Shift-invert and Cayley transforms for the detection of eigenvalues with largest real part of nonsymmetric matrices

Karl Meerbergen[*]
Alastair Spence[†]
Dirk Roose[‡]
February 1994

Abstract

This manuscript is concerned with the determination of the right-most eigenvalues of large sparse real nonsymmetric matrices. Specifically, the use of subspace iteration preconditioned by the Cayley transform and/or shift-invert is discussed. The convergence properties of subspace iteration are used to construct a strategy to validate the right-most eigenvalue, which is computed by an iterative method. The motivation behind this paper is that rational preconditioners are very reliable in general but they can miss right-most eigenvalues with large imaginary part. Numerical examples are given to illustrate the theory.

[*] Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200A, 3001 Heverlee, Belgium, Karl.Meerbergen@cs.kuleuven.ac.be

[†] University of Bath, School of Mathematical Sciences, Claverton Down, Bath, BA2 7AY, United Kingdom, A.Spence@maths.bath.ac.uk

[‡] Katholieke Universiteit Leuven, Department of Computer Science, Dirk.Roose@cs.kuleuven.ac.be

1. Introduction

Consider the eigenvalue problem

$$Av - \mu v = 0 \tag{1.1}$$

where A is a large sparse nonsymmetric real $N \times N$-matrix with the eigenvalues $\mu_1, \mu_2, \ldots, \mu_N$ ordered by decreasing real part i.e. $i > j \Rightarrow \mathrm{Re}(\mu_i) \le \mathrm{Re}(\mu_j)$. The problem discussed in this paper is that of finding the right-most eigenvalue(s) of A. The motivation of this work lies in the determination of the linearised stability of a nonlinear system of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, \lambda) \quad f : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N\,,\ x \in \mathbb{R}^N\,,\ \lambda \in \mathbb{R} \tag{1.2}$$

with large N. Typically f arises from the spatial discretisation of a partial differential equation. A steady state solution (x^*, λ^*) is stable if the eigenvalues of the Jacobian matrix $A = f_x(x^*, \lambda^*)$ have negative real parts. The stability of a steady state may change as the parameter λ varies at a bifurcation point. A strategy for the detection of bifurcation points of (1.2) at both stable and unstable steady states is to monitor a few right-most eigenvalues.

The simplest approach to the problem would be to apply Arnoldi's method or subspace iteration directly to A, but this can only be regarded as a preliminary attack. Recent approaches have employed a 'preconditioning' of A with the aim of transforming the dangerous eigenvalues of A into dominant eigenvalues of a preconditioned A, which can then be found by Arnoldi's method or subspace iteration. This idea was used by Saad [17, 18] and Duff and Scott [6] who applied a Chebyshev polynomial preconditioner. Sorensen [21] developed a restarted Arnoldi method with implicit application of polynomial filters. Christodoulou and Scriven [3], Garratt [7], Garratt, Moore and Spence [9, 8] applied a rational transformation known as the Cayley transform and Ruhe [16] developed rational Krylov methods with the aim of using several invert shifts in one rational transformation rather than a single one in Arnoldi's method with shift-invert. When (1.2) arises from (a system of) partial differential equations, one would usually expect the matrix A to have eigenvalues with large negative real parts, as might arise from a discretised Laplacian operator. Shifted and scaled Chebyshev polynomials behave poorly in this case, because a polynomial cannot map the large number of unwanted eigenvalues very close to zero and at the same time map the rightmost eigenvalues very far from the origin. This results in slow convergence, even with high degree polynomials. Rational preconditioners like the Cayley transform and shift-invert only use information about rightmost eigenvalues. Experience indicates that rational preconditioners have faster convergence behaviour and need less work than polynomials [9, 7]. However, as described in Section 2, these rational preconditioners might fail to detect rightmost eigenvalues with large imaginary part.

The aim of this paper is to present a strategy for validating an eigenvalue $\hat{\mu}_1$, i.e. checking that $\hat{\mu}_1$, say found using shift-invert or some other iterative method, is indeed the right-most eigenvalue. The validation has two steps and the philosophy is one of robustness rather than speed, though a user has the option of only performing the first step.

The plan of the paper is as follows. In Section 2 we discuss the application of subspace iteration to matrices preconditioned by shift-invert and the Cayley transform. We introduce the concept of a 'domain of confidence' which is defined as a region in the complex plane, in which all eigenvalues are expected to be computed by the preconditioned method. Section 3 describes our 'two step' technique to validate a candidate rightmost eigenvalue and Section 4 gives some numerical examples. Section 5 summarises the paper and presents some conclusions.

2. The subspace iteration method

In this section we give a brief overview of the subspace iteration or simultaneous iteration algorithm and its convergence properties. We then discuss the algorithm and its convergence properties applied to shift-invert and Cayley transform preconditioners.

Subspace iteration is the block power method improved by a Galerkin projection step [1, 22, 19], as shown in Algorithm 2.1. The matrix $A \in \mathbb{R}^{N\times N}$ has eigenvalues $\mu_1, \mu_2, \dots \mu_N$ and the initial set of vectors $Q_m^{(0)} = [q_1^{(0)} \dots q_m^{(0)}] \in \mathbb{R}^{N\times m}$ is orthogonal.

ALGORITHM 2.1. Subspace iteration algorithm

procedure $\{(\hat\mu_i, \hat u_i), i = 1, \dots, m\}$ = subspace_iteration(A, k, $Q_m^{(0)}$)
 for $l = 1$ **to** k **do**
  Perform one block power step : $V_m^{(l)} = AQ_m^{(l-1)}$.
  Orthonormalise by a QR factorisation : $Q_m^{(l)}R_m^{(l)} = V_m^{(l)}$, ${Q_m^{(l)}}^H Q_m^{(l)} = I$.
 end for
 Solve the projected residual ${Q_m^{(k)}}^H AQ_m^{(k)} z = \hat\mu z$ by the QR-method.
 Set $\hat u_i = Q_m^{(k)} z_i, i = 1, \dots, m$, the approximate eigenvectors.
end procedure

During the iteration process, the subspace span$\{q_1^{(l)}, q_2^{(l)}, \dots, q_m^{(l)}\}$ tends to the subspace spanned by the eigenvectors corresponding the eigenvalues of largest moduli, i.e. the dominant subspace of the matrix A, if the initial $Q_m^{(0)}$, which is chosen randomly, is not deficient in the dominant subspace. The computed eigenpairs $(\hat\mu_i^{(k)}, \hat u_i), i = 1 \dots m$ are approximate dominant eigenpairs of A for which the following convergence theorem holds [22, 1, 2].

THEOREM 2.1. *Suppose the N eigenvalues of A are simple and ordered by decreasing modulus as follows : $|\mu_1| \geq |\mu_2| \geq \dots \geq |\mu_m| > |\mu_{m+1}| \geq \dots \geq |\mu_N|$. Under mild assumptions on the initial subspace $Q_m^{(0)}$ and if k is large enough, then there exists $\hat\mu_i^{(k)}$ for $i = 1, \dots, m$ such that*

$$|\hat\mu_i^{(k)} - \mu_i| = O\left(\left|\frac{\mu_i}{\mu_{m+1}}\right|^{-k}\right).$$

Roughly speaking, the convergence of approximations to the m dominant eigenvalues of A is determined by the quantities $\rho_i = |\mu_i|/|\mu_{m+1}|, i = 1 \dots, m$, which we call the convergence ratios. Thus, even if only one eigenvalue is wanted, we choose m larger than 1 to attain faster convergence. Since $\rho_1 \geq \rho_2 \geq \dots \geq \rho_m$, it is expected that μ_1 converges first, then μ_2 and so on. This is usually the case when k is large enough. More generally, if an eigenvalue μ_i has been computed to an acceptable accuracy, we may expect that all eigenvalues μ_j with $j < i$ have been computed as well, which is confirmed in practice.

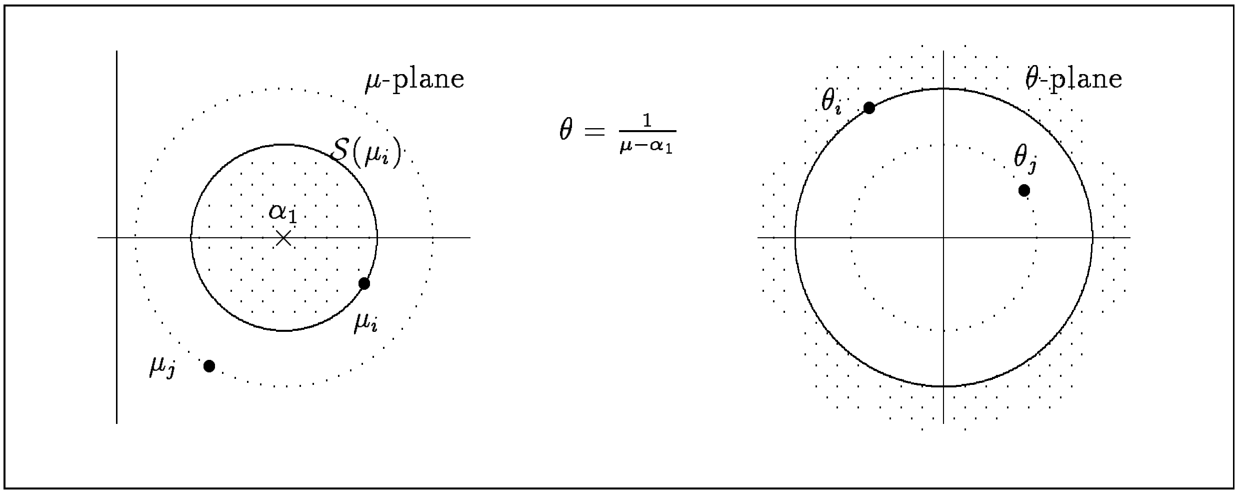
In our application, we apply subspace iteration on a rational transformation $T(A)$ with eigenvalues $\theta_i = T(\mu_i)$, $i = 1, \ldots, N$, that maps the rightmost eigenvalue(s) of A to the dominant eigenvalue(s) of $T(A)$. If the dominant eigenvalues of $T(A)$ are found by subspace iteration, the right-most eigenvalues of A can be recovered. We denote the real part of $\mu \in \mathbb{C}$ by $\mathrm{Re}(\mu)$ and the imaginary part of μ by $\mathrm{Im}(\mu)$.

The first transformation we consider is $T(A) = S(A) = (A - \alpha_1 I)^{-1}$ with $\alpha_1 \in \mathbb{R}$, which is called the shift-invert preconditioner or shift-invert for short. In a practical implementation of subspace iteration, $A - \alpha_1 I$ is factorised and $V_m^{(l)} = S(A)Q_m^{(l-1)}$ is computed by doing m backsolves. Shift-invert maps eigenvalues close to α_1 far from the origin and eigenvalues far from α_1 close to zero, which follows from Lemma 2.2.

LEMMA 2.2. *When $\mu_i \neq \alpha_1$, the circle with centre α_1 and radius $|\mu_i - \alpha_1|$ in the μ-plane is mapped to a circle in the θ-plane with centre 0 and radius $|\theta_i| = 1/|\mu_i - \alpha_1|$.*

This lemma shows that circles are mapped to circles and as can be seen in Figure 2.1, the dotted area $\{\theta : |\theta| > |\theta_i|\}$ in the θ-plane is mapped to a disk $\mathcal{S}(\mu_i)$ centred at α_1 with radius $|\mu_i - \alpha_1|$ in the μ-plane with the following meaning : if an eigenvalue $\mu_i = S^{-1}(\theta_i)$ has converged using subspace iteration on $S(A)$ starting with a random $Q_m^{(0)}$, then all eigenvalues in $\mathcal{S}(\mu_i)$ are expected to have converged as well. We call $\mathcal{S}(\mu_i)$ the *domain of confidence* of shift-invert.

FIG. 2.1. Domain of Confidence of shift-invert



The generalised Cayley transform [8] is defined by

$$T(A) = C(A) = (A - \alpha_1 I)^{-1}(A - \alpha_2 I) = I + (\alpha_1 - \alpha_2)S(A)$$

with α_1 and $\alpha_2 \in \mathbb{R}$ and $\alpha_1 > \alpha_2$. This rational transformation can be evaluated in a similar way as the shift-invert transformation by factorising $(A - \alpha_1 I) = LU$ and doing several backsolves in $V_m^{(l)} = C(A)Q_m^{(l-1)} = Q_m^{(l-1)} - (\alpha_2 - \alpha_1)(LU)^{-1}Q_m^{(l-1)}$. Note that with $\mu_i \neq \alpha_1$ the eigenvalues of $C(A)$ are $(\mu_i - \alpha_2)/(\mu_i - \alpha_1)$ and several important properties follow from this relation (cf. the Möbius transformation in [20]). The Cayley transform maps eigenvalues of A close to α_1 far from the unit circle and eigenvalues close to α_2 to eigenvalues with small modulus. The most interesting property is that eigenvalues to the left side of the centre line of α_1 and α_2, denoted by $\mathcal{L}(\alpha_1, \alpha_2) = \{\mu : \mathrm{Re}(\mu) = \frac{1}{2}(\alpha_1 + \alpha_2)\}$, are mapped inside the unit circle, and eigenvalues to the right side of this line are mapped outside the unit circle. Well chosen parameters α_1 and α_2 map all eigenvalues except the rightmost ones inside the unit circle. The rightmost eigenvalues become dominant eigenvalues of $C(A)$ and are readily found by subspace iteration on $C(A)$. Details of the mapping are given in following lemmas.

LEMMA 2.3. *Suppose $v \neq \alpha_1$ then the line $Re(\mu) = v$ in the μ-plane is mapped to a circle in the θ-plane with centre $1 + \frac{\alpha_1 - \alpha_2}{2(v - \alpha_1)}$, radius $\frac{\alpha_1 - \alpha_2}{2(v - \alpha_1)}$ and going through +1.*

PROOF. See Silverman [20] and Garratt [7]. ∎

LEMMA 2.4. *Suppose $c \in \mathbb{R}^+ \setminus \{1\}$ then the circle with centre*

$$\frac{\alpha_1 c^2 - \alpha_2}{c^2 - 1} \tag{2.1}$$

and radius

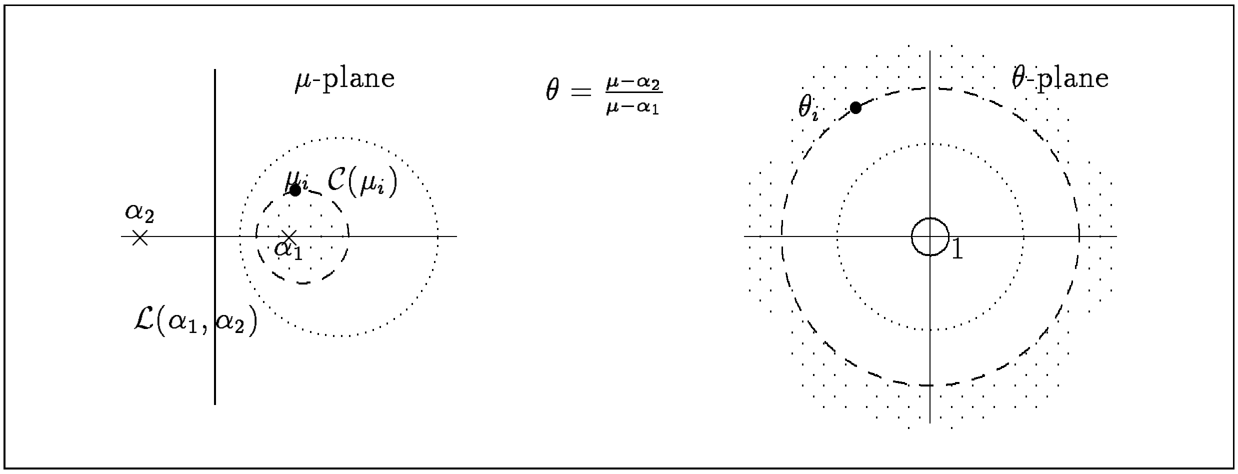
$$\frac{(\alpha_1 - \alpha_2)c}{|c^2 - 1|} \tag{2.2}$$

in the μ-plane is mapped to the circle with centre 0 and radius c in the θ-plane. Moreover $\frac{\alpha_1 c^2 - \alpha_2}{c^2 - 1}$ is larger than α_1 for $c > 1$ and smaller than α_2 for $c < 1$.

PROOF. Let $\mu = \mu_r + i\mu_i$ then $((\mu_r - \alpha_2)^2 + \mu_i^2)/(\mu_r - \alpha_1)^2 + \mu_i^2) = c^2$ is the equation of the circle in the μ-plane and $|\theta| = c$ represents the circle in the θ-plane. ∎

From Lemma 2.4 we can derive that the area $\{\theta : |\theta| > |\theta_i|\}$ in the θ-plane corresponds to a disk with centre $\frac{\alpha_1 |\theta_i|^2 - \alpha_2}{|\theta_i|^2 - 1}$ and radius $\frac{(\alpha_1 - \alpha_2)|\theta_i|}{||\theta_i|^2 - 1|}$ in the μ-plane, called the *domain of confidence* and denoted by $\mathcal{C}(\mu_i)$, see Figure 2.2.

FIG. 2.2. Domain of Confidence of the Cayley transform



If A arises from a discretised PDE, then A often has many eigenvalues with large negative real part. The shift-invert preconditioner $S(A)$ maps eigenvalues far from α_1

close to the origin. Of course this includes both the unwanted eigenvalues with large negative real parts and any eigenvalue with large imaginary part. So it is possible that subspace iteration applied to $S(A)$ will not converge to a rightmost eigenvalue with large imaginary part. A similar conclusion holds for $C(A)$ because of the linear relation between the eigenvalues of $S(A)$ and $C(A)$. Eigenvalues of A with large modulus are mapped very close to 1 by the Cayley transform. Hence if the number of eigenvalues μ_i to the right of $\mathcal{L}(\alpha_1, \alpha_2)$ is smaller than m, the convergence rate ρ_i of such an eigenvalue μ_i can then be approximated by $\rho_i \simeq |C(\mu_i)|/1$. The next section proposes a technique that allows us to find an eigenvalue to the right of a known eigenvalue $\hat{\mu}_1$, provided such an eigenvalue exists.

Note that the use of subspace iteration here is crucial. If Arnoldi's method were used as eigenvalue solver then there is no advantage in distinguishing between $S(A)$ and $C(A)$. This follows from the fact that $C(A) = I + (\alpha_1 - \alpha_2)S(A)$, which is a translated transform $S(A)$, and the fact that Arnoldi's method is translation-invariant, see following Lemma.

LEMMA 2.5. *Suppose the approximate eigenvalues of A computed by Arnoldi's method applied on A, starting with initial vector v_1, are $\hat{\mu}_i^{(1)}, i = 1, \ldots, m$ then the approximate eigenvalues of $(A - \sigma I)$ computed by Arnoldi's method on $(A - \sigma I)$ starting with the same initial vector v_1, are $\hat{\mu}_i^{(2)} = \hat{\mu}_i^{(1)} - \sigma, i = 1, \ldots, m$. The corresponding eigenvectors are the same.*

PROOF. The proof follows from the fact that A and $A - \sigma I$ build the same Krylov subspace (see [14, p.238] and [13, Appendix B]. ■

3. A detection strategy

Suppose s approximate eigenvalues $\hat{\mu}_1, \ldots, \hat{\mu}_s$ with $\mathrm{Re}(\hat{\mu}_1) \geq \cdots \geq \mathrm{Re}(\hat{\mu}_s)$ have been computed by an iterative method. It is realistic to assume that $s > 1$ since most methods compute several approximate eigenvalues at a time. In this section we describe a technique to validate the rightmost eigenvalue $\hat{\mu}_1$ of a real nonsymmetric matrix A. This validation or detection is equivalent to checking if there is an eigenvalue of A in the region $\mathcal{R} = \{\mu : \mathrm{Re}(\mu) \geq \mathrm{Re}(\hat{\mu}_1)\}$. The approach of this paper is an extension and refinement of Algorithm 3.5.1 in [7]. The region $\mathcal{R}$ can perhaps be reduced by use of Gerschgorin's theorem or some other bound on the eigenvalues. In our algorithm we expect $\mathcal{R}$ to be a circle segment and in our examples $\mathcal{R}$ is determined by Gerschgorin's theorem. The validation of the rightmost eigenvalue $\hat{\mu}_1$ is split into two steps.

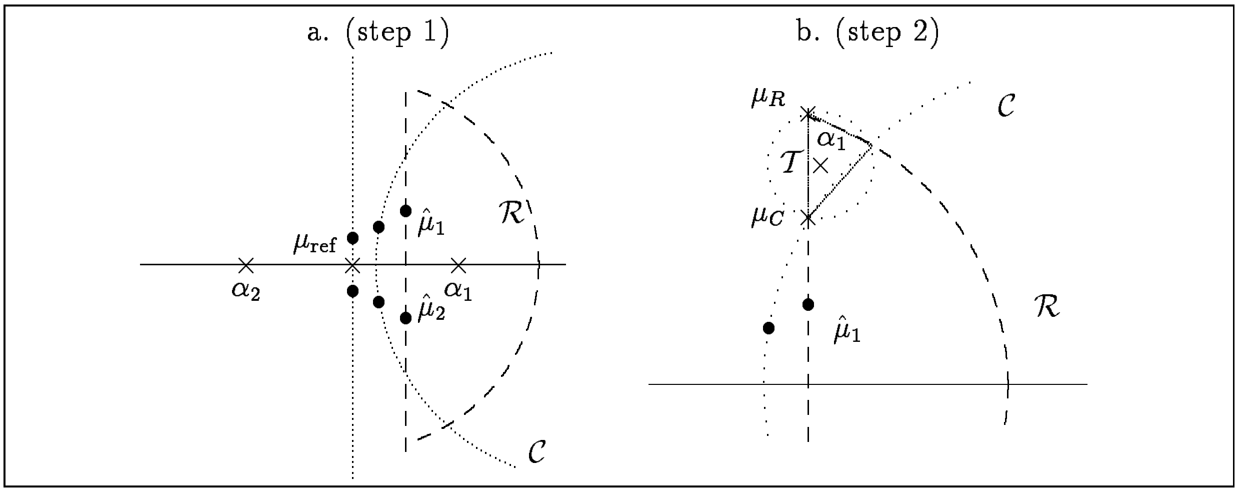
1. Validate $\hat{\mu}_1$ by subspace iteration on $C(A) = (A - \alpha_2 I)(A - \alpha_1 I)^{-1}$. In fact, this step looks for eigenvalues 'close' to the real axis by use of the Cayley transform.
2. Perform a final validation by subspace iteration on $S(A)$ with complex shift α_1 to scan for eigenvalues with large imaginary part.

Each of these steps will be explained in the next subsections.

3.1. Step 1 : Cayley validation

A large part of $\mathcal{R}$ can be scanned by shift-invert or the Cayley transform. There are several reasons to prefer the Cayley transform. We are not interested in a large number

FIG. 3.1. Validation strategy



of eigenvalues to the left of $\hat{\mu}_1$. Shift-invert might find uninteresting eigenvalues along the real axis to the left of $\hat{\mu}_1$ (i.e. outside $\mathcal{R}$) rather than eigenvalues with large imaginary part in $\mathcal{R}$ (if any exist). The Cayley transform, on the other hand, maps uninteresting eigenvalues to the left of the line $\mathcal{L}(\alpha_1, \alpha_2)$ inside the unit circle and does not converge to them easily. The second and more important reason is that the Cayley transform gives rise to a larger domain of confidence for the same α_1, as shown in following lemma.

LEMMA 3.1. *Suppose* $S(A) = (A - \alpha_1 I)^{-1}$ *and* $C(A) = (A - \alpha_2 I)(A - \alpha_1 I)^{-1}$ *with* $\alpha_1 > \alpha_2$ *and* $\frac{1}{2}(\alpha_1 + \alpha_2) < Re(\hat{\mu}_i) \leq Re(\hat{\mu}_1)$. *Denote* $\mathcal{R} = \{\mu : Re(\mu) \geq Re(\hat{\mu}_1)\}$ *then* $\mathcal{S}(\hat{\mu}_i) \bigcap \mathcal{R}$ *is contained in* $\mathcal{C}(\hat{\mu}_i) \bigcap \mathcal{R}$.

PROOF. The centre of $\mathcal{S}(\hat{\mu}_i)$ is α_1. The centre of $\mathcal{C}(\hat{\mu}_i)$ is $\frac{\alpha_1 |C(\hat{\mu}_i)|^2 - \alpha_2}{|C(\hat{\mu}_i)|^2 - 1} > \alpha_1$. ■

The Cayley validation of the candidate rightmost eigenvalue $\hat{\mu}_1$ is outlined in Algorithm 3.1 and Figure 3.1.a. We need to make choices for the two free parameters α_1 and α_2. These are decided as follows :

1. $\alpha_1 + \alpha_2 = 2\,\mu_{\text{ref}}$ where $\mu_{\text{ref}} = \hat{\mu}_{r+1}$. This is in effect a choice of r $(1 \leq r \leq s-1)$. The effect of varying r is described in the examples in Section 4.

2. $|C(\hat{\mu}_1)| = |\hat{\theta}_1| = \rho_{\text{user}}$, a user prescribed parameter. Typical values are $\rho_{\text{user}} = 1.2$ or 1.5.

We now discuss the reasons for these choices. The choice $\mu_{\text{ref}} = \hat{\mu}_{r+1}$ ensures that only $\hat{\mu}_1, \ldots, \hat{\mu}_r$ are to the right of $\mathcal{L}(\alpha_1, \alpha_2)$. The second degree of freedom is determined such that the domain of confidence $\mathcal{C}$ encloses a large part of $\mathcal{R}$. In applications from partial differential equations most eigenvalues of A will lie to the left of μ_{ref}, hence they are mapped inside the unit circle though many of them are mapped close to 1 because they lie far from α_1 and α_2. Eigenvalues close to α_1 are mapped far from the unit circle and hence subspace iteration on $C(A)$ converges to these eigenvalues. Convergence towards at least one eigenvalue is wanted otherwise the domain of confidence cannot be computed, so α_1 should not lie too far from $\hat{\mu}_1$. On the other hand, a rather large domain of confidence $\mathcal{C}$ is wanted to enclose a big part of $\mathcal{R}$. The region $\mathcal{C}$ becomes larger when α_1 shifts to the right, but then, we run the risk that *no* eigenvalues converge. We have found that a good choice for α_1, α_2 is that the convergence rate $\rho(\hat{\mu}_1) \simeq 1.2$, hence the introduction of the user defined parameter ρ_{user} in (2) above. Note that making the choice $|\hat{\theta}_1| = \rho_{\text{user}}$ is the same as asking that $\rho(\hat{\mu}_1) = \rho_{\text{user}}$. With these choices for α_1 and α_2, subspace iteration is performed on an orthonormal random set of vectors until the residuals $\epsilon_i = \|A\hat{u}_i - \hat{\mu}_i\hat{u}_i\|/\|\hat{u}_i\| < \epsilon_{\text{tol}}, i = 1 \dots r$. A typical value of ϵ_{tol} in bifurcation analysis is 10^{-3} since in the determination of stability of steady state solutions of (1.2), we are only interested in the sign of the real part of the eigenvalue rather than accurate estimates. When the Cayley validation finds an eigenvalue to the right of $\hat{\mu}_1$, then the procedure can be restarted with a new set of approximate eigenvalues. If the region $\mathcal{R}$ is contained in $\mathcal{C}$ then no further validation is needed, otherwise we continue with step 2.

ALGORITHM 3.1. Step 1 : Cayley transform validation

```
procedure Cayley_Validation (A, {μ̂_i, i = 1,...,r}, R, r, ρ_user)
    {Step 1, see Figure 3.1.a}
    Choose Q_m^(0) randomly and orthonormalise.
    Determine the line V = {μ|Re(μ) = Re(μ̂_1)}.
    Determine α_1 and α_2 such that ½(α_1 + α_2) = Re(μ_{r+1}) and |C(μ̂_1)| = ρ_user.
    Factorise LU = (A − α_1 I).
    for i = 1 to convergence of r eigenvalues do
        ((θ̂_i, û_i), i = 1,...,m) = subspace_iteration(C(A), k, Q_m^(i−1))
        μ̂_i = C^{−1}(θ̂_i), i = 1,...,m
        orthonormalise [û_1,...,û_m] ⟶ Q_m^(i)
        Order the eigenvalues in decreasing real part.
    end for
    Compute the domain of confidence C of μ̂_1,...,μ̂_r.
    Check if R is contained in C.
end procedure
```

3.2. Step 2 : Complex validation

In general, the region $\mathcal{R} \setminus \mathcal{C}$ is not empty (see Figure 3.1.b) and must be scanned for possible eigenvalues with large imaginary part. It is likely to be inefficient to perform this scan using real choices for α_1 and α_2. In this subsection the idea of a complex shift is exploited. The region $\mathcal{R} \setminus \mathcal{C}$ in Figure 3.1.b is nearly triangular or can be approximated easily by a triangle $\mathcal{T}$ with left vertices μ_R and μ_C, where μ_C is the intersection of $\partial\mathcal{C}$ and the line $\mathcal{V} = \{\mu|\text{Re}(\mu) = \text{Re}(\hat{\mu}_1)\}$ and μ_R is the crossing of the circle $\partial\mathcal{R}$ and the line $\mathcal{V}$. Complex shift-invert is well-suited to checking for eigenvalues in $\mathcal{T}$ (Algorithm 3.2). The following lemma allows α_1 to be determined.

LEMMA 3.2. *Consider a triangle $\mathcal{T}$ with vertices β_1, β_2 and β_3, then the centre α_1 of the circle with smallest radius enclosing $\mathcal{T}$ makes $M(\alpha_1) = \min_{\mu\in\mathcal{T}} |\mu - \alpha_1|^{-1}$ maximal. There are two cases :*

- *The minimum of $M(\alpha_1)$ for $\mu \in \mathcal{T}$ is reached at β_1, β_2 and β_3 so $|\beta_1 - \alpha_1| = |\beta_2 - \alpha_1| = |\beta_3 - \alpha_1|$ and then α_1 is the centre of the circle going through β_1, β_2*

ALGORITHM 3.2. Complex shift-invert Validation

```
procedure Complex_Validation(A, {μ̂_i, i = 1,...,r}, T, n_max)
    {Step 2, see Figure 3.1.b}
    Determine the triangle T and the optimal centre α_1 of T by use of Lemma 3.2.
    Choose Q_m^(0) randomly and orthonormalise.
    Factorise (A − α_1 I) = LU in complex arithmetic.
    for i = 1 to convergence of 1 eigenvalue and i ≤ n_max do
        Do ((θ̂_i, û_i), i = 1,...,m) = subspace_iteration(S(A), k, Q_m^(i−1)).
        μ̂_i = S^{-1}(θ̂_i), i = 1,...,m
        Orthonormalise [û_1,...,û_m] ⟶ Q_m^(i)
        Order the eigenvalues in decreasing real part.
    end for
end procedure
```

and β_3.

- *The minimum of* $M(\alpha_1)$ *for* $\mu \in \mathcal{T}$ *is reached at* β_i *and* β_j *with* $i \neq j$ *and* $i, j \in \{1, 2, 3\}$ *so* $|\beta_i - \alpha_1| = |\beta_j - \alpha_1|$ *and then* $\alpha_1 = \frac{1}{2}(\beta_i + \beta_j)$.

Proof. The proof is given in [13] using the Alternative Theorem [12, 11]. □

Shift-invert will rapidly converge to any eigenvalue in $\mathcal{T}$. The complex arithmetic to compute the transformations and perform subspace iteration can be reduced partly by using techniques from Parlett and Saad [15]. However, these are not considered here because they change the transformations in a rather complicated way or destroy possibly advantageous sparsity in the applications we have in mind.

The iterations are stopped when $\epsilon_i < \epsilon_{\mathrm{tol}}, i = 1 \ldots, r$ but if $\mathcal{T}$ does not contain eigenvalues, subspace iteration will converge very slowly to an eigenvalue outside $\mathcal{T}$, e.g. $\hat{\mu}_1$. So, it is essential to restrict the number of iterations. Several upper bounds on the number of iterations $n_{\max}$ can be suggested. Note that the complex validation can be seen as a final test on the validity of $\hat{\mu}_1$ and that a small number of iterations is enough to know if $\mathcal{T}$ contains eigenvalues or not. A first upper bound, we propose, limits the computation time to p times the work of one LU-factorisation with p defined by the user. Suppose A is an $N \times N$-matrix with half bandwidth b then the work of LU-factorisation with row-pivoting is approximately $N(2b)b$. The work for subspace iteration on $S(A)$ (see Algorithm 2.1) is $mkN3b$ where one $S(A)v$ matrix-vector multiplication costs $3Nb$ work. So the maximum number of iterations we perform is $pN2b^2/(mkN3b) = 2pb/(3mk)$. Similarly the work can be limited to p times the work of the Cayley validation. Another bound — which can be used in tandem with the previous one — is based on the convergence of the subspace iteration method. Suppose, we accept $\hat{\mu}_i, i = 1, \ldots, r$ to be the rightmost eigenvalues to the left of $\hat{\mu}_1$, then the convergence rate of an eigenvalue in $\mathcal{T}$ is probably larger than $\rho_{\mathrm{complex}} = \max_{i=1\ldots r} |S(\mu_R)|/|S(\hat{\mu}_i)|$. Suppose that the error tolerance is ϵ_{tol}, then the maximum number of iterations expected is $n_M = n_{\min} - \log(\epsilon_{\mathrm{tol}})/\log(\rho_{\mathrm{complex}})/k$ with for example $n_{\min} = 3$. If no eigenvalues are computed after this number of iterations, we may conclude that there is probably no eigenvalue in $\mathcal{T}$.

In our examples the number of complex validation iterations is restricted to $n_{\max} = \min(n_C, n_M)$, with n_C the number of Cayley validation iterations in Algorithm 3.1.

4. Examples

EXAMPLE 1. *The first example is the computation of the rightmost eigenvalue of a* 400×400*-matrix with half bandwidth 10 and eigenvalues* $0\pm 52i$ *and* $-k+2, k = 3, \ldots, 400$. *This example is generated artificially but its spectrum is similar in 'shape' to the double diffusive convection problem in Cliffe, Garratt and Spence [4] and Cliffe and Winters [5]. Figure 4.1 shows the rightmost part of the spectrum of* A*, and the boundary of* $\mathcal{R}$*, enclosing all eigenvalues of* A*.*

Suppose $\hat{\mu}_1, \ldots, \hat{\mu}_5 = -1, -2, -3, -4, -5$ are computed by shift-invert with $\alpha_1 = 0$. The validation procedure is run to validate $\hat{\mu}_1$ and the results are presented in Table 4.1. The parameters α_1 and α_2 for the Cayley transform are chosen such that $\frac{1}{2}(\alpha_1 + \alpha_2) = \mu_{\text{ref}} = \mu_{r+1}$ for $r = 1$, 2 and 4 and $|C(\hat{\mu}_1)| = \rho_{\text{user}}$. Remind that r determines the choice of the reference value μ_{ref}. In both the Cayley and the complex shift-invert validation the subspace iteration parameters $m = 10$ and $k = 10$ are used. In Table 4.1, 5 cases are shown. Cases 1–3 compare the results for different r, cases 3–5 compare the influence of different ρ_{user}. The corresponding domains of confidence are denoted by $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$, $\mathcal{C}_4$ and $\mathcal{C}_5$. In the first case the domain of confidence $\mathcal{C}_1$ of the Cayley validation is so small that $\mathcal{R}$ is not reduced very well, see also the dotted circle in Figure 4.1. In this case complex shift-invert validation is not very useful since the domain of confidence $\mathcal{C}_1$ does not enclose the right part of $\mathcal{R}$. One can expect that a large r introduces eigenvalues μ_j to the left of $\hat{\mu}_1$ and that $\mathcal{C}(\mu_j)$ is very large. But in this example, the influence of r is not really clear. Note that a large r is only meaningful when μ_{ref} lies to the left of μ_r, otherwise convergence is very slow. For smaller ρ_{user} a larger domain of confidence is observed but of course, slower convergence to $\hat{\mu}_1$, see cases 3, 4 and 5. In case 5, the domain $\mathcal{C}_5$ encloses $\mathcal{R}$ totally, but the extra computation time to reach this result is large compared to the real and complex execution times for cases 3 and 4. It is very interesting to see that case 4 converges much faster than case 3. The reason is that the convergence rate of μ_1 and μ_2 is larger in case 4. Note that in cases 2 to 5 the number of computed eigenvalues is larger than r. For the complex validation step the maximum number of iterations is set to $n_{\max} = \min(n_C, n_M)$ (see subsection 3.2), where $n_M = n_{\min} - \log(10^{-3})/\log(\rho_{\text{complex}})/k$ with $\rho_{\text{complex}} = |\hat{\mu}_1 - \alpha_1|/|\mu_R - \alpha_1| = 1.87$, $n_{\min} = 3$ and $k = 10$ so $n_M = 4.1$. Note that ρ_{complex} is a guess of the convergence rate. The exact convergence rate is $|\mu_{11} - \alpha_1|/|\mu_R - \alpha_1| = 1.88$.

EXAMPLE 2. *This example is a model of the tubular reactor [10, 7]. The equations modelling the conservation of rectant and energy in a homogeneous tube of length L in dimensionless form are :*

$$\frac{L}{v}\dot{y} + \frac{1}{Pe_m}\frac{\partial^2 y}{\partial s^2} - \frac{\partial y}{\partial s} - Dy\exp(\gamma - \gamma T^{-1}) = 0,$$

$$\frac{L}{v}\dot{T} + \frac{1}{Pe_h}\frac{\partial^2 T}{\partial s^2} - \frac{\partial T}{\partial s} - \beta(T - T_0) + BDy\exp(\gamma - \gamma T^{-1}) = 0$$

with spatial coordinate $s \in [0, 1]$ *where* $y(s)$ *represents the concentration and* $T(s)$ *the temperature. We consider boundary conditions* $y'(0) = Pe_m y(0)$, $T'(0) = Pe_h T(0)$ *and*

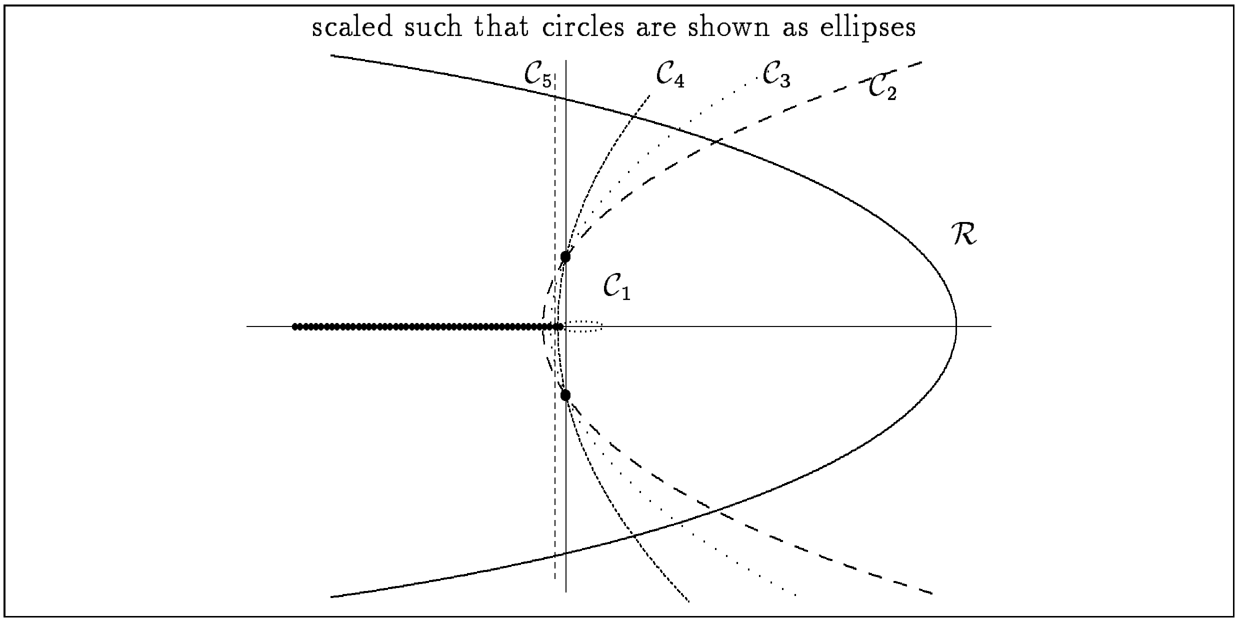
TABLE 4.1. Validation results for example 1

	case 1	case 2	case 3	case 4	case 5
	Validation — step 1 (Cayley)				
r	1	4	2	2	2
μ_{ref}	$\hat{\mu}_2$	$\hat{\mu}_5$	$\hat{\mu}_3$		
ρ_{user}	1.5	1.5	1.5	1.1	1.02
α_1	1	15	7	39	199
α_2	-5	-25	-13	-45	-205
iterations	4	8	16	5	27
time (sec) (DEC5000/240)	9	16	35	11	54
eigenvalues	-1	$0 \pm 52i$ $-1, -2$ $-3, -4$	$0 \pm 52i$ $-1, -2$	$0 \pm 52i$ $-1, -2$	$0 \pm 52i$ $-1, -2$
	Validation — step 2 (complex shift-invert)				
α_1		$0 + 111.5i$	$0 + 111.5i$	$0 + 111.5i$	/
iterations		2	2	2	/
time (sec) (DEC5000/240)		8	8	8	/
eigenvalues		$0 + 52i$	$0 + 52i$	$0 + 52i$	/

$y'(1) = 0$, $T'(1) = 0$. For $x^T = [y_1, T_1, y_2, T_2, \ldots y_n, T_n]^T$ the equations can be written as $\dot{x} = f(x)$. The parameters in the differential equation are set to $Pe_h = Pe_m = 5, B = 0.5, \gamma = 25, \beta = 3.5$. Central differences are used to discretise the steady state solutions. The Jacobian matrix $A = \frac{\partial f}{\partial x}$ has bandwidth 5 and in this example $N = 400$. The Damkohler number D is varied in a continuation process to detect Hopf bifurcation points. For the details, see Garratt, Moore and Spence [9] and Heinemann and Poore [10]. We computed the rightmost eigenvalues $\hat{\mu}_{1,2} = 0.173168 \pm 1.412567$, $\hat{\mu}_{3,4} = -4.019198 \pm 2.531755$ and $\hat{\mu}_{5,6} = -9.896724 \pm 2.511188$ of the discretised Jacobian for $D = 0.2656$ by use of Arnoldi's method on A^{-1}. The rightmost eigenvalues for $D = 0.2662$ are wanted. The region $\mathcal{R} = \{\mu | \text{Re}(\mu) \geq 0.173, |\mu - (-15844)| \leq 15869\}$ is computed by use of Gerschgorin's theorem and is very large.

Table 4.2 contains the results of three runs of the validation algorithm. The real and complex step are both run with $m = 10$ and $k = 10$. At the end of subsection 3.1, we suggested restarting the procedure once an eigenvalue at the right of $\hat{\mu}_1$ is found. In this example, we continue with the complex validation step but with the new $\hat{\mu}_1 = 0.2044 \pm 1.382i$, $\mu_R = 0.204 + 895i$ and $\mu_C = \hat{\mu}_1$ to determine $\mathcal{T}$. The number of complex shift-invert interations is limited by n_C (see subsection 3.2) and by $n_M = n_{\min} - \log(10^{-3})/\log(\rho_{\text{complex}})/k \simeq 11.1$ with $\rho_{\text{complex}} = |\hat{\mu}_1 - \alpha_1|/|\mu_R - \alpha_1| \simeq 1.8$ for both $\alpha_1 = 0.204 + 464i$ and $\alpha_1 = 0.204 + 469i$. The upper bound on the number of iterations is reached without convergence : the rightmost approximate eigenvalue (in brackets) computed by complex shift-invert validation has a residual of order 1 and lies far from α_1. Note that in cases 1 and 2 the rightmost eigenvalues μ_1 and μ_2 converge in the same

FIG. 4.1. Illustration of the validation for example 1



number of iterations. Cases 2 and 3 only differ in the number of converged eigenvalues. It appears that in this example the choice of r does not determine the number of iterations to converge to μ_1.

5. Conclusions

In this paper, a technique is proposed to detect eigenvalues to the right of a vertical line in the complex plane. The concept of domains of confidence is introduced for subspace iteration with shift-invert and Cayley transforms, and proves to be reliable in the examples considered. It is not clear how to use a similar concept for Arnoldi's method. The detection of eigenvalues with large imaginary parts is very efficient with complex shift-invert but a practical problem arises when no eigenvalues with large imaginary part exist. In most examples, however, the real Cayley transform does not miss eigenvalues very often, and complex validation, which is expensive, is often unnecessary. Therefore we suggest three possible strategies, each involving increased computational effort, to validate a computed eigenvalue $\hat{\mu}_1$:

- Step 1 of the validation, namely, a Cayley validation with α_1, α_2 chosen as in Section 3.
- Steps 1 and 2 of the validation, but with a limited number of complex iterations in step 2.
- If an absolute guarantee is required, do not limit the number of steps in step 2.

A user could decide on which strategy to employ based on the particular problem, and the amount of effort that was considered appropriate.

TABLE 4.2. Validation results of example 2

	case 1	case 2	case 3
	Validation — step 1 (Cayley)		
r	2	4	
μ_{ref}	$\hat{\mu}_3$	$\hat{\mu}_5$	
ρ_{user}	1.5	1.5	
α_1	16.8	40.4	
α_2	-24.9	-60.2	
stop if	convergence of 2 eigenvalues	convergence of 2 eigenvalues	convergence of 4 eigenvalues
iterations	5	5	7
time (sec) DEC5000/240	10.6	11	14
eigenvalues	$0.2044 \pm 1.382i$	$0.2044 \pm 1.382i$	$0.2044 \pm 1.382i$ $-4.049 \pm 2.546i$
	Validation — step 2 (complex shift-invert)		
α_1	$0.204 + 464i$	$0.204 + 464i$	$0.204 + 469i$
iterations	5	5	7
time (sec) DEC5000/240	20	20	28
eigenvalues	$[1.2 + 6.3i]$	$[1.2 + 6.3i]$	$[1.7 + 6.2i]$

We stress the importance of a well defined region $\mathcal{R}$ enclosing the eigenvalues of A to the right of the approximate rightmost eigenvalues. Sometimes $\mathcal{R}$ can be determined based on sharp physical or mathematical bounds on the eigenvalues. Variations on the techniques proposed in this paper are possible. E.g. one can use a real shift-invert and a complex Cayley transform and the shifts α_1 and α_2 can be determined in other ways. In the description of the algorithm, we assumed that $\hat{\mu}_1$ is accurate but in fact a rough guess is sufficient since eigenvalues close to $\hat{\mu}_1$ will be found readily (see also Example 2).

6. Acknowledgements

The authors are grateful to the British Council in Brussels and the Belgian National Fund for Scientific Research (NFWO) for supporting the collaboration between the Katholieke Universiteit Leuven and the University of Bath.

REFERENCES

[1] F. Chatelin. *Spectral Approximation of Linear Operators.* Academic Press, New York, 1983.

[2] F. Chatelin. *Eigenvalues of matrices.* Pure and applied mathematics. John Wiley and Sons, 1993.

[3] K.N. Christodoulou and L.E. Scriven. *Finding leading modes of a viscous free surface flow : an asymmetric generalized eigenproblem.* PhD thesis, Center for Interfacial Engineering and Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, 1988.

[4] K.A. Cliffe, T.J. Garratt, and A. Spence. Calculation of eigenvalues of the discretised Navier-Stokes and related equations. In J.R. Whiteman, editor, *The mathematics of finite elements and applications VII MAFELAP*, pages 119–123. Academic Press, 1990.

[5] K.A. Cliffe and K.H. Winters. Convergence properties of the finite element method for Bénard convection in an infinite layer. *J. Comput. Phys.*, 60:346–351, 1985.

[6] I.S. Duff and J.A. Scott. Computing selected eigenvalues of large sparse unsymmetric matrices using subspace iteration. *ACM Trans. Math. Softw.*, 19:137–159, 1993.

[7] T.J. Garratt. *The Numerical Detection of Hopf Bifurcations in Large Systems Arising in Fluid Mechanics.* PhD thesis, University of Bath, UK, 1991.

[8] T.J. Garratt, G.Moore, and A. Spence. A Generalised Cayley transform for the numerical detection of Hopf bifurcations in large systems. Technical report, School of Mathematical Sciences, University of Bath, 1992.

[9] T.J. Garratt, G. Moore, and A. Spence. Two methods for the numerical detection of Hopf bifurcations. In R. Seydel, F.W. Schneider, and H. Troger, editors, *Bifurcation and Chaos : Analysis, Algorithms, Applications*, pages 119–123. Birkhäuser, 1991.

[10] R.F. Heinemann and A.B. Poore. Multiplicity, stability, and oscillatory dynamics of a tubular reactor. *Chem. Eng. Sci.*, 36:1411–1419, 1981.

[11] T. Manteuffel. An iterative method for solving nonsymmetric linear systems with dynamic estimation of parameters. Technical Report UIUCDCS-R-75-758, Department of Computer Science, University of Illinois, Urbana-Champaign, Urbana, Illinois, 1975.

[12] T. Manteuffel. The Tchebyshev Iteration for Nonsymmetric Linear Systems. *Numer. Math.*, 28:307–327, 1977.

[13] K. Meerbergen, A. Spence, and D. Roose. Shift-invert and Cayley transforms for the detection of eigenvalues wi th largest real part of nonsymmetric matrices. Report TW200, Department of Computing Science, KU Leuven, 1993.

[14] B.N. Parlett. *The symmetric eigenvalue problem.* Prentice Hall series in computational mathematics. Prentice-Hall Englewood Cliffs (N.J.), 1980.

[15] B.N. Parlett and Y. Saad. Complex Shift and Invert Strategies for Real Matrices. Technical Report YALEU/DCS/RR-424, University of California at Berkeley, Mathematics Department, 1985.

[16] A. Ruhe. Rational krylov algorithms for nonsymmetric eigenvalue problems. In G. Golub, M. Luskin, and A. Greenbaum, editors, *IMA volumes in Mathematics and its Applications*, volume 60, pages 149–164. Springer, 1993.

[17] Y. Saad. Chebyshev Acceleration Techniques for Solving Nonsymmetric Eigenvalue Problems. *Math. Comp.*, 42(166):567–588, 1984.

[18] Y. Saad. Numerical solution of large nonsymmetric eigenvalue problems. *Computer Physics Communications*, 53:71–90, 1989.

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

[20] R.A. Silverman. *Introductory complex analysis.* Prentice-Hall, Englewood Cliffs, 1967.

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

[22] G.W. Stewart. Simultaneous Iteration for Computing Invariant Subspaces of Non-Hermitian Matrices. *Numer. Math.*, 25:123–136, 1976.