Linear Algebra and its Applications 436 (2012) 3839–3863

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

An integral method for solving nonlinear eigenvalue problems

Wolf-Jürgen Beyn [1]

Department of Mathematics, Bielefeld University, P.O. Box 100131, D-33501 Bielefeld, Germany

ARTICLE INFO

Article history:
Received 27 October 2010
Accepted 15 March 2011
Available online 20 April 2011

Submitted by V. Mehrmann

Dedicated to Heinrich Voss on the occasion of his 65th birthday

AMS classification:
65F15
47J10

Keywords:
Nonlinear eigenvalue problems
Numerical methods
Contour integrals

ABSTRACT

We propose a numerical method for computing all eigenvalues (and the corresponding eigenvectors) of a nonlinear holomorphic eigenvalue problem that lie within a given contour in the complex plane. The method uses complex integrals of the resolvent operator, applied to at least k column vectors, where k is the number of eigenvalues inside the contour. The theorem of Keldysh is employed to show that the original nonlinear eigenvalue problem reduces to a linear eigenvalue problem of dimension k. No initial approximations of eigenvalues and eigenvectors are needed. The method is particularly suitable for moderately large eigenvalue problems where k is much smaller than the matrix dimension. We also give an extension of the method to the case where k is larger than the matrix dimension. The quadrature errors caused by the trapezoid sum are discussed for the case of analytic closed contours. Using well known techniques it is shown that the error decays exponentially with an exponent given by the product of the number of quadrature points and the minimal distance of the eigenvalues to the contour.

1. Introduction

We consider nonlinear eigenvalue problems of the form

$$T(z)v = 0, \quad v \in \mathbb{C}^m,\ v \neq 0,\ z \in \Omega, \tag{1}$$

where $T : \Omega \to \mathbb{C}^{m,m}$ is assumed to be holomorphic in some domain $\Omega \subset \mathbb{C}$. The computation of all eigenvalues and eigenvectors inside Ω usually requires the solution of two problems (see [24,4] for recent reviews):

E-mail address: beyn@math.uni-bielefeld.de

[1] Supported by CRC 701 'Spectral Structures and Topological Methods in Mathematics'.

0024-3795/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.laa.2011.03.030

1. Approximate localization and separation of eigenvalues in suitable domains.
2. Accurate computation of eigenvalues and associated eigenvectors by an iterative method.

The global problem of localization can be substantially simplified if minimum–maximum characterizations similar to the linear symmetric case hold [36,33]. Voss and co-workers have combined these principles with locally convergent methods of Arnoldi or Jacobi–Davidson type (see [34,6,35]), and in this way provided an effective means for computing all eigenvalues.

Another case where both problems can be solved is for polynomials

$$T(z) = \sum_{j=0}^{p} T_j (z - z_0)^j, \quad T_j \in \mathbb{C}^{m,m}.$$

This eigenvalue problem can be reduced to a linear eigenvalue problem of dimension pm, and this is the path taken by the MATLAB routine *polyeig*. Quite a few papers in the literature pursue and analyze this linearization approach, see [21] for a survey. Another approach is to generalize methods for linear eigenvalue problems directly to polynomial problems by using its internal structure. An example of this type is the SOAR method (Second Order ARnoldi) for quadratic eigenvalue problems [3,23].

In the general holomorphic case we just have a power series near each $z_0 \in \Omega$,

$$T(z) = \sum_{j=0}^{\infty} T_j (z - z_0)^j, \quad |z - z_0| \text{ small}, \ T_j \in \mathbb{C}^{m,m}.$$

One may then use polynomial truncation and a polynomial eigenvalue solver for getting good initial estimates of the eigenvalues (see e.g. [19]). However, the success of this method strongly depends on the radius of convergence and on the decay of the coefficient matrices. Also, it may be necessary to compute power series at many different points in Ω.

Finally, we refer to the recent approach of Kressner [20], who uses the fact that any holomorphic matrix function can be written as

$$T(z) = \sum_{j=1}^{p} f_j(z) T_j, \quad T_j \in \mathbb{C}^{m,m}$$

with holomorphic functions $f_j : \Omega \mapsto \mathbb{C}$ (such a representation always exists for some $p \leqslant m^2$). Then a Newton-type iteration is devised in [20] that allows one to compute a group of eigenvalues and an associated subspace. By construction the method has local convergence properties.

In this paper we tackle the global problem by using contour integrals, which seem to be the only available tool in the general holomorphic case. The idea is to use the theorem of Keldysh [17,18], which provides an expansion of $T(z)^{-1}$ in a neighborhood $\mathcal{U} \subset \Omega$ of an eigenvalue $\lambda \in \Omega$ as follows:

$$T(z)^{-1} = \sum_{j=-\kappa}^{\infty} S_j (z - \lambda)^j, \quad z \in \mathcal{U} \setminus \{\lambda\}, \ S_j \in \mathbb{C}^{m,m}, \ S_{-\kappa} \neq 0. \tag{2}$$

More specifically, Keldysh' theorem gives a representation of the singular part in (2) in terms of generalized eigenvectors of $T(z)$ and its Hermitian transpose $T^H(z)$. A good reference for the underlying theory is [26], which we briefly review in Section 2.

Numerical methods based on contour integrals seem not to have attracted much attention in the past. Notable exceptions are exponential integrators and, more recently, approaches for computing analytic functions of matrices via suitably transformed contour integrals ([15]; [16, 13.3.2]). An application of computing spectral projectors via contour integrals appears in [28].

Our goal is to compute all eigenvalues and the associated eigenvectors that lie within a given closed contour Γ in Ω. The main algorithm is described in Section 3. Suppose that $k \leqslant m$ eigenvalues of (1) lie inside Γ. Then our method reduces the nonlinear eigenvalue problem to a linear one of dimension

k by evaluating the contour integrals

$$A_p = \frac{1}{2\pi i}\int_\Gamma z^p T(z)^{-1}\hat{V}\,dz, \quad p = 0, 1. \tag{3}$$

Here $\hat{V} \in \mathbb{C}^{m,k}$ is generally taken as a random matrix. The contour integrals in (3) are calculated approximately by the trapezoid sum. If N quadrature points are used, this requires one to compute N LU-decompositions and to solve Nk linear systems, which is the main numerical effort. As a consequence, our method is limited to moderately large nonlinear eigenvalue problems for which a fast (sparse) direct solver is available.

In Section 4 we apply the algorithm to several examples, showing that a moderate number of quadrature nodes ($N \approx 25$) is usually sufficient to get good estimates of eigenvalues and eigenvectors. Based on [9], we prove in Section 4 that the quadrature error decays exponentially with an exponent that depends on the product of the number of quadrature nodes and the smallest distance of the eigenvalues to the contour.

In Section 5 we deal with two problems that are typical for nonlinear eigenvalue problems and that do not occur in the linear case: First, there can be many more eigenvalues than the matrix dimension (e.g. characteristic functions for delay equations) and, second, eigenvectors belonging to different eigenvalues can be linearly dependent, even if the number of eigenvalues is less than the matrix dimension. In Section 5 we extend our integral method such that it applies to the case $k > m$ and that it can also handle rank defects of eigenspaces. For the extended integral method it is necessary to evaluate A_p from (3) for indices $0 \leqslant p \leqslant 2\left\lceil \frac{k}{m} \right\rceil - 1$. Numerical examples show that this extension is suitable for solving both aforementioned problems.

2. Nonlinear eigenvalues and Keldysh' Theorem

The material in this section is largely based on the monograph [26]. It contains a general study of meromorphic operator functions that have values in spaces of Fredholm operators of index 0. For our purposes it is sufficient to consider matrix valued mappings

$$T : \Omega \subset \mathbb{C} \to \mathbb{C}^{m,m}$$

that are holomorphic in some open domain Ω. We write this as $T \in H(\Omega, \mathbb{C}^{m,m})$. For a matrix A we denote by $R(A)$ and $N(A)$ its range and nullspace, respectively.

Definition 2.1. A number $\lambda \in \Omega$ is called an *eigenvalue* of $T(\cdot)$ if $T(\lambda)v = 0$ for some $v \in \mathbb{C}^m,\ v \neq 0$. The vector v is then called a (*right*) *eigenvector*. By $\sigma(T)$ we denote the set of all eigenvalues and by $\rho(T) = \Omega \setminus \sigma(T)$ we denote the *resolvent set*.

The eigenvalue λ is called *simple* if

$$N(T(\lambda)) = \operatorname{span}\{v\},\ v \neq 0, \quad T'(\lambda)v \notin R(T(\lambda)).$$

Throughout the paper we assume that the resolvent set is nonempty, i.e. $\det(T(z))$ does not vanish identically.

Theorem 2.2. *Every eigenvalue $\lambda \in \sigma(T)$ of $T \in H(\Omega, \mathbb{C}^{m,m})$ is isolated, i.e. $\mathcal{U} \setminus \{\lambda\} \subset \rho(T)$ for some neigborhood $\mathcal{U}$ of λ.*

Moreover, $T(z)^{-1}$ is meromorphic at λ, i.e. there exist $\kappa \in \mathbb{N}$ and $S_j \in \mathbb{C}^{m,m}$ for $j \geqslant -\kappa$ such that $S_{-\kappa} \neq 0$ and

$$T(z)^{-1} = \sum_{j=-\kappa}^{\infty} S_j (z-\lambda)^j, \quad z \in \mathcal{U} \setminus \{\lambda\}. \tag{4}$$

Remark 2.3. The number κ is uniquely determined and called the order of the pole at λ.

The Theorem of Keldysh (see Theorem 2.6) gives a representation of the singular part

$$\sum_{j=-\kappa}^{-1} S_j(z-\lambda)^j$$

in terms of (generalized) eigenvectors of T and T^H. It goes back to Keldysh [17] with a proof given in [18]. Generalizations of Keldysh' theorem were derived by Trofimov [32], who introduced the concept of root polynomials, and by Markus and Sigal [22] and Gohberg and Sigal [12], who used factorizations of operator functions. A simple direct proof was found by Mennicken and Möller [25], who later gave a concise approach to the whole theory in [26].

For the motivation of the algorithm in the next section it is instructive to first state Keldysh' theorem for simple eigenvalues. In this case Definition 2.1 implies for the adjoint $T^H(z)$

$$N(T^H(\lambda)) = \operatorname{span}\{w\} \text{ for some } w \in \mathbb{C}^m, \quad w \neq 0,$$
$$w^H T'(\lambda) v \neq 0.$$

Without loss of generality we can normalize v and w such that

$$w^H T'(\lambda) v = 1. \tag{5}$$

Then we are still free to further normalize either $|w| = 1$ or $|v| = 1$.

Theorem 2.4. *Assume $\lambda \in \Omega$ is a simple eigenvalue of $T \in H(\Omega, \mathbb{C}^{m,m})$ with eigenvectors normalized as in (5). Then there is a neighborhood $\mathcal{U} \subset \Omega$ of λ and a holomorphic function $R \in H(\mathcal{U}, \mathbb{C}^{m,m})$ such that*

$$T(z)^{-1} = \frac{1}{z-\lambda} v w^H + R(z), \quad z \in \mathcal{U} \setminus \{\lambda\}. \tag{6}$$

Moreover, let $\mathcal{C} \subset \Omega$ be a compact subset that contains only simple eigenvalues λ_n, $n = 1, \ldots, k$ with eigenvectors v_n, w_n satisfying

$$T(\lambda_n) v_n = 0, \quad w_n^H T(\lambda_n) = 0, \quad w_n^H T'(\lambda_n) v_n = 1. \tag{7}$$

Then there is a neighborhood $\mathcal{U}$ of $\mathcal{C}$ in Ω and a holomorphic function $R \in H(\mathcal{U}, \mathbb{C}^{m,m})$ such that

$$T(z)^{-1} = \sum_{n=1}^{k} \frac{1}{z-\lambda_n} v_n w_n^H + R(z), \quad z \in \mathcal{U} \setminus \{\lambda_1, \ldots, \lambda_k\}. \tag{8}$$

Proof. The first part is a special case of Theorem 2.6. For the second part, note that eigenvalues are isolated and hence we can choose a neighborhood $\mathcal{C} \subset \mathcal{U} \subset \Omega$ such that $\sigma(T) \cap \mathcal{U} = \{\lambda_1, \ldots, \lambda_k\}$. Then the function

$$R(z) = T(z)^{-1} - \sum_{n=1}^{k} \frac{1}{z-\lambda_n} v_n w_n^H$$

is holomorphic in $\mathcal{U} \cap \rho(T)$ and by the first part it is also holomorphic in suitable neighborhoods of λ_n, $n = 1, \ldots, k$. □

Definition 2.5. Let $T \in H(\Omega, \mathbb{C}^{m,m})$ and $\lambda \in \Omega$.

(i) A function $v \in H(\Omega, \mathbb{C}^m)$ is called a *root function of T at λ* if

$$v(\lambda) \neq 0, \quad T(\lambda) v(\lambda) = 0.$$

The order of the zero $z = \lambda$ of $T(z)v(z)$ is called the *multiplicity of v at λ* and denoted by $s(v)$.

(ii) A tuple $(v_0, \dots, v_{n-1}) \in (\mathbb{C}^m)^n$, $n \geqslant 1$ is called a *chain of generalized eigenvectors (CGE) of T at λ* if $v(z) = \sum_{j=0}^{n-1} (z-\lambda)^j v_j$ is a root function of T at λ of multiplicity $s(v) \geqslant n$.

(iii) For a given $v_0 \in N(T(\lambda))$, $v_0 \neq 0$ the number

$$r(v_0) = \max\{s(v) : v \text{ is a root function of } T \text{ at } \lambda \text{ with } v(\lambda) = v_0\}$$

is finite and called the *rank of v_0*.

(iv) A system of vectors in $\mathbb{C}^m$

$$V = \left(v_j^\ell, 0 \leqslant j \leqslant m_\ell - 1, 1 \leqslant \ell \leqslant L\right)$$

is called a *canonical system of generalized eigenvectors (CSGE) of T at λ* if the following conditions hold:

(a) The vectors $v_0^1, \dots, v_0^L$ form a basis of $N(T(\lambda))$.

(b) The tuple $(v_0^\ell, \dots, v_{m_\ell - 1}^\ell)$ is a CGE of T at λ for $\ell = 1, \dots, L$.

(c) $m_\ell = \max\{r(v_0) : v_0 \in N(T(\lambda)) \setminus \operatorname{span}\{v_0^\nu : 0 \leqslant \nu < \ell\}\}$ for $\ell = 1, \dots, L$.

One can show that a CSGE always exists and that the numbers m_ℓ are ordered according to

$$m_1 \geqslant m_2 \geqslant \cdots \geqslant m_L. \tag{9}$$

They are called the *partial multiplicities of T at λ*. As an example take $v(z) = v_0 + zv_1 + \frac{1}{2}z^2 v_2$ and note that this is a root function of order 3 at $\lambda = 0$ if the equalities

$$T(0)v_0 = T'(0)v_0 + T(0)v_1 = T''(0)v_0 + 2T'(0)v_1 + T(0)v_2 = 0$$

and the inequality $T'''(0)v_0 + 3T''(0)v_1 + 3T'(0)v_2 \neq 0$ hold.

With the notions from the above definition we can state the following general theorem, see [26, Theorem 1.6.5].

Theorem 2.6 (Keldysh). *Let $T \in H(\Omega, \mathbb{C}^{m,m})$ be given with $\rho(T) \neq \emptyset$. For $\lambda \in \sigma(T)$ let*

$$V = \left(v_j^\ell, 0 \leqslant j \leqslant m_\ell - 1, 1 \leqslant \ell \leqslant L\right)$$

be a CSGE of T at λ. Then there exists a CSGE

$$W = \left(w_j^\ell, 0 \leqslant j \leqslant m_\ell - 1, 1 \leqslant \ell \leqslant L\right)$$

of T^H at λ, a neighborhood $\mathcal{U}$ of λ and a function $R \in H(\mathcal{U}, \mathbb{C}^{m,m})$ such that

$$T(z)^{-1} = \sum_{\ell=1}^{L} \sum_{j=1}^{m_\ell} (z-\lambda)^{-j} \sum_{\nu=0}^{m_\ell - j} v_\nu^\ell w_{m_\ell - j - \nu}^{\ell H} + R(z), \quad z \in \mathcal{U} \setminus \{\lambda\}. \tag{10}$$

The system W for which (10) holds is the unique CSGE of T^H at λ that satisfies the following conditions

$$r(w_0^\ell) = m_\ell,$$

$$\sum_{\alpha=0}^{j} \sum_{\beta=1}^{m_\nu} w_{j-\alpha}^{\ell H} T_{\alpha+\beta} v_{m_\nu - \beta}^\nu = \delta_{\nu\ell}\delta_{0j}, \quad 0 \leqslant j \leqslant m_\ell - 1,\ 1 \leqslant \ell, \nu \leqslant L, \tag{11}$$

where

$$T_j = \frac{1}{j!} T^{(j)}(\lambda), \quad j \geqslant 0. \tag{12}$$

Remark 2.7. Rather than using generalized eigenvectors one can also write $T(z)^{-1}$ in terms of left and right root functions, see [26, Theorem 1.5.4].

The representation (10) and the ordering (9) of multiplicities shows that the order κ of the pole in (4) is given by

$$\kappa = m_1 = \max\{m_\ell : \ell = 1, \dots, L\}.$$

Further, the number $L = \dim(N(T(\lambda)))$ is the geometric multiplicity while $\sum_{\ell=1}^{L} m_\ell$ is the algebraic multiplicity of λ. In the semi-simple case $m_\ell = 1,\ l = 1, \dots, L$, Eqs. (10) and (11) simplify to

$$T(z)^{-1} = (z-\lambda)^{-1} \sum_{\ell=1}^{L} v_0^\ell w_0^{\ell H} + R(z),$$

$$w_0^{\ell H} T'(\lambda) v_0^\nu = \delta_{\nu\ell}, \quad 1 \leqslant \ell,\ \nu \leqslant L,$$

which in case $L = 1$ further simplify to (6) and (5).

Consider now all eigenvalues inside a compact set $\mathcal{C} \subset \Omega$. In the same way as (8) followed from (6), we obtain from Theorem 2.6 the following corollary.

Corollary 2.8. *Let $\mathcal{C} \subset \Omega$ be compact and $T \in H(\Omega, \mathbb{C}^{m,m})$. Then $\mathcal{C}$ contains at most finitely many eigenvalues $\lambda_n,\ n = 1, \dots, n(\mathcal{C})$ with corresponding CSGEs*

$$V_n = \left(v_j^{\ell,n}, 0 \leqslant j \leqslant m_{\ell,n} - 1, 1 \leqslant \ell \leqslant L_n\right), \quad n = 1, \dots, n(\mathcal{C}).$$

Let

$$W_n = \left(w_j^{\ell,n}, 0 \leqslant j \leqslant m_{\ell,n} - 1, 1 \leqslant \ell \leqslant L_n\right), \quad n = 1, \dots, n(\mathcal{C})$$

be the corresponding CSGEs of T^H such that

$$r(w_0^{\ell,n}) = m_{\ell,n}$$

and with $T_{j,n} = \frac{1}{j!} T^{(j)}(\lambda_n)$

$$\sum_{\alpha=0}^{j} \sum_{\beta=1}^{m_{\nu,n}} w_{j-\alpha}^{\ell,nH} T_{\alpha+\beta,n} v_{m_{\nu,n}-\beta}^{\nu,n} = \delta_{\nu\ell}\delta_{0j}, \quad 0 \leqslant j \leqslant m_{\ell,n} - 1,\ 1 \leqslant \ell, \nu \leqslant L_n.$$

Then there exists a neighborhood $\mathcal{C} \subset \mathcal{U} \subset \Omega$ and a function $R \in H(\mathcal{U}, \mathbb{C}^{m,m})$ such that for all $z \in \mathcal{U} \setminus \{\lambda_1, \dots, \lambda_{n(\mathcal{C})}\}$

$$T(z)^{-1} = \sum_{n=1}^{n(\mathcal{C})} \sum_{\ell=1}^{L_n} \sum_{j=1}^{m_{\ell,n}} (z-\lambda_n)^{-j} \sum_{\nu=0}^{m_{\ell,n}-j} v_\nu^{\ell,n} w_{m_{\ell,n}-j-\nu}^{\ell,nH} + R(z).$$

Consider now a contour $\Gamma \subset \Omega$, i.e. a simple closed curve that has its interior $\mathrm{int}(\Gamma)$ in Ω. An easy consequence of the residue theorem is the following result.

Theorem 2.9. *Let $T \in H(\Omega, \mathbb{C}^{m,m})$ have no eigenvalues on the contour $\Gamma \subset \Omega$ and denote by $\lambda_n,\ n = 1, \dots, n(\Gamma)$ the eigenvalues in the interior $\mathrm{int}(\Gamma) \subset \Omega$. Then with the CSGEs from Corollary 2.8 we have for any $f \in H(\Omega, \mathbb{C})$*

$$\frac{1}{2\pi i} \int_\Gamma f(z) T(z)^{-1} dz = \sum_{n=1}^{n(\Gamma)} \sum_{\ell=1}^{L_n} \sum_{j=1}^{m_{\ell,n}} \frac{f^{(j-1)}(\lambda_n)}{(j-1)!} \sum_{\nu=0}^{m_{\ell,n}-j} v_\nu^{\ell,n} w_{m_{\ell,n}-\nu-j}^{\ell,nH}. \tag{13}$$

If all eigenvalues are simple the formula reads

$$\frac{1}{2\pi i}\int_\Gamma f(z)T(z)^{-1}dz = \sum_{n=1}^{n(\Gamma)} f(\lambda_n)v_n w_n^H, \tag{14}$$

where v_n, w_n are left and right eigenvectors corresponding to λ_n and normalized according to

$$w_n^H T'(\lambda_n)v_n = 1, \quad n = 1, \ldots, n(\Gamma). \tag{15}$$

Proof. Corollary 2.8 applies to $\mathcal{C} = \mathrm{int}(\Gamma) \cup \Gamma$, where the function $f(z)T(z)^{-1}$ has residues at λ_j given by the right-hand side of (13). The special case $L_n = 1,\ m_{0n} = 1,\ n = 1, \ldots, n(\Gamma)$ yields Eq. (14). □

3. The algorithm for a few eigenvalues

In the following we set up an algorithm for computing all eigenvalues of $T \in H(\Omega, \mathbb{C}^{m,m})$ inside a given contour Γ in Ω. We assume that the sum of all algebraic multiplicities

$$k = \sum_{n=1}^{n(\Gamma)} \sum_{\ell=1}^{L_n} m_{\ell,n} \tag{16}$$

is not larger than the system dimension m. For the opposite case we refer to Section 5. In large-scale problems we actually expect to have $k \ll m$.

3.1. Simple eigenvalues inside the contour

As in the second part of Theorem 2.9, let us assume that all eigenvalues $\lambda_1, \ldots, \lambda_{n(\Gamma)}$ in $\mathrm{int}(\Gamma)$ are simple so that $k = n(\Gamma)$. We introduce the matrices

$$V = \begin{pmatrix} v_1 \ldots v_k \end{pmatrix}, W = \begin{pmatrix} w_1 \ldots w_k \end{pmatrix} \in \mathbb{C}^{m,k}.$$

We assume that we have chosen a matrix

$$\hat{V} \in \mathbb{C}^{m,l}, \quad k \leqslant l \leqslant m,$$

such that

$$W^H\hat{V} \in \mathbb{C}^{k,l} \text{ has rank } k. \tag{17}$$

Note that this requires $\mathrm{rank}(W) = k$. In applications we choose $\hat{V}$ at random (see Section 4), so that (17) can be expected to hold in a generic sense if $\mathrm{rank}(W) = k$. We note that (in contrast to linear eigenvalue problems) it is easy to construct nonlinear eigenvalue problems for which W is rank deficient. However, this seems to be a nongeneric situation for typical applications. In addition to (17) we assume

$$\mathrm{rank}(V) = k, \tag{18}$$

which again is expected to hold in generic cases.

Next we compute the two integrals

$$A_0 = \frac{1}{2\pi i}\int_\Gamma T(z)^{-1}\hat{V}dz \in \mathbb{C}^{m,l}, \tag{19}$$

$$A_1 = \frac{1}{2\pi i}\int_\Gamma zT(z)^{-1}\hat{V}dz \in \mathbb{C}^{m,l}. \tag{20}$$

The evaluation of these integrals by quadrature rules is by far the most expensive part of the algorithm and will be discussed below. Note also, that in the linear case $T(z) = zI - A$ the matrix A_0 is obtained by applying to $\hat{V}$ the Riesz projector onto the invariant subspace associated with all eigenvalues inside Γ.

By (14) we obtain

$$A_0 = \sum_{n=1}^{k} v_n w_n^H \hat{V} = V W^H \hat{V}. \tag{21}$$

Similarly,

$$A_1 = \sum_{n=1}^{k} \lambda_n v_n w_n^T \hat{V} = V \Lambda W^H \hat{V}, \quad \Lambda = \mathrm{diag}(\lambda_n, n = 1, \ldots, k). \tag{22}$$

In the next step we compute the singular value decomposition (SVD) of $A_0 \in \mathbb{C}^{m,l}$ in reduced form

$$V W^H \hat{V} = A_0 = V_0 \Sigma_0 W_0^H, \tag{23}$$

where $V_0 \in \mathbb{C}^{m,k}$, $\Sigma_0 = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$, $W_0 \in \mathbb{C}^{l,k}$, $V_0^H V_0 = I_k$, $W_0^H W_0 = I_k$. Note that the rank conditions (17), (18) show that $\mathrm{rank}(A_0) = k$, hence A_0 has singular values

$$\sigma_1 \geqslant \cdots \geqslant \sigma_k > 0 = \sigma_{k+1} = \cdots = \sigma_l.$$

By the rank condition (18) we have

$$R(A_0) = R(V) = R(V_0).$$

Since both V_0 and V are $m \times k$ matrices and V_0 has orthonormal columns, we obtain

$$V = V_0 S, \quad S = V_0^H V \in \mathbb{C}^{k,k} \text{ nonsingular.} \tag{24}$$

With (21), (24) we find $V_0 S W^H \hat{V} = V_0 \Sigma_0 W_0^H$ and thus

$$W^H \hat{V} = S^{-1} \Sigma_0 W_0^H.$$

This relation is used to eliminate $W^H \hat{V}$ from $A_1 = V_0 S \Lambda W^H \hat{V}$. We obtain

$$V_0^H A_1 = S \Lambda W^H \hat{V} = S \Lambda S^{-1} \Sigma_0 W_0^H,$$

which upon multiplication by $W_0 \Sigma_0^{-1}$ from the right finally gives

$$S \Lambda S^{-1} = V_0^H A_1 W_0 \Sigma_0^{-1}. \tag{25}$$

Note that the matrix on the right-hand side can be computed from the integrals A_0, A_1 without a priori information about eigenvalues and eigenvectors. The matrix is diagonalizable and has as eigenvalues exactly the eigenvalues of T inside the contour. We summarize the result in a theorem.

Theorem 3.1. *Suppose that $T \in H(\Omega, \mathbb{C}^{m,m})$ has only simple eigenvalues $\lambda_1, \ldots, \lambda_k$ inside the contour Γ in Ω with left and right eigenvectors normalized as in (15). Moreover, let a matrix $\hat{V} \in \mathbb{C}^{m,l}$ be given such that $k \leqslant l \leqslant m$ and the rank conditions (17), (18) are satisfied. Then the matrix*

$$B = V_0^H A_1 W_0 \Sigma_0^{-1} \in \mathbb{C}^{k.k}, \tag{26}$$

given by (19), (20) and the SVD (23), is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_k$. From the eigenvectors $s_1, \ldots, s_k \in \mathbb{C}^k$ of B one obtains the eigenvectors of T through

$$v_n = V_0 s_n, \quad n = 1, \ldots, k.$$

Remarks 3.2. (a) The method proposed here is very close to the method of [2]. In their block method the authors of [2] premultiply A_0 and A_1 by another test matrix $U^H \in \mathbb{C}^{k,m}$ and then perform an SVD and solve an eigenvalue problem. The derivation of the algorithm is quite different and uses the Smith normal form [11,10].

(b) For reasons of numerical stability we may replace A_1 by

$$\tilde{A}_1 = \frac{1}{2\pi i}\int_\Gamma (z-z_0)T(z)^{-1}\hat{V}\,dz = A_1 - z_0 A_0.$$

For example, in case of a circle Γ, one can take z_0 as its center. Then (22) holds with $\Lambda - z_0$ instead of Λ and the matrix $\tilde{B} = V_0^H \tilde{A}_1 W_0 \Sigma_0^{-1}$ has eigenvalues $\lambda_n - z_0$. Therefore, the eigenvalues of T are found by adding z_0 to the eigenvalues of $\tilde{B}$. This is also proposed in [1,2].

(c) If the rank of either V or W or both is not maximal then we have a degenerate situation that does not occur for linear eigenvalue problems. In principle such a case can be detected from a rank drop of Σ_0 provided we know the number k from (16) beforehand. However, this information is usually not available. Instead, we propose to include a test of residuals in order to detect this situation (see the algorithm in Section 3.3 for more details and, in particular, Example 4.12 in Section 4.2). A general cure for this rank deficient case is provided by the generalized algorithm in Section 5 which uses the higher order moments from (3). Finally, a rank drop of $W^H\hat{V}$ due to the random choice of $\hat{V}$ hardly ever occurs in practice.

(d) An alternative to solve the eigenvalue problem for B is not to invert Σ_0, but rather to solve the generalized eigenvalue problem for the matrix pencil $(V_0^H A_1 W_0,\ \Sigma_0)$. Generally this will improve the conditioning of the eigenvalue computations if Σ_0 contains small singular values. However, it does not help with the rank deficient case discussed in (c) above, see also Section 3.3. One could go even further and try to avoid SVD's completely. Note that by the representations (21) and (22) the eigenvalue problem $\det(\Lambda - \lambda I_k) = 0$ is equivalent to the rectangular eigenvalue problem $\operatorname{rank}(A_1 - \lambda A_0) < k$ (see [7] for the numerical treatment). However, reliable algorithms for such problems seem not to be available.

3.2. Multiple eigenvalues inside the contour

Let us consider the general case where $T \in H(\Omega, \mathbb{C}^{m,m})$ has no eigenvalues on the contour Γ but may have multiple eigenvalues inside. We apply Corollary 2.8 to the compact set $\mathcal{C} = \Gamma \cup \operatorname{int}(\Gamma)$ and assume that the matrix composed of all CSGEs that belong to eigenvalues inside Γ,

$$V = \left(v_j^{\ell,n},\ 0 \leqslant j \leqslant m_{\ell,n}-1,\ 1 \leqslant \ell \leqslant L_n,\ 1 \leqslant n \leqslant n(\Gamma)\right), \tag{27}$$

has rank k, cf. (16). Then, using Theorem 2.9 with $f(z) = 1$ shows that A_0, as defined in (19), satisfies

$$A_0 = \sum_{n=1}^{n(\Gamma)} \sum_{\ell=1}^{L_n} \sum_{\nu=0}^{m_{\ell,n}-1} v_\nu^{\ell,n} w_{m_{\ell,n}-1-\nu}^{\ell,nH} \hat{V}.$$

Further, we assume that the matrix

$$W^H\hat{V} \in \mathbb{C}^{k,l} \tag{28}$$

has maximum rank k, where

$$W = \left(w_{m_{\ell,n}-1-\nu}^{\ell,n},\ 0 \leqslant \nu \leqslant m_{\ell,n}-1,\ 1 \leqslant \ell \leqslant L_n,\ 1 \leqslant n \leqslant n(\Gamma)\right) \in \mathbb{C}^{m,k}, \tag{29}$$

is normalized as in Theorem 2.6. With Theorem 2.9 we then find

$$A_1 = \sum_{n=1}^{n(\Gamma)} \sum_{\ell=1}^{L_n} \left[\lambda_n \sum_{\nu=0}^{m_{\ell,n}-1} v_\nu^{\ell,n} w_{m_{\ell,n}-1-\nu}^{\ell,nH} + \sum_{\nu=0}^{m_{\ell,n}-2} v_\nu^{\ell,n} w_{m_{\ell,n}-2-\nu}^{\ell,nH}\right] \hat{V} = V\Lambda W^H\hat{V},$$

where Λ has Jordan normal form

$$\Lambda = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_{n(\Gamma)} \end{pmatrix}, \quad J_n = \begin{pmatrix} J_{n,1} & & \\ & \ddots & \\ & & J_{n,L_n} \end{pmatrix},$$
$$J_{n,\ell} = \begin{pmatrix} \lambda_n & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_n & 1 \\ & & & \lambda_n \end{pmatrix} \in \mathbb{C}^{m_{\ell,n}, m_{\ell,n}}. \tag{30}$$

As in Section 3.1 the next steps are the SVD (23) for A_0 and the computation of $B = V_0^H A_1 W_0 \Sigma_0^{-1} \in \mathbb{C}^{k.k}$. Then B has eigenvalues $\lambda_1, \ldots, \lambda_{n(\Gamma)}$ and its Jordan normal form has the same partial multiplicities as $T(z)$.

Theorem 3.3. *Suppose that $T \in H(\Omega, \mathbb{C}^{m,m})$ has no eigenvalues on the contour Γ in Ω and pairwise distinct eigenvalues $\lambda_n,\ n = 1, \ldots, n(\Gamma)$ inside Γ with partial multiplicities $m_{1,n} \geqslant \cdots \geqslant m_{L_n,n},\ n = 1, \ldots, n(\Gamma)$. Moreover, assume that the matrix of generalized eigenvectors from (27) and the matrix $W^H \hat{V}$ from (28) have rank k with k given by (16). Then the matrix $B \in \mathbb{C}^{k,k}$ from (26) has Jordan normal form (30) with the same eigenvalues λ_n and partial multiplicities $m_{\ell,n}$ ($\ell = 1, \ldots, L_n, n = 1, \ldots, n(\Gamma)$). Suitable CSGEs for T can be obtained from corresponding CSGEs $s_j^{\ell,n}$ for B via*

$$v_j^{\ell,n} = V_0 s_j^{\ell,n}, \quad 0 \leqslant j \leqslant m_{\ell,n} - 1,\ 1 \leqslant \ell \leqslant L_n,\ 1 \leqslant n \leqslant n(\Gamma). \tag{31}$$

Remark 3.4. Essentially, the theorem reduces the nonlinear problem for eigenvalues inside a contour to a linear eigenvalue problem for a $k \times k$-matrix. The linear eigenvalue problem inherits the multiplicity structure of the nonlinear problem.

3.3. Quadrature and numerical realization

The major step in the algorithm consists of evaluating the integrals (19) and (20) by numerical quadrature and by solving the linear systems involved in the evaluation of the integrand. We assume that Γ has a 2π-periodic smooth parameterization

$$\varphi \in C^1(\mathbb{R}, \mathbb{C}), \quad \varphi(t + 2\pi) = \varphi(t) \quad \forall t \in \mathbb{R}.$$

Of particular interest is the real analytic case $\varphi \in C^\omega(\mathbb{R}, \mathbb{C})$. Taking equidistant nodes $t_j = \frac{2j\pi}{N}$, $j = 0, \ldots, N$ and using the trapezoid sum, we find the following approximations

$$\begin{aligned} A_0 &= \frac{1}{2\pi i} \int_0^{2\pi} T(\varphi(t))^{-1} \hat{V} \varphi'(t) dt \approx \\ A_{0,N} &= \frac{1}{iN} \sum_{j=0}^{N-1} T(\varphi(t_j))^{-1} \hat{V} \varphi'(t_j), \end{aligned} \tag{32}$$

where we used $\varphi(t_0) = \varphi(t_N)$. Similarly,

$$A_1 \approx A_{1,N} = \frac{1}{iN} \sum_{j=0}^{N-1} T(\varphi(t_j))^{-1} \hat{V} \varphi(t_j) \varphi'(t_j). \tag{33}$$

In order to compute $A_{0,N}$ we need to solve Nl linear systems with N different matrices $T(\varphi(t_j))$, $j = 0, \ldots, N-1$ and with l different right-hand sides each. Note that we can use the solutions of these linear systems to compute $A_{1,N}$ at almost no extra cost. For the special case of a circle $\varphi(t) = \mu + Re^{it}$ we obtain the formulas

$$A_{0,N} = \frac{R}{N} \sum_{j=0}^{N-1} T(\varphi(t_j))^{-1} \hat{V} \exp\left(\frac{2\pi ij}{N}\right),$$

$$A_{1,N} = \mu A_{0,N} + \frac{R^2}{N} \sum_{j=0}^{N-1} T(\varphi(t_j))^{-1} \hat{V} \exp\left(\frac{4\pi ij}{N}\right).$$

The algorithm can be summarized as follows:

Integral algorithm 1

Step 1: Choose an index $l \leqslant m$ and a matrix $\hat{V} \in \mathbb{C}^{m,l}$ at random.
Step 2: Compute $A_{0,N}, A_{1,N}$ from (32), (33).
Step 3: Compute the SVD $A_{0,N} = V\Sigma W^H$, where $V \in \mathbb{C}^{m,l}$, $W \in \mathbb{C}^{l,l}$, $V^HV = W^HW = I_l$, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_l)$.
Step 4: Perform a rank test for Σ, i.e. find $0 < k \leqslant l$ such that $\sigma_1 \geqslant \ldots \geqslant \sigma_k > \mathrm{tol}_{\mathrm{rank}} > \sigma_{k+1} \approx \ldots \approx \sigma_l \approx 0$. If $k = l$ then increase l and go to Step 1. Else let $V_0 = V(1:m, 1:k)$, $W_0 = W(1:l, 1:k)$ and $\Sigma_0 = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$.
Step 5: Compute $B = V_0^H A_{1,N} W_0 \Sigma_0^{-1} \in \mathbb{C}^{k,k}$.
Step 6: Solve the eigenvalue problem for B. If all eigenvalues $\lambda_j, j = 1, \ldots, k$ are well-conditioned with corresponding eigenvectors s_j then accept λ_j and the eigenvector $v_j = V_0 s_j$ provided $\lambda_j \in \mathrm{int}(\Gamma)$ and $\|T(\lambda_j)v_j\| \leqslant \mathrm{tol}_{\mathrm{res}}$.
Otherwise compute a Schur decomposition $BQ = QU$ with Q unitary and U upper triangular. Reorder eigenvalues such that eigenvalues inside Γ occur first and discard eigenvalues outside Γ and the corresponding columns of Q. Block diagonalize U such that diagonal blocks belong to different eigenvalues. Let λ_j be the diagonal entry of the jth block and determine the corresponding eigenvector $s_j \in \mathbb{C}^k$ of B from the first column of the jth block in U. Accept $v_j = V_0 s_j$ as eigenvector and λ_j as eigenvalue if $\|T(\lambda_j)v_j\| \leqslant \mathrm{tol}_{\mathrm{res}}$. For details we refer to MATLAB routines `eig, condeig, schur, ordschur` and the remarks below.

Remarks 3.5. (a) If we find $k = l$ positive singular values in Step 4 then we take this as an indication that there may be more than l eigenvalues (including multiplicities) inside Γ. We then increase l until a rank drop is detected in Step 4. If l reaches the value m then a warning advises the user to try the extended algorithm in Section 5.
(b) Note that $A_{1,N}$ is computed from the same columns as $A_{0,N}$, just new coefficients appear. Therefore, it is generally not necessary to store factorizations of $T(\varphi(t_j))$ if a direct solver is used. However, when l is increased in Step 4, new columns of $\hat{V}$ appear and one can take advantage of previous factorizations for computing the new columns of $A_{0,N}$, $A_{1,N}$. If storage space is crucial or iterative solvers are used or if l is increased many times it can be advantageous to shift the computation of $A_{1,N}$ to Step 5 (when k has been determined) and solve the linear systems with multiple right-hand sides in a second pass.
(c) According to Theorem 3.3 the matrix B retains the complete multiplicity structure of eigenvalues inside the contour. Therefore, a Schur decomposition followed by reordering and block diagonalization is used for the case of multiple eigenvalues in Step 6 (cf. [13, Chapter 7.6]) . However, currently only the first column vector belonging to any diagonal block is taken when testing the residual. No attempt is made to recover non-constant root functions from the remaining columns. This is an interesting topic for further investigations.
(d) The examples in Section 4.2 will show that the choice of the three parameters N, $\mathrm{tol}_{\mathrm{rank}}$, $\mathrm{tol}_{\mathrm{res}}$ is crucial for the algorithm. If the rank defect mentioned in Remark 3.2(c) occurs then eigenvalues inside the contour may be missed completely. In this situation other eigenvalues found inside the contour

typically have low accuracy which is indicated by $\mathrm{tol}_{\mathrm{res}}$, see Example 4.12. If eigenvalues lie close to the contour and N is not large enough then quadrature errors become dominant. In this case the rank test produces k-values larger than the number of eigenvalues inside and Step 6 finds eigenvalues outside but close to the contour, see Example 4.10. The accuracy of these outside eigenvalues is still reasonable unless $\mathrm{tol}_{\mathrm{rank}}$ has been chosen such that extremely small singular values are included in Step 4, cf. Example 4.10.

(e) As mentioned in Remark 3.2 (d) the conditioning of the eigenvalue problem improves if Σ_0 is not inverted but a generalized eigenvalue problem is solved instead. In the examples below this did not make a noticeable difference, probably due to a conservative choice of the rank tolerance in Step 4.

4. Error analysis and numerical examples

4.1. Error analysis

Standard results on the trapezoid sum for holomorphic periodic integrands imply exponential convergence at a rate that depends on the number of nodes times the width of the horizontal strip of holomorphy, see [8]; [9, 4.6.5]. Applications of these results to the computation of matrix functions via contour integrals appear in [15].

Theorem 4.1. *Let $f \in H(S(d_-, d_+), \mathbb{C})$ be 2π-periodic on the strip*

$$S(d_-, d_+) = \{z \in \mathbb{C} : -d_- < \operatorname{Im} z < d_+\}, \quad d_\pm > 0.$$

Then the error of the trapezoid sum

$$E_N(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx - \frac{1}{N}\sum_{j=0}^{N-1} f\left(\frac{2\pi j}{N}\right)$$

satisfies for all $0 < r_- < d_-,\ 0 < r_+ < d_+$

$$|E_N(f)| \leqslant \max_{\operatorname{Im}(z)=r_+} |f(z)|G(e^{-Nr_+}) + \max_{\operatorname{Im}(z)=r_-} |f(z)|G(e^{-Nr_-}),$$

where $G(x) = \frac{x}{1-x},\ x \neq 1$.

Remark 4.2. Note that Theorem 4.1 is a slight variation of [9, 4.6.5] since f is not assumed to be real on $[0, 2\pi]$ and the strip $S(d_-, d_+)$ can be unsymmetric, in general.

In the proof of [9, 4.6.5] the real line is transformed into a closed circle via the exponential function and then a Laurent expansion is used. Since we need such a result for closed contours in the sequel, we give an explicit formulation and, for better readability, include the rather short proof.

Theorem 4.3. *Let $f \in H(A(a_-, a_+), \mathbb{C})$ be holomorphic on the annulus*

$$A(a_-, a_+) = \left\{z \in \mathbb{C} : \frac{1}{a_-} < \frac{|z|}{R} < a_+\right\}, \quad a_\pm > 1,$$

for some $R > 0$. Then the error of the trapezoid sum

$$E_N(f) = \frac{1}{2\pi i}\int_{|z|=R} f(z)dz - \frac{R}{N}\sum_{j=0}^{N-1} f(R\omega_N^j)\omega_N^j, \quad \omega_N = \exp\left(\frac{2\pi i}{N}\right), \tag{34}$$

satisfies for all $1 < \rho_- < a_-,\ 1 < \rho_+ < a_+$

$$|E_N(f)| \leqslant \max_{|z|=\rho_+ R} |f(z)| G(\rho_+^{-N}) + \max_{\rho_-|z|=R} |f(z)| G(\rho_-^{-N}). \tag{35}$$

Proof. We use the Laurent expansion of f (see e.g. [14])

$$f(z) = \sum_{k=-\infty}^{\infty} f_k z^k, \quad f_k = \frac{1}{2\pi i} \int_{|z|=R} f(z) z^{-k-1} dz, \tag{36}$$

which converges uniformly on compact subdomains of the annulus. By a simple computation,

$$E_N(z^k) = \begin{cases} -R^{\ell N}, & k+1 = \ell N,\ \ell \in \mathbb{Z} \setminus \{0\}, \\ 0, & \text{otherwise.} \end{cases}$$

Applying E_N to (36) leads to

$$E_N(f) = -\sum_{\ell=1}^{\infty} (f_{\ell N} R^{\ell N} + f_{-\ell N} R^{-\ell N}). \tag{37}$$

By Cauchy's Theorem we can shift the contour from $|z| = R$ to $|z| = \rho_+ R$ and obtain by a standard estimate

$$\begin{aligned} |f_{\ell N} R^{\ell N}| &= \left| \tfrac{R^{\ell N}}{2\pi i} \textstyle\int_{|z|=R} f(z) z^{-\ell N-1} dz \right| \\ &= R^{\ell N} \left| \tfrac{1}{2\pi i} \textstyle\int_{|z|=\rho_+ R} f(z) z^{-\ell N-1} dz \right| \\ &\leqslant \tfrac{R^{\ell N}}{2\pi} 2\pi \rho_+ R \max_{|z|=\rho_+ R} |f(z)|\, (\rho_+ R)^{-\ell N-1} \\ &= \max_{|z|=\rho_+ R} |f(z)| \rho_+^{-\ell N}. \end{aligned}$$

In a similar way,

$$|f_{-\ell N} R^{-\ell N}| \leqslant \max_{\rho_-|z|=R} |f(z)|\ \rho_-^{-\ell N}.$$

Using these estimates in (37) completes the proof. □

The proof shows that the ρ_--term can be discarded in (35) if the principal term in the Laurent expansion vanishes (i.e. $f_k = 0$ for $k \leqslant -1$). Likewise, the ρ_+-term disappears when $f_k = 0$ for $k \geqslant 0$. For the function

$$f(z) = (z-\lambda)^{-j}, \quad j \geqslant 1, \tag{38}$$

the principal term vanishes for $|\lambda| > R$ while the secondary term vanishes for $|\lambda| < R$. Example (38) is crucial for the application to the meromorphic functions from Section 3. Therefore, we note the following explicit formula.

Lemma 4.4. *The error of the trapezoid sum* (34) *for the function* (38) *in case* $N \geqslant j$ *is given as follows,*

$$E_N((z-\lambda)^{-j}) = \frac{(-1)^{j-1}\lambda^{-j}}{(j-1)!} \begin{cases} \frac{d^{j-1}}{dx^{j-1}} (x^{j-1} G(x^{-N}))_{|x=\frac{R}{\lambda}}, & |\lambda| < R, \\ \frac{d^{j-1}}{dx^{j-1}} (x^{j-1} G(x^{N}))_{|x=\frac{R}{\lambda}}, & |\lambda| > R. \end{cases} \tag{39}$$

In particular,

$$E_N((z-\lambda)^{-j}) = \begin{cases} \mathcal{O}\left(|\lambda|^{-j}\left(\frac{|\lambda|}{R}\right)^{N-j+1}\right), & |\lambda| < R, \\ \mathcal{O}\left(|\lambda|^{-j}\left(\frac{R}{|\lambda|}\right)^{N+j-1}\right), & |\lambda| > R. \end{cases} \tag{40}$$

Remark 4.5. If $f \in H(A(a_-, a_+), \mathbb{C})$ is meromorphic on an open neighborhood of the closed annulus $A(a_-, a_+)^c$, then the estimate (35) can be sharpened as follows

$$E_N(f) = \mathcal{O}(a_+^{-N} + a_-^{-N}).$$

In order to see this, first consider the singular part that belongs to poles on the boundary of $A(a_-, a_+)$, and use Lemma 4.4. Then apply Theorem 4.3 to the remaining part on a slightly larger annulus.

Consider a general contour Γ in Ω with 2π-periodic parametrization $\varphi(t)$, $t \in [0, 2\pi]$. Moreover, assume that φ has a 2π-periodic holomorphic extension to a strip

$$\varphi \in H(S(d_-, d_+), \Omega), \quad \varphi(z + 2\pi) = \varphi(z). \tag{41}$$

For definiteness, we also assume that

$$\varphi(z) \begin{cases} \in \operatorname{int}(\Gamma), & 0 < \operatorname{Im}(z) < d_+, \\ \notin \operatorname{int}(\Gamma), & -d_- < \operatorname{Im}(z) < 0. \end{cases} \tag{42}$$

Common examples are circles $\varphi(z) = z_0 + Re^{iz}$ with $z \in \mathbb{C}$ and ellipses $\varphi(z) = a\cos(z) + ib\sin(z)$ with $|\operatorname{Im}(z)| < \operatorname{artanh}\left(\min\left(\frac{a}{b}, \frac{b}{a}\right)\right)$.

Consider $g \in H(\Omega, \mathbb{C})$. Then the error of the trapezoid sum for $f(z) = g(\varphi(z))\varphi'(z), z \in S(d_-, d_+)$ is

$$E_N(g) = \frac{1}{2\pi i}\int_\Gamma g(z)dz - \frac{1}{iN}\sum_{j=0}^{N-1} g\left(\varphi\left(\frac{2\pi j}{N}\right)\right)\varphi'\left(\frac{2\pi j}{N}\right). \tag{43}$$

From Theorem 4.1 we obtain an estimate

$$|E_N(g)| \leqslant \Phi(r_+)G(e^{-Nr_+}) + \Phi(r_-)G(e^{-Nr_-}), \tag{44}$$

where $0 < r_- < d_-$, $0 < r_+ < d_+$ and $\Phi(r) = \max_{\operatorname{Im}(z)=r} |\varphi'(z)\|g(\varphi(z))|$. The following lemma gives a rough estimate of the right-hand sides for the pole function $g(z) = (z-\lambda)^{-j}$, $\lambda \in \Omega$.

Lemma 4.6. *Let Ω be bounded and let φ satisfy conditions (41), (42). Then there exist constants $C_1, C_2, C_3 > 0$ (depending on φ, j but not on N or $\lambda \in \Omega$) such that for* $\operatorname{dist}(\lambda, \Gamma) \leqslant C_3$,

$$|E_N((\cdot - \lambda)^{-j})| \leqslant C_1 \operatorname{dist}(\lambda, \Gamma)^{-j} \exp\left(-C_2 N \operatorname{dist}(\lambda, \Gamma)\right). \tag{45}$$

Proof. For a fixed $0 < q < 1$ there are bounds $|\varphi'(z)| \leqslant M_+$ for $0 \leqslant \operatorname{Im}(z) \leqslant qd_+$ and $|\varphi'(z)| \leqslant M_-$ for $0 \leqslant -\operatorname{Im}(z) \leqslant qd_-$. Let $C_3 = \max(M_+d_+, M_-d_-)$ and define $r_+ = \frac{q\operatorname{dist}(\lambda,\Gamma)}{M_+}$. Then there exists some $z_+ = s_+ + ir_+$, $0 \leqslant s_+ < 2\pi$ such that

$$\begin{aligned} \min_{\operatorname{Im}(z)=r_+} |\lambda - \varphi(z)| &= |\lambda - \varphi(z_+)| \geqslant |\lambda - \varphi(s_+)| - |\varphi(s_+) - \varphi(z_+)| \\ &\geqslant \operatorname{dist}(\lambda, \Gamma) - M_+r_+ = (1-q)\operatorname{dist}(\lambda, \Gamma). \end{aligned}$$

The first term in (44) can be estimated as follows

$$|\Phi(r_+)|G(e^{-Nr_+}) \leqslant M_+ \max_{\mathrm{Im}\, z = r_+} |(\varphi(z)-\lambda)^{-j}|G(e^{-Nr_+})$$
$$\leqslant C(1-q)^{-j} M_+ \mathrm{dist}(\lambda, \Gamma)^{-j} \exp\left(-N\mathrm{dist}(\lambda, \Gamma)\frac{q}{M_+}\right).$$

The second term is treated analogously. □

As a consequence of Lemmas 4.4 and 4.6 we obtain an exponential estimate for the errors in (32) and (33).

Theorem 4.7. *For $T \in H(\Omega, \mathbb{C})$ let T^{-1} have poles of order at most κ, cf. Theorem 2.2. Further, let Γ be a simple closed contour in Ω with $\sigma(T) \cap \Gamma = \emptyset$ and such that the parametrization φ satisfies (41) and (42). Then there exist constants $C_1, C_2 > 0$ (depending on T and $\hat{V}$ but not on N) such that the matrices from (32), (33) satisfy*

$$\|A_p - A_{p,N}\| \leqslant C_1 d(T)^{-\kappa} e^{-C_2 N d(T)}, \quad p = 0, 1,$$

where $d(T) = \min_{\lambda\in\sigma(T)} \mathrm{dist}(\lambda, \Gamma)$ and $d(T) = 1$ if $\sigma(T) = \emptyset$. If Γ is a circle with parametrization $\varphi(t) = z_0 + Re^{it}$, then the following estimate holds

$$\|A_p - A_{p,N}\| \leqslant C_1 \left[\rho_-^{N-\kappa+1} + \rho_+^{N+\kappa-1}\right], \quad p = 0, 1,$$

where

$$\rho_- = \max_{\lambda\in\sigma(T), |\lambda-z_0|<R} \frac{|\lambda - z_0|}{R}, \quad \rho_+ = \max_{\lambda\in\sigma(T), |\lambda-z_0|>R} \frac{R}{|\lambda - z_0|}.$$

Combining these estimates with the well-known perturbation theory for singular value decompositions [31] we find that the integral algorithm detects the correct rank k of $A_{0,N}$ if N is sufficiently large. Further, the perturbation theory for simple eigenvalues [31] leads to the following corollary.

Corollary 4.8. *Let the assumptions of Theorems 3.1 and 4.7 be satisfied. Let $\lambda_1, \ldots, \lambda_k$ be the eigenvalues of T inside Γ and let $\lambda_{1,N}, \ldots, \lambda_{k,N}$ be the eigenvalues from Step 6 of the integral algorithm. With the notation from Theorem 4.7 we then have the error estimates*

$$\max_{j=1,\ldots,n(\Gamma)} |\lambda_j - \lambda_{j,N}| \leqslant C_1 d(T)^{-\kappa} e^{-C_2 N d(T)},$$

in case of a general curve satisfying (41), (42), and

$$\max_{j=1,\ldots,n(\Gamma)} |\lambda_j - \lambda_{j,N}| \leqslant C\left[\rho_-^{N-\kappa+1} + \rho_+^{N+\kappa-1}\right]$$

in case of a circle with radius R and center z_0.

4.2. Numerical examples

Example 4.9. For the first test we choose a real quadratic polynomial

$$T(z) = T_0 + zT_1 + z^2 T_2, \quad T_j \in \mathbb{R}^{60,60},\ j = 0, 1, 2, \tag{46}$$

where T_0, T_1, T_2 are taken at random (*rand* from MATLAB). In this case we can compare with the spectrum $\sigma_{\mathrm{polyeig}}$ resulting from MATLAB's *polyeig*.

Fig. 1(left) shows the result from polyeig (open circles) and the eigenvalues from Integral algorithm 1 (filled boxes) for the data

$$\varphi(t) = Re^{it},\ t \in [0, 2\pi],\ R = 0.33,\ \mathrm{tol}_{\mathrm{rank}} = 10^{-4},\ \mathrm{tol}_{\mathrm{res}} = 10^{-1}. \tag{47}$$

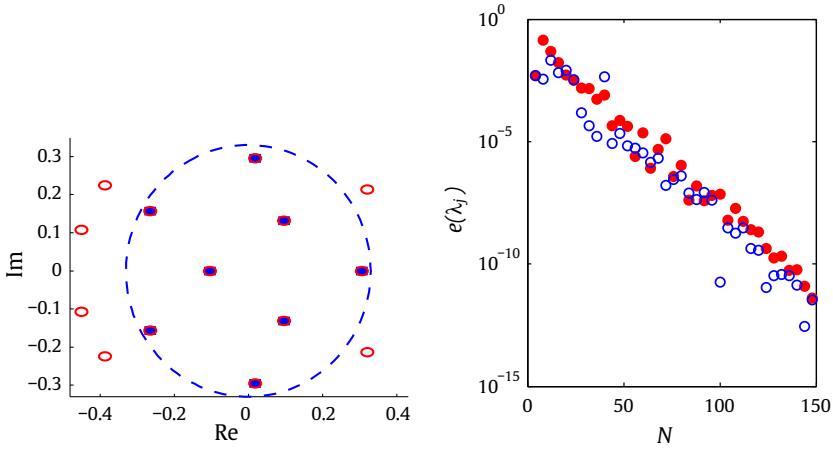

Fig. 1. Example 4.9: Eigenvalues of a quadratic eigenvalue problem from polyeig (open circles) and Integral algorithm 1 (filled squares) with $N = 150$ (left). Difference $e(\lambda_j)$ of eigenvalues $\lambda_1 \approx 0.30578$ (filled circles) and $\lambda_2 \approx 0.0961 - 0.1315i$ (open circles) between polyeig and the integral algorithm versus the number of nodes N (right).

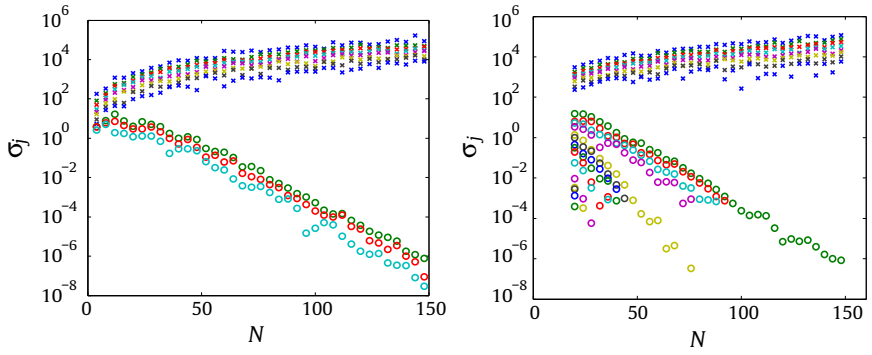

Fig. 2. Example 4.9: Singular values versus N for a fixed number of $l = 11$ columns in the integral algorithm (left), reduction of the number of singular values by the rank test of the adaptive algorithm versus N (right).

The eight eigenvalues inside the circle are detected and well approximated by the integral algorithm. Fig. 1(right) shows the errors

$$e(\lambda_j) = \min\{|\lambda_j - \mu| : \mu \in \sigma_{\text{polyeig}}\}$$

for two characteristic eigenvalues inside the circle. Both show exponential decay with respect to N at approximately the same rate.

While Fig. 1(left) results from the integral algorithm with an adaptive number l of columns (which yields $l = 8$ at $N = 150$), the computations in Fig. 1(right) are done with a fixed number of $l = 11$ columns. For this case we show the behavior of the 11 largest singular values of $A_{0,N}$ in Fig. 2(left). Sufficient separation of singular values already occurs at values $N \approx 25$, much smaller than 150. Fig. 2(right) shows how the adaptive algorithm reduces the number of singular values from $l = 23$ at $N = 20$ to $l = 8$ for $N \geqslant 95$.

Example 4.10. For the next experiment we take random complex entries in (46), a fixed number $l = 10$ of columns, and the same circle as in (47). Again, the six eigenvalues inside the circle from *polyeig* are well approximated by the integral algorithm, see Fig. 3(left).

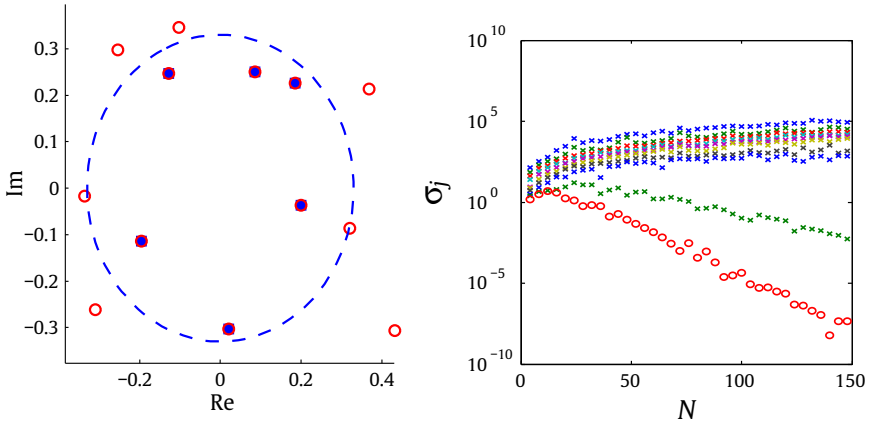

Fig. 3. Example 4.10: Eigenvalues from polyeig (open circles) and eigenvalues from the integral algorithm for a random quadratic complex matrix polynomial (left), singular values of integral algorithm with $l = 10$ columns versus the number N of quadrature nodes for the same example (right).

But this time the singular values do not separate as well as in Fig. 2(left). Two of them decay rather slowly, while two others, due to eigenvalues very close but outside the contour, remain of order one. However, this behavior does not result in spurious eigenvalues. On the contrary, if we keep $k = l = 10$ for the eigenvalue computation, then this yields the 6 eigenvalues inside and in addition the four eigenvalues lying closest to the contour, but outside. Errors and residuals of approximate eigenvalues decay exponentially (as N increases to 150) to values of $\approx 10^{-11}$ for inside and $\approx 10^{-6}$ for outside eigenvalues. Therefore, in this case only the contour test in Step 6 discards the eigenvalues outside the contour. Such a behavior is also suggested by our error analyis in Section 4.1, according to which the principle error term depends on the distance of eigenvalues to the contour, both for eigenvalues inside and outside.

Computational experience shows that spurious eigenvalues only arise when one uses large values of l and very small singular values. For example, $l = 30$, $\mathrm{tol}_{\mathrm{res}} = 10^{-15}$ leads to a spurious eigenvalue at $-0.0060 + 0.0294i$ with residual 3.426 which is easily avoided by the residual test in Step 6.

Example 4.11. This example, taken from [30,20], is a finite element discretization of a nonlinear boundary eigenvalue problem

$$-u''(x) = \lambda u(x), \quad 0 \leqslant x \leqslant 1, \quad u(0) = 0 = u'(1) + \frac{\lambda}{\lambda - 1} u(1).$$

The matrix function is $T(z) = T_1 + \frac{1}{1-z} e_m e_m^T - zT_3$, where

$$T_1 = m \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & 2 & -1 \\ & & -1 & 1 \end{pmatrix}, \quad T_3 = \frac{1}{6m} \begin{pmatrix} 4 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & 4 & 1 \\ & & 1 & 2 \end{pmatrix}.$$

We use $m = 400$ and compute five eigenvalues in the interval $[2, 298]$. Again Fig. 4(left) shows the real eigenvalues in the circle which agree with those from [20]. Note that we avoided the singularity of T at $z = 1$. The residuals of the computed eigenvectors and eigenvalues decay exponentially as expected, see Fig. 4, but not as smoothly as in the previous examples.

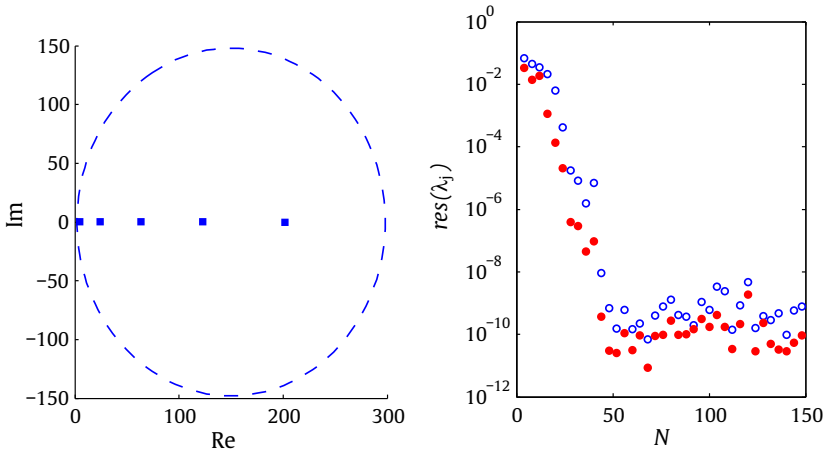

Fig. 4. Example 4.11: Eigenvalues from the integral algorithm for the finite element discretization of a nonlinear boundary eigenvalue problem (left), decay of residuals $\mathrm{res}(\lambda_j) = \|T(\lambda_j)(v_j)\|$ for $\lambda_1 \approx 24$ (open circles), $\lambda_2 \approx 123$ (filled circles) versus the number N of quadrature nodes for the same example (right).

Example 4.12. Consider the quadratic polynomial

$$T(z) = T_0 + (z-a)(b-z)T_1, \quad a < b \in \mathbb{R}, \quad T_0, T_1 \in \mathbb{R}^{15,15}, \tag{48}$$

where T_0 has zeroes in the first column. All other entries of T_0, T_1 are chosen at random. Then $T(z)$ has different eigenvalues a and b with the same eigenvector $e^1 \in \mathbb{R}^m$. This is a critical case since the rank condition (18) is violated. In Fig. 5(left) we show the results of *polyeig* and of the integral algorithm (with $l = 5$ and the data from (47)).

There are three eigenvalues inside the circle. Both eigenvalues $a = -0.2$ and $b = 0.1$ are missed by the integral method, while the third one is found, though at lower accuracy than in the previous examples ($\mathrm{res}(\lambda_3) = 0.0183$, $e(\lambda_3) = 0.0189$ at $N = 150$). Fig. 5 shows that only one singular value stays of order one when N is increased. If we use $l = k = 3$ in Step 6 nevertheless, then we find a bad approximation of an eigenvalue outside ($\lambda_2 = 0.598$, $\mathrm{res}_2 = 1.07$) as well as a spurious eigenvalue ($\lambda_3 \approx 10^{10}$, $\mathrm{res}_3 \approx 10^{20}$). Both will be discarded by the residual as well as the contour test. Solving the generalized eigenvalue problem in this case (see Remark 3.2(d)) leads to very similar results. This example will be reconsidered in Section 5.

5. The algorithm for many eigenvalues

In this section we show how the method from Section 3 can be extended to nonlinear eigenvalue problems with more eigenvalues than the dimension of the system, i.e. $m < k$, and to the rank deficient cases, see Remark 3.2 and Example 4.12. Similar to Section 3, the approach in [1,2] differs from ours by multiplying the block Hankel matrices below with suitable test matrices from the left and then taking singular values.

5.1. Construction of algorithm

In case $m < k$ condition (18) is always violated and there is no matrix $\hat{V}$ satisfying (17). Therefore, we compute more integrals of type (19),(20), namely

$$A_p = \frac{1}{2\pi i}\int_\Gamma z^p T(z)^{-1}\hat{V}\,dz \in \mathbb{C}^{m,l}, \quad p \in \mathbb{N}.$$

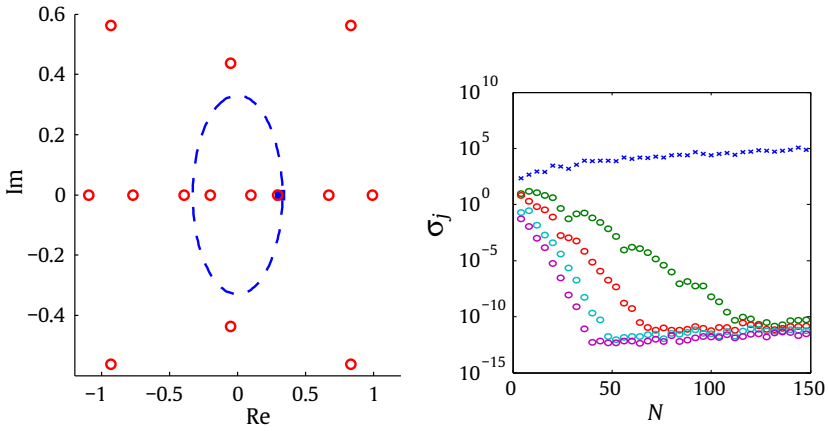

Fig. 5. Example 4.12: Eigenvalues from polyeig (open circles) and eigenvalues from the integral algorithm for a quadratic matrix polynomial with rank defect (left), singular values of integral algorithm with $l = 5$ columns versus the number N of quadrature nodes for the same example (right).

Here we assume that $\hat{V} \in \mathbb{C}^{m,l}$ with $l \leqslant m$. In fact, in case $k > m$ we set $\hat{V} = I_m$ instead of making a random choice.

From Theorem 2.9 we obtain

$$A_p = V\Lambda^p W^H \hat{V}, \quad p \in \mathbb{N}, \tag{49}$$

where $V, W \in \mathbb{C}^{m,k}$ are given by (27) and (29) and Λ has the normal form (30).

Now we choose $K \in \mathbb{N}$, $K \geqslant 1$ and form the $Km \times Kl$ matrices

$$B_0 = \begin{pmatrix} A_0 & \cdots & A_{K-1} \\ \vdots & & \vdots \\ A_{K-1} & \cdots & A_{2K-2} \end{pmatrix}, \quad B_1 = \begin{pmatrix} A_1 & \cdots & A_K \\ \vdots & & \vdots \\ A_K & \cdots & A_{2K-1} \end{pmatrix}. \tag{50}$$

From (49) we find the representations

$$B_0 = \begin{pmatrix} V \\ \vdots \\ V\Lambda^{K-1} \end{pmatrix} \left(W^H\hat{V} \ \cdots \ \Lambda^{K-1}W^H\hat{V} \right), \tag{51}$$

and

$$B_1 = \begin{pmatrix} V \\ \vdots \\ V\Lambda^{K-1} \end{pmatrix} \Lambda \left(W^H\hat{V} \ \cdots \ \Lambda^{K-1}W^H\hat{V} \right). \tag{52}$$

We assume that K has been chosen such that the following rank condition holds

$$\operatorname{rank} \begin{pmatrix} V \\ \vdots \\ V\Lambda^{K-1} \end{pmatrix} = k. \tag{53}$$

The smallest index having this property is called the minimality index in [20]. In case $k > m$ this can be expected to hold if we choose

$$(K-1)m < k \leqslant Km.$$

In case $k \leqslant m$ with $\operatorname{rank}(V) < k$ (see Remark 3.2(b)) the following lemma shows that (53) holds for K larger than the sum of the maximal ranks at all eigenvalues.

Lemma 5.1. *Let the assumptions of Corollary 2.8 be satisfied. Then the rank condition (53) holds with k as defined in (16) for*

$$K \geqslant \sum_{n=1}^{n(\mathcal{C})} \max_{1\leqslant \ell \leqslant L_n} m_{\ell,n}. \tag{54}$$

Proof. By the ordering (9) we have $m_{1,n} = \max_{1\leqslant \ell\leqslant L_n} m_{\ell,n}$ for all n. Assume that $V\Lambda^j x = 0$, $j = 0,\dots,K-1$ for K satisfying (54) and for some $x \in \mathbb{C}^k$. For any $n \in \{1,\dots,n(\mathcal{C})\}$ and $0 \leqslant \beta \leqslant m_{1,n}-1$ define the polynomial

$$P_{n,\beta}(z) = (z-\lambda_n)^\beta \prod_{r=1,r\neq n}^{n(\mathcal{C})} (z-\lambda_r)^{m_{1,r}}.$$

By our assumption these polynomials have at most degree $K-1$ and, therefore, satisfy $0 = VP_{n,\beta}(\Lambda)x$. We partition V into columns and x into blocks compatible with the Jordan structure (30)

$$V = \begin{pmatrix} V_1 & \cdots & V_{n(\mathcal{C})} \end{pmatrix}, \quad V_n = \begin{pmatrix} V_{1,n} & \cdots & V_{L_n,n} \end{pmatrix}, \quad V_{\ell,n} = \begin{pmatrix} v_0^{\ell,n} & \cdots & v_{m_{\ell,n}-1}^{\ell,n} \end{pmatrix}$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_{n(\mathcal{C})} \end{pmatrix}, \quad x_n = \begin{pmatrix} x_{1,n} \\ \vdots \\ x_{L_n,n} \end{pmatrix}, \quad x_{\ell,n} = \begin{pmatrix} x_0^{\ell,n} \\ \vdots \\ x_{m_{\ell,n}-1}^{\ell,n} \end{pmatrix}.$$

Using this in $0 = VP_{n,\beta}(\Lambda)x$ leads to

$$0 = \sum_{j=1}^{n(\mathcal{C})} V_j (J_j-\lambda_n)^\beta \prod_{r=1,r\neq n}^{n(\mathcal{C})} (J_j-\lambda_r)^{m_{1,r}} x_j.$$

Since $(J_r-\lambda_r)^{m_{1,r}} = 0$ we obtain

$$0 = V_n (J_n-\lambda_n)^\beta \prod_{r=1,r\neq n}^{n(\mathcal{C})} (J_n-\lambda_r)^{m_{1,r}} x_n.$$

Expanding into columns again and using $(J_{n,\ell}-\lambda_n)^\beta = 0$ for indices $m_{\ell,n} \leqslant \beta$ gives

$$0 = \sum_{\substack{\ell=1 \\ \beta\leqslant m_{\ell,n}-1}}^{L_n} V_{\ell,n} \left[\prod_{r=1,r\neq n}^{n(\mathcal{C})} (J_{n,\ell}-\lambda_r)^{m_{1,r}} \right] (J_{n,\ell}-\lambda_n)^\beta x_{\ell,n}. \tag{55}$$

We use this equation to prove for any given n and by induction on $\beta = m_{1,n}-1,\dots,0$ the following

$$x_\nu^{n,\ell} = 0, \text{ for indices } \beta \leqslant \nu \leqslant m_{\ell,n}-1, \quad \ell \in \{1,\dots,L_n\}. \tag{56}$$

For $\beta = m_{1,n} - 1$, Eq. (55) reads

$$\begin{aligned} 0 &= \sum_{\substack{\ell=1 \\ m_{\ell,n}=m_{1,n}}}^{L_n} \left(v_0^{\ell,n} \cdots v_{m_{\ell,n}-1}^{\ell,n}\right) \prod_{\substack{r=1 \\ r\neq n}}^{n(\mathcal{C})} (J_{n,\ell} - \lambda_r)^{m_{1,r}} \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_0^{\ell,n} \\ \vdots \\ x_{m_{\ell,n}-1}^{\ell,n} \end{pmatrix} \\ &= \prod_{\substack{r=1 \\ r\neq n}}^{n(\mathcal{C})} (\lambda_n - \lambda_r)^{m_{1,r}} \sum_{\substack{\ell=1 \\ m_{\ell,n}=m_{1,n}}}^{L_n} v_0^{\ell,n} x_{m_{\ell,n}-1}^{\ell,n}. \end{aligned}$$

Thus condition (56) follows for $\beta = m_{1,n} - 1$ from the linear independence of the vectors $v_0^{\ell,n}$ (cf. Definition 2.5(iv)). For the induction step we use (55) with $\beta - 1$ instead of β. Taking the hypothesis (56) into account we find in a similar way

$$\begin{aligned} 0 &= \sum_{\substack{\ell=1 \\ m_{\ell,n}=m_{1,n}}}^{L_n} \left(v_0^{\ell,n} \cdots v_{m_{\ell,n}-1}^{\ell,n}\right) \prod_{\substack{r=1 \\ r\neq n}}^{n(\mathcal{C})} (J_{n,\ell} - \lambda_r)^{m_{1,r}} \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & 0 & \ddots & \\ 0 & \cdots & 0 & \cdots & 1 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_0^{\ell,n} \\ \vdots \\ x_{\beta-1}^{\ell,n} \\ \vdots \\ 0 \end{pmatrix} \\ &= \prod_{r=1, r\neq n}^{n(\mathcal{C})} (\lambda_n - \lambda_r)^{m_{1,r}} \sum_{\substack{\ell=1 \\ m_{\ell,n}\geqslant\beta}}^{L_n} v_0^{\ell,n} x_{\beta-1}^{\ell,n}. \end{aligned}$$

Therefore, condition (56) holds for $\beta - 1$. In summary, we have shown $x = 0$ and this finishes the proof. □

The computational procedure is now a straightforward generalization of Section 3.1. First compute $B_0, B_1 \in \mathbb{C}^{Km,Kl}$ from (50). In addition to (53), assume

$$\operatorname{rank}\left(W^H\hat{V} \ \cdots \ \Lambda^{K-1}W^H\hat{V}\right) = k. \tag{57}$$

Let us abbreviate

$$V_{[K]} = \begin{pmatrix} V \\ \vdots \\ V\Lambda^{K-1} \end{pmatrix} \in \mathbb{C}^{Km,k}, \quad W_{[K]}^H = \left(W^H\hat{V} \ \cdots \ \Lambda^{K-1}W^H\hat{V}\right) \in \mathbb{C}^{k,Kl}.$$

Compute the SVD

$$V_{[K]}W_{[K]}^H = B_0 = V_0\Sigma_0W_0^H,$$

where $V_0 \in \mathbb{C}^{Km,k}$, $V_0^HV_0 = I_k$, $\Sigma_0 = \operatorname{diag}(\sigma_1, \ldots, \sigma_k) \in \mathbb{C}^{k,k}$, and $W_0 \in \mathbb{C}^{Kl,k}$, $W_0^HW_0 = I_k$. From the rank conditions (53), (57),

$$\sigma_1 \geqslant \cdots \sigma_k > 0 = \sigma_{k+1} = \cdots = \sigma_{Kl}.$$

The rank condition (53) also implies

$$R(B_0) = R(V_{[K]}) = R(V_0).$$

Thus the matrix $S = V_0^H V_{[K]} \in \mathbb{C}^{k,k}$ is nonsingular and satisfies

$$V_{[K]} = V_0 S. \tag{58}$$

With (51), (58) we find

$$W_{[K]}^H = S^{-1} \Sigma_0 W_0^H,$$

and then from (52)

$$B_1 = V_{[K]} \Lambda W_{[K]}^H = V_0 S \Lambda S^{-1} \Sigma_0 W_0^H.$$

Finally, this leads to

$$D := V_0^H B_1 W_0 \Sigma_0^{-1} = S \Lambda S^{-1}. \tag{59}$$

Therefore, the analog of Theorem 3.3 is

Theorem 5.2. *Suppose that $T \in H(\Omega, \mathbb{C}^{m,m})$ has no eigenvalues on the contour Γ in Ω and pairwise distinct eigenvalues λ_n, $n = 1, \ldots, n(\Gamma)$ inside Γ with partial multiplicities $m_{1,n} \geqslant \cdots \geqslant m_{L_n,n}$, $n = 1, \ldots, n(\Gamma)$. Assume that the rank conditions (53), (57) are satisfied with k given by (16). Then the matrix $D \in \mathbb{C}^{k,k}$ from (59) has Jordan normal form (30) with the same eigenvalues λ_n and partial multiplicities $m_{\ell,n}$ ($\ell = 1, \ldots, L_n$, $n = 1, \ldots, n(\Gamma)$). Suitable CSGEs for T can be obtained from corresponding CSGEs $s_j^{\ell,n}$ for D via*

$$v_j^{\ell,n} = V_0^{[1]} s_j^{\ell,n}, \quad 0 \leqslant j \leqslant m_{\ell,n} - 1, \, 1 \leqslant \ell \leqslant L_n, \, 1 \leqslant n \leqslant n(\Gamma),$$

where $V_0^{[1]}$ is the upper $m \times k$ block in

$$V_0 = \begin{pmatrix} V_0^{[1]} \\ \vdots \\ V_0^{[K]} \end{pmatrix}. \tag{60}$$

Remark 5.3. In a sense this generalization is similar to linearizing a polynomial eigenvalue problem by increasing the dimension. Note, however, that this only becomes necessary if there are too many eigenvalues inside the contour, or if rank defects occur that are not present in linear eigenvalue problems.

The generalization of the algorithm from Section 3.3 is the following.

Integral algorithm 2

Step 1: Choose numbers $l \leqslant m, K \geqslant 1$ and a matrix $\hat{V} \in \mathbb{C}^{m,l}$ at random. If more than m eigenvalues are expected inside Γ, let $l = m$, $\hat{V} = I_m$.

Step 2: Compute $A_{p,N} = \frac{1}{iN} \sum_{j=0}^{N-1} T(\varphi(t_j))^{-1} \hat{V} \varphi(t_j)^p \varphi'(t_j)$, $p = 0, \ldots, 2K-1$, and form $B_{0,N}$, $B_{1,N}$ as in (50).

Step 3: Compute the SVD $B_{0,N} = V \Sigma W^H$, where $V \in \mathbb{C}^{Km,Kl}$, $W \in \mathbb{C}^{Kl,Kl}$, $V^H V = W^H W = I_{Kl}$, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{Kl})$.

Step 4: Perform a rank test for Σ, i.e. find $0 < k \leqslant Kl$ such that $\sigma_1 \geqslant \cdots \geqslant \sigma_k > \sigma_{k+1} \approx \cdots \approx \sigma_{Kl} \approx 0$. If $k = Kl$ then increase l or K and go to Step 1. Else let $V_0 = V(1 : Km, 1 : k)$, $W_0 = W(1 : Kl, 1 : k)$ and $\Sigma_0 = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$.

Step 5: Compute $D = V_0^H B_{1,N} W_0 \Sigma_0^{-1} \in \mathbb{C}^{k,k}$.

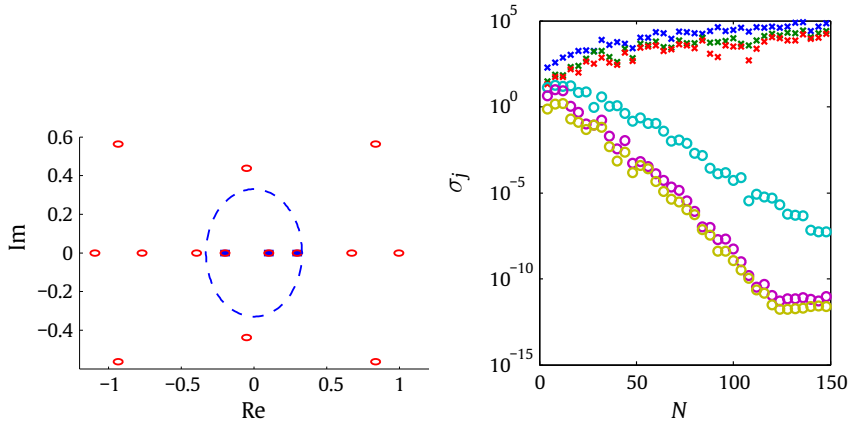

Fig. 6. Example 5.4: Eigenvalues from polyeig (open circles) and eigenvalues from the integral algorithm 2 ($K = 2$, filled boxes) for a quadratic matrix polynomial with rank defect (left), singular values of integral algorithm 2 with $l = 3$ columns versus the number N of quadrature nodes for the same example (right).

Step 6: Solve the eigenvalue problem for D. If all eigenvalues $\lambda_j, j = 1, \dots, k$ are well-conditioned with corresponding eigenvectors $s_j \in \mathbb{C}^k$ then accept λ_j and the eigenvector $v_j = V_0^{[1]} s_j$ provided $\lambda_j \in \mathrm{int}(\Gamma)$ and $\|T(\lambda_j)v_j\| \leqslant \mathrm{tol}_{\mathrm{res}}$. Otherwise compute a Schur decomposition $DQ = QT$ with Q unitary and T upper triangular. Reorder eigenvalues such that eigenvalues inside Γ occur first and discard eigenvalues outside Γ and corresponding columns of Q. Block diagonalize T such that diagonal blocks belong to different eigenvalues. Let λ_j be the diagonal entry of the jth block and let $s_j \in \mathbb{C}^k$ be the first column vector from the corresponding block. Accept $v_j = V_0^{[1]} s_j$ as eigenvector and λ_j as eigenvalue if $\|T(\lambda_j)v_j\| \leqslant \mathrm{tol}_{\mathrm{res}}$.

Similar remarks as in Remark 3.5 apply. In particular, it would be advantageous to have an approximation of higher order root functions that generalize the invariant pairs for matrix polynomials in [5].

5.2. Numerical examples

Example 5.4. We apply the Integral algorithm 2 to the rank deficient example (48), where $K = 2$, $l = 3$ and the contour is the circle from (47). Now the eigenvalues $a = -0.2$ and $b = 1$ are reproduced correctly (see Fig. 6(left)), and three singular values survive as expected (Fig. 6(right)).

Example 5.5. Consider the characteristic equation of a delay system $\dot{x} = T_0 x(t) + T_1 x(t - \tau)$ from [27,20, Section 2.4.2], given by

$$T(z) = zI - T_0 - T_1 e^{-z\tau}, \quad T_0 = \begin{pmatrix} -5 & 1 \\ 2 & -6 \end{pmatrix}, \quad T_1 = \begin{pmatrix} -2 & 1 \\ 4 & -1 \end{pmatrix}. \tag{61}$$

In case $\tau = 1$ there are more than two eigenvalues inside the circle $\varphi(t) = z_0 + Re^{it}, z_0 = -1, R = 6$. We set $l = 2, \hat{V} = I_2$ and $K = 3$ for the integral algorithm 2 and obtain with $N = 150$ five eigenvalues inside the circle, (see Fig. 7(left)), which coincide with the computed ones in [20]. Much smaller values than $N = 150$ give sufficient accuracy, since there is a good separation of singular values and a fast decay of residuals, see Fig. 7(right).

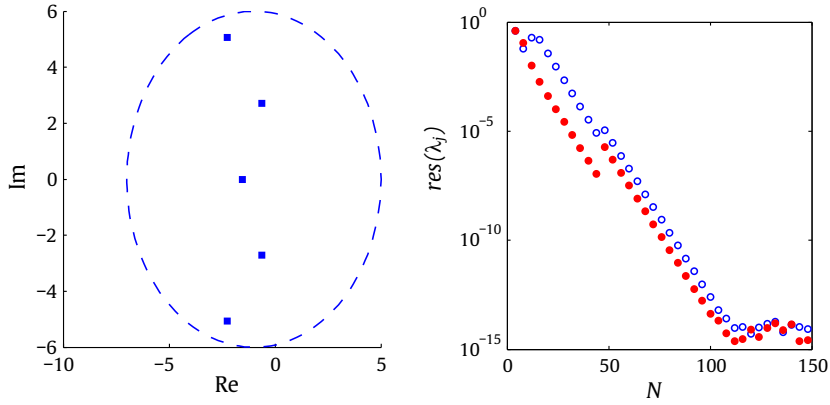

Fig. 7. Example 5.5: Eigenvalues of the characteristic equation (61) inside a circle of radius 6 and with center -1, computed with the integral algorithm 2 with $K = 3$, $l = 2$.(left), residuals $\|T(\lambda_j)v_j\|$ for $\lambda_1 \approx -0.6 + 2.71i$, $\lambda_2 \approx -2.27 + 5.07i$ versus the number N of quadrature nodes for the same example (right).

Note added in proof

One of the referees called my attention to the integral method of Asakura, Sakurai, Tadano, Ikegami and Kimura [2,1] that extended earlier work of Sakurai and Sugiura [29] on the generalized eigenvalue problem. The block method in [2] differs only slightly from ours in premultiplying the matrices A_p in (3) by another random matrix $\hat{U}^H \in \mathbb{C}^{k,m}$ before taking singular values.

The theoretical approach of [1,2], however, is quite different since it uses the Smith normal form rather than Keldysh' Theorem. So far, the results in [1,2] assume the eigenvalues inside the contour to be simple and nondegenerate (note that linear independence of eigenvectors for different eigenvalues is not automatic in the nonlinear case). Also the effect of quadrature errors on the eigenvalue computations is not considered in [1,2].

Acknowledgements

The author thanks Ingwar Petersen for the support with the numerical experiments. He is also grateful to the referees for several constructive criticisms and further references that improved the first version of the paper.

References

[1] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, K. Kimura, A numerical method for polynomial eigenvalue problems using contour integral, Technical Report CS-TR-08-15, Department of Computer Science, University of Tsukuba, 2008.
[2] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, K. Kimura, A numerical method for nonlinear eigenvalue problems using contour integrals, JSIAM Lett. 1 (2009) 52–55.
[3] Z. Bai, Y. Su, A second-order Arnoldi method for the solution of the quadratic eigenvalue problem, SIAM J. Matrix Anal. Appl. 26 (3) (2005) 640–659.
[4] T. Betcke, N.J. Higham, V. Mehrmann, C. Schröder, F. Tisseur, NLEVP: a collection of nonlinear eigenvalue problems, Technical Report 2008.40, MIMS, University of Manchester, April 2008.
[5] T. Betcke, D. Kressner, Perturbation, computation and refinement of invariant matrix pairs, Technical Report 2009-21, SAM, ETH Zürich, 2009.
[6] T. Betcke, H. Voss, A Jacobi–Davidson type projection method for nonlinear eigenvalue problems, Future Gener. Comput. Syst. 20 (2004) 363–372.
[7] D. Chu, G. Golub, On a generalized eigenvalue problem for nonsquare pencils, SIAM J. Matrix Anal. Appl. 28 (2006) 770–787.
[8] P.J. Davis, On the numerical integration of periodic analytic functions, in: R.E. Langer (Ed.), Proceedings of a Symposium On Numerical Approximation, Madison, April 21–23, 1958, Publication No. 1 of the Mathematics Research Center, US Army, the University of Wisconsin, The University of Wisconsin Press, Madison, 1959, pp. 45–59.
[9] P.J. Davis, P. Rabinowitz, Methods of Numerical Integration, Dover Publications Inc., Mineola, NY, 2007., Corrected reprint of the second (1984) edition.

[10] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, Academic Press Inc., New York, 1982.
[11] I. Gohberg, L. Rodman, Analytic matrix functions with prescribed local data, J. Anal. Math. 40 (1982) (1981) 90–128.
[12] I.C. Gohberg, E.I. Sigal, An operator generalization of the logarithmic residue theorem and Rouché's theorem, Mat. Sb. 84 (126) (1971) 607–629.
[13] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.
[14] R.E. Greene, S.G. Krantz, Function Theory of One Complex Variable, Graduate Studies in Mathematics, vol. 40, third ed., American Mathematical Society, Providence, RI, 2006.
[15] N. Hale, N.J. Higham, L. Trefethen, Computing A^{α}, $\log(A)$, and related matrix functions by contour integrals, SIAM J. Numer. Anal. 46 (2008) 2505–2523.
[16] N.J. Higham, Functions of Matrices, SIAM, 2008.
[17] M.V. Keldysh, On the characteristic values and characteristic functions of certain classes of non-self-adjoint equations, Dokl. Akad. Nauk SSSR (N.S.) 77 (1951) 11–14.
[18] M.V. Keldysh, The completeness of eigenfunctions of certain classes of nonselfadjoint linear operators, Uspehi Mat. Nauk 26 (4160) (1971) 15–41.
[19] S. Kirkup, S. Amini, Soluton of the Helmholtz eigenvalue problem via the boundary element method, Internat. J. Numer. Methods Engrg. 36 (1993) 321–330.
[20] D. Kressner, A block Newton method for nonlinear eigenvalue problems, Numer. Math. 114 (2009) 355–372.
[21] D.S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, Vector spaces of linearizations for matrix polynomials, SIAM J. Matrix Anal. Appl. 28 (4) (2006) 971–1004. (electronic).
[22] A.S. Markus, E.I. Sigal, The multiplicity of the characteristic number of an analytic operator function, Mat. Issled. 5 (3(17)) (1970) 129–147.
[23] K. Meerbergen, The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem, SIAM J. Matrix Anal. Appl. 30 (4) (2008) 1463–1482.
[24] V. Mehrmann, H. Voss, Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods, GAMM Mitt. 27 (2004)
[25] R. Mennicken, M. Möller, Root functions, eigenvectors, associated vectors and the inverse of a holomorphic operator function, Arch. Math. (Basel) 42 (5) (1984) 455–463.
[26] R. Mennicken, M. Möller, Non-Self-Adjoint Boundary Eigenvalue Problems, North-Holland Mathematics Studies, vol. 192, North-Holland Publishing Co., Amsterdam, 2003.
[27] W. Michiels, S.-I. Niculescu, Stability and Stabilization of Time-delay Systems, Advances in Design and Control, vol. 12, Society for Industrial and Applied Mathematics (SIAM), 2007.
[28] E. Polizzi, Density-matrix-based algorithms for solving eigenvalue problems, Phys. Rev. B (2009) 115112
[29] T. Sakurai, H. Sugiura, A projection method for generalized eigenvalue problems, J. Comput. Appl. Math. 159 (2003) 119–128.
[30] S. Solovev, Preconditioned iterative methods for a class of nonlinear eigenvalue problems, Linear Algebra Appl. 415 (2006) 210–229.
[31] G. Stewart, J.G. Sun, Matrix Perturbation Theory, Academic Press Inc., Boston, MA, 1990.
[32] V.P. Trofimov, The root subspaces of operators that depend analytically on a parameter, Mat. Issled. 3 (vyp. 3 (9)) (1968) 117–125.
[33] H. Voss, A maxmin principle for nonlinear eigenvalue problems with application to a rational spectral problem in fluid–solid vibration, Appl. Math 48 (2003) 607–622.
[34] H. Voss, An Arnoldi method for nonlinear eigenvalue problems, BIT 44 (2004) 387–401.
[35] H. Voss, A Jacobi–Davidson method for nonlinear and nonsymmetric eigenvalue problems, Comput. Struct. 85 (2007) 1284–1292.
[36] H. Voss, B. Werner, A minmax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems, Math. Methods Appl. Sci. 4 (1982) 415–424.