Chebyshev interpolation for nonlinear eigenvalue problems

Cedric Effenberger · Daniel Kressner

Abstract This work is concerned with numerical methods for matrix eigenvalue problems that are nonlinear in the eigenvalue parameter. In particular, we focus on eigenvalue problems for which the evaluation of the matrix-valued function is computationally expensive. Such problems arise, e.g., from boundary integral formulations of elliptic PDE eigenvalue problems and typically exclude the use of established nonlinear eigenvalue solvers. Instead, we propose the use of polynomial approximation combined with non-monomial linearizations. Our approach is intended for situations where the eigenvalues of interest are located on the real line or, more generally, on a pre-specified curve in the complex plane. A first-order perturbation analysis for nonlinear eigenvalue problems is performed. Combined with an approximation result for Chebyshev interpolation, this shows exponential convergence of the obtained eigenvalue approximations with respect to the degree of the approximating polynomial. Preliminary numerical experiments demonstrate the viability of the approach in the context of boundary element methods.

Keywords Nonlinear eigenvalue problem · Chebyshev interpolation · Linearization · Krylov subspace method · Boundary element formulation

Mathematics Subject Classification 47J10 · 35P30 · 41A10

C. Effenberger (✉) · D. Kressner
ANCHP, MATHICSE, EPF Lausanne, Station 8, 1015 Lausanne, Switzerland
e-mail: cedric.effenberger@epfl.ch

D. Kressner
e-mail: daniel.kressner@epfl.ch

1 Introduction

We consider a nonlinear eigenvalue problem of the form

$$T(\lambda)x = 0, \quad x \neq 0, \tag{1.1}$$

for a holomorphic matrix-valued function $T : \Omega \to \mathbb{C}^{n \times n}$. Any pair (x, λ) satisfying (1.1) is called an eigenpair of T, consisting of the eigenvector x and the eigenvalue λ.

Nonlinear eigenvalue problems arise in a number of applications. In the context of partial differential equations (PDEs), the nonlinearity is usually caused by frequency-dependent boundary conditions or material parameters; see [6, 25]. In the context of the finite element method (FEM), additional nonlinearities can be introduced by the use of frequency-dependent basis functions. In a similar manner, a recently proposed boundary element method (BEM) for elliptic PDE eigenvalue problems [29] also leads to nonlinear eigenvalue problems. This will be illustrated for the 3D Laplace eigenvalue problem with Dirichlet boundary conditions:

$$\begin{aligned} -\Delta u &= \lambda^2 u \quad \text{in } \mathcal{D} \subset \mathbb{R}^3, \\ u &= 0 \qquad \text{on } \mathcal{B} := \partial\mathcal{D}. \end{aligned} \tag{1.2}$$

By means of the representation formula for the Helmholtz operator, the problem (1.2) can be reformulated [29] as the boundary integral equation

$$\frac{1}{4\pi} \int_{\mathcal{B}} \frac{e^{i\lambda\|\xi-\eta\|}}{\|\xi-\eta\|} u_n(\eta) \, dS(\eta) = 0 \quad \text{for all } \xi \in \mathcal{B}, \tag{1.3}$$

where u_n is the exterior normal derivative of u. Discretizing a weak formulation of (1.3) by a Galerkin approach eventually leads to a nonlinear eigenvalue problem (1.1). In particular, when using a boundary element space of piecewise constant functions on a surface mesh consisting of triangles $\Delta_1, \ldots, \Delta_n$, the entries of $T(\lambda)$ take the form

$$[T(\lambda)]_{ij} = \frac{1}{4\pi} \int_{\Delta_i} \int_{\Delta_j} \frac{e^{i\lambda\|\xi-\eta\|}}{\|\xi-\eta\|} \, dS(\eta) \, dS(\xi). \tag{1.4}$$

The assembly of T is far more expensive than in typical FEM, due to the nonlocality and singularity of the integral kernel; see, e.g., [27, 28].

The solution of nonlinear eigenvalue problems has been extensively discussed in the literature; see [25] for an overview of classic methods and [8, 14, 20, 31] for more recent developments. However, most existing methods are not appropriate for situations as (1.4), where the evaluation of the matrix-valued function $T(\lambda)$ is very expensive. For example, many methods rely on frequent evaluations of the residual $T(\tilde{\lambda})\tilde{x}$ for approximate eigenpairs $(\tilde{x}, \tilde{\lambda})$. A possible exception are methods based on contour integrals [2, 3, 9]. In this paper, we consider a rather different approach to deal with such nonlinear eigenvalue problems. All evaluations of $T(\lambda)$ are performed once in a preprocessing step to construct a polynomial approximation $P(\lambda)$ to $T(\lambda)$ in the eigenvalue region of interest. This has two advantages: (i) the evaluation of

$P(\lambda)$ is considerably cheaper; (ii) a standard linear eigensolver can be used after an appropriate linearization. Similar approaches have been proposed in [10, 11, 14, 15, 24]; we will discuss their relation to our approach in Sect. 2.1.

The remainder of this paper is structured as follows. In Sect. 2, we present our approach and discuss various details, such as the choice of the polynomial approximation, the linearization, and the linear eigenvalue solver. Also, the concept of invariant pairs is introduced, which provides a robust way of handling several eigenpairs. In Sect. 3, we analyze the impact of the error incurred by the polynomial approximation on the accuracy of the computed invariant pairs. Finally, Sect. 4 contains preliminary numerical experiments, demonstrating the viability of our approach for solving the 3D Laplace eigenvalue problem (1.2) by the BEM.

2 Derivation of the method

2.1 Polynomial approximation

In the following, we constrain ourselves to the situation that the eigenvalue region of interest is the interval $[-1, 1]$. This covers general finite intervals or even prescribed smooth curves in the complex plane through an appropriate analytic reparameterization.

The main idea of our approach for solving the nonlinear eigenvalue problem (1.1) is to replace T by a polynomial approximant P. More specifically, for fixed interpolation nodes $\lambda_0, \lambda_1, \ldots, \lambda_d \in [-1, 1]$, we replace T by the unique matrix polynomial P of degree at most d, satisfying the interpolation conditions

$$P(\lambda_j) = T(\lambda_j), \quad j = 0, \ldots, d \tag{2.1}$$

This leads to the polynomial eigenvalue problem

$$P(\lambda)x = 0. \tag{2.2}$$

We expect that a small interpolation error will lead to a small error in the eigenpairs. This expectation is confirmed by an error analysis in Sect. 3. Standard choices of interpolation nodes include Chebyshev nodes of the first kind,

$$\lambda_j = \cos\left(\frac{j + \frac{1}{2}}{d+1}\pi\right), \quad j = 0, \ldots, d, \tag{2.3}$$

and of the second kind,

$$\lambda_j = \cos\left(\frac{j}{d}\pi\right), \quad j = 0, \ldots, d. \tag{2.4}$$

As is well known and shown for our particular situation in Proposition 3.1 below, the interpolation error of such a Chebyshev interpolant decays exponentially with d, and hence we expect that a moderate polynomial degree will be sufficient to ensure good accuracy.

Remark 2.1 The idea of using polynomial approximations to solve nonlinear eigenvalue problems is not new. In particular, for boundary integral formulations of PDE eigenvalue problems, the use of Taylor approximations is a well-known approach in applications [15–19, 26]. Polynomial and Padé approximation techniques are frequently used in deriving absorbing boundary conditions for Helmholtz and Schrödinger eigenvalue problems; see, e.g., [10]. In this context, Botchev et al. [11] have proposed a polynomial approximation technique based on empirical interpolation. Very recently, Meerbergen [24] has proposed a polynomial approximation technique with a heuristic selection of interpolation nodes based on Ritz values. While this approach is more general, it is likely to require more evaluations of T than our approach and there is no convergence analysis available to the best of our knowledge.

Finally, we mention that a very different approach has been proposed in [14], which constructs a polynomial approximation in the course of applying an Arnoldi-like method. This approach, however, relies on the construction of an integral operator representing (1.1).

2.2 Linearization of the polynomial eigenproblem

Once we have substituted the interpolating polynomial P for the nonlinear function T, we are facing the need to solve the resulting polynomial eigenvalue problem (2.2). A popular way of solving polynomial eigenvalue problems is to transform them into an equivalent (generalized) linear eigenvalue problem and then apply standard techniques. This transformation is not at all unique [23]. A common choice are companion linearizations based on an expansion of the polynomial P in the monomial basis. However, there is a number of inconveniences associated with the use of the monomial basis. First of all, the coefficient matrices of P with respect to the monomial basis are not readily available from the construction in Sect. 2.1. Moreover, especially for higher degrees of P, this transformation may cause numerical difficulties. Therefore we employ a different linearization scheme described in [1], which is based on an expansion of P in the polynomial basis formed by the first $d + 1$ Chebyshev polynomials (of the first kind),

$$P(\lambda) = P_0 \tau_0(\lambda) + \cdots + P_d \tau_d(\lambda). \tag{2.5}$$

Combining the expansion (2.5) with the interpolation conditions (2.1), through which P is defined, leads to

$$T(\lambda_j) = \sum_{i=0}^{d} P_i \cos \frac{i(j + \frac{1}{2})\pi}{d + 1}, \quad j = 0, \ldots, d$$

if the Chebyshev nodes of the first kind in (2.3) are used as interpolation nodes, and to

$$T(\lambda_j) = \sum_{i=0}^{d} P_i \cos \frac{ij\pi}{d}, \quad j = 0, \ldots, d$$

for the Chebyshev nodes of the second kind in (2.4). In both cases, the coefficient matrices $P_0, \ldots, P_d$ can be efficiently computed by a sequence of inverse discrete

cosine transforms of type III or type I, respectively. For details, the reader is referred to, e.g., [4].

For the sake of completeness, let us recall the linearization technique from [1] for the polynomial eigenvalue problem

$$\big(P_0\tau_0(\lambda) + \cdots + P_d\tau_d(\lambda)\big)x = 0 \tag{2.6}$$

expressed in the Chebyshev basis. Introducing the vectors $x_k := \tau_k(\lambda)x$, the polynomial eigenvalue problem (2.6) can be rewritten as

$$P_0 x_0 + \cdots + P_d x_d = 0. \tag{2.7}$$

Furthermore, the three-term recurrence for the Chebyshev polynomials τ_k yields

$$x_1 = \lambda x_0 \quad \text{and} \quad x_k = 2\lambda x_{k-1} - x_{k-2}, \quad k = 2, \ldots, d.$$

By means of the preceding identities, we can eliminate x_d from the polynomial eigenvalue problem (2.7). The remaining equation,

$$P_0 x_0 + \cdots + P_{d-3} x_{d-3} + (P_{d-2} - P_d)x_{d-2} + P_{d-1}x_{d-1} + 2\lambda P_d x_{d-1} = 0,$$

can be reformulated as the equivalent (generalized) linear eigenvalue problem

$$\mathcal{L}_0 y = \lambda \mathcal{L}_1 y \tag{2.8}$$

with $y = [x_0^\mathsf{T}, \ldots, x_{d-1}^\mathsf{T}]^\mathsf{T}$ and

$$
\mathcal{L}_0 = \begin{bmatrix}
0 & I & & & & \\
I & 0 & I & & & \\
& \ddots & \ddots & & \ddots & \\
& & I & 0 & & I \\
-P_0 & \cdots & -P_{d-3} & P_d - P_{d-2} & -P_{d-1}
\end{bmatrix},
$$

$$
\mathcal{L}_1 = \begin{bmatrix}
I & & & & \\
& 2I & & & \\
& & \ddots & & \\
& & & 2I & \\
& & & & 2P_d
\end{bmatrix}.
$$

$$\tag{2.9}$$

It has been shown in [1] that (2.8)–(2.9) is a strong linearization of the polynomial eigenvalue problem (2.6).

2.3 Solution of the linearized eigenproblem

The resulting linearizations (2.8)–(2.9) are typically large. Their size is equal to the size of the original nonlinear eigenvalue problem times the degree of the interpolating polynomial P. The eigenvalues of interest are those lying in or close to the real

interval $[-1, 1]$. As these are likely to be interior eigenvalues of the problem, we pursue a shift-and-invert strategy for their computation. A natural choice for the shift is the center of the interval, i.e., zero. This choice leads us to the computation of a few eigenvalues of largest magnitude for the matrix $\Phi = \mathcal{L}_0^{-1}\mathcal{L}_1$, which can be easily accomplished using Krylov subspace methods, such as the implicitly restarted Arnoldi algorithm [21].

Krylov subspace methods crucially depend on repeated matrix-vector multiplication with the matrix Φ, which, in our case, can be broken up into successive multiplications by $\mathcal{L}_1$ and $\mathcal{L}_0^{-1}$. Whereas the multiplication by the block diagonal matrix $\mathcal{L}_1$ can be performed efficiently in a straightforward manner, the question of how to invert $\mathcal{L}_0$ is more subtle and will be treated subsequently.

The linear system $\mathcal{L}_0 x = y$ has the block structure

$$
\begin{bmatrix}
0 & I & & & & \\
I & 0 & I & & & \\
& \ddots & \ddots & & \ddots & \\
& & I & 0 & & I \\
-P_0 & \cdots & -P_{d-3} & P_d - P_{d-2} & & -P_{d-1}
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_1 \\
\vdots \\
x_{d-2} \\
x_{d-1}
\end{bmatrix}
=
\begin{bmatrix}
y_0 \\
y_1 \\
\vdots \\
y_{d-2} \\
y_{d-1}
\end{bmatrix},
\qquad (2.10)
$$

where we have partitioned the vectors x and y in accordance with $\mathcal{L}_0$. The odd-numbered block rows of (2.10) amount to the recursion

$$
x_1 = y_0, \quad x_{2j+1} = y_{2j} - x_{2j-1}, \quad j = 1, 2, 3, \ldots,
$$

which permits us to compute the entries $x_1, x_3, x_5, \ldots$ of the solution. In a similar fashion, the even-numbered block rows give

$$
x_{2j} = \hat{y}_{2j-1} + (-1)^j x_0, \quad j = 1, 2, 3, \ldots, \qquad (2.11)
$$

where the vectors $\hat{y}_{2j-1}$ are determined by the recurrence

$$
\hat{y}_1 = y_1, \quad \hat{y}_{2j+1} = y_{2j+1} - \hat{y}_{2j-1}, \quad j = 1, 2, 3, \ldots.
$$

Inserting identity (2.11) into the last block row of the linear system (2.10), we arrive at the equation

$$
(-P_0 + P_2 - P_4 + P_6 - \cdots + \cdots)x_0
$$
$$
= (P_1 x_1 + P_3 x_3 + P_5 x_5 + \cdots) + (P_2 \hat{y}_1 + P_4 \hat{y}_3 + P_6 \hat{y}_5 + \cdots), \quad (2.12)
$$

which needs to be solved for x_0. An LU factorization of the system matrix $(-P_0 + P_2 - P_4 + P_6 - \cdots + \cdots)$ should be computed once in a preprocessing step before the actual Krylov subspace method is invoked. In this way, each application of $\mathcal{L}_0^{-1}$ requires only one pair of forward and backward solves. After x_0 has been computed, the remaining components $x_2, x_4, x_6, \ldots$ are determined via (2.11).

2.4 Invariant pairs

Often, we are interested in computing more than one eigenvalue and eigenvector. As discussed in [9, 20], the concept of invariant pairs provides a robust way of representing and computing several eigenpairs simultaneously.

Definition 2.1 A pair $(X, \Lambda) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ is called an *invariant pair* of the nonlinear eigenvalue problem (1.1) if

$$\mathbf{T}(X, \Lambda) := \frac{1}{2\pi \mathbf{i}} \int_\Gamma T(z) X (zI - \Lambda)^{-1} \, \mathrm{d}z = 0, \tag{2.13}$$

where Γ is a contour containing the eigenvalues of Λ in its interior.

Any holomorphic matrix-valued function T can be expressed as

$$T(\lambda) = T_1 f_1(\lambda) + \cdots + T_K f_K(\lambda) \tag{2.14}$$

with scalar holomorphic functions $f_1, \ldots, f_K$ and constant coefficient matrices $T_1, \ldots, T_K$. For example, one may consider the entries of T individually to arrive at a representation (2.14) with $K = n^2$ terms. Using (2.14), the characterization (2.13) becomes equivalent to

$$T_1 X f_1(\Lambda) + \cdots + T_K X f_K(\Lambda) = 0, \tag{2.15}$$

where $f_1(\Lambda), \ldots, f_K(\Lambda)$ are to be understood as matrix functions in the usual sense [13].

Equation (2.13), or equivalently (2.15), need to be complemented with a normalization condition to avoid degenerate situations, such as $X = 0$. As discussed in [20], a suitable condition is to require that there is an integer $\ell \geq 1$ such that the matrix

$$\mathbf{V}_\ell(X, \Lambda) = \begin{bmatrix} X \\ X\Lambda \\ \vdots \\ X\Lambda^{\ell-1} \end{bmatrix} \in \mathbb{C}^{\ell n \times m}$$

has full column rank. An invariant pair satisfying this condition is called *minimal*. For any minimal invariant pair (X, Λ), the eigenvalues of Λ are eigenvalues of the nonlinear eigenvalue problem (1.1); see, e.g., [20]. If additionally the algebraic multiplicities of the eigenvalues of Λ match those of (1.1), then (X, Λ) is called a *simple invariant pair*.

2.5 Extraction of invariant pairs

An invariant pair (X, Λ) of the matrix polynomial (2.5) satisfies

$$P_0 X \tau_0(\Lambda) + \cdots + P_d X \tau_d(\Lambda) = 0. \tag{2.16}$$

In the following, we will show how such an invariant pair can be obtained from the corresponding linearization $\mathcal{L}_0 - \lambda\mathcal{L}_1$ defined in (2.8)–(2.9). A linear eigensolver, such as the Arnoldi method discussed in Sect. 2.3, applied to $\mathcal{L}_0 - \lambda\mathcal{L}_1$ yields an invariant pair $(Y, \Lambda) \in \mathbb{C}^{dn \times m} \times \mathbb{C}^{m \times m}$. For this special case, the characterization above takes the form

$$\mathcal{L}_0 Y = \mathcal{L}_1 Y \Lambda. \tag{2.17}$$

Note that span(Y) is usually called an invariant subspace.

Partitioning $Y = [X_0^\mathsf{T}, \ldots, X_{d-1}^\mathsf{T}]^\mathsf{T}$ with $X_j \in \mathbb{C}^{n \times m}$ and exploiting the block structure of $\mathcal{L}_0, \mathcal{L}_1$, the first $d-1$ block rows of (2.17) amount to

$$X_1 = X_0 \Lambda, \quad X_{k-2} + X_k = 2X_{k-1}\Lambda, \quad k = 2, \ldots, d-1. \tag{2.18}$$

A simple induction using the three-term recurrence for the Chebyshev polynomials τ_k shows that (2.18) implies $X_k = X_0 \tau_k(\Lambda)$ for $k = 0, \ldots, d-1$. Inserting these relations into the last block row of (2.17),

$$-\sum_{k=0}^{d-1} P_k X_k + P_d X_{d-2} = 2P_d X_{d-1}\Lambda,$$

and rearranging terms yields

$$-\sum_{k=0}^{d-1} P_k X_0 \tau_k(\Lambda) = P_d X_0 \big(2\tau_{d-1}(\Lambda)\Lambda - \tau_{d-2}(\Lambda)\big).$$

Exploiting once more the three-term recursion finally shows that the pair (X_0, Λ) satisfies (2.16) and is therefore an invariant pair of the polynomial (2.5).

Remark 2.2 The discussion above suggests a simple extraction procedure: Given an invariant pair (Y, Λ) of the linearization, an invariant pair of the polynomial is obtained as (X_0, Λ), where X_0 is the first block component of Y. In finite-precision arithmetic, this relation is affected by roundoff error. Numerical aspects of such extraction procedures have been discussed in [7] for the class of so-called L_1 linearizations. Moreover, alternative algorithms for extraction have been proposed in [7], which turn out to be numerically more robust in certain situations. However, in our particular setting, $|\tau_k(\lambda)| \leq 1$ for all eigenvalues λ of interest. Hence, we expect that X_0 is a dominant component of Y and therefore a numerically reasonable choice. This is confirmed by our numerical experiments, which also show that suitable adaptions of the alternative algorithms mentioned above do not result in significant accuracy improvements.

The accuracy of the extracted invariant pair (X_0, Λ) can be further refined by applying a Newton iteration as described in [7, 20].

3 Error analysis

Instead of the original nonlinear eigenvalue problem (1.1), our method from Sect. 2 solves the perturbed problem

$$(T + \Delta T)(\lambda)x = 0,$$

where the perturbation $\Delta T = P - T$ amounts to the interpolation error. It is therefore important to analyze the impact of such a perturbation on the eigenvalues or, more generally, on the invariant pairs. For this purpose, we will derive a general perturbation result for nonlinear eigenvalue problems. This will be combined with a polynomial approximation result to establish convergence rates for our method.

3.1 First-order perturbation theory

In the following, let T_0 be a holomorphic function on some domain $\Omega \subset \mathbb{C}$ with values in $\mathbb{C}^{n \times n}$. Furthermore, we will assume that T_0 is bounded on Ω with respect to the Frobenius norm $\|\cdot\|_{\mathsf{F}}$ and regular; i.e., $\det T_0(\lambda)$ does not vanish identically for all $\lambda \in \Omega$.

Let $(X_0, \Lambda_0) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ be a minimal invariant pair of T_0 such that all eigenvalues of Λ_0 are contained inside Ω. Then the triple (X_0, Λ_0, T_0) constitutes a solution of the nonlinear equation

$$F(X, \Lambda, T) = 0 \tag{3.1}$$

with

$$F : \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m} \times B(\Omega) \to \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m},$$

$$(X, \Lambda, T) \mapsto \left(\frac{1}{2\pi \mathbf{i}} \int_{\Gamma} T(z) X (zI - \Lambda)^{-1} \, \mathrm{d}z, \; W^{\mathsf{H}} \big[\mathbf{V}_{\ell}(X, \Lambda) - \mathbf{V}_{\ell}(X_0, \Lambda_0) \big] \right). \tag{3.2}$$

Here, Γ is a contour in Ω containing the eigenvalues of Λ_0 in its interior, and $B(\Omega)$ denotes the Banach space of all bounded, holomorphic, $\mathbb{C}^{n \times n}$-valued functions on $\Omega \subset \mathbb{C}$, equipped with the supremum norm

$$\|\cdot\|_{\infty} : B(\Omega) \to \mathbb{R}, \quad T \mapsto \|T\|_{\infty} := \sup_{\lambda \in \Omega} \|T(\lambda)\|_{\mathsf{F}}.$$

Note that the convergence of functions in the supremum norm amounts to uniform convergence. The first term in (3.2) characterizes the invariance of the pair (X_0, Λ_0), whereas the second term characterizes minimality, provided that the normalization matrix $W \in \mathbb{C}^{\ell n \times m}$ is chosen such that $W^{\mathsf{H}} \mathbf{V}_{\ell}(X_0, \Lambda_0)$ is invertible.

Lemma 3.1 *The mapping F defined above is continuously Fréchet differentiable in a neighborhood of (X_0, Λ_0, T_0).*

Proof As a norm in the space $\mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m} \times B(\Omega)$, we employ

$$\|(\Delta X, \Delta \Lambda, \Delta T)\| := \|\Delta X\|_{\mathsf{F}} + \|\Delta \Lambda\|_{\mathsf{F}} + \|\Delta T\|_{\infty}.$$

Since the continuous differentiability of the second component of F is easily seen, we will only treat the first component $F^{(1)}$ and demonstrate that its derivative is given by the map

$$\mathrm{D}F^{(1)}(X, \Lambda, T)(\Delta X, \Delta \Lambda, \Delta T)$$
$$= \mathrm{D}_X F^{(1)}(X, \Lambda, T)(\Delta X) + \mathrm{D}_\Lambda F^{(1)}(X, \Lambda, T)(\Delta \Lambda) + \mathrm{D}_T F^{(1)}(X, \Lambda, T)(\Delta T)$$

with

$$\mathrm{D}_X F^{(1)}(X, \Lambda, T)(\Delta X) = \frac{1}{2\pi \mathbf{i}} \int_\Gamma T(z) \Delta X (zI - \Lambda)^{-1} \, \mathrm{d}z,$$

$$\mathrm{D}_\Lambda F^{(1)}(X, \Lambda, T)(\Delta \Lambda) = \frac{1}{2\pi \mathbf{i}} \int_\Gamma T(z) X (zI - \Lambda)^{-1} \Delta \Lambda (zI - \Lambda)^{-1} \, \mathrm{d}z,$$

$$\mathrm{D}_T F^{(1)}(X, \Lambda, T)(\Delta T) = \frac{1}{2\pi \mathbf{i}} \int_\Gamma \Delta T(z) X (zI - \Lambda)^{-1} \, \mathrm{d}z.$$

For this purpose, let $(X, \Lambda, T) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m} \times B(\Omega)$ be fixed. We assume Λ to be sufficiently close to Λ_0 so that its eigenvalues still lie inside Ω. Consequently, there exists a contour Γ in Ω which contains all eigenvalues of Λ in its interior. Because the number of eigenvalues is finite, we can w.l.o.g. assume the contour to possess a finite length L. Let $\gamma : [0, 1] \to \Gamma$ be a parametrization of the contour Γ. As Γ touches none of the eigenvalues of Λ, the mapping $\varphi \mapsto \|(\gamma(\varphi)I - \Lambda)^{-1}\|_\mathsf{F}$ is continuous and therefore bounded on the compact interval $[0, 1]$ by some $M > 0$.

Now suppose $\|(\Delta X, \Delta \Lambda, \Delta T)\| < M^{-1}$, implying, in particular, $\|\Delta \Lambda\|_\mathsf{F} < M^{-1}$. Thus, $\|\Delta \Lambda (zI - \Lambda)^{-1}\|_\mathsf{F} < 1$ for arbitrary $z \in \Gamma$, and the Neumann series gives

$$\left[zI - (\Lambda + \Delta \Lambda)\right]^{-1} = (zI - \Lambda)^{-1} + (zI - \Lambda)^{-1} \Delta \Lambda (zI - \Lambda)^{-1} + O\!\left(\|\Delta \Lambda\|_\mathsf{F}^2\right),$$

where the constant implicitly contained in the $O(\|\Delta \Lambda\|_\mathsf{F}^2)$ term is independent of z. Altogether, we obtain

$$\big\| F^{(1)}(X + \Delta X, \Lambda + \Delta \Lambda, T + \Delta T) - F^{(1)}(X, \Lambda, T)$$
$$\qquad - \mathrm{D}F^{(1)}(X, \Lambda, T)(\Delta X, \Delta \Lambda, \Delta T) \big\|_\mathsf{F}$$
$$= \left\| \frac{1}{2\pi \mathbf{i}} \int_\Gamma [\Delta T(z) X + T(z) \Delta X + \Delta T(z) \Delta X](zI - \Lambda)^{-1} \Delta \Lambda (zI - \Lambda)^{-1} \right.$$
$$\qquad \left. + \Delta T(z) \Delta X (zI - \Lambda)^{-1} + O\!\left(\|\Delta \Lambda\|_\mathsf{F}^2\right) \mathrm{d}z \right\|_\mathsf{F}$$
$$= O\!\left(\|(\Delta X, \Delta \Lambda, \Delta T)\|^2\right)$$

confirming the claim that $F^{(1)}$ is differentiable with the derivative $\mathrm{D}F^{(1)}$ stated above. The continuity of $\mathrm{D}F^{(1)}$ can be established by a similar estimate. $\square$

The next result concerns the derivative of F at (X_0, Λ_0, T_0) only with respect to X and Λ, but not T. The corresponding linear operator from $\mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ onto itself will be denoted by $\mathrm{D}_{(X,\Lambda)} F(X_0, \Lambda_0, T_0)$.

Theorem 3.1 [20, Theorem 10] *Let (X_0, Λ_0) be a minimal invariant pair of T_0. Then the derivative $D_{(X,\Lambda)}F$ at (X_0, Λ_0, T_0) is an automorphism on $\mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ if and only if (X_0, Λ_0) is simple.*

In combination with Lemma 3.1, Theorem 3.1 permits us to apply the Implicit Function Theorem to the nonlinear equation (3.1) in the vicinity of a simple invariant pair (X_0, Λ_0) of T_0. This yields the existence of continuously differentiable functions $X : B(\Omega) \to \mathbb{C}^{n \times m}$ and $\Lambda : B(\Omega) \to \mathbb{C}^{m \times m}$ with $X(T_0) = X_0$ and $\Lambda(T_0) = \Lambda_0$ such that

$$F\big(X(T), \Lambda(T), T\big) = 0$$

for all T in a neighborhood of T_0. Moreover, the derivatives with respect to T of these two functions are given by

$$\begin{bmatrix} D_T X(T_0) \Delta T_0 \\ D_T \Lambda(T_0) \Delta T_0 \end{bmatrix} = -\Big(\big[D_{(X,\Lambda)}F(X_0, \Lambda_0, T_0)\big]^{-1} \circ D_T F(X_0, \Lambda_0, T_0)\Big)\Delta T_0,$$

where $[D_{(X,\Lambda)}F(X_0, \Lambda_0, T_0)]^{-1}$ refers to the inverse of the bijective linear operator $D_{(X,\Lambda)}F(X_0, \Lambda_0, T_0) : \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m} \to \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ as established by Theorem 3.1. Setting $T = T_0 + \Delta T_0$, we conclude that the perturbed problem $(T_0 + \Delta T_0)(\lambda)x = 0$ has an invariant pair (X, Λ) satisfying

$$\begin{bmatrix} X \\ \Lambda \end{bmatrix} = \begin{bmatrix} X_0 \\ \Lambda_0 \end{bmatrix} - \Big(\big[D_{(X,\Lambda)}F(X_0, \Lambda_0, T_0)\big]^{-1} \circ D_T F(X_0, \Lambda_0, T_0)\Big)\Delta T_0 + o\big(\|\Delta T_0\|_\infty\big).$$
(3.3)

The main result of this section is summarized in the subsequent theorem.

Theorem 3.2 *Let T_0 and ΔT_0 be bounded, holomorphic, $\mathbb{C}^{n \times n}$-valued functions on some domain $\Omega \subset \mathbb{C}$ and suppose that T_0 is regular. Let $(X_0, \Lambda_0) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ be a simple invariant pair of T_0. If $\|\Delta T_0\|_\infty$ is sufficiently small, then there exists an invariant pair $(X, \Lambda) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}$ of the perturbed problem $T_0 + \Delta T_0$ satisfying (3.3) with F defined as in (3.2).*

3.2 Convergence rates

We will now apply Theorem 3.2 to analyze the method from Sect. 2. To this end, we assume that T is a regular, analytic function on the real interval $[-1, 1]$ with values in $\mathbb{C}^{n \times n}$, and hence can be extended to a holomorphic function on a neighborhood of this interval in the complex plane. For simplicity, this extension will also be referred to as T. In particular, we can choose $\rho > \rho_0 > 1$ such that the Bernstein ellipse

$$E_\rho := \big\{\cos(t - \mathbf{i}\ln \bar\rho) : t \in [0, 2\pi], \ \bar\rho \in [1, \rho]\big\}$$

is contained in the analyticity domain of T. Moreover, the holomorphic extension of T obviously inherits its regularity. With this notation, we obtain the following convergence result.

Corollary 3.1 *Let T be as above and let $P^{(d)}$ denote the interpolating polynomial of degree d for T with respect to the Chebyshev nodes of either the first or the second kind. Let (X, Λ) be a simple invariant pair of T such that all eigenvalues of Λ lie in the real interval $[-1, 1]$. Then there exists a sequence (X_d, Λ_d) of invariant pairs belonging to the polynomials $P^{(d)}$, which converges to (X, Λ) exponentially as $d \to \infty$.*

Proof Choose $\Omega = E_{\rho_0}$ with $\rho > \rho_0 > 1$ as above, and set $\Delta T^{(d)} := P^{(d)} - T$, $d \in \mathbb{N}_0$. Because Ω is compact, we have $T, \Delta T^{(d)} \in B(\Omega)$ for all $d \in \mathbb{N}_0$. Consequently, by Theorem 3.2, there exists an invariant pair of $P^{(d)}$ satisfying

$$\left\| \begin{bmatrix} X_d \\ \Lambda_d \end{bmatrix} - \begin{bmatrix} X \\ \Lambda \end{bmatrix} \right\|_{\mathsf{F}} = O\big(\|\Delta T^{(d)}\|_\infty\big).$$

Since the interpolation error $\Delta T^{(d)}$ converges to zero exponentially according to Proposition 3.1 below, the assertion follows. □

As we have seen, the proof of Corollary 3.1 relies on convergence estimates of the Chebyshev interpolant inside the Bernstein ellipse E_{ρ_0}. These will be covered by the subsequent proposition, which is a variation of classical polynomial approximation results [22].

Proposition 3.1 *Let $T : U \to \mathbb{C}^{n \times m}$ be holomorphic in a neighborhood U of the Bernstein ellipse E_ρ with $\rho > \rho_0 > 1$ and let $P^{(d)}$ denote the interpolating polynomial of degree d for T with respect to the Chebyshev nodes of either the first or the second kind. Then there exists a constant $C > 0$ depending only on T, ρ, and ρ_0 such that for all $\lambda \in E_{\rho_0}$*

$$\|T(\lambda) - P^{(d)}(\lambda)\|_{\mathsf{F}} \le C \left(\frac{\rho_0}{\rho} \right)^d.$$

Proof Depending on what kind of Chebyshev nodes are used, we define $(Q_d)_{d \in \mathbb{N}_0}$ to be the sequence of Chebyshev polynomials of the first or second kind, respectively. In either case, the interpolation nodes are the zeroes of the polynomials Q_d.

In the following, we will show the claim for the Chebyshev nodes of the first kind. In this case, it is well known that

$$Q_d(\cos \theta) = \cos(d\theta). \tag{3.4}$$

The statement for the Chebyshev nodes of the second kind then follows by similar arguments using the identity

$$Q_d(\cos \theta) \sin \theta = \sin\big((d + 1)\theta\big)$$

instead and is therefore omitted.

Let $\lambda \in E_{\rho_0}$. If λ is identical with one of the interpolation nodes, the claimed inequality trivially holds for any $C > 0$. Thus, we may assume w.l.o.g. that $Q_d(\lambda) \ne 0$. For fixed $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, we consider the (i, j)-th component T_{ij}

of T. By applying the residue theorem to the function $z \mapsto \frac{T_{ij}(z)}{(z-\lambda)Q_d(z)}$ and exploiting that all roots of its denominator are simple, one shows that the interpolation error satisfies

$$T_{ij}(\lambda) - P_{ij}^{(d)}(\lambda) = \frac{Q_d(\lambda)}{2\pi \mathbf{i}} \int_{\partial E_\rho} \frac{T_{ij}(z)}{(z-\lambda)Q_d(z)}\, dz. \tag{3.5}$$

We proceed by individually estimating the factors on the right-hand side of (3.5). To begin with, we notice that $z \in \partial E_\rho$ can be expressed as $z = \cos(t - \mathbf{i} \ln \rho)$ for some $t \in [0, 2\pi]$, and hence $Q_d(z) = \cos(dt - \mathbf{i}d \ln \rho)$ due to (3.4). A simple calculation then reveals that

$$|Q_d(z)|^2 = \tfrac{1}{4}(\rho^d - \rho^{-d})^2 + \cos^2(dt),$$

implying the estimate

$$\tfrac{1}{2}(\rho^d - \rho^{-d}) \le |Q_d(z)| \le \tfrac{1}{2}(\rho^d + \rho^{-d})$$

because dt is real. Analogously, $\lambda \in E_{\rho_0}$ can be written as $\lambda = \cos(s - \mathbf{i} \ln \bar\rho)$ for some $s \in [0, 2\pi]$, $\bar\rho \in [1, \rho_0]$, and we conclude that

$$|Q_d(\lambda)| \le \tfrac{1}{2}(\bar\rho^d + \bar\rho^{-d}) \le \tfrac{1}{2}(\rho_0^d + \rho_0^{-d}).$$

Furthermore, $|z - \lambda|$ is bounded from below by the minimal distance between E_{ρ_0} and ∂E_ρ, which is given by $\mathrm{dist}(E_{\rho_0}, \partial E_\rho) = \tfrac{1}{2}[\rho + \rho^{-1} - (\rho_0 + \rho_0^{-1})]$ due to geometric considerations. Finally, $|T_{ij}(z)| \le \|T(z)\|_\mathsf{F} \le \|T\|_\infty$. Taking absolute values in (3.5) and inserting the above estimates, we obtain the bound

$$\left| T_{ij}(\lambda) - P_{ij}^d(\lambda) \right| \le \frac{L_\rho \|T\|_\infty}{2\pi\, \mathrm{dist}(E_{\rho_0}, \partial E_\rho)} \cdot \frac{\rho_0^d + \rho_0^{-d}}{\rho^d - \rho^{-d}},$$

where L_ρ is the circumference of the Bernstein ellipse E_ρ. The proof is completed by combining the bounds for all components of T and taking into account that $\rho_0^{-d} \to 0$ and $\rho^{-d} \to 0$ as $d \to \infty$. □

3.3 Spurious eigenvalues

Based on real interpolation nodes, the interpolating polynomials $P^{(d)}$ tend to be accurate only in the vicinity of the real axis. Away from the real axis, the approximation quality quickly deteriorates. This might cause the appearance of spurious eigenvalues; i.e., eigenvalues of the interpolating polynomial which do not approximate any eigenvalue of the original nonlinear eigenvalue problem, in the sense that the associated residual is large. However, the subsequent result shows that this problem does not occur for sufficiently large degree d.

Corollary 3.2 *Let the assumptions of Proposition 3.1 hold. Then for every $\lambda \in E_{\rho_0}$ such that $T(\lambda)$ is non-singular (i.e., λ is not an eigenvalue of T), there exists $d_0 \in \mathbb{N}_0$ such that $P^{(d)}(\lambda)$ is non-singular (i.e., λ is not an eigenvalue of $P^{(d)}$) for all $d \ge d_0$.*

Proof Let $\lambda \in E_{\rho_0}$ be fixed. According to Proposition 3.1, we can choose $d_0 \in \mathbb{N}_0$ such that the Frobenius norm of the interpolation error $\Delta T^{(d)}(\lambda) = P^{(d)}(\lambda) - T(\lambda)$ is strictly bounded above by $\|T(\lambda)^{-1}\|_{\mathsf{F}}^{-1}$ for all $d \geq d_0$. Thus, $\|-\Delta T^{(d)}(\lambda)T(\lambda)^{-1}\|_{\mathsf{F}} < 1$ for $d \geq d_0$, and by a Neumann series argument, $P^{(d)}(\lambda) = T(\lambda) + \Delta T^{(d)}(\lambda)$ is invertible. $\qquad\square$

Corollary 3.2 states that in the limit $d \to \infty$, $\lambda \in E_{\rho_0}$ can only be an eigenvalue of $P^{(d)}$ if it is also an eigenvalue of T. Thus, asymptotically, there will be no spurious eigenvalues inside the Bernstein ellipse E_{ρ_0}. Since the interval $[-1, 1]$ is enclosed by E_{ρ_0}, we expect spurious eigenvalues to occur only in some distance to the interval. This motivates the following mechanism for detecting spurious eigenvalues: An eigenvalue is discarded as spurious if its real part lies outside the interval $[-1, 1]$ or its imaginary part exceeds a certain threshold in magnitude. Particularly for nonlinear eigenvalue problems resulting from boundary integral formulations, see (1.4), this strategy leads to considerable computational savings over the conventional approach of checking the residuals for all computed eigenvalues.

4 Numerical Experiments

To assess the performance of the method developed in Sect. 2, we have applied it to a set of test problems. All computations have been performed under MATLAB 7.10 (R2010a) on a cluster of 24 Intel Xeon X5650 processors with 72 GB of shared memory. The reported computing times are averages over 20 identical runs. For the solution of the linearized eigenvalue problems, we have utilized a MATLAB implementation of the Arnoldi algorithm with Krylov-Schur restarting [30]. Our preliminary implementation can be found under http://www.math.ethz.ch/~kressner/chebapprox/.

In our first experiment, we consider the Laplace eigenvalue problem (1.2) on the unit cube $\mathcal{D} = [0, 1]^3$ with homogeneous Dirichlet boundary conditions. The eigenvalues and eigenfunctions of this problem are known to be given by

$$\lambda_{j_1, j_2, j_3} = \pi \sqrt{j_1^2 + j_2^2 + j_3^2},$$
$$u_{j_1, j_2, j_3}(x_1, x_2, x_3) = \sin(j_1 \pi x_1) \sin(j_2 \pi x_2) \sin(j_3 \pi x_3).$$

The 6 smallest eigenvalues are summarized in Table 1(a). The occurence of multiple eigenvalues is due to the symmetry of the domain.

We construct a boundary formulation of the problem as described in Sect. 1 and solve the resulting nonlinear eigenvalue problem by the method developed in Sect. 2. To capture the 6 smallest distinct eigenvalues (17, counting multiplicities), we select $[5, 12]$ as the interval of interest. Furthermore, we set the degree of the interpolating polynomial to 12 and compute 20 Ritz values with the implicitly restarted Arnoldi algorithm. The result for a uniform boundary mesh with 2400 triangles is depicted in Fig. 1. The plot also reveals a small number of spurious eigenvalues (marked by crosses). However, as predicted by the results in Sect. 3.3, these spurious eigenvalues are well-separated from the true eigenvalues close to the real axis and can be easily identified.

Table 1 (**a**) The 6 smallest eigenvalues of the negative Laplace operator on the unit cube. (**b**) Reference eigenvalues for the Fichera corner computed by the method from Sect. 2 using an interpolating polynomial of degree 30 on a uniform boundary mesh with 2400 triangles

No.	Eigenvalue	Multiplicity	No.	Eigenvalue
1	5.441398	1	1	6.484702318577543
2	7.695299	3	2	8.142495692472265
3	9.424778	3	3	8.142499335034771
4	10.419484	3	4	9.053846829423080
5	10.882796	1	5	9.716892649192921
6	11.754763	6	6	9.716894006586880
	(a)			(b)

Table 2 Execution times for determining all eigenvalues within the interval of interest using an interpolating polynomial of degree 30 for increasingly fine boundary meshes of the unit cube and the Fichera corner, respectively

h	No. of triangles	Average execution time (s)	
		Unit cube	Fichera corner
$\frac{1}{6}$	864	409	361
$\frac{1}{8}$	1536	1154	1054
$\frac{1}{10}$	2400	2722	2451

Fig. 1 Ritz values obtained by the method in Sect. 2 for the Laplace eigenproblem on the unit cube using a uniform boundary mesh with 2400 triangles. The circles mark Ritz values corresponding to true eigenvalues whereas the crosses indicate spurious eigenvalues
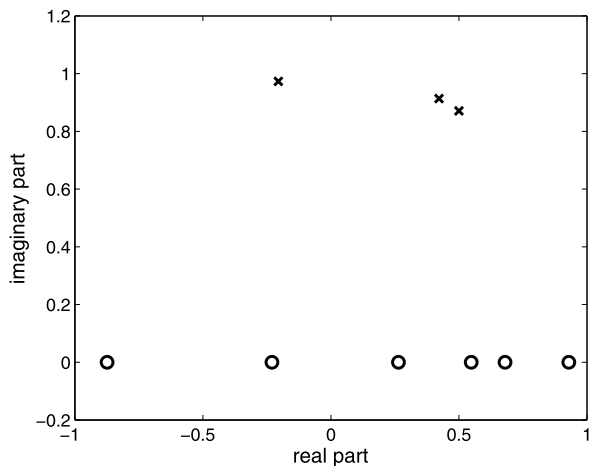

We have experimented with different levels of mesh refinement and different degrees of the interpolating polynomial. Figure 2 shows the spectral convergence of the smallest eigenvalue, computed on a boundary mesh with 2400 triangles ($h = \frac{1}{10}$), towards a reference solution obtained with polynomial degree 30. The convergence behavior for the larger eigenvalues and/or different mesh sizes has been found to be identical, supporting the exponential convergence of eigenvalues predicted by Corollary 3.1. The execution times for the reference solutions are reported in the third column of Table 2. These include the time for setting up the interpolation polynomial,

Fig. 2 Spectral convergence of the smallest eigenvalue for the unit cube using a boundary mesh with $h = \frac{1}{10}$ consisting of 2400 triangles
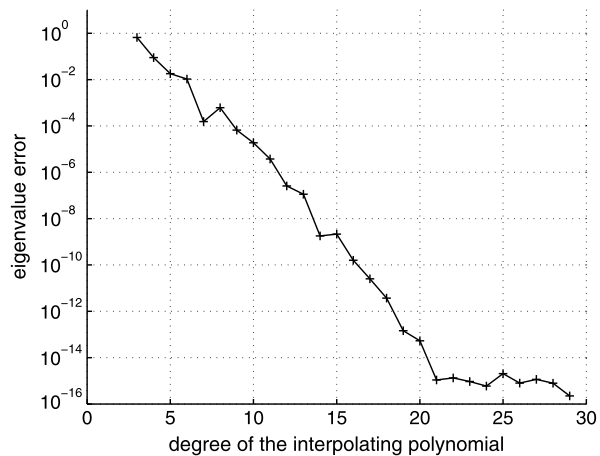
Fig. 3 Relative residual of an invariant pair representing the first 11 eigenvalues of the Laplace eigenvalue problem on the unit cube during three steps of Newton-based iterative refinement. Each *curve* represents a different level of mesh refinement: $h = \frac{1}{6}$, 864 triangles (*dash-dotted* with *circles*), $h = \frac{1}{8}$, 1536 triangles (*dashed* with *crosses*), $h = \frac{1}{10}$, 2400 triangles (*solid* with *plusses*)
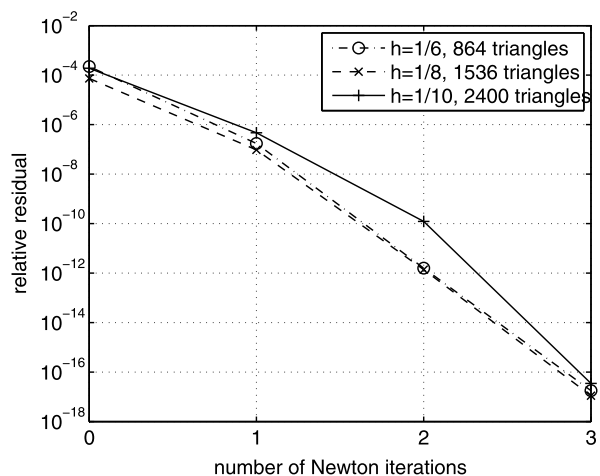

which is the dominating part, as well as the time for the solution of the polynomial eigenvalue problem.

Furthermore, we have implemented and tested the extraction scheme for invariant pairs and their subsequent refinement using Newton iterations outlined in Remark 2.2. For different levels of mesh refinement, we apply the Arnoldi method to compute an approximate invariant pair representing the first 11 eigenvalues employing a degree-20 interpolating polynomial. These initial invariant pairs have relative residuals of about 10^{-4}. We then perform three Newton steps. The result is depicted in Fig. 3. Already after two steps, the relative residual has decreased to an order between 10^{-10} and 10^{-12}. Finally, after the third step, the residual reaches machine accuracy.

As a second experiment, we consider the Laplace eigenvalue problem (1.2) with homogeneous Dirichlet boundary conditions for the Fichera corner $\mathcal{D} = [0, 1]^3 \setminus [\frac{1}{2}, 1]^3$. The boundary element formulation and the resulting nonlinear eigenvalue problem for this case are again obtained as outlined in Sect. 1. However, this time there is no analytic expression for the eigenvalues available.

Fig. 4 Spectral convergence of the smallest eigenvalue for the Fichera corner using a boundary mesh with $h = \frac{1}{10}$ consisting of 2400 triangles
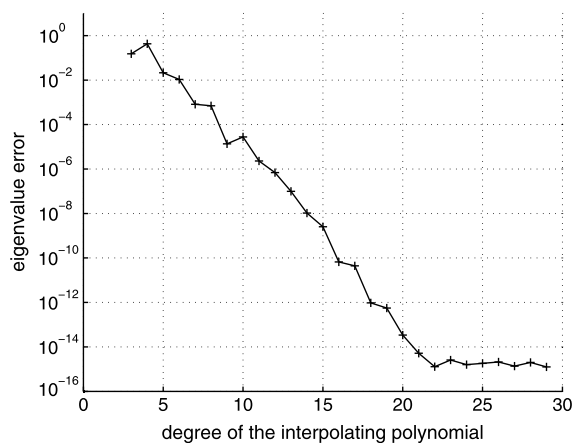
Fig. 5 Relative residual of an invariant pair representing the first 6 eigenvalues of the Laplace eigenvalue problem for the Fichera corner during two steps of Newton-based iterative refinement. Each *curve* represents a different level of mesh refinement: $h = \frac{1}{6}$, 864 triangles (*dash-dotted* with *circles*), $h = \frac{1}{8}$, 1536 triangles (*dashed* with *crosses*), $h = \frac{1}{10}$, 2400 triangles (*solid* with *plusses*)
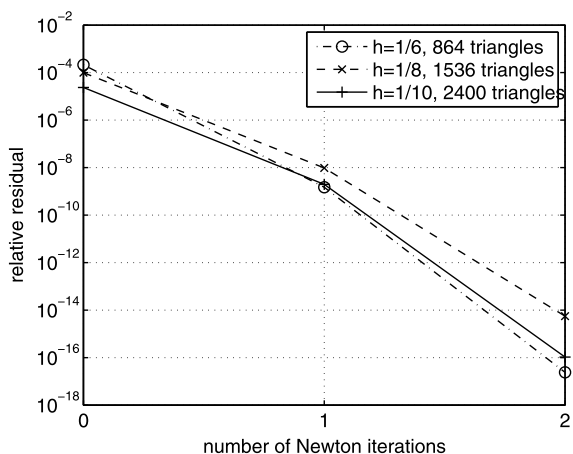
On a uniform boundary mesh with 2400 triangles and with an interpolation polynomial of degree 30, our new method computes the approximate eigenvalues listed in Table 1(b). The spectral convergence of the first of these eigenvalues towards a reference solution computed with polynomial degree 30 is illustrated in Fig. 4. The convergence behavior for the other eigenvalues as well as for different levels of mesh refinement turns out to be the same. As before, exponential convergence of the eigenvalues is observed, in agreement with the statement of Corollary 3.1. The fourth column of Table 2 summarizes the computing times for the reference solutions, consisting of the time to set up the interpolating polynomial and the time to solve the resulting polynomial eigenvalue problem.

Also in this case, we have applied the extraction scheme and Newton-based iterative refinement from Remark 2.2. Starting from an approximate invariant pair with relative residual 10^{-4} obtained from a degree-20 interpolating polynomial, the first refinement step brings the residual down to about 10^{-9}. Already after the second step, the relative residual approaches the level of the machine accuracy. The results are visualized in Fig. 5.

Remark 4.1 When considering the execution times in Table 2, one should take into account that we did not use a highly optimized BEM code for our computations. Possible improvements include, e.g., the exploitation of the inherent parallelism in the computation of the matrix entries (1.4) as well as the use of hierarchical matrix techniques; see, e.g., [5, 12].

5 Conclusions

Based on Chebyshev interpolation, we have proposed a polynomial-based method for solving nonlinear eigenvalue problems. In comparison to existing approaches using polynomial approximation, our approach admits a rigorous error analysis for the obtained approximate eigenpairs. However, it should be emphasized that we have restricted ourselves to problems with eigenvalues of interest in (or close to) a real interval. On the other hand, our approach also allows to deal with very complicated settings where the evaluation of the matrix-valued function is very expensive. Such settings typically impose severe limitations on most existing methods.

Our preliminary numerical experiments indicate that the newly proposed method is effective for nonlinear eigenvalue problems from boundary integral formulations. Its potential for other applications remains to be explored. To address more challenging problems, an eventual implementation should also take advantage of the techniques discussed in Remark 4.1.

References

1. Amiraslani, A., Corless, R.M., Lancaster, P.: Linearization of matrix polynomials expressed in polynomial bases. IMA J. Numer. Anal. **29**(1), 141–157 (2009)
2. Asakura, J., Sakurai, T., Tadano, H., Ikegami, T., Kimura, K.: A numerical method for nonlinear eigenvalue problems using contour integrals. JSIAM Lett. **1**, 52–55 (2009)
3. Asakura, J., Sakurai, T., Tadano, H., Ikegami, T., Kimura, K.: A numerical method for polynomial eigenvalue problems using contour integral. Jpn. J. Ind. Appl. Math. **27**(1), 73–90 (2010)
4. Battles, Z., Trefethen, L.N.: An extension of Matlab to continuous functions and operators. SIAM J. Sci. Comput. **25**(5), 1743–1770 (2004)
5. Bebendorf, M.: Hierarchical LU decomposition-based preconditioners for BEM. Computing **74**(3), 225–247 (2005)
6. Betcke, T., Higham, N.J., Mehrmann, V., Schröder, C., Tisseur, F.: NLEVP: A collection of nonlinear eigenvalue problems. MIMS EPrint 2010.98, Manchester Institute for Mathematical Sciences, The University of Manchester, UK (2010)
7. Betcke, T., Kressner, D.: Perturbation, extraction and refinement of invariant pairs for matrix polynomials. Linear Algebra Appl. **435**, 514–536 (2011)
8. Betcke, T., Voss, H.: A Jacobi-Davidson type projection method for nonlinear eigenvalue problems. Future Gener. Comput. Syst. **20**, 363–372 (2004)
9. Beyn, W.-J.: An integral method for solving nonlinear eigenvalue problems. Linear Algebra Appl. **436**(10), 3839–3863 (2012)
10. Bindel, D.: Analyzing resonances via nonlinear eigenvalues. In: Talk at 7th International Congress on Industrial and Applied Mathematics (ICIAM) in Vancouver, BC, Canada (2011)

11. Botchev, M.A., Sleijpen, G.L.G., Sopaheluwakan, A.: An SVD-approach to Jacobi-Davidson solution of nonlinear Helmholtz eigenvalue problems. Linear Algebra Appl. **431**(3–4), 427–440 (2009)
12. Hackbusch, W., Grasedyck, L., Börm, S.: An introduction to hierarchical matrices. In: Proceedings of EQUADIFF, 10, Prague, 2001, vol. 127, pp. 229–241 (2002)
13. Higham, N.J.: Functions of Matrices: Theory and Computation. SIAM, Philadelphia (2008)
14. Jarlebring, E., Michiels, W., Meerbergen, K.: A linear eigenvalue algorithm for the nonlinear eigenvalue problem. Numer. Math. (2012). To appear. doi:10.1007/s00211-012-0453-0
15. Kamiya, N., Andoh, E., Nogae, K.: Eigenvalue analysis by the boundary element method: new developments. Eng. Anal. Bound. Elem. **12**, 151–162 (1993)
16. Kamiya, N., Andoh, E., Nogae, K.: A new complex-valued formulation and eigenvalue analysis of the Helmholtz equation by boundary element method. Adv. Eng. Softw. **26**(3), 219–227 (1996)
17. Karma, O.: Approximation in eigenvalue problems for holomorphic Fredholm operator functions II (Convergence rate). Numer. Funct. Anal. Optim., **17**(3-4) 389–408 (1996)
18. Kirkup, S.M., Amini, S.: Solution of the Helmholtz eigenvalue problem via the boundary element method. Int. J. Numer. Methods Eng. **36**, 321–330 (1993)
19. Kitahara, M.: Boundary integral equation methods in eigenvalue problems of elastodynamics and thin plates. In: Studies in Applied Mechanics, vol. 10. Elsevier, Amsterdam (1985)
20. Kressner, D.: A block Newton method for nonlinear eigenvalue problems. Numer. Math. **114**(2), 355–372 (2009)
21. Lehoucq, R.B., Sorensen, D.C., Yang, C.: ARPACK Users' Guide. SIAM, Philadelphia (1998)
22. Lorentz, G.G.: Approximation of Functions, 2nd edn. Chelsea, New York (1986)
23. Mackey, D.S., Mackey, N., Mehl, C., Mehrmann, V.: Vector spaces of linearizations for matrix polynomials. SIAM J. Matrix Anal. Appl. **28**, 971–1004 (2006)
24. Meerbergen, K.: The solution of a nonlinear eigenvalue problem using polynomial eigenvalue solvers. In: Talk at 17th ILAS Conference in Braunschweig, Germany (2011)
25. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods. GAMM Mitt. Ges. Angew. Math. Mech. **27**(2), 121–152 (2004)
26. Nowak, A.J., Brebbia, C.A.: The multiple-reciprocity method. A new approach for transforming BEM domain integrals to the boundary. Eng. Anal. Bound. Elem. **6**, 164–168 (1989)
27. Sauter, S.A., Schwab, C.: Boundary Element Methods. Springer Series in Computational Mathematics, vol. 39. Springer, Berlin (2011)
28. Steinbach, O.: Numerical Approximation Methods for Elliptic Boundary Value Problems. Finite and Boundary Elements. Springer, New York (2008)
29. Steinbach, O., Unger, G.: A boundary element method for the Dirichlet eigenvalue problem of the Laplace operator. Numer. Math. **113**, 281–298 (2009)
30. Stewart, G.W.: A Krylov-Schur algorithm for large eigenproblems. SIAM J. Matrix Anal. Appl. **23**(3), 601–614 (2001)
31. Voss, H.: An Arnoldi method for nonlinear eigenvalue problems. BIT Numer. Math. **44**, 387–401 (2004)