Numer. Math. (2011) 119:489–516
DOI 10.1007/s00211-011-0392-1
Numerische Mathematik

Continuation of eigenvalues and invariant pairs for parameterized nonlinear eigenvalue problems

Wolf-Jürgen Beyn · Cedric Effenberger · Daniel Kressner

Received: 23 February 2011 / Revised: 7 June 2011 / Published online: 10 July 2011

Abstract Invariant pairs have been proposed as a numerically robust means to represent and compute several eigenvalues along with the corresponding (generalized) eigenvectors for matrix eigenvalue problems that are nonlinear in the eigenvalue parameter. In this work, we consider nonlinear eigenvalue problems that depend on an additional parameter and our interest is to track several eigenvalues as this parameter varies. Based on the concept of invariant pairs, a theoretically sound and reliable numerical continuation procedure is developed. Particular attention is paid to the situation when the procedure approaches a singularity, that is, when eigenvalues included in the invariant pair collide with other eigenvalues. For the real generic case, it is proven that such a singularity only occurs when two eigenvalues collide on the real axis. It is shown how this situation can be handled numerically by an appropriate expansion of the invariant pair. The viability of our continuation procedure is illustrated by a numerical example.

Mathematics Subject Classification (2000) 65F15 · 65H17 · 47A15 · 65P30

W.-J. Beyn and D. Kressner were supported by CRC 701 'Spectral Analysis and Topological Methods in Mathematics'. C. Effenberger was supported by the SNF research module *Robust numerical methods for solving nonlinear eigenvalue problems* within the SNF ProDoc *Efficient Numerical Methods for Partial Differential Equations*.

W.-J. Beyn
Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

C. Effenberger
Seminar for Applied Mathematics, D-MATH, ETH Zurich, Raemistrasse 101,
8092 Zurich, Switzerland
e-mail: ece@math.ethz.ch

D. Kressner (✉)
MATHICSE, EPF Lausanne, Station 8, 1015 Lausanne, Switzerland
e-mail: daniel.kressner@epfl.ch

1 Introduction

This paper is concerned with eigenvalue problems that are nonlinear in the eigenvalue parameter,

$$T(\lambda)x = 0, \quad x \neq 0, \tag{1}$$

with the matrix-valued function $T : \Omega \to \mathbb{C}^{n\times n}$ having entries that are holomorphic functions on some domain $\Omega \subseteq \mathbb{C}$. In applications, this type of (non-polynomial) nonlinearity usually originates from PDE models with λ-dependent boundary conditions [11,29], λ-dependent material coefficients [31], or from the use of special basis functions in the discretization [5,21,32]; see also [28] for an overview of applications leading to (1). Another prominent example are characteristic functions of delay differential equations [19,30].

Unlike for linear eigenvalue problems, there may be infinitely many eigenvalues λ satisfying (1). In practice, one is typically interested in only a few eigenvalues closest to a target point or a line in the complex plane. Frequently, the matrix-valued function T also depends on one or more real parameters and the goal is to compute and track the eigenvalues of interest as these parameters vary.

The numerical continuation of *one* eigenvalue can be considered a classical topic in numerical analysis; see, e.g., [23,26]. In contrast, the numerical continuation of several eigenvalues has not been investigated to a large extent for nonlinear eigenvalue problems, with the exception of the work [4,9] on polynomial eigenvalue problems. In principle, one could continue several eigenvalues individually, but this approach bears the risk of undetected eigenvalue collisions, does not allow for eigenvalues of higher multiplicity, and can be expected to become quite challenging to implement in a robust manner. For linear eigenvalue problems, the notion of invariant subspaces offers a more convenient, elegant, and robust approach to handling several eigenvalues [8,10,13]. An appropriate extension of this notion to nonlinear eigenvalue problems is nontrivial and has been considered only rather recently for numerical purposes [4,6,9,25]. This extension is most conveniently defined if T takes the form

$$T(\lambda) = f_1(\lambda)T_1 + f_2(\lambda)T_2 + \cdots + f_K(\lambda)T_K \tag{2}$$

for holomorphic functions $f_j : \Omega \to \mathbb{C}$ and matrices $T_j \in \mathbb{C}^{n\times n}$. Then $(X, \Lambda) \in \mathbb{C}^{n\times m} \times \mathbb{C}^{m\times m}$ is called an *invariant pair* if

$$\mathbf{T}(X, \Lambda) := T_1 X f_1(\Lambda) + T_2 X f_2(\Lambda) + \cdots + T_K X f_K(\Lambda) = 0, \tag{3}$$

where $f_j(\Lambda)$ is a matrix function of Λ in the sense of [18]. It is immediate to see that (X, Λ) becomes an eigenvector/eigenvalue pair for $m = 1$, provided that $X \neq 0$. Also for $m > 1$, it can be shown that the eigenvalues of Λ are eigenvalues of the nonlinear eigenvalue problem (1) and X contains the corresponding (generalized) eigenvectors, provided that a certain minimality condition holds; see Sect. 2.2. In fact, Sect. 2.3 will reveal a one-to-one correspondence between invariant pairs and the notions of root functions and eigenvector chains popularized in [15,29].

Although any matrix-valued function can be written in the form (2) with $K \leq n^2$ terms, this form may lead to practical inconveniences for problems with $K \gg 1$. As we will show in Sect. 2.4, an equivalent expression for $\mathbf{T}(X, \Lambda)$ is

$$\mathbf{T}(X, \Lambda) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} T(z) X (zI - \Lambda)^{-1} \, \mathrm{d}z, \tag{4}$$

where Γ is a contour (i.e., a simply closed curve) in Ω containing the spectrum of Λ in its interior. Note that since the eigenvalues are isolated points in Ω, the contour Γ can always be chosen such that its interior is completely contained in Ω. This expression is motivated by the approach in [6]; it has the advantage of being independent of the particular representation (2) and generalizes naturally to operators. However, (4) appears to be more cumbersome for numerical purposes and creates theoretical inconveniences for parameter-dependent problems when eigenvalues cross the contour. We have therefore chosen to base our developments on (2).

The remainder of this paper is organized as follows. In Sect. 2, we briefly review basic facts on invariant pairs and more closely examine their relation to the spectral properties of the nonlinear eigenvalue problem (1). Section 3 is devoted to some theoretical results vital to the continuation method we propose in Sect. 4. We conclude with a numerical example in Sect. 5.

2 Preliminaries

In this section, we briefly review the concepts of root functions and Jordan chains for nonlinear eigenvalue problems, and describe their intimate relationship to invariant pairs. Throughout this paper, we assume that the matrix-valued function T is *regular*, that is, $\det T \not\equiv 0$.

2.1 Root functions and Jordan chains

A holomorphic, vector-valued function $x : \Omega \to \mathbb{C}^n$ is called a *root function* of T at $\lambda_0 \in \Omega$ if

$$T(\lambda_0) x(\lambda_0) = 0, \qquad x(\lambda_0) \neq 0.$$

The order of the zero of $T(\lambda)x(\lambda)$ at $\lambda = \lambda_0$ is called the *multiplicity of the root function* x and will be denoted by $\nu(x)$. Note that the regularity of T implies $\nu(x) < \infty$. Since x is holomorphic, it admits an expansion of the form

$$x(\lambda) := \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j x_j, \qquad x_0 \neq 0. \tag{5}$$

Trivially, x_0 is an eigenvector of T.

Definition 2.1 (*Keldyš* [22]) Given a root function (5) of multiplicity $\nu(x)$, any vector sequence of the form

$$x_0, x_1, \ldots, x_{m-1}, \qquad 1 \le m \le \nu(x), \tag{6}$$

is called a *Jordan chain at* λ_0.

Note that Definition 2.1 coincides with the usual notion of Jordan chains in the case of linear eigenvalue problems, $T(\lambda) = \lambda I - A$.

Any Jordan chain associated with a root function (5) gives rise to a so called *root polynomial*

$$\tilde{x}(\lambda) = \sum_{j=0}^{m-1} (\lambda - \lambda_0)^j x_j, \qquad 1 \le m \le \nu(x),$$

of multiplicity at least m. In particular, taking $m = \nu(x)$ shows that it suffices to work with root polynomials, in general.

We refer to [29, Chapter 1] for a broader introduction into the theory of root functions and Jordan chains in the nonlinear case. For our purpose, it will be sufficient to know that for every eigenvalue λ_0, there is a so called canonical system of root functions $x^{(1)}(\lambda), \ldots, x^{(p)}(\lambda)$ such that their multiplicities sum up to the algebraic multiplicity of λ_0 and $x^{(1)}(\lambda_0), \ldots, x^{(p)}(\lambda_0)$ are linearly independent eigenvectors. The integer p is called the *geometric multiplicity* of λ and equals the dimension of the null space of $T(\lambda_0)$.

2.2 Minimal invariant pairs

Jordan chains are conceptually elegant but fragile under perturbations and, therefore, not well suited for numerical purposes; see [33] for a recent discussion. In a computational setting, it is therefore recommended to replace Jordan chains by the more robust concept of invariant pairs. To describe this replacement, we first remark that the definition (3) of an invariant pair (X, Λ) needs to be complemented by the notion of minimality to exclude degenerate cases, such as $X = 0$.

Definition 2.2 A pair $(X, \Lambda) \in \mathbb{C}^{n\times m} \times \mathbb{C}^{m\times m}$ is called *minimal* if there is an $\ell \in \mathbb{N}$ such that rank $\mathbf{V}_\ell(X, \Lambda) = m$ for

$$\mathbf{V}_\ell(X, \Lambda) := \begin{bmatrix} X \\ X\Lambda \\ \vdots \\ X\Lambda^{\ell-1} \end{bmatrix}. \tag{7}$$

The smallest such ℓ is called the *minimality index* of (X, Λ).

As will be seen below, minimality of an invariant pair (X, Λ) implies that the eigenvalues of Λ are eigenvalues of the nonlinear eigenvalue problem (1). In the case that

the spectrum of Λ consists of a single eigenvalue λ_0, minimal invariant pairs coincide with the notion of *eigenpairs* [16]. Moreover, if Λ is in Jordan canonical form, then (X, Λ) is usually called a *Jordan pair*.

Basic results from mathematical systems theory [20] show that a pair (X, Λ) is minimal if and only if

$$\operatorname{rank} \begin{bmatrix} \lambda I_m - \Lambda \\ X \end{bmatrix} = m \tag{8}$$

for every $\lambda \in \mathbb{C}$ (or, equivalently, for every eigenvalue λ of Λ).

2.3 Equivalence of Jordan chains and invariant pairs

Assuming that the nonlinear eigenvalue problem takes the form (3), a Jordan chain x_0, x_1 at an eigenvalue λ_0 satisfies

$$\begin{aligned} 0 &= T(\lambda_0)x_0 = \sum_{k=1}^{K} f_k(\lambda_0) T_k x_0, \\ 0 &= T'(\lambda_0)x_0 + T(\lambda_0)x_1 = \sum_{k=1}^{K} f_k'(\lambda_0) T_k x_0 + f_k(\lambda_0) T_k x_1. \end{aligned} \tag{9}$$

Noting that

$$f_k\left(\begin{bmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{bmatrix}\right) = \begin{bmatrix} f(\lambda_0) & f'(\lambda_0) \\ 0 & f(\lambda_0) \end{bmatrix},$$

the relations (9) are found to be equivalent to

$$0 = \sum_{k=1}^{K} T_k \, [x_0, \, x_1] \, f_k\left(\begin{bmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{bmatrix}\right).$$

In other words, $(X, J_2(\lambda_0))$ with $X = [x_0, x_1]$ and $J_2(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{bmatrix}$ is an invariant pair. Since $x_0 \neq 0$, $(X, J_2(\lambda_0))$ is minimal. This construction can be extended to Jordan chains of arbitrary length.

Proposition 2.3 ([14, Lemma 2.1]) *Let λ_0 be an eigenvalue of T and consider a matrix $X = \begin{bmatrix} x_0, \ldots, x_{m-1} \end{bmatrix} \in \mathbb{C}^{n \times m}$ with $x_0 \neq 0$. Then $x_0, \ldots, x_{m-1}$ is a Jordan chain at λ_0 if and only if $(X, J_m(\lambda_0))$ is an invariant pair of T, where $J_m(\lambda_0)$ denotes the $m \times m$ Jordan block belonging to λ_0.*

The above result may be generalized to the case of multiple chains of generalized eigenvectors.

Proposition 2.4 *Let λ_0 be an eigenvalue of T and consider a matrix*

$$X = [X^{(1)}, \ldots, X^{(p)}], \qquad X^{(i)} = \left[x_0^{(i)}, \ldots, x_{m_i-1}^{(i)}\right],$$

with $x_0^{(i)} \neq 0$. Then every $x_0^{(i)}, \ldots, x_{m_i-1}^{(i)}$ for $i = 1, \ldots, p$ is a Jordan chain at λ_0 if and only if $\left(X, J_{\lambda_0}\right)$ with $J_{\lambda_0} := \operatorname{diag}\left(J_{m_1}(\lambda_0), \ldots, J_{m_p}(\lambda_0)\right)$ is an invariant pair of T. Moreover, $\left(X, J_{\lambda_0}\right)$ is minimal if and only if $x_0^{(1)}, \ldots, x_0^{(p)}$ are linearly independent.

Proof The first statement follows directly from Proposition 2.3 using the fact that

$$f_k(J_{\lambda_0}) = \operatorname{diag}\left(f(J_{m_1}(\lambda_0)), \ldots, f(J_{m_p}(\lambda_0))\right).$$

The second statement can be easily deduced from the characterization (8) of minimality. □

A similar but slightly less general version of Proposition 2.4 has also been given in [14, Theorem 2.3]. The result of Proposition 2.4 can be extended in a straightforward manner to Jordan chains belonging to different eigenvalues. Again, X contains the vectors of the Jordan chains and Λ is a block diagonal matrix containing the Jordan blocks associated with each Jordan chain on the diagonal. Using (8), it then follows that (X, Λ) is minimal if and only if the first vectors in the Jordan chains (i.e., the eigenvectors) belonging to the *same* eigenvalue are linearly independent. Note that eigenvectors belonging to different eigenvalues are allowed to be linearly dependent; see, e.g., [25] for examples.

To summarize the discussion above: Jordan chains can be stacked into invariant pairs. For the opposite direction, to turn an invariant pair (X, Λ) into Jordan chains, we mention that $(XT, T^{-1}\Lambda T)$ for any invertible T of appropriate size is also an invariant pair. Choosing T such that $T^{-1}\Lambda T$ is in Jordan canonical form, Proposition 2.4 implies that XT contains Jordan chains.

For our considerations, we also need some notion of multiplicity for invariant pairs. In the sequel, we let $\operatorname{alg}_\Lambda \lambda_0$ and $\operatorname{alg}_T \lambda_0$ denote the algebraic multiplicity of an eigenvalue λ_0 for a matrix Λ and a matrix-valued function $T(\lambda)$, respectively. Recall that in the latter case, $\operatorname{alg}_T \lambda_0$ is defined to be the multiplicity of λ_0 as a root of $\det T(\lambda)$. The subsequent definition has been proposed in [8,9] for linear and quadratic eigenvalue problems.

Definition 2.5 The *multiplicity* of a minimal invariant pair (X, Λ) is defined as

$$1 + \sum_{\lambda_0 \in \sigma(\Lambda)} \left(\operatorname{alg}_T \lambda_0 - \operatorname{alg}_\Lambda \lambda_0\right),$$

where $\sigma(\Lambda)$ denotes the spectrum of Λ. A minimal invariant pair of multiplicity 1 is called *simple*.

In view of the discussion above, a simple invariant pair (X, Λ) necessarily contains the entire canonical system of root functions for every eigenvalue of Λ.

2.4 A contour integral representation

To conclude this section, we show that the definition (3) of an invariant pair is equivalent to the contour integral representation (4). This connection will be helpful in some of the subsequent theoretical developments.

Proposition 2.6 *Let* $(X, \Lambda) \in \mathbb{C}^{n\times m} \times \mathbb{C}^{m\times m}$. *Then*

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

where $\Gamma \subset \Omega$ *is a contour with the spectrum of* Λ *in its interior.*

Proof By the contour integral representation of matrix functions [18],

$$f_k(\Lambda) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} f_k(z)(zI - \Lambda)^{-1}\, \mathrm{d}z,$$

and hence

$$\begin{aligned}\mathbf{T}(X, \Lambda) &= \sum_{k=1}^{K} T_k X \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} f_k(z)(zI - \Lambda)^{-1}\, \mathrm{d}z \\ &= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} \sum_{k=1}^{K} f_k(z) T_k X (zI - \Lambda)^{-1}\, \mathrm{d}z \\ &= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} T(z)X(zI - \Lambda)^{-1}\, \mathrm{d}z.\end{aligned}$$

□

As a consequence of Proposition 2.6, we obtain the following formula for the total derivative of $\mathbf{T}$ at a pair (X, Λ) in direction $(\triangle X, \triangle\Lambda)$:

$$\mathbf{DT}(X, \Lambda)(\triangle X, \triangle\Lambda) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} T(z) \left(\triangle X + X(zI - \Lambda)^{-1}\triangle\Lambda\right)(zI - \Lambda)^{-1}\, \mathrm{d}z. \tag{10}$$

3 Characterization of invariant pairs

This section will provide the theoretical foundations of our continuation method. One of the major tools will be the characterization of invariant pairs as solutions to certain nonlinear equations.

3.1 Characterization of simple invariant pairs

By its definition (3), an invariant pair $(X_0, \Lambda_0) \in \mathbb{C}^{n\times m} \times \mathbb{C}^{m\times m}$ satisfies the nonlinear equation

$$\mathbf{T}(X, \Lambda) := T_1 X f_1(\Lambda) + T_2 X f_2(\Lambda) + \cdots + T_K X f_K(\Lambda) = 0. \tag{11}$$

Minimality of a pair (X, Λ) is characterized by the full column rank of the matrix $\mathbf{V}_\ell(X, \Lambda)$ defined in (7), where ℓ is chosen sufficiently large, i.e., not smaller than the minimality index. This motivates the normalization condition

$$\mathbf{N}(X, \Lambda) := W^T \big(\mathbf{V}_\ell(X, \Lambda) - \mathbf{V}_\ell(X_0, \Lambda_0)\big) = 0, \tag{12}$$

where

$$W = \begin{bmatrix} W_0 \\ \vdots \\ W_{\ell-1} \end{bmatrix} \in \mathbb{C}^{\ell n\times m}, \qquad W_0, W_1, \ldots, W_{\ell-1} \in \mathbb{C}^{n\times m}, \tag{13}$$

is chosen such that $W^T \mathbf{V}_\ell(X_0, \Lambda_0)$ is invertible. In [25, Theorem 10] it has been shown that a minimal invariant pair (X_0, Λ_0) is simple if and only if it is a regular solution of

$$\mathbf{F}(X, \Lambda) := \begin{bmatrix} \mathbf{T}(X, \Lambda) \\ \mathbf{N}(X, \Lambda) \end{bmatrix} = 0 \tag{14}$$

in the sense that the total derivative **DF** at (X_0, Λ_0) is a bijective linear operator.

3.2 Characterization of non-simple invariant pairs

In the following, we restrict ourselves to *real* nonlinear eigenvalue problems:

$$T(\lambda) = f_1(\lambda)T_1 + f_2(\lambda)T_2 + \cdots + f_K(\lambda)T_K, \qquad T_k \in \mathbb{R}^{n\times n}, \tag{15}$$

where the holomorphic functions $f_j : \Omega \to \mathbb{C}$ satisfy $f_j(\bar{\lambda}) = \overline{f_j(\lambda)}$ for all $\lambda \in \Omega$, and Ω is supposed to be closed under complex conjugation. In particular, this implies $T(\bar{\lambda}) = \overline{T(\lambda)}$ and hence also the spectrum of $T(\lambda)$ is closed under complex conjugation.

When considering an invariant pair (X, Λ) for (15), it often makes sense to include for every non-real eigenvalue of Λ also its complex conjugate into Λ. We can, therefore, assume that the invariant pair under consideration is real: $(X, \Lambda) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$.

Intuitively, the most likely situation for a real invariant pair to become non-simple is when a real eigenvalue contained in the pair meets a real eigenvalue not contained in the pair; see Fig. 1. This intuition has been made mathematically rigorous for linear eigenvalue problems already in the classic works by Arnol'd [2,3].

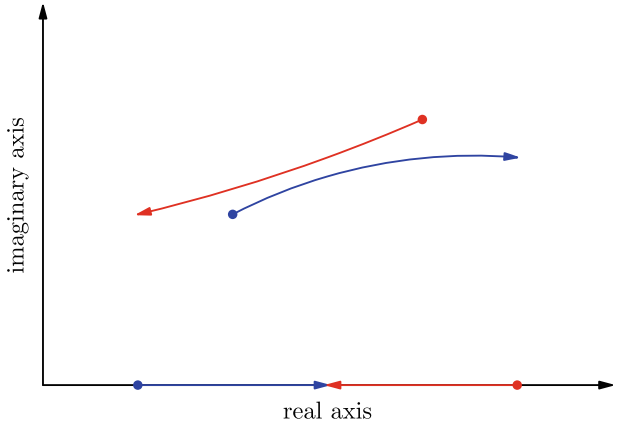

Fig. 1 Illustration of the typical movement of eigenvalues under one-parameter variation: Eigenvalues on the real axis collide while eigenvalues in the complex plane miss each other

From a more general perspective, it is well known that singular solutions in one-parameter systems occur at limit points (see [1,7,17]), where the tangent of the branch is vertical with respect to the parameter coordinate. In a generic sense, these limit points are *quadratic turning points* that are defined by three nondegeneracy conditions. Applied to the nonlinear equations (14), the first two conditions read as follows.

There is $0 \neq (\triangle X_0, \triangle \Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$ such that

$$\ker \mathbf{DF}(X_0, \Lambda_0) = \operatorname{span}\left\{(\triangle X_0, \triangle \Lambda_0)\right\} \tag{TP1}$$

and

$$\mathbf{D}^2\mathbf{F}(X_0, \Lambda_0)(\triangle X_0, \triangle \Lambda_0)^2 \notin \operatorname{im} \mathbf{DF}(X_0, \Lambda_0). \tag{TP2}$$

The third condition, which will be discussed in Sect. 4.1.1, describes transversality with respect to the parameter.

Theorem 3.1 *Let* $(X_0, \Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$ *be a minimal invariant pair of a real, regular nonlinear eigenvalue problem* (15)*. Then the turning point conditions* (TP1) *and* (TP2) *are equivalent to the following set of conditions.*

(J1) *The pair* (X_0, Λ_0) *has multiplicity* 2.
(J2) $T(\lambda)$ *has a real eigenvalue* μ *of geometric multiplicity* 1*, and*

$$\operatorname{alg}_T \mu = 2, \quad \operatorname{alg}_{\Lambda_0} \mu = 1.$$

Proof Set

$$\ell := \sum_{i=1}^{r} \max\left\{3 \operatorname{alg}_{\Lambda_0} \lambda_i, \operatorname{alg}_T \lambda_i + 1\right\}, \tag{16}$$

where $\lambda_1, \ldots, \lambda_r$ are the mutually distinct eigenvalues of Λ_0. Since the functions f_k in the representation (2) of $T(\lambda)$ are holomorphic, there exist polynomials p_k of degree not exceeding ℓ such that

$$p_k^{(j)}(\lambda_i) = f_k^{(j)}(\lambda_i), \qquad i = 1, \ldots, r, \quad j = 0, \ldots, \max\left\{3 \operatorname{alg}_{\Lambda_0} \lambda_i - 1, \operatorname{alg}_T \lambda_i\right\}.$$

Because the eigenvalues of Λ_0 are closed under complex conjugation, p_k inherits the symmetry of f_k with respect to the real axis and, therefore, has real coefficients. Consequently, $P(\lambda) := \sum_{k=1}^{K} T_k\, p_k(\lambda)$ is a matrix-valued polynomial of degree at most ℓ in λ with real coefficient matrices in the monomial expansion. By [27, Theorem 4.1] and well known properties of matrix functions [18], we have

$$\begin{aligned}
&\begin{bmatrix} p_k(\Lambda_0) & \mathbf{D}p_k(\Lambda_0)(\triangle\Lambda_0) & \frac{1}{2}\mathbf{D}^2 p_k(\Lambda_0)(\triangle\Lambda_0)^2 \\ & p_k(\Lambda_0) & \mathbf{D}p_k(\Lambda_0)(\triangle\Lambda_0) \\ & & p_k(\Lambda_0) \end{bmatrix} = p_k\left(\begin{bmatrix} \Lambda_0 & \triangle\Lambda_0 & 0 \\ & \Lambda_0 & \triangle\Lambda_0 \\ & & \Lambda_0 \end{bmatrix}\right) \\
&= f_k\left(\begin{bmatrix} \Lambda_0 & \triangle\Lambda_0 & 0 \\ & \Lambda_0 & \triangle\Lambda_0 \\ & & \Lambda_0 \end{bmatrix}\right) = \begin{bmatrix} f_k(\Lambda_0) & \mathbf{D}f_k(\Lambda_0)(\triangle\Lambda_0) & \frac{1}{2}\mathbf{D}^2 f_k(\Lambda_0)(\triangle\Lambda_0)^2 \\ & f_k(\Lambda_0) & \mathbf{D}f_k(\Lambda_0)(\triangle\Lambda_0) \\ & & f_k(\Lambda_0) \end{bmatrix},
\end{aligned}$$

which implies

$$\begin{aligned}
\mathbf{P}(X_0, \Lambda_0) &= \mathbf{T}(X_0, \Lambda_0), \\
\mathbf{DP}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0) &= \mathbf{DT}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0), \\
\mathbf{D}^2\mathbf{P}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0)^2 &= \mathbf{D}^2\mathbf{T}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0)^2.
\end{aligned}$$

Moreover,

$$\frac{\mathbf{d}^{m_i}}{\mathbf{d}\lambda^{m_i}} \det P(\lambda_i) = \frac{\mathbf{d}^{m_i}}{\mathbf{d}\lambda^{m_i}} \det T(\lambda_i) \neq 0,$$

for $i = 1, \ldots, r$ with $m_i = \operatorname{alg}_T \lambda_i$, which shows that P inherits the regularity of T.

Every occurrence of T in (TP1) and (TP2) can now be replaced by the interpolating polynomial. Hence, the statement of the theorem follows from Lemma A.2. □

Remark 3.2 Depending on the size of the invariant pair (X_0, Λ_0) under consideration and the algebraic multiplicities of the eigenvalues it contains, the degree ℓ of the interpolating polynomial in the proof of Theorem 3.1 (see (16)) may become quite large. Using such a large value of ℓ in the normalization condition (12) can be computationally expensive. However, we have some freedom in the choice of the normalization matrix W. As a consequence, we can work with an $\tilde{\ell}$ as small as the minimality index of (X_0, Λ_0) by choosing the bottom part of W to be zero, without violating the requirement that $W^T\mathbf{V}_\ell(X_0, \Lambda_0)$ be invertible. Usually, the minimality index is small; in fact, except in the case $m > n$, it is often equal to one.

The conditions (J1) and (J2) in Theorem 3.1 state that there exists a Jordan chain x_0, x_1 of length two belonging to a real eigenvalue μ whose first vector x_0 (the eigenvector) is represented in the invariant pair (X_0, Λ_0), but whose second vector x_1 (the associated generalized eigenvector) is not. Adding x_1 to (X_0, Λ_0) yields an enlarged invariant pair which is simple. In fact, the null space of the total derivative **DF** at (X_0, Λ_0) provides all the necessary information to carry out this enlargement.

Theorem 3.3 *Let* $(X_0, \Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$ *be a minimal invariant pair of a real, regular nonlinear eigenvalue problem* (15)*, and let the conditions* (J1), (J2) *of Theorem* 3.1*, and equivalently,* (TP1), (TP2) *be satisfied. Then, the null space of* $\mathbf{DF}(X_0, \Lambda_0)$ *is spanned by a pair* $(\triangle X_0, \triangle\Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$ *having the form*

$$\triangle X_0 = xv^T, \quad \triangle\Lambda_0 = uv^T, \qquad u, v \in \mathbb{R}^m, \quad x \in \mathbb{R}^n, \tag{17}$$

where $v^T\Lambda_0 = \mu v^T$ *and* $x = x_1 + X_0 c$ *for a generalized eigenvector* x_1 *belonging to the real eigenvalue* μ *in* (J2) *and some vector* $c \in \mathbb{R}^m$*. Furthermore, the extended matrices*

$$\hat{X}_0 = \begin{bmatrix} X_0 \, x \end{bmatrix}, \qquad \hat{\Lambda}_0 = \begin{bmatrix} \Lambda_0 & u \\ 0 & \mu \end{bmatrix}$$

constitute a simple invariant pair.

Proof We will explicitly construct a pair having the form (17). According to (J2), there is a Jordan chain x_0, x_1 with $x_0 \neq 0$ satisfying

$$T(\mu)x_1 + T'(\mu)x_0 = 0. \tag{18}$$

Moreover, there is a right eigenvector $\tilde{u} \neq 0$ belonging to the eigenvalue μ of Λ_0 such that $x_0 = X_0\tilde{u}$. For our construction, we also need a corresponding left eigenvector $v \neq 0$. Because μ is a simple eigenvalue of Λ_0, there exists an invertible matrix $U = [\tilde{u}, U^{(1)}]$ with the property

$$U^{-1}\Lambda_0 U = \begin{bmatrix} \mu & 0 \\ 0 & \Lambda_0^{(1)} \end{bmatrix}$$

for a suitably chosen $\Lambda_0^{(1)} \in \mathbb{R}^{(m-1)\times(m-1)}$.

Now consider the pair $(\triangle X_0, \triangle\Lambda_0) = \left(xv^T, uv^T\right)$, where

$$x = x_1 + X_0 c, \qquad u = \tilde{u} - (\mu I - \Lambda_0)c \tag{19}$$

for some $c \in \mathbb{R}^m$. Since $U^{-1}u =: \left[\begin{smallmatrix}1\\ u^{(1)}\end{smallmatrix}\right]$, the vector u, and hence $\triangle\Lambda_0$, are nonzero. From the integral representation (10) of **DT**, we find

$$\begin{aligned}
&\mathbf{DT}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0)\\
&= \frac{1}{2\pi \mathrm{i}} \int_\Gamma T(z) \left(x v^T + X_0 (zI - \Lambda_0)^{-1} u v^T \right) (zI - \Lambda_0)^{-1} \, \mathrm{d}z\\
&= \frac{1}{2\pi \mathrm{i}} \int_\Gamma T(z) \left((x_1 + X_0 c) + X_0 (zI - \Lambda_0)^{-1} \left(\tilde{u} + (\Lambda_0 - \mu I) c\right) \right) v^T \frac{\mathrm{d}z}{z - \mu}\\
&= \frac{1}{2\pi \mathrm{i}} \int_\Gamma T(z) x_1 v^T \frac{\mathrm{d}z}{z - \mu} + \frac{1}{2\pi \mathrm{i}} \int_\Gamma T(z) x_0 v^T \frac{\mathrm{d}z}{(z - \mu)^2}\\
&\quad + \frac{1}{2\pi \mathrm{i}} \int_\Gamma T(z) X_0 \left(I + (zI - \Lambda_0)^{-1} (\Lambda_0 - \mu I) \right) c v^T \frac{\mathrm{d}z}{z - \mu}\\
&= T(\mu) x_1 v^T + T'(\mu) x_0 v^T + \frac{1}{2\pi \mathrm{i}} \int_\Gamma T(z) X_0 (zI - \Lambda_0)^{-1} c v^T \, \mathrm{d}z\\
&= \left(T(\mu) x_1 + T'(\mu) x_0 \right) v^T + \mathbf{T}(X_0, \Lambda_0)\, c v^T = 0.
\end{aligned}$$

Now, define the matrix polynomial $W(z) = W_0 + zW_1 + \cdots + z^{\ell-1} W_{\ell-1} \in \mathbb{R}^{n \times m}$ associated with the partitioning (13) of the matrix W in the normalization condition (12). Then, an analogous derivation shows

$$\mathbf{DN}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0) = \left\{ W(\mu)^T x_1 + W'(\mu)^T x_0 + W^T \mathbf{V}_\ell(X_0, \Lambda_0) c \right\} v^T.$$

Since $W^T \mathbf{V}_\ell(X_0, \Lambda_0)$ is invertible, we may set

$$c = -\left\{ W^T \mathbf{V}_\ell(X_0, \Lambda_0) \right\}^{-1} \left\{ W(\mu)^T x_1 + W'(\mu)^T x_0 \right\}$$

to obtain $\mathbf{DN}(X_0, \Lambda_0)(\triangle X_0, \triangle\Lambda_0) = 0$. This proves the first part of the theorem.

Concerning the extended invariant pair $\left(\hat{X}_0, \hat{\Lambda}_0\right)$, we first confirm its invariance by a direct calculation using the contour integral representation of **T** similarly as above:

$$\begin{aligned}
&\mathbf{T}\left(\hat{X}_0, \hat{\Lambda}_0\right)\\
&= \frac{1}{2\pi \mathbf{i}} \int_\Gamma T(z) \begin{bmatrix} X_0, & x \end{bmatrix} \begin{bmatrix} zI - \Lambda_0 & -u \\ 0 & z - \mu \end{bmatrix}^{-1} \mathrm{d}z\\
&= \frac{1}{2\pi \mathbf{i}} \int_\Gamma T(z) \begin{bmatrix} X_0 (zI - \Lambda_0)^{-1}, & X_0 (zI - \Lambda_0)^{-1} u (z - \mu)^{-1} + x (z - \mu)^{-1} \end{bmatrix} \mathrm{d}z\\
&= \left[\mathbf{T}(X_0, \Lambda_0), \; \frac{1}{2\pi \mathbf{i}} \int_\Gamma T(z) \left\{ X_0 (zI - \Lambda_0)^{-1} \left(\tilde{u} - (\mu I - \Lambda_0) c\right) + x_1 + X_0 c \right\} \frac{\mathrm{d}z}{z - \mu} \right]\\
&= \left[0, \; T(\mu) x_1 + T'(\mu) x_0 + \mathbf{T}(X_0, \Lambda_0) c \right] = 0.
\end{aligned}$$

To verify the minimality of $\left(\hat{X}_0, \hat{\Lambda}_0\right)$, we employ criterion (8) and demonstrate that the matrix

$$\begin{bmatrix} \lambda I - \Lambda_0 & -u \\ 0 & \lambda - \mu \\ X_0 & x \end{bmatrix}. \tag{20}$$

has full column rank for all $\lambda \in \mathbb{C}$. Due to the minimality of (X_0, Λ_0), the first block column of (20) has full column rank for all $\lambda \in \mathbb{C}$. Therefore, it suffices to show that the last column of (20) is linearly independent from the rest. For $\lambda \neq \mu$, this is easily seen by considering the second block row of (20), and for $\lambda = \mu$, by looking at the first block row of the factorization

$$\begin{bmatrix} \mu I - \Lambda_0 & -u \\ 0 & 0 \\ X_0 & x \end{bmatrix} = \begin{bmatrix} U & & \\ & 1 & \\ & & I \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & \mu I - \Lambda_0^{(1)} & -u^{(1)} \\ 0 & 0 & 0 \\ x_0 & X_0 U^{(1)} & x \end{bmatrix} \begin{bmatrix} U^{-1} & \\ & 1 \end{bmatrix}.$$

The statement about the simplicity of $\left(\hat{X}_0, \hat{\Lambda}_0\right)$ follows immediately from the fact that the algebraic multiplicity of μ is raised by one in the transition from Λ_0 to $\hat{\Lambda}_0$ while the multiplicities of the other eigenvalues remain the same. □

4 Continuation of invariant pairs

In the following, we consider a real nonlinear eigenvalue problem

$$T(\lambda, s)x = 0, \qquad x \neq 0 \tag{21}$$

depending on a real, scalar parameter s. The dependence on s is assumed to be sufficiently smooth. Recall that (21) is called real if $T\left(\overline{\lambda}, s\right) = \overline{T(\lambda, s)}$ holds for all λ and all s of interest. In analogy to the parameter-free case, we assume T to be of the form

$$T(\lambda, s) = f_1(\lambda, s)T_1(s) + \cdots + f_K(\lambda, s)T_K(s), \tag{22}$$

where $T_k(s) \in \mathbb{R}^{n \times n}$ and $f_k\left(\overline{\lambda}, s\right) = \overline{f_k(\lambda, s)}$ to enforce realness of the problem.

Let (X_0, Λ_0) be a minimal invariant pair of the nonlinear eigenvalue problem (21) at a fixed parameter value $s = s_0$. The goal is now to continue this invariant pair as the parameter s varies. By the considerations in Sect. 3, locally, the continuation of (X_0, Λ_0) as an invariant pair amounts to the continuation of (X_0, Λ_0) as a solution of the parameterized nonlinear equation

$$\mathbf{F}(X, \Lambda, s) = 0, \tag{23}$$

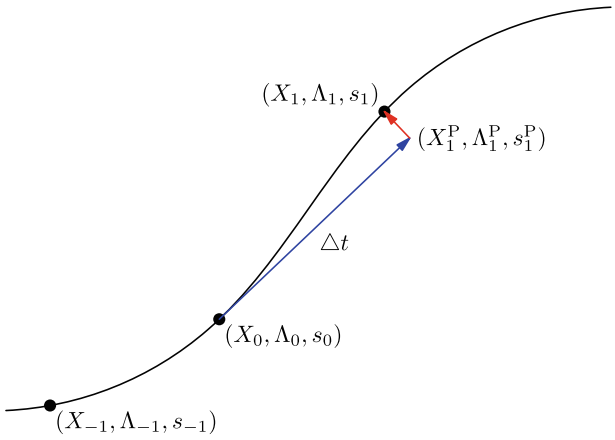

Fig. 2 Illustration of the pseudo-arclength continuation algorithm

where $\mathbf{F}(X, \Lambda, s) = \begin{bmatrix} \mathbf{T}(X, \Lambda, s) \\ \mathbf{N}(X, \Lambda) \end{bmatrix}$ with the normalization condition $\mathbf{N}(X, \Lambda)$ defined as in (12) and

$$\mathbf{T}(X, \Lambda, s) := T_1(s) X f_1(\Lambda, s) + \cdots + T_K(s) X f_K(\Lambda, s). \tag{24}$$

4.1 Pseudo-arclength continuation

To continue an invariant pair, we implement a standard pseudo-arclength continuation technique [12,17]. For this purpose, a reparameterization of the problem (23) is required: We now consider X, Λ, and s as being smoothly dependent on a new parameter t and look for a solution curve $\big(X(t), \Lambda(t), s(t)\big)$ such that

$$\mathbf{F}\big(X(t), \Lambda(t), s(t)\big) = 0.$$

Setting $\big(X(0), \Lambda(0), s(0)\big) := (X_0, \Lambda_0, s_0)$, the continuation of (X, Λ, s) with respect to t proceeds in two steps:

Predictor. Take a step of length $\triangle t$ along the tangent of the solution curve at the current iterate.

Corrector. Determine the next iterate as a point on the solution curve close to the prediction.

This procedure is visualized in Fig. 2.

4.1.1 Predictor

To simplify the notation, we introduce the abbreviations

$$\mathbf{D}_X \mathbf{F}_0 := \mathbf{D}_X \mathbf{F}(X_0, \Lambda_0, s_0), \quad \mathbf{D}_\Lambda \mathbf{F}_0 := \mathbf{D}_\Lambda \mathbf{F}(X_0, \Lambda_0, s_0),$$
$$\mathbf{D}_s \mathbf{F}_0 := \mathbf{D}_s \mathbf{F}(X_0, \Lambda_0, s_0)$$

and denote derivatives with respect to t by dots. To determine the direction $\left(\dot{X}_0, \dot{\Lambda}_0, \dot{s}_0\right)$ of the tangent to the solution curve at the current iterate, we differentiate (23) and obtain the linear system

$$\mathbf{D}_X\mathbf{F}_0\left(\dot{X}_0\right) + \mathbf{D}_\Lambda\mathbf{F}_0\left(\dot{\Lambda}_0\right) + \mathbf{D}_s\mathbf{F}_0(\dot{s}_0) = 0.$$

This needs to be combined with the normalization condition

$$\langle \dot{X}_{-1}, \dot{X}_0\rangle + \langle \dot{\Lambda}_{-1}, \dot{\Lambda}_0\rangle + \langle \dot{s}_{-1}, \dot{s}_0\rangle = 1, \tag{25}$$

where $\left(\dot{X}_{-1}, \dot{\Lambda}_{-1}, \dot{s}_{-1}\right)$ is the tangential direction at the previous iterate. If there is no previous iterate, we simply use $\left(\dot{X}_{-1}, \dot{\Lambda}_{-1}, \dot{s}_{-1}\right) = (0, 0, 1)$ to continue s in positive direction or $\left(\dot{X}_{-1}, \dot{\Lambda}_{-1}, \dot{s}_{-1}\right) = (0, 0, -1)$ to continue s in negative direction. The inner products are trace inner products, weighted by the number of entries, i. e.,

$$\langle \dot{X}_{-1}, \dot{X}_0\rangle = \tfrac{1}{nm}\operatorname{tr}\dot{X}_{-1}^T\dot{X}_0, \qquad \langle \dot{\Lambda}_{-1}, \dot{\Lambda}_0\rangle = \tfrac{1}{m^2}\operatorname{tr}\dot{\Lambda}_{-1}^T\dot{\Lambda}_0, \qquad \langle \dot{s}_{-1}, \dot{s}_0\rangle = \dot{s}_{-1}\dot{s}_0. \tag{26}$$

In summary, we obtain the following linear system in block operator form:

$$\begin{bmatrix} \mathbf{D}_X\mathbf{F}_0(\cdot) & \mathbf{D}_\Lambda\mathbf{F}_0(\cdot) & \mathbf{D}_s\mathbf{F}_0(\cdot) \\ \langle \dot{X}_{-1}, \cdot\rangle & \langle \dot{\Lambda}_{-1}, \cdot\rangle & \langle \dot{s}_{-1}, \cdot\rangle \end{bmatrix} \begin{bmatrix} \dot{X}_0 \\ \dot{\Lambda}_0 \\ \dot{s}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{27}$$

This system has a unique solution provided that the transversality condition

$$\mathbf{D}_s\mathbf{F}_0 \notin \operatorname{im}\left[\mathbf{D}_X\mathbf{F}_0,\ \mathbf{D}_\Lambda\mathbf{F}_0\right] \tag{TP3}$$

holds. By definition, (TP3) is fulfilled at turning points.

Once the tangential direction $\left(\dot{X}, \dot{\Lambda}, \dot{s}\right)$ of the solution curve has been computed from the linear system (27), a first-order prediction of the next iterate is given by

$$\left(X_1^{\mathrm{P}}, \Lambda_1^{\mathrm{P}}, s_1^{\mathrm{P}}\right) = (X_0, \Lambda_0, s_0) + \frac{\triangle t}{\eta}\left(\dot{X}_0, \dot{\Lambda}_0, \dot{s}_0\right), \tag{28}$$

where (X_0, Λ_0, s_0) is the current iterate and $\eta = \left[\langle \dot{X}_0, \dot{X}_0\rangle + \langle \dot{\Lambda}_0, \dot{\Lambda}_0\rangle + \langle \dot{s}_0, \dot{s}_0\rangle\right]^{1/2}$.

4.1.2 Corrector

Using the prediction (28) as initial guess, the continued invariant pair (X_1, Λ_1, s_1) is found by applying Newton's method to the nonlinear equation (23) combined with the normalization condition

$$\langle \dot{X}_0, \triangle X\rangle + \langle \dot{\Lambda}_0, \triangle\Lambda\rangle + \langle \dot{s}_0, \triangle s\rangle = 0,$$

which yields a correction $(\triangle X, \triangle\Lambda, \triangle s)$ orthogonal to the tangential direction.

4.1.3 Solving the linear systems

Both stages, predictor and corrector, require the solution of a linear system of the form

$$\begin{bmatrix} \mathbf{D}_X\mathbf{T}_0(\cdot) & \mathbf{D}_\Lambda\mathbf{T}_0(\cdot) & \mathbf{D}_s\mathbf{T}_0(\cdot) \\ \mathbf{D}_X\mathbf{N}_0(\cdot) & \mathbf{D}_\Lambda\mathbf{N}_0(\cdot) & \mathbf{D}_s\mathbf{N}_0(\cdot) \\ \langle \dot{X}, \cdot \rangle & \langle \dot{\Lambda}, \cdot \rangle & \langle \dot{s}, \cdot \rangle \end{bmatrix} \begin{bmatrix} \triangle X \\ \triangle \Lambda \\ \triangle s \end{bmatrix} = \begin{bmatrix} R \\ S \\ \triangle t \end{bmatrix}. \tag{29}$$

For this purpose, we first transform Λ_0 to upper triangular form by a (complex) Schur decomposition. The structure of (29) is preserved under the corresponding unitary transformation, and we can therefore assume w.l.o.g. that the matrix Λ_0 in (29) is already upper triangular.

Under this assumption, it is easily seen from the contour integral representation (10) that the j-th columns of $\mathbf{D}_X\mathbf{T}_0(\triangle X)$ and $\mathbf{D}_\Lambda\mathbf{T}_0(\triangle\Lambda)$, in the following denoted by $\left[\mathbf{D}_X\mathbf{T}_0(\triangle X)\right]_j$ and $\left[\mathbf{D}_\Lambda\mathbf{T}_0(\triangle\Lambda)\right]_j$, only depend on the first j columns of $\triangle X$ and $\triangle\Lambda$, denoted by $\triangle X_i$ and $\triangle\Lambda_i, i = 1, \ldots, j$. The same is true for the j-th columns of $\mathbf{D}_X\mathbf{N}_0(\triangle X)$ and $\mathbf{D}_\Lambda\mathbf{N}_0(\triangle\Lambda)$, $\left[\mathbf{D}_X\mathbf{N}_0(\triangle X)\right]_j$ and $\left[\mathbf{D}_\Lambda\mathbf{N}_0(\triangle\Lambda)\right]_j$, by an analogous consideration. In other words, for suitable linear operators $\left[\mathbf{D}_X\mathbf{T}_0\right]_{ij}$, $\left[\mathbf{D}_\Lambda\mathbf{T}_0\right]_{ij}$, $\left[\mathbf{D}_X\mathbf{N}_0\right]_{ij}$, and $\left[\mathbf{D}_\Lambda\mathbf{N}_0\right]_{ij}$,

$$\left[\mathbf{D}_X\mathbf{T}_0(\triangle X)\right]_j = \sum_{i=1}^{j} \left[\mathbf{D}_X\mathbf{T}_0\right]_{ij} \triangle X_i, \qquad \left[\mathbf{D}_\Lambda\mathbf{T}_0(\triangle\Lambda)\right]_j = \sum_{i=1}^{j} \left[\mathbf{D}_\Lambda\mathbf{T}_0\right]_{ij} \triangle\Lambda_i,$$

$$\left[\mathbf{D}_X\mathbf{N}_0(\triangle X)\right]_j = \sum_{i=1}^{j} \left[\mathbf{D}_X\mathbf{N}_0\right]_{ij} \triangle X_i, \qquad \left[\mathbf{D}_\Lambda\mathbf{N}_0(\triangle\Lambda)\right]_j = \sum_{i=1}^{j} \left[\mathbf{D}_\Lambda\mathbf{N}_0\right]_{ij} \triangle\Lambda_i.$$

This fact suggests a columnwise forward substitution scheme to solve (29).

More specifically, we will adapt the bordered Bartels–Stewart algorithm from [9,25] to our setting, where $\mathbf{T}$ is given by (24). Using the notation introduced above, we define the matrices

$$L_j = \begin{bmatrix} \left[\mathbf{D}_X\mathbf{T}_0\right]_{jj} & \left[\mathbf{D}_\Lambda\mathbf{T}_0\right]_{jj} & \mathbf{D}_s\mathbf{T}_0 \\ \left[\mathbf{D}_X\mathbf{N}_0\right]_{jj} & \left[\mathbf{D}_\Lambda\mathbf{N}_0\right]_{jj} & 0 \\ \frac{1}{nm}\dot{X}_j & \frac{1}{m^2}\dot{\Lambda}_j & \frac{1}{m}\dot{s}_0 \end{bmatrix}$$

and solve the linear systems

$$L_j \begin{bmatrix} \triangle X_j^0 \\ \triangle\Lambda_j^0 \\ \triangle s_j^0 \end{bmatrix} = \begin{bmatrix} R_j - \sum_{i=1}^{j-1} \left(\left[\mathbf{D}_X\mathbf{T}_0\right]_{ij} \triangle X_i^0 + \left[\mathbf{D}_\Lambda\mathbf{T}_0\right]_{ij} \triangle\Lambda_i^0 \right) \\ S_j - \sum_{i=1}^{j-1} \left(\left[\mathbf{D}_X\mathbf{N}_0\right]_{ij} DX_i^0 + \left[\mathbf{D}_\Lambda\mathbf{N}_0\right]_{ij} \triangle\Lambda_i^0 \right) \\ 0 \end{bmatrix},$$

$$L_j \begin{bmatrix} \triangle X_j^j \\ \triangle \Lambda_j^j \\ \triangle s_j^j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

as well as

$$L_j \begin{bmatrix} \triangle X_j^k \\ \triangle \Lambda_j^k \\ \triangle s_j^k \end{bmatrix} = \begin{bmatrix} -\sum_{i=k}^{j-1} \left(\left[\mathbf{D}_X \mathbf{T}_0 \right]_{ij} \triangle X_i^k + \left[\mathbf{D}_\Lambda \mathbf{T}_0 \right]_{ij} \triangle \Lambda_i^k \right) \\ -\sum_{i=k}^{j-1} \left(\left[\mathbf{D}_X \mathbf{N}_0 \right]_{ij} DX_i^k + \left[\mathbf{D}_\Lambda \mathbf{N}_0 \right]_{ij} \triangle \Lambda_i^k \right) \\ 0 \end{bmatrix}, \quad k = 1, \ldots, j-1.$$

The total number of systems to be solved is $\frac{1}{2}m(m+3)$. The j-th column of the solution to the linear system (29) is then given by the linear combination

$$\triangle X_j = \triangle X_j^0 + \sum_{k=1}^{j} \alpha_k \triangle X_j^k, \quad \triangle \Lambda_j = \triangle \Lambda_j^0 + \sum_{k=1}^{j} \alpha_k \triangle \Lambda_j^k, \quad \triangle s = \triangle s_1^0 + \alpha_1 \triangle s_1^1,$$

where the coefficients $\alpha_1, \ldots, \alpha_m$ satisfy

$$\begin{bmatrix} \triangle s_2^1 - \triangle s_1^1 & \triangle s_2^2 & & \\ \vdots & \vdots & \ddots & \\ \triangle s_m^1 - \triangle s_1^1 & \triangle s_m^2 & \cdots & \triangle s_m^m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} \triangle s_1^0 - \triangle s_2^0 \\ \vdots \\ \triangle s_1^0 - \triangle s_m^0 \\ \triangle t \end{bmatrix}.$$

4.1.4 Step size control

We apply a heuristic step size control to avoid slow or no convergence of the Newton corrector in Sect. 4.1.2. More specifically, the step size is halved if Newton's method does not converge within 5 iterations. On the other hand, the step size is increased by 50% if 3 or less Newton iterations are needed to attain convergence.

4.2 Turning points

The pseudo-arclength continuation method is robust and reliable even in the presence of quadratic turning points as characterized by the conditions (TP1), (TP2), and (TP3). However, it is well known [17, Sect. 2.2] that at a quadratic turning point, $\dot{s}$ switches signs, which corresponds to a reversal of the direction of s at the turning point. In the following, we discuss how this undesirable behavior can be avoided.

4.2.1 Augmenting a nonsimple invariant pair

By Theorem 3.1, a quadratic turning point occurs when a real eigenvalue included in the invariant pair collides with a real eigenvalue not (yet) included. Theorem 3.3 provides the foundation for augmenting the invariant pair to also include the latter eigenvalue.

At a quadratic turning point (X_*, Λ_*, s_*), the component $\dot{s}_*$ of the corresponding tangential direction $\left(\dot{X}_*, \dot{\Lambda}_*, \dot{s}_*\right)$ determined from the linear system (27) is zero. Consequently, $\left(\dot{X}_*, \dot{\Lambda}_*\right)$ spans the null space of $\mathbf{D}_{(X,\Lambda)}\mathbf{F}$ at (X_*, Λ_*, s_*). By Theorem 3.3, $\dot{\Lambda}_*$ is a rank-one matrix of the form (17). We augment (X_*, Λ_*) via the update

$$\hat{X}_* = \left[\, X_* \;\; \dot{X}_* v_1 \,\right], \qquad \hat{\Lambda}_* = \begin{bmatrix} \Lambda_* & u_1 \sigma_1 \\ 0 & v_1^T \Lambda_* v_1 \end{bmatrix},$$

where σ_1 is the largest singular value of $\dot{\Lambda}_*$ and u_1, v_1 are corresponding left and right singular vectors.

After the update, a few steps of Newton's iteration should be executed starting from $\left(\hat{X}_*, \hat{\Lambda}_*, s_*\right)$ to make sure that the new pair is truly invariant. The algorithm is then restarted with the enlarged invariant pair. In particular, in the first step after the restart, we again use $\left(\dot{X}_{-1}, \dot{\Lambda}_{-1}, \dot{s}_{-1}\right) = (0, 0, \pm 1)$ in the normalization condition (25) for the predictor.

4.2.2 Detecting and computing quadratic turning points

Turning points are detected by monitoring $\dot{s}$ computed in the prediction stage during the continuation process. If the values of $\dot{s}(t_0)$ and $\dot{s}(t_1)$ in two consecutive continuation steps are found to have opposite signs, the Intermediate Value Theorem implies the existence of a $t_* \in (t_0, t_1)$ such that $\dot{s}(t_*) = 0$. That is, a potential turning point $(X_*, \Lambda_*, s_*) = \left(X(t_*), \Lambda(t_*), s(t_*)\right)$ has been detected.

To compute t_*, we proceed as described in [12, pp. 259–261]. First, a cubic Hermite interpolating polynomial is constructed, matching the values and derivatives of the curve $\left(X(t), \Lambda(t), s(t)\right)$ at $t = t_0$ as well as $t = t_1$. This yields a first estimate of the turning point by choosing the value $\left(X_*^{\mathrm{P}}, \Lambda_*^{\mathrm{P}}, s_*^{\mathrm{P}}\right)$ of the polynomial at the point $t_*^{\mathrm{P}} \in (t_0, t_1)$ for which s_*^{P} attains an extremum; see Fig. 3. In most cases, this estimate or, more precisely, the derivative $\left(\dot{X}_*^{\mathrm{P}}, \dot{\Lambda}_*^{\mathrm{P}}, \dot{s}_*^{\mathrm{P}}\right)$ at t_*^{P} predicted by the polynomial will be enough to carry out the enlargement of the invariant pair outlined in Sect. 4.2.1. In fact, this has been the case in all of our experiments. In the rare event that this estimate is not sufficiently accurate, it can be further refined by a bisection approach with respect to t. To this end, the estimate is corrected back to the solution curve by the Newton corrector in Sect. 4.1.2, and the procedure is repeated with either t_0 or t_1 replaced by t_*^{P}.

4.3 Overall algorithm

In summary, we obtain Algorithm 1 for continuing an invariant pair.

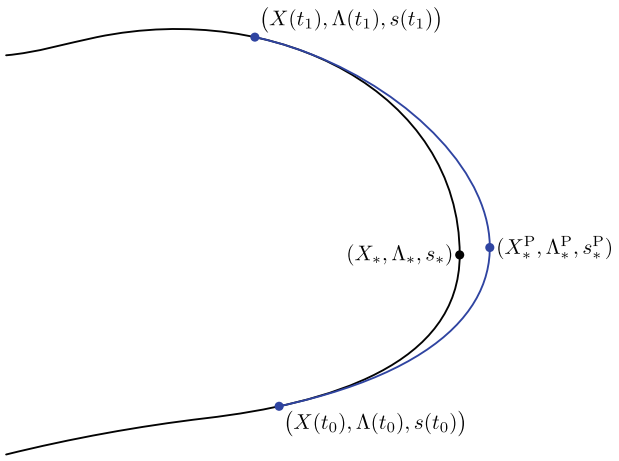

Fig. 3 An estimate of the turning point is obtained from the extremal point of a cubic interpolating polynomial (*blue*)

Algorithm 1 (Continuation of invariant pairs for nonlinear eigenvalue problems)

Input: $T(\lambda, s)$ in the form (22), initial parameter s_0, (approximate) initial invariant pair $\left(X_0^{\mathrm{P}}, \Lambda_0^{\mathrm{P}}\right)$ at $s_0^{\mathrm{P}} = s_0$, initial step size $\triangle t$.
Output: continued invariant pairs (X_i, Λ_i), $i = 0, 1, \ldots$ at parameter values $s_0 < s_1 < s_2 < \cdots$.

% Initialization
$\dot{X}_0 := 0, \quad \dot{\Lambda}_0 := 0, \quad \dot{s}_0 := 1$
$W_k := X_0^{\mathrm{P}} \left(\Lambda_0^{\mathrm{P}}\right)^k$ for $k = 0, \ldots, \ell$
% Continuation
for $i = 0, 1, \ldots$ **do**
  % Corrector
  Apply Newton method from Section 4.1.2 to obtain invariant pair (X_i, Λ_i) at parameter value s_i from estimate $\left(X_i^{\mathrm{P}}, \Lambda_i^{\mathrm{P}}, s_i^{\mathrm{P}}\right)$.
  if Newton process does not converge **then**
    Reduce step size $\triangle t$. Return to predictor if sensible and terminate otherwise.
  end if
  % Handling of turning points and predictor
  Update $W_k := X_i \Lambda_i^k$ for $k = 0, \ldots, \ell$.
  Compute tangential direction $\left(\dot{X}_i, \dot{\Lambda}_i, \dot{s}_i\right)$ at (X_i, Λ_i, s_i) from (27).
  if $\dot{s}_i < 0$ **then**
    Compute turning point as described in Section 4.2.2.
    Augment invariant pair according to Section 4.2.1 and store result in $\left(X_{i+1}^{\mathrm{P}}, \Lambda_{i+1}^{\mathrm{P}}, s_{i+1}^{\mathrm{P}}\right)$.
  else
    Determine $\left(X_{i+1}^{\mathrm{P}}, \Lambda_{i+1}^{\mathrm{P}}, s_{i+1}^{\mathrm{P}}\right)$ by taking a step of length $\triangle t$ along the computed tangent.
  end if
end for

5 Numerical experiment

To verify our implementation of the numerical continuation method detailed in Sect. 4, we have applied it to an academic test problem. We consider a parabolic partial differential equation with a time delay τ:

$$\frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) + a_0 u(x,t) + a_1(x)u(x,t-\tau)$$
$$u(0,t) = u(\pi,t) = 0,$$

with $a_0 = 20$, $a_1(x) = -4.1 + x\big(1 - \mathbf{e}^{x-\pi}\big)$. This example is taken from [19, Sect. 2.4.1], which in turn is a modification of [34, Chapter 3, Example 1.12]. A spatial discretization by finite differences with the uniform grid size $h = \frac{\pi}{n+1}$ yields the delay differential equation

$$\dot{v}(t) = A_0 v(t) + A_1 v(t-\tau) \tag{30}$$

of dimension n, where $v(t) = \big[u(x_1,t),\ldots,u(x_n,t)\big]^T$ with $x_i = \frac{i}{n+1}\pi$, $i = 1,\ldots,n$, and

$$A_0 = \Big(\frac{n+1}{\pi}\Big)^2 \begin{bmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix} + \begin{bmatrix} a_0 & & \\ & \ddots & \\ & & a_0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} a_1(x_1) & & \\ & \ddots & \\ & & a_1(x_n) \end{bmatrix}.$$

For the stability analysis of the delay differential equation (30), one is interested in a few eigenvalues with largest real part of the nonlinear eigenvalue problem

$$\big(-\lambda I + A_0 + \mathbf{e}^{-\tau\lambda} A_1\big)v = 0, \tag{31}$$

which depends on the delay τ as a parameter. In the special case $\tau = 0$, i. e., when there is no delay, the eigenvalue problem (31) is, in fact, linear and symmetric. Therefore, its eigenvalues can be easily computed by standard methods and turn out to be all real. When increasing the delay, several eigenvalues remain real while others collide and form complex conjugate pairs. We apply our continuation algorithm for $n = 100$ to the six eigenvalues with largest real part at $\tau = 0$ and continue them until $\tau = 0.4$. On two occasions eigenvalues collide, the first collision takes place at $\tau \approx 0.051$ and the other one at $\tau \approx 0.078$. In both cases, the step size is decreased and the invariant pair is enlarged. Figure 4 illustrates the obtained results.

6 Conclusions

We have developed a scheme for simultaneously continuing several eigenvalues and (generalized) eigenvectors of a nonlinear eigenvalue problem. The concept of invariant pairs has been proven to be a suitable nonlinear substitute for the concept of invariant subspaces, which is a well-established tool in the linear case and has been successfully used for numerical continuation. On the theoretical side, we prove that in the course of the continuation, turning points only occur upon eigenvalue collisions (Theorem 3.1). Moreover, we show how such collisions can be handled by incorporating additional information into the invariant pair (Theorem 3.3). Based on these results, a numerical

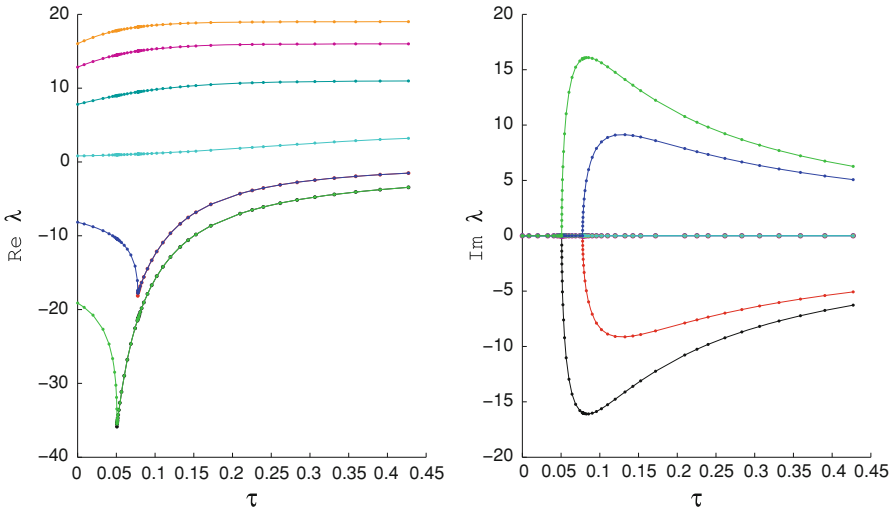

Fig. 4 Continued eigenvalues versus delay τ for the delay eigenvalue problem (31). *Left* real part. *Right* imaginary part

algorithm has been proposed and verified for an example related to delay differential equations. However, algorithmic aspects for large-scale problems as well as the adaptation to a wider range of practical problems remains future work.

Acknowledgments We thank the referees for very helpful comments on an earlier version of this paper.

7 Appendix A: Turning point conditions for linear and polynomial eigenvalue problems

The aim of this appendix is to show the statement of Theorem 3.1 for linear and polynomial eigenvalue problems, which is needed as a basis to prove the statement for the nonlinear case.

7.1 Generalized linear eigenvalue problems

In the linear case, $T(\lambda) = A - \lambda B$, an invariant pair $(\Phi_0, \Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$ satisfies $A\Phi_0 - B\Phi_0\Lambda_0 = 0$. Minimality of the pair simply means that the matrix Φ_0 has full column rank. Hence, the nonlinear equations (14) simplify to

$$\mathbf{F}_L(\Phi, \Lambda) := \begin{bmatrix} A\Phi - B\Phi\Lambda \\ W^T(\Phi - \Phi_0) \end{bmatrix} = 0, \tag{32}$$

where $W \in \mathbb{R}^{n\times m}$ is to be chosen such that $W^T\Phi_0$ is nonsingular.

Theorem 3.1 has been shown for standard linear eigenvalue problems, that is, when B is nonsingular, in [8, Theorem 3]. The following lemma does not assume nonsin-

gularity of B; also, its proof is conceptually somewhat simpler than the proof given in [8].

Lemma A.1 *The statement of Theorem 3.1 holds for real regular linear eigenvalue problems.*

Proof In the linear case, the turning point conditions (TP1)–(TP2) state that there exists a unique $0 \neq (\triangle\Phi_0, \triangle\Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$, up to scaling, such that

$$A\,\triangle\Phi_0 - B\,\triangle\Phi_0\Lambda_0 - B\Phi_0\,\triangle\Lambda_0 = 0, \quad W^T\triangle\Phi_0 = 0, \tag{TP1L}$$

as well as that the equation

$$A\,\triangle\Phi - B\,\triangle\Phi\Lambda_0 - B\Phi_0\,\triangle\Lambda = -2B\,\triangle\Phi_0\,\triangle\Lambda_0, \quad W^T\triangle\Phi = 0. \tag{TP2L}$$

admits no solution $(\triangle\Phi, \triangle\Lambda)$. After suitable transformations, we may, w.l.o.g., assume Λ_0 to be in Jordan canonical form and (A, B) to be in (permuted) Kronecker canonical form such that $\Phi_0 = [\frac{I}{0}]$. Then, the above equations read

$$\begin{bmatrix}\Lambda_0 & A_{12}\\ 0 & A_{22}\end{bmatrix}\begin{bmatrix}\triangle\Phi_{0,1}\\ \triangle\Phi_{0,2}\end{bmatrix} - \begin{bmatrix}I & 0\\ 0 & B_{22}\end{bmatrix}\begin{bmatrix}\triangle\Phi_{0,1}\\ \triangle\Phi_{0,2}\end{bmatrix}\Lambda_0 - \begin{bmatrix}\triangle\Lambda_0\\ 0\end{bmatrix} = 0,$$
$$W_1^T\triangle\Phi_{0,1} + W_2^T\triangle\Phi_{0,2} = 0, \tag{TP1L$'$}$$

and

$$\begin{bmatrix}\Lambda_0 & A_{12}\\ 0 & A_{22}\end{bmatrix}\begin{bmatrix}\triangle\Phi_1\\ \triangle\Phi_2\end{bmatrix} - \begin{bmatrix}I & 0\\ 0 & B_{22}\end{bmatrix}\begin{bmatrix}\triangle\Phi_1\\ \triangle\Phi_2\end{bmatrix}\Lambda_0 - \begin{bmatrix}\triangle\Lambda\\ 0\end{bmatrix} = -2\begin{bmatrix}I & 0\\ 0 & B_{22}\end{bmatrix}\begin{bmatrix}\triangle\Phi_{0,1}\\ \triangle\Phi_{0,2}\end{bmatrix}\triangle\Lambda_0,$$
$$W_1^T\,\triangle\Phi_1 + W_2^T\,\triangle\Phi_2 = 0, \tag{TP2L$'$}$$

where we have partitioned $\triangle\Phi_0 = \begin{bmatrix}\triangle\Phi_{0,1}\\ \triangle\Phi_{0,2}\end{bmatrix}$, $W = \begin{bmatrix}W_1\\ W_2\end{bmatrix}$, and $\triangle\Phi = \begin{bmatrix}\triangle\Phi_1\\ \triangle\Phi_2\end{bmatrix}$ conformally. The invertibility of $W^T\Phi_0$ yields the invertibility of W_1, implying that (TP1L$'$) is equivalent to

$$\triangle\Phi_{0,1} = \underbrace{-\left(W_1^T\right)^{-1}W_2^T}_{=:G}\triangle\Phi_{0,2}, \quad \triangle\Lambda_0 = \Lambda_0 G\,\triangle\Phi_{0,2} + A_{12}\,\triangle\Phi_{0,2} - G\,\triangle\Phi_{0,2}\Lambda_0, \tag{33}$$

$$A_{22}\,\triangle\Phi_{0,2} - B_{22}\,\triangle\Phi_{0,2}\Lambda_0 = 0. \tag{34}$$

As a consequence of (33), fixing $\triangle\Phi_{0,2}$ uniquely determines both $\triangle\Phi_{0,1}$ and $\triangle\Lambda_0$. Known results [24] concerning null spaces of generalized Sylvester equations show that (34) (and consequently (TP1L$'$)) has a one-dimensional solution space if and only if

1. Λ_0 and the pair (A_{22}, B_{22}) share exactly one eigenvalue $\mu \in \mathbb{R}$;
2. Λ_0 and (A_{22}, B_{22}) each have exactly one Jordan block belonging to μ, denoted by $J_{n_1}(\mu)$ and $J_{n_2}(\mu)$, respectively, with $\min\{n_1, n_2\} = 1$, where n_1, n_2 are the sizes of the blocks.

Let us now assume that (TP1L′) holds and, consequently, both conditions are fulfilled. By applying a suitable permutation, we may assume

$$\Lambda_0 = \begin{bmatrix} J_{n_1}(\mu) & 0 \\ 0 & \Lambda_{0,2} \end{bmatrix}, \quad A_{22} = \begin{bmatrix} J_{n_2}(\mu) & 0 \\ 0 & A_{22,2} \end{bmatrix}, \quad B_{22} = \begin{bmatrix} I_{n_2} & 0 \\ 0 & B_{22,2} \end{bmatrix}.$$

By [24, Theorem 4], every solution of the generalized Sylvester equation (34) is $\triangle\Phi_{0,2} = e_1 e_{n_1}^T$ or a scalar multiple thereof. From (33), it then follows that

$$\triangle\Phi_{0,1} = G e_1 e_{n_1}^T, \quad \triangle\Lambda_0 = (\Lambda_0 - \mu I) G e_1 e_{n_1}^T + A_{12} e_1 e_{n_1}^T. \tag{35}$$

Hence, the right-hand side of (TP2L′) becomes

$$\begin{aligned} -2 \begin{bmatrix} I & 0 \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} \triangle\Phi_{0,1} \\ \triangle\Phi_{0,2} \end{bmatrix} \triangle\Lambda_0 &= -2 \begin{bmatrix} G e_1 \\ B_{22} e_1 \end{bmatrix} \big(e_{n_1}^T (\Lambda_0 - \mu I) G + e_{n_1}^T A_{12} \big) e_1 e_{n_1}^T \\ &= -2 \begin{bmatrix} G e_1 \\ e_1 \end{bmatrix} e_{n_1}^T A_{12} e_1 e_{n_1}^T. \end{aligned}$$

Let us now also assume that there is no solution to (TP2L′). This immediately allows us to exclude the case $A_{12} e_1 = 0$ (meaning the two blocks $J_{n_2}(\mu)$, $J_{n_1}(\mu)$ are uncoupled in the permuted Kronecker form of (A, B)) since this would entail that (TP2L′) has zero right-hand side and thus admits the solution $(\triangle\Phi, \triangle\Lambda) = (0, 0)$. Hence, the two blocks $J_{n_2}(\mu)$, $J_{n_1}(\mu)$ are necessarily coupled and, therefore, $A_{12} e_1 = e_{n_1}$. This leads to (TP2L′) taking the form

$$\begin{aligned} \begin{bmatrix} \Lambda_0 & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \triangle\Phi_1 \\ \triangle\Phi_2 \end{bmatrix} - \begin{bmatrix} I & 0 \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} \triangle\Phi_1 \\ \triangle\Phi_2 \end{bmatrix} \Lambda_0 - \begin{bmatrix} \triangle\Lambda \\ 0 \end{bmatrix} &= -2 \begin{bmatrix} G e_1 e_{n_1}^T \\ e_1 e_{n_1}^T \end{bmatrix} \\ W_1^T \triangle\Phi_1 + W_2^T \triangle\Phi_2 &= 0. \end{aligned} \tag{36}$$

If $n_1 \geq 2$, one easily verifies that a solution of (36) is given by

$$\triangle\Phi = 2 \begin{bmatrix} G e_1 \\ e_1 \end{bmatrix} e_{n_1 - 1}^T, \quad \triangle\Lambda = 2\big((\Lambda_0 - \mu I) G e_1 + e_{n_1} \big) e_{n_1 - 1}^T.$$

Similarly, for $n_2 \geq 2$, a solution of (36) is given by

$$\triangle\Phi = -2 \begin{bmatrix} G e_2 \\ e_2 \end{bmatrix} e_{n_1}^T, \quad \triangle\Lambda = 2\big(G e_1 - (\Lambda_0 - \mu I) G e_2 \big) e_{n_1}^T.$$

Consequently, for (36) having no solution, $n_1 = n_2 = 1$, and hence μ has algebraic multiplicity 2 and geometric multiplicity 1. This shows one direction of the statement of Theorem 3.1.

For the opposite direction, assume that $A-\lambda B$ has a real eigenvalue μ of geometric multiplicity 1 and algebraic multiplicity 2 such that $\mathrm{alg}_{\Lambda_0}\mu = 1$. Then, in the permuted Kronecker form of (A, B), we have

$$\Lambda_0 = \begin{bmatrix} \mu & 0 \\ 0 & \Lambda_{0,2} \end{bmatrix}, \quad A_{22} = \begin{bmatrix} \mu & 0 \\ 0 & A_{22,2} \end{bmatrix}, \quad B_{22} = \begin{bmatrix} 1 & 0 \\ 0 & B_{22,2} \end{bmatrix}.$$

Consequently, the generalized Sylvester equation (34) has, up to scaling, the unique solution $\triangle\Phi_{0,2} = e_1 e_1^T$, implying (TP1L′). Furthermore, using this solution, the lower part of equation (TP2L′) becomes

$$\begin{bmatrix} \mu & 0 \\ 0 & A_{22,2} \end{bmatrix} \triangle\Phi_2 - \begin{bmatrix} 1 & 0 \\ 0 & B_{22,2} \end{bmatrix} \triangle\Phi_2 \begin{bmatrix} \mu & 0 \\ 0 & \Lambda_{0,2} \end{bmatrix} = -2e_1 e_1^T, \tag{37}$$

which—by considering the (1, 1)-element—clearly has no solution. □

7.2 Polynomial eigenvalue problems

In the polynomial case, $P(\lambda) = \sum_{k=0}^{p} \lambda^k P_k$, an invariant pair (X_0, Λ_0) satisfies

$$\mathbf{P}(X_0, \Lambda_0) := \sum_{k=0}^{p} P_k X_0 \Lambda_0^k = 0.$$

If (X_0, Λ_0) is minimal, its minimality index cannot exceed p (see [4]); we can, therefore, use $\ell = p$ in the normalization condition (12). The nonlinear equations (14) thus amount to

$$\mathbf{F}_P(X, \Lambda) := \begin{bmatrix} \mathbf{P}(X, \Lambda) \\ W^T\big(\mathbf{V}_p(X, \Lambda) - \mathbf{V}_p(X_0, \Lambda_0)\big) \end{bmatrix} = 0, \tag{38}$$

where W must be chosen such that $W^T\mathbf{V}_p(X_0, \Lambda_0)$ is invertible.

The proof of the following lemma follows from the proof of [9, Theorem 2.5] for quadratic eigenvalue problems, with the notable difference that the leading coefficient P_p is not assumed to be invertible.

Lemma A.2 *The statement of Theorem 3.1 holds for real regular polynomial eigenvalue problems.*

Proof By linearization, the polynomial eigenvalue problem $P(\lambda) = \sum_{k=0}^{p} \lambda^k P_k$ is equivalent to the linear eigenvalue problem

$$A - \lambda B := \begin{bmatrix} 0 & I & & \\ & \ddots & \ddots & \\ & & 0 & I \\ -P_0 & \cdots & -P_{p-2} & -P_{p-1} \end{bmatrix} - \lambda \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & P_p \end{bmatrix}, \tag{39}$$

in the sense that (X_0, Λ_0) is an invariant pair of the polynomial eigenvalue problem if and only if (Φ_0, Λ_0) with $\Phi_0 := \mathbf{V}_p(X_0, \Lambda_0)$ is an invariant pair for (39). Letting (J1P)+(J2P) and (J1L)+(J2L) denote the eigenvalue conditions of Theorem 3.1 for $P(\lambda)$ and $A-\lambda B$, respectively, it is well known [15] that (J1P)+(J2P) $\Leftrightarrow$ (J1L)+(J2L). Moreover, Lemma A.1 shows (TP1L)+(TP2L) $\Leftrightarrow$ (J1L)+(J2L). Hence in order to show the statement of the theorem, (TP1P)+(TP2P) $\Leftrightarrow$ (J1P)+(J2P), it suffices to show (TP1P)+(TP2P) $\Leftrightarrow$ (TP1L)+(TP2L), i.e., the equivalence of the turning point conditions for the polynomial and the linearized eigenvalue problem.

(TP1P) $\Leftrightarrow$ (TP1L). The first turning point condition for $P(\lambda)$ states that there exists a unique $0 \neq (\triangle X_0, \triangle \Lambda_0) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times m}$, up to scaling, such that

$$\mathbf{DF}_P(X_0, \Lambda_0)\,(\triangle X_0, \triangle \Lambda_0) = \begin{bmatrix} \sum_{k=0}^{p} P_k \mathbf{D}\big(X_0\Lambda_0^k\big)(\triangle X_0, \triangle \Lambda_0) \\ \sum_{k=0}^{p-1} W_k^T \mathbf{D}(X_0, \Lambda_0)(\triangle X_0, \triangle \Lambda_0) \end{bmatrix} = 0, \quad \text{(TP1P)}$$

where $\mathbf{D}(X_0\Lambda_0)(\triangle X_0, \triangle \Lambda_0)$ denotes the Fréchet derivative of the mapping $(X, \Lambda) \mapsto X\Lambda^k$ at $(X, \Lambda) = (X_0, \Lambda_0)$ in the direction $(\triangle X_0, \triangle \Lambda_0)$. From the product rule, one can easily derive the recursion

$$\mathbf{D}\Big(X_0\Lambda_0^{k+1}\Big)(\triangle X_0, \triangle \Lambda_0) = \mathbf{D}\Big(X_0\Lambda_0^k\Big)(\triangle X_0, \triangle \Lambda_0)\Lambda_0 + X_0\Lambda_0^k\triangle\Lambda_0, \quad (40)$$

which will be useful later on. The equivalence of (TP1P) and (TP1L) is contained in the proof of Theorem 7 in [4]; its detailed proof is therefore omitted. We only note that (TP1P) implies that $(\triangle\Phi_0, \triangle\Lambda_0)$ with

$$\triangle\Phi_0 = \begin{bmatrix} \triangle X_0 \\ \mathbf{D}\,(X_0\Lambda_0)\,(\triangle X_0, \triangle \Lambda_0) \\ \vdots \\ \mathbf{D}\Big(X_0\Lambda_0^{p-1}\Big)\,(\triangle X_0, \triangle \Lambda_0) \end{bmatrix} \quad (41)$$

satisfies (TP1L). In the opposite direction, for any $(\triangle\Phi_0, \triangle\Lambda_0)$ satisfying (TP1L), the matrix $\triangle\Phi_0$ necessarily takes the form (41) and $(\triangle X_0, \triangle \Lambda_0)$ satisfies (TP1P).

(TP2P) $\Leftrightarrow$ (TP2L). Now, assume $(\triangle X_0, \triangle \Lambda_0) \neq 0$ fulfills (TP1P) or, equivalently, $(\triangle\Phi_0, \triangle\Lambda_0) \neq 0$ with $\triangle\Phi_0$ of the form (41) fulfills (TP1L). The second turning point condition for $P(\lambda)$ states that

$$\mathbf{DF}_P(X_0, \Lambda_0)(\triangle X_1, \triangle \Lambda_1) = \mathbf{D}^2\mathbf{F}_P(X_0, \Lambda_0)(\triangle X_0, \triangle \Lambda_0)^2, \quad \text{(TP2P)}$$

with

$$\mathbf{D}^2\mathbf{F}_P(X_0,\Lambda_0)(\triangle X_0,\triangle\Lambda_0)^2=\begin{bmatrix}\sum_{k=0}^{p}P_k\mathbf{D}^2(X_0\Lambda_0)(\triangle X_0,\triangle\Lambda_0)^2\\ \sum_{k=0}^{p-1}W_k^T\mathbf{D}^2(X_0\Lambda_0)(\triangle X_0,\triangle\Lambda_0)^2\end{bmatrix}$$

has no solution $(\triangle X_1,\triangle\Lambda_1)$. Here, $\mathbf{D}^2(X_0\Lambda_0)(\triangle X_0,\triangle\Lambda_0)^2$ denotes the second Fréchet derivative of the mapping $(X,\Lambda)\mapsto X\Lambda^k$ at $(X,\Lambda)=(X_0,\Lambda_0)$ twice in the direction $(\triangle X_0,\triangle\Lambda_0)$. Differentiating (40) yields the recursion

$$\begin{aligned}\mathbf{D}^2\Big(X_0\Lambda_0^{k+1}\Big)(\triangle X_0,\triangle\Lambda_0)^2&=\mathbf{D}^2\Big(X_0\Lambda^k\Big)(\triangle X_0,\triangle\Lambda_0)^2\Lambda_0\\&\quad+2\mathbf{D}\Big(X_0\Lambda_0^k\Big)(\triangle X_0,\triangle\Lambda_0)\triangle\Lambda_0.\end{aligned}\tag{42}$$

Now, let us assume that there is $(\triangle\Phi_1,\triangle\Lambda_1)$ satisfying (TP2L). Partitioning

$$\triangle\Phi_1=\begin{bmatrix}\triangle\Phi_{1,0}\\ \vdots\\ \triangle\Phi_{1,p-1}\end{bmatrix}\tag{43}$$

in accordance with the block structure of A and B, the relation (TP2L) is easily seen to be equivalent to the following set of equations:

$$\begin{aligned}&0\le k\le p-2:\quad \triangle\Phi_{1,k+1}-\triangle\Phi_{1,k}\Lambda_0-\Phi_{0,k}\triangle\Lambda_1=-2\triangle\Phi_{0,k}\triangle\Lambda_0,\\&-\sum_{k=0}^{p-1}P_k\triangle\Phi_{1,k}-P_p\triangle\Phi_{1,p-1}\Lambda_0-P_p\Phi_{0,p-1}\triangle\Lambda_1=-2P_p\triangle\Phi_{0,p-1}\triangle\Lambda_0,\\&\qquad\qquad\sum_{k=0}^{p-1}W_k^T\triangle\Phi_{1,k}=0.\end{aligned}\tag{44}$$

Combined with the recurrences (40) and (42), induction on k shows that (44) is equivalent to

$$\triangle\Phi_{1,k}=\mathbf{D}\Big(X_0\Lambda_0^k\Big)(\triangle\Phi_{1,0},\triangle\Lambda_1)-\mathbf{D}^2\Big(X_0\Lambda_0^k\Big)(\triangle X_0,\triangle\Lambda_0)^2\tag{45}$$

for $k=0,\dots,p-1$. If we set $\triangle X_1:=\triangle\Phi_{1,0}$, this equation implies, again with the aid of the recurrences (40) and (42),

$$\begin{aligned}&\sum_{k=0}^{p}P_k\mathbf{D}\Big(X_0\Lambda_0^k\Big)(\triangle X_1,\triangle\Lambda_1)-\sum_{k=0}^{p}P_k\mathbf{D}^2\Big(X_0\Lambda_0^k\Big)(\triangle X_0,\triangle\Lambda_0)^2\\&\quad=\sum_{k=0}^{p-1}P_k\triangle\Phi_{1,k}+P_p\big[\triangle\Phi_{1,p-1}\Lambda_0+\Phi_{0,p-1}\triangle\Lambda_1-2\triangle\Phi_{0,p-1}\triangle\Lambda_0\big]=0\end{aligned}$$

and

$$\sum_{k=0}^{p-1} W_k^T \mathbf{D}\Big(X_0\Lambda_0^k\Big)(\triangle X_1, \triangle\Lambda_1) - \sum_{k=0}^{p-1} W_k^T \mathbf{D}^2\Big(X_0\Lambda_0^k\Big)(\triangle X_0, \triangle\Lambda_0)^2 = \sum_{k=0}^{p-1} W_k^T \triangle\Phi_{1,k} = 0.$$

Hence, $(\triangle X_1, \triangle\Lambda_1)$ satisfies the relation (TP2P).

Conversely, if $(\triangle X_1, \triangle\Lambda_1)$ satisfies (TP2P), we can set $\triangle\Phi_{1,0} := \triangle X_1$, and then $(\triangle\Phi_1, \triangle\Lambda_1)$ with $\triangle\Phi_1$ defined by (43) and (45) will satisfy (TP2L). □

References

1. Allgower, E.L., Georg, K.: Numerical continuation methods. In: Springer Series in Computational Mathematics, vol. 13. An Introduction. Springer, Berlin (1990)
2. Arnol'd, V.I.: Matrices depending on parameters. Uspehi Mat. Nauk **26**(2(158)), 101–114 (1971)
3. Arnol'd, V.I.: Lectures on bifurcations and versal families. Uspehi Mat. Nauk **27**(5(167)), 119–184 A series of articles on the theory of singularities of smooth mappings (1972)
4. Betcke, T., Kressner, D.: Perturbation, extraction and refinement of invariant pairs for matrix polynomials. Linear Algebra Appl. **435**(3), 514–536 (2011)
5. Betcke, T., Trefethen, L.N.: Reviving the method of particular solutions. SIAM Rev. **47**(3), 469–491 (2005)
6. Beyn, W.-J.: An integral method for solving nonlinear eigenvalue problems. Linear Algebra Appl. arXiv:1003.1580 (2010)
7. Beyn, W.-J., Champneys, A., Doedel, E., Govaerts, W., Kuznetsov, Y.A., Sandstede, B.: Numerical continuation, and computation of normal forms. In: Handbook of Dynamical Systems, vol. 2, pp. 149–219. North-Holland, Amsterdam (2002)
8. Beyn, W.-J., Kleß, W., Thümmler, V.: Continuation of low-dimensional invariant subspaces in dynamical systems of large dimension. In: Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems, pp. 47–72. Springer, Berlin (2001)
9. Beyn, W.-J., Thümmler, V.: Continuation of invariant subspaces for parameterized quadratic eigenvalue problems. SIAM J. Matrix Anal. Appl. **31**(3), 1361–1381 (2009)
10. Bindel, D. S., Demmel, J.W., Friedman, M.: Continuation of invariant subspaces in large bifurcation problems. SIAM J. Sci. Comput. **30**(2), 637–656 (2008)
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. Deuflhard,P.: Newton methods for nonlinear problems—affine invariance and adaptive algorithms. In: Springer Series in Computational Mathematics, vol. 35. Springer, Berlin (2006). Corr. 2nd printing edition
13. Dieci, L., Friedman, M.J.: Continuation of invariant subspaces. Numer. Linear Algebra Appl. **8**(5), 317–327 (2001)
14. Gohberg, I., Kaashoek, M.A., van Schagen, F.: On the local theory of regular analytic matrix functions. Linear Algebra Appl. **182**, 9–25 (1993)
15. Gohberg, I., Lancaster, P., Rodman, L.: Matrix Polynomials. Academic Press, New York (1982)
16. Gohberg, I., Rodman, L.: Analytic matrix functions with prescribed local data. J. Anal. Math. **40**, 90–128 (1982)
17. Govaerts, W.: Numerical Methods for Bifurcations of Dynamical Equilibria. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2000)
18. Higham, N.J.: Functions of Matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia (2008)
19. Jarlebring, E.: The spectrum of delay-differential equations: numerical methods, stability and perturbation. PhD thesis, Inst. Comp. Math, TU Braunschweig (2008)

20. Kailath, T.: Linear Systems. In: Prentice-Hall Information and System Sciences Series. Prentice-Hall Inc., Englewood Cliffs (1980)
21. Kamiya, N., Wu, S.T.: Generalized eigenvalue formulation of the Helmholtz equation by the Trefftz method. Eng. Comput. **11**, 177–186 (1994)
22. Keldyš, M.V.: On the characteristic values and characteristic functions of certain classes of non-self-adjoint equations. Doklady Akad. Nauk SSSR (N.S.) **77**, 11–14 (1951)
23. Keller, H.B.: Numerical solution of bifurcation and nonlinear eigenvalue problems. In: Applications of Bifurcation Theory (Proc. Advanced Sem., Univ. Wisconsin, Madison, Wis., 1976), pp. 359–384. Publ. Math. Res. Center, No. 38. Academic Press, New York (1977)
24. Košir, T.: Kronecker bases for linear matrix equations, with application to two-parameter eigenvalue problems. Linear Algebra Appl. **249**, 259–288 (1996)
25. Kressner, D.: A block Newton method for nonlinear eigenvalue problems. Numer. Math. **114**, 355–372 (2009)
26. Lui, S.H., Keller, H.B., Kwok, T.W.C.: Homotopy method for the large, sparse, real nonsymmetric eigenvalue problem. SIAM J. Matrix Anal. Appl. **18**(2), 312–333 (1997)
27. Mathias, R.: A chain rule for matrix functions and applications. SIAM J. Matrix Anal. Appl. **17**(3), 610–620 (1996)
28. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods. GAMM Mitteilungen 27 (2004)
29. Mennicken, R., Möller, M.: Non-Self-Adjoint Boundary Eigenvalue Problems. In: North-Holland Mathematics Studies, vol. 192. North-Holland, Amsterdam (2003)
30. Michiels, W., Niculescu, S.I.: Stability and Stabilization of Time-Delay Systems. In: Advances in Design and Control, vol. 12. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2007)
31. Moreno, E., Erni, D., Hafner, C.: Band structure computations of metallic photonic crystals with the multiple multipole method. Phys. Rev. B **65**(15), 155120 (2002)
32. Steinbach, O., Unger, G.: A boundary element method for the Dirichlet eigenvalue problem of the Laplace operator. Numer. Math. **113**(2), 281–298 (2009)
33. Wilkening, J.: An algorithm for computing Jordan chains and inverting analytic matrix functions. Linear Algebra Appl. **427**(1), 6–25 (2007)
34. Wu, J.: Theory and Applications of Partial Functional Differential Equations. In: Applied Mathematical Sciences, vol. 119. Springer, New York (1996)