Implicit QR Algorithms for Palindromic and Even Eigenvalue Problems

Daniel Kressner*, Christian Schröder†, David S. Watkins‡

January 17, 2008

Abstract
In the spirit of the Hamiltonian QR algorithm and other bidirectional chasing algorithms, a structure-preserving variant of the implicit QR algorithm for palindromic eigenvalue problems is proposed. This new palindromic QR algorithm is strongly backward stable and requires less operations than the standard QZ algorithm, but is restricted to matrix classes where a preliminary reduction to structured Hessenberg form can be performed. By an extension of the implicit Q theorem, the palindromic QR algorithm is shown to be equivalent to a previously developed explicit version. Also, the classical convergence theory for the QR algorithm can be extended to prove local quadratic convergence. We briefly demonstrate how even eigenvalue problems can be addressed by similar techniques.

Keywords: palindromic eigenvalue problem, implicit QR algorithm, bulge chasing, bulge exchange, (even eigenvalue problem, convergence theory)
MSC2000: 65F15, 15A18, 15A22, 15A23

1 Introduction

Consider the *palindromic* eigenvalue problem

$$Ax = \lambda A^\star x, \tag{1}$$

eq:axlatx

where $A \in \mathbb{C}^{n\times n}$. The operator $(\cdot)^\star$ denotes either the transpose, or the conjugate transpose of a matrix. We treat both cases in a unified way.

Palindromic eigenvalue problems of the form (1) are (up to a sign) the linear case of polynomial palindromic eigenvalue problems

$$P(\lambda)x = \left(\sum_{i=0}^{k} A_i\lambda^i\right)x = 0, \text{ where } A_{k-i}^\star = A_i \in \mathbb{C}^{n\times n}, \ i = 0,\dots,k. \tag{2}$$

eq:polyeig

The name palindromic stems from the invariance of the matrix polynomial $P(\lambda)$ under reversing the order of the coefficients and (conjugate) transposing the individual matrices A_i, $i = 0,\dots,k$, a property that is shared by palindromes, i.e., words like 'rotor'. Polynomial palindromic eigenvalue problems were introduced and analyzed in [16] and arise for example in the vibration analysis of rail tracks [9]. In the following, we only consider linear palindromic problems of the form (1). From a theoretical point of view this is not a severe restriction as polynomial problems of the

*kressner@math.ethz.ch, Seminar für angewandte Mathematik, ETH Zürich, Switzerland.
†corresponding author, schroed@math.tu-berlin.de, Institut für Mathematik, MA 4-5, Technische Universität Berlin, Germany. Supported by Deutsche Forschungsgemeinschaft through MATHEON, the DFG Research Center *Mathematics for key technologies* in Berlin.
‡watkins@math.wsu.edu, Department of Mathematics, Washington State University, Pullman, WA, 99164, USA; Partly supported by Deutsche Forschungsgemeinschaft through MATHEON, the DFG Research Center *Mathematics for key technologies* in Berlin.

form (2) can, under mild assumptions, be reformulated as equivalent but larger linear palindromic problems, see [16].

Another class of structured eigenvalue problems are symmetric/skew symmetric problems of the form

$$Mx = \lambda N x, \text{ with } M = M^\star, \; N = -N^\star. \tag{3}$$

Again, there is a natural generalization to polynomial problems,

$$P(\lambda)x = \left(\sum_{i=0}^{k} A_i \lambda^i \right) x = 0, \text{ where } A_i^\star = (-1)^i A_i \in \mathbb{C}^{n \times n}, \; i = 0, \ldots, k. \tag{4}$$

Because of the relation $P(-\lambda) = P(\lambda)^\star$, these polynomials are called *even*. As for the palindromic case, a polynomial even eigenvalue problem can, under mild assumptions, be reformulated as a linear even eigenvalue problem (3), see [16].

Even and palindromic eigenvalue problems can be related to each other via the generalized Cayley transformation [16], which maps a generalized eigenvalue problem $Ax = \lambda Bx$ into

$$(A + B)x = \frac{\lambda + 1}{\lambda - 1}(A - B)x.$$

The associated matrix pairs (A, B) and $(A + B, A - B)$ have the same eigenvectors and the eigenvalues are transformed as $\lambda \mapsto \frac{\lambda+1}{\lambda-1}$. The generalized Cayley transform of an even problem, $(M + N)x = \frac{\lambda+1}{\lambda-1}(M - N)x$, is palindromic and vice versa.

The structure in the coefficient matrices of (1) and (3) induces symmetries in the spectrum. Indeed, transposing the palindromic problem (1) yields $x^\star A = \frac{1}{\lambda^\star} x^\star A^\star$. Hence, if λ is an eigenvalue with (right) eigenvector x then also $\frac{1}{\lambda^\star}$ is an eigenvalue with left eigenvector $x^\star$. This pairing also holds for a zero eigenvalue – its counterpart is an infinite eigenvalue. From here on, we use the convention $\frac{1}{0} = \infty$ in order to unify the treatment of finite and infinite eigenvalues. Also the number and sizes of Jordan blocks corresponding to the eigenvalues λ and $\frac{1}{\lambda^\star}$ coincide. This follows from the structured Kronecker canonical forms for palindromic eigenvalue problems presented in [23] or (as canonical forms under congruence) in [10, 11, 21].

Analogously, (conjugate) transposing (3) yields $x^\star M = -\lambda x^\star N$. Hence, the eigenvalues of an even problem come in pairs $(\lambda, -\lambda^\star)$. Canonical forms of symmetric/ skew symmetric [30], or Hermitian skew-Hermitian [29] pencils under congruence show that also the number and sizes of Jordan blocks corresponding to λ and $-\lambda^\star$ coincide.

The existence of symmetries in the spectrum of palindromic and even eigenvalue problems calls for algorithms that preserve these symmetries. Due to rounding errors, the eigenvalues computed by a general-purpose method (such as the QZ algorithm [8]) are in general *not* paired. To avoid this effect, a structure-preserving method is needed. Ideally, the method is strongly backward stable, i.e., it returns computed eigenvalues that are the exact ones of a slightly perturbed *structured* eigenvalue problem.

There are a number of structure preserving-methods for palindromic eigenvalue problems available. Implicitely based on the generalized Cayley transform, the URV-like algorithm described in [25] is suitable for computing eigenvalues but not immediately for computing eigenvectors or invariant subspaces. The Jacobi-style method in [9] is strongly backward stable for both eigenvalues and eigenvectors, but requires much more computing time than the QZ algorithm; a disadvantage that is common to all nonsymmetric Jacobi methods [18]. Alternatively, a variant of the Laub trick [17] can be used to post-processes the output of the QZ algorithm, enforcing a small structured backward error. In the presence of near-exceptional eigenvalue, this computationally quite inexpensive approach breaks down and must be combined with a strongly backward stable method [17]. Other alternatives arise from the fact that palindromic eigenvalue problems of even size can be transformed into symplectic eigenvalue problems; for example the structured doubling algorithm can be adapted to also solve (1), see [6].

None of the methods described above is ideal in the sense of being strongly backward stable *and* requiring less computing time than the QZ algorithm. The goal of this paper is to approach this ideal by continuing the work from [24], where an explicit, structure-preserving QR-like algorithm was presented. We derive an implicit version of this algorithm, bringing the complexity from $\mathcal{O}(n^4)$ down to $\mathcal{O}(n^3)$. Our implicit QR-like algorithm is much in the spirit of Byers' Hamiltonian QR algorithm [4] and its symplectic variant [19]. It also shares their fundamental limitation of being only applicable to problems for which a structured Hessenberg(-triangular) form is known a priori or can be computed within reasonable complexity.

The rest of this paper is structured as follows. Section 2 revisits the explicit palindromic QR algorithm. Section 3 introduces a Hessenberg-like matrix form. Our implicit algorithm, which preserves this form, is presented in Section 4. In Section 5 we show that both the explicit and implicit algorithms produce essentially the same result in exact arithmetic. The existing convergence theory for the QZ algorithm is extended in Section 6 to cover our algorithm. In Section 7, we discuss modifications to adequately address real palindromic problems. The extension to even eigenvalue problems is briefly described in Section 8.

1.1 Notation

A matrix A is called *antitriangular*, if $a_{ij} = 0$, whenever $i + j \leq n$. An antitriangular matrix is depicted by $A =$ ◿. The identity matrix of order n is denoted by I_n, its columns, the unit vectors, are denoted by $e_1, e_2, \ldots, e_n$. $F_n = [e_n, \ldots, e_2, e_1]$ denotes the flip matrix. The subscript is omitted if clear from context. The set of eigenvalues of a matrix A or a pencil (B, C) is denoted by $\Lambda(A)$ and $\Lambda(B, C)$, respectively.

We use MATLAB notation for submatrices, e.g., $A(i : j, :)$ denotes the rows i to j, whereas $A(:, k : l)$ selects the columns k through l. Moreover, $A(i : j, k : l)$ denotes the rows i to j of the columns k through l.

2 The palindromic Schur form and the palindromic QR step

For the rest of this paper, we will assume that the matrix pair $(A, A^\star)$ is regular, i.e., $\det(A - \lambda A^\star) \not\equiv 0$. The following result provides a structured Schur form which is the basis and aim of the palindromic QR algorithm. Note that we call an eigenvalue λ *exceptional* if $\lambda = 1/\lambda^\star$.

th:palschur **Theorem 1** *Let $A \in \mathbb{C}^{n,n}$ be such that $(A, A^\star)$ is a regular matrix pair with at most one exceptional eigenvalue. Then there exists a unitary matrix Q such that*

$$Q^\star A Q = T = \text{◿}. \tag{5}$$

eq:palschur

Proof: A proof can be found in [24]. Note that the case $\star = T$ case was first proved in [17]. ■

Given a decomposition (5), the eigenvalues of $(A, A^\star)$ coincide with those of $(T, T^\star)$ and can thus be read off from the antidiagonal of T:

$$\lambda_i = \frac{t_{n+1-i,i}}{t^\star_{i,n+1-i}}, \quad \text{for } i = 1, \ldots, n. \tag{6}$$

eq:lamratio

Further, for each $k = 1, 2, \ldots, n$ it holds that the first k columns of Q span a right deflating subspace [28] belonging to the eigenvalues $\lambda_1, \ldots, \lambda_k$. The corresponding left deflating subspace is spanned by the last k columns of $Q^{-\star}$. Note that, after a suitable reordering, the matrix Q in (5) can be chosen such that any prespecified eigenvalue order $\lambda_1, \lambda_2 \ldots \lambda_n$ appears on the antidiagonal of T, provided of course that $\lambda_i^\star = \frac{1}{\lambda_{n+1-i}}$ is satisfied for $i = 1, \ldots, n$.

Definition 2 *Let $A \in \mathbb{C}^{n,n}$ be such that $(A, A^\star)$ is regular. A matrix $T \in \mathbb{C}^{n,n}$ is called a* palindromic Schur form *of A, if T is antitriangular and there exists a unitary $Q \in \mathbb{C}^{n,n}$ such that $T = Q^\star A Q$.*

The main algorithm for solving the dense standard eigenvalue problem $Ax = \lambda x$ is the QR algorithm. In its simplest form one step consists of computing a QR factorization $A = QR$ and forming the next iterate $A_+ = Q^H AQ$. In [24] the QR algorithm was adapted to the palindromic problem (1). In the basic palindromic QR step the next iterate A_+ is formed as $A_+ = Q^\star AQ$, where the unitary matrix Q stems from an anti-QR factorization $A = Q^{-\star}R$ with $R =$. Note that, because Q is unitary, $Q^{-\star}$ is simply Q or $\bar{Q}$, respectively. In the case that all eigenvalues have different absolute value the iteration converges linearly to palindromic Schur form [34].

Shifts can be incorporated as follows. Given numbers $(\kappa_1, \ldots, \kappa_k)$, none of which is an exact eigenvalue of (1), define the *shift polynomial* $p(\lambda)$, its $\star$-reversal $q(\lambda)$ and the rational function $r(\lambda)$ via

$$p(\lambda) = \prod_{i=1}^{k}(\lambda - \kappa_i), \quad q(\lambda) = \prod_{i=1}^{k}(\kappa_i^\star \lambda - 1), \quad r(\lambda) = p(\lambda)/q(\lambda). \tag{7}$$ eq:starreversal

Note that $r(A^{-\star}A) = \prod_{i=1}^{k}(\kappa_i^\star A - A^\star)^{-1}(A - \kappa_i A^\star)$ and hence the expression $r(A^{-\star}A)$ and its inverse are always well-defined, even if A itself is not invertible. We define an *explicit palindromic QR step* with shifts $(\kappa_1, \ldots, \kappa_k)$ applied to A as

$$A_+ = Q^\star AQ, \tag{8}$$ eq:expalqrstep

where the unitary matrix Q stems from the QR decomposition

$$r(A^{-\star}A) = QR, \quad R = \ . \tag{9}$$ eq:q1explicit

Note that in [24] $p(\lambda)$ and $q(\lambda)$ were admitted to have independent coefficients, and the above step was called a "palindromic QR step with the $2k$ shifts $\kappa_1, \kappa_1, \ldots, \kappa_k, \kappa_k$". Forcing q to be the $\star$-reversal of p gives the relation

$$r(A^{-\star}A) = r(AA^{-\star})^{-\star}, \tag{10}$$ eq:rataraat

as is easily checked. Together with (9) this implies

$$r(AA^{-\star}) = (r(A^{-\star}A))^{-\star} = (QR)^{-\star} = Q^{-\star}R^{-\star} = Q^{-\star}L,$$

where $L = R^{-\star}$ is lower triangular. Thus, the palindromic QR step effects a standard QR step on $A^{-\star}A$ and a standard QL step on $AA^{-\star}$ driven by the same function r. This mixed behavior can be explained by the fact that if A is in palindromic Schur form, i.e., it is antitriangular, then $A^{-\star}A$ is upper triangular, whereas $AA^{-\star}$ is lower triangular.

It follows from (10) that

$$A_+ = Q^\star AQ = R^{-\star}r(A^{-\star}A)^\star Ar(A^{-\star}A)R^{-1} = R^{-\star}Ar(AA^{-\star})^\star r(A^{-\star}A)R^{-1} \overset{(10)}{=} R^{-\star}AR^{-1}. \tag{11}$$ eq:rtar

Thus, if A has a structure that is invariant under $\star$-congruence transformations $\tilde{R}^\star A\tilde{R}$ with arbitrary upper triangular matrices $\tilde{R}$, then this structure is preserved by palindromic QR steps. One such structure is the following variant of Hessenberg matrices.

Definition 3 *A matrix $A \in \mathbb{C}^{n,n}$ is called an* anti-Hessenberg matrix*, if $a_{ij} = 0$ whenever $i + j < n$. An anti-Hessenberg matrix is depicted by $A =$.*

In this paper we only consider the case that A is already in anti-Hessenberg form. In contrast to the standard Hessenberg from, there is no direct method known that reduces a general matrix to anti-Hessenberg form by a unitary $\star$-congruence transformation. However, several applications where this matrix shape arises naturally or can be computed with reasonable complexity are listed in [24].

3 Anti-Hessenberg-triangular matrices

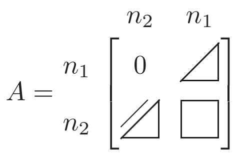
sec:semianti

Somewhat surprisingly, an anti-Hessenberg matrix can be transformed into an even further condensed form that remains invariant under palindromic QR steps. In this section, we introduce such a form along with an algorithm for transforming an anti-Hessenberg matrix into this form.

def:semiantihess

Definition 4 *An anti-Hessenberg matrix $A \in \mathbb{C}^{n,n}$ is called an* anti-Hessenberg-triangular matrix*, if $a_{i,n-i} = 0$ for $i = 1, \ldots, n_1 := \lfloor \frac{n-1}{2} \rfloor$. It is called* unreduced*, if $a_{i,n-i} \neq 0$ for $i = n_1+1, \ldots, n-1$ and $a_{i,n-i+1} \neq 0$ for $i = 1, \ldots, n_1$.*

An anti-Hessenberg-triangular matrix can be depicted by



Lemma 5 *The anti-Hessenberg-triangular structure is invariant under palindromic QR steps.*

Proof: This structural invariance also follows from (11), as the anti-Hessenberg-triangular structure is invariant under $\star$-congruence transformations with an upper triangular matrix. ■

We now show that any matrix in anti-Hessenberg form can be transformed to anti-Hessenberg-triangular form by a unitary $\star$-congruence transformation. To this end, let us assume that A is in anti-Hessenberg form and we have already annihilated the $m-1 \leq n_1 - 2$ super-antidiagonal elements $a_{i,n-i}$, $i = 1, \ldots, m-1$, in the top right corner. To explain the idea of the reduction, we consider only the following submatrix in the middle of A:

$$A(m-1:n-m+2, m-1:n-m+2) = \begin{bmatrix} & & & & 0 & x \\ & & & x & x & x \\ & & \iddots & & & \vdots \\ & x & & x & x & x \\ x & x & & x & x & x \\ x & x & \cdots & x & x & x \end{bmatrix}.$$

Our aim is to annihilate the next super-antidiagonal element $a_{m,n-m}$. This can be achieved by applying a Givens rotation to columns $n-m, n-m+1$. However, in order to preserve $\star$-congruence, we have to apply this rotation also to rows $n-m, n-m+1$, resulting in an additional nonzero entry at position $(n-m, m-1)$:

$$A(m-1:n-m+2, m-1:n-m+2) \leftarrow \begin{bmatrix} & & & & & x \\ & & & \mathbf{0} & x & x \\ & & \iddots & & & \vdots \\ + & x & & x & x & x \\ x & x & & x & x & x \\ x & x & \cdots & x & x & x \end{bmatrix}.$$

This fill-in is annihilated by a rotation in columns $m-1, m$. The corresponding row rotation introduces a nonzero at $(n-m+1, m)$.

$$A(m-1:n-m+2, m-1:n-m+2) = \begin{bmatrix} & & & & + & x \\ & & & 0 & x & x \\ & & \iddots & & & \vdots \\ \mathbf{0} & x & & x & x & x \\ x & x & & x & x & x \\ x & x & \cdots & x & x & x \end{bmatrix}$$

Thus, both rotations together have moved the unwanted nonzero one position to the top right corner. Hence, repeated application of this process moves the nonzero $(1, n-1)$ position, where it can be annihilated by a rotation in the last two columns. The corresponding row rotation introduces no fill-in.

The described procedure leads to the following algorithm.

lg:antihess2semi

Algorithm 1 Reduction to anti-Hessenberg-triangular form

Input: $A \in \mathbb{C}^{n,n}$ in anti-Hessenberg form

Output: A is overwritten by $\hat{A} = Q^\star A Q$ in anti-Hessenberg-triangular form

for $m = 1 : \lfloor \frac{n-1}{2} \rfloor$ **do**
  for $j = m : -1 : 2$ **do**
    define rotation G such that $A(j, n-j : n-j+1)G = [0, *]$
    $A(j : n, n-j : n-j+1) \leftarrow A(j : n, n-j : n-j+1)G$
    $A(n-j : n-j+1, j-1 : n) \leftarrow G^\star A(n-j : n-j+1, j-1 : n)$
    define rotation G such that $A(n-j, j-1 : j)G = [0, *]$
    $A(n-j : n, j-1 : j) \leftarrow A(n-j : n, j-1 : j)G$
    $A(j-1 : j, n-j : n) \leftarrow G^\star A(j-1 : j, n-j : n)$
  end for
  define rotation G such that $A(1, n-1 : n)G = [0, *]$
  $A(:, n-1 : n) \leftarrow A(:, n-1 : n)G$
  $A(n-1 : n, :) \leftarrow G^\star A(n-1 : n, :)$
end for

Algorithm 1 requires approximately $\frac{3}{2}n^3$ flops. Forming Q as the product of all employed Givens rotations takes another $\frac{1}{2}n^3$ flops, provided that the nonzero pattern of Q is exploited during the accumulation. In contrast, updating an existing full matrix Q by the Givens rotations costs $\frac{3}{2}n^3$ flops.

The above reduction to anti-Hessenberg-triangular form resembles the reduction of a pair of Hessenberg matrices to Hessenberg-triangular form GolV96 [8]. To see this, set $H = FA(\lfloor \frac{n}{2} \rfloor + 1 : n, 1 : \lceil \frac{n}{2} \rceil)$ and $R_1 = FA(1 : \lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor + 1 : n)^\star$, where F denotes the flip matrix of dimension $\lceil \frac{n}{2} \rceil$. Note that the submatrices H and R_1 overlap in the case of odd n. Since A is in anti-Hessenberg form, both, H and R_1, are in upper Hessenberg form. The pencil (H, R_1) can be transformed to Hessenberg-triangular form by subsequently chasing the subdiagonal entries in R_1 to the top left corner, where they are annihilated without introducing additional fill-in. Interpreting these transformations on H and R_1 as transformations on A yields the anti-Hessenberg-triangular reduction.

Example 1 We illustrate Algorithm 1 for the reduction of a 5-by-5 anti-Hessenberg matrix and the corresponding transformations on the flipped submatrices H and R_1.

$$A \quad \begin{bmatrix} & & & x & x \\ & & x & x & x \\ & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \rightarrow \begin{bmatrix} & & & \mathbf{0} & x \\ & & x & x & x \\ & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \rightarrow \begin{bmatrix} & & & & x \\ & & \mathbf{0} & x & x \\ + & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \rightarrow \begin{bmatrix} & & & + & x \\ & & & x & x \\ \mathbf{0} & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \rightarrow \begin{bmatrix} & & & \mathbf{0} & x \\ & & & x & x \\ & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix}$$

$$H \quad \begin{bmatrix} x & x & x \\ x & x & x \\ & x & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ x & x & x \\ & x & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ x & x & x \\ + & x & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ x & x & x \\ \mathbf{0} & x & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ x & x & x \\ & x & x \end{bmatrix}$$

$$R_1 \quad \begin{bmatrix} x & x & x \\ x & x & x \\ & x & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ \mathbf{0} & x & x \\ & x & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ & x & x \\ & \mathbf{0} & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ + & x & x \\ & & x \end{bmatrix} \qquad \begin{bmatrix} x & x & x \\ \mathbf{0} & x & x \\ & & x \end{bmatrix}$$

□

The described relationship already indicates some resemblance between the palindromic QR algorithm applied to A and the QZ algorithm applied (H, R_1). This relation will allow us to extend existing results on the QZ algorithm partially to the palindromic QR algorithm. However, it is important remark that – due to the coupling nonzero element in the middle of the antidiagonal of A – the eigenvalue problems $(A, A^\star)$ and (H, R_1) are *not* equivalent.

sec:deflation

3.1 Deflation

An anti-Hessenberg-triangular matrix that is not unreduced in the sense of Definition 4 can be deflated to yield unreduced eigenvalue problems of smaller order. This is easily seen when $a_{k,n-k} = 0$ for some $k \in [n_1+1, n-1]$, since A can then be partitioned as

$$
A = \begin{array}{c} \\ l \\ 2k-n \\ n-k \end{array}
\begin{array}{c} \begin{array}{ccc} n-k & 2k-n & n-k \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{array}\right] \end{array},
$$

and the eigenvalues of $(A, A^\star)$ are obtained from the generalized eigenvalue problem $(A_{31}, A_{13}^\star)$ and the palindromic eigenvalue problem $(A_{22}, A_{22}^\star)$.

In the case of a zero antidiagonal element, $a_{k,n-k+1} = 0$ for $k = 1, \dots, n_1$, we can adapt the standard algorithm for the deflation of infinite eigenvalues GolV96 [8] (making use of the relation to the QZ algorithm described above), to move this zero to position $k = 1$ by a unitary $\star$-congruence transformation and additionally force $a_{n-1,1} = 0$. Partitioning

$$
A = \begin{array}{c} \\ l \\ n-1 \\ 1 \end{array}
\begin{array}{c} \begin{array}{ccc} 1 & n-1 & 1 \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{array}\right] \end{array},
$$

the eigenvalues of (A, A^*) are given by $0, \infty$, and the eigenvalues of $A_{22}, A_{22}^\star$.

In practice, it will be a rare event that the entries $a_{k,n-k}$ or $a_{k,n-k+1}$ become exactly zero. Following Wil88,KagK06 [35, 12], we regard such an entry as zero if it is not larger than machine precision times the sum of the absolute values of the neighbouring elements.

4 The Implicit palindromic QR step

The explicit palindromic QR step (8) involves the formation and QR factorization of $r(A^{-\star}A)$, where r is a rational function. Both operations are prohibitive for stability and efficiency reasons. In Sch07 [24] a method was presented that does not require inverses, but works directly with the factors $A - \kappa_i A^\star$. Still, the anti-Hessenberg structure of A is only implicitely preserved during the algorithm. This implies that roundoff error will intrude the zero part of A, yielding either computational complexity $\mathcal{O}(n^4)$ or, if the top left part of A is explicitely set to zero after each iteration, loss of numerical stability. In the following, we present an implicit version of the palindromic QR step that avoids this drawback.

Consider a set of k shifts $\kappa_1, \dots, \kappa_k \in \mathbb{C}$, $k < \frac{n}{2}$, that is assumed to be $\star$-reciprocal free, i.e., $\kappa_i \neq 1/\kappa_j^\star$ for any $1 \le i, j \le k$. Set $p(\lambda) = \prod_{i=1}^k (\lambda - \kappa_i)$. Let us illustrate the proposed implicit algorithm for $k = 2$ shifts. Only the bottom left and top right 5×5 corners of A are displayed:

$$
A: \left[\begin{array}{cccccc}
 & & & & \iddots & \iddots \\
 & & & x & x & \\
 & & x & x & x & \\
 & x & x & x & x & \\
x & x & x & x & x & \\
x & x & x & x & x & \cdots
\end{array}\right]
\left[\begin{array}{cccccc}
 & & & & & x \\
 & & & & x & x \\
 & & & x & x & x \\
 & & x & x & x & x \\
 & x & x & x & x & x \\
\iddots & & & & & \vdots
\end{array}\right]
$$

Consider the first column of the shift polynomial,

$$
x = A^\star p(A^{-\star}A)e_1 = p(AA^{-\star})A^\star e_1 = \alpha p(AA^{-\star})e_n
$$

for some constant α. Since the last $\lfloor \frac{n-1}{2} \rfloor$ columns of $AA^{-\star}$ are in lower Hessenberg form, only the last $k+1$ elements of x are nonzero. Let Q_0 be a unitary matrix (e.g., a Householder transformation) such that $Q_0^\star x$ is mapped to a scalar multiple of e_n and apply the congruence $A_0 = Q_0^\star A Q_0$.

$$A_0 : \left[\begin{array}{cccccc} & & & & \iddots & \iddots \\ & & & x & x & \\ & & x & x & x & \\ + & x & x & x & x & \\ x & x & x & x & x & \\ x & x & x & x & x & \cdots \end{array}\right] \left[\begin{array}{cccccc} & & & + & + & x \\ & & & + & x & x \\ & & & x & x & x \\ & & x & x & x & x \\ & x & x & x & x & x \\ \iddots & & & & & \vdots \end{array}\right]. \tag{12}$$

eq:a0bulges

Following the usual convention, the caused fill-in is referred to as *bulges*. We will now show how to chase these bulges towards the middle of the antidiagonal.

To this end, let $\tilde{Q}_{1/2}$ be a unitary matrix, such that the first row of $\tilde{Q}_{1/2}^\star A_0(1:3, n-2:n)$ is a multiple of $e_3^\star$. This could be the unitary factor of the anti QR factorization of $A_0(1:3, n-1:n)$ or an opposite Householder transformation, which was introduced in [34], and shown to be backward stable in [14, Section 2.3.3]. Imbedding $\tilde{Q}_{1/2}$ into $Q_{1/2} = \tilde{Q}_{1/2} \oplus I_{n-3}$ and applying the congruence $A_{1/2} = Q_{1/2}^\star A_0 Q_{1/2}$ yields

$$A_{1/2} : \left[\begin{array}{cccccc} & & & & \iddots & \iddots \\ & & & x & x & \\ + & + & x & x & x & \\ x & x & x & x & x & \\ x & x & x & x & x & \\ x & x & x & x & x & \cdots \end{array}\right] \left[\begin{array}{cccccc} & & & \mathbf{0} & \mathbf{0} & x \\ & & & x & x & x \\ & & & x & x & x \\ & & x & x & x & x \\ & x & x & x & x & x \\ \iddots & & & & & \vdots \end{array}\right].$$

Next, the first column of $A_{1/2}$ is reduced. To this end, let $\tilde{Q}_1$ be a unitary matrix such that $\tilde{Q}_1^\star A_{1/2}(n-4:n-1)$ is a multiple of e_3. Applying this $\tilde{Q}_1$ to rows and columns $n-4:n-1$ results in the following nonzero pattern:

$$A_1 : \left[\begin{array}{cccccc} & & & & \iddots & \iddots \\ & & & x & x & \\ \mathbf{0} & x & x & x & x & \\ \mathbf{0} & x & x & x & x & \\ x & x & x & x & x & \\ x & x & x & x & x & \cdots \end{array}\right] \left[\begin{array}{cccccc} & & & & & x \\ & & + & x & x & x \\ & & + & x & x & x \\ & & x & x & x & x \\ & x & x & x & x & x \\ \iddots & & & & & \vdots \end{array}\right].$$

Comparing A_0 with A_1 indicates that the two transformations just described moved the bulges indeed one step to the middle of the antidiagonal. Repeating the described procedure effects another step:

$$A_{1+1/2} : \left[\begin{array}{ccccccc} & & & & & \iddots & \iddots \\ & + & + & x & x & & \\ & x & x & x & x & & \\ & x & x & x & x & & \\ x & x & x & x & x & & \\ x & x & x & x & x & \cdots & \end{array}\right] \left[\begin{array}{cccccc} & & & & & x \\ & & \mathbf{0} & \mathbf{0} & x & x \\ & & x & x & x & x \\ & & x & x & x & x \\ & x & x & x & x & x \\ \iddots & & & & & \vdots \end{array}\right]$$

$$A_2 : \left[\begin{array}{ccccccc} & & & & & \iddots & \iddots \\ & \mathbf{0} & x & x & x & & \\ & \mathbf{0} & x & x & x & & \\ & x & x & x & x & & \\ x & x & x & x & x & & \\ x & x & x & x & x & \cdots & \end{array}\right] \left[\begin{array}{cccccc} & & & & & x \\ & & & & x & x \\ & + & x & x & x & x \\ & + & x & x & x & x \\ & x & x & x & x & x \\ \iddots & & & & & \vdots \end{array}\right]$$

After $n_1 - k$ steps the bulges arrive at the center. Note the structural difference depending on whether n is even or odd:

$$
A_{n_1-k}: \quad
\underset{}{\overset{\underline{n \text{ even}}}{\left[\begin{array}{ccccccccccc}
 & & & & & & & & & \iddots \\
 & & & & & & & 0 & x & \cdots \\
 & & & & & x & x & x & x & \\
 & & & & & x & x & x & x & \\
 & & & & x & x & x & x & x & \\
 & & x & x & x & x & x & x & x & \\
 & & x & x & x & x & x & x & x & \\
 & x & x & x & x & x & x & x & x & \\
\iddots & x & x & x & x & x & x & x & x & \cdots \\
\iddots & \vdots & & & & & & & \vdots & \ddots
\end{array}\right]}}
\quad
\overset{\underline{n \text{ odd}}}{\left[\begin{array}{cccccccccc}
 & & & & & & & & \iddots \\
 & & & & & & 0 & x & \cdots \\
 & & & & x & x & x & x & \\
 & & & & x & x & x & x & \\
 & & x & x & x & x & x & x & \\
 & & x & x & x & x & x & x & \\
 & x & x & x & x & x & x & x & \\
\iddots & x & x & x & x & x & x & x & \cdots \\
\iddots & \vdots & & & & & & \vdots & \ddots
\end{array}\right]}
\tag{13}
$$

eq:centerbulge

At this point the bulge chase cannot be continued as before, as the bulges would interfer with each other. In order to explain how to go on, we require the notion of bulge pairs.

We have seen that the sequence of transformed matrices A_i contains two bulges. One bulge is located in the submatrix

$$B_i := A_i(n-k-i : n-1-i, i+1 : i+k) \tag{14}$$

eq:defb

and moves upwards. The other bulge is located at

$$C_i := A_i(i+1 : i+k, n-k-i : n-1-i) \tag{15}$$

eq:defc

and moves downwards. Considering the matrix pair $(A_i, A_i^\star)$ we thus have an upwards moving bulge pair $(B_i, C_i^\star)$ and a downwards moving bulge pair $(C_i, B_i^\star)$. The bulge pairs show an important invariance property.

lem:bulges

Lemma 6 *Assume that the anti-Hessenberg-triangular matrix A is unreduced. Then the eigenvalues of $(B_i, C_i^\star)$ are given by the shifts $\kappa_1, \dots, \kappa_k$ for all $i = 0, \dots, n_1 - k$. Consequently, the eigenvalues of $(C_i, B_i^\star)$ are $1/\kappa_1^\star, \dots, 1/\kappa_k^\star$ for all $i = 0, \dots, n_1 - k$.*

Proof: Let $H = FA(\lfloor \frac{n}{2} \rfloor + 1 : n, 1 : \lceil \frac{n}{2} \rceil)$ and $R_1 = FA(1 : \lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor + 1 : n)^\star$, where F denotes the flip matrix of appropriate dimension. Then (H, R_1) is a matrix pair in unreduced Hessenberg-triangular form. Note that the submatrices H and R_1 include a common element from A for odd n. This does not pose a problem, however, since the bulge chase is stopped before the bulges touch this overlap, see (13). Analogous to the discussion in Section 3, the palindromic bulge chase applied to A can be shown to be equivalent to a bulge chase in the implict QZ step applied to (H, R_1) stopped just before squeezing out the bulges, see [32] for more details. Also, the bottom part of x can be written as $x(\lfloor \frac{n}{2} \rfloor + 1 : n) = FR_1 p(R_1^{-1} H)e_1 = \alpha^\star F p(HR_1^{-1})e_1$, the flipped first column of the shift polynomial in the QZ step.

Because of this equivalence, the statement of the lemma follows from the corresponding result for the QZ algorithm, see [32, Section 7.1]. Note that in the terminology of [32], the pencils $(B_i, C_i^\star)$ and $(C_i, B_i^\star)$ are *intermediate* bulge pencils. ∎

Lemma 6 shows that the upward moving bulge carries the shifts and the downward moving bulge their (conjugated) reciprocals. In order to allow this shift transport mechanism to continue, the bulges have to be “passed through each other”, i.e., we need a matrix A_{n_1+k-1} that has the same shape as A_{n_1-k} such that the eigenvalues of $(B_{n_1+k-1}, C^\star_{n_1+k-1})$, which is the upper bulge now, are still the shifts. Of course this also implies that the eigenvalues of the lower bulge $(C_{n_1+k-1}, B^\star_{n_1+k-1})$ are $1/\kappa_1^\star, \dots, 1/\kappa_k^\star$. Note that the jump in the index from A_{n_1-k} to A_{n_1+k-1}

is needed in order to be consistent with the definition (14)–(15) of B_i, C_i. We postpone the description of a method for computing such a matrix A_{n_1+k-1} to Section 4.1.

Once the bulges are exchanged, they have to be chased out again. This process is exactly the opposite of the process presented above and we thus refrain from discussing it in detail.

Summarizing the discussion so far yields the following algorithm.

alg:palqrstep

Algorithm 2 Implicit Palindromic QR step

Input: $A \in \mathbb{C}^{n,n}$ in unreduced anti-Hessenberg-triangular form, $\star$-reciprocal free shifts $\kappa_1, \ldots, \kappa_k$, $k < \frac{n}{2}$
Output: one palindromic QR step is applied to A

$n_1 = \lfloor \frac{n-1}{2} \rfloor$
compute $x = A^\star p(A^{-\star}A)e_1$ *% create bulges*
define unitary $\tilde{Q}$ with $\tilde{Q}^\star x(n-k:n) = \alpha e_{k+1}$
$A(n-k:n,:) \leftarrow \tilde{Q}^\star A(n-k:n,:)$
$A(:,n-k:n) \leftarrow A(:,n-k:n)\tilde{Q}$
for $i = 0 : n_1 - k - 1$ **do** *% chase bulges towards middle*
    define unitary $\tilde{Q}$ with $e_1^T \tilde{Q}^\star A(i+1:i+k+1, n-k-i:n-i) = \alpha e_{k+1}^T$
    $A(i+1:i+k+1, n-k-i:n) \leftarrow \tilde{Q}^\star A(i+1:i+k+1, n-k-i:n)$
    $A(n-k-i-1:n, i+1:i+k+1) \leftarrow A(n-k-i-1:n, i+1:i+k+1)\tilde{Q}$
    define unitary $\tilde{Q}$ with $\tilde{Q}^\star A(n-k-1-i:n-1-i, i+1) = \alpha e_{k+1}$
    $A(n-k-1-i:n-1-i, i+1:n) \leftarrow \tilde{Q}^\star A(n-k-1-i:n-1-i, i+1:n)$
    $A(i+2:n, n-k-1-i:n-1-i) \leftarrow A(i+2:n, n-k-1-i:n-1-i)\tilde{Q}$
end for
apply Algorithm 3 to $A(n_1-k+1:n-1-n_1+k, n_1-k+1:n-1-n_1+k)$ *% exchange bulges*
$A(n_1-k+1:n-1-n_1+k, n_1-k+1:n) \leftarrow \tilde{Q}^\star A(n_1-k+1:n-1-n_1+k, n_1-k+1:n)$
$A(n_1-k+1:n, n_1-k+1:n-1-n_1+k) \leftarrow A(n_1-k+1:n, n_1-k+1:n-1-n_1+k)\tilde{Q}$
for $i = n_1 - k - 1 : -1 : 0$ **do** *% chase bulges outwards*
    define unitary $\tilde{Q}$ with $A(2+i:2+k+i, n-k-1-i:n-1-i)\tilde{Q}e_1 = e_{k+1}$
    $A(2+i:n, n-k-1-i:n-1-i) \leftarrow A(2+i:n, n-k-1-i:n-1-i)\tilde{Q}$
    $A(n-k-1-i:n-1-i, 1+i:n) \leftarrow \tilde{Q}^\star A(n-k-1-i:n-1-i, 1+i:n)$
    define unitary $\tilde{Q}$ with $A(n-k-1-i, 1+i:k+1+i)\tilde{Q} = \alpha e_{k+1}^T$
    $A(n-k-1-i:n, 1+i:k+1+i) \leftarrow A(n-k-1-i:n, 1+i:k+1+i)\tilde{Q}$
    $A(1+i:k+1+i, n-k-i:n) \leftarrow \tilde{Q}^\star A(1+i:k+1+i, n-k-i:n)$
end for
for $i = k : -1 : 1$ **do** *% squeeze out bulges*
    define unitary $\tilde{Q}$ with $A(1:i+1, n-i:n)\tilde{Q}e_1 = e_{i+1}$
    $A(:,n-i:n) \leftarrow A(:,n-i:n)\tilde{Q}$
    $A(n-i:n,:) \leftarrow \tilde{Q}^\star A(n-i:n,:)$
    if $i > 1$ **then**
        define unitary $\tilde{Q}$ with $A(n-i, 1:i)\tilde{Q} = \alpha e_i^T$
        $A(n-i:n, 1:i) \leftarrow A(n-i:n, 1:i)\tilde{Q}$
        $A(1:i, n-i+1:n) \leftarrow \tilde{Q}^\star A(1:i, n-i+1:n)$
    end if
end for

Algorithm 2 requires $(8k+6)n^2 + \mathcal{O}(k^3 n)$ flops to transform A and the same amount to update a given $n \times n$ unitary matrix Q by the employed unitary transformations. This is the same flop count required by one implicit QZ iteration for an $n \times n$ matrix pair Kre05 [14].

It remains to determine the shifts. A reasonable choice is to take the eigenvalues of the pencil $(A(1:k, n+1-k:n), A(n+1-k:n, 1:k)^\star)$, as this corresponds to Rayleigh quotient shifts. Using Algorithm 2 with these shifts typically results in rapid decay of one or more super antidiagonal elements of A close to the bottom left corner, see Section 6 for a more detailed convergence discussion. Once such an element becomes negligible we can deflate the eigenvalue problem, see Section 3.1. As the eigenvalue problem is deflated simultaneously at both corners, whereas the QZ algorithm typically yields deflations only at one corner, it can be expected that the implicit palindromic QR algorithm yields a faster reduction of the problem size and is more efficient than the implicit QZ algorithm applied to $(A, A^\star)$. Numerical experiments from Section 9 will confirm this expectation.

ec:bulgeexchange

4.1 Bulge exchange

To complete the description of Algorithm 2, we still need to show how to exchange bulges in the middle of the antidiagonal. To avoid technicalities, let us only consider the submatrix of A_{n_1-1} that contains the bulges (cf. the highlighted area in (13)) and partition it as follows:

$$A = \begin{array}{c c} & \begin{array}{ccc} k & m & k \end{array} \\ \begin{array}{c} k \\ m \\ k \end{array} & \begin{bmatrix} & & A_{13} \\ & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \end{array}. \tag{16}$$

eq:exchangein

Note that $m = 1$ if n is even and $m = 0$ otherwise, but the discussion below is for general m. Exchanging the bulges consists of finding a unitary matrix Q such that

$$\tilde{A} = Q^\star A Q = \begin{array}{c c} & \begin{array}{ccc} k & m & k \end{array} \\ \begin{array}{c} k \\ m \\ k \end{array} & \begin{bmatrix} & & \tilde{A}_{13} \\ & \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{31} & \tilde{A}_{32} & \tilde{A}_{33} \end{bmatrix} \end{array}, \tag{17}$$

eq:exchangeout

and $\Lambda(\tilde{A}_{31}, \tilde{A}_{13}^\star) = \Lambda(A_{13}, A_{31}^\star)$.

For the moment, we allow non-unitary transformations. Note that, if $Y, Z^\star \in \mathbb{C}^{k\times m}$ satisfy

$$A_{31}Y + Z^\star A_{22} = -A_{32}, \tag{18}$$

eq:azya1

$$A_{13}^\star Y + Z^\star A_{22}^\star = -A_{23}^\star, \tag{19}$$

eq:azya2

and $X \in \mathbb{C}^{k\times k}$ solves

$$A_{31}X + X^\star A_{13} = -(A_{33} + A_{32}Z + Z^\star A_{23} + Z^\star A_{22} Z), \tag{20}$$

eq:axxtaaz

then the following $\star$-congruence transformation achieves our goal,

$$\begin{bmatrix} X^\star & Z^\star & I \\ Y^\star & I & \\ I & & \end{bmatrix} \begin{bmatrix} & & A_{13} \\ & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} X & Y & I \\ Z & I & \\ I & & \end{bmatrix} = \begin{bmatrix} & & A_{31} \\ & A_{22} & \\ A_{13} & & \end{bmatrix}.$$

The system (18), (19) is a generalized Sylvester equation and thus has a unique solution if and only if $\Lambda(A_{31}, A_{13}^\star) \cap \Lambda(A_{22}, A_{22}^\star) = \emptyset$, see [27] (Ste72) or [32, Theorem 6.6.8] (Wat07). The solvability condition for (20) is provided by the following lemma.

lem:bxxcd

Lemma 7 *Let $B, C \in \mathbb{C}^{k\times k}$. Then the matrix equation*

$$BX + X^\star C = D \tag{21}$$

eq:bxxcd

has a unique solution X for every right hand side $D \in \mathbb{C}^{k\times k}$ if and only if the following conditions hold:

1) the pencil $(B, C^\star)$ is regular, and

2a) if $\star = T$, $\Lambda(B, C^T) \setminus \{1\}$ is T-reciprocal free and if 1 is an eigenvalue, it has algebraic multiplicity 1, or

2b) if $\star = H$, $\Lambda(B, C^H)$ is H-reciprocal free.

Proof: The case $\star = T$ is proved in [5] (ByeK05).

For the case $\star = H$ note that the matrix operator $S : (\mathrm{Re}(X), \mathrm{Im}(X)) \mapsto (\mathrm{Re}(BX+X^H C), \mathrm{Im}(BX+X^H C))$ is linear and thus is injective if and only if its surjective.

We consider the following cases:

Case 1: Assume that (B, C^H) is regular and its spectrum is H-reciprocally free. Then $\Lambda(B, C^H) \cap \Lambda(C, B^H) = \emptyset$ and the generalized Sylvester equation

$$BX + YC = D, \quad C^H X + Y B^H = D^H$$

has a unique solution (X, Y). By symmetry, (Y^H, X^H) is also a solution, so $X = Y^H$. Thus, X is a solution of (21).

Case 2: Assume there is a nonzero vector x and λ on the unit circle such that $Bx = \lambda C^H x$, i.e., (B, C^H) is singular or has an eigenvalue on the unit circle. Then $X_1 := \sqrt{-\bar{\lambda}} x x^H C \neq 0$ gives $S(\mathrm{Re}(X_1), \mathrm{Im}(X_1))$=0.

Case 3: Analogously, if λ and $1/\bar{\lambda}$ are distinct eigenvalues with corresponding eigenvectors x and y, then choosing $X_2 = xy^H C - yx^H B^H \neq 0$ results in $S(\mathrm{Re}(X_2), \mathrm{Im}(X_2)) = 0$. ■

Hence, the matrix equations (18)–(20) are uniquely solvable under rather mild conditions on the shifts. To solve these equations we apply Gaussian elimination with partial pivoting to the corresponding vectorized linear systems. Note that for the case of $m = 0$, Y, Z are void and (20) reduces to $A_{31}X + X^\star A_{13} = -A_{33}$.

In order to obtain a unitary transformation let us consider the QR factorization

$$\begin{bmatrix} X & Y & I \\ Z & I & \\ I & & \end{bmatrix} = Q \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ & R_{22} & R_{23} \\ & & R_{33} \end{bmatrix}, \tag{22}$$

eq:xyzqr

where $R_{ii}, i = 1, 2, 3$ are non-singular, since the left hand matrix is. Then

$$\tilde{A} = Q^\star \begin{bmatrix} & & A_{13} \\ & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} Q = \begin{bmatrix} & & R_{11}^{-\star} A_{31} R_{33}^{-1} \\ & R_{22}^{-\star} A_{22} R_{22}^{-1} & \tilde{A}_{23} \\ R_{33}^{-\star} A_{13} R_{11}^{-1} & \tilde{A}_{32} & \tilde{A}_{33} \end{bmatrix}, \tag{23}$$

eq:tildeA

achieves the desired exchange.

The described exchange procedure is summarized in the following algorithm.

alg:swap

Algorithm 3 Palindromic bulge exchange

Input: A in form (16) with $\Lambda(A_{13}, A_{31}^\star)$ $\star$-reciprocal free, (in the T-case possibly with the exception of an eigenvalue 1 having algebraic multiplicity 1) and $\Lambda(A_{22}, A_{22}^\star) \cap \Lambda(A_{13}, A_{31}^\star) = \emptyset$.

Output: $\tilde{A}, Q$ satisfying (17) with $\Lambda(\tilde{A}_{31}, \tilde{A}_{13}^\star) = \Lambda(A_{13}, A_{31}^\star)$

1: solve (18),(19) for Y, Z
2: solve (20) for X
3: compute QR factorization (22)
4: compute $\tilde{A} = Q^\star A Q$

In finite-precision arithmetic, roundoff error pollutes the block antitriangular form (23) of $\tilde{A}$ and elements above the block antidiagonal need to be explicitely set to zero. It might happen that the elements become significantly larger than unit roundoff times $\|A\|_F$, in which case the exchange must be rejected in the interest of numerical backward stability. Then all transformations of the current QR step need to be undone and a new attempt with different shifts has to be started. In several numerical experiments with random matrices, we have not encountered a single instance of such a rejection, even though the equations (18)–(20) become very ill-conditioned as convergence sets in. This surprisingly robust behavior bears similarities to the well-known robustness of swapping algorithms for (unstructured) Schur forms BaiD93,KagP96 [1, 13].

4.2 Aggressive early deflaiton

Modern implementations of the standard QR and QZ algorithms employ several strategies to considerably speed up the computation, in particular multi-shift and advanced deflation techniques BraBM02;BraBM02a,KagK06 [2, 3, 12] These strategies can be extended to the palindromic QR algorithm in a rather straightforward way. We demonstrate this claim with the aggressive early deflation technique introduced in [3] BraBM02a, which often detects converged eigenvalues much earlier than the classical deflation described in Section 3.1.

Given A in anti-Hessenberg-triangular form, we select a window size w with $1 < w \ll n$ and partition A as

$$A = \begin{array}{l} \\ w \\ n-2w-1 \\ 1 \\ w \end{array} \begin{array}{c} \begin{array}{ccc} w & n-2w & w \end{array} \\ \begin{bmatrix} 0 & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \\ A_{41} & A_{42} & A_{43} \end{bmatrix} \end{array}.$$

By the generalized Schur form, there are unitary matrices U, V such that $U^\star A_{23} V$ and $V^\star A_{41} U$ are anti-triangular. Setting $Q = \mathrm{diag}(U, I_{n-2w}, V)$ thus yields the transformed matrix

$$Q^\star A Q = \begin{array}{l} \\ w \\ n-2w-1 \\ 1 \\ w \end{array} \begin{array}{c} \begin{array}{ccc} w & n-2w & w \end{array} \\ \begin{bmatrix} 0 & 0 & \tilde{A}_{13} \\ 0 & A_{22} & \tilde{A}_{23} \\ \tilde{A}_{31} & A_{32} & \tilde{A}_{33} \\ \tilde{A}_{41} & \tilde{A}_{42} & \tilde{A}_{43} \end{bmatrix} \end{array} = \begin{bmatrix} & & \triangle \\ *{-}* & \square & \square \\ \triangle & \square & \square \end{bmatrix}.$$

Provided that the first entry of the so-called spike $\tilde{A}_{31} \in \mathbb{C}^{1,n}$ is negligible, it can be set to zero and the eigenvalue residing in the outer corners is deflated. If the entry is not negligible, any other eigenvalue λ of $(\tilde{A}_{13}, \tilde{A}_{14}^\star)$, can be tested by reordering λ to the outer corners. This process can be repeatedly applied until no further deflatable eigenvalue can be found. As the entire procedure is numerically equivalent to aggressive early deflation in the QZ algorithm in the sense of the discussion in Section 3, we refer to [12] KagK06 for more details.

5 Equivalence of explicit and implicit procedures

In this section we show that the implicit palindromic QR step carries its name rightfully, i.e., it effects the same transformation as the explicit palindromic QR step. As for the standard eigenvalue problem, an implicit Q theorem will play a prominent role in the proof.

Theorem 8 (palindromic implicit Q Theorem) *Let $A \in \mathbb{C}^{n,n}$. Let Q, V be unitary matrices such that $Q^\star A Q$ and $V^\star A V$ are both in unreduced anti-Hessenberg-triangular form. Then, if $Qe_1 = const \cdot Ve_1$ or $Qe_n = const \cdot Ve_n$, there exists a unitary diagonal matrix D such that $V = QD$.*

Proof: For $n \leq 2$ there is nothing to show, so let $n \geq 3$. Set $n_1 = \lfloor \frac{n-1}{2} \rfloor$ and $n_2 = n - n_1$. We will show that the knowledge of the first or the last column of Q fixes every column up to a constant of norm 1.

We have $Q^\star A Q = H$, with H in unreduced anti-Hessenberg-triangular form. Inverting and (conjugate-) transposing gives

$$Q^{-1} A^{-\star} Q^{H\star} = H^{-\star} =: G = \begin{array}{l} \\ n_1 \\ n_2 \end{array} \begin{array}{c} \begin{array}{cc} n_2 & n_1 \end{array} \\ \begin{bmatrix} \square & \diagup \\ \square & 0 \end{bmatrix} \end{array},$$

so multiplication by (the conjugate of) Q yields

$$\begin{aligned} AQ &= Q^{H\star} H, && (24) \quad \text{eq:aqqh} \\ A^{-\star} Q^{H\star} &= QG. && (25) \quad \text{eq:aqqg} \end{aligned}$$

Evaluating the last column of (25) gives $A^{-\star}q_n^{H\star} = q_1 g_{1,n}$. So, if q_1 is given, then this determines q_n up to a multiple of norm one, and vice versa. Hence, at this point q_1 and q_n are known. Evaluating the first column of (24) yields $Aq_1 = q_n^{H\star}h_{n,1} + q_{n-1}^{H\star}h_{n-1,1}$. Since q_1 and q_n are known and q_{n-1} is orthogonal to q_n, this fixes q_{n-1} up to a multiple of norm one.

The remainder follows by induction. Suppose the first $k-1$ and the last k columns of Q are known. The $(n+1-k)$th column of (25) is

$$A^{-\star}q_{n+1-k}^{H\star} = \sum_{i=1}^{k} q_i g_{i,n+1-k}.$$

This determines q_k up to a multiple of norm one. The kth column of (24) is given by

$$Aq_k = \sum_{i=1}^{k+1} q_{n+1-i}^{H\star} h_{n+1-i,k}.$$

This fixes q_{n-k} up to a multiple of norm one.

If n is odd, all columns of Q are determined in this manner. If n is even, the column $\frac{n}{2}$ is still missing. It cannot be determined by (25), as the column n_2 of G is full. However, $q_{\frac{n}{2}}$ has to span the orthogonal complement of the known $(n-1)$ columns and is thus fixed up to a constant of modulus one. ■

So, as is the case with the reduction to standard Hessenberg form (e.g., GolV96 [8]), the first column essentially fixes the whole transformation matrix. But in case of the palindromic eigenvalue problem this is not enough to prove that the explicit and implicit QR steps yield the same result, because the first column of Q is manipulated not only in the beginning, but also in the very end of the implicit QR step. More work is necessary to derive the desired result.

Let V be the unitary matrix generated by the implicit QR algorithm, i.e., $\hat{A} = V^\star AV$. Consider the enlarged matrix

$$\left[\begin{array}{c|c} 0 & 0 \\ \hline x & A \end{array}\right]. \text{ Then } \begin{bmatrix} 1 & \\ & V \end{bmatrix}^\star \left[\begin{array}{c|c} 0 & 0 \\ \hline x & A \end{array}\right] \begin{bmatrix} 1 & \\ & V \end{bmatrix} = \left[\begin{array}{c|c} 0 & 0 \\ \hline \hat{x} & \hat{A} \end{array}\right] \quad (26)$$ eq:extendedpencil

where $x = A^\star p(A^{-\star}A)e_1$, $\hat{x} = V^\star x$.

Analyzing what happens to x during an implicit palindromic QR step, we note that the first transformation reduces x to a multiple of e_n. Then it stays untouched until the last phase of the process when the bulges are squeezed out again. Then the last $k+1$ elements are transformed. Since the eigenvalues of the bulge that was squeezed out at the lower left corner of A were $1/\kappa_1^\star, \ldots, 1/\kappa_k^\star$, and because of the following lemma, $\hat{x}$ satisfies a relation similar to x. Indeed,

$$\hat{x} \quad = \quad \hat{A}^\star q(\hat{A}^{-\star}\hat{A})e_1, \quad (27)$$ eq:hatx

where $q(\lambda) = \prod_{i=1}^{k}(\kappa_i\lambda - 1)$ is the reversal of $p(\lambda)$.

Lemma 9 *Let $k < \frac{n}{2}$, and let $x \in \mathbb{C}^n$ have the property that $x_i = 0$ for $i = 1, \ldots, n-k-1$ and $x_{n-k} \neq 0$. Then there exists a unique monic polynomial p of degree exactly k such that $x = \alpha A^\star p(A^{-\star}A)e_1$ for some nonzero scalar α.*

Proof: The result follows from an analogous result for the QZ algorithm, see Wat07 [32, Section 7.1]. ■

Combining (26) with (27) yields

$$A^\star p(A^{-\star}A)e_1 = x = V^{-\star}\hat{x} = V^{-\star}\hat{A}^\star q(\hat{A}^{-\star}\hat{A})e_1 = A^\star Vq(\hat{A}^{-\star}\hat{A})e_1 = A^\star q(A^{-\star}A)Ve_1 \quad (28)$$

So,

$$Ve_1 = q(A^{-\star}A)^{-1}p(A^{-\star}A)e_1 = r(A^{-\star}A)e_1$$

with $r = p/q$. By (9) the first column of the transformation matrix used in the explicit palindromic QR step also satisfies

$$Qe_1 = const \cdot r(A^{-\star}A)e_1.$$

Thus, by the palindromic implicit Q Theorem, V and Q are essentially equal. We have thus proven the following central result.

theo:impexp

Theorem 10 *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular and unreduced anti-Hessenberg-triangular. Let $\kappa_1, \dots, \kappa_k$ with $k < \frac{n}{2}$ be a $\star$-reciprocal free set of shifts, that are not exact eigenvalues of $(A, A^\star)$. Let A_+ be the result of an explicit palindromic QR step (8), applied to A. Let $\hat{A}$ be the result of the implicit palindromic QR step applied to A. Then there exists a unitary diagonal matrix D such that $\hat{A} = D^\star A_+ D$.*

6 Convergence theory

sec:convergence

In the following, we show local quadratic convergence of the palindromic QR iteration, provided that Rayleigh-quotient shifts are chosen, see the discussion after Algorithm 2. Because of the equivalence shown in Theorem 10, we can restrict our convergence proof to the more accessible explicit palindromic QR algorithm. Moreover, by deflation, we may assume without loss of generality that A is invertible.

Let us recall that for rational functions $r_1, r_2, \dots$ the palindromic QR iteration is defined as

$$r_i(A_{i-1}^{-\star}A_{i-1}) = Q_i R_i, \qquad \text{(QR decomposition)} \tag{29}$$ eq:qr1

$$A_i \leftarrow Q_i^\star A_{i-1} Q_i. \tag{30}$$ eq:qr2

The aim is to show that the sequence A_i converges to block anti-triangular form

$$A_i \to \begin{array}{c c} & \begin{array}{ccc} k & n-2k & k \end{array} \\ \begin{array}{c} k \\ m \\ k \end{array} & \left[\begin{array}{ccc} & & A_{13} \\ & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{array}\right] \end{array}. \tag{31}$$ eq:convpart

for some $1 \le k \le \lfloor \frac{n}{2} \rfloor$. Since $A_i = \hat{Q}_i^\star A \hat{Q}_i$ with $\hat{Q}_i = Q_1 Q_2 \cdots Q_i$, this is equivalent to the convergence of the linear subspaces $\mathcal{S}_i$ and $\mathcal{T}_i$ spanned by the first k columns of $\hat{Q}_i$ and the last k columns of $\hat{Q}_i^{-\star}$ to right and left deflating subspaces $\mathcal{X}_1$ and $\mathcal{Y}_1$, respectively. Note that $\mathcal{X}_1$ and $\mathcal{Y}_1$ are also invariant subspaces of $A^{-\star}A$ and $AA^{-\star}$, respectively.

Assuming that no zero or pole of r_i coincides with an eigenvalue of $(A, A^\star)$, all R_i are invertible, which together with (29)–(30) implies

$$\mathcal{S}_i = \hat{Q}_{i-1} r_i(A_{i-1}^{-\star}A_{i-1})\mathcal{S}_0 = r_i(A^{-\star}A)\mathcal{S}_{i-1} = \cdots = \hat{r}_i(A^{-\star}A)\mathcal{S}_0,$$

where $\mathcal{S}_0 = \operatorname{span}\{e_1, \dots, e_k\}$ and $\hat{r}_i = r_1 r_2 \cdots r_i$. Similarly,

$$\mathcal{T}_i = \hat{r}_i(A^{-\star}A)^{-\star}\mathcal{T}_0, \qquad \mathcal{T}_0 = \operatorname{span}\{e_{n-k+1}, \dots, e_n\}.$$

The convergence of $\hat{r}_i(A^{-\star}A)\mathcal{S}_0$ to an invariant subspace of $A^{-\star}A$ is covered by the standard convergence theory for the QR algorithm WatE91 [33], in principle only for the case of polynomial $\hat{r}_i$. However, as noted in Wat96 [31], the Cayley-Hamilton theorem allows us to extend this theory to any analytic function $\hat{f}_i$. In particular, we obtain the following result as a corrollary of Theorem 5.4 in WatE91 [33].

theo:conv1

Corollary 11 *Let r be a rational function and $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of $A^{-\star}A$, ordered so that $|r(\lambda_1)| \ge |r(\lambda_2)| \ge \cdots \ge |r(\lambda_n)|$. Suppose there is an integer $k < \lfloor n/2 \rfloor$ such that $|r(\lambda_k)| > |r(\lambda_{k+1})|$ and let $\rho = |r(\lambda_{k+1})|/|r(\lambda_k)|$. Consider a sequence of rational functions $r_1, r_2, \dots$ such that $r_i \to r$. Let $\mathcal{X}_1$ and $\mathcal{X}_2$ be the invariant subspaces of $A^{-\star}A$ associated with $\lambda_1, \dots, \lambda_k$ and $\lambda_{k+1}, \dots, \lambda_n$, respectively, and assume $\mathcal{S}_0 \cap \mathcal{X}_2 = \{0\}$. Then for every $\hat{\rho}$ satisfying $\rho < \hat{\rho} < 1$ there is a constant $\hat{C}$ not depending on $\mathcal{S}_0$ such that*

$$d(\mathcal{S}_i, \mathcal{X}_1) \le \hat{C} \frac{d(\mathcal{S}_0, \mathcal{X}_1)}{\sqrt{1 - d(\mathcal{S}_0, \mathcal{X}_1)^2}} \hat{\rho}^i,$$

where d denotes the gap between two subspaces GolV96 *[8].*

An analogous result holds for $\mathcal{T}_i$, with the most notable difference that a gap between $r(\lambda_{n-k+1})$, $r(\lambda_{n-k})$ is required and ρ is replaced by $\tilde{\rho} = |r(\lambda_{n-k+1})|/|r(\lambda_{n-k})|$. Let us now restrict ourselves to the particular choice $r = p/q$ for some polynomial p and its $\star$-reversal q, see (7). Then $\frac{1}{|r(\mu)|} = |r(1/\mu^\star)|$, implying that we can choose $\lambda_{n-j+1} = 1/\lambda_j^\star$ in the ordering of Corrolary 11 and $\tilde{\rho} = |r(\lambda_{k+1}^\star)|/|r(\lambda_k^\star)|$.

We thus have shown superlinear convergence of the palindromic QR iteration if the zeros and poles of r_i tend to $\lambda_1, \dots, \lambda_k$ and $1/\lambda_1^\star, \dots, 1/\lambda_k^\star$, respectively, provided of course that the set $\lambda_1, \dots, \lambda_k$ is $\star$-reciprocal free. Even better convergence results are obtained when choosing

$$p_i(z) = \prod_{\kappa_p \in \Sigma} (z - \kappa_p), \quad q_i(z) = \prod_{\kappa_p \in \Sigma} (\kappa_p^\star z - 1), \quad r_i(z) = \frac{p_i(z)}{q_i(z)}. \tag{32}$$

eq:francis

with $\Sigma = \Lambda((A_{13}^{(i-1)})^{-\star} A_{31}^{(i-1)})$ and the partitioning

$$A_{i-1} = \hat{Q}_{i-1}^* A \hat{Q}_{i-1} = \begin{array}{c} \\ k \\ n-k \\ k \end{array} \!\!\! \begin{array}{c} \begin{array}{ccc} k & n-2k & k \end{array} \\ \begin{bmatrix} A_{11}^{(i-1)} & A_{12}^{(i-1)} & A_{13}^{(i-1)} \\ A_{21}^{(i-1)} & A_{22}^{(i-1)} & A_{23}^{(i-1)} \\ A_{31}^{(i-1)} & A_{32}^{(i-1)} & A_{33}^{(i-1)} \end{bmatrix} \end{array}. \tag{33}$$

eq:itpart

This parallels Francis' choice of shifts for the standard QR algorithm and can be expected to result in rapid convergence, since $A_{31}^{(i-1)}$ and $A_{13}^{(i-1)}$ contain increasingly good approximations to eigenvalues as the iterations proceeds. The main result of this section shows that this choice actually results in local quadratic convergence.

Theorem 12 *Assume $(A, A^\star)$ has distinct finite eigenvalues and let $\hat{r}_i = r_1 r_2 \cdots r_i$, where $r_j = p_j/q_j$ with p_j and q_j defined as in (32). Suppose the corresponding subspace sequences*

$$\mathcal{S}_i = \hat{r}_i(A^{-\star}A)\mathcal{S}_0, \qquad \mathcal{T}_i = \hat{r}_i(A^{-\star}A)^{-\star}\mathcal{T}_0,$$

with $\mathcal{S}_0$ and $\mathcal{T}_0$ as above, converge to a pair of right and left deflating subspaces of $(A, A^\star)$. Then both converge quadratically.

Proof: The proof follows the ideas of WatE91 [33, Thm 6.3]; it is shown that for sufficiently small

$$\varepsilon = d(\mathcal{S}_{i-1}, \mathcal{X}_1) + d(\mathcal{T}_{i-1}, \mathcal{Y}_1) \tag{34}$$

eq:boundsubspaces

the distances of the next iterates, $d(\mathcal{S}_i, \mathcal{X}_1)$ and $d(\mathcal{T}_i, \mathcal{Y}_1)$, are proportional to ε^2.

For this purpose, consider the partitioning (31) of A and let $\Lambda_1 = \Lambda(A_{13}^{-\star} A_{31})$, $\Lambda_3 = \Lambda(A_{31}^{-\star} A_{13})$ denote the eigenvalues associated with $\mathcal{X}_1, \mathcal{Y}_1$. Set $\Lambda_3 = \Lambda(A) \backslash (\Lambda_1 \cup \Lambda_2)$. Consider the analogous partition (33) of A_{i-1}. From (34) it follows that $\|[A_{11}^{(i-1)} \; A_{12}^{(i-1)}]\|_2 \le \sqrt{2}\varepsilon\|A\|_2$ as well as $\left\| \left[\begin{smallmatrix} A_{11}^{(i-1)} \\ A_{21}^{(i-1)} \end{smallmatrix} \right] \right\|_2 \le \sqrt{2}\varepsilon\|A\|_2$, see WatE94 [34, Lemma 3.1]. Analogously, $d(\mathcal{T}_{i-1}, \mathcal{X}_q) \le \varepsilon$ implies $\left\| \left[\begin{smallmatrix} A_{21}^{(i-1)} \\ A_{31}^{(i-1)} \end{smallmatrix} \right] \right\|_2 \le$ $\sqrt{2}\varepsilon\|A\|_2$. Since $(A, A^\star)$ is diagonalizable, the generalized Bauer-Fike theorem ElsS82 [7] yields bounds for the eigenvalues of the truncation of the iterate $(A_{i-1}, A_{i-1}^\star)$:

Christian: changed atop to smallmatrix

Christian: changed atop to smallmatrix

$$\max_{\lambda \in \Lambda_j, \tilde{\lambda} \in \tilde{\Lambda}_j} |\lambda - \tilde{\lambda}| \le c\varepsilon\|A\|_2, \tag{35}$$

eq:dist

where c only depends on A and $\tilde{\Lambda}_1 = \Lambda((A_{13}^{(i-1)})^{-\star} A_{31}^{(i-1)}) = \Sigma$, $\tilde{\Lambda}_2 = \Lambda((A_{22}^{(i-1)})^{-\star} A_{22}^{(i-1)})$, $\tilde{\Lambda}_3 = \Lambda((A_{31}^{(i-1)})^{-\star} A_{13}^{(i-1)})$. Here, we implicitely assumed that ε is sufficiently small such that none of the $\tilde{\Lambda}_j$ contains an infinite value. Thus

$$\max\{|p_i(\lambda_p)| : \lambda_p \in \Lambda_1\} \le c_p\varepsilon$$

with a constant c_p only depending on k and A. By (35) we may assume for sufficiently small ε that the distance between Λ_1 and $\tilde{\Lambda}_3$ (containing the roots of q_i) is at least $\delta_{\Lambda_1,\Lambda_3}/2$, i.e., half the distance between Λ_1 and Λ_3. This yields

$$\min\{|q_i(\lambda_p)| : \lambda_p \in \Lambda_1\} \cdot \prod_{\kappa_p \in \Sigma} |\kappa_p| \geq (\delta_{\Lambda_1,\Lambda_3}/2)^k.$$

Similarly,

$$\min\{|p_i(\lambda)| : \lambda \in \Lambda_2 \cup \Lambda_3\} \geq (\delta_{\Lambda_1,\Lambda_2\cup\Lambda_3}/2)^k$$

and

$$\max\{|q_i(\lambda)| : \lambda \in \Lambda_2 \cup \Lambda_3\} \leq c_q \tag{36}$$

eq:mq

for some constant c_q only depending on A and k. Hence,

$$\begin{aligned}
\rho &:= \frac{\max\{|r_i(\lambda_p)| : \lambda_p \in \Lambda_1\}}{\min\{|r_i(\lambda)| : \lambda \in \Lambda_2 \cup \Lambda_3\}} \\
&= \frac{\max\{|p_i(\lambda_p)| : \lambda_p \in \Lambda_1\} \cdot \max\{|q_i(\lambda)| : \lambda \in \Lambda_2 \cup \Lambda_3\}}{\min\{|p_i(\lambda)| : \lambda \in \Lambda_2 \cup \Lambda_3\} \cdot \min\{|q_i(\lambda_p)| : \lambda_p \in \Lambda_1\}} \\
&\leq c_p c_q (2/\delta_{\Lambda_1,\Lambda_2\cup\Lambda_3})^k (2/\delta_{\Lambda_1,\Lambda_3})^k \prod_{\kappa_p \in \Sigma} |\kappa_p| \cdot \varepsilon =: C_p \varepsilon.
\end{aligned}$$

For sufficiently small ε, the conditions of Corollary 11 are clearly satisfied. Setting $\hat{\rho} = \sqrt{2} C_p \varepsilon$ and applying Corollary 11 to one step of the iteration results in

$$d(\mathcal{S}_i, \mathcal{X}_1) \leq \sqrt{2} \hat{C} C_p \frac{\varepsilon^2}{\sqrt{1-\varepsilon^2}} \leq 2\hat{C} C_p \varepsilon^2,$$

where the latter inequality holds for $\varepsilon \leq 1/\sqrt{2}$. The proof for $d(\mathcal{T}_i, \mathcal{Y}_1)$ is analogous. ■

Remark 13 In the practically not so relevant case $2k = n$ the inequality (36) can be improved to

$$\max\{|q_i(\lambda)| : \lambda \in \Lambda_1 \cup \Lambda_2\} \leq c_q \varepsilon.$$

This yields cubic convergence for $\mathcal{S}_i$ and $\mathcal{T}_i$. □

7 The real implicit palindromic QR algorithm

Often the problems arising in physical applications are real. In these cases one wants to keep the computations in real arithmetic, mainly because complex computations need twice as much memory and are up to four times slower.

There is a real version of the palindromic Schur form.

Theorem 14 (real palindromic Schur form, Sch07 [24]) *Let $A \in \mathbb{R}^{n,n}$ be such that (A, A^T) is a regular pencil with at most one eigenvalue on the unit circle. Then there exists an orthogonal Q such that*

$$Q^T A Q = \begin{bmatrix} & & T_{13} \\ & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix},$$

where $T_{31}, T_{13} \in \mathbb{R}^{m\times m}$, $m = \lfloor \frac{n}{2} \rfloor$ are such that (FT_{31}, FT_{13}^T) is in real generalized Schur form. The eigenvalues of (A, A^T) are given by the eigenvalues of (FT_{31}, FT_{13}^T), their reciprocals, and, if n is odd, additionally 1. Moreover, Q can be chosen such that the eigenvalues of (FT_{31}, FT_{13}^T) appear in any order.

Furthermore, the first k columns of Q span a right deflating subspace with the corresponding left deflating subspace spanned by the last k columns of $Q^{-\star}$, belonging to the eigenvalues of $(FT(n-k+1:n, 1:k), FT(1:k, n-k+1:n)^T)$ for all k satisfying $t_{n-k,k} = 0 = t_{k,n-k}$.

If A in (1) is real then every algorithm discussed in this paper stays in real arithmetic provided that the vector $x = \alpha A^T p(A^{-T}A)e_1$ is real. This is the case if the shifts are closed under complex conjugation. In this case x can also be computed completely avoiding complex operations. Note that for complex conjugate sets, T- and H-reciprocal freeness are equivalent.

8 The even implicit QR algorithm

In this section we briefly show how the results of this paper can be extended to the even eigenvalue problem

$$Mx = \lambda N x, \quad \text{where } M = M^\star,\ N = -N^\star.$$

One way to solve an even eigenvalue problem is to transform it to a palindromic problem by a Cayley transform. However, it is also possible to work directly on the matrices M, N. In the following we show that the palindromic QR algorithm has an even counterpart.

We will assume that the pencil (M, N) is given in anti-Hessenberg form, i.e., both, M and N, are anti-Hessenberg. As in the palindromic case, this form arises in some applications, but there is no direct method known to reduce a general even pencil to this form.

The first step is to transform the pencil into anti-Hessenberg-triangular form, in which M is anti-Hessenberg and N is antitriangular. This form can be achieved by chasing the super-antidiagonal elements of N, one at a time, out to the $(1, n-1)$ and $(n-1, 1)$ position, where they can be annihilated by a $\star$-congruence rotation in the last two rows and columns. Algorithm 1 can be adapted to even pencils as follows. In the lines 3 and 10 'A' is replaced by 'M', whereas 'A' is replaced by 'N' in line 6. Moreover, the transformations are applied to both, M and N.

An *even implicit QR step* with the shifts $\kappa_1, \dots, \kappa_k$ has the following form: compute the vector $x = Np(N^{-1}M)e_1$ where $p(\lambda) = \prod_{i=1}^{k}(\lambda - \kappa_i)$. Note that $N^{-1}M$ is upper Hessenberg. Thus, only the last $k+1$ elements of x are nonzero. Applying a transformation that reduces the nonzero part of x to a multiple of e_{k+1} to the last $k+1$ rows and columns of M and N creates two bulges in the pencil. One bulge appears at the lower left corner of (M, N) carrying the shifts as eigenvalues. Because of symmetry there is another bulge at the upper right corner, having $-\kappa_1^\star, \dots, -\kappa_k^\star$ as eigenvalues. By alternating elimination of a row or column of the bulge these bulges are chased towards the center. There, the bulges are passed through each other by a technique similar to palindromic bulge exchange. This is discussed in the following subsection. Then the bulges are chased out again and squeezed out at the corners.

An algorithm for the even QR step can be obtained from Algorithm 2 upon replacing line 2 by $x = Np(N^{-1}M)e_1$; replacing 'A' by 'M' in lines 10, 21, and 30; replacing 'A' by 'N' in lines 7, 18, and 26; and applying all transformations to M and N. If only one half of M and N are stored, it has the same flop counts as Algorithm 2.

Note, that this algorithm is equivalent to the bidirectional QZ algorithm Wat07 [32] applied to (FM, FN) with top shifts $\kappa_1, \dots, \kappa_k$ and bottom shifts $-\kappa_1^\star, \dots, -\kappa_k^\star$. Also, if n is even and $N = \left[\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right]$, the algorithm essentially reduces to the Hamiltonian QR algorithm Bye86 [4]. Christian: changed smatrix to smallmatrix

8.1 Bulge exchange

A $\star$-congruence transformation $(\tilde{M}, \tilde{N}) = W^\star(M, N)W$ with

$$W = \begin{bmatrix} X & Y & I \\ Z & I & \\ I & & \end{bmatrix}$$

block diagonalizes the 3-by-3 block antitriangular even pencil (M, N) if and only if $Y, Z^\star$ satisfy

$$M_{31}Y + Z^\star M_{22} = -M_{32}, \qquad (37) \quad \boxed{\texttt{eq:myzmm}}$$

$$N_{31}Y + Z^\star N_{22} = -N_{32}, \qquad (38) \quad \boxed{\texttt{eq:nyznn}}$$

and X is a solution of

$$\begin{aligned} M_{31}X + X^\star M_{31}^\star &= -(M_{33} + M_{32}Z + Z^\star M_{32}^\star + Z^\star M_{22}Z^\star), & (39) \\ N_{31}X - X^\star N_{31}^\star &= -(N_{33} + N_{32}Z - Z^\star N_{32}^\star + Z^\star N_{22}Z^\star). & (40) \end{aligned}$$

eq:mxxm eq:nxxn

The system (37), (38) is a generalized Sylvester equation and thus has a unique solution if and only if $\Lambda(M_{31}, N_{31}) \cap \Lambda(M_{22}, N_{22}) = \emptyset$. In order to assess the solvability of (39), (40), we note that their sum

$$(M_{31} + N_{31})X + X^\star (M_{31} - N_{31})^\star = rhs \tag{41}$$

eq:mnxxmn

(with rhs the sum of the right hand sides of (39), (40)) is of the form (20) and that (39) and (40) can be recovered as the symmetric and skew symmetric parts of (41), respectively. Thus, by applying Lemma 7 to (41), the system of matrix equations (39), (40) have a unique solution, if and only if $\Lambda(M_{31}+N_{31}, M_{31}-N_{31})$ is $\star$-reciprocal free (in the T-case possibly with the exception of an eigenvalue 1 of algebraic multiplicity 1). Because $(M_{31}+N_{31}, M_{31}-N_{31})$ is the Cayley transform of (M_{31}, N_{31}), this is the case if and only if $\Lambda(M_{31}, N_{31})$ contains no pairs of the form $(\lambda, -\lambda^\star)$ (in the T-case possibly with the exception of an eigenvalue ∞ of algebraic multiplicity 1).

The actual solution of (37), (38), and (39), (40) amounts to small vectorized linear systems, which can be solved by Gaussian elimination with partial pivoting.

The material of this subsection is summarized in the following algorithm.

alg:evenswap

Algorithm 4 Even bulge exchange

Input: even 3-by-3 block antitriangular pencil (M, N) such that $\Lambda(M_{31}, N_{31})$ contains no pairs of the form $(\lambda, -\lambda^\star)$ (in the T-case except possibly for an eigenvalue ∞ of algebraic multiplicity 1) and $\Lambda(M_{31}, N_{31}) \cap \Lambda(M_{22}, N_{22}) = \emptyset$.

Output: $\tilde{A}, Q$ satisfying (17) with $\Lambda(\tilde{A}_{31}, \tilde{A}_{13}^\star) = \Lambda(A_{13}, A_{31}^\star)$

1: solve (37),(37) for Y, Z via Kronecker product formulation
2: solve (39),(40) for X via Kronecker product formulation
3: compute QR factorization (22)
4: compute $(\tilde{M}, \tilde{N}) = Q^\star (M, N) Q$

9 Numerical experiments

sec:experiments

In the following, we provide examples indicating that the presented palindromic QR algorithm can be both more accurate and more efficient than the standard QZ algorithm applied to the matrix pair $(A, A^\star)$. For the first example, we have used a straightforward Matlab implementation of Algorithm 2.

ex:rump

Example 2 The following matrix is a variation of Rump's example [22] (Rum06)

$$A = \begin{bmatrix} 1 & -1+\varphi & 0 \\ 1-\varphi & 1 & \imath \\ 0 & -\imath & 1 \end{bmatrix}, \quad 0 < \varphi \leq 1. \tag{42}$$

eq:rump

A simple hand computation reveals that the eigenvalues of the matrix pair (A, A^T) are given by

$$\lambda_1 = 1, \quad \lambda_2 = \frac{1+\sqrt{2\varphi - \varphi^2}}{1-\sqrt{2\varphi-\varphi^2}}, \quad \lambda_3 = \frac{1-\sqrt{2\varphi-\varphi^2}}{1+\sqrt{2\varphi-\varphi^2}}.$$

As $\varphi \to 0$, the structured eigenvalue condition numbers tend to become significantly lower than the unstructured ones, see Figure 1. (For computing the structured eigenvalue condition numbers, we have used [15, Thm. 4.7] (KrePM06).) Therefore, it can be expected that a structure-preserving method delivers higher accuracy than the QZ algorithm. As shown in Figure 1, this expectation is met. Note that we applied the palindromic QR algorithm with a single shift. In order to do so, we first need to transform A to anti-Hessenberg form by anti-triangularizing the leading 2-by-2 principal submatrix of A. □

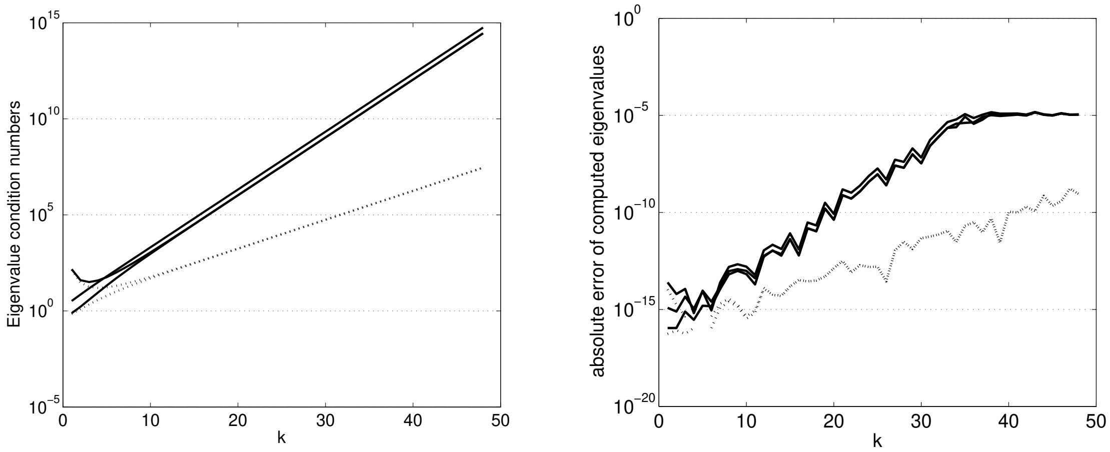


fig:rump

Figure 1: Rump's example (42) with $\varphi = 2^{-k}$ for $k = 1, \ldots, 48$. *Left:* Structured (dotted lines) and unstructured eigenvalue condition numbers of Example 2. Being zero, the structured condition numbers for $\lambda_1 = 1$ are not displayed. *Right:* Absolute errors of eigenvalues computed by the palindromic QR algorithm (dotted lines) and by the QZ algorithm (solid lines). Being zero, the errors of $\lambda_1 = 1$ computed by the palindromic QR algorithm are not displayed.

The second example aims at demonstrating the efficiency of Algorithm 2 compared to the current LAPACK 3.1 implementation of the QZ algorithm. For this purpose, we have created a prototype Fortran 77 implementation of the palindromic QR algorithm with a double shift (Algorithm 2) that accepts and preserves real input matrices A, see `http://www.math.ethz.ch/~kressner/palindromic.php`. To allow for a fair comparison, no aggressive early deflation or multishift techniques have been used, as the LAPACK routine `DHGEQZ` has neither of them. The tests were performed on an x86 Intel 1.2Ghz processor with 512 MByte RAM, using the Compaq Visual Fortran compiler 6.5 with the Compaq Extended Math Library and optimization level 3.

Example 3 Consider a linear discrete-time optimal control problem of the form

$$\begin{aligned} \min \qquad & \sum_{k=0}^{\infty} \begin{bmatrix} u_k \\ x_k \end{bmatrix}^T \begin{bmatrix} R & S^T \\ S & Q \end{bmatrix} \begin{bmatrix} u_k \\ x_k \end{bmatrix}, \text{ subject to} && (43) \\ E_d x_{k+1} \quad = \quad & A_d x_k + B_d u_k, \quad x_0 \text{ given}, && (44) \end{aligned}$$

eq:opcon1 eq:opcon2

where $E_d, A_d, Q \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$, and $B, S \in \mathbb{R}^{n\times m}$. Under certain conditions, the solution to this optimal control problem can obtained from the eigenvalues and deflating subspaces of the palindromic matrix pair (A, A^T) with

$$A = \begin{bmatrix} 0 & B_d & A_d \\ 0 & R & S^T \\ E_d^T & S & Q \end{bmatrix}, \qquad (45)$$

eq:palfromopcon

see [26] (Sch08) for more details. In the case of a single-input system, it is possible to reduce A to anti-Hessenberg-triangular form by computing the controller Hessenberg form of (44). One first computes a Householder reflector Q_0 such that $Q_0 B_d = \text{const} \cdot e_1$ and then applies a variant of Hessenberg-triangular reduction to compute orthogonal matrices $Q = \text{diag}(1, \tilde{Q})$ and Z such that $(Q^T Q_0 E_d Z, Q^T Q_0 A_d Z)$ has Hessenberg-triangular form. Setting $U = \text{diag}(Q_0 Q F, 1, Z)$, with F being the flip matrix, then gives an anti-Hessenberg-triangular matrix $U^T A U$.

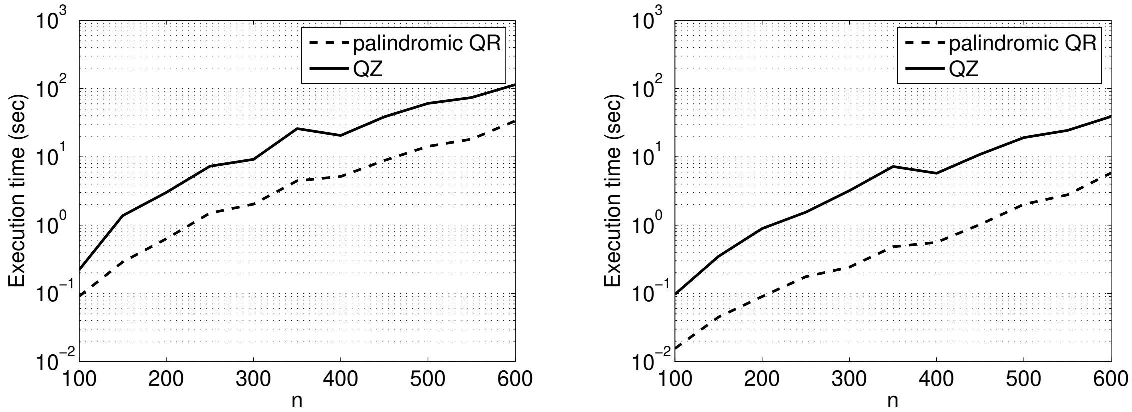


fig:times Figure 2: Execution times needed for solving the palindromic eigenvalue problem (45) with coefficients (47). The figures do not include the time needed for reduction to Hessenberg-triangular form. Note that A is $(2n+1) \times (2n+1)$. *Left:* The complete palindromic/generalized Schur form of (A, A^T) is computed.
Right: Only the eigenvalues of (A, A^T) are computed.

As a simple scalable example, we consider the discretized 1D-Laplace with boundary control

$$A_c = \frac{1}{h^2} \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix}, \quad B_c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{46}$$ eq:cont

and $h = 1/n$. This is turned into a discrete-time problem by the Cayley transform:

$$E_d = I_n - A_c, \quad A_d = I_n + A_c, \quad B_d = \sqrt{2} B_c. \tag{47}$$ eq:disc

The weighting matrices are simply chosen to be $R = 1, S = 0, Q = I_n$. Figure 2 shows the execution times needed for solving the corresponding palindromic eigenvalue problem. It turns out that the palindromic QR algorithm requires roughly between 10% and 25% of the time needed by the QZ algorithm, depending on whether only the eigenvalues or the complete Schur form is to be computed.

In general, however, the attainable reduction depends on the speed of convergence and hence on the matrix entries. For example, when replacing (46) by the steel rail cooling problem from LYAPACK [20] (order $n = 821$), the QZ algorithm requires 163 seconds while the palindromic QR algorithm requires 136 seconds. This much less impressive reduction is due to the slow convergence of the palindromic QR algorithm caused by the fact that the eigenvalues of the corresponding palindromic matrix pair form a tight cluster around -1. $\square$

10 Conclusions

We have developed variants of the implicit QR algorithm that preserve the structure of palindromic and even pencils. Each of these starts bulges at both ends of the pencil, chases them towards each other, and passes them through each other.

Although the palindromic QR algorithm solves a generalized eigenvalue problem, it has some of the features of the standard QR algorithm: it operates on one n-by-n matrix only, and one step costs the same as one step of the bidirectional QR algorithm. The flop count and memory requirement of the even QR iteration is the same. Both are strongly backwards stable. Moreover, modern techniques, such as multiple bulge chases and aggressive early deflation BraBM02,BraBM02a [2, 3] can be adapted to these variants in a straightforward way. Thus, these are the methods of choice if the problem is given in anti-Hessenberg form. Unfortunately, the reduction of a general palindromic or even pencil to anti-Hessenberg form remains elusive.

References

BaiD93 [1] Zhaojun Bai and James W. Demmel. On swapping diagonal blocks in real Schur form. *Linear Algebra Appl.*, 186:73–95, 1993.

BraBM02 [2] Karen Braman, Ralph Byers, and Roy Mathias. The multishift QR algorithm. I. Maintaining well-focused shifts and level 3 performance. *SIAM J. Matrix Anal. Appl.*, 23(4):929–947, 2002.

BraBM02a [3] Karen Braman, Ralph Byers, and Roy Mathias. The multishift QR algorithm. II. Aggressive early deflation. *SIAM J. Matrix Anal. Appl.*, 23(4):948–973, 2002.

Bye86 [4] Ralph Byers. A Hamiltonian QR algorithm. *SIAM J. Sci. Statist. Comput.*, 7(1):212–229, 1986.

ByeK05 [5] Ralph Byers and Daniel Kressner. Structured condition numbers for invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 28(2):326–347, 2006.

ChuHLW07 [6] Eric King-Wah Chu, Tsung-Min Hwang, Wen-Wei Lin, and Chin-Tien Wu. Vibration of Fast Trains, Palindromic Eigenvalue Problems and Structure-Preserving Doubling Algorithms. Technical report, School of Mathematical Sciences, Monash University, Australia, January 2007.

ElsS82 [7] Ludwig Elsner and Ji-Guang Sun. Perturbation theorems for the generalized eigenvalue problem. *Linear Algebra Appl.*, 48:341–357, 1982.

GolV96 [8] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

HilMM04 [9] Andreas Hilliges, Christian Mehl, and Volker Mehrmann. On the solution of palindromic eigenvalue problems. In *Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS)*. Jyväskylä, Finland, 2004. CD-ROM.

HorS06 [10] Roger A. Horn and Vladimir V. Sergeichuk. Canonical forms for complex matrix congruence and *congruence. *Linear Algebra Appl.*, 416(2-3):1010–1032, 2006.

HorS06a [11] Roger A. Horn and Vladimir V. Sergeichuk. A regularization algorithm for matrices of bilinear and sesquilinear forms. *Linear Algebra Appl.*, 412(2-3):380–395, 2006.

KagK06 [12] Bo Kågström and Daniel Kressner. Multishift variants of the QZ algorithm with aggressive early deflation. *SIAM J. Matrix Anal. Appl.*, 29(1):199–227, 2006/07.

KagP96 [13] Bo Kågström and Peter Poromaa. Computing eigenspaces with specified eigenvalues of a regular matrix pair (A, B) and condition estimation: theory, algorithms and software. *Numer. Algorithms*, 12(3-4):369–407, 1996.

Kre05 [14] Daniel Kressner. *Numerical methods for general and structured eigenvalue problems*, volume 46 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005.

KrePM06 [15] Daniel Kressner, Maria J. Peláez, and Julio Moro. Structured Hölder condition numbers for multiple eigenvalues. Uminf report, Department of Computing Science, Umea University, Sweden, October 2006.

MacMMM05b [16] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM Journal on Matrix Analysis and Applications*, 28(4):1029–1051, 2006.

MacMMM07 [17] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Numerical methods for palindromic eigenvalue problems: Computing the anti-triangular Schur form. Technical Report 409, DFG Research Center MATHEON, *Mathematics for key technologies* in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2007. url: http://www.matheon.de, submitted.

Meh07 [18] Christian Mehl. On asymptotic convergence of nonsymmetric Jacobi algorithms. Preprint 393, MATHEON - DFG Research Center "Mathematics for key techonolgies", Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, FRG, 2007. To appear in SIAM J. Matrix Anal. Appl.

Meh88 [19] Volker Mehrmann. A symplectic orthogonal method for single input or single output discrete time optimal quadratic control problems. *SIAM J. Matrix Anal. Appl.*, 9(2):221–247, 1988.

Pen00 [20] Thilo Penzl. Lyapack users guide. Technical report SFB393/00-33, Sonderforschungsbereich 393 Numerische Simulation auf massiv parallelen Rechnern, TU Chemnitz, 09107 Chemnitz, FRG, 2000.

Rod05 [21] Leiba Rodman. Bounded and stably bounded palindromic difference equations of first order. *ELA*, 15:22–49, January 2006.

Rum06 [22] S. M. Rump. Eigenvalues, pseudospectrum and structured perturbations. *Linear Algebra Appl.*, 413(2-3):567–593, 2006.

Sch06 [23] Christian Schröder. A canonical form for palindromic pencils and palindromic factorizations. Preprint 316, TU Berlin, MATHEON, Germany, 2006.

Sch07 [24] Christian Schröder. A QR-like algorithm for the palindromic eigenvalue problem. Preprint 388, TU Berlin, MATHEON, Germany, 2007. submitted.

Sch07a [25] Christian Schröder. URV decomposition based structured methods for palindromic and even eigenvalue problems. Preprint 375, TU Berlin, MATHEON, Germany, March 2007. submitted.

Sch08 [26] Christian Schröder. *Palindromic and Even Eigenvalue Problems - Analysis and Numerical Methods*. PhD thesis, Technical University Berlin, Germany, 2008. in preparation.

Ste72 [27] G. W. Stewart. On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$. *SIAM J. Numer. Anal.*, 9:669–686, 1972.

SteS90 [28] G. W. Stewart and Ji-Guang Sun. *Matrix perturbation theory*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1990.

Tho76 [29] Robert C. Thompson. The characteristic polynomial of a principal subpencil of a Hermitian matrix pencil. *Linear Algebra Appl.*, 14:135–177, 1976.

Tho91 [30] Robert C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.

Wat96 [31] David S. Watkins. QR-like algorithms—an overview of convergence theory and practice. In *The mathematics of numerical analysis (Park City, UT, 1995)*, volume 32 of *Lectures in Appl. Math.*, pages 879–893. Amer. Math. Soc., Providence, RI, 1996.

Wat07 [32] David S. Watkins. *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods*. SIAM, Philadelphia, 2007.

WatE91 [33] David S. Watkins and Ludwig Elsner. Convergence of algorithms of decomposition type for the eigenvalue problem. *Linear Algebra Appl.*, 143:19–47, 1991.

WatE94 [34] David S. Watkins and Ludwig Elsner. Theory of decomposition and bulge-chasing algorithms for the generalized eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 15(3):943–967, 1994.

Wil88 [35] James H. Wilkinson. *The Algebraic Eigenvalue Problem.* Monographs on Numerical Analysis. Oxford Science Publications, Clarendon Press, Oxford, 1988.