# Convergence factors of Newton methods for nonlinear eigenvalue problems

Elias Jarlebring [a]

[a] *Department of Computer Science, K.U. Leuven, Celestijnenlaan 200 A, 3001 Leuven-Heverlee, Belgium*

**Abstract**

Consider a complex sequence $\{\lambda_k\}_{k=0}^{\infty}$ convergent to $\lambda_* \in \mathbb{C}$ with order $p \in \mathbb{N}$. The *convergence factor* is typically defined as the fraction $c_k := (\lambda_{k+1} - \lambda_*)/(\lambda_k - \lambda_*)^p$ in the limit $k \to \infty$. In this paper we prove formulas characterizing $c_k$ in the limit $k \to \infty$ for two different Newton-type methods for nonlinear eigenvalue problems. The formulas are expressed in terms of the left and right eigenvectors.

The two treated methods are called the *method of successive linear problems* (MSLP) and *augmented Newton* and are widely used in the literature. We prove several explicit formulas for $c_k$ for both methods. Formulas for both methods are found for simple as well as double eigenvalues. In some cases, we observe in examples that the limit $c_k$ as $k \to \infty$ does not exist. For cases where this limit appears to not exist, we prove other limiting expressions such that a characterization of $c_k$ in the limit is still possible.

*Key words:* nonlinear eigenvalue problems, Newton's method, convergence factors

## 1 Introduction

Consider the very general problem of finding a scalar $\lambda \in \mathbb{C}$ such that a given parameter-dependent matrix is singular. That is, find $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n \backslash \{0\}$ such that

$$T(\lambda)v = 0, \tag{1}$$

or equivalently, find $\lambda \in \mathbb{C}$ and $w \in \mathbb{C}^n \backslash \{0\}$ such that

$$w^H T(\lambda) = 0. \tag{2}$$

*Email address:* `elias.jarlebring cs.kuleuven.be` (Elias Jarlebring).

As usual we will call $v$ a *right eigenvector* and $w$ a *left eigenvector* associated with the *eigenvalue* $\lambda$.

The problem to solve (1) is often called the *nonlinear eigenvalue problem.* It is very general and includes many of the fundamental problems of numerical linear algebra. For instance, it includes the standard eigenvalue problem, $T(\lambda) = A - \lambda I$, the generalized eigenvalue problem, $T(\lambda) = A - \lambda B$, the quadratic eigenvalue problem, $T(\lambda) = M\lambda^2 + C\lambda + K$ (see e.g. [29]), the polynomial eigenvalue problem, $T(\lambda) = A_0 + A_1\lambda + \cdots + A_m\lambda^m$ (see e.g. [15]) and the delay eigenvalue problem, $T(\lambda) = -\lambda I + A_0 + A_1 e^{-\tau\lambda}$ (see e.g. [18] or [10]). Moreover, even though not often used in practice, it also includes the problem of solving a linear system of equations $Ax = b$. If we let $T(\lambda) = A - \lambda bb^H$ and assume $\lambda b^H v \neq 0$, then $x = v/(\lambda b^H v)$ solves $Ax = b$. We will assume that the elements of $T$ depend analytically on $\lambda$, which is often the case, in particular for all the above mentioned sub-problems.

It is not surpising that this somewhat fundamental problem has received a lot of attention in the literature. See the surveys of numerical methods [17, 24] and the recent problem collection [2].

Let $\{\lambda_k\}_{k=0}^{\infty}$ be a sequence convergent to $\lambda_*$ generated by a numerical method and denote

$$c_k := \frac{\lambda_{k+1} - \lambda_*}{(\lambda_k - \lambda_*)^p},$$

where $p$ is the convergence order of the sequence. If the limit exists and $c_k \to c \in \mathbb{C}$ then it defines the *convergence factor* $c$ (or sometimes $|c|$). In this paper we are investigating this limit. We give the convergence factor for several cases, and in other cases we characterize $c_k$ with explicit limit expressions. In several cases we observe in examples that $c_k$ does not converge, indicating that the convergence factor does not always exist. Despite the indication that the convergence factor does not always exist, the formulas allow us to make statements about the limit behavior of $c_k$.

The nonlinear eigenvalue problem is a nonlinear equation with a special structure. As for nonlinear equations, the Newton method is often used to find accurate solutions when some approximation is available, as in e.g., [4].

The results of this paper are for two Newton-type methods, the *method of successive linear problems* (MSLP) and *augmented Newton* (or just Newton). The found explicit formulas for the asymptotic behavior of $c_k$ are in terms of the left and right eigenvectors and include simple as well as non-semisimple double eigenvalues. Here, semisimple and non-semisimple refers to the definitions induced by the multiplicity concepts of the eigenvalues of nonlinear eigenvalue problems and Jordan chains in [6, Section 1.4] (and [7]). This definition of

Jordan chains is a consistent way to define a Jordan structure for nonlinear eigenvalue problems. The formulas for double non-semisimple eigenvalues will be expressed in terms of the generalized eigenvectors $\dot{v}$. Again, following the terminology in [6, Section 1.4] a generalized eigenvector $\dot{v}$ is a solution to the singular linear system

$$T(\lambda_*)\dot{v} = T'(\lambda_*)v. \tag{3}$$

The left generalized eigenvector $\dot{w}$ is defined analogously.

The new contributions for the two methods can be summarized as follows.

- The method of successive linear problems (Section 2):
  - For simple eigenvalues, the convergence factor exists and we prove a formula for it.
  - For double eigenvalues, $c_k$ asymptotically fulfills a quadratic equation. Observations in the examples indicate that $c_k$ does not approach only one of the solutions of the quadratic equation but oscillates between the two.
  - The formula allows us to characterize a situation where the convergence factor is large and MSLP hence likely to be inefficient.
- Augmented Newton (Section 3)
  - For simple eigenvalues we give an asymptotic expression for $c_k$.
  - For a special normalization vector the convergence factor exists and is equal to that of MSLP. For other normalization vectors we do not observe convergence of $c_k$.
  - For double eigenvalues, the convergence factor is $c = 1/2$, similar to the scalar case.

A number of convergence results for variants of Newton methods applied to nonlinear eigenvalue problems can be found in the literature. To the author's knowledge, formulas for convergence factors expressed in terms of eigenvectors have not been addressed in the general setting. For instance, superlinear convergence was already mentioned in an early publication [30]. Sufficient conditions for quadratic convergence together with expressions for the convergence region is given in [23]. The approximation error is bounded by norms in [25]. The connection with the convergence of inverse iteration is analyzed in [22]. This applies to this context, since inverse iteration (with Rayleigh shifts) has indeed been analyzed using eigenvectors [21], but mostly with bounds and not for the nonlinear eigenvalue problem. See also [5] and [8] for further convergence results of inverse iteration. The existence of a region of quadratic convergence was also proved in [31, Theorem 1]. See [1] for further bounds and a proof of quadratic convergence.

There is also a general classical theory for Newton methods. For instance, the standard references [20] and [26] (see also [25]). The classical theory is often in terms of norms, and will hence only give bounds of the convergence and the convergence factors. Moreover, they are typically not adapted for eigenvalue computations since the formulas are normally not in terms of eigenvectors.

Finally, we mention some other methods which are sometimes based on derivations similar to the Newton method but beyond the scope of this paper. There is a residual inverse iteration [19] and Jacobi-Davidson type methods [27], [3], [28]. There is also the approach of Kublanovskaya [9, 13] and similar methods for repeated eigenvalues [16]. Some methods are also based on the nonlinear Rayleigh functional [14]. More recently, a block version of Newton was generalized to nonlinear eigenvalue problems [12].

## 2 Convergence factors for MSLP

Consider the Taylor expansion of $T(\lambda_*)$ applied to (1),

$$0 = T(\lambda_*)v = T(\lambda_k)v + (\lambda_* - \lambda_k)T'(\lambda_k)v + O(\lambda_* - \lambda_k)^2.$$

In the classical derivation of the *method of successive linear problems* (presented in [24]) we neglect the higher order terms and replace $\lambda_* \approx \lambda_{k+1}$. Hence, one step of MSLP consists of solving the generalized eigenvalue problem,

$$T(\lambda_k)v_k = \mu T'(\lambda_k)v_k, \tag{4}$$

and setting $\lambda_{k+1} = \lambda_k - \mu$. From this reasoning we expect that $(\lambda_k, v_k)$ approximate $(\lambda_*, v)$ and that the approximation error is decreasing with $k$ if $\lambda_k$ is sufficiently close to $\lambda_*$. There are similar methods and derivations in the literature, e.g., [32]. Existence of a region of quadratic convergence region was proved in [31, Theorem 3]. The following two theorems contain (unlike [31]) exact formulas for the convergence factors for simple as well as double eigenvalues.

**Theorem 1 (MSLP, simple eigenvalue)** *Let $\{\lambda_k\}_{k=0}^{\infty}$ be a sequence generated by MSLP convergent to the semisimple eigenvalue $\lambda_* \in \mathbb{C}$, and $\{v_k\}_{k=0}^{\infty}$ the corresponding sequence of vectors (fulfilling (4)) convergent to right eigenvector $v$. Suppose $\lambda_k \neq \lambda_*$, $k \in \mathbb{N}$. Let $w \in \mathbb{C}^n \backslash \{0\}$ be a left eigenvector corresponding to the eigenvale $\lambda_*$ and suppose $w^H T'(\lambda_*)v \neq 0$. Then*

$$c := \lim_{k\to\infty} \frac{\lambda_{k+1} - \lambda_*}{(\lambda_k - \lambda_*)^2} = \frac{1}{2} \frac{w^H T''(\lambda_*)v}{w^H T'(\lambda_*)v}.$$

**Proof.** Let

$$\varphi_k(\lambda) := \lambda - \frac{w^H T(\lambda)v_k}{w^H T'(\lambda)v_k},$$

and note that the iteration fulfills $\lambda_{k+1} = \varphi_k(\lambda_k)$. Moreover, $\varphi_k(\lambda_*) = \lambda_*$ and $\varphi_k'(\lambda_*) = 0$ for all $k$ and

$$\varphi_k''(\lambda) = \frac{w^H T''(\lambda) v_k}{w^H T'(\lambda) v_k} + \varphi_k'(\lambda) w^H T'''(\lambda) w_k - 2\varphi_k'(\lambda) \frac{(w^H T''(\lambda) v_k)^2}{w^H T'(\lambda) v_k}.$$

Consider the Taylor expansion

$$\varphi_k(\lambda_k) = \varphi_k(\lambda_*) + (\lambda_k - \lambda_*)\varphi_k'(\lambda_*) + \frac{1}{2}(\lambda_k - \lambda_*)^2 \varphi_k''(\lambda_*) + O((\lambda_k - \lambda_*)^3).$$

From the Taylor expansion and the properties of $\varphi_k$, the convergence factor can now be expressed in terms of $\varphi_k''$,

$$\frac{\lambda_{k+1} - \lambda_*}{(\lambda_k - \lambda_*)^2} = \frac{\varphi_k(\lambda_k) - \varphi_k(\lambda_*)}{(\lambda_k - \lambda_*)^2} = \frac{1}{2}\varphi_k''(\lambda_*) + O(\lambda_k - \lambda_*).$$

The proof is completed by noting that

$$\varphi_k''(\lambda_*) = \frac{w^H T''(\lambda_*) v_k}{w^H T'(\lambda_*) v_k} \to \frac{w^H T''(\lambda_*) v}{w^H T'(\lambda_*) v} \text{ as } k \to \infty.$$

□

**Remark 2 (The degeneracy condition)** *In the theorem above (Theorem 1) we assumed that $w^H T'(\lambda_*) v \neq 0$. If the eigenvalue is simple, then the condition $w^H T'(\lambda_*) v \neq 0$ is always fulfilled. For semisimple eigenvalues of multiplicity greater than one, there are situations where the left and right eigenvectors, which are elements of the left and right null-space of $T(\lambda_*)$, yield $w^H T'(\lambda_*) v = 0$. However, given a sequence convergent to the right eigenvector $v$, there is always a left eigenvector which can be used in Theorem 1 unless the left null-space of $T(\lambda_*)$ is orthogonal to $T'(\lambda_*) v$. This is a degenerate situation.*

The following characterization of $c_k$ is possible using generalized eigenvectors defined by (3). Note that in the terminology of [6,7], a non-semisimple double eigenvalue always has a generalized eigenvector $\dot{v}$.

**Theorem 3 (MSLP, double eigenvalue)** *Let $\{\lambda_k\}_{k=0}^{\infty}$ be a sequence generated by MSLP convergent to a non-semisimple double eigenvalue $\lambda_* \in \mathbb{C}$ and let $\{v_k\}_{k=0}^{\infty}$ be the corresponding sequence of vectors (fulfilling (4)) convergent to right eigenvector $v$. Suppose $w^H T''(\lambda_*) v \neq 0$. Let $\dot{w} \in \mathbb{C}^n$ and $\dot{v} \in \mathbb{C}^n$ be generalized left eigenvector and right eigenvector of the double eigenvalue, i.e., $\dot{w}^H T(\lambda_*) = w^H T'(\lambda_*)$ and $T(\lambda_*)\dot{v} = T'(\lambda_*) v$. Suppose $\lambda_k \neq \lambda_*$, $k \in \mathbb{N}$. Moreover, let*

$$c_k := \frac{\lambda_{k+1} - \lambda_*}{\lambda_k - \lambda_*}.$$

*Then, in the limit, $c_k$ fulfills a quadratic equation,*

$$\lim_{k\to\infty}\left(\frac{\dot{w}^H T(\lambda_*)\dot{v}}{w^H T''(\lambda_*)v}c_k^2 - c_k + \frac{1}{2}\right) = 0. \tag{5}$$

**Proof.** From the definition of the sequence, it holds that

$$(1-c_k)(\lambda_k-\lambda_*)T'(\lambda_k)v_k = T(\lambda_k)v_k, \tag{6}$$

and in particular

$$(1-c_k)(\lambda_k-\lambda_*)w^H T'(\lambda_k)v_k = w^H T(\lambda_k)v_k. \tag{7}$$

The rest of the proof essentially consists of forming several Taylor expansions of the left and right hand side of (7). First consider the Taylor expansion of the right hand side of (7) and use $w^H T'(\lambda_*) = \dot{w}^H T(\lambda_*)$,

$$\begin{aligned}
&w^H T(\lambda_k)v_k = \\
&w^H T(\lambda_*)v_k + (\lambda_k-\lambda_*)w^H T'(\lambda_*)v_k + \frac{1}{2}(\lambda_k-\lambda_*)^2 w^H T''(\lambda_*)v_k + O(\lambda_k-\lambda_*)^3 \\
&\quad = (\lambda_k-\lambda_*)\dot{w}^H T(\lambda_*)v_k + \frac{1}{2}(\lambda_k-\lambda_*)^2 w^H T''(\lambda_*)v_k + O(\lambda_k-\lambda_*)^3.
\end{aligned}$$

We now again consider a Taylor expansion but only of the first term and in the other direction, i.e., the expansion point $\lambda_k$,

$$\begin{aligned}
&\dot{w}^H T(\lambda_*)v_k = \dot{w}^H T(\lambda_k)v_k - (\lambda_k-\lambda_*)\dot{w}^H T'(\lambda_k)v_k + O(\lambda_k-\lambda_*)^2 = \\
&(1-c_k)(\lambda_k-\lambda_*)\dot{w}^H T'(\lambda_k)v_k - (\lambda_k-\lambda_*)\dot{w}^H T'(\lambda_k)v_k + O(\lambda_k-\lambda_*)^2 = \\
&\qquad - c_k(\lambda_k-\lambda_*)\dot{w}^H T'(\lambda_k)v_k + O(\lambda_k-\lambda_*)^2,
\end{aligned} \tag{8}$$

where we used the iteration expression (6). Now consider the Taylor expansion of

$$w^H T'(\lambda_k)v_k = w^H T'(\lambda_*)v_k + (\lambda_k-\lambda_*)w^H T''(\lambda_*)v_k + O(\lambda_k-\lambda_*)^2.$$

For the first term we again have $w^H T'(\lambda_*)v_k = \dot{w}^H T(\lambda_*)v_k$ and can use (8) to establish that

$$w^H T'(\lambda_k)v_k = (\lambda_k-\lambda_*)(-c_k\dot{w}^H T'(\lambda_k)v_k + w^H T''(\lambda_*)v_k) + O(\lambda_k-\lambda_*)^2. \tag{9}$$

By combining (7), (8) and (9) we find that,

$$\begin{aligned}
&(1-c_k)(-c_k\dot{w}^H T'(\lambda_*)v_k + w^H T''(\lambda_*)v_k) = \\
&\qquad \frac{1}{2}w^H T''(\lambda_*)v_k - c_k\dot{w}^H T'(\lambda_*)v_k + O(\lambda_k-\lambda_*).
\end{aligned}$$

One term cancels and since the limit $v_k \to v$ exists, we have in the limit that

$$c_k^2 \dot{w}^H T'(\lambda_*)v - c_k w^H T''(\lambda_*)v + \frac{1}{2} w^H T''(\lambda_*)v \to 0.$$

By using that $T'(\lambda_*)v = T(\lambda_*)\dot{v}$ we have completed the proof.

□

**Remark 4 (Existence of convergence factors)** *In Theorem 3 we demonstrated that a limit expression containing $c_k$ exists and is zero. Note that this result does not imply or depend on the existence of $\lim_{k\to\infty} c_k$. In fact, in Example 2 we observe that $c_k$ oscillates between the two roots of the quadratic equation (5). This indicates that the limit does not always exist, which also seems to be the generic case.*

**Remark 5 (Numerical instability)** *A numerical method is likely to be inefficient if the convergence factor is large. Such a situation can be identified from Theorem 3. If $\dot{w}^H T(\lambda_*)\dot{v}/w^H T''(\lambda_*)v = 1/2$ then both roots will be unity. Hence, $|c_k| \to 1$ and MSLP will have slow convergence.*

## 3 Convergence factors for augmented Newton

The nonlinear eigenvalue problem (1) is equivalent to the set of nonlinear equations with the unknowns $x^H = (v^H, \lambda_*^*)$,

$$F(x) := \begin{pmatrix} T(\lambda)v \\ d^H v - 1 \end{pmatrix} = 0, \tag{10}$$

for a given normalization vector $d \in \mathbb{C}^n \backslash \{0\}$ which must be chosen in such a way that it is not orthogonal to the eigenvector. The Newton method applied to (10) is often called *augmented Newton*. For the theoretical reasoning in this section we need a different formulation. Some manipulations, which are often attributed to [30], (see also [25, Proposition 4.4]) result in the iteration

$$v_k = \alpha_k T(\lambda_k)^{-1} T'(\lambda_k) v_{k-1}. \tag{11}$$

with

$$\lambda_{k+1} = \lambda_k - \alpha_k, \;\; \alpha_k^{-1} = d^H T(\lambda_k)^{-1} T'(\lambda_k) v_{k-1}.$$

There are formulations with better numerical properties in terms of rounding errors. The formulation (11) is however suitable for our purposes and in this paper we only consider exact arithmetic, i.e., no rounding errors. In the next two theorems we give formulas for $c_k$ for the iteration (11). More precisely, for simple eigenvalues, we find that the convergence factor exists for a special normalization vector. If we do not have the special normalization vector, we still have a limit expression which now contains the quantity $\Delta v_k$, defined as the quotient of the difference between two consecutive eigenvector approximations and the error in the eigenvalue.

**Theorem 6 (Augmented Newton, simple eigenvalue)** *Let $\{\lambda_k\}_{k=0}^{\infty}$ be a sequence generated by augmented Newton convergent to a semisimple eigenvalue $\lambda_* \in \mathbb{C}$ and let $\{v_k\}_{k=0}^{\infty}$ be the corresponding sequence of vectors convergent to a right eigenvector $v$. Let $w \in \mathbb{C}^n \backslash \{0\}$ be a left eigenvector corresponding to the eigenvale $\lambda_*$ and suppose $w^H T'(\lambda_*)v \neq 0$. Suppose $\lambda_k \neq \lambda_*$, $k \in \mathbb{N}$. Let*

$$\Delta v_k := \frac{v_k - v_{k-1}}{\lambda_k - \lambda_*}.$$

*and*

$$c_k := \frac{\lambda_{k+1} - \lambda_*}{(\lambda_k - \lambda_*)^2}.$$

*Then*

$$\lim_{k\to\infty} \left( c_k + \frac{w^H T'(\lambda_*)}{w^H T'(\lambda_*)v} \Delta v_k \right) = \frac{1}{2} \frac{w^H T''(\lambda_*)v}{w^H T'(\lambda_*)v}. \tag{12}$$

*If, moreover, the normalizalization vector is $d^H = w^H T'(\lambda_*)$, then the convergence factor exists and is given by*

$$c := \lim_{k\to\infty} c_k = \frac{1}{2} \frac{w^H T''(\lambda_*)v}{w^H T'(\lambda_*)v}. \tag{13}$$

**Proof.** Similar to the proof of MSLP we can define

$$\varphi_k(\lambda) := \lambda - \frac{w^H T(\lambda) v_k}{w^H T'(\lambda) v_{k-1}}.$$

Note that $\varphi_k(\lambda_*) = \lambda_*$ and $\lambda_{k+1} = \varphi_k(\lambda_k)$. However,

$$\varphi_k'(\lambda_*) = 1 - \frac{w^H T'(\lambda_*) v_k}{w^H T'(\lambda_*) v_{k-1}} = -(\lambda_k - \lambda_*) \frac{w^H T'(\lambda_*) \Delta v_k}{w^H T'(\lambda_*) v_{k-1}},$$

and (unlike the corresponding iteration for MSLP) the derivative does not vanish for any $k$, i.e., generically, $\varphi_k'(\lambda_*) \neq 0$. The Taylor expansion is

$$\varphi_k(\lambda_k) = \lambda_* - (\lambda_k - \lambda_*)(\lambda_k - \lambda_*)\frac{w^H T'(\lambda_*)\Delta v_k}{w^H T'(\lambda_*) v_{k-1}} + \frac{1}{2}(\lambda_k - \lambda_*)^2 \varphi_k''(\lambda_*) + O((\lambda_k - \lambda_*)^3). \quad (14)$$

Hence,

$$c_k = \frac{\lambda_{k+1} - \lambda_*}{(\lambda_k - \lambda_*)^2} = \frac{\varphi_k(\lambda_k) - \lambda_*}{(\lambda_k - \lambda_*)^2} = -\frac{w^H T'(\lambda_*)\Delta v_k}{w^H T'(\lambda_*) v_{k-1}} + \frac{1}{2}\varphi_k''(\lambda_*) + O(\lambda_k - \lambda_*). \quad (15)$$

We now solve (15) for $\frac{1}{2}\varphi_k''(\lambda_*) + O(\lambda_k - \lambda_*)$ and take the limit on both sides. Note that

$$\lim_{k\to\infty} \varphi_k''(\lambda_*) = \frac{w^H T''(\lambda_*) v}{w^H T'(\lambda_*) v},$$

and $w^H T'(\lambda_*)\Delta v_k$ is bounded. We have shown (12).

It remains to show (13). Since the normalization vector is the same for all iterations, we have $d^H v_k = 1$ for all $k > 1$. Now note that (14) can be simplified since,

$$w^H T'(\lambda_*)(v_k - v_{k-1}) = d^H(v_k - v_{k-1}) = 0.$$

The proof is completed by repeating the steps leading to (15) which now does not contain the term $\Delta v_k$. □

**Remark 7 (Existence)** *In the theorem we used the quantity $\Delta v_k := (v_k - v_{k-1})/(\lambda_k - \lambda_*)$. If $w^H T'(\lambda_*)\Delta v_k$ is unbounded then convergence would not be quadratic, hence, by contradiction with the fact that Newton converges quadratically, the sequence $w^H T'(\lambda_*)\Delta v_k$ is bounded. However, in the examples section we do not observe convergence of $w^H T'(\lambda_*)\Delta v_k$. It also seems to be the generic case that the sequence does not converge and hence $c_k$ also does not converge.*

**Remark 8 (Convergence factor MSLP vs. augmented Newton)** *For the very special normalization vector $d^H = w^H T'(\lambda_*)$, the convergence factor exists. Moreover, for this normalization vector the convergence factor is actually equal to the convergence factor of MSLP in Theorem 1. The quantity $w^H T'(\lambda_*)$ is of course not available in practice and in general (unless information about the left eigenvector and the eigenvalue is available) $\frac{w^H T'(\lambda_*)}{w^H T'(\lambda_*) v}\Delta v_k$ will not vanish. Hence, since the convergence factor of MSLP will converge (for simple eigenvalues) it is in a sense more predictable than augmented Newton.*

Finally, we prove that the convergence factor exists for double eigenvalues and that it is given by $c = 1/2$.

**Theorem 9 (Augmented Newton, double eigenvalue)** *Let* $\{\lambda_k\}_{k=0}^{\infty}$ *be a sequence generated by augmented Newton convergent to a non-semisimple double eigenvalue* $\lambda_* \in \mathbb{C}$ *and* $w^H T''(\lambda_*)v - 2\dot{w}^H T'(\lambda_*)v \neq 0$. *Suppose* $\lambda_k \neq \lambda_*$, $k \in \mathbb{N}$ *and let,*

$$c_k := \frac{\lambda_{k+1} - \lambda_*}{\lambda_k - \lambda_*}.$$

*Then,*

$$\lim_{k\to\infty} c_k = \frac{1}{2}. \tag{16}$$

**Proof.** Note that the augmented Newton iteration fulfills,

$$(1 - c_k)(\lambda_k - \lambda_*)T'(\lambda_k)v_{k-1} = T(\lambda_k)v_k, \tag{17}$$

and in particular

$$(1 - c_k)(\lambda_k - \lambda_*)w^H T'(\lambda_k)v_{k-1} = w^H T(\lambda_k)v_k. \tag{18}$$

The proof follows the same ideas as in the proof of the convergence MSLP to double eigenvalues (Theorem 3), but different terms cancel and we use that $u_k := (v_{k-1} - v_k) \to 0$.

First note that by Taylor expansion around $\lambda_k$ and using (17) we have that,

$$T(\lambda_*)v_k = -c_k(\lambda_k - \lambda_*)T'(\lambda_k)v_{k-1} + (\lambda_k - \lambda_*)T'(\lambda_k)u_k + O(\lambda_k - \lambda_*)^2. \tag{19}$$

Taylor expansion and (19), yields

$$\begin{aligned}
&w^H T(\lambda_k)v_k = \\
&w^H T(\lambda_*)v_k + (\lambda_k - \lambda_*)\dot{w}^H T(\lambda_*)v_k + \frac{1}{2}(\lambda_k - \lambda_*)^2 w^H T''(\lambda_*)v_k + O(\lambda_k - \lambda_*)^3 = \\
&(\lambda_k - \lambda_*)^2 \left(-c_k \dot{w}^H T'(\lambda_k)v_{k-1} + \dot{w}^H T'(\lambda_k)u_k + \frac{1}{2} w^H T''(\lambda_*)v_k\right) + \\
&\qquad O(\lambda_k - \lambda_*)^3
\end{aligned} \tag{20}$$

The expression $w^H T'(\lambda_k)v_{k-1}$ is now expanded in $\lambda_*$. After using (19) for $k-1$ instead of $k$ we have that,

$$\begin{aligned}
&w^H T'(\lambda_k)v_{k-1} = \\
&(\lambda_k - \lambda_*)(-\dot{w}^H T'(\lambda_{k-1})v_{k-2} + (1/c_{k-1})\dot{w}^H T'(\lambda_{k-1})u_{k-1} + w^H T''(\lambda_*)v_{k-1}) + \\
&\qquad O(\lambda_k - \lambda_*)^2.
\end{aligned} \tag{21}$$

We insert (21) and (20) into (18) and cancel the quadratic term

$$(1-c_k)(-\dot{w}^H T'(\lambda_{k-1})v_{k-2}+(1/c_{k-1})\dot{w}^H T'(\lambda_{k-1})u_{k-1}+w^H T''(\lambda_*)v_{k-1}) = \\ -c_k\dot{w}^H T'(\lambda_k)v_{k-1}+\dot{w}^H T'(\lambda_k)u_k+\frac{1}{2}w^H T''(\lambda_*)v_k+O(\lambda_k-\lambda_*).$$

Note that $1/c_k$ is bounded since the convergence is otherwise superlinear and $u_k \to 0$. By forming the limit and removing the vanishing terms, we find that

$$0 = \lim_{k\to\infty}\Big((1-c_k)(-\dot{w}^H T'(\lambda_*)v + w^H T''(\lambda_*)v)+ \\ c_k\dot{w}^H T'(\lambda_*)v - \frac{1}{2}w^H T''(\lambda_*)v\Big).$$

We arrive at (16) by rearrangement of terms and using the assumption that $w^H T''(\lambda_*)v - 2\dot{w}^H T'(\lambda_*)v \neq 0$. □

## 4 Examples

**Example 1 (Simple eigenvalue)** *Consider the polynomial eigenvalue problem with*

$$T(\lambda) = A_0 + A_1\lambda + A_2\lambda^2 + A_3\lambda^3,$$

*where*

$$A_0 = \begin{pmatrix} -16 & -4 & 7 \\ -14 & 7 & 13 \\ 6 & 8 & 7 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 2 & -6 & 1 \\ -2 & 22 & 11 \\ 7 & -1 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} -4 & 3 & 12 \\ -17 & -11 & 0 \\ 1 & -1 & 3 \end{pmatrix}.$$

*We consider the iterations where we start the methods in such a way that the simple eigenvalue $\lambda_* \approx 0.0257 + 0.4701i$ is found. We let $d = (5, 3, 0)^T =: d_1^T$, $\lambda_0 = -0.4 + 0.6i$ and $v_0 = (3, -2, 0.1)^T$.*

*We observe in Figure 1 that $c_k$ for MSLP quickly converges to the expected value whereas the corresponding value for augmented Newton (with $d = d_1$) has not yet converged after 15 iterations.*

*In order to illustrate that the convergence factor of augmented Newton and MSLP coincide if $d^H = w^H T'(\lambda_*) =: d_*^H$, we also computed the left eigenvector*

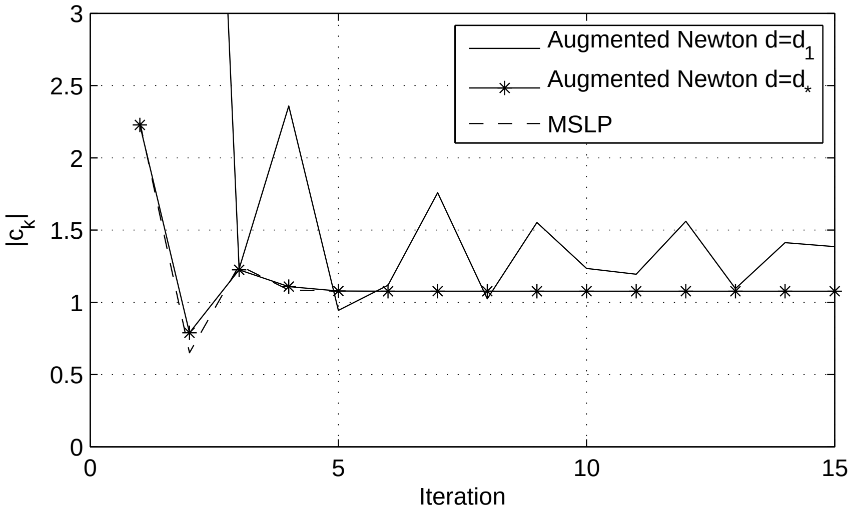


Fig. 1. The convergence factor for Example 1.

*and used the corresponding vector $d_*$ as a normalization vector in augmented Newton. In Figure 1, we observe, as expected from Theorem 6, that the convergence factors for augmented Newton with $d = d_*$ equals the convergence factor of MSLP. In fact, the difference is small, already after a few iterations.*

*Note that we are not considering rounding errors in this work. We have used software for (very) high precision arithmetic in order to carry out the numerical experiments in such a way that rounding errors are not influencing the plots.*

**Example 2 (Double eigenvalue)** *Consider the delay eigenvalue problem*

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

*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*

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

*This time-delay system which is presented in [11], has a double non-semisimple eigenvalue in* $\lambda = 3\pi i$*. As expected, we observe linear convergence of both augmented Newton and MSLP in Figure 2. We also see that augmented Newton converges faster. This can be explained by the fact that one of the roots of the quadratic equation* (5) *is considerably larger than* $1/2$ *which is the convergence factor for Newton. In Figure 3 we observe that for Newton,* $c_k$ *quickly converges to* $c = 1/2$ *whereas for MSLP,* $c_k$ *is alternating in an irregular way between the two roots of quadratic equation.*

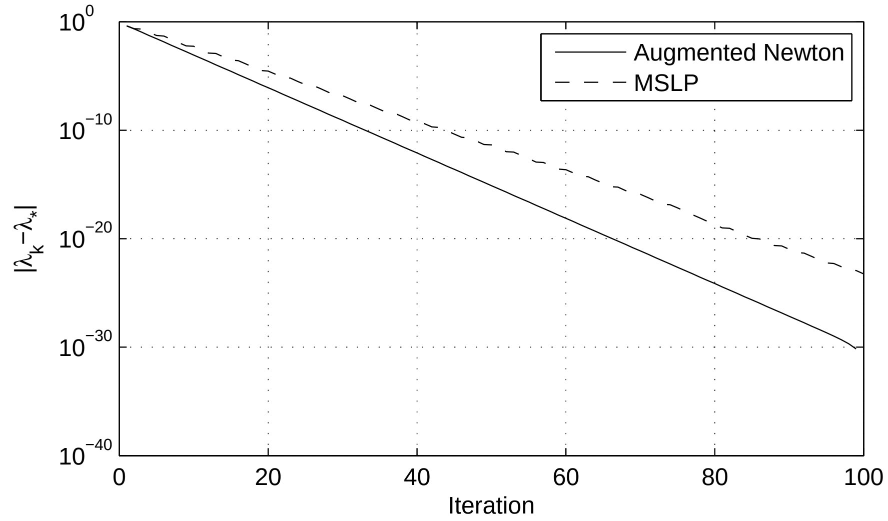


Fig. 2. The convergence for Example 2.

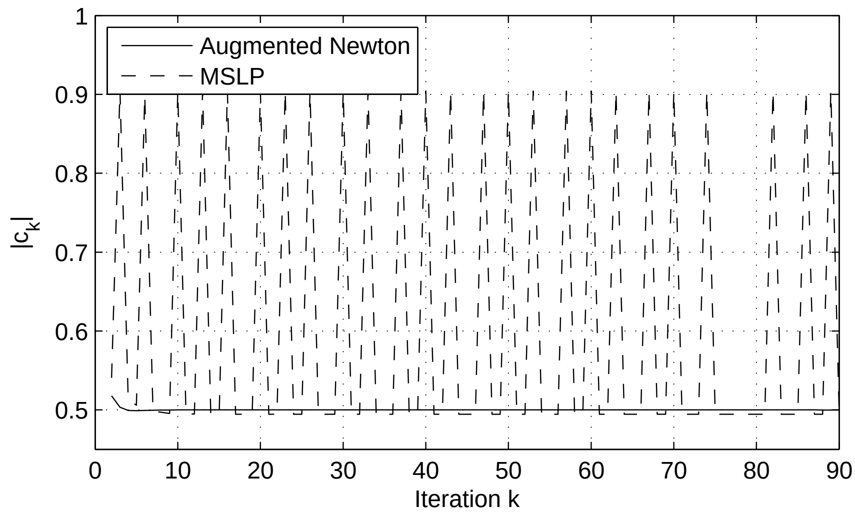


Fig. 3. The convergence factor for Example 2.

## 5 Conclusions

The purpose of this paper has been to provide explicit formulas for the fraction $c_k := (\lambda_{k+1} - \lambda_*)/(\lambda_k - \lambda_*)^p$ in the limit when $k \to \infty$, where $\lambda_k$ are generated by two different methods for nonlinear eigenvalue problems. We have shown formulas for simple as well as double eigenvalues. Even though the limit $c_k$ apparently does not always exist, we find expressions which can be used to analyze the asymptotic behavior.

Finally, we comment on extensions and further interpretations of the presented results. Convergence factors $c := \lim_{k\to\infty} c_k$ are often used to understand and improve numerical methods. Hence, our results open up a possibility to construct error indicators and further characterization of numerical instabilities. We have also provided further understanding to the two methods MSLP and augmented Newton by showing that for special normalization vectors they have the same convergence factors.

## Acknowledgements

The author would like to dedicate this paper to Prof. Dr. Heinrich Voss who introduced the author to the field of nonlinear eigenvalue problems and has always provided us with new interesting ideas, theory, problems and methods to solve problems in the field of numerical linear algebra. We are also grateful that the anonymous referees provided technically detailed comments which improved the presentation and clarity of the article in general and in particular the understanding of the degenerate situation in Remark 2.

This article present 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, and the project STRT1-09/33 of the K.U.Leuven Research Foundation. A part of this work was conducted at T.U. Braunschweig, Germany.

## References

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

[2] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, F. Tisseur, NLEVP: A collection of nonlinear eigenvalue problems, Tech. rep., Manchester Institute for

Mathematical Sciences (2008).

[3] T. Betcke, H. Voss, A Jacobi-Davidson type projection method for nonlinear eigenvalue problems, Future Generation Computer Systems 20 (3) (2004) 363–372.

[4] K. Engelborghs, T. Luzyanina, D. Roose, Numerical bifurcation analysis of delay differential equations using DDE-BIFTOOL, ACM Trans. Math. Softw. 28 (1) (2002) 1–24.

[5] M. Freitag, A. Spence, Convergence of inexact inverse iteration with application to preconditioned iterative solves, BIT 47 (2007) 27–44.

[6] I. Gohberg, P. Lancaster, L. Rodman, Matrix polynomials, Academic press, 1982.

[7] R. Hryniv, P. Lancaster, On the perturbation of analytic matrix functions, Integral Equations Oper. Theory 34 (3) (1999) 325–338.

[8] I. C. F. Ipsen, Computing an eigenvector with inverse iteration, SIAM Rev. 39 (2) (1997) 254–291.

[9] N. K. Jain, K. Singhal, On Kublanovskaya's approach to the solution of the generalized latent value problem for functional lambda-matrices, SIAM J. Numer. Anal. 20 (1983) 1062–1070.

[10] E. Jarlebring, The spectrum of delay-differential equations: numerical methods, stability and perturbation, Ph.D. thesis, TU Braunschweig (2008).

[11] E. Jarlebring, W. Michiels, Invariance properties in the root sensitivity of time-delay systems with double imaginary roots, Automatica 46 (2010) 1112–1115.

[12] D. Kressner, A block Newton method for nonlinear eigenvalue problems, Numer. Math. 114 (2) (2009) 355–372.

[13] V. Kublanovskaya, On an approach to the solution of the generalized latent value problem for $\lambda$-matrices, SIAM J. Numer. Anal. 7 (1970) 532–537.

[14] P. Lancaster, A generalized Rayleigh quotient iteration for lambda-matrices, Arch. Ration. Mech. Anal. 8 (1961) 309–322.

[15] P. Lancaster, Lambda-matrices and vibrating systems, Mineola, NY: Dover Publications, 2002.

[16] R.-C. Li, Compute multiply nonlinear eigenvalues, J. Comput. Math. 10 (1) (1992) 1–20.

[17] V. Mehrmann, H. Voss, Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods, GAMM Mitteilungen 27 (2004) 121–152.

[18] W. Michiels, S.-I. Niculescu, Stability and Stabilization of Time-Delay Systems: An Eigenvalue-Based Approach, Advances in Design and Control 12, SIAM Publications, Philadelphia, 2007.

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

[20] J. Ortega, W. Rheinboldt, Iterative solution of nonlinear equations in several variables, SIAM, Society for Industrial and Applied Mathematics, 2000.

[21] A. Ostrowski, On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. I, II, Arch. Ration. Mech. Anal. 1 (1959) 233–241.

[22] G. Peters, J. Wilkinson, Inverse iterations, ill-conditioned equations and Newton's method, SIAM Rev. 21 (1979) 339–360.

[23] L. Rall, A note on the convergence of Newton's method, SIAM J. Numer. Anal. 11 (1974) 34–36.

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

[25] K. Schreiber, Nonlinear eigenvalue problems: Newton-type methods and nonlinear Rayleigh functionals, Ph.D. thesis, TU Berlin (2008).

[26] J. Schröder, Über das Newtonsche Verfahren, Arch. Ration. Mech. Anal. 1 (1957) 154–180.

[27] H. Schwetlick, K. Schreiber, A primal-dual Jacobi-Davidson-like method for nonlinear eigenvalue problems, Tech. Rep. ZIH-IR-0613, pp. 1–20, Techn. Univ. Dresden, Zentrum für Informationsdienste und Hochleistungsrechnen (2006).

[28] G. L. Sleijpen, A. G. Booten, D. R. Fokkema, H. A. van der Vorst, Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems, BIT 36 (3) (1996) 595–633.

[29] F. Tisseur, K. Meerbergen, The quadratic eigenvalue problem, SIAM Rev. 43 (2) (2001) 235–286.

[30] H. Unger, Nichtlineare Behandlung von Eigenwertaufgaben, Z. Angew. Math. Mech. 30 (1950) 281–282, english translation: http://www.math.tu-dresden.de/~schwetli/Unger.html.

[31] H. Voss, 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.

[32] W. H. Yang, A method for eigenvalues of sparse lambda-matrices, Int. J. Numer. Methods Eng. 19 (1983) 943–948.