The implicit application of a rational filter in the RKS method

Gorik De Samblanx *Karl Meerbergen* *Adhemar Bultheel*

Report TW 239, Revised December, 1996

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

The implicit application of a rational filter in the RKS method

Gorik De Samblanx *Karl Meerbergen*[*] *Adhemar Bultheel*

Report TW 239, Revised December, 1996

Department of Computer Science, K.U.Leuven

Abstract

The implicitly restarted Arnoldi method implicitly applies a polynomial filter to the Arnoldi vectors by use of orthogonal transformations. In this paper, an implicit filtering by rational functions is proposed for the rational Krylov method. This filtering is performed in an efficient way. Two applications are considered. The first one is the filtering of unwanted eigenvalues using exact shifts. This approach is related to the use of exact shifts in the implicitly restarted Arnoldi method. Second, eigenvalue problems can have an infinite eigenvalue without physical relevance. This infinite eigenvalue can corrupt the eigensolution. An implicit filtering is proposed for avoiding such corruptions.

AMS Subject Classification. 65F15
Keywords: Rational Krylov method, implicitly restarted Arnoldi, generalised eigenvalue problem, shift-invert

[*]LMS Numerical Technologies, Interleuvenlaan 70, 3001 Heverlee, Belgium, km@lmsnit.be

1 Introduction

In this paper, we study the rational Krylov sequence method (RKS), developed by Ruhe [19, 20, 21, 22]. This method is used for the calculation of a few eigenvalues of $Ax = \lambda Bx$ with A and B real $n \times n$ matrices, and where B may be singular. These eigenvalues can be the rightmost eigenvalues, the eigenvalues near the origin, eigenvalues in a certain region of the complex plane... These will be referred to as the 'wanted' eigenvalues. The method is named after the fact that it builds, step by step, an orthonormal basis for the space spanned by the rational Krylov sequence

$$v_1,\ S_1v_1,\ S_2S_1v_1, \ldots,\ S_kS_{k-1}\ldots S_1v_1, \tag{1}$$

where $S_j = (A - \mu_j B)^{-1}B$. This is the sequence computed by inverse iteration or Rayleigh quotient iteration. When all $\mu_j = \mu$, (1) is a standard Krylov sequence and the space spanned by (1) is a Krylov space of $S = (A - \mu B)^{-1}B$. The shift-invert Arnoldi method computes eigenvector estimates in this Krylov space. This method has been studied extensively [3, 2, 15, 17, 14, 23, 18, 12]. The RKS method differs from the shift-invert Arnoldi method in that the pole, μ_j, may change at any iteration. The rational Krylov method is a natural extension of the Arnoldi method, so it inherits many properties. For example, the Arnoldi method applied to S is characterised by the recurrence relation $SV_k = V_{k+1}H_k$ where H_k is a $k+1 \times k$ upper Hessenberg matrix. The rational Krylov method produces a similar recurrence relation $AV_{k+1}H_k = BV_{k+1}K_k$ where H_k and K_k are both $k+1 \times k$ upper Hessenberg matrices.

In a practical situation, it might happen that, for both Arnoldi's method and rational Krylov, the basis required for the accurate calculation of the wanted eigenvalues becomes too large to be stored by the computer. For the Arnoldi method, this problem was solved in an elegant manner by Sorensen. His implicitly restarted Arnoldi method (IRA) [25] applies p implicitly shifted QR steps on the Arnoldi Hessenberg matrix H_k resulting in a $k-p+1 \times k-p$ upper Hessenberg matrix H^+_{k-p}. Then it applies the corresponding orthogonal transformations to the columns of V_{k+1}. The resulting Arnoldi basis, V^+_{k-p+1}, has a reduced dimension and satisfies a new recurrence relation $AV^+_{k-p+1} = V^+_{k-p}H^+_{k-p}$. After this shrinking phase, the basis can be extended again. The method uses orthogonal transformations of small dimension, which leads to an efficient and reliable algorithm. One can show [12, 10] that IRA is an efficient implementation of polynomial subspace iteration, applied to an Arnoldi basis. The zeros of the polynomial are called shifts. Implicitly restarted Arnoldi acts as a filter that pushes away eigenvector components that correspond to the eigenvalues near the shifts.

In this paper, we present a similar idea for the rational Krylov method. The aim is to throw away unwanted directions from the rational Krylov space by applying a rational filter function to its basis. Since two matrices, H_k and K_k, are involved in the recurrence relation of RKS, we apply p QZ steps to the pencil (K_k, H_k) rather than p QR steps on a single matrix H_k. The QZ steps can be organised such that the Hessenberg structure of H_k and K_k is preserved. This operation is called the Implicitly Filtered Rational Krylov method (IFRKS). The zeros of the rational function are called shifts. They are chosen freely but different from the poles μ_j in (1). The poles of the rational filter are fixed values among the poles in (1), so they cannot be chosen by the user.

For the Arnoldi method, shifts are called 'exact shifts' when they are picked as the unwanted eigenvalues computed from H_k. For IRA, Morgan [13] gives a nice interpretation of the Krylov space obtained by exact shifts. He concludes that the implicitly restarted Arnoldi method with exact shifts is the 'optimal' way to restart an Arnoldi process. A procedure, similar to exact shifts, can be employed on RKS.

It is well known that the implicitly shifted QR algorithm is numerically backward stable, but may loose forward stability [16, 26], and so may do the implicitly restarted Arnoldi method. Lehoucq and Sorensen [9] presented a locking and purging strategy for Arnoldi's method to remove unwanted (converged) eigenvectors from the Krylov space while keeping the wanted ones. It is a numerically more stable alternative to an implicit restart. For rational Krylov, Ruhe [22] developed a similar approach. Theorem 8.1 in Lehoucq [10] shows, assuming exact arithmetic, an equivalence between purging and implicitly restarting Arnoldi's method. We prove a similar relation between IFRKS and purging in Theorem 5.2. The comparison of Ruhe's purging with our implementation of IFRKS (with exact shifts) shows that, at least for the example presented in §6.1, both methods are equivalent.

The connection between implicitly restarted Arnoldi and subspace iteration was used by Meerbergen and Spence [12] in order to avoid spurious eigenvalues that may arise, due to a singular B. In this case, $Ax = \lambda Bx$ has an infinite eigenvalue without physical relevance. The infinite eigenvalue can disturb the calculation of finite eigenvalues and can lead to so-called spurious eigenvalues. It is shown in [12] how the implicitly restarted Arnoldi method with zero shift applied to the shift-invert transformation decreases the possibility for spurious eigenvalues. We give an example where a similar filtering is employed in rational Krylov.

The plan of the paper is as follows. In §2, we present the RKS method and derive its recurrence relation. In §3, we collect some simple properties of Hessenberg matrices and in particular of the pencil (K_k, H_k) which will turn out to be useful in the remainder of the paper. In §4, the main theory and the algorithm for IFRKS are developed. Section 5 discusses the use of exact shifts and shows the relation with purging. Section 6 presents a comparison between IFRKS with exact shifts and the full rational Krylov process, i.e. without shrinking the subspace. It also illustrates how spurious eigenvalues can be avoided when B is singular. Finally, §7 closes the paper with the main conclusions and remaining questions.

Throughout the paper, matrices are denoted by upper case characters. Subscripts usually denote the number of columns of a matrix, e.g. $H_k \in \boldsymbol{C}^{k+1\times k}$. However, if the dimensions are clear from the context, then we drop the subscripts. Elements of H are denoted by $h_{i,j}$. Vectors are written as lower case Roman characters, and scalars by Greek characters. $\bar{\alpha}$ denotes the complex conjugate of α. By e_i, we denote the i-th unit vector of appropriate dimension. If $H_k \in \boldsymbol{C}^{k+1\times k}$ is a rectangular matrix, then $\underline{H}_k$ denotes the $k \times k$ upper part of H_k. x^T denotes the transpose and x^* the Hermitian transpose. We use $\|\cdot\|$ to denote the Euclidian norm as well as the spectral norm.

2 The rational Krylov method

The rational Krylov method differs from the shift-invert Arnoldi method at two separate points. The most important difference is that with RKS, the pole μ may change at every iteration. For Arnoldi, the pole must be fixed. Also, RKS collects the information about the eigenvalues not in one matrix H that is related to the shifted and inverted matrix $(A-\mu B)^{-1}B$, but in a matrix pair (K, H) that is directly related to (A, B).

In this section, we recall the basic facts about the RKS algorithm of Ruhe. We give the algorithm, the RKS relation and discuss the computation of the approximate eigenvalue pairs.

Algorithm 2.1 (Rational Krylov (RKS))
0. Given $v_1 \in \boldsymbol{C}^n$, $\|v_1\| = 1$.
   Let $V_1 = [v_1]$.
1. **for** $j = 1, ..., k$ **do**
   1.1. Select a pole μ_j and a continuation vector $t_j \neq 0 \in \boldsymbol{C}^j$.

1.2. Form $w = (A - \mu_j B)^{-1} B V_j t_j$.
1.3. Orthogonalise w against the columns of V_j and let $\underline{h}_j = V_j^* w$.
1.4. Normalise $v_{j+1} = w / h_{j+1,j}$ with $h_{j+1,j} = \|w\|$.
1.5. Compute the approximate eigenpair (θ, x) (see further).

Typically, in Step 1.2, $(A - \mu_j B)^{-1} B V_j t_j$ is computed by Gaussian elimination of $A - \mu_j B$. When this factorisation is expensive, it is recommended to keep successive μ_j unchanged for a few steps. Fixing μ_j can slow down the convergence in terms of 'iteration steps' but not necessarily in overall execution time. (We give an example in §6.1.) When Gaussian elimination is not feasible, one can use the inexact RKS method [10] with an iterative method. The orthogonalisation in Step 1.3 is performed by the (modified) Gram-Schmidt method with reorthogonalisation [1]. Reorthogonalisation turns out to be indispensable in order to obtain 'numerically' orthonormal basis vectors. Common choices for the *continuation vector* t_j are $t_j = e_j$, or t_j such that $V_j t_j = x$, the approximate eigenvector from Step 1.5.

Let us now derive the RKS relation. By eliminating w at step j of Algorithm 2.1, we have

$$V_{j+1} h_j = (A - \mu_j B)^{-1} B V_j t_j$$

with $h_j^T = [\underline{h}_j^T \quad h_{j+1,j}]^T$. This is equivalent to

$$A V_{j+1} h_j \quad = \quad B V_{j+1} \left(h_j \mu_j + \begin{bmatrix} t_j \\ 0 \end{bmatrix} \right) \; .$$

If we merge the corresponding equations from the previous iterations, we obtain after k steps

$$A V_{k+1} H_k = B V_{k+1} K_k \; ,$$

where

$$K_k = H_k M_k + T_k \; ,$$

$$\begin{aligned} H_k &= \begin{bmatrix} H_{k-1} & \underline{h}_k \\ 0 & h_{k+1,k} \end{bmatrix} \in \boldsymbol{C}^{k+1 \times k} && \text{is upper Hessenberg} \; , \\ M_k &= \mathrm{diag}(\mu_1, \ldots, \mu_k) \in \boldsymbol{C}^{k \times k} && \text{and} \\ T_k &= \begin{bmatrix} T_{k-1} & t_k \\ 0 & 0 \end{bmatrix} \in \boldsymbol{C}^{k+1 \times k} && \text{is upper triangular} \; . \end{aligned}$$

When the value of k is clear from the context, we drop the indices and the RKS relation is written as

$$AVH = BVK \; . \tag{2}$$

We may assume that $h_{j+1,j} \neq 0$ for $j = 1, \ldots, k$. If $h_{k+1,k} = 0$, then $\mathrm{Range}(V_k)$ is an invariant subspace and

$$A V_k \underline{H}_k = B V_k \underline{K}_k, \tag{3}$$

where $\underline{H}_k$ and $\underline{K}_k$ are the $k \times k$ upper parts of H_k and K_k respectively. At this point, the Gram-Schmidt process fails and the algorithm will stop. In the remainder of this paper, it is assumed that $h_{j+1,j} \neq 0$ for $j = 1, \ldots, k$. H_k is then called *unreduced*.

Definition 2.1 *An upper Hessenberg matrix is called unreduced when its subdiagonal elements are all nonzero.*

An RKS process is characterised by its matrices V, H and K. In order to make referring to the rational Krylov process easier, we introduce the term *RKS triple.*

Definition 2.2 *(V, H, K) with $V \in \boldsymbol{C}^{n\times k+1}$ and $K, H \in \boldsymbol{C}^{k+1\times k}$ is called an* RKS triple *of order k for (A, B) iff*

(A) $AVH = BVK$,

(B) K and H are upper Hessenberg matrices with H unreduced and

(C) none of the $\mu_j = k_{j+1,j}/h_{j+1,j}$, $j = 1, \ldots, k$ is an eigenvalue of (A, B).

It is clear that if V, H and K are produced by Algorithm 2.1, then (V, H, K) is an RKS triple. The pole μ_j is, of course, no eigenvalue of (A, B), since otherwise, $A - \mu_j B$ is not invertible in $\boldsymbol{C}^n$. There are various ways for the extraction of approximate eigenpairs from (2). In this paper, we use the following one.

Definition 2.3 *Given an RKS triple (V, H, K), then (θ, x) is called an* approximate eigenpair *if $x = V_k\underline{H}y$ and $\underline{K}y = \theta\underline{H}y$.*

A connection with harmonic Ritz values [24] in Range(V_k) is given in [22]. Observe that, as in the Arnoldi method, the residual

$$\begin{aligned} Ax - \theta Bx &= AV_k\underline{H}y - \theta BV_k\underline{H}y \\ &= AV_k\underline{H}y - BV_k\underline{K}y \\ &= -(A - \mu_k B)v_{k+1}h_{k+1,k}e_k^*y \end{aligned} \tag{4}$$

is small when $|h_{k+1,k}|$ and/or $|e_k^*y|$ are small.

Rather than computing eigenpairs, Ruhe [22] suggests the computation of a partial Schur form

$$A(V_kY_k) = B(V_kY_k)T_k \tag{5}$$

where $\underline{K}\,\underline{H}^{-1}Y_k = Y_kT_k$ is a Schur form of $\underline{K}\,\underline{H}^{-1}$. Y_k is a $k \times k$ unitary matrix, whose columns span the Schur basis, and T_k is $k \times k$ upper triangular. Its main diagonal elements are the eigenvalues of $\underline{K}y = \theta\underline{H}y$. In §5, we prove the connection between this partial Schur form and the RKS triple obtained by the implicit application of a rational filter with exact shifts. The major difference is that the implicit application of the filter results in H and K that are upper Hessenberg. The formation of $\underline{H}^{-1}$ in $\underline{K}\,\underline{H}^{-1}$ is not required. We also give a numerical example in §6.

We now proceed with some other properties of the Hessenberg pencil (K, H).

3 A note on Hessenberg matrices and RKS triples

In order to keep the elaboration in the next sections easy to understand, we review some properties of Hessenberg matrices, pencils and RKS triples. Most of these properties are obvious or well known. Because they will be used (frequently) in the rest of the paper and because they lead to some assumptions that are made implicitly, we bring them together in this section.

For a pair of upper Hessenberg matrices (K, H), we define operators $M(H, K)$ and $T(H, K)$ as follows.

Definition 3.1 *Let $H,\ K \in \boldsymbol{C}^{k+1\times k}$ be upper Hessenberg matrices with H unreduced. Define $M := M(H,K)$ a $k\times k$ diagonal matrix as $M = \operatorname{diag}(k_{j+1,j}/h_{j+1,j})$ and $T := T(H,K)$ a $k+1\times k$ upper triangular matrix as*

$$T = K - HM. \tag{6}$$

Since H is unreduced, $T(H,K)$ and $M(H,K)$ are uniquely defined. If (V,H,K) is an RKS triple of order k then $M := M(H,K)$ has the poles $\mu_1,...,\mu_k$ on its main diagonal and $T := T(H,K)$ collects the continuation vectors t_j.

Using Definition 3.1, we show the following properties.

Lemma 3.1 *Let F be an upper Hessenberg matrix in $\boldsymbol{C}^{k+1\times k}$. Consider the QR factorisation*

$$F = \begin{bmatrix} Q & q \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = QR,$$

with $Q \in \boldsymbol{C}^{k+1\times k}$ upper Hessenberg, $q \in \boldsymbol{C}^{k+1}$, $[Q \quad q]$ orthogonal and $R \in \boldsymbol{C}^{k\times k}$ an upper triangular matrix. Then:

(A) $q^*F = 0$.

(B) F is unreduced iff Q is unreduced and R has full rank.

(C) If F has full rank but $\underline{F}$ is singular, then

$$[Q\ q] = \begin{bmatrix} \tilde{Q} & 0 & \underline{q} \\ 0 & \gamma & 0 \end{bmatrix},$$

with $|\gamma| = 1$ and $\underline{q}^*\tilde{Q} = 0$.

(D) Suppose that F has full rank, then $q^*e_1 \neq 0$ iff F is unreduced.

Proof

(A–B) are trivial observations.

(C) If F has full rank and $\underline{F}$ is singular, then there is a $y \neq 0$ such that $Fy = \eta e_{k+1}$, with $\eta \neq 0$. Since Q is orthogonal and upper Hessenberg, $Ry = Q^*Fy = \tilde{\eta}e_k$, $\tilde{\eta} \neq 0$ since R is nonsingular. From $Q(Ry) = \eta e_{k+1}$, it follows that $Qe_k = (\eta/\tilde{\eta})e_{k+1} = \gamma e_{k+1}$.

(D) Since $q \perp \operatorname{Range}(F)$ and $\operatorname{Range}(F) + \operatorname{Range}(q) = \boldsymbol{C}^{k+1}$, we have $q^*e_1 \neq 0$ iff $e_1 \notin \operatorname{Range}(F)$.

If F is unreduced, there is no linear combination of the columns of F, that results in e_1, so $e_1 \notin \operatorname{Range}(F)$.

Let F be reduced, with $f_{j+1,j} = 0$. Decompose

$$F = \begin{bmatrix} \underline{F}_j & * \\ 0 & * \end{bmatrix},$$

with $\underline{F}_j$ the leading $j\times j$ submatrix of F.

$$F \begin{bmatrix} \underline{F}_j^{-1} \\ 0 \end{bmatrix} q = \begin{bmatrix} q \\ 0 \end{bmatrix} \in \operatorname{Range}(F)$$

for any $q \in \boldsymbol{C}^j$. It is easy to see that q can be chosen such that $[q^T \quad 0]^T = e_1$.

This shows the lemma. $\square$

The following Lemma collects some results on upper Hessenberg matrices that originate from an RKS triple.

Lemma 3.2 *Let (V_{k+1}, H_k, K_k) be an RKS triple of order k. Then*

(A) Let $\underline{K}_j$ and $\underline{H}_j$ be the $j+1 \times j$ leading submatrices of K_k, H_k respectively. If $M(H_k, K_k) = \mathrm{diag}(\mu_1, \ldots, \mu_k)$, then

$$K_k - \mu_j H_k = \begin{bmatrix} \underline{K}_j - \mu_j \underline{H}_j & * \\ 0 & * \end{bmatrix},$$

is not unreduced and

$$(A - \mu_j B)V_{j+1}H_j = BV_j(\underline{K}_j - \mu_j \underline{H}_j).$$

(B) For all $\alpha,\ \beta \in \boldsymbol{C}$, $|\alpha| + |\beta| \neq 0$ implies that $F = \alpha K - \beta H$ has full rank.

(C) For any $q \neq 0 \in \boldsymbol{C}^{k+1}$ for which $q^* H = 0$, we have $q^* K e_1 \neq 0$.

Proof

(A) is a trivial observation.

(B) can be proven as follows. If $\alpha\mu_j \neq \beta$, for $j = 1, \ldots, k$ then F is unreduced so it has full rank. If $\alpha\mu_j = \beta$, then F is not unreduced but we can rewrite (2) as

$$\alpha(A - \mu_j B)V_{k+1}H_k = \alpha B V_{k+1}(K_k - \mu_j H_k). \tag{7}$$

Note that if $|\alpha| + |\beta| \neq 0$ and $\alpha\mu_j = \beta$, then $\alpha \neq 0$. The left hand side in (7) has rank k, so $K_k - \mu_j H_k$ must also have rank k, even if B is singular.

(C) If $q^* H = 0$ with $q \neq 0$, then $\mathrm{Range}(H) \cup \mathrm{Range}(q) = \boldsymbol{C}^{k+1}$. Suppose that $Ke_1 \perp q$, so $Ke_1 \in \mathrm{Range}(H)$. Following (A), there is a γ such that $(K - \mu_1 H)e_1 = \gamma e_1$. Following (B), $\gamma \neq 0$. So, $Ke_1 = \mu_1 H + \gamma e_1$. Following Lemma 3.1(D), $\gamma e_1 \notin \mathrm{Range}(H)$, so $q^* K e_1 \neq 0$.

$\square$

Property (C) of the former lemma will be used along with Lemma 3.1(D) to show that the restarting procedure in §4 is always well defined.

In the remainder of this paragraph, we derive some properties of a 'shifted' RKS triple defined by the following transformation.

Definition 3.2 *Given $\alpha, \beta \in \boldsymbol{C}$, with $|\alpha| + |\beta| \neq 0$. Define the function $\gamma_{\alpha,\beta}(\mu) = (\bar{\beta}\mu + \bar{\alpha})/(\alpha\mu - \beta)$ for $\mu \in \{\mu \in \boldsymbol{C} \mid \alpha\mu \neq \beta\}$, and its inverse $\gamma^{-1}_{\alpha,\beta}(\nu) = (\bar{\alpha} + \beta\nu)/(\alpha\nu - \bar{\beta})$. Similarly, given two matrices H, K that have matching dimensions, define the mapping $(G, F) = \Gamma_{\alpha,\beta}(K, H)$ by*

$$F = \alpha K - \beta H \quad \text{and} \quad G = \bar{\beta} K + \bar{\alpha} H.$$

The inverse mapping $\Gamma^{-1}_{\alpha,\beta}(G, F) = (K, H)$ is given by

$$H = (|\alpha|^2 + |\beta|^2)^{-1}(\alpha G - \bar{\beta} F) \quad \text{and} \quad K = (|\alpha|^2 + |\beta|^2)^{-1}(\beta G + \bar{\alpha} F).$$

The transformation $\Gamma_{\alpha,\beta}$ defines a one-to-one correspondence between the sets of matrices (H, K) and (F, G). It defines a similar correspondence between the eigenvalues of square matrix pencils. If (λ, x) is an eigenpair of (A, B) and $\alpha\lambda \neq \beta$, then it is easily seen that $(\gamma_{\alpha,\beta}(\lambda), x)$ is an eigenpair of $\Gamma_{\alpha,\beta}(A, B)$. A similar result exists for 'shifted' RKS triples.

Lemma 3.3 *Let $\mu_1, \ldots, \mu_k$ and $\nu_1, \ldots, \nu_k$ be scalars such that $\alpha\mu_j \neq \beta$ and $\gamma_{\alpha,\beta}(\mu_j) = \nu_j$, $j = 1 \ldots, k$. Let $(G, F) = \Gamma_{\alpha,\beta}(K, H)$. Then (V, H, K) is an RKS triple for (A, B) with $M(H, K) = \mathrm{diag}(\mu_1, \ldots, \mu_k)$, iff (V, F, G) is an RKS triple for $\Gamma_{\alpha,\beta}(A, B)$ with $M(F, G) = \mathrm{diag}(\nu_1, \ldots, \nu_k)$.*

Proof We first show that if (V, H, K) is an RKS triple for (A, B), then (V, F, G) is an RKS triple for $(C, D) = \Gamma_{\alpha,\beta}(A, B)$. We show the three defining properties of Definition 2.2.

(A) From $AVH = BVK$, it follows that

$$(\bar{\beta}A + \bar{\alpha}B)V(\alpha K - \beta H) = (\alpha A - \beta B)V(\bar{\beta}K + \bar{\alpha}H),$$

so $CVF = DVG$.

(B) It is clear that F and G are upper Hessenberg. F is unreduced, since $\alpha\mu_j \neq \beta$ for $j = 1, \ldots, k$.

(C) The poles that correspond to (F, G) can be derived from

$$\begin{aligned} M(F, G) &= (\alpha M(H, K) - \beta I)^{-1}(\bar{\beta}M(H, K) + \bar{\alpha}I) \\ &= \mathrm{diag}(\gamma_{\alpha,\beta}(\mu_1), \ldots, \gamma_{\alpha,\beta}(\mu_k)) = \mathrm{diag}(\nu_1, \ldots, \nu_k). \end{aligned}$$

If λ is an eigenpair of (A, B), then $\mu_j \neq \lambda$ for $j = 1, \ldots, k$. Consequently, $\nu_j = \gamma_{\alpha,\beta}(\mu_j) \neq \gamma_{\alpha,\beta}(\lambda)$ is not an eigenvalue of (C, D).

Note that $\mu_1, \ldots, \mu_k$ and $\nu_1, \ldots, \nu_k$ are defined such that $\gamma_{\alpha,\beta}$ and $\Gamma_{\alpha,\beta}$ are invertible, so the lemma holds in the other direction as well. □

4 The implicit application of a rational filter

In this section, we derive an elegant algorithm for the reduction of the order of an RKS triple. The RKS triple (V, H, K) of order k is reduced to an RKS triple $(V^+ = VQ, H^+ = Q^*HZ, K^+ = Q^*KZ)$, of order $k - p$, where Q and Z are orthogonal matrices of suitable dimensions. In this paper, (V^+, H^+, K^+) is called a QZ transformation of (V, H, K). This section is concerned with a specific choice of transformation. (In Section 5, we consider another choice derived from Ruhe's purging strategy [22]. We show a connection with the transformation from this section.) More precisely, we shall prove the following result from Theorem 4.2. An RKS triple (V, H, K) can, given some parameters α, β, be reduced into a set (V^+, H^+, K^+) such that :

(A) (V^+, H^+, K^+) *is an RKS triple of order $k - 1$ for (A, B).*

(B) If $M(H, K) = \mathrm{diag}(\mu_1, \ldots, \mu_k)$, then $M(H^+, K^+) = \mathrm{diag}(\mu_2, \ldots, \mu_k)$.

(C) $\mathrm{Range}(V_j^+) = \mathrm{Range}((A - \mu_j B)^{-1}(\alpha A - \beta B)V_j)$ *for $j = 1, \ldots, k$.*

This result shows that the QZ transformation reduces the order of the RKS triple by one. It also shows an equivalence of one step of subspace iteration on the $\mathrm{Range}(V_j)$. The remainder of this section is devoted to the proof and interpretation of this result. In Lemma 4.1, we prove properties of a simple QZ transformation. This allows for a more elegant proof of Theorem 4.2. A natural extension of this theorem leads to Theorem 4.3 and Algorithm 4.1, which reduces a triple of order k to order $k-p$.

We first deduce a simple QZ transformation of an RKS triple in the following lemma.

Lemma 4.1 *Let (V, F, G) be an RKS triple of order k for (C, D). Consider the QR factorisation*

$$F = \begin{bmatrix} Q & q \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = QR \ ,$$

with $Q \in C^{k+1\times k}$, $q \in C^{k+1}$, $[Q \quad q]$ unitary, and $R \in C^{k\times k}$ upper triangular. Define a full rank upper Hessenberg matrix $Z \in C^{k\times k-1}$ for which $q^*GZ = 0$. Define $V^+ = VQ$, $F^+ = Q^*FZ$, and $G^+ = Q^*GZ$, then (V^+, F^+, G^+) is an RKS triple of order $k-1$. Moreover, if $M(F, G) = \mathrm{diag}(\nu_1, \ldots, \nu_k)$, then $M(F^+, G^+) = \mathrm{diag}(\nu_2, \ldots, \nu_k)$.

Proof We prove that (V^+, F^+, G^+) satisfies the defining properties in Definition 2.2.

(A) By plugging $I = QQ^* + qq^*$ in $CVF = DVG$ and by multiplying on the right by Z, we obtain

$$\begin{aligned} CV(QQ^* + qq^*)FZ &= DV(QQ^* + qq^*)GZ \\ CVQQ^*FZ &= DVQQ^*GZ, \end{aligned}$$

since $q^*FZ = 0 = q^*GZ$. This shows that $CV^+F^+ = DV^+G^+$.

(B) Let $M = M(F, G)$ and $T = T(F, G)$. Recall that $G = FM + T = QRM + T$. Since $q^*GZ = q^*TZ = 0$, TZ lies in the range of Q, i.e. there exist $S \in C^{k\times k-1}$ such that $TZ = QS$. Since Q and TZ are upper Hessenberg, S is upper triangular. As a result

$$G^+ = Q^*GZ = RMZ + S \text{ and } F^+ = Q^*FZ = RZ$$

imply that G^+ and F^+ are upper Hessenberg. Let $g = G^*q$. Following Lemma 3.2(C), it holds that $g^*e_1 \neq 0$. Since $g^*Z = 0$ and Z is of full rank, it follows from Lemma 3.1(D) that Z must be unreduced. Hence, F^+ is unreduced.

(C) Recall that $F^+ = RZ$ and $G^+ = RMZ + S$. The i-th subdiagonal elements of F^+ and G^+ are $f^+_{i+1,i} = r_{i+1,i+1}z_{i+1,i}$ and $g^+_{i+1,i} = (r_{i+1,i+1}\nu_{i+1})z_{i+1,i}$ respectively. Since F^+ is unreduced, we have $\nu^+_i = g^+_{i+1,i}/f^+_{i+1,i} = \nu_{i+1}$ for $i = 1, \ldots, k-1$. These are no eigenvalues of (C, D), since (V, F, G) is an RKS triple.

□

The lemma shows how an RKS triple can be transformed into a new RKS triple (V^+, F^+, G^+). It is easy to see that an upper Hessenberg Z always exists. Moreover, if we normalise Z to be a unitary upper Hessenberg matrix with real positive subdiagonal elements, then Z is uniquely defined. We do not explain here how to compute Z in practice. There are various possibilities. Algorithm 6.1 in §6 presents the algorithm that we used in our numerical experiments.

We are now ready to extend these results to a more general QZ transformation. We shall use the definition of $\Gamma_{\alpha,\beta}$ and lemma 3.3.

Theorem 4.2 *Let (V, H, K) be an RKS triple of order k for (A, B). Let $\alpha M(H, K) - \beta I$ be nonsingular, and $(G, F) = \Gamma_{\alpha,\beta}(K, H)$. Consider the QR factorisation*

$$F = \begin{bmatrix} Q & q \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = QR \ ,$$

with $Q \in \mathbf{C}^{k+1\times k}$, $q \in \mathbf{C}^{k+1}$, $[Q \quad q]$ unitary, and $R \in \mathbf{C}^{k\times k}$ upper triangular. Suppose that Z is a full rank upper Hessenberg matrix such that $q^* G Z = 0$. Define $V^+ = VQ$, $H^+ = Q^* H Z$ and $K^+ = Q^* K Z$. Then:

(A) (V^+, H^+, K^+) *is an RKS triple of order $k-1$ for (A, B).*

(B) If $M(H, K) = \mathrm{diag}(\mu_1, \ldots, \mu_k)$, *then* $M(H^+, K^+) = \mathrm{diag}(\mu_2, \ldots, \mu_k)$.

(C) $\mathrm{Range}(V_j^+) = \mathrm{Range}((A - \mu_j B)^{-1}(\alpha A - \beta B) V_j)$ *for $j = 1, \ldots, k$.*

Proof

(A) Following Lemma 3.3, (V, F, G) is an RKS triple of order k for $\Gamma_{\alpha,\beta}(A, B)$. Following Lemma 4.1 (V^+, F^+, G^+) is an RKS triple of order $k-1$ for $\Gamma_{\alpha,\beta}(A, B)$. Following Lemma 3.3, (V^+, H^+, K^+) is an RKS triple of order $k-1$ for (A, B).

(B) From Lemmas 4.1 and 3.3, it follows that $\mu_j^+ = \gamma_{\alpha,\beta}^{-1}(\nu_j^+) = \gamma_{\alpha,\beta}^{-1}(\nu_{j+1}) = \mu_{j+1}$, for $j = 1, \ldots, k-1$.

(C) Let R_j denote the $j \times j$ upper left part of R and let Q_j be the $j+1 \times j$ upper left part of Q. Then using $AVH = BVK$ and $F = \alpha K - \beta H = QR$, combined with Lemma 3.2(A), we derive

$$\begin{aligned} (A - \mu_j B) V_{k+1} (\alpha K_k - \beta H_k) &= (\alpha A - \beta B) V_{k+1} (K_k - \mu_j H_k) \\ (A - \mu_j B) V_{j+1} Q_j R_j &= (\alpha A - \beta B) V_j (\underline{K}_j - \mu_j \underline{H}_j) \\ V_j^+ &= (A - \mu_j B)^{-1} (\alpha A - \beta B) V_j [(\underline{K}_j - \mu_j \underline{H}_j) R_j^{-1}]. \end{aligned}$$

Note that following Lemma 3.2(B), $(\underline{K}_j - \mu_j \underline{H}_j) R_j^{-1}$ is of full rank.

□

It follows from (C) that the space spanned by V_j^+ can be regarded as having been computed by subspace iteration with a Cayley (or Möbius) transform applied to V_j. From a practical point of view, this is quite interesting since the QZ step provides us with a cheap implementation of subspace iteration. The zero of the Cayley transform depends on α and β and can be chosen freely. The only restriction is that $\alpha\mu_j \neq \beta$ for $j = 1, ..., k$, i.e. a pole of the Cayley transform cannot be chosen as shift. The condition $\alpha\mu_j \neq \beta$ is required to obtain a new RKS triple. One can check that if $\alpha\mu_j = \beta$ for some $1 \leq j \leq k$, then $\mathrm{Range}(V_j^+) = \mathrm{Range}(V_j)$. The recurrence relation $AV^+H^+ = BV^+K^+$ still holds, but H^+ and K^+ are no longer guaranteed to be upper Hessenberg and H^+ unreduced.

A second interpretation of Theorem 4.2 is that the RKS triple (V^+, H^+, K^+) can be viewed as having been computed by an RKS process with the poles $\mu_2, ..., \mu_k$ (the first pole μ_1 disappears), without the need for a complete recalculation of the RKS process.

A similar interpretation led to the name of the implicitly restarted Arnoldi method. If $A = I$, $B = S$ and $K = I$, then the RKS relation becomes $V_{k+1}H = SV_k$, which is the recurrence relation of Arnoldi's method applied to S. With $\alpha = -\sigma$, and $\beta = -1$, we have

$F = H - \sigma I$ and $G = I + \bar{\sigma} H$, so $q^* G Z = 0$ when $Z = \tilde{Q}$ where $\tilde{Q}$ denotes the $k \times k-1$ upper left part of Q. Following Theorem 4.2, we have that $V^+ = VQ$ and $H^+ = Q^* H \tilde{Q}$ satisfy $V_k^+ H^+ = S V_{k-1}$ and $\mathrm{Range}(V_j^+) = \mathrm{Range}((S - \sigma I) V_j)$ for $j = 1, ..., k$. This result corresponds with Theorem 3 in [12] and Lemma 6.1 in [10].

After one QZ step, the order of the RKS triple is reduced from k to $k-1$. This process can be repeated in order to reduce the order to $k-p$. The following theorem shows the accumulated effect.

Theorem 4.3 *Let (V, H, K) be an RKS triple of order k for (A, B) with $M(H, K) = \mathrm{diag}(\mu_1, ..., \mu_k)$. Given shift pairs (α_j, β_j) with $\alpha_j M(H, K) - \beta_j I$ nonsingular for $j = 1, ..., p$. Then Algorithm 4.1 produces an RKS triple (V^+, H^+, K^+) of order $k-p$ for (A, B), such that $M(H^+, K^+) = \mathrm{diag}(\mu_{p+1}, ..., \mu_k)$ and*

$$\mathrm{Range}(V_{k-p+1}^+) = \mathrm{Range}\left(S_p S_{p-1} \cdots S_1 V_{k-p+1}\right), with S_j = (A - \mu_{k-p+j} B)^{-1} (\alpha_j A - \beta_j B) \ . \quad (8)$$

Proof Let us first prove that $M(H^+, K^+) = \mathrm{diag}(\mu_{p+1}, ..., \mu_k)$. Following Theorem 4.2, we have, after one restart, $M(H^+, K^+) = \mathrm{diag}(\mu_2, ..., \mu_k)$, independent of the choice of shift. So, we have that after p restarts $M(H^+, K^+) = \mathrm{diag}(\mu_{p+1}, ..., \mu_k)$.

Let us now show (8) for $p = 2$. Given V_{k+1}, H_k, K_k. After the first restart, we have, with $j = k-1$ in Theorem 4.2,

$$\begin{aligned} \mathrm{Range}(V_{k-1}^+) &= \mathrm{Range}([(A - \mu_{k-1} B)^{-1} (\alpha_1 A - \beta_1 B)] V_{k-1}) \ , \\ \mu_{k-1}^+ &= \mu_k \ . \end{aligned}$$

Call $W_k \equiv V_k^+$, then after the second restart, we have

$$\begin{aligned} \mathrm{Range}(W_{k-1}^+) &= \mathrm{Range}([(A - \mu_{k-1}^+ B)^{-1} (\alpha_2 A - \beta_2 B)] W_{k-1}) \ , \\ &= \mathrm{Range}\left([(A - \mu_k B)^{-1} (\alpha_2 A - \beta_2 B)][(A - \mu_{k-1} B)^{-1} (\alpha_1 A - \beta_1 B)] V_{k-1}\right) \ . \end{aligned}$$

By further induction, we obtain the proof. □

The shifts (α_j, β_j) determine a rational transformation

$$\psi(\theta) = \prod_{j=1}^{p} \frac{\alpha_j \theta - \beta_j}{\theta - \mu_{k-p+j}} \quad (9)$$

on the eigenvalues of $Ax = \lambda Bx$. With a good choice of the shift pairs (α_j, β_j), we are able to filter away any part of the spectrum and to enhance the regions near the last p shifts.

Algorithm 4.1 (Implicitly Filtered Rational Krylov (IFRKS))

0. Given (V, H, K).
1. Select $p < k$, $\alpha_1, \beta_1, ..., \alpha_p, \beta_p$.
2. Set $Q_1 = I_{k+1}$.
3. **for** $j = 1, ..., p$ **do**
   3.1. Compute the QR factorisation $[Q \ \ q] \begin{bmatrix} R \\ 0 \end{bmatrix} = \alpha_j K - \beta_j H$.
   3.2. Determine Z such that $q^* (\bar{\beta}_j K + \bar{\alpha}_j H) Z = 0$ (e.g. by Algorithm 6.1).
   3.3. Let $Q_1 \leftarrow Q_1 Q$.
   3.4. Let $H \leftarrow Q^* H Z$.
   3.5. Let $K \leftarrow Q^* K Z$.
4. Let $V \leftarrow V Q_1$

5 The choice of the filter ψ

The analysis of problems coming from fluid dynamics often lead to eigenvalue problems that have a singular B. If B is singular but $A - \mu_j B$ is not, then infinity is an eigenvalue of (A, B). Since this eigenvalue is of no physical interest, its computation is undesirable. Moreover, it can disturb the convergence of the finite eigenvalues of (A, B). A similar situation occurs in the analysis of buckling problems [5, 6]. In this case, the matrix A is singular, and the eigenvalues near zero are wanted. The zero eigenvalue itself is unwanted.

These are only special cases of the more general observation that the filter can be chosen to avoid convergence to any known eigenvalue that we do not want to compute. Indeed, by playing with the shift pairs (α_j, β_j) in the rational filter $\psi(\theta)$ from (9), we can filter away any unwanted eigenspace. An example of a known eigenvalue that is unwanted, is an infinite eigenvalue (if B is singular) or a zero eigenvalue (if A is singular).

Suppose that we are not interested in the calculation of a known eigenvalue σ of $Ax = \lambda Bx$. Then this calculation can be avoided by applying the matrix $\Psi = (A - \mu_k B)^{-1}(A - \sigma B)$ to V_k. The eigenspace corresponding to σ is equal to the nullspace of Ψ and thus it will disappear from Range(ΨV_k). This filtering can be done by Algorithm 4.1 using a shift pair $(\alpha, \beta) = (1, \sigma)$. An illustration is given in §6.2.

The mapping from (V, H, K) to (V^+, H^+, K^+) can also be used to reduce the order of the RKS triple. This is extremely useful, when the wanted eigenvalues are not yet computed accurately enough, but increasing the order of the RKS triple is not feasible, due to increasing computer memory cost, for example. By the selection of a good rational filter, only the wanted eigenvector components are kept in the reduced Krylov space and the convergence of the RKS algorithm will hopefully not slow down too much. Suppose that we have computed a number of approximate eigenvalues $\theta_1, \ldots, \theta_k$, but that we are not interested in $\theta_1, \ldots, \theta_p$. Then, we can throw away the latter eigenvalues by IFRKS using shift pairs $(\alpha_j, \beta_j) = (1, \theta_j)$, $j = 1, \ldots, p$. For example, when we are interested in the rightmost eigenvalues, we will throw away the eigenspaces of the leftmost approximate eigenvalues of (V, H, K). We call $(\alpha_j, \beta_j) \equiv (1, \theta_j)$ an *exact shift (pair)* by analogy with exact shifts in the implicitly restarted Arnoldi method [25]. We shall show in Theorem 5.2 that IFRKS with exact shifts indeed purges $\theta_1, \ldots, \theta_p$ and leaves the other approximate eigenvalues untouched. There is also a nice interpretation in terms of the generalised Schur form of $(\underline{K}, \underline{H})$. This relation suggests a link with the purging strategy by Ruhe [22].

Consider the following generalised Schur form of $(\underline{K}, \underline{H})$ [4, p.396] :

$$\underline{K}[U_{k-p} \quad U_p] = [Y_{k-p} \quad Y_p] \begin{bmatrix} T_{k-p} & * \\ 0 & T_p \end{bmatrix}, \tag{10a}$$

$$\underline{H}[U_{k-p} \quad U_p] = [Y_{k-p} \quad Y_p] \begin{bmatrix} S_{k-p} & * \\ 0 & S_p \end{bmatrix}, \tag{10b}$$

with Y, U unitary matrices and S, T square upper triangular. We assume that $\underline{H}$ is nonsingular. Let the Schur form be ordered such that $[\theta_1, \ldots, \theta_p] = \text{diag}(T_p)/\text{diag}(S_p)$ and $[\theta_{p+1}, \ldots, \theta_k] = \text{diag}(T_{k-p})/\text{diag}(S_{k-p})$. We will show that implicit filtering with exact shifts is equivalent to deleting the right part of this Schur form.

We first show the following lemma for the shifted problem.

Lemma 5.1 *Let (V, F, G) be an RKS triple for (C, D). Suppose that (α, β) are such that*

$$\underline{G}[U \quad u] = [Y \quad y] \begin{bmatrix} S & s \\ 0 & \sigma \end{bmatrix}$$

$$\underline{F}[U \quad u] \;=\; [Y \quad y]\begin{bmatrix} T & t \\ 0 & 0 \end{bmatrix}$$

is a generalised Schur form of $(\underline{G}, \underline{F})$, *where* $U,\, Y \in \boldsymbol{C}^{k\times k-1}$ $u,\, y \in \boldsymbol{C}^k$, $[U \quad u]$ *and* $[Y \quad y]$ *are unitary matrices,* $T,\, S \in \boldsymbol{C}^{k-1\times k-1}$ *are upper triangular and* $t,\, s \in \boldsymbol{C}^{k-1}$ *and* $\sigma \neq 0$. *With the definitions of Lemma 4.1,*

$$\underline{G}^+ U^+ = Y^+ S \qquad \underline{F}^+ U^+ = Y^+ T \ ,$$

is a generalised Schur form of $(\underline{G}^+, \underline{F}^+)$, *where* U^+ *and* Y^+ *are* $k-1 \times k-1$ *unitary matrices.*

Proof It is clear that if F is unreduced, but $\underline{F}$ is singular, then $F = QR$ satisfies Lemma 3.1(C). Since $\underline{F} = \tilde{Q}R$, we have $\underline{F}[U \quad u] = \tilde{Q}R[U \quad u] = Y[T \quad t]$. So, $y^*\tilde{Q} = 0$ and $\mathrm{Range}(\tilde{Q}) = \mathrm{Range}(Y)$. From $\underline{q}^*\underline{G}Z = 0$, it follows that

$$\begin{aligned} \underline{q}^*\underline{G}Z &= \underline{q}^*[Y \quad y]\begin{bmatrix} S & s \\ 0 & \sigma \end{bmatrix}\begin{bmatrix} U^* \\ u^* \end{bmatrix} Z = 0 \ , \ \text{so} \\ \sigma u^* Z &= 0 \ . \end{aligned}$$

Finally, it follows from $y^*\tilde{Q} = 0$ and $u^*Z = 0$ that

$$\begin{aligned} \underline{G}^+ &= \tilde{Q}^*\underline{G}Z = \tilde{Q}^*[Y \quad y]\begin{bmatrix} S & s \\ 0 & \sigma \end{bmatrix}\begin{bmatrix} U^* \\ u^* \end{bmatrix} Z \\ &= (\tilde{Q}^*Y)S(U^*Z) \ , \end{aligned}$$

and similarly, $F^+ = (\tilde{Q}^*Y)T(U^*Z)$. Since $\mathrm{Range}(\tilde{Q}) = \mathrm{Range}(Y)$, we have that $Y^+ = \tilde{Q}^*Y$ is unitary. Similarly, it follows from $\mathrm{Range}(U) = \mathrm{Range}(Z)$ that $U^+ = Z^*U$ is unitary, which completes the proof. □

Recursive application of this lemma with $\alpha = 1$ and $\beta = \theta_j$ leads to the important connection between the truncation of the generalised Schur form (10) and the result in Theorem 4.3.

Theorem 5.2 *Let* (V, H, K) *be an RKS triple of order* k. *Let* $\underline{H}$ *be nonsingular. Consider the generalised Schur form (10) with the given ordering. Assume that* $M(H,K) - \theta_j I$, $j = 1, \ldots, p$ *is nonsingular. By applying IFRKS with shifts* $(\alpha_j, \beta_j) = (1, \theta_j)$, *we obtain* (K^+, H^+) *with the generalised Schur form*

$$\underline{K}^+ U^+ = Y^+ T_{k-p} \quad , \quad \underline{H}^+ U^+ = Y^+ S_{k-p} \ .$$

for some unitary matrices U^+ *and* Y^+. *Moreover,* $V_k Y_{k-p} = V_{k-p}^+ Y^+$ *and* $V_k \underline{H} U_{k-p} = V_{k-p}^+ \underline{H}^+ U^+$.

Proof The proof follows by induction on p and by using Lemma 5.1. For $p = 1$, for example, note that $\alpha_1 T_p - \beta_1 S_p = 0$ and $\sigma = \bar{\beta}_1 T_p + \bar{\alpha}_1 S_p = (|\theta_1|^2 + 1) S_p \neq 0$ such that Lemma 5.1 applies.

We show the equivalence of the Schur bases for $p = 1$. Recall from the proof of Lemma 5.1 that $\mathrm{Range}(\tilde{Q} = \mathrm{Range}(Y)$. This implies $V_k Y_{k-p} = V_k \tilde{Q}\tilde{Q}^* Y_{k-p} = V_{k-p}^+ Y^+$. From the generalised Schur forms, we derive that $V_{k-p}^+ \underline{H}^+ U^+ = V_{k-p}^+ Y^+ S_{k-p}$ and $V_k \underline{H} U_{k-p} = V_k Y_{k-p} S_{k-p}$. This implies that $V_k \underline{H} U_{k-p} = V_{k-p}^+ \underline{H}^+ U^+$. □

The conclusion of this theorem is that the Schur form of (V^+, H^+, K^+) is contained in the Schur form of (V, H, K). The approximate eigenvectors corresponding to $\theta_{p+1}, ..., \theta_k$ are the same. This theorem has a clear link to the purging idea by Ruhe [22]. This idea is related to the work by Lehoucq and Sorensen [9] for the Arnoldi method. Ruhe proposes to compute the Schur form

$$\underline{K}\underline{H}^{-1}[W_{k-p} \;\; W_p] = [W_{k-p} \;\; W_p] \begin{bmatrix} R_{k-p} & * \\ 0 & R_p \end{bmatrix} , \tag{11}$$

ordered such that $\text{diag}(R_p) = [\theta_1, ..., \theta_p]$ and $\text{diag}(R_{k-p}) = [\theta_{p+1}, ..., \theta_k]$. It is easy to see that the Schur form (10) can be transformed into (11) with $W_i = Y_i$ and $R_i = T_i S_i^{-1}$ for $i = p, k-p$. The unwanted eigenvalues $\theta_1, ..., \theta_p$ are purged by dropping the last block column in (11). To do this within an RKS framework, Ruhe suggests the computation of $V^+ = VQ$, $H^+ = Q^* HZ$ and $K^+ = Q^* KZ$ with

$$Q = \begin{bmatrix} W_{k-p} & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad Z = \underline{H}^{-1} W_{k-p} \ . \tag{12}$$

The following theorem shows the effect of this operation.

Proposition 5.3 *Let (V, H, K) be an RKS triple of order k. Assume that $\underline{H}$ is nonsingular and consider the Schur form (11). Using (12), define $V^+ = VQ$, $H^+ = Q^* HZ$ and $K^+ = Q^* KZ$. Then $AV^+H^+ = BV^+K^+$, but H^+ and K^+ are no longer upper Hessenberg and H^+ unreduced. Moreover, $\underline{K}^+(\underline{H}^+)^{-1}W^+ = W^+R_{k-p}$ is a Schur form. In addition, the Schur bases $V_k W_{k-p}$ and $V_{k-p}^+ W^+$ are the same.*

Proof First note that $I = QQ^* + \begin{bmatrix} W_p \\ 0 \end{bmatrix} \begin{bmatrix} W_p^* & 0 \end{bmatrix}$. By plugging this matrix in $AVH = BVK$, and multiplying on the right by Z, we have

$$AV_{k+1}QQ^*H_kZ + AV_kW_pW_p^*\underline{H}_kZ = BV_{k+1}QQ^*K_kZ + BV_kW_pW_p^*\underline{K}_kZ \ .$$

It is easy to see that $W_p^*\underline{H}Z = 0 = W_p^*\underline{K}Z$ such that indeed $AV_{k-p+1}^+H_{k-p}^+ = BV_{k-p+1}^+K_{k-p}^+$.

Direct substitution gives

$$\underline{K}^+(\underline{H}^+)^{-1} = (W_{k-p}^*\underline{K}Z)(I)^{-1} = W_{k-p}^*\underline{K}\underline{H}^{-1}W_{k-p} = R_{k-p}$$

and $W^+ = I$. Finally, from $V_{k-p+1}^+ = V_{k+1}Q$, it follows that $V_{k-p}^+W^+ = V_kW_{k-p}$, which completes the proof. □

The equivalence between (11) and (10) automatically implies the relation with IFRKS.

6 Numerical examples

In this section, we present two numerical examples. The first one illustrates the use of exact shifts. The second one shows the filtering of the infinite eigenvalue.

We first give the algorithm for the computation of Z, used in our calculations. As was mentioned before, there are various ways for the calculation of Z. Among our implementations, the following algorithm performed best in all tests.

Algorithm 6.1 (Computation of Z_{k-1})
0. Given $G \in C^{k+1 \times k}$, $q \in C^{k+1}$.
1. Compute $g = [g_1, \ldots, g_k] = q^* G$, with $g_1 \neq 0$.
2. Set $\tilde{Z}_1 \leftarrow I_{2,1}$.
3. **for** $j = 2, \ldots, k-1$ **do**
   3.1. Compute $l \; : \; |g_l| = \max\{|g_i| \; : i \leq j\}$.
   3.2. Set $\tilde{Z}_j = \begin{bmatrix} \tilde{Z}_{j-1} & -g_{j+1} e_l \\ 0 & g_l \end{bmatrix}$ $\quad$ ($e_l \in C^j$ is the l-th unit vector).
4. Set $Z_{k-1} = \tilde{Z}_{k-1} U$, with $U \in C^{k-1 \times k-1}$ upper triangular such that $Z_{k-1}^* Z_{k-1} = I$.

The algorithm constructs Z_{k-1} explicitly such that $q^* G Z = g Z_{k-1} = g \tilde{Z}_{k-1} U = 0$. The index l is such that the largest elements of each column of Z is on the subdiagonal. Following Lemma 3.1(D), we have that $g_1 \neq 0$, which implies that Z_{k-1} is unreduced. (See also Lemma 4.1.)

6.1 The use of exact shifts

The example compares the IFRKS method with exact shifts and the full RKS method, i.e. without implicit filtering. Morgan [13] argues that implicitly restarted Arnoldi may converge as fast as full Arnoldi. His conclusion is that implicitly restarted Arnoldi is the optimal procedure for restarting the Arnoldi process. Here we give an example that shows similar behaviour for the rational Krylov method.

The Brusselator equations [7],

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{D_u}{L^2}\left[\frac{\partial^2 u}{\partial X^2} + \frac{\partial^2 u}{\partial Y^2}\right] - (B+1)u + u^2 v + C \\ \frac{\partial v}{\partial t} &= \frac{D_v}{L^2}\left[\frac{\partial^2 v}{\partial X^2} + \frac{\partial^2 v}{\partial Y^2}\right] - u^2 v + Bu \ , \end{aligned}$$

for u and $v \in (0,1)\times(0,1)$ with homogeneous Dirichlet boundary conditions form a 2D reaction-diffusion model where u and v represent the concentrations of two reactants. The equations are discretised with central differences with grid-sizes $h_u = h_v = 1/(m+1)$ with $m = (n/2)^{1/2}$. For $x^* = [u_{1,1}, v_{1,1}, u_{1,2}, v_{1,2}, \ldots, u_{m,m}, v_{m,m}]$, the discretised equations can be written as $\dot{x} = f(x)$. We wish to compute the rightmost eigenvalues λ_1 and λ_2 of the Jacobian matrix $A = \partial f / \partial x$ for $B = 5.45$, $C = 2$, $D_u = 0.004$, $D_v = 0.008$ and $L = 1$ for $u(X,Y) = C$ and $v(X,Y) = B/C$. The matrix A that we used has dimension 968.

In order to make a fair comparison, we consider three related algorithms. The experiments were run in Matlab on a DEC5000/33. The three methods use the pole $\mu = 10$ in the first four iterations and the same initial vector, v_1, chosen randomly using the Matlab function `randn`. Figure 1 shows, for the computed eigenpair (θ, x), the residual norm $\|Ax - \theta x\|/\|x\|$ in function of the RKS iteration and Figure 2 shows the number of Matlab flops in function of the residual norm.

Method 1 First, we performed the 'full RKS algorithm' (see the circles in Figures 1 and 2). From the fifth iteration on, the pole was set to the rightmost eigenvalue of $(\underline{K}, \underline{H})$ and the continuation vector equal to the corresponding eigenvector. Full RKS appears to converge quadratically, as is shown in Figure 1.

Method 2 We used the sparse matrix features of Matlab for efficient sparse factorisation. Nevertheless, factorisations are still relatively expensive. Therefore, we did four RKS steps

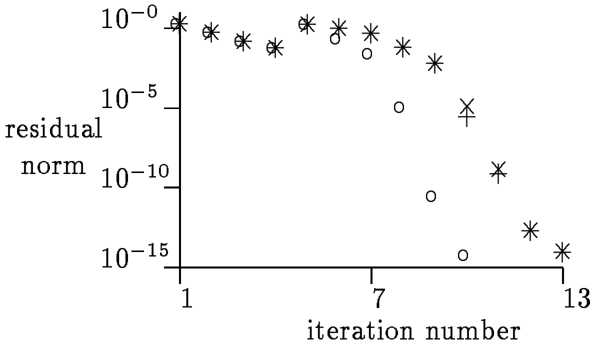

Figure 1: Residual norms for the rightmost eigenvalue of the Brusselator model versus iteration number. Legend: ∘ : Method 1 (full RKS); + : Method 2 (RKS with one factorisation per four iterations); × : Method 3 (four iterations of RKS alternated with four IFRKS steps).

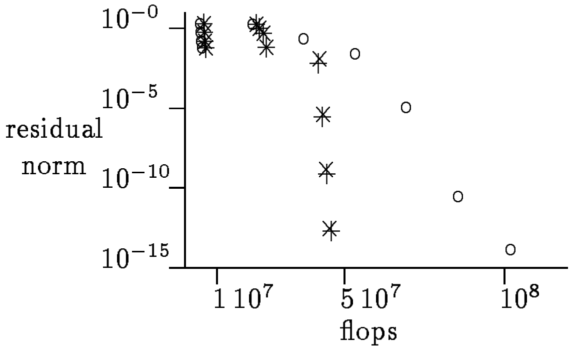

Figure 2: Residual norms versus number of MATLAB flops. Same legend as in Figure 1. The figure illustrates the difference between convergence in terms of iteration steps and computational effort.

per factorisation, rather than a factorisation on each iteration in Method 1. Method 2 requires more iterations than Method 1 for the same order of accuracy, but only 3 factorisations were required instead of 7 for Method 1. Figure 2 shows that this has implications on the overall computation time : Method 2 appears to be less expensive than Method 1.

Method 3 Here, we used the same strategy for the selection of poles as in Method 2, but from $k = 8$ on, we performed, after each set of four RKS steps (with the same pole), four IFRKS steps with exact shifts, selected as the four leftmost eigenvalues of $\underline{K}y = \theta \underline{H}y$. So, the order of the RKS triple remains limited to eight. Figure 1 shows a minor difference with Method 2. The cost for the orthogonalisation and the computation of the eigenvalues of $(\underline{K}, \underline{H})$ are smaller when the size of V_k is restricted, so, one would expect an additional iteration step to be cheaper for Method 3 than for Method 2. For Method 3, there is, however, the additional cost for the QR factorisation of F and the calculation of V^+, H^+ and K^+. Figure 2 shows that in this example, both methods have the same order of computational work, though Method 2 was slightly more expensive. It is pleasing that Method 3 has about the same convergence and timings results as Method 2, though requires much less basis vectors. When the Krylov space dimension becomes larger in Method 2, the difference in computational cost for one iteration with Method 3 can dramatically increase. However, it is important to note that Method 3 works with a small subspace, which can lead to much slower convergence than Method 2, in terms of iterations but also of execution time. We repeated this experiment with Ruhe's purging scheme (see Proposition 5.3) and found the same results.

6.2 Filtering the infinite eigenvalue

The second example comes from a model of viscous free-surface fluid flow on a tilted plane [11]. The Navier-Stokes equations were discretised by a finite element approach leading to an eigenvalue problem $Ax = \lambda Bx$ of size $n = 536$. The matrices A and B are nonsymmetric, B is singular (B has rank 429) and A is not. The goal is to find the rightmost eigenvalues used for the stability analysis of a steady state solution of the Navier-Stokes equations.

The main problem is that B is singular. This means that $Ax = \lambda Bx$ has an infinite eigenvalue. Moreover, for this example, this eigenvalue is defective. When we apply an iterative eigenvalue solver to this problem, this infinite eigenvalue may emerge as finite spurious eigenvalues. The calculation of spurious eigenvalues can be avoided as follows (see the work in [15, 2, 18, 12] for the shift-invert Arnoldi method). Let $\mathcal{R}$ denote the eigenspaces (and the generalised eigenspaces if the eigenvalues are defective) corresponding to the finite eigenvalues and $\mathcal{N}$ the eigenspace (and generalised eigenspace) corresponding to the infinite eigenvalue. Since $\boldsymbol{R}^n = \mathcal{R} + \mathcal{N}$, it is sufficient to remove the $\mathcal{N}$ component in V, in order to avoid the calculation of 'infinity'. Since an infinite eigenvalue of $Ax = \lambda Bx$ corresponds to a zero eigenvalue of $A^{-1}B$, $\mathcal{N} = \text{nullspace}((A^{-1}B)^\nu)$ where ν is the index of this eigenvalue. If $\nu = 1$, then the eigenvalue is non-defective. Thus, the calculation of spurious eigenvalues due to singular B can be avoided by filtering

$$V^+ \leftarrow \text{orth}((A^{-1}B)^\nu V)$$

or even by shifting the problem, e.g. as

$$V^+ \leftarrow \text{orth}((A - \mu_\nu B)^{-1} B \cdots (A - \mu_1 B)^{-1} BV) .$$

Such a filtering is easily done by the IFRKS method with shift pair $(0, 1)$.

For this example, the rightmost eigenvalues are known, so it was easy to verify after the calculation whether the computed rightmost eigenvalue was spurious. The rightmost eigenvalues are $\lambda_1 = -9.4883$, $\lambda_{2,3} = -11.6062 \pm 14.6602i$, and $\lambda_{4,5} = -15.9689 \pm 3.2343i$. The index of the infinite eigenvalue is not known, but we found that a small number of QZ steps were sufficient to avoid the calculation of spurious eigenvalues in this problem. We ran the following algorithm in Matlab on a Sun SPARC Station 4.

Algorithm 6.2

0. Let $v_1 = [1, \ldots, 1]^* / \sqrt{n}$, $V = [v_1]$
   Let $\mu = 0$ and $t_1 = [1]$.
1. **for** $l = 1$ **to** 3 **do**
   1.1. Factorise $LU = A - \mu B$
   1.2. **for** $j = 1 + 3(l-1)$ **to** $3l$ **do**
      1.2.1. Perform an RKS iteration with pole μ and continuation vector t_j
      1.2.2. Let the next continuation vector be $t_{j+1} = e_{j+1}$
   1.3. Filter away the infinite eigenvalue :
      1.3.1. Perform an RKS iteration with pole μ and continuation vector e_{j+1}
      1.3.2. Perform an IFRKS step with shift $(0, 1)$
   1.4. Compute eigenvalues of $\underline{K}y = \theta \underline{H}y$
   1.5. Order the eigenvalues following decreasing real part
   1.6. Extract the eigenvector y_1 corresponding to θ_1
   1.7. Let $\mu = \theta_1$ and next $t_{j+1} = [(\underline{H}y_1)^* \;\; 0]^*$

The algorithm consists of 9 RKS iterations. After the 3rd, the 6th and the 9th iteration, we performed one additional RKS step and one IFRKS step with shift $(0, 1)$. A new pole is selected according to this filtered RKS triple. Table 6.2 shows that the filter step is useful. The second column shows for Algorithm 6.2 the successive values of eigenvalue estimates, which are used for the selection of the poles, μ. It is clear that $\mu = \theta_1$ converges to $\lambda_1 = -9.4883$. The third column shows the sequence of eigenvalue estimates when Step 1.3 is omitted from Algorithm 6.2. The algorithm diverges to infinity, which is certainly not what we want. IFRKS filtering pulls the eigenvalues of $\underline{K}y = \theta \underline{H}y$ to the finite eigenvalues of $Ax = \lambda Bx$.

Table 1: Eigenvalue estimates, θ_1, in the RKS method for the tilted plane problem from §6.2

iteration	with IFRKS	without IFRKS
3	-8.4677	8.432
6	-9.4883	19.751
9	-9.48831	74.83

7 Conclusions

In this section, we summarise the main ideas of this work and pose some open questions.

This paper presented an implicit application of a rational polynomial to a rational Krylov space. It is a generalisation of the implicitly restarted Arnoldi method. In fact, the latter method is a special case of IFRKS and, as we have shown in §4, can be derived from Theorem 4.2. For our numerical examples, IFRKS appears to be efficient and reliable. This is due o the application of orthogonal QZ transformations on the RKS triple. The poles of the rational function are fixed. The zeros can be chosen freely, but different from the poles of the RKS method.

There are several interpretations of this method. First, IFRKS turns out to be an efficient implementation of rational subspace iteration. This property can be used for filtering unwanted eigenspaces, e.g. the infinite eigenvalue when B is singular. Second, when exact shifts are used, it is equivalent to a purging operation of a Schur form of the RKS triple.

The theory and the numerical example in §6.1 indicate that IFRKS is very helpful when one wishes to calculate accurate eigenvalue estimates in cases where the order of the RKS triple is bounded, e.g., due to memory limitations, or increasing cost for the Gram Schmidt orthogonalisation or the solution of the small eigenvalue problem.

The difference between Ruhe's purging strategy and IFRKS with exact shift pairs lies in the fact that purging does not preserve the Hessenberg structure of H^+ and K^+ and uses the inverse of $\underline{H}$. When $\mu_j \equiv \mu$ for $j = 1, ..., k$, the RKS method reduces to the shift-invert Arnoldi method. The implicitly restarted Arnoldi method works with a single Hessenberg matrix, H, instead of a pencil, and uses a QR step instead of a QZ one. So this method is much simpler than IFRKS. For this reason, it is wise to use the implicitly restarted Arnoldi method when the poles do not change. Arnoldi's method can no longer be used when poles do change and rational Krylov is the only choice.

There are still a number of open questions. The comparison between Ruhe's purging (Proposition 5.3) and IFRKS in §6.1 confirms, at least for this example and our implementation, the equivalence between both methods. An implicit restart in the Arnoldi method may fail due to forward instability of the implicitly shifted QR step [16, 26] on the Hessenberg matrix. Purging [9] appears to be more robust. In Arnoldi's method, the error on the basis vectors can substantially grow when the shifted Hessenberg matrix has a small singular value [12]. The numerical stability of implicit filtering presented in this paper and Ruhe's purging is still an open question. Second, the matrix Z is not uniquely defined. We chose Z to be orthogonal, but other possibilities could be tried.

8 Acknowledgements

The authors are grateful to Gerard Sleijpen who provided us with useful suggestions that improved the notations in this paper. We also thank the referees for the many useful suggestions.

The work of Gorik De Samblanx and Adhemar Bultheel was supported by the National Fund for Scientific Research (NFWO), project Lanczos, grant #2.0042.93 and by the Human Capital and Mobility project ROLLS of the European Community under contract ERBCHRXCT930416. The research by Karl Meerbergen was supported by the Belgian programme on Interuniversity Poles of Attraction (IUAP 17), initiated by the Belgian State — Prime Minister's Service — Federal Office for Scientific, Technical and Cultural Affairs and the project Iterative Methods in Scientific Computing, contract number HCM network CHRC-CT93-0420, coordinated by CERFACS, Toulouse, France.

References

[1] J.W. Daniel, W.B. Gragg, L. Kaufman, and G.W. Stewart. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Math. Comp.*, 30:772–795, 1976.

[2] T. Ericsson. A generalised eigenvalue problem and the Lanczos algorithm. In J. Cullum and R.A. Willoughby, editors, *Large Scale Eigenvalue Problems*, pages 95–119. Elsevier Science Publishers BV, 1986.

[3] T. Ericsson and A. Ruhe. The spectral transformation Lanczos method for the numerical solution of large sparse generalised symmetric eigenvalue problems. *Math. Comp.*, 35:1251–1268, 1980.

[4] G. Golub and C. Van Loan. *Matrix computations.* The Johns Hopkins University Press, 2nd edition, 1989.

[5] R.G. Grimes, J.G. Lewis, and H.D. Simon. Eigenvalue problems and algorithms in structural engineering. In J. Cullum and R.A. Willoughby, editors, *Large Scale eigenvalue problems*, pages 81–93. Elsevier Science Publishers B.V., 1986.

[6] R.G. Grimes, J.G. Lewis, and H.D. Simon. A shifted block Lanczos algorithm for solving sparse symmetric generalised eigenproblems. *SIAM J. Matrix Anal. Applic.*, 15:228–272, 1994.

[7] B.D. Hassard, N. Kazarinoff, and Y.H. Wan. *Theory and Applications of Hopf Bifurcation.* Cambridge University Press, Cambridge, UK, 1981.

[8] R.B. Lehoucq. Restarting an Arnoldi reduction. Preprint MCS-P591-0496, Argonne National Laboratory, Argonne, IL 60439, 1996. Submitted to SISC.

[9] R.B. Lehoucq and D.C. Sorensen. Deflation techniques within an implicitly restarted Arnoldi iteration. *SIAM J. Matrix Anal. Applic.*, 1996. To Appear.

[10] R.B. Lehoucq and K. Meerbergen. The inexact rational Krylov sequence method. Preprint 978, Utrecht University, Department of Mathematics, Utrecht, The Netherlands, 1996.

[11] K. Meerbergen and D. Roose. Matrix transformations for computing rightmost eigenvalues of real nonsymmetric matrices. *IMA J. Numer. Anal.*, 16:297–346, 1996.

[12] K. Meerbergen and A. Spence. Implicitly restarted Arnoldi and purification for the shift-invert transformation. *Math. Comp.*, 1997. To appear.

[13] R.B. Morgan. On restarting the Arnoldi method for large nonsymmetric eigenvalue problems. *Math. Comp.*, 1996. To appear.

[14] R. Natarajan. An Arnoldi-based iterative scheme for nonsymmetric matrix pencils arising in finite element stability problems. *J. Comput. Phys.*, 100:128–142, 1992.

[15] B. Nour-Omid, B.N. Parlett, T. Ericsson, and P.S. Jensen. How to implement the spectral transformation. *Math. Comp.*, 48:663–673, 1987.

[16] B.N. Parlett and J. Le. Forward instability of tridiagonal QR. *SIAM J. Matrix Anal. Applic.*, 14:279–316, 1993. 316.

[17] B.N. Parlett and Y. Saad. Complex shift and invert strategies for real matrices. *Linear Alg. Appl.*, 88/89:575–595, 1987.

[18] B. Philippe and M. Sadkane. Improving the spectral transformation block Arnoldi method. In P.S. Vassilevski and S.D. Margenov, editors, *Second IMACS Symposium on Iterative Methods in Linear Algebra*, volume 3 of *IMACS Series in Computational and Applied Mathematics*, pages 57–63. IMACS Symposium on Iterative Methods in Linear Algebra, 1996.

[19] A. Ruhe. Rational Krylov sequence methods for eigenvalue computation. *Linear Alg. Appl.*, 58:391–405, 1984.

[20] A. Ruhe. Rational Krylov algorithms for nonsymmetric eigenvalue problems, II: Matrix pairs. *Linear Alg. Appl.*, 197/198:283–296, 1994.

[21] A. Ruhe. The Rational Krylov algorithm for nonsymmetric eigenvalue problems, III: Complex shifts for real matrices. *BIT*, 34:165–176, 1994.

[22] A. Ruhe. Rational Krylov, a practical algorithm for large sparse nonsymmetric matrix pencils. Technical Report UCB/CSD-95-871, Computer Science Division, University of California, Berkeley, CA, USA, 1995.

[23] Y. Saad. *Numerical methods for large eigenvalue problems.* Algorithms and Architectures for Advanced Scientific Computing. Manchester University Press, Manchester, UK, 1992.

[24] G.L.G. Sleijpen and H.A. van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Applic.*, 17:401–425, 1996.

[25] D.C. Sorensen. Implicit application of polynomial filters in a k-step Arnoldi method. *SIAM J. Matrix Anal. Applic.*, 13:357–385, 1992.

[26] D.S. Watkins. Forward stability and transmission of shifts in the QR algorithm. *SIAM J. Matrix Anal. Applic.*, 16:469–487, 1995.