given diagonal matrix J having $+1$ and -1 entries in the diagonal. Unfortunately this nice group O_J is not suitable for such an algorithm, because the set of real square matrices which split up into a product of an $H\in O_J$ and an upper triangular matrix is too small. Brebner, Grad, and Vrečko [1, 2] proposed the set G_J of all real square matrices H with the property that there exists a permutation matrix P such that $H^TJH=P^TJP$ for a given J as defined above. Elsner [6] has shown that the direct product of G_J and the set of real upper triangular matrices with positive diagonal elements is dense in the set of all real square matrices, that is to say, almost every real square matrix A has a decomposition $A=HR$, the so-called HR decomposition with respect to J, where $H\in G_J$ and R is an upper triangular matrix with positive diagonal elements.

In this paper we study the algorithm which is based on the HR decomposition of an arbitrary complex matrix. The aim is to point out the connection between this algorithm and some well-known methods for the computation of matrix eigenvalues.

In Sec. 2 we briefly discuss the existence, uniqueness, and construction of the HR decomposition of a complex matrix and introduce the HR algorithm with respect to J. It is easily seen that for the choice J equal to the identity we get the QR algorithm as a special case.

In Sec. 3 we give a short proof of convergence for the HR algorithm in its basic form and point out that the algorithm can be interpreted as a modified power method. It is shown that the HR algorithm with shifts contains a modified inverse iteration method.

In Sec. 4 we examine how the HR algorithm acts on the special class of pseudo-Hermitian matrices which are matrices A with the property $A^*J=JA$ for a given J. The pseudo-Hermitian form is invariant under the HR algorithm. This yields in particular that after a slight modification the tridiagonal form of a real matrix can be preserved throughout the algorithm even for nonsymmetric matrices. Finally we show that for pseudo-Hermitian matrices the LR and HR algorithms are very closely related. The $(2i+1)$th iterate of the LR algorithm and the $(i+1)$th iterate of the HR algorithm differ only by a similarity transformation with a diagonal matrix.

2. THE *HR* ALGORITHM

We denote

by $\mathrm{GL}_n(\mathbf{C})$ and $\mathrm{GL}_n(\mathbf{R})$ the set of all nonsingular complex and real $n\times n$ matrices respectively,

by $T_n(\mathbf{C})$ and $T_n(\mathbf{R})$ the set of all nonsingular complex and real upper triangular $n\times n$ matrices respectively, and

by $T_n^+(\mathbf{C})$ and $T_n^+(\mathbf{R})$ all matrices of $T_n(\mathbf{C})$ and $T_n(\mathbf{R})$ respectively which in addition have a real positive diagonal.

By $\mathrm{diag}_k^n(\pm 1)$ we denote the set of all $n\times n$ diagonal matrices with $+1$ and -1 entries on the diagonal where k is the exact number of negative entries.

Definition 2.1. Let $J_1, J_2 \in \mathrm{diag}_k^n(\pm 1)$. Then $H\in \mathrm{GL}_n(\mathbf{C})$ is called (J_1, J_2)-unitary if $H^*J_1H=J_2$.

By $U_n(J_1, J_2)$ we denote the set of all (J_1, J_2)-unitary matrices.

Remark 2.2. If $J_1\in \mathrm{diag}_k^n(\pm 1)$, $J_2\in \mathrm{diag}_m^n(\pm 1)$, and $k\neq m$, then obviously, by a version of Sylvester's law of inertia for Hermitian forms [9], $U_n(J_1, J_2)$ is empty. For $k=m$ it is easy to verify that $U_n(J_1,J_2)\cap T_n^+(\mathbf{C})$ contains only the identity matrix I if $J_1=J_2$, and that it is empty if $J_1\neq J_2$. Also, $U_n(I,I)=U_n(-I,-I)$ is the set of all unitary matrices.

The following theorem shows that for almost every matrix there exists a decomposition into a product HR where $R\in T_n(\mathbf{C})$ and $H\in U_n(J_1,J_2)$ for a given J_1 and a suitable J_2. This decomposition is uniquely determined if in addition we demand $R\in T_n^+(\mathbf{C})$.

Theorem 2.3. *Let* $A\in GL_n(\mathbf{C})$ *and* $J_1, J_2\in \mathrm{diag}_k^n(\pm 1)$.

(i) *There exist* $H\in U_n(J_1,J_2)$ *and* $R\in T_n(\mathbf{C})$ *with* $A=HR$ *if and only if no principal minor of* A^*J_1A *vanishes and the product of the first* i *diagonal elements of* J_2 *coincides with the sign of the* i*th principal minor of* A^*J_1A *for all* $i\in\{1,\dots,n\}$.

(ii) *Let* $H_1, H_2\in U_n(J_1,J_2)$, $R_1, R_2\in T_n^+(\mathbf{C})$. *If* $A=H_1R_1=H_2R_2$, *then* $H_1=H_2$ *and* $R_1=R_2$.

Proof. (i): A^*J_1A has all principal minors nonzero if and only if it has an LR decomposition $A^*J_1A=L\tilde{R}$, where $\tilde{R}, L^*\in T_n(\mathbf{C})$ and L has a unit diagonal. The product of the first i diagonal elements of $\tilde{R}$ coincides with the ith principal minor of A^*J_1A for all $i \in \{1,\dots,n\}$ [14, pp. 201–205].

As A^*J_1A is Hermitian, all its principal minors are real, and therefore it is easily seen that all diagonal elements of $\tilde{R}$ are real. Let J_2 be a diagonal matrix with $+1$ and -1 entries on the diagonal such that for all $i \in \{1,\dots,n\}$ the product of the first i diagonal elements is just the sign of the ith principal minor of A^*J_1A. Then obviously the ith diagonal element of J_2 is the sign of the ith diagonal element of $\tilde{R}$, and we can find a real diagonal matrix D such that $\hat{R} = D^{-1}J_2D^{-1}\tilde{R}$ has a unit diagonal. As A^*J_1A is Hermitian, we have

$$A^*J_1A = L\tilde{R} = LDJ_2D\hat{R} = \hat{R}^*DJ_2DL^*,$$

and the uniqueness of the LR decomposition of A^*J_1A implies $L = \hat{R}^*$. Defining $R = D\hat{R}$, we have now

$$A^*J_1A = R^*J_2R. \tag{2.3.1}$$

[Note that by Sylvester's law of inertia $J_2 \in \mathrm{diag}_k^n(\pm 1)$.] From (2.3.1) we get $A = J_1A^{-*}RJ_2R$, and for $H = J_1A^{-*}R^*J_2$ we find $H \in U_n(J_1, J_2)$.

On the other hand, if $A = HR$ with $R \in T_n(\mathbf{C})$ and $H \in U_n(J_1, J_2)$, then $A^*J_1A = R^*H^*J_1HR = R^*J_2R$, which means that A^*J_1A has an LR decomposition. By examining the principal minors of R^*J_2R it is easily seen that the ith principal minor of J_2 is just the sign of the ith principal minor of A^*J_1A.

(ii): From $H_1R_1 = H_2R_2$ we get $R_1R_2^{-1} = H_1^{-1}H_2 \in U_n(J_2, J_2) \cap T_n^+(\mathbf{C})$. From Remark 2.2 it follows that $R_1R_2^{-1} = H_1^{-1}H_2 = I$. ■

By this theorem we are justified in defining the following.

Definition 2.4. Let $A \in \mathrm{GL}_n(\mathbf{C})$ and $J_1 \in \mathrm{diag}_k^n(\pm 1)$. Let A^*J_1A have no vanishing principal minor, and let $J_2 \in \mathrm{diag}_k^n(\pm 1)$ such that for all $i \in \{1,\dots,n\}$ the product of the first i diagonal elements of J_2 coincides with the sign of the ith principal minor of A^*J_1A. The factorization $A = HR$ with $H \in U_n(J_1, J_2)$, $R \in T_n^+(\mathbf{C})$ is called the HR decomposition of A with respect to J_1.

Remark 2.5.

(i) For $J_1 = I$ or $J_1 = -I$ we get the QR decomposition of A. By Theorem 2.3 the condition for this decomposition to exist is that no principal minor of A^*A vanishes, which is always true.

(ii) As the ith principal minor of $(A-sI)^*J_1(A-sI)$ is a polynomial in s of degree $2i$ there are at most $n(n+1)$ values $\tilde{s}$ for which $(A-\tilde{s}I)$ has no *HR* decomposition with respect to J_1. Therefore a nondecomposable matrix can always be shifted into a decomposable one even if we confine ourselves to real shifts.

(iii) If $A\in \mathrm{GL}_n(\mathbf{R})$ and $A=HR$ is the *HR* decomposition of A with respect to J_1, then it can be shown that both matrices H and R are real. This becomes more evident by looking at the construction of such a decomposition.

For given $A\in \mathrm{GL}_n(\mathbf{C})$ and $J\in \mathrm{diag}_k^n(\pm 1)$ the *HR* decomposition with respect to J can be constructed by computing H^{-1}, the matrix which reduces A to upper triangular form, as a product of elementary elimination matrices and permutation matrices. To eliminate an element in position (m,i), $i<m$, we can use matrices $H_1=(h_{l_i})$ defined by

$$\left.\begin{array}{ll} h_{ii}=e^{i\alpha}\cos\phi, & h_{im}=e^{i\beta}\sin\phi, \\ h_{mi}=-e^{-i\beta}\sin\phi, & h_{mm}=e^{-i\alpha}\cos\phi, \\ h_{pp}=1 \quad \text{for } p\neq i,m, & h_{pq}=0 \quad \text{otherwise} \end{array}\right\} \quad \text{if } j_i=j_m$$

and

$$\left.\begin{array}{ll} h_{ii}=e^{i\alpha}\cosh\phi, & h_{im}=e^{i\beta}\sinh\phi, \\ h_{mi}=e^{-i\beta}\sinh\phi & h_{mm}=e^{-i\alpha}\cosh\phi, \\ h_{pp}=1 \quad \text{for } p\neq i,m, & h_{pq}=0 \quad \text{otherwise} \end{array}\right\} \quad \text{if } j_i=-j_m.$$

To eliminate all elements of a whole column but the first, we can use matrices H_1 defined by

$$H_1=I-2J\mathbf{v}\mathbf{v}^* \qquad \text{where} \quad \mathbf{v}\in\mathbf{C}^n,\ \mathbf{v}^*J\mathbf{v}=1.$$

In any case we have $H_1\in U_n(J,J)$.

In [1] and [2] these transformations are discussed for real matrices, and in [3] Brebner and Grad discuss the danger of severe cancellation errors when calculating elements of these transformation matrices. In the ith step of reduction using the second class of elimination matrices we proceed as follows: Let $A^{(i)} = H_{i-1} \cdots H_1 A$ already be a matrix for which all elements below the diagonal in the first $i-1$ columns vanish and $H_{i-1} \cdots H_1 \in U_n(J, J_{i-1})$. If $J_{i-1} = \operatorname{diag}(j_1, \ldots, j_n)$, then we denote $\tilde{J} = \operatorname{diag}(j_i, \ldots, j_n)$. Let $\mathbf{a} \in \mathbf{C}^{n-i+1}$ be the vector formed by the last $n-i+1$ elements of the ith column of $A^{(i)}$, and assume $\mathbf{a}^* \tilde{J} \mathbf{a} \neq 0$. Let $j_m \in \{j_i, \ldots, j_n\}$ be chosen so that $j_m \mathbf{a}^* \tilde{J} \mathbf{a}$ is positive, and let P be the $(n-i+1)$-dimensional permutation matrix which interchanges row 1 and $m-i+1$. Then for $\mathbf{b} = P\mathbf{a}$ and $\hat{J} = P\tilde{J}P = \operatorname{diag}(\hat{j}_1, \ldots, \hat{j}_{n-i+1})$ we have $\hat{j}_1 \mathbf{b}^* \hat{J} \mathbf{b} > 0$.

Now

$$H^{-1} = I - \frac{\hat{j}_1}{K}(\mathbf{b} - z\mathbf{e}_1)(\mathbf{b} - z\mathbf{e}_1)^* \hat{J}$$

with

$$z = -\operatorname{sgn}(b_1)\left(\hat{j}_1 \mathbf{b}^* \hat{J} \mathbf{b}\right)^{1/2} \quad \text{and} \quad K = z(z - b_1)$$

[b_1 is the first component of $\mathbf{b}$, and $\operatorname{sgn} b_1 = b_1/|b_1|$; $\mathbf{e}_1$ is the $(n-i+1)$-dimensional first unit vector] is a $(\hat{J}, \hat{J})$-unitary matrix with the property

$$H^{-1}P\mathbf{a} = H^{-1}\mathbf{b} = z\mathbf{e}_1.$$

$H^{-1}P$ is enlarged to give an n-dimensional matrix

$$H_i = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & H^{-1}P \end{array}\right].$$

Defining $J_i = H_i^* J_{i-1} H_i$, we get $H_i H_{i-1} \cdots H_1 \in U_n(J, J_i)$, and in $A^{(i+1)} = H_i A^{(i)}$ we have one more column with zero elements below the diagonal. For the special case $J = I$ the reduction using the first class of elimination matrices is just Givens's method, and the reduction using the second class of elimination matrices is just Householder's method to compute the QR decomposition of a matrix.

An algorithm based on the HR decomposition can now be defined. Let $A \in \mathrm{GL}_n(\mathbf{C})$ and $J \in \operatorname{diag}_k^n(\pm 1)$. The HR algorithm with respect to J produces a sequence of matrices $\{A_i\}_{i \in \mathbf{N}}$ in such a way that, starting with

$$A_1 = A, \qquad J_1 = J,$$

in the ith step the *HR* decomposition of A_i with respect to J_i is computed:

$$A_i = H_i R_i.$$

Subsequently A_{i+1} is constructed and J_{i+1} is defined by

$$A_{i+1} = R_i H_i, \qquad J_{i+1} = H_i^* J_i H_i.$$

REMARK 2.6.

(i) For all $i \in \mathbf{N}_+$ it is easily shown that:

(a) $A_{i+1} = H_i^{-1} A_i H_i = R_i A_i R_i^{-1}$.
(b) $A_{i+1} = (H_1 \cdots H_i)^{-1} A (H_1 \cdots H_i) = (R_i \cdots R_1) A (R_i \cdots R_1)^{-1}$.
(c) $A^i = H_1 \cdots H_i R_i \cdots R_1$.
(d) From the second equality in (b) it follows that for an upper Hessenberg A all A_i are upper Hessenberg.

(ii) Obviously we have, for all $i \in \mathbf{N}_+$ and all $m \in \{1, \dots, i-1\}$, that $(H_m \cdots H_i)^* J_m (H_m \cdots H_i) = J_{i+1}$. Therefore all J_i produced in the algorithm are contained in $\operatorname{diag}_k^n(\pm 1)$.

(iii) For $J = I$ or $J = -I$ this is just the well-known *QR* algorithm in its basic form.

(iv) For $J \neq I$ and $J \neq -I$ it may occur that an iterate A_i has no *HR* decomposition with respect to J_i. Then we say that the *HR* algorithm with respect to J (without shifts) is not constructible for A.

(v) If A is a real matrix, then so are all matrices occurring in the algorithm.

3. CONVERGENCE PROPERTIES

For matrices with eigenvalues of distinct moduli we give a proof of convergence for the *HR* method similar to the proofs known for the *LR* and *QR* algorithms in this case.

THEOREM 3.1. *Let $A \in \mathrm{GL}_n(\mathbf{C})$, $J \in \operatorname{diag}_k^n(\pm 1)$, $Y \in \mathrm{GL}_n(\mathbf{C})$, and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, where $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ and $A = YDY^{-1}$. Let Y^{-1} have an LR decomposition, and let Y have an HR decomposition with respect to J. Let the HR algorithm with respect to J be constructible for A. If*

$A_s=(a_{ij}^{(s)})$ *denotes the sth iterate in the algorithm, then*

$$\lim_{s\to\infty} a_{ij}^{(s)}=0 \qquad \textit{for} \quad i<j, \quad i,j\in\{1,\dots,n\},$$

and

$$\lim_{s\to\infty} a_{ii}^{(s)}=\lambda_i \qquad \textit{for all} \quad i\in\{1,\dots,n\}.$$

Proof. Let $Y^{-1}=L_Y R_Y$ denote the LR decomposition of Y^{-1}, $Y=H_Y\tilde{R}$ the HR decomposition of Y with respect to J, and $L_Y=(l_{ij})$, $J'=H^*JH$. For $s\in\mathbf{N}$ the matrix $D^sL_YD^{-s}$ is lower triangular with a unit diagonal. As $(\lambda_i/\lambda_j)^s l_{ij}$ is the (i,j)th element for $i<j$, there exists a sequence of matrices $\{E_s\}_{s\in\mathbf{N}}$ which tends to the zero matrix for $s\to\infty$ and $D^sL_YD^{-s}=I+E_s$.

Because the HR decomposition is unique and continuous and $\lim_{s\to\infty}(I+\tilde{R}_YE_s\tilde{R}_Y^{-1})=I$, for sufficiently large s there exists the HR decomposition with respect to J': $I+\tilde{R}_YE_s\tilde{R}_Y^{-1}=\hat{H}_s\hat{R}_s$ with $\hat{H}_s^*J'\hat{H}_s=J_s$ and $\lim_{s\to\infty}\hat{H}_s=I$. Therefore for sufficiently large s we have

$$\begin{aligned}A^s&=YD^sY^{-1}=YD^sL_YD^{-s}D^sR_Y=Y(I+E_s)D^sR_Y=H_Y\tilde{R}_Y(I+E_s)D^sR_Y\\&=H_Y(I+\tilde{R}_YE_s\tilde{R}_Y^{-1})\tilde{R}_YD^sR_Y=H_Y\hat{H}_s\hat{R}_s\tilde{R}_YD^sR_Y.\end{aligned}$$

Let $\tilde{D}_s$ be a unitary diagonal matrix such that $\tilde{D}_s\hat{R}_s\tilde{R}_YD^sR_Y\in T_n^+(\mathbf{C})$. Now $H_Y\hat{H}_s\tilde{D}_s^{-1}\in U_n(J,J_s)$ and with Remark 2.6(i) (c) we find two HR decompositions with respect to J of A^s:

$$A^s=H_Y\hat{H}_s\tilde{D}_s^{-1}\tilde{D}_s\hat{R}_s\tilde{R}_YD^sR_Y \quad\text{and}\quad A^s=H_1\cdots H_sR_s\cdots R_1.$$

Because of uniqueness $H_Y\hat{H}_s\tilde{D}_s^{-1}=H_1\cdots H_s$ must hold, and with Remark 2.6 (i) (b) this yields

$$\begin{aligned}A_{s+1}&=\left(H_Y\hat{H}_s\tilde{D}_s^{-1}\right)^{-1}AH_Y\hat{H}_s\tilde{D}_s^{-1}\\&=\tilde{D}_sH_s^{-1}\hat{H}_Y^{-1}YDY^{-1}H_Y\hat{H}_s\tilde{D}_s^{-1}=\tilde{D}_s^{-1}\hat{H}_s^{-1}\tilde{R}_YD\tilde{R}_Y^{-1}\hat{H}_s\tilde{D}_s^{-1}.\end{aligned}$$

As $\lim_{s\to\infty}\hat{H}_s=I$ and $\tilde{D}_s$, $\tilde{D}_s^{-1}$ are bounded, we get the statement of the theorem by this last equation. ■

According to this theorem the diagonal elements $a_{ii}^{(s)}$ converge to the eigenvalues λ_i of A essentially as fast as $(\lambda_i/\lambda_{i+1})^s$ converges to 0 for $s\to\infty$. But this is just the rate of convergence we get when using the power method for computing λ_i. Parlett and Poole [10] pointed out how the *QR* algorithm is connected to the power method. The following lemma proves that the *HR* algorithm can likewise be interpreted as a nested sequence of n power methods starting with the subspaces spanned by $\{\mathbf{e}_1\}$, $\{\mathbf{e}_1,\mathbf{e}_2\},\dots,$ $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$, where $\mathbf{e}_i$ denotes the ith unit vector.

For $A\in \mathrm{GL}_n(\mathbf{C})$ let $\mathcal{A}$ denote the operator on $\mathbf{C}^n$ defined by A. For an $n\times q$ matrix X let $\mathbf{X}$ denote the subspace of $\mathbf{C}^n$ spanned by the q columns of X.

Starting with an $n\times q$ matrix X_0 corresponding to the subspace $\mathbf{X}_0$, the power method creates a sequence of subspaces $\mathbf{X}_k=\mathcal{A}\mathbf{X}_{k-1}$. For each $k\in\mathbf{N}_+$ the subspace $\mathbf{X}_k$ is represented by a suitable $n\times q$ matrix X_k, which means that in each step X_k is gained by suitably normalizing AX_{k-1}. Under certain conditions the sequence $\{\mathbf{X}_k\}_{k\in\mathbf{N}}$ "converges" to the dominant q-dimensional invariant subspace of $\mathcal{A}$ [10].

Now looking at the *HR* algorithm with respect to $J\in\mathrm{diag}_k^n(\pm1)$ applied to A, with the usual notation $A_s=H_sR_s$, $H_s^*J_sH_s=J_{s+1}$ for the sth step and with $P_s=H_1\cdots H_s$, the following holds:

LEMMA 3.2. *For each $s\in\mathbf{N}_+$ and $q\in\{1,\dots,n\}$, the first q columns of P_s span the subspace $\mathbf{X}_s$ of the power method applied to the starting subspace spanned by the first q unit vectors.*

Proof. According to Remark 2.6(i) (b) we have $A_{s+1}=P_s^{-1}AP_s$ or $AP_s=P_sA_{s+1}$, which leads to

$$AP_sR_{s+1}^{-1}=P_sA_{s+1}R_{s+1}^{-1}=P_sH_{s+1}R_{s+1}R_{s+1}^{-1}=P_{s+1}. \tag{3.2.1}$$

Therefore the first q columns of P_{s+1} are linear combinations of the first q columns of AP_s. If $\mathbf{P}_{s,q}$ denotes the subspace spanned by the first q columns of P_s, then

$$\mathcal{A}\mathbf{P}_{s,q}=\mathbf{P}_{s+1,q}\text{ for all }s\in\mathbf{N}_+$$

with $P_1=H_1=AR_1^{-1}=AIR_1^{-1}$. ■

In particular, for $q=1$ the *HR* algorithm in its basic form contains the modified power method:

THEOREM 3.3. *Let* $\mathbf{p}_s$ *denote the first column of* P_s. *The sequence* $\{\mathbf{p}_s\}_{s\in\mathbf{N}}$ *is then created by the modified power method:* $\mathbf{p}_0=\mathbf{e}_1$, $\mathbf{p}_{s+1}=(1/r_{s+1})A\mathbf{p}_s$ *with* $r_{s+1}=(|\mathbf{p}_s^*A^*JA\mathbf{p}_s|)^{1/2}$ *for all* $s\in\mathbf{N}_+$. *The scalar* r_{s+1} *can be considered as an approximation at the* $(s+1)$*th step to the modulus of the dominant eigenvalue of* A.

Proof. If we denote by r the first diagonal element of R_{s+1}, then (3.2.1) yields

$$\mathbf{p}_{s+1}=\frac{1}{r}A\mathbf{p}_s.$$

Let $\mathbf{a}_{s+1}$ be the first column of A_{s+1}. Then $A_{s+1}=H_{s+1}R_{s+1}$ and $H_{s+1}^*J_{s+1}H_{s+1}=J_{s+2}$ imply

$$\left|\left(\frac{1}{r}\mathbf{a}_{s+1}\right)^*J_{s+1}\left(\frac{1}{r}\mathbf{a}_{s+1}\right)\right|=1$$

and therefore $r^2=|\mathbf{a}_{s+1}^*J_{s+1}\mathbf{a}_{s+1}|$. In addition, we see from $A_{s+1}=P_s^{-1}AP_s$ and $P_s^*JP_s=J_{s+1}$ that

$$\mathbf{a}_{s+1}^*J_{s+1}\mathbf{a}_{s+1}=\left(P_s^{-1}A\mathbf{p}_s\right)^*J_{s+1}\left(P_s^{-1}A\mathbf{p}_s\right)=\mathbf{p}_s^*A^*JA\mathbf{p}_s.$$

Therefore $r=r_{s+1}$ holds.

According to Theorem 3.1, under suitable conditions $\mathbf{a}_{s+1}$ tends to $\lambda\mathbf{e}_1$, λ being the dominant eigenvalue of A. Hence $(|\mathbf{a}_{s+1}^*J_{s+1}\mathbf{a}_{s+1})|)^{1/2}$ can be considered as an approximation to $|\lambda|$ at the $(s+1)$th step, since $\lim_{s\to\infty}(|\mathbf{a}_{s+1}^*J_{s+1}\mathbf{a}_{s+1}|)^{1/2}=\lim_{s\to\infty}r_{s+1}=|\lambda|$ holds for strictly dominant λ. ∎

To accelerate convergence, shifts can be used in the algorithm. The *HR* algorithm with respect to $J\in\operatorname{diag}_k^n(\pm 1)$ and the sequence $\{k_i\}_{i\in\mathbf{N}_+}$ of shift parameters reads:

$$A_1=A,\qquad J_1=J,$$

$$(A_i-k_iI)=H_iR_i\qquad\text{with}\quad H_i^*J_iH_i=J_{i+1}\text{ and }R_i\in T_n^+(\mathbf{C}),$$

$$A_{i+1}=R_iH_i+k_iI.$$

For these A_i, H_i, R_i and J_i the statements of Remark 2.6 except 2.6(i) (c) remain valid.

Allowing shifts, we can also assure that the algorithm is constructible, because a nondecomposable or almost nondecomposable matrix can always be shifted away from this dangerous region.

Stewart [12] demonstrated that the *QR* algorithm contains a modified inverse iteration method. We shall see how this carries over to its generalization. If in the *HR* algorithm with respect to J we take the last diagonal element of A_i as shift parameter k_i, with the abovementioned notation and defining $P_i = H_1, \cdots, H_i$, we have

THEOREM 3.4. *Let* $\mathbf{p}_s$ *denote the last column of* P_s. *The sequence* $\{\mathbf{p}_s\}_{s\in\mathrm{N}}$ *is then created by the modified inverse iteration method*:

$$\mathbf{p}_0 = \mathbf{e}_n, \qquad k_1 = \mathbf{p}_0^T A \mathbf{p}_0, \quad j_1 = \mathrm{sgn}(\mathbf{p}_0^* J \mathbf{p}_0),$$

and for all $s \in \mathrm{N}$,

$$\mathbf{p}_{s+1} = \hat{\mathbf{p}}_{s+1} j_{s+1} j_{s+2} r_{s+1}, \qquad k_{s+2} = j_{s+2} \mathbf{p}_{s+1}^* JA \mathbf{p}_{s+1} \quad j_{s+2} = \mathrm{sgn}(\hat{\mathbf{p}}_{s+1}^* J \hat{\mathbf{p}}_{s+1}).$$

with

$$\hat{\mathbf{p}}_{s+1} = J(A - k_{s+1} I)^{-*} J \mathbf{p}_s, \qquad r_{s+1} = \left(|\hat{\mathbf{p}}_{s+1}^* J \hat{\mathbf{p}}_{s+1}|\right)^{-1/2}.$$

Proof. $AP_s = P_s A_{s+1}$ implies $P_{s+1} = P_s H_{s+1} R_{s+1} R_{s+1}^{-1} = P_s(A_{s+1} - k_{s+1} I) R_{s+1}^{-1} = (A - k_{s+1} I) P_s R_{s+1}^{-1}$ and therefore $P_{s+1}^{-*} = (A - k_{s+1} I)^{-*} P_s^{-*} R_{s+1}^*$. Together with $P_{s+1}^* J P_{s+1} = J_{s+2}$, this yields $P_{s+1} = J(A - k_{s+1} I)^{-*} J P_s J_{s+1} R_{s+1}^* J_{s+2}$. If we denote by r the last diagonal element of R_{s+1} and by j and $\tilde{j}$ the last diagonal elements of J_{s+1} and J_{s+2} respectively, then $\mathbf{p}_{s+1} = J(A - k_{s+1} I)^{-*} J \mathbf{p}_s r j \tilde{j}$ follows. From $(A_{s+1} - k_{s+1} I)^{-*} R_{s+1}^* = H_{s+1}^{-*}$ we find, with an argument similar to the one used for Theorem 3.3, that $r = r_{s+1}$. That $j = j_{s+1}$ and $\tilde{j} = j_{s+2}$ is a trivial consequence of $P_i J P_i = J_{i+1}$ for all $i \in \mathrm{N}$. ∎

For real upper Hessenberg A two successive steps of the *HR* algorithm can be performed together, avoiding complex arithmetic if two real or complex conjugate shifts are used for such a two-step iteration. This is possible because the transformation of a matrix to upper Hessenberg form using (J, J')-unitary matrices is essentially unique:

LEMMA 3.5. *Let* $A \in \mathrm{GL}_n(\mathrm{C})$, $J_1, J_2, J_3 \in \mathrm{diag}_k^n(\pm 1)$, *and* $H_1 \in U_n(J_1, J_2)$, $H_2 \in U_n(J_1, J_3)$. *Let* $H_1^{-1} A H_1$ *and* $H_2^{-1} A H_2$ *be upper Hessenberg with at least* *one having all subdiagonal elements nonzero. If H_1 and H_2 have the same first column, then there exists a unitary diagonal matrix D such that $H_1 = DH_2$.*

Proof. Let $B = H_1^{-1}AH_1 = (b_{ij})$ have all subdiagonal elements nonzero. With $C = H_2^{-1}AH_2 = (c_{ij})$ and $H = H_2^{-1}H_1$, we have

$$CH = HB. \tag{3.5.1}$$

If $\mathbf{h}_m$ denotes the mth column of H, then $\mathbf{h}_1 = \mathbf{e}_1$. It can now be shown by induction using (3.5.1) and

$$H^*J_2H = J_3 \tag{3.5.2}$$

that $\mathbf{h}_i = z_i\mathbf{e}_i$ with $|z_i| = 1$.

From (3.5.2), $0 = \mathbf{h}_2^* J \mathbf{h}_1 = \mathbf{h}_2^* J_2 \mathbf{e}_1$ follows, which means that the first component of $\mathbf{h}_2$ vanishes. Now (3.5.1) yields $C\mathbf{e}_1 = b_{11}\mathbf{e}_1 + b_{21}\mathbf{h}_2$ or $\mathbf{h}_2 = (1/b_{21})[(c_{11} - b_{11}, c_{21}, 0, \ldots, 0)^{\mathrm{T}}]$, which leads to $b_{11} = c_{11}$, $\mathbf{h}_2 = (c_{21}/b_{21})\mathbf{e}_2$. For $z_1 := c_{21}/b_{21}$ we find from (3.5.2) that $|z_1| = 1$. Proceeding in the same way we get the statement. ■

The two-step iteration technique with complex conjugate shifts or two real shifts is very well known for the QR algorithm, and with the foregoing lemma the arguments there carry over to the HR algorithm. Therefore we shall not go into further details.

Usually the shifts k_{2i+1}, k_{2i+2} are taken to be the two eigenvalues of the 2×2 matrix in the bottom right-hand corner of the current A_{2i+1}. The first column of $H \in U_n(J_{2i+1}, J_{2i+3})$ is computed, where HR is the HR decomposition with respect to J_{2i+1} of the real matrix $(A_{2i+1} - k_{2i+1}I)(A_{2i+1} - k_{2i+2}I)$. Then a (J_{2i+1}, J_{2i+3})-unitary matrix is constructed which has just this first column and transforms A_{2i+1} into an upper Hessenberg A_{2i+3}. In the following small examples the eigenvalues were computed using the HR algorithm with respect to J in this form.

EXAMPLE 3.6. We give the results computed by the HR algorithm with respect to J for several J. For each J we list the computed eigenvalues, the total number of two-step iterations, and the actual computing time on the TR 440 at the University of Bielefeld.

(i) $$\begin{bmatrix} 4 & 1 & 1 \\ 2 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} \qquad \text{Eigenvalues: } \begin{matrix} 6 \\ 3 \\ 3 \end{matrix}$$

	Computed Eigenvalues	Iterations	Time (10^{-5} sec)
(a)	$J=I$		
	6.00000000012 3.00000894629 2.99999105372	4	37.03
(b)	$J=\operatorname{diag}(-1,-1,1)$		
	6.00000000064 $3.00000000018+i0.00000190734$ $3.00000000018-i0.00000190734$	4	41.09
(c)	$J=\operatorname{diag}(1,-1,1)$		
	6.00000000006 3.00000000006 3.00000000000	1	11.12

(ii) $$\begin{bmatrix} 0.00 & 0.07 & 0.27 & -0.33 \\ 1.31 & -0.36 & 1.21 & 0.41 \\ 1.06 & 2.86 & 1.49 & -1.34 \\ -2.64 & -1.84 & -0.24 & -2.01 \end{bmatrix} \quad \text{Eigenvalues: } \begin{array}{r} 0.03 \\ 3.03 \\ -1.97+i \\ -1.97-i \end{array}$$

Before applying the *HR* algorithm the matrix was transformed into upper Hessenberg form.

	Computed Eigenvalues	Iterations	Time (10^{-5} sec)
(a)	$J=I$		
	3.03000000009 $-1.96999999991\pm i1.00000000006$ 0.02999999989	12	74.79
(b)	$J=\operatorname{diag}(1,-1,-1,1)$		
	3.03000000044 0.03000000113 $-1.97000000143\pm i1.00000000087$	16	209.19
(c)	$J=\operatorname{diag}(-1,1,-1,-1)$		
	3.03000000073 $-1.97000000032\pm i1.00000000023$ 0.02999999995	14	187.19

Note that for both cases (a) the method is just the *QR* algorithm.

4. PSEUDO-HERMITIAN MATRICES

The Hermitian form of a matrix is preserved under similarity transformations with unitary matrices. It is well known that advantage can be taken of this fact for an economical application of the QR algorithm to Hermitian or symmetric matrices. For the generalization studied here we can observe that the so-called pseudo-Hermitian form of a matrix is invariant under similarity transformations with (J, J')-unitary matrices.

DEFINITION 4.1. Let $A\in \mathrm{GL}_n(\mathbf{C})$ and $J\in \mathrm{diag}_k^n(\pm 1)$. A is called

J-Hermitian if $A^*J=JA$, and
J-symmetric if in addition A is real.

A is called pseudo-Hermitian (pseudo-symmetric) if there exist a $k\in\{1,\dots,n\}$ and a $J\in \mathrm{diag}_k^n(\pm 1)$ such that A is J-Hermitian (J-symmetric).

EXAMPLE 4.2.

(i) For a Hermitian or symmetric nonsingular matrix B we can find a nonsingular matrix M and a $J\in \mathrm{diag}_k^n(\pm 1)$ such that $B=M^*JM$, where J contains the signs of the eigenvalues of B. If we have to solve the problem $A\mathbf{x}=\lambda B\mathbf{x}$ where A too is Hermitian or symmetric, then we may transform this equation to $JM^{-*}AM^{-1}M\mathbf{x}=\lambda M\mathbf{x}$ or $C\mathbf{y}=\lambda\mathbf{y}$, where $\mathbf{y}=M\mathbf{x}$, and $C=JM^{-*}AM^{-1}$ is a J-Hermitian or J-symmetric matrix. Some examples for the symmetric case are given in [1].

(ii) Specifically, for real matrices A each eigenvalue problem $A\mathbf{x}=\lambda\mathbf{x}$ can be transformed into problems $T\mathbf{y}=\lambda\mathbf{y}$ with T pseudosymmetric and tridiagonal. For it is known [11] that any real matrix A is similar to a real tridiagonal matrix $\tilde{T}$, which may be obtained from A by the Lanczos method [8] with suitable starting vectors for instance. It is easily seen that by similarity transformation with a diagonal matrix, each $\tilde{T}$ with all codiagonal elements nonzero can be transformed into a tridiagonal T for which corresponding codiagonal elements have the same absolute value. Such a T is J-symmetric with

$$J=\mathrm{diag}\left(1,\mathrm{sgn}\frac{t_{12}}{t_{21}},\dots,\mathrm{sgn}\left(\frac{t_{12}}{t_{21}}\cdots\frac{t_{n-1,n}}{t_{n,n-1}}\right)\right).$$

Note that for complex matrices we have no analogous transformation into a pseudo-Hermitian tridiagonal matrix, because $A=M^{-1}TM$ for a J-Hermitian T would imply $A=M^{-1}JJTM=M^{-1}JM^{-*}M^*JTM$. So A has to be a product of two Hermitian matrices, namely $M^{-1}JM^{-*}$ and M^*JTM, which is not true for arbitrary $A\in \mathrm{GL}_n(\mathbf{C})$ but holds for arbitrary real matrices (see [7], [13]).

Now if $J, J' \in \mathrm{diag}_k^n(\pm 1)$ and $H \in U_n(J, J')$, then for a J-Hermitian A the matrix $H^{-1}AH$ is J'-Hermitian, because

$$(H^{-1}AH)^*J' = H^*A^*H^{-*}J' = H^*A^*JH = H^*JAH = J'H^{-1}AH.$$

In particular, if we apply the *HR* algorithm with respect to J to a J-Hermitian matrix A, then because

$$A_i = H_{i-1}^{-1} \cdots H_1^{-1}AH_1 \cdots H_{i-1} \quad \text{and} \quad H_1 \cdots H_{i-1} \in U_n(J, J_i),$$

each iterate A_i is J_i-Hermitian.

A pseudo-Hermitian upper Hessenberg matrix is obviously tridiagonal. Therefore, because of the invariance of the upper Hessenberg form under the *HR* algorithm, we find that starting with a tridiagonal J-Hermitian or J-symmetric matrix A, in the *HR* algorithm with respect to J all iterates are tridiagonal.

This is of special interest in the case of real matrices which are tridiagonal but not symmetric or which can be transformed into such a matrix in a stable manner. According to Example 4.2(ii) this tridiagonal matrix can be easily modified to a J-symmetric tridiagonal matrix, and unlike the *QR* algorithm, the *HR* algorithm with respect to J preserves this tridiagonal form even if this starting matrix is not symmetric.

In the special case of tridiagonal pseudosymmetric matrices convergence can be proved [4] under much weaker conditions than in Theorem 3.1.

Finally, for J-Hermitian matrices it can be shown that the *HR* algorithm with respect to J converges twice as fast as the *LR* algorithm if no shifts are used, because we find:

THEOREM 4.3. *For $J \in \mathrm{diag}_k^n(\pm 1)$ let $A \in \mathrm{GL}_n(\mathbf{C})$ be J-Hermitian. Let the HR algorithm with respect to J be denoted by*

$$A_1 = A, \qquad J_1 = J,$$

$$A_i = H_i R_i,$$

$$A_{i+1} = R_i H_i, \qquad J_{i+1} = H_i^* J_i H_i,$$

and the LR algorithm by

$$\tilde{A}_1 = A,$$

$$\tilde{A}_i = L_i \tilde{R}_i,$$

$$\tilde{A}_{i+1} = \tilde{R}_i L_i.$$

If the LR algorithm is constructible, then so is the HR algorithm with respect to J, and for each $i \in \mathbf{N}$ there exists a diagonal matrix $D_i \in \mathrm{GL}_n(\mathbf{C})$ such that

$$A_{i+1} = D_i^{-1} \tilde{A}_{2i+1} D_i.$$

Proof. If the LR algorithm is constructible, which means that for each iterate the LR decomposition exists, then there also exists the LR decomposition of A^m for all $m \in \mathbf{N}_+$, namely

$$A^m = L_1 \cdots L_m \tilde{R}_m \cdots \tilde{R}_1. \tag{4.3.1}$$

We prove the statement by induction taking advantage of the fact that for all $m \in \mathbf{N}$, A^m is J-Hermitian.

By (4.3.1) the LR decomposition of $A^*JA = JA^2$ exists, which according to Theorem 2.3 yields that the HR decomposition of $A_1 = A$ with respect to $J_1 = J$ exists. Now assume that for $i_0 \in \mathbf{N}$ we have the following: for all $i \in \{1, \dots, i_0\}$ the HR decomposition of A_i with respect to J_i exists. Then for $i \in \{1, \dots, i_0\}$, by (4.3.1) we find

$$A^{2i} = L_1 \cdots L_{2i} \tilde{R}_{2i} \cdots \tilde{R}_1,$$

and further

$$A^i = H_1 \cdots H_i R_i \cdots R_1.$$

Because A^i is J-Hermitian and $H_1 \cdots H_i \in U_n(J, J_{i+1})$,

$$JA^{2i} = A^{i*} J A^i = R_1^* \cdots R_i^* J_{i+1} R_i \cdots R_1.$$

Hence A^{2i} has a decomposition

$$A^{2i} = JR_1^* \cdots R_i^* J_{i+1} R_i \cdots R_1$$

into a product of a lower and an upper triangular matrix.

Now if D_i is the diagonal matrix for which $JR_1^* \cdots R_i^* J_{i+1} D_i^{-1}$ has a unit diagonal, then

$$A^{2i} = JR_1^* \cdots R_i^* J_{i+1} D_i^{-1} D_i R_i \cdots R_1,$$

or

$$A^{2i} = L_1 \cdots L_{2i} \tilde{R}_{2i} \cdots \tilde{R}_1,$$

is the LR decomposition of A^{2i}. Because the LR decomposition is unique, this yields

$$JR_1^*\cdots R_i^* J_{i+1} D_i^{-1} = L_1\cdots L_{2i}.$$

Therefore we have

$$\begin{aligned} D_i^{-1}\tilde{A}_{2i+1}D_i &= D_i^{-1}L_{2i}^{-1}\cdots L_1^{-1}AL_1\cdots L_{2i}D_i \\ &= J_{i+1}R_i^{-*}\cdots R_1^{-*}JAJR_1^*\cdots R_i^*J_{i+1} \\ &= J_{i+1}R_i^{-*}\cdots R_1^{-*}A^*R_1^*\cdots R_i^*J_{i+1} \\ &= J_{i+1}A_{i+1}^*J_{i+1} = A_{i+1}. \end{aligned}$$

It remains to prove that the HR decomposition of A_{i_0+1} with respect to J_{i_0+1} exists. From the last equality we get in particular $A_{i_0+1} = D_{i_0}^{-1}\tilde{A}_{2i_0+1}D_{i_0}$. Now $A_{i_0+1}^*J_{i_0+1}A_{i_0+1} = J_{i_0+1}A_{i_0+1}^2 = J_{i_0+1}D_{i_0}^{-1}\tilde{A}_{2i_0+1}^2D_{i_0}$, and because

$$\begin{aligned} \tilde{A}_{2i_0+1}^2 &= L_{2i_0+1}\tilde{R}_{2i_0+1}L_{2i_0+1}\tilde{R}_{2i_0+1} = L_{2i_0+1}\tilde{A}_{2i_0+2}R_{2i_0+1} \\ &= L_{2i_0+1}L_{2i_0+2}\tilde{R}_{2i_0+2}R_{2i_0+1}, \end{aligned}$$

we find that the LR decomposition of $A_{i_0+1}^*J_{i_0+1}A_{i_0+1}$ exists. According to Theorem 2.3 this completes the proof. ■

In particular, for $J=I$ or $J=-I$, i.e. for the QR algorithm applied to Hermitian matrices, this theorem gives a generalization of the relationship between the QR and the Cholesky LR algorithm pointed out by Wilkinson [14, p. 545].

It is easy to construct J-Hermitian matrices for which the HR algorithm with respect to J is constructible but the LR algorithm is not, even if $J\neq \pm I$. If for example we apply the LR algorithm to the $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$-symmetric matrix

$$A = \begin{bmatrix} 3 & -1 \\ 1 & 21 \end{bmatrix},$$

we see that the third iterate

$$A_3 = \begin{bmatrix} 0 & -1 \\ 64 & 24 \end{bmatrix}$$

has no LR decomposition.

Now if we examine how the *HR* algorithm with respect to $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ works on A, we find that the first diagonal element a_{i+1} of the $(i+1)$th iterate A_{i+1} satisfies

$$a_{i+1}=a_i-\frac{24(a_i^2-24a_i+64)}{24a_i-64}.$$

The quantity $|24a_i-64|$ is the absolute value of the first principal minor of $A_i^T J_i A_i$, and because $a_1=3$ it can be shown by studying the recurrence relation for a_i that this value does not vanish. According to Theorem 2.3 this means that the algorithm is constructible.

REFERENCES

1 M. A. Brebner, J. Grad, and J. Vrečko, Eigenvalues of $Ax=\lambda Bx$ for real symmetric matrices A and B computed by the *HR* process, Research Paper No. 238, Dept. of Mathematics, Statistics and Computing Science, Univ. of Calgary, July 1974.

2 M. A. Brebner and J. Grad, Similarity transformations for pseudosymmetric matrices with particular reference to the *HR* method, Research Paper No. 245, Dept. of Mathematics, Statistics and Computing Science, Univ. of Calgary, Aug. 1974.

3 M. A. Brebner and J. Grad, The general form and certain properties of similarity transformations for pseudo-symmetric matrices, Research Paper No. 281, Dept. of Mathematics, Statistics and Computing Science, Univ. of Calgary, June 1975.

4 A. Bunse-Gerstner, Der *HR*-Algorithmus zur numerischen Bestimmung der Eigenwerte einer Matrix, Dissertation, Univ. Bielefeld, 1978.

5 J. Della-Dora, Sur quelques algorithmes de recherche de valeurs propres, Thèse, L'université scientifique et medicale de Grenoble, July 1973.

6 L. Elsner, On some algebraic problems in connection with general eigenvalue algorithms, *Linear Algebra and Appl.*, to appear.

7 F. G. Frobenius, Über die mit einer Matrix vertauschbaren Matrizen, *S.-B. Kgl. Preuss. Akad. Wiss.*: 3–15 (1910).

8 C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards* 45: 255–282 (1950).

9 S. Lang, *Algebra*, Addison-Wesley, Reading, Mass., 1965, pp. 374–375.

10 B. N. Parlett and W. G. Poole, A geometric theory for the *QR*, *LU* and power method, *SIAM J. Numer. Anal.* 10: 389–412 (1973).

11 H. Rutishauser, Beiträge zur Kenntnis des Biorthogonalisierungs-Algorithmus von Lanczos, *Z. Angew. Math. Phys.* 4: 35–56 (1953).

12 G. W. Stewart, *Introduction to Matrix Computations*, Academic, New York, 1973, pp. 352–353.

13 O. Taussky and H. Zassenhaus, On the similarity transformation between a matrix and its transpose, *Pacific J. Math*. 9: 893–896 (1959).

14 J. H. Wilkinson, *The Algebraic Eigenvalue problem*, Clarendon Press, Oxford, 1967.

Received 3 August 1979; revised 19 December 1979