Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Journal of Computational and Applied Mathematics 218 (2008) 482–491

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

A *QZ*-method based on semiseparable matrices☆

Yvette Vanberghen, Raf Vandebril, Marc Van Barel*

K.U.Leuven, Department of Computer Science, Celestijnenlaan 200A, 3001 Heverlee (Leuven), Belgium

Received 30 September 2006; received in revised form 22 June 2007

Abstract

This manuscript focusses on an alternative method for computing the eigenvalues of a pencil of two matrices, based on semiseparable matrices. An effective reduction of a matrix pair to lower semiseparable, upper triangular form will be presented as well as a *QZ*-iteration for this matrix pair. Important to remark is that this reduction procedure also inherits a kind of nested subspace iteration as was the case when solving the standard eigenvalue problem with semiseparable matrices. It will also be shown, that the *QZ*-iteration for a semiseparable-triangular matrix pair is closely related to the *QZ*-iteration for a Hessenberg-triangular matrix pair.

Keywords: Semiseparable; *QZ*-algorithm; Generalized eigenvalue problem

1. Introduction

The generalized eigenvalue problem for a pencil of two n by n matrices A and B consists of finding nontrivial solutions of the equation

$$Ax = \lambda Bx.$$

A standard procedure to solve this eigenvalue problem can be found for example in [7,8,4]. One first reduces, via unitary transformations the pencil $A - \lambda B$ to Hessenberg-triangular form. Following this reduction an iterative procedure, named the *QZ*-iteration is performed on this transformed couple, to bring both matrices to upper triangular form.

☆ The research was partially supported by the Research Council K.U.Leuven, OT/05/40 (Large rank structured matrix computations), Center of Excellence: Optimization in Engineering, by the Fund for Scientific Research–Flanders (Belgium), projects G.0078.01 (SMA: Structured Matrices and their Applications), G.0176.02 (ANCILA: Asymptotic aNalysis of the Convergence behavior of Iterative methods in numerical Linear Algebra), G.0184.02 (CORFU: Constructive study of Orthogonal Functions) G.0455.0 (RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation),G.0423.05 (RAM: Rational modelling: optimal conditioning and stable algorithms), and by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, project IAP VI/4 DYSCO: Dynamical Systems, Control and Optimization. The second author has a grant of "Post-doctoral Onderzoeker" from the Fund of Scientific Research Flanders (FWO-Vlaanderen).

* Corresponding author.

E-mail addresses: yvette.vanberghen@cs.kuleuven.be (Y. Vanberghen), raf.vandebril@cs.kuleuven.be (R. Vandebril), marc.vanbarel@cs.kuleuven.be (M. Van Barel)

URLs: http://www.cs.kuleuven.be/~raf (R. Vandebril), http://www.cs.kuleuven.be/~marc (M. Van Barel).

doi:10.1016/j.cam.2007.07.032

The generalized eigenvalue problem is a special case of more general polynomial matrix eigenvalue problems. A nice overview of the quadratic eigenvalue problem, solution methods and applications can be found in the overview article by Tisseur and Meerbergen [9]

Nowadays a lot of attention is paid to finding alternative approaches for solving eigenvalue and related problems, based on structured rank matrices. The first publications w.r.t. eigenvalue computations and structured rank matrices focussed on the effective computation of the eigenvalues of such matrices by reducing them to upper Hessenberg form or for symmetric matrices to tridiagonal form [1,5]. There are divide and conquer approaches [6] and methods for directly performing *QR*-steps on these structured rank matrices [11,2]. Due to the strong relation between lower semiseparable, a special kind of structured rank matrix, and Hessenberg matrices, one started to search for possible translations of the reduction to Hessenberg form towards a reduction step based on semiseparable matrices. Several algorithms were deduced, for reducing the matrices to structured rank form [10] and for performing the *QR*-algorithms on these matrices. The new 'structured rank' approach involved slightly more operations in the reduction as well as in the *QR*-step. However, this increased complexity does not necessarily lead to a global larger complexity when computing the whole or part of the spectrum. These extra involved operations create an extra convergence behavior, which can lead to deflation already in the reduction to structured rank form, leading to an immediate reduction in complexity of the following *QR*-method. A study of these results, the convergence properties inside the reduction and a global cost comparison was presented in [14].

The aim of this paper is to present an alternative method for solving the generalized eigenvalue problem, based on structured rank matrices. Instead of a Hessenberg and an upper triangular matrix as intermediate matrices, a matrix which is lower semiseparable and an upper triangular matrix will be used. We will numerically show that we can recover the specific (advantageous) convergence behavior from the standard eigenvalue problem based on structured rank matrices. Having a reduction algorithm is not sufficient for computing all the generalized eigenvalues. Hence we present also the associated *QZ*-iteration. It will be shown that there is close relation between this *QZ*-iteration and the *QZ*-iteration performed on a Hessenberg-triangular pair.

The manuscript is organized as follows. In Section 2 an algorithm is presented to reduce a pair of two general matrices (A, B) to a lower semiseparable, upper triangular matrix pair. Section 4 describes the *QZ*-iteration for a lower semiseparable-triangular matrix pair. Some numerical experiments are discussed in Section 5 followed by some conclusions.

1.1. Semiseparable matrices and notation

Before continuing the analysis of the generalized eigenvalue problem, we will briefly define semiseparable matrices, and present some notations.

Definition 1. A matrix $A \in \mathbb{C}^{n\times n}$ is called a lower semiseparable matrix, if all matrices taken out of the lower triangular part of the matrix A (including the diagonal), have rank at most 1.

The following property explains the connection between lower semiseparable and Hessenberg matrices, which will play a vital role in the remainder of this manuscript.

Proposition 2. *The inverse of an invertible lower semiseparable matrix A is a Hessenberg matrix H.*

The proof of this proposition can, for example, be found in [12,3].

Moreover, as the inverse of a nonsingular upper triangular matrix, is again an upper triangular matrix, one might want to replace the role of the Hessenberg matrix, when solving the generalized eigenvalue problem, by a lower semiseparable matrix.

In this article if we use a capital letter to denote a matrix (e.g., A) then the corresponding lower case letter with subscript *ij* refers to the (ij) entry of that matrix (e.g., a_{ij}). We will use a Matlab-like notation to denote submatrices, for example $M(i : j, k : l)$ is the submatrix consisting of the rows $i \dots j$ and columns $k \dots l$ taken out of the matrix M.

In solving the generalized eigenvalue problem $A - \lambda B$ we will need to perform several Givens transformations on the matrices. The Givens transformations that are applied to the left of the matrices, will be denoted by Q_{ij} and

Givens transformations, that are applied to the right, by Z_{kl}. Here the subscript indicates which rows or columns are involved in the transformation. When applying consecutive Givens transformations to a matrix M, we add a subscript and a superscript to the notation of the matrix in the following way $M_{ij}^{(k)}$. This means that we have performed i Givens transformations to the left and j Givens ransformations to the right of the matrix M in the kth step of the algorithm. Note that we will use M and $M^{(0)}$ or $M_{00}^{(k)}$ and $M^{(k)}$ interchangeably.

Furthermore this manuscript includes several figures of matrices. In these figures we will denote elements of the matrix by $\times$, elements that satisfy the semiseparable structure by $\boxtimes$ and elements that will be annihilated by the next transformation by $\otimes$.

2. The reduction to lower semiseparable-triangular form

The reduction of a pair of matrices A and B, to Hessenberg-triangular form, via unitary transformations Q and Z is well known, and can be found for example in [4].

We will now provide a constructive proof for the following theorem.

Theorem 3. *Suppose two matrices A, $B \in \mathbb{C}^{n\times n}$ are given. Then there exist, two unitary matrices Q and Z such that Q^*AZ is a lower semiseparable matrix, and Q^*BZ is an upper triangular matrix.*

Proof. The proof is constructive. An algorithm will be proposed for reducing the pair A and B to lower semiseparable, upper triangular form. We will consider the $n = 6$ case, as it illustrates the general case. We will start by reducing the matrix B to upper triangular form using unitary transformations. These transformations are also applied to the matrix A, in order to preserve the eigenvalues. The resulting matrices are depicted below, where $A^{(0)} = A$, $B^{(0)} = B$ and U^* denotes the transformation responsible for the triangularization of B.

$$A_{00}^{(1)} = U^* A_{00}^{(0)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times & \times \end{pmatrix}, \quad B_{00}^{(1)} = U^* B_{00}^{(0)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

Since $B^{(1)}$ already has the desired upper triangular form, it suffices to reduce the matrix $A^{(1)}$ to lower semiseparable form, while preserving the form of $B^{(1)}$. First we determine a Givens transformation to annihilate a_{61}. In multiplying both matrices to the left by this Givens transformation, we destroy the upper triangular form of the matrix $B^{(1)}$, as shown below.

$$A_{10}^{(1)} = Q_{56}^{(1)*} A_{00}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{pmatrix}, \quad B_{10}^{(1)} = Q_{56}^{(1)*} B_{00}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \otimes & \times \end{pmatrix}.$$

We can annihilate the nonzero element in the lower triangular part of $B^{(1)}$, by applying a suitable Givens transformation to the right of both matrices

$$A_{11}^{(1)} = A_{10}^{(1)} Z_{56}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{pmatrix}, \quad B_{11}^{(1)} = B_{10}^{(1)} Z_{56}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

In a similar manner, we can create additional zeros in the first column of $A^{(1)}$, while preserving the upper triangular structure of the matrix $B^{(1)}$, such that eventually we get the following result:

$$A_{54}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{pmatrix}, \quad B_{54}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

The final Givens transformation of this step, $Z_{12}^{(1)}$ annihilates b_{21} and simultaneously creates the first nontrivial rank one block in the matrix $A^{(2)}$:

$$A_{00}^{(2)} = A_{54}^{(1)} Z_{12}^{(1)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \end{pmatrix}, \quad B_{00}^{(2)} = B_{54}^{(1)} Z_{12}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

Now assume by induction, that the first three columns of A already satisfy the lower semiseparable structure and we want to create the next low rank block (i.e., add a column to the lower semiseparable structure):

$$A_{00}^{(3)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \otimes & \otimes & \otimes & \times & \times & \times \end{pmatrix}, \quad B_{00}^{(3)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

As we did before, we create zeros in the matrix $A^{(3)}$, while preserving the upper triangular form of the matrix $B^{(3)}$. Because of the rank structure already present in the matrix $A^{(3)}$, applying the first Givens transformation to the left will annihilate three elements in its bottom row, resulting in the following matrix pair:

$$A_{10}^{(3)} = Q_{56}^{(3)*} A_{00}^{(3)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{pmatrix}, \quad B_{10}^{(3)} = Q_{56}^{(3)*} B_{00}^{(3)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \otimes & \times \end{pmatrix}.$$

As in the previous step we should eliminate the nonzero element in the lower triangular part of B by applying the Givens transformation $Z_{56}^{(3)}$. The reader can verify that we can obtain the following matrix pair by applying $Z_{56}^{(3)}$ followed by the application of three additional Givens transformations $Q_{45}^{(3)*}$, $Z_{45}^{(3)}$ and $Q_{34}^{(3)*}$:

$$A_{32}^{(3)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{pmatrix}, \quad B_{32}^{(3)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \otimes & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

In eliminating the last nonzero element in the lower triangular part of the matrix $B^{(3)}$ we destroy previously created rank one blocks in the matrix $A^{(3)}$. But at the same time we do fill up part of the block of zeros in $A^{(3)}$ hereby creating part of the desired semiseparable structure.

$$A_{33}^{(3)} = A_{32}^{(3)} Z_{34}^{(3)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \otimes & \otimes & \times & \times & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times & \times \end{pmatrix}, \quad B_{33}^{(3)} = B_{32}^{(3)} Z_{34}^{(3)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

We know that, in continuing to eliminate elements in the first column(s) of $A^{(3)}$, we will destroy the upper triangular form of $B^{(3)}$. The Givens transformations that restore the upper triangular form of $B^{(3)}$ also restore the semiseparable structure of the lower triangular part of $A^{(3)}$. The following figures illustrate the restoration of part of the semiseparable structure:

$$A_{43}^{(3)} = Q_{23}^{(3)*} A_{33}^{(3)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times & \times \\ 0 & 0 & \boxtimes & \boxtimes & \times & \times \end{pmatrix}, \quad B_{43}^{(3)} = Q_{23}^{(3)*} B_{33}^{(3)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \otimes & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

$$A_{44}^{(3)} = A_{43}^{(3)} Z_{23}^{(3)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times & \times \\ 0 & \boxtimes & \boxtimes & \times & \times & \times \\ 0 & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ 0 & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ 0 & \boxtimes & \boxtimes & \boxtimes & \times & \times \end{pmatrix}, \quad B_{44}^{(3)} = B_{43}^{(3)} Z_{23}^{(3)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

Applying suitable Givens transformations $Q_{12}^{(3)*}$ and $Z_{12}^{(3)}$ concludes this step of the reduction and gives us matrices $A^{(4)}$ and $B^{(4)}$ as pictured below:

$$A^{(4)} = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \end{pmatrix}, \quad B^{(4)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

Finally we can create the last nontrivial low rank block in a similar way and hereby get the desired reduction. Compared to the reduction to Hessenberg-triangular form, the reduction to semiseparable-triangular form uses an extra number of Givens transformations to restore the semiseparable structure. This leads to an extra cost of $3n^3 + \mathrm{O}(n^2)$ operations.

At the end of each reduction step an extra column has been added to the semiseparable structure of the matrix A, while B remains upper triangular. Because of its upper triangular form, multiplying A with B^{-1} preserves the structure already present in A. This means that a reduction step of the pair (A, B) towards (semiseparable, upper triangular) form, can be seen as a reduction step of the matrix AB^{-1} towards semiseparable form. Although the reduction towards semiseparable form requires more operations than a reduction towards a Hessenberg matrix, these extra operations create an added convergence behavior, such that gaps in the spectrum are revealed during the reduction. This implies that, in the case one is only interested in the dominant eigenvalues, the reduction can be stopped once these eigenvalues are revealed. Since our reduction is in fact also a reduction of AB^{-1} to semiseparable form, it also inherits this special convergence behavior. An example of this convergence behavior will be given in Section 5. More information about the convergence properties inside the reduction and a global cost comparison can be found in [14]. In that article the results are presented for the symmetric case, but they can be directly applied to the nonsymmetric case. The only difference is that the name Lanczos-Ritz values should be replaced by Arnoldi-Ritz values.

Notice that from now on we will denote the lower semiseparable matrix A by S and the upper triangular matrix B by R. □

3. Deflation

Normally, when describing the *QZ*-iteration applied to a Hessenberg-triangular pair, one assumes that the Hessenberg matrix is unreduced. If this is not the case, deflation can be performed. We will need a similar notion of unreducibility for a lower semiseparable matrix. To this end we introduce generator representable lower semiseparable matrices, which are defined as follows.

Definition 4. A matrix $S \in \mathbb{C}^{n \times n}$ is called a generator representable lower semiseparable matrix, if there exist column vectors u and v such that the following relation is satisfied:

$$\mathrm{tril}(S) = \mathrm{tril}(uv^*).$$

In other words, we can say that a lower semiseparable matrix is generator representable, if its lower triangular part comes from a rank 1 matrix.

In this and in the following sections, we will assume that S is a generator representable lower semiseparable matrix. If this is not the case the matrix can be divided into several different blocks that are generator representable [12, Proposition 3, p. 845]. When using a Hessenberg matrix H instead of a lower semiseparable matrix, the matrices are split up if a subdiagonal element of H equals zero, because in that case an entire subdiagonal block of H is zero. In our approach, we should consider the norm of the whole subdiagonal block to determine whether or not the problem can be divided into two subproblems. To this end we use the normal deflation criterion described in [11]. A fast method to compute the norms is also described in this article. Below we will illustrate the easiest case, in which the matrix S has a zero block of dimension $(n-k) \times k$ in the lower left position. In this case $S - \kappa R$, can be divided as follows:

$$S - \kappa R = \begin{pmatrix} S_{11} - \kappa R_{11} & S_{12} - \kappa R_{12} \\ 0 & S_{22} - \kappa R_{22} \end{pmatrix}.$$

It should be clear that because of the size of the zero block, $S_{11} - \kappa R_{11}$ is a $k \times k$ block and $S_{22} - \kappa R_{22}$ a $(n-k) \times (n-k)$ block. Now it suffices to solve the two smaller problems $S_{11} - \kappa R_{11}$ and $S_{22} - \kappa R_{22}$.

Furthermore we will assume that R is nonsingular. If not, the matrices can be split up again. We will illustrate this process by means of an example. Consider the following situation, where S and R are both 6×6 matrices and the element r_{33} equals zero:

$$S = \begin{pmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes & \boxtimes \end{pmatrix}, \quad R = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}.$$

Similar to the reduction in the previous subsection, we will start by creating zeros in the matrix S, from bottom to top, all the while preserving the upper triangular form of R. We will need a total of six Givens transformations to create the situation depicted below. Three of which will be responsible for the creation of the zeros in the structured lower triangular part of S, while the remaining three see to it that R remains upper triangular and create the semiseparable structure in the lower left part of S:

$$S^{(1)} = \left(\begin{array}{ccc|ccc} \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \hline 0 & 0 & 0 & \boxtimes & \times & \times \\ 0 & 0 & 0 & \boxtimes & \boxtimes & \times \\ 0 & 0 & 0 & \boxtimes & \boxtimes & \boxtimes \end{array}\right), \quad R^{(1)} = \left(\begin{array}{ccc|ccc} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ \hline 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{array}\right).$$

We recognize the above mentioned example, where we could divide the matrices due to a block of zeros in the lower left part of S. Lets call

$$\begin{aligned}\tilde{S}_{11} &= S(1:3,1:3),\\ \tilde{S}_{22} &= S(4:6,4:6),\\ \tilde{R}_{11} &= R(1:3,1:3),\\ \tilde{R}_{22} &= R(4:6,4:6).\end{aligned}$$

The pair $(\tilde{S}_{22}, \tilde{R}_{22})$ is already in the desired form. However $\tilde{R}_{11}$ is still singular. We continue to create zeros in S, while preserving the form of R, but this time we only apply the Givens transformation to $\tilde{S}_{11}$ and $\tilde{R}_{11}$. Consider this initial situation:

$$\tilde{S}_{11} = \begin{pmatrix} \boxtimes & \times & \times \\ \otimes & \boxtimes & \times \\ \otimes & \boxtimes & \boxtimes \end{pmatrix}, \quad \tilde{R}_{11} = \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & 0 \end{pmatrix}.$$

We apply a Givens transformation to the first and the second column of $\tilde{S}_{11}$ and $\tilde{R}_{11}$. As a result the element in position $(2,1)$ of $\tilde{R}_{11}$ fills up. A Givens transformation $\tilde{Q}_{12}$ will be designed to annihilate this element:

$$\tilde{S}_{11}^{(1)} = \tilde{S}_{11}\tilde{Z}_{12} = \begin{pmatrix} \boxtimes & \times & \times \\ 0 & \boxtimes & \times \\ 0 & \boxtimes & \boxtimes \end{pmatrix}, \quad \tilde{R}_{11}^{(1)} = \tilde{R}_{11}\tilde{Z}_{12} = \begin{pmatrix} \times & \times & \times \\ \otimes & \times & \times \\ 0 & 0 & 0 \end{pmatrix},$$

$$\tilde{S}_{11}^{(2)} = \tilde{Q}_{12}^{*}\tilde{S}_{11}^{(1)} = \begin{pmatrix} \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \times \\ 0 & \otimes & \boxtimes \end{pmatrix}, \quad \tilde{R}_{11}^{(2)} = \tilde{Q}_{12}^{*}\tilde{R}_{11}^{(1)} = \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & 0 \end{pmatrix}.$$

One additional Givens transformation $\tilde{Z}_{23}$ suffices to annihilate the element in position $(3,2)$ of $\tilde{S}_{11}$, resulting in the following matrix pair:

$$\tilde{S}_{11}^{(3)} = \tilde{S}_{11}^{(2)}\tilde{Z}_{23} = \left(\begin{array}{cc|c} \boxtimes & \times & \times \\ \boxtimes & \boxtimes & \times \\ \hline 0 & 0 & \times \end{array}\right), \quad \tilde{R}_{11}^{(3)} = \tilde{R}_{11}^{(2)}\tilde{Z}_{23} = \left(\begin{array}{cc|c} \times & \times & \times \\ 0 & \times & \times \\ \hline 0 & 0 & 0 \end{array}\right).$$

The last row of $\tilde{R}_{11}^{(2)}$ consists entirely of zeros, while in the corresponding row of $\tilde{S}_{11}^{(2)}$ only the diagonal element differs from zero. Therefore we have found an infinite generalized eigenvalue. Now we can again deflate and continue with $\tilde{S}_{11}^{(2)}(2:3,2:3)$ and $\tilde{R}_{11}^{(2)}(2:3,2:3)$. The case where S is singular can be solved in a similar manner.

4. The *QZ*-iteration

In this section we will deduce a *QZ*-iteration for a pair of matrices S and R, where S is a lower semiseparable matrix and R is an upper triangular matrix. Performing one step of the *QZ*-iteration corresponds to performing one *QR*-step on the matrix $M = SR^{-1}$. Since S is a lower semiseparable matrix and R is an upper triangular matrix we know that the matrix M is lower semiseparable. To perform a *QR*-step on a semiseparable matrix we need $2n-2$ Givens transformations. These Givens transformations are applied in two consecutive steps. The first step consists of transforming the first column of $(M - \kappa I)$ to a multiple of e_1, where κ represents a chosen shift (e.g., Rayleigh, Wilkinson). The matrix $\hat{Q}$, responsible for the transformation $\hat{Q}^*(M-\kappa I)e_1 = \beta e_1$, is then applied to M, namely $\hat{Q}^* M \hat{Q}$. In the second step a matrix Q is determined implicitly such that $Qe_1 = e_1$ and $Q^*\hat{Q}^* M \hat{Q} Q$ is again a lower semiseparable matrix. Due to the implicit Q-theorem for semiseparable matrices [13], we obtain that we performed a *QR*-step on M. For more information concerning the implicit *QR*-method based on semiseparable matrices we refer the reader to [11].

There are two main goals that have to be satisfied by our *QZ*-iteration. The first one consists of performing a *QR*-step on SR^{-1}, without explicitly forming this matrix product. The second one requires that performing a *QZ*-step preserves the structure of the matrix pair (S, R). Notice that we restricted ourselves to a single shift approach, since discussing the multishift approach would be a whole manuscript in itself. Therefore it will be addressed in a future manuscript.

4.1. Part I

We want to reduce the first column of $(SR^{-1} - \kappa I)$, to a multiple of e_1. This column can be written as follows:

$$\begin{aligned} v &= (SR^{-1} - \kappa I)e_1 \\ &= Sr_{11}^{-1}e_1 - \kappa e_1 \\ &= \begin{pmatrix} r_{11}^{-1}s_{11} - \kappa \\ r_{11}^{-1}s_{21} \\ r_{11}^{-1}s_{31} \\ \vdots \\ r_{11}^{-1}s_{n-1,1} \\ r_{11}^{-1}s_{n1} \end{pmatrix}. \end{aligned}$$

It is clear that the last $n-1$ elements of v are just a multiple of the last $n-1$ elements of the first column of S. Due to the semiseparable structure of S we can easily determine Givens transformations $G_1, G_2 \ldots G_{n-1}$ such that $G_{n-1}^* \ldots G_1^* S = \hat{R}$, with $\hat{R}$ an upper triangular matrix (c.[12]). Applying these Givens transformations to the column v gives us

$$v = (\times \quad \times \quad 0 \quad \ldots \quad 0)^*.$$

By applying these Givens transformations to the matrix SR^{-1} in the following manner $G_{n-1}^* \ldots G_1^*(SR^{-1})G_1 \ldots G_{n-1}$, we have in fact performed a step of *QR* without shift on the matrix SR^{-1}. Therefore SR^{-1} is again lower semiseparable. Now, we just need one additional Givens transformation G_n to obtain the desired form, i.e., $G_n^* G_{n-1}^* \ldots G_1^* v = \beta e_1$. We do not actually compute the product SR^{-1}, rather we apply the Givens transformations $G_1, G_2 \ldots G_{n-1}$ to S and R separately, resulting in a triangular-Hessenberg pair $(S^{(1)}, R^{(1)})$. Applying the Givens transformation G_n to both $S^{(1)}$ and $R^{(1)}$ introduces a bulge in the lower triangular part of $S^{(1)}$. The chasing of the bulge happens during the next part of the *QZ*-step.

4.2. Part II

In this subsection we will again refer to the Givens transformations using Q_{ij} and Z_{kl}. The following figure illustrates the matrix pair after applying the first n Givens transformations from the first part of the *QZ*-step. We notice a disturbance in the lower triangular part of S_{01}.

$$S_{10}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}, \quad R_{10}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{pmatrix}.$$

The bulge in the lower triangular part of $S^{(1)}$, will be annihilated by the next Givens transformation as illustrated below:

$$S_{11}^{(1)} = S_{10}^{(1)} Z_{12}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}, \quad R_{11}^{(1)} = R_{10}^{(1)} Z_{12}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{pmatrix}.$$

Note that while $Z_{12}^{(1)}$ annihilated the bulge in $S^{(1)}$, it created a new bulge in $R^{(1)}$, which will have to be annihilated next

$$S_{21}^{(1)} = Q_{23}^{(1)*} S_{11}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{pmatrix}, \quad R_{21}^{(1)} = Q_{23}^{(1)*} R_{11}^{(1)} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{pmatrix}.$$

By multiplying both matrices with $Q_{23}^{(1)*}$ a new bulge is created in $S^{(1)}$ and we are back at our starting position, only now the bulge is situated lower than the original bulge. The next series of Givens transformations will chase the bulges in $S^{(1)}$ and $R^{(1)}$ down so that in the end we get the triangular-Hessenberg matrix pair $(S^{(2)}, R^{(2)})$. Lastly we have to reduce $R^{(2)}$ to upper triangular form. We create the necessary zeros by applying $n-1$ Givens transformations to the right of the matrices, from back to front. In doing so, we also create the semiseparable structure of the matrix $S^{(2)}$. This gives us a semiseparable-triangular matrix pair and ends one *QZ*-step. Such a *QZ*-step costs $18n^2 + \mathrm{O}(n)$ flops.

4.3. A second look at the QZ-algorithm

The *QZ*-iteration as developed in this section is based on the *QR*-iteration for the corresponding SR^{-1} problem. Taking a closer look at the *QZ*-iteration above, we notice that performing the first sequence of transformations from bottom to top transforms our (S, R) pair into a pair (U, H), with U and H, respectively, upper triangular and Hessenberg. Following this first sequence of Givens transformations, in fact a kind of *QZ*-step for the traditional case is performed on the matrices (U, H). Finally the last sequence of Givens transformations transforms the resulting matrices back to the semiseparable-triangular form.

In fact one can change from the (S, R) pair to a pair (U, H) and compute the eigenvalues of this pair via the traditional *QZ*-iteration. This means that we can combine the advantages of the reduction to semiseparable form and the speed and knowledge of the traditional *QZ*-iteration. One has to be careful however because w.r.t. the traditional *QZ*-iteration applied on a Hessenberg-triangular pair we work here on a triangular-Hessenberg pair, which causes in some sense an inverse convergence behavior.

To conclude we might say that the *QZ*-iteration for semiseparable-triangular matrices also uses as an intermediate step the *QZ*-iteration for Hessenberg-triangular (or triangular-Hessenberg) matrices. We remark that this is not the case in the traditional *QR*-iteration for semiseparable matrices, w.r.t. the *QR*-iteration for tridiagonal matrices. This insight opens the possibility for combining the advantages of the reduction method with the speed and knowledge of the standard *QZ*-iteration. This is work in progress.

5. Numerical experiments

In this section two numerical experiments are described. The experiments were executed in Matlab $R2006a$ on a Linux workstation. For reasons of clarity we explained the algorithms of the previous sections using full matrices. However in implementing these algorithms we used a Givens-vector representation, as defined in [12], to represent the semiseparable part of the matrices. This ensures us that the semiseparable structure of the matrix A is maintained in the presence of rounding errors. The first experiment illustrates the convergence properties of the reduction algorithm. In the second experiment, the accuracy of the eigenvalues found with our *QZ*-algorithm are compared to those found by Matlab.

For the first experiment we constructed a matrix pair (A, B) such that the generalized eigenvalues equal $1, 2, \ldots, 128, 256, 257, 258, 512, 513, 514, 515$. Notice that there are two gaps in the spectrum. After reduction, we get the semiseparable-triangular matrix pair (S, R). In the left graph of Fig. 1 we plotted the norms of the subdiagonal blocks of S. The norms of $S(129:135, 1:128)$ and $S(132:135, 1:131)$ are negligible. This implies that $S(129:131, 129:131)$ and $S(132:135, 132:135)$ can already be deflated before performing the first step of the *QZ*, this means that the reduction revealed the gaps in the spectrum. In fact if one were only interested in the largest eigenvalues, 70 steps of the algorithm would suffice to deflate $S(132:135, 132:135)$.

A set of test matrices was generated of dimensions 10×2^k for $k = 0, \ldots, 6$ for the second experiment. First we computed the generalized eigenvalues in extended precision in Matlab. Then we considered these eigenvalues to be the

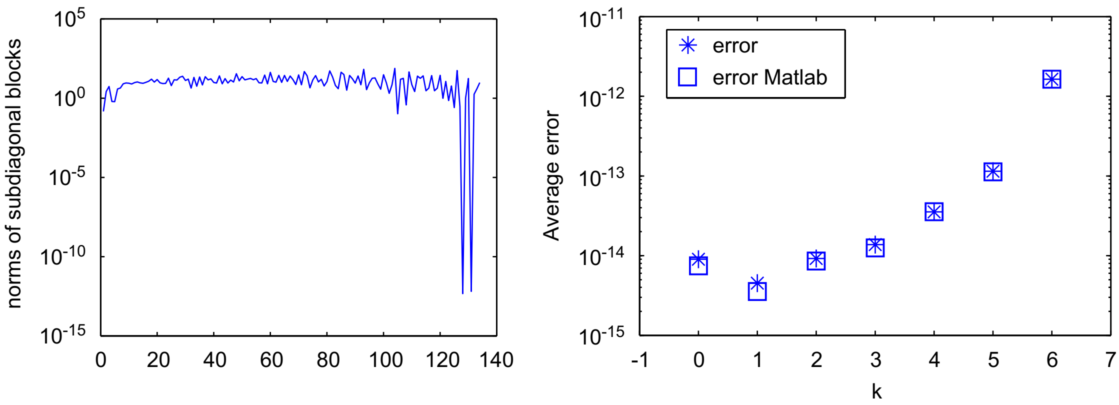

Fig. 1. The figure on the left contains the norms of the subdiagonal blocks of the matrix S. The relative errors of the eigenvalues for matrices of dimension 10.2^k are pictured in the figure on the right.

exact eigenvalues. We compared the eigenvalues found by our algorithm and the eigenvalues found by Matlab (using regular precision) to these exact eigenvalues. The average relative error of the eigenvalues for the different dimensions is pictured in Fig. 1 on the right. As can be seen on the picture the accuracy of our algorithm is comparable to the accuracy of the eig-routine of Matlab.

6. Conclusions

In this manuscript we proposed a new method for computing the generalized eigenvalues of a pair of matrices. The newly described method makes use of a semiseparable-triangular matrix pair instead of a Hessenberg-triangular matrix pair. The algorithm consists naturally of two steps. In a first step a reduction to semiseparable-triangular form is presented. Important in this reduction is a kind of inherited nested subspace iteration, which is performed on the matrices. The second step consists of applying the QZ-iteration to this pair.

References

[1] S. Chandrasekaran, M. Gu, A divide and conquer algorithm for the eigendecomposition of symmetric block-diagonal plus semi-separable matrices, Numer. Math. 96 (4) (2004) 723–731.
[2] Y. Eidelman, I.C. Gohberg, V. Olshevsky, The QR iteration method for Hermitian quasiseparable matrices of an arbitrary order, Linear Algebra Appl. 404 (2005) 305–324.
[3] M. Fiedler, Structure ranks of matrices, Linear Algebra and its Appl. 179 (1993) 119–127.
[4] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., The Johns Hopkins University Press, 1996.
[5] N. Mastronardi, S. Chandrasekaran, S. Van Huffel, Fast and stable algorithms for reducing diagonal plus semiseparable matrices to tridiagonal and bidiagonal form, BIT 41 (1) (2003) 149–157.
[6] N. Mastronardi, M. Van Barel, E. Van Camp, Divide and conquer algorithms for computing the eigendecomposition of symmetric diagonal-plus-semiseparable matrices, Numer. Algorithms 39 (4) (2005) 379–398.
[7] C.B. Moler, G.W. Stewart, An algorithm for generalized matrix eigenvalue problems, SIAM J. Numer. Anal. 10 (2) (1973) 241–256.
[8] G. W. Stewart, Matrix Algorithms, Vol. II Eigensystems. SIAM, 1999.
[9] F. Tisseur, K. Meerbergen, The quadratic eigenvalue problem, SIAM Rev. 43 (2) (2001) 235–286.
[10] M. Van Barel, R. Vandebril, N. Mastronardi, An orthogonal similarity reduction of a matrix into semiseparable form, SIAM J. Matrix Anal. Appl. 27 (1) (2005) 176–197.
[11] R. Vandebril, M. Van Barel, N. Mastronardi, An implicit QR-algorithm for symmetric semiseparable matrices, Numer. Linear Algebra Appl. 12 (7) (2005) 625–658.
[12] R. Vandebril, M. Van Barel, N. Mastronardi, A note on the representation and definition of semiseparable matrices, Numer. Linear Algebra Appl. 12 (8) (2005) 839–858.
[13] R. Vandebril, M. Van Barel, N. Mastronardi, An implicit Q theorem for Hessenberg-like matrices, Mediterranean J. Math. 2 (2005) 59–275.
[14] R. Vandebril, E. Van Camp, M. Van Barel, N. Mastronardi, On the convergence properties of the orthogonal similarity transformations to tridiagonal and semiseparable (plus diagonal) form, Numer. Math. 104 (2006) 205–239.