Technische Universität Graz

Convergence orders of iterative methods for nonlinear eigenvalue problems

O. Steinbach, G. Unger

Berichte aus dem Institut für Numerische Mathematik

Bericht 2011/1

Technische Universität Graz

Convergence orders of iterative methods for nonlinear eigenvalue problems

O. Steinbach, G. Unger

Berichte aus dem Institut für Numerische Mathematik

Bericht 2011/1

Technische Universität Graz
Institut für Numerische Mathematik
Steyrergasse 30
A 8010 Graz

WWW: `http://www.numerik.math.tu-graz.at`

Convergence orders of iterative methods for nonlinear eigenvalue problems

O. Steinbach[1], G. Unger[2]

[1]Institute of Computational Mathematics, TU Graz,
Steyrergasse 30, A 8010 Graz, Austria

o.steinbach@tugraz.at

[2]Johann Radon Institute of Computational and Applied Mathematics,
Altenberger Str. 69, A 4040 Linz, Austria

gerhard.unger@ricam.oeaw.ac.at

Abstract

The convergence analysis of iterative methods for nonlinear eigenvalue problems is in the most cases restricted either to algebraic simple eigenvalues or to polynomial eigenvalue problems. In this paper we consider two classical methods for general holomorphic eigenvalue problems, namely the nonlinear generalized Rayleigh quotient iteration (NGRQI) and the augmented Newton method. For both methods we prove local quadratic convergence for semi–simple eigenvalues. For defective eigenvalues local linear convergence is shown for the NGRQI. The key tool of our analysis is the representation of the eigenvalues as poles of the resolvent which is a classical result in operator theory. The convergence orders of the mentioned methods depend on the order of the poles of the resolvent. In numerical experiments the theoretical results are verified.

1 Introduction

We consider nonlinear eigenvalue problems for holomorphic functions $T : \Lambda \to \mathbb{C}^{n\times n}$, where $\Lambda \subset \mathbb{C}$ is a domain, of the following form: Find $\lambda \in \Lambda$ and $v \in \mathbb{C}^n \setminus \{0\}$ such that

$$T(\lambda)v = 0. \tag{1.1}$$

Typically, a pair (λ, v) which fulfills the eigenvalue problem (1.1), is called eigenpair, λ is called eigenvalue, and v eigenvector. In the following we assume $\det T(\cdot) \not\equiv 0$ on Λ. A comprehensive review about applications of nonlinear eigenvalue problems and their numerical solution is presented in [4, 19].

In this paper we focus on the refinement of already existing approximations of eigenpairs of nonlinear eigenvalue problems. For the approximate localization of eigenvalues recently the contour integral method was proposed [3, 5]. This method allows to find approximations of all eigenvalues in a given domaind and related eigenvectors without requiring initial approximations. In the case of general holomorphic eigenvalue problems the combination of the contour integral method with refinement methods is a reasonable approach.

For the refinement of approximations of eigenpairs usually iterative methods are proposed which are based on the solution of a sequence of linearized problems, such as the nonlinear generalized Rayleigh quotient iteration (NGRQI) [15, 16, 17, 27], augmented Newton–type methods [1, 2, 22, 23, 26, 34], the method of successive linear problems [25], or methods which utilize QR decompositions [6, 14, 18]. For comprehensive reviews we refer to [19, 26].

In this paper we present a convergence analysis for the nonlinear generalized Rayleigh quotient iteration and for an augmented Newton–type method. For both methods we prove that they have a local quadratic convergence order for semi–simple eigenvalues. For defective eigenvalues we show that the NGRQI is locally linear convergent. As main theoretical tool for our analysis we use the theory of eigenvalue problems for holomorphic Fredholm operator–valued functions [8, 13, 20]. This concept provides an extension of the theory of linear eigenvalue problems and it is based on the characterization of the eigenvalues as poles of the resolvent.

The NGRQI was introduced for polynomial eigenvalue problems in [16] as generalization of the two–sided Rayleigh quotient iteration [24], and local quadratic convergence was shown for semi–simple eigenvalues. In [15], this result was extended to polynomial eigenvalue problems in arbitrarily dimensional Hilbert spaces. For general holomorphic eigenvalue problems the NGRQI was analyzed in [17] under the assumption that the eigenvalues are isolated and that they can be locally characterized as poles of the resolvent. For eigenvalue problems with holomorphic matrix–valued functions $T : \Lambda \to \mathbb{C}^{n\times n}$ this assumption is always satisfied if $\det T(\cdot) \not\equiv 0$ on Λ, see, e.g., [7]. The convergence rate of the NGRQI depends on the order which an eigenvalue has as pole of the resolvent [17]. Since semi–simple eigenvalues are simple poles of the resolvent, the NGRQI converges locally quadratically. In the defective case local linear convergence is obtained. A modified version of the NGRQI is suggested and analyzed in [27, 29] where for algebraic simple eigenvalues local quadratic convergence is shown [27].

There are several variants of augmented Newton–type methods for nonlinear eigenvalue problems available. The classical approach is to apply Newton's method to the system of nonlinear equations consisting of the nonlinear eigenvalue problem and of an additional normalization condition for the eigenvector [1, 2, 34]. This approach is called the augmented Newton method or the inverse iteration for nonlinear eigenvalue problems. Different modifications of this approach are suggested in order to reduce the costs of the computations [22, 26], and to increase the convergence rate of the iteration [23, 25, 26]. In all of the mentioned publications the convergence results for the augmented Newton–type methods are restricted to algebraic simple eigenvalues. In this case, the derivative of the augmented form is non–singular at an eigenvalue and the classical argument of the proof for the convergence of Newton's method can be applied. If the algebraic multiplicity of an eigenvalue is not simple, this argumentation is not possible since the derivative of the augmented form is singular at the eigenvalue. In this paper we show that for semi–simple eigenvalues local quadratic convergence is still obtained where we utilize that the eigenvalues are simple poles of the resolvent. The defective case will be not considered. In [10], the convergence factors for semi–simple and for double non–semi–simple eigenvalues are analyzed. However, the existence of the convergence is assumed.

This paper is organized as follows. In the next section we outline the concept of eigenvalue problems for holomorphic matrix–valued functions and present some important characterizations and properties of the eigenvalues, the eigenvectors, and the resolvent. In Section 3 we review the derivation and the convergence analysis of the NGRQI. The augmented Newton method is analyzed in Section 4 where it is shown that it converges locally quadratically in the case of semi–simple eigenvalues. Finally, we present some numerical experiments and demonstrate the applicability of the NGRQI and the augmented Newton method in combination with the contour integral method.

In this paper we will use $(\cdot,\cdot)$ as standard inner product, i.e., $(x,y) := y^H x$ for all $x, y \in \mathbb{C}^n$, which generates the Euclidean norm $\|x\| := \sqrt{(x,x)}$. The norm for matrices is always the subordinated spectral norm.

2 Basics of holomorphic eigenvalue problems

In this section we introduce notations and properties of eigenvalue problems for holomorphic Fredholm operator–valued functions [13, 20]. Here we restrict our presentation to the case of matrix–valued functions. We denote by $\sigma(T)$ the set of all eigenvalues of T in the domain Λ, and by $\rho(T) = \Lambda \setminus \sigma(T)$ the resolvent set. Recall that we assume $\det T(\cdot) \not\equiv 0$ on Λ which implies that the resolvent set $\rho(T)$ is not empty. The dimension of the nullspace $\ker T(\lambda)$ of an eigenvalue λ is called the geometric multiplicity of λ. An ordered collection of vectors $v_{0,1}, v_{0,2}, \ldots, v_{0,m}$ in $\mathbb{C}^n$ is a Jordan chain of λ of length m if $v_{0,1}$ is an eigenvector corresponding to λ and if

$$\sum_{j=0}^{k-1} \frac{1}{j!} T^{(j)}(\lambda) v_{0,k-j} = 0 \quad \text{for } k = 1, \ldots, m \tag{2.1}$$

is satisfied, where $T^{(j)}$ is the j–th derivative. The maximal length of a Jordan chain of an eigenvalue λ is denoted by $\varkappa(T, \lambda)$. An eigenvalue λ is called semi–simple if the maximal length of a Jordan chain of λ is one, otherwise it is called defective. If in addition the geometric multiplicity of an eigenvalue is one then it is called an algebraic simple eigenvalue. This definition of an algebraic simple eigenvalue is equivalent to the common one that $\det T(\lambda) = 0$ [13]. Other equivalent definitions of the multiplicities are possible by using the Smith form [13], or by using root functions [13, 20].

The first result shows that the resolvent $T(\cdot)^{-1} : \Lambda \setminus \sigma(T) \to \mathbb{C}^{n\times n}$ can be represented as a meromorphic function where the eigenvalues are the poles. The order of the poles coincides with the maximal length of the Jordan chains of the eigenvalues.

Theorem 2.1. [7, Cor.8.4], [32] *Let Λ be an open subset of $\mathbb{C}$ and let $T : \Lambda \to \mathbb{C}^{n\times n}$ be a holomorphic matrix–valued function with $\det T(\cdot) \not\equiv 0$. Then, every eigenvalue λ of T is isolated, i.e., there is some neighborhood U of λ such that $U \setminus \{\lambda\} \subset \rho(T)$. Moreover, the resolvent admits a representation as*

$$T(\mu)^{-1} = \sum_{k=-r}^{-1} (\mu-\lambda)^k B_k + F(\lambda), \quad \mu \in U \setminus \{\lambda\}, \tag{2.2}$$

with $B_{-r} \neq 0$, where $r = \varkappa(T,\lambda)$ and $F : \Lambda \to \mathbb{C}^{n\times n}$ is holomorphic.

A characterization of the matrices B_k of the principal part of the resolvent (2.2) in terms of generalized eigenvectors of T and of the adjoint matrix function T^H provides the Theorem of Keldysh [12], [13, Thm. A.10.2]. The adjoint function $T^H : \{\lambda : \overline{\lambda} \in \Lambda\} \to \mathbb{C}^{n\times n}$ is defined by

$$T^H(\lambda) := (T(\overline{\lambda}))^H.$$

We cite the Theorem of Kelydsh for semi–simple eigenvalues which we need in the following. For the general version we refer to [13, Thm. A.10.2].

Theorem 2.2. [13, Thm. A.10.1] *Let the assumptions of Theorem 2.1 be satisfied. Suppose that $\lambda \in \sigma(T)$ is semi–simple and that $\{v^1,\dots,v^J\}$ is a basis of the eigenspace $\ker T(\lambda)$. Then there exists a unique basis $\{w^1,\dots,w^J\}$ of $\ker T^H(\overline{\lambda})$ such that in some neighborhood U of λ*

$$T(\mu)^{-1} = \sum_{j=1}^{J} \frac{1}{\mu-\lambda}(\,\cdot\,, w^j)v^j + F(\mu), \quad \mu \in U \setminus \{\lambda\}, \tag{2.3}$$

where $F : U \to \mathbb{C}^{n\times n}$ is holomorphic. Moreover, the following biorthogonality relation

$$\frac{1}{\mu-\lambda}(T(\mu)v^k, w^j) = \delta_{kj} + \mathcal{O}(\mu-\lambda) \quad \text{as } \mu \to \lambda \tag{2.4}$$

holds for $k, j = 1,\dots,J$.

From the representation (2.3) of the resolvent and the biorthogonality relation (2.4) some important properties for the derivative $T'(\lambda)$ and the eigenvectors follow. These results are needed later for the analysis of the augmented Newton method.

Corollary 2.3. *Let the assumptions of Theorem 2.1 be satisfied. Suppose that λ is a semi–simple eigenvalue and let*

$$\{v^1,\dots,v^J\} \quad \text{and} \quad \{w^1,\dots,w^J\}$$

be a basis of the eigenspaces $\ker T(\lambda)$ and $\ker T^H(\overline{\lambda})$, respectively, such that the resolvent $T(\mu)^{-1}$ admits the representation (2.3). Then:

i. For $k, j = 1,\dots,J$ we have

$$(T'(\lambda)v^k, w^j) = \delta_{kj}. \tag{2.5}$$

ii.

$$\sum_{j=1}^{J}(T'(\lambda)v, w^j)v^j = v \quad \textit{for all } v \in \ker T(\lambda). \tag{2.6}$$

iii. For F as given in (2.3),

$$T(\lambda)F(\lambda)T'(\lambda)v^j = 0, \quad j = 1, \ldots, J,$$

holds.

Proof.

i. Inserting $T(\lambda)v^k = 0$ in (2.4) gives

$$\frac{1}{\mu - \lambda}([T(\mu) - T(\lambda)]v^k, w^j) = \delta_{kj} + \mathcal{O}(\lambda - \mu) \quad \text{as } \mu \to \lambda.$$

Taking the limit $\mu \to \lambda$ yields the assertion.

ii. follows immediately from (2.5).

iii. By using (2.3) we have

$$F(\mu) = T(\mu)^{-1} - \sum_{j=1}^{J} \frac{1}{\mu - \lambda}(\cdot, w^j)v^j.$$

Multiplying by $T(\mu)$ and adding $\sum_{j=1}^{J} \frac{1}{\mu - \lambda}(\cdot, w^j)T(\lambda)v^j = 0$ this gives

$$\begin{aligned} T(\mu)F(\mu) &= I_n - \sum_{j=1}^{J} \frac{1}{\mu - \lambda}(\cdot, w^j)[T(\mu) - T(\lambda)]v^j \\ &\to I_n - \sum_{j=1}^{J}(\cdot, w^j)T'(\lambda)v^j \quad \text{as } \mu \to \lambda. \end{aligned}$$

With *i.* we obtain $T(\lambda)F(\lambda)T'(\lambda)v^j = 0$ for $j = 1, \ldots, J$.

□

3 Nonlinear generalized Rayleigh quotient iteration

In this section we review the NGRQI and present the derivation and the convergence properties where we follow [17]. This approach is based on the construction and analysis of a scalar function ψ which has the eigenvalues as zeros.

Let $\lambda \in \sigma(T)$ be an arbitrary but fixed eigenvalue of (1.1). By Theorem 2.1, there exists a neighborhood U of λ such that

$$T(\mu)^{-1} = \sum_{k=-r}^{\infty} (\mu - \lambda)^k B_k \quad \text{for } \mu \in U \setminus \{\lambda\},$$

with $B_{-r} \neq 0$. Let $a, b \in \mathbb{C}^n$ arbitrary but fixed vectors with $\|a\| = \|b\| = 1$ and which satisfy

$$(B_{-r}a, b) \neq 0. \tag{3.1}$$

Define the function $\varphi : U \to \mathbb{C}$ by

$$\varphi(\mu) := (T(\mu)^{-1}a, b),$$

then φ is obviously holomorphic on $U \setminus \{\lambda\}$. Since $|\varphi(\mu)| \to \infty$ as $\mu \to \lambda$, there exists a neighborhood $U_1 \subset U$ of λ and a constant $c_1 > 0$ such that

$$c_1 \leq |\varphi(\mu)| \quad \text{for all } \mu \in U_1 \setminus \{\lambda\}.$$

Hence, we may define the function $\psi : U_1 \to \mathbb{C}$ by

$$\psi(\mu) := \begin{cases} \dfrac{1}{\varphi(\mu)} & \text{for } \mu \neq \lambda, \\ 0 & \text{for } \mu = \lambda. \end{cases} \tag{3.2}$$

The function ψ is holomorphic on U_1 and allows the Taylor series expansion

$$\psi(\mu) = \frac{(\mu - \lambda)^r}{(B_{-r}a, b)} - (\mu - \lambda)^{r+1} \frac{(B_{-r+1}a, b)}{(B_{-r}a, b)^2} + \mathcal{O}\left((\mu - \lambda)^{r+2}\right),$$

which shows that λ is a zero of ψ with multiplicity r. This characterization of the eigenvalues as zero of a holomorphic scalar function is the essential idea of the approach in [17]. The use of Newton's method to determine the zero of the function ψ yields the NGRQI. By using the Banach fixed point theorem the following convergence result follows immediately.

Theorem 3.1 ([17, Satz 3]). *Let $s \in \mathbb{N}$, where $s \leq r = \varkappa(T, \lambda)$. Then, there exists a $\delta > 0$ such that the iteration*

$$\lambda_{i+1} = \lambda_i - s\frac{\psi(\lambda_i)}{\psi'(\lambda_i)} \quad \textit{for } i = 0, 1, 2, \ldots \tag{3.3}$$

converges for any $\lambda_0 \in U_\delta(\lambda) \setminus \{\lambda\}$ *to* λ. *If* $s = r$, *then the convergence is quadratic and we have*

$$\frac{\lambda_{i+1} - \lambda}{(\lambda_i - \lambda)^2} \to \frac{(B_{-r+1}a, b)}{r(B_{-r}a, b)} \quad \textit{as } i \to \infty. \tag{3.4}$$

If $s < r$, *then the convergence is linear and we have*

$$\frac{\lambda_{i+1} - \lambda}{\lambda_i - \lambda} \to \frac{r - s}{r} \quad \textit{as } i \to \infty. \tag{3.5}$$

If $\varkappa(T, \lambda)$ is not known a priori, then $s = 1$ can be chosen for the iteration (3.3), which yields the classical Newton's method for $\psi(\lambda) = 0$ and which ensures at least local linear convergence. For a semi–simple eigenvalue the choice $s = 1$ gives quadratic convergence.

Let us now consider the computational steps for the NGRQI. Recalling the definition (3.2) of ψ,

$$\psi(\mu) = \frac{1}{\varphi(\mu)} = \frac{1}{(T(\mu)^{-1}a, b)} \quad \text{for } \mu \in U_1 \setminus \{\lambda\},$$

we get, by using the representation of

$$\frac{d}{d\mu} T(\mu)^{-1} = -T(\mu)^{-1} T'(\mu) T(\mu)^{-1},$$

see, e.g. [11, p. 32],

$$\frac{\psi(\mu)}{\psi'(\mu)} = \frac{(T(\mu)^{-1}a, b)}{(T(\mu)^{-1}T'(\mu)T(\mu)^{-1}a, b)} = \frac{(a, [T(\mu)^{-1}]^H b)}{(T'(\mu)T(\mu)^{-1}a, [T(\mu)^{-1}]^H b)}.$$

Let us denote by $v_i \in \mathbb{C}^n$ and $w_i \in \mathbb{C}^n$ the solutions of

$$T(\lambda_i) v_i = a \quad \text{and} \quad T(\lambda_i)^H w_i = b, \tag{3.6}$$

then we can write the iteration (3.3) as

$$\lambda_{i+1} = \lambda_i - \frac{(T(\lambda_i)v_i, w_i)}{(T'(\lambda_i)v_i, w_i)}, \tag{3.7}$$

where we have set $s = 1$. In this form the NGRQI was introduced in [16] and the right hand side of (3.7) is called the generalized Rayleigh quotient of (λ_i, v_i, w_i). An analysis of the generalized Rayleigh quotient as a functional and a comparison with other Rayleigh functionals are presented in [28].

The NGRQI approximates a right and left eigenvector by v_i and w_i, respectively. Here, we cite the result for a right eigenvector. The same holds for a left eigenvector.

Lemma 3.2 ([17, Satz 4]). *Let* v_i *be defined by (3.6) and* λ_i *by (3.7). Then, there exists a* $i_0 \in \mathbb{N}$ *such that*

$$\inf_{v \in \ker T(\lambda)} \left\| v - \frac{v_i}{\|v_i\|} \right\| \le c\,|\lambda_i - \lambda|$$

for all $i \ge i_0$, *where* $c > 0$ *is a constant which is independent of* i.

For Hermitian eigenvalue problems, the choice $b = a$ as input vectors for the NGRQI is suggested since then only one system of linear equations has to be solved in each iteration step. An analogous simplification of the iteration is obtained in the case of complex symmetric eigenvalue problems, i.e., if $T(\lambda) = T(\lambda)^\top$. Provided that a and b are chosen such that $b = \overline{a}$, the solution w_i of the second equation in (3.6) is the complex conjugate of the solution v_i of the first equation, since

$$T(\lambda_i)v_i = a \quad \Leftrightarrow \quad T(\lambda_i)^\top v_i = a \quad \Leftrightarrow \quad \overline{T(\lambda_i)^\top v_i} = \overline{a} \quad \Leftrightarrow \quad T(\lambda_i)^H \overline{v_i} = \overline{a}.$$

In practical applications bordered systems for the update of λ_{i+1} are used to minimize rounding errors and ensure stability [27]. Let us consider the bordered systems

$$\begin{pmatrix} T(\lambda_i) & a \\ b^H & 0 \end{pmatrix} \begin{pmatrix} s_i \\ \mu_i \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha \end{pmatrix}, \quad \begin{pmatrix} T(\lambda_i)^H & b \\ a^H & 0 \end{pmatrix} \begin{pmatrix} t_i \\ \nu_i \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha \end{pmatrix}, \tag{3.8}$$

where $\alpha \in \mathbb{R} \setminus \{0\}$ is a scaling factor. Then we obtain $\mu_i v_i = -s_i$ and $\nu_i w_i = -t_i$. Using that $\mu_i = \overline{\nu_i}$, one gets for the update

$$\lambda_{i+1} = \lambda_i - \frac{(T(\lambda_i)s_i, t_i)}{(T'(\lambda_i)s_i, t_i)}.$$

In [27, 29] a modified version of the NGRQI was proposed where the vector a and b are updated in every iteration step by w_i and v_i, respectively. The motivation for this modification is that in the case of an algebraic simple eigenvalue λ the norm of the inverse of the bordered matrices in (3.8) at λ are minimized if for a the corresponding left eigenvector and for b the corresponding right eigenvector is chosen [27]. For this modified version local quadratic convergence was shown for algebraic simple eigenvalues in [27].

From a theoretical point of view the convergence order of the NGRQI for semi–simple eigenvalues does not depend on their geometric multiplicity. But there is an important difference between an eigenvalue with simple geometric multiplicity and an eigenvalue with multiple multiplicity with respect to the conditioning of the linear systems (3.8). For multiple eigenvalues the linear systems get very ill–conditioned close to an eigenvalue and they are singular in the limiting case. However, the error which is made due to this ill–conditioning points in the direction of the eigenspace and only slightly effects the performance of the algorithm. In numerical experiments still a quadratic convergence behavior for semi–simple eigenvalues with multiple geometric multiplicity is obtained, see Examples 5.1 and 5.3.

4 Augmented Newton method

One of the classical approaches for the solution of the nonlinear eigenvalue problem (1.1) is to apply Newton's method to the augmented system

$$F(v, \lambda) := \begin{pmatrix} T(\lambda)v \\ d^H v - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{4.1}$$

where the second equation is a normalization constraint with some chosen vector $d \in \mathbb{C}^n \setminus \{0\}$ for which it is assumed that it is not orthogonal to the eigenspace $\ker T(\lambda)$. In many publications [2, 23, 25, 26, 35], this approach was analyzed for algebraic simple eigenvalues. Utilizing that the derivative of the augmented form

$$F'(v,\lambda) = \begin{pmatrix} T(\lambda) & T'(\lambda)v \\ d^H & 0 \end{pmatrix} \tag{4.2}$$

for an eigenpair (λ, v) of an algebraic simple eigenvalue is non–singular, the standard arguments of the proof for the convergence of Newton's method can be applied to show local quadratic convergence. Different modifications are proposed in order to reduce the costs of the computations [22, 26] and to increase the convergence rate of the iteration [23, 25, 26]. However, the convergence analysis in all of these publications are restricted to algebraic simple eigenvalues and are based on the regularity of the derivative of the augmented form (4.2) which is for a multiple eigenvalue not regular anymore.

For our theoretical analysis it is suitable to write the augmented Newton method in the form as given in Algorithm 1. However, in practical computations it is recommended to perform it in the augmented form by reasons of rounding errors and stability.

Algorithm 1 Augmented Newton method

1: **Input:** λ_0, v_0, d such that $d^H v_0 = 1$
2: **for** $i = 0, 1, 2, \ldots$ until convergence **do**
3: solve $T(\lambda_i)s_{i+1} = T'(\lambda_i)v_i$ for s_{i+1}
4: $\lambda_{i+1} = \lambda_i - (d^H v_i)/(d^H s_{i+1})$
5: $v_{i+1} = s_{i+1}/d^H s_{i+1}$
6: **end for**

In the following we will consider an algorithm which is similar to Algorithm 1 but where we choose a normalization condition which needs no assumption in advance. For d in step 4 of Algorithm 1 we choose $d = s_{i+1}$ and in step 5 we normalize $v_{i+1} = s_{i+1}/\|s_{i+1}\|$ which gives Algorithm 2. Note that for linear eigenvalue problems with $T(\lambda) = A - \lambda I$ Algorithm 2 is the Rayleigh quotient iteration.

For a given approximation (λ_i, v_i) of an eigenpair $(\lambda_i + \widetilde{\Delta}\lambda_i, v_i + \widetilde{\Delta}v_i)$ with $\|v_i\| = 1$ one step of Algorithm 2 can be interpreted as linearization of

$$\begin{aligned} 0 &= [(v_i + \widetilde{\Delta}v_i)^H v_i] T(\lambda_i + \widetilde{\Delta}\lambda_i)(v_i + \widetilde{\Delta}v_i) \\ &= [(v_i + \widetilde{\Delta}v_i)^H v_i] T(\lambda_i)v_i + \widetilde{\Delta}\lambda_i T'(\lambda_i)v_i + [(v_i + \widetilde{\Delta}v_i)^H v_i] T(\lambda_i)\widetilde{\Delta}v_i + r(\widetilde{\Delta}\lambda_i, \widetilde{\Delta}v_i), \end{aligned}$$

where $r(\widetilde{\Delta}\lambda_i, \widetilde{\Delta}v_i)$ contains only terms of at least quadratic order with respect to $\widetilde{\Delta}\lambda_i$ and $\widetilde{\Delta}v_i$. Neglecting $r(\widetilde{\Delta}\lambda_i, \widetilde{\Delta}v_i)$ yields the equation

$$[(v_i + \Delta v_i)^H v_i] T(\lambda_i)(v_i + \Delta v_i) = -\Delta\lambda_i T'(\lambda_i)v_i$$

for the new corrections $\Delta\lambda_i = \lambda_{i+1} - \lambda_i$ and $\Delta v_i = v_{i+1} - v_i$. Suppose that s_{i+1} is a solution of $T(\lambda_i)s_{i+1} = T'(\lambda_i)v_i$, then $[(v_i + \Delta v_i)^H v_i](v_i + \Delta v_i) = -\Delta\lambda_i s_{i+1}$. Enforcing the normalization condition $\|v_i + \Delta v_i\| = 1$ gives $v_i + \Delta v_i = s_{i+1}/\|s_{i+1}\|$, and finally

$$-\Delta\lambda_i s_{i+1}^H s_{i+1} = [(v_i + \Delta v_i)^H v_i] s_{i+1}^H (v_i + \Delta v_i) = \frac{s_{i+1}^H v_i}{\|s_{i+1}\|} \frac{s_{i+1}^H s_{i+1}}{\|s_{i+1}\|} = s_{i+1}^H v_i$$

which corresponds to step 4 of Algorithm 2.

Algorithm 2 Modified augmented Newton method

1: **Input:** λ_0, v_0 such that $v_0^H v_0 = 1$
2: **for** $i = 0, 1, 2, \ldots$ until convergence **do**
3: solve $T(\lambda_i)s_{i+1} = T'(\lambda_i)v_i$ for s_{i+1}
4: $\lambda_{i+1} = \lambda_i - (s_{i+1}^H v_i)/(s_{i+1}^H s_{i+1})$
5: $v_{i+1} = s_{i+1}/\|s_{i+1}\|$
6: **end for**

In the following we will present a rigorous convergence analysis of Algorithm 2 in the case of semi–simple eigenvalues. Let (μ, z) be an approximation of an eigenpair (λ, v). We first derive an error estimate for the new eigenvector approximation of Algorithm 2 with respect to the errors of μ and z.

Lemma 4.1. *Let $\lambda \in \sigma(T)$ be a semi–simple eigenvalue. Then, there exist $\varepsilon > 0$, $\tau > 0$, and $c_\lambda > 0$ such that for all*

$$z \in \left\{ y \in \mathbb{C}^n : \|y\| = 1, \min_{v \in \ker T(\lambda)} \|y - v\| \le \varepsilon \right\}$$

and for all μ with $0 < |\mu - \lambda| \le \delta$

$$\min_{v \in \ker T(\lambda)} \left\| \frac{1}{\|T(\mu)^{-1}T'(\mu)z\|} T(\mu)^{-1}T'(\mu)z - v \right\| \le c_\lambda |\mu - \lambda| \big(|\mu - \lambda| + \|z - v_z\| \big) \qquad (4.3)$$

holds, where v_z is the best approximation of z in $\ker T(\lambda)$.

Proof. Let $\lambda \in \sigma(T)$ be a semi–simple eigenvalue and let $\{v^1, \ldots, v^J\}$ be a basis of the eigenspace $\ker T(\lambda)$. By Theorem 2.2, there exists a neighborhood U of λ such that $T(\mu)^{-1}$ admits a representation by

$$T(\mu)^{-1} = (\mu - \lambda)^{-1} B_{-1} + F(\mu), \qquad \mu \in U \setminus \{\lambda\}, \qquad (4.4)$$

with a holomorphic function $F : U \to \mathbb{C}^{n \times n}$ and with

$$B_{-1} = \sum_{j=1}^{J} (\,\cdot\,, w^j) v^j, \qquad (4.5)$$

where $\{w^1, \ldots, w^J\}$ is an appropriate basis of $\ker T^H(\overline{\lambda})$. Choose $\tau_* > 0$ such that the closed disk $\overline{U}_{\tau_*}(\lambda)$ with center λ and radius τ_* is a subset of U. For $z \in \mathbb{C}^n$ with $\|z\| = 1$ we denote by $s(\mu, z)$ the solution of

$$T(\mu)s(\mu, z) = T'(\mu)z.$$

Using the representation (4.4), we can write

$$s(\mu, z) = \frac{1}{(\mu - \lambda)} B_{-1}T'(\mu)z + F(\mu)T'(\mu)z, \quad \mu \in U \setminus \{\lambda\}. \tag{4.6}$$

For the norm of $s(\mu, z)$ we get

$$\begin{aligned} \|s(\mu, z)\|^2 &= (s(\mu, z), s(\mu, z)) \\ &= \frac{\|B_{-1}T'(\mu)z\|^2}{|\mu - \lambda|^2} + 2\,\mathrm{Re}\frac{(B_{-1}T'(\mu)z, F(\mu)T'(\mu)z)}{(\mu - \lambda)} + \|F(\mu)T'(\mu)z\|^2 \\ &= \frac{\|B_{-1}T'(\mu)z\|^2}{|\mu - \lambda|^2}\chi(\mu, z), \end{aligned} \tag{4.7}$$

with

$$\chi(\mu, z) := 1 + \frac{2\mathrm{Re}[(\mu - \lambda)(B_{-1}T'(\mu)z, F(\mu)T'(\mu)z)]}{\|B_{-1}T'(\mu)z\|^2} + \frac{|\mu - \lambda|^2\|F(\mu)T'(\mu)z\|}{\|B_{-1}T'(\mu)z\|^2}.$$

We first show that the function χ is well defined in a neighborhood of λ provided that z is sufficiently close to the eigenspace $\ker T(\lambda)$. Let $v_z \in \mathbb{C}^n$ be the best approximation of z in $\ker T(\lambda)$ and let

$$\delta z := z - v_z.$$

Using the Taylor series expansion of $T'(\mu)$ in λ and the property (2.6), i.e. $B_{-1}T'(\lambda)v = v$ for all $v \in \ker T(\lambda)$, we get

$$\begin{aligned} B_{-1}T'(\mu)z &= B_{-1}T'(\mu)(v_z + \delta z) \\ &= B_{-1}T'(\lambda)v_z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z \\ &= v_z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z \end{aligned} \tag{4.8}$$

for some $\widetilde{\mu} \in U_{\tau_*}(\lambda)$. With $v_z = z - \delta z$ we have

$$\begin{aligned} \|B_{-1}T'(\mu)z\|^2 &= \|z - \delta z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z\|^2 \\ &= (z, z) - (z, \delta z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z) \\ &\quad +(\delta z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z, B_{-1}T'(\mu)z). \end{aligned}$$

Since $B_{-1}T'(\mu)$ and $B_{-1}T''(\mu)$ are bounded in $\overline{U}_{\tau_*}(\lambda)$, and since $\|z\| = 1$, $\|v_z\| \leq 1$, there exists a constant $c_{\tau_*} > 0$ such that

$$\|B_{-1}T'(\mu)z\|^2 \leq 1 + c_{\tau_*}(\|\delta z\| + |\mu - \lambda|) \tag{4.9}$$

for all $\mu \in \overline{U}_{\tau_*}(\lambda)$. This implies that there exists a $\tau_1 > 0$ such that the function $\chi(\mu, z)$ is well defined for all μ with $|\mu - \lambda| \leq \tau_1$, and for all z with $\|\delta z\| \leq \tau_1$.

Since $B_{-1}T'(\mu)z$ and $F(\mu)T'(\mu)z$ are uniformly bounded on the compact set $\overline{U}_{\tau_1}(\lambda) \times \{z \in \mathbb{C}^n : \|z\| = 1\}$, we can write

$$\|B_{-1}T'(\mu)z\|^2\,\chi(\mu, z) = 1 + \nu(\mu, z), \tag{4.10}$$

where

$$|\nu(\mu, z)| \leq c_\nu(\|\delta z\| + |\mu - \lambda|) \tag{4.11}$$

for all μ with $|\mu - \lambda| \leq \tau_1$ and for all $\|\delta z\| \leq \tau_1$.

Using (4.6) and (4.7) we obtain

$$\frac{s(\mu, z)}{\|s(\mu, z)\|} = \frac{|\mu - \lambda|}{\|B_{-1}T'(\mu)z\|\,\chi(\mu, z)^{1/2}}\left((\mu - \lambda)^{-1}B_{-1}T'(\mu)z + F(\mu)T'(\mu)z\right). \tag{4.12}$$

The vector $B_{-1}T'(\mu)z$ is an element of the eigenspace $\ker T(\lambda)$ due to the construction of B_{-1}, see (4.5). The Taylor series expansion of $F(\mu)T'(\mu)$ in λ gives

$$F(\mu)T'(\mu)(v_z + \delta z) = F(\lambda)T'(\lambda)v_z + (\mu - \lambda)\frac{d}{d\mu}[F(\mu)T'(\mu)]_{|\mu=\hat{\mu}}v_z + F(\mu)T'(\mu)\delta z \tag{4.13}$$

for some $\hat{\mu} \in U_{\tau_1}(\lambda)$. By Corollary 2.3, *iii.*, we have $F(\lambda)T'(\lambda)v \in \ker T(\lambda)$ for all $v \in \ker T(\lambda)$. Hence,

$$\widetilde{z}(\mu, z) := \frac{|\mu - \lambda|}{\|B_{-1}T'(\mu)z\|\chi(\mu, z)^{1/2}}\left((\mu - \lambda)^{-1}B_{-1}T'(\mu)z + F(\lambda)T'(\lambda)v_z\right)$$

is an element of $\ker T(\lambda)$. Thus, we get from (4.12) with (4.13) that

$$\begin{aligned}\inf_{v \in \ker T(\lambda)} \big\|\|s(\mu, z)\|^{-1}s(\mu, z) - v\big\| &\leq \big\|\|s(\mu, z)\|^{-1}s(\mu, z) - \widetilde{z}(\mu, z)\big\| \\ &= \frac{|\mu - \lambda|}{\|B_{-1}T'(\mu)z\|\chi(\mu, z)^{1/2}}\Big\|(\mu - \lambda)\frac{d}{d\mu}[F(\mu)T'(\mu)]_{|\mu=\widetilde{\mu}}v_z + F(\mu)T'(\mu)\delta z\Big\|.\end{aligned}$$

Using (4.10) and (4.11), there exist $0 < \varepsilon \leq \tau_1$ and $0 < \tau \leq \tau_1$ such that for all z with $\|\delta z\| \leq \varepsilon$ and for all μ with $0 < |\mu - \lambda| \leq \tau$ we have

$$\|B_{-1}T(\mu)z\|\chi(\mu, z)^{1/2} \geq \frac{1}{2}.$$

Since $\frac{d}{d\mu}[F(\mu)T'(\mu)]$ and $F(\mu)T'(\mu)$ are bounded in $\overline{U}_\tau(\lambda)$, there exists a constant $c_\lambda > 0$ such that

$$\min_{v \in \ker T(\lambda)} \big\|\|s(\mu, z)\|^{-1}s(\mu, z) - v\big\| \leq c_\lambda|\mu - \lambda|\big(|\mu - \lambda| + \|\delta z\|\big)$$

for all z with $\|\delta z\| \leq \varepsilon$ and for all μ with $0 < |\mu - \lambda| \leq \tau$. □

In the next theorem we prove the convergence of Algorithm 2.

Theorem 4.2. *Let $\lambda \in \sigma(T)$ be a semi–simple eigenvalue. Then, there exist an $\varepsilon_0 > 0$ and a $\tau_0 > 0$ such that for all*

$$v_0 \in \left\{ z \in \mathbb{C}^n : \|z\| = 1, \min_{v \in \ker T(\lambda)} \|z - v\| \leq \varepsilon_0 \right\}$$

and for all λ_0 with $0 < |\lambda_0 - \lambda| \leq \tau_0$ the sequence $\{\lambda_i\}_{i \in \mathbb{N}}$ defined by Algorithm 2 converges to λ. Moreover, there exists a constant $c > 0$ such that

$$|\lambda_{i+1} - \lambda| + \min_{v \in \ker T(\lambda)} \|v_{i+1} - v\| \leq c\,|\lambda_i - \lambda| \left(|\lambda_i - \lambda| + \|\delta v_i\| \right)$$

for all $i \in \mathbb{N}$, where δv_i is the best approximation error of v_i in $\ker T(\lambda)$*.*

Proof. Let $\lambda \in \sigma(T)$ be a semi–simple eigenvalue, then we can choose $\tau_* > 0$ such that for all $\mu \in \overline{U}_{\tau_*}(\lambda) \setminus \{\lambda\}$ the resolvent $T(\mu)^{-1}$ admits a representation by

$$T(\mu)^{-1} = (\mu - \lambda)^{-1} B_{-1} + F(\mu)$$

where $F : \overline{U}_{\tau_*}(\lambda) \to \mathbb{C}^{n \times n}$ is continuous and holomorphic in $U_{\tau_*}(\lambda)$, and where B_{-1} is defined as in the proof of Lemma 4.1. Let $z \in \mathbb{C}^n$, $\|z\| = 1$, then we can write

$$\frac{[T(\mu)^{-1}T'(\mu)z]^H z}{\|T(\mu)^{-1}T'(\mu)z\|^2} = (\mu - \lambda) \frac{[B_{-1}T'(\mu)z + (\mu - \lambda)F(\mu)T'(\mu)z]^H z}{\|B_{-1}T'(\mu)z\|^2\,\chi(\mu, z)}, \tag{4.14}$$

where χ is defined as in (4.7). Let us denote again by $v_z \in \mathbb{C}^n$ the best approximation of z in $\ker T(\lambda)$, and let

$$\delta z := z - v_z.$$

Then, by (4.8) we have

$$B_{-1}T'(\mu)z = v_z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z$$

for some $\widetilde{\mu} \in U_{\tau_*}(\lambda)$. Further, we get

$$\begin{aligned}
[B_{-1}T'(\mu)z]^H z &= [v_z + (\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z]^H [v_z + \delta z] \\
&= v_z^H v_z + [(\mu - \lambda)B_{-1}T''(\widetilde{\mu})v_z + B_{-1}T'(\mu)\delta z]^H v_z + [B_{-1}T'(\mu)z]^H \delta z \\
&= 1 + \alpha(\mu, z) + (\overline{\mu - \lambda})\beta(\mu, z),
\end{aligned}$$

where

$$\begin{aligned}
\alpha(\mu, z) &:= -\delta z^H \delta z + [B_{-1}T'(\mu)\delta z]^H v_z + [B_{-1}T'(\mu)z]^H \delta z, \\
\beta(\mu, z) &:= [B_{-1}T''(\widetilde{\mu})v_z]^H v_z.
\end{aligned}$$

Let us introduce

$$\gamma(\mu, z) := [F(\mu)T'(\mu)z]^H z,$$

then we have

$$\frac{[T(\mu)^{-1}T'(\mu)z]^H z}{\|T(\mu)^{-1}T'(\mu)z\|^2} = (\mu-\lambda)\frac{1+\alpha(\mu,z)+(\overline{\mu-\lambda})\big(\beta(\mu,z)+\gamma(\mu,z)\big)}{\|B_{-1}T'(\mu)z\|^2\,\chi(\mu,z)}. \tag{4.15}$$

Recall the representation (4.10),

$$\|B_{-1}T'(\mu)z\|^2\,\chi(\mu,z) = 1+\nu(\mu,z),$$

where $|\nu(\mu,z)| \leq c_\nu(\|\delta z\|+|\mu-\lambda|)$ for all μ with $|\mu-\lambda|\leq\tau_\nu$ and for all $\|\delta z\|\leq\tau_\nu$ for a sufficiently small $0<\tau_\nu\leq\tau_*$. Thus, we can write

$$\frac{[T(\mu)^{-1}T'(\mu)z]^H z}{\|T(\mu)^{-1}T'(\mu)z\|^2} = (\mu-\lambda)\left(1+\frac{\alpha(\mu,z)+(\overline{\mu-\lambda})\big(\beta(\mu,z)+\gamma(\mu,z)\big)-\nu(\mu,z)}{1+\nu(\mu,z)}\right).$$

Since $T'(\mu)$, $T''(\mu)$ and $F(\mu)$ are continuous in $\overline{U}_{\tau_\nu}(\lambda)$, $\|z\|=1$, $\|\delta z\|\leq 1$, and $\|v_z\|\leq 1$, we have

$$\begin{aligned}
|\alpha(\mu,z)| &\leq \|\delta z\|(\|\delta z\|+2\|B_{-1}T'(\mu)\|)\leq c_1\|\delta z\|,\\
|\beta(\mu,z)| &\leq \|T''(\widetilde{\mu})\|\leq c_2,\\
|\gamma(\mu,z)| &\leq \|F(\mu)T'(\mu)\|\leq c_3
\end{aligned}$$

for all $\mu\in\overline{U}_{\tau_\nu}(\lambda)$. Hence, for sufficiently small $\tau>0$ and $\varepsilon>0$ there exists a constant $\tilde{c}>0$ such that

$$\begin{aligned}
\left|\mu-\lambda-\frac{[T(\mu)^{-1}T'(\mu)z]^H z}{\|T(\mu)^{-1}T'(\mu)z\|^2}\right| &= \left|(\mu-\lambda)\frac{\alpha(\mu,z)+(\overline{\mu-\lambda})\big(\beta(\mu,z)+\gamma(\mu,z)\big)-\nu(\mu,z)}{1+\nu(\mu,z)}\right|\\
&\leq \tilde{c}\,|\mu-\lambda|\big(|\mu-\lambda|+\|\delta z\|\big) \qquad (4.16)
\end{aligned}$$

for all z with $\|\delta z\|\leq\varepsilon$ and for all $\mu\in\overline{U}_\tau(\lambda)\setminus\{\lambda\}$. Let us assume that τ and ε are chosen sufficiently small such that also the estimate (4.3) holds, i.e.

$$\min_{v\in\ker T(\lambda)}\left\|\frac{1}{\|T(\mu)^{-1}T'(\mu)z\|}T(\mu)^{-1}T'(\mu)z-v\right\|\leq c_\lambda\,|\mu-\lambda|\,\big(|\mu-\lambda|+\|\delta z\|\big). \tag{4.17}$$

Consider now Algorithm 2. The first update λ_1 for an initial pair (λ_0,v_0) is given by

$$\lambda_1=\lambda_0-\frac{[T(\lambda_0)^{-1}T'(\lambda_0)v_0]^H v_0}{\|T(\lambda_0)^{-1}T'(\lambda_0)v_0\|^2}.$$

Using (4.16) and (4.17) we get

$$\begin{aligned}
|\lambda_1-\lambda| &\leq \tilde{c}\,|\lambda_0-\lambda|\,\big(|\lambda_0-\lambda|+\|\delta v_0\|\big),\\
\min_{v\in\ker T(\lambda)}\|v_1-v\| &\leq c_\lambda\,|\lambda_i-\lambda|\,\big(|\lambda_0-\lambda|+\|\delta v_0\|\big)
\end{aligned}$$

for all v_0 with $\|\delta v_0\| \le \varepsilon$ and for $\lambda_0 \in \overline{U}_\tau(\lambda) \setminus \{\lambda\}$. Choose $\varepsilon_0 > 0$ and $\tau_0 > 0$ such that

$$\varepsilon_0 < \min\left\{\frac{1}{2\tilde{c}}, \varepsilon\right\} \quad \text{and} \quad \tau_0 < \min\left\{\frac{\varepsilon_0}{c_\lambda(\tau+\varepsilon_0)}, \frac{1}{2\tilde{c}}, \tau\right\}.$$

This implies that $\eta := \tilde{c}(\tau_0 + \varepsilon_0) < 1$ and

$$\begin{aligned} |\lambda_1 - \lambda| &\le \tilde{c}\,|\lambda_0 - \lambda|\big(|\lambda_0 - \lambda| + \|\delta v_0\|\big) \le |\lambda_0 - \lambda|\,\eta < \tau_0 < \tau, \\ \|\delta v_1\| &\le c_\lambda \tau_0(\tau_0 + \varepsilon_0) \le c_\lambda \tau_0(\tau + \varepsilon_0) \le \varepsilon_0 < \varepsilon \end{aligned}$$

for all λ_0 with $|\lambda_0 - \lambda| \le \tau_0$ and for all v_0 with $\|\delta v_0\| \le \varepsilon_0$. Thus, by induction we obtain with (4.16) and (4.17)

$$|\lambda_{i+1} - \lambda| \le \tilde{c}\,|\lambda_i - \lambda|\,\big(|\lambda_i - \lambda| + \|\delta v_i\|\big) \le \eta^i\,|\lambda_0 - \lambda| \to 0 \tag{4.18}$$

as $i \to \infty$ and

$$\min_{v \in \ker T(\lambda)} \|v_{i+1} - v\| \le c_\lambda\,|\lambda_i - \lambda|\,\big(|\lambda_i - \lambda| + \|\delta v_i\|\big), \tag{4.19}$$

which proves the assertions. □

From the error estimates (4.18) and (4.19) it follows that the sequence $\{(\lambda_i - \lambda, \delta v_i)\}_{i\in\mathbb{N}}$ converges quadratically to $0 \in \mathbb{C}^{n+1}$.

Remark 4.3. *The convergence results of Algorithm 2 can be derived in the same way also for Algorithm 1. However, it has to be assumed that the normalization vector d is not orthogonal to the eigenspace* $\ker T(\lambda)$, *and that the set $\{v \in \ker T(\lambda) : v/d^H v = 1\}$ is bounded. The second assumption ensures that the the sequence $\{v_i\}_{i\in\mathbb{N}}$ remains bounded.*

For defective eigenvalues, numerical examples indicate that the convergence of Algorithm 1 and Algorithm 2 is linear. However, to the best of our knowledge, a proof of this conjecture is not available. By considering the representation of the principal part of the resolvent in terms of generalized eigenvectors, which provides the Theorem of Keldysh, a proof of this conjecture might be possible. In [10], the convergence factor for double non–semi–simple eigenvalues is analyzed. Provided that the sequence $\{\lambda_i\}_{i\in\mathbb{N}}$ converges to an eigenvalue, it is shown that the convergence factor is 1/2.

As for the NGRQI, also for the augmented Newton method the linear system which has to be solved in every iteration step is ill–conditioned close to a multiple eigenvalue. However, again this slightly effects the iteration in practical computations. In numerical experiments still a quadratic convergence behavior for semi–simple eigenvalues with multiple geometric multiplicity is obtained, see Examples 5.1 and 5.3.

5 Examples

Example 5.1 We first consider a nonlinear eigenvalue problem of the form

$$T(\lambda) = e^\lambda F D(\lambda) G - \lambda I,$$

where $D(\lambda) = \text{diag}(\sin\lambda, e^\lambda - 1, 3, \ldots, n)$ with $n = 100$, and where $F, G \in \mathbb{R}^{n\times n}$ are taken as random with full rank. $\lambda = 0$ is a semi–simple eigenvalue of T with geometric multiplicity 2. We observe a quadratic convergence order of the NGRQI and of the augmented Newton method for the approximation $\lambda = 0$, see Figure 1. This confirms the theoretical results of Theorem 3.1 and of Theorem 4.2.

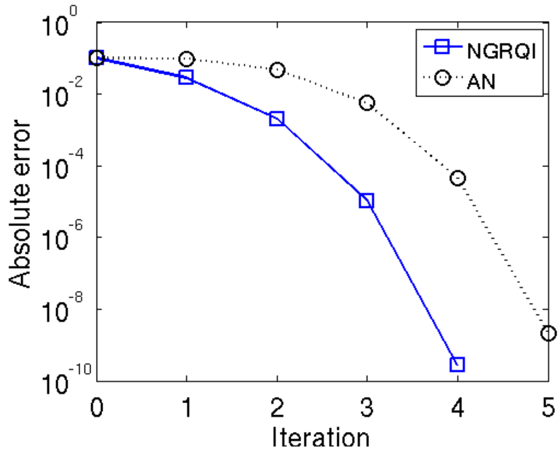

Figure 1: Convergence for Example 1.

Example 5.2 Next we consider the delay eigenvalue problem [10, Example 2]

$$T(\lambda) = -\lambda I + A_0 + A_1 e^{-\lambda},$$

where

$$A_0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -b_3 & -b_2 & -b_1 \end{pmatrix},$$

and

$$a_1 = \frac{2}{5}\frac{65\pi + 32}{8 + 5\pi}, \qquad a_2 = \frac{9\pi^2(13 + 5\pi)}{8 + 5\pi}, \quad a_3 = \frac{324}{5}\frac{pi^2(5\pi + 4)}{8 + 5\pi},$$
$$b_1 = \frac{260\pi + 128 + 225\pi 2}{80 + 50\pi}, \quad b_2 = \frac{45\pi 2}{8 + 5\pi}, \qquad b_3 = \frac{81(\pi^2(40\pi + 32 + 25\pi^2)}{80 + 50\pi}.$$

This eigenvalue problem has a double non–semi–simple eigenvalue $\lambda = 3\pi i$ [10]. According to Theorem 3.1, the NGRQI converges linearly with the convergence factor 1/2 which is confirmed by the computations, see Figure 2. Also the augmented Newton method converges linearly and the convergence factor is 1/2, as it was already demonstrated in [10].

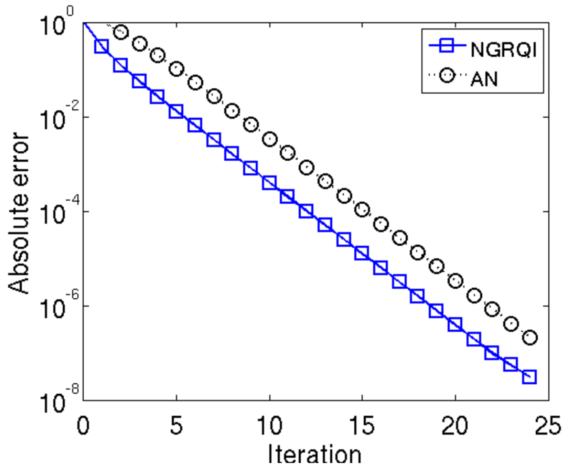

Figure 2: Convergence for Example 2.

Example 5.3 Finally we consider the boundary element discretization of the interior and exterior Dirichlet Laplacian eigenvalue problem

$$-\Delta u(x) = \lambda u(x) \quad \text{for } x \in \Omega_i, \quad u(x) = 0 \quad \text{for } x \in \partial\Omega_i, \quad i = 1, 2, \tag{5.1}$$

where for the interior eigenvalue problem we consider $\Omega_1 = \Omega = \{x \in \mathbb{R}^3 : \|x\| < 1\}$, while we have $\Omega_2 = \mathbb{R}^3 \setminus \overline{\Omega}$ for the exterior eigenvalue problem. For the exterior eigenvalue problem in addition an outgoing radiation condition for the eigenfunctions is assumed, see [9]. Both eigenvalue problems can be represented in terms of boundary integral equations [30, 31, 33]: If (κ^2, u) is an eigenpair of either the interior or exterior eigenvalue problem, then $(\kappa, \frac{\partial}{\partial n}u)$, where $\frac{\partial}{\partial n}u$ is the normal derivative of u on the boundary $\partial\Omega$, is an eigenpair of the nonlinear eigenvalue problem

$$\frac{1}{4\pi}\int_{\partial\Omega} \frac{e^{i\kappa|x-y|}}{|x-y|}\frac{\partial}{\partial n_y}u(y)ds_y = 0 \quad \text{for } x \in \partial\Omega. \tag{5.2}$$

The real eigenvalues of (5.2) correspond to the eigenvalues of the interior problem whereas the non–real ones correspond to the eigenvalues of the exterior problem. The exact eigenvalues of the interior problem are the squares of the zeros of the Bessel functions $J_{m+1/2}$, $m \in \mathbb{N}_0$. For the exterior problem the exact eigenvalues are the squares of the zeros of the spherical Hankel functions of the first kind [21]. For the discretization of the eigenvalue problem (5.2) we have approximated the boundary $\partial\Omega$ by n_h planar triangles τ_ℓ. As ansatz space for the eigenfunctions we have chosen the space of piecewise constant functions. The Galerkin discretization of the eigenvalue problem (5.2) takes the form [30]: Find $(\kappa_h, v_h) \in \mathbb{C} \times \mathbb{C}^{n_h} \setminus \{0\}$ such that

$$T_h(\kappa_h)v_h = 0, \tag{5.3}$$

where

$$T_h(\kappa_h)[k, \ell] := \frac{1}{4\pi}\int_{\tau_\ell}\int_{\tau_k} \frac{e^{i\kappa_h|x-y|}}{|x-y|}ds_y ds_x \quad \text{for } k, \ell = 1, \dots, n_h.$$

In order to get coarse approximations for the eigenpairs we have used the contour integral method [5] on the discretization level $L = 2$ with $n_h = 320$ boundary elements.

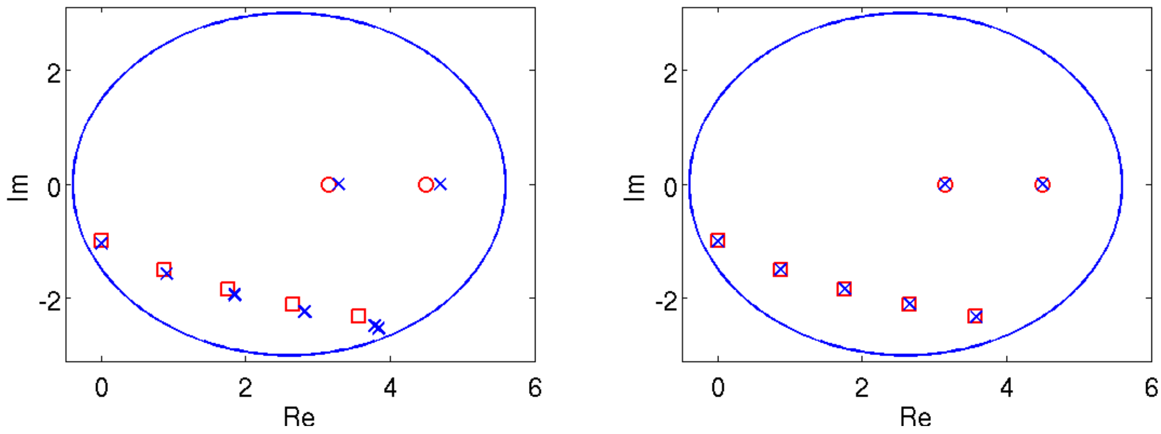

Figure 3: Numerical results for Example 5.3. *Left plot*: Approximations (*cross*) of the exact interior (*circle*) and exterior (*square*) eigenvalues inside the ellipse by the contour integral method on level $L = 2$. *Right plot*: Refinements by the NGRQI and the augmented Newton method on level $L = 5$.

Figure 3 shows the approximations of the eigenvalues inside the chosen ellipse and their refinements which we have got by the NGRQI and the augment Newton method on level $L = 5$ with $n_h = 5140$ boundary elements. The convergence behavior of both methods for the two smallest eigenvalues in modulus of the interior and exterior eigenvalue problem, respectively, are given in Figure 4 where for each eigenvalue a quadratic convergence can be observed. As solver for the linear systems we have used a GMRES methos without preconditioning.

On the continuous level the smallest interior eigenvalue is an algebraic simple eigenvalue whereas the others are semi–simple eigenvalues. By the discretization each semi–simple eigenvalue splits into several simple discrete eigenvalues which are very close to each other. The maximal absolute difference of the considered discrete eigenvalues on level $L = 4$ which are associated with a continuous eigenvalue is smaller than 10^{-5}. Thus, the results of this example rather demonstrate that both methods exhibit in practice also in the case of clustered simple eigenvalues a quadratic convergence behavior.

6 Conclusions

In this paper we have reviewed and extended the convergence results of two classical iterative methods for holomorphic eigenvalue problems which are usually restricted either to algebraic simple eigenvalues or to polynomial eigenvalue problems. We have considered the nonlinear generalized Rayleigh quotient iteration and the augmented Newton method

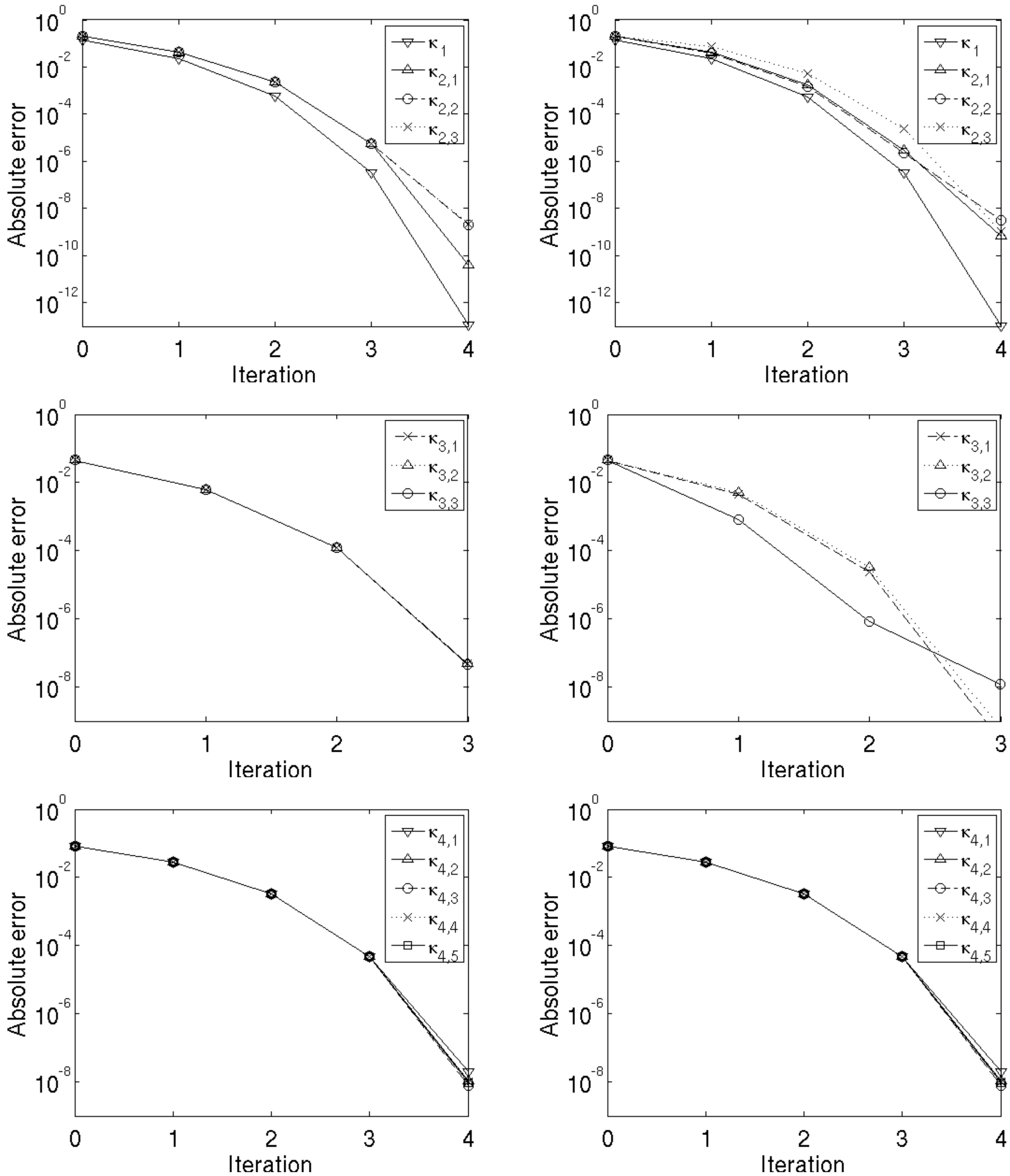

Figure 4: Numerical results for Example 5.3. Obtained eigenvalue approximations in the course of the NGRQI (*left column*) and the augmented Newton method (*right column*) on level $L = 5$. *First row*: Approximation of the interior eigenvalues. *Second and third row*: Approximation of the exterior eigenvalues.

which can be used for a more accurate approximation of eigenpairs. In our convergence analysis we utilize the representation of the resolvent as a meromorphic matrix–valued function which has the eigenvalues as poles. The convergence behavior of both methods depends on the order of the poles of the resolvent. For semi–simple eigenvalues, which are simple poles of the resolvent, local quadratic convergence has been shown for both methods. In the case of defective eigenvalues, local linear convergence for the nonlinear generalized Rayleigh quotient iteration has been shown.

The computational costs of both methods differ for non–Hermitian and non–symmetric eigenvalue problems. If the augmented Newton method is used, only one linear system has to be solved per iteration step whereas for the nonlinear generalized Rayleigh quotient iteration additionally an adjoint problem has to be solved.

Numerical experiments confirm the theoretical convergence results even then when the linear systems get ill–conditioned in the case of semi–simple eigenvalues with multiple geometric multiplicity and in the case of clustered eigenvalues. However, especially for these cases, appropriate preconditioning techniques are recommended.

References

[1] A. L. Andrew, K. E. Chu, P. Lancaster: On the numerical solution of nonlinear eigenvalue problems. Computing 55 (1995) 91–111.

[2] P. M. Anselone, L. B. Rall: The solution of characteristic value–vector problems by Newton's method. Numer. Math. 11 (1968) 38–45.

[3] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, K. Kimura: A numerical method for nonlinear eigenvalue problems using contour integrals. JSIAM Letters 1 (2009) 52–55.

[4] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, F. Tisseur: Nlevp: A collection of nonlinear eigenvalue problems. MIMS EPrint, Manchester Institute for Mathematical Sciences, The University of Manchester, 2008.

[5] W. J. Beyn: An integral method for solving nonlinear eigenvalue problems. Technical Report sfb701b3 10–007, Arbeitsgemeinschaft Numerische Analysis Dynamischer Systeme, Universität Bielefeld, 2010.

[6] H. Dai, P. Lancaster: Numerical methods for finding multiple eigenvalues of matrices depending on parameters. Numer. Math. 76 (1997) 189–208.

[7] I. Gohberg, S. Goldberg, M. A. Kaashoek: Classes of Linear Operators. Vol. I. Birkhäuser, Basel, 1990.

[8] I. C. Gohberg, E. I. Sigal: An operator generalization of the logarithmic residue theorem and Rouché's theorem. Math. USSR–Sb. 13 (1971) 603–625.

[9] T. Hohage, L. Nannen: Hardy space infinite elements for scattering and resonance problems. SIAM J. Numer. Anal. 47 (2009) 972–996.

[10] E. Jarlebring: Convergence factors of Newton methods for nonlinear eigenvalue problems. Linear Algebra Appl., 2010. (doi:10.1016/j.laa.2010.08.045.)

[11] T. Kato: Perturbation Theory for Linear Operators. Die Grundlehren der mathematischen Wissenschaften, Band 132. Springer, New York, 1966.

[12] M. V. Keldysh: On the eigenvalues and eigenfunctions of certain classes of non–selfadjoint operators. Dokl. Akad. Nauk SSSR 77 (1951) 11–14.

[13] V. Kozlov, V. Mazýa: Differential Equations with Operator Coefficients with Applications to Boundary Value Problems for Partial Differential Equations. Springer Monographs in Mathematics. Springer, Berlin, 1999.

[14] V. N. Kublanovskaya: On an approach to the solution of the generalized latent value problem for λ–matrices. SIAM J. Numer. Anal. 7 (1970) 532–537.

[15] H. Kummer: Zur praktischen Behandlung nichtlinearer Eigenwertaufgaben abgeschlossener linearer Operatoren. Mitt. Math. Sem. Giessen 62, 1964.

[16] P. Lancaster: A generalized Rayleigh quotient iteration for lambda–matrices. Arch. Rational Mech. Anal. 8 (1961) 309–322.

[17] U. Langer: Untersuchungen zum Kummerschen Verfahren zur numerischen Behandlung nichtlinearer Eigenwertaufgaben. Beiträge Numer. Math. 6 (1977) 97–110.

[18] R. C. Li: Compute multiple nonlinear eigenvalues. J. Comput. Math. 10 (1992) 1–20.

[19] V. Mehrmann, H. Voss: Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods. GAMM Mitt. Ges. Angew. Math. Mech. 27 (2005) 121–152.

[20] R. Mennicken, M. Möller: Non–self–adjoint boundary eigenvalue problems. North–Holland Mathematics Studies, vol. 192. North–Holland Publishing Co., Amsterdam, 2003.

[21] L. Nannen: Hardy–Raum Methoden zur Lösung von Streu– und Resonanzproblemen auf unbeschränkten Gebieten. PhD thesis, University Göttingen, 2008.

[22] A. Neumaier: Residual inverse iteration for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. 22 (1985) 914–923.

[23] M. R. Osborne: Inverse iteration, Newton's method and nonlinear eigenvalue problems. In: The contributions of Dr. J. H. Wilkinson to Numerical Analysis, Symp. Proc. Series 19, pp. 21–53, London, 1978.

[24] A. M. Ostrowski: On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. IV. (Generalized Rayleigh quotient for nonlinear elementary divisors). Arch. Rational Mech. Anal. 3 (1959) 341–347.

[25] A. Ruhe: Algorithms for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. 10 (1973) 674–689.

[26] K. Schreiber: Nonlinear Eigenvale Problems: Newton–type Methods and Nonlinear Rayleigh Functionals. PhD thesis, TU Berlin, 2008.

[27] H. Schwetlick, K. Schreiber: A primal–dual Jacobi–Davidson–like method for nonlinear eigenvalue problems. ePrints 2006–011, Institut für Mathematik, TU Berlin, 2006.

[28] H. Schwetlick, K. Schreiber: Nonlinear Rayleigh functionals. Linear Algebra Appl., 2010. (doi:10.1016/j.laa.2010.6.048.)

[29] A. Spence, C. Poulton: Photonic band structure calculations using nonlinear eigenvalue techniques. J. Comput. Phys. 204 (2005) 65–81.

[30] O. Steinbach, G. Unger: Convergence analysis of a Galerkin boundary element method for the Dirichlet Laplacian eigenvalue problem. Berichte aus dem Institut für Numerische Mathematik, Bericht 2010/9, TU Graz, 2010.

[31] M. E. Taylor: Partial Differential Equations. II. Springer, New York, 1996.

[32] V. P. Trofimov: The root subspaces of operators that depend analytically on a parameter. Mat. Issled. 3(1968) 117–125.

[33] G. Unger: Analysis of boundary element methods for Laplacian eigenvalue problems. Doctoral thesis, TU Graz, 2009.

[34] H. Unger: Nichtlineare Behandlung von Eigenwertaufgaben. Z. Angew. Math. Mech. 30 (1950) 281–282.

[35] H. Voss: Numerical methods for sparse nonlinear eigenvalue problems. In: Proc. of the 15th Summer School on Software and Algorithms of Numerical Mathematics, Hejnice, Czech Republic, pp. 133–160. University of West Bohemia, 2004.

Erschienene Preprints ab Nummer 2010/01

2010/1	G. Of, T. X. Phan, O. Steinbach: Boundary element methods for Dirichlet boundary control problems.
2010/2	P. D. Ledger, S. Zaglmayr: hp–Finite element simulation of three–dimensional eddy current problems on multiply connected domains.
2010/3	O. Steinbach, P. Urthaler: On the unique solvability of boundary integral equations in poroelasticity.
2010/4	S. Engleder, O. Steinbach: Boundary integral formulations for the forward problem in magnetic induction tomography.
2010/5	Z. Andjelic, G. Of, O. Steinbach, P. Urthaler: Direct and indirect boundary element methods for magnetostatic field problems.
2010/6	M. Neumüller, O. Steinbach: A space–time discontinuous Galerkin method for parabolic initial boundary value problems
2010/7	L. Tchoualag, G. Of, O. Steinbach: Circulant matrix methods for Yukawa type linear elasticity problems.
2010/8	A. Klawonn, U. Langer, L. F. Pavarino, O. Steinbach, O. B. Widlund (eds.): Workshop on Domain Decomposition Solvers for Heterogeneous Field Problems, Book of Abstracts.
2010/9	O. Steinbach, G. Unger: Convergence analysis of a Galerkin boundary element method for the Dirichlet Laplacian eigenvalue problem.
2010/10	O. Steinbach (ed.): Workshop on Computational Electromagnetics, Book of Abstracts.
2010/11	S. Beuchler, V. Pillwein, S. Zaglmayr: Sparsity optimized high order finite element functions for H(div) on simplices.
2010/12	U. Langer, O. Steinbach, W. L. Wendland (eds.): 8th Workshop on Fast Boundary Element Methods in Industrial Applications, Book of Abstracts.