The periodic Schur decomposition. Algorithms and applications

Adam Bojanczyk
Cornell University, Dept. Electrical Engineering
Ithaca, NY 14853-3801

Gene Golub
Stanford University, Dept. Computer Science
Stanford, CA 94305

Paul Van Dooren
University of Illinois at Urbana-Champaign, Coordinated Science Laboratory
1308 W. Main Str., Urbana, IL 61801

Abstract.

In this paper we derive a unitary eigendecomposition for a sequence of matrices which we call the *periodic Schur decomposition*. We prove its existence and discuss its application to the solution of periodic difference equations arising in control. We show how the classical QR algorithm can be extended to provide a stable algorithm for computing this generalized decomposition. We apply the decomposition also to cyclic matrices and two point boundary value problems.

Key words. Numerical algorithms, linear algebra, periodic systems, K-cyclic matrices, two-point boundary value problems

1 Introduction

In the study of time-varying control systems in (generalized) state space form :

$$\left\{ \begin{array}{rl} E_k \cdot z_{k+1} & = F_k \cdot z_k + G_k \cdot u_k \\ y_k & = H_k \cdot z_k + J_k \cdot u_k \end{array} \right. \tag{1}$$

the *periodic coefficients* case has always been considered the simplest extension of the time-invariant case. Here the coefficients satisfy, for some $K > 0$ the periodicity conditions $E_k = E_{k+K}$, $F_k = F_{k+K}$, $G_k = G_{k+K}$, $H_k = H_{k+K}$. $J_k = J_{k+K}$. The last few years there has been a renewed interest in the area because such systems arise naturally in multi-rate sampling of continuous time systems [1]. Several papers were devoted to the algebraic structure of periodic discrete time systems and it appears that a lot of the algebra indeed carries over from the time-invariant case [9]. For period $K = 1$ one has the time invariant case $E_k = E$, $F_k = F$, $G_k = G$, $H_k = H$, $J_k = J$, and it is well-known that the generalized eigenvalues of particular pencils derived from these matrices then determine the behaviour of these difference equations [13]. In the case $K > 1$ one can derive a set of K time-invariant subsampled systems [2], [9] that describe the behaviour of the periodic system. Problems of pole placement, optimal control and robust control can then be solved via these K subsampled systems.

During the last few decades linear algebra has played an important role in advances being made in the area of systems and control [16]. The most profound impact has been in the computational and implementational aspects, where numerical linear algebraic algorithms have strongly influenced the ways in which problems are being solved. The most reliable numerical linear algebra methods proposed for particular control problems are related to particular eigenvalue and singular value decompositions of "special" matrices, such as special Schur decompositions for solving Riccati equations [10], [14]. Here we present a new decomposition called the *periodic Schur form* that has important applications in control theoretic problems of periodic systems. We present a few of these applications and predict that several other uses will be found.

The decomposition has also a direct application to K-cyclic matrices and pencils, which occur in the study of Markov chains and the solution of two point boundary value problems. We show how the periodic Schur form naturally decomposes the underlying $n \times n$ matrix problem into n scalar poblems with the same structure. This can then directly be used for the solution of Markov chains and two point boundary value problems in an elegant manner. The relation with K-cyclic pencils also allows to completely characterize the singular matrix case and give conditions for the existence of solutions in the singular case.

2 Periodic Schur decomposition

Consider the set of (homogenous) difference equations

$$B_i \cdot x_{i+1} = A_i \cdot x_i, \quad i = 1, \dots \tag{2}$$

with *periodic coefficients* $A_i = A_{i+K}$, $B_i = B_{i+K}$. For period $K = 1$ one has the constant coefficient case $A_i = A$, $B_i = B$ and it is well-known that the generalized eigenvalues of the pair A, B yield important information about the system (2). When $K > 1$ one derives from (2) a set of K time invariant systems which describe completely the behavior of (2). For simplicity we first assume all B_i to be invertible. Then define the matrices $S_i = B_i^{-1} A_i$ yielding the system :

$$x_{i+1} = B_i^{-1} A_i \cdot x_i = S_i \cdot x_i, \quad i = 1, \dots \tag{3}$$

which is an explicit system of difference equations in x_i, again with periodic coefficients $S_i = S_{i+K}$.

One can now consider *subsampled systems* which describe the evolution of (3) over K steps, and since the coefficient matrices of (3) are K-periodic, one may expect these subsampled systems to be *time invariant*. Indeed, defining the matrices

$$S^{(k)} = S_{k+K-1} \cdot \dots S_{k+1} \cdot S_k, \; k = 1, ..., K. \tag{4}$$

then one obtains from (3), (4) the set of K *subsampled systems* :

$$\begin{aligned} x_{1+(i+1)K} &= S^{(1)} \cdot x_{1+iK}, \quad i = 0, 1, 2, \dots \\ x_{2+(i+1)K} &= S^{(2)} \cdot x_{2+iK}, \quad i = 0, 1, 2, \dots \\ &\vdots \\ x_{K+(i+1)K} &= S^{(K)} \cdot x_{K+iK}, \quad i = 0, 1, 2, \dots \end{aligned} \tag{5}$$

One easily checks that the above set of difference equations, initialized with the vectors $x_i, i = 1, \dots, K$ yields the same solution as (3). In order to describe the behaviour of these systems one thus requires the eigenvalues and eigenvectors of the *periodic matrix products* $S^{(k)}$. It is known from similar decompositions [11], [4], that explicitly forming the matrices $S^{(k)}$ ought to be avoided if possible. An implicit decomposition of these matrices is now obtained in the following theorem.

Theorem 1 *Let the matrices A_i, B_i, $i = 1, \ldots, K$ be all $n \times n$ and complex. Then there exist unitary matrices Q_i, Z_i, $i = 1, ..., K$ such that :*

$$\begin{array}{llll} \hat{B}_1 & = Z_1^* \cdot B_1 \cdot Q_2 & \hat{A}_1 & = Z_1^* \cdot A_1 \cdot Q_1 \\ \hat{B}_2 & = Z_2^* \cdot B_2 \cdot Q_3 & \hat{A}_2 & = Z_2^* \cdot A_2 \cdot Q_2 \\ & \vdots & & \\ \hat{B}_{K-1} & = Z_{K-1}^* \cdot B_{K-1} \cdot Q_K & \hat{A}_{K-1} & = Z_{K-1}^* \cdot A_{K-1} \cdot Q_{K-1} \\ \hat{B}_K & = Z_K^* \cdot B_K \cdot Q_1 & \hat{A}_K & = Z_K^* \cdot A_K \cdot Q_K \end{array} \tag{6}$$

where now all matrices $\hat{B}_i$, $\hat{A}_i$ are upper triangular. Moreover if the matrices B_i are invertible then each Q_i puts the matrix $S^{(i)}$ in upper Schur form, i.e. $Q_i^* S^{(i)} Q_i$ is upper triangular.

Proof : Because of its simplicity and constructive derivation, we give here a simple proof assuming all matrices A_i and B_i are non-singular, except possibly A_1. The more complex case of singular matrices is proven in section 3.2.

If all matrices B_i are invertible then all matrices $S^{(i)}$ exist. Compute the upper Schur form of $S^{(1)}$:

$$Q_1^* S^{(1)} Q_1 = \hat{S}^{(1)}.$$

This defines the matrix Q_1 and one can thus consider the matrix $B_K \cdot Q_1$ and its QR decomposition :

$$Z_K \cdot \hat{B}_K = [B_K Q_1]$$

which defines the unitary factor Z_K and upper-triangular factor $\hat{B}_K$. In turn, one then considers the matrix $Z_K^* \cdot A_K$ and its RQ decomposition (i.e. dual to the QR decomposition) :

$$\hat{A}_K \cdot Q_K^* = [Z_K^* A_K]$$

which defines the unitary factor Q_K and upper-triangular factor $\hat{A}_K$. Repeating this for all subsequent matrices defines :

- Z_i and $\hat{B}_i$ from the QR factorization of $B_i \cdot Q_{i+1}$ for $i = K, ..., 1$ and
- Q_i and $\hat{A}_i$ from the RQ factorization of $Z_i^* \cdot A_i$ for $i = K, ..., 2$.

Notice that each of these decompositions in fact corresponds to one of the equations in (6), starting from bottom to top. By now all transformation matrices Q_i and Z_i are defined but we have not proved that the last matrix $\hat{A}_1$ is upper-triangular, since in the equation

$$\hat{A}_1 = Z_1^* \cdot A_1 \cdot Q_1$$

the matrix Q_1 was already defined. But consider now the product

$$Q_1^* S^{(1)} Q_1 = [Q_1^* B_K^{-1} Z_K][Z_K^* A_K Q_K] \cdots [Q_3^* B_2^{-1} Z_2][Z_2^* A_2 Q_2][Q_2^* B_1^{-1} Z_1][Z_1^* A_1 Q_1] \tag{7}$$

or

$$\hat{S}^{(1)} = \hat{B}_K^{-1} \hat{A}_K \cdots \hat{B}_2^{-1} \hat{A}_2 \hat{B}_1^{-1} [Z_1^* A_1 Q_1]. \tag{8}$$

Now since all "hat" matrices in both sides of equation (8) are upper-triangular and invertible, this must also hold for the matrix $\hat{A}_1 = Z_1^* A_1 Q_1$. This completes the constructive proof of the existence of (6).

Notice that the proof shows how to derive all matrices Q_i and Z_i from just one of them. Moreover, by periodically interchanging the products in (7) one easily sees that also

$$Q_i^* S^{(i)} Q_i = \hat{S}^{(i)} = \hat{B}_{i-1}^{-1} \hat{A}_{i-1} \cdots \hat{B}_1^{-1} \hat{A}_1 \hat{B}_K^{-1} \hat{A}_K \cdots \hat{B}_{i+1}^{-1} \hat{A}_{i+1} \hat{B}_i^{-1} \hat{A}_i \tag{9}$$

is upper triangular and hence a Schur decomposition. So all Schur forms are actually dependent on one another via (6). ■

Corollary 1 *Let the matrices $A_i,\ B_i,\ i = 1,\ldots,K$ be all $n \times n$ and real. Then there exist orthogonal matrices $Q_i,\ Z_i,\ i = 1,...,K$ such that the above decomposition (6) holds and all but one of the matrices $\hat{B}_i,\ \hat{A}_i$ are upper triangular. This last one is in quasi-upper triangular form with 1×1 and 2×2 diagonal blocks.*

Proof : Assume that all matrices are invertible except, say, A_1 (see section 3.2 for the general case). The proof then goes as before. Pick a *real* transformation Q_1 that puts $S^{(1)}$ in *real* Schur form $\hat{S}^{(1)} = Q_1^T S^{(1)} Q_1$. Then perform all QR factorizations as above to define the remaining transformation matrices $Z_i,\ i = K...,1$ and $Q_i,\ i = K,...,2$ in decreasing order (these are *real* transformations, of course). In (8) $\hat{B}_K,\ i = K,...,1$, $\hat{A}_K,\ i = K,...,2$ (and their inverses) are upper triangular, and $\hat{S}^{(1)}$ is quasi upper-triangular. From this it follows that $\hat{A}_1$ must be of the same form as $\hat{S}^{(1)}$. If one would have started the definition of the transformations Z_i and Q_i from the other side (i.e. the QR factorization of A_1Q_1 instead of B_KQ_1) then $\hat{B}_K$ (and its inverse) would have the same form as $\hat{S}^{(1)}$. Finally, by starting the above reasoning with a different index i it is clear that one can pick any matrix $\hat{A}_i$ or $\hat{B}_i$ to have the quasi-triangular shape. It is easy to move it around as well via a "post-processing" using updating Givens rotations. ■

In fact the matrices Q_i transform the vectors x_i to $\hat{x}_i = Q_i^* \cdot x_i$ and the difference equations (2) to the *equivalent system* :

$$Z_i^* B_i Q_{i+1} \cdot Q_{i+1}^* x_{i+1} = Z_i^* A_i Q_i \cdot Q_i^* x_i, \quad i = 1,\ldots \tag{10}$$

or

$$\hat{B}_i \cdot \hat{x}_{i+1} = \hat{A}_i \cdot \hat{x}_i, \quad i = 1,\ldots \tag{11}$$

with *periodic coefficients* $\hat{A}_i = \hat{A}_{i+K},\ \hat{B}_i = \hat{B}_{i+K}$ which are now all *upper triangular* (except one quasi triangular one in the real case). The same transformations can of course be applied to the non-homogenous case, and this will be used later on.

An elegant consequence of the above theorem is the following corollary.

Corollary 2 *All periodic products $S^{(i)}$ have equal eigenvalues and their Schur forms $\hat{S}^{(i)}$ given by the implicit decomposition (6) have the same eigenvalues on diagonal.*

Proof : It is trivially seen that $S^{(i)}$ and $S^{(1)}$ have equal eigenvalues since

$$S^{(i)} = M_1 M_2, \ \ S^{(1)} = M_2 M_1$$

with

$$M_2 = S_K \cdot \ldots \cdot S_i, \ \ M_1 = S_{i-1} \cdot \ldots \cdot S_1.$$

Equality of spectrum indeed follows immediately from this. The Schur forms of the matrices $S^{(i)}$ will thus have the same diagonal elements, *up to their ordering.* But the Schur forms constructed by (6) have the additional property that the diagonal elements of the $\hat{S}^{(i)}$ matrices are all actually equal. Indeed, they are the products of the diagonal elements of the upper triangular matrices $\hat{B}_i^{-1}\hat{A}_i$. So, if one matrix $\hat{S}^{(i)}$ has a particular ordering of eigenvalues then all other matrices $\hat{S}^{(j)}$ have the same ordering of eigenvalues. ■

We give in the next section an algorithm to compute the above decomposition *implicitly*, i.e. without ever forming the products $S^{(i)}$. Moreover we show how to reorder the eigenvalues of these Schur forms. We call this the *periodic QR algorithm* as related to the above periodic Schur decomposition.

3 Periodic QR algorithm

We now consider the computation of the periodic Schur decomposition. Here we will not require the invertibility of the matrices A_i, B_i. In order to have a periodic QR algorithm we need the following ingredients to make the algorithm work :

1. a reduction to some kind of Hessenberg form
2. a direct deflation of the singular case
3. a shift calculation procedure
4. a method for performing QR steps
5. a procedure for reordering eigenvalues.

In the above list one should try to do as much as possible implicitly, i.e. without ever *constructing* the products $S^{(i)}$. Moreover one would like the total complexity of the algorithm to be comparable to the cost of K Schur decompositions, since this is what we implicitly compute. This means that the complexity should be $O(Kn^3)$ for the whole process. Notice that this indeed precludes the construction of the products $S^{(i)}$ since this would already require $O(K^2n^3)$ operations. We now derive such implicit solutions for each item. Below $\mathcal{H}(i,j)$ denotes the group of *Householder* transformations whereby (i,j) is the range of rows/columns they operate on. Similarly $\mathcal{G}(i,i+1)$ denotes the group of *Givens* transformations operating on rows/columns i and $i+1$.

3.1 Hessenberg-triangular reduction

We first consider the case where all B_i are the identity. We thus only have a product of matrices A_i and in order to illustrate the procedure we show its evolution on a product of 3 matrices only, i.e. $A_3A_2A_1$. Below is a sequence of "snapshots" of the evolution of the Hessenberg-triangular reduction. Each snapshot indicates the pattern of zeros ('0') and nonzeros ('x') in the three matrices.

First perform a Householder transformation $Q_3 \in \mathcal{H}(1,n)$ on the rows of A_2 and the columns of A_3. Choose Q_3 to annihilate all but one element in the first column of A_2 :

$$\begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \end{bmatrix}.$$

Then perform a Householder transformation $Q_1 \in \mathcal{H}(1,n)$ on the rows of A_3 and the columns of A_1. Choose Q_1 to annihilate all but one element in the first column of A_3 :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \end{bmatrix}.$$

Then perform a Householder transformation $Q_2 \in \mathcal{H}(2,n)$ on the rows of A_1 and the columns of A_2. Choose Q_2 to annihilate all but two element in the first column of A_1 :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix}.$$

Notice that this third transformation did not destroy any of the previously created elements in A_2 because it did not transform its first column. A similar set of three transformations yields the following three snapshots :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix}$$

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & x & x & x & x & x \end{bmatrix}$$

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & x & x & x & x \end{bmatrix},$$

and this continues until we reach the Hessenberg-triangular form :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}.$$

When the matrices B_i are not the identity matrix, one starts with transforming each of them to triangular form. Then one proceeds with a similar reduction procedure for the matrices A_i as above. While the zero elements are being created in the matrices A_i one preserves the matrices B_i in upper triangular form at each step. Therefore, one can not make use of Householder transformations anymore. Indeed, applying a Householder transformation in $\mathcal{H}(k,n)$ (left or right) to a triangular matrix B_i fills it in and one can not find a Householder transformation in the same class operating on the other side of B_i, that will restore its triangular shape. On the other hand, this is easily done when using a Givens transformation in $\mathcal{G}(k,k+1)$ since then only the element $B_i(k+1,k)$ fills in below the diagonal and this can immediately be annihilated again using another Givens transformation in $\mathcal{G}(k,k+1)$ operating on the other side of

B_i. The above procedure of creating zeros in A_i, while maintaining the matrices B_i in upper triangular form, can thus go through. Notice that for the case $K = 1$ one retrieves *exactly* the Hessenberg-triangular reduction of the QZ algorithm [11]. Operation counts for this Hessenberg-triangular reduction are given in section 5.1.

3.2 Direct deflation of the singular case

In this section we show how to perform *direct deflations* in the Hessenberg-triangular form when either of the *pivot elements* is zero. With pivot element we mean the elements on the diagonal of each triangular matrix A_i, $i = 2, ..., K$, B_i, $i = 1, ..., K$ and below the diagonal in the Hessenberg matrix A_1. Below we treat three different cases and show how direct deflations can be performed to yield one or several subproblems of smaller dimensions where now all pivot elements are nonzero. This corresponds to subproblems without eigenvalues at zero or ∞.

Case 1. When an element below the diagonal of A_1 is zero, the problem trivially decomposes in two lower dimensional problems, as shown below for matrices B_2, A_2, B_1, A_1 where the $(4, 3)$ element in A_1 is zero :

$$\left[\begin{array}{ccc|ccc} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|ccc} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|ccc} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|ccc} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{array}\right].$$

This reduction is identical to what happens in the single matrix case and clearly can be repeated until one obtains smaller dimensional matrices A_1 with non-zero subdiagonals (i.e. unreduced Hessenberg forms). Moreover the reduction does not involve any transformation but only a partitioning. The next two cases are zero diagonal elements in any of the remaining matrices. One first deflates the zeros in the first matrix in the sequence $B_2, A_3, B_3, ..., A_K, B_K$, i.e. one first treats the "closest" matrix to A_1.

Case 2. If the closest matrix to A_1 with zero diagonal elements is A_i, then the partial product $A_i B_{i-1}^{-1} A_{i-1} ... B_1^{-1} A_1$ again decomposes in a block diagonal matrix, as indicated below with the sequence $A_2 B_1^{-1} A_1$ where A_2 has a zero diagonal in position $(4, 4)$:

$$\left[\begin{array}{cccc|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{cccc|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right]^{-1} \left[\begin{array}{ccc|cccc} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{array}\right]$$

$$= \left[\begin{array}{ccc|cccc} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{array}\right].$$

Moreover the bottom block is rank 3 only and one ought to be able to extract a zero eigenvalue. We now show how a sequence of Givens transformations can be generated to obtain a deflated and decomposed form of the type :

$$\cdots \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{array}\right].$$

We first apply the row transformation $Z_1^* = G_3.G_2.G_1$ to A_1, where the Givens transformations $G_1 \in \mathcal{G}(1,2)$, $G_2 \in \mathcal{G}(2,3)$ and $G_3 \in \mathcal{G}(3,4)$ are chosen to annihilate the elements 0_1, 0_2 and 0_3, respectively, as given below. Propagating these through the intermediate triangular matrices (here only B_1) this results in the column transformation $Q_2 = G_3.G_4.G_5$ applied to A_2, where the Givens transformations $G_4 \in \mathcal{G}(1,2)$ and $G_5 \in \mathcal{G}(2,3)$ respectively create the nonzero elements x_4 and x_5 ($G_6 \in \mathcal{G}(3,4)$ does not create any element) :

$$\cdots \left[\begin{array}{cccc|ccc} x & x & x & x & x & x & x \\ x_4 & x & x & x & x & x & x \\ 0 & x_5 & x & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{cccc|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|cccc} x & x & x & x & x & x & x \\ 0_1 & x & x & x & x & x & x \\ 0 & 0_2 & x & x & x & x & x \\ 0 & 0 & 0_3 & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{array}\right].$$

Then the two elements x_4 and x_5 are annihilated again by Givens transformations $G_7 \in \mathcal{G}(1,2)$ and $G_8 \in \mathcal{G}(2,3)$ as part of the row transformation $Z_2^* = G_8.G_7$ acting on A_2 (this yields 0_7 and 0_8, respectively). Propagating these through the intermediate triangular matrices left of A_2 and then back to A_1, this results in the column transformation $Q_1 = G_9.G_{10}$ acting on A_1. Here the Givens transformations $G_9 \in \mathcal{G}(1,2)$ and $G_{10} \in \mathcal{G}(2,3)$ create the elements x_9 and x_{10}, respectively :

$$\cdots \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ 0_7 & x & x & x & x & x & x \\ 0 & 0_8 & x & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|cccc} x & x & x & x & x & x & x \\ x_9 & x & x & x & x & x & x \\ 0 & x_{10} & x & x & x & x & x \\ 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{array}\right].$$

This subsequence of matrices is now already closer to the desired result. The next steps are dual to the ones above and are just indicated below by the sequence of annihilated and created elements. Just as above, everything is done via appropriate Givens rotations :

$$\cdots \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x_{15} & x & x \\ 0 & 0 & 0 & 0 & 0 & x_{14} & x \end{array}\right] \left[\begin{array}{ccc|c|ccc} x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x \end{array}\right] \left[\begin{array}{ccc|cccc} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x \\ \hline 0 & 0 & 0 & x & x & x & x \\ \hline 0 & 0 & 0 & 0_{13} & x & x & x \\ 0 & 0 & 0 & 0 & 0_{12} & x & x \\ 0 & 0 & 0 & 0 & 0 & 0_{11} & x \end{array}\right],$$

and finally :

$$
\left[\begin{array}{ccc|c|ccc}
x & x & x & x & x & x & x \\
0 & x & x & x & x & x & x \\
0 & 0 & x & x & x & x & x \\
\hline
0 & 0 & 0 & 0 & x & x & x \\
\hline
0 & 0 & 0 & 0 & x & x & x \\
0 & 0 & 0 & 0 & 0_{18} & x & x \\
0 & 0 & 0 & 0 & 0 & 0_{17} & x
\end{array}\right]
\left[\begin{array}{ccc|c|ccc}
x & x & x & x & x & x & x \\
0 & x & x & x & x & x & x \\
0 & 0 & x & x & x & x & x \\
\hline
0 & 0 & 0 & x & x & x & x \\
\hline
0 & 0 & 0 & 0 & x & x & x \\
0 & 0 & 0 & 0 & 0 & x & x \\
0 & 0 & 0 & 0 & 0 & 0 & x
\end{array}\right]
\left[\begin{array}{ccc|c|ccc}
x & x & x & x & x & x & x \\
x & x & x & x & x & x & x \\
0 & x & x & x & x & x & x \\
\hline
0 & 0 & 0 & x & x & x & x \\
\hline
0 & 0 & 0 & 0 & x & x & x \\
0 & 0 & 0 & 0 & x_{20} & x & x \\
0 & 0 & 0 & 0 & 0 & x_{19} & x
\end{array}\right],
$$

which is precisely the desired form. Notice that all this requested about n Givens rotations on each side of each condensed matrix. As a result a zero eigenvalue was deflated and moreover a block reduction was obtained as the same time (see section 5.1 for more details on the operation count).

Case 3. We now consider the case where the closest matrix with a zero diagonal element occurs in a matrix B_i. Without loss of generality we may assume that it is the matrix B_1, since we can always associate the subproduct $A_i B_{i-1}^{-1} A_{i-1} \ldots B_1^{-1} A_1$ with the matrix A_1 (this subproduct indeed exists and is unreduced Hessenberg). Below we thus take the example ... B_1, A_1 where B_1 has a zero diagonal in position $(4,4)$:

$$
\cdots
\begin{bmatrix}
x & x & x & x & x & x \\
0 & x & x & x & x & x \\
0 & 0 & x & x & x & x \\
0 & 0 & 0 & 0 & x & x \\
0 & 0 & 0 & 0 & x & x \\
0 & 0 & 0 & 0 & 0 & x
\end{bmatrix}
\begin{bmatrix}
x & x & x & x & x & x \\
x & x & x & x & x & x \\
0 & x & x & x & x & x \\
0 & 0 & x & x & x & x \\
0 & 0 & 0 & x & x & x \\
0 & 0 & 0 & 0 & x & x
\end{bmatrix}.
$$

We first perform a row transformation $Z_1^* = G_1$ on both B_1 and A_1 where $G_1 \in \mathcal{G}(4,5)$ is chosen to annihilate the element 0_1 in B_1. At the same time a nonzero element x_1 is created in A_1 :

$$
\cdots
\begin{bmatrix}
x & x & x & x & x & x \\
0 & x & x & x & x & x \\
0 & 0 & x & x & x & x \\
0 & 0 & 0 & 0 & x & x \\
0 & 0 & 0 & 0 & 0_1 & x \\
0 & 0 & 0 & 0 & 0 & x
\end{bmatrix}
\begin{bmatrix}
x & x & x & x & x & x \\
x & x & x & x & x & x \\
0 & x & x & x & x & x \\
0 & 0 & x & x & x & x \\
0 & 0 & x_1 & x & x & x \\
0 & 0 & 0 & 0 & x & x
\end{bmatrix}.
$$

Then a column transformation $Q_1 = G_2$ with $G_2 \in \mathcal{G}(3,4)$ is applied to A_1 to annihilate the element x_1 again (yielding 0_2). Propagating this over all triangular matrices back to B_1 yields a column transformation $Q_2 \in \mathcal{G}(3,4)$ that does not create any fill in :

$$
\cdots
\begin{bmatrix}
x & x & x & x & x & x \\
0 & x & x & x & x & x \\
0 & 0 & x & x & x & x \\
0 & 0 & 0 & 0 & x & x \\
0 & 0 & 0 & 0 & 0 & x \\
0 & 0 & 0 & 0 & 0 & x
\end{bmatrix}
\begin{bmatrix}
x & x & x & x & x & x \\
x & x & x & x & x & x \\
0 & x & x & x & x & x \\
0 & 0 & x & x & x & x \\
0 & 0 & 0_2 & x & x & x \\
0 & 0 & 0 & 0 & x & x
\end{bmatrix}.
$$

After this step the B_1 matrix has two consecutive zero diagonal elements. The next pair of steps move these zero diagonals one elements down while keeping A_1 Hessenberg. First apply a row transformation

$Z_1^* = G_3$ on both B_1 and A_1 where $G_3 \in \mathcal{G}(5,6)$ annihilates 0_3 in B_1 and creates x_3 in A_1 :

$$\dots \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 & 0 & 0_3 \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & x_3 & x & x \end{bmatrix}.$$

Then apply the column transformation $Q_1 = G_4$ with $G_4 \in \mathcal{G}(4,5)$ on A_1 to annihilate the element x_3 again (yielding 0_4). Propagating this over all triangular matrices back to B_1 yields a column transformation $Q_2 \in \mathcal{G}(4,5)$ that creates the element x_4 :

$$\dots \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x_4 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0_4 & x & x \end{bmatrix}.$$

With the two consecutive zero diagonals now at the bottom of B_1, we finally apply a column transformation $Q_1 = G_5$ with $G_5 \in \mathcal{G}(5,6)$ on A_1 to annihilate its bottom off diagonal element (yielding 0_5). Propagating this back to B_1 yields a column transformation $Q_2 \in \mathcal{G}(5,6)$ that creates the element x_5 :

$$\dots \left[\begin{array}{ccccc|c} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x_5 & x \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{ccccc|c} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ \hline 0 & 0 & 0 & 0 & 0_5 & x \end{array}\right].$$

The above form can now be deflated as indicated above. Notice that again the number of Givens transformations applied to each matrix is at most of the order of n for one deflated eigenvalue at ∞.

Summary. The above three cases indicate that any zero pivot element can be deflated with $O(n)$ Givens transformations per matrix, until a (set of) lower dimensional problem(s) is obtained where now all triangular matrices are invertible and A_1 is unreduced Hessenberg. In the proof of Theorem 1 and Corollary 1 the general case can thus be "pretreated" by the Hessenberg-triangular reduction followed by the direct deflation described above. Theorem 1 and Corollary 1 can then be applied to these "nonsingular" cases, which implicitly yields a proof of these theorems for the general case where any B_i or A_i may be singular. Moreover, since the above procedure allows us to reduce the general problem to the nonsingular case, we only need to consider this simpler case in the sequel.

3.3 Shift calculation and QR step construction

Since we have now a Hessenberg-triangular form with all lower order matrices invertible and unreduced, the corresponding products $B_K^{-1}A_K \dots B_2^{-1}A_2B_1^{-1}A_1$ exist and are unreduced Hessenberg. In the QR algorithm applied to an unreduced Hessenberg matrix, the shift is typically computed from the bottom 2×2 submatrix. For the above sequence, this is of the form

$$\begin{bmatrix} b^{(K)}_{n-1,n-1} & b^{(K)}_{n-1,n} \\ 0 & b^{(K)}_{n,n} \end{bmatrix}^{-1} \begin{bmatrix} a^{(K)}_{n-1,n-1} & a^{(K)}_{n-1,n} \\ 0 & a^{(K)}_{n,n} \end{bmatrix} \dots \begin{bmatrix} b^{(1)}_{n-1,n-1} & b^{(1)}_{n-1,n} \\ 0 & b^{(1)}_{n,n} \end{bmatrix}^{-1} \begin{bmatrix} a^{(1)}_{n-1,n-1} & a^{(1)}_{n-1,n} \\ a^{(1)}_{n,n-1} & a^{(1)}_{n,n} \end{bmatrix}. \tag{12}$$

Notice that the triangular 2×2 inverses can be replaced by their adjoints up to a scalar factor. The eigenvalues of this 2×2 matrix are thus easily computed and are used for calculating the shift of the QR-step.

The transformation Q_1 of the QR step applied to the Hessenberg matrix

$$B_K^{-1} A_K ... B_2^{-1} A_2 B_1^{-1} A_1$$

is now completely defined by its first column. In the case of a single shift λ, this first column has only two nonzero elements, corresponding to the normalized version of the 2-vector :

$$\begin{bmatrix} b_{1,1}^{(K)} & b_{1,2}^{(K)} \\ 0 & b_{2,2}^{(K)} \end{bmatrix}^{-1} \begin{bmatrix} a_{1,1}^{(K)} & a_{1,2}^{(K)} \\ 0 & a_{2,2}^{(K)} \end{bmatrix} \cdots \begin{bmatrix} b_{1,1}^{(1)} & b_{1,2}^{(1)} \\ 0 & b_{2,2}^{(1)} \end{bmatrix}^{-1} \begin{bmatrix} a_{1,1}^{(1)} \\ a_{2,1}^{(1)} \end{bmatrix} - \begin{bmatrix} \lambda \\ 0 \end{bmatrix}.$$

Since the matrices Q_i and Z_i are all defined by one another through the constraint that updates on $B_i,\ i = 1..., K$ and $A_i,\ i = 2, ..., K$ must be upper triangular, one could as well compute any other matrix than Q_1. It turns out that the simplest one to contruct is Z_1. It performs a QR step on the unreduced Hessenberg matrix

$$A_H \doteq A_1 B_K^{-1} A_K ... B_2^{-1} A_2 B_1^{-1}$$

and is again defined by its first column, consisting of only two nonzero elements. Now this 2-vector is the normalized version of :

$$\begin{bmatrix} a_{1,1}^{(1)} \\ a_{2,1}^{(1)} \end{bmatrix} - \begin{bmatrix} \lambda \\ 0 \end{bmatrix} \frac{b_{1,1}^{(1)} \cdot ... \cdot b_{1,1}^{(K)}}{a_{1,1}^{(2)} \cdot ... \cdot a_{1,1}^{(K)}}$$

which involves much less computations.

In the implicit double shift one determines the first column of the real matrix $(A_H - \lambda_1)(A_H - \lambda_2)$ where λ_1 and λ_2 are the two eigenvalues of (12). In order to avoid complex arithmetic when $\lambda_i,\ i = 1, 2$ are complex conjugate one constructs the first column of $A_H^2 - s \cdot A_H + p \cdot I$ where $s = (\lambda_1 + \lambda_2)$ and $p = \lambda_1 . \lambda_2$ are *real*. This vector has only three nonzero elements and is up to a constant :

$$\begin{bmatrix} a_{1,1}^{(1)} & a_{1,2}^{(1)} \\ a_{2,1}^{(1)} & a_{2,2}^{(1)} \\ 0 & a_{3,2}^{(1)} \end{bmatrix} \begin{bmatrix} b_{1,1}^{(K)} & b_{1,2}^{(K)} \\ 0 & b_{2,2}^{(K)} \end{bmatrix}^{-1} \begin{bmatrix} a_{1,1}^{(K)} & a_{1,2}^{(K)} \\ 0 & a_{2,2}^{(K)} \end{bmatrix} \cdots \begin{bmatrix} b_{1,1}^{(1)} & b_{1,2}^{(1)} \\ 0 & b_{2,2}^{(1)} \end{bmatrix}^{-1} \begin{bmatrix} a_{1,1}^{(1)} \\ a_{2,1}^{(1)} \end{bmatrix}$$

$$-s \begin{bmatrix} a_{1,1}^{(1)} \\ a_{2,1}^{(1)} \\ 0 \end{bmatrix} + \begin{bmatrix} p \\ 0 \\ 0 \end{bmatrix} \frac{b_{1,1}^{(1)} \cdot ... \cdot b_{1,1}^{(K)}}{a_{1,1}^{(2)} \cdot ... \cdot a_{1,1}^{(K)}}.$$

3.4 Periodic QR step

Again for simplicity we only consider the product of four matrices $B_2^{-1} A_2 B_1^{-1} A_1$ and the case of a single shift in order to explain the general idea. The first three matrices are upper triangular. The last matrix A_1 is upper Hessenberg.

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}.$$

Apply first $Z_1^* \in \mathcal{G}(1,2)$ to annihilate the bottom element in the 2-vector determined above. Applying this to the rows of B_1 and A_1 yields :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x_1 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}.$$

Then construct the column transformation $Q_2 \in \mathcal{G}(1,2)$ to annihilate again x_1 in B_1 but also apply this transformation to the columns of A_2, creating x_2 :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x_2 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0_2 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}$$

Then apply the row transformation $Z_2^* \in \mathcal{G}(1,2)$ to B_2 and A_2 annihilating x_2 but creating x_3 :

$$\begin{bmatrix} x & x & x & x & x & x \\ x_3 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0_3 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}.$$

Finally close the loop with the column transformation $Q_2 \in \mathcal{G}(1,2)$ applied to B_2 and A_1 to annihilate again x_3 but creating a "bulge" x_4 in A_1 :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0_4 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x_4 & x & x & x & x & x \\ 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}.$$

Repeating this process chases the bulge one step down at each sequence of Givens transformations, untill it finally dissapears at the bottom of the Hessenberg matrix A_1. Basically the same procedure applies to the implicit double shift for real matrices except that then the bulge chasing transformations are 3×3 unitary matrices, realized by a product of Householder transformations or Givens transformations.

3.5 Reordering eigenvalues

We assume now that an upper triangular decomposition was obtained upon convergence of the above QR steps (blow there is only one 2×2 block in A_1). Then we want to permute the two (real) eigenvalues corresponding to the diagonal elements x_1 and x_2 :

$$\begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x_1 & x & x & x \\ 0 & 0 & 0 & x_2 & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x_1 & x & x & x \\ 0 & 0 & 0 & x_2 & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ 0 & x & x & x & x & x \\ 0 & 0 & x_1 & x & x & x \\ 0 & 0 & 0 & x_2 & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & 0 & x_1 & x & x & x \\ 0 & 0 & 0 & x_2 & x & x \\ 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix}$$

One then computes the product of the corresponding 2×2 matrices and computes from there the requested updating Givens transformations that will perform the swapping. Care has to be taken to implement this in a numerically stable manner as was e.g. the case for the QZ reordering in [14]. This especially applies to the swapping of two 2×2 blocks which is a much more delicate problem.

4 Applications of the periodic Schur form

4.1 Periodic control systems

The application of this decomposition to control theory is apparent. Periodic discrete time systems naturally arise when performing multirate sampling of continuous time systems [1]. In optimal control of such a periodic system one considers the problem :

$$\begin{array}{c} Minimize\ J = \sum_{i=1}^{\infty} z_i^T Q_i z_i + u_i^T R_i u_i \\ subject\ to\ E_i z_{i+1} = F_i z_i + G_i u_i \end{array} \tag{13}$$

where the matrices Q_i, R_i, E_i, F_i, G_i are periodic with period K. The Hamiltonian equations are periodic homogenous systems of difference equations (2) in the state z_i and co-state λ_i of the system. The correspondences with (2) are :

$$x_i \doteq \begin{bmatrix} \lambda_i \\ z_i \end{bmatrix}, B_i \doteq \begin{bmatrix} -G_i R_i^{-1} G_i^T & E_i \\ F_i^T & 0 \end{bmatrix}, A_i \doteq \begin{bmatrix} 0 & F_i \\ E_i^T & Q_i \end{bmatrix}. \tag{14}$$

For finding the periodic solutions to the underlying periodic Riccati equation one has to find the stable invariant subspaces of matrices $S^{(i)}$ as above, which happen to be simplectic in the discrete time case (one has to assume here that E_i, F_i and R_i are invertible and eliminate implicitly E_i [7]). Clearly the Schur form is useful here as well as the reordering of eigenvalues [10], [14].

In pole placement of periodic systems [9], again the periodic Schur form and reordering is useful when one wants to extend Varga's pole placement algorithm [17] to periodic systems. Consider the system

$$\begin{array}{c} B_i z_{i+1} = A_i z_i + D_i u_i \\ with\ state\ feedback\ u_i = F_i z_i + v_i \end{array} \tag{15}$$

where the matrices A_i, B_i, D_i, F_i are periodic with period K. This results in the closed loop system

$$B_i z_{i+1} = (A_i + D_i F_i) z_i + D_i v_i \tag{16}$$

of which the underlying time invariant eigenvalues are those of the matrix :

$$S_F^{(1)} \doteq B_K^{-1}(A_K + D_K F_K) \cdots B_2^{-1}(A_2 + D_2 F_2) B_1^{-1}(A_1 + D_1 F_1). \tag{17}$$

In the above equation it is not apparent at all how to choose the matrices F_i to assign particular eigenvalues of $S_F^{(1)}$. Yet when the matrices A_i, B_i are in the triangular form (6), one can choose the F_i matrices to have only nonzero elements in the last column. This will preserve the triangular form of the matrices $A_i + D_iF_i$ and it is then trivial to choose e.g. one such column vector to assign one eigenvalue. In order to assign the other eigenvalues one needs to *reorder* the diagonal elements in the periodic Schur form and each time assign another eigenvalue with the same technique. This algorithm will of course fail when the periodic system is not controllable, but this very procedure can in fact be adapted to precisely construct the controllable subspace of the periodic system.

4.2 K-cyclic matrix problems

Here we consider the following pencils of matrices :

$$\lambda\mathcal{B} - \mathcal{A} \doteq \lambda \begin{bmatrix} B_K & 0 & \cdots & \cdots & 0 \\ 0 & B_1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & B_{K-2} & 0 \\ 0 & 0 & \cdots & 0 & B_{K-1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & \cdots & 0 & A_K \\ A_1 & 0 & 0 & \cdots & 0 \\ \vdots & A_2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & A_{K-1} & 0 \end{bmatrix}. \tag{18}$$

If the B_i matrices here are invertible one can divide them out by columns transformation, yielding :

$$\lambda I_{nK} - \mathcal{B}^{-1}\mathcal{A} \doteq \lambda I_{nK} - \mathcal{S} \doteq \lambda \begin{bmatrix} I_n & 0 & \cdots & \cdots & 0 \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & I_n & 0 \\ 0 & 0 & \cdots & 0 & I_n \end{bmatrix} - \begin{bmatrix} 0 & 0 & \cdots & 0 & S_K \\ S_1 & 0 & 0 & \cdots & 0 \\ \vdots & S_2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & S_{K-1} & 0 \end{bmatrix}$$

where the matrices $S_i = B_i^{-1}A_i$ are as defined earlier. The matrix $\mathcal{S}$ is now known as a K-cyclic matrix, and by extension we will call $\lambda\mathcal{B} - \mathcal{A}$ a K-cyclic pencil. It is well-known that the eigenvalues of $\mathcal{S}$ are the K-th roots of those of the matrix $\mathcal{S}^K$, but the latter is easily checked to be block diagonal :

$$\mathcal{S}^K \doteq \begin{bmatrix} S^{(1)} & 0 & \cdots & \cdots & 0 \\ 0 & S^{(2)} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & S^{(K-1)} & 0 \\ 0 & 0 & \cdots & 0 & S^{(K)} \end{bmatrix}$$

where again the matrices $S^{(i)}$ are as defined earlier. This shows the relation between the two problems. We now show that the decomposition (6) actually yields a block Schur decomposition of the above pencil as well. Indeed the orthogonal transformations $\mathcal{Z} \doteq diag\{Z_K,\ Z_1,\ \ldots Z_{K-1}\}$ and $\mathcal{Q} \doteq diag\{Q_1,\ \ldots Q_{K-1},\ Q_K\}$ yield a pencil $\mathcal{Z}^* \cdot (\lambda\mathcal{B} - \mathcal{A}) \cdot \mathcal{Q}$ which after appropriate reordering becomes *upper block triangular* with on diagonal pencils of the type :

$$\lambda \begin{bmatrix} b^{(i)}_{K,K} & 0 & \cdots & \cdots & 0 \\ 0 & b^{(i)}_{1,1} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & b^{(i)}_{K-2,K-2} & 0 \\ 0 & 0 & \cdots & 0 & b^{(i)}_{K-1,K-1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & \cdots & 0 & a^{(i)}_{K,K} \\ a^{(i)}_{1,1} & 0 & 0 & \cdots & 0 \\ \vdots & a^{(i)}_{2,2} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & a^{(i)}_{K-1,K-1} & 0 \end{bmatrix}$$

where $^{(i)}$ indicates that the element belongs to the triangular matrices $\hat{A}_i$ or $\hat{B}_i$. For this reason the pencil $\lambda\mathcal{B} - \mathcal{A}$ is nonsingular iff $\Pi_i a^{(i)}_{j,j}/\Pi_i b^{(i)}_{j,j}$ is well defined, i.e. iff there are no zero by zero divides in two *corresponding* elements on the diagonals of the decomposition (6).

4.3 Two point boundary value problems

In the solution of two point boundary problems (not necessarily periodic), one encounters inversions of matrices of cyclic type $(\mathcal{B}+\mathcal{A})x = u$ where $\mathcal{A}$ and $\mathcal{B}$ are as above (18). Again we can apply the orthogonal transformations $\mathcal{Z}^*$ and $\mathcal{Q}$ to obtain the system of equations $\mathcal{Z}^*(\mathcal{B}+\mathcal{A})\mathcal{Q}(\mathcal{Q}^*x) = \mathcal{Z}^*u$ which essentially decomposes in n *scalar* TPBV problems. The big advantage of this is that increasing and decreasing solution in the TPBV problem have been decoupled. The periodic Schur form in fact "aligns" stable and unstable solutions at each step. The decomposition could also be computed at a coarse mesh and then "extrapolated" at finer meshes in order to avoid too much work. This is still under investigation.

5 Numerical aspects

The use of Householder and Givens transformations for all operations in the periodic QR algorithm guarentees that the obtained matrices $\hat{A}_i$ and $\hat{B}_i$ in fact correspond to slightly perturbed data as follows (indices are taken modulo K) :

$$\hat{A}_i = \overline{Z}_i^*(A_i + \delta A_i)\overline{Q}_i, \quad \hat{B}_i = \overline{Z}_i^*(B_i + \delta B_i)\overline{Q}_{i+1},$$

where $\overline{Q}_i$ and $\overline{Z}_i$ are *exactly unitary* matrices and where $\|\overline{Q}_i - Q_i\|$, $\|\overline{Z}_i - Z_i\|$, $\|\delta A_i\|/\|A_i\|$ and $\|\delta B_i\|/\|B_i\|$ are all of the order of the machine precision ϵ. This is obvious for the Hessenberg-triangular reduction and the direct deflation since each element transformed to zero can indeed be put equal to zero without affecting the ϵ bound (see [18], [8]). Things are different with the QR steps, since there one puts off-diagonal elements in A_1 equal to zero *only when these elements have converged to sufficiently small elements.* Convergence of the QR process is thus needed to guarentee stability as well. Finally, for the reordering one needs to *prove* that the swapping transformations indeed result in strictly upper triangular matrices with reversed order of eigenvalues. This is the subject of another report.

6 Concluding remarks

The above decomposition has clearly many applications and we expect that additional ones will be found in the future (e.g. in robust control of periodic systems). The above decomposition is also related to [4] which computes the Jordan chains of sequences as considered here. This *generalized QR decomposition* in fact plays the role of the rank determination (via QR or SVD) needed to reconstruct the Jordan/Kronecker structure of pencils of the type (18). This could be used as a preprocessing to eliminate the chains at $\lambda = 0$ or $\lambda = \infty$ and extract in this manner a set of smaller but invertible matrices A_i, B_i as was also done in section 3.2 via direct deflation. The advantage of this new approach is that it also identifies the structural indices at these two eigenvalues. Moreover, the generalized QR decomposition allows for *non-square matrices* as well, and one can thus consider systems of the type (2) with $m \times n$ matrices A_i and B_i.

Similar unpublished ideas are being pursued by John Hench, UC Santa Barbara (personal communication), who arrives at the same decomposition (6) with a different algorithm. His condensed form essentially consists of all A_i matrices in Hessenberg form and all B_i matrices in triangular form. We feel that the connection with the QR algorithm then fails to go through, although he reports a good convergence of that algorithm as well. Possible application to periodic continuous control systems are also being considered by him.

Acknowledgements

Part of this research was performed while the authors were visiting the Institute of Mathematics and Applications of the University of Minnesota, Minneapolis, during the summer quarter of the Applied Linear Algebra Year organized there. We greatly appreciated the hospitality and the productive atmosphere of that institute. Bojanczyk was partially supported by the Joint Services Electronics Program (Grant F49620–90–C–0039 monitored by AFOSR). G. Golub was partially supported by the Army Research (Grant DAAL03–90–G–0105) and ARGOSystems (Grant 59613 Dept. Air Force). P. Van Dooren was partially supported by the Research Board of the University of Illinois at Urbana-Champaign (Grant P 1–2–68114) and by the National Science Foundation (Grant CCR 9209349).

References

[1] B. Francis, T. Georgiou, Stability theory for linear time-invariant plants with periodic digital controllers, *IEEE Trans. Aut. Contr.* **33** (1988) 820-832.

[2] S. Bittanti, P. Colaneri, G. de Nicolao, The difference periodic Riccati equation for the periodic prediction problem, *IEEE Trans. Aut. Contr.* **33** (1988) 706-712.

[3] A. Bojanczyk, G. Golub, P. Van Dooren, The periodic Schur form. Algorithms and Applications, SCCM Intern. Rept. NA-92-07, Stanford University, August 1992.

[4] B. De Moor, P. Van Dooren, Generalizations of the singular value and QR decomposition, *SIAM Matr. Anal. & Applic.* **13** (1992).

[5] J. Doyle, B. Francis and A. Tannenbaum, *Feedback Control Theory*, McMillan, 1992.

[6] D. Flamm, A new shift-invariant representation for periodic linear systems, *Proceedings American Control Conference*, May 1990, San Diego CA, 1510-1515.

[7] J. Gardiner, A. Laub, A generalization of the matrix-sign-function solution to the algebraic Riccati equations, *Int. Journal Control* **44** (1986) 823-832.

[8] G. Golub, C. Van Loan, *Matrix Computations* 2nd edition, The Johns Hopkins University Press, Baltimore, Maryland, 1989.

[9] O. Grasselli, S. Longhi, The geometric approach for linear periodic discrete-time systems, *Lin. Algebra & Applic.* **158** (1991) 27-60.

[10] A. Laub, Invariant subspace methods for the numerical solution of Riccati equations, in *The Riccati equation*, Eds. Bittanti, Laub, Willems, Springer Verlag, 1990.

[11] C. Moler, G. Stewart, An algorithm for the generalized matrix eigenvalue problem, *SIAM Numer. Anal.* **10** (1973) 241-256.

[12] A. Sage, C. White, *Optimum Systems Control*, 2nd Ed., Prentice-Hall, New Jersey, 1977.

[13] P. Van Dooren, The generalized eigenstructure problem in linear system theory, *IEEE Trans. Aut. Contr.* **26** (1981) 111-129.

[14] P. Van Dooren, A generalized eigenvalue approach for solving Riccati equations, *SIAM Sci. & Stat. Comp.* **2** (1981) 121-135.

[15] P. Van Dooren, M. Verhaegen, On the use of unitary state-space transformations, in *Special Issue of Contemporary Mathematics on Linear Algebra and its Role in Linear System Theory*, AMS, 1985.

[16] P. Van Dooren, Numerical aspects of system and control algorithms, *Journal A* **30** (1989) 25-32.

[17] A. Varga, A Schur method for pole assignment, *IEEE Trans. Aut. Contr.* **26** (1981) 517-519.

[18] J. H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon press, Oxford, 1965.