manuscript No.
(will be inserted by the editor)

Analyzing the convergence factor of residual inverse iteration

Elias Jarlebring · **Wim Michiels**

Received: date / Accepted: date

Abstract We will establish here a formula for the convergence factor of the method called *residual inverse iteration*, which is a method for nonlinear eigenvalue problems and a generalization of the well-known *inverse iteration*. The formula for the convergence factor is explicit and only involves quantities associated with the eigenvalue to which the iteration converges, in particular the eigenvalue and eigenvector. Residual inverse iteration allows for some freedom in the choice of a vector w_k and we can use the formula for the convergence factor to analyze how it depends on the choice of w_k. We also use the formula to illustrate the convergence when the shift is close to the eigenvalue. Finally, we explain the slow convergence for double eigenvalues by showing that under generic conditions, the convergence factor is one, unless the eigenvalue is semisimple. If the eigenvalue is semisimple, it turns out that we can expect convergence similar to the simple case.

Keywords Nonlinear eigenvalue problems · Residual inverse iteration · Convergence factors

Mathematics Subject Classification (2000) 65F15 · 65H17 · 15A18

1 Introduction

Suppose $M : \Omega \to \mathbb{C}^{n\times n}$ corresponds to a matrix depending on the parameter $\lambda \in \mathbb{C}$ such that it is analytic in an open set $\Omega \subseteq \mathbb{C}$ and consider the problem of finding $\lambda_* \in \Omega$ and $v_* \in \mathbb{C}^n \backslash \{0\}$ such that,

$$M(\lambda_*)v_* = 0. \tag{1.1}$$

This equation is sometimes called the nonlinear eigenvalue problem and includes many of the fundamental problems of numerical linear algebra as special cases. See, e.g. the survey papers [13,18] and the more recent problem collection [4].

The method known as *residual inverse iteration* is an iterative method for finding one solution of (1.1) and can be summarized as follows. See [14] for a thorough introduction. Given an approximation of the eigenvector v_k and a vector w_k (further discussed later) compute a solution $\lambda \in \mathbb{C}$ to the scalar nonlinear equation

$$w_k^H M(\lambda) v_k = 0. \tag{1.2}$$

We assign the next eigenvalue iterate the value of a solution to (1.2), i.e., $\lambda_{k+1} \leftarrow \lambda$, and compute a new eigenvector approximation v_{k+1},

$$v_{k+1} = \beta_k (v_k - M(\sigma)^{-1} M(\lambda_{k+1}) v_k), \tag{1.3}$$

where the normalization constant $\beta_k \in \mathbb{C}$ is chosen such that $c^H v_{k+1} = 1$, with $c \in \mathbb{C}^n \backslash \{0\}$ and $c^H v_* \neq 0$. The complex number $\sigma \in \mathbb{C}$ is called the shift and is considered given by the user. The iteration consists of repeating the two steps corresponding to (1.2), i.e., solving the scalar equation, and (1.3), i.e., updating the eigenvector approximation. An attractive property of residual inverse iteration is that the matrix in the linear system $M(\sigma)$ remains constant throughout the iteration and can hence be factorized as a precomputation.

Several choices for the vector w_k have been proposed in the literature. In [14], Neumaier proposes to set

i) $w_k^H := c^H M(\sigma)^{-1}$ if the problem has no particular symmetry; and
ii) $w_k := v_k$ if $M(\lambda_*)$ is Hermitian and $\lambda_* \in \mathbb{R}$.

Schreiber [19] also proposes to use a different type of approximation of the left eigenvector. Although these specific recommendations on how to select w_k are available, the understanding of the impact of the choice of w_k is not complete in the general situation. In order to constructively study the impact of the choice of w_k, we derive the theory for the general choice

$$w_k := w(v_k) \to w(v_*) =: w_*.$$

The function $w : \mathbb{C}^n \to \mathbb{C}^n$ is assumed to be infinitely differentiable and such that $w_*^H M'(\lambda_*) v_* \neq 0$. This is related to the definition of the Rayleigh functional further discussed in Section 2.

The iteration converges linearly to simple eigenvalues. Let $\Delta v_k := v_k - v_*$ and $\Delta \lambda_k := \lambda_k - \lambda_*$ denote the error in the eigenvector and eigenvalue correspondingly. In this paper we will study the quantities $\alpha_{v,k} := \|v_{k+1} - v_*\|_2 / \|v_k - v_*\|_2$ and $\alpha_{\lambda,k} := |\lambda_{k+1} - \lambda_*| / |\lambda_k - \lambda_*|$ in the limit $k \to \infty$, which are referred to as the *eigenvector convergence factor* $\alpha_{v,*}$ and *eigenvalue convergence factor* $\alpha_{\lambda,*}$. This will be achieved by first (in Section 2) restating the iteration as a fixed point iteration $v_{k+1} = F(v_k)$, which is an iteration only in *eigenvector approximations*. We show (in Section 3) that, when close to a fixed point, the eigenvector error is given by

$$\Delta v_{k+1} = F'(v_*) \Delta v_k + O(\Delta v_k)^2,$$

where $F'(v_*)$ is a matrix depending on quantities involving the fixed point. Since the generic situation is that the iteration speed is limited by the largest eigenvalue of $F'(v_*)$, we can use $\rho(F'(v_*))$ to analyze the (eigenvector) convergence factor.

We use the formula to analyze several properties of the algorithm. In Section 4 we use it to analyze the situation where σ is close to the eigenvalue λ_*. In particular, we derive a formula for the first-order expansion when σ is close to λ_* and see that for any choice of w_k, the convergence factor approaches zero as σ approaches λ_*. We also analyze the case where w_k approaches the left eigenvector and establish that the eigenvalue convergence factor is smaller for this case, but show by example that it is not always optimal when σ is close to λ_*.

Neumaier conjectures by experiments that the method also works for double eigenvalues. However, observations in an example [17] indicate that the method can indeed fail for double eigenvalues. The formula allows us to completely explain the convergence to double eigenvalues. We show (in Section 6) that the iteration converges slowly or not at all for the generic situation (non-semisimple double eigenvalues). On the other hand, if the double eigenvalue is semisimple, we show that we can expect convergence similar to the simple case.

The method has been successfully combined with and used in modern methods for large-scale nonlinear eigenvalue problems [12]. The relation with the nonlinear Arnoldi method [22] is particularly noteworthy, since the residual inverse iteration is the motivation for the subspace expansion in [22] and it has been successfully used to solve many different types of nonlinear eigenvalue problems; see, e.g., [24,3,2, 23] and related works. Residual inverse iteration has been used in [17] to study the stability of a time-delay system with periodic coefficients. An extended version of the method in [22] was presented in [11] and a two-sided variation is given and analyzed in [19, Section 4.2.2]. Despite the extensive use, no progress has to our knowledge been made in terms of qualitative understanding of the convergence, apart from the original paper [14] and some notes in [19].

Finally, we wish to point out that residual inverse iteration can be seen as a variant of Newton's method. There are many variations of Newton's method for nonlinear eigenvalue problems. We have, e.g., the method of successive linear problems [18], block Newton [10], Rayleigh function iteration [20]. Some of the convergence properties of Newton's method (for nonlinear eigenvalue problems) are available [15, Section 5] and [1] including convergence factor analysis in [8]. The derivation and analysis of the methods differ substantially and we have not found any direct use of the analysis of these methods in this work.

2 A fixed point formulation

In order to characterize the convergence in the following sections we will need to formulate the iteration in a compact way. One component of the iteration consists of solving a nonlinear scalar equation, which depends on the eigenvector iterate $v \in \mathbb{C}^n$,

$$w(v)^H M(\lambda)v = 0. \tag{2.1}$$

Now suppose (λ_*, v_*) is a solution and

$$w_*^H M'(\lambda_*)v_* \neq 0. \tag{2.2}$$

The implicit function theorem now yields that there is an open neighborhood (which we denote $N(v_*)$) of v_* such that (2.1) defines a function $p : N(v_*) \to \Omega$, for which,

$$w(v)^H M(p(v))v = 0 \quad \text{for all } v \in N(v_*).$$

This is a common generic assumption for Rayleigh functionals. See [16,6] and [19, Chapter 3] for more on nonlinear Rayleigh functionals.

Since we are only interested in a local convergence analysis in this work, we can now assume that the iteration is contained in the neighborhood. An important aspect in our analysis, is that the unique definition of p allows us to reformulate the iteration (1.2) and (1.3) as a fixed point iteration in the eigenvector approximation,

$$v_{k+1} = F(v_k), \tag{2.3}$$

where

$$F(v) := \beta(v)(v - M(\sigma)^{-1}M(p(v))v), \tag{2.4}$$

and $\beta(v)$ is such that $c^H F(v) = 1$, i.e., $\beta(v) := 1/c^H(v - M(\sigma)^{-1}M(p(v))v)$. The elimination of the λ_k iteration variable is illustrated in Fig. 2.1.

In the following result we show that the fixed points of the iteration (2.3) are indeed solutions to the nonlinear eigenvalue problem. Moreover, the fixed points are isolated if $M(p(v))$ has a null space of dimension one, which is the case for simple eigenvalues. The case where the null space has rank greater than one corresponds to a non-isolated fixed point and will be further analyzed in the section on double eigenvalues (Section 6).

Proposition 2.1 (Fixed point equivalence) *Consider a vector $v \in \mathbb{C}^n$ with $c^H v \neq 0$ such that* (2.1) *has a unique solution $\lambda = p(v)$. Then the following statements are equivalent:*

i) $F(v) = v$.
ii) $M(p(v))v = 0$ *and* $c^H v = 1$.

Proof Suppose i) holds. Then, $F(v) = v$ and $1 = c^H F(v) = c^H v$. It follows from the definition of β and $F(v)$, i.e., (2.4), that $\beta(v) = 1$ and

$$v = F(v) = v - M(\sigma)^{-1}M(p(v))v,$$

from which we see that ii) holds. The converse follows analogously from direct manipulations.

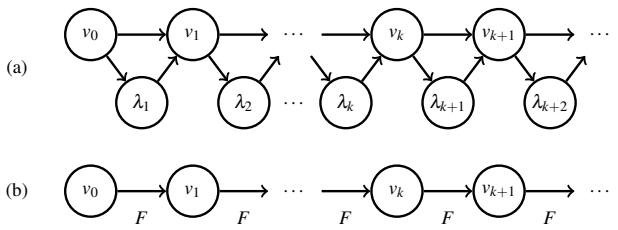


Fig. 2.1 The computation tree for residual inverse iteration (part (a)) and the fixed point reformulation (2.3) is shown in part (b).

3 The convergence factor

We are now ready to study the error of the residual inverse iteration using the fixed point formulation in the previous section. We first show how the fixed point iteration behaves close to a fixed point.

Theorem 3.1 (Linearization of fixed point iteration) *Consider a solution* (λ_*, v_*) *of the nonlinear eigenproblem* (1.1) *and a function w such that* (2.2) *is satisfied. Then, there exists a neighborhood of* v_* *denoted* $N(v_*)$*, such that for any pair* $v_k, v_{k+1} \in N(v_*)$ *corresponding to one step of residual inverse iteration* $v_{k+1} = F(v_k)$*, the eigenvector error* $\Delta v_k = v_k - v_*$ *satisfies*

$$\Delta v_{k+1} = (I - v_* c^H) A \Delta v_k + O(\Delta v_k)^2, \tag{3.1}$$

where

$$A := I - M(\sigma)^{-1} M(\lambda_*) + \frac{1}{w_*^H M'(\lambda_*) v_*} M(\sigma)^{-1} M'(\lambda_*) v_* w_*^H M(\lambda_*). \tag{3.2}$$

Proof Let $z(v) := v - M(\sigma)^{-1} M(p(v))v$. Then, the fixed point iteration can be written as

$$F(v) = \frac{1}{c^H z(v)} z(v).$$

Note that at the fixed point we have $c^H z(v_*) = 1$. By using the chain rule we now find that the Jacobian (evaluated in a fixed point $v_* = F(v_*)$) is

$$\begin{aligned} F'(v_*) &= (I - v_* c^H) z'(v_*) = \\ &\quad (I - v_* c^H) \left(I - M(\sigma)^{-1} M'(p(v_*)) v_* p'(v_*) - M(\sigma)^{-1} M(p(v_*)) \right), \end{aligned} \tag{3.3}$$

where $p'(v) \in \mathbb{C}^n$ denotes the (row) vector corresponding to the Jacobian of p. The Jacobian of p can be directly computed from the definition (2.1) and the implicit function theorem. Since the fixed point is a solution to the eigenvalue problem in the sense that $M(p(v_*))v_* = 0$ the term involving the Jacobian of w vanishes and it follows that

$$p'(v_*) = -\frac{w_*^H M(p(v_*))}{w_*^H M'(p(v_*)) v_*}. \tag{3.4}$$

We now express the eigenvector error and form a Taylor expansion of $F(v_k)$ around v_*,

$$\Delta v_{k+1} = v_{k+1} - v_* = F(v_k) - F(v_*) = F'(v_*) \Delta v_k + O(\|\Delta v_k\|^2). \tag{3.5}$$

The first-order expansion (3.1) follows from the combination of (3.3), (3.4) and (3.5).

Now note that the first-order characterization of the eigenvector error in (3.1) contains the information necessary to describe the error of the iteration, when v_k is close to v_*.

We will use the traditional approach called the *principle of linearization* roughly stating that the behavior of $v_{k+1} = F(v_k)$ when v_k is close to v_* is governed by the first-order approximation,

$$\Delta \tilde{v}_{k+1} = F'(v_*) \Delta \tilde{v}_k = (I - v_* c^H) A \Delta \tilde{v}_k. \tag{3.6}$$

In order to characterize the convergence factor of $v_{k+1} = F(v_k)$ we will consider the convergence factor of the linearized equation (3.6). We define

$$\tilde{\alpha}_{v,k} := \frac{\|\Delta\tilde{v}_{k+1}\|_2}{\|\Delta\tilde{v}_k\|_2},$$

and consider the limit as $k \to \infty$. Note that $\alpha_{v,k}$ is expected to coincide with $\tilde{\alpha}_{v,k}$ when sufficiently close to the fixed point. Consider one step of (3.1) and one step of (3.6) started with the same vector $\Delta v_k = \Delta\tilde{v}_k$. Then, the difference is

$$\alpha_{v,k} - \tilde{\alpha}_{v,k} = \frac{\|\Delta v_{k+1}\|_2\|\Delta\tilde{v}_k\|_2 - \|\Delta\tilde{v}_{k+1}\|_2\|\Delta v_k\|_2}{\|\Delta v_k\|_2\|\Delta\tilde{v}_k\|_2} = O(\Delta v_k).$$

In this way, $\alpha_{v,k}$ corresponding to step of the equation (3.1) is asymptotically the same as $\tilde{\alpha}_{v,k}$ corresponding to one step of the linearized equation (3.6). If started appropriately, the convergence factor for (3.6) is,

$$\rho((I - v_* c^H)A), \tag{3.7}$$

where $\rho : \mathbb{C}^{n\times n} \to \mathbb{R}$ denotes the spectral radius, i.e., the modulus of the largest eigenvalue. Formally, we have the following result which is easily derived, e.g., from [21, Theorem 27.1].

Lemma 3.2 (Convergence of linearized iteration map) *Let $\gamma_1, \ldots, \gamma_n$ be the eigenvalues of $(I - v_* c^H)A$ with corresponding eigenvectors $x_1, \ldots, x_n$. Suppose $|\gamma_1| > |\gamma_i|$ for all $i \neq 1$ and $x_1^H \Delta v_0 \neq 0$. Then, the iterates $\{\Delta\tilde{v}_i\}_{i\in\mathbb{N}}$ satisfy*

$$\frac{\Delta\tilde{v}_k}{\|\Delta\tilde{v}_k\|_2} \to x_1 \text{ as } k \to \infty, \tag{3.8}$$

and

$$\tilde{\alpha}_{v,k} := \frac{\|\tilde{\Delta}v_{k+1}\|_2}{\|\tilde{\Delta}v_k\|_2} \to |\gamma_1| = \rho((I - v_* c^H)A) \text{ as } k \to \infty. \tag{3.9}$$

Remark 3.1 (The matrix A and the iteration matrix) The construction of the iteration matrix

$$(I - v_* c^H)A$$

can be interpreted as follows. The operation of multiplying the matrix A from the left by $(I - v_* c^H)$ only changes one of the eigenvalues of A. Note that the matrix A always has an eigenvalue equal to one since $Av_* = v_*$. Moreover, we have that $c^H v_* = 1$ and the multiplication from the left with $(I - v_* c^H)$ corresponds to transforming the eigenvalue of v_* to zero and leaving all other eigenvalues unchanged. Hence, the eigenvalues of $(I - v_* c^H)A$ are the same as the eigenvalues of A except for the eigenvalue 1 corresponding to eigenvector v_*.

This is consistent with the fact that the iteration is independent of c (as long as $c^H v_k \neq 0$). Although the iteration matrix in (3.1) depends on c, the convergence factor given by (3.7) is also the modulus of an eigenvalue of A which is independent of c.

The convergence factor of the fixed point iteration (2.3) is completely characterized by (3.7). Since the fixed point iteration is an iteration in the eigenvector approximations v_k, the result in Theorem 3.1 does not directly imply anything about the convergence factor corresponding the eigenvalue iterates λ_k. We will now characterize eigenvalue iterates also using the matrix A.

For the eigenvalue convergence factor we need to distinguish between two cases. We essentially establish that the convergence factor is the same as in Theorem 3.1 for one case and square that of Theorem 3.1 for the other. Note that in this corollary we assume that $\Delta v_k/\|\Delta v_k\| \to x_1$, which is the case for the linearized iteration in Lemma 3.2 and it is the generic situation also for the nonlinear case.

Corollary 3.3 (Eigenvalue convergence factor) *Consider a solution (λ_*, v_*) of the nonlinear eigenproblem* (1.1) *and a sequence (λ_k, v_k) generated by residual inverse iteration convergent to (λ_*, v_*). Suppose the assumptions of Theorem 3.1 are satisfied for all k. Let $x_1 \in \mathbb{C}^n$ be as in Lemma 3.2. Moreover, suppose $\Delta v_k/\|\Delta v_k\| \to x_1$ as $k \to \infty$. Suppose $w : \mathbb{C}^n \to \mathbb{C}^n$ is infinitely differentiable in $v = v_*$.*

i) If $w_*^H M(\lambda_*)x_1 \neq 0$, then

$$\lim_{k\to\infty} \frac{|\Delta\lambda_{k+1}|}{|\Delta\lambda_k|} = \rho((I - v_* c^H)A). \tag{3.10}$$

ii) If, on the other hand, $w_*^H M(\lambda_*) = 0$ but $x_1^H p''(v_*)x_1 \neq 0$, then,

$$\lim_{k\to\infty} \frac{|\Delta\lambda_{k+1}|}{|\Delta\lambda_k|} = \rho((I - v_* c^H)A)^2. \tag{3.11}$$

Proof We will consider two steps and the corresponding expansions for $\lambda_{k+1} = p(v_k)$ and $\lambda_k = p(v_{k-1})$,

$$\lambda_k = p(v_{k-1}) = p(v_* + \Delta v_{k-1}) = \\ p(v_*) + p'(v_*)\Delta v_{k-1} + \frac{1}{2}\Delta v_{k-1}^H p''(v_*)\Delta v_{k-1} + O(\Delta v_{k-1})^3, \tag{3.12}$$

and

$$\lambda_{k+1} = p(v_* + \Delta v_k) = p(v_*) + p'(v_*)(I - v_* c^H)A\Delta v_{k-1} + \\ \frac{1}{2}\Delta v_{k-1}^H A^H (I - v_* c^H)^H p''(v_*)(I - v_* c^H)A\Delta v_{k-1} + O(\Delta v_k)^3, \tag{3.13}$$

where we used Theorem 3.1 to derive (3.13). Now we subtract $\lambda_* = p(v_*)$ from both equations (3.12) and (3.13) and form the quotient $|\lambda_{k+1} - \lambda_*|/|\lambda_k - \lambda_*|$. Then we take the limit and note that $\Delta v_{k-1}/\|\Delta v_{k-1}\| \to x_1$ which is an eigenvector of $(I - v_* c^H)A$ (corresponding to the largest eigenvalue). When using the formula for $p(v_*)$, (3.4), the first term is non-zero in case i and we directly establish (3.10). Case ii, follows similarly by noting that if $p'(v_*)x_1 = 0$, i.e., the first term cancels and (3.13) follows from the second term.

Remark 3.2 (Eigenvalue convergence factor and $w_*^H M(\lambda_*)x_1 = 0$) If w_* is a left eigenvector we have that $w_*^H M(\lambda_*) = 0$ and we have case ii) in Corollary 3.3. This is the case in particular when $M(\lambda_*)$ is Hermitian and λ_* real and we set (as Neumaier proposes) $w_k = v_k$.

Note that squaring the convergence factor does not imply a higher convergence order. The convergence is still linear, but faster. As an illustration we consider the nonlinear eigenvalue problem

$$M(\lambda) = -\lambda I + \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 6 \end{pmatrix} e^{-\lambda}, \tag{3.14}$$

which is symmetric and has a real eigenvalue $\lambda_* \approx 3.18581$. We carry out residual inverse iteration with $c = (1,1,1)^T$ and $\sigma = 3$, for different choices of w_k. The error is illustrated in Figure 3.1. We clearly see that the convergence factor for the eigenvalue is the essentially the square of the convergence factor for the eigenvector if $w_k = v_k$. The eigenvector convergence, however, does not appear faster than for the choice $w_k = c$.

Note that in all simulations in this paper we used software for high precision arithmetic, with such precision that the round-off error does not influence the figure.

4 Convergence factor for shift close to the eigenvalue

Due to the interpretation of σ as a shift, it is to be expected that we have fast convergence if it is close to an eigenvalue. This was already proven in Neumaier [14, page 919] for the choices of the vector $w_k^H = c^H M(\sigma)^{-1}$ and $w_k = v_k$, where it is shown that the convergence factor grows linearly with the distance between the shift and the eigenvalue. We will in this section see that this statement holds for arbitrary w_k.

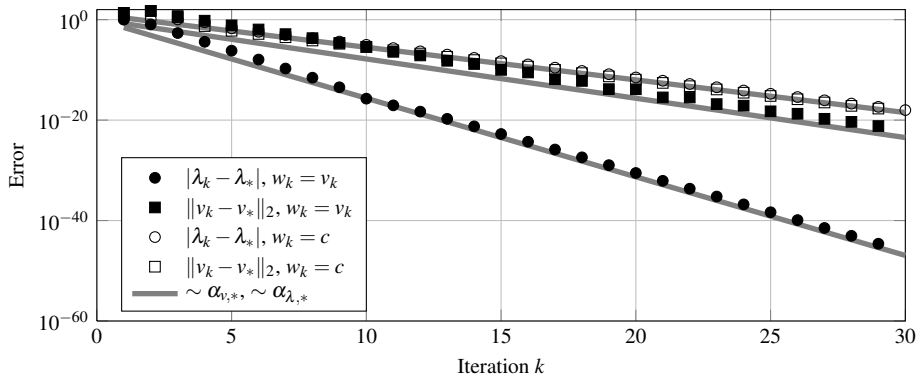


Fig. 3.1 The difference between the exact eigenvalue and the eigenvalue iterate for different choices of w_k, when applied to problem (3.14). The limits of $\alpha_{v,k}$ and $\alpha_{\lambda,k}$ predicted by (3.7), (3.10) and (3.11) are consistent with the numerical experiment.

Moreover, we derive explicit formulas for the growth of convergence factor as a function of the shift-eigenvalue distance. With the explicit formulas we can also characterize what is an optimal choice of w_k in terms of *growth rate*, here defined as the modulus of the derivative of the convergence factor as a function of the shift. We use an example to show that the optimal choice is not necessarily the left eigenvector.

In order to derive the main result we need the following lemma, which is essentially an extension of the zeroth-order expansion [14, Proposition 1] to a first-order expansion. The proof is available in Appendix A.

Lemma 4.1 *Let $\lambda_* \in \Omega$ be a solution to the nonlinear eigenvalue problem* (1.1) *with corresponding left and right eigenvectors u and v. Suppose $M(\lambda_*)$ has a null space of dimension one and $u^H M'(\lambda_*)v \neq 0$. Then, for any $\sigma \in \Omega$,*

$$M(\lambda_*)M(\sigma)^{-1} = B_0 + (\sigma - \lambda_*)B_1 + O(\sigma - \lambda_*)^2, \tag{4.1}$$

where $B_0 \in \mathbb{C}^{n\times n}$ is given by

$$B_0 = I - \frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v}$$

and $B_1 \in \mathbb{C}^{n\times n}$ is

$$B_1 = B_0\left(-M'(\lambda_*)M(\lambda_*)^+B_0 - \frac{M''(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right)\left(I + \frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right)^{-1},$$

where $(\cdot)^+$ denotes the pseudo inverse.

It turns out that when the shift is close to the eigenvalue, the convergence factor is also small.

Theorem 4.2 (Shift close to the eigenvalue) *Suppose $\lambda_* \in \Omega$ is a solution to the nonlinear eigenvalue problem* (1.1) *with corresponding normalized left and right eigenvectors u and v. Consider residual inverse iteration with shift $\sigma \in \Omega$. Suppose the largest eigenvalue (in modulus) of the iteration matrix $(I - v_* c^H)A$ is unique and simple and suppose* 1 *is a simple eigenvalue of A. Then, the convergence factor as a function of σ is to first order given by,*

$$\rho((I - v_* c^H)A) = |\sigma - \lambda_*|\rho\left(\left(I - \frac{M'(\lambda_*)v_* w_*^H}{w_*^H M'(\lambda_*)v_*}\right)B_1\right) + o(\sigma - \lambda_*) \tag{4.2}$$

and in particular if $w_* = u_*$,

$$\rho((I - v_* c^H)A) = |\sigma - \lambda_*|\rho(B_1) + O(\sigma - \lambda_*)^2. \tag{4.3}$$

Proof Consider the matrix $M(\sigma)^{-1}AM(\sigma)$, which has the same eigenvalues as A. It is straightforward to verify from the definition of A that w_* is a left eigenvector of $M(\sigma)^{-1}AM(\sigma)$ corresponding to the (simple) eigenvalue one. Let the columns of $V \in \mathbb{C}^{n\times(n-1)}$ be an orthogonal basis of the subspace perpendicular to w_*, i.e., $w_*^H V = 0$. Since, according to Remark 3.1, the eigenvalues of A are the the eigenvalues of $(I - v_* c^H)A$ except the eigenvalue at one, we can now study $\rho(I - v_* c^H)A) = \rho(V^H M(\sigma)^{-1}AM(\sigma)V)$.

In the following, we will need the following relation several times,

$$(I-\frac{M'(\lambda_*)vr^H}{r^HM'(\lambda_*)v})B_0 = I-M'(\lambda_*)v\left(\frac{r^H}{r^HM'(\lambda_*)v}+\frac{u^H}{u^HM'(\lambda_*)v}-\frac{r^H}{r^HM'(\lambda_*)v}\frac{M'(\lambda_*)vu^H}{u^HM'(\lambda_*)v}\right) = I-\frac{M'(\lambda_*)vr^H}{r^HM'(\lambda_*)v}. \quad (4.4)$$

Now form the Taylor expansion of A (as a function of σ) and simplify,

$$\begin{aligned} M(\sigma)AM(\sigma)^{-1} &= I-\left(I-\tfrac{M'(\lambda_*)v_*w_*^H}{w_*^HM'(\lambda_*)v_*}\right)M(\lambda_*)M(\sigma)^{-1} \\ &\quad I-\left(I-\tfrac{M'(\lambda_*)v_*w_*^H}{w_*^HM'(\lambda_*)v_*}\right)\big(I+(\sigma-\lambda_*)\,\big(-M'(\lambda_*)M(\lambda_*)^+B_0 \\ &\quad -\tfrac{M''(\lambda_*)v_*u^H}{u^HM'(\lambda_*)v_*}\Big)\left(I+\tfrac{M'(\lambda_*)v_*u^H}{u^HM'(\lambda_*)v_*}\right)^{-1}+O(\sigma-\lambda_*)^2\Big) \\ &= T_1+T_2(\sigma-\lambda_*)+O((\sigma-\lambda)^2), \end{aligned} \quad (4.5)$$

where

$$T_1=\frac{M'(\lambda_*)v_*w_*^H}{w_*^HM'(\lambda_*)v_*},\quad T_2=\left(I-\frac{M'(\lambda_*)v_*w_*^H}{w_*^HM'(\lambda_*)v_*}\right)B_1.$$

The projection now takes the form

$$V^HM(\sigma)AM(\sigma)^{-1}V = V^HT_1V+V^HT_2V(\sigma-\lambda_0)+O((\sigma-\lambda_*)^2).$$

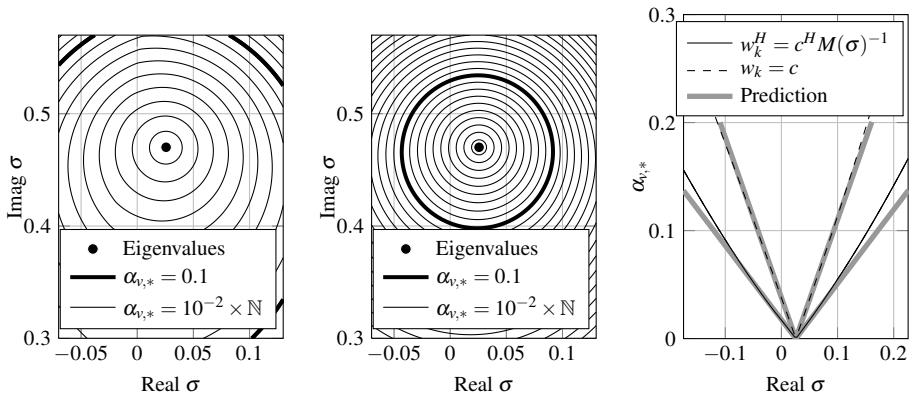
We have $V^HT_1V=0$. Furthermore, we have $\rho(V^HT_2V)=\rho(T_2)$, because $w_*^HT_2=0$, which implies on its turn that $\sigma(V^TT_2V)=\sigma(T_2)\cup\{0\}$. Assertion (4.3) follows. Formula (4.3) is proven analogously, by directly applying Lemma 4.1.

Remark 4.1 (The choice of w_k) Neumaier [14] and Schreiber [19] propose to choose w_k as an approximation of the left eigenvector. Although this is a natural choice since the eigenvalue convergence factor is expected to be small (see Remark 3.2), we will now see that it is not necessarily optimal in terms of growth rate. Consider the following cubic polynomial eigenvalue problem,

$$M(\lambda):=\begin{pmatrix}-16&-4&7\\-14&7&13\\6&8&7\end{pmatrix}+\lambda^2\begin{pmatrix}2&-6&1\\-2&22&11\\7&-1&1\end{pmatrix}+\lambda^3\begin{pmatrix}-4&3&12\\-17&-11&0\\1&-1&3\end{pmatrix}. \quad (4.6)$$

In Figure 4.1, where we illustrate the convergence factor and the growth rate for this example, we see the growth rate for $c=(1,1,1)$ and for $w_k^H=c^HM(\sigma)^{-1}$.

The growth rate for $w_k=u$ is $\rho'\approx 0.685$. By sampling with random w in the region $[-1,1]^3$ we find a better growth rate $\rho'\approx 0.658$ for $w_k=(0.630,-0.754,0.185)^T$. In a sense, this contradicts the idea that Neumaiers choice is optimal in terms of growth rate of the convergence factor. We note that at least for this example, the difference between the found optimum and the growth rate corresponding to the left eigenvector is small.



(a) $w_k^H = c^H M(\sigma)^{-1}$

(b) $w_k = c = (1,1,1)^T$

(c) Predicted growth-rate (4.2) and exact convergence factor (3.7).

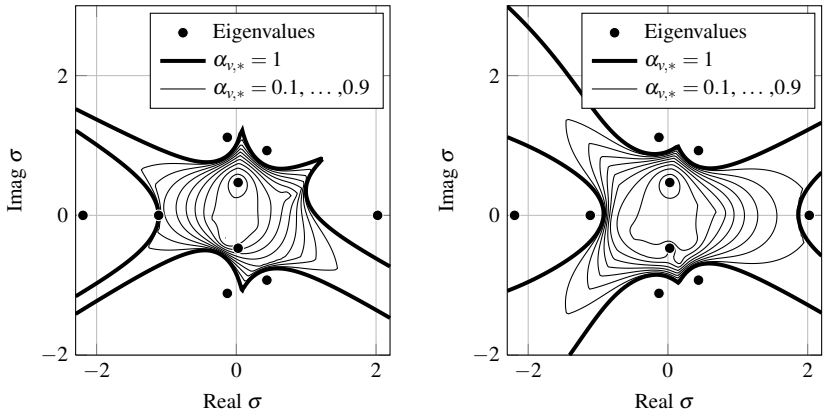
Fig. 4.1 The convergence factor $\rho((I - v_* c^H A)$, in (4.2), as a function of the shift, when the shift is close to the eigenvalue for the nonlinear eigenvalue problem (4.6).

5 Convergence basin as a function of the shift

We saw above that the formula for the convergence factor in Section 3 could be used when the shift-eigenvalue distance is small. We will now illustrate the use of the formula for the convergence factor not only locally. In this non-local analysis we can also illustrate the different choices of w_k.

We again consider the nonlinear eigenvalue problem (4.6). The convergence factor corresponding to the case where the iteration is started sufficiently close to $\lambda_* \approx 0.5i$ is shown in Figure 5.1 for two different choices of w_k. The left subfigure (Figure 5.1a) shows the convergence factor for the choice proposed by Neumaier. The convergence factor when w_k is chosen as the left eigenvector is illustrated in the right subfigure (Figure 5.1b). The region enclosed by the bold curve, i.e., a convergence factor equal to one, corresponds to possible shifts where the iteration will converge when started with a vector sufficiently close to the eigenvector corresponding to $\lambda_* \approx 0.5i$.

The figure indicates that in this case, the acceptable choices of σ is larger if w_k is close to a left eigenvector, supporting Schreiber's suggestion [19, page 82] that an approximation of the left eigenvector is a good choice. Note however that this example is not strictly conclusive, since there is a region of the complex plane where the choice of Neumaier (Figure 5.1a) is acceptable but the convergence factor corresponding to the left eigenvector is not.



(a) $w^H = c^H M(\sigma)^{-1}$ (b) $w = u$

Fig. 5.1 Level contours of the convergence factor (given by (3.7)) as a function of shift σ.

6 Double eigenvalues

In order characterize the situation where λ_* is a double eigenvalue we will use the concepts and generalizations of Jordan chains and generalized eigenvectors for nonlinear eigenvalue problems used in [5, Section 1.4] and [7]. The classification can be summarized as follows. A double eigenvalue can be either semisimple or non-semisimple. The matrix $M(\lambda_*)$ has a null space of dimension (exactly) two if and only if the eigenvalue is semisimple. Correspondingly, $M(\lambda_*)$ has a null space of dimension (exactly) one if the double eigenvalue is non-semisimple. Moreover, a double eigenvalue is non-semisimple if and only if there is a vector $v_{*,1} \in \mathbb{C}^n$ called a *generalized eigenvector*, associated with the eigenvector v_*, such that

$$M'(\lambda_*)v_* + M(\lambda_*)v_{*,1} = 0. \tag{6.1}$$

6.1 A double non-semisimple eigenvalue

Suppose λ_* is a double non-semisimple eigenvalue, which is the generic situation for double eigenvalues of nonlinear eigenvalue problems without any particular structure. Note that this does not change the fixed point formulation (in Section 2). In particular, since a non-semisimple double eigenvalue only has one eigenvector, we have from Proposition 2.1 that the fixed point corresponding to a double eigenvalue is isolated.

Suppose w_* is not a left eigenvector. We know from Remark 3.1 that the matrix A always has an eigenvalue equal to one corresponding to eigenvector v_*. It is now easy to verify that $v_{*,1}$ is also an eigenvector of A corresponding to the eigenvalue one, i.e., one is a double eigenvalue of A. It follows from the definition of the generalized

eigenvector (6.1) that,

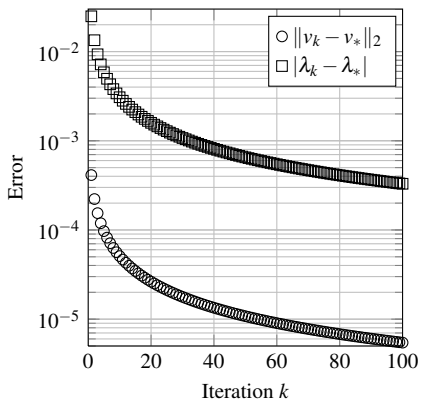
$$Av_{*,1} = v_{*,1} - M(\sigma)^{-1}M(\lambda_*)v_{*,1} + \frac{M(\sigma)^{-1}M'(\lambda_*)v_* w_*^H M(\lambda_*)v_{*,1}}{w_*^H M'(\lambda_*)v_*} = v_{*,1}. \quad (6.2)$$

The iteration matrix $(I - v_* c^H)A$ has the same eigenvalues as A except for the eigenvalue corresponding to v_* which is transformed to zero (see Remark 3.1). Loosely speaking, we now have that one of the two eigenvalues of A at one is removed when instead considering $(I - v_* c^H)A$, and one eigenvalue remains. The eigenvector corresponding to eigenvalue one is explicitly $\hat{v} := v_{*,1} - (c^H v_{*,1})v_*$. That is,

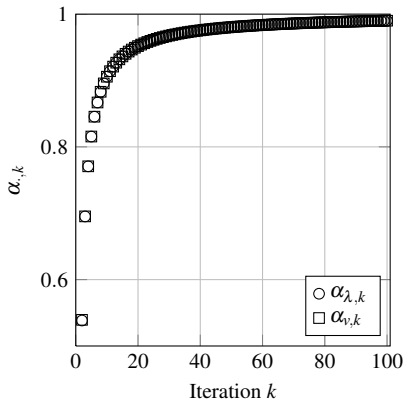
$$(I - v_* c^H)A\hat{v} = (I - v_* c^H)Av_{*,1} = (I - v_* c^H)v_{*,1} = \hat{v}, \quad (6.3)$$

where we used that $(I - v_* c^H)Av_* = 0$. From the fact that the iteration matrix $(I - v_* c^H)A$ always has an eigenvalue one we predict that for double non-semisimple eigenvalues, the convergence will be slow or the iteration will not converge at all. This is consistent with the application of residual inverse iteration in [17], where one bifurcation curve, which corresponds to a double eigenvalue, can not be accurately followed.

We here assumed that w_* is not a left eigenvector. If this is the case, the iteration is no longer well posed since the Rayleigh functional $p(\cdot)$ no longer uniformly defines a unique solution. That is, the condition (2.2) is violated. Once v_k is close to the solution, the Rayleigh functional will have (at least) two solutions close to the exact eigenvalues λ_* and the iteration does not uniquely define a next eigenvalue iterate.



(a) The convergence is slow for the double eigenvalue.



(b) The values $\alpha_{v,k}$ and $\alpha_{\lambda,k}$ both approach 1.

Fig. 6.1 Illustration of the convergence for the Example 6.1 which has a double non-semisimple eigenvalue.

Example 6.1 (Double non-semisimple eigenvalue) Consider as in [9],

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

where

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

and

$$\begin{aligned} a_1 &= \frac{2}{5}\frac{(65\pi+32)}{8+5\pi} \approx 3.98,\ a_2 = \frac{9\pi^2(13+5\pi)}{8+5\pi} \approx 108, \\ a_3 &= \frac{324}{5}\frac{\pi^2(5\pi+4)}{8+5\pi} \approx 531,\ b_1 = \frac{260\pi+128+225\pi^2}{10(8+5\pi)} \approx 13.6, \\ b_2 &= \frac{45\pi^2}{8+5\pi} \approx 18.7 \text{ and } b_3 = \frac{81\pi^2(40\pi+32+25\pi^2)}{10(8+5\pi)} \approx 1363. \end{aligned}$$

This nonlinear eigenvalue problem has a double non-semisimple eigenvalue for $\lambda = 3\pi i$.

We apply residual inverse iteration to this problem with the parameters $c = (1,1,1)^T$, $\sigma = 9.4i$, $w = M(\sigma)^{-H}c$ and $v_0 = v_* + (10^{-3}, 10^{-3}, 10^{-3})^T$. In the convergence diagram in Figure 6.1a we see that the convergence stagnates although the shift as well as the starting vector is very close to the exact solution. Figure 6.1b shows the estimate of the convergence factor $c_k := \|v_{k+1} - v_*\|/\|v_k - v_*\|$. It is clear that $\alpha_{v,k}$ approaches one, as predicted by (6.3).

6.2 A double semisimple eigenvalue

Now suppose λ_* is a double semisimple eigenvalue, i.e., the null space of $M(\lambda_*)$ has dimension two. The following lemma characterizes the fixed points of the iteration map F and follows directly from Proposition 2.1.

Lemma 6.1 *Let λ_* be a double semisimple eigenvalue. Let $\{v_1, v_2\}$ be a basis of the null space of $M(\lambda_*)$. Then, the set of fixed points of F corresponding to the eigenvalue λ_*, is given by*

$$V := \left\{ v \in \mathbb{C}^n \setminus \{0\} :\ v \in \operatorname{span}\{v_1, v_2\},\ \ c^H v = 1 \right\}.$$

Note that if either $c^H v_1 \neq 0$ or $c^H v_2 \neq 0$, then V is a one-dimensional linear space. Hence, in the generic case, the fixed point is not isolated. The properties of the iteration matrix are described in next theorem.

Theorem 6.2 *Let λ_*, v_1, v_2 and V be defined as in Lemma 6.1 and assume that either $c^H v_1 \neq 0$ or $c^H v_2 \neq 0$. Consider any fixed point $v_* \in V$. Then, the corresponding iteration matrix*

$$(I - v_* c^H)A, \tag{6.4}$$

with A defined by (3.2), always has an eigenvalue equal to one. If this eigenvalue is simple then the corresponding eigenvector $\tilde{v}$ satisfies

$$(v_* + \varepsilon\tilde{v}) \in V, \quad \forall \varepsilon \in \mathbb{R}. \tag{6.5}$$

Proof Take $\tilde{v} \in \operatorname{span}\{v_1, v_2\} \setminus \{0\}$ such that

$$c^H \tilde{v} = 0. \tag{6.6}$$

This is always possible because $c^H v_1 \neq 0$ or $c^H v_2 \neq 0$. We have

$$(I - v_* c^H)A\tilde{v} = (I - v_* c^H)\tilde{v} = \tilde{v}.$$

Thus, matrix (6.4) has an eigenvalue equal to one and $\tilde{v}$ is a corresponding eigenvector. The assertion (6.5) follows from $c^H v_* = 1$ and (6.6).

We have hence shown that, similarly to the non-semisimple case, the iteration matrix $(I - v_* c^H)A$ always has an eigenvalue at one. However, in contrast to the non-semisimple case, the eigenvalue one of the iteration matrix for the semisimple case does *not* affect the convergence rate of residual inverse iteration whenever it is simple. This follows from (6.5): a perturbation of the fixed point v_* in the direction of $\tilde{v}$, for which the linearized analysis is inconclusive about the recovery (eigenvalue one), corresponds to a perturbation *along* the line of fixed points.

The conclusion of the above reasoning is that for double semisimple eigenvalues one should study

$$\sigma((I - v_* c^H)A) \setminus \{1\} \tag{6.7}$$

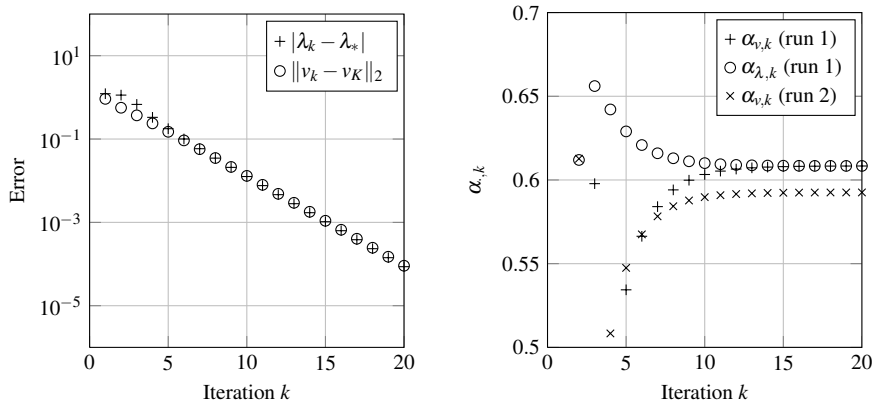
in order to establish the convergence factor. The convergence factor depends on the fixed point, or equivalently, the eigenvector, under consideration. Because there is a one-dimensional subspace of (normalized) eigenvectors V, the method will, when converging, not necessarily converge to a *specific* eigenvector. In theory, the asymptotics of the iteration could also exhibit a drift along the subspace of eigenvectors V. The latter has however not been observed in our numerical experiments.

Example 6.2 (A double semisimple eigenvalue) The nonlinear eigenvalue problem corresponding to

$$M(\lambda) := \begin{pmatrix} 1-\lambda^2 & 0 & \lambda \\ 0 & 2(1-\lambda^2) & 0 \\ 1-\lambda & 0 & \lambda^3 + 2\lambda^2 + 1 \end{pmatrix},$$

has a double semisimple eigenvalue for $\lambda = 1$. It is easy to verify that two corresponding right eigenvectors are $v = e_1$ and $v = e_2$. We fix $w = (1,1,2)^T$, $c = (1,1,1)^T$ and start the iteration with $v_0 = (1,2,1)^T$.

We first observe in Figure 6.2a that the convergence is linear. The figure shows the eigenvalue error $|\lambda_k - \lambda_*| = |\lambda_k - 1|$ and the difference $|v_k - v_K|$, where v_k is the eigenvector iterate and v_K an eigenvector iterate approximating a converged eigenvector ($K = 50$). We use this type of construction since, any vector in $\operatorname{span}(e_1, e_2)$ is an eigenvector, and it is not known a priori to which it converges (if any at all).



(a) Eigenvalue error ($|\lambda_k - \lambda_*|$) and eigenvector error ($\|v_k - v_K\|$) vs. iteration (for run 1)

(b) Convergence factor estimates: $\alpha_{\lambda,k} := |\lambda_{k+1} - \lambda_*|/|\lambda_k - \lambda_*|$ and $\alpha_{v,k} := \|v_{k+1} - v_K\|_2/\|v_k - v_K\|_2$.

Fig. 6.2 Visualization of the (linear) convergence for the semisimple eigenvalue in Example 6.2.

In a second run, we start the iteration with a different vector $v_0 = \frac{1}{2}(2,1,2)^T$. Note in Figure 6.2b that convergence factor changes with starting value. Although Theorem 3.1 gives an asymptotic expression for the eigenvector error when v_k is close to an eigenvector v_*, the expression depends on v_* which is not unique. Different eigenvectors of the manifold V will yield different A and hence different convergence factors.

7 Conclusions and outlook

An explicit formula for convergence factor, such as the one we have derived, can be used in many different ways. In this paper we used it to illustrate the convergence behavior when the shift is close to the eigenvalue and how the convergence factor depends on the choice of w_k. We also used it characterize the convergence for double eigenvalues.

We finally wish to point out that the use is not only limited to these concepts. A number of important quantities and concepts associated with an iterative method can be analyzed with the convergence factor. For instance, if the convergence factor can be accurately and cheaply estimated, it can be used to accelerate the method or even increase the convergence order by predicting the error of the iterate, similar to the technique of extrapolation. Accurate stopping criteria can also be derived using formulas for convergence factors. In this case, it could also be used to establish at what points in the iteration the shift should be updated.

A The proof of Lemma 4.1

Throughout this derivation we will need the following two formulas. Suppose E is a singular matrix with a null space of dimension one, with left and right null vectors u and v and $u^H v \neq 0$. Then,

$$\operatorname{adj}(E) = -\frac{\|\operatorname{adj}(E)\|_2}{u^H v} v u^H =: \beta v u^H. \tag{A.1}$$

We will also use the Jacobi formula for $\frac{d}{d\sigma}\det(M(\sigma))$,

$$\left(\frac{d}{d\sigma}\det(M(\sigma))\right)_{\sigma=\lambda_*} = \operatorname{Tr}(\operatorname{adj}(M(\lambda_*))M'(\lambda_*)) = \beta \operatorname{Tr}(v u^H M'(\lambda_*)) = \beta u^H M'(\lambda_*) v. \tag{A.2}$$

By using (A.1) and (A.2), the quantity we seek can now be expressed as

$$M(\lambda_*)M(\sigma)^{-1} = I + (M(\lambda_*) - M(\sigma))\frac{\operatorname{adj}(M(\sigma))}{\det(M(\sigma))} = I - (\sigma-\lambda_*)M'(\lambda_*)\frac{\operatorname{adj}(M(\sigma))}{\det(M(\sigma))} + O(\sigma-\lambda_*)^2 = I - \frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v} + O(\sigma-\lambda_*)^2, \tag{A.3}$$

where in the last step we expanded $\det(M(\sigma)) = \det(M(\lambda_*)) + (\sigma-\lambda_*)(\frac{d}{d\lambda}\det(M(\lambda)))_{\lambda=\lambda_*} + O(\sigma-\lambda_*)^2$ and applied (A.2). This proves the formula for B_0 in the expansion (4.1).

We will derive the formula for B_1 by first noting that

$$B_1 = \left(\frac{d}{d\sigma}M(\lambda_*)M(\sigma)^{-1}\right)_{\sigma\to\lambda_*}.$$

Hence,

$$\frac{d}{d\sigma}M(\lambda_*)M(\sigma)^{-1} = -M(\lambda_*)M(\sigma)^{-1}M'(\sigma)M(\sigma)^{-1} =$$

$$-(B_0 + (\sigma-\lambda_*)B_1)(M'(\lambda_*) + (\sigma-\lambda_*)M''(\lambda_*))\frac{\operatorname{adj}(M(\lambda_*)) + (\sigma-\lambda_*)(\frac{d}{d\lambda}\operatorname{adj}(M(\lambda)))_{\lambda=\lambda_*}}{(\sigma-\lambda_*)\beta u^H M'(\lambda_*)v} + O(\sigma-\lambda_*),$$

where we inserted the (still unknown) expansion $M(\lambda_*)M(\sigma)^{-1} = B_0 + (\sigma-\lambda_*)B_1 + O(\sigma-\lambda_*)^2$. Note that $B_0 M'(\lambda_*)\operatorname{adj}(M(\lambda_*)) = 0$. We now again use that $\operatorname{adj}(M(\lambda_*)) = \beta v u^H$ to find that

$$B_1 = \frac{-B_1 M'(\lambda_*)\beta v u^H - B_0 M''(\lambda_*)\beta v u^H - B_0 M'(\lambda_*)\frac{d}{d\lambda_*}\operatorname{adj}(M(\lambda_*))}{\beta u^H M'(\lambda_*)v}$$

By solving for B_1, we have that

$$B_1 = \left(-\frac{B_0 M'(\lambda_*)(\frac{d}{d\lambda}\operatorname{adj}(M(\lambda)))_{\lambda=\lambda_*}}{\beta u^H M'(\lambda_*)v} - \frac{B_0 M''(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right)\left(I + \frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right)^{-1}. \tag{A.4}$$

In the last step we will now compute the derivative of $M(\lambda_*)\operatorname{adj}(M(\sigma)) = M(\lambda_*)M(\sigma)^{-1}\det(M(\sigma)) = (B_0 + (\sigma-\lambda_*)B_1 + O(\sigma-\lambda_*)^2)\det(M(\sigma))$. By applying the product rule, we find that,

$$M(\lambda_*)\left(\frac{d}{d\sigma}\operatorname{adj}(M(\sigma))\right)_{\sigma=\lambda_*} = B_0(I\beta u^H M'(\lambda_*)v - \beta M'(\lambda_*)vu^H) = \beta(u^H M'(\lambda_*)v)B_0^2 = \beta(u^H M'(\lambda_*)v)B_0,$$

to which the potential solutions are parameterized by the variable $x \in \mathbb{C}^n$, such that

$$\left(\frac{d}{d\sigma}\operatorname{adj}(M(\sigma))\right)_{\sigma=\lambda_*} = \beta(u^H M'(\lambda_*)v)M(\lambda_*)^+ B_0 + vx^H. \tag{A.5}$$

Now note that the derivative of the adjoint matrix only appears in combination with the product

$$B_0 M'(\lambda_*)\frac{d}{d\lambda_*}\operatorname{adj}(M(\lambda_*))$$

in (A.4). In this combination, the free variable term vx^H in (A.5) vanishes since $B_0 M'(\lambda_*)v = 0$. This fact and insertion of (A.5) into (A.4) completes the proof.

Acknowledgements This article presents results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, the Optimization in Engineering Centre OPTEC of the K.U.Leuven, the project STRT1-09/33 of the K.U.Leuven Research Foundation and G0712.11 of the Research Council Flanders (FWO). We would like to thank Ed Lee of the department of mathematics, KU Leuven for carefully proofreading the manuscript.

References

1. Anselone, P., Rall, L.: The solution of characteristic value-vector problems by Newton's method. Numer. Math. **11**, 38–45 (1968)
2. Betcke, M.: Iterative projection methods for symmetric nonlinear eigenvalue problems with applications. Ph.D. thesis, Technical University Hamburg-Harburg (2007)
3. Betcke, M., Voss, H.: Stationary Schrödinger equations governing electronic states of quantum dots in the presence of spinorbit splitting. Appl. Math. **52**, 267 – 284 (2007)
4. Betcke, T., Higham, N.J., Mehrmann, V., Schröder, C., Tisseur, F.: NLEVP: A collection of nonlinear eigenvalue problems. Tech. rep., Manchester Institute for Mathematical Sciences (2008)
5. Gohberg, I., Lancaster, P., Rodman, L.: Matrix polynomials. Academic press (1982)
6. Hadeler, K.: Mehrparametrige und nichtlineare Eigenwertaufgaben. Arch. Ration. Mech. Anal. **27**, 306–328 (1967)
7. Hryniv, R., Lancaster, P.: On the perturbation of analytic matrix functions. Integral Equations Oper. Theory **34**(3), 325–338 (1999)
8. Jarlebring, E.: Convergence factors of Newton methods for nonlinear eigenvalue problems. Linear Algebra Appl. (2010). In press, http://dx.doi.org/10.1016/j.laa.2010.08.045
9. Jarlebring, E., Michiels, W.: Invariance properties in the root sensitivity of time-delay systems with double imaginary roots. Automatica **46**, 1112–1115 (2010)
10. Kressner, D.: A block Newton method for nonlinear eigenvalue problems. Numer. Math. **114**(2), 355–372 (2009)
11. Liao, B.S., Bai, Z., Lee, L.Q., Ko, K.: Nonlinear Rayleigh-Ritz iterative method for solving large scale nonlinear eigenvalue problems. Taiwanese Journal of Mathematics **14**(3), 869–883 (2010)
12. Meerbergen, K.: The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem. SIAM J. Matrix Anal. Appl. **30**(4), 1463–1482 (2008)
13. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. GAMM Mitteilungen **27**, 121–152 (2004)
14. Neumaier, A.: Residual inverse iteration for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. **22**, 914–923 (1985)
15. Peters, G., Wilkinson, J.: Inverse iterations, ill-conditioned equations and Newton's method. SIAM Rev. **21**, 339–360 (1979)
16. Rogers, E.: A minimax theory for overdamped systems. Arch. Ration. Mech. Anal. **16**, 89–96 (1964)
17. Rott, O., Jarlebring, E.: An iterative method for the multipliers of periodic delay-differential equations and the analysis of a PDE milling model. In: Proceedings of the 9th IFAC workshop on time-delay systems, Prague, pp. 1–6 (2010)
18. Ruhe, A.: Algorithms for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. **10**, 674–689 (1973)
19. Schreiber, K.: Nonlinear eigenvalue problems: Newton-type methods and nonlinear Rayleigh functionals. Ph.D. thesis, TU Berlin (2008)
20. Schwetlick, H., Schreiber, K.: A Jacobi-Davidson like method for nonlinear eigenvalue problems based on singularity theory. PAMM **7**(1), 1020,401–1020,402 (2007)
21. Trefethen, L.N., Bau, D.I.: Numerical linear algebra. SIAM, Philadelphia (1997)
22. Voss, H.: An Arnoldi method for nonlinear eigenvalue problems. BIT **44**, 387 – 401 (2004)
23. Voss, H.: Numerical methods for sparse nonlinear eigenvalue problems. In: Proc. XVth Summer School on Software and Algorithms of Numerical Mathematics, Hejnice, Czech Republic (2004). Report 70. Arbeitsbereich Mathematik, TU Hamburg-Harburg
24. Voss, H.: Iterative projection methods for computing relevant energy states of a quantum dot. J. Comput. Phys. **217**(2), 824–833 (2006)