

COMPUTING THE JORDAN STRUCTURE OF AN EIGENVALUE

NICOLA MASTRONARDI* AND PAUL VAN DOOREN†

Abstract. In this paper we revisit the problem of finding an orthogonal similarity transformation that puts an $n \times n$ matrix A in a block upper-triangular form that reveals its Jordan structure at a particular eigenvalue λ_0 . The obtained form in fact reveals the dimensions of the null spaces of $(A - \lambda_0 I)^i$ at that eigenvalue via the sizes of the leading diagonal blocks, and from this the Jordan structure at λ_0 is then easily recovered. The method starts from a Hessenberg form that already reveals several properties of the Jordan structure of A . It then updates the Hessenberg form in an efficient way to transform it to a block-triangular form in $\mathcal{O}(mn^2)$ floating point operations, where m is the total multiplicity of the eigenvalue. The method only uses orthogonal transformations and is backward stable. We illustrate the method with a number of numerical examples.

Key words. Jordan structure, staircase form, Hessenberg form

March 4, 2019

1. Introduction. Finding the eigenvalues and their corresponding Jordan structure of a matrix A is one of the most studied problems in numerical linear algebra. This structure plays an important role in the solution of *explicit* differential equations, which can be modeled as

$$\lambda x(t) = Ax(t), \quad x(0) = x_0, \quad A \in \mathbb{R}^{n \times n} \quad (1.1)$$

where λ stands for the differential operator. The structure of the solutions of (1.1) depend heavily on the Jordan structure of A at each of its eigenvalues. A particular point is the origin, since if A has an eigenvalue at $\lambda_0 = 0$, then its Jordan structure also plays an important role in the construction of the Drazin inverse. We refer to [?, Chap. 4] and [5] for further details. When considering systems of *implicit* differential equations, modeled as

$$\lambda Bx(t) = Ax(t), \quad x(0) = x_0, \quad A, B \in \mathbb{R}^{n \times n} \quad (1.2)$$

one can reduce the problem to the explicit system of equations $\lambda x(t) = B^{-1}Ax(t)$ if $\det(B) \neq 0$ and then the Jordan structure of the eigenvalues of $B^{-1}A$ play a similar role. But one often imposes the weaker assumption that the system is *regular*, i.e. $\det(\lambda B - A) \neq 0$ for at least one value of λ . In this case one defines the *generalized* eigenvalues as the roots of the polynomial $\det(\lambda B - A)$ and the Jordan form is then replaced by the so-called Weierstrass form, which also allows to define the structure of the eigenvalue $\lambda_0 = \infty$. The structure of the eigenvalue at $\lambda_0 = \infty$ corresponds to the so-called *impulsive* solutions of (1.2) and the so-called *index* is the size of the corresponding largest Jordan block. We refer to [7, Chap. 7], [14, Chap. 2] for further details.

*Istituto per le Applicazioni del Calcolo “M. Picone”, Consiglio Nazionale delle Ricerche, sede di Bari, Italy (n.mastronardi@ba.iac.cnr.it). The work of this author is partly supported by GNCS-INDAM. This author would like to acknowledge the hospitality of the the Catholic University of Louvain-la-Neuve, Belgium, where part of this work took place.

†Department of Mathematical Engineering, Catholic University of Louvain, Louvain-la-Neuve, Belgium (paul.vandooren@uclouvain.be). The work of this author is partly supported by the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office and by CNR under the Short Term Mobility Program. The scientific responsibility rests with its authors.

Both these problems are of course intimately related, hence this work will focus on the Jordan structure of a single matrix A and will indicate how to extend this to the more general cases.

Let us suppose that the matrix A has $\lambda_0 = 0$ as an eigenvalue and define the spaces \mathcal{N}_i as the null spaces of the powers A^i . Clearly these null spaces are nested and one defines the *index* k of this eigenvalue, as the smallest integer i for which the dimensions $n_k := \dim(\mathcal{N}_k)$ of these spaces do not change anymore :

$$\mathcal{N}_1 \subset \mathcal{N}_2 \subset \cdots \subset \mathcal{N}_k = \mathcal{N}_{k+1} \quad (1.3)$$

$$n_1 < n_2 < \cdots < n_k = n_{k+1}. \quad (1.4)$$

The dimensions n_i of these spaces can be shown to define uniquely the Jordan structure at the eigenvalue $\lambda_0 = 0$. A proof of this follows implicitly from the block triangular forms constructed by Kublanovskaya [12], Ruhe [15] and Golub&Wilkinson [9] and which were generalized to so-called *staircase forms* in [17]. Their basic result is the construction of a unitary matrix V partitioned as

$$V = [V_1 \mid V_2 \mid \cdots \mid V_k \mid V_{k+1}] \quad (1.5)$$

where

$$\mathcal{N}_i = \text{Im}([V_1 \mid V_2 \mid \cdots \mid V_i]), \quad i = 1, \dots, k$$

i.e. V_i completes the orthogonal basis of \mathcal{N}_{i-1} to an orthogonal basis for the larger space \mathcal{N}_i . If we define the integers r_i as the increments of the dimensions n_i :

$$r_1 = n_1, \quad r_i = n_i - n_{i-1}, \quad i = 2, \dots, k, \quad (1.6)$$

then this unitary matrix V transforms A to the following staircase form

$$\tilde{A} = V^T A V = \left[\begin{array}{cccc|cc} 0_{r_1} & \tilde{A}_{1,2} & \cdots & \tilde{A}_{1,k-1} & \tilde{A}_{1,k} & \tilde{A}_{1,k+1} \\ 0 & 0_{r_2} & \tilde{A}_{2,3} & \cdots & \tilde{A}_{2,k} & \tilde{A}_{2,k+1} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & 0_{r_{k-1}} & \tilde{A}_{k-1,k} & \tilde{A}_{k-1,k+1} \\ 0 & 0 & \cdots & \cdots & 0_{r_k} & \tilde{A}_{k,k+1} \\ \hline 0 & 0 & 0 & \cdots & 0 & \tilde{A}_{k+1,k+1} \end{array} \right] \quad (1.7)$$

where

- the diagonal blocks 0_{r_i} of \tilde{A} are square and of dimension r_i , for $i = 1, \dots, k$,
- the blocks $\tilde{A}_{i-1,i}$ are of full column rank r_i , for $i = 2, \dots, k$,
- the block $\tilde{A}_{k+1,k+1}$ is nonsingular (provided it is not empty).

It follows also from these properties that the index set $\{r_1, \dots, r_k\}$ (also known as the Weyr characteristic of A at $\lambda_0 = 0$ [16], [19, Chap. 2]) is non-increasing :

$$r_1 \geq r_2 \geq \cdots \geq r_k.$$

These indices then define the Jordan structure of the corresponding eigenvalue via the following rules :

there are exactly $r_i - r_{i+1}$ Jordan blocks of size i for $i = 1, \dots, k$

where we have assumed $r_{k+1} = 0$.

The original papers [12, 15] for computing the above form for a given eigenvalue (say, $\lambda_0 = 0$), had a worst-case complexity of $\mathcal{O}(rn^3)$ flops for $r = \sum_{i=1}^k r_i$ recovered eigenvalues, where unitary transformations were used throughout the calculations, which implied that the computed form corresponded exactly to a slightly perturbed matrix A . Software implementing this decomposition was also provided in [11]. This procedure was extended in [17] and later in [6] to cover also the Jordan structure of a regular pencil $\lambda B - A$, but there also the complexity was of the same order for r recovered generalized eigenvalues. This was because all these methods used full matrix decompositions for each rank determination, and there could possibly be $\mathcal{O}(r)$ such rank tests during the process of constructing the staircase form (1.7). The first algorithm that had at most cubic complexity was proposed by Golub&Wilkinson [9] but it used non-orthogonal bases for the representation of some intermediate calculations, therefore precluding the same kind of backward stability result. The same issue also applies to the method of Anstreicher&Rothblum [1], which uses Gauss-Jordan elimination, which is potentially unstable. The first backward stable methods with cubic complexity were proposed in [2], [3], for the general type of staircase form as well as for the special case of Jordan canonical form described above. These algorithms used either updating of echelon forms or updating of QR decompositions to maintain cubic complexity. More recently, Guglielmi et al. [10] revisited this problem and produced a new algorithm of cubic complexity. They also did extensive testing to illustrate the performance of their algorithm and discussed the sensitivity of the decomposition.

In this paper we propose yet another algorithm of overall cubic complexity, which has several advantages. The starting point is a special Hessenberg form, obtained in $10n^3/3$ flops, which turns out to have a number of additional properties, when the matrix has several Jordan blocks at the same eigenvalue (the so-called *derogatory case*). These properties are then exploited to yield a method of $\mathcal{O}(rn^2)$ complexity to retrieve the staircase form at a given eigenvalue. The resulting form can be exploited to compute the Jordan structure at any other given eigenvalue, again with $\mathcal{O}(\hat{r}\hat{n}^2)$ complexity, where now $\hat{n} = (n - r)$ and \hat{r} is the multiplicity of this second eigenvalue. Continuing like this for all eigenvalues of A , yields eventually a staircase form in $\mathcal{O}(n^3)$ flops.

For the sake of simplicity, we consider only real matrices. The extension to complex matrices is straightforward. We will use the following notations. Matrices and submatrices are denoted by capital letters, i.e., A, B, H . The entry (i, j) of the matrix A is denoted by the lower case letter $a_{i,j}$. Vectors are denoted by bold letters, i.e., $\mathbf{a}, \mathbf{b}, \dots$. The identity matrix of order n is denoted by I_n and its i -th column by $\mathbf{e}_i^{(n)}$, or, if there is no ambiguity, simply by I and \mathbf{e}_i , respectively. Generic entries different from zero in matrices or vectors are denoted by “ \times ”. The machine precision is denoted by ε . We denote a Givens rotation by

$$G_i = \begin{bmatrix} I_{i-1} & & & & \\ & c & -s & & \\ & s & c & & \\ & & & I_{n-i-1} & \\ & & & & \end{bmatrix}, \quad \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix}^T = I_2.$$

The entries to be annihilated by the product of a matrix or a vector by a Givens rotations will be deduced from the context and graphically denoted by the symbol “ \otimes ” in the pictures.

The paper is organized as follows. The new algorithm is described and analyzed in Section 2. A number of numerical experiments illustrating the accuracy of our method are reported in Section 3. We then end the paper with a number of concluding remarks.

2. A new algorithm. In this section we develop the new algorithm for finding the Weyr characteristic at a particular eigenvalue λ_0 . Without loss of generality, we will assume that $\lambda_0 = 0$, because it simplifies the description of the algorithm.

The *first step* of the algorithm consists in a *reduction to Hessenberg form*. Since the matrix we are interested in has a non-trivial Weyr characteristic, it must have several eigenvalues at the origin. In this case, we will see that this preliminary reduction almost always yields a Hessenberg form with several smaller unreduced blocks and we will show theoretically that the sizes of these unreduced blocks can be deduced in terms of the Jordan structure of the matrix. This preliminary reduction is computationally important since it significantly reduces the computational complexity of the algorithm.

At the *second step*, we check if each irreducible Hessenberg block is singular with a technique inspired by Guglielmi et al. [10]. Then, in each singular block the eigenvalue $\lambda_0 = 0$ is moved to the top of each Hessenberg block by a backward variant of the *QR*-step with perfect shift $\lambda_0 = 0$, called in the sequel “*backward QR-step*”, requiring first the computation of the corresponding eigenvector in order to guarantee backward stability. For the full analysis of this backward *QR*-step with perfect shift, we refer to another paper [13].

The *third step* consists of moving all deflated zero eigenvalues to the top of the matrix by a special echelon reduction. That part is strongly related to the results described in [2], [3]. An important difference here is that we show how to exploit the Hessenberg form of the different blocks in order to keep the complexity as low as possible. At the end of the third step, the computed matrix H has a block Hessenberg form with the first r_1 columns zeroed, where r_1 is the first element of the Weyr characteristic. The whole procedure is then recursively applied to $H(r_1 + 1 : n, r_1 + 1 : n)$ until the considered submatrix is either nonsingular or empty.

2.1. Hessenberg reduction. The Hessenberg reduction of a square matrix is the first step of the algorithm. In this section we analyze the Hessenberg form of a matrix $A \in \mathbb{R}^{n \times n}$ with given Jordan structure at the origin.

THEOREM 2.1. *Let $m(\lambda) = \prod_{i=1}^{\ell} (\lambda - \lambda_i)^{k_i}$ be the monic minimal polynomial of a matrix A , and let $d := \sum_{i=1}^{\ell} k_i$ be its degree. Then the Krylov subspace $\mathcal{K}_k(A, b) = \text{Im} [b, Ab, \dots, A^{k-1}b]$ has dimension bounded by $\min(k, d)$. Moreover*

- (i) *this upper bound is reached for almost any vector b , i.e. it is generic,*
- (ii) *such a Krylov subspace of maximal dimension d is an invariant subspace of A corresponding to a largest Jordan block of each eigenvalue, and the vector b is its cyclic generator¹.*

Proof. It is well known that since $m(\lambda)$ has degree d , then $A^d = \sum_{i=0}^{d-1} c_i A^i$. Therefore $A^{d+j}b$ (for all $j \geq 0$) is a linear combination of the previous vectors in the Krylov subspace, and, by induction, also of $\mathcal{K}_d(A, b)$. We use this to prove the genericity results (i) and (ii) for a matrix A which is nilpotent (no eigenvalues except $\lambda = 0$). For this special case we can give a simple algebraic characterization of the genericity condition. We can assume without loss of generality that the matrix A is in

¹A vector b is a cyclic generator if $\mathcal{K}_k(A, b)$ is of full dimension [20, Chap. 0].

its staircase form \tilde{A} (provided we transform the vector b as well), since the dimension of the Krylov subspace is invariant to similarity transformations.

$$\tilde{A} = \begin{bmatrix} 0_{r_1} & \tilde{A}_{1,2} & \cdots & \tilde{A}_{1,k-1} & \tilde{A}_{1,k} \\ 0 & 0_{r_2} & \tilde{A}_{2,3} & \cdots & \tilde{A}_{2,k} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & 0_{r_{k-1}} & \tilde{A}_{k-1,k} \\ 0 & 0 & \dots & \cdots & 0_{r_k} \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \vdots \\ \tilde{b}_k \end{bmatrix}.$$

In that coordinate system, any vector $\tilde{b} = V^T b$ that has a nonzero subvector \tilde{b}_k in its last r_k components, will yield a Krylov subspace that grows by one at each step to reach full dimension k . This condition can also be written as $\tilde{A}^{k-1} \tilde{b} \neq 0$ or, equivalently, $A^{k-1} b \neq 0$, where k is the degree of the minimal polynomial. Notice that meanwhile we also proved (ii) for a nilpotent matrix, since the Jordan form of A , restricted to $\mathcal{K}(A, b)$, is a single Jordan block of length k . That follows from the fact that the dimension of the Krylov subspace is k .

If we now consider a matrix with several different eigenvalues λ_i , the argument could be repeated for each eigenvalue, but we give here a shorter geometric proof. We want to show that in the generic case, $\dim \mathcal{K}_j(A, b) = \min(j, d)$. If this is *not* the case, then there must exist an index $j < d$ such that $\dim \mathcal{K}_j(A, b) = \dim \mathcal{K}_{j+1}(A, b)$. Then $\mathcal{K}_j(A, b)$ is an invariant subspace of A and b belongs to it. Let V be an $n \times j$ matrix spanning $\mathcal{K}_j(A, b)$, then there exist a matrix $\tilde{A} \in \mathbb{R}^{j \times j}$ and a vector $\tilde{b} \in \mathbb{R}^j$ such that

$$AV = V\tilde{A}, \quad b = V\tilde{b}.$$

Let $\tilde{m}(\lambda)$ be the monic minimal polynomial of \tilde{A} . Then it follows that

$$\tilde{m}(\tilde{A}) = 0 \implies \tilde{m}(A)V = V\tilde{m}(\tilde{A}) = 0 \implies \tilde{m}(A)V\tilde{b} = \tilde{m}(A)b = 0.$$

So b must belong to the kernel of $\tilde{m}(A)$ for all monic polynomials $\tilde{m}(\lambda)$ of degree j that divide $m(\lambda)$. There is a finite number of such polynomials and the matrices $\tilde{m}(A)$ are non-zero since otherwise $m(\lambda)$ would not be the minimal polynomial of A . Therefore, a random vector b in $\mathcal{K}_j(A, b)$ almost always satisfies $\tilde{m}(A)b \neq 0$ for all $\tilde{m}(\lambda)$. The generic result is thus that the minimal polynomial $m(\lambda)$ of A is also the minimal polynomial of A restricted to $\mathcal{K}_d(A, b)$. Since $\dim \mathcal{K}_j(A, b) = j$ for $j \leq d$ the vector b is a cyclic generator of $\mathcal{K}_d(A, b)$ and the corresponding restriction can have only one Jordan block for each eigenvalue λ_i . \square

We illustrate this Theorem with a simple example.

EXAMPLE 2.1. *Let the Jordan structure of A be described by its list of Jordan blocks $J_i(\lambda_j)$ of size i at eigenvalue λ_j and the corresponding Weyr characteristic at eigenvalue λ_i :*

<i>Eigenvalue</i>	<i>Jordan blocks</i>	<i>Weyr characteristic</i>
$\lambda_1 = 0$	$J_4(0), J_2(0), J_1(0)$	$3, 2, 1, 1$
$\lambda_2 = 1$	$J_3(1)$	$1, 1, 1$
$\lambda_3 = 2$	$J_2(2), J_1(2)$	$2, 1$

For this structure the generic Krylov subspace will have dimension $d = 4 + 3 + 2 = 9$ and the Jordan structure of A restricted to that eigenspace will be $J_4(0), J_3(1), J_2(2)$. If we now apply the Hessenberg reduction with a starting vector b that is random,

we will deflate an invariant subspace with that corresponding eigenstructure. The remaining part of the matrix A will therefore have a Jordan structure corresponding to the leftover blocks, i.e. $J_2(0), J_1(0), J_1(2)$. \square

COROLLARY 2.2. Let A be an arbitrary matrix with minimal polynomial $m(\lambda) = \prod_{i=1}^{\ell} (\lambda - \lambda_i)^{k_i}$ of degree $d := \sum_{i=1}^{\ell} k_i$. Then for almost any “seed vector” b , the Hessenberg reduction of A yields an irreducible deflated matrix whose Jordan blocks are $J_{k_i}(\lambda_i)$:

$$\hat{A} = V^T A V = \left[\begin{array}{c|c} rdH & \times \\ \hline 0 & B \end{array} \right], \quad \text{where} \quad H := \begin{bmatrix} \times & \cdots & & \cdots & \times \\ \times & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \times \\ 0 & \cdots & 0 & \times & \times \end{bmatrix}$$

is unreduced and has characteristic and minimal polynomial $m(\lambda)$.

Proof. The proof follows essentially from the point (ii) of the previous Theorem.

\square

If we now repeat the above procedure and choose a new seed for the remaining submatrix B then we find the following *generic* Hessenberg reduction,

$$\hat{A} = V^T A V = \begin{bmatrix} H_1 & \times & \cdots & \times \\ 0 & H_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ 0 & \cdots & 0 & H_k \end{bmatrix} \quad (2.1)$$

where each diagonal block H_j is an unreduced Hessenberg matrix and its Jordan blocks are the j -th Jordan block of each eigenvalue λ_i of A . Therefore, these matrices are of decreasing size. We revisit the above example, in order to clarify this.

EXAMPLE 2.2. For the matrix A given in Example 2.1 there would be three Hessenberg matrices H_j of dimension d_j and their respective Jordan structures would be

j	Jordan blocks of H_j	d_j
1	$J_4(0), J_3(1), J_2(2)$	9
2	$J_2(0), J_1(2)$	3
3	$J_1(0)$	1

\square

REMARK 2.1. We believe that the above result is new, although a similar Hessenberg reduction was obtained in [4], but based on the Euclidean algorithm and with no reference to genericity. The fact that a non-derogatory matrix almost always leads to a complete Krylov subspace, can be found in [20, Chap. 1], where it is also pointed out that the vector b is then a cyclic generator. But the recursive Hessenberg reduction and the link to the minimal polynomial are not in [20, Chap. 1].

REMARK 2.2. We observe that, from a numerical point of view, the Hessenberg reduction often is not able to reveal the structure of the matrix described in this paragraph, due to the ill-conditioning of the eigenvalues of the matrix. Nevertheless, the computed Hessenberg form will be the starting step of the proposed algorithm and the form (1.7) will be computed in the other steps.

The first step of our algorithm uses these results to compute a “backward” upper Hessenberg form $H = Q_H A Q_H^T$, with $Q_H \in \mathbb{R}^{n \times n}$ orthogonal, and H in upper Hessenberg form. This Hessenberg reduction differs from the usual one [8, p. 368] because it starts from the bottom–right corner of the matrix A and proceeds backward to the upper–left one. We choose such a reduction to Hessenberg form because it is more suitable to compute the form (1.7). This reduction is accomplished in $n - 2$ steps. In the first step, the entries $1, 2, \dots, n - 3, n - 2$ of row n are annihilated by the right-multiplication with a “reverse” Householder matrix Q_1 ; we then apply the similarity transformation $Q_1 A Q_1^T$. At the second step, we then construct the updated transformation $Q_2 Q_1 A Q_1^T Q_2^T$ with Q_2 a “reverse” Householder matrix chosen to annihilate the entries $1, 2, \dots, n - 3$, of the row $n - 1$. Proceeding in this way, at the step $n - 2$, the entry in position $(3, 1)$ is annihilated and we have the final transformation $H := Q_{n-2} \cdots Q_2 Q_1 A Q_1^T Q_2^T \cdots Q_{n-2}^T$.

We observe that the matrix H has the structure described in (2.1), with the sizes of the Hessenberg diagonal blocks in reversed order.

Computational complexity. The number of flops required to compute the backward Hessenberg reduction is $10n^3/3$, the same number of flops required by the “forward” Hessenberg reduction.

2.2. Deflating the 0 eigenvalues. As described above, the computed Hessenberg matrix H has the structure described in (2.1). The natural choice to check whether 0 is an eigenvalue of each H_j , $j = 1, \dots, k$, is to apply one step of the backward QR method with zero shift to each submatrix H_j . In exact arithmetic, if 0 is an eigenvalue of the irreducible submatrix $H_j \in \mathbb{R}^{n_j \times n_j}$, then one backward QR -step with zero shift, yields a new Hessenberg submatrix $\hat{H}_j = W_j H_j W_j^T$, with its first column zeroed [19, Chap. 4]. Notice that $W_j \in \mathbb{R}^{n_j \times n_j}$ is a product of $n_j - 1$ Givens rotations G_{n_j-i} , $i = 1, \dots, n_j - 1$, and is also in Hessenberg form.

Unfortunately, in finite precision this may not be the case anymore because of the effect known as “blurred shift” [18] and the ill-conditioning of the zero eigenvalue (see Example 2.3). The first column of the computed \hat{H}_j might therefore be far from the zero column and its norm will depend on the condition number of the zero eigenvalue.

Therefore, we need to consider an alternative approach. We compute the eigenvector corresponding to the smallest eigenvalue of H_j by inverse iteration. If 0 is an eigenvalue up to a fixed tolerance τ , one step of inverse iteration is enough to compute the corresponding eigenvector. Moreover, even though the submatrices H_j could be very ill-conditioned, it is well known that the errors of the computed eigenvector by inverse iteration are in the same direction as the true eigenvector [8, Chap. 7].

Let $(\lambda, \mathbf{x}, \mathbf{y})$ be the triple consisting of an eigenvalue and the corresponding right and left eigenvector of an irreducible Hessenberg matrix $H \in \mathbb{R}^{n \times n}$. In Theorem A.1, it is shown that one backward (forward) QR -step with perfect shift, i.e., with shift equal to λ , can be constructed considering the sequence of Givens rotations Q reducing the corresponding right (left) eigenvector \mathbf{x} (\mathbf{y}) to \mathbf{e}_1 (e_n), i.e., $Q\mathbf{x} = \pm\mathbf{e}_1$ ($Q\mathbf{y} = \pm\mathbf{e}_n$). Then, in exact arithmetic, $QH Q^T$ is still in upper Hessenberg form with the first column (last row) zeroed. We call this procedure modified backward (forward) QR -step. A modified backward QR -step is graphically depicted in Picture 2.2. We observe that, in floating point arithmetic, the resulting matrix after such a step may not be in Hessenberg form anymore and a $O(n^3)$ algorithm is proposed in [?] to recover the Hessenberg structure.

However, in Appendix A, we also show that, under very reasonable assumptions, this version of the so-called backward QR -step with (perfect) zero shift, produces the

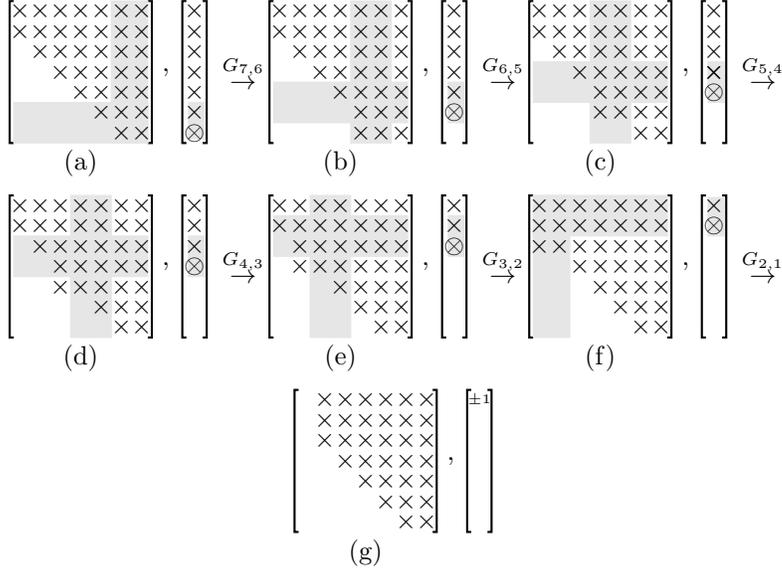


FIG. 2.1. Modification of the matrix and the eigenvector associated to the zero eigenvalue

desired result. The possible loss and recovery of the accuracy in computing the new Hessenberg matrix with the first column zeroed described in this section is shown in the following example.

EXAMPLE 2.3. Let $A \in \mathbb{R}^{18 \times 18}$ be the second of the two matrices considered in [10, Appendix B] arising in surface subdivision. The two smallest singular values, computed by `matlab` are $\sigma_{17} = 9.16 \times 10^{-11}$ and $\sigma_{18} = 2.57 \times 10^{-16}$ while the smallest eigenvalue λ_{18} computed by `eig` of `matlab` has modulus equal to 3.34×10^{-5} with eigenvalue condition number $\kappa(\lambda_{18}) = 1.8088 \times 10^{11}$.

Let \tilde{H} be the Hessenberg matrix computed by the backward Hessenberg reduction and \tilde{H}_1 the one obtained from \tilde{H} by applying one (standard) step of the backward QR algorithm with zero shift. In exact arithmetic, the first column of H_1 is $\mathbf{0}$. However, the norm of the first column of the computed \tilde{H}_1 is 6.25×10^{-2} . The matrix \tilde{H}_2 is then obtained applying one forward step of the QR method with zero shift to \tilde{H} . The norm of the last row of the upper Hessenberg matrix \tilde{H}_2 is 3.32×10^{-15} . Hence, by considering Theorem A.1 in a forward fashion, the left eigenvector, associated to the zero eigenvalue, is the vector \mathbf{e}_{18} , the 18-th vector of the canonical basis of \mathbb{R}^{18} . Therefore, $1/x_n^{(2)}$ gives an estimate of the condition number of $\lambda = 0$ [8, Ch. 7].

For this, we compute $\mathbf{x}^{(2)}$, the eigenvector associated to the smallest eigenvalue in modulus of \tilde{H}_2 , via inverse iteration. Its last two components are respectively, $x_{n-1}^{(2)} = -7.77 \times 10^{-2}$ and $x_n^{(2)} = -3.68 \times 10^{-12}$. This high sensitivity also explains the blurring of the shift in the usual QR step. On the other hand, it is shown in the Appendix A that if the subvector $[x_{n-1}^{(2)}, x_n^{(2)}]$ is not too small (e.g. bounded below by a number μ of the order of 1), then, forcing the part of the matrix below the first subdiagonal to be zero after a modified backward QR-step using the vector $\mathbf{x}^{(2)}$, will introduce a perturbation in the entries of the rest of the Hessenberg matrix of size ε/μ .

If the bound

$$\|[x_{n-1}^{(2)}, x_n^{(2)}]\|_2 > \mu \quad (2.2)$$

on the subvector of $\mathbf{x}^{(2)}$ is not met, we propose to perform a *forward QR*-step with zero shift in order to obtain a transformed eigenvector $\mathbf{x}^{(2)}$ for which the bound is met. We summarize this procedure in the following scheme.

```

Apply one step of inverse iteration to the matrices  $H_j$ ,  $j = 1, \dots, k$ .
if 0 is an eigenvalue and  $\mathbf{x}^{(j)}$  the corresponding eigenvector of  $H_j$ ,  $j \in \{1, \dots, k\}$ ,
  if  $\|[x_{n_j-1}^{(j)}, x_{n_j}^{(j)}]\|_2 > \mu$ ,
    apply the modified backward zero shift QR-step to  $H_j$  and  $\mathbf{x}^{(j)}$ 
  else
    apply one step of the QR method with zero shift to  $\hat{H}_j = W_j^T H_j W_j$ 
    compute  $\hat{\mathbf{x}}^{(j)} = W_j^T \mathbf{x}^{(j)}$ ;
    apply the modified backward zero shift QR-step to  $\hat{H}_j$  and  $\hat{\mathbf{x}}^{(j)}$ ;
  end if
end if

```

If 0 is an eigenvalue of $H_j \in \mathbb{R}^{n_j \times n_j}$, then at this stage, the computed \hat{H}_j is still in Hessenberg form with the first column zeroed.

If (2.2) is not satisfied, one can check whether

$$\|[y_{n-1}^{(2)}, y_n^{(2)}]\|_2 > \mu. \quad (2.3)$$

If the latter bound is satisfied one modified forward *QR*-step is applied in the same fashion. In this case, the new computed Hessenberg matrix has its last row negligible. Then with a technique similar to the one described in Section 2.3 we can compute a similar Hessenberg matrix with the first column negligible. When both (2.2) and (2.3) are not satisfied, a sufficiently large entry of $\mathbf{x}^{(2)}$ closer to $x_n^{(2)}$ is sought and moved to the last position of the vector, permuting the Hessenberg matrix accordingly. The modified backward *QR*-step on the latter vector will need to chase a larger bulge in the Hessenberg matrix in this case.

Computational complexity. For each matrix $H_j \in \mathbb{R}^{n_j \times n_j}$, $j = 1, \dots, k$, $\sum_{j=1}^k n_j = n$, the computation of the eigenvector corresponding to the zero eigenvalue requires $4n_j^2$ floating point operations. One (forward or backward) *QR*-step and the procedure described in the Appendix A requires $6n_j^2$ floating point operations. Moreover, the cost of updating the whole Hessenberg matrix by the product of $n_j - 1$ Givens rotations is $6n_j(n - n_j)$ floating point operations.

2.3. Echelon reduction and the Weyr characteristic. Let H be the block Hessenberg matrix computed after the previous two steps

$$H = \begin{bmatrix} H_k & \times & \cdots & \times \\ & H_{k-1} & \ddots & \vdots \\ & & \ddots & \times \\ & & & H_1 \end{bmatrix},$$

where $H_i \in \mathbb{R}^{n_i \times n_i}$, $i = 1, \dots, k$, and $\sum_{i=1}^k n_i = n$.

In this step the matrix H is column reduced to echelon form by an orthogonal similarity transformation. A detailed description of the reduction of a general matrix

to echelon form via orthogonal transformations can be found in [2]. Here we develop this for the special case of Hessenberg matrices.

Let us suppose H_1 is singular otherwise we can restrict ourselves to the reduction to echelon form of the submatrix

$$\begin{bmatrix} H_k & \times & \cdots & \times \\ & H_{k-1} & \ddots & \vdots \\ & & \ddots & \times \\ & & & H_2 \end{bmatrix}.$$

The aim is to first compress to the right the Hessenberg matrix H by right multiplication with a matrix Q_c^T made by the product of a sequence of Givens rotations,

$$\hat{H} = H Q_c^T. \quad (2.4)$$

If the first Weyr characteristic is r_1 , then the first r_1 column of \hat{H} will be zero.

This step is graphically displayed in Figure 2.3 for the matrix considered in the Example 3.2 after the first two steps, i.e., after the reduction to Hessenberg form and the deflation of the zero eigenvalue in each Hessenberg block along the diagonal. These Hessenberg blocks are displayed in Figure 2.3 (a) with blue boxes.

We now denote by $G_j^{(i)T}$ the Givens rotation acting on the columns j and $j+1$ of H annihilating the entry in position (j, i) in the product $H G_j^{(i)T}$. The red entries denote rank-one submatrices present or created during the computation. In particular, the submatrix of indexes $(1 : 2, 2 : 3)$ is already of rank-one from the beginning of this step, while the submatrix of indexes $(3 : 4, 4 : 5)$ is of rank-one after the multiplication by $G_5^{(5)T}$ ((g) \rightarrow (h)). Moreover, the submatrix $(1 : 2, 2 : 4)$ becomes of rank-one after the multiplication with $G_4^{(4)T}$ ((h) \rightarrow (i)). When an entry of a rank-one submatrix is annihilated, the other entries in the same column of the submatrix are annihilated too ((h) \rightarrow (i), (i) \rightarrow (j) and (j) \rightarrow (k)). After this transformation the first r_1 columns are zeroed, where r_1 is the first Weyr characteristic.

Let Q_c^T be the product of all the Givens matrices applied to the right of H . Then the Hessenberg structure is restored multiplying (2.4) to the left by Q_c obtaining

$$\hat{H}^{(2)} = Q_c H Q_c^T, \quad (2.5)$$

such that

$$\hat{H}^{(2)}[\mathbf{e}_1, \dots, \mathbf{e}_{r_1}] = \mathbf{0}.$$

This step is graphically displayed in Figure 2.3 for the matrix considered in the Example 3.2.

After computing $\hat{H}^{(2)}$ in (2.5), the whole procedure is recursively applied to

$$H^{(2)} = \hat{H}^{(2)}(r_1 + 1 : n, r_1 + 1 : n)$$

(the part below and to the right of the red cross in Figure 2.3 (h)), until $H^{(k)}$, for a $k \geq 1$ is either empty or nonsingular.

The blue boxes displayed in Figure 2.3 (h) denote the block Hessenberg submatrices of $H^{(2)}$, whose sizes are obtained by subtracting one from those of the diagonal blocks of H in Figure 2.2 (a).

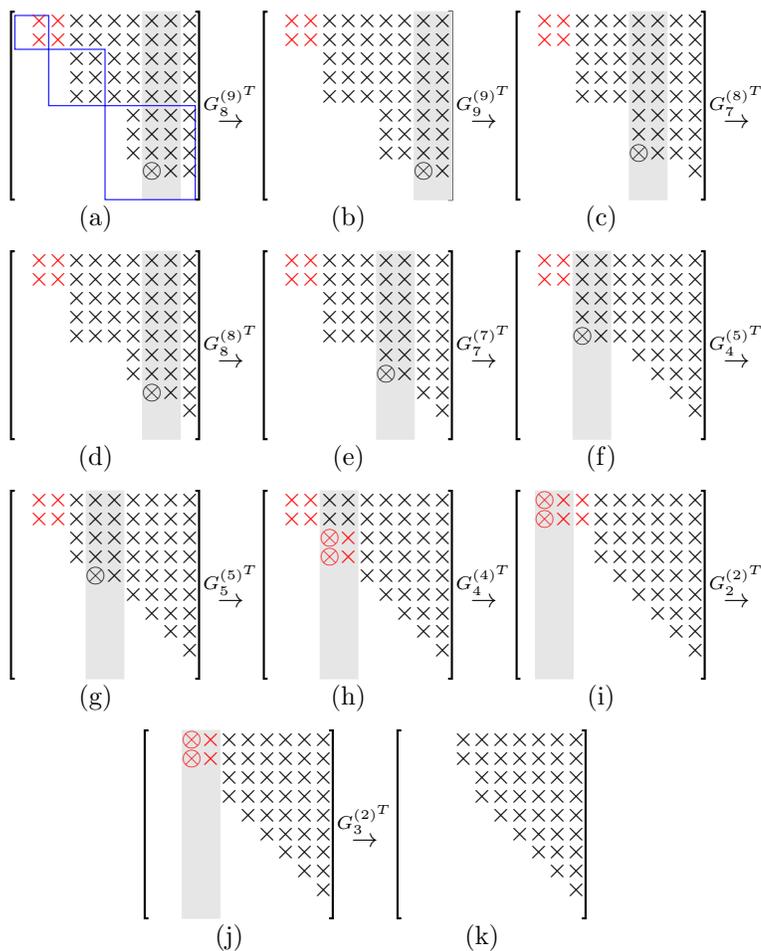


FIG. 2.2. Reduction of the Hessenberg matrix of Example 3.2 to column echelon form by a sequence of Givens rotations applied to the right.

Computational complexity. Suppose all the Hessenberg block H_i , $i = 1, \dots, k$, are singular. Then the number of Givens rotations in order to compute \hat{H} in (2.4) is less than $2 \sum_{i=1}^k n_i$. Hence, about $6n^2$ floating point operations are required to compute \hat{H} and about the same number of floating point operations are required to compute $\hat{H}^{(2)}$ in (2.5).

3. Numerical experiments. In this section we consider some numerical examples. The described algorithm, denoted in the sequel by **MVD**, was implemented in **matlab** and compared to the one described in [10], denoted in the sequel by **GOS**, and available at epubs.siam.org/doi/suppl/10.1137/140956737.

Let $A = Q_1 B_1 Q_1^T$ and $A = Q_2 B_2 Q_2^T$ be the decomposition computed by **GOS** and **MVD**, respectively. The tolerance considered in all the examples is $\tau = 10^{-13}$.

EXAMPLE 3.1. In this example we construct a matrix with the same Weyr char-

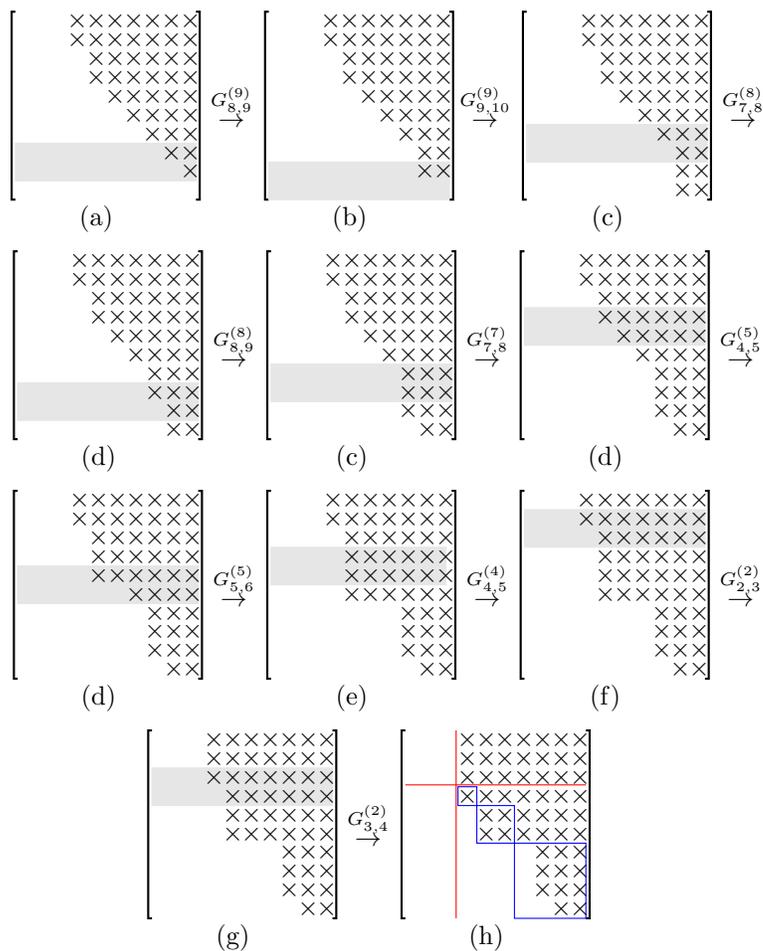


FIG. 2.3. Restoring the Hessenberg matrix from the column reduced echelon form.

acteristic of the matrix of Example 2.2. Therefore, the considered matrix

$$A_0 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix},$$

A is given by $A = QA_0Q^T$, where $Q \in \mathbb{R}^{n \times n}$ is an orthogonal random matrix generated by the matlab function gallery('qmult', n) with $n = 13$.

	GOS	MVD
$\frac{\ A-QBQ^T\ _2}{\ A\ _2}$	9.34×10^{-16}	1.66×10^{-15}

TABLE 3.1

Relative residual of the factorizations computed by GOS (left) and MVD (right).

The relative residuals of the two considered factorizations are displayed in Table 3.1. Both algorithms yield the same Weyr characteristic, i.e., $[3, 2, 1, 1]$.

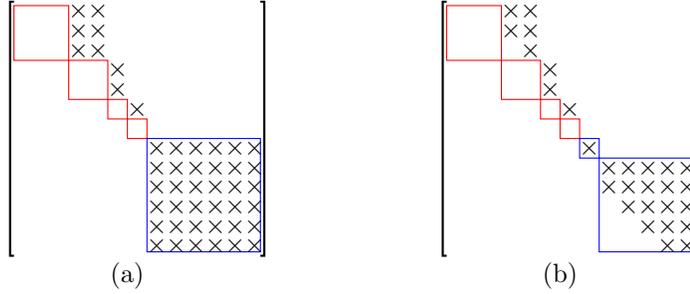


FIG. 3.1. Nonzero structure of the computed matrices in Example 3.1 by GOS (a) and by MVD (b).

The nonzero structure of the matrices computed by GOS and MVD is depicted in Figure 3.1 ((a) and (b), respectively). The red boxes in the pictures denote the Weyr structure, while the blue boxes denote the remaining diagonal block structure, which is Hessenberg for MVD.

EXAMPLE 3.2. Let $A \in \mathbb{R}^{10 \times 10}$ be the first of the two matrices considered in [10, Appendix B.1]. Such a matrix arises in the design of surfaces using subdivision algorithms [10]. This matrix is singular and the zero eigenvalue has multiplicity 4, with one Jordan block of order 2 and two of order 1. Moreover, the Weyr characteristic at the zero eigenvalue is $[3, 1]$.

	GOS	MVD
$\frac{\ A-QBQ^T\ _2}{\ A\ _2}$	5.75×10^{-16}	9.14×10^{-16}

TABLE 3.2

Relative residual of the factorizations computed by GOS (left) and MVD (right).

EXAMPLE 3.3. The matrix $A \in \mathbb{R}^{18 \times 18}$ of this example was considered in [10, Appendix B.1]. The relative residuals and nonzero structure of the matrices computed by GOS and MVD and they are depicted in Figure 3.3 and Table 3.3, respectively. The red box in the pictures denote the Weyr structure, while the blue box denotes the other diagonal block, which is Hessenberg for MVD. Both methods compute the same Weyr characteristic equal to $[1]$.

REMARK 3.1. A phenomenon called the “ κ^2 effect” affects the algorithm GOS presented in [10]. In particular, given $A = XJX^{-1} + E \equiv A_{true} + E$, where J is the original Jordan structure, X a nonsingular matrix with condition number κ , and E is the error matrix such that $\|E\| = \rho \|A_{true}\|$. Since $\|J\| = 1$, then $\|A_{true}\| \leq \|X\| \|X^{-1}\| = \kappa$. If now one sets $X^{-1}AX = J + F$, it follows that $F = X^{-1}EX$.

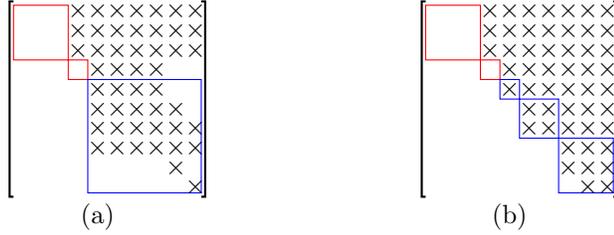


FIG. 3.2. Nonzero structure of the computed matrices in Example 3.2 by GOS (a) and by MVD (b).

	GOS	MVD
$\frac{\ A - QBQ^T\ _2}{\ A\ _2}$	5.80×10^{-16}	9.93×10^{-16}

TABLE 3.3

Relative residual of the factorizations computed by GOS (left) and MVD (right).

Therefore

$$\|F\| \leq \kappa \|E\| \leq \kappa^2 \rho.$$

The same analysis applies to the algorithm presented in this paper and a similar behavior as the GOS algorithm was noticed in the performed numerical experiments. We therefore omit the details here.

4. Conclusions. In this paper we discussed the problem of finding an orthogonal similarity transformation that puts an $n \times n$ matrix A in a block upper-triangular form revealing its Jordan structure at the eigenvalue $\lambda = 0$. But clearly, the method can be applied to any particular eigenvalue of the original matrix A , by just a shift to $A - \lambda_0 I$. The proposed numerical method yields a transformed matrix whose nonzero structure reveals the dimensions of the null spaces of $(A - \lambda_0 I)^i$ at that eigenvalue via the sizes of the leading diagonal blocks. The size of these blocks is the Weyr characteristic of the matrix A , which also reveals the Jordan structure of A at λ_0 . The method only uses orthogonal transformations and is backward stable with $O(n^3)$ as worst case computational cost. The method compares favourably with earlier methods because it reduces first the matrix to Hessenberg form, which is then exploited in the subsequent steps to minimize the computational cost.

Even though the method was applied here to the standard eigenvalue problem $\lambda I - A$ only, it can easily be extended to the generalized eigenvalue problem $\lambda B - A$, provided the pencil $(\lambda B - A)$ is regular, i.e. it has a determinant that is not identically zero. In this case one performs orthogonal transformations Q and Z that transform the pencil to $(\lambda \hat{B} - \hat{A}) = Q^T (\lambda B - A) Z$ where \hat{B} is triangular and \hat{A} is in a staircase form that reveals the Weyr characteristic of the matrix $\hat{B}^{-1} \hat{A} = Z^T (B^{-1} A) Z$, provided B is non-singular. But even when B is singular, this procedure can be applied to a regular pencil $\lambda B - A$ in the same vein as the QZ algorithm of Moler & Stewart [?]. The form obtained then is described in [2] and can also be applied to find the Weyr characteristic of the infinite eigenvalue for the pencil.

An important side result of this paper is the derivation of what could be called a “generic” Hessenberg form of matrices with nontrivial minimal polynomials. We showed that the largest unreduced Hessenberg submatrix contains exactly the largest Jordan block of each eigenvalue and that all other unreduced Hessenberg submatrices

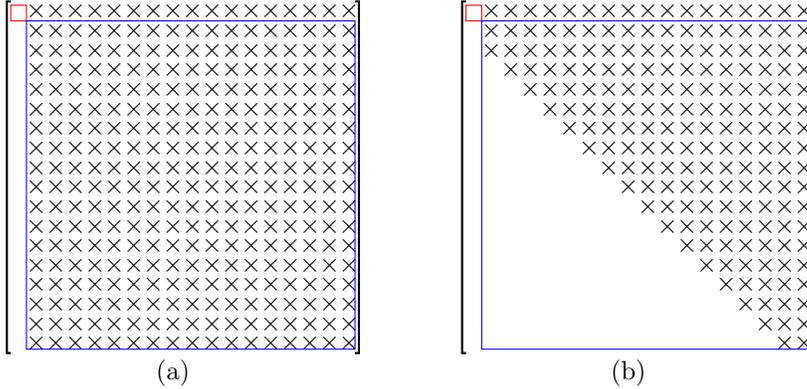


FIG. 3.3. Nonzero structure of the computed matrix in Example 3.3 by GOS (a) and by MVD (b).

then contain, in turn, the next largest one, and so on. This form can be useful to better understand the properties of generic Krylov sequences of an (A, b) pair. Last but not least, we believe to have found a way to perform perfect shifts in the QR algorithm, without suffering from the well-known blurring effect. A more detailed analysis of this will be the subject of a forthcoming report.

REFERENCES

- [1] K. M. ANSTREICHER AND U. G. ROTHBLUM, *Using Gauss-Jordan elimination to compute the index, generalized nullspaces, and Drazin inverse*, Linear Algebra and its Applications, 85 (1987), pp. 221–239.
- [2] T. BEELEN AND P. VAN DOOREN, *An improved algorithm for the computation of Kronecker’s canonical form of a singular pencil*, Linear Algebra and its Applications, 105 (1988), pp. 9–65.
- [3] ———, *Computational aspects of the Jordan canonical form*, Oxford Univ. Press, New York, 1990, pp. 57–72.
- [4] L. BRUGNANO AND D. TRIGIANTE, *Polynomial roots: the ultimate answer?*, Linear Algebra and its Applications, 225 (1995), pp. 207–219.
- [5] S. L. CAMPBELL AND C. D. MEYER, *Continuity properties of the Drazin pseudoinverse*, Linear Algebra and its Applications, 10 (1975), pp. 77–83.
- [6] J. DEMMEL AND B. KÅGSTRÖM, *Computing stable eigendecompositions of matrix pencils*, Linear Algebra & Applications, 88/89 (1987), pp. 139–186.
- [7] F. R. GANTMACHER, *Theory of Matrices*, Chelsea, New York, 1959.
- [8] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 2013.
- [9] G. H. GOLUB AND J. H. WILKINSON, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Rev., 36 (1976), pp. 578–619.
- [10] N. GUGLIELMI, M. OVERTON, AND G. STEWART, *An efficient algorithm for computing the generalized null space decomposition*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 38–54.
- [11] B. KÅGSTRÖM AND A. RUHE, *An algorithm for the numerical computation of the Jordan normal form of a complex matrix*, ACM Trans. Math. Software, 6 (1980), pp. 398–419.
- [12] V. N. KUBLANOVSKAYA, *On solving the complete eigenvalue problem for a degenerate matrix*, USSR Computational Math. and Math. Phys., 6 (1968), pp. 1–14.
- [13] N. MASTRONARDI AND P. VAN DOOREN, *The QR steps with perfect shifts*, Siam J. Matrix Anal. Appl., (2018), p. to appear.
- [14] V. MEHRMANN, *The Autonomous Linear Quadratic Control Problem*, Springer, Berlin, 1991.
- [15] A. RUHE, *An algorithm for numerical determination of the structure of a general matrix*, BIT, 10 (1970), pp. 196–216.
- [16] H. SHAPIRO, *The Weyr characteristic*, American Mathematical Monthly, 106 (1999), pp. 919–929.

- [17] P. VAN DOOREN, *The computation of Kronecker's canonical form of a singular pencil*, Linear Algebra and its Applications, 27 (1979), pp. 103–140.
- [18] D. S. WATKINS, *The transmission of shifts and shift blurring in the QR algorithm*, Linear Algebra and its Applications, 241–243 (1996), pp. 877–896.
- [19] ———, *The Matrix Eigenvalue Problem*, SIAM, Philadelphia, 2007.
- [20] W. M. WONHAM, *Linear Multivariable Control*, Springer-Verlag, New York, third ed., 1985.

Appendix A. Implicit Q theorem revisited. In this appendix we briefly revisit the implicit Q theorem and analyze the possible effect of rounding of implementing a so-called *perfect zero shift*. We show that under reasonable assumptions, the modified backward QR -step using the intermediate computation of the right eigenvector, produces the desired result.

THEOREM A.1. (*Implicit Q theorem revisited*). *Let $H \in \mathbb{R}^{n \times n}$ be an irreducible Hessenberg matrix with eigenvalue λ_0 . Then*

1. H has an “essentially unique” normalized eigenvector \mathbf{x} corresponding to λ_0 :

$$H\mathbf{x} = \lambda_0\mathbf{x}, \quad \|\mathbf{x}\|_2 = 1$$

and its last component $x_n \neq 0$;

2. there is an “essentially unique” sequence of Givens rotations $G_{n-1,n}, \dots, G_{1,2}$ whose product

$$Q := G_{1,2}G_{2,3} \cdots G_{n-1,n}$$

transforms the pair (H, \mathbf{x}) to a similar one

$$(\tilde{H}, \tilde{\mathbf{x}}) := (QHQ^T, Q\mathbf{x})$$

where

$$\tilde{\mathbf{x}} = \pm \mathbf{e}_1, \quad \tilde{H}\mathbf{e}_1 = \lambda_0\mathbf{e}_1, \quad \tilde{H} \text{ in Hessenberg form.}$$

Proof.

1. The fact that \mathbf{x} is unique (up to a scaling factor ± 1) follows from the equation

$$(H - \lambda_0 I)\mathbf{x} = 0, \quad \|\mathbf{x}\|_2 = 1,$$

where $(H - \lambda_0 I)$ has rank $n - 1$ since it is irreducible Hessenberg. For the same reason $x_n \neq 0$.

2. The reduction of \mathbf{x} to $\tilde{\mathbf{x}} = Q\mathbf{x} = \pm \mathbf{e}_1$ requires a sequence of Givens rotations

$$G_{i-1,i} \in \mathbb{R}^{n \times n}, \quad i = n, n-1, \dots, 2,$$

in order to eliminate the entries x_i , $i = n, n-1, \dots, 2$ of the vector \mathbf{x} . The only degree of freedom lies in a left and right diagonal scaling with elements of modulus 1. These are the same rotations that reduce

$$(H - \lambda_0 I)Q^T = \begin{bmatrix} \square & & \\ & \square & \\ & & \square \end{bmatrix} = R$$

to upper triangular form, and

$$Q(H - \lambda_0 I)Q^T = \begin{bmatrix} \square & & \\ & \square & \\ & & \square \end{bmatrix} = \tilde{H} - \lambda_0 I$$

to Hessenberg form again. Moreover, since $\mathbf{x} = \pm Q^T \mathbf{e}_1$, we also have

$$R\mathbf{e}_1 = 0, \quad (\tilde{H} - \lambda_0 I)\mathbf{e}_1 = 0$$

from which it follows that $\tilde{H}\mathbf{e}_1 = \lambda_0\mathbf{e}_1$.

□

REMARK A.1. *The implicit Q theorem is closely related to this lemma. It explains that the transformation Q can also be determined from the first rotation $G_{n-1,n}$ that computes*

$$[h_{n,n-1}, h_{n,n} - \lambda_0] G_{n-1,n}^T = [0 \quad \times]$$

and from the fact that $QHQT$ is still Hessenberg. This is known as “chasing the bulge” [19].

Let us now perturb H (and therefore also \mathbf{x}) and see what we can still prove. We assume from now on that $\lambda_0 = 0$.

Hence we look for

$$H\mathbf{x} = \mathbf{0}, \quad \|\mathbf{x}\|_2 = 1.$$

But since H is not exactly singular we use the singular value decomposition of H to find approximation \mathbf{v} of \mathbf{x} :

$$H\mathbf{v} = \underline{\sigma}\mathbf{u}, \quad \|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1, \quad \underline{\sigma} = \sigma_{\min}(H),$$

where \mathbf{u} and \mathbf{v} are respectively the left and right singular vector corresponding to $\underline{\sigma}$.

This then yields the minimum norm solution ΔH such that

$$(H + \Delta H)\mathbf{v} = \mathbf{0}, \quad \Delta H = -\underline{\sigma}\mathbf{u}\mathbf{v}^T.$$

But this does not satisfy the implicit Q theorem and hence is not appropriate. What we need is the minimum norm solution

$$\Delta H = \begin{bmatrix} \delta h_{1,1} & \delta h_{2,1} & \cdots & \delta h_{1,n} \\ \delta h_{2,1} & \delta h_{2,2} & \cdots & \delta h_{2,n} \\ & \ddots & \cdots & \vdots \\ & & \delta h_{n,n-1} & \delta h_{n,n} \end{bmatrix}$$

of Hessenberg structure such that

$$(H + \Delta H)\mathbf{v} = \mathbf{0}$$

and also

$$\Delta H\mathbf{v} = -\underline{\sigma}\mathbf{u}.$$

Since this is a linear set of equations, it has a minimum norm solution which we can solve row by row.

Let row i of this equation be

$$[0, \dots, 0, \underbrace{\delta \mathbf{h}_i^T}_{\min(n, n-i+2)}] \mathbf{v} = -\underline{\sigma}u_i,$$

with

$$\delta \mathbf{h}_i^T = [\delta h_{i,i-1}, \delta h_{i,i}, \dots, \delta h_{i,n}].$$

Then

$$\delta \mathbf{h}_i^T = -\underline{\sigma} \frac{u_i \mathbf{v}_i^T}{\|\mathbf{v}_i\|_2^2}, \quad \|\delta \mathbf{h}_i\|_2 = \underline{\sigma} \frac{|u_i|}{\|\mathbf{v}_i\|_2},$$

where

$$\mathbf{v}_i^T = [v_{i-1}, v_i, \dots, v_{n-1}, v_n], \quad i = n, n-1, \dots, 3, 2.$$

It is easy to see that if $\|\mathbf{v}_n\|_2 > \mu \approx 1$, where $\mathbf{v}_n = \begin{bmatrix} v_{n-1} \\ v_n \end{bmatrix}$, then all the subsequent norms are larger and the Frobenius norm of ΔH is bounded by

$$\|\Delta H\|_F \leq \underline{\sigma} / \mu.$$

It is worth pointing out that this can also be solved recursively by only considering problems on \mathbb{R}^2 . In the first step we have that

$$\delta \mathbf{h}_{n-1}^T = \underline{\sigma} \frac{u_n \mathbf{v}_n^T}{\|\mathbf{v}_n\|_2}.$$

But after this calculation we can apply the corresponding Givens transformation $G_{n-1,n}$ and consider the effect of the first ΔH :

$$(H + \Delta H)G_{n-1,n}^T = \begin{bmatrix} \times & \times & \cdots & \cdots & \times \\ \times & \times & \cdots & \cdots & \times \\ & \ddots & & & \vdots \\ & & \times & \times & \times \\ & & & \hat{\times} & \hat{\times} \end{bmatrix} G_{n-1,n}^T = \left[\begin{array}{cccc|c} \times & \times & \cdots & \cdots & \times \\ \times & \times & \cdots & \cdots & \times \\ & \ddots & & & \vdots \\ & & \times & \times & \times \\ \hline & & & 0 & r \end{array} \right].$$

Since now

$$\begin{bmatrix} I_{n-2} & \\ & G_{n-1,n} \end{bmatrix} \mathbf{v} = \begin{bmatrix} \times \\ \vdots \\ \times \\ \hline 0 \end{bmatrix},$$

we need only to consider the “deflated” problem of order $n-1$ which is again in Hessenberg form and the deflated null-vector \mathbf{v} which correspond to it.

We also point out that we can do slightly better than this solution by allowing \mathbf{v} to be perturbed as well, as long as

$$(H + \Delta H)(\mathbf{v} + \delta \mathbf{v}) = \mathbf{0}.$$

but we will not elaborate on this here.